VMware vSphere: Optimize and Scale

Lecture Manual
ESXi 6.7 and vCenter Server 6.7

VMware® Education Services
VMware, Inc.
www.vmware.com/education

https://t.me/learningnets

VMware vSphere:
Optimize and Scale
Lecture Manual
ESXi 6.7 and vCenter Server 6.7
Part Number EDU-EN-VSOS67-LECT (7/2018)

Copyright © 2018 VMware, Inc. All rights reserved. This manual and its accompanying materials are protected by U.S. and international copyright and intellectual property laws. VMware products are covered by one or more patents listed at http://www.vmware.com/go/patents. VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

The training material is provided “as is,” and all express or implied conditions, representations, and warranties, including any implied warranty of merchantability, fitness for a particular purpose or noninfringement, are disclaimed, even if VMware, Inc., has been advised of the possibility of such claims. This training material is designed to support an instructor-led training course and is intended to be used for reference purposes in conjunction with the instructor-led training course. The training material is not a standalone training tool. Use of the training material for self-study without class attendance is not recommended.

These materials and the computer programs to which it relates are the property of, and embody trade secrets and confidential information proprietary to, VMware, Inc., and may not be reproduced, copied, disclosed, transferred, adapted or modified without the express written approval of VMware, Inc.

www.vmware.com/education

https://t.me/learningnets

CONTENTS

Module 1 Course Introduction 1
1-2 Importance 2
1-3 Learner Objectives 3
1-4 You Are Here 4
1-5 Typographical Conventions 5
1-6 References (1) 6
1-7 References (2) 7
1-8 VMware Online Resources 8
1-9 Review of vSphere User Interfaces 9
1-10 About Scalability 10
1-11 About Optimization 11
1-12 About Security 12
1-13 VMware Education Overview 13
1-14 VMware Certification Overview 14

Module 2 Network Scalability 15
2-2 You Are Here 16
2-3 Importance 17
2-4 Module Lessons 18
2-5 Lesson 1: Creating vSphere Distributed Switches 19
2-6 Learner Objectives 20
2-7 About Distributed Switches 21
2-8 Distributed Switch Architecture 22
2-9 Standard Switch and Distributed Switch Feature Comparison (1) 23
2-10 Standard Switch and Distributed Switch Feature Comparison (2) 24
2-11 Load-Based NIC Teaming 25
2-12 Discovery Protocols 26
2-13 Configuring CDP or LLDP 27
2-14 About Port Binding 28
2-15 About the Traffic Filtering and Marking Policy 29
2-16 Creating Network Traffic Rules 30
2-17 Network Traffic Rule: Example 1 31
2-18 Network Traffic Rule: Example 2 32
2-19 Marking Network Traffic 33
2-20 Activity: Distributed Switch Features (1) 34
2-21 Activity: Distributed Switch Features (2) 35
2-22 About Health Check for Distributed Switches 36
2-23 Health Check Example 37
2-24 Enabling Health Check 38
2-25 Monitoring Health Check Results 39
2-26 Backing Up and Restoring Configurations 40
2-27 Exporting Distributed Switch Configurations 41
2-28 Restoring and Importing Distributed Switch Configurations 42

https://t.me/learningnets

2-29 Rollback and Recovery of the Management Network . . . 43
2-30 Rollback Details . . . 44
2-31 Recovery Through the DCUI . . . 45
2-32 Lab 1: Using vSphere Distributed Switches . . . 46
2-33 Review of Learner Objectives . . . 47
2-34 Lesson 2: Distributed Switch Feature: Network I/O Control . . . 48
2-35 Learner Objectives . . . 49
2-36 About Network I/O Control Version 3 . . . 50
2-37 Network I/O Control Versions 2 and 3 . . . 51
2-38 About the Bandwidth Allocation Model for System Traffic . . . 52
2-39 Configuring Bandwidth Allocations for System Traffic . . . 53
2-40 Reserving Bandwidth for System Traffic (1) . . . 54
2-41 Reserving Bandwidth for System Traffic (2) . . . 55
2-42 Bandwidth Aggregation for Network Resource Pools . . . 56
2-43 Creating Network Resource Pools . . . 57
2-44 Bandwidth Allocation for Virtual Machine Traffic . . . 58
2-45 Bandwidth Allocation for Individual Virtual Machines . . . 59
2-46 Bandwidth Admission Control in vSphere DRS (1) . . . 60
2-47 Bandwidth Admission Control in vSphere DRS (2) . . . 61
2-48 Bandwidth Admission Control in vSphere HA . . . 62
2-49 Activity: Network I/O Control (1) . . . 63
2-50 Activity: Network I/O Control (2) . . . 64
2-51 Review of Learner Objectives . . . 65
2-52 Lesson 3: Other Distributed Switch Features . . . 66
2-53 Learner Objectives . . . 67
2-54 About LACP . . . 68
2-55 About LAG Architecture . . . 69
2-56 Example of LACP Deployment with Two LAGs . . . 70
2-57 Creating LAGs . . . 71
2-58 Configuring LAGs Without Losing Connectivity (1) . . . 72
2-59 Configuring LAGs Without Losing Connectivity (2) . . . 73
2-60 Configuring LAGs Without Losing Connectivity (3) . . . 74
2-61 About NetFlow . . . 75
2-62 About Network Flows . . . 76
2-63 Network Flow Analysis . . . 77
2-64 Configuring NetFlow on Distributed Switches . . . 78
2-65 Enabling NetFlow on a Distributed Port Group . . . 79
2-66 About Port Mirroring . . . 80
2-67 Port Mirroring Session Types . . . 81
2-68 Port Mirroring Session Properties . . . 82
2-69 Source and Destination in a Port Mirroring Session . . . 83
2-70 Lab 2: Using Port Mirroring . . . 84
2-71 Review of Learner Objectives . . . 85
2-72 Key Points . . . 86

https://t.me/learningnets

Module 3 Storage Scalability 87
3-2 You Are Here 88
3-3 Importance 89
3-4 Module Lessons 90
3-5 Lesson 1: VMFS Datastores 91
3-6 Learner Objectives 92
3-7 VMFS Datastores (1) 93
3-8 VMFS Datastores (2) 94
3-9 VMFS5 and VMFS6 Datastores 95
3-10 Storage Device Formats 96
3-11 Storage Devices Supported by VMFS Datastores 97
3-12 Limitations of Using 4Kn Drives 98
3-13 Snapshot Formats for VMFS 99
3-14 Migrating Your Data from VMFS5 to VMFS6 100
3-15 Activity: VMFS6 Features (1) 101
3-16 Activity: VMFS6 Features (2) 102
3-17 Review of Learner Objectives 103
3-18 Lesson 2: Storage APIs 104
3-19 Learner Objectives 105
3-20 About vSphere Storage APIs - Array Integration 106
3-21 Hardware Acceleration APIs for Block Storage 107
3-22 Hardware Acceleration for NAS 108
3-23 Array Thin Provisioning APIs 109
3-24 About Space Reclamation (1) 110
3-25 About Space Reclamation (2) 111
3-26 Space Reclamation with VMFS 112
3-27 Performing Manual Space Reclamation on a VMFS5 Datastore 113
3-28 Enabling Automatic Space Reclamation on a VMFS6 Datastore 114
3-29 Configuring the Automatic Unmap Rate 115
3-30 Space Reclamation Support for Guest Operating Systems (1) 116
3-31 Space Reclamation Support for Guest Operating Systems (2) 117
3-32 Space Reclamation Support for Guest Operating Systems (3) 118
3-33 Activity: vSphere Storage APIs - Array Integration (1) 119
3-34 Activity: vSphere Storage APIs - Array Integration (2) 120
3-35 About vSphere API for Storage Awareness 121
3-36 Benefits of Storage Providers 122
3-37 Comparison of Storage Provider Versions 123
3-38 Storage Provider Categories 124
3-39 Registering a Storage Provider 125
3-40 About vSphere APIs for I/O Filtering 126
3-41 Types of I/O Filters 127
3-42 How I/O Filters Work (1) 128
3-43 How I/O Filters Work (2) 129

https://t.me/learningnets

3-44 I/O Filter Providers . . . 130
3-45 Displaying I/O Filter Provider Information . . . 131
3-46 Activity: I/O Filters (1) . . . 132
3-47 Activity: I/O Filters (2) . . . 133
3-48 Review of Learner Objectives . . . 134
3-49 Lesson 3: Storage Policy-Based Management . . . 135
3-50 Learner Objectives . . . 136
3-51 About Storage Policy-Based Management (1) . . . 137
3-52 About Storage Policy-Based Management (2) . . . 138
3-53 About VM Storage Policies . . . 139
3-54 Storage Policy Example (1) . . . 140
3-55 About Storage Policy Rules . . . 141
3-56 Storage Policy Example (2) . . . 142
3-57 About Storage Policy Components . . . 143
3-58 Storage Policy Component Example . . . 144
3-59 Activity: Storage Policy Rules (1) . . . 145
3-60 Activity: Storage Policy Rules (2) . . . 146
3-61 Creating and Managing Storage Policies: Overview . . . 147
3-62 Step 1: Configure Storage That Will Be Used for Storage Policies . . . 148
3-63 Step 2: Create Storage Policy Components . . . 149
3-64 Step 3: Create VM Storage Policies . . . 150
3-65 Example: Creating a Silver Tier Storage Policy (1) . . . 151
3-66 Example: Creating a Silver Tier Storage Policy (2) . . . 152
3-67 Example: Creating a Silver Tier Storage Policy (3) . . . 153
3-68 Example: Creating a Silver Tier Storage Policy (4) . . . 154
3-69 Step 4: Apply the Storage Policy to the VM . . . 155
3-70 Step 5: Check Compliance for the VM Storage Policy . . . 156
3-71 Lab 3: Policy-Based Storage . . . 157
3-72 Review of Learner Objectives . . . 158
3-73 Lesson 4: Storage Policies for vSAN and Virtual Volumes . . . 159
3-74 Learner Objectives . . . 160
3-75 About vSAN (1) . . . 161
3-76 About vSAN (2) . . . 162
3-77 vSAN Disk Architecture . . . 163
3-78 Object-Based Storage . . . 164
3-79 vSAN Storage Providers . . . 165
3-80 vSAN Storage Policies . . . 166
3-81 Creating vSAN Storage Policies . . . 167
3-82 vSAN Capabilities (1) . . . 168
3-83 vSAN Capabilities (2) . . . 169
3-84 About vSphere Virtual Volumes . . . 170
3-85 vSphere Virtual Volumes Architecture . . . 171
3-86 vSphere Virtual Volumes Storage Providers . . . 172

https://t.me/learningnets

3-87 Protocol Endpoints . . . 173
3-88 Storage Containers and vSphere Virtual Volumes Datastores . . . 174
3-89 Virtual Volumes . . . 175
3-90 Mapping Virtual Machine Files to Virtual Volumes . . . 176
3-91 vSphere Virtual Volumes Replication . . . 177
3-92 Virtual Volume Storage Policies . . . 178
3-93 Virtual Volume Storage Policy: Examples . . . 179
3-94 Creating Virtual Volumes Storage Policies (1) . . . 180
3-95 Creating Virtual Volumes Storage Policies (2) . . . 181
3-96 Activity: Displaying Storage Capabilities (1) . . . 182
3-97 Activity: Displaying Storage Capabilities (2) . . . 183
3-98 Lab 4: Creating vSAN Storage Policies . . . 184
3-99 Review of Learner Objectives . . . 185
3-100 Lesson 5: Storage I/O Control . . . 186
3-101 Learner Objectives . . . 187
3-102 About Storage I/O Control . . . 188
3-103 Enabling Storage I/O Control . . . 189
3-104 Monitoring Latency . . . 190
3-105 Automatic Threshold Detection . . . 191
3-106 Storage Policy Components for Storage I/O Control . . . 192
3-107 Configuring Storage I/O Control . . . 193
3-108 Storage I/O Control Requirements . . . 194
3-109 Activity: Storage I/O Control (1) . . . 195
3-110 Activity: Storage I/O Control (2) . . . 196
3-111 Review of Learner Objectives . . . 197
3-112 Lesson 6: Datastore Clusters and vSphere Storage DRS . . . 198
3-113 Learner Objectives . . . 199
3-114 About Datastore Clusters . . . 200
3-115 Datastore Cluster Requirements . . . 201
3-116 Host Cluster to Datastore Cluster Relationships . . . 202
3-117 vSphere Storage DRS Overview . . . 203
3-118 Initial Disk Placement . . . 204
3-119 vSphere Storage DRS Affinity Rules . . . 205
3-120 Storage Migration Recommendations . . . 206
3-121 Datastore Correlation Detector . . . 207
3-122 Configuring vSphere Storage DRS Runtime Settings . . . 208
3-123 vSphere Storage DRS Maintenance Mode . . . 209
3-124 Backups and vSphere Storage DRS . . . 210
3-125 Activity: vSphere Storage DRS (1) . . . 211
3-126 Activity: vSphere Storage DRS (2) . . . 212
3-127 vSphere Storage DRS and vSphere Technology Compatibility . . . 213
3-128 Interoperability with Site Recovery Manager . . . 214
3-129 Interoperability with vSphere Replication . . . 215

https://t.me/learningnets

3-130 Enabling vSphere Storage DRS to Recognize Storage Policies 216
3-131 vSphere Storage DRS and Array Feature Compatibility 217
3-132 Deduplication Awareness 218
3-133 Thin-Provisioned Datastore Awareness 219
3-134 Array-Based Autotiering Awareness 220
3-135 vSphere Storage DRS and Storage I/O Control 221
3-136 Lab 5: Managing Datastore Clusters 222
3-137 Review of Learner Objectives 223
3-138 Key Points 224

Module 4 Host and Management Scalability 225
4-2 You Are Here 226
4-3 Importance 227
4-4 Module Lessons 228
4-5 Lesson 1: Content Libraries 229
4-6 Learner Objectives 230
4-7 About Content Libraries 231
4-8 Content Library Benefits 232
4-9 Content Library Types 233
4-10 Storing Content in Content Libraries 234
4-11 Creating Content Libraries 235
4-12 Populating Content Libraries 236
4-13 Using the Content Library to Deploy Virtual Machines (1) 237
4-14 Using the Content Library to Deploy Virtual Machines (2) 238
4-15 Updating Content Library Items 239
4-16 Publishing Content Libraries for External Use 240
4-17 Content Library Subscriptions 241
4-18 Published and Subscribed Content Libraries 242
4-19 Synchronizing and Versioning Content Libraries 243
4-20 Content Library Maximums 244
4-21 Lab 6: Creating a Content Library 245
4-22 Review of Learner Objectives 246
4-23 Lesson 2: Host Profiles 247
4-24 Learner Objectives 248
4-25 Host Configuration Overview 249
4-26 About Host Profiles 250
4-27 Basic Workflow to Implement Host Profiles 251
4-28 Monitoring for Compliance 252
4-29 Applying Host Profiles 253
4-30 Customizing Hosts 254
4-31 Managing Host Customizations 255
4-32 Standardization Across Multiple vCenter Server Instances 256
4-33 Review of Learner Objectives 257

https://t.me/learningnets

4-34 Lesson 3: vSphere ESXi Image Builder . . . 258
4-35 Learner Objectives . . . 259
4-36 ESXi Images . . . 260
4-37 vSphere Installation Bundles . . . 261
4-38 ESXi Image Deployment . . . 262
4-39 vSphere ESXi Image Builder Overview . . . 263
4-40 vSphere ESXi Image Builder Architecture . . . 264
4-41 Building ESXi Images . . . 265
4-42 Step 1: Starting the Image Builder Service . . . 266
4-43 Step 2: Connecting to a Software Depot . . . 267
4-44 Step 3: Creating an Image Profile . . . 268
4-45 Step 4: Generating the New ESXi Image . . . 269
4-46 Review of Learner Objectives . . . 270
4-47 Lesson 4: vSphere Auto Deploy . . . 271
4-48 Learner Objectives . . . 272
4-49 About vSphere Auto Deploy . . . 273
4-50 vSphere Auto Deploy Modes . . . 274
4-51 vSphere Auto Deploy Mode: Stateless Caching . . . 275
4-52 vSphere Auto Deploy Mode: Stateful Installation . . . 276
4-53 Locations for Configuration and State Information . . . 277
4-54 vSphere Auto Deploy Architecture . . . 278
4-55 Rules Engine Basics . . . 279
4-56 PXE Boot Infrastructure Setup . . . 280
4-57 Initial Boot of an Autodeployed ESXi Host (1) . . . 281
4-58 Initial Boot of an Autodeployed ESXi Host (2) . . . 282
4-59 Initial Boot of an Autodeployed ESXi Host (3) . . . 283
4-60 Initial Boot of an Autodeployed ESXi Host (4) . . . 284
4-61 Initial Boot of an Autodeployed ESXi Host (5) . . . 285
4-62 Subsequent Boot of a Stateless ESXi Host (1) . . . 286
4-63 Subsequent Boot of a Stateless ESXi Host (2) . . . 287
4-64 Subsequent Boot of a Stateless ESXi Host (3) . . . 288
4-65 Subsequent Boot of a Stateless ESXi Host (4) . . . 289
4-66 Activity: vSphere Auto Deploy Modes (1) . . . 290
4-67 Activity: vSphere Auto Deploy Modes (2) . . . 291
4-68 Running Custom Scripts on Stateless Hosts . . . 292
4-69 Booting Many Stateless Hosts . . . 293
4-70 Stateless Caching Host Profile Configuration . . . 294
4-71 Subsequent Boot of a Stateless Caching Host . . . 295
4-72 Stateful Installation Host Profile Configuration . . . 296
4-73 Subsequent Boot of a Stateful Installation Host . . . 297
4-74 Configuring a vSphere Auto Deploy Environment . . . 298
4-75 Step 1: Preparing the DHCP Server . . . 299
4-76 Step 2: Starting the vSphere Auto Deploy Service . . . 300

https://t.me/learningnets

4-77 Step 3: Preparing the TFTP Server . . . 301
4-78 Step 4: Creating Deployment Rules . . . 302
4-79 Step 5: Activating Deployment Rules . . . 303
4-80 Managing the vSphere Auto Deploy Environment . . . 304
4-81 Using vSphere Auto Deploy with vSphere Update Manager . . . 305
4-82 Introduction to Lab 7: Using vSphere Auto Deploy . . . 306
4-83 Lab 7: Using vSphere Auto Deploy . . . 307
4-84 Review of Learner Objectives . . . 308
4-85 Key Points . . . 309

Module 5 CPU Optimization . . . 311
5-2 You Are Here . . . 312
5-3 Importance . . . 313
5-4 Module Lessons . . . 314
5-5 Lesson 1: CPU Virtualization Concepts . . . 315
5-6 Learner Objectives . . . 316
5-7 About the CPU Scheduler . . . 317
5-8 About Virtual Machine Worlds . . . 318
5-9 CPU Scheduler Features . . . 319
5-10 CPU Scheduler Support for SMP Virtual Machines . . . 320
5-11 Relaxed Co-Scheduling . . . 321
5-12 CPU Performance Factors . . . 322
5-13 Warning Sign: Ready Time . . . 323
5-14 Transitioning Between World States . . . 324
5-15 Activity: CPU Scheduler (1) . . . 325
5-16 Activity: CPU Scheduler (2) . . . 326
5-17 Review of Learner Objectives . . . 327
5-18 Lesson 2: NUMA Architecture in vSphere . . . 328
5-19 Learner Objectives . . . 329
5-20 NUMA-Aware ESXi . . . 330
5-21 How the ESXi NUMA Scheduler Works . . . 331
5-22 Terminology: vCPUs, pCPUs, Cores, and More . . . 332
5-23 About the NUMA Client . . . 333
5-24 Wide VM NUMA Support . . . 334
5-25 Virtual NUMA Topology . . . 335
5-26 Optimizing the Number of Cores per Socket for VMs (1) . . . 336
5-27 Optimizing the Number of Cores per Socket for VMs (2) . . . 337
5-28 Optimizing the Number of Cores per Socket for VMs (3) . . . 338
5-29 Activity: VM Cores Per Socket (1) . . . 339
5-30 Activity: VM Cores Per Socket (2) . . . 340
5-31 Memory Considerations with vNUMA (1) . . . 341
5-32 Memory Considerations with vNUMA (2) . . . 342
5-33 Memory Considerations with vNUMA (3) . . . 343

https://t.me/learningnets

5-34 Memory Considerations with vNUMA (4) . . . 344
5-35 vNUMA Considerations . . . 345
5-36 Review of Learner Objectives . . . 346
5-37 Lesson 3: Monitoring CPU Activity . . . 347
5-38 Learner Objectives . . . 348
5-39 About esxtop (1) . . . 349
5-40 About esxtop (2) . . . 350
5-41 Getting Help in esxtop . . . 351
5-42 Customizing Fields in esxtop . . . 352
5-43 CPU Key Performance Indicators for ESXi Hosts . . . 353
5-44 CPU Key Performance Indicators for VMs . . . 354
5-45 Using esxtop to Monitor ESXi Host CPU . . . 355
5-46 Using esxtop to View CPU Metrics per VM . . . 356
5-47 Using esxtop to View Single CPU Statistics . . . 357
5-48 Using esxtop to Monitor Virtual Machine vCPU . . . 358
5-49 Important Metrics to Monitor . . . 359
5-50 Example: Identifying CPU Overcommitment . . . 360
5-51 Host CPU Saturation . . . 361
5-52 Resolving Host CPU Saturation . . . 362
5-53 Activity: CPU Metrics (1) . . . 363
5-54 Activity: CPU Metrics (2) . . . 364
5-55 Guest CPU Saturation . . . 365
5-56 Using One vCPU in an SMP VM . . . 366
5-57 Low Guest CPU Usage . . . 367
5-58 Introduction to Lab 8: Monitoring CPU Performance . . . 368
5-59 Lab 8: Monitoring CPU Performance . . . 369
5-60 Review of Lab 8: Monitoring CPU Performance . . . 370
5-61 Review of Learner Objectives . . . 371
5-62 Key Points . . . 372

Module 6 Memory Optimization . . . 373

6-2 You Are Here . . . 374
6-3 Importance . . . 375
6-4 Module Lessons . . . 376
6-5 Lesson 1: Memory Virtualization Concepts . . . 377
6-6 Learner Objectives . . . 378
6-7 Virtual Memory Overview . . . 379
6-8 Application and Guest Operating System Memory Management . . . 380
6-9 Virtual Machine Memory Management . . . 381
6-10 Memory Reclamation . . . 382
6-11 Reason for Reclaiming Memory . . . 383
6-12 Guest Operating System Memory Terminology . . . 384
6-13 Virtual Machine Memory Reclamation Techniques . . . 385

https://t.me/learningnets

6-14 Transparent Page Sharing (1) . . . 386
6-15 Transparent Page Sharing (2) . . . 387
6-16 Impact of Hardware MMU on Transparent Page Sharing . . . 388
6-17 Impact of NUMA Architecture on Transparent Page Sharing . . . 389
6-18 Transparent Page Sharing and Intra-VM Memory Sharing . . . 390
6-19 Using Salting with Transparent Page Sharing (1) . . . 391
6-20 Using Salting with Transparent Page Sharing (2) . . . 392
6-21 Memory Ballooning in the Guest Operating System . . . 393
6-22 Reclaiming Memory with Ballooning . . . 394
6-23 Memory Compression . . . 395
6-24 Reclaiming Memory with Host Swapping . . . 396
6-25 Activity: Identifying Memory Reclamation Features (1) . . . 397
6-26 Activity: Identifying Memory Reclamation Features (2) . . . 398
6-27 Sliding Scale Mem.MemMinFreePct . . . 399
6-28 Criteria for Reclaiming Host Memory . . . 400
6-29 Transitioning from One Memory State to Another (1) . . . 401
6-30 Transitioning from One Memory State to Another (2) . . . 402
6-31 Review of Learner Objectives . . . 403
6-32 Lesson 2: Monitoring Memory Activity . . . 404
6-33 Learner Objectives . . . 405
6-34 Memory Usage Metrics in the Guest Operating System . . . 406
6-35 About Consumed Host Memory and Active Guest Memory . . . 407
6-36 Using esxtop to Monitor Memory Usage . . . 408
6-37 Host Ballooning Activity in esxtop (1) . . . 409
6-38 Host Ballooning Activity in esxtop (2) . . . 410
6-39 Host Memory Compression Activity in esxtop . . . 411
6-40 Monitoring Host Cache Swapping . . . 412
6-41 Host Swapping Activity in esxtop: Memory Screen . . . 413
6-42 Host Swapping Activity in esxtop: CPU Screen . . . 414
6-43 Causes of Active Host-Level Swapping . . . 415
6-44 Resolving Host-Level Swapping . . . 416
6-45 Activity: Identifying esxtop Fields (1) . . . 417
6-46 Activity: Identifying esxtop Fields (2) . . . 418
6-47 Reducing Memory Overcommitment . . . 419
6-48 Enabling Balloon Driver in Virtual Machines . . . 420
6-49 Reducing Virtual Machine Memory Reservation . . . 421
6-50 Dedicating Memory to Critical Virtual Machines . . . 422
6-51 Introduction to Lab 9: Monitoring Memory Performance . . . 423
6-52 Lab 9: Monitoring Memory Performance . . . 424
6-53 Review of Lab 9: Monitoring Memory Performance . . . 425
6-54 Review of Learner Objectives . . . 426
6-55 Key Points . . . 427

https://t.me/learningnets

Module 7 Storage Optimization 429
7-2 You Are Here 430
7-3 Importance 431
7-4 Module Lessons 432
7-5 Lesson 1: Storage Virtualization Concepts 433
7-6 Learner Objectives 434
7-7 Storage Performance Overview 435
7-8 Storage Protocol Performance 436
7-9 About SAN Configuration 437
7-10 About Storage Queues 438
7-11 Network Storage: iSCSI and NFS 439
7-12 Review of Learner Objectives 440
7-13 Lesson 2: Monitoring Storage Activity 441
7-14 Learner Objectives 442
7-15 Where Storage Problems Can Occur 443
7-16 Storage Key Performance Indicators 444
7-17 Monitoring Disk Throughput by Storage Adapter (1) 445
7-18 Monitoring Disk Throughput by Storage Adapter (2) 446
7-19 Correlating Storage Devices to Datastores (1) 447
7-20 Correlating Storage Devices to Datastores (2) 448
7-21 Monitoring Path Statistics for a Specific Storage Device (1) 449
7-22 Monitoring Path Statistics for a Specific Storage Device (2) 450
7-23 Monitoring Individual VMDK Statistics for a VM (1) 451
7-24 Monitoring Individual VMDK Statistics for a VM (2) 452
7-25 Disk Latency Metrics 453
7-26 Monitoring Disk Latency with esxtop 454
7-27 Monitoring Commands and Command Queuing 455
7-28 Example: Disk Latency and Queuing 456
7-29 Example: Bad Disk Throughput 457
7-30 Activity: Hardware Disk Latency (1) 458
7-31 Activity: Hardware Disk Latency (2) 459
7-32 Overloaded Storage and Command Aborts 460
7-33 Unexpected Increase in I/O Latency on Shared Storage 461
7-34 Storage Performance Best Practices 462
7-35 Introduction to Lab 10: Monitoring Storage Performance 463
7-36 Lab 10: Monitoring Storage Performance 464
7-37 Review of Lab 10: Monitoring Storage Performance 465
7-38 Review of Learner Objectives 466
7-39 Key Points 467

https://t.me/learningnets

Module 8 Network Optimization 469
8-2 You Are Here 470
8-3 Importance 471
8-4 Module Lessons 472
8-5 Lesson 1: Networking Virtualization Concepts 473
8-6 Learner Objectives 474
8-7 Network I/O Virtualization Overhead 475
8-8 VMXNET Network Adapters 477
8-9 Virtual Network Adapters 478
8-10 About VMXNET3 Optimization 480
8-11 Configuring a Single CPU Thread per vNIC 481
8-12 Configuring Multiple CPU Threads per vNIC 482
8-13 Virtual Switch and Physical NIC Optimization 483
8-14 About TCP Segmentation Offload 484
8-15 About Jumbo Frames 485
8-16 About SplitRx Mode 486
8-17 About SplitTx Mode 487
8-18 Activity: vSphere Networking Features (1) 488
8-19 Activity: vSphere Networking Features (2) 489
8-20 Review of Learner Objectives 490
8-21 Lesson 2: Monitoring Network I/O Activity 491
8-22 Learner Objectives 492
8-23 Monitoring for Performance in the Network I/O Stack 493
8-24 Network Capacity Metrics 494
8-25 esxtop Networking Statistics 495
8-26 esxtop Network Output 496
8-27 Dropped Network Packets 497
8-28 Dropped Packets in vNICs 498
8-29 Dropped Received Packets in Physical NICs 499
8-30 Dropped Transmit Packets in Physical NICs 500
8-31 Unexpected Increases in Data Transfer Rate 501
8-32 Networking Best Practices 502
8-33 Introduction to Lab 11: Monitoring Network Performance (1) 503
8-34 Introduction to Lab 11: Monitoring Network Performance (2) 504
8-35 Lab 11: Monitoring Network Performance 505
8-36 Review of Lab 11: Monitoring Network Performance 506
8-37 Review of Learner Objectives 507
8-38 Key Points 508

https://t.me/learningnets

Module 9 vCenter Server Performance Optimization 509
9-2 You Are Here 510
9-3 Importance 511
9-4 Learner Objectives 512
9-5 vCenter Server Services 513
9-6 Communication Between Management Services 515
9-7 vCenter Server Performance Considerations 517
9-8 Performance Considerations for Platform Services Controllers 518
9-9 Concurrent vCenter Server Operations 519
9-10 vCenter Server CPU Performance 520
9-11 vCenter Server Memory Performance 521
9-12 Monitoring CPU and Memory Usage with VAMI 522
9-13 Using vimtop to Monitor vCenter Server Appliance Resources 523
9-14 vimtop Commands and Columns 524
9-15 Monitoring CPU Usage with vimtop 525
9-16 Activity: CPU and Memory Usage (1) 526
9-17 Activity: CPU and Memory Usage (2) 527
9-18 Monitoring Memory Usage with vimtop 528
9-19 Viewing Service Memory Allocation with cloudvm-ram-size 529
9-20 Viewing Service Heap Size 530
9-21 Changing the Service Heap Size 531
9-22 vCenter Server Network Performance 532
9-23 Monitoring Network Activity with VAMI 533
9-24 Monitoring Network Activity with vimtop 534
9-25 vCenter Server Database Performance (1) 535
9-26 vCenter Server Database Performance (2) 536
9-27 Effects of Changing Statistics Level on Database Traffic 537
9-28 vCenter Server Appliance Disk Usage (1) 538
9-29 vCenter Server Appliance Disk Usage (2) 539
9-30 Monitoring Disk Usage with VAMI 540
9-31 Monitoring Disk Usage with the df Command 541
9-32 Monitoring Disk Activity with vimtop 542
9-33 Monitoring Database Activity with VAMI 543
9-34 Activity: Network Performance (1) 544
9-35 Activity: Network Performance (2) 545
9-36 Activity: vCenter Server Tools (1) 546
9-37 Activity: vCenter Server Tools (2) 547
9-38 Review of Learner Objectives 548
9-39 Key Points 549

https://t.me/learningnets

Module 10 vSphere Security 551
10-2 You Are Here 552
10-3 Importance 553
10-4 Module Lessons 554
10-5 Lesson 1: Configuring ESXi Host Access and Authentication 555
10-6 Learner Objectives 556
10-7 About ESXi Host Security Properties 557
10-8 Configuring the ESXi Firewall 558
10-9 Configuring ESXi Services 559
10-10 About Lockdown Mode (1) 560
10-11 About Lockdown Mode (2) 561
10-12 Normal Lockdown Mode Behavior 562
10-13 Strict Lockdown Mode Behavior 563
10-14 Integrating ESXi with Active Directory 564
10-15 VMware vSphere Authentication Proxy Service 565
10-16 Lab 12: Configuring Lockdown Mode 566
10-17 Review of Learner Objectives 567
10-18 Lesson 2: Securing vSphere 568
10-19 Learner Objectives 569
10-20 About Securing vSphere 570
10-21 About the vSphere Security Configuration Guide (1) 571
10-22 About the vSphere Security Configuration Guide (2) 572
10-23 About the vSphere Security Configuration Guide (3) 573
10-24 Controlling Access to vCenter Server Systems 574
10-25 vCenter Server Access Control 575
10-26 Securing vCenter Server with TLS 1.2 576
10-27 Securing vCenter Server Systems (1) 577
10-28 Securing vCenter Server Systems (2) 578
10-29 Securing vSphere Management Networks 579
10-30 Securing ESXi Hosts 580
10-31 UEFI Secure Boot for ESXi Hosts 581
10-32 Secure Boot Sequence for ESXi Hosts 582
10-33 TPM 2.0 on ESXi Hosts 583
10-34 About Remote Attestation 584
10-35 Remote Attestation Process 585
10-36 Requirements for Using TPM 2.0 586
10-37 Activity: ESXi Secure Boot (1) 587
10-38 Activity: ESXi Secure Boot (2) 588
10-39 vSphere Compliance with FIPS 140-2 (1) 589
10-40 vSphere Compliance with FIPS 140-2 (2) 590
10-41 FIPS 140-2 Architecture 591
10-42 Enabling and Disabling FIPS 140-2 Mode on ESXi Hosts 592
10-43 Enabling and Disabling FIPS 140-2 Mode on vCenter Server Appliance 593

https://t.me/learningnets

10-44 Virtual Machine Protection Overview . . . 594
10-45 Activity: Security Technologies (1) . . . 595
10-46 Activity: Security Technologies (2) . . . 596
10-47 Review of Learner Objectives . . . 597
10-48 Lesson 3: VMware Certificate Management . . . 598
10-49 Learner Objectives . . . 599
10-50 Reviewing Public Key Certificates . . . 600
10-51 About CAs . . . 601
10-52 vSphere Certificate Management . . . 602
10-53 About VMware CA . . . 603
10-54 About the VMware Endpoint Certificate Store . . . 604
10-55 vSphere Certificate Types . . . 605
10-56 vSphere CA Certificates . . . 606
10-57 Chain of Trust (1) . . . 607
10-58 Chain of Trust (2) . . . 608
10-59 Chain of Trust (3) . . . 609
10-60 Solution Endpoints Before vSphere 6.0 . . . 610
10-61 vSphere 6.x Reverse Proxy . . . 611
10-62 vSphere Solution User Certificates . . . 612
10-63 vSphere Machine Certificates . . . 613
10-64 vSphere Certificate Usage Summary . . . 614
10-65 Activity: VMware Endpoint Certificate Store (1) . . . 615
10-66 Activity: VMware Endpoint Certificate Store (2) . . . 616
10-67 VMware CA Modes . . . 617
10-68 VMware CA Mode: Root CA . . . 618
10-69 VMware CA Mode: Subordinate CA . . . 619
10-70 Certificate Replacement Options . . . 620
10-71 Certificate Manager . . . 621
10-72 Replacing the VMware CA Certificate in Root CA Mode . . . 622
10-73 Replacing the VMware CA Certificate with an Enterprise CA Certificate . . . 623
10-74 Replacing All VMware CA Certificates with Custom Certificates . . . 624
10-75 ESXi Certificate Replacement Options . . . 625
10-76 Installing the VMware CA Root Certificate in Your Browser . . . 626
10-77 VMware CA Availability . . . 627
10-78 Lab 13: Working with Certificates . . . 628
10-79 Review of Learner Objectives . . . 629
10-80 Lesson 4: Securing Virtual Machines . . . 630
10-81 Learner Objectives . . . 631
10-82 Business Use Case: Securing Virtual Machines . . . 632
10-83 About VM Encryption . . . 633
10-84 Advantages of VM Encryption . . . 634
10-85 Virtual Machine Encryption Architecture . . . 635
10-86 About the KMS . . . 636

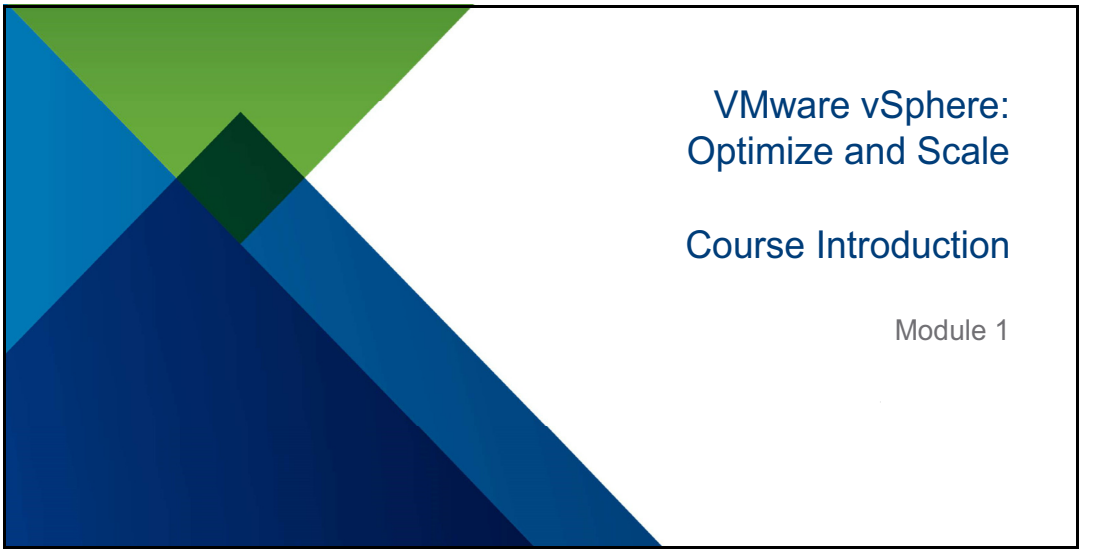
https://t.me/learningnets

10-87 Key Management Clusters . . . 637
10-88 Role of vCenter Server in Virtual Machine Encryption . . . 638
10-89 KMS and vCenter Server Communication . . . 639
10-90 KMIP Client Certificate . . . 640
10-91 Making an ESXi Host Cryptographically Safe . . . 641
10-92 VM Encryption Process Flow (1) . . . 642
10-93 VM Encryption Process Flow (2) . . . 643
10-94 Summary of Keys . . . 644
10-95 Activity: VM Keys (1) . . . 645
10-96 Activity: VM Keys (2) . . . 646
10-97 Managing Virtual Machine Encryption . . . 647
10-98 vCenter Server Role: No Cryptography Administrator . . . 648
10-99 VM Encryption Prerequisites . . . 649
10-100 Encrypting New Virtual Machines . . . 650
10-101 Encrypting Existing Virtual Machines . . . 651
10-102 Backing Up Encrypted Virtual Machines (1) . . . 652
10-103 Backing Up Encrypted Virtual Machines (2) . . . 653
10-104 Encrypted VM Operations . . . 654
10-105 Unlocking an Encrypted VM (1) . . . 655
10-106 Unlocking an Encrypted VM (2) . . . 656
10-107 Activity: Cryptographic Privileges (1) . . . 657
10-108 Activity: Cryptographic Privileges (2) . . . 658
10-109 About Encrypted Core Dumps . . . 659
10-110 How Core Dumps Are Encrypted . . . 660
10-111 Providing Passwords for Encrypted Core Dumps . . . 661
10-112 About Encrypted vSphere vMotion . . . 662
10-113 Configuring Encrypted vSphere vMotion . . . 663
10-114 Encrypted vSphere vMotion Requirements . . . 664
10-115 Virtual Machine Protection with UEFI Secure Boot . . . 665
10-116 Using vTPM in a VM (1) . . . 666
10-117 Using vTPM in a VM (2) . . . 667
10-118 Prerequisites to Add vTPM . . . 668
10-119 Virtualization-Based Security for Windows Guest Operating Systems . . . 669
10-120 VBS Requirements and Limitations . . . 670
10-121 Enabling VBS . . . 671
10-122 What's Next . . . 672
10-123 Lab 14: Virtual Machine Encryption . . . 673
10-124 Review of Learner Objectives . . . 674
10-125 Key Points . . . 675

https://t.me/learningnets

MODULE 1

Course Introduction



https://t.me/learningnets

1-2 Importance

Administrators must have the advanced skills for configuring and maintaining a highly available and scalable VMware vSphere® environment.

You must know how to optimize, scale, and secure VMware ESXi™ hosts and VMware vCenter Server® instances in your environment.

https://t.me/learningnets

1-3 Learner Objectives

By the end of this course, you should be able to meet the following objectives:

- Configure and manage vSphere networking and storage for a large and sophisticated enterprise
- Use VMware vSphere® Client™, VMware vSphere® Web Client, and VMware vSphere® ESXi™ Shell to manage vSphere
- Create a content library for deploying virtual machines
- Use VMware vSphere® Auto Deploy™ and host profiles to provision ESXi hosts and manage ESXi host compliance
- Monitor and analyze key performance indicators for compute, storage, and networking resources for ESXi hosts
- Optimize the performance of ESXi and vCenter Server
- Harden the vSphere environment against security threats
- Encrypt virtual machines for additional security

https://t.me/learningnets

1-4 You Are Here

1. **Course Introduction**
2. Network Scalability
3. Storage Scalability
4. Host and Management Scalability
5. CPU Optimization
6. Memory Optimization
7. Storage Optimization
8. Network Optimization
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

1-5 Typographical Conventions

The following typographical conventions are used in this course.

`Monospace`	Filenames, folder names, path names, command names, and code: Navigate to the `VMS` folder.
`Monospace Bold`	What the user inputs: Enter **`ipconfig /release`**.
Boldface	User interface controls: Click the **Configuration** tab.
Italic	Book titles and placeholder variables: • *vSphere Virtual Machine Administration* • *ESXi_host_name*

https://t.me/learningnets

1-6 References (1)

Title	Location
VMware ESXi Installation and Setup	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf
vCenter Server Installation and Setup	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-vcenter-server-67-installation-guide.pdf
vCenter Server and Host Management	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-host-management-guide.pdf
VMware Configuration Maximums	https://configmax.vmware.com/home
VMware Compatibility Guide	http://www.vmware.com/resources/compatibility
vSphere Central	https://vspherecentral.vmware.com
Storage and Availability Technical Documents	https://storagehub.vmware.com

https://t.me/learningnets

1-7 References (2)

Title	Location
vSphere Networking	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf
vSphere Storage	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-storage-guide.pdf.pdf
vSphere Resource Management	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-resource-management-guide.pdf
vSphere Host Profiles	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-host-profiles-guide.pdf
vSphere Security	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-security-guide.pdf
vSphere Monitoring and Performance	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-monitoring-performance-guide.pdf

https://t.me/learningnets

1-8 VMware Online Resources

VMware vSphere Blog: http://blogs.vmware.com/vsphere

VMware Communities: http://communities.vmware.com

VMware Support: http://www.vmware.com/support

VMware Education: http://www.vmware.com/education

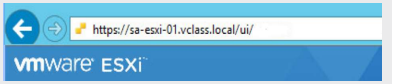
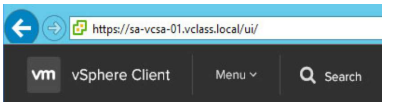
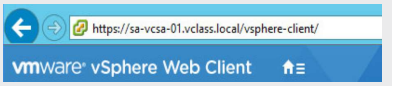
VMware Certification: http://www.vmware.com/certification

VMware Education and Certification Blog: http://blogs.vmware.com/education

https://t.me/learningnets

1-9 Review of vSphere User Interfaces

vSphere provides the following graphical user interfaces.

User Interface	Description	Display
VMware Host Client™	• HTML5-based client • Connects directly to an ESXi host • https://*ESXi host*/ui	
vSphere Client	• HTML5-based client • Connects to vCenter Server • Preferred client • https://*vCenter Server*/ui	
vSphere Web Client	• Flash-based client • Connects to vCenter Server • https://*vCenter Server*/vsphere-client • Will no longer be available starting with the next numbered release of vSphere	

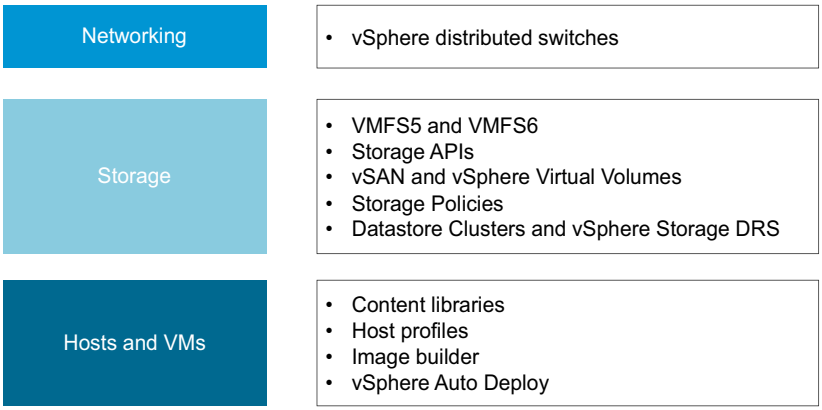
VMware vSphere® Web Client should still be used for the few functions that are not yet implemented in VMware vSphere® Client™.

For information on functionality support in vSphere Client, see https://docs.vmware.com/en/VMware-vSphere/6.5/rn/vsphere-client-65-html5-functionality-support.html.

https://t.me/learningnets

1-10 About Scalability

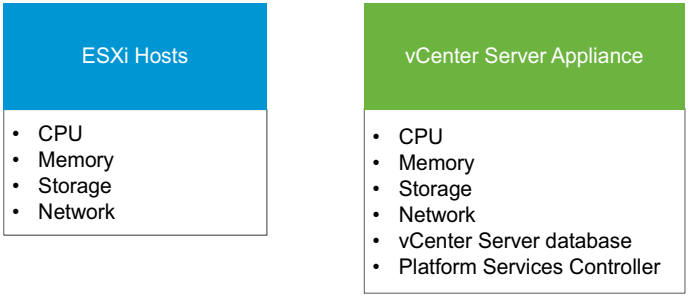
vSphere consists of technology that builds a foundation for a truly scalable infrastructure.



https://t.me/learningnets

1-11 About Optimization

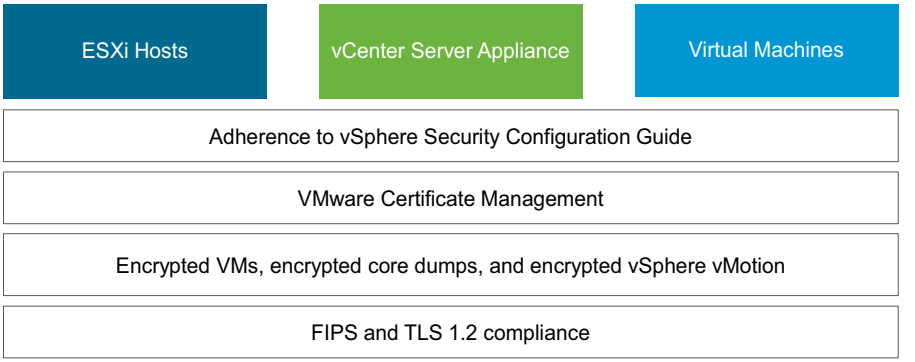
Identify areas in your vSphere environment where resource use can be optimized.



https://t.me/learningnets

1-12 About Security

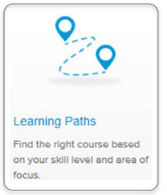
Your security strategy should include securing the different layers in the vSphere infrastructure.



https://t.me/learningnets

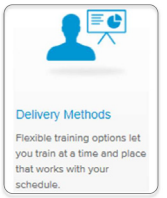
1-13 VMware Education Overview

VMware Education provides training and certification programs to grow your skills with the VMware technology.



Learning paths help you find the course you need based on product, your role, and your level of experience.

For example, this course is part of the Data Center Virtualization Infrastructure learning path.



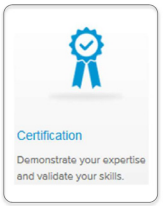
Examples of delivery methods:

- On-Demand
  - Self-paced learning solution that combines modular training with hands-on practice labs, and is an alternative to traditional classroom training.
- Lab Connect
  - Self-paced, technical lab environment designed to enhance your learning experience.
  - These cloud-based, on-demand, labs let you practice skills learned during instructor-led training, and get extra hands-on practice before applying your new skills in an operational environment.

https://t.me/learningnets

1-14 VMware Certification Overview

VMware Certification sets the standard for IT professionals and validates critical skills with VMware technology.





This course can be used to fulfill the training requirement for the VMware Certified Professional 6.5 – Data Center Virtualization (VCP6.5-DCV) certification.



VMware Learning Zone provides a cloud-based video training library that can help you prepare efficiently and give you confidence to pass the certification exam.

https://t.me/learningnets

Network Scalability



https://t.me/learningnets

2-2 You Are Here

1. Course Introduction
2. **Network Scalability**
3. Storage Scalability
4. Host and Management Scalability
5. CPU Optimization
6. Memory Optimization
7. Storage Optimization
8. Network Optimization
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

2-3 Importance

As you scale your vSphere environment, you must be aware of the VMware vSphere® Distributed Switch™ features and functions that help you manage networking in your environment.

https://t.me/learningnets

2-4 Module Lessons

Lesson 1:	Creating vSphere Distributed Switches
Lesson 2:	Distributed Switch Feature: Network I/O Control
Lesson 3:	Other Distributed Switch Features

https://t.me/learningnets



https://t.me/learningnets

2-6 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Describe the benefits and features of distributed switches:
  - Discovery protocols
  - Port binding
  - Traffic filtering and marking policy
  - Automatic rollback and recovery of networking configurations
- Create a distributed switch
- Perform a distributed switch health check
- Back up and restore a distributed switch configuration

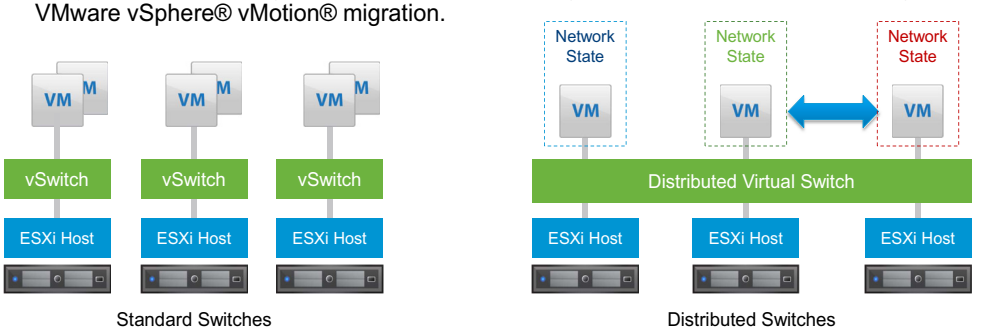
https://t.me/learningnets

2-7 About Distributed Switches

A distributed switch functions as a single virtual switch across all associated hosts.

Distributed switches have several benefits over standard switches:

- They simplify data center administration.
- They enable networking statistics and policies to migrate with virtual machines during a VMware vSphere® vMotion® migration.



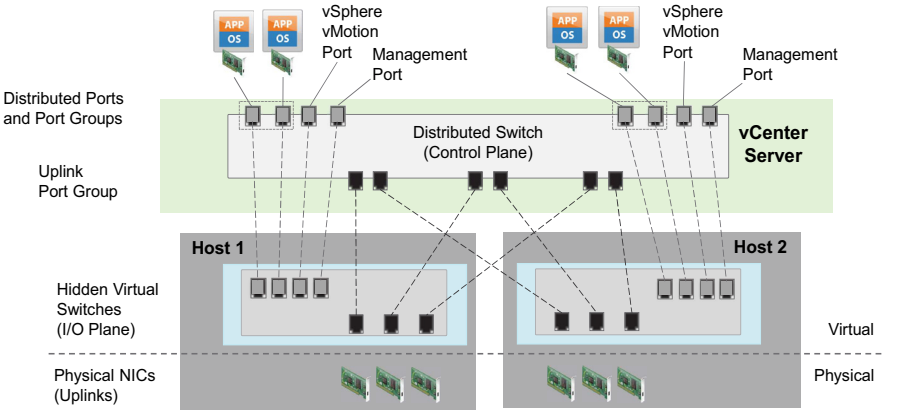
Having the network configuration at the data center level (VMware vSphere® Distributed Switch™), not at the host level (standard switch), offers the following advantages:

- Data center setup and administration are simplified by centralizing network configuration. For example, adding a host to a cluster and making it compatible with VMware vSphere® vMotion® is much easier.
- Distributed ports migrate with their clients. For example, when you migrate a virtual machine with vSphere vMotion, the distributed port statistics and policies move with the virtual machine, thus simplifying debugging and troubleshooting.

https://t.me/learningnets

2-8 Distributed Switch Architecture

A distributed switch moves network management components to the data center level.



Managed by VMware vCenter Server®, a distributed switch is a logical entity that is used to create and maintain a consistent virtual networking configuration throughout all of your VMware ESXi™ hosts.

The distributed switch architecture consists of the control plane and the I/O plane.

The control plane resides in vCenter Server. The control plane is responsible for configuring distributed switches, distributed port groups, distributed ports, uplinks, NIC teaming, and so on. The control plane also coordinates the migration of the ports and is responsible for the switch configuration.

The I/O plane is implemented as a hidden virtual switch in the VMkernel of each ESXi host. The I/O plane manages the I/O hardware on the host and is responsible for forwarding packets. vCenter Server oversees the creation of these hidden virtual switches.

Each distributed switch includes distributed ports. You can connect any networking entity, such as a virtual machine or a VMkernel interface, to a distributed port. vCenter Server stores the state of distributed ports in the vCenter Server database.

A distributed port group enables you to logically group distributed ports to simplify configuration. A distributed port group specifies port configuration options for each member port on a distributed switch. Ports can also exist without port groups.

Uplinks are abstractions of vmnics from multiple hosts to a single distributed switch. An uplink is to a distributed switch what a vmnic is to a standard switch. Two virtual machines on different hosts can communicate with each other only if both virtual machines have uplinks in the same broadcast domain.

https://t.me/learningnets

2-9 Standard Switch and Distributed Switch Feature Comparison (1)

The table provides a summary of capabilities present in standard switches and distributed switches.

Feature	Standard Switch	Distributed Switch
Layer 2 switch	✓	✓
VLAN segmentation	✓	✓
IPv6 support	✓	✓
802.1Q tagging	✓	✓
NIC teaming	✓	✓
Outbound traffic shaping	✓	✓
Inbound traffic shaping		✓
Configuration backup and restore		✓
Private VLANs		✓
Link Aggregation Control Protocol		✓

https://t.me/learningnets

2-10 Standard Switch and Distributed Switch Feature Comparison (2)

The table provides a summary of capabilities present in distributed switches.

Feature	Standard Switch	Distributed Switch
Data center level management		✓
vSphere vMotion migration of network state		✓
Network I/O Control		✓
Per-port policy settings		✓
Port state monitoring		✓
NetFlow		✓
Port mirroring		✓

During a vSphere vMotion migration, a distributed switch tracks the virtual networking state (for example, counters and port statistics) as the virtual machine moves from host to host. The tracking provides a consistent view of a virtual network interface, regardless of the virtual machine location or vSphere vMotion migration history. Tracking simplifies network monitoring and troubleshooting activities where vSphere vMotion is used to migrate virtual machines between hosts.

https://t.me/learningnets

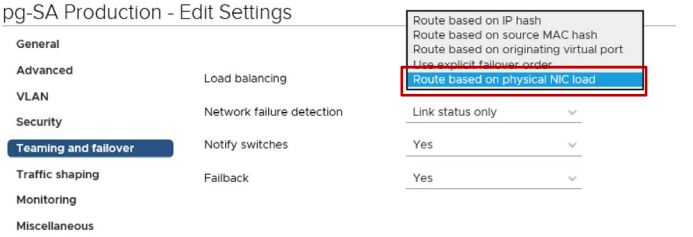
2-11 Load-Based NIC Teaming

Load-based NIC teaming ensures that physical NIC capacity in a NIC team is optimized.

Load-based NIC teaming moves I/O flows among uplinks: A flow is moved only when the mean send or receive utilization on an uplink exceeds 75 percent of the capacity over a 30-second period.

Load-based NIC teaming is supported only on distributed switches.

To use load-based NIC teaming, select **Route based on physical NIC load**.



Load-based NIC teaming checks the real load of the uplinks and takes steps to reduce the load on overloaded uplinks. No changes on the physical switch are required.

The distributed switch calculates uplinks for virtual machines by taking the virtual machine port ID and the number of uplinks in the NIC team. The distributed switch tests the uplinks every 30 seconds. If the load of an uplink exceeds 75 percent of usage, the port ID of the virtual machine with the highest I/O is moved to a different uplink.

Load-based NIC teaming is not the default teaming policy, and so you must configure the policy to be able to use it.

https://t.me/learningnets

2-12 Discovery Protocols

Switch discovery protocols help network administrators determine the capabilities of a network device.

vSphere supports the following discovery protocols:

- Cisco Discovery Protocol (CDP)
- Link Layer Discovery Protocol (LLDP)

CDP is available for vSphere standard switches and distributed switches connected to Cisco physical switches.

LLDP is a vendor-neutral protocol that is available only for distributed switches.

You can use CDP and LLDP to gather configuration and connection information about a physical or virtual switch. Such information might help in troubleshooting network problems.

Discovery protocols are available in vSphere virtual switches.

Cisco Discovery Protocol (CDP) was developed by Cisco Systems to broadcast information at network layer 2 about connected devices. CDP has been supported in vSphere since version 4.0.

Link Layer Display Protocol (LLDP) was introduced in vSphere 5.0. LLDP supports the standards-based IEEE 802.1AB discovery protocol.

CDP or LLDP is used by network devices for advertising their identity, capabilities, and neighbors on a network.

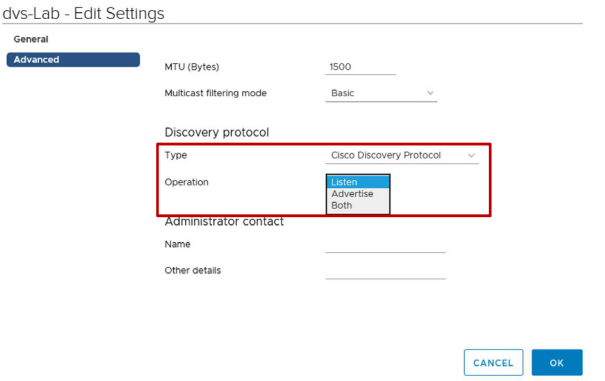
https://t.me/learningnets

2-13 Configuring CDP or LLDP

With CDP or LLDP enabled, the virtual switch can be configured for different modes of operation:

- Listen: Information is received from the physical switches.
- Advertise: Information is sent to the physical switches.
- Both: Information is both sent to and received from the physical switches.

CDP is enabled in listen mode by default.



CDP and LLDP enable vSphere Client to identify properties of a physical switch, such as switch name, port number, and port speed/duplex settings. You can also configure CDP or LLDP so that information about physical adapters and ESXi host names is passed to the CDP- or LLDP-compatible switches.

You can configure the discovery protocol to use one of the following modes of operation:

- Listen (default): ESXi detects and displays information about the associated physical switch port, but information about the virtual switch is not available to the physical switch administrator.
- Advertise: ESXi provides information about the virtual switch to the physical switch administrator but does not detect and display information about the physical switch.
- Both: ESXi detects and displays information about the associated physical switch and provides information about the virtual switch to the physical switch administrator.

You can use the `esxcli` command to enable CDP on a standard switch.

https://t.me/learningnets

2-14 About Port Binding

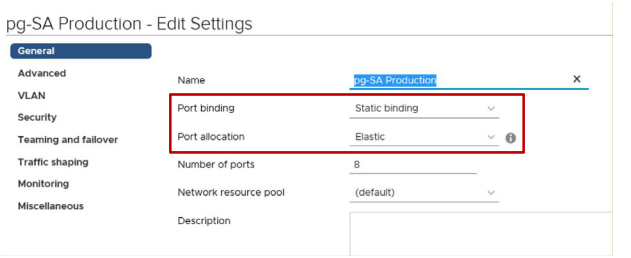
Port binding determines when and how a virtual machine virtual NIC is assigned to a virtual switch port.

The following port binding options are available:

- Static binding: This is the default setting.
- Ephemeral: No binding occurs.

For static binding, these port allocation options are available:

- Elastic (default): When all ports are assigned, a new set of eight ports is created.
- Fixed: No additional ports are created when all ports are assigned.



Port binding is configured at the distributed port group level.

When you connect a virtual machine to a port group configured with static binding, a port is immediately assigned and reserved for the virtual machine, guaranteeing connectivity at all times. The port is disconnected only when the virtual machine is removed from the port group. Static binding is recommended for general use.

If static binding is selected, the default number of ports is set to eight. Elastic is the default port allocation setting.

With ephemeral binding, a port is created and assigned to a virtual machine when the virtual machine is powered on and its NIC is in a connected state. The port is deleted when the virtual machine is powered off or the virtual machine NIC is disconnected.

Ephemeral port assignments can be made through ESXi as well as vCenter Server, giving you the flexibility to manage virtual machine connections through the host when vCenter Server is down. Although only an ephemeral binding allows you to modify virtual machine network connections when vCenter Server is down, network traffic is unaffected by a vCenter Server failure, regardless of port binding type.

Ephemeral port groups should be used only for recovery purposes when you want to provision ports directly on an ESXi host, bypassing vCenter Server, but not for any other case.

https://t.me/learningnets

2-15 About the Traffic Filtering and Marking Policy

vSphere distributed switches provide a traffic filtering and marking policy. This policy enables you to protect your virtual network from unwanted traffic and security attacks, and to manage network traffic to meet specified SLA targets.

The traffic filtering and marking policy has the following features:

- It can permit or deny specific types of traffic.
- It can apply a QoS tag to mark a certain type of traffic.
- It is equivalent to the access control list feature available on physical switches.

The traffic filtering and marking policy consists of one or more network traffic rules, defined at the distributed port group or uplink port group level.

Use the traffic filtering and marking policy to create a set of rules for security and QoS tagging of packets flowing through distributed switch ports.

The distributed switch applies rules on traffic at different places in the data stream. The distributed switch can apply network traffic rules on the data path between the virtual machine network adapter and the distributed port. The distributed switch can also apply network traffic rules between the uplink port and the physical network adapter.

The traffic filtering and marking policy is supported only on distributed switches.

https://t.me/learningnets

2-16 Creating Network Traffic Rules

You can define network traffic rules for processing traffic related to virtual machines or to physical adapters.

A network traffic rule consists of the following elements:

- Action
- Traffic direction
- Traffic qualifiers:
  - IP
  - MAC
  - System traffic



In general, a network traffic rule consists of a qualifier for traffic and of an action for restricting or prioritizing the matching traffic.

In a network traffic rule, the Allow action allows traffic to pass through the distributed switch port or uplink port, and the Drop action blocks the traffic. The Tag action marks (or tags) traffic passing through the distributed switch port or uplink port.

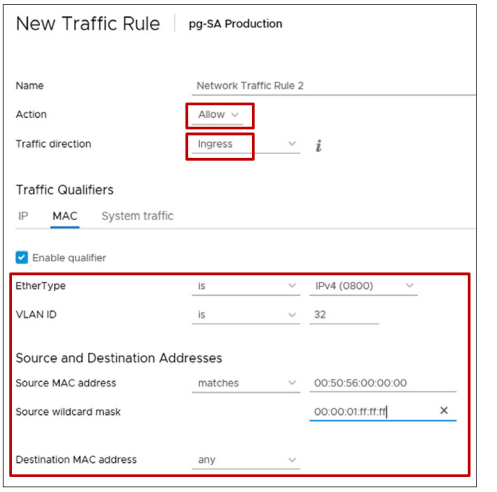
Traffic direction is with respect to the distributed switch. The direction can be ingress (traffic entering the distributed switch), egress (traffic leaving the distributed switch), or both. The direction also influences how you identify the traffic source and destination.

A qualifier represents a set of matching criteria related to a networking layer. You can match traffic based on system traffic type, layer 2 traffic properties, or layer 3 traffic properties. You can use the qualifier for a specific networking layer, or you can combine qualifiers to match packets more precisely.

https://t.me/learningnets

2-17 Network Traffic Rule: Example 1

In this example, this rule allows incoming traffic from systems on VLAN 32 that have a MAC address starting with 00:50:56.



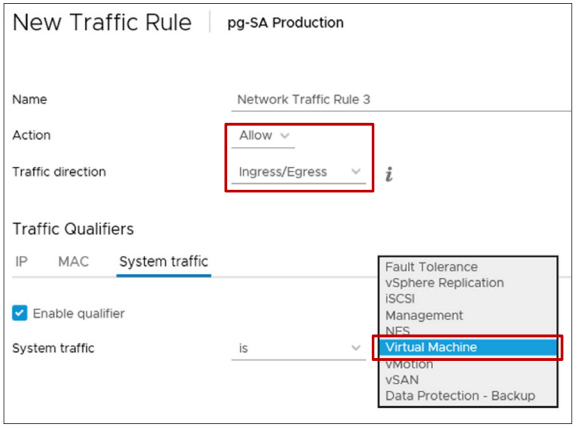
Use a MAC traffic qualifier to define matching criteria for the layer 2 (data link layer) properties of packets such as MAC address, VLAN ID, and a protocol that consumes the frame payload (IPv4, IPv6, or Address Resolution Protocol).

Locating traffic with a VLAN ID on a distributed port group works with Virtual Guest Tagging. To match traffic to VLAN ID if Virtual Switch Tagging is active, use a rule on an uplink port group or uplink port.

https://t.me/learningnets

2-18 Network Traffic Rule: Example 2

In this example, this rule allows incoming and outgoing virtual machine traffic.



Use the system traffic qualifier to match packets to the type of virtual infrastructure data that is flowing through the ports of the distributed port group.

In this example, the rule enforces that the traffic through the pg-SA Production port group must be VM traffic. The distributed switch drops any other types of traffic that are sent to this port group.

You can select the type of traffic through the ports of the group that carries system data, such as traffic for management from vCenter Server, iSCSI storage, and vSphere vMotion.

By using qualifiers for the system data type, both layer 2 and layer 3 packet attributes set the properties that packets must have to match the rule.

https://t.me/learningnets

2-19 Marking Network Traffic

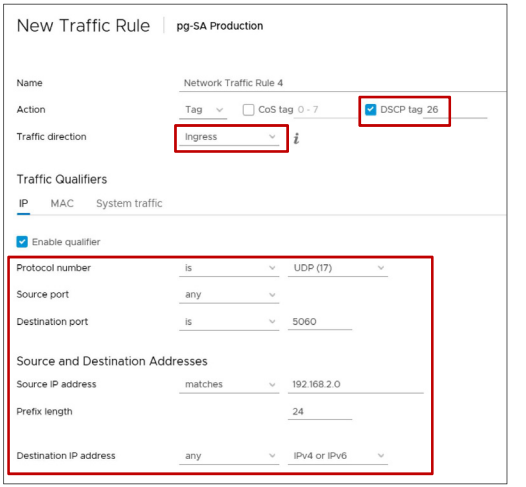
You can assign priority tags to traffic that has higher networking requirements for bandwidth, low latency, and so on.

You can mark the traffic with a CoS tag in layer 2 or a DSCP tag in layer 3.

Marking traffic has the following benefits:

- Important traffic can be tagged so that these packets are not dropped by the physical network during congestion.
- End-to-end QoS and SLA requirements can be provided.

In this example, this rule marks incoming SIP UDP packets from subnet 192.168.2.0/24.



Marking, or priority tagging, is a mechanism to mark traffic that has higher QoS demands. Priority tagging enables the network to recognize different classes of traffic. The network devices can handle the traffic from each class according to its priority and requirements.

Similar to traffic filtering, marking also requires classifying the traffic first, based on these qualifiers: system, MAC, and IP. After you define your traffic qualifiers, you can decide how to mark your traffic.

You can prioritize traffic by using a CoS priority tag, a DSCP tag, or both.

Traffic can be marked with a CoS priority tag in the layer 2 packet header. Accepted values are 0 through 7.

Traffic can be marked with a DSCP tag in the layer 3 packet header. Accepted values are 0 through 63.

You can assign priority tags to traffic, such as VoIP and streaming video, that has higher networking requirements for bandwidth, low latency, and so on.

This example marks VoIP traffic. VoIP flows have special requirements for QoS in terms of low loss and delay. The traffic related to Session Initiation Protocol for VoIP usually has a DSCP tag equal to 26.

https://t.me/learningnets

2-20 Activity: Distributed Switch Features (1)

With the Route Based on Physical NIC Load teaming policy, uplinks in the port group are tested every 30 seconds. If the load on the uplink exceeds 75 percent of usage, then the VM with the highest I/O is moved to a different uplink.

- ❑ True
- ❑ False

https://t.me/learningnets

2-21 Activity: Distributed Switch Features (2)

With the Route Based on Physical NIC Load teaming policy, uplinks in the port group are tested every 30 seconds. If the load on the uplink exceeds 75 percent of usage, then the VM with the highest I/O is moved to a different uplink.

- ✓ True
- ❑ False

Load-based NIC teaming is a distributed switch feature that is aware of the load on uplinks and takes care to reduce the load if needed.

https://t.me/learningnets

2-22 About Health Check for Distributed Switches

The health check support helps you identify and troubleshoot configuration errors in a vSphere distributed switch.

Health check regularly examines certain settings on the distributed and physical switches to identify common configuration errors:

- Mismatched VLAN trunks between the distributed switch and physical switch
- Mismatched MTU settings between the distributed switch, physical adapter, and physical switch ports
- Mismatched virtual switch teaming policies for the physical switch port-channel settings

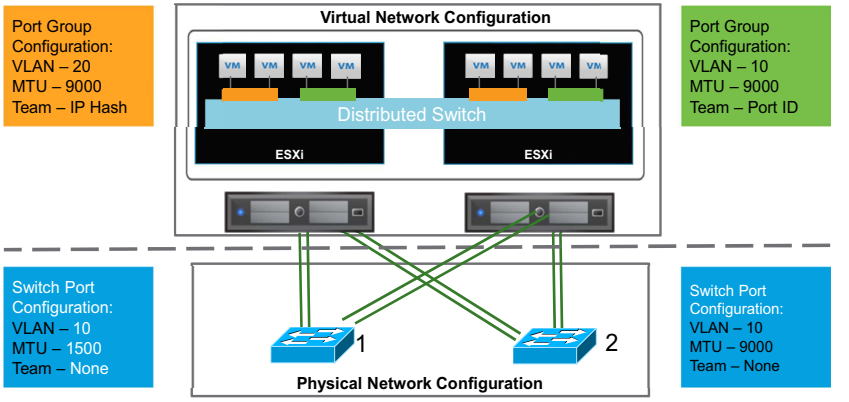
Health check is a feature that detects certain inconsistencies between the physical and the virtual networks. Key parameters, such as VLAN tags, MTUs, and NIC teaming configuration, must be configured consistently on the physical and the virtual switches. An inconsistent configuration can lead to network connectivity problems.

Health check searches for configuration inconsistencies and reports them to the administrator. The default check interval is one minute.

https://t.me/learningnets

2-23 Health Check Example

In this example, two ESXi hosts use one distributed switch with two distributed port groups.



In addition to the distributed switch, two physical switches with their switch port configurations are shown.

Comparing the green virtual port group with the settings on physical switch 2, the VLAN ID and the MTU settings are identical. The virtual port group is configured with the port ID teaming selection. But because the configuration is internal to the virtual switch and requires no configuration on the physical switch, the virtual and the physical settings are consistent and should therefore experience no connectivity issues.

Next, compare the yellow virtual port group to the settings on physical switch 1. The VLAN IDs and the MTU settings are different, and the yellow teaming setting is set to IP hash. For IP hash to operate properly, EtherChannel or Link Aggregation Control Protocol (LACP) must be configured on the physical switch.

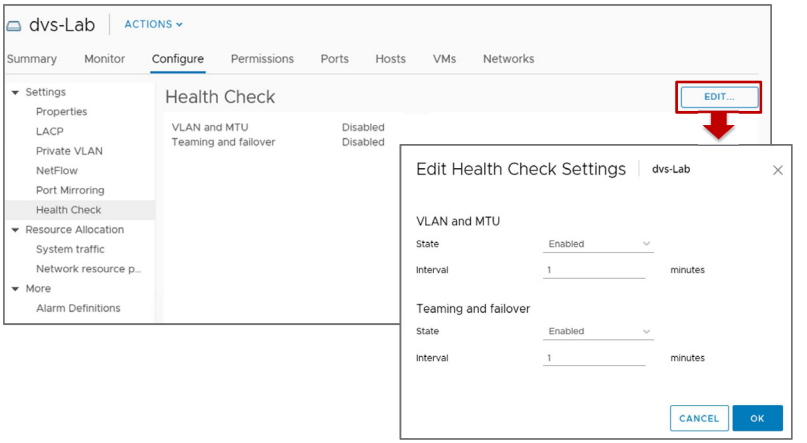
Health check can detect and report the configuration differences between the port group and the switch port configuration by using layer 2 Ethernet packets. At one-minute intervals (by default), request and acknowledge packets are sent back and forth between the virtual interface and the switch. When packets are dropped, a configuration warning appears in vSphere Client.

https://t.me/learningnets

2-24 Enabling Health Check

Health check can be performed on the following network components:

- VLAN and MTU
- Teaming and failover



Health check has the following requirements:

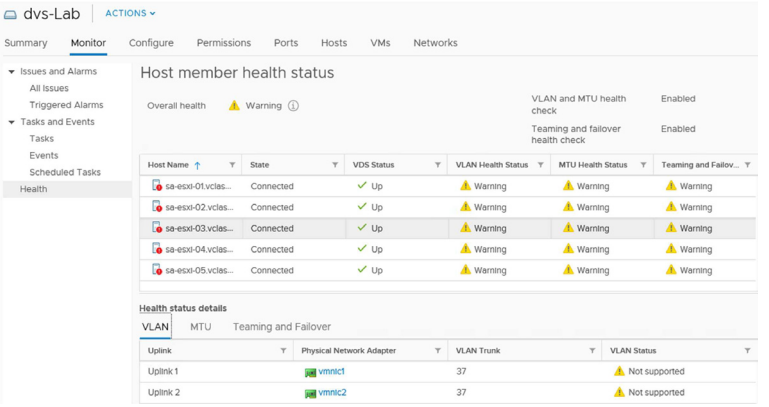
- For VLAN and MTU checks, at least two physical uplinks must be connected to the distributed virtual switch.
- For the teaming policy check, at least two active uplinks must be in the team, and at least two ESXi hosts must be connected to the virtual distributed switch.

See *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

https://t.me/learningnets

2-25 Monitoring Health Check Results

After health check runs for a few minutes, you can monitor the results in the Health pane in vSphere Client.



Health status is organized by VLAN, MTU, and Teaming and Failover.

https://t.me/learningnets

2-26 Backing Up and Restoring Configurations

You can back up and restore the configuration of your distributed switch, distributed port groups, and uplink port groups for deployment, rollback, and sharing purposes.

The following operations are supported:

- Back up the configuration on disk
- Restore the switch and port group configuration from a backup
- Create a new switch or port group from the backup
- Revert to a previous port group configuration after changes are made

You perform these operations by using the export, import, and restore functions available for distributed switches.

You have other available options if the switch configuration is lost, for example, in the case of vCenter Server database corruption, or if the virtual switch or port group settings are recently misconfigured. Ways to restore the virtual switch include restoring the database completely or rebuilding the switch. Although both of these measures restore the switch, they can be time-consuming.

https://t.me/learningnets

2-27 Exporting Distributed Switch Configurations

You can export distributed switch and distributed port group configurations to a file.

Exporting enables you to perform the following tasks:

- Make a backup of your distributed switch configuration
- Create a template of a distributed switch configuration
- Create a revision control system for your distributed switch configuration

You export the distributed switch and distributed port group configuration to a file on the system that is running vSphere Client. The file preserves valid network configurations, enabling distribution of these configurations to other deployments.

Making periodic backups with the export function enables you to preserve your distributed switch configuration in case of a failure, such as the corruption of the vCenter Server database.

You can use the template created from the export function to create similar distributed switch configurations on other vCenter Server systems.

You can keep revisions by saving the distributed switch configuration after each change. By keeping revisions, you can restore the current configuration to an older configuration if necessary.

You can automate this task with the following VMware PowerCLI™ cmdlets:

- `Export-VDSwitch`: Exports the configuration of a specified vSphere distributed switch to a `.zip` file.
- `Export-VDPortGroup`: Exports the configuration of a specified distributed port group to a specified `.zip` file.

https://t.me/learningnets

2-28 Restoring and Importing Distributed Switch Configurations

After you export a distributed switch configuration, you can restore or import a configuration:

- Restoring resets the configuration of an existing distributed switch from an exported configuration file.
- Importing creates a new distributed switch from an exported configuration file.

You can use the restore function to reset a distributed switch configuration that has become corrupted.

You can use the import function to create a new distributed switch, for example, on a different vCenter Server system.

With the restore and import functions, you can recreate a distributed switch configuration.

Restoring a distributed switch configuration overwrites the current settings of the distributed switch and its port groups. Full functionality is restored in instances of network settings failure or vCenter Server database corruption. The restore function does not delete port groups that are not part of the configuration file.

When you import a distributed switch configuration, you can create multiple copies of an existing deployment.

https://t.me/learningnets

2-29 Rollback and Recovery of the Management Network

Rollback prevents the accidental misconfiguration and loss of connectivity to vCenter Server by rolling back to the previous valid management network configuration.

Rollback provides the following options to recover from management network misconfigurations:

- Automatic rollback if misconfiguration is detected
- Direct Console User Interface (DCUI) to recover the management network

Automatic rollback and recovery is a vSphere feature that helps prevent management network outages. You protect the management network in the following ways:

- The automatic rollback feature detects any configuration changes on the management network. If the host cannot reach vCenter Server, the changes are not permitted to take effect.
- You can also reconfigure the management network on the distributed virtual switch by using the Direct Console User Interface (DCUI). Using the DCUI to change the management network parameters of the switch changes them for all hosts connected to the switch.

https://t.me/learningnets

2-30 Rollback Details

By rolling back configuration changes, vSphere protects hosts from losing connection to vCenter Server due to misconfiguration of the management network.

Updates that trigger a rollback:

- Host-level rollback: Triggered by a change in the host networking configurations, such as a physical NIC speed change, a change in MTU configuration, or a change in IP settings.
- Distributed switch-level rollback: Occurs after the user updates distributed switch-related objects, such as a port group or distributed ports.

Rollback is enabled by default.

The following types of updates can trigger a rollback:

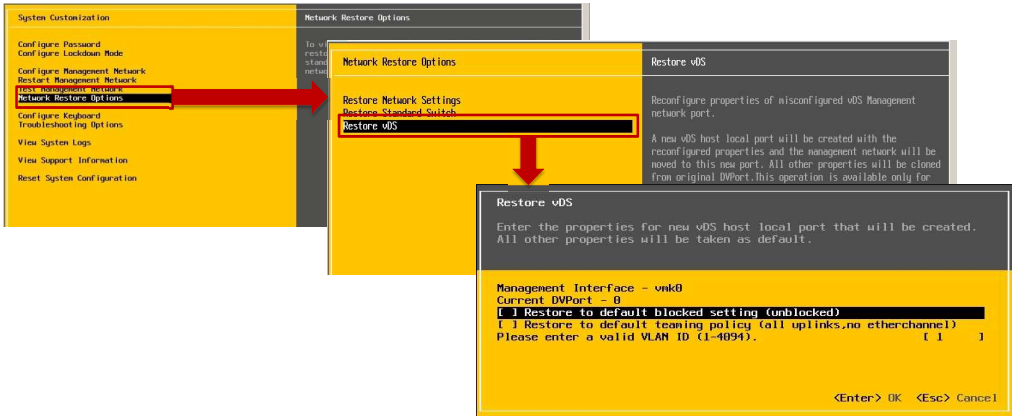
- If a change in the host networking configuration causes connectivity problems between the host and the vCenter Server system, a rollback occurs. Examples of misconfigurations include a mismatch in the speed and duplex of the physical NIC, a mismatch in the jumbo frame's configuration, and a misconfiguration in the IP settings of the management network.
- Misconfiguration of the distributed virtual switch that causes connectivity issues between the hosts and the vCenter Server system can also cause a rollback. Port group and distributed virtual port misconfigurations can cause network downtime.

You can disable rollback by changing the vCenter Server `config.vpxd.network.rollback` advanced setting to false. After you modify the advanced setting, you must restart the vCenter Server system for the new setting to take effect.

https://t.me/learningnets

2-31 Recovery Through the DCUI

If rollback is disabled, the DCUI provides an easy way for the user to connect directly to the host and fix the distributed switch properties. DCUI recovery must be performed per host.



To roll back the configuration manually from the DCUI:

1. In the DCUI, press the F2 key.
2. When prompted, log in to the DCUI with the appropriate credentials.
3. Use the down arrow key to scroll to Network Restore Options and press Enter.

   The Network Restore Options screen appears.
4. Select **Restore vDS** to roll back the distributed network settings.

   The Restore vDS screen appears.

   From the Restore vDS screen, you can either restore the host settings to the default blocked setting or restore the default teaming policy of the host.
5. Select the appropriate option and press Enter.

https://t.me/learningnets

2-32 Lab 1: Using vSphere Distributed Switches

Create, configure, and back up a distributed switch

1. Access Your Lab Environment
2. Verify That the vSphere Licenses Are Valid
3. Assign Valid vSphere Licenses
4. Create a Distributed Switch
5. Add ESXi Hosts to the New Distributed Switch
6. Examine Your Distributed Switch Configuration
7. Migrate the Virtual Machines to a Distributed Switch Port Group
8. Enable the Distributed Switch Health Check
9. Back Up the Distributed Switch Configuration

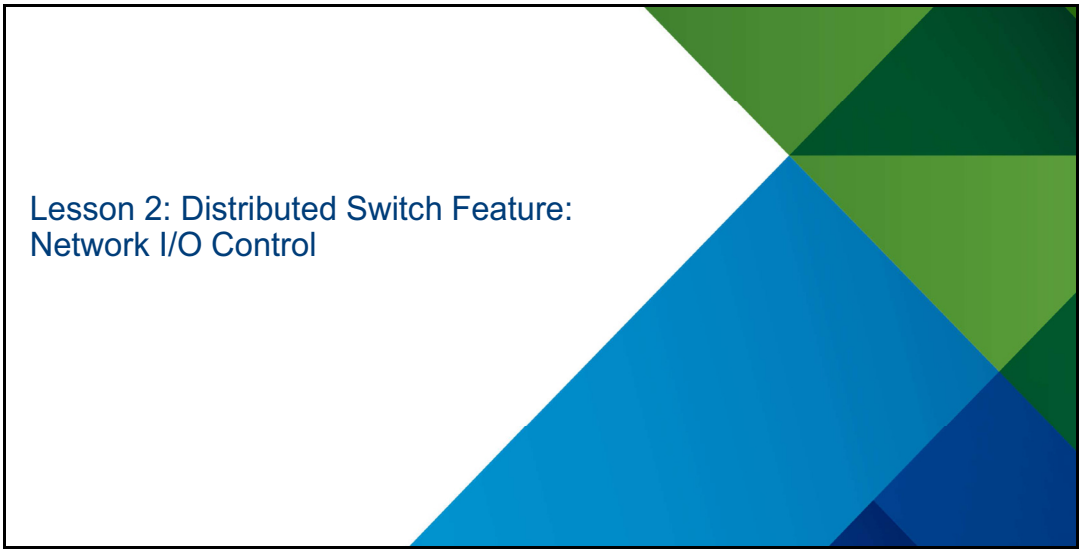
https://t.me/learningnets

2-33 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe the benefits and features of distributed switches:
  - Discovery protocols
  - Port binding
  - Traffic filtering and marking policy
  - Automatic rollback and recovery of networking configurations
- Create a distributed switch
- Perform a distributed switch health check
- Back up and restore a distributed switch configuration

https://t.me/learningnets



https://t.me/learningnets

2-35 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

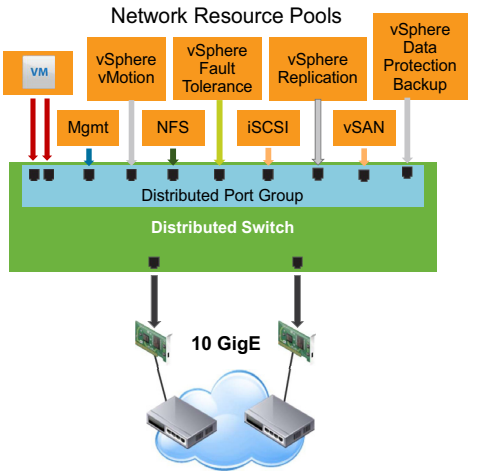
- Describe how VMware vSphere® Network I/O Control enhances performance
- Describe how to use Network I/O Control to allocate bandwidth to different types of system traffic
- Describe how to use Network I/O Control to allocate bandwidth to virtual machines

https://t.me/learningnets

2-36 About Network I/O Control Version 3

Network I/O Control version 3 can be used to allocate network bandwidth to business-critical applications and to resolve situations where several types of traffic compete for common resources.

Network I/O Control version 3 allocates network bandwidth over distributed switches by using network resource pools for virtual machine and system traffic.



VMware vSphere® Network I/O Control version 3 introduces a mechanism to reserve bandwidth for system traffic based on the capacity of the physical adapters on a host. It enables fine-grained resource control at the VM network adapter level, similar to the model that you use for allocating CPU and memory resources.

https://t.me/learningnets

2-37 Network I/O Control Versions 2 and 3

vSphere 6.x provides Network I/O Control version 3.

Version 3 has the following features:

- You can reserve bandwidth for system traffic and virtual machine traffic based on the capacity of the physical adapters on a host.
- It enables detailed resource control at the virtual machine network adapter level, similar to the model that you use for allocating CPU and memory resources. Shares, reservations, and limits are used to control bandwidth.

vSphere 6.x also supports Network I/O Control version 2, available in distributed switch versions 5.1 and 5.5:

- Version 2 and version 3 use different models for allocating bandwidth to virtual machines and system traffic.
- vCenter Server can support up to 128 distributed switches, each of which can be configured with either version 2 or version 3.

You can upgrade Network I/O Control from version 2 to version 3 by upgrading the distributed switch from version 5.1.0 or 5.5.0 to version 6.x. However, the upgrade is disruptive. Certain functionality is available only in Network I/O Control version 2 and is removed during the upgrade to version 3.

The following functionality is removed from Network I/O Control version 3:

- User-defined network resource pools, including all associations between them and existing distributed port groups.
  You can preserve certain resource allocation settings by transferring the shares from the user-defined network resource pools to shares for individual network adapters. Before you upgrade to Network I/O Control version 3, ensure that the upgrade does not greatly affect the bandwidth allocation that is configured for virtual machines in Network I/O Control version 2.
- Existing associations between ports and user-defined network resource pools.

For more information about the upgrade procedure, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

https://t.me/learningnets

2-38 About the Bandwidth Allocation Model for System Traffic

Network I/O Control version 3 uses several configuration parameters to allocate bandwidth to system traffic (such as management, iSCSI, VMware vSAN™, and VMs).

Bandwidth Parameter	Description
Shares	The relative priority of a system traffic type against other system traffic types that are active on the same physical adapter. Use the following values to define the number of shares: • Low: 25 • Normal: 50 • High: 100 • Custom: A user-defined value (from 1 through 100)
Reservations	The minimum bandwidth, in Mbps, that must be guaranteed on a single physical adapter.
Limits	The maximum bandwidth, in Mbps or Gbps, that a system traffic type can consume on a single physical adapter.

The amount of bandwidth available to a system traffic type is determined by its relative shares and the amount of data that the other system features transmit.

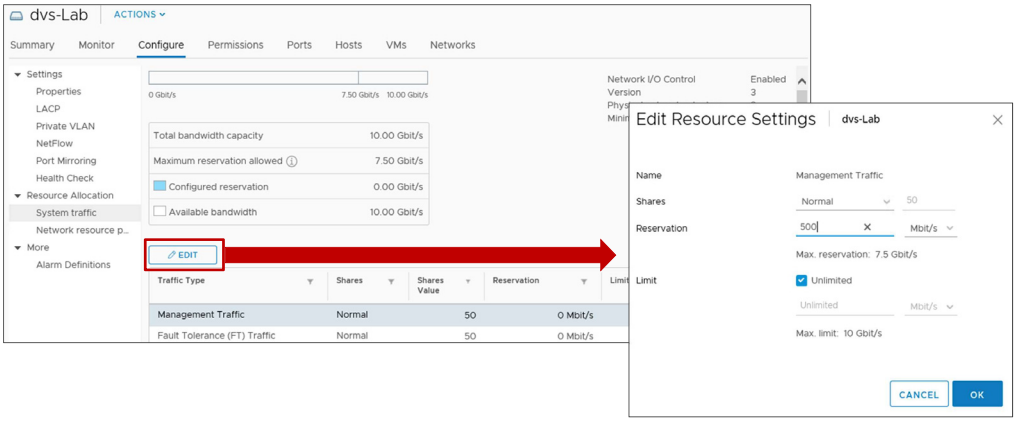
Reserved bandwidth that is unused becomes available to other types of system traffic. However, Network I/O Control version 3 does not redistribute the capacity that system traffic does not use to virtual machine placement.

For example, you configure a reservation of 2 Gbps for iSCSI. Because iSCSI uses a single path, the distributed switch might never impose this reservation on a physical adapter. In this case, the vacant bandwidth is not allocated to virtual machine system traffic. Thus, Network I/O Control version 3 cannot safely meet a potential need for bandwidth for system traffic. This failure to reallocate unused reserved bandwidth to virtual machine system traffic is especially true in the case of a new iSCSI path where you must provide bandwidth to a new VMkernel adapter.

https://t.me/learningnets

2-39 Configuring Bandwidth Allocations for System Traffic

To use Network I/O Control, you can configure shares, bandwidth reservation, and limits for each type of system traffic.



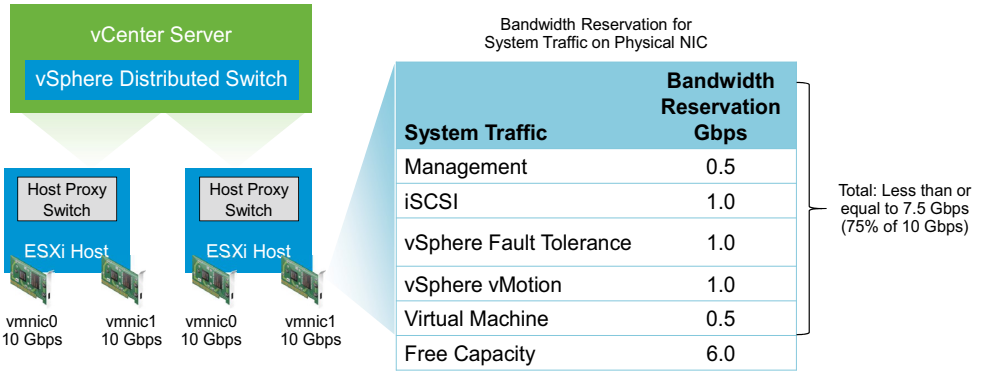
Network I/O Control version 3 can be used to configure bandwidth allocation for the traffic that is related to the main features of vSphere:

- Virtual machines
- Management
- vSphere vMotion
- NFS
- VMware vSphere® Fault Tolerance
- iSCSI
- VMware vSphere® Replication™
- VMware vSAN™
- VMware vSphere® Data Protection™

https://t.me/learningnets

2-40 Reserving Bandwidth for System Traffic (1)

Network I/O Control version 3 allocates the requested bandwidth on each physical network adapter. You can reserve no more than 75 percent of the bandwidth of a physical network adapter.



You can guarantee minimum bandwidth to a system feature for optimal operation according to the capacity of the physical adapters.

If you have a distributed switch that is connected to ESXi hosts with 10 GbE network adapters, you can configure a reservation to guarantee 1 Gbps for vSphere vMotion traffic, 1 Gbps for vSphere Fault Tolerance, 0.5 Gbps for virtual machine traffic, and so on.

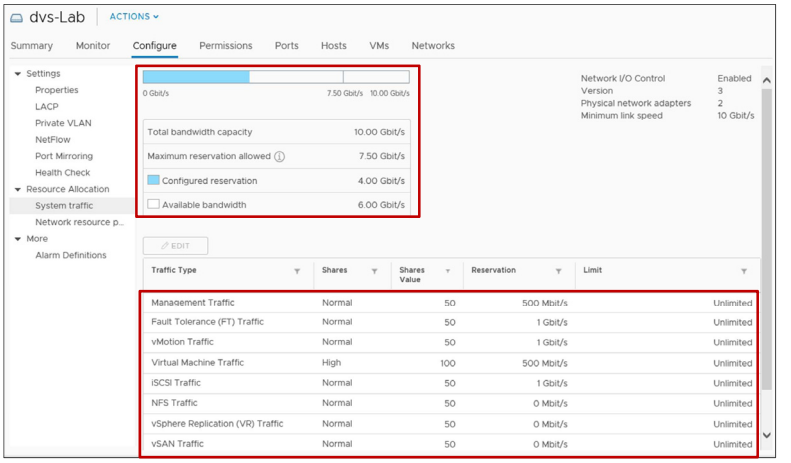
You might leave more capacity unreserved to let the host dynamically allocate bandwidth according to shares, limits, and use, and to reserve only bandwidth that is enough for the operation of a system feature.

You can reserve up to 75 percent of the bandwidth of a physical network adapter. For example, if your ESXi host uses a 40 GbE network adapter, then 30 Gbps (.75 x 40) is reserved.

https://t.me/learningnets

2-41 Reserving Bandwidth for System Traffic (2)

vSphere Client shows bandwidth usage information as well as the share, reservation, and limit values for each traffic type.

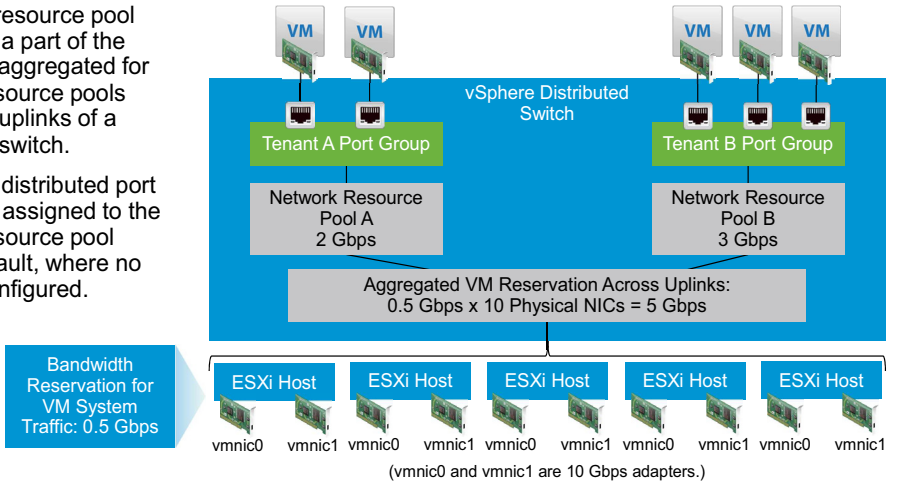


https://t.me/learningnets

2-42 Bandwidth Aggregation for Network Resource Pools

A network resource pool represents a part of the bandwidth aggregated for network resource pools across the uplinks of a distributed switch.

By default, distributed port groups are assigned to the network resource pool named default, where no quota is configured.



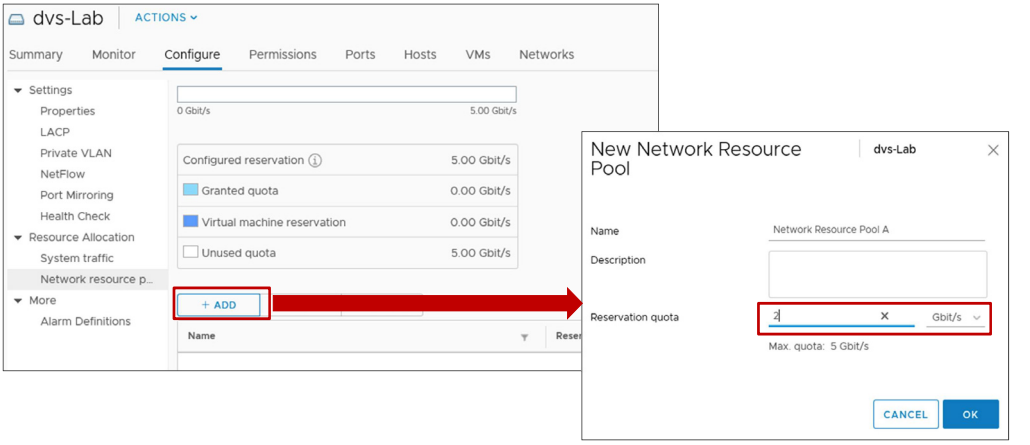
Network I/O Control version 3 allocates bandwidth for virtual machines across the entire distributed switch and on the physical adapter carrying the virtual machine traffic.

To use Network I/O Control version 3 to enable bandwidth allocation for virtual machines, configure the virtual machine system traffic. The bandwidth reservation for virtual machine traffic is also used in admission control. When you power on a virtual machine, admission control verifies that enough bandwidth is available.

https://t.me/learningnets

2-43 Creating Network Resource Pools

Create a network resource pool to reserve bandwidth for a set of virtual machines.



For example, if the virtual machine system traffic has 0.5 Gbps reserved on each 10 GbE uplink on a distributed switch that has 10 uplinks, then the total aggregated bandwidth available for virtual machine reservation on this switch is 5 Gbps. Each network resource pool can reserve a quota of this 5 Gbps capacity.

The bandwidth quota that is dedicated to a network resource pool is shared among the distributed port groups associated with the pool. A virtual machine receives bandwidth from the pool through the distributed port group that the virtual machine is connected to.

https://t.me/learningnets

2-44 Bandwidth Allocation for Virtual Machine Traffic

Bandwidth is allocated to individual virtual machines based on configured shares, reservations, and limits for the virtual machine network adapters.

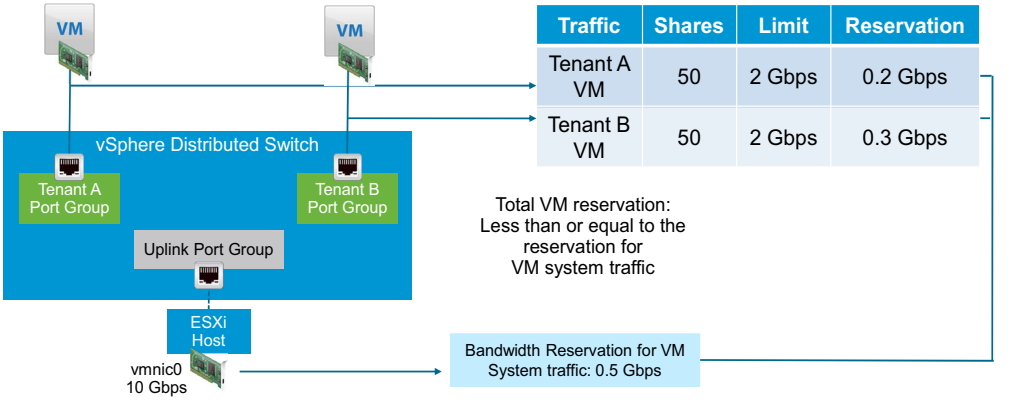
Bandwidth Parameter	Description
Shares	The relative priority of the traffic through the virtual machine network adapter against the capacity of the physical adapter that is carrying the virtual machine traffic to the network. Use the following values to define the number of shares: • Low: 25 • Normal: 50 • High: 100 • Custom: A user-defined value (from 1 through 100)
Reservations	The minimum bandwidth, in Mbps, that the virtual machine network adapter must receive on the physical adapter.
Limits	The maximum bandwidth on the virtual machine network adapter for traffic to other virtual machines on the same or on another host.

A network resource pool provides a reservation quota to virtual machines. The quota represents a portion of the bandwidth that is reserved for virtual machine system traffic on the physical adapters connected to the distributed switch. You can set aside bandwidth from the quota for the virtual machines that are associated with the pool. The reservation from the network adapters of powered-on VMs that are associated with the pool must not exceed the quota of the pool.

https://t.me/learningnets

2-45 Bandwidth Allocation for Individual Virtual Machines

The total bandwidth reservation of the virtual machines on a host cannot exceed the reserved bandwidth that is configured for the virtual machine system traffic.



To guarantee bandwidth, Network I/O Control version 3 implements a traffic placement engine that becomes active if a virtual machine has bandwidth reservation configured. The distributed switch tries to direct the traffic from a virtual machine network adapter to a physical adapter that can supply the required bandwidth and is in the scope of the active teaming policy.

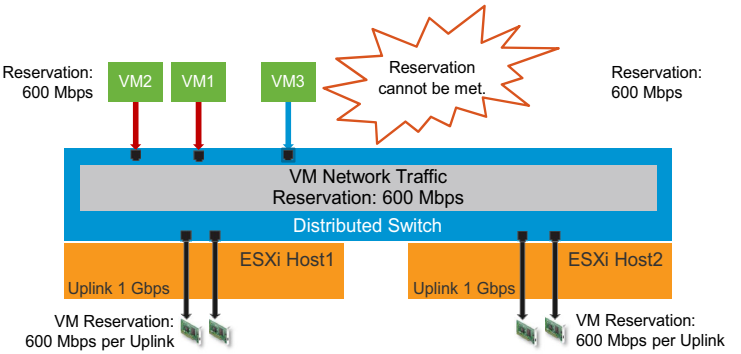
The real limit and reservation also depend on the traffic-shaping policy on the distributed port group that the adapter is connected to. For example, if a virtual machine network adapter asks for a limit of 200 Mbps and the average bandwidth configured in the traffic-shaping policy is 100 Mbps, then the effective limit becomes 100 Mbps.

https://t.me/learningnets

2-46 Bandwidth Admission Control in vSphere DRS (1)

Bandwidth admission control verifies that the virtual machine reservation can be met.

If the reservation cannot be met on the current host, then VMware vSphere® Distributed Resource Scheduler™ places the virtual machine on a host that has the capacity to guarantee the bandwidth reserved for the virtual machine.



If you power on a virtual machine that is in a cluster, VMware vSphere® Distributed Resource Scheduler™ places that virtual machine on a host that has the capacity to guarantee the bandwidth reserved for the virtual machine, according to the active teaming policy.

To use admission control in vSphere DRS, you must perform the following tasks:

- Configure bandwidth allocation for the virtual machine system traffic on the distributed switch.
- Configure the bandwidth requirements of a virtual machine that is connected to the distributed switch.

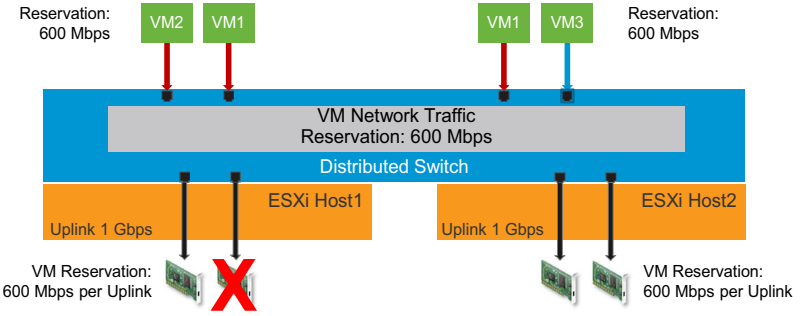
If a virtual machine reservation cannot be met, the virtual machine is not powered on.

https://t.me/learningnets

2-47 Bandwidth Admission Control in vSphere DRS (2)

vSphere DRS migrates a VM to another host to satisfy the bandwidth reservation of a VM in the following situations:

- The reservation is changed to a value that the initial host can no longer satisfy.
- A physical adapter that carries traffic from the VM is offline.

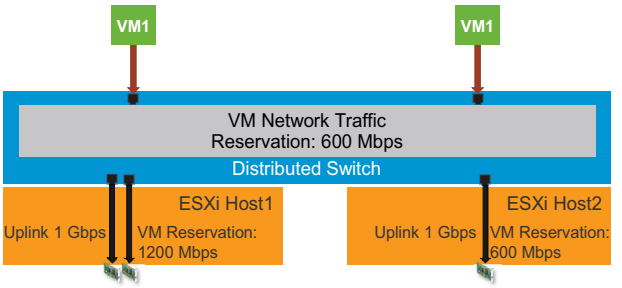


In this example, Host1 loses an uplink, which leaves Host1 with one working uplink. Because Host2 still has both of its uplinks working, vSphere DRS migrates either VM1 or VM2 to Host2 in order to meet the VM's network reservation requirement.

https://t.me/learningnets

2-48 Bandwidth Admission Control in vSphere HA

When a host fails, VMware vSphere® High Availability powers on the failed virtual machines on another host in the cluster according to the bandwidth reservation and teaming policy.



Bandwidth admission control prevents a virtual machine from being started if the bandwidth reservation for that virtual machine cannot be met.

To use admission control in VMware vSphere® High Availability, you must perform the following tasks:

- Allocate bandwidth for the virtual machine system traffic
- Configure the bandwidth requirements of a virtual machine that is connected to the distributed switch.

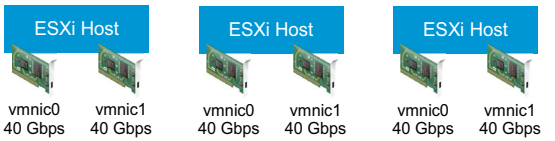
https://t.me/learningnets

2-49 Activity: Network I/O Control (1)

You have a distributed port group named Prod 1. You want to guarantee that the VMs connected to Prod 1 can access 25 percent of the available reserved bandwidth for VM traffic.

Using the bandwidth reservation example shown, what must you configure to provide the appropriate bandwidth to the VMs on Prod 1?

System Traffic	Bandwidth Reservation Gbps
Management	4
iSCSI	12
Virtual Machine	8



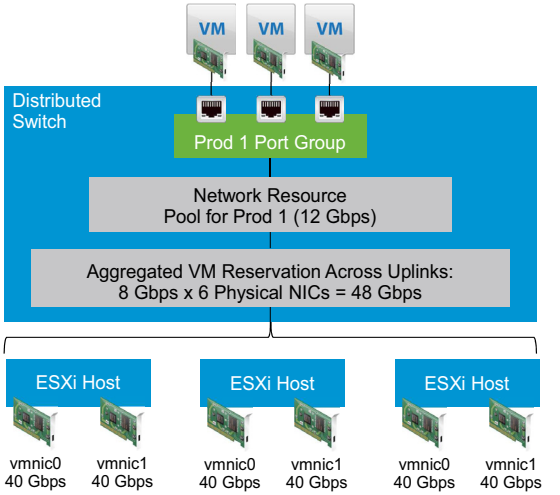
https://t.me/learningnets

2-50 Activity: Network I/O Control (2)

The aggregated amount of reserved bandwidth for VM traffic is 48 Gbps.

25 percent of the aggregated amount is 12 Gbps.

Therefore, you must create a network resource pool for Prod 1 with a 12 Gbps reservation.



https://t.me/learningnets

2-51 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe how Network I/O Control enhances performance
- Describe how to use Network I/O Control to allocate bandwidth to different types of system traffic
- Describe how to use Network I/O Control to allocate bandwidth to virtual machines

https://t.me/learningnets



https://t.me/learningnets

2-53 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Describe how Link Aggregation Control Protocol (LACP) enhances network availability and performance
- Describe how NetFlow can be used to monitor network security and performance
- Configure port mirroring on a distributed switch

https://t.me/learningnets

2-54 About LACP

LACP is a standards-based (802.3ad) link aggregation method that is supported on distributed switches.

LACP has the following features:

- Enables you to connect ESXi hosts to physical switches that use dynamic link aggregation
- Detects link failures and cabling mistakes
- Automatically negotiates link aggregation properties between virtual and physical switches

You use LACP to increase network bandwidth and redundancy.

LACP support is provided by using a link aggregation group (LAG). A LAG aggregates the bandwidth of physical NICs on ESXi hosts that are connected to LACP port channels.

LACP sends frames on all links that have the protocol enabled. If LACP finds a device on the other end of the link that also has LACP enabled, LACP sends frames along the same set of links. This action enables the units to detect multiple links between themselves and combine them into a single logical channel.

LACP uses a heartbeat between the endpoints to detect link failures and cabling mistakes. LACP also automatically reconfigures the broken links.

Using the LACP support on a distributed switch, network devices can negotiate the automatic bundling of links by sending LACP packets to a peer. However, some limitations are imposed when you use LACP with a distributed switch:

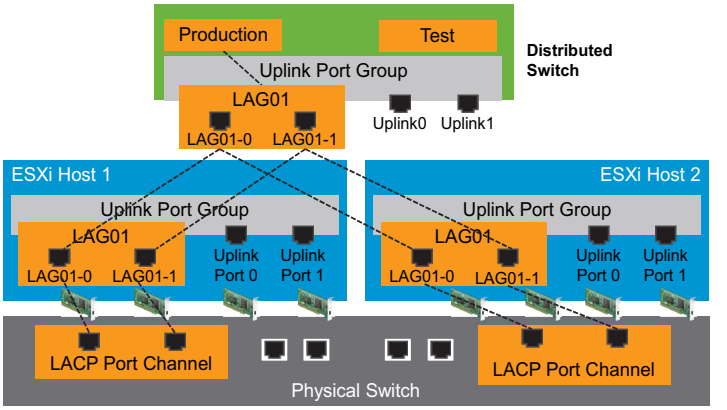
- The LACP support is not compatible with software iSCSI multipathing.
- The LACP support settings do not exist in host profiles.
- The LACP support does not work with port mirroring.

For information about LACP support, including the complete list of limitations, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

https://t.me/learningnets

2-55 About LAG Architecture

You configure the same number of ports for a LAG as the number of ports on the LACP port channels on the physical switch.



LAG ports have the same function as standalone uplinks and are teamed within the LAG.

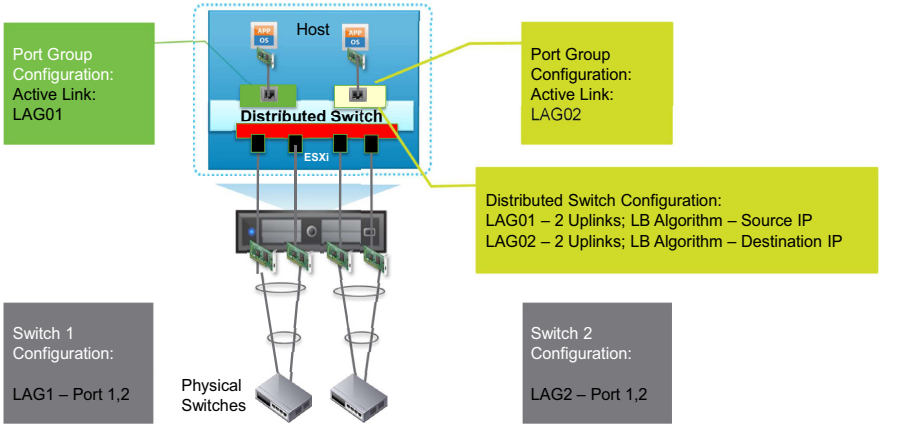
A representation of a LAG is available on the distributed switch and on the proxy switch on every host that is connected to the distributed switch. For example, if you create LAG01 with two ports, LAG01 with the same number of ports is created on every host that is connected to the distributed switch.

On the host side, you can connect one physical NIC to each LAG port. On the distributed switch, one LAG port can have multiple physical NICs from hosts connected to it. The physical NICs on a host that you connect to the LAG ports must be connected to links that participate in an LACP port channel on the physical switch.

https://t.me/learningnets

2-56 Example of LACP Deployment with Two LAGs

This example shows a vSphere host deployment with four uplinks, connected to two physical switches.



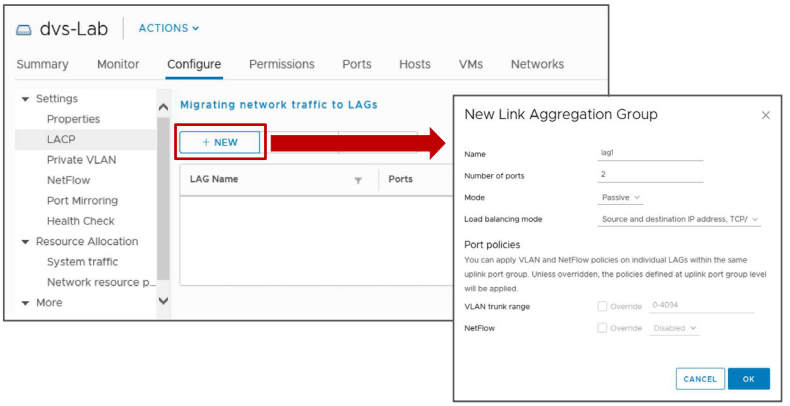
LAGs are created by combining two uplinks on the physical and virtual switches. The LACP configuration on the vSphere host is performed on the distributed switch and the port groups.

First, the LAGs and the associated uplinks are configured on the distributed switch. Then, the port groups are configured to use those LAGs. In the example, the green port group is configured with LAG1 and the yellow port group is configured with LAG2. All the traffic from virtual machines connected to the green port group follow the LAG1 path.

https://t.me/learningnets

2-57 Creating LAGs

When you create a new LAG on the distributed switch, the new LAG does not yet have physical NICs assigned to its ports.



Set the same amount of ports to the LAG as the amount of ports in the LACP port channel on the physical switch. A LAG port has the same function as an uplink on the distributed switch. All LAG ports form a NIC team in the context of the LAG.

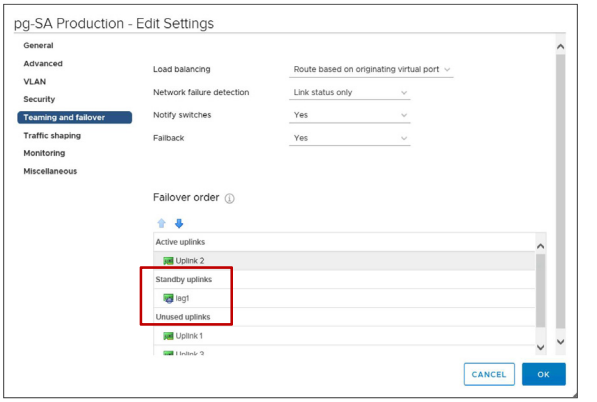
https://t.me/learningnets

2-58 Configuring LAGs Without Losing Connectivity (1)

The first step is to make the LAG a standby uplink:

- Because only one LAG or only standalone uplinks can be active for distributed port groups, you must create an intermediate teaming and failover configuration where the LAG is a standby uplink.

This configuration allows you to migrate physical NICs to the LAG ports by keeping the network connectivity up.



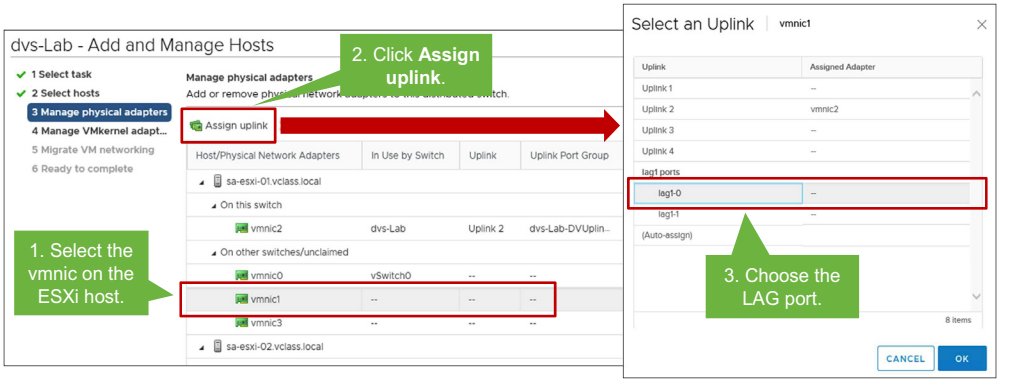
By default, the new LAG is unused in the teaming and failover order of distributed port groups.

https://t.me/learningnets

2-59 Configuring LAGs Without Losing Connectivity (2)

Next, add physical NICs to the LAG:

- With the LAG in standby mode, you can safely migrate the physical NICs from standalone uplinks to the LAG ports without losing network connectivity.



For example, if you have a LAG with two ports, you configure a physical NIC to each LAG port in the Add and Manage Hosts wizard.

https://t.me/learningnets

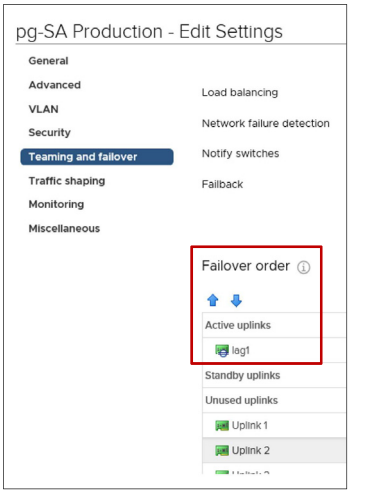
2-60 Configuring LAGs Without Losing Connectivity (3)

Finally, use the arrow keys to move the LAG to the Active uplinks section:

- You can now use the LAG to handle traffic for a distributed port group by setting the LAG as active in the group's teaming and failover order.

Also, move all standalone uplinks to the Unused uplinks section.

The network traffic is load balanced between the LAG ports: All load balancing algorithms of LACP are supported by the distributed switch.



You can create up to 64 LAGs on an ESXi host. However, the number of LAGs that you can use depends on the capabilities of the underlying physical environment and the topology of the virtual network.

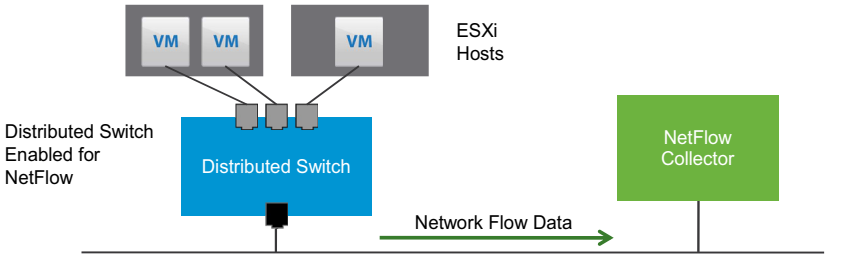
https://t.me/learningnets

2-61 About NetFlow

NetFlow is a network analysis tool for monitoring the network and viewing virtual machine traffic flowing through a distributed switch.

NetFlow can be used for profiling, intrusion detection and prevention, networking forensics, and compliance.

The vSphere distributed switch supports IPFIX (NetFlow version 10).



NetFlow is a protocol that Cisco Systems developed for analyzing network traffic. NetFlow has since become an industry-standard specification for collecting types of network data for monitoring and reporting. The data sources are network devices, such as switches and routers. For ESXi deployments, NetFlow enables detailed monitoring and analysis of virtual machine network traffic.

Standard switches do not support NetFlow.

NetFlow collectors are available from third-party providers.

https://t.me/learningnets

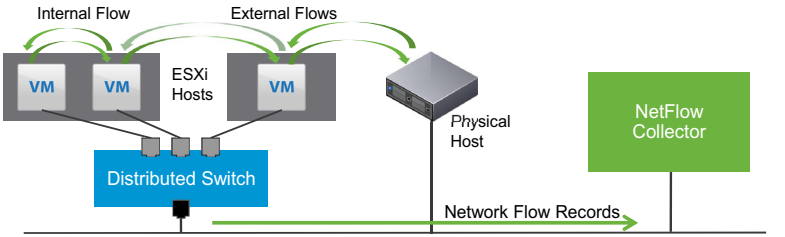
2-62 About Network Flows

A network flow is a unidirectional sequence of packets, with each packet sharing a common set of properties.

NetFlow captures the following types of flows:

- Internal flow: Represents intrahost virtual machine traffic.
- External flow: Represents interhost virtual machine traffic and physical machine-to-virtual machine traffic.

Flow records are sent to a NetFlow collector for analysis.



Network flows give you a complete view of virtual machine traffic, which can be collected for historical views and used for multiple purposes.

Internal flows are generated from intra-host virtual machine traffic, that is, traffic between virtual machines on the same hosts.

External flows are generated from traffic between virtual machines located on different hosts or virtual machines on different distributed switches. External flows are also generated from traffic between physical machines and virtual machines.

A flow is a sequence of packets that share the following properties, which include source and destination IP addresses, source and destination ports, input and output interface IDs, and protocol.

A flow is unidirectional. Flows are processed and stored as flow records by supported network devices, such as a distributed switch. The flow records are then sent to a NetFlow collector for additional analysis.

Although flow processing is an efficient method, NetFlow can strain the distributed switch. NetFlow requires both additional processing and storage on the host for the flow records to be processed and exported.

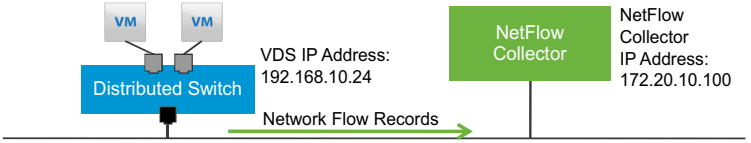
https://t.me/learningnets

2-63 Network Flow Analysis

Network flow data is sent to a third-party NetFlow collector, which accepts and stores network flow records.

A NetFlow collector has the following features:

- Includes a storage system for long-term storage of flow-based data:
  - You can investigate and isolate excessive network bandwidth use, bottlenecks, and unexpected application traffic.
  - You can view historical records to diagnose the cause outages or breaches.
- Mines, aggregates, and reports on the collected data:
  - You can analyze network traffic by rate, volume, and utilization.
  - You can analyze trends in virtual machine and host traffic.



NetFlow sends aggregated network flow data to a NetFlow collector. Third-party vendors have NetFlow collector products.

A NetFlow collector accepts and stores the completed network flow records. NetFlow collectors vary in functionality by vendor. A NetFlow collector provides the following features:

- Analysis software to mine, aggregate, and report on the collected data
- A storage system to enable long-term storage so that you can archive the network flow data
- A customized user interface, often based on a web browser

The NetFlow collector can report on various kinds of networking information:

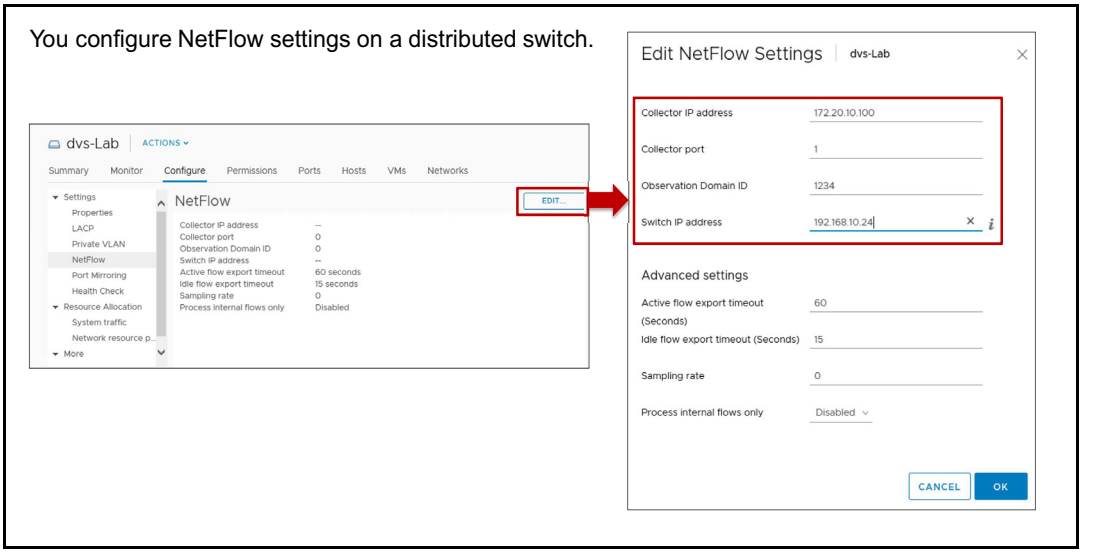
- The current top network flows consuming the most bandwidth in a particular (virtual) switch
- The IP addresses that are behaving irregularly
- The number of bytes that a particular virtual machine has sent and received in the past 24 hours

With NetFlow data, you can investigate the causes of excessive use of network bandwidth, bottlenecks, and unexpected application traffic. The historical records that you put in long-term storage can help you diagnose what might have caused these outages or breaches.

Because NetFlow data comes from NetFlow-enabled network devices, additional network probes to collect the flow-based data are not needed. NetFlow collectors and analyzers can provide a detailed set of network performance data. Given enough storage on the NetFlow collector, flow data can be archived for a long time, providing a long-term record of network behavior.

https://t.me/learningnets

2-64 Configuring NetFlow on Distributed Switches



To configure NetFlow on a distributed switch, specify the following NetFlow settings:

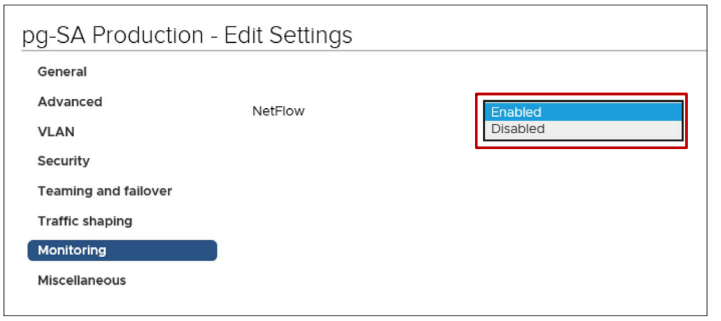
- Collector IP address and Collector port: The IP address and port number used to communicate with the NetFlow collector system. The values are assigned by the collector.
- Observation Domain ID: A value that identifies the information related to the switch.
- Switch IP address: An optional IP address that is used to identify the source of the network flow to the NetFlow collector. The IP address is not associated with a network port, and it does not have to be pingable. This IP address is used to fill the source IP field of NetFlow packets. This IP address enables the NetFlow collector to interact with the distributed switch as a single switch, rather than seeing a separate, unrelated switch for each associated host. If this IP address is not configured, the host's management IP address is used instead.

For a description of the advanced NetFlow settings, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

https://t.me/learningnets

2-65 Enabling NetFlow on a Distributed Port Group

After configuring NetFlow on the distributed switch, you either enable or disable NetFlow on a distributed port group, a specific port, or at the uplink.



https://t.me/learningnets

2-66 About Port Mirroring

Port mirroring is a technology that duplicates network packets from a source to a destination.

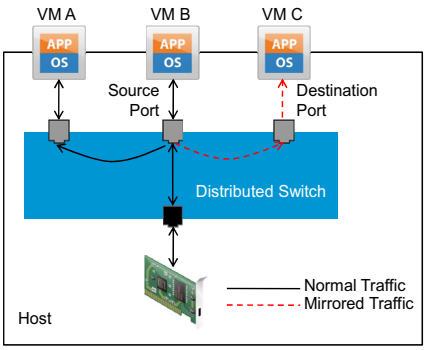
Port mirroring has the following uses:

- Assistance with troubleshooting
- Input for network analysis appliances

Many network switch vendors implement port mirroring in their products.

vSphere port mirroring supports the following switch functionality:

- Switched Port Analyzer (SPAN)
- Remote SPAN (RSPAN)
- Encapsulated RSPAN



On physical switches, administrators often have to mirror traffic to special ports to troubleshoot network-related problems. Port mirroring is commonly used for network appliances that require the monitoring of network traffic, such as intrusion detection systems.

Many network switch vendors implement port mirroring in their products. For example, port mirroring on a Cisco Systems switch is usually called Switched Port Analyzer (SPAN).

Port mirroring eliminates the need to enable promiscuous mode on a distributed switch in order to troubleshoot network issues. If you enable promiscuous mode on a distributed port, the port sees all the network traffic passing through the distributed switch. You cannot select the port or port group that a promiscuous port is allowed to see. The promiscuous port sees all traffic that is on the broadcast domain.

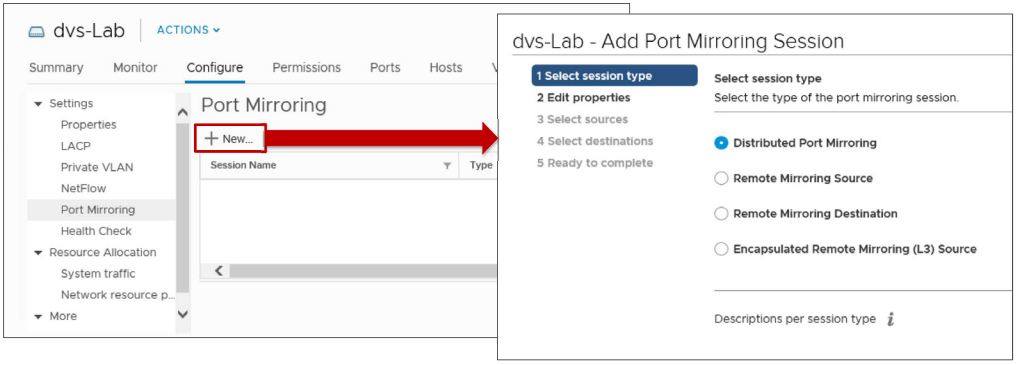
Port mirroring supports Cisco Remote Switch Port Analyzer (RSPAN) and Encapsulated Remote Switch Port Analyzer (ERSPAN). With RSPAN, mirrored traffic can be directed to a remote monitor. The RSPAN session can span multiple switch hops on a network. With ERSPAN, the session can span an IP network.

https://t.me/learningnets

2-67 Port Mirroring Session Types

You create a port mirroring session to mirror distributed switch traffic to ports, uplinks, and remote IP addresses.

You must select a port mirroring session type.



The following port mirroring session options are available:

- Distributed Port Mirroring: Port mirroring between distributed ports can only be local. If the source and the destination ports are on different hosts due to vSphere vMotion migration, mirroring between them does not work. However, if the source and destination ports move to the same host, port mirroring works.
- Remote Mirroring Source: When a source distributed port is moved from host A to host B, the original mirroring path from the source port to A's destination uplink is removed on A. Also, a new mirroring path from the source port to B's destination uplink is created on B. Which uplink is used is determined by the uplink name specified in the session.
- Remote Mirroring Destination: When a destination distributed port is moved from host A to host B, all original mirroring paths from source VLANs to the destination port are moved from A to B.
- Encapsulated Remote Mirroring (L3) Source: When a source distributed port is moved from host A to host B, all original mirroring paths from the source port to the destination IP addresses are moved from A to B.

The type of source and destination involved in port mirroring depends on the session type that is selected for port mirroring.

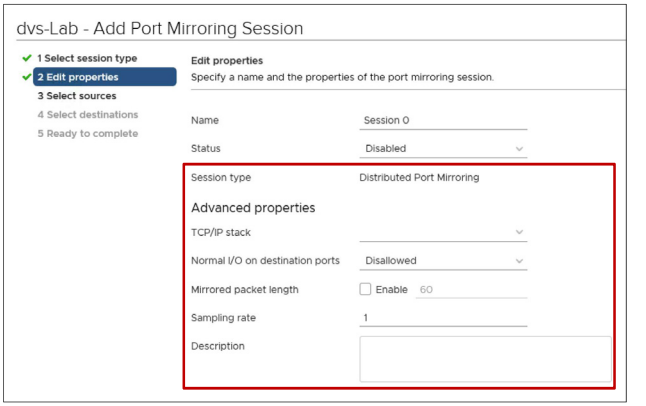
https://t.me/learningnets

2-68 Port Mirroring Session Properties

Based on the port mirroring session type that you select, you can configure one or more advanced properties.

For example, the following properties exist for a distributed port mirroring type session:

- TCP/IP stack type (default or mirror)
- Whether to allow or disallow normal I/O on destination ports
- Mirrored packet length (in bytes)
- Rate at which packets are sampled
- Description



Every port mirroring session is uniquely identified by its name. A session can also have a description.

For a description of the advanced session properties, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

https://t.me/learningnets

2-69 Source and Destination in a Port Mirroring Session

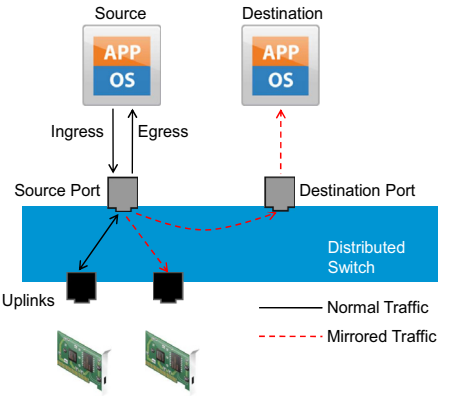
A port mirroring session requires details about the source and the destination.

Source details include:

- Traffic direction
- One or more port IDs

Destination details include:

- One or more port IDs
- One or more uplinks
- IP address



When creating a port mirroring session, you must configure the source and the destination. When configuring the source, you must specify the traffic direction.

Traffic direction is categorized as ingress or egress. Ingress traffic direction is traffic flowing from the source virtual machine into the distributed switch. Egress traffic direction is traffic flowing from the distributed switch into the source virtual machine.

To avoid flooding the network with mirrored traffic, port mirroring has the following restrictions:

- In a session, a port cannot be a source and a destination.
- A port cannot be a destination for more than one session.
- A promiscuous port cannot be an egress source or destination.
- An egress source cannot be a destination for sessions, to avoid cycles of mirroring paths.

https://t.me/learningnets

2-70 Lab 2: Using Port Mirroring

Configure port mirroring and capture network traffic on a distributed switch

1. Prepare to Capture Mirrored Network Traffic
2. Configure Port Mirroring on the Distributed Switch
3. Verify That Port Mirroring Is Capturing Traffic
4. Restore the Distributed Switch Configuration

https://t.me/learningnets

2-71 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe how LACP enhances network availability and performance
- Describe how NetFlow can be used to monitor network security and performance
- Configure port mirroring on a distributed switch

https://t.me/learningnets

2-72 Key Points

- A distributed switch provides functions that are similar to a standard switch. But the distributed switch defines a single configuration that is shared across all associated hosts.
- Network I/O Control version 3 allocates bandwidth to each type of system traffic by using shares, reservations, and limits.
- Distributed switch and distributed port group configurations can be backed up and restored.
- The use of LACP increases network bandwidth and redundancy.
- Distributed switches support the use of network analysis and troubleshooting tools, specifically NetFlow and port mirroring.

Questions?

https://t.me/learningnets

Storage Scalability



https://t.me/learningnets

3-2 You Are Here

1. Course Introduction
2. Network Scalability
3. **Storage Scalability**
4. Host and Management Scalability
5. CPU Optimization
6. Memory Optimization
7. Storage Optimization
8. Network Optimization
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

3-3 Importance

As the enterprise grows, scalability features and software-defined storage configurations in vSphere enable the infrastructure to handle growth efficiently.

Datastore growth and balancing issues can be addressed automatically with VMware vSphere® Storage DRS™.

Storage policy-based management is a major element of your software-defined storage environment. This framework helps align storage with the application demands of your VMs.

https://t.me/learningnets

3-4 Module Lessons

Lesson 1:	VMFS Datastores
Lesson 2:	Storage APIs
Lesson 3:	Storage Policy-Based Management
Lesson 4:	Storage Policies for vSAN and Virtual Volumes
Lesson 5:	Storage I/O Control
Lesson 6:	Datastore Clusters and vSphere Storage DRS

https://t.me/learningnets

3-5 Lesson 1: VMFS Datastores



https://t.me/learningnets

3-6 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Explain why VMware vSphere® VMFS is a high-performance, scalable file system
- Describe the performance improvements of VMFS6 over VMFS5
- Describe the migration procedure from VMFS5 to VMFS6

https://t.me/learningnets

3-7 VMFS Datastores (1)

VMware vSphere VMFS is a high-performance, cluster file system that is optimized for VMs.

vSphere 6.5 and later supports VMFS5 and VMFS6 datastores.

In vSphere 6.7, VMFS3 datastores are automatically upgraded to VMFS5 when mounted.

VMFS5 and VMFS6 datastores are highly scalable:

- The maximum size for both a datastore and a single extent is 64 TB.
- The maximum virtual disk size is 62 TB.
- The file system block size used is 1 MB, which supports files up to 62 TB in size.
- Small files are efficiently stored:
  - File system blocks are divided into subblocks to efficiently store small files.
  - Data of very small files (less than or equal to 1 KB) is stored directly in the file descriptor.
- A virtual disk can be hot-extended to any size, up to 62 TB.

For more information on the hot extend feature, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-storage-guide.pdf.

https://t.me/learningnets

3-8 VMFS Datastores (2)

Certain ESXi host versions support VMFS5 and VMFS6 datastores.

Host Version	VMFS3	VMFS5	VMFS6	Maximum Datastores Per Host	Number of Paths
ESXi 6.0	Can be mounted, but not created.	Yes	No	256	1024
ESXi 6.5	Can be mounted, but not created.	Yes	Yes	512	2048
ESXi 6.7	Automatically upgraded to VMFS5 on first mount.	Yes	Yes	1024	4096

https://t.me/learningnets

3-9 VMFS5 and VMFS6 Datastores

VMFS5 and VMFS6 datastores support key vSphere features:

- vSphere vMotion
- vSphere Storage vMotion across different datastore types
- vSphere HA and VMware vSphere® Fault Tolerance
- vSphere DRS and vSphere Storage DRS
- Raw device mappings:
  - In physical compatibility mode, the RDM has a maximum size of 64 TB.
  - In virtual compatibility mode, the RDM has a maximum size of 62 TB.

VMFS6 provides improvements in scalability and performance over VMFS5:

- Faster file-creation times
- Parallel disk operations and concurrent transactions
- For more information about VMFS6 enhancements, see https://storagehub.vmware.com/t/vsphere-storage/vmware-vsphere-vmfs.

https://t.me/learningnets

3-10 Storage Device Formats

The storage industry is moving towards advanced format drives to provide large capacity drives to servers and storage arrays.

ESXi supports storage devices with traditional and advanced sector formats:

- 512n: Traditional sector format
- 512e (512-byte emulation): An advanced sector format that can support legacy applications and guest operating systems
- 4Kn (4K native): An advanced sector format that has greater capacity density, improved availability, and improved performance over the 512-byte sector drive

Sector Format	Logical Sector Size	Physical Sector Size
512n	512	512
512e	512	4,096
4Kn	4,096	4,096

4K drives have advantages over 512-byte sector drives:

- Improved total cost of ownership: 4K drives require less space for error correction codes (ECCs) than regular 512-byte sector drives. This space saving results in greater data density on 4K drives, which provides an improved total cost of ownership.
- Improved availability: 4K drives have a larger ECC field, which improves error correction and data integrity.
- Improved performance: 4K drives are expected to have better performance than 512-byte sector drives. However, this is true only when the guest operating system is configured to issue I/Os that are aligned to the 4K sector size.

Because legacy applications and operating systems do not support 4K native drives, the drive industry has introduced 4K sector size drives in 512-byte emulation mode to support legacy applications. These drives have a physical sector size of 4K, but the logical sector size is 512 bytes. These drives are called 512e drives, where the letter e stands for emulation mode. 512e drives will help drive adoption of 4K drives with the end goal of moving to 4K native drives.

https://t.me/learningnets

3-11 Storage Devices Supported by VMFS Datastores

You can deploy VMFS5 and VMFS6 datastores on 512n and 512e storage devices.

With vSphere 6.7, VMFS6 datastores can be deployed on 4Kn direct-attached storage devices.

ESXi continues to expose 512-byte sector virtual disk files to the guest operating system and therefore, ESXi emulates 4Kn drives as 512e drives.

Storage Device	Supported by VMFS5	Supported by VMFS6
512n	Yes.	Yes.
512e	Yes. However VMFS5 is not supported on local 512e devices.	Yes.
4Kn	No.	Yes, but virtual disks are presented to the guest operating system as a 512e device.

For more information about vSphere support for 512e and 4K native drives, see VMware knowledge base article 2091600 at http://kb.vmware.com/kb/2091600.

https://t.me/learningnets

3-12 Limitations of Using 4Kn Drives

4Kn drives have the following limitations when used in a vSphere environment:

- Only local SAS, SATA hard disk drives are supported.
- 4Kn SSD drives, 4Kn NVMe drives, and 4Kn drives as RDM are not supported.
- Booting from 4Kn drives is supported with UEFI only.
- Third-party multipathing plug-ins are not supported.

https://t.me/learningnets

3-13 Snapshot Formats for VMFS

On the VMFS datastore, the delta disk of a VM snapshot is a sparse disk.

Depending on the type of VMFS datastore, delta disks use different sparse formats.

Sparse Format	Used by VMFS5	Used by VMFS6
VMFSsparse	Yes, for virtual disks smaller than 2 TB.	No.
SEsparse (space-efficient sparse)	Yes, for virtual disks equal to or larger than 2 TB.	Yes, it is the default format.

The SEsparse format is similar to VMFSsparse, with some enhancements. This format is space efficient and supports space reclamation.

SEsparse is the recommended disk format for virtual desktop infrastructure workloads.

When you take a snapshot, the state of the virtual disk is preserved, which prevents the guest operating system from writing to it. A delta or child disk is created. The delta disk represents the difference between the current state of the virtual disk and the state that existed when the previous snapshot was taken.

VMFSsparse is implemented on top of VMFS. I/Os issued to a snapshot virtual machine are processed by the VMFSsparse layer. Technically, VMFSsparse is a redo-log that starts empty, immediately after a virtual machine snapshot is taken. The redo-log grows to the size of its base virtual machine disk (VMDK), when the entire VMDK is rewritten with new data after the virtual machine snapshots are taken. This redo-log is only a file in the VMFS datastore. After taking the snapshot, the base VMDK attached to the virtual machine is changed to the newly-created sparse VMDK.

With SEsparse space reclamation, blocks that are deleted by the guest operating system are marked, and commands are issued to the SEsparse layer in the hypervisor to unmap those blocks. This operation helps reclaim the space allocated by SEsparse after the guest operating system has deleted that data.

SEsparse is the recommended disk format for VMware Horizon® environments. In these environments, reclamation of storage space is critical because of the many tenants sharing storage.

https://t.me/learningnets

3-14 Migrating Your Data from VMFS5 to VMFS6

VMware does not support upgrading from VMFS5 to VMFS6.

A VMFS6 datastore can be created only as a new datastore.

You can use VMware vSphere® Storage vMotion® to migrate data from your VMFS5 datastore to a VMFS6 datastore.



https://t.me/learningnets

3-15 Activity: VMFS6 Features (1)

What scalability and performance features are available only with VMFS6 datastores? Select all that apply.

- ❑ Support for 4Kn drives
- ❑ Faster file-creation times
- ❑ SEsparse as a snapshot format
- ❑ Maximum datastore size of 64 TB
- ❑ Hot extending a virtual disk to any size up to 62 TB
- ❑ Parallel disk operations and concurrent transactions

https://t.me/learningnets

3-16 Activity: VMFS6 Features (2)

What scalability and performance features are available only with VMFS6 datastores? Select all that apply.

- ✓ Support for 4Kn drives
- ✓ Faster file-creation times
- ❑ SEsparse as a snapshot format
- ❑ Maximum datastore size of 64 TB
- ❑ Hot extending a virtual disk to any size up to 62 TB
- ✓ Parallel disk operations and concurrent transactions

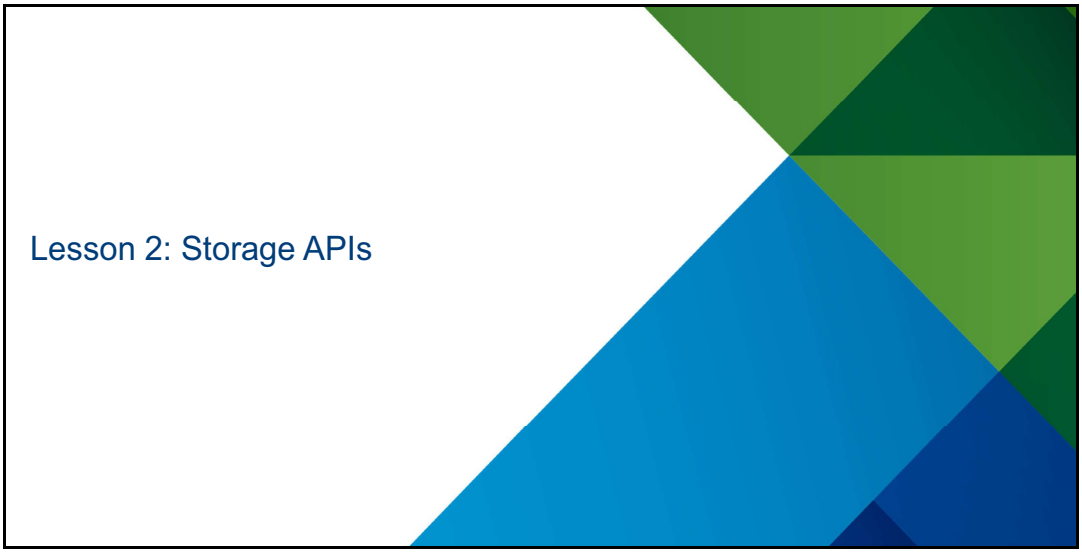
https://t.me/learningnets

3-17 Review of Learner Objectives

You should be able to meet the following objectives:

- Explain why VMFS is a high-performance, scalable file system
- Describe the performance improvements of VMFS6 over VMFS5
- Describe the migration procedure from VMFS5 to VMFS6

https://t.me/learningnets



https://t.me/learningnets

3-19 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Describe how VMware vSphere® Storage APIs - Array Integration helps storage arrays integrate with vSphere
- Explain how VMware vSphere® API for Storage Awareness™ can ensure that a VM's storage requirements are met
- Explain how vSphere APIs for I/O Filtering (VAIO) enables vendors to create data services for VMs

https://t.me/learningnets

3-20 About vSphere Storage APIs - Array Integration

vSphere Storage APIs - Array Integration helps storage vendors provide hardware assistance to accelerate vSphere I/O operations that are more efficiently accomplished in the storage hardware.

vSphere Storage APIs - Array Integration is a set of protocol interfaces and VMkernel APIs between ESXi and storage arrays:

- Hardware Acceleration APIs enable arrays to integrate with vSphere to transparently offload certain storage operations to the array.
- Array Thin Provisioning APIs help prevent out-of-space conditions and perform space reclamation.

VMware vSphere Storage® APIs is a family of APIs used by third-party hardware, software, and storage vendors to develop components that enhance several vSphere features and solutions.

In a virtualized environment, virtual disks are files on a VMFS datastore. Disk arrays cannot interpret the VMFS datastore's on-disk data layout, so the VMFS datastore cannot use hardware functions per virtual machine or per virtual disk file. VMware vSphere® Storage APIs - Array Integration plug-ins can improve data transfer performance and are transparent to the end user.

https://t.me/learningnets

3-21 Hardware Acceleration APIs for Block Storage

Hardware Acceleration APIs enable the ESXi host to offload specific virtual machine and storage management operations to storage hardware. Use of these APIs significantly reduces the CPU overhead on the host.

Hardware acceleration is supported by block storage devices such as Fibre Channel and iSCSI devices.

Hardware acceleration for block storage devices supports the following array operations.

Array Operation	Use Cases
Full copy, also called clone blocks, copy offload, or XCOPY	Used by vSphere Storage vMotion, cloning VMs, and deploying a VM from a templates
Block zeroing, also called write same	Used when creating eager-zeroed thick virtual disks
Hardware-assisted locking, also called atomic test and set	Improves performance for VMFS metadata changes

With the storage hardware assistance, your host performs the listed operations faster and consumes less CPU, memory, and storage fabric bandwidth.

ESXi hardware acceleration for block storage devices supports the following array operations:

- Full copy: Enables the storage arrays to make full copies of data in the array without having the host read and write the data. This operation reduces the time and network load when cloning virtual machines, provisioning from a template, or migrating with vSphere Storage vMotion.
- Block zeroing: Enables storage arrays to zero out a large number of blocks to provide newly allocated storage, free of previously-written data. This operation reduces the time and network load when creating virtual machines and formatting virtual disks. This operation is activated when the Thick Provision Eager Zero option is used for a disk.
- Hardware-assisted locking: Supports discrete virtual machine locking without use of SCSI reservations. This operation allows disk locking per sector, unlike SCSI reservations, which locks the entire logical unit number (LUN).

The hardware acceleration features are enabled by default and are used on storage arrays that support these features. It is possible, but not recommended, to disable these features. Older arrays, or those with outdated firmware, might not support hardware acceleration. In such cases, the host falls back on older and more CPU-intensive techniques to achieve the same tasks.

For more information about vSphere Storage APIs - Array Integration, see VMware knowledge base article 1021976 at http://kb.vmware.com/kb/1021976.

https://t.me/learningnets

3-22 Hardware Acceleration for NAS

Hardware acceleration for NAS enables NAS arrays to integrate with vSphere to offload certain storage operations to the array, such as offline cloning. This integration reduces CPU overhead on the host.

Hardware acceleration for NAS supports the following NAS operations:

- Full file clone: The entire file, instead of file segments, is cloned.
- Reserve space: Space for a virtual disk is allocated in thick format.
- Native snapshot support: VMware Horizon® can offload the creation of linked clones to a NAS array.
- Extended file statistics

Hardware acceleration also allows your host to integrate with NAS devices and use several hardware operations that NAS storage provides:

- Full file clone: This operation is similar to VMFS block cloning, except that NAS devices clone entire files instead of file segments.
- Reserve space: Enables storage arrays to allocate space for a virtual disk file in thick format.

  Typically, when you create a virtual disk on an NFS datastore, the NAS server determines the allocation policy. The default allocation policy on most NAS servers is thin and does not guarantee backing storage to the file. However, the reserve space operation can instruct the NAS device to use vendor-specific mechanisms to reserve space for a virtual disk. As a result, you can create thick virtual disks on the NFS datastore.
- Native snapshot support: Enables View to offload creation of linked clones to a NAS array.
- Extended file statistics: Enables storage arrays to accurately report space utilization.

Hardware acceleration for NAS is implemented through vendor-specific NAS plug-ins. These plug-ins are typically created and supported by vendors and are distributed as vSphere installation bundles (VIBs) through a webpage.

https://t.me/learningnets

3-23 Array Thin Provisioning APIs

Array Thin Provisioning APIs enable the host to integrate with physical storage and become aware of space usage in thin-provisioned LUNs:

- A VMFS datastore that you deploy on a thin-provisioned LUN can detect only the logical size of the LUN.

  For example, if an array reports 2 TB of storage but provides only 1 TB, the datastore considers 2 TB to be the LUN's size.

- Using thin-provisioning integration, the host can perform these tasks:
  - Monitor the use of space on thin-provisioned LUNs to avoid running out of physical space.
  - Inform the array about datastore space that is freed when the following actions take place:
    - Virtual disks and VM files are deleted from the datastore.
    - Virtual disks and VM files are migrated off of the datastore by vSphere Storage vMotion.
    - VM snapshots are deleted and consolidated.

Traditional LUNs that arrays present to the ESXi host are thick-provisioned. The entire physical space needed to back each LUN is allocated in advance. ESXi also supports thin-provisioned LUNs. When a LUN is thin-provisioned, the storage array reports the LUN logical size, which might be larger than the real physical capacity backing that LUN.

With this integration, as the datastore expands, or if you use vSphere Storage vMotion to migrate virtual machines to a thin-provisioned LUN, the host communicates with the LUN and warns you about breaches in physical space and about out-of-space conditions.

No installation steps are required for the Array Thin Provisioning extensions. Array Thin Provisioning works on all VMFS5 and VMFS6 volumes. Device firmware enabled for this API is required to take advantage of the Array Thin Provisioning features. ESXi continuously checks for firmware that is compatible with Array Thin Provisioning. After the firmware is upgraded, ESXi starts using the Array Thin Provisioning features.

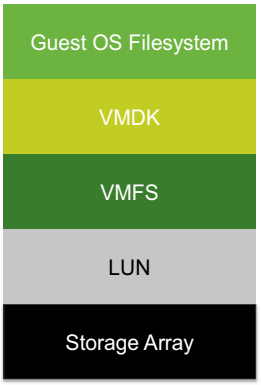
https://t.me/learningnets

3-24 About Space Reclamation (1)

In a virtualized environment, different levels of abstraction exist at the storage layer:

- The user files reside at the guest operating system layer. The guest operating system manages these files.
- The guest operating system resides in the VMDK file, which is managed by the VMFS datastore.
- The VMFS datastore resides on the LUN, which is presented as the storage device to the ESXi host.

Data deletion at one of the layers does not automatically translate into free space at the layers below it.



https://t.me/learningnets

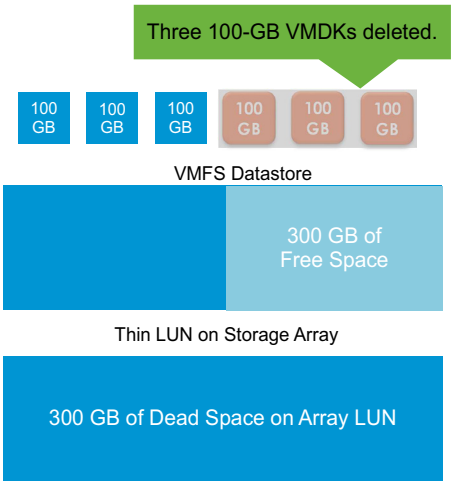
3-25 About Space Reclamation (2)

Data is deleted from the VMFS datastore in the following ways:

- Data is deleted within the guest operating system.
- A VMDK file is deleted from the datastore.

In either case, the space on the array is not freed. This space is referred to as dead space or dirty space.

The array cannot reuse this space unless it is specifically told to do so, such as by the SCSI UNMAP command.



Deleting or removing files from a VMFS datastore frees space within the file system. This free space is mapped to a storage device until the file system releases or unmaps it.

https://t.me/learningnets

3-26 Space Reclamation with VMFS

ESXi supports the reclamation of free space, which is also called the SCSI UNMAP command.

This command enables an ESXi host to inform the storage array to reclaim free space.

With a VMFS5 datastore, you can manually send the UNMAP command with the ESXCLI command.

With a VMFS6 datastore, the ESXi host automatically sends the UNMAP command to the storage array.

Feature	VMFS5	VMFS6
Automatic space reclamation	No	Yes
Manual space reclamation through ESXCLI	Yes	Yes

The UNMAP command enables an ESXi host to inform the storage array that the files of virtual machines have been moved or deleted from a thin-provisioned VMFS datastore. This notification enables the array to reclaim the freed blocks.

https://t.me/learningnets

3-27 Performing Manual Space Reclamation on a VMFS5 Datastore

To manually reclaim space on a VMFS5 datastore, you can use the `esxcli` command:

```
esxcli storage vmfs unmap
```

Command Option	Description
`-l | --volume-label=`*`volume_label`*	The label of the VMFS volume to unmap. A mandatory argument. Use either the `-l` or the `-u` option, not both.
`-u | --volume-uuid=`*`volume_uuid`*	The UUID of the VMFS volume to unmap. A mandatory argument. Use either the `-l` or the `-u` option, not both.
`-n | --reclaim-unit=`*`number`*	Number of VMFS blocks to unmap per iteration. An optional argument. The default value is 200.

You can automate the process by using this command in a VMware PowerCLI™ script and scheduling the script to run during off-hours.

On VMFS5 datastores, the `esxcli storage vmfs unmap` command must be run manually. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported.

This command can be run without any maintenance window:

- Reclaim size can be specified in blocks, instead of a percentage value, to make reclaim size more intuitive to calculate.
- Dead space is reclaimed in increments, instead of all at once, to avoid possible performance issues.

When you use the command, keep in mind that it might send many unmap requests at a time, which can lock some of the resources during this operation.

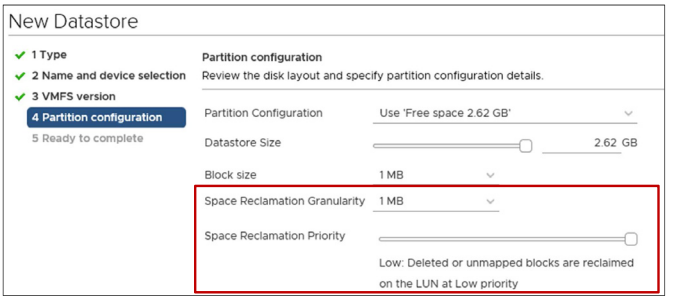
For more information about this command, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-storage-guide.pdf.

https://t.me/learningnets

3-28 Enabling Automatic Space Reclamation on a VMFS6 Datastore

When you create a VMFS6 datastore, you can modify the default settings for automatic asynchronous space reclamation:

- Space Reclamation Granularity: Defines the unit size for the UNMAP command. The default is 1 MB.
- Space Reclamation Priority: Controls the rate at which deleted or unmapped blocks are reclaimed on the LUNs backing the datastore. The default priority is Low, which uses an unmap rate of 25 MBps.



By default, the space reclamation granularity equals the block size, which is 1 MB. Storage sectors smaller than 1 MB are not reclaimed.

By default, the LUN performs the space reclamation operation at a low rate. You can also set the space reclamation priority to None to disable the operation for the datastore.

https://t.me/learningnets

3-29 Configuring the Automatic Unmap Rate

vSphere 6.7 allows you to set a fixed unmap rate for the datastore. The unmap rate is the rate at which automatic UNMAP commands are processed.

Change the unmap rate to match the storage array's capabilities.



https://t.me/learningnets

3-30 Space Reclamation Support for Guest Operating Systems (1)

Both VMFS5 and VMFS6 datastores provide support to the UNMAP command that originates from the guest operating system.

Some guest operating systems, such as Windows Server 2012 R2, have the ability to send UNMAP commands to the array.

The guest UNMAP commands are translated down to the array and space that can be reclaimed.

For thin-provisioned VMDK files, the virtual disk is shrunk down by the amount of space reclaimed.

ESXi supports the UNMAP commands issued directly from a guest operating system to reclaim storage space.

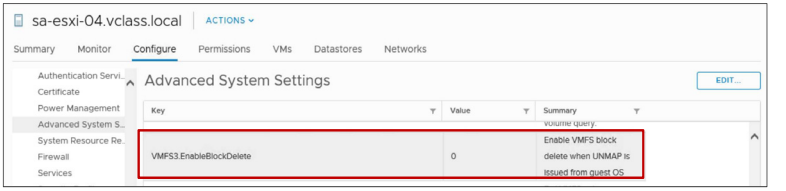
https://t.me/learningnets

3-31 Space Reclamation Support for Guest Operating Systems (2)

VMFS5 supports the automatic space reclamation requests for a limited number of guest operating systems.

To send unmap requests from the guest operating system to the array, the VM must meet the following prerequisites:

- The virtual disk must be thin-provisioned.
- The guest operating system must be able to identify the virtual disk as thin-provisioned.
- Virtual machine hardware must be at version 11 (ESXi 6.0) or later.
- The EnableBlockDelete advanced ESXi host setting must be set to **1**.



https://t.me/learningnets

3-32 Space Reclamation Support for Guest Operating Systems (3)

With VMFS6, use the Space Reclamation Priority setting on the datastore. The EnableBlockDelete option is ignored.

When you use space reclamation with VMFS6 datastores, the following considerations apply:

- VMFS6 processes the unmap request from the guest operating system only when the space to reclaim equals 1 MB or is a multiple of 1 MB.
- For VMs with snapshots in the default SEsparse format, VMFS6 supports the automatic space reclamation only on ESXi hosts version 6.7 or later.

VMFS6 generally supports automatic space reclamation requests that generate from the guest operating systems, and passes these requests to the array. Many guest operating systems can send the unmap command and do not require any additional configuration. The guest operating systems that do not support automatic unmaps might require user intervention. For information about guest operating systems that support automatic space reclamation on VMFS6, contact your vendor.

https://t.me/learningnets

3-33 Activity: vSphere Storage APIs - Array Integration (1)

In what ways can an ESXi host benefit from using storage devices that support vSphere Storage APIs – Array Integration? Select all that apply.

- ❑ The ESXi host can inform the storage array to reclaim free space in a VMFS datastore.
- ❑ The ESXi host can offload certain storage operations to a NAS array, such as offline cloning.
- ❑ The ESXi host can retrieve information about a storage array, such as storage topology and capabilities.
- ❑ The ESXi host can offload specific storage operations to a block storage device, such as Fibre Channel and iSCSI.

https://t.me/learningnets

3-34 Activity: vSphere Storage APIs - Array Integration (2)

In what ways can an ESXi host benefit from using storage devices that support vSphere Storage APIs – Array Integration? Select all that apply.

- ✓ The ESXi host can inform the storage array to reclaim free space in a VMFS datastore.
- ✓ The ESXi host can offload certain storage operations to a NAS array, such as offline cloning.
- ❑ The ESXi host can retrieve information about a storage array, such as storage topology and capabilities.
- ✓ The ESXi host can offload specific storage operations to a block storage device, such as Fibre Channel and iSCSI.

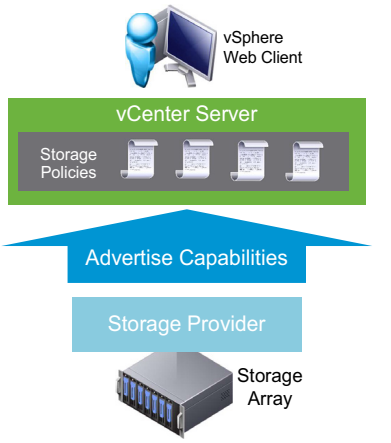
https://t.me/learningnets

3-35 About vSphere API for Storage Awareness

vSphere API for Storage Awareness enables a storage vendor to develop a software component, called a storage provider, for its storage arrays.

Through this API, a storage provider obtains information from the storage array about available storage topology, capabilities, and state.

vCenter Server connects to a storage provider. Information from this storage provider appears in vSphere Client.



In vCenter Server, vSphere administrators do not have access to the storage capabilities of the storage array on which their virtual machines are stored. Virtual machines are provisioned to a storage black box. All that the vSphere administrator sees of the storage is a LUN identifier, such as a Network Address Authority ID (NAA ID) or a T10 identifier.

A storage vendor can use VMware vSphere® API for Storage Awareness™ to provide information about its storage array.

The storage provider is written by the storage array vendor. The storage provider can exist on either the storage array processor or on a standalone host. The decision is made by the storage vendor.

The storage provider acts as a server in the vSphere environment. vCenter Server connects to the provider to obtain information about available storage topology, capabilities, and state. The information is viewed in vSphere Client. A storage provider can report information about one or more storage devices. A storage provider can support connections to a single vCenter Server instance or to multiple vCenter Server instances.

For information about the vSphere API for Storage Awareness program, go to https://developercenter.vmware.com/web/dp/other-programs/storage/vasa.

https://t.me/learningnets

3-36 Benefits of Storage Providers

A storage provider supplies capability information about storage configuration, status, and storage data services offered in your environment.

Storage providers benefit vSphere administrators by granting them additional functionality which enables them in the following ways:

- To be aware of the topology, capabilities, and state of the physical storage devices on which their virtual machines are located
- To monitor the health and usage of their physical storage devices
- To configure virtual machine storage policies that choose the correct storage in terms of space, performance, and service-level agreement requirements

vSphere Storage DRS uses the information supplied by a storage provider to make VM placement and migration decisions that are compatible with the storage system.

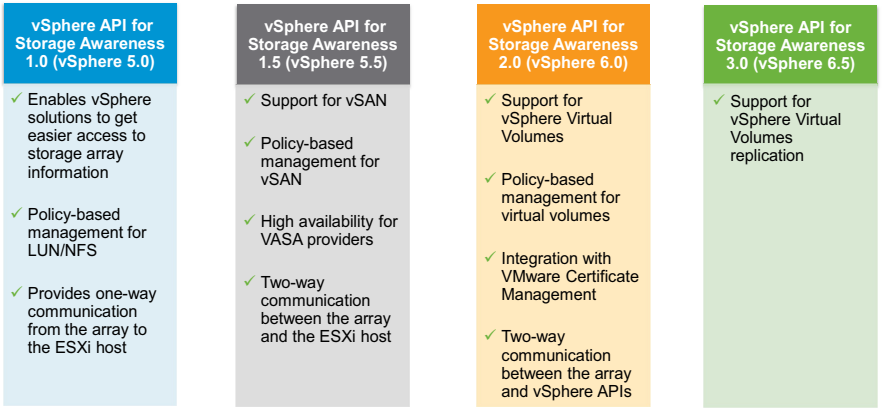
A storage provider supplies capability information that includes the following storage characteristics:

- Storage capabilities: Information about characteristics and services that the underlying storage offers, such as the number and type of spindles for a volume, the I/O operations in megabytes per second, the type of compression used, and whether thick-provisioned format is used.
- Storage status: Status of various storage entities, which also includes alarms and events for notifying of configuration changes.

https://t.me/learningnets

3-37 Comparison of Storage Provider Versions

The storage provider has evolved since its introduction in vSphere 5.0.



https://t.me/learningnets

3-38 Storage Provider Categories

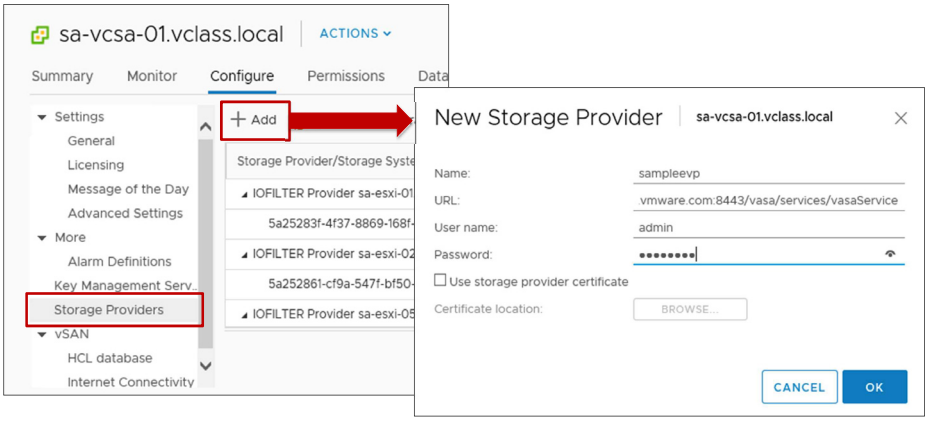
In general, vCenter Server and ESXi use storage providers to obtain information about storage configuration, status, and storage data services offered in your environment.

Category	Description	Examples
Persistence storage provider	• Manages arrays and storage abstractions • Can provide data services, such as replication	• vSAN storage provider • VMware vSphere® Virtual volumes™ storage provider
Data service provider (or I/O filter storage provider)	• Offers data services, such as host-based caching, compression, and encryption	• Storage providers that support vSphere APIs for I/O Filtering
Built-in storage provider	• Offered by VMware, and typically does not require registration	• Storage providers that support vSphere APIs for I/O Filtering, which are automatically registered
Third-party storage provider	• Offered by a third party, and typically requires registration	• vSphere Virtual Volumes storage provider • vSphere APIs for Storage Awareness storage providers

https://t.me/learningnets

3-39 Registering a Storage Provider

In vSphere Client, select **Configure** > **Storage Providers** to register and manage a storage provider.



The Storage Providers pane enables you to register a storage provider. All system storage capabilities that are presented by the storage provider appear in vSphere Client.

To register a storage provider, the storage vendor provides a URL, a login account, and a password. Users log in to the provider to get array information. vCenter Server must trust the provider host, so a security certificate from the provider must be installed on the vCenter Server system.

For information about registration procedures, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-storage-guide.pdf.pdf.

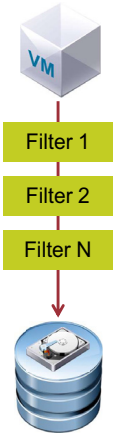
https://t.me/learningnets

3-40 About vSphere APIs for I/O Filtering

VAIO provides a framework for creating and implementing I/O filters in a virtual machine data stream:

- I/O filters intercept and modify virtual machine disk I/O within the VMkernel before the data is sent to physical storage.
- I/O filters are independent of storage topology.

VAIO enables VMware and third-party vendors to create data services such as caching and replication.



I/O filters can gain direct access to the virtual machine I/O path. You can enable the I/O filter for an individual virtual disk.

VMware offers certain categories of I/O filters. In addition, third-party vendors can create the I/O filters. Typically, these I/O filters are distributed as packages that provide an installer to deploy the filter components on vCenter Server and ESXi host clusters.

I/O filters can support all datastore types:

- VMFS
- NFS 3
- NFS 4.1
- VMware vSphere® Virtual Volumes™
- vSAN

https://t.me/learningnets

3-41 Types of I/O Filters

The types of I/O filters that can be applied to virtual machines are grouped into filter categories.

Category	Examples
Replication	Replicates all write I/O operations to an external target location, such as another host or cluster
Encryption	Provides encryption mechanisms for VMs
Caching	Implements a cache for virtual disk data
Storage I/O Control	Prioritizes storage I/O that is allocated to VMs during periods of I/O contention

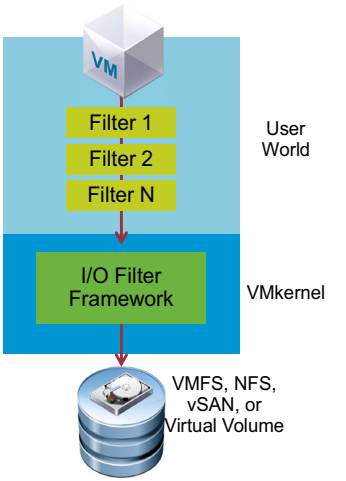
VMware provides certain categories of I/O filters that are installed on your ESXi hosts. In addition, VMware partners can create the I/O filters through the vSphere APIs for I/O Filtering developer program. The I/O filters can serve multiple purposes.

https://t.me/learningnets

3-42 How I/O Filters Work (1)

The I/O filter framework within the VMkernel is responsible for intercepting I/O and applying the filters associated with a virtual machine.

I/O filters are run in the virtual machine user world to ensure that any filter problems affect that virtual machine and not the entire host.



Filters operate by intercepting data within the VMkernel before the data is sent out to physical storage.

The I/O filter framework exists outside the virtual SCSI layer, which means that offline I/O to the disk (that is, modifications when a virtual machine is powered off) can also be intercepted. The framework defines which filters should be applied and in which order.

I/O filters run in the virtual machine's user world. A world is an execution context that is scheduled on a processor. A world is like a process in conventional operating systems.

https://t.me/learningnets

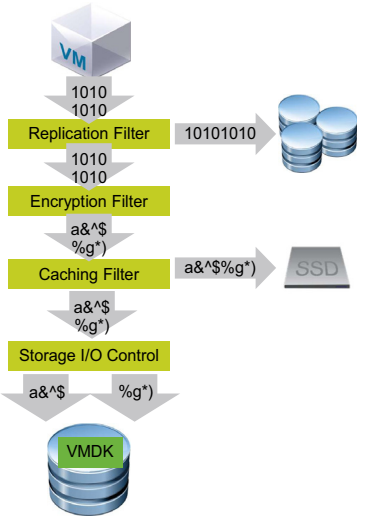
3-43 How I/O Filters Work (2)

The I/O filter framework applies I/O filters in the following order:

1. Replication
2. Encryption
3. Caching
4. Storage I/O Control

Data passes from the filters in that order before being forwarded to the storage on which the virtual machine files or objects reside.

You can install several filters from the same category on your ESXi host. However, you can have only one filter from the same category per virtual disk.



The order in which I/O passes through these filter categories is predefined. Understanding this order is important, especially if you plan to implement a replication filter and an encryption filter. Because the replication filter comes before the encryption filter, the data that is replicated is done in plain text. Thus, additional security measures must be taken to protect these transmissions and the data on the replication target.

https://t.me/learningnets

3-44 I/O Filter Providers

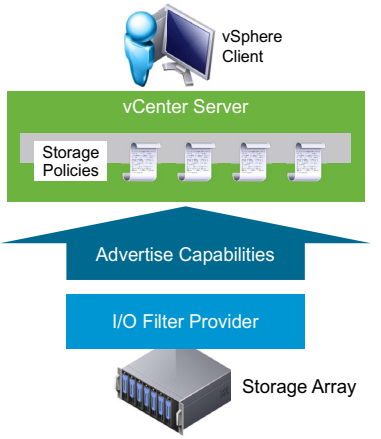
ESXi has the following preinstalled I/O filters:

- VM encryption
- Storage I/O Control

Third-party I/O filters must be installed according to the vendor instructions.

After I/O filters are installed on ESXi hosts, the I/O filter framework configures and registers a storage provider, also called an IOFilter provider, for each host.

The IOFilter provider advertises the new filter capabilities to vCenter Server. These capabilities can be used to build storage policies.



Storage providers for I/O filtering are software components that are offered by vSphere. They integrate with I/O filters and report data service capabilities that I/O filters support to vCenter Server.

If you use I/O filters provided by third parties, you must install the I/O filters in an ESXi host cluster. You cannot install the filters on selected hosts. The filter packages are typically distributed as VIBs which can include I/O filter daemons, CIM providers, and other associated components.

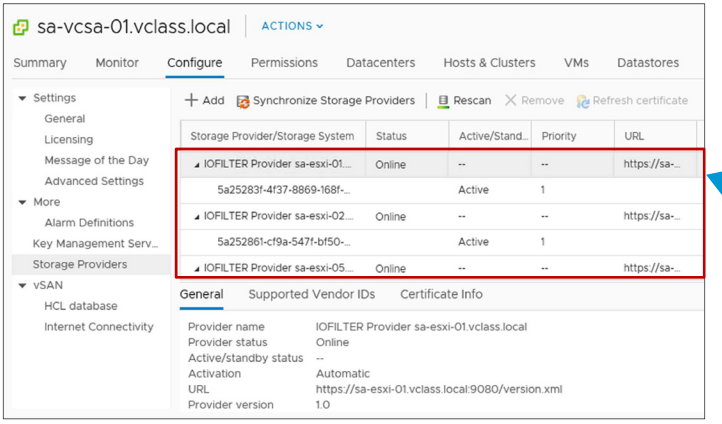
Typically, to deploy the filters, you run installers provided by vendors.

The capabilities populate the VM Storage Policies pane in vSphere Client and can be referenced in a virtual machine storage policy. You then apply this policy to virtual disks so that the I/O filters can process I/O for the disks.

https://t.me/learningnets

3-45 Displaying I/O Filter Provider Information

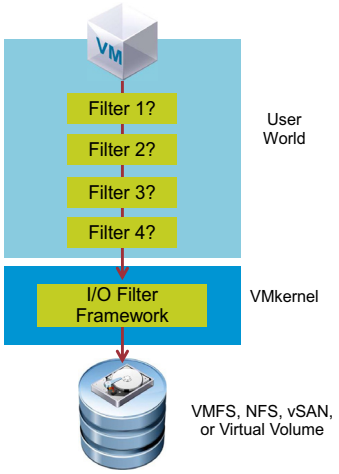
To view the I/O filter providers for each host, select the vCenter Server object in the inventory, and select **Configure > Storage Providers**.



https://t.me/learningnets

3-46 Activity: I/O Filters (1)

Using the diagram, identify the I/O filter types and the order in which these filter types are applied to the VM's I/O stream.

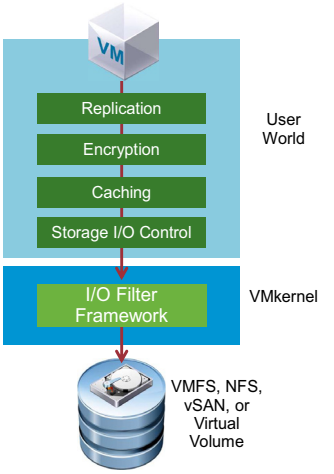


https://t.me/learningnets

3-47 Activity: I/O Filters (2)

The I/O filters are applied by the I/O filter framework in the following order:

1. Replication
2. Encryption
3. Caching
4. Storage I/O Control



https://t.me/learningnets

3-48 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe how vSphere Storage APIs - Array Integration helps storage arrays integrate with vSphere
- Explain how vSphere API for Storage Awareness can ensure that a VM's storage requirements are met
- Explain how VAIO enables vendors to create data services for VMs

https://t.me/learningnets



https://t.me/learningnets

3-50 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Explain storage policy-based management
- Configure and use virtual machine storage policies

https://t.me/learningnets

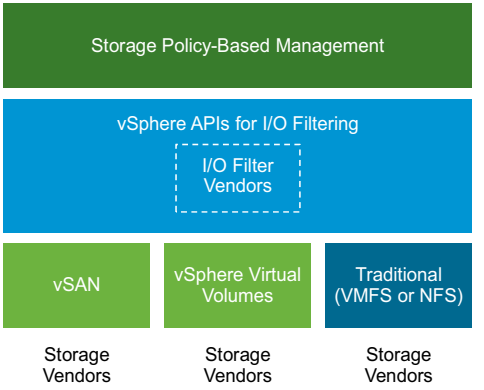
3-51 About Storage Policy-Based Management (1)

Storage policy-based management helps you ensure that VMs use storage that guarantees a specified level of capacity, performance, availability, redundancy, and so on.

Storage policy-based management abstracts storage and data services delivered by vSAN, vSphere Virtual Volumes, I/O filters, or traditional storage.

VMware partners and storage vendors can provide support for these storage and data services.

Rather than integrating with each individual vendor or type of storage and data service, storage policy-based management provides a universal framework for many types of storage entities.



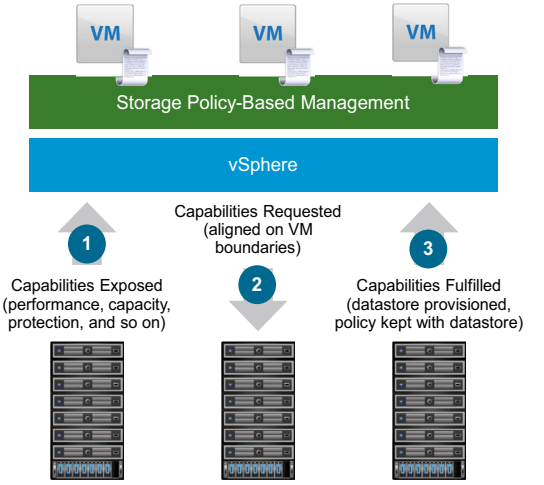
When an application is to be provisioned, vSphere leverages the Storage Policy-Based Management module to match the policy requirements against configured storage container capabilities to determine where the VMDK files should exist.

https://t.me/learningnets

3-52 About Storage Policy-Based Management (2)

Storage policy-based management offers the following mechanisms:

- Advertisement of storage capabilities and data services that storage arrays and other entities, such as I/O filters, offer.
- Bidirectional communication between ESXi and vCenter Server on one side, and storage arrays and entities on the other.
- Virtual machine provisioning based on virtual machine storage polices.



Virtual machine storage policies minimize the amount of storage planning that you must do for each virtual machine. For example, you can use virtual machine storage policies to create basic storage tiers. Datastores with similar capabilities are tagged to form gold, silver, and bronze tiers. Redundant, high-performance storage might be tagged as the gold tier. Nonredundant, medium-performance storage might be tagged as the bronze tier.

https://t.me/learningnets

3-53 About VM Storage Policies

VM storage policies are essential to VM provisioning.

A VM storage policy controls which type of storage is provided for the VM and how the VM is placed within storage. It also determines data services that the VM can use.

VM storage policies help you meet the following goals:

- Categorize datastores based on certain levels of service.
- Provision a VM's disks on the correct storage.

VM storage policies capture storage characteristics that VM home files and virtual disks require to run applications on a VM:

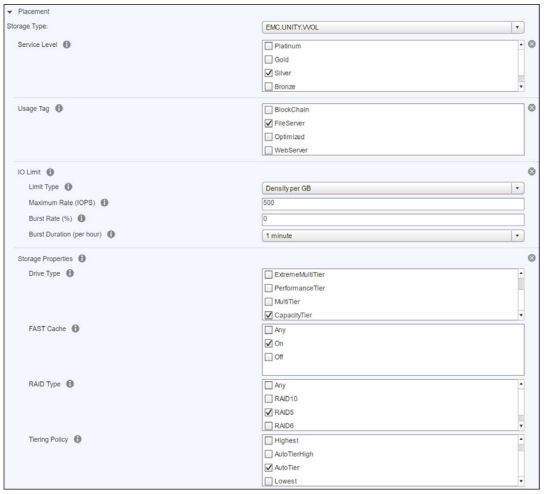
- These storage characteristics are defined in a set of rules.

Virtual machine storage policies can be used during the provisioning of a virtual machine to ensure that a virtual machine's disks are placed on the storage that is best for its situation. For example, virtual machine storage policies can help you ensure that the virtual machine running a critical I/O-intensive database is placed in the gold tier. Ideally, you want to create the best match of predefined virtual machine storage requirements with available physical storage properties.

https://t.me/learningnets

3-54 Storage Policy Example (1)

When you create a storage policy, you can choose one or more rules (which define storage characteristics) that represent service levels demanded by virtual machines.



Each array vendor has an identified namespace. The namespace contains the storage container identifiers and assigned capabilities. The capabilities appear as options in the client interface. In this example, vSphere Web Client is used.

https://t.me/learningnets

3-55 About Storage Policy Rules

Rules are the basic elements of a storage policy. Each rule is a statement that describes a single requirement for virtual machine storage and data services.

Category	Description	Examples
Capability-based placement rules	Specify a particular storage requirement for the VM, which guides the decision about VM placement.	• RAID level • Number of failures to tolerate
Tag-based placement rules	Reference datastore tags to define VM placement.	• Gold Tier tag • Silver Tier tag • Bronze Tier tag
Data service rules	Do not define storage placement and VM storage requirements but activate specific data services for the VM.	• Caching • Replication

Capability-based placement rules describe how the virtual machine storage objects are allocated within the datastore to receive the required level of service. For example, the rules can list virtual volumes as a destination and define the maximum recovery point objective for the virtual volume objects. Storage policy-based management finds the virtual volumes datastores that can match the rules and satisfy the storage requirements of the virtual machine.

Data service rules activate specific data services. Storage systems or other entities can provide these services. They can also be installed on your hosts and vCenter Server. You include the data service rules in the storage policy components.

Tag-based placement rules reference datastore tags. You can use the tag-based rules to fine-tune your virtual machine placement request further. For example, you can exclude datastores with the Palo Alto tag from the list of your virtual volumes datastores.

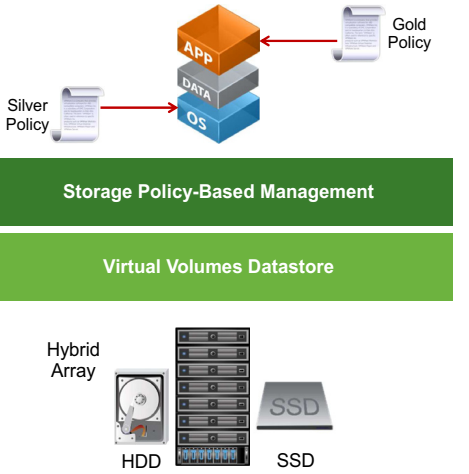
https://t.me/learningnets

3-56 Storage Policy Example (2)

VM storage policies ensure that VMs run on storage with the correct performance and availability characteristics.

For example, you want to place a database server's disks on appropriate storage:

- Assign the Silver policy to the operating system disk:
  - The Silver policy is set up to use only HDDs.
- Assign the Gold policy to the data disk:
  - The Gold policy uses only SSDs.



For more information about storage policy-based management, see https://blogs.vmware.com/virtualblocks/2017/01/16/understanding-storage-policy-based-management.

https://t.me/learningnets

3-57 About Storage Policy Components

A VM storage policy can include one or several reusable and interchangeable building blocks, called storage policy components.

Each component describes one type of service from a storage provider.

The services vary depending on the storage provider that you use but they generally belong to one of the following categories:

- Replication
- Encryption
- Caching
- Storage I/O Control

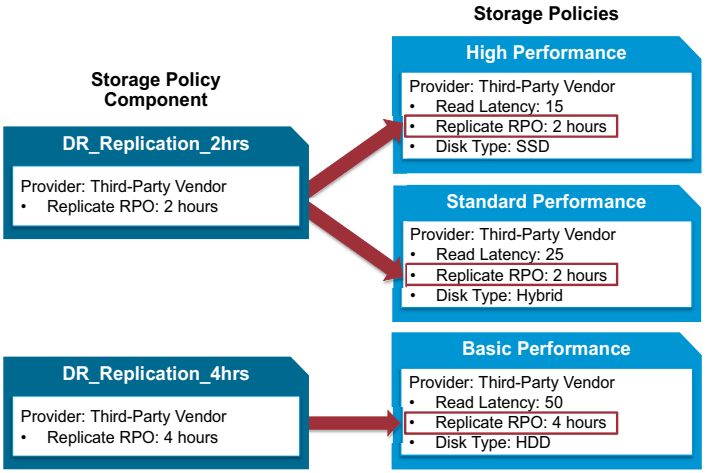
The service provider can be a storage system, an I/O filter, or another entity.

https://t.me/learningnets

3-58 Storage Policy Component Example

When you create the storage policy component, you define the rules for one specific type and grade of service.

You must add the component to the storage policy, then assign the policy to the VM. You cannot assign a component directly to a VM or virtual disk.



You can define the policy components in advance and associate them with multiple VM storage policies.

https://t.me/learningnets

3-59 Activity: Storage Policy Rules (1)

Data service rules are used to help storage policy-based management place VMs on the proper storage.

- ❑ True
- ❑ False

https://t.me/learningnets

3-60 Activity: Storage Policy Rules (2)

Data service rules are used to help storage policy-based management place VMs on the proper storage.

- ❑ True
- ✓ False

Capability-based and tag-based placement rules guide decisions about VM placement, not data service rules. Data service rules activate specific data services (such as caching or replication) for the VM.

https://t.me/learningnets

3-61 Creating and Managing Storage Policies: Overview

The entire process of creating and managing storage policies typically includes several steps:

1. Configure storage that will be used for storage policies.
2. Create storage policy components.
3. Create VM storage policies.
4. Apply the storage policy to the VM.
5. Check compliance for the VM storage policy.

https://t.me/learningnets

3-62 Step 1: Configure Storage That Will Be Used for Storage Policies

When creating storage policies, vSphere Client is populated with information based on the configuration of datastores and data services in your environment.

This information is obtained from storage providers and datastore tags.

For entities represented by storage providers, ensure that the appropriate provider is registered:

- Storage provider entities include vSAN, vSphere Virtual Volumes, and I/O filters.
- Some providers are self-registered. Other providers must be manually registered.

For datastores that are not represented by storage providers, you create datastore tags:

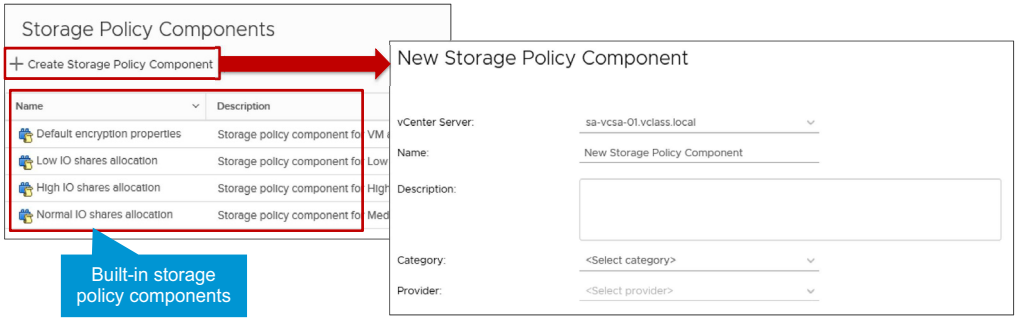
- Use tags to indicate a property that is not communicated through the storage providers, such as geographical location or administrative group.

https://t.me/learningnets

3-63 Step 2: Create Storage Policy Components

Define storage policy components in advance so that you can use them when you create a storage policy.

vCenter Server provides a few built-in storage policy components, specifically for encryption and VMware vSphere® Storage I/O Control.



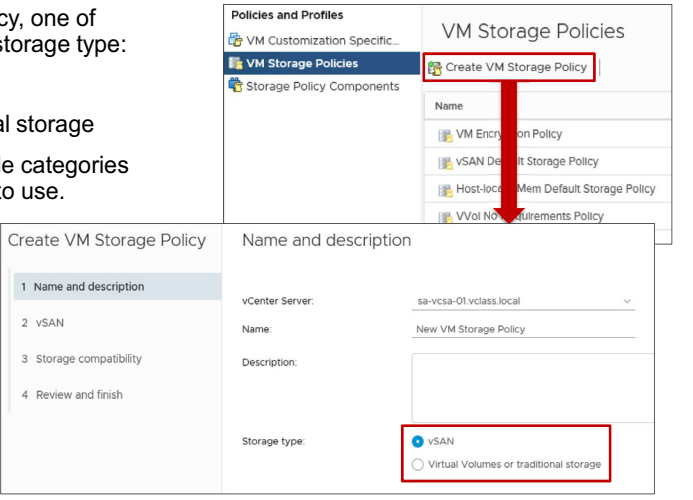
https://t.me/learningnets

3-64 Step 3: Create VM Storage Policies

When you create a storage policy, one of the first things to specify is the storage type:

- vSAN
- Virtual volumes or traditional storage

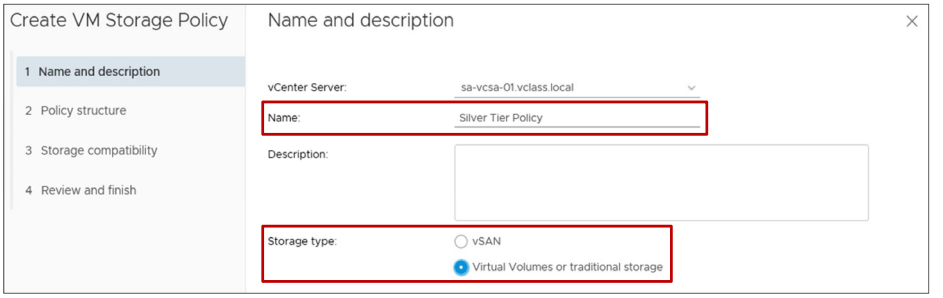
The storage type defines the rule categories available for the storage policy to use.



https://t.me/learningnets

3-65 Example: Creating a Silver Tier Storage Policy (1)

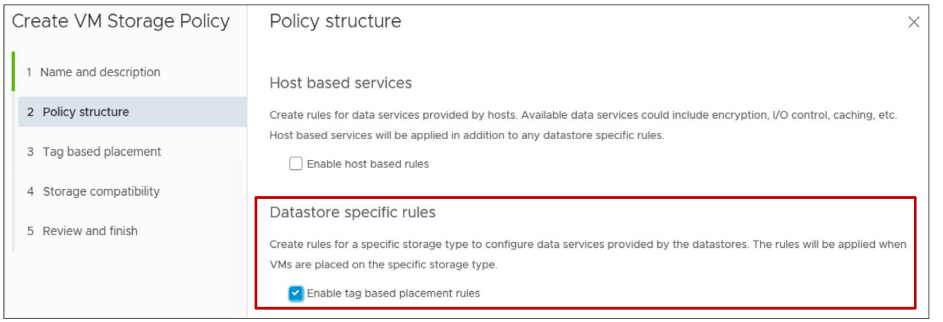
In this example, you create a storage policy named Silver Tier Policy. The storage type is traditional storage, that is, a VMFS datastore.



https://t.me/learningnets

3-66 Example: Creating a Silver Tier Storage Policy (2)

Silver Tier Policy uses a tag-based placement rule.

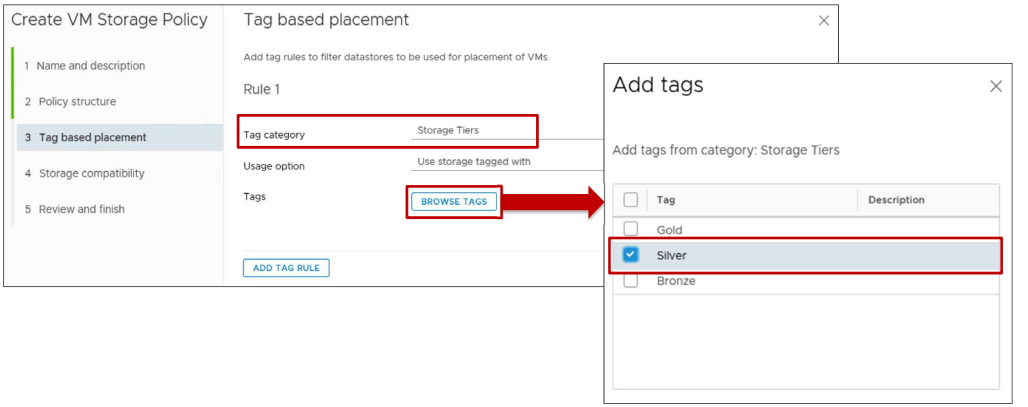


https://t.me/learningnets

3-67 Example: Creating a Silver Tier Storage Policy (3)

Choose the Storage Tiers tag category. Then, choose the tag named Silver.

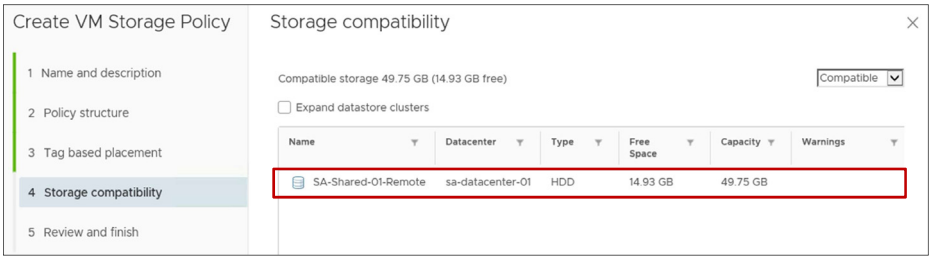
Silver Tier Policy will use a datastore that is associated with the tag Silver.



https://t.me/learningnets

3-68 Example: Creating a Silver Tier Storage Policy (4)

The Storage compatibility pane lists the datastores that are compatible with Silver Tier Policy.



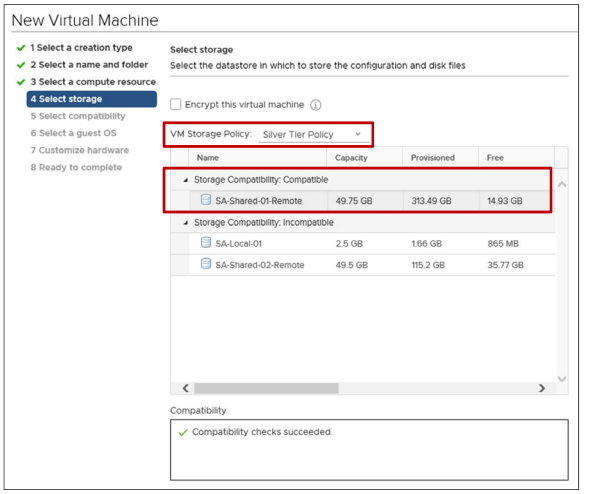
https://t.me/learningnets

3-69 Step 4: Apply the Storage Policy to the VM

After creating the storage policy, you can assign the policy to a VM when you create, clone, or migrate a VM.

If you assign the Silver Tier Policy to a VM, then the list of compatible datastores is displayed.

In this example, there is only one compatible datastore, named SA-Shared-01-Remote.



When you select a virtual machine storage policy, vSphere Client displays the datastores that are compatible with capabilities of that policy. You can then select a datastore or a datastore cluster. If you select a datastore that does not match the storage policy, vSphere Client shows that the virtual machine is using noncompliant storage.

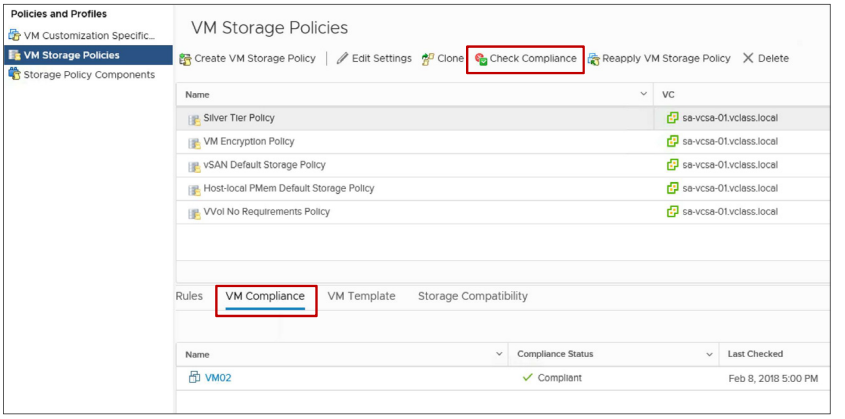
When a virtual machine storage policy is selected, datastores are divided into two categories: compatible and incompatible. You can still choose other datastores outside of the virtual machine storage policy, but these datastores put the virtual machine into a noncompliant state.

By using virtual machine storage policies, you can easily see which storage is compatible or incompatible. You can eliminate the need to ask the SAN administrator, or refer to a spreadsheet of NAA IDs, every time you deploy a virtual machine.

https://t.me/learningnets

3-70 Step 5: Check Compliance for the VM Storage Policy

You can check if virtual machines use datastores that are compliant with the storage policy.



Virtual machine storage policies can be used during the ongoing management of the virtual machines. You can periodically check whether a virtual machine has been migrated to or created on inappropriate storage, potentially making it noncompliant. Storage information can also be used to monitor the health and usage of the storage and report to you if the virtual machine's storage is not compliant.

https://t.me/learningnets

3-71 Lab 3: Policy-Based Storage

Use policy-based storage to create tiered storage

1. Add Datastores for Use by Policy-Based Storage
2. Use vSphere Storage vMotion to Migrate a Virtual Machine to the Gold Datastore
3. Configure Storage Tags
4. Create Virtual Machine Storage Policies
5. Assign Storage Policies to Virtual Machines

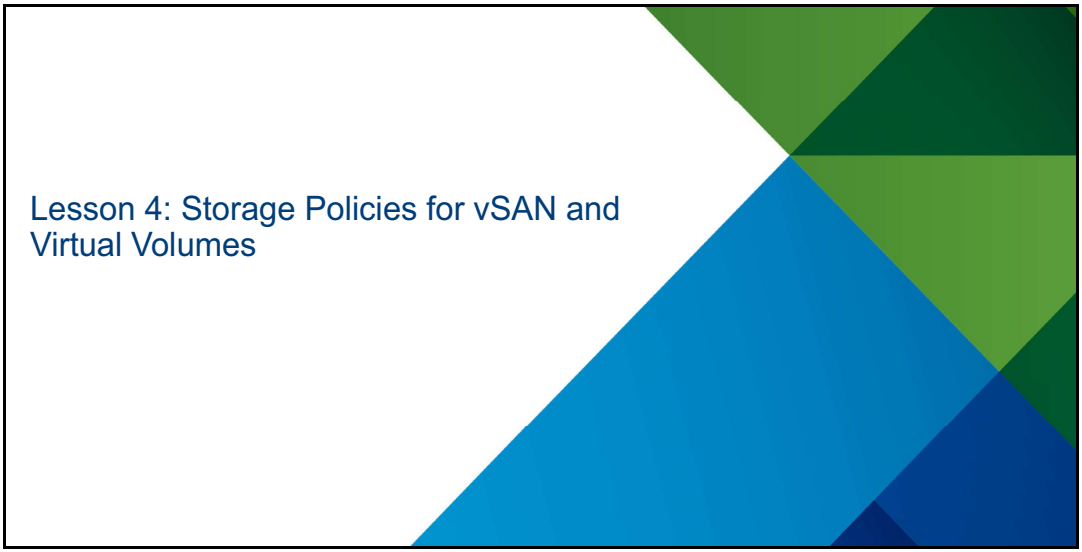
https://t.me/learningnets

3-72 Review of Learner Objectives

You should be able to meet the following objectives:

- Explain storage policy-based management
- Configure and use virtual machine storage policies

https://t.me/learningnets



https://t.me/learningnets

3-74 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

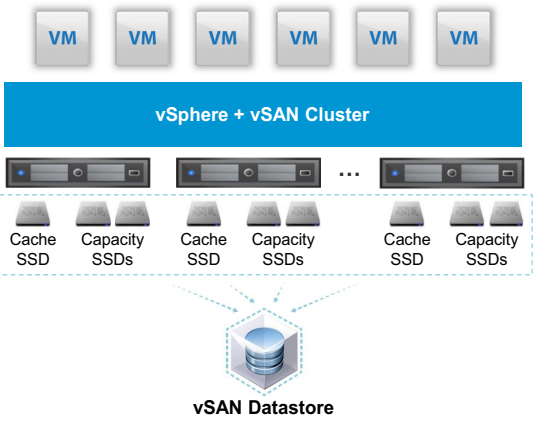
- Describe the vSAN architecture
- Create vSAN storage policies
- Describe the vSphere Virtual Volumes architecture
- Explain the rule set for creating vSphere Virtual Volumes storage policies

https://t.me/learningnets

3-75 About vSAN (1)

vSAN is software-defined storage built into vSphere that has the following features:

- Aggregates the direct-attached storage of ESXi hosts to create a storage pool that can be used by virtual machines
- Enables VM-centric data operations and policy-driven management



In a vSAN environment, ESXi hosts are configured to form a vSAN cluster. All of the ESXi hosts communicate through a dedicated network. At least three hosts in the cluster must have local storage devices. While not a best practice, hosts without local storage can share their compute resources and use the clustered storage resources.

vSAN datastores help administrators use software-defined storage in the following ways:

- Storage policy per virtual machine architecture: Multiple policies per datastore enable each virtual machine to have different storage configurations in terms of availability, performance, and capacity.
- vSphere and vCenter Server integration: vSAN capability is built in to the VMkernel and requires no appliance. You create a vSAN cluster in vCenter Server, as you would in a vSphere HA or vSphere DRS cluster.
- Scale-out storage: Up to 64 ESXi hosts can be in a cluster. You scale out by populating new hosts in the cluster.
- Built-in resiliency: A default policy mirrors all objects for virtual machines that are configured for vSAN.

https://t.me/learningnets

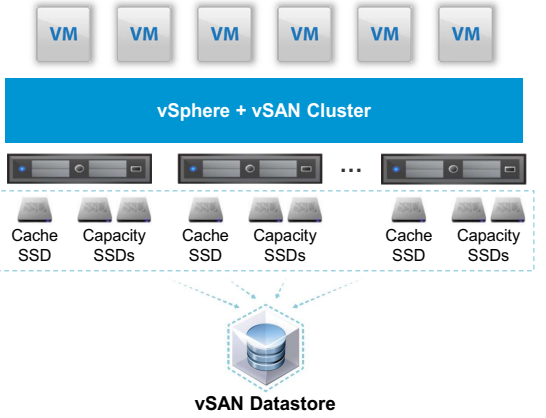
3-76 About vSAN (2)

vSAN is available in these architectures:

- All-Flash: Uses flash devices for the capacity and cache layers.
- Hybrid: Uses magnetic disks for the capacity layer and a flash device for the cache layer.

All-flash configurations provide additional features:

- Deduplication and compression
- RAID 5/6 (Erasure coding), in addition to RAID 0, 1, and 10 which is supported by both architectures



All-flash configurations can use features that are not available in hybrid configurations:

- Erasure coding: This feature provides the same levels of redundancy as mirroring (RAID 1), while consuming less storage capacity. Use of erasure coding reduces capacity consumption by as much as 50 percent versus mirroring at the same fault tolerance level.
- Deduplication and compression are space-saving features that can reduce the amount of storage consumption by as much as seven times. Deduplication and compression results vary based on the types of data stored on vSAN storage.

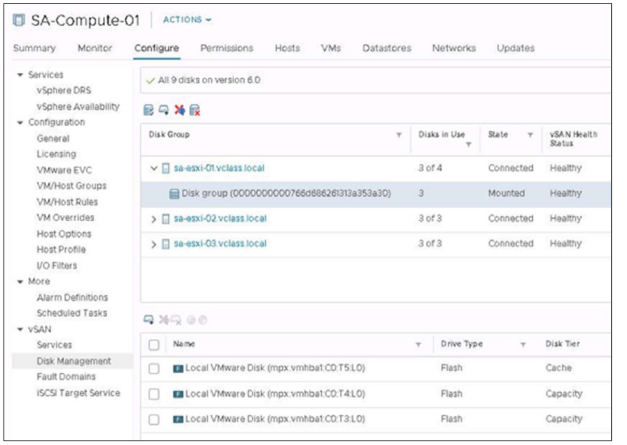
https://t.me/learningnets

3-77 vSAN Disk Architecture

Disks from up to 64 ESXi hosts are grouped to form a vSAN datastore.

Disk groups containing one cache device and up to seven capacity devices are created on each host to determine the capacity of the vSAN datastore.

Each host can support a maximum of five disk groups.



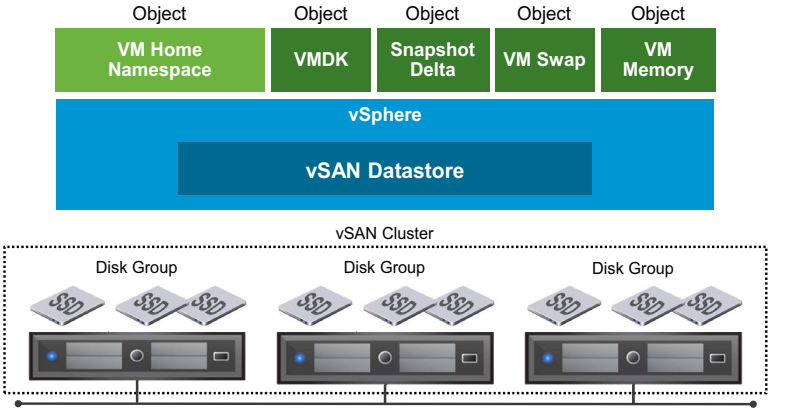
The local storage components of the ESXi hosts in the vSAN cluster are combined to create a vSAN datastore. Only one datastore can be created for each vSAN cluster.

Local storage is combined on each host to form a disk group. A disk group is a vSAN management construct that includes one cache device and one to seven capacity devices.

https://t.me/learningnets

3-78 Object-Based Storage

vSAN stores and manages data in the form of flexible data containers called objects. A virtual machine is a collection of objects.



The object store file system manages data as objects. Each object includes its own data, part of the metadata, and a unique ID. By using this unique ID, the object can be globally addressed by more than the filename and path. The use of objects enables a detailed level of configuration on the object level, for example, RAID type or disk usage at a level higher than the physical disk blocks.

When you provision a virtual machine on a vSAN datastore, a set of objects is created. These objects are of the following types:

- VM home namespace: Stores the virtual machine metadata (configuration files)
- VMDK: Virtual machine disk
- Snapshot delta: Created when snapshots of the virtual machine are taken
- VM swap: Virtual machine swap file, which is created when the virtual machine is powered on
- VM memory: A virtual machine's memory state when the virtual machine is suspended or when a snapshot is taken of a virtual machine and its memory state is preserved

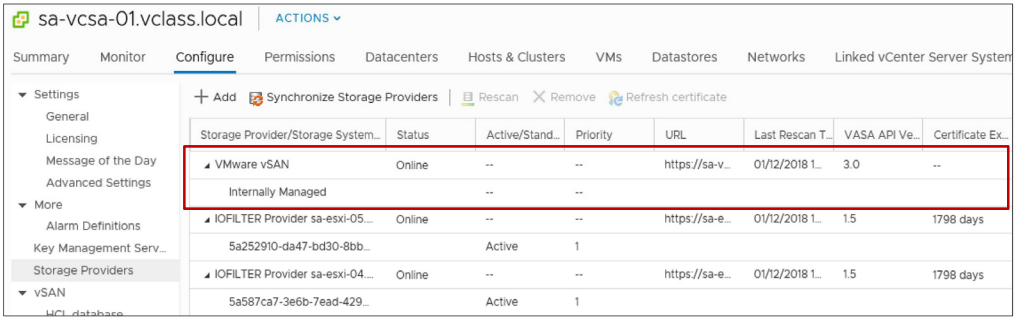
https://t.me/learningnets

3-79 vSAN Storage Providers

The vSAN datastore is represented by a storage provider.

Through the storage provider, the vSAN datastore advertises its capabilities, data services, and other characteristics in vSphere Client.

You use these characteristics to define capability-based placement rules for your storage policies.



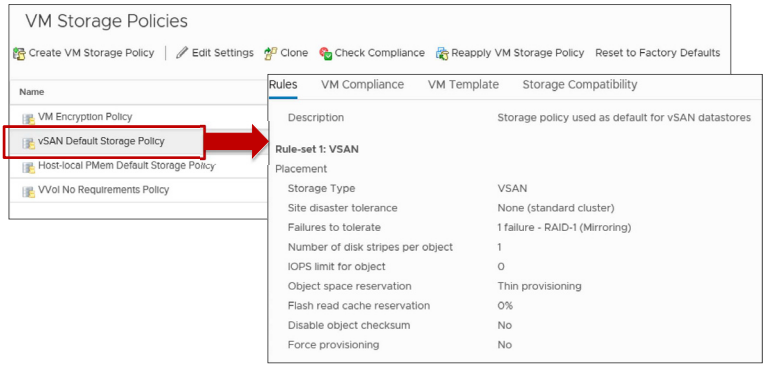
The vSAN storage providers report a set of underlying storage capabilities to vCenter Server. They also communicate with the vSAN layer to report the storage requirements of the virtual machines.

https://t.me/learningnets

3-80 vSAN Storage Policies

Multiple VM storage policies can be created for use by a single vSAN datastore:

- vSAN has a default VM storage policy.
- Custom storage policies can be created that leverage vSAN capabilities.
- The default vSAN storage policy is used, unless a different storage policy is selected.
- Use vSphere Client to view, create, and modify policies.

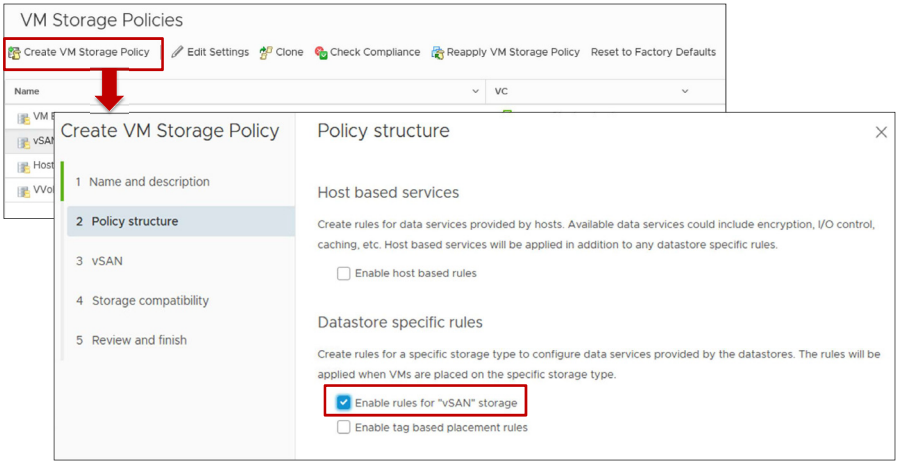


vSAN ensures that the virtual machines deployed to vSAN datastores are assigned at least one virtual machine storage policy. If a storage policy is not explicitly assigned to the virtual machine that is provisioned, a default storage policy is applied to the virtual machine from the datastore. If a custom policy has not been applied to the vSAN datastore, then the vSAN default storage policy is used.

https://t.me/learningnets

3-81 Creating vSAN Storage Policies

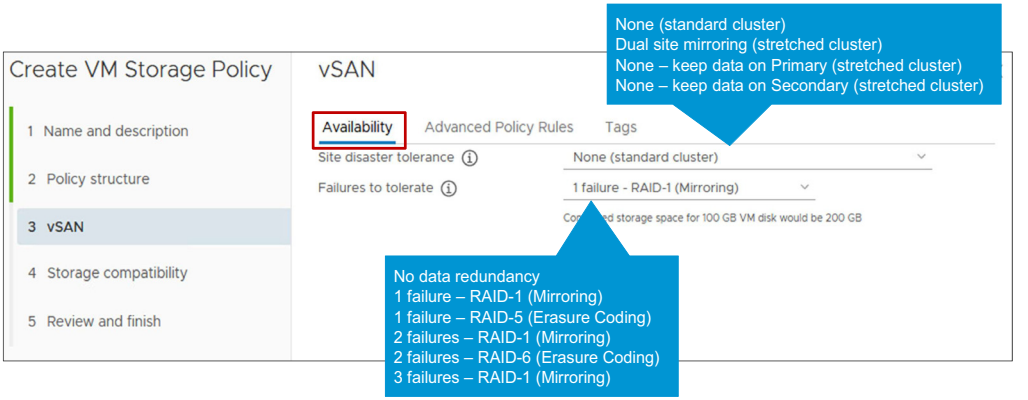
To create a vSAN storage policy, you must use rules for the vSAN storage type.



https://t.me/learningnets

3-82 vSAN Capabilities (1)

You define the rule set or set of capabilities. The first category is Availability, where you can define site disaster tolerance and number of failures to tolerate.

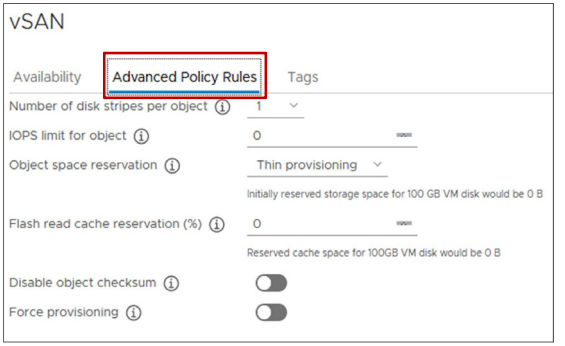


https://t.me/learningnets

3-83 vSAN Capabilities (2)

The other categories are Advanced Policy Rules and Tags:

- You can configure one or more advanced vSAN options.
- You can create a tag rule that uses storage associated with a certain tag.





For more information about vSAN advanced options, see Administering VMware vSAN at https://docs.vmware.com/en/VMware-vSphere/6.7/vsan-67-administration-guide.pdf.

https://t.me/learningnets

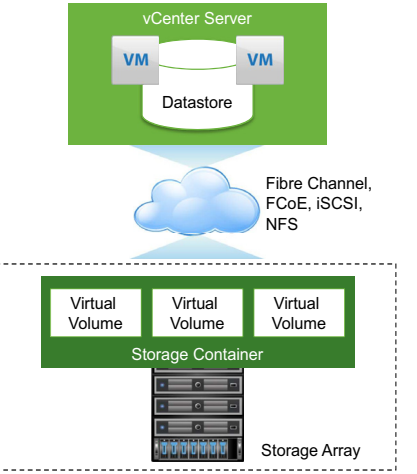
3-84 About vSphere Virtual Volumes

vSphere Virtual Volumes functionality changes the storage management paradigm from managing space inside datastores to managing abstract storage objects handled by storage arrays.

Abstract storage containers replace traditional storage volumes based on LUNs or NFS shares.

Rather than arranging storage around features of a storage system, vSphere Virtual Volumes arranges storage around the needs of individual virtual machines.

vSphere Virtual Volumes supports Fibre Channel, FCoE, iSCSI, and NFS.



Traditionally, vSphere storage management used a datastore-centric approach. With this approach, storage administrators and vSphere administrators discuss in advance the underlying storage requirements for virtual machines. The storage administrator then sets up LUNs or NFS shares and presents them to ESXi hosts. The vSphere administrator creates datastores based on LUNs or NFS shares and uses these datastores as virtual machine storage. Typically, the datastore is the lowest level at which data management occurs, from a storage perspective. However, a single datastore contains multiple virtual machines which might have different requirements.

With the traditional approach, differentiation per virtual machine is difficult.

The functionality of VMware vSphere® Virtual Volumes™ enables you to differentiate virtual machine services per application by offering a new approach to storage management.

https://t.me/learningnets

3-85 vSphere Virtual Volumes Architecture

Virtual and physical components interact with one another to provide vSphere Virtual Volumes functionality.

The following components constitute the vSphere Virtual Volumes architecture:

- vSphere Virtual Volumes storage providers
- Protocol endpoints
- Storage containers
- Virtual volumes datastores
- Virtual volumes

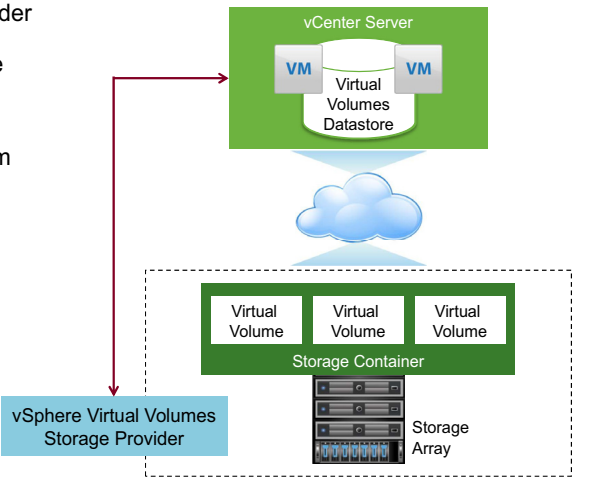
The architecture of vSphere Virtual Volumes exists on the storage itself, as well as on various components in the vSphere environment.

https://t.me/learningnets

3-86 vSphere Virtual Volumes Storage Providers

A vSphere Virtual Volumes storage provider is a software component that mediates communication between vSphere on one side and a storage system on the other.

The vSphere Virtual Volumes storage provider delivers storage information from the underlying storage container so that storage container capabilities appear in vCenter Server and vSphere Client.



A vSphere Virtual Volumes storage provider is implemented through vSphere API for Storage Awareness and is used to manage all aspects of vSphere Virtual Volumes storage. The storage provider integrates with the Storage Monitoring Service, shipped with vSphere, to communicate with vCenter Server and ESXi hosts.

The storage provider communicates virtual machine storage requirements, which you define in a storage policy, to the storage layer. This integration process ensures that a virtual volume created in the storage layer meets the requirements outlined in the policy.

Typically, vendors are responsible for supplying storage providers that can integrate with vSphere and provide support to virtual volumes. Every storage provider must be certified by VMware and properly deployed. For information about deploying the vSphere Virtual Volumes storage provider, contact your storage vendor.

For information about vSphere Virtual Volumes storage partners, see the vSphere Virtual Volumes product page at http://www.vmware.com/products/vsphere/features/virtual-volumes.html.

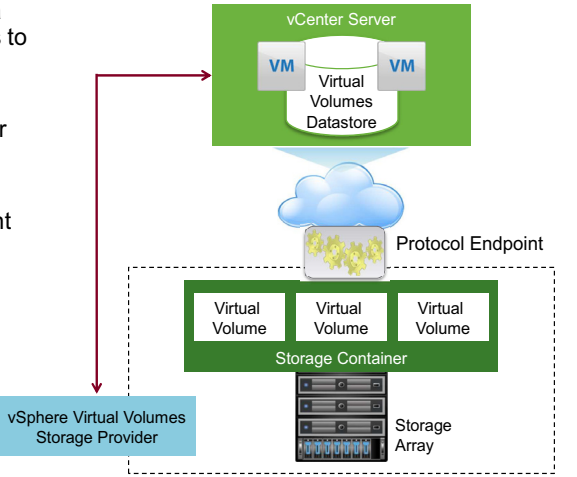
https://t.me/learningnets

3-87 Protocol Endpoints

ESXi uses protocol endpoints to establish a data path on demand from virtual machines to their respective virtual volumes.

Protocol endpoints are configured on the storage system side, and one or several per storage container.

Each virtual volume is bound to a specific protocol endpoint. A single protocol endpoint can connect to many virtual volumes.



Although storage systems manage all aspects of virtual volumes, ESXi hosts have no direct access to virtual volumes on the storage side. Instead, ESXi hosts use a logical I/O proxy, called the protocol endpoint, to communicate with virtual volumes and virtual disk files that virtual volumes encapsulate.

When the SCSI-based protocol is used, the protocol endpoint represents a LUN defined by a T10-based LUN World Wide Name. For the NFS protocol, the protocol endpoint is a mount point, such as an IP address or a DNS name and a share name.

When a virtual machine on the host performs an I/O operation, the protocol endpoint directs the I/O to the appropriate virtual volume. Typically, a storage system requires a very small number of protocol endpoints.

A storage administrator configures protocol endpoints. Protocol endpoints are a part of the physical storage fabric and are exported, along with associated storage containers, by the storage system through a storage provider. After you map a storage container to a virtual volumes datastore, protocol endpoints are discovered by ESXi and appear in vSphere Client. Protocol endpoints can also be discovered during a storage rescan.

https://t.me/learningnets

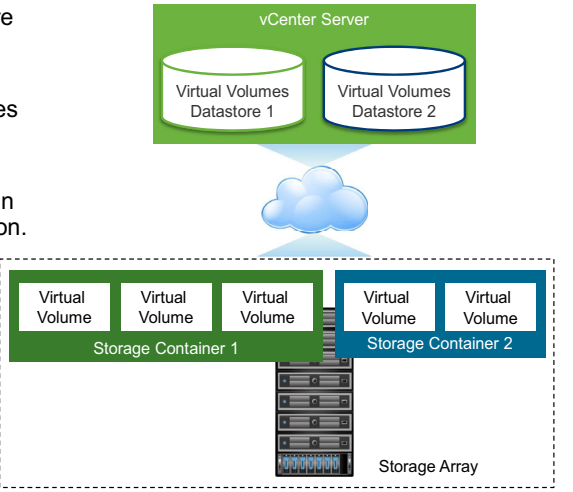
3-88 Storage Containers and vSphere Virtual Volumes Datastores

Virtual volumes are grouped into one or more storage containers.

A storage container is part of the underlying hardware and logically groups virtual volumes based on management and administrative needs.

For example, a storage container can contain all virtual volumes for a particular organization.

A virtual volumes datastore maps to a specific storage container.



Unlike traditional LUN-based and NFS-based vSphere storage, vSphere Virtual Volumes functionality does not require preconfigured volumes on a storage side. Instead, vSphere Virtual Volumes uses a storage container, which is a pool of raw storage capacity or an aggregation of storage capabilities that a storage system can provide to virtual volumes.

Typically, a storage administrator on the storage side defines storage containers. The number of storage containers and their capacity depend on a vendor-specific implementation, but at least one container for each storage system is required.

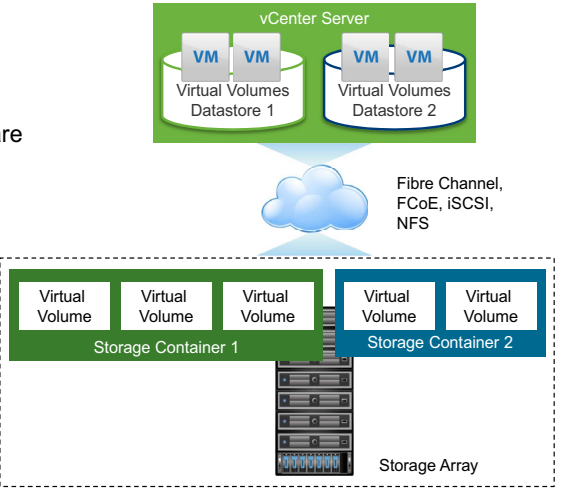
After vCenter Server discovers storage containers exported by storage systems, you must map a storage container to a virtual volumes datastore. The virtual volumes datastore that you create corresponds directly to the specific storage container and becomes the container's representation in vCenter Server and vSphere Client.

https://t.me/learningnets

3-89 Virtual Volumes

Virtual volumes are objects that are stored natively inside a storage system that is connected through an Ethernet or a Fibre Channel network.

Virtual volumes are not preprovisioned but are created automatically when you perform virtual management operations.



Virtual volumes are encapsulations of virtual machine files, virtual disks, and their derivatives.

Virtual volumes are exported as objects by a compliant storage system and are managed entirely by hardware on the storage side. Typically, a unique GUID identifies a virtual volume.

https://t.me/learningnets

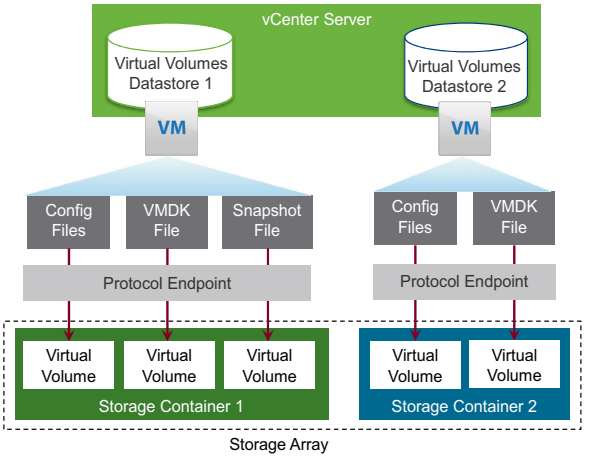
3-90 Mapping Virtual Machine Files to Virtual Volumes

The storage system creates virtual volumes for the core elements that make up the virtual machine:

- Configuration virtual volumes
- Data virtual volumes

Additional virtual volumes can be created for other VM components, such as snapshots and swap files.

vSphere Virtual Volumes maps VM files directly to virtual volumes, enabling vSphere to offload intensive storage operations, such as taking snapshots, cloning, and replication, to the storage system.



A data virtual volume corresponds directly to each virtual disk (`.vmdk`) file. Like virtual disk files on traditional datastores, virtual volumes are presented to virtual machines as SCSI disks.

A configuration virtual volume, or a home directory, represents a small directory that contains metadata files for a virtual machine. Metadata files include a `.vmx` file, descriptor files for virtual disks, log files, and so on. The configuration virtual volume is formatted with a file system. When ESXi uses the SCSI protocol to connect to storage, configuration virtual volumes are formatted with VMFS. With the NFS protocol, configuration virtual volumes are presented as an NFS directory.

Additional virtual volumes can be created for other virtual machine components and virtual disk derivatives, such as clones, snapshots, and replicas. These virtual volumes include a swap virtual volume to hold virtual machine swap files and a virtual memory volume to hold the contents of virtual machine memory for a snapshot.

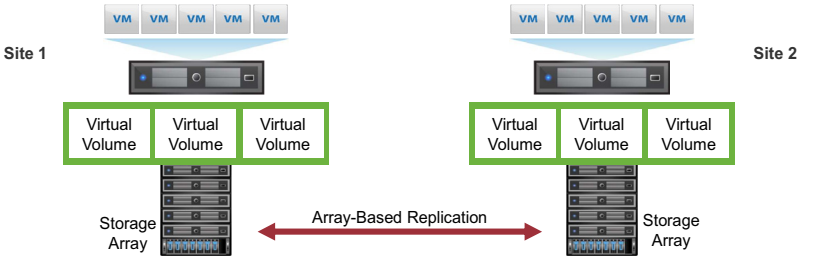
https://t.me/learningnets

3-91 vSphere Virtual Volumes Replication

vSphere Virtual Volumes supports replication. This feature enables you to offload replication of virtual machines to your storage array and use the full replication capabilities of the array.

vSphere Virtual Volumes replication:

- Is available with vSphere API for Storage Awareness version 3.0
- Is available for ESXi hosts, version 6.5 and later
- Requires the storage array to be replication-capable



Implementation of vSphere Virtual Volumes replication depends on your array and might be different for storage vendors. The following requirements generally apply to all vendors:

- The storage arrays that you use to implement replication must be compatible with virtual volumes.
- The arrays must integrate with the version of the storage (VASA) provider compatible with vSphere Virtual Volumes replication.
- The storage arrays must be replication capable and configured to use vendor-provided replication mechanisms. Typical configurations usually involve one or two replication targets. Any required configurations, such as pairing of the replicated site and the target site, must also be performed on the storage side.
- When applicable, replication groups and fault domains for virtual volumes must be preconfigured on the storage side.

https://t.me/learningnets

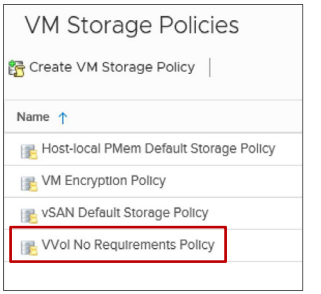
3-92 Virtual Volume Storage Policies

A VM that runs on a virtual volumes datastore requires a VM storage policy.

If you do not create the VM storage policy compatible with the virtual volumes datastore, the system uses the No Requirements default policy.

The No Requirements policy is a generic virtual volumes policy that contains no rules or storage specifications.

With the No Requirements policy, storage arrays can determine the optimum placement for the VM objects.



https://t.me/learningnets

3-93 Virtual Volume Storage Policy: Examples

By using different virtual volumes for different VM components, you can apply storage policies at the finest granularity level.

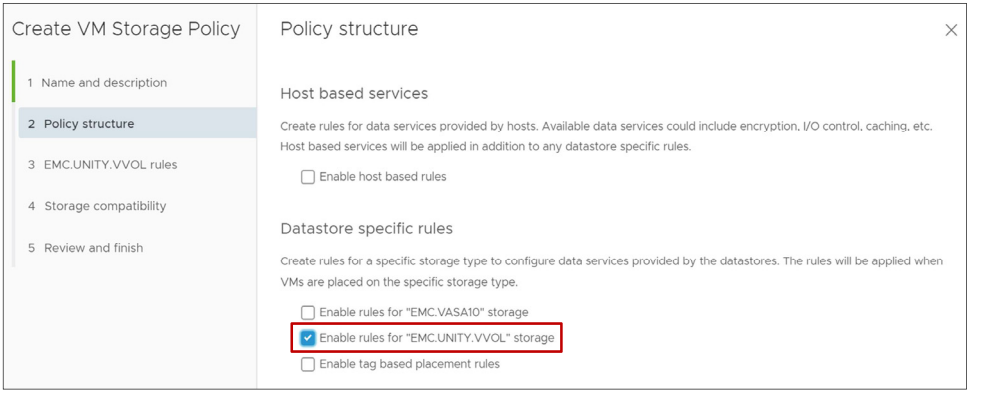
For example:

- A virtual volume that contains a virtual data disk can have a richer set of services than the virtual volume for the VM boot disk.
- Similarly, a snapshot virtual volume can use a different storage tier compared to a data virtual volume.
- If information about replication capabilities of the virtual volumes storage array appears in vCenter Server, you can activate replication for your VMs, for example, your business critical VMs.

https://t.me/learningnets

3-94 Creating Virtual Volumes Storage Policies (1)

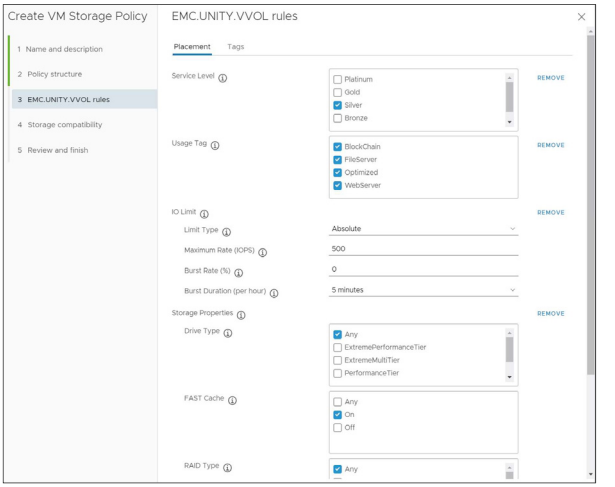
A virtual volumes storage policy can include host-based service rules and datastore-specific rules.



https://t.me/learningnets

3-95 Creating Virtual Volumes Storage Policies (2)

In this example, datastore-specific rules are listed for the EMC Unity virtual volumes storage.



https://t.me/learningnets

3-96 Activity: Displaying Storage Capabilities (1)

If the capabilities of your virtual volumes storage array do not appear in the storage policy interface in vSphere Client, what should you check?

- ❑ Verify that tags are created for the virtual volumes storage array.
- ❑ Verify that the virtual volumes storage policy components are preconfigured.
- ❑ Verify that the virtual volumes storage provider is registered with vCenter Server.
- ❑ Verify that the Virtual Volumes No Requirements storage policy has not been deleted.

https://t.me/learningnets

3-97 Activity: Displaying Storage Capabilities (2)

If the capabilities of your virtual volumes storage array do not appear in the storage policy interface in vSphere Client, what should you check?

- ❑ Verify that tags are created for the virtual volumes storage array.
- ❑ Verify that the virtual volumes storage policy components are preconfigured.
- ✓ Verify that the virtual volumes storage provider is registered with vCenter Server.
- ❑ Verify that the Virtual Volumes No Requirements storage policy has not been deleted.

https://t.me/learningnets

3-98 Lab 4: Creating vSAN Storage Policies

Create and review vSAN storage policies

1. Examine the Default Storage Policy
2. Create a Custom Policy with No Failure Tolerance
3. Assign the Custom Policy to a Virtual Machine
4. Make the Virtual Machine Compliant
5. Create an Invalid Storage Policy

https://t.me/learningnets

3-99 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe the vSAN architecture
- Create vSAN storage policies
- Describe the vSphere Virtual Volumes architecture
- Explain the rule set for creating Virtual Volumes storage policies

https://t.me/learningnets



https://t.me/learningnets

3-101 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

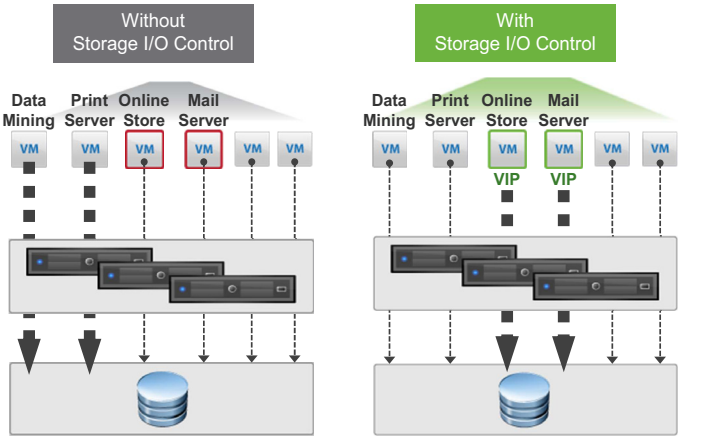
- Describe how Storage I/O Control balances the I/O load on a datastore
- Configure Storage I/O Control

https://t.me/learningnets

3-102 About Storage I/O Control

Storage I/O Control enables cluster-wide storage I/O prioritization, which enables better workload consolidation and helps reduce extra costs associated with overprovisioning.

With Storage I/O Control, you can balance I/O load in a datastore cluster that is enabled for vSphere Storage DRS.



During High I/O from a Noncritical Application

VMware vSphere® Storage I/O Control extends the constructs of shares, limits, and reservations to handle storage I/O resources. Storage I/O Control is a proportional-share, I/O operations per second (IOPS) scheduler that, under contention, throttles IOPS. You can control the amount of storage I/O that is allocated to virtual machines during periods of I/O congestion. Controlling storage I/O ensures that more important virtual machines get preference over less important virtual machines for I/O resource allocation.

You can use Storage I/O Control with or without vSphere Storage DRS. Two thresholds exist: one for Storage I/O Control and one for vSphere Storage DRS. For vSphere Storage DRS, latency statistics are gathered by Storage I/O Control for an ESXi host, sent to vCenter Server, and stored in the vCenter Server database. With these statistics, vSphere Storage DRS can decide whether or not a virtual machine should be migrated to another datastore.

For more information about Storage I/O Control, see *vSphere Resource Management* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-resource-management-guide.pdf.

https://t.me/learningnets

3-103 Enabling Storage I/O Control

With Storage I/O Control, you can ensure that the most important virtual machines get adequate I/O resources even during times of congestion.

Storage I/O Control is disabled by default. When Storage I/O Control is enabled on a datastore, all VMs have a default of 1000 disk shares, no IOPS limit, and no reservation.

Define the number of shares, an upper limit of I/O operations per second (IOPS), and the number of reserved IOPS for virtual machines.

Examples: Two Virtual Machines Running Iometer

	IOPS	Iometer Latency
VM1	1,500	20 ms
VM2	1,500	21 ms

This table compares VM1 and VM2 using the default settings (1000 shares and no IOPS limit). No shares, limits, or reservations are set.

	IOPS	Iometer Latency
VM1	1,080	31 ms
VM2	1,900	16 ms

This table compares VM1 and VM2. VM1 still uses 1000 shares and VM2 uses 2000 shares. Shares or limits might be set, but reservations are not set.

Storage I/O Control provides quality-of-service capabilities for storage I/O in the form of I/O shares, limits, and reservations that are enforced across all virtual machines accessing a datastore, regardless of which host they are running on.

When you enable Storage I/O Control on a datastore, ESXi begins to monitor the device latency that hosts observe when communicating with that datastore. When device latency exceeds a threshold, the datastore is considered to be congested, and each virtual machine that accesses that datastore is allocated I/O resources in proportion to its shares.

When you allocate storage I/O resources, you can limit the IOPS that are allowed for a virtual machine. By default, the number of IOPS allowed for a virtual machine is unlimited. If the limit that you want to set for a virtual machine is in megabytes per second instead of IOPS, you can convert megabytes per second into IOPS based on the typical I/O size for that virtual machine. For example, a backup application has a typical I/O size of 64 KB. To restrict a backup application to 10 MB per second, set a limit of 160 IOPS: 10 MB per second divided by 64 KB I/O size = 160 IOPS.

On the slide, virtual machines VM1 and VM2 are running an I/O load generator called Iometer. Each virtual machine is running on a different host, but they are running the same type of workload which is 16 KB random reads. The shares of VM2 are set to twice as many shares as VM1, which implies that VM2 is more important than VM1. With Storage I/O Control disabled, the IOPS and the I/O latency that each virtual machine achieves are identical. However, with Storage I/O Control enabled, the IOPS achieved by the virtual machine with more shares (VM2) is greater than the IOPS of VM1. The example assumes that each virtual machine is running enough load to cause a bottleneck on the datastore.

https://t.me/learningnets

When you enable Storage I/O Control, ESXi monitors datastore latency and throttles the I/O load if the datastore average latency exceeds the threshold.

By default, Storage I/O Control uses an injector-based model to automatically detect the latency threshold.

The benefit of using the injector-based model is that Storage I/O Control determines the best threshold for a datastore.

Storage I/O Control is disabled by default when new datastores are configured. When Storage I/O Control is enabled, performance statistics are collected and used to improve vSphere Storage DRS behavior.

Storage I/O Control automatically determines the optimal latency threshold by using injector-based models. However, you can also override this threshold by setting a specific latency value. This default latency setting might be fine for some storage devices, but other devices might reach their latency threshold well before or after the default setting is reached.

For example, solid-state drives (SSDs) typically reach their contention point sooner than the default setting protects against. Because not all devices are equal, the injector model is the preferred option.

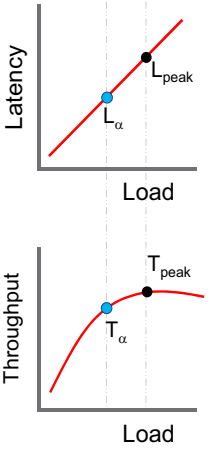
https://t.me/learningnets

3-105 Automatic Threshold Detection

With the injector-based model, Storage I/O Control determines the peak throughput of the device and sets the threshold to 90 percent of that value:

- Automatic threshold detection works well when a range of disk arrays and datastores are configured.
- The threshold varies according to the performance of each datastore.

You can also manually set the threshold value for a datastore. If this option is selected, the default latency setting is 30 ms.



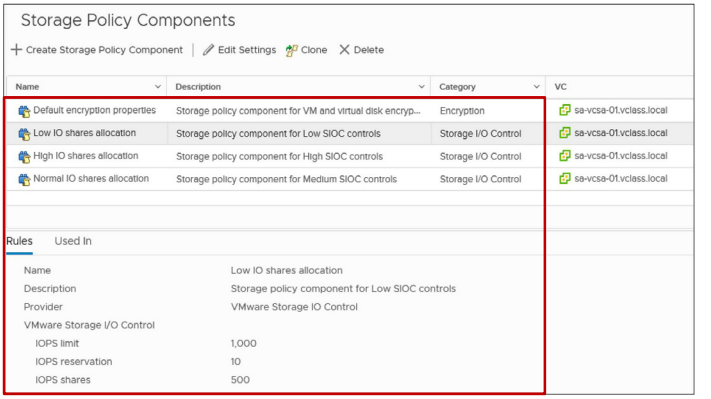
The I/O injector model determines the peak throughput of a datastore. The resulting peak throughput measurement can be used to determine the peak latency of a datastore.

https://t.me/learningnets

3-106 Storage Policy Components for Storage I/O Control

Storage I/O Control is implemented as an I/O filter. The IOFilter provider advertises Storage I/O Control capabilities to vCenter Server.

Built-in storage policy components exist that define Storage I/O Control capabilities.

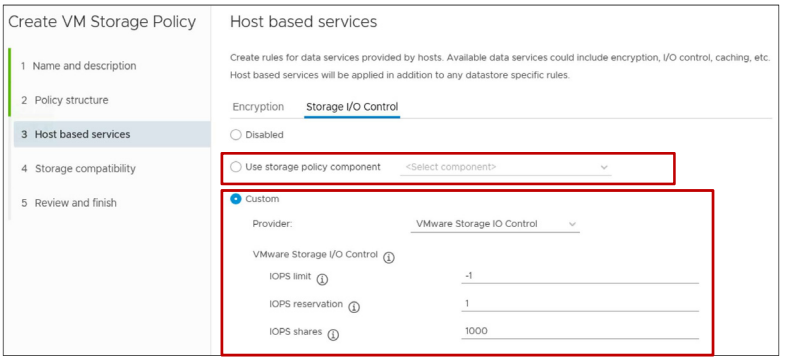


https://t.me/learningnets

3-107 Configuring Storage I/O Control

You can use a storage policy component for Storage I/O Control. Or, you can create a custom storage policy component, which defines the IOPS limit, IOPS reservation, and IOPS shares.

You then use the component in the storage policy and assign the policy to a VM.



https://t.me/learningnets

3-108 Storage I/O Control Requirements

Storage I/O Control requirements are as follows:

- Storage I/O Control is supported only on datastores that are managed by a single vCenter Server system.
- Storage I/O Control is supported for Fibre Channel, iSCSI, and NFS storage:
  - Raw device mapping is not supported.
  - vSAN and virtual volumes datastores are not supported.
- Storage I/O Control does not support datastores with multiple extents.
- If your datastores are backed by arrays with automated storage tiering capabilities, your array must be certified as compatible with Storage I/O Control.

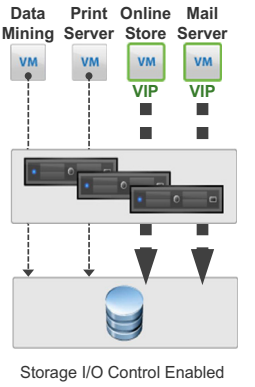
Automated storage tiering is the ability of an array (or a group of arrays) to migrate LUNs, virtual volumes, or parts of LUNs or virtual volumes to different types of storage media (SSD, Fibre Channel, SAS, SATA, and so on) based on user-set policies and current I/O patterns. No special certification is required for arrays that do not have these automatic migration or tiering features, including those that provide the ability to manually migrate data between different types of storage media.

To verify that your automated tiered storage array is certified as compatible with Storage I/O Control, see the VMware Compatibility Guide at http://www.vmware.com/resources/compatibility.

https://t.me/learningnets

3-109 Activity: Storage I/O Control (1)

On a datastore with Storage I/O Control enabled, list the steps that you would perform to ensure that the online store VM and the mail server VM have higher IOPS than the other VMs during periods of contention.



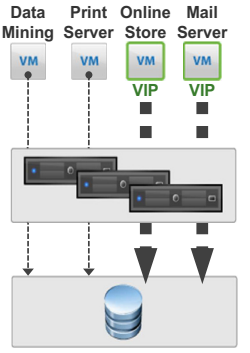
Storage I/O Control Enabled

https://t.me/learningnets

3-110 Activity: Storage I/O Control (2)

The following steps describe one way to ensure that the online store VM and mail server VM have higher throughput than other VMs during contention:

1. Create a storage policy that uses the predefined storage policy component called High IO Shares Allocation. High IO Shares Allocation has a disk shares value of 2000.
2. Assign the storage policy to only the online store VM and the mail server VM. All other VMs on the datastore use the default of 1000 disk shares.



Storage I/O Control Enabled

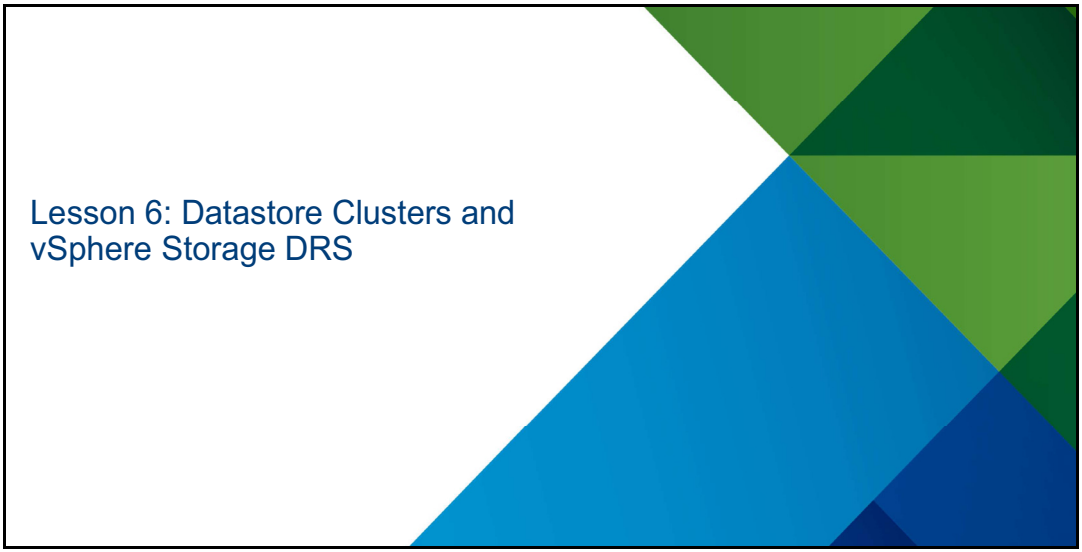
https://t.me/learningnets

3-111 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe how Storage I/O Control balances the I/O load on a datastore
- Configure Storage I/O Control

https://t.me/learningnets



https://t.me/learningnets

3-113 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

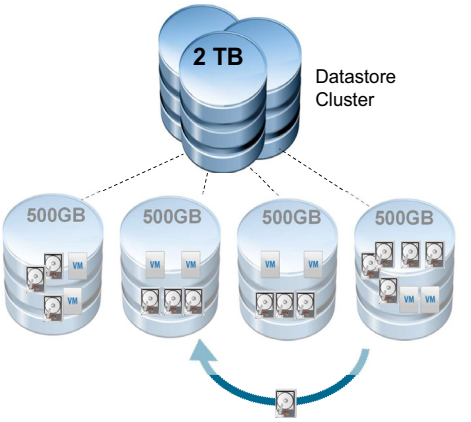
- Create a datastore cluster
- Configure vSphere Storage DRS
- Explain how Storage I/O Control and vSphere Storage DRS complement each other

https://t.me/learningnets

3-114 About Datastore Clusters

A datastore cluster is a collection of datastores that are grouped together without functioning together.

A datastore cluster enabled for vSphere Storage DRS is a collection of datastores working together to balance capacity and I/O latency.



The datastore cluster serves as a container or folder. You can store datastores in the container but the datastores work as separate entities.

A datastore cluster that is enabled for vSphere Storage DRS is a collection of datastores designed to work as a single unit.

https://t.me/learningnets

3-115 Datastore Cluster Requirements

Datastores and hosts that are associated with a datastore cluster must meet certain requirements to successfully use datastore cluster features.

Follow these guidelines when you create a datastore cluster:

- Datastore clusters must contain similar or interchangeable datastores, with the following exceptions:
  - All datastores in a datastore cluster must be in the same format, either VMFS or NFS.
  - vSAN and virtual volumes datastores are not supported in a datastore cluster.
  - Replicated datastores cannot be combined with nonreplicated datastores in the same datastore cluster enabled for vSphere Storage DRS:
    - This applies only to storage arrays that do not support vSphere API for Storage Awareness.
- Datastores shared across multiple data centers cannot be included in a datastore cluster.
- As a best practice, do not include datastores that have hardware acceleration enabled in the same datastore cluster as datastores that do not have hardware acceleration enabled.

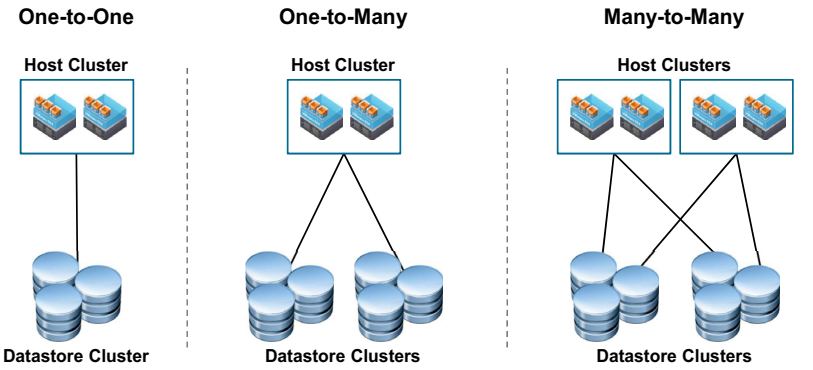
A datastore cluster can contain a mix of datastores of different sizes and I/O capacities, as well as datastores from different arrays and vendors. However, LUNs with different performance characteristics can cause performance problems.

vSphere Storage DRS cannot move virtual machines between NFS and VMFS datastores.

https://t.me/learningnets

3-116 Host Cluster to Datastore Cluster Relationships

The relationship between a vSphere host cluster and a datastore cluster can be one-to-one, one-to-many, or many-to-many.



Host clusters and datastore clusters can coexist in the virtual infrastructure. A host cluster is a vSphere DRS or a vSphere HA cluster.

Load balancing by vSphere DRS and vSphere Storage DRS can occur at the same time. vSphere DRS balances virtual machines across hosts based on CPU usage, memory usage, and network usage (the latter if Network I/O Control version 3 is used). vSphere Storage DRS load-balances virtual machines across storage, based on storage capacity and IOPS.

A standalone host (one that is not part of a host cluster) can also use a datastore cluster.

https://t.me/learningnets

3-117 vSphere Storage DRS Overview

vSphere Storage DRS enables you to manage the aggregated resources of a datastore cluster.

vSphere Storage DRS provides the following functions:

- Initial placement of virtual machines based on storage capacity, and optionally on I/O latency
- Use of vSphere Storage vMotion to migrate virtual machines based on storage capacity and, optionally, I/O latency
- Configuration in either manual mode or fully automated mode
- Use of affinity and anti-affinity rules to govern virtual disk location
- Use of datastore maintenance mode to clear a LUN of virtual machine files

vSphere Storage DRS manages the placement of virtual machines in a datastore cluster, based on the space usage of the datastores. It tries to keep usage as even as possible across datastores in the datastore cluster.

vSphere Storage DRS uses vSphere Storage vMotion to migrate virtual machines to maintain the balance across datastores.

Optionally, the user can configure vSphere Storage DRS to balance I/O latency across members of the datastore cluster as a way to help mitigate performance problems that are caused by I/O latency.

vSphere Storage DRS can be set up to work in either manual mode or fully automated mode:

- Manual mode presents migration and placement recommendations to the user, but nothing is executed until the user accepts the recommendation.
- Fully automated mode automatically handles initial placement and migrations based on runtime rules.

https://t.me/learningnets

3-118 Initial Disk Placement

Initial placement occurs when vSphere Storage DRS selects a datastore in a datastore cluster on which to place a virtual machine disk.

When virtual machines are created, cloned, or migrated, you select a datastore cluster rather than a single datastore. vSphere Storage DRS selects a member datastore based on capacity and, optionally, on IOPS load.

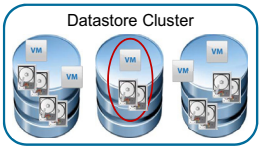
By default, a virtual machine's files are placed on the same datastore in the datastore cluster. This behavior can be changed by using vSphere Storage DRS anti-affinity rules.

When a virtual machine is created, cloned, or migrated, and the user selects a datastore cluster, vSphere Storage DRS chooses a member datastore in the datastore cluster based on storage use. vSphere Storage DRS tries to keep the member datastores evenly used.

https://t.me/learningnets

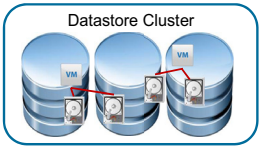
3-119 vSphere Storage DRS Affinity Rules

You can create vSphere Storage DRS anti-affinity rules to control which virtual disks should not be placed on the same datastore in a datastore cluster.



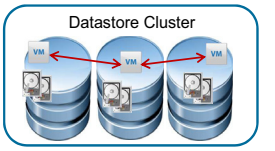
Intra-VM VMDK Affinity Rule (default):

- Keep virtual machine VMDKs together on the same datastore.
- Maximize virtual machine availability when all disks are needed to run.



Intra-VM VMDK Anti-Affinity Rule:

- Keep virtual machine VMDKs on different datastores.
- Rule can be applied to all or a subset of virtual machine disks.



VM Anti-Affinity Rule:

- Keep virtual machines on different datastores.
- Rule is similar to the vSphere DRS anti-affinity rule.
- Maximize availability of a set of redundant virtual machines.

By default, all of a virtual machine's disks can be on the same datastore.

A user might want the virtual disks on different datastores. For example, a user can place a system disk on one datastore and the data disks on another. In this case, the user can set up a VMDK anti-affinity rule which keeps a virtual machine's virtual disks on separate datastores.

Virtual machine anti-affinity rules keep virtual machines on separate datastores. This rule is useful when redundant virtual machines must always be available.

https://t.me/learningnets

3-120 Storage Migration Recommendations

vCenter Server displays migration recommendations on the vSphere Storage DRS Recommendations page for datastore clusters that have manual mode set.

Migration recommendations are executed under the following circumstances:

- When the I/O latency threshold is exceeded
- When the space utilization threshold is exceeded

If enabled, IOPS load history is checked every eight hours by default.

vSphere Storage DRS selects a datastore based on utilization and, optionally, I/O latency.

Load balancing is based on IOPS workload, which ensures that no datastore exceeds a particular VMkernel I/O latency level.

vSphere Storage DRS provides as many recommendations as necessary to balance the space and, optionally, the IOPS resources of the datastore cluster.

Reasons for migration recommendations include balancing space usage in the datastore, reducing datastore I/O latency, and balancing datastore IOPS load.

vSphere Storage DRS can also make mandatory recommendations based on the following conditions:

- A datastore is out of space.
- vSphere Storage DRS anti-affinity rules or affinity rules are being violated.
- A datastore is entering maintenance mode.

vSphere Storage DRS also considers moving powered-off virtual machines to other datastores in order to achieve optimal balance of virtual machines across datastores in the datastore cluster.

https://t.me/learningnets

3-121 Datastore Correlation Detector

Datastore correlation refers to datastores that are created on the same physical set of spindles.

vSphere Storage DRS detects datastore correlation by performing the following operations:

- Measuring individual datastore performance
- Measuring combined datastore performance

If latency increases on multiple datastores when load is placed on one datastore, then the datastores are considered to be correlated.

Correlation is determined by a long-running background process.

Anti-affinity rules can use correlation detection to ensure that the virtual machines or virtual disks are on different spindles.

Datastore correlation is enabled by default.

The purpose of datastore correlation is to help the decision-making process in vSphere Storage DRS when deciding where to move a virtual machine. For example, you gain little advantage by moving a virtual machine from one datastore to another if both datastores are backed by the same set of physical spindles on the array.

The datastore correlation detector uses the I/O injector to determine whether a source and a destination datastore are using the same back-end spindles.

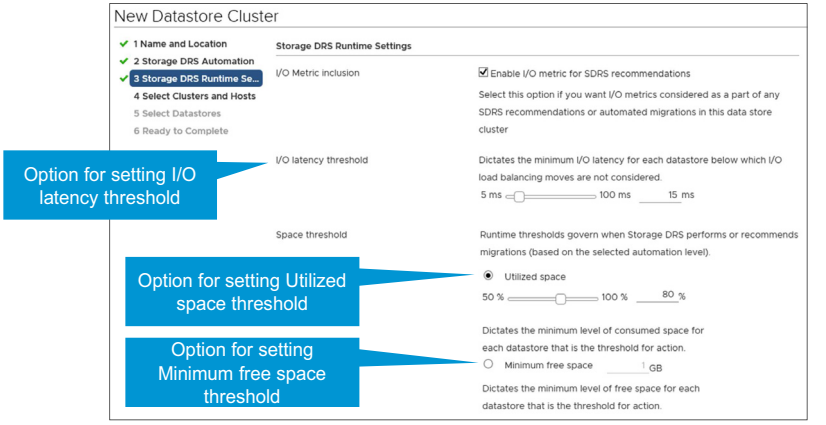
The datastore correlation detector works by monitoring the load on one datastore and monitoring the latency on another. If latency increases on other datastores when a load is placed on one datastore, the datastores are correlated.

The datastore correlation detector can also be used for anti-affinity rules, ensuring that virtual machines and virtual disks are not only kept on separate datastores but also kept on different spindles on the back end.

https://t.me/learningnets

3-122 Configuring vSphere Storage DRS Runtime Settings

vSphere Storage DRS thresholds can be configured to determine when vSphere Storage DRS performs or recommends migrations.



You use the **Enable I/O metric for SDRS recommendations** check box on the Storage DRS Runtime Settings page to enable or disable IOPS metric inclusion.

When I/O load balancing is enabled, Storage I/O Control is enabled for all the datastores in the datastore cluster if it is not already enabled for the cluster.

When this option is deselected, you disable the following functions:

- IOPS load balancing among datastores in the datastore cluster.
- Initial placement for virtual disks based on the IOPS metric. Space is the only consideration when placement and balancing recommendations are made.

vSphere Storage DRS runtime settings include the following options:

- I/O latency threshold: Indicates the minimum I/O latency for each datastore below which I/O load-balancing moves are not considered. This setting is applicable only if the **Enable I/O metric for SDRS recommendations** check box is selected.
- Space threshold: Determines the minimum levels of consumed space and free space for each datastore that is the threshold for action.

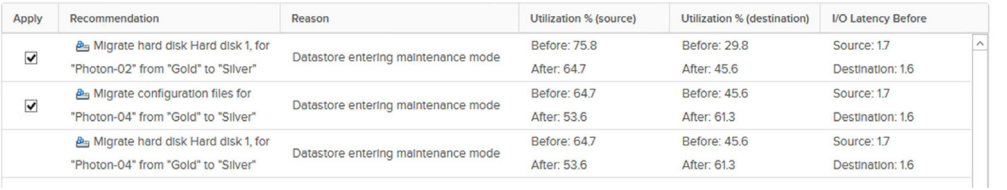
https://t.me/learningnets

3-123 vSphere Storage DRS Maintenance Mode

vSphere Storage DRS maintenance mode enables you to take a datastore out of use to service it.

vSphere Storage DRS maintenance mode evacuates virtual machines from a datastore placed in maintenance mode:

- Registered virtual machines (on or off) are moved.
- Templates, unregistered virtual machines, ISO images, and nonvirtual machine files are not moved.



Apply	Recommendation	Reason	Utilization % (source)	Utilization % (destination)	I/O Latency Before
☑	Migrate hard disk Hard disk 1, for "Photon-02" from "Gold" to "Silver"	Datastore entering maintenance mode	Before: 75.8 After: 64.7	Before: 29.8 After: 45.6	Source: 1.7 Destination: 1.6
☑	Migrate configuration files for "Photon-04" from "Gold" to "Silver"	Datastore entering maintenance mode	Before: 64.7 After: 53.6	Before: 45.6 After: 61.3	Source: 1.7 Destination: 1.6
	Migrate hard disk Hard disk 1, for "Photon-04" from "Gold" to "Silver"	Datastore entering maintenance mode	Before: 64.7 After: 53.6	Before: 45.6 After: 61.3	Source: 1.7 Destination: 1.6

vSphere Storage DRS enables you to place a datastore in maintenance mode. vSphere Storage DRS maintenance mode is available to datastores in a datastore cluster that is enabled for vSphere Storage DRS. Standalone datastores cannot be placed in maintenance mode.

The datastore does not enter maintenance mode until all files on the datastore are moved. You must manually move these files off the datastore so that the datastore can enter vSphere Storage DRS maintenance mode.

If the datastore cluster is placed in fully automated mode, virtual machines are automatically migrated to other datastores.

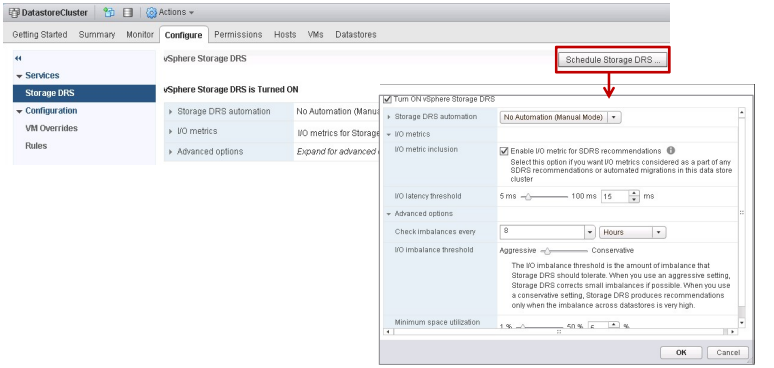
If the datastore cluster is placed in manual mode, migration recommendations appear in vSphere Client. The virtual disks cannot be moved until the recommendations are accepted.

https://t.me/learningnets

3-124 Backups and vSphere Storage DRS

Backing up virtual machines can add latency to a datastore.

Using vSphere Web Client, you can schedule a task to disable vSphere Storage DRS behavior for the duration of the backup.

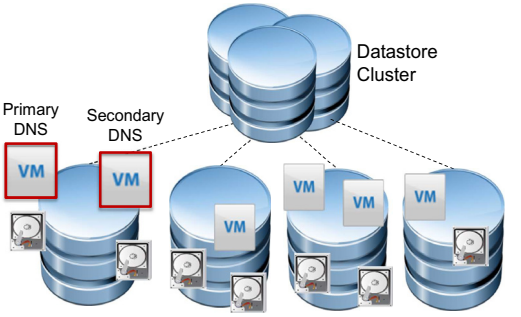


Scheduled tasks can be configured to change vSphere Storage DRS behavior. Scheduled tasks can be used to change the vSphere Storage DRS configuration of the datastore cluster to match enterprise activity. For example, if the datastore cluster is configured to perform migrations based on I/O latency, you might disable the use of I/O metrics by vSphere Storage DRS during the backup window. You can reenable I/O metrics use after the backup window closes.

https://t.me/learningnets

3-125 Activity: vSphere Storage DRS (1)

In the datastore cluster below, how would you configure vSphere Storage DRS to eliminate the single point of failure for the primary DNS and secondary DNS servers?

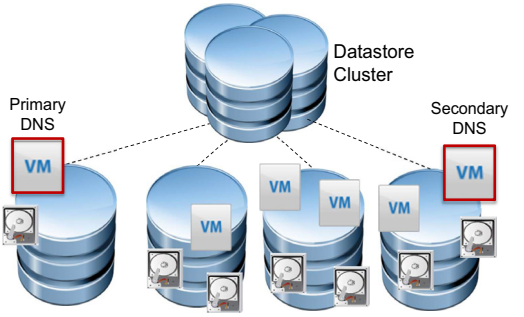


https://t.me/learningnets

3-126 Activity: vSphere Storage DRS (2)

You create a VM anti-affinity rule for the primary DNS and secondary DNS servers.

A VM anti-affinity rule ensures that the VMs listed in the rule are always placed on different datastores.



https://t.me/learningnets

3-127 vSphere Storage DRS and vSphere Technology Compatibility

Several vSphere technologies are supported with vSphere Storage DRS. Each technology has a recommended migration method.

Feature or Product	Supported / Not Supported	Migration Recommendation
VMware snapshots	Supported	Fully Automated
Raw device mapping pointer files	Supported	Fully Automated
VMware thin-provisioned disks	Supported	Fully Automated
VMware vSphere linked clones	Supported	Fully Automated
vSphere Storage Metro Cluster	Supported	Manual
Site Recovery Manager Server	Supported	Fully Automated (from protected site)
vCloud Director	Supported	Fully Automated
vSphere Replication	Supported	Fully Automated (from protected site)
VM Storage Policies	Supported	Fully Automated (from protected site)

For more information about vSphere Storage DRS interoperability, see http://www.yellow-bricks.com/2013/05/13/vsphere-5-1-storage-drs-interoperability and http://www.yellow-bricks.com/2015/02/09/what-is-new-for-storage-drs-in-vsphere-6-0.

https://t.me/learningnets

3-128 Interoperability with Site Recovery Manager

vSphere Storage DRS makes decisions or recommendations based on replication policy:

- vSphere Storage DRS recommendations are appropriate to a particular replication policy. Some datastores might have asynchronous replication and other datastores might have synchronous replication.
- vSphere Storage DRS ensures that each virtual machine is placed or balanced on a datastore with the same policy.
- When moving a virtual machine to a datastore that does not have the same policy, vSphere Storage DRS alerts the datastore that the virtual machine might experience a temporary loss of replication, for example, when the datastore is in maintenance mode.

vSphere DRS can process and honor VMware Site Recovery Manager™ storage policies. For example, some datastores might have asynchronous replication and others might have synchronous replication.

vSphere Storage DRS ensures that a virtual machine is placed and balanced on a datastore with the same policy. A recommendation to move a virtual machine to another datastore that lacks the same policy or that is in maintenance mode might exist. vSphere Storage DRS alerts that the datastore to which the virtual machine is being moved might result in a temporary loss of replication.

https://t.me/learningnets

3-129 Interoperability with vSphere Replication

In VMware vSphere® Replication™, replica disks are created on the secondary site:

- vSphere Storage DRS recognizes the replica disks on the secondary site and can determine the space usage.
- Disks on the primary and secondary sites can be balanced in the same way that standard virtual machine files are balanced.

Through vSphere API for Storage Awareness, vSphere Storage DRS can examine replica disks to determine the space used. For example, replica disks are instantiated on the secondary site. Because vSphere Storage DRS understands the space usage of replica disks, it can be used for replica disks on the secondary site and the replica disks can be balanced in the same way as standard virtual machine files.

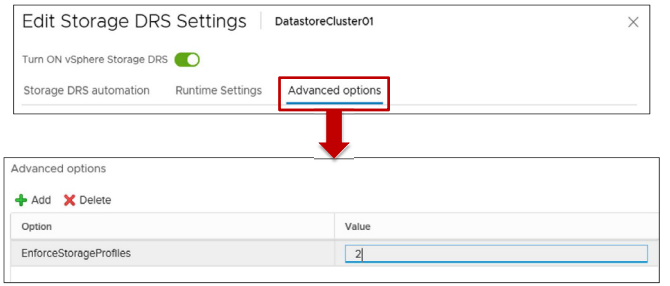
https://t.me/learningnets

3-130 Enabling vSphere Storage DRS to Recognize Storage Policies

If VMs have storage policies associated with them, vSphere Storage DRS can enforce placement based on underlying datastore capabilities.

You can enforce storage policies in a datastore cluster by setting the EnforceStorageProfiles advanced option:

- 0: Default. Do not enforce storage policies.
- 1: Soft. vSphere Storage DRS violates storage policy compliance if it is required to do so.
- 2: Hard. vSphere Storage DRS does not violate storage policy compliance.



https://t.me/learningnets

3-131 vSphere Storage DRS and Array Feature Compatibility

Several storage array features are supported with vSphere Storage DRS and are dependent on vSphere API for Storage Awareness. Your storage vendor must provide an updated storage provider. Also, the storage administrator must register the storage provider before vSphere Storage DRS can perform these operations.

Feature or Product	Initial Placement	Recommended Migration Method	
		Without vSphere API for Storage Awareness 2.0	With vSphere API for Storage Awareness 2.0
Array-based snapshots	Supported	Manual	Automated
Array-based deduplication	Supported	Manual	Automated
Array-based thin provisioning	Supported	Manual	Automated
Array-based autotiering	Supported	Manual (only capacity load balancing)	Automated
Array-based replication	Supported	Manual	Automated

Storage vendors can use vSphere API for Storage Awareness to provide information to vSphere about specific storage arrays for tighter integration between storage and the virtual infrastructure. The shared information includes details on storage virtualization, such as health status, configuration, capacity, and thin provisioning. This level of detail can be passed through vCenter Server to the user.

https://t.me/learningnets

3-132 Deduplication Awareness

vSphere Storage DRS can use vSphere API for Storage Awareness to determine if a datastore is being deduplicated:

- If a datastore is being deduplicated, vSphere Storage DRS can determine whether one or more datastores share common storage pools (aggregated physical storage resources).
- vSphere Storage DRS avoids moving virtual machines between datastores that are not in the same deduplication pool.

vSphere API for Storage Awareness is used to determine active deduplication state on a datastore. This information helps vSphere Storage DRS to prevent moving virtual machines across datastores that are not in the same deduplication pool.

Physical storage resources are aggregated into storage pools, from which the logical storage is created. More storage systems, which might be heterogeneous in nature, can be added when needed. The virtual storage space scales up by the same amount. This process is fully transparent to the applications that are using the storage infrastructure.

https://t.me/learningnets

3-133 Thin-Provisioned Datastore Awareness

vSphere Storage DRS uses vSphere API for Storage Awareness to discover the common backing pool shared by multiple datastores:

- The API enables vSphere Storage DRS to make recommendations based on the real available capacity in the shared storage pool, not on the reported capacity of the datastore.
- vSphere Storage DRS avoids migrating virtual machines between two thin-provisioned datastores backed by the same pool.
- vSphere Storage DRS can also provide remediation when the free space in the storage pool is running out, by moving virtual machines away from datastores that are sharing the same common storage pool.

vSphere Storage DRS uses vSphere API for Storage Awareness to examine the capacity of the backing pool on the array by using that capacity in all subsequent calculations.

This feature enables vSphere Storage DRS to avoid migrating virtual machines between two thin-provisioned datastores that share the same backing pool. vSphere Storage DRS can make recommendations based on the real available space in the shared backing pool, rather than the reported capacity of the datastore, which might be larger.

https://t.me/learningnets

3-134 Array-Based Autotiering Awareness

Increasingly, vendors are offering autotiering arrays that require varied amounts of performance data on which to base decisions.

Each vendor solution works differently. Sometimes, decisions are made in real time. At other times, 24 hours might be required.

vSphere Storage DRS uses vSphere API for Storage Awareness to identify autotiering storage arrays.

After vSphere Storage DRS determines that an array has autotiering capabilities, the array is treated appropriately for performance modeling.

https://t.me/learningnets

3-135 vSphere Storage DRS and Storage I/O Control

vSphere Storage DRS and Storage I/O Control are complementary solutions:

- Storage I/O Control is set to statistics-only mode by default:
  - vSphere Storage DRS works to avoid I/O bottlenecks.
  - Storage I/O Control manages unavoidable I/O bottlenecks.
- Storage I/O Control works in real time.
- vSphere Storage DRS does not use real-time latency to calculate load balancing.
- vSphere Storage DRS and Storage I/O Control provide the performance that you need in a shared environment, without having to significantly overprovision storage.

Both vSphere Storage DRS and Storage I/O Control work with IOPS, and they should be used together. Storage I/O Control is automatically enabled on each datastore in the datastore cluster when the I/O Metric feature of vSphere Storage DRS is enabled. Storage I/O Control is used to manage, in real time, unavoidable IOPS bottlenecks, such as short, intermittent bottlenecks and congestion on every datastore in the datastore cluster.

Storage I/O Control continuously checks for latency and controls I/O accordingly.

vSphere Storage DRS uses IOPS load history to determine migrations. vSphere Storage DRS runs infrequently and performs analysis to determine long-term load balancing.

Storage I/O Control monitors the I/O metrics of the datastores. vSphere Storage DRS uses this information to determine whether a virtual machine should be moved from one datastore to another.

https://t.me/learningnets

3-136 Lab 5: Managing Datastore Clusters

Create a datastore cluster and work with vSphere Storage DRS

1. Create a Datastore Cluster That Is Enabled for vSphere Storage DRS
2. Use Datastore Maintenance Mode to Evacuate a Datastore
3. Run vSphere Storage DRS and Apply Migration Recommendations
4. Clean Up for the Next Lab

https://t.me/learningnets

3-137 Review of Learner Objectives

You should be able to meet the following objectives:

- Create a datastore cluster
- Configure vSphere Storage DRS
- Explain how Storage I/O Control and vSphere Storage DRS complement each other

https://t.me/learningnets

3-138 Key Points

- With a VMFS6 datastore, the ESXi host automatically sends the SCSI UNMAP command to the storage array. The unmap rate is configurable.
- vSphere API for Storage Awareness enables storage vendors to provide information about the capabilities of their storage arrays to vCenter Server.
- VAIO enables VMware and third-party vendors to create data services such as caching and replication.
- Storage policy-based management abstracts storage and data services delivered by vSAN, vSphere Virtual Volumes, I/O filters, or traditional storage.
- Storage I/O Control enables cluster-wide storage I/O prioritization.
- A datastore cluster enabled for vSphere Storage DRS is a collection of datastores working together to balance storage capacity and I/O latency.

Questions?

https://t.me/learningnets

Host and Management Scalability



https://t.me/learningnets

4-2 You Are Here

1. Course Introduction
2. Network Scalability
3. Storage Scalability
4. **Host and Management Scalability**
5. CPU Optimization
6. Memory Optimization
7. Storage Optimization
8. Network Optimization
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

4-3 Importance

As you scale your vSphere environment, you must be aware of the vSphere features and functions that help you manage hosts in your environment.

https://t.me/learningnets

4-4 Module Lessons

Lesson 1:	Content Libraries
Lesson 2:	Host Profiles
Lesson 3:	vSphere ESXi Image Builder
Lesson 4:	vSphere Auto Deploy

https://t.me/learningnets



https://t.me/learningnets

4-6 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Summarize the purpose of content libraries in a vSphere environment
- Discuss the vSphere requirements for content libraries
- Create a local content library
- Subscribe to a published content library
- Deploy virtual machines from a content library

https://t.me/learningnets

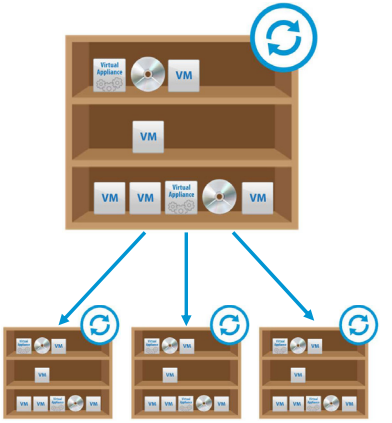
4-7 About Content Libraries

vSphere content libraries provide storage, versioning, and synchronization of files across sites and vCenter Server instances.

They provide simple and effective management for templates, vApps, OVFs, ISO images, and scripts.

Content libraries include powerful publish and subscribe features to replicate content.

Content libraries are stored on vSphere datastores or vCenter Server file systems.



Content libraries provide a simple and effective way to manage multiple content types in a small vSphere environment or across several vCenter Server instances. A content library can be synchronized across remotely-located sites and against remotely-located vCenter Server instances.

Content libraries allow subscription to published data set, which enable the use of consistent data across multiple sites. Publication is controlled by a vSphere administrator who maintains consistent data and security on that data. Data can be synchronized automatically across sites or on demand when needed.

The type of storage used for content libraries depends on the use for the data in the library.

https://t.me/learningnets

4-8 Content Library Benefits

A content library has many benefits for vSphere administrators:

- Simple management and sharing of templates and other file types
- Consistent content between multiple sites
- Consistent file security between multiple sites
- Efficient use of content storage
- Synchronization between local and remote vCenter Server instances
- Synchronization between vCenter Server and VMware vCloud Director®

Storage and consistency is a key reason to install and use a content library. Sharing content and ensuring that this content is kept up to date is a major task. For example, you might have a main vCenter Server instance, where you might create a central content library to store the master copies of OVF templates, ISO images, or other file types.

You can publish this content library to allow other libraries that are located across the world to subscribe and download an exact copy of the data. If an OVF template is added, modified, or deleted from the published catalog, the subscriber synchronizes with the publisher, and the libraries are updated with the latest content.

Because content libraries are globally available for subscription, security might be a concern. To resolve this security concern, content libraries can be password-protected during publication. This password is a static password and no integration occurs with VMware vCenter® Single Sign-On or Active Directory in the current release of vCenter Server.

https://t.me/learningnets

4-9 Content Library Types

Content is stored in one of the content library types:

- Local: Content controlled by the administrator.
- Published: A local library that is available for subscription.
- Optimized published: A published library to ensure lower CPU usage and faster streaming of content over HTTP.
- Subscribed: A library that synchronizes with a published library.

Content in a local or published content library can be added, modified, or deleted by administrators.

Subscribed content library content cannot be changed by a user who has subscribed to it.

A subscribed content library cannot be published.

Use an optimized published content library when the subscribed libraries reside on a remote vCenter Server system and enhanced linked mode is not used.

It is not possible to deploy VMs from templates in an Optimized library. A subscribed library must first be created before VMs can be deployed from the subscribed library.

A local library is the simplest form of library. A published library is a local library that is available for subscription. Version changes are tracked by the content library, using the vCenter Server database. No option to use previous versions of content is available.

A subscribed library is configured to subscribe to a published library or an optimized published library. An administrator cannot directly change the contents of the subscribed library. But the administrator can control how the data is synchronized to the subscribed content library.

An optimized published library is a library that is optimized to ensure lower CPU usage and faster streaming of the content over HTTP. Use this library as a main content depot for your subscribed libraries. You cannot deploy virtual machines from an optimized library. You must first configure a subscribed library to subscribe to the optimized published library and then deploy virtual machines from the subscribed library.

https://t.me/learningnets

4-10 Storing Content in Content Libraries

The following options for content library storage are available:

- Using mounted remote storage on a vCenter Server instance provides access for content library synchronization and deployment of OVF files to ESXi hosts.
- Using a vSphere datastore provides deployment of OVF files to ESXi hosts and synchronization and download of other content.

Whichever content library storage option is used, content can be stored only on a single file system mount or datastore. After a content library is created, it cannot be moved to a different type of storage.

Factors that can affect library synchronization performance are:

- Type of data path: direct connection versus streaming
- Network capacity
- Data compression or decompression
- Data encryption or decryption

Content library storage can be a datastore in your vSphere inventory or it can be a local file system on a vCenter Server system.

Content library storage can also be a remote storage location. If you use a Windows-based vCenter Server system, then the storage can be an SMB server. If you use VMware vCenter® Server Appliance™, then the storage can be an NFS server.

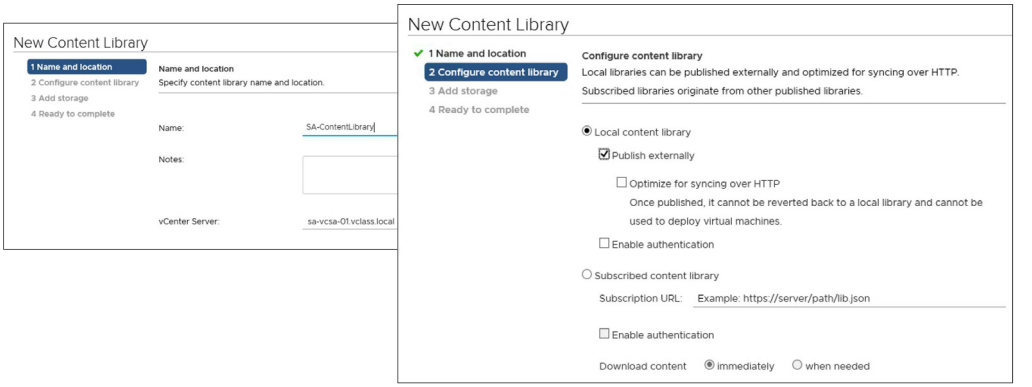
For more information about how to increase the performance of content libraries, see the VMware performance team blog How to Efficiently Synchronize, Import and Export Content in VMware vSphere Content Library at https://blogs.vmware.com/performance/2015/09/efficiently-synchronize-import-export-content-vmware-vsphere-content-library.html.

https://t.me/learningnets

4-11 Creating Content Libraries

When you create a content library, you select the content library type.

You also select where the library will be stored, either in a datastore on a host or in a mounted filesystem on a vCenter Server.



The **Optimize for syncing over HTTP** check box controls how a published library stores and synchronizes content. When this check box is selected, the content is stored in a compressed format. Compression reduces the synchronization time between vCenter Server instances that are not using Enhanced Linked Mode.

It is not possible to deploy virtual machines directly from templates stored on an optimized published library. You must first create a subscribed library which will synchronize with the optimized published library and then you can deploy virtual machines from the subscribed library.

https://t.me/learningnets

4-12 Populating Content Libraries

You populate a content library with templates that you can use to provision new virtual machines.

To add templates to a content library, use one of the following methods:

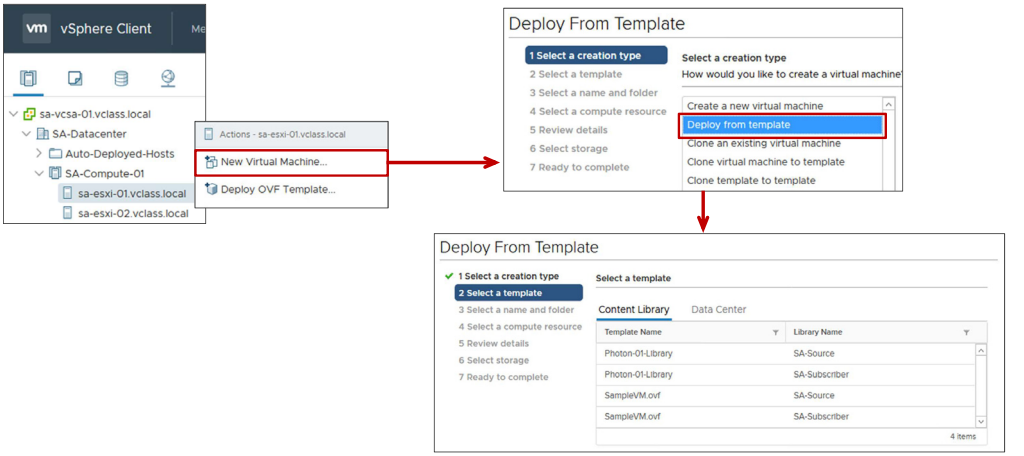
- Clone a virtual machine to a template in the content library.
- Clone a template from the vSphere inventory or from another content library.
- Clone a vApp.
- Import a template from a URL.
- Import an OVF file from your local file system.

Your source to import items into content libraries can be a file stored on your local machine or a file stored on a web server. You can add an item that resides on your local system to a content library.

https://t.me/learningnets

4-13 Using the Content Library to Deploy Virtual Machines (1)

You can deploy a virtual machine from a template in the content library.

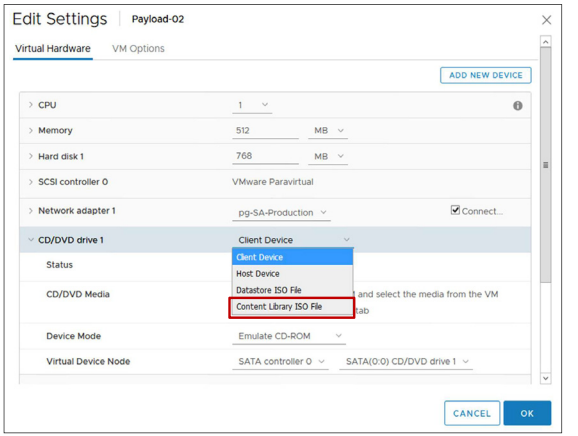


You can use a virtual machine template from a content library to deploy a virtual machine to a host or a cluster in your vSphere inventory. You can also apply a guest operating system customization specification to the virtual machine.

https://t.me/learningnets

4-14 Using the Content Library to Deploy Virtual Machines (2)

You can also create a virtual machine and mount an ISO file from the content library.

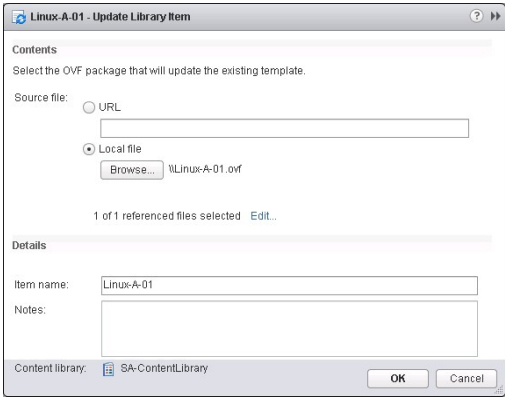


To install a guest operating system and its applications on a new virtual machine, you can connect the CD/DVD device to an ISO file that is stored in a content library.

https://t.me/learningnets

4-15 Updating Content Library Items

You can directly update a virtual machine template, for example, to add a patch to it instead of deleting the existing template and creating a new one.

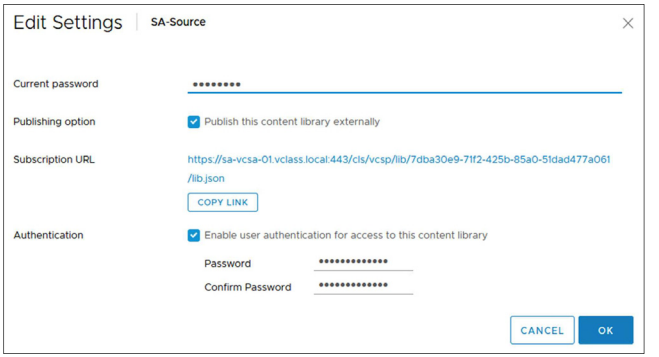


Managing and keeping your virtual environment up to date might require you to update the content of a library item.

https://t.me/learningnets

4-16 Publishing Content Libraries for External Use

You can publish a content library for external use and add password protection by editing the content library settings. Users access the library through the subscription URL that is system generated.

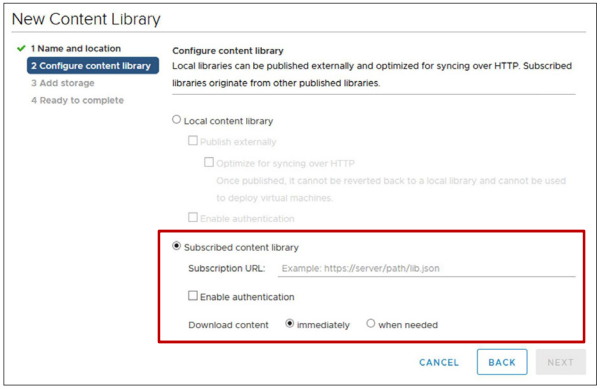


https://t.me/learningnets

4-17 Content Library Subscriptions

You subscribe to a content library by configuring the path to the subscription URL:

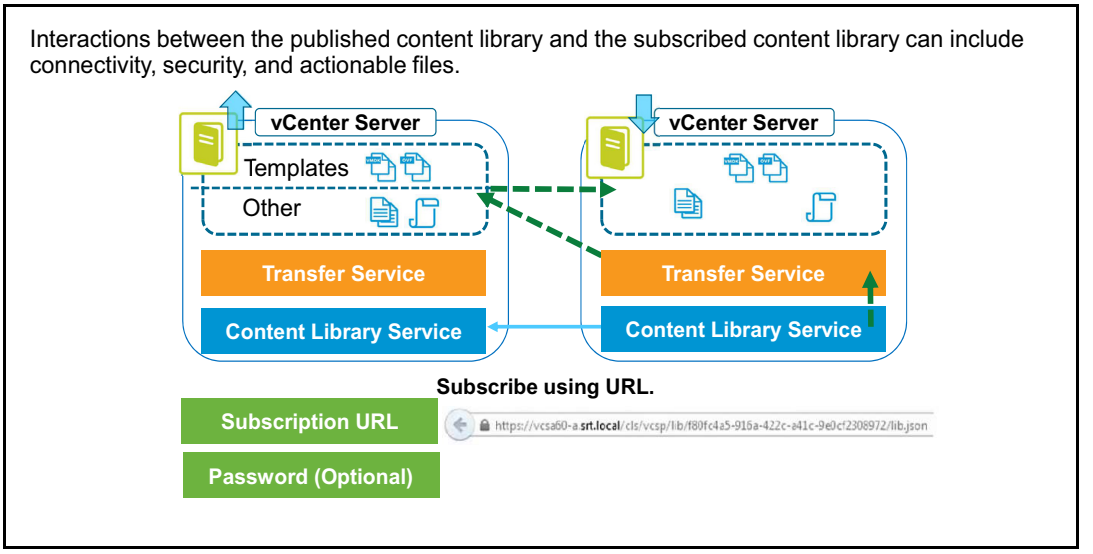
- You can immediately download all library content to the storage location that you configure.
- To save space, store only metadata for items until they are used.



When a content library subscription is created, the administrator selects how the content will be synchronized with the published content library. Content can be downloaded immediately if space for the content is not a concern. Synchronization can be set to on-demand so that only the metadata is copied when the subscription is created. The full payload of the content library is downloaded only as it is used.

https://t.me/learningnets

4-18 Published and Subscribed Content Libraries



The slide shows two vCenter Server instances, with both the Content Library Service and the Transfer Service installed and running on them. The user creates a content library on the first vCenter Server instance.

The content in the content library is divided into two categories:

- Templates: This category contains only OVF templates. You can deploy them directly from the content library as virtual machines to hosts, clusters, virtual data centers, and so on.
- Other: This category contains all other file types, such as ISO images. You can connect the CD/DVD device to an ISO file that is stored in a content library, for example, to install a guest operating system on a new virtual machine.

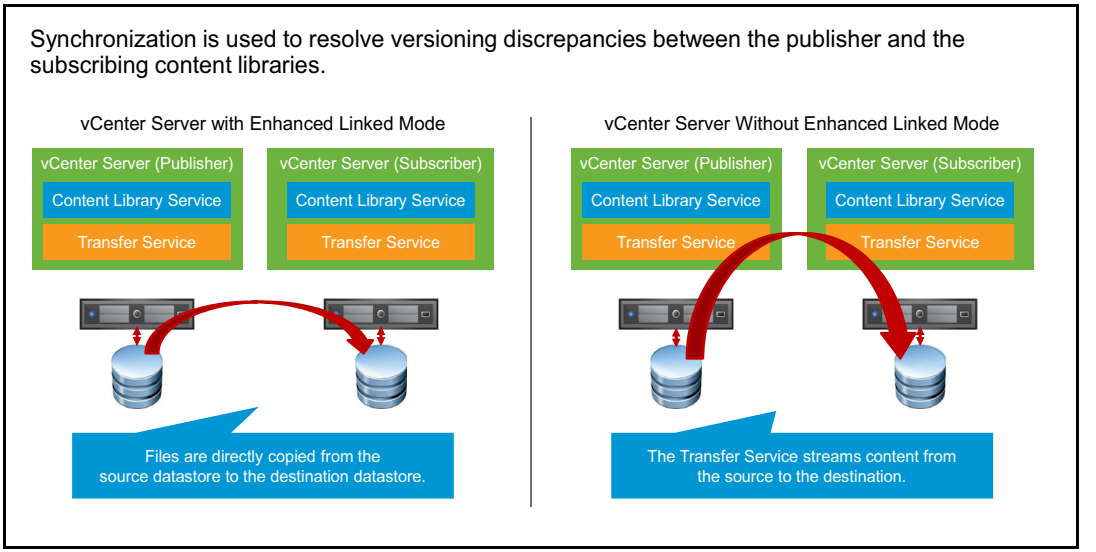
The administrator can publish the library. Publishing generates a URL that points to the `lib.json` file of that library. As an option, the administrator can enable authentication and assign a password. By default, the user name is vcsp. You cannot set a user name during the creation of the content library but you can change the user name in the content library properties after creation.

Another content library is created on a separate vCenter Server instance and subscribes to the publisher using the URL and password through the Content Library Service. The Content Library Service then calls the Transfer Service. The Transfer Service is responsible for the import and export of content from the publisher to the subscriber over HTTP NFC.

See *vSphere Virtual Machine Administration* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-virtual-machine-admin-guide.pdf.

https://t.me/learningnets

4-19 Synchronizing and Versioning Content Libraries



Synchronization occurs after a library is subscribed. Synchronization occurs in an environment that has two content libraries. Both libraries have the same content, but a new ISO image is added.

When any item is changed in a content library, the item number is incremented to note the modification. The version number for the content library as a whole is also incremented. This process is called simple versioning. The content library service determines whether synchronization should occur by using simple versioning. Simple versioning does not store multiple versions or offer rollback features, it is purely a numerical value assigned to the content library and its content.

The content library's version number is checked first. If this number matches the one in the published and subscribed libraries, no synchronization is needed. If there is a discrepancy, the version number of each item in the library is checked. If any discrepancies are found, only items with such discrepancies are synchronized.

When multiple vCenter Server instances are configured in an Enhanced Linked Mode environment, the contents of the published library can be synchronized to a subscribed library by directly copying the files from the source datastore to the destination datastore. When Enhanced Linked Mode is not available, the contents of a published library must be streamed through the Transfer Service component. In such a case, three sub-scenarios exist, based on the storage configuration:

1. Both published and subscribed libraries reside on a datastore.
2. Both published and subscribed libraries reside on an NFS file system mounted on the vCenter Server instances.
3. The published library resides on an NFS file system while the subscribed library resides on a datastore.

https://t.me/learningnets

4-20 Content Library Maximums

The content library has the following configuration maximums.

Configuration Item	Maximum
Content library item size	1 TB
Total items per library	1,000
Total library items per vCenter Server instance (across all libraries)	2,000
Maximum number of concurrent sync operations on the published library's vCenter Server instance	16
Total number of libraries per vCenter Server instance	1,000

The content library can be backed by a datastore or stored to available storage on vCenter Server. Regardless of the option chosen, the content library can be backed only by a single file system or datastore.

The maximum size of a library item is 1 TB. A content library can hold a maximum of 1,000 items and a total of 2,000 items across all libraries in a vCenter Server instance.

The maximum number of concurrent synchronization operations on the published library's vCenter Server instance is 16.

Automatic synchronization occurs once every 24 hours by default, but the time and frequency can be configured through the API. The administrator can synchronize an entire content library or an individual item at any time through vSphere Client.

https://t.me/learningnets

4-21 Lab 6: Creating a Content Library

Create a multisite content library

1. Create a Content Library
2. Upload Data to the New Content Library
3. Create a Subscriber Content Library
4. Clone a Template to the Source Library
5. Synchronize the Content Libraries
6. Deploy a Virtual Machine from the Library

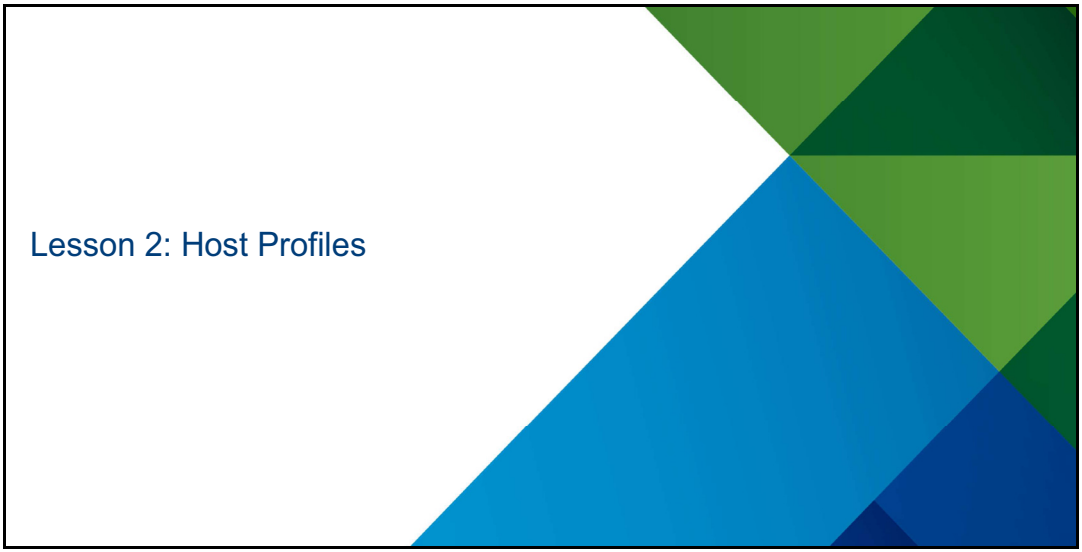
https://t.me/learningnets

4-22 Review of Learner Objectives

You should be able to meet the following objectives:

- Summarize the purpose of content libraries in a vSphere environment
- Discuss the vSphere requirements for content libraries
- Create a local content library
- Subscribe to a published content library
- Deploy virtual machines from a content library

https://t.me/learningnets



https://t.me/learningnets

4-24 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Describe the components of host profiles
- Describe the benefits of host profiles
- Explain how host profiles operate
- Use a host profile to manage ESXi configuration compliance
- Describe the benefits of exporting and importing host profiles

https://t.me/learningnets

4-25 Host Configuration Overview

ESXi host configurations include the following elements:

- Storage
- Networking
- Licensing
- Date and time configuration
- DNS and routing
- Security settings

Existing processes for modifying a single host's configuration are either manual or automated through scripts.

Over time, your vSphere environment expands and changes. New hosts are added to the data center and must be configured to coexist with other hosts in the cluster:

- Storage: Datastores such as VMFS and NFS, iSCSI initiators and targets, and multipathing.
- Networking: Virtual switches, port groups, physical NIC speed, security, NIC teaming policies, and so on.
- Licensing: Edition selection, license key input, and so on.
- Date and time configuration: Network Time Protocol server
- Security settings: Firewall settings, Active Directory configuration, ESXi services, and so on.
- DNS and routing: DNS server, IP default gateway, and so on.

Processes already exist to modify the configuration of a single host:

- Manually, using vSphere Client or the command prompt
- Automatically, using scripts

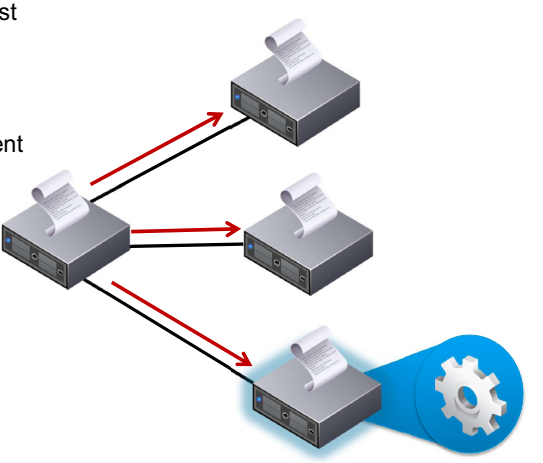
https://t.me/learningnets

4-26 About Host Profiles

Host profiles eliminate configurations per host and maintain configuration consistency and correctness across the entire data center.

Host profiles have many benefits:

- Simplified setup and change management for ESXi hosts
- Easy detection of noncompliance with standard configurations
- Automated remediation



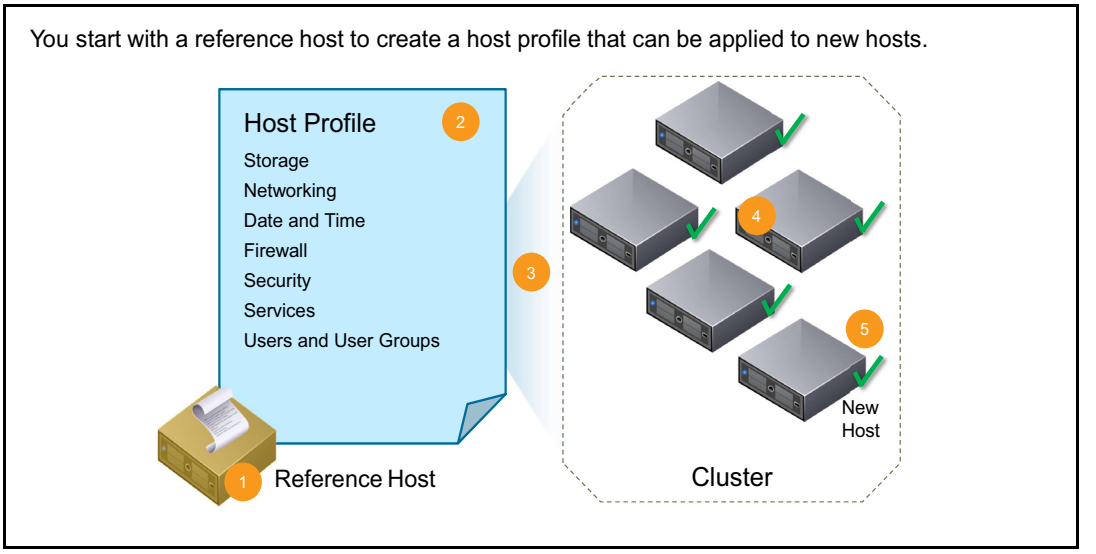
The Host Profiles feature enables you to export configuration settings from a master reference host and save them as a portable set of policies, called a host profile. You can use this profile to quickly configure other hosts in the data center. Configuring hosts with this method drastically reduces the setup time of new hosts. Multiple steps are reduced to a single click. Host profiles also eliminate the need for specialized scripts to configure hosts.

The vCenter Server system uses host profiles as host configuration baselines so that you can monitor changes to the host configuration, detect discrepancies, and fix them.

As a licensed feature of vSphere, host profiles are available only with Enterprise Plus licensing. If you see errors, ensure that you have the appropriate vSphere licensing for your hosts.

https://t.me/learningnets

4-27 Basic Workflow to Implement Host Profiles



Follow this basic workflow to implement host profiles:

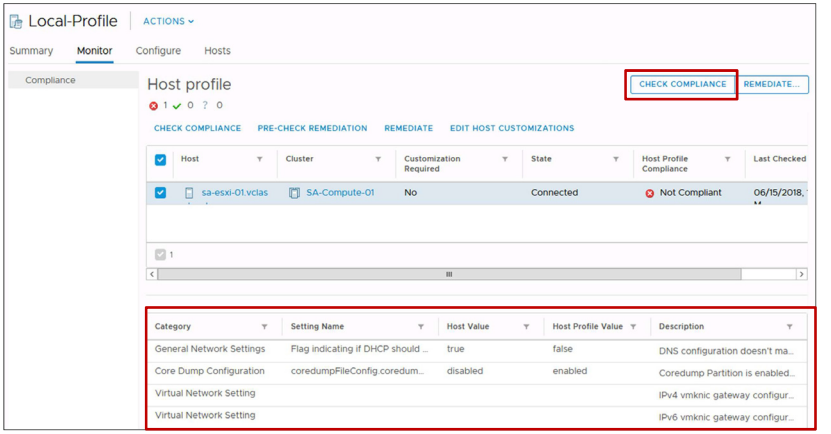
1. Set up and configure a host to be used as the reference host. A reference host is the host from which the profile is created.
2. Use the Create Profile wizard to create a profile from the designated reference host. vCenter Server retrieves the configuration settings of the reference host, including IP addressing, NFS mounts, and vSphere vMotion ports. You can edit the profile to further customize the policies in the profile.
3. Attach the host profile to a host or a cluster of hosts. You can attach a profile to standalone hosts, or to one or more clusters of hosts.
4. Check host compliance to the reference host's profile. If all hosts are compliant with the reference host, the hosts are correctly configured.
5. Check new hosts for compliance against the reference host's profile. You can easily apply the host profile of the reference host to other hosts or clusters of hosts that are not compliant.

You can also import and export a profile file to a host profile that is in the VMware profile format (`.vpf`). This feature is useful for sharing host profiles across multiple vCenter Server instances.

https://t.me/learningnets

4-28 Monitoring for Compliance

When a new host is added to a cluster, the host is checked for compliance against the host profile that was applied. A cluster can also be manually checked for compliance against a host profile.



After the host profile is created and associated with a set of hosts or clusters, you can check the compliance status in various places in vSphere Client:

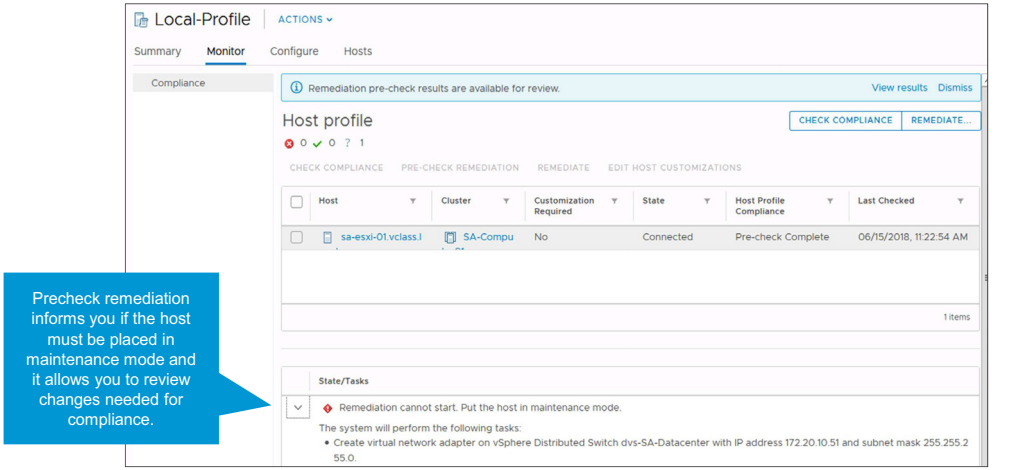
- Host profile's main view: Displays the compliance status of hosts.
- Host **Summary** tab: Displays the compliance status of the selected host.
- Host profile's **Monitor** tab: After you check compliance, this view displays the compliance status of the selected hosts. This view also displays any inconsistencies between the host values and the host profile values (shown on the slide).

Compliance can also be checked for individual components or settings. These built-in cluster compliance checks are useful for vSphere DRS and VMware vSphere® Distributed Power Management™.

https://t.me/learningnets

4-29 Applying Host Profiles

Hosts and clusters are brought into compliance by remediating them against a host profile.



After attaching the profile to an entity (a host or a cluster of hosts), configure the entity by applying the configuration as contained in the profile. Before you apply the changes, notice that vCenter Server displays a detailed list of actions that will be performed on the host to configure it and bring it into compliance. This list of actions enables the administrator to cancel applying the host profile if the parameters shown are not wanted.

Certain host profile policy configurations require that the host be rebooted after remediation. In those cases, you are prompted to place the host into maintenance mode. You might be required to place hosts into maintenance mode before remediation. Hosts that are in a fully-automated vSphere DRS cluster are placed in maintenance mode at remediation. For other cases, the remediation process stops if the host is not placed in maintenance mode when it is needed to remediate a host.

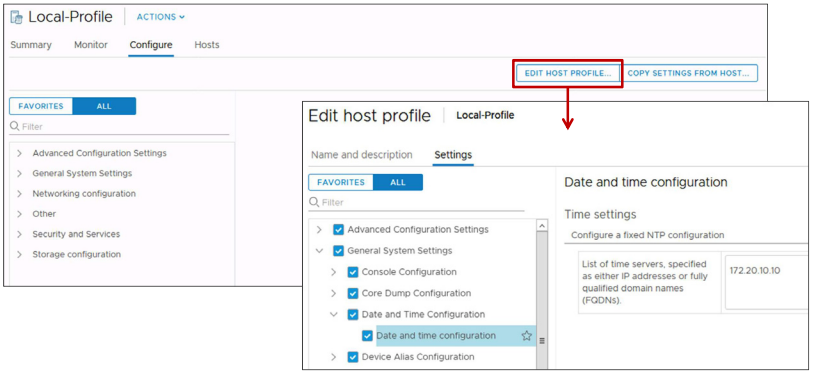
https://t.me/learningnets

4-30 Customizing Hosts

A host customization is created when the profile is first applied to a particular host.

A host customization contains settings and configuration values that are unique to a host, such as IP addresses, MAC addresses, and iSCSI qualified name.

The host customization is useful when provisioning hosts with vSphere Auto Deploy.



For hosts provisioned with VMware vSphere® Auto Deploy™, vCenter Server manages the entire host configuration, which is captured in a host profile. In most cases, the host profile information is sufficient to store all configuration information. Sometimes, the user is prompted for input when the host provisioned with vSphere Auto Deploy boots. The host customization mechanism manages such cases. After the user specifies the information, the system generates a host-specific customization and stores it with the vSphere Auto Deploy cache and the vCenter Server host object.

NOTE

In some documentation, you might see references to a host profile answer file. Answer file is equivalent to host customization.

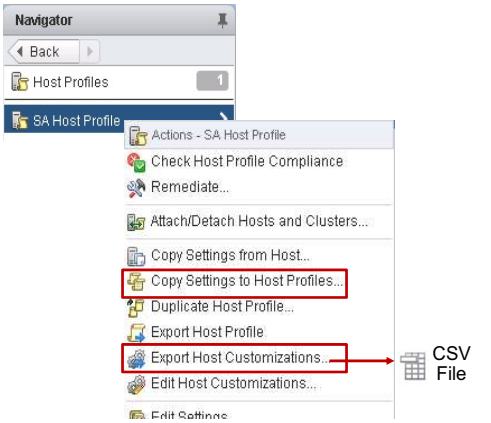
You can check the status of a host customization. If the host customization has all of the required user input answers, then the status is Complete. If the host customization is missing some of the required user input answers, then the status is Incomplete. You can also update or change the user input parameters for the host profile's policies in the host customization.

https://t.me/learningnets

4-31 Managing Host Customizations

You can copy settings across host profiles. This functionality enables you to maintain a golden master host profile for global settings that are common across all host profiles.

You can edit host customizations by exporting the data to a CSV file and then editing the CSV file. This functionality enables you to create host customizations offline.

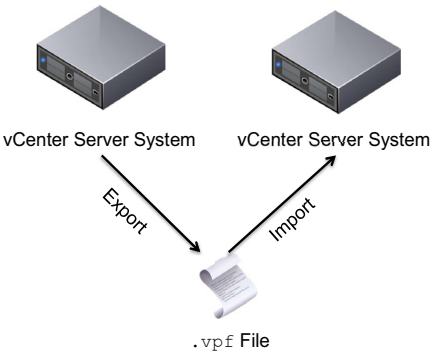


If a host profile contains any customized attributes, you can export them to a CSV file on your desktop. For security, sensitive data such, as passwords, is not exported. After the file is saved to your desktop, you can manually edit the file and save it to apply the customizations later.

https://t.me/learningnets

4-32 Standardization Across Multiple vCenter Server Instances

Organizations that span multiple vCenter Server instances can export the host profile from one vCenter Server instance and import the host profile for use in other vCenter Server instances.



Some organizations might have large environments that span many vCenter Server instances. If you want to standardize a single host configuration across multiple vCenter Server environments, you can export the host profile from one vCenter Server instance so that other vCenter Server instances can import the profile for use in their environments. Host profiles are not replicated, shared, or kept synchronized across vCenter Server instances connected by Enhanced Linked Mode.

Files are exported in the VMware profile format (`.vpf`).

NOTE

When a host profile is exported, administrator and user profile passwords are not exported. This restriction is a security measure that stops passwords from being exported in plain text. After the profile is imported, you are prompted to reenter values for the password and the password is applied to a host.

https://t.me/learningnets

4-33 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe the components of host profiles
- Describe the benefits of host profiles
- Explain how host profiles operate
- Use a host profile to manage ESXi configuration compliance
- Describe the benefits of exporting and importing host profiles

https://t.me/learningnets



https://t.me/learningnets

4-35 Learner Objectives

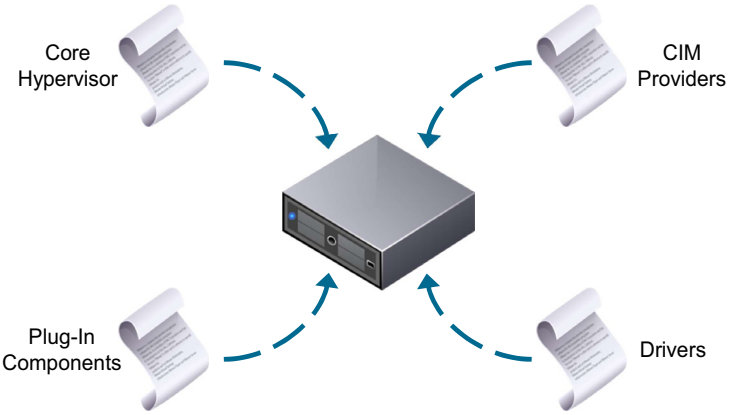
By the end of this lesson, you should be able to meet the following objectives:

- Explain the purpose of vSphere ESXi Image Builder
- Use vSphere Web Client to create an ESXi image

https://t.me/learningnets

4-36 ESXi Images

An ESXi image is a software bundle that consists of four main components.



An ESXi image contains all the software necessary to run on an ESXi host and includes the following components:

- The base ESXi software, also called the core hypervisor
- Specific hardware drivers
- Common Information Model (CIM) providers
- Specific applications or plug-in components

An ESXi image can be installed on a local or a SAN-based boot device, or it can be delivered at boot time using vSphere Auto Deploy.

https://t.me/learningnets

4-37 vSphere Installation Bundles

vSphere installation bundles (VIBs) are software packages that are added to an ESXi image.

A VIB is used to package any of the following ESXi image components:

- ESXi base image
- Drivers
- CIM providers
- Plug-ins and other components

A VIB specifies relationships with other VIBs:

- VIBs that the VIB depends on
- VIBs that the VIB conflicts with

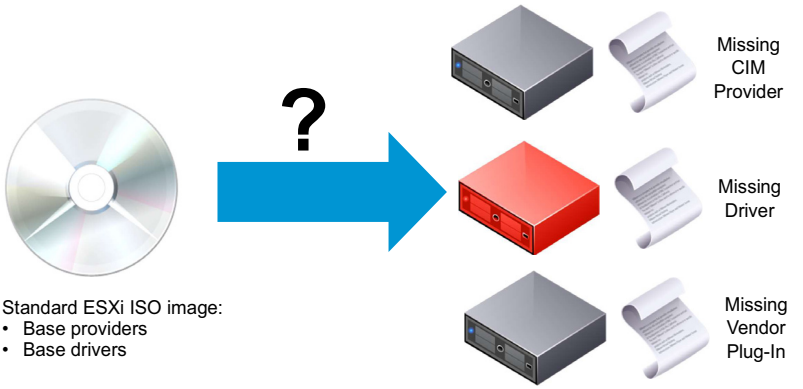
An ESXi image includes one or more vSphere installation bundles (VIBs). A VIB is an ESXi software package. VMware and its partners package solutions, drivers, CIM providers, and applications that extend the ESXi platform in VIBs.

An ESXi image should always contain one base VIB. Other VIBs can be added to include additional drivers, CIM providers, updates, patches, and applications.

https://t.me/learningnets

4-38 ESXi Image Deployment

The challenge of using a standard ESXi image is that the image might be missing desired functionality.



Standard ESXi images are provided by VMware and are available on the VMware website. ESXi images can also be provided by VMware partners.

The challenge that administrators face when using the standard ESXi image provided by VMware is that the standard image is sometimes limited in functionality. For example, the standard ESXi image might not contain all the drivers or CIM providers for a specific set of hardware, or the standard image might not contain vendor-specific plug-in components.

To create an ESXi image that contains custom components, use vSphere Web Client.

https://t.me/learningnets

4-39 vSphere ESXi Image Builder Overview

vSphere ESXi Image Builder lets you create and manage image profiles:

- An image profile is a group of VIBs that are used to create an ESXi image.

vSphere ESXi Image Builder enables the administrator to build customized ESXi boot images:

- Used for booting disk-based ESXi installations
- Used by vSphere Auto Deploy to boot an ESXi host from the network

You can access vSphere ESXi Image Builder by using the following interfaces:

- vCenter Web Client. You must first start the Image Builder service on the vCenter Server instance.
- PowerCLI cmdlets for vSphere ESXi Image.

vSphere ESXi Image Builder is a utility for customizing ESXi images. You access vSphere ESXi Image Builder through vSphere Web Client. vSphere ESXi Image Builder is used to create and manage VIBs, image profiles, software depots, and software channels.

For more information about vSphere ESXi Image Builder and its prerequisite software, see *VMware ESXi Installation and Setup* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

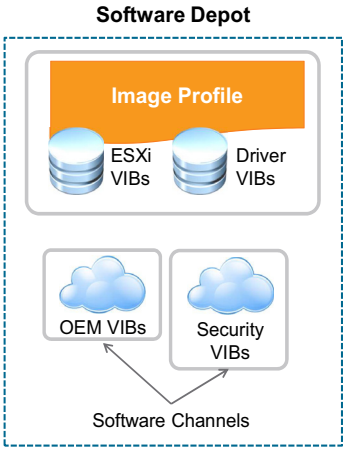
For more information about PowerCLI, see the documentation page at https://www.vmware.com/support/developer/PowerCLI.

https://t.me/learningnets

4-40 vSphere ESXi Image Builder Architecture

vSphere ESXi Image Builder consists of several components:

- VIB:
  - Is an ESXi software package that is provided by VMware and its partners
- Image profile:
  - Defines an ESXi image
  - Consists of one or more VIBs
- Software depot:
  - Is a collection of VIBs and image profiles
  - Can be online or offline:
    - The online depot is available through an HTTP URL.
    - The offline depot is available through a ZIP file.



VIBs are software packages that consist of packaged solutions, drivers, CIM providers, and applications that extend the ESXi platform. VIBs are available in software depots.

An image profile always includes a base VIB and typically includes other VIBs. You use vSphere Web Client to examine and define an image profile.

The software depot is a hierarchy of files and folders that can be available through an HTTP URL (online depot) or a ZIP file (offline depot or bundle). VMware has depots for you to use. The offline bundle is available on the Downloads page on the VMware website. Go to the Downloads page for VMware vSphere Hypervisor (ESXi) 6.7. The online software depot is located at https://hostupdate.vmware.com/software/VUM/PRODUCTION/main/vmw-depot-index.xml.

VMware partners also make software depots available. Companies with large VMware installations might also create internal depots to provision ESXi hosts with vSphere Auto Deploy or to export an ISO image for ESXi installation.

https://t.me/learningnets

4-41 Building ESXi Images

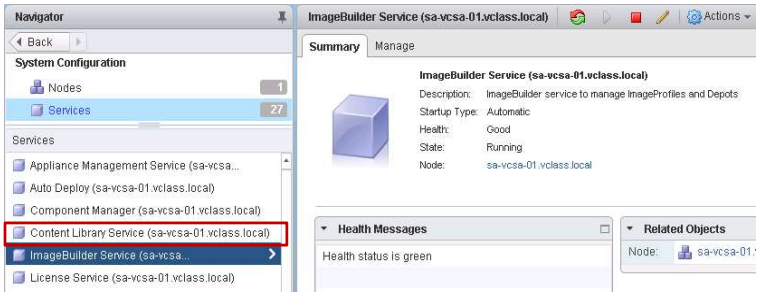
Building an ESXi image includes the following steps:

1. Starting the Image Builder service
2. Connecting to a software depot
3. Creating an image profile
4. Generating the new ESXi image

https://t.me/learningnets

4-42 Step 1: Starting the Image Builder Service

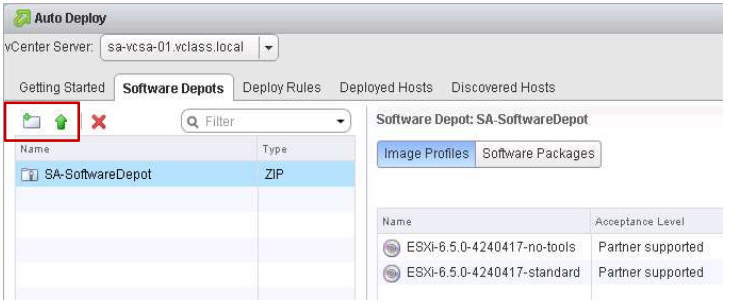
The Image Builder service enables you to manage and customize image profiles through vSphere Web Client.



https://t.me/learningnets

4-43 Step 2: Connecting to a Software Depot

You add or import a software depot, and verify that the software depot can be read.



Before you can create or customize an ESXi image, vSphere ESXi Image Builder must be able to access one or more software depots.

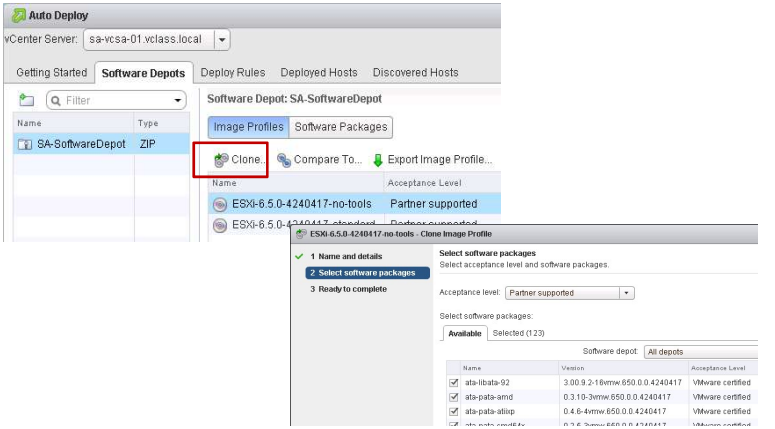
Use vSphere Web Client to add an ESXi software depot or offline depot ZIP file to vCenter Server. You can then create new image profiles and generate ISO files from the image profiles and VIBs in the depots.

After adding the software depot to vSphere ESXi Image Builder, verify that you can view the software packages found in the depot.

https://t.me/learningnets

4-44 Step 3: Creating an Image Profile

You clone and modify an image profile.

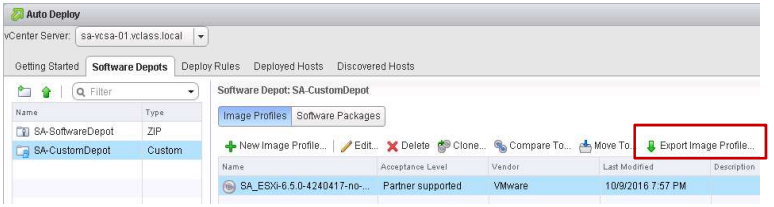


Cloning a published profile is the easiest way to create a custom image profile. Cloning a profile is useful when you want to remove a few VIBs from a profile. Cloning is also useful when you want to use hosts from different vendors and use the same basic profile, with the addition of vendor-specific VIBs. VMware partners or customers with large installations might consider creating a profile from scratch.

https://t.me/learningnets

4-45 Step 4: Generating the New ESXi Image

You generate a new ESXi image as either an ISO file or a ZIP file.



After creating an image profile, you can generate an ESXi image. You can export an image profile as an ISO image or ZIP file. An ISO image can be used to boot a host to install ESXi. A ZIP file can be used by VMware vSphere® Update Manager™ for remediating ESXi hosts. The exported image profile can also be used with vSphere Auto Deploy to boot ESXi hosts.

https://t.me/learningnets

4-46 Review of Learner Objectives

You should be able to meet the following objectives:

- Explain the purpose of vSphere ESXi Image Builder
- Use vSphere Web Client to create an ESXi image

https://t.me/learningnets



https://t.me/learningnets

4-48 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Explain the purpose of vSphere Auto Deploy
- Configure vSphere Auto Deploy
- Use vSphere Auto Deploy to deploy ESXi hosts

https://t.me/learningnets

4-49 About vSphere Auto Deploy

You use vSphere Auto Deploy to deploy ESXi hosts:

- The ESXi host's state and configuration run entirely in memory.
- When the ESXi host is shut down, the state information is cleared from memory.
- Preboot execution environment (PXE) is used to load the ESXi image.
- Works with vSphere ESXi Image Builder, vCenter Server, and host profiles.

vSphere Auto Deploy has the following benefits:

- Many ESXi hosts can be deployed quickly and easily.
- A standard ESXi image can be shared across many hosts.
- The host image is decoupled from the physical server which simplifies host recovery.

With vSphere Auto Deploy, the ESXi image is streamed across the network to the host and loaded directly into memory. All changes made to the state of the host are stored in memory and synchronized to the vCenter Server. When the host is shut down, the memory-based state of the host is lost but can be streamed into memory again from the vCenter Server when the host is powered back on.

vSphere Auto Deploy does not store the ESXi state on the host disk. vCenter Server stores and manages ESXi updates and patching through an image profile and, optionally, the host configuration through a host profile.

vSphere Auto Deploy enables the rapid deployment of many hosts. vSphere Auto Deploy simplifies ESXi host management by eliminating the necessity to maintain a separate boot image for each host. A standard ESXi image can be shared across many hosts.

Because the host image is decoupled from the physical server, the host can be recovered without having to recover the hardware or restore from a backup.

vSphere Auto Deploy can eliminate the need for a dedicated boot device, thus freeing the boot device for other uses.

For more information about vSphere Auto Deploy, see *VMware ESXi Installation and Setup* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

https://t.me/learningnets

4-50 vSphere Auto Deploy Modes

You can autodeploy hosts by using the following modes:

- Stateless
- Stateless caching
- Stateful install

Stateless is a mode of execution for ESXi hosts with no local storage:

- Every stateless ESXi boot is like a first boot.
- The host configuration is abstracted into a host profile.

Stateless mode has the following advantages:

- Easy replacement, removal, and addition of failed hardware
- Improved ease in scaling hardware deployment

The stateless caching and stateful install modes help minimize the dependency on the vSphere Auto Deploy service (on vCenter Server) and the PXE boot environment.

With stateless mode, the ESXi host boots with networking connectivity to vCenter Server through the built-in standard switch. If the host profile specifies distributed switch membership, vCenter Server joins the ESXi host to the distributed switch.

Using the stateless caching or the stateful installation mode greatly lessens the dependency on other systems for booting ESXi hosts. Hosts can be rebooted even when the vSphere Auto Deploy service, the vCenter Server, or the DHCP and TFTP servers are unavailable.

https://t.me/learningnets

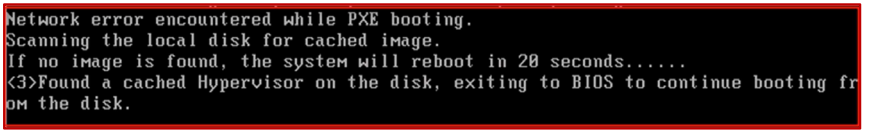
4-51 vSphere Auto Deploy Mode: Stateless Caching

vSphere Auto Deploy stateless caching saves the image and configuration to a local disk, but the host continues to perform stateless reboots.

Requirements include a dedicated boot device.

If the host is unable to reach the PXE host or the vSphere Auto Deploy service on vCenter Server, the host boots using the local image:

- The image on the local disk is used as a backup.
- Stateless caching can be configured to overwrite or preserve existing VMFS datastores.



vSphere Auto Deploy stateless caching PXE boots the ESXi host and loads the image in memory like stateless ESXi hosts. However, when the host profile is applied to the ESXi host, the image running in memory is copied to a boot device. The saved image acts as a backup in case the PXE infrastructure, the vCenter Server, or vSphere Auto Deploy service are unavailable. If the host needs to reboot and it cannot contact the DHCP, TFTP, or vSphere Auto Deploy service, the network boot will time out and the host will reboot using the cached disk image.

Although stateless caching might help ensure availability of an ESXi host by enabling the host to boot in the event of an outage affecting the DHCP, TFTP, or vCenter Servers, or the vSphere Auto Deploy service, stateless caching does not guarantee that the image is current or that the vCenter Server system is available after booting. The primary benefit of stateless caching is that it enables you to boot the host to help troubleshoot and resolve problems that prevent a successful PXE boot.

Unlike stateless ESXi hosts, stateless caching requires that a dedicated boot device be assigned to the host.

https://t.me/learningnets

4-52 vSphere Auto Deploy Mode: Stateful Installation

The ESXi host initially boots using vSphere Auto Deploy. All subsequent reboots use local disks.

vSphere Auto Deploy stateful installations enable you to quickly and efficiently provision hosts. After these hosts are provisioned, no further requirement is placed on the PXE host and the vSphere Auto Deploy service.

Using stateful installations has disadvantages:

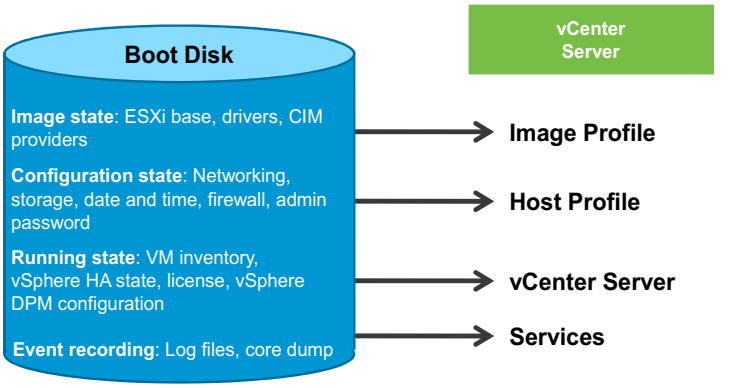
- Over time, the configuration might become out of sync with the vSphere Auto Deploy image.
- Patching and updating ESXi hosts must be done using traditional methods.

Setting up stateful installation is similar to configuring stateless caching. But instead of attempting to use vSphere Auto Deploy service on the vCenter Server to boot the host, the host does a one-time PXE boot to install ESXi. After the image is cached to disk, the host boots from the disk image on all subsequent reboots.

https://t.me/learningnets

4-53 Locations for Configuration and State Information

Because stateless vSphere Auto Deploy hosts are configured without a boot disk, all host configuration and state information is stored in or managed by vCenter Server.



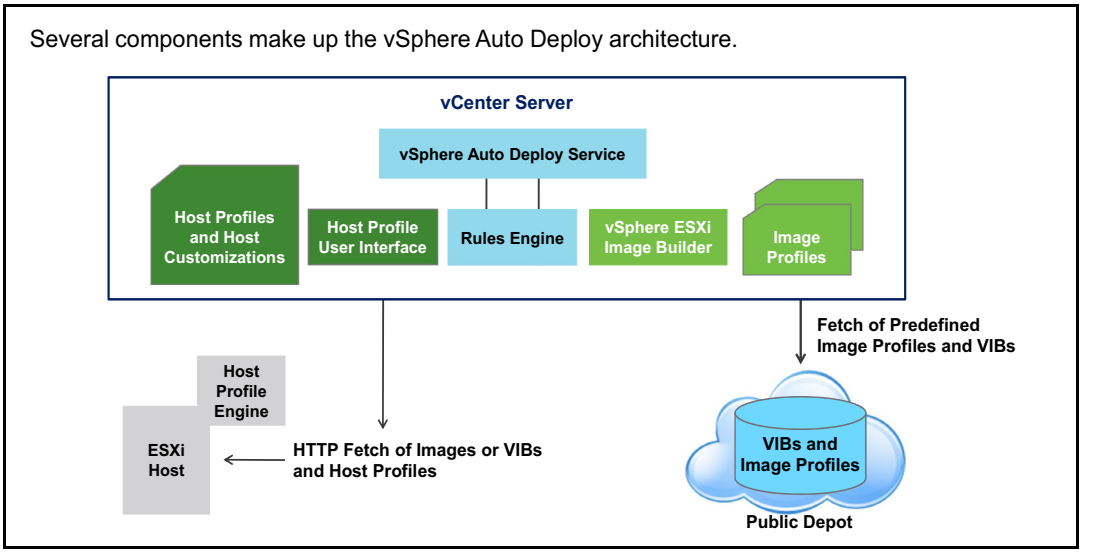
Without the use of vSphere Auto Deploy, the ESXi host's image (binaries and VIBs), configuration, state, and log files are stored on the boot device.

With Stateless vSphere Auto Deploy, a boot device no longer holds all of the host's information. The host information that is stored includes the following items:

- Image state: Executable software to run on the ESXi host. The information is part of the image profile, which can be created and customized with vSphere Web Client or vSphere ESXi Image Builder.
- Configuration state: Configurable settings that determine how the host is configured. Examples include virtual switch settings, boot parameters, and driver settings. Host profiles are created using the host profile user interface in vSphere Web Client.
- Running state: Settings that apply while the ESXi host is up and running. This state also includes the location of the virtual machine in the inventory and the virtual machine autostart information. This state information is managed by the vCenter Server instance.
- Event recording: Information found in log files and core dumps. This information can be managed by vCenter Server, using services like VMware vSphere® ESXi™ Dump Collector and VMware vSphere® Syslog Collector. vSphere Syslog Collector uses Linux rsyslogd.

https://t.me/learningnets

4-54 vSphere Auto Deploy Architecture



The vSphere Auto Deploy infrastructure consists of the following components:

- vSphere Auto Deploy service: Serves images and host profiles to ESXi hosts. The service, which runs on the vCenter Server, is at the heart of the vSphere Auto Deploy infrastructure.
- vSphere Auto Deploy rules engine: Tells the vSphere Auto Deploy service which image and which host profiles to serve to a host. You use vSphere Web Client to define the rules that assign image profiles and host profiles to hosts.
- Image profiles: Define the set of VIBs with which to boot ESXi hosts. VMware and its partners make image profiles and VIBs available in public software depots. Use vSphere Web Client to examine the depot and the vSphere Auto Deploy rules engine to specify which image profile to assign to a host. You can create a custom image profile based on the public image profiles and VIBs in the depot and apply that image profile to the host.
- Host profiles: Templates that define an ESXi host's configuration, such as networking or storage setup. You can save the host profile for an individual host and reuse this host profile to reprovision that host. You can save the host profile of a template host and use that profile for other hosts.
- Host customizations: Stores host-specific information. A host customization must be created for each host.

https://t.me/learningnets

4-55 Rules Engine Basics

vSphere Auto Deploy has a rules engine that determines which ESXi image profile and host profile can be used for each host.

The rules engine maps software profiles and host profiles to hosts, based on attributes of the host. For example, a rule can apply to all hosts or to a host with a specific IP or MAC address.

For new hosts, the vSphere Auto Deploy service checks with the rules engine before serving an image profile and host profile to a host.

The rules engine includes rules and rule sets.

The rules engine includes the following information:

- Rules: Rules can assign image profiles and host profiles to a set of hosts or specify the inventory location (folder or cluster) of a host on the target vCenter Server system. A rule can identify target hosts by boot MAC address, SMBIOS asset tag, BIOS UIID, or fixed DHCP IP address. In most cases, rules apply to multiple hosts. You use vSphere Web Client to create rules. After you create a rule, you must add it to a rule set. After you add a rule to a rule set, you cannot edit it.
- Active rule set: When a newly started host contacts the vSphere Auto Deploy service with a request for an image, the server checks the active rule set for matching rules. The image profile, host profile, and vCenter Server inventory location that are mapped by matching rules are then used to boot the host. If more than one item is mapped by the rules, the vSphere Auto Deploy service uses the item that is first in the rule set.
- Working rule set: The working rule set enables you to test changes to rules before making them active. For example, you can use vSphere Web Client to test compliance with the working rule set.

https://t.me/learningnets

4-56 PXE Boot Infrastructure Setup

vSphere Auto Deploy requires a PXE boot infrastructure:

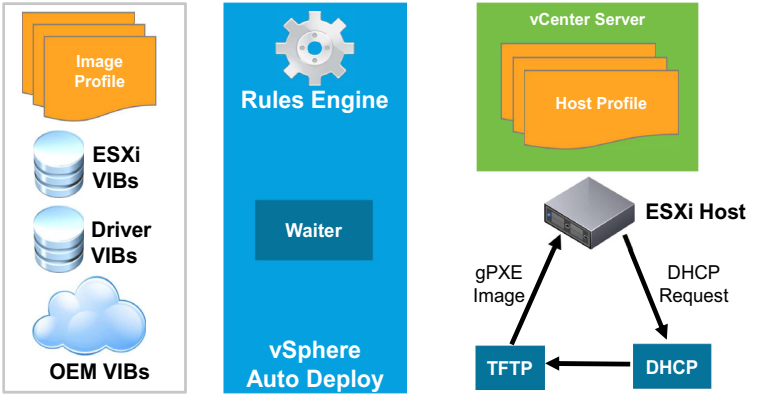
- The ESXi host must get its IP address from a DHCP server.
- The DHCP server must be configured to direct the ESXi host to a Trivial File Transfer Protocol (TFTP) server to download the PXE file.
- The following DHCP services can be used:
  - The organization's current DHCP server
  - A new DHCP server installed for vSphere Auto Deploy use
- A third-party TFTP server must be accessible from the DHCP server and the vCenter Server instance.
- Although vCenter Server Appliance has built-in DHCP and TFTP services, VMware does not support the use of these services with vSphere Auto Deploy.

A DHCP server and a TFTP server must be configured. The DHCP server assigns IP addresses to each autodeployed host on startup and points the host to a TFTP server to download the gPXE configuration files. The ESXi hosts can either use the infrastructure's existing DHCP and TFTP servers or new DHCP and TFTP servers can be created for use with vSphere Auto Deploy. Any DHCP server that supports the `next-server` and `filename` options can be used.

https://t.me/learningnets

4-57 Initial Boot of an Autodeployed ESXi Host (1)

The ESXi host boots from the PXE boot server.



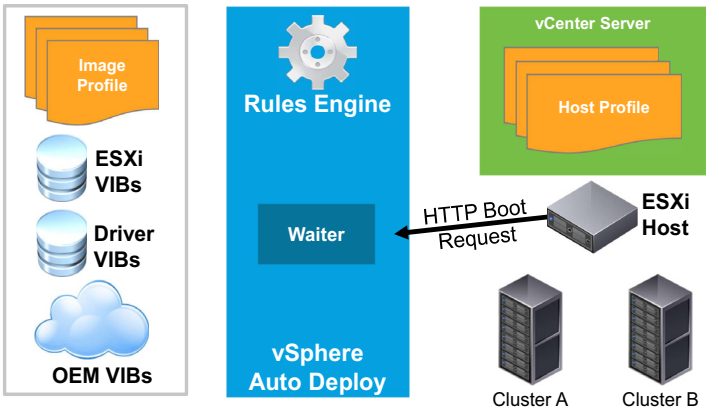
When an autodeployed host boots for the first time, the following events occur:

1. When the ESXi host is powered on, the ESXi host starts a PXE boot sequence.
2. The DHCP server assigns an IP address to the ESXi host and instructs the host to contact the TFTP server.
3. The ESXi host downloads the gPXE image file and the gPXE configuration file from the TFTP server.

https://t.me/learningnets

4-58 Initial Boot of an Autodeployed ESXi Host (2)

The ESXi host contacts the vSphere Auto Deploy service.

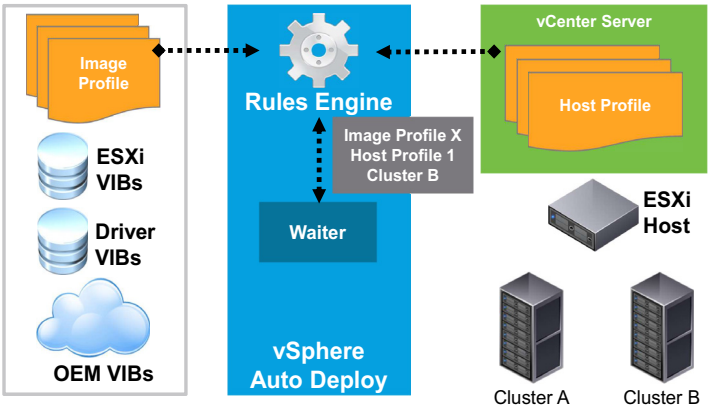


The gPXE configuration file instructs the host to make an HTTP boot request to the vSphere Auto Deploy service.

https://t.me/learningnets

4-59 Initial Boot of an Autodeployed ESXi Host (3)

The host image profile, host profile, and cluster are determined.



The vSphere Auto Deploy service queries the rules engine for the following information about the host:

- The image profile to use
- The host profile to use
- Which cluster the host belongs to, if any

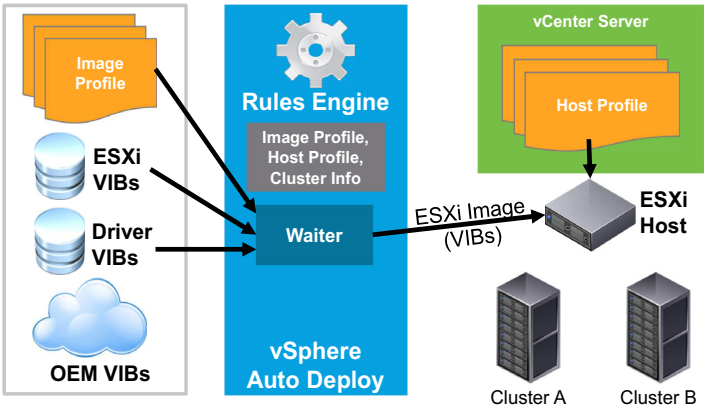
The rules engine maps software and configuration settings to hosts, based on the attributes of the host. For example, you can deploy image profiles or host profiles to two clusters of hosts by writing two rules. One rule specifies the host's location as cluster A, and the other rule specifies the host's location as cluster B.

For hosts that are not yet added to vCenter Server, vSphere Auto Deploy checks with the rules engine before serving image profiles and host profiles to hosts.

https://t.me/learningnets

4-60 Initial Boot of an Autodeployed ESXi Host (4)

The image is pushed to the host and the host profile is applied.



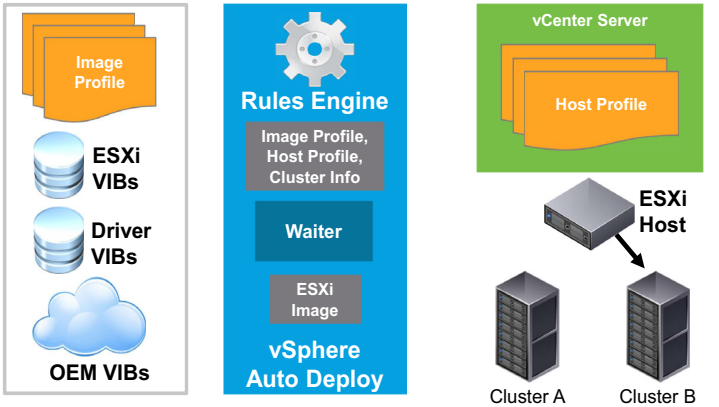
The vSphere Auto Deploy service streams VIBs specified in the image profile to the host.

The host boots based on the image profile and the host profile received from vSphere Auto Deploy. vSphere Auto Deploy also assigns the host to the vCenter Server instance with which it is registered.

https://t.me/learningnets

4-61 Initial Boot of an Autodeployed ESXi Host (5)

The host is placed in the appropriate cluster, if specified by a rule. The ESXi image and information on the image profile, host profile, and cluster to use are held on the vCenter Server instance for use by the vSphere Auto Deploy service.



If a rule specifies a target folder or a cluster on the vCenter Server instance, the host is added to that location. If no rule exists, vSphere Auto Deploy adds the host to the first data center.

If a host profile was used without a host customization and the profile requires specific information from the user, the host is placed in maintenance mode. The host remains in maintenance mode until the user reapplies the host profile and answers any outstanding questions.

If the host is part of a fully-automated vSphere DRS cluster, the cluster rebalances itself based on the new host, moving virtual machines onto the new host.

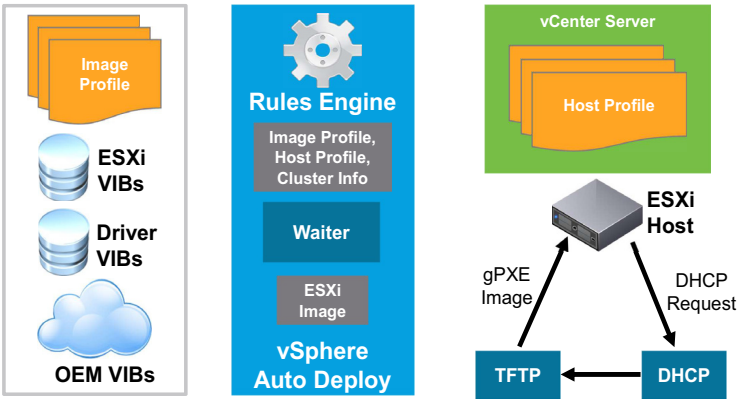
To make subsequent boots quicker, the vSphere Auto Deploy service stores the ESXi image as well as the following information:

- The image profile to use
- The host profile to use
- The location of the host in the vCenter Server inventory

https://t.me/learningnets

4-62 Subsequent Boot of a Stateless ESXi Host (1)

The autodeployed host is rebooted and the PXE boot sequence starts.



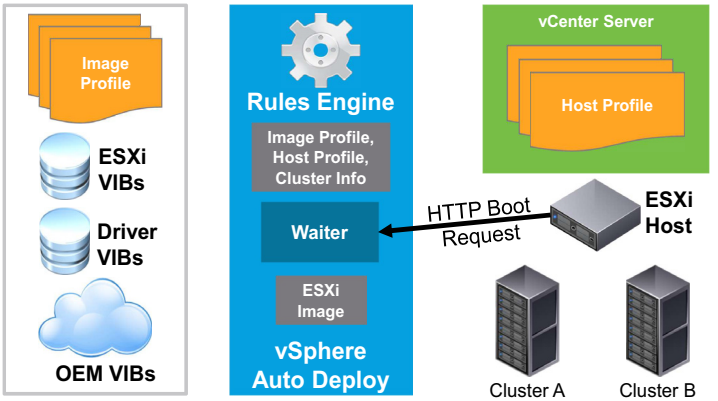
When an autodeployed ESXi host is rebooted, a slightly different sequence of events occurs:

1. The ESXi host goes through the PXE boot sequence, as it does in the initial boot sequence.
2. The DHCP server assigns an IP address to the ESXi host and instructs the host to contact the TFTP server.
3. The host downloads the gPXE image file and the gPXE configuration file from the TFTP server.

https://t.me/learningnets

4-63 Subsequent Boot of a Stateless ESXi Host (2)

The vSphere Auto Deploy service is contacted by the ESXi host.

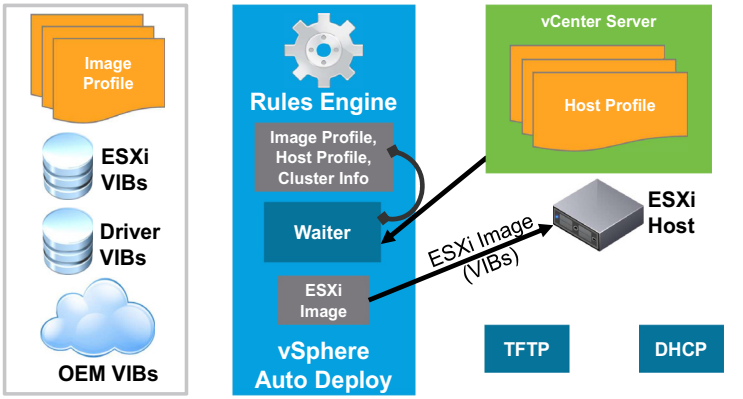


As in the initial boot sequence, the gPXE configuration file instructs the host to make an HTTP boot request to the vSphere Auto Deploy service.

https://t.me/learningnets

4-64 Subsequent Boot of a Stateless ESXi Host (3)

The ESXi image and the host profile are downloaded from vCenter Server to the host.



The subsequent boot sequence differs from the initial boot sequence.

When an ESXi host is booted for the first time, vSphere Auto Deploy queries the rules engine for information about the host. The information about the host's image profile, host profile, and cluster is stored on the vSphere Auto Deploy service.

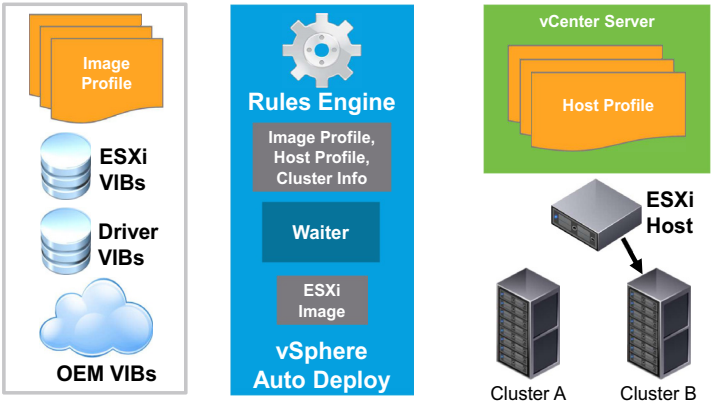
On subsequent reboots, the vSphere Auto Deploy service uses the saved information instead of using the rules engine to determine this information. Using the saved information saves time during subsequent boots because the host does not have to be checked against the rules in the active rule set. vSphere Auto Deploy checks the host against the active rule set only once and that is during the initial boot.

The vSphere Auto Deploy service delivers the image profile and host profile to the host.

https://t.me/learningnets

4-65 Subsequent Boot of a Stateless ESXi Host (4)

The host is placed in the appropriate cluster.



The ESXi host is placed in its assigned cluster on the vCenter Server instance.

https://t.me/learningnets

4-66 Activity: vSphere Auto Deploy Modes (1)

The stateful install and stateless caching modes help minimize the dependency on vCenter Server and the PXE boot environment.

- ❑ True
- ❑ False

https://t.me/learningnets

4-67 Activity: vSphere Auto Deploy Modes (2)

The stateful install and stateless caching modes help minimize the dependency on vCenter Server and the PXE boot environment.

- ✓ True
- ❑ False

https://t.me/learningnets

4-68 Running Custom Scripts on Stateless Hosts

You can configure one or more hosts by associating a bundle of custom scripts with a vSphere Auto Deploy rule. Scripts enable you to perform additional tasks whenever a stateless host boots.

The script bundle is in .tgz format, has a maximum size of 10 MB, and must be written in Python or the BusyBox ash scripting language.

Use the vSphere Auto Deploy PowerCLI cmdlet `Add-ScriptBundle`. This cmdlet adds a script bundle to the vSphere Auto Deploy inventory.

Then create a rule that assigns the script bundle to the appropriate hosts.

The scripts run in alphabetical order after the initial ESXi boot workflow of the host.

For more information about how to add a script bundle, see *VMware ESXi Installation and Setup* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

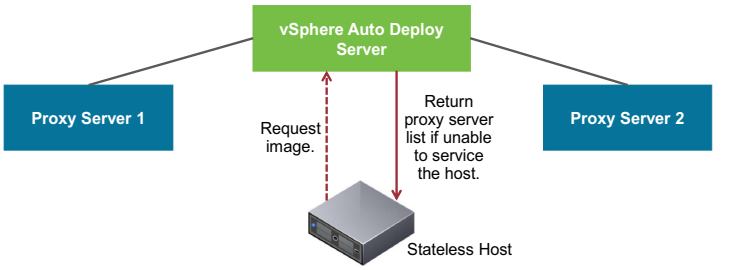
https://t.me/learningnets

4-69 Booting Many Stateless Hosts

Simultaneously booting many stateless hosts places a significant load on vCenter Server.

You can load balance the requests between the vSphere Auto Deploy service on vCenter Server and one or more cache proxy servers that you register with vSphere Auto Deploy.

Use the vSphere Auto Deploy PowerCLI cmdlet `Add-ProxyServer` to add a cache proxy server.



When a stateless host boots, the host requests an image from the vSphere Auto Deploy service. Under normal circumstances, the vSphere Auto Deploy service can serve an image to the host. However, if the network is saturated because hundreds of stateless hosts are booting at the same time, the vSphere Auto Deploy might take a long time to respond to a request.

Proxy servers (which are off-the-shelf proxy servers) help lessen the load on the vSphere Auto Deploy service which is on vCenter Server. If you have proxy servers fronting the vSphere Auto Deploy service, then in an extreme situation where vSphere Auto Deploy cannot serve the image, the vSphere Auto Deploy service sends a list of proxy servers to the stateless host. The stateless host tries to boot from the first proxy server on the list. If this proxy server is available, then the image is served to the host. But if the first proxy server is overloaded or unavailable, then the stateless host requests an image from the second proxy server in the list.

For information about how to register a caching proxy server address with vSphere Auto Deploy, see *VMware ESXi Installation and Setup* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

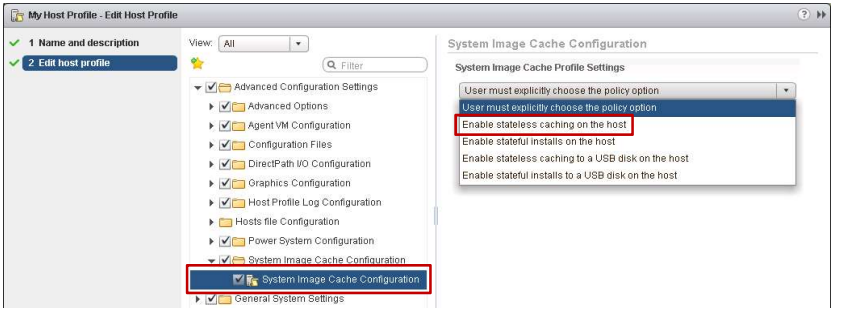
https://t.me/learningnets

4-70 Stateless Caching Host Profile Configuration

You use host profiles to configure auto deploy hosts for stateless caching:

- Create a host profile with stateless caching configured.
- Use vSphere Auto Deploy to boot the ESXi host. The host profile is applied and the ESXi image is cached to disk.

The host runs stateless under normal operations.

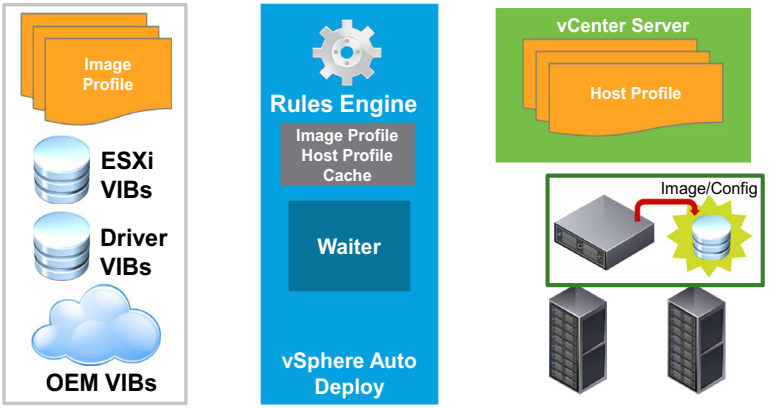


Stateless caching is configured using host profiles. When configuring stateless caching, you can choose to save the image to a local boot device or to a USB disk. You can also leave the VMFS datastore on the boot device.

https://t.me/learningnets

4-71 Subsequent Boot of a Stateless Caching Host

The host copies the state locally. Reboots are stateless when the PXE and Auto Deploy infrastructure is available.



vSphere Auto Deploy caches the image when you apply the host profile if Enable stateless caching on the host is selected in the System Cache Configuration host profile. When you reboot the host, it continues to use the vSphere Auto Deploy infrastructure to retrieve its image. If the vSphere Auto Deploy service is not available, the host uses the cached image.

If the ESXi hosts that run your virtual machines lose connectivity to the vSphere Auto Deploy service or to the vCenter Server system, some limitations apply the next time that you reboot the host:

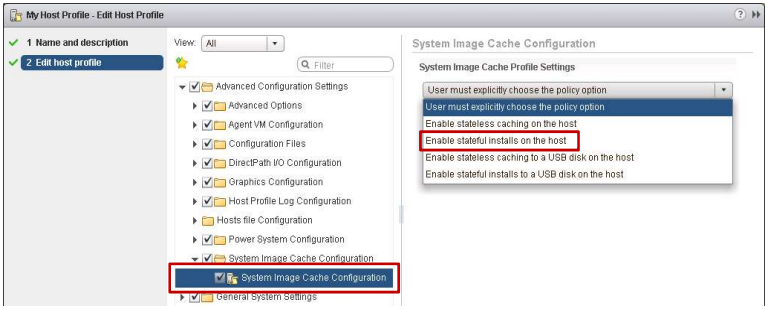
- If the vCenter Server system is available but the vSphere Auto Deploy service is unavailable, hosts do not connect to the vCenter Server system automatically. You can manually connect the hosts to the vCenter Server system or wait until the vSphere Auto Deploy service is available again.
- If the vCenter Server system is unavailable (which implies that the vSphere Auto Deploy service is unavailable), you can connect to each ESXi host by using VMware Host Client™ and add virtual machines to each host.
- If the vCenter Server system is not available, vSphere DRS does not work. The vSphere Auto Deploy service cannot add hosts to vCenter Server. You can connect to each ESXi host by using VMware Host Client and add virtual machines to each host.
- If you make changes to your setup while connectivity is lost, the changes are lost when the connection to the vSphere Auto Deploy service is restored.

https://t.me/learningnets

4-72 Stateful Installation Host Profile Configuration

To configure a stateful installation host profile:

1. Create the host profile with stateful installation configured.
2. Use vSphere Auto Deploy to boot the ESXi host. The ESXi image is saved to disk and the host profile is applied.
3. Direct future boots to the local disk. No future attempts are made to contact the PXE server or the vSphere Auto Deploy service.

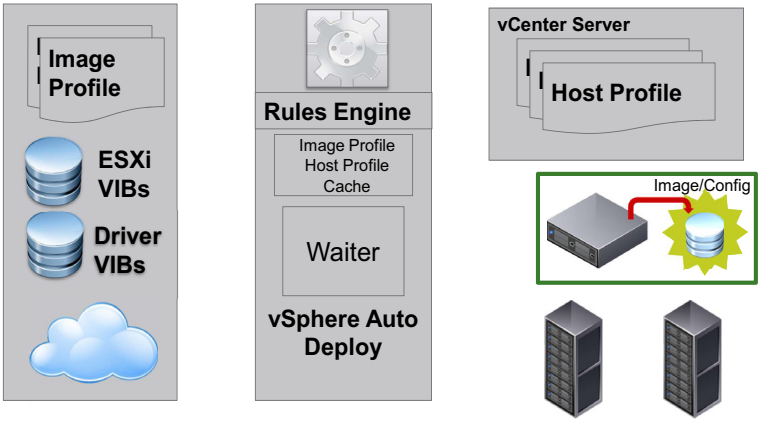


Stateful installation is configured using host profiles. When configuring stateful installation, you can choose to save the image to a local boot device or to a USB device. You can also leave the VMFS datastore on the boot device.

https://t.me/learningnets

4-73 Subsequent Boot of a Stateful Installation Host

Initial bootup uses vSphere Auto Deploy to install the image on the server. Subsequent reboots are performed from local storage.



With stateful installation, the vSphere Auto Deploy service is used to provision new ESXi hosts. The first time that the host boots, it uses the PXE host, the DHCP server, and the vSphere Auto Deploy service like a stateless host. All subsequent reboots use the image that is saved to the local disk.

After the image is written to the host's local storage, the image is used for all future reboots. The initial boot of a system is the equivalent of an ESXi host installation to a local disk. Because no attempt is made to update the image, the image can become stale over time. Manual processes must be implemented to manage configuration updates and patches, losing most of the key benefits of vSphere Auto Deploy.

https://t.me/learningnets

4-74 Configuring a vSphere Auto Deploy Environment

Before configuring vSphere Auto Deploy, ensure that host profiles and image profiles for use by autodeployed hosts are created.

Configuring a vSphere Auto Deploy environment includes the following steps:

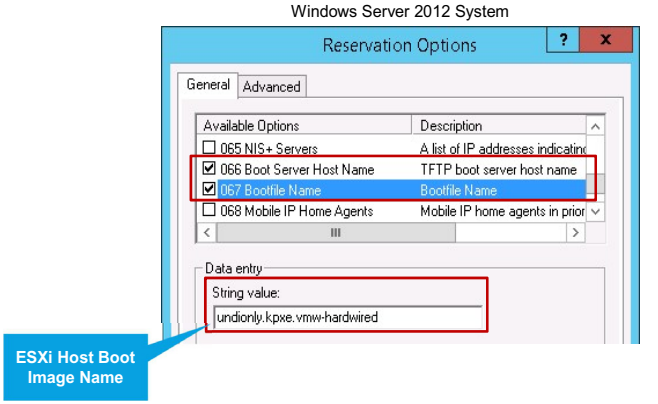
1. Preparing the DHCP server
2. Starting the vSphere Auto Deploy service
3. Preparing the TFTP server
4. Creating deployment rules
5. Activating deployment rules

You need a DHCP server and a TFTP server to PXE boot the ESXi installer. To PXE boot the ESXi installer, the DHCP server must send the address of the TFTP server and the filename of the initial boot loader to the ESXi host.

https://t.me/learningnets

4-75 Step 1: Preparing the DHCP Server

You set up the DHCP server to point the ESXi hosts to the TFTP server and the ESXi host boot image.



You must set up the DHCP server to serve each target ESXi host with an iPXE binary. In this example, the DHCP server is included with Windows Server 2012.

The `undionly.kpxe.vmw-hardwired` file is an iPXE binary that is used to boot the ESXi hosts.

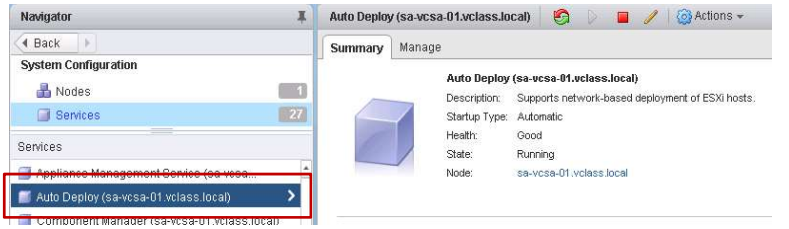
See *VMware ESXi Installation and Setup* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

https://t.me/learningnets

4-76 Step 2: Starting the vSphere Auto Deploy Service

vSphere Auto Deploy is part of the vCenter Server collection of services.

Ensure that the vSphere Auto Deploy service is enabled.

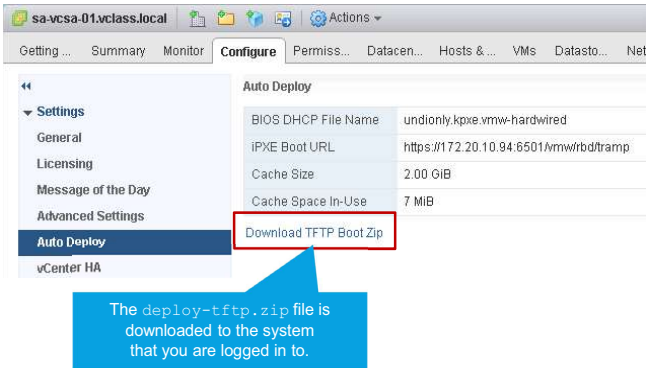


Before you can use vSphere Auto Deploy, you must configure the vSphere Auto Deploy service startup type in the vCenter Server system that you plan to use for managing the hosts that you provision.

https://t.me/learningnets

4-77 Step 3: Preparing the TFTP Server

You download the TFTP boot ZIP file from vCenter Server to the TFTP server's root directory (defined by `TFTP_Root`). The ZIP file contains `undionly.kpxe.vmw-hardwired`.



After you prepare the DHCP server, you must start the Auto Deploy service on vCenter Server and configure the TFTP server. You must download a TFTP boot ZIP file from your vSphere Auto Deploy service. The TFTP server serves the boot images that vSphere Auto Deploy provides.

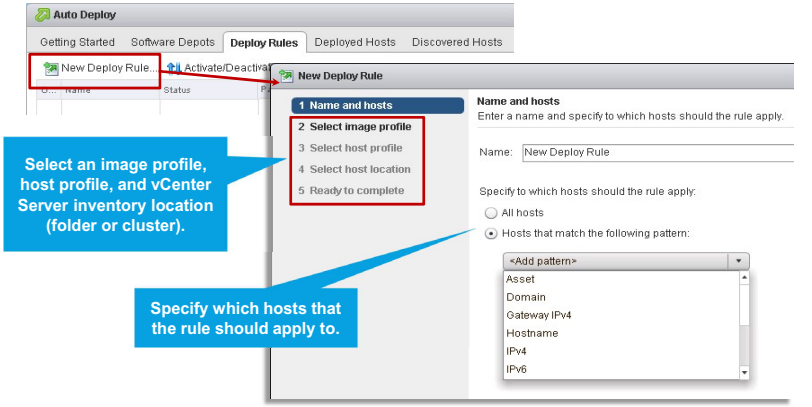
When you click **Download TFTP Boot Zip**, the `deploy-tftp.zip` file is downloaded to the system on which vSphere Web Client is running. You unzip this file and copy its contents to the TFTP server's root directory. The TFTP root directory is defined by the `TFTP_Root` parameter on the TFTP server.

For more information about how to configure the TFTP environment for vSphere Auto Deploy provisioning, see *VMware ESXi Installation and Setup* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

https://t.me/learningnets

4-78 Step 4: Creating Deployment Rules

You create a rule that identifies the target hosts, and the image profile, host profile, and inventory location to use when autodeploying hosts.



Rules can assign image profiles and host profiles to a set of hosts, or specify the location (folder or cluster) of a host on the target vCenter Server system. A rule can identify target hosts by boot MAC address, SMBIOS information, BIOS UUID, vendor, model, or fixed DHCP IP address. In most cases, rules apply to multiple hosts.

After you create a rule, you must add it to a rule set. Only two rule sets are supported:

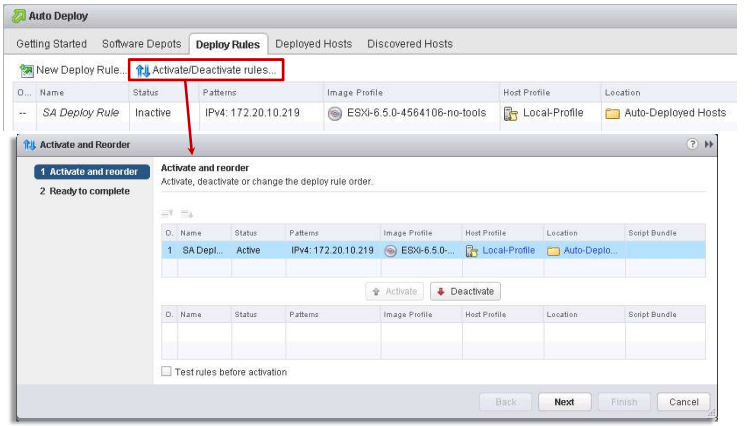
- The active rule set
- The working rule set.

A rule can belong to both sets (the default) or to only the working rule set. After you add a rule to a rule set, you can no longer change the rule. Instead, you copy the rule and replace items or patterns in the copy. If you are managing vSphere Auto Deploy with vSphere Web Client, you can edit a rule if it is in the inactive state.

https://t.me/learningnets

4-79 Step 5: Activating Deployment Rules

After you create a rule, you must activate the rule for it to take effect.



After you create a vSphere Auto Deploy rule, the rule is in the inactive state. You must activate the rule for it to take effect. You can use the Activate and Reorder wizard to activate, deactivate, and change the order of the rules.

https://t.me/learningnets

Consider the following guidelines for managing your vSphere Auto Deploy environment:

- If you change a rule set, unprovisioned hosts use the new rules automatically when they are booted. For all other hosts, vSphere Auto Deploy applies new rules only when you test a host's rule compliance and perform remediation.
- vCenter Server resides in your management cluster. The hosts in this cluster must not be managed by vSphere Auto Deploy.
- vCenter Server does not need to be available for reboots of stateless caching hosts or stateful install hosts.

You can change a rule set, for example, to require a host to boot from a different image profile. You can also require a host to use a different host profile. Unprovisioned hosts that you boot are automatically provisioned according to these modified rules. For all other hosts, vSphere Auto Deploy applies the new rules only when you test their rule compliance and perform remediation.

If vCenter Server becomes unavailable, stateless hosts that are already autodeployed remain up and running. However, you will be unable to boot or reboot hosts. VMware recommends installing vCenter Server in a virtual machine and placing the virtual machine in a vSphere HA cluster to keep the virtual machine available.

For information about the procedure and the commands used for testing and repairing rule compliance, see *VMware ESXi Installation and Setup* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

https://t.me/learningnets

4-81 Using vSphere Auto Deploy with vSphere Update Manager

VMware vSphere® Update Manager™ supports hosts that boot using vSphere Auto Deploy.

vSphere Update Manager can patch hosts but cannot update the ESXi image used to boot the host.

vSphere Update Manager can remediate only patches that do not require a reboot (live install). Patches requiring reboot cannot be installed.

The workflow for patching includes the following steps:

1. Manually update the image that vSphere Auto Deploy uses with patches. If rebooting is possible, rebooting is all that is required to update the host.
2. If a reboot cannot be performed, create a baseline in vSphere Update Manager and remediate the host.

VMware vSphere® Update Manager™ can be used to upgrade hosts that boot using vSphere Auto Deploy. Image profiles must be patched and modified using vSphere Web Client or vSphere ESXi Image Builder.

Only live-install patches can be remediated with vSphere Update Manager. Any patch that requires a reboot cannot be installed on a PXE host. The live-install patches can be from VMware or from a third party.

https://t.me/learningnets

4-82 Introduction to Lab 7: Using vSphere Auto Deploy

The diagram shows the systems used in this lab.



The TFTP and DHCP services available with vCenter Server Appliance are not supported for vSphere Auto Deploy in production environments.

In this lab, you use the TFTP service on sa-vcsa-01.vclass.local. Use the TFTP service that is available in vCenter Server Appliance only for testing purposes. You use the DHCP service on dc.vclass.local, a Windows Server 2012 system. The host to autodeploy is sa-esxi-04.vclass.local.

https://t.me/learningnets

4-83 Lab 7: Using vSphere Auto Deploy

Configure vSphere Auto Deploy on vCenter Server Appliance to boot stateless hosts

1. Create a Folder for Autodeployed Hosts
2. Start the vSphere Auto Deploy Service
3. Start the vSphere ESXi Image Builder Service
4. Import a Software Depot and Create a Custom Depot
5. Create a Custom Image Profile and Export the Image Profile
6. Create and Activate a Deployment Rule
7. Configure DHCP
8. Start the TFTP Service on vCenter Server Appliance
9. Review the Autodeployment Preparation Steps
10. Prepare to Monitor ESXi Bootup During the Autodeploy Process
11. Power On the ESXi Host and Monitor the Bootup Process
12. Verify Host Profile Compliance of the Autodeployed Host
13. Review the Noncompliant Items

https://t.me/learningnets

4-84 Review of Learner Objectives

You should be able to meet the following objectives:

- Explain the purpose of vSphere Auto Deploy
- Configure vSphere Auto Deploy
- Use vSphere Auto Deploy to deploy ESXi hosts

https://t.me/learningnets

4-85 Key Points

- The Host Profiles feature enables you to export configuration settings from a master reference host and save them as a host profile. You can use the host profile to quickly configure other hosts in the data center.
- Content libraries provide a centralized, easily-managed repository for files used with vSphere.
- You use vSphere Auto Deploy to deploy large numbers of ESXi hosts quickly and easily.

Questions?

https://t.me/learningnets

https://t.me/learningnets

MODULE 5

CPU Optimization



https://t.me/learningnets

5-2 You Are Here

1. Course Introduction
2. Network Scalability
3. Storage Scalability
4. Host and Management Scalability
5. **CPU Optimization**
6. Memory Optimization
7. Storage Optimization
8. Network Optimization
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

5-3 Importance

In a vSphere environment, multiple virtual machines run on the same host. To prevent a situation where virtual machines are allocated insufficient CPU resources, you must know how to monitor both host and virtual machine CPU usage.

https://t.me/learningnets

5-4 Module Lessons

Lesson 1:	CPU Virtualization Concepts
Lesson 2:	NUMA Architecture in vSphere
Lesson 3:	Monitoring CPU Activity

https://t.me/learningnets



https://t.me/learningnets

5-6 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Discuss the CPU scheduler features
- Discuss what affects CPU performance

https://t.me/learningnets

5-7 About the CPU Scheduler

The CPU scheduler is crucial to providing good performance in a consolidated environment.

The CPU scheduler has multiple functions:

- Schedules virtual CPUs (vCPUs) on physical CPUs
- Enforces the proportional-share algorithm for CPU usage
- Supports symmetric multiprocessing (SMP) virtual machines
- Uses relaxed co-scheduling for SMP virtual machines

The CPU scheduler is aware of non-uniform memory access (NUMA), processor, and cache topology.

Because most modern processors are equipped with multiple cores per processor, building a system with tens of cores running hundreds of virtual machines is easy. In such a large system, allocating CPU resources efficiently and fairly is critical.

https://t.me/learningnets

5-8 About Virtual Machine Worlds

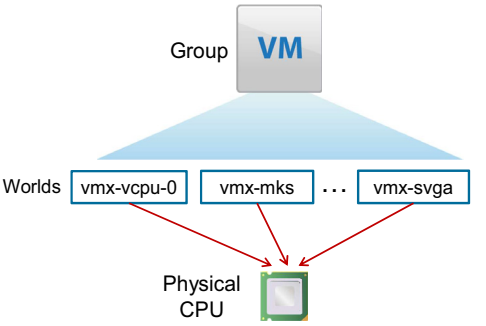
A world is an execution context that is scheduled on a processor. A world is like a process in conventional operating systems.

A virtual machine is a collection, or group, of worlds.

Different types of virtual machine worlds exist. Here are some examples:

- One for each vCPU (vmx-vcpu-0, vmx-vcpu-1, and so on)
- One for the virtual machine's mouse, keyboard, and screen (vmx-mks)
- One for the SVGA device (vmx-svga)

The CPU scheduler chooses which world to schedule on a processor.



The role of the CPU scheduler is to assign execution contexts to processors in a way that meets system objectives such as responsiveness, throughput, and usage. On conventional operating systems, the execution context corresponds to a process or a thread. On ESXi hosts, the execution context corresponds to a world.

Non-virtual machine worlds exist as well. These non-virtual machine worlds are VMkernel worlds, which are used to perform various system tasks. Examples of these non-virtual machine worlds include the `idle` and `vmotionServer` worlds.

https://t.me/learningnets

5-9 CPU Scheduler Features

The CPU scheduler allocates CPU resources and coordinates CPU usage.

The CPU scheduler uses dynamic and transparent CPU resource allocation:

- Checks physical CPU use every 2 to 30 milliseconds and migrates vCPUs as necessary

The CPU scheduler allows a vCPU to run on a physical CPU for 50 milliseconds (the default time slice) before another vCPU of the same priority gets scheduled.

While a vCPU is running on a physical CPU, other vCPUs must wait their turn on that physical CPU, which can introduce queuing.

The CPU scheduler enforces the proportional-share algorithm for CPU usage:

- Hosts time-slice physical CPUs across all virtual machines when CPU resources are overcommitted
- Prioritizes each vCPU by resource allocation settings: shares, reservations, and limits

One of the main tasks of the CPU scheduler is to choose which world is to be scheduled to a processor. If the target processor is already occupied, the scheduler must decide whether or not to preempt the currently-running world on behalf of the chosen world.

A world migrates from a busy processor to an idle processor. A world migration can be initiated either by a physical CPU that becomes idle or by a world that becomes ready to be scheduled.

An ESXi host implements the proportional-share-based algorithm. When CPU resources are overcommitted, the ESXi host time-slices the physical CPUs across all virtual machines so that each virtual machine runs as if it had its specified number of virtual processors. The ESXi host associates each world with a share of CPU resource.

This association of resources is called entitlement. Entitlement is calculated from user-provided resource specifications such as shares, reservation, and limits. When making scheduling decisions, the ratio of the consumed CPU resource to the entitlement is used as the priority of the world. If a world has consumed less than its entitlement, the world is considered high priority and is likely to be chosen to run next.

https://t.me/learningnets

5-10 CPU Scheduler Support for SMP Virtual Machines

ESXi uses a form of co-scheduling that is optimized to run SMP virtual machines efficiently.

Co-scheduling is a technique for scheduling related processes to run on different processors at the same time. At any time, each vCPU might be scheduled, descheduled, preempted, or blocked while waiting for an event.

The CPU scheduler takes skew into account when scheduling vCPUs in an SMP virtual machine:

- Skew is the difference in execution rates between vCPUs in an SMP virtual machine.
- A vCPU's skew increases when it is not making progress but one of its sibling vCPUs is.
- A vCPU is considered to be skewed if its cumulative skew exceeds a threshold.

A symmetric multiprocessing (SMP) virtual machine presents the guest operating system and applications with the illusion that they are running on a dedicated physical multiprocessor. An ESXi host implements this illusion by supporting co-scheduling of the virtual CPUs (vCPUs) in an SMP virtual machine.

Without co-scheduling, the vCPUs associated with an SMP virtual machine would be scheduled independently, breaking the guest's assumptions regarding uniform progress.

The progress of each vCPU in an SMP virtual machine is tracked individually. The skew is measured as the difference in progress between the slowest vCPU and each of the other vCPUs.

The ESXi scheduler maintains a detailed cumulative skew value for each vCPU in an SMP virtual machine. A vCPU is considered to be making progress if it consumes CPU at the guest level or if it halts. The time spent in the hypervisor is excluded from the progress. This exclusion means that the hypervisor execution might not always be co-scheduled. This behavior is acceptable because not all operations in the hypervisor benefit from being co-scheduled. When co-scheduling is beneficial, the hypervisor makes explicit co-scheduling requests to achieve good performance.

https://t.me/learningnets

5-11 Relaxed Co-Scheduling

Relaxed co-scheduling is a technique for scheduling a subset of a virtual machine's vCPUs simultaneously after skew is detected.

Relaxed co-scheduling lowers the number of required physical CPUs for the virtual machine to co-start and increases CPU use.

An idle vCPU has no scheduling overhead.

With relaxed co-scheduling, only the vCPUs that are skewed must be co-started. Relaxed co-scheduling ensures that when any vCPU is scheduled, all other vCPUs that are behind are also scheduled, reducing skew. This approach is called relaxed co-scheduling because only a subset of a virtual machine's vCPUs must be scheduled simultaneously after skew is detected.

The vCPUs that advanced too much are individually stopped. After the lagging vCPUs catch up, the stopped vCPUs can start individually. Co-scheduling all vCPUs is still attempted to maximize the performance benefit of co-scheduling.

An idle CPU has no co-scheduling overhead because the skew does not increase when a vCPU halts. An idle vCPU does not accumulate skew and is treated as if it were running for co-scheduling purposes. This optimization ensures that idle guest vCPUs do not waste physical processor resources, which can instead be allocated to other virtual machines.

For example, an ESXi host with two physical cores might be running one vCPU each from two different virtual machines, if their sibling vCPUs are idling, without incurring co-scheduling overhead.

https://t.me/learningnets

5-12 CPU Performance Factors

The following factors impact CPU performance:

- Idling virtual machines: Consider the overhead of delivering guest timer interrupts.
- CPU affinity: CPU affinity constrains the scheduler and can cause an imbalanced load.
- SMP virtual machines: Some co-scheduling overhead is incurred.
- Insufficient CPU resources for demand: If CPU contention exists, the scheduler forces vCPUs of lower-priority virtual machines to queue their CPU requests in deference to higher-priority virtual machines.

You might think that idle virtual machines do not cost anything in terms of performance. However, timer interrupts still must be delivered to these virtual machines. Lower timer interrupt rates can help a guest operating system's performance.

Try to use as few vCPUs in your virtual machine as possible to reduce the amount of timer interrupts necessary, as well as to reduce any co-scheduling overhead that might be incurred. Also, use SMP kernels, and not uniprocessor kernels, in SMP virtual machines.

CPU capacity is a finite resource. Even on a server that allows additional processors to be configured, the number of processors that can be installed always has a maximum. As a result, performance problems might occur when insufficient CPU resources are available to satisfy demand.

https://t.me/learningnets

5-13 Warning Sign: Ready Time

Ready time is the amount of time that the vCPU waits for the physical CPU to become available.

vCPUs are allocated CPU cycles on an assigned physical CPU based on the proportional share algorithm enforced by the CPU scheduler:

- If a vCPU tries to execute a CPU instruction while no cycles are available on the physical CPU, the request is queued.
- A physical CPU with no available cycles might be due to high load on the physical CPU, or a higher-priority vCPU receiving preference on that physical CPU.

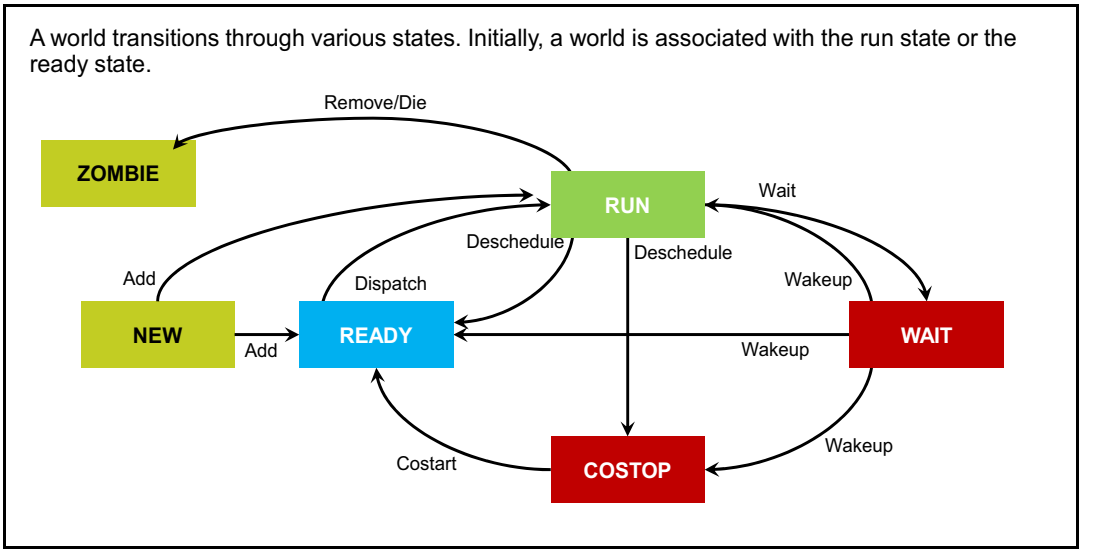
Ready time can affect performance of the guest operating system and its applications in a virtual machine.

To achieve best performance in a consolidated environment, you must consider ready time. Ready time is the time that a virtual machine must wait in the queue in a ready-to-run state before it can be scheduled on a CPU.

When multiple processes are trying to use the same physical CPU, that CPU might not be immediately available, and a process must wait before the ESXi host can allocate a CPU to it. The CPU scheduler manages access to the physical CPUs on the host system.

https://t.me/learningnets

5-14 Transitioning Between World States



When first added, a world is either in the run state or the ready state, depending on the availability of a physical CPU. A world in the ready state is dispatched by the CPU scheduler and enters the run state. It can be later descheduled and enters either the ready state or the costop state. The co-stopped world is co-started later and enters the ready state. A world in the run state enters the wait state by blocking on a resource. It is woken up when the resource becomes available.

A world becoming idle enters the wait_idle state, a special type of wait state, although it is not explicitly blocking on a resource. An idle world is woken up whenever it is interrupted.

https://t.me/learningnets

5-15 Activity: CPU Scheduler (1)

What tasks does the CPU scheduler perform? Select all that apply.

- ❑ Schedules vCPUs on physical CPUs
- ❑ Creates a world for each virtual machine to run in
- ❑ Determines which worlds are entitled to CPU time
- ❑ Migrates worlds from busy processors to idle processors

https://t.me/learningnets

5-16 Activity: CPU Scheduler (2)

What tasks does the CPU scheduler perform? Select all that apply.

- ✓ Schedules vCPUs on physical CPUs
- ❑ Creates a world for each virtual machine to run in
- ✓ Determines which worlds are entitled to CPU time
- ✓ Migrates worlds from busy processors to idle processors

https://t.me/learningnets

5-17 Review of Learner Objectives

You should be able to meet the following objectives:

- Discuss the CPU scheduler features
- Discuss what affects CPU performance

https://t.me/learningnets



https://t.me/learningnets

5-19 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Explain how ESXi hosts are NUMA-aware
- Explain the virtual NUMA topology
- Describe how to optimize the number of cores per socket for a VM

https://t.me/learningnets

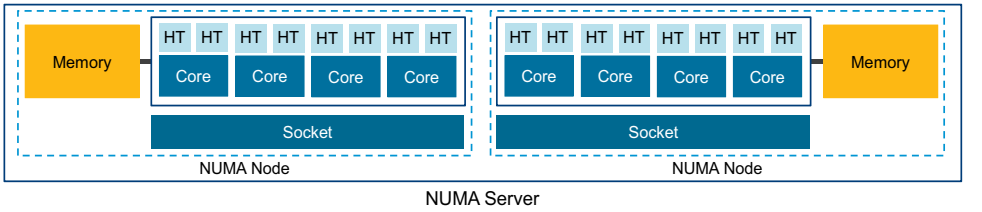
5-20 NUMA-Aware ESXi

ESXi supports memory access optimization for processors in server architectures that support NUMA.

Each CPU on a NUMA node has local memory directly connected by one or more local memory controllers.

Processes running on a CPU can access this local memory faster than memory on a remote CPU on the same server.

The ESXi NUMA scheduler dynamically balances processor load and attempts to maintain good NUMA locality (where a high percentage of a VM's memory is local).



A non-uniform memory access (NUMA) node contains processors and memory, much like a small SMP system. However, an advanced memory controller allows a node to use memory on all other nodes, creating a single system image.

In a NUMA host, the delay incurred when accessing memory varies for different memory locations. When poor NUMA locality occurs, the VM's performance might be less than if its memory were all local.

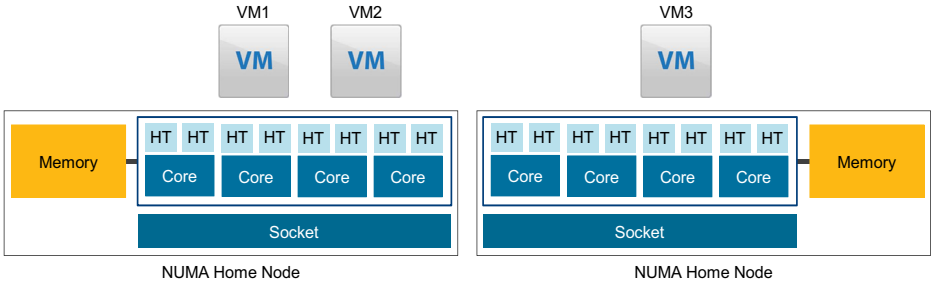
https://t.me/learningnets

5-21 How the ESXi NUMA Scheduler Works

When a VM is powered on, the ESXi NUMA scheduler assigns the VM to a home node.

When memory is allocated to a VM, the ESXi host preferentially allocates it from the home node. The vCPUs of the VM are constrained to run on the home node to maximize memory locality.

The ESXi NUMA scheduler can dynamically change a VM's home node to respond to changes in system load.



A home node is one of the system's NUMA nodes. In selecting a home node for a VM, the ESXi NUMA scheduler attempts to keep both the VM and its memory located on the same node, thereby maintaining good NUMA locality.

The ESXi NUMA scheduler might migrate a VM to a new home node to reduce processor load imbalance. Because this event might cause more of the VM's memory to be remote, the scheduler might migrate the VM's memory dynamically to its new home node to improve memory locality. The NUMA scheduler might also swap VMs between nodes when this improves overall memory locality.

https://t.me/learningnets

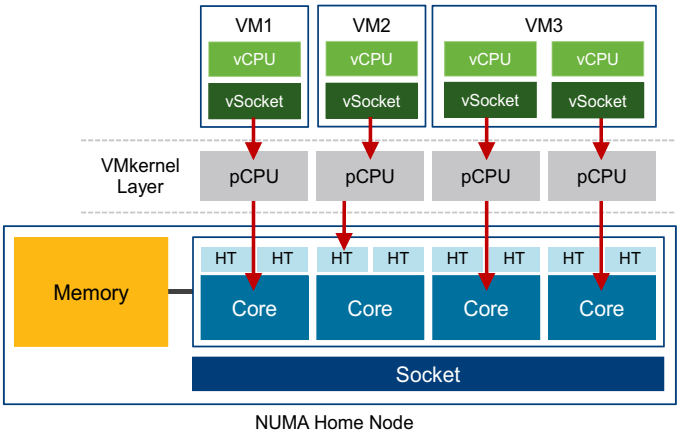
5-22 Terminology: vCPUs, pCPUs, Cores, and More

ESXi presents vCPUs and virtual sockets to each VM.

A virtual socket is mapped to one or more pCPUs, depending on the number of vCPUs.

A pCPU is a VMkernel construct.

Each pCPU can consume a full core or a hyperthreading thread (also known as a logical processor) in a NUMA node.



In order to understand how NUMA functions in vSphere, a review of CPU components at the virtual and physical layers is helpful.

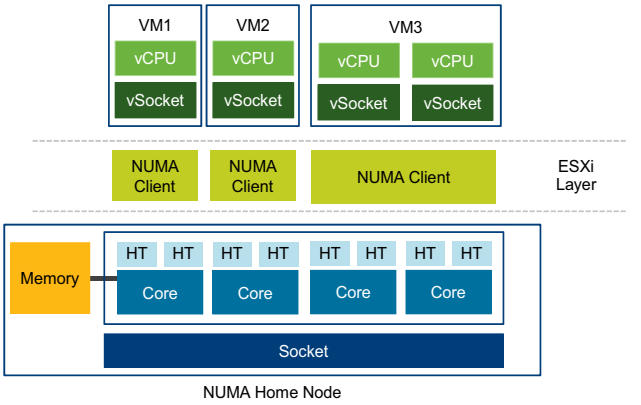
https://t.me/learningnets

5-23 About the NUMA Client

ESXi groups the vCPUs and memory of a VM into a single unit called a NUMA client.

A NUMA client fits inside a NUMA home node.

The ESXi NUMA scheduler uses the NUMA client to initially place a VM on a node and to perform ongoing load-balancing operations.



When a VM is powered on, the number of vCPUs in a VM are compared with the number of physical cores in a NUMA node. If enough physical cores exist in the NUMA node to satisfy the vCPU count, then a single NUMA client is created. A uniform memory address space is presented to the VM.

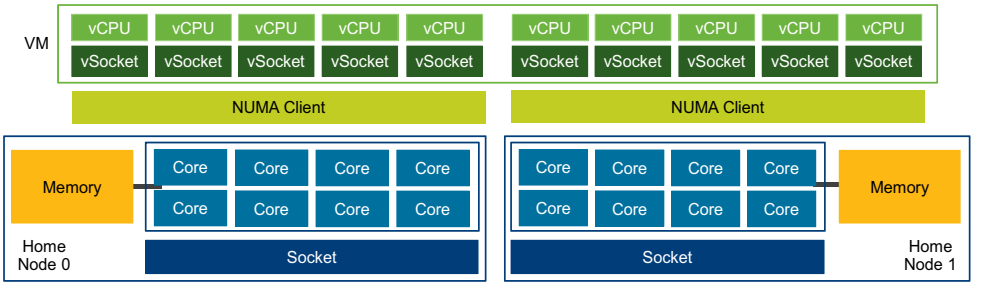
https://t.me/learningnets

5-24 Wide VM NUMA Support

If a VM has more vCPUs than the number of cores in a NUMA node, the VM is split into a number of NUMA clients and it is referred to as a wide VM.

For example, a 10-vCPU VM is considered a wide VM on a dual-socket, eight cores per socket system:

- Each NUMA client is assigned a home node.
- Only the cores count. Hyperthreading threads do not count.



Two NUMA clients are created for this VM and the vCPUs are equally distributed across the two NUMA clients.

Wide VMs are assigned two or more NUMA nodes and are preferentially allocated memory local to those NUMA nodes.

Because vCPUs in these wide VMs might sometimes need to access memory outside their own NUMA node, the VM might experience higher average memory access latencies than VMs that fit entirely within a NUMA node.

https://t.me/learningnets

5-25 Virtual NUMA Topology

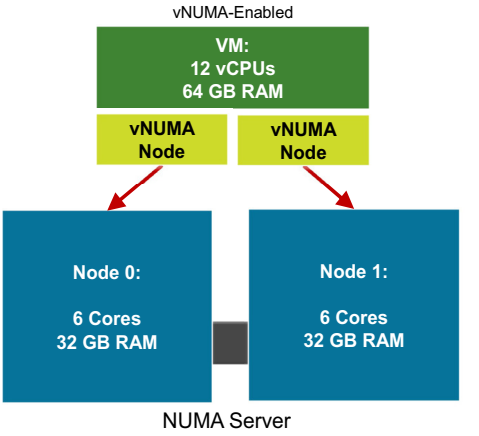
When a wide VM is created, a virtual NUMA (vNUMA) topology is created and presented to the guest operating system.

vNUMA enables NUMA-aware guest operating systems and applications to make efficient use of the underlying hardware's NUMA architecture.

The vNUMA topology presents a vNUMA node to the VM.

vNUMA topology is exposed to the VM if two conditions are met:

- The VM contains nine or more vCPUs.
- The vCPU count exceeds the core count of the physical NUMA node.



Memory latency can be mitigated by appropriately configuring virtual NUMA (vNUMA) on the VMs. vNUMA enables the guest operating system to assume part of the memory-locality management task.

vNUMA can provide significant performance benefits, although the benefits depend heavily on the level of NUMA optimization in the guest operating system and applications.

You can obtain the maximum performance benefits from vNUMA if your vSphere clusters are composed entirely of hosts with matching NUMA architecture. When a VM that is enabled for vNUMA is powered on, its vNUMA topology is set based in part on the NUMA topology of the underlying physical host. After a VM's vNUMA topology is initialized, it does not change unless the number of vCPUs in that VM is changed. Thus, if a vNUMA VM is moved to a host with a different NUMA topology, the VM's vNUMA topology might no longer be optimal. This move might result in reduced performance.

For more information on NUMA and vNUMA, see "NUMA Deep Dive Part 5: ESXi VMkernel NUMA Constructs" at http://frankdenneman.nl/2016/08/22/numa-deep-dive-part-5-esxi-vmkernel-numa-constructs.

https://t.me/learningnets

5-26 Optimizing the Number of Cores per Socket for VMs (1)

The number of vCPUs and cores per socket for a VM affect the ability of guest operating systems and applications to optimize their cache usage.

You can configure the number of vCPUs and the number of cores per socket for a VM by editing the VM's CPU properties.



To optimize guest operating system memory behavior, ensure that the number of virtual cores per socket does not exceed the number of physical cores per socket on your ESXi host.

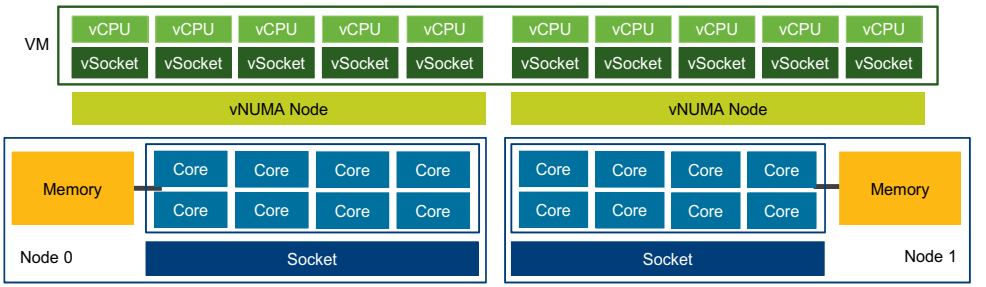
When you create a new VM, the number of vCPUs that you specify is divided by the cores per socket value to give you the number of sockets. The default cores per socket value is 1.

To ensure an optimal vNUMA topology and optimal performance, regardless of what vSphere version you are using, ensure that the VM vCPU count does not exceed the physical core count of a single physical NUMA node.

https://t.me/learningnets

5-27 Optimizing the Number of Cores per Socket for VMs (2)

For example, you have a VM that is currently configured with 10 vCPUs and 1 core per socket (10 sockets). This VM runs on a dual-socket, eight cores per socket system.



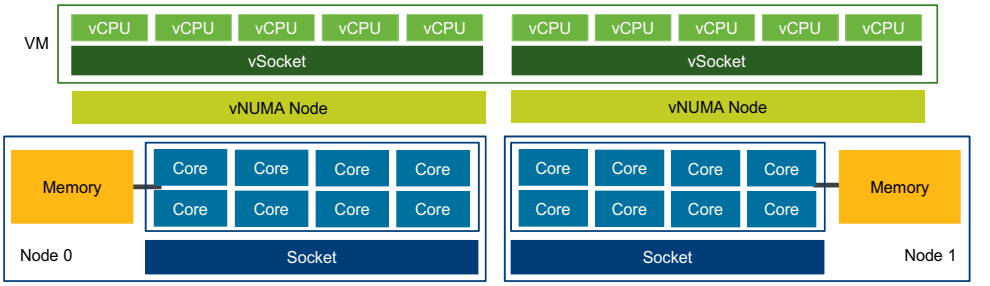
Dual-Socket, Eight Cores per Socket System

https://t.me/learningnets

5-28 Optimizing the Number of Cores per Socket for VMs (3)

To optimize guest operating system memory behavior for this VM, change the number of cores per socket to five, which results in two sockets.

With this new configuration, the vNUMA topology presented to the VM aligns with the NUMA system on which the VM resides.

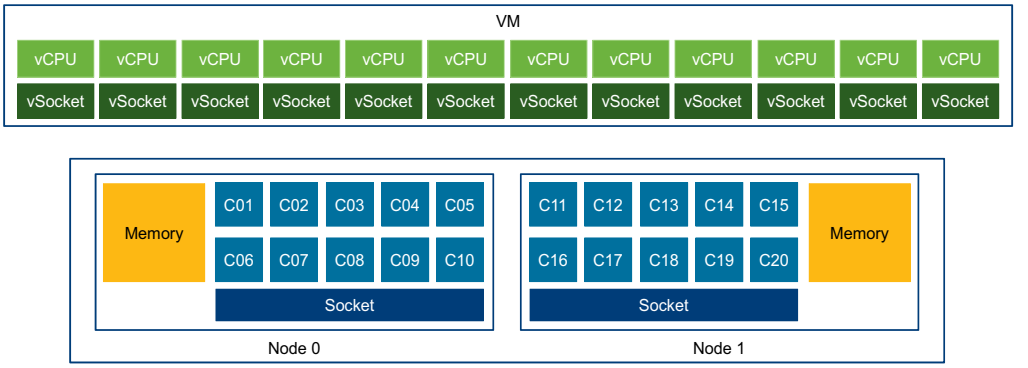


Dual-Socket, Eight Cores per Socket System

https://t.me/learningnets

5-29 Activity: VM Cores Per Socket (1)

For optimal guest operating system performance, what should the number of cores per socket be for a 12-vCPU VM running on a dual-socket, 10 cores per socket system?

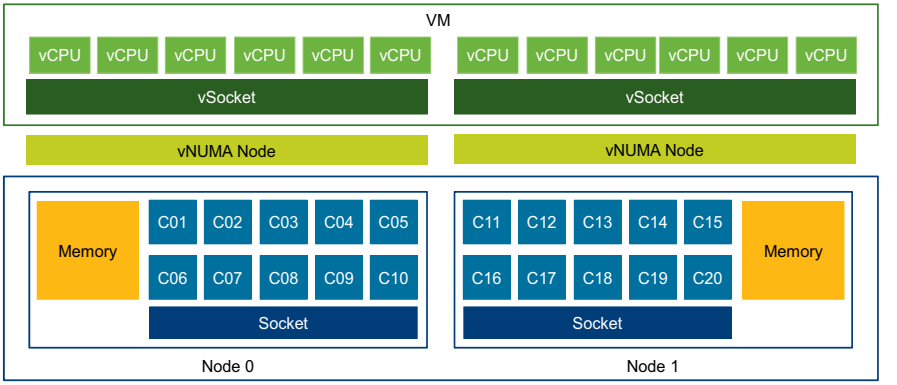


Dual-Socket, 10 Cores per Socket System

https://t.me/learningnets

5-30 Activity: VM Cores Per Socket (2)

The VM should be configured with 6 cores per socket (2 sockets). This configuration aligns with the physical characteristics of this NUMA system.



Dual-Socket, 10 Cores per Socket System

https://t.me/learningnets

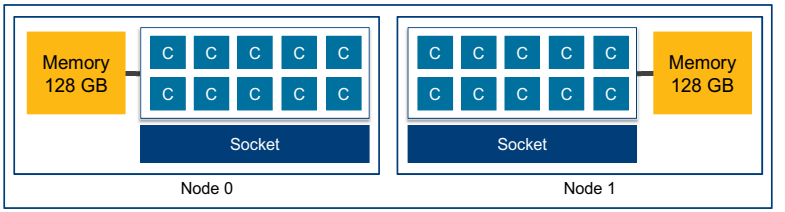
5-31 Memory Considerations with vNUMA (1)

When a vNUMA topology is calculated, only the compute dimensions are considered.

The calculation does not consider the amount of memory configured for the VM, or the amount of memory available within each physical NUMA node.

Therefore, you must manually account for memory.

For example, an ESXi host has two sockets, 10 cores per socket, and 128 GB of RAM per NUMA node for a total of 256 GB of RAM on the host.

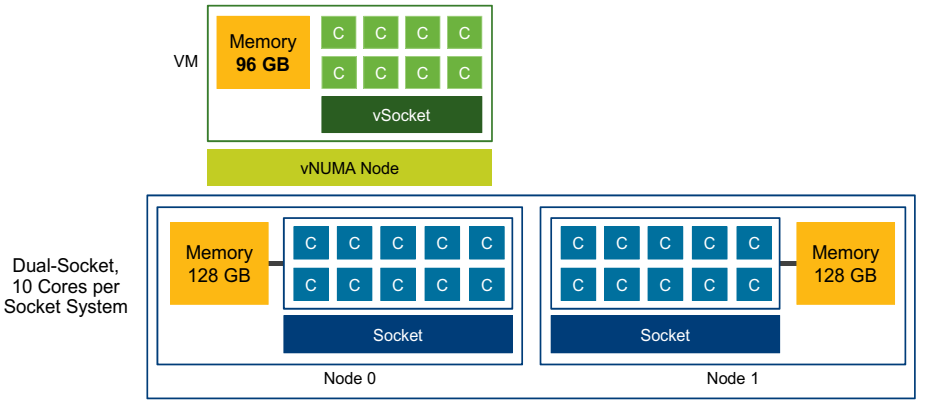


Dual-Socket, 10 Cores per Socket System

https://t.me/learningnets

5-32 Memory Considerations with vNUMA (2)

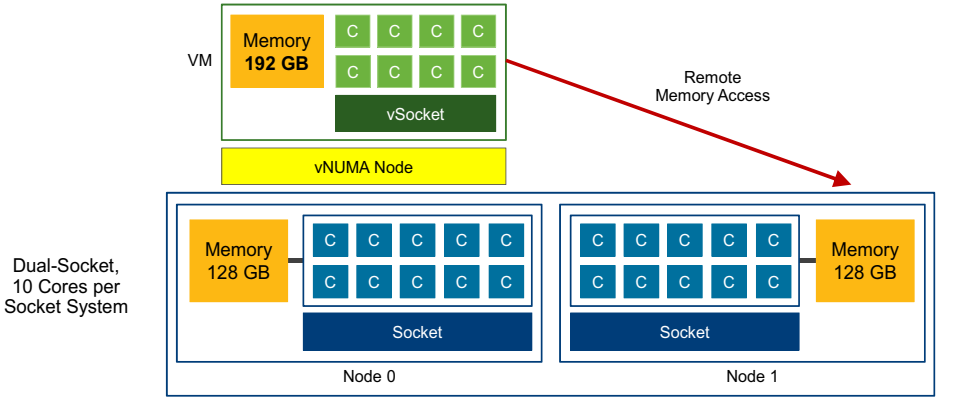
Continuing with the example, you create a VM with 96 GB of RAM, 1 socket, and 8 cores per socket. For this VM, ESXi creates a single vNUMA node. The VM fits into a single physical NUMA node.



https://t.me/learningnets

5-33 Memory Considerations with vNUMA (3)

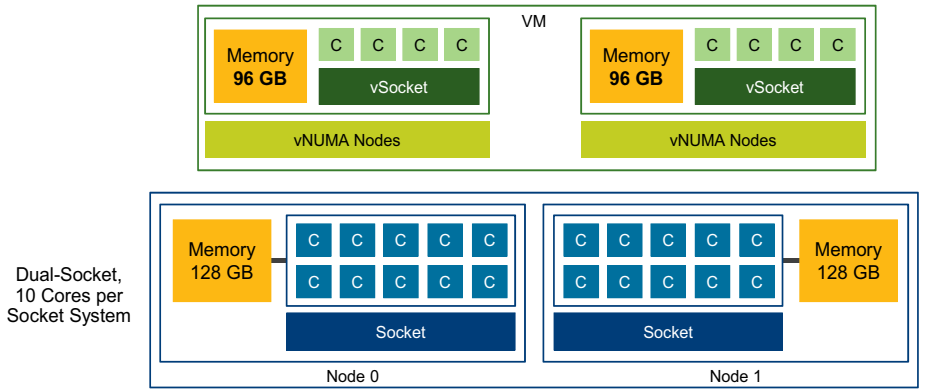
If you create a VM with 192 GB of RAM, 1 socket, and 8 cores per socket, ESXi still creates a single vNUMA node. However, the VM must span 2 physical NUMA nodes, resulting in remote memory access.



5-34 Memory Considerations with vNUMA (4)

If the VM is memory-intensive, the optimal configuration for this VM is 2 sockets and 4 cores per socket. You must also set the `numa.vcpu.maxPerMachineNode` VM to 4.

With this configuration, ESXi creates 2 vNUMA nodes and distributes 96 GB of RAM to each of the nodes.



https://t.me/learningnets

5-35 vNUMA Considerations

Consider these guidelines to ensure that the optimal vNUMA configuration is implemented:

- Always configure the VM vCPU count to be reflected as cores per socket, until you exceed either the physical core count or the total memory available on a single physical NUMA node.
- When you need to configure more vCPUs than there are physical cores in the NUMA node, or if you assign more memory than a NUMA node contains, evenly divide the vCPU count across the minimum number of NUMA nodes.
- If you enable the CPU Hot Add function in a VM, the vNUMA topology is disabled:
  - The VM is started without vNUMA and will instead use uniform memory access (UMA), also known as SMP.
- While several advanced vNUMA settings exist, do not use them unless instructed to do so by VMware Global Support Services.

For more information about vNUMA rightsizing and vNUMA topology, see “Virtual Machine vCPU and vNUMA Rightsizing - Rules of Thumb” at https://blogs.vmware.com/performance/2017/03/virtual-machine-vcpu-and-vnuma-rightsizing-rules-of-thumb.html as well as Decoupling of Cores per Socket from Virtual NUMA Topology in vSphere 6.5 at http://frankdenneman.nl/2016/12/12/decoupling-cores-per-socket-virtual-numa-topology-vsphere-6-5.

https://t.me/learningnets

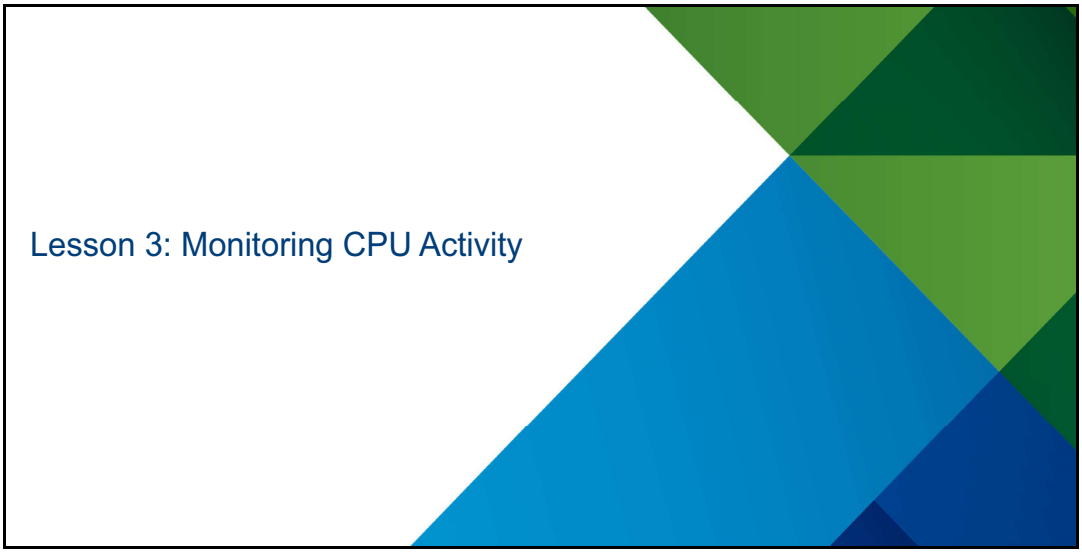
5-36 Review of Learner Objectives

You should be able to meet the following objectives:

- Explain how ESXi hosts are NUMA-aware
- Explain the virtual NUMA topology
- Describe how to optimize the number of cores per socket for a VM

https://t.me/learningnets

5-37 Lesson 3: Monitoring CPU Activity



https://t.me/learningnets

5-38 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Identify key CPU metrics to monitor
- Use metrics in `esxtop`
- Monitor CPU usage

https://t.me/learningnets

5-39 About esxtop (1)

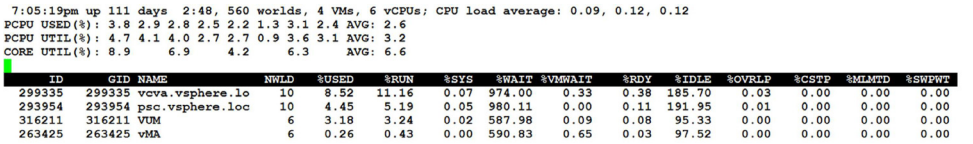
`esxtop` is the primary real-time performance monitoring tool for vSphere:

- This tool can be run from a host's local vSphere ESXi Shell as `esxtop`.
- This tool can also be run remotely from VMware vSphere® Command-Line Interface as `resxtop`.

`esxtop` provides much of the same information as vSphere Client but supplements the output with more detailed information.

`esxtop` works like the `top` performance utility found in Linux operating systems.

`esxtop` displays information in percentages of sample time (5 seconds by default).



You can run the `esxtop` utility using VMware vSphere® ESXi™ Shell to communicate with the management interface of the ESXi host. You must have root user privileges.

https://t.me/learningnets

5-40 About esxtop (2)

You view key performance indicators on individual resource screens by entering the appropriate keys. Commands are case-sensitive.

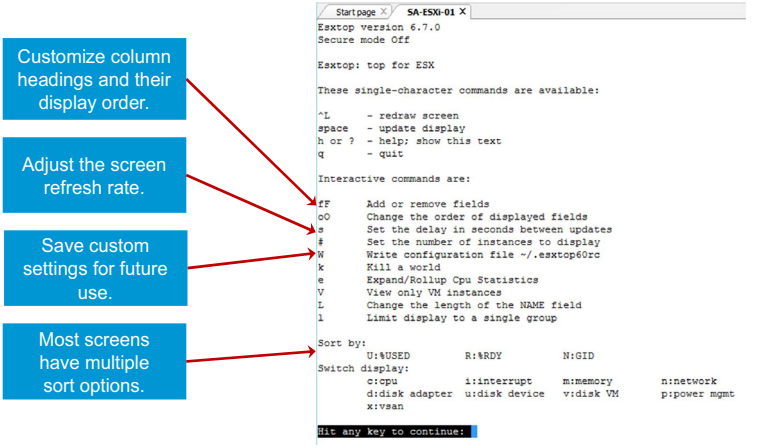
`c`	CPU screen (default)
`m`	Memory screen
`d`	Disk (adapter) screen
`u`	Disk (device) screen
`v`	Virtual disk view (lowercase v)
`n`	Network screen
`f/F`	Add or remove statistic columns
`V`	Virtual machine view (uppercase V)
`h`	Help
`q`	Quit

You use lowercase and uppercase letters to specify which fields appear in which order on the CPU, memory, storage adapter, storage device, virtual machine storage, and network panels. You can add or remove columns containing data in the respective `esxtop` panels.

https://t.me/learningnets

5-41 Getting Help in esxtop

The `esxtop` help utility lists the available commands. Enter **`h`** or **`?`** to display the help screen.



https://t.me/learningnets

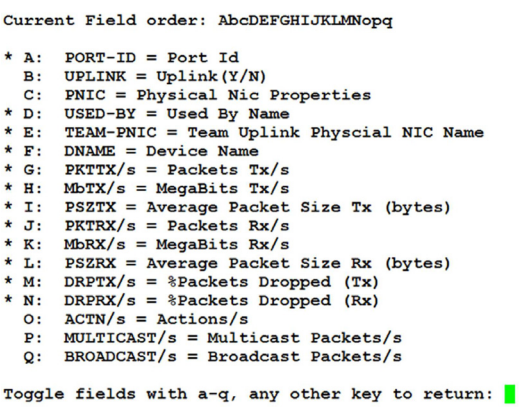
5-42 Customizing Fields in esxtop

You can customize any `esxtop` screen by entering **`f`** or **`F`**.

Each resource screen has multiple columns that you can display.

Press the corresponding key to toggle the display of a column.

An asterisk (*) indicates that the column will be displayed.



For more information about the `esxtop` utility, see *vSphere Monitoring and Performance* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-monitoring-performance-guide.pdf.

https://t.me/learningnets

5-43 CPU Key Performance Indicators for ESXi Hosts

To accurately determine how well the CPU is running, you review key performance indicators in `esxtop`.

Use the following CPU key performance indicators for ESXi hosts:

Key Performance Indicator	esxtop Counter
High PCPU utilization	PCPU UTIL (%)
Low PCPU utilization	PCPU UTIL (%)
High core utilization	CORE UTIL (%)
CPU load average	CPU load average

The values obtained from `esxtop` are gathered based on different sampling intervals. By default, `esxtop` uses a sampling interval of 5 seconds.

https://t.me/learningnets

5-44 CPU Key Performance Indicators for VMs

In `esxtop`, a group is a resource pool, a running VM, or a non-VM world. For worlds belonging to a VM, statistics for the running VM are shown and are identified by the VM's group ID (GID).

Use the following CPU key performance indicators for VMs:

Key Performance Indicator	esxtop Counter
High vCPU usage	%USED
Low vCPU usage	%USED
Ready time	%RDY
Co-stop time	%CSTP
Max limited time	%MLMTD
VM wait time	%VMWAIT
Swap wait time	%SWPWT
Number of members in a VM world	NWLD

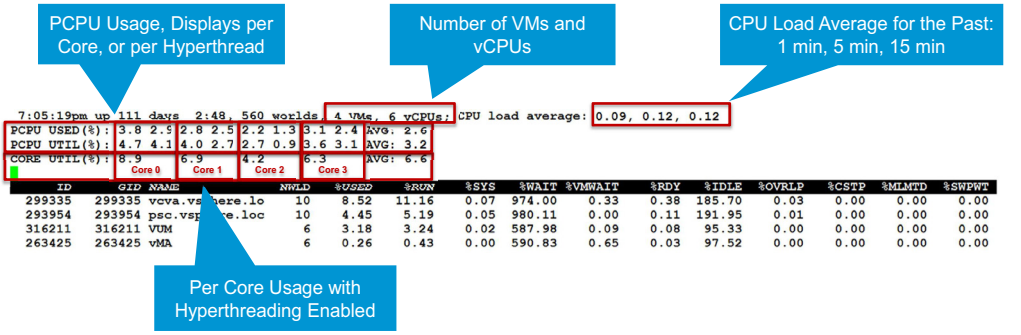
The following CPU statistics are shown for a virtual machine:

- %USED: CPU usage. The percentage of physical CPU core cycles used by a group of worlds running virtual machines, resource pools, or other worlds.
- %RDY: Percentage of time that the virtual machine was ready to run but was not provided CPU resources on which to execute.
- %CSTP: Percentage of time that the vCPUs of a virtual machine spent in the costop state, waiting to be co-started. This value gives an indication of the co-scheduling overhead incurred by the virtual machine. If this value is low, then any performance problems should be attributed to other issues and not to the co-scheduling of the virtual machine's vCPUs.
- %MLMTD: Percentage of time that the VMkernel did not run the virtual machine because that would violate the virtual machine's limit setting.
- %VMWAIT: Percentage of time that the virtual machine spent in a blocked state waiting for events.
- %SWPWT: Percentage of time that the virtual machine spends waiting for the VMkernel to swap memory.
- NWLD: Number of members in the virtual machine of the world that is running.

https://t.me/learningnets

5-45 Using esxtop to Monitor ESXi Host CPU

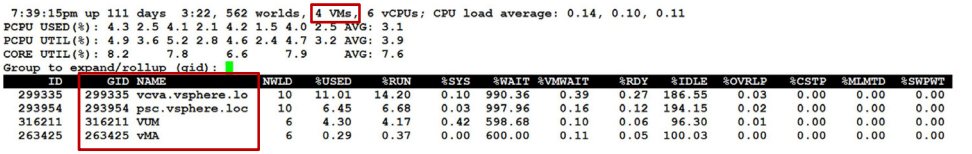
Enter **c** to display CPU information about the ESXi host.



https://t.me/learningnets

5-46 Using esxtop to View CPU Metrics per VM

Enter **v** to filter the output to show only virtual machines.



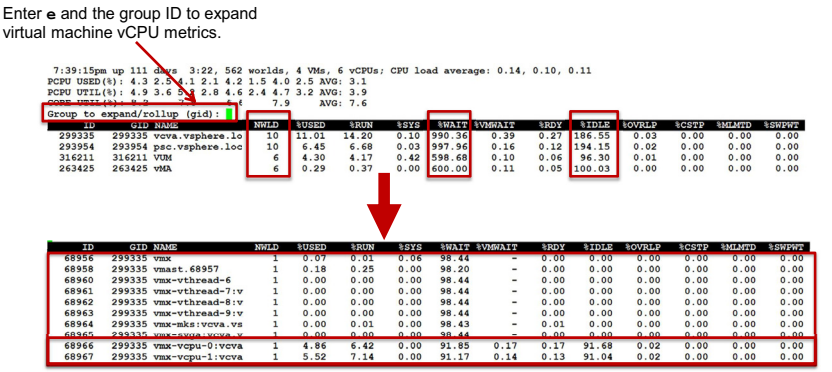
`esxtop` often shows more information than you need for the specific performance problem that you are troubleshooting. You can issue commands in the interactive mode to customize the display. The following commands are the most useful when monitoring CPU performance:

- Press the spacebar: Immediately updates the current screen.
- Enter **s**: Prompts you for the delay between updates, in seconds. The default value is 5 seconds. The minimum value is 2 seconds.
- Enter **v**: Displays virtual machine instances only. Do not confuse this command with lowercase v. Lowercase v shifts the view to the storage virtual machine resource usage screen. If you accidentally use lowercase, enter **c** to return to the CPU resource usage panel.
- Enter **e**: Toggles between CPU statistics being displayed expanded or unexpanded. The expanded display includes CPU resource usage statistics broken down by individual worlds belonging to a virtual machine. All percentages of the individual worlds are percentages of a single physical CPU.

https://t.me/learningnets

5-47 Using esxtop to View Single CPU Statistics

Enter **e** to show all the worlds associated with a single virtual machine.



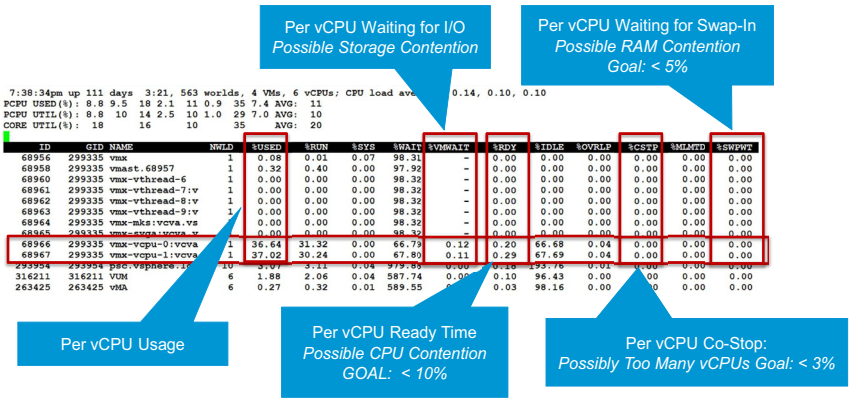
The default `esxtop` screen shows aggregate information for all vCPUs in a virtual machine. Enter **e** and enter the group ID to expand the group and display all the worlds associated with that group.

In the example, the virtual machine group with group ID 299335 is expanded to show all the worlds associated with that group.

https://t.me/learningnets

5-48 Using esxtop to Monitor Virtual Machine vCPU

Virtual machine vCPU information can be monitored with `esxtop`.



CPU usage, ready time, wait time, and co-stop information for each vCPU can be tracked with this `esxtop` panel.

https://t.me/learningnets

5-49 Important Metrics to Monitor

Monitoring both high-usage values and ready time is important to accurately depict CPU performance.

High-usage values alone do not always indicate poor performance:

- This value is often an indicator of high usage.
- This value is a goal of many organizations.

Ready time reflects the idea of queuing in the CPU scheduler:

- Ready time is the best indicator of possible CPU performance problems.
- Ready time occurs when more CPU resources are being requested by virtual machines than can be scheduled on the given physical CPUs.

High usage is not necessarily an indication of poor CPU performance, although it is often mistakenly understood as such. Administrators in the physical space view high usage as a warning sign, due to the limitations of physical architectures. However, one of the goals of virtualization is to efficiently use the available resources. Detecting high CPU usage, along with queuing, might indicate a performance problem.

As an analogy, consider a freeway. If every lane in the freeway is packed with cars but traffic is moving at a reasonable speed, you have high use but good performance. The freeway is doing what is was designed to do: moving a large number of vehicles at its maximum capacity.

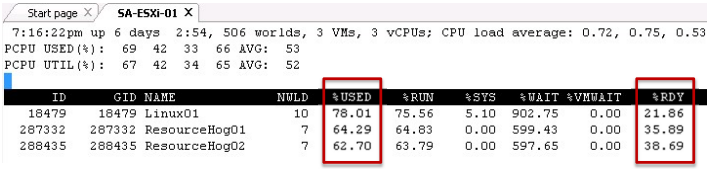
Now, imagine the same freeway at rush hour. More cars than the maximum capacity of the freeway are trying to share the limited space. Traffic speed reduces to the point where cars line up (queue) at the on-ramps while they wait for shared space to open up so that they can merge onto the freeway. This condition indicates high use with queuing and that a performance problem exists.

Ready time reflects the idea of queuing in the CPU scheduler. Ready time is the amount of time that a vCPU was ready to run but was asked to wait due to an overcommitted physical CPU.

https://t.me/learningnets

5-50 Example: Identifying CPU Overcommitment

The %USED and %RDY columns of the `esxtop` command output indicate CPU overcommitment.

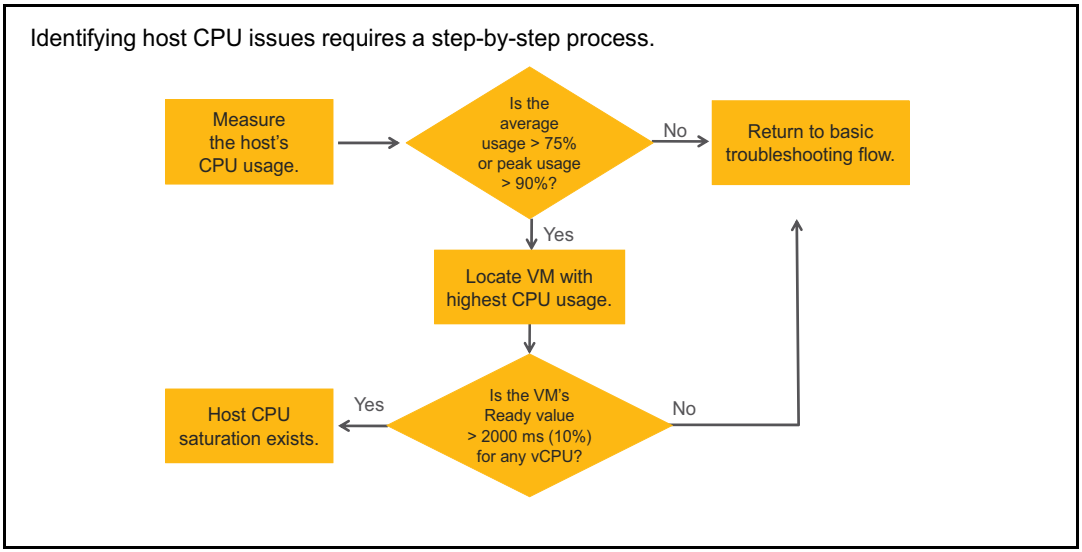


The example uses `esxtop` to detect CPU overcommitment. Looking at the PCPU line near the top of the screen, you can determine that the host has four CPUs, three of which are moderately used.

Three active virtual machines are shown: Linux01, ResourceHog01, and ResourceHog02. These virtual machines are active because they have relatively high values in the %USED column. The values in the %USED column alone do not necessarily indicate that the CPUs are overcommitted. In the %RDY column, you see that the three active virtual machines have high values. Values greater than 10 percent are considered high. High %RDY values, plus high %USED values, are a sure indicator that your CPU resources are overcommitted.

https://t.me/learningnets

5-51 Host CPU Saturation



In `esxtop`, ready time is reported in a percentage format. A figure of 5 percent means that the virtual machine spent 5 percent of its last sample period waiting for available CPU resources.

Using `esxtop` values for ready time, you can interpret the value as follows:

- A ready time that is less than or equal to 5 percent is normal. Very small single-digit numbers result in minimal effects on users.
- If ready time is between 5 and 10 percent, ready time merits attention.
- If ready time is greater than 10 percent, even though some systems continue to meet expectations, double-digit ready time percentages often mean that action is required to address performance problems.

https://t.me/learningnets

5-52 Resolving Host CPU Saturation

Several options are available to resolve host CPU saturation:

- Reduce the number of virtual machines running on the host.
- Increase the available CPU resources by adding the host to a vSphere DRS cluster.
- Increase the efficiency with which virtual machines use CPU resources.
- Use resource controls to direct available resources to critical virtual machines.

Populating an ESXi host with too many virtual machines running compute-intensive applications can make supplying sufficient resources to any individual virtual machine impossible.

When the host has a small number of virtual machines with high CPU demand, you might solve the problem by first increasing the efficiency of CPU usage in the virtual machines. Virtual machines with high CPU demands are those most likely to benefit from simple performance tuning. In addition, unless care is taken with virtual machine placement, moving virtual machines with high CPU demands to a new host might move the problem to that host.

When the host has many virtual machines with moderate CPU demands, reducing the number of virtual machines on the host by rebalancing the load is the simplest solution. Performance tuning on virtual machines with low CPU demands yields smaller benefits. It is also time-consuming if the virtual machines are running different applications.

When the host has a mix of virtual machines with high and low CPU demand, the appropriate solution depends on available resources and skill sets. If additional ESXi hosts are available, rebalancing the load might be the best approach. If additional hosts are not available or if expertise is available for tuning the high-demand virtual machines, then increasing efficiency might be the best approach.

https://t.me/learningnets

5-53 Activity: CPU Metrics (1)

Which of the following metrics in `esxtop` can give you an indication of memory contention?

- ❑ %RDY
- ❑ %CSTP
- ❑ %USED
- ❑ %SWPWT
- ❑ %VMWAIT

https://t.me/learningnets

5-54 Activity: CPU Metrics (2)

Which of the following metrics in `esxtop` can give you an indication of memory contention?

- ❑ %RDY
- ❑ %CSTP
- ❑ %USED
- ✓ %SWPWT
- ❑ %VMWAIT

%SWPWT is the percentage of time that a world spends waiting for the VMkernel to swap memory.

https://t.me/learningnets

5-55 Guest CPU Saturation

Monitor virtual machine CPU usage to identify guest CPU saturation.

If the CPU average usage is greater than 75 percent or if CPU usage peaks above 90 percent, then guest CPU saturation exists.

Cause	Solution
Guest operating system and applications use all of the CPU resources.	Increase the CPU resources provided to the application.
	Increase the efficiency with which the VM uses the CPU resources.

Guest CPU saturation occurs when the application and operating system running in a virtual machine use all of the CPU resources that the ESXi host is providing to that virtual machine. The occurrence of guest CPU saturation does not necessarily indicate that a performance problem exists. Compute-intensive applications commonly use all available CPU resources. Even less-intensive applications might experience periods of high CPU demand without experiencing performance problems. However, if a performance problem exists while guest CPU saturation occurs, steps should be taken to eliminate the condition.

Adding CPU resources is often the easiest choice, particularly in a virtualized environment. However, this approach misses inefficient behavior in the guest application and operating system that might be wasting CPU resources or, in the worst case scenario, that might be an indication of an error condition in the guest. If a virtual machine continues to experience CPU saturation even after adding CPU resources, the tuning and behavior of the application and operating system should be investigated.

https://t.me/learningnets

5-56 Using One vCPU in an SMP VM

Monitor virtual machine usage for all vCPU objects in an SMP virtual machine.

If usage for all vCPUs except one is close to zero, the SMP virtual machine is using only one vCPU.

Cause	Solution
Guest operating system is configured with a uniprocessor kernel.	Upgrade to an SMP kernel or a hardware abstraction layer.
Application is pinned to a single core in the guest operating system.	Remove OS-level controls or reduce the number of vCPUs.
Application is single-threaded.	Reduce the number of vCPUs to one.

When a virtual machine that is configured with more than one vCPU actively uses only one of those vCPUs, resources that could be used to perform useful work are being wasted.

The guest operating system might be configured with a uniprocessor kernel (with Linux) or hardware abstraction layer (HAL) (with Windows). For a virtual machine to take advantage of multiple vCPUs, the guest operating system running in the virtual machine must be able to recognize and use multiple-processor cores. Follow the documentation provided by your operating system vendor to check for the type of kernel or HAL that is being used by the guest operating system.

The application might be pinned to a single core in the guest operating system. Many modern operating systems provide controls for restricting applications to run on only a subset of available processor cores. If these controls were applied in the guest operating system, the application might run on only vCPU0. To determine whether this is the case, inspect the commands used to start the application or other operating system-level tunings applied to the application. The inspection of these values must be accompanied by an understanding of the operating system vendor's documentation regarding restricting CPU resources.

The application might be single-threaded. Many applications, particularly older applications, are written with only a single thread of control. These applications cannot take advantage of more than one processor core. Running a single-threaded application on an SMP virtual machine might lead to low guest CPU usage, instead of running a single-threaded application on vCPU0 only. Without knowledge of the application design, determining whether an application is single-threaded might be difficult. Observing the behavior of the application in an SMP virtual machine might help in determining whether the application is single-threaded.

https://t.me/learningnets

5-57 Low Guest CPU Usage

Monitor a virtual machine to identify low CPU usage.

If average usage is less than 75 percent, guest CPU usage is low.

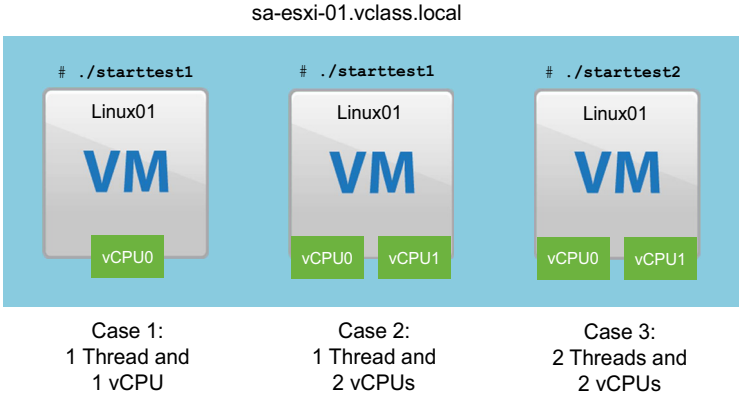
Cause	Solution
High storage response times.	Troubleshoot slow or overloaded storage.
High response times from external systems.	Solution is application- and environment-dependent.
Poor application or operating system tuning.	
Application pinned to cores in guest operating system.	Remove OS-level controls or reduce the number of vCPUs.
Too many configured vCPUs.	Reduce the number of vCPUs.
Restrictive resource allocations.	Modify VM resource settings.

To determine whether the problem is low guest CPU usage, use vSphere Web Client to check the CPU usage of the virtual machines on the host.

https://t.me/learningnets

5-58 Introduction to Lab 8: Monitoring CPU Performance

The diagram shows the tests that you run.



In this lab, you compare the performance of a MySQL database running on a single-vCPU virtual machine with a dual-vCPU virtual machine. You run three tests. Each test measures CPU usage, ready time, idle time, and operations per minute. You record the data after each test:

- Case 1: On a single-vCPU virtual machine, you run the `starttest1` program. `starttest1` simulates a single-threaded application accessing the database.
- Case 2: On a dual-vCPU virtual machine, you run the `starttest1` program.
- Case 3: On a dual-vCPU virtual machine, you run the `starttest2` program. `starttest2` simulates a dual-threaded application accessing the database.

https://t.me/learningnets

5-59 Lab 8: Monitoring CPU Performance

Use the `esxtop` command to monitor CPU performance

1. Run a Single-Threaded Program in a Single-vCPU Virtual Machine
2. Start `esxtop` and View Statistics
3. Record Statistics for Case 1: Single Thread and Single vCPU
4. Run a Single-Threaded Program in a Dual-vCPU Virtual Machine
5. Record Statistics for Case 2: One Thread and Two vCPUs
6. Run a Dual-Threaded Program in a Dual-vCPU Virtual Machine
7. Record Statistics for Case 3: Two Threads and Two vCPUs

https://t.me/learningnets

5-60 Review of Lab 8: Monitoring CPU Performance

Organize your results for lab 8.

CPU Counters	Case 1 (Task 4)	Case 2 (Task 6)		Case 3 (Task 8)	
	Linux01	vCPU0	vCPU1	vCPU0	vCPU1
%USED					
%RDY					
%IDLE					
OPM					

https://t.me/learningnets

5-61 Review of Learner Objectives

You should be able to meet the following objectives:

- Identify key CPU metrics to monitor
- Use metrics in `esxtop`
- Monitor CPU usage

https://t.me/learningnets

5-62 Key Points

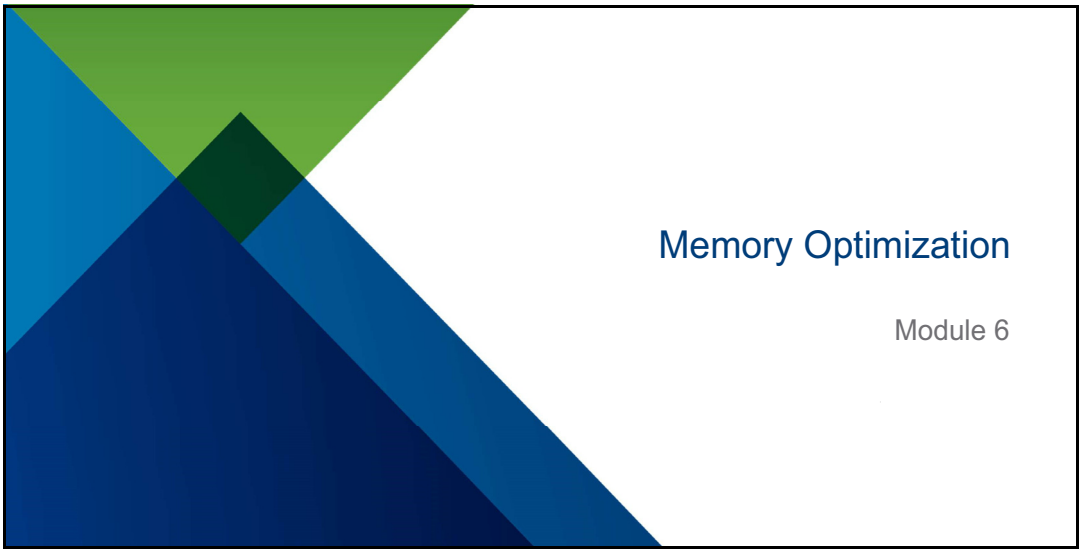
- CPU problems are usually caused by host CPU resources that are insufficient to satisfy vCPU demand.
- High CPU usage values with high ready time often indicate CPU performance problems.
- For CPU-related problems, you should check host CPU saturation and guest CPU saturation.

Questions?

https://t.me/learningnets

MODULE 6

Memory Optimization



https://t.me/learningnets

6-2 You Are Here

1. Course Introduction
2. Network Scalability
3. Storage Scalability
4. Host and Management Scalability
5. CPU Optimization
6. **Memory Optimization**
7. Storage Optimization
8. Network Optimization
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

6-3 Importance

Although vSphere employs various mechanisms to efficiently allocate memory, you might still encounter a situation in which virtual machines are allocated insufficient physical memory.

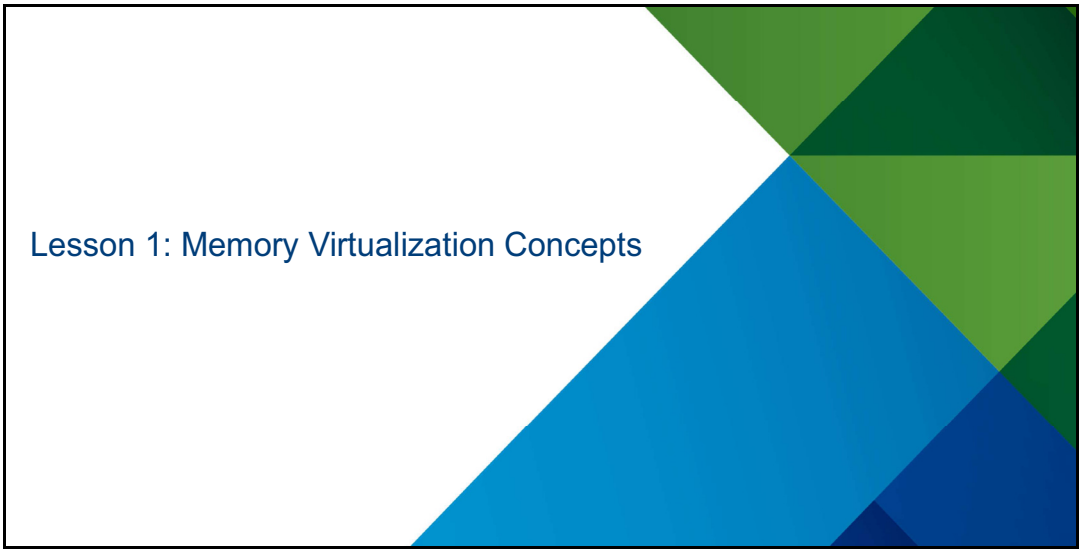
You should know how to monitor memory usage of the host and the virtual machines.

https://t.me/learningnets

6-4 Module Lessons

Lesson 1:	Memory Virtualization Concepts
Lesson 2:	Monitoring Memory Activity

https://t.me/learningnets



https://t.me/learningnets

6-6 Learner Objectives

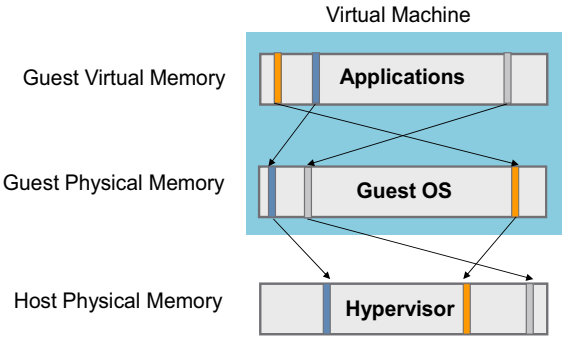
By the end of this lesson, you should be able to meet the following objectives:

- Describe how memory is used in a virtualized environment
- Explain each memory reclamation technique
- Explain how memory overcommitment affects performance

https://t.me/learningnets

6-7 Virtual Memory Overview

A running virtual machine has several levels of memory.



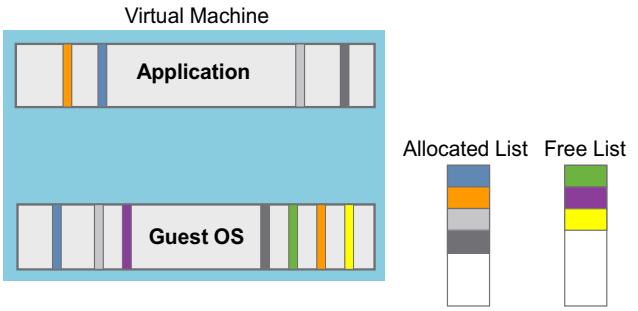
The following levels of memory exist when running a virtual machine:

- Guest virtual memory: Memory visible to the applications running in the virtual machine. This memory is a continuous virtual address space presented by the guest operating system to applications.
- Guest physical memory: Memory visible to the guest operating system running in the virtual machine. The guest operating system sees this memory as physical memory.
- Host physical memory: Also called machine memory, it is the physical memory that is available to the hypervisor.

https://t.me/learningnets

6-8 Application and Guest Operating System Memory Management

Applications start with no memory. During execution, they use the interfaces provided by the operating system to explicitly allocate or deallocate virtual memory.



In a nonvirtual environment, the operating system owns all physical memory in the system. The hardware does not provide interfaces for the operating system to explicitly allocate or free physical memory. The operating system establishes the definitions of allocated or free physical memory. Operating systems have different implementations to realize this abstraction. So whether a physical page is free depends on which list the page resides in.

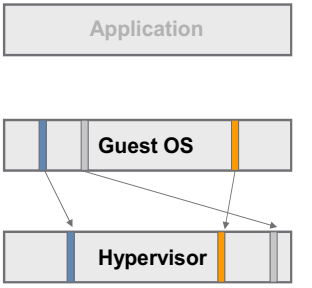
https://t.me/learningnets

6-9 Virtual Machine Memory Management

Virtual machines start with no host physical memory allocated to them.

Host physical memory is allocated on demand:

- The guest operating system does not explicitly allocate memory.
- Memory is allocated on the virtual machine's first access to memory (read or write).



Because a virtual machine runs a guest operating system and one or more applications, the virtual machine memory management properties combine both application and operating system memory management properties. Like an application, when a virtual machine first starts, it has no preallocated host physical memory. Like an operating system, the virtual machine cannot explicitly allocate host physical memory through standard interfaces.

The hypervisor intercepts the virtual machine's memory accesses and allocates host physical memory for the virtual machine on its first access to the memory. The hypervisor always writes zeros to the host physical memory before assigning it to a virtual machine.

https://t.me/learningnets

6-10 Memory Reclamation

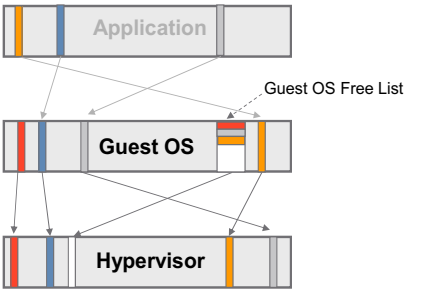
Guest physical memory is not freed in the typical sense:

- Changed guest physical memory is moved to the free list.
- Mapped host physical memory does not change.

The hypervisor is not aware of when the guest frees the memory:

- The hypervisor has no access to the guest's free list.
- The virtual machine can accrue lots of host physical memory.

The hypervisor cannot directly reclaim host physical memory that is in the guest operating system free list.



Virtual machine memory deallocation acts like an operating system, such that the guest operating system frees guest physical memory by adding memory page numbers to the guest free list. However, the data of the freed memory might not be modified at all. As a result, when a portion of guest physical memory is freed, the mapped host physical memory does not usually change its state. Only the guest free page list is changed.

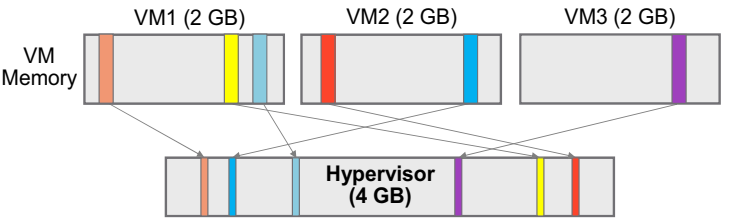
It is difficult for the hypervisor to know when to free host physical memory when guest physical memory is deallocated, or freed, because the guest operating system free list is not accessible to the hypervisor. The hypervisor is completely unaware of which pages are free or allocated in the guest operating system. As a result, the hypervisor cannot reclaim host physical memory when the guest operating system frees guest physical memory.

https://t.me/learningnets

6-11 Reason for Reclaiming Memory

The hypervisor reclaims memory to support ESXi memory overcommitment.

A host's memory is overcommitted when the total amount of guest physical memory is greater than the amount of host physical memory.



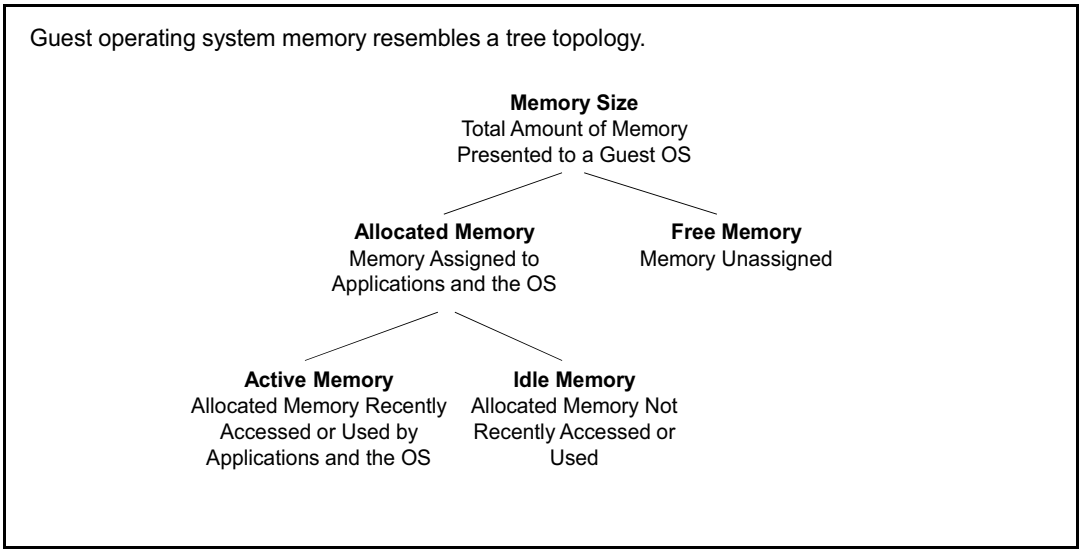
With memory overcommitment, ESXi ensures that the host physical memory is consumed by active guest memory as much as possible. Typically, some virtual machines might be lightly loaded compared to others, and so, for much of the time, their memory sits idle. Memory overcommitment allows the hypervisor to use memory-reclamation techniques to take the inactive or unused host physical memory away from the idle virtual machines and give it to other virtual machines that will actively use it.

With memory overcommitment, each virtual machine has a smaller footprint in host physical memory, making it possible to fit more virtual machines on the host while still achieving good performance for all virtual machines. In the example, you can enable a host with 4 GB of physical memory to run three virtual machines with 2 GB of VM memory each.

This action assumes that all virtual machines are using the default setting for memory reservation, which is 0. In such a case, all the virtual machines would power on. If all the virtual machines had a memory reservation of 2 GB each, without memory overcommitment, then only one virtual machine can be run. The hypervisor cannot reserve host physical memory for more than one virtual machine, because each virtual machine has overhead memory as well.

https://t.me/learningnets

6-12 Guest Operating System Memory Terminology



The memory used by the guest operating system in the virtual machine can be described as follows:

- Memory size is the total amount of memory that is presented to the guest. By default, the memory size corresponds to the memory configured when the virtual machine is created.
- The total amount of memory (memory size) can be divided into two parts:
  - Free memory: Memory that is not assigned to the guest operating system or to applications.
  - Allocated memory: Memory that is assigned to the guest operating system or to applications.
- Allocated memory can be further subdivided into two types:
  - Active memory: Memory that was recently used by applications.
  - Idle memory: Memory that was not recently used by applications.

  To maximize virtual machine performance, keeping a virtual machine's active memory in host physical memory is vital.

These terms are used to describe the guest operating system's memory. The ESXi host does not know about a guest's free, active, or idle memory.

https://t.me/learningnets

6-13 Virtual Machine Memory Reclamation Techniques

The hypervisor relies on the following memory reclamation techniques to free host physical memory:

- Transparent page sharing
- Ballooning
- Memory compression
- Host swap cache
- Host-level (hypervisor) swapping

When multiple virtual machines are running, some virtual machines might have identical sets of memory content. Opportunities exist for sharing memory across virtual machines and in a single virtual machine. With transparent page sharing (TPS), the hypervisor can reclaim redundant memory page copies. Only one read-only copy is kept in the host physical memory. The page is shared by multiple virtual machines. When a virtual machine writes to the page, a copy of the page is created for the virtual machine before the write operation is complete.

Ballooning, memory compression, host-level swapping, and host swap cache are the other memory-reclamation techniques.

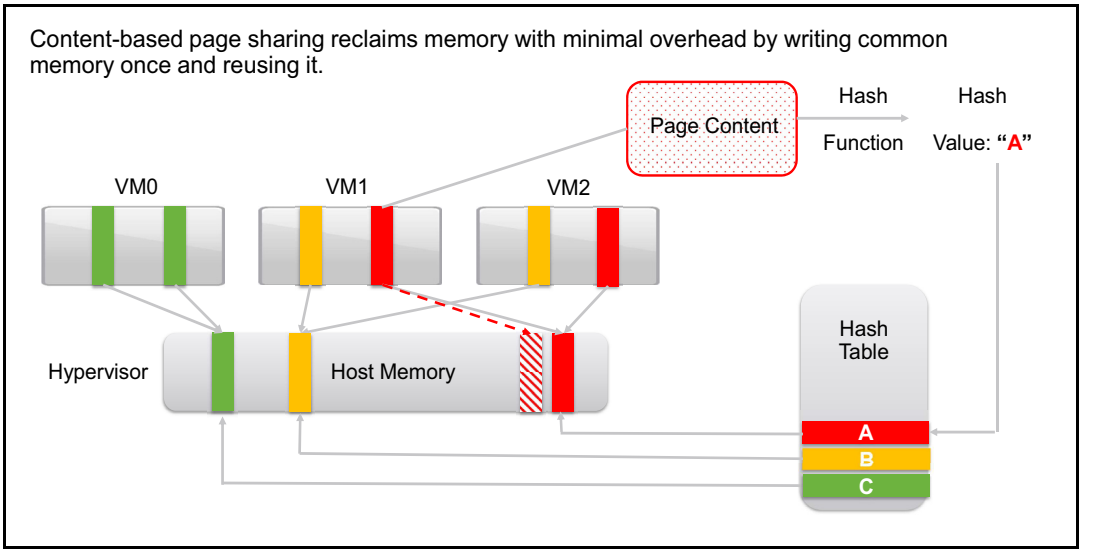
Hypervisor swapping is used as a last resort when ballooning and TPS are not sufficient to reclaim memory. If needed, the hypervisor will swap out guest physical memory to a virtual swap (vswp) file.

Host swap cache is an optional memory reclamation technique that uses local flash storage to cache a virtual machine's memory pages. By using local flash storage, the virtual machine avoids the latency associated with a storage network that would be used if it swapped memory pages to vswp files.

When there is severe memory pressure and the hypervisor needs to swap memory pages to disk, the hypervisor will swap to a host swap cache rather than to a vswp file. When a host runs out of space on the host cache, a virtual machine's cached memory will be migrated to a virtual machine's regular vswp file. Each host will need to have its own host swap cache configured.

https://t.me/learningnets

6-14 Transparent Page Sharing (1)



Many physical memory pages on a host often store identical contents. For example, if many virtual machines are running Windows 2012R2, each virtual machine has the same executable files in its memory. TPS is a background process that scans and hashes the contents of guest physical memory pages. Pages that generate the same hash value are scanned, byte by byte. If the pages are truly identical, they are single instanced and shared between the relevant virtual machines.

A hash value is generated based on the candidate guest physical page content. The hash value is used as a key to find a global hash table. In this table, each entry records a hash value and the physical page number of a shared page. If the hash value of the candidate guest physical page matches an existing entry, a full comparison of the page contents is performed to exclude a false match. If the candidate guest physical page's content is confirmed to match the content of an existing shared host physical page, then the guest physical-to-host physical mapping of the candidate guest physical page is changed to the shared host physical page. The redundant host memory copy (the page pointed to by the dashed arrow) is reclaimed. This remapping is invisible to the virtual machine and inaccessible to the guest operating system. Because of this invisibility, sensitive information cannot be leaked from one virtual machine to another.

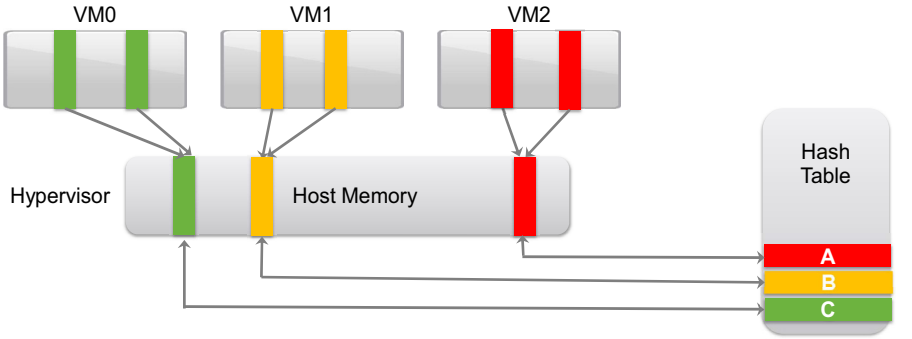
A standard copy-on-write technique is used to handle writes to the shared host physical pages. An attempt to write to the shared pages generates a minor page fault. In the page fault handler, the hypervisor transparently creates a private copy of the page for the virtual machine and remaps the affected guest physical page to this private copy. In this way, virtual machines can safely modify the shared pages without disrupting other virtual machines sharing that memory. Writing to a shared page incurs overhead compared to writing to nonshared pages due to the extra work performed in the page fault handler.

https://t.me/learningnets

6-15 Transparent Page Sharing (2)

Inter-VM transparent page sharing (TPS) is disabled by default and page sharing is restricted to intra-VM memory sharing.

For example, page sharing across VM0, VM1, and VM2 (inter-VM) is disabled by default. Page sharing within VM0 (intra-VM) is enabled by default.



For information about the security risks associated with inter-VM TPS, see VMware knowledge base article 2080735 at http://kb.vmware.com/kb/2080735.

https://t.me/learningnets

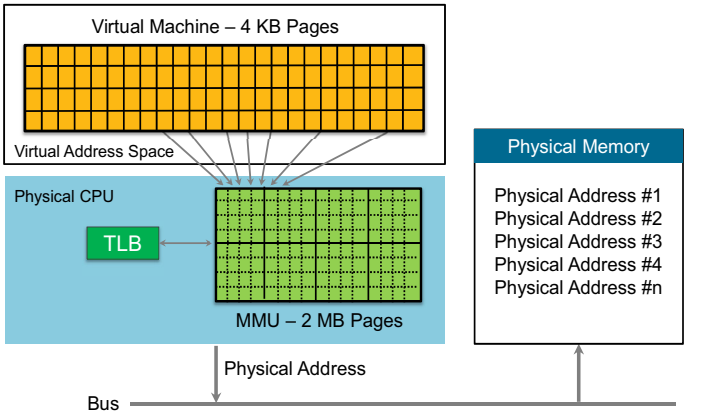
6-16 Impact of Hardware MMU on Transparent Page Sharing

All modern x86-64 systems use a memory management unit (MMU). The MMU performs virtual memory management, memory protection, cache control, and bus arbitration.

TPS is designed to scan and share small pages (4 KB in size).

The MMU uses large pages (2 MB in size).

The VMkernel does not attempt to share large pages.



In modern x86-64 systems, the memory management unit (MMU) is integrated into the central processing unit (CPU). MMU is a computer hardware unit having all memory references passed through itself, primarily performing the translation of virtual memory addresses to physical addresses. The translation lookaside buffer (TLB) caches the most recent translations of virtual-to-physical memory addresses.

Intel EPT Hardware Assist and AMD RVI Hardware Assist are the second generation of MMU. When using EPT or RVI, ESXi backs guest physical pages with large host physical pages. These physical pages have a 2 MB contiguous memory region instead of 4 KB, for regular pages for better performance due to fewer translation look-aside misses.

In such systems, ESXi does not share those large pages for the following reasons:

- The probability of finding two large pages with identical contents is low.
- The overhead of doing a bit-by-bit comparison for a 2 MB page is much larger than that for a 4 KB page.

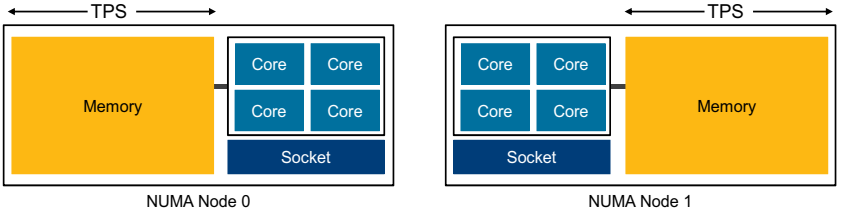
When using EPT or RVI, `esxtop` might show zero or few shared pages, as TPS uses small pages (4 KB), and EPT and RVI use large pages (2 MB). An ESXi host tracks what pages can be shared. If memory resources become overcommitted, ESXi breaks the large pages into small pages and begins sharing memory between VMs.

https://t.me/learningnets

6-17 Impact of NUMA Architecture on Transparent Page Sharing

TPS for ESXi systems are optimized for use on non-uniform memory access (NUMA) systems:

- On NUMA systems, pages are shared within the NUMA node, resulting in each NUMA node maintaining its own local copy of shared pages.
- When virtual machines use shared pages, they do not access memory remotely and thereby do not incur the latency penalty associated with remote memory access.

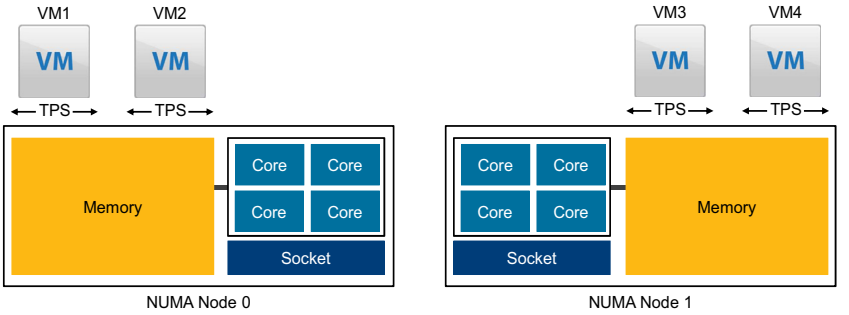


https://t.me/learningnets

6-18 Transparent Page Sharing and Intra-VM Memory Sharing

The default behavior of TPS focuses on intra-VM memory sharing:

- TPS only scans and collapses small pages that contain identical content within a VM.
- Zero-filled pages are the primary focus of intra-VM memory sharing.



https://t.me/learningnets

6-19 Using Salting with Transparent Page Sharing (1)

Salting enables the management of groups of VMs participating in inter-VM TPS, depending on the `Mem.ShareForceSalting` host setting and the `sched.mem.pshare.salt` VM setting.

Host Setting	VM Setting	vc.uuid	Salt Value	Inter-VM TPS	Intra-VM TPS
0	Ignored	Ignored	0	Yes, among all VMs on host.	Yes
1	Present	Ignored	sched.mem.pshare.salt	Only among VMs with same salt.	Yes
1	Not Present	Ignored	0	Yes, among all VMs.	Yes
2	Present	Ignored	sched.mem.pshare.salt	Only among VMs with same salt.	Yes
2	Not Present (default)	Present (default)	vc.uuid	No Inter-VM TPS.	Yes
2	Not Present	Not Present	random number	No Inter-VM TPS.	Yes

The concept of salting was introduced to address the security concerns around using TPS. With salting, VMs can share pages only when the pages are identical and the salt values for the VMs are the same.

Intra-VM TPS is always enabled. The host configuration option `Mem.ShareForceSalting` is used to force the concept of salting. Each VM can have a unique salt value set in the virtual machine's `.vmx` file called `sched.mem.pshare.salt`. When salting is enforced, VMs with unique salt values have inter-VM TPS disabled. However, any VMs with the same salt value configured have inter-VM TPS enabled as a group.

The available settings for `Mem.ShareForceSalting` are 0, 1, and 2:

- When set to 0: Salting is not enforced and all VMs have inter-VM TPS enabled in the host.
- When set to 1: Salting is enforced for only the VMs that have their salt parameter configured. Because VMs do not have a salt parameter by default, most VMs have inter-VM TPS enabled. Only the VMs with a unique salt value set have inter-VM TPS disabled.
- When set to 2: This is the default setting. Salting is enforced and uses a different VM parameter to determine if inter-VM TPS will be enabled. The `vc.uuid` parameter is a unique value set by vCenter Server when the VM is created. Because `vc.uuid` is unique for every VM, inter-VM TPS is disabled for all VMs by default. Only VMs that have the `sched.mem.pshare.salt` set to the same value will have inter-VM TPS enabled. In this case, the `sched.mem.pshare.salt` value overrides the `vc.uuid` setting and permits different groups of VMs with inter-VM TPS enabled for only their group. If a group of VMs can be trusted to share pages, they can be assigned a common salt value.

https://t.me/learningnets

6-20 Using Salting with Transparent Page Sharing (2)

A host setting is used to achieve a particular goal.

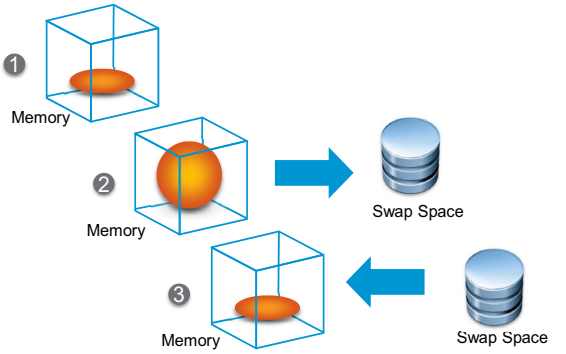
Host Setting	VM Setting	vc.uuid	Salt Value	Inter-VM TPS	Intra-VM TPS
0	**Goal: Inter-VM TPS for All VMs**				
1	**Goal: Inter-VM TPS for All VMs as Default**				
1	Permits Some VMs with No Inter-VM TPS				
2	**Goal: No Inter-VM TPS for All VMs as Default**				
2	Permits Some VMs with Inter-VM TPS				
2					

For more information about TPS capabilities, see VMware knowledge base article 2097593 at http://kb.vmware.com/kb/2097593.

https://t.me/learningnets

6-21 Memory Ballooning in the Guest Operating System

The memory balloon driver (`vmmemctl`) collaborates with the virtual machine to reclaim pages that are considered least valuable by the guest operating system. `vmmemctl` is the only way to reclaim unused memory from a guest operating system.



The `vmmemctl` driver is loaded with VMware Tools™. The driver uses a proprietary ballooning technique that provides predictable performance that closely matches the behavior of a native system under similar memory constraints. This technique increases or decreases memory pressure on the guest operating system, causing the guest to use its own native memory management algorithms. When memory is tight, the guest operating system determines which pages to reclaim and, if necessary, swaps them to its own swap file, for example, to `pagefile.sys` in Windows.

The diagram shows the following stages of operation:

1. Under normal operation, the VMkernel does not need to reclaim memory from the virtual machine. Memory pressure does not exist and ballooning is not active.
2. The balloon driver artificially creates memory pressure on the virtual machine. The guest operating system responds to the memory pressure by swapping out the optimal pages according to its analysis.
3. Memory pressure is relieved, ballooning is inactive, and the guest operating system swaps pages into memory as needed.

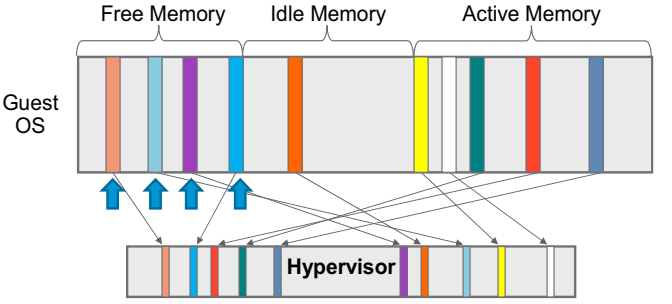
If necessary, you can limit the amount of memory that `vmmemctl` reclaims by setting the `sched.mem.maxmemctl` parameter for a specific virtual machine. This option specifies the maximum amount of memory that can be reclaimed from a virtual machine in megabytes.

https://t.me/learningnets

6-22 Reclaiming Memory with Ballooning

The goal of ballooning is to make the guest operating system aware of the low memory status of the host so that the guest operating system can free some of its memory.

Ballooning preferentially selects free or idle virtual machine memory. But if asked to reclaim too much memory, ballooning eventually starts reclaiming active memory.



Due to the virtual machine's isolation, the guest operating system is not aware that it is running in a virtual machine and is not aware of the states of other virtual machines on the same host. When the hypervisor runs multiple virtual machines, the total amount of free host physical memory might become low. A guest operating system cannot detect the host physical memory shortage, so none of the virtual machines free guest physical memory.

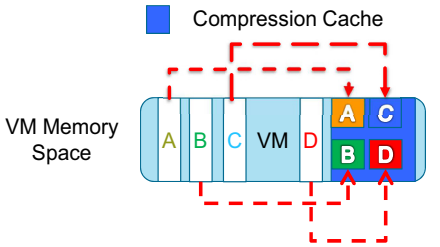
The guest operating system determines whether it needs to page out guest physical memory to satisfy the balloon driver's allocation requests. If the virtual machine has plenty of free or idle guest physical memory, inflating the balloon does not induce guest-level paging and does not affect guest performance. However, if the guest is already under memory pressure, the guest operating system decides which guest physical pages are to be paged to satisfy the balloon driver's allocation requests.

https://t.me/learningnets

6-23 Memory Compression

With memory compression, the ESXi host stores pages in a compression cache in the host's physical memory.

If a 4 KB memory page does not compress to less than 2 KB, the page remains uncompressed and the page is swapped out instead.



The compression cache is located in the VM's memory space. Memory compression outperforms host swapping because the next access to the compressed page causes only a page decompression, which can be an order of magnitude faster than the disk access. ESXi determines whether a page can be compressed by checking the compression ratio for the page. Memory compression occurs when the page's compression ratio is greater than 50 percent. Otherwise, the page is swapped out. Only pages that are to be swapped out are chosen as candidates for memory compression. This specification means that ESXi does not compress guest pages when host swapping is unnecessary. Memory compression does not affect workload performance when host memory is not overcommitted.

On the slide, assume that ESXi needs to reclaim 8 KB physical memory (that is, two 4 KB pages) from a virtual machine. With memory compression, a swap candidate page is compressed and stored using 2 KB of space in a per-virtual machine compression cache. Each compressed page yields 2 KB memory space for ESXi to reclaim. In order to reclaim 8 KB of physical memory, four swap candidate pages must be compressed. The page compression is much faster than the normal page swap-out operation, which involves a disk I/O.

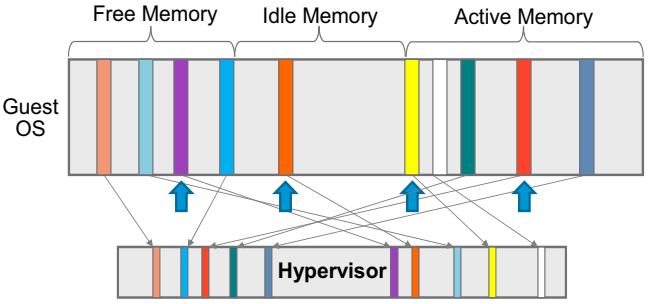
The compression cache size starts at 0 and by default, can grow to a maximum of 10 percent of the VM's memory size. You can control the cache size by using the `Mem.MemZipMaxPct` ESXi advanced system setting. Since this setting is applied at the host level, the setting impacts all VMs running on that host.

https://t.me/learningnets

6-24 Reclaiming Memory with Host Swapping

ESXi uses host-level swapping when TPS, ballooning, and memory compression are insufficient to reclaim memory.

Host-level swapping randomly selects guest physical memory to reclaim. This memory might be a virtual machine's active memory.



When starting a virtual machine, the hypervisor creates a separate swap file (`.vswp`) for the virtual machine to support host-level swapping. Then, if necessary, the hypervisor can directly swap out guest physical memory to the swap file, which frees host physical memory for other virtual machines.

Host-level, or hypervisor, swapping is a guaranteed technique to reclaim a specific amount of memory within a specific amount of time. However, host-level swapping might severely penalize guest performance. Penalization occurs when the hypervisor has no knowledge of which guest physical pages should be swapped out and the swapping might cause unintended interactions with the native memory management policies in the guest operating system. For example, the guest operating system will never page out its kernel pages, because those pages are critical to ensure guest kernel performance. The hypervisor, however, cannot identify those guest kernel pages, so it might swap out those physical pages.

NOTE

ESXi swapping is distinct from the swapping performed by a guest operating system due to memory pressure in a virtual machine. Guest operating system level swapping might occur even when the ESXi host has ample resources.

https://t.me/learningnets

6-25 Activity: Identifying Memory Reclamation Features (1)

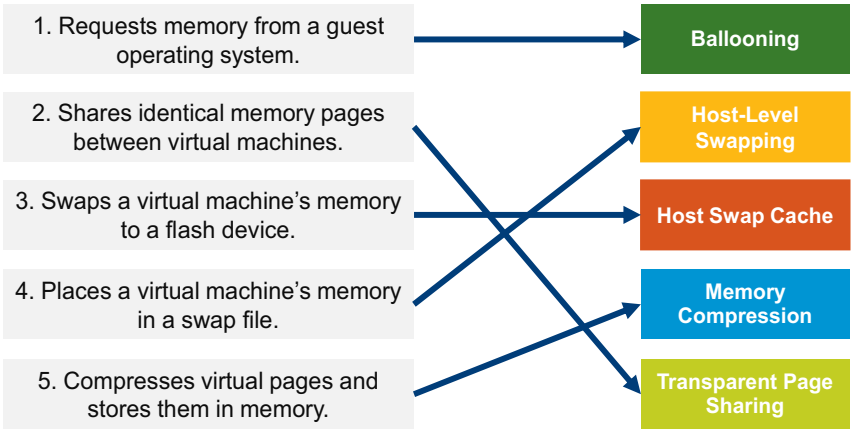
Match each description on the left with the memory reclamation feature on the right.

1. Requests memory from a guest operating system.	Ballooning
2. Shares identical memory pages between virtual machines.	Host-Level Swapping
3. Swaps a virtual machine's memory to a flash device.	Host Swap Cache
4. Places a virtual machine's memory in a swap file.	Memory Compression
5. Compresses virtual pages and stores them in memory.	Transparent Page Sharing

https://t.me/learningnets

6-26 Activity: Identifying Memory Reclamation Features (2)

Match each description on the left with the memory reclamation feature on the right.



https://t.me/learningnets

6-27 Sliding Scale Mem.MemMinFreePct

`Mem.MemMinFreePct` is the amount of memory that the VMkernel should keep free.

The VMkernel uses a sliding scale to determine the `Mem.MemMinFreePct` threshold based on the amount of memory installed on the host.

Memory Installed	Free State Threshold Range	Example: 128 GB Installed on Host
0 through 4 GB (first 4 GB)	6%	246 MB (6% of 4,096 MB)
4 through 12 GB (next 8 GB)	4%	328 MB (4% of 8,192 MB)
12 through 28 GB (next 16 GB)	2%	328 MB (2% of 16,384 MB)
Remaining memory	1%	1,024 MB (1% of 102,400 MB)
Sum of all percentages	Mem.MemMinFreePct	1,926 MB (246+328+328+1,024)

ESXi uses a sliding scale for determining the `Mem.MemMinFreePct` threshold. By using a sliding scale, a sensible value is calculated regardless of the memory size of the ESXi host.

For example, assume a server is configured with 128 GB of RAM. With the sliding-scale technique, the scale allocates 245.76 MB of RAM or 6 percent of the memory up to 4 GB. The scale then allocates 327.68 MB of memory out of the 4 GB through 12 GB range, 327.68 MB out of the 12 GB through 28 GB range, and 1024 MB out of the remaining memory, for a total of 1925.12 MB to calculate the `Mem.MemMinFreePct` value.

https://t.me/learningnets

6-28 Criteria for Reclaiming Host Memory

ESXi host physical memory is reclaimed based on these states: High, clear, soft, hard, and low.

Memory State	Memory Reclamation Methods Performed
High	Large pages are scanned and small page candidates for TPS are identified.
Clear	Break large pages and actively call TPS to collapse pages.
Soft	TPS + Balloon
Hard	TPS + Compress + Swap
Low	TPS + Compress + Swap + Block

ESXi enables page sharing by default and reclaims host physical memory with little overhead.

In the high state, the aggregate virtual machine guest memory usage is less than the host physical memory size. Normal TPS occurs.

The clear memory state was introduced in vSphere 6.0. In this state, large pages are broken into smaller pages, and TPS is actively called to collapse pages.

If the host free memory drops below the soft threshold, the hypervisor starts to reclaim memory by using ballooning. Ballooning happens before free memory reaches the soft threshold because it takes time for the balloon driver to allocate guest physical memory. Usually, the balloon driver can reclaim memory in a timely fashion so that the host free memory stays above the soft threshold.

If ballooning is not sufficient to reclaim memory or the host free memory drops below the hard threshold, the hypervisor starts to use memory compression and swapping. Through memory compression, the hypervisor should be able to quickly reclaim memory and bring the host memory state back to the soft state.

If the host free memory drops below the low threshold, the hypervisor reclaims memory through compression and swapping. The hypervisor also blocks the execution of all virtual machines that consume more memory than their target memory allocations.

https://t.me/learningnets

6-29 Transitioning from One Memory State to Another (1)

When memory pressure increases (that is, when free memory is decreasing), a set of thresholds are used to transition the ESXi host from one memory state to another.

For example, an ESXi host has 128 GB (131,072 MB) of RAM and the minimum free memory (Mem.MemMinFreePct, or MinFree for short) equals 1926 MB.

Memory State	Transition Threshold	Example
High		The ESXi host starts in High state.
Clear	300% of MinFree	When memory usage is greater than 125,294 MB, the ESXi host enters Clear state.
Soft	64% of MinFree	When memory usage is greater than 129,839 MB, the ESXi host enters Soft state.
Hard	32% of MinFree	When memory usage is greater than 130,456 MB, the ESXi host enters Hard state.
Low	16% of MinFree	When memory usage is greater than 130,764 MB, the ESXi host enters Low state.

In this example, the ESXi host transitions from High state to Clear state when memory usage is greater than 125,294 MB. This value is calculated as follows:

- The host has 128 GB, or 131,072 MB, of RAM. Using the transition threshold for Clear state, the value of 300% of MinFree is 3 * 1,926 MB, which equals 5,778 MB. Memory usage is 131,072 MB minus 5,778 MB, which equals 125,294 MB.

ESXi transitions from Clear state to Soft state when memory usage is greater than 129,839 MB. This value is calculated using the transition threshold of 64% of MinFree:

- 0.64 multiplied by 1,926 MB equals 1,233 MB. 131,072 MB minus 1,233 MB equals 129,839 MB.

ESXi transitions from Soft state to Hard state when memory usage is greater than 130,456 MB. This value is calculated using the transition threshold of 32% of MinFree:

- 0.32 multiplied by 1,926 MB equals 616 MB. 131,072 MB minus 616 MB equals 130,456 MB.

ESXi transitions from Hard state to Low state when memory usage is greater than 130,764 MB. This value is calculated using the transition threshold of 16% of MinFree:

- 0.16 multiplied by 1,926 MB equals 308 MB. 131,072 MB minus 308 MB equals 130,764 MB.

For a detailed discussion on memory reclamation, see VMware vSphere 6.5 Host Resources Deep Dive at http://frankdenneman.nl/publications.

https://t.me/learningnets

6-30 Transitioning from One Memory State to Another (2)

When memory pressure decreases (that is, when free memory is increasing), a different set of thresholds are used to transition the ESXi host from one memory state to another.

For example, an ESXi host has 128 GB (131,072 MB) of RAM and the minimum free memory (Mem.MemMinFreePct, or MinFree for short) equals 1926 MB.

Memory State	Transition Threshold	Example
Low		When memory usage is greater than 130,610 MB, the ESXi host remains in Low state.
Hard	24% of MinFree	When memory usage is less than 130,610 MB, the ESXi host enters Hard state.
Soft	48% of MinFree	When memory usage is less than 130,148 MB, the ESXi host enters Soft state.
Clear	100% of MinFree	When memory usage is less than 129,146 MB, the ESXi host enters Clear state.
High	400% of MinFree	When memory usage is less than 123,368 MB, the ESXi host enters High state.

In this example, the ESXi host transitions from Low state to Hard state when memory usage is less than 130,610 MB. This value is calculated using the transition threshold of 24% of MinFree:

- 0.24 multiplied by 1,926 MB equals 462 MB. 131,072 MB minus 462 MB equals 130,610 MB.

ESXi transitions from Hard state to Soft state when memory usage is less than 130,148 MB. This value is calculated using the transition threshold of 48% of MinFree:

- 0.48 multiplied by 1,926 MB equals 924 MB. 131,072 MB minus 924 MB equals 130,148 MB.

ESXi transitions from Soft state to Clear state when memory usage is less than 129,146 MB. This value is calculated using the transition threshold of 100% of MinFree:

- 1 multiplied by 1,926 MB equals 1,926 MB. 131,072 MB minus 1,926 MB equals 129,146 MB.

ESXi transitions from Clear state to High state when memory usage is less than 123,368 MB. This value is calculated using the transition threshold of 400% of MinFree:

- 4 multiplied by 1,926 MB equals 7,704 MB. 131,072 MB minus 7,704 MB equals 123,368 MB.

For a detailed discussion on memory reclamation, see VMware vSphere 6.5 Host Resources Deep Dive at http://frankdenneman.nl/publications.

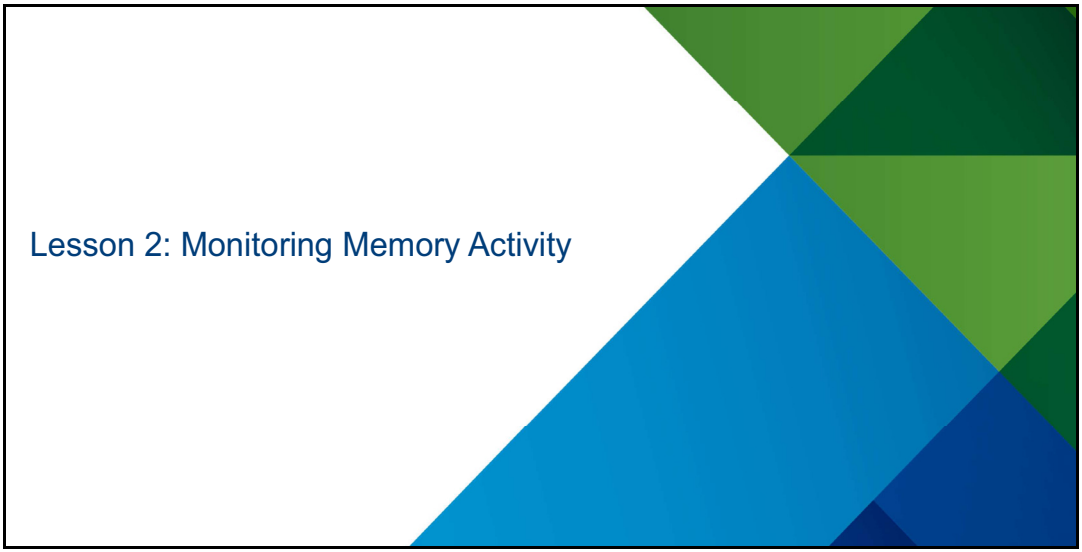
https://t.me/learningnets

6-31 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe how memory is used in a virtualized environment
- Explain each memory reclamation technique
- Explain how memory overcommitment affects performance

https://t.me/learningnets



https://t.me/learningnets

6-33 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Identify key memory metrics to monitor
- Use metrics in `esxtop`
- Monitor host memory usage and virtual machine memory usage
- Monitor host swapping activity
- Monitor host ballooning activity

https://t.me/learningnets

6-34 Memory Usage Metrics in the Guest Operating System

Guest or host memory usage might appear different from what is seen in the operating system:

- The guest operating system has better information about real memory usage than the host when estimating active memory.
- The ESXi active memory estimate technique can take time to converge.
- Host memory usage does not correspond to any memory metric in the guest.
- Host memory usage size is based on a virtual machine's relative priority on the physical host and memory usage by the guest.

For guest memory usage, the values reported in vSphere Web Client might be different from the active memory usage reported by the guest operating system. The first reason for this difference is that the guest operating system generally has a better idea than the hypervisor of what memory is active. The guest operating system knows what applications are running and how it has allocated memory. The second reason is that the method employed by the hypervisor to estimate active memory usage takes time to converge. So the guest operating system's estimate might be more accurate than the ESXi host's if the memory workload is fluctuating.

For host memory usage, the host memory usage metric has no meaning inside the guest operating system. It has no meaning because the guest operating system does not know that it is running in a virtual machine or that other virtual machines exist on the same physical host.

https://t.me/learningnets

6-35 About Consumed Host Memory and Active Guest Memory

Consumed host memory can be higher than active guest memory:

- The consumed host memory represents the highest amount of memory used by the guest.
- If memory is not overcommitted, this combination is good.

Consumed host memory can be less than or equal to active guest memory:

- Active guest memory does not completely reside in host physical memory.
- This combination points to potential performance degradation.

Consumed host memory is greater than active guest memory because, for physical hosts that are not overcommitted on memory, consumed host memory represents the highest amount of memory usage by a virtual machine. In the past, this virtual machine might have actively used a large amount of memory.

Because the host physical memory is not overcommitted, the hypervisor has no reason to invoke ballooning or host-level swapping to reclaim memory. Thus, you can find cases where the active guest memory use is low, but the amount of host physical memory assigned to it is high. This combination is a normal situation.

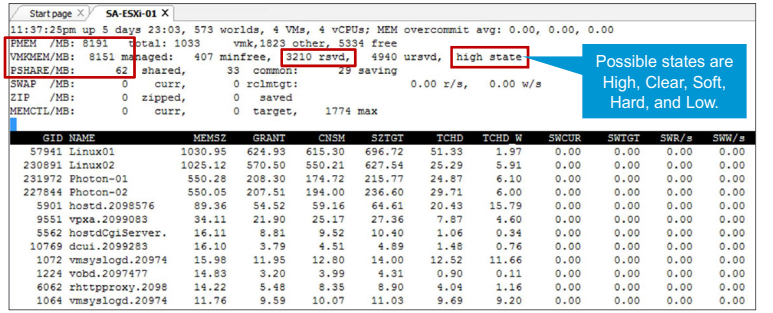
Consumed host memory might be less than or equal to active guest memory because the active guest memory of a virtual machine does not completely reside in the host physical memory. Consumed host memory might be less if a guest's active memory is reclaimed by the balloon driver or if the virtual machine is swapped out by the hypervisor. Either case of lower consumed host memory is probably due to high memory overcommitment.

https://t.me/learningnets

6-36 Using esxtop to Monitor Memory Usage

`esxtop` offers several options for monitoring memory usage.

Enter **m** to display the memory screen.



The following metrics are useful in monitoring memory usage:

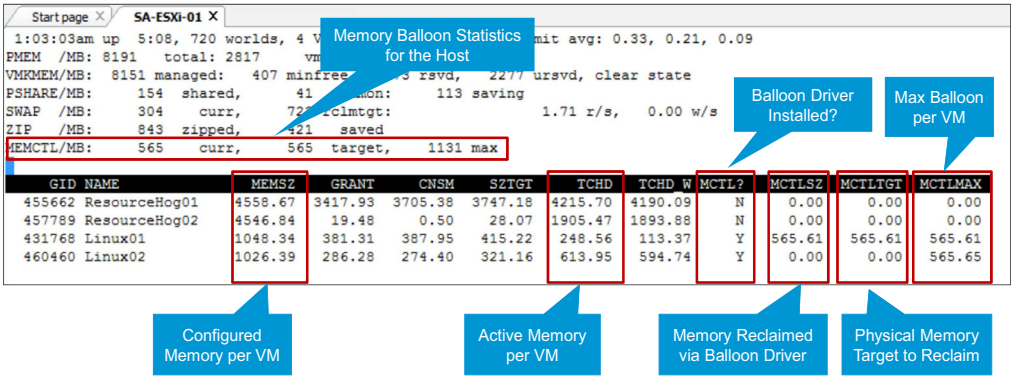
- PMEM: The total amount of physical memory on your host in megabytes.
- VMKMEM: The memory that is managed by the VMkernel in megabytes.
- PSHARE: The saving field shows you how much memory you saved because of TPS. In the example, two virtual machines are running, with a memory savings of 62 MB.
- Memory state: This value can be high, clear, soft, hard, or low. This value refers to the current state of the VMkernel's memory. This value also tells you whether the VMkernel has enough free memory to perform its critical operations.

  If the state is high, the VMkernel has sufficient memory to perform critical tasks. However, if the state is clear, soft, hard, or low, the VMkernel is unable to maintain adequate free memory. Typically, you should see only a state of high, even if you are overcommitting your memory. ESXi employs efficient mechanisms for reserving memory for the VMkernel. However, if your host is severely overcommitted, you might see a state other than high.
- A PCI hole is the memory reserved by the PCI devices on the host. This memory is used solely by the PCI devices for I/O purposes, such as DMA ring buffers. This memory cannot be used by the VMkernel. Reserving memory by PCI devices is common in all operating systems. The amount of memory used for PCI devices is part of the other value, in the PMEM/MB line.

https://t.me/learningnets

6-37 Host Ballooning Activity in esxtop (1)

On the `esxtop` memory screen, enter **`f`** to display the fields window and then enter **`j`** to add the memory control fields used to monitor virtual machine ballooning activity.



The following metrics are useful in monitoring ballooning activity:

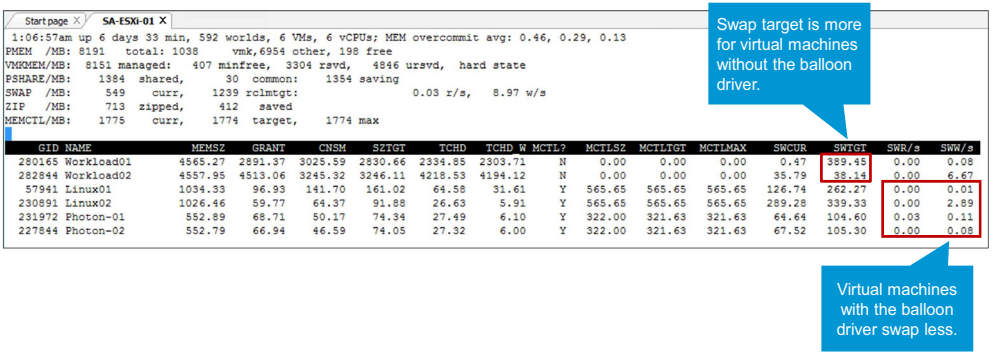
- MEMCTL/MB: This line displays the memory balloon statistics for the entire host. All numbers are in megabytes:
  - curr is the total amount of physical memory reclaimed using the balloon driver (`vmmemctl`).
  - target is the total amount of physical memory ESXi wants to reclaim with the balloon driver.
  - max is the maximum amount of physical memory that the ESXi host can reclaim with the balloon driver.
- MCTL?: This value, which is either Y (for yes) or N (for no), indicates whether the balloon driver is installed in the virtual machine.
- MCTLSZ: This value is reported for each virtual machine and represents the amount of physical memory that the balloon driver is holding for use by other virtual machines.
- MCTLTGT: This value is the amount of physical memory that the host wants to reclaim from the virtual machine through ballooning.

https://t.me/learningnets

6-38 Host Ballooning Activity in esxtop (2)

The `esxtop` memory screen shows where ballooning is used and where swapping is used.

Enter **`f`** to display the fields window and then enter **`k`** to add the swap statistics fields.



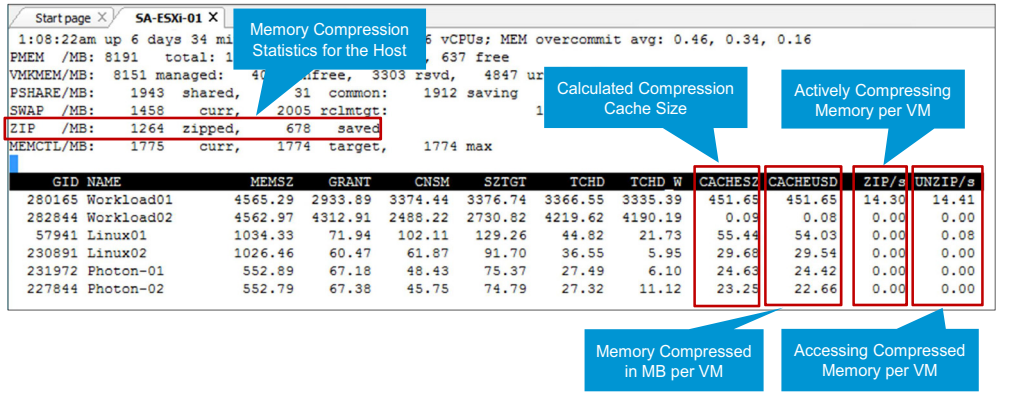
The example shows why ballooning is preferable to swapping. Multiple virtual machines are running, and all the virtual machines are running a memory-intensive application. The MCTL? field shows that the virtual machines named Workload01 and Workload02 do not have the balloon driver installed. The balloon driver is installed on all the other virtual machines. If the balloon driver is not installed, typically VMware Tools was not installed on the virtual machine.

The virtual machines with no balloon driver installed usually have a higher swap target (SWTGT) than the virtual machines that have the balloon driver installed. Virtual machines that have the balloon driver installed can reclaim memory by using the balloon driver. If ESXi cannot reclaim memory through ballooning, then it must resort to other methods, such as host-level swapping.

https://t.me/learningnets

6-39 Host Memory Compression Activity in esxtop

On the `esxtop` memory screen, enter **`f`** to display the fields window and then enter **`q`** to add the memory compression fields to monitor virtual machine compression activity.



ESXi provides a memory compression cache to improve virtual machine performance when you use memory overcommitment. Memory compression is enabled by default. When a host's memory becomes overcommitted, ESXi compresses virtual pages and stores them in memory.

Because accessing compressed memory is faster than accessing memory that is swapped to disk, memory compression in ESXi enables you to overcommit memory without significantly hindering performance. When a virtual page needs to be swapped, ESXi first tries to compress the page. Pages that can be compressed to 2 KB or smaller are stored in the virtual machine's compression cache, increasing the capacity of the host.

You can monitor the following values in `esxtop`:

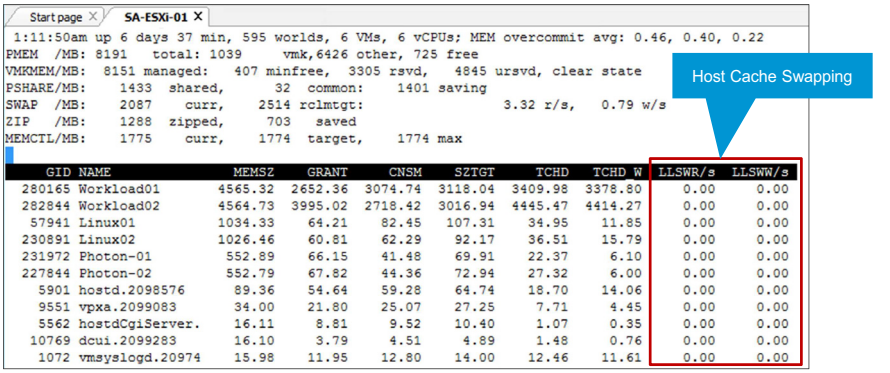
- CACHESZ (MB): Compression memory cache size
- CACHEUSD (MB): Used compression memory cache
- ZIP/s (MB/s): Compressed memory per second
- UNZIP/s (MB/s): Decompressed memory per second

https://t.me/learningnets

6-40 Monitoring Host Cache Swapping

On the `esxtop` memory screen, enter **`f`** to display the fields window and then enter **`l`** to add the swap statistics fields to monitor host cache swapping activity.

`esxtop` reports on host cache swapping in the LLSWR/s and LLSWW/s fields.

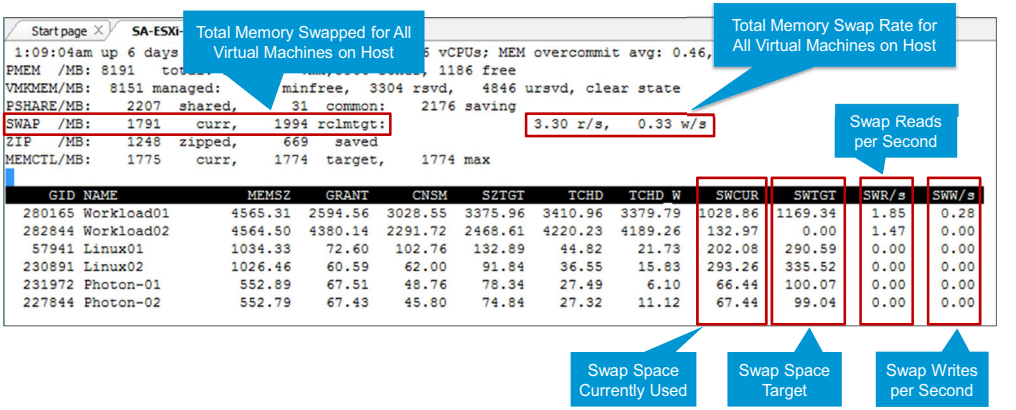


In `esxtop`, LLSWR/s is the rate (in MB) at which memory is read from the host cache. LLSWW/s is the rate (in MB) at which memory is written to the host cache.

https://t.me/learningnets

6-41 Host Swapping Activity in esxtop: Memory Screen

On the `esxtop` memory screen, enter **`f`** to display the fields window and then enter **`k`** to add the swap statistics fields to monitor virtual machine swapping activity.



A useful metric in the memory screen is SWAP/MB. This metric represents total swapping for all the virtual machines on the host. In the example, the value of curr is 1791 MB, which means that 1,791 MB of swap space is currently used. The rclmtgt value (1994 in the example) is where the ESXi system expects the reclaimed memory to be. Memory can be reclaimed by swapping or by compression.

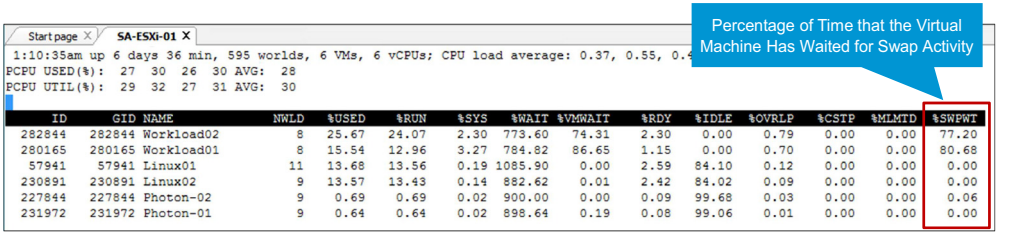
To monitor host-level swapping activity per virtual machine, enter uppercase **`v`** in the `esxtop` window to display only the virtual machines. The following metrics are useful:

- SWR/s and SWW/s: Measured in megabytes, these counters represent the rate at which the ESXi host is swapping memory in from disk (SWR/s) and swapping memory out to disk (SWW/s).
- SWCUR: The amount of swap space currently used by the virtual machine.
- SWTGT: The amount of swap space that the host expects the virtual machine to use.

https://t.me/learningnets

6-42 Host Swapping Activity in esxtop: CPU Screen

In `esxtop`, enter **c** to display the CPU screen for virtual machine worlds and then enter **V** to display only the virtual machines.



The CPU screen has a metric named %SWPWT. This metric is the best indicator of a performance problem due to wait time experienced by the virtual machine. This metric represents the percentage of time that the virtual machine is waiting for memory pages to be swapped in.

The ready time (%RDY) is low because no competition exists for the CPU resource until the pages are swapped in.

https://t.me/learningnets

6-43 Causes of Active Host-Level Swapping

The basic cause of host-level swapping is memory overcommitment from using memory-intensive virtual machines whose combined configured memory is greater than the amount of host physical memory available.

The causes of active host-level swapping include excessive memory overcommitment, memory overcommitment with memory reservations, and balloon drivers in virtual machines not running or disabled.

The following conditions can cause host-level swapping:

- Excessive memory overcommitment: In this condition, the level of memory overcommitment is too high for the combination of memory capacity, virtual machine allocations, and virtual machine workload. In addition, the amount of overhead memory must be taken into consideration when using memory overcommitment.
- Memory overcommitment with memory reservations: In this condition, the total amount of balloonable memory might be reduced by memory reservations. If a large portion of a host's memory is reserved for some virtual machines, the host might be forced to swap the memory of other virtual machines. Swapping memory with other virtual machines might occur even though virtual machines are not actively using all of their reserved memory.
- Balloon drivers in virtual machines not running or disabled: In this condition, if the balloon driver is not running or has been deliberately disabled in some virtual machines, the amount of balloonable memory on the host will be decreased. As a result, the host might be forced to swap even though idle memory is available that might have been reclaimed through ballooning.

https://t.me/learningnets

6-44 Resolving Host-Level Swapping

You have several ways to resolve performance problems that are caused by active host-level swapping:

- Reduce the level of memory overcommitment:
  - Use vSphere vMotion.
- Enable the balloon driver in all virtual machines:
  - Install VMware Tools™.
- Add memory to the host.
- Reduce memory reservations:
  - Configured virtual machine memory can be overcommitted, not reserved virtual machine memory.
- Use resource controls to dedicate memory to critical virtual machines:
  - Use resource controls only as a last resort.

In most situations, reducing memory overcommitment levels is the proper approach for eliminating swapping on an ESXi host. However, you should consider as many factors as possible to ensure that the reduction is adequate to eliminate swapping. If other approaches are used, monitor the host to ensure that swapping was eliminated.

https://t.me/learningnets

6-45 Activity: Identifying esxtop Fields (1)

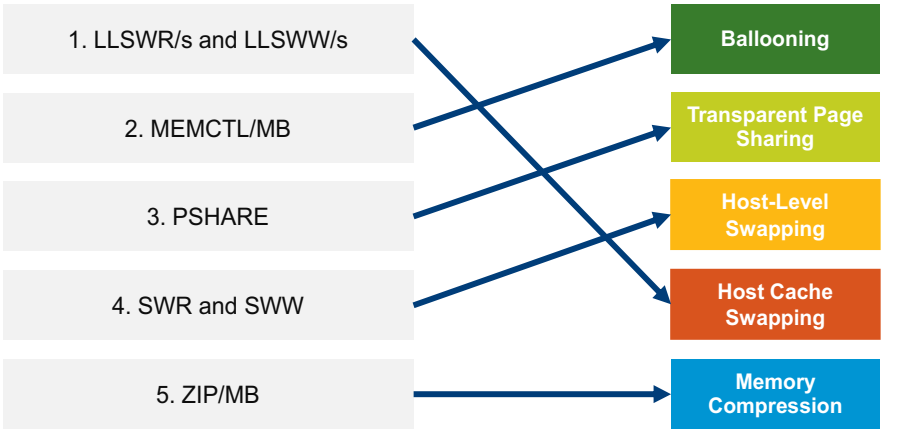
Match the `esxtop` fields on the left with the memory reclamation technique that each monitors on the right.

1. LLSWR/s and LLSWW/s	Ballooning
2. MEMCTL/MB	Transparent Page Sharing
3. PSHARE	Host-Level Swapping
4. SWR and SWW	Host Cache Swapping
5. ZIP/MB	Memory Compression

https://t.me/learningnets

6-46 Activity: Identifying esxtop Fields (2)

Match the `esxtop` fields on the left with the memory reclamation technique that each monitors on the right.



https://t.me/learningnets

6-47 Reducing Memory Overcommitment

Methods to reduce the level of memory overcommitment include the following actions:

- Add physical memory to the ESXi host.
- Reduce the number of virtual machines running on the ESXi host.
- Increase available memory resources by adding the host to a vSphere DRS cluster.

To reduce the level of memory overcommitment, perform one or more of the following actions:

- Add physical memory to the ESXi host: Adding physical memory to the host reduces the level of memory overcommitment and might eliminate the memory pressure that caused swapping to occur.
- Reduce the number of virtual machines running on the ESXi host: One way to reduce the number of virtual machines is to use vSphere vMotion to migrate virtual machines to hosts with available memory resources. To use vSphere vMotion successfully, you should use the vSphere Web Client performance charts to determine the memory usage of each virtual machine on the host and the available memory resources on the target hosts. You must ensure that the migrated virtual machines will not cause swapping to occur on the target hosts.

  If additional ESXi hosts are not available, you can reduce the number of virtual machines by powering off noncritical virtual machines. Reducing the number of running virtual machines makes additional memory resources available for critical applications. Even idle virtual machines consume some memory resources.
- Increase available memory resources by adding the host to a vSphere DRS cluster: Using a vSphere DRS cluster is similar to the previous solution. However, in a vSphere DRS cluster, load rebalancing can be performed automatically. You do not have to manually compute the compatibility of specific virtual machines and hosts or account for peak usage periods.

https://t.me/learningnets

6-48 Enabling Balloon Driver in Virtual Machines

To maximize the ability of ESXi to recover idle memory from virtual machines, enable the balloon driver in all virtual machines.

To enable the balloon driver in a virtual machine, install VMware Tools.

If a virtual machine has critical memory needs, use resource controls to satisfy those needs.

A best practice is to always keep the balloon driver enabled.

If a virtual machine has critical memory needs, then reservations and other resource controls should be used to ensure that those needs are satisfied. The balloon driver is enabled when VMware Tools is installed on the virtual machine.

However, if excessive memory overcommitment exists, ballooning only delays the onset of swapping. Ballooning is therefore not a sufficient solution. The level of memory overcommitment should be reduced as well.

The balloon driver should never be deliberately disabled in a virtual machine. Disabling the balloon driver might cause unintended performance problems, such as host-level swapping. It also makes tracking down memory-related problems more difficult. vSphere provides other mechanisms, such as memory reservations, for controlling the amount of memory available to a virtual machine.

https://t.me/learningnets

6-49 Reducing Virtual Machine Memory Reservation

Reevaluate the memory reservation of a virtual machine if this reservation causes the host to swap virtual machines without reservations.

Reduce the memory reservation of a virtual machine if the virtual machine is not using its full reservation.

If the reservation cannot be reduced, memory overcommitment must be reduced.

https://t.me/learningnets

6-50 Dedicating Memory to Critical Virtual Machines

Use memory reservations to prevent critical virtual machines from swapping.

Be conservative with using memory reservations, because you might move the swapping problem to other virtual machines.

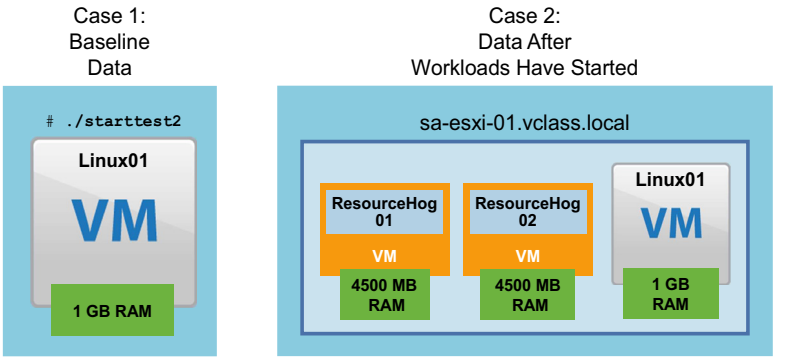
When reducing the level of memory overcommitment is impossible, configure your performance-critical virtual machines with sufficient memory reservations to prevent them from swapping.

However, using memory reservations only moves the swapping problem to other virtual machines whose performance will be severely degraded. In addition, the swapping of other virtual machines might still affect the performance of virtual machines with reservations, due to added disk traffic and memory management overhead. This approach should be used only with caution when no other options exist.

https://t.me/learningnets

6-51 Introduction to Lab 9: Monitoring Memory Performance

The diagram shows the test cases that you perform.



In this lab, you perform two tests:

- Test case 1: Run `starttest2` on your test virtual machine and record baseline performance data.
- Test case 2: Start the ResourceHog01 and ResourceHog02 VMs to generate memory contention. After the VMs start up, record performance data.

Answer questions posed in the lab and compare the baseline data with the performance data after memory contention is generated.

https://t.me/learningnets

6-52 Lab 9: Monitoring Memory Performance

Use the `esxtop` command to monitor memory performance under load

1. Generate Database Activity in the Test Virtual Machine
2. Check for Overcommitment of Virtual Machine Memory
3. Configure `esxtop` to Report Virtual Machine Memory Statistics
4. Observe Memory Statistics
5. Start a Memory Test on ResourceHog01 and ResourceHog02
6. Record Memory Statistics
7. Clean Up for the Next Lab

https://t.me/learningnets

6-53 Review of Lab 9: Monitoring Memory Performance

Your instructor will review these questions with you:

- For Linux01, did the values for MCTLSZ and MCTLTGT converge?
- For Linux01, did the values for SWCUR and SWTGT converge?
- Which virtual machines do not have the balloon driver installed?
- Which virtual machines experienced swapping activity? Did the %SWPWT value exceed 5 percent for any of these virtual machines?
- What is the memory state (high, clear, soft, hard, or low)?
- Based on the OPM values that you recorded in the lab, did the performance of the starttest2 script degrade?

Memory Counters	Resource Hog02	Resource Hog01	Linux01
MCTL?			
MCTLSZ			
MCTLTGT			
SWCUR			
SWTGT			
SWR/s			
SWW/s			
%SWPWT			

https://t.me/learningnets

6-54 Review of Learner Objectives

You should be able to meet the following objectives:

- Identify key memory metrics to monitor
- Use metrics in `esxtop`
- Monitor host memory usage and virtual machine memory usage
- Monitor host swapping activity
- Monitor host ballooning activity

https://t.me/learningnets

6-55 Key Points

- The hypervisor uses memory reclamation techniques to reclaim host physical memory. TPS, ballooning, memory compression, and host-level swapping are used to reclaim memory.
- Host swap rates and ballooning activity are key memory performance metrics.
- The basic cause of host memory swapping is a combination of memory overcommitment and memory-intensive virtual machines.

Questions?

https://t.me/learningnets

https://t.me/learningnets

MODULE 7

Storage Optimization



https://t.me/learningnets

7-2 You Are Here

1. Course Introduction
2. Network Scalability
3. Storage Scalability
4. Host and Management Scalability
5. CPU Optimization
6. Memory Optimization
7. **Storage Optimization**
8. Network Optimization
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

7-3 Importance

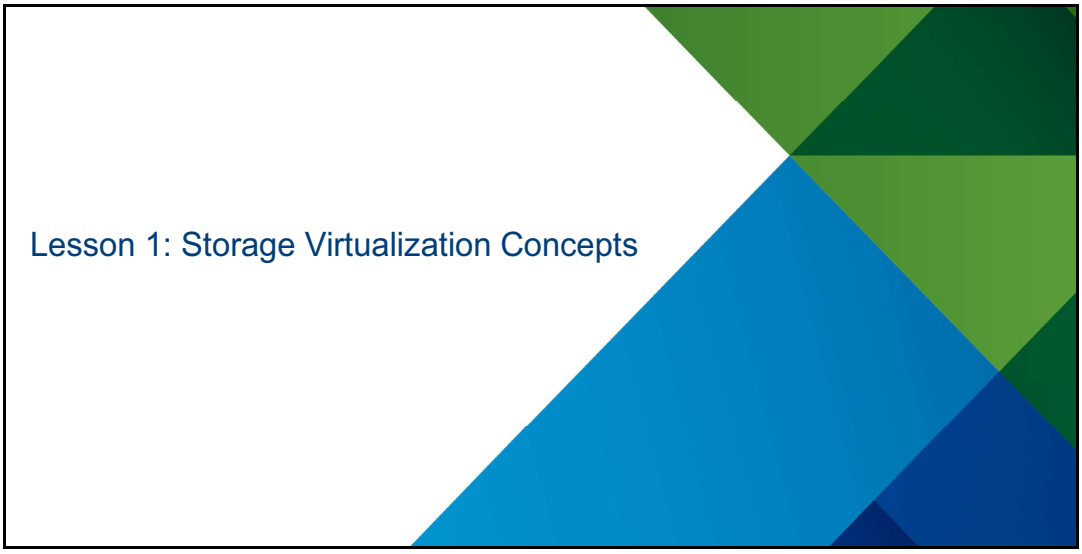
Storage can limit the performance of enterprise workloads. You should know how to monitor a host's storage throughput.

https://t.me/learningnets

7-4 Module Lessons

Lesson 1:	Storage Virtualization Concepts
Lesson 2:	Monitoring Storage Activity

https://t.me/learningnets



https://t.me/learningnets

7-6 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

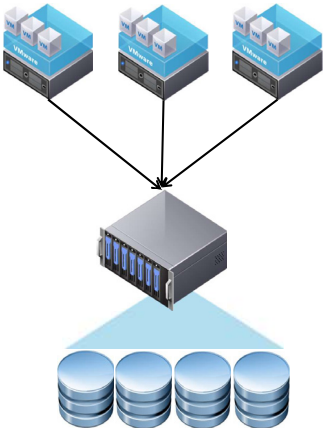
- Describe factors that affect storage performance

https://t.me/learningnets

7-7 Storage Performance Overview

Performance of shared storage in vSphere hinges on several factors:

- Proper configuration of your storage devices
- Load balancing across available storage
- Storage queues



ESXi enables multiple hosts to reliably share the same physical storage through its optimized storage stack. Shared storage of virtual machines can be accomplished by using VMFS, NFS, vSAN, and vSphere Virtual Volumes. Shared storage enables virtualization capabilities such as vSphere vMotion, vSphere DRS, and vSphere HA.

To get the most from shared storage, you must understand the storage performance limits of a given physical environment. Understanding limits helps ensure that you do not overcommit resources.

https://t.me/learningnets

7-8 Storage Protocol Performance

ESXi supports Fibre Channel, Fibre Channel over Ethernet, hardware iSCSI, software iSCSI, and NFS.

All storage protocols are capable of delivering high throughput performance. When CPU resources are not causing a bottleneck, software iSCSI and NFS can be part of a high-performance solution.

ESXi hosts provide support for high-performance hardware features:

- 32 Gb Fibre Channel
- Software and hardware iSCSI and NFS support for jumbo frames:
  - Using 1, 10, 25, 40, 50 and 100 Gb Ethernet NICs
  - Using 10, 20, 25 and 40 Gb iSCSI hardware initiators

For Fibre Channel, Fibre Channel over Ethernet (FCoE), and hardware iSCSI, a major part of the protocol processing is offloaded to the host bus adapter (HBA). Consequently, the cost of each I/O is very low.

For software iSCSI and NFS, host CPUs are used for protocol processing, which increases cost. Furthermore, the cost of NFS and software iSCSI is higher with larger block sizes, such as 64 KB. This cost is due to the additional CPU cycles needed for each block for tasks, such as check summing and blocking. Software iSCSI and NFS are more efficient at smaller blocks. Both are capable of delivering high throughput performance when CPU resource is not causing a bottleneck.

https://t.me/learningnets

7-9 About SAN Configuration

Proper SAN configuration can help eliminate performance problems.

Each LUN should have the correct RAID level and storage characteristics for the applications in virtual machines that use it.

The choice of path selection policy greatly affects SAN performance:

- Most recently used
- Fixed
- Round robin
- Optional third-party path selection policy

Storage performance is a vast topic that depends on workload, hardware, vendor, RAID level, cache size, stripe size, and so on. Consult the appropriate VMware documentation, as well as storage vendor documentation, for information on how to configure your storage devices appropriately.

Because each application running in your vSphere environment has different requirements, you can achieve high throughput and minimal latency by choosing the appropriate RAID level and path selection policy for applications running in the virtual machines.

By default, active-passive storage arrays use the Most Recently Used path policy. To avoid LUN thrashing, do not use the Fixed path policy for active-passive storage arrays.

By default, active-active storage arrays use the Fixed path policy. When using this policy, you can maximize the use of your bandwidth to the storage array by designating preferred paths to each LUN through different storage controllers.

The Round Robin policy uses an automatic path selection by rotating through all available paths and enabling the distribution of load across the configured paths:

- For active-passive storage arrays, only the paths to the active controller are used in the Round Robin policy.
- For active-active storage arrays, all paths are used in the Round Robin policy.

See *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-storage-guide.pdf.pdf.

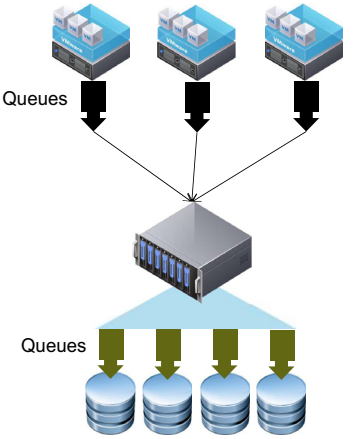
https://t.me/learningnets

7-10 About Storage Queues

Several storage queue types exist:

- Queuing at the host:
  - The device driver queue controls the number of active commands on the LUN at any time.
  - The VMkernel queue is an overflow queue for device driver queue.
- Queuing at the storage array:
  - Queuing occurs when the number of active commands to a LUN is too high for the storage array to handle.
  - The queue is located at the LUN for VMFS datastores, and at the protocol endpoint for vSphere Virtual Volumes datastores.

Latency increases with excessive queuing at the host or storage array.



The device driver queue is used for low-level interaction with the storage device. This queue controls how many active commands can be on a LUN at the same time. This number is effectively the concurrency of the storage stack. Set the device queue to 1, and each storage command becomes sequential, that is, each command must complete before the next starts.

The kernel queue can be thought of as an overflow queue for the device driver queues. A kernel queue includes features that optimize storage. These features include multipathing for failover and load balancing, prioritization of storage activities based on virtual machine and cluster shares, and optimizations to improve efficiency for long sequential operations.

SCSI device drivers have a configurable parameter, called the LUN queue depth, that determines how many commands can be active at a time on a given LUN. If the total number of outstanding commands from all virtual machines exceeds the LUN queue depth, the excess commands are queued in the ESXi kernel, which increases latency.

In addition to queuing at the ESXi host, command queuing can also occur at the storage array.

https://t.me/learningnets

7-11 Network Storage: iSCSI and NFS

Avoid oversubscribing links to storage:

- Applications or systems that write large amounts of data to storage should not share Ethernet links to a storage device.
- Oversubscribed links cause packet drops, which degrade performance.
- VLANs do not solve the problem of oversubscription, but isolating iSCSI traffic and NFS traffic on private VLANs is helpful.

For software iSCSI and NFS, protocol processing uses CPU resources on the host.

Applications and systems such as data acquisition or transaction logging systems perform best with multiple connections to storage devices.

For iSCSI and NFS, ensure that your network topology does not contain Ethernet bottlenecks. Bottlenecks where multiple links are routed through fewer links can result in oversubscription and dropped network packets. Any time that a number of links transmitting near capacity are switched to a smaller number of links, such oversubscription becomes possible. Recovering from dropped network packets results in large performance degradation.

Using VLANs or VPNs does not provide a suitable solution to the problem of link oversubscription in shared configurations. However, creating separate VLANs for NFS and iSCSI is beneficial. This separation minimizes network interference from other packet sources.

Finally, with software-initiated iSCSI and NFS, the network protocol processing takes place on the host system and thus can require more CPU resources than other storage options.

https://t.me/learningnets

7-12 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe factors that affect storage performance

https://t.me/learningnets



https://t.me/learningnets

7-14 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Determine which disk metrics to monitor
- Identify metrics in `esxtop`
- Demonstrate how to monitor disk throughput

https://t.me/learningnets

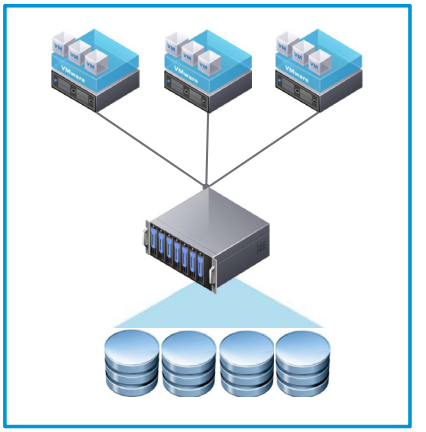
7-15 Where Storage Problems Can Occur

A storage solution network is only as fast as the slowest device in the network.

Problems can arise at several levels:

- Virtual machine: Virtual machine disk (VMDK) file
- ESXi host: HBA or NIC
- Physical storage:
  - Storage network switches
  - Storage processors
  - Storage array network interfaces

You can identify storage performance problems by using `esxtop` and understanding the storage key performance indicators.



https://t.me/learningnets

7-16 Storage Key Performance Indicators

To identify disk-related performance problems, determine the available bandwidth on your host and compare it with your expectations.

In vSphere, the key storage performance indicators to monitor are as follows:

- Disk throughput
- Latency (device, kernel, and so on)
- Number of aborted disk commands
- Number of active disk commands
- Number of disk commands queued

Several questions can be asked when looking at disk performance issues:

- Are you getting the expected I/Os per second (IOPS)?
- Are you getting the expected bandwidth (read/write)?

The same questions apply to disk latency. Check disk latency and compare it with your expectations. Are the latencies higher than you expected? Compare performance with the hardware specifications of the particular storage subsystem.

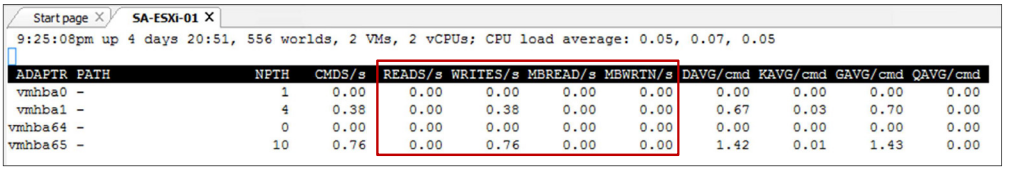
Disk bandwidth and latency help determine whether storage is overloaded or slow.

https://t.me/learningnets

7-17 Monitoring Disk Throughput by Storage Adapter (1)

Use the `esxtop` command to monitor disk throughput.

Enter **d** to display disk throughput by storage adapter.



Disk throughput appears in these columns:

- READS/s and WRITES/s
- MBREAD/s and MBWRTN/s

Disk throughput can also be monitored using `esxtop`:

- READS/s: Number of disk reads per second
- WRITES/s: Number of disk writes per second

The sum of reads per second and writes per second equals IOPS. IOPS is a common benchmark for storage subsystems and can be measured with tools like Iometer.

If you prefer, you can monitor throughput by using the following metrics instead:

- MBREAD/s: Number of megabytes read per second
- MBWRTN/s: Number of megabytes written per second

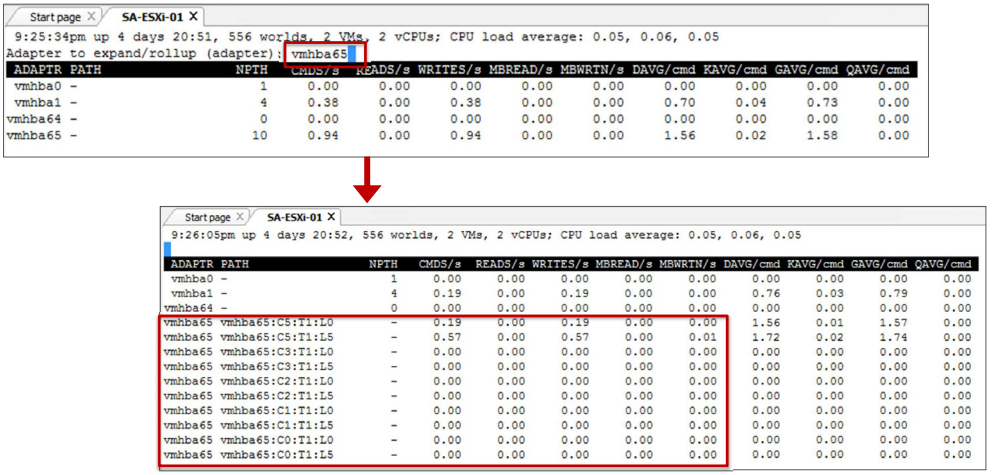
All of these metrics can be monitored per HBA (vmhba#).

The `esxtop` screen on the slide shows disk activity per disk adapter. To display the set of fields, enter **f**. Ensure that the following fields are selected: A (adapter name), B (path name), C (number of paths), E (I/O statistics), and G (overall latency statistics). If necessary, select fields to display by entering **a**, **b**, **c**, **e**, or **g**.

https://t.me/learningnets

7-18 Monitoring Disk Throughput by Storage Adapter (2)

Enter **e** and the adapter name to display disk/LUN throughput for a specific storage adapter.



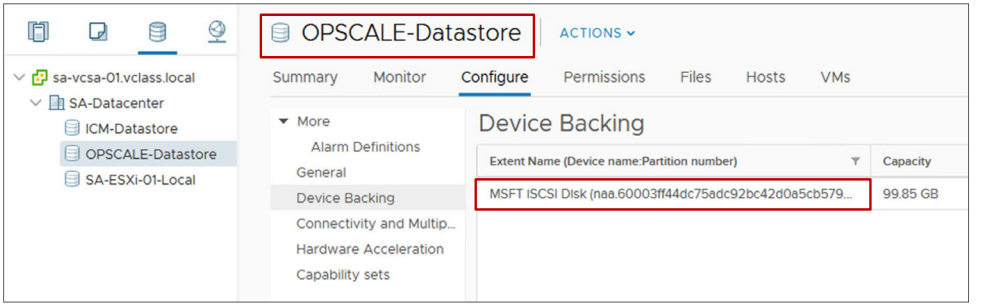
In this example, vmhba65 is entered in the `esxtop` window. `esxtop` displays disk/LUN throughput information for each of the adapter's paths.

https://t.me/learningnets

7-19 Correlating Storage Devices to Datastores (1)

Use vSphere Client to identify which storage devices (LUNs) are associated with a datastore:

- View the Device Backing panel for a particular datastore.
- Record the device name and its unique identifier, for example, the NAA ID.



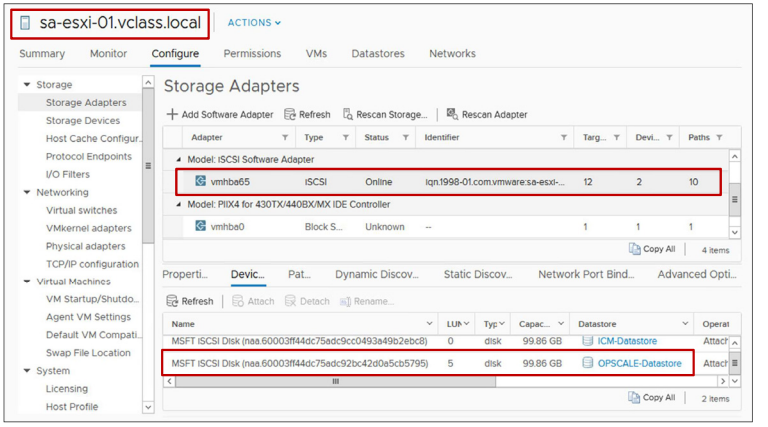
Storage administrators usually keep a record (for example, in a spreadsheet) of each storage device, its unique ID, and its use.

You can use vSphere Client to correlate the storage device with a datastore. In the example, the LUN identified as naa.60003ff44dc75adc92bc42d0a5cb5795 is a Microsoft iSCSI disk that backs the OPSCALE-Datastore datastore.

https://t.me/learningnets

7-20 Correlating Storage Devices to Datastores (2)

View the Storage Adapters panel for the ESXi host that is connected to the datastore. Read the device information and identify the LUN by its unique identifier.



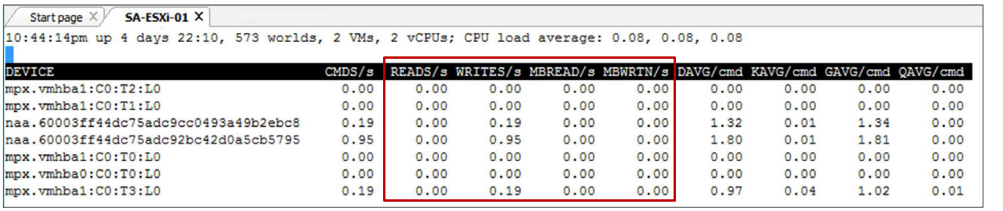
You can also identify what ESXi host or hosts can access that particular LUN. In this example, the naa.60003ff44dc75adc92bc42d0a5cb5795 LUN is accessed by the sa-esxi-01.vclass.local host.

In addition, you can view which ESXi hosts can access a particular datastore by selecting the datastore in the Navigator pane and selecting **Configure** > **Connectivity and Multipathing**.

https://t.me/learningnets

7-21 Monitoring Path Statistics for a Specific Storage Device (1)

In the `esxtop` window, enter **`u`** to display disk throughput by storage device.



A storage device can be associated with a particular VMFS datastore.

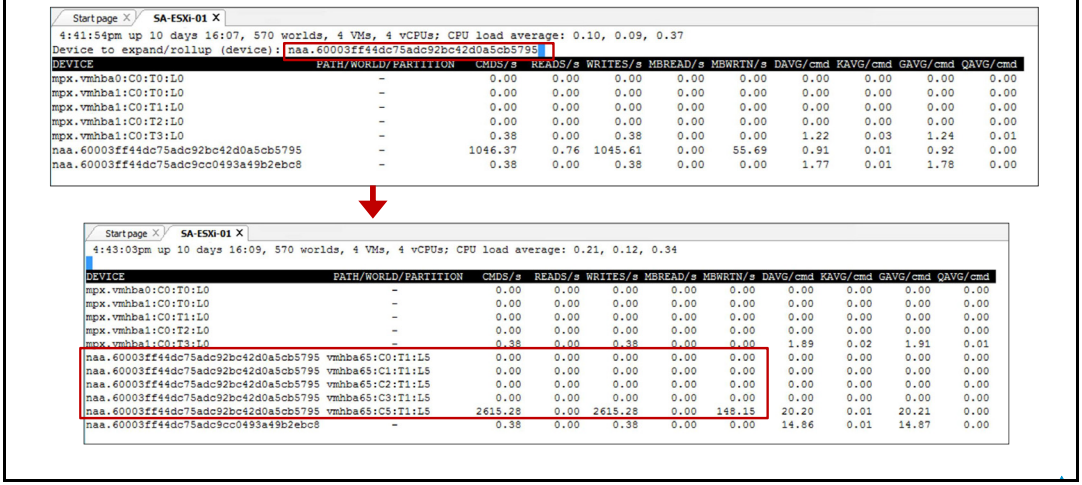
Storage devices (or LUNs) are identified by a unique identifier, such as a Network Address Authority ID (NAA ID) or a T10 identifier.

This `esxtop` screen shows disk activity by storage device. To display the set of fields, enter **`f`**. Ensure that the following fields are selected: A (device name), B (path, world, partition ID), F (queue statistics), G (I/O statistics), and I (overall latency statistics). If necessary, select fields to display by entering **`a`**, **`b`**, **`f`**, **`g`**, or **`i`**.

https://t.me/learningnets

7-22 Monitoring Path Statistics for a Specific Storage Device (2)

Enter **p** and the device ID to display path statistics to a specific storage device.

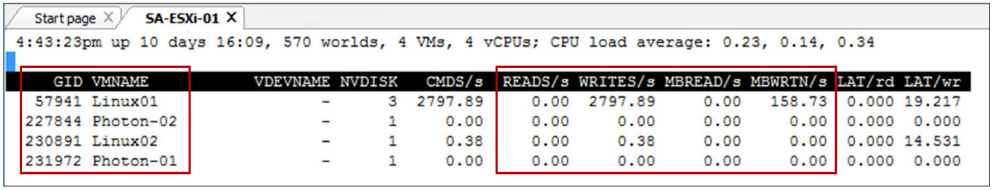


In this example, naa.60003ff44dc75adc92bc42d0a5cb5795 is entered in the `esxtop` window. `esxtop` displays path statistics for this particular device.

https://t.me/learningnets

7-23 Monitoring Individual VMDK Statistics for a VM (1)

On the `esxtop` screen, enter **`v`** to display disk throughput by virtual machine.

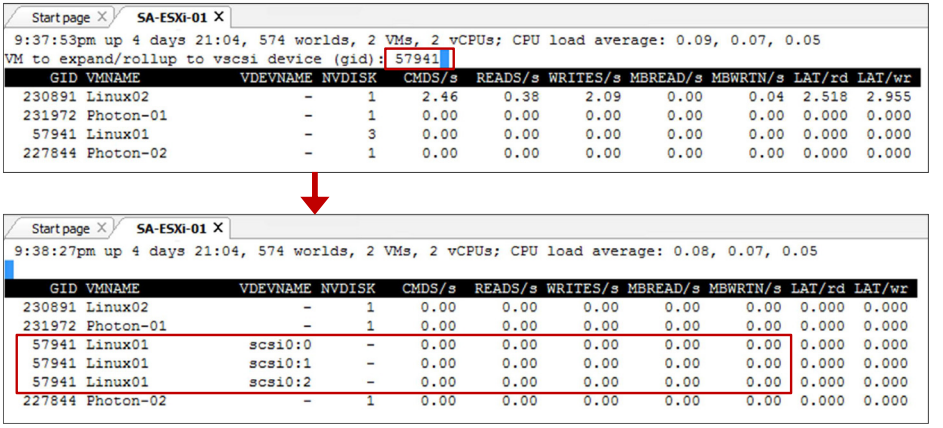


This `esxtop` screen shows disk activity per virtual machine. To display the set of fields, enter **`f`**. Ensure that the following fields are selected: B (group ID), C (VM name), D (virtual device name), E, (number of virtual disks), I (I/O statistics), J (read latency statistics), and K (write latency statistics). If necessary, select fields to display by entering **`b`**, **`c`**, **`d`**, **`e`**, **`i`**, **`j`**, or **`k`**.

https://t.me/learningnets

7-24 Monitoring Individual VMDK Statistics for a VM (2)

Enter **e** and the VM's GID to display individual VMDK statistics for that VM.



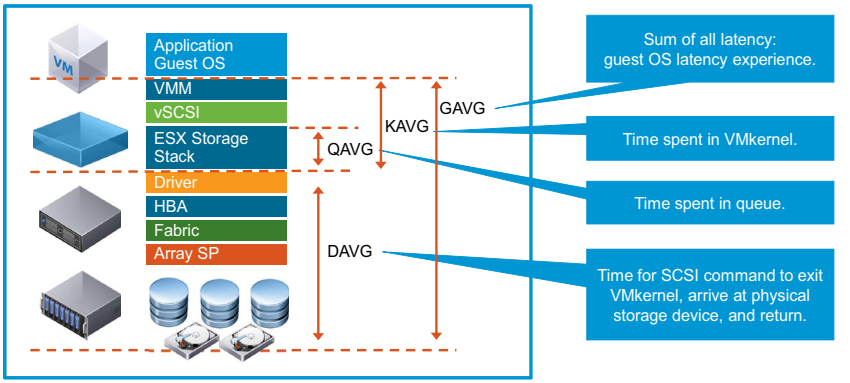
In this example, the GID for the Linux01 VM (57941) is entered in the `esxtop` window. `esxtop` displays disk throughput statistics for each VMDK connected to the VM.

https://t.me/learningnets

7-25 Disk Latency Metrics

Disk latency is the time (measured in milliseconds) that a SCSI command spends in transit, from the source to the destination and back.

The following disk latency values can be monitored with `esxtop`.



Latency statistics are measured at the different layers of the ESXi storage stack.

GAVG is the round-trip latency that the guest operating system sees for all I/O requests sent to the virtual storage device.

KAVG tracks the latencies due to the ESXi VMkernel commands. The KAVG value should be very small in comparison to the DAVG value and it should be close to zero. When a lot of queuing occurs in ESXi, KAVG can be as high as or higher than DAVG. If this situation occurs, check the queue statistics.

DAVG is the latency seen at the device driver level. It includes the round-trip time between the HBA and the storage. DAVG is a good indicator of performance of the back-end storage. If I/O latencies are suspected of causing performance problems, DAVG should be examined.

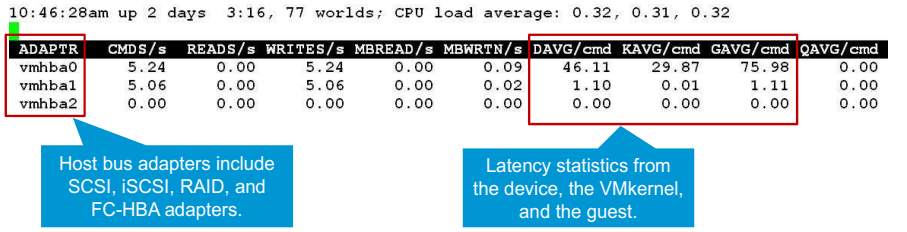
Compare I/O latencies with corresponding data from the storage array. If they are close, check the array for misconfiguration or faults. If they are not close, compare DAVG with corresponding data from points in between the array and the ESXi host, for example, Fibre Channel switches. If this intermediate data also matches DAVG values, the storage is likely underconfigured for the application. Adding disk spindles or changing the RAID level might help in such cases.

QAVG is the average queue latency. QAVG is part of KAVG. Response time is the sum of the time spent in queues in the storage stack and the service time spent by each resource in servicing the request. The largest component of the service time is the time spent retrieving data from physical storage. If QAVG is high, consider examining the queue depths at each level in the storage stack.

https://t.me/learningnets

7-26 Monitoring Disk Latency with esxtop

In addition to disk throughput, the `esxtop` disk adapter screen (enter **d** in the window) enables you to monitor disk latency.



DAVG/cmd: Average latency (ms) of the device (LUN).

KAVG/cmd: Average latency (ms) in the VMkernel, also called queuing time.

GAVG/cmd: Average latency (ms) in the guest. GAVG = DAVG + KAVG.

ADAPTR is the name of the host bus adapter (vmhba#), which includes SCSI, iSCSI, RAID, Fibre Channel, and FCoE adapters.

DAVG/cmd is the average amount of time it takes a device, which includes the HBA, the storage array, and everything in between, to service a single I/O request (read or write).

A value of less than 10 indicates that the system is healthy. A value in the range of 11 through 20 indicates that you must monitor the value more frequently. A value greater than 20 indicates a problem.

KAVG/cmd is the average amount of time it takes the VMkernel to service a disk operation. This number represents the time taken by the CPU to manage I/O. Because processors are much faster than disks, this value should be close to zero. A value of 1 or 2 is considered high for this metric.

GAVG/cmd is the total latency seen from the virtual machine when performing an I/O request. GAVG is the sum of DAVG plus KAVG.

https://t.me/learningnets

7-27 Monitoring Commands and Command Queuing

Metrics are available for monitoring the number of active disk commands and the number of disk commands that are queued.

Performance Metric	Name in esxtop
Number of active commands: I/O operations currently active	ACTV
Number of commands queued: I/O operations that require processing	QUED

The metrics in the table provide information about your disk performance. They are often used to further interpret the latency values that you might be observing:

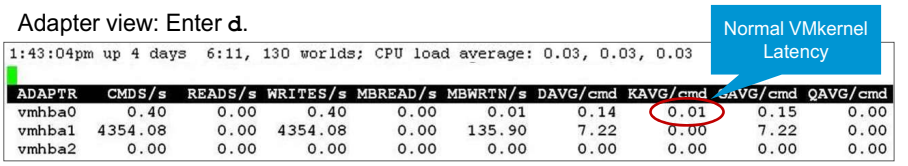
- Number of active commands: This metric represents the number of I/O operations that are currently active. This number includes operations for which the host is processing. This metric can serve as a quick view of storage activity. If the value of this metric is close to or at zero, the storage subsystem is not being used. If the value is a nonzero number, sustained over time, then constant interaction with the storage subsystem is occurring.
- Number of commands queued: This metric represents the number of I/O operations that require processing but have not yet been addressed. Commands are queued and awaiting management by the kernel when the driver's active command buffer is full. Occasionally, a queue forms and results in a small, nonzero value for QUED. However, any significant (double-digit) average of queued commands means that the storage hardware is unable to keep up with the host's needs.

https://t.me/learningnets

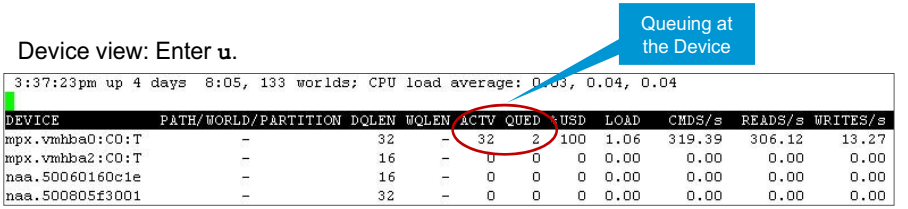
7-28 Example: Disk Latency and Queuing

You use `esxtop` disk latency readings to indicate device queueing.

Adapter view: Enter **d**.



Device view: Enter **u**.



This is an example of monitoring the kernel latency value, KAVG/cmd. This value is being monitored for the vmhba0 device. In the first `esxtop` screen (enter **d** in the window), the kernel latency value is 0.01 milliseconds. This value is good because it is nearly zero.

In the second `esxtop` screen (enter **u** in the window), 32 active I/Os (ACTV) and 2 I/Os are being queued (QUED). Some queuing is happening at the VMkernel level.

Queuing happens if I/O to the device is excessive and the LUN queue depth setting is insufficient. The default LUN queue depth is 32. The default value depends on the HBA. QLogic HBAs have a default queue depth of 64 in vSphere 5.x and 6.x.

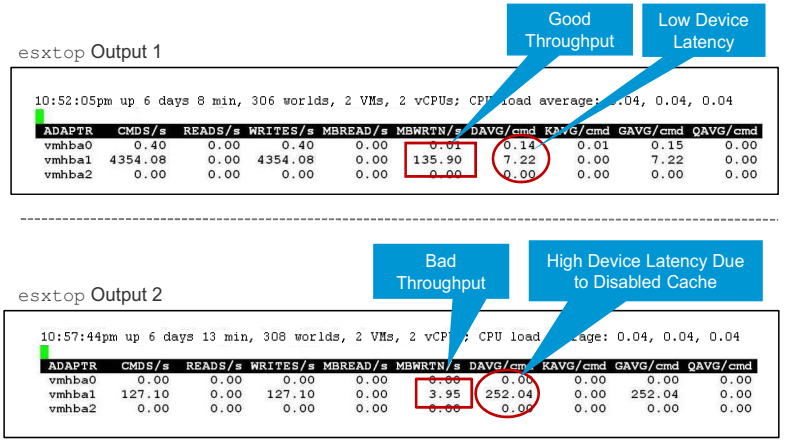
If too many I/Os (that is, more than 32) exist simultaneously, the device gets bottlenecked to only 32 outstanding I/Os at a time. To resolve this problem, you might change the queue depth of the device driver.

For information about changing the queue depth of the device driver, see *vSphere Monitoring and Performance* at https://docs.vmware.com/en/VMware-vSphere/6.0/vsphere-esxi-vcenter-server-601-monitoring-performance-guide.pdf.

https://t.me/learningnets

7-29 Example: Bad Disk Throughput

View latency values to determine if a performance problem truly exists.



In the example, the first `esxtop` screen (enter **d** in the window) shows the disk activity of a host. Looking at the vmhba1 adapter, you see that very good throughput exists: 135.90 MB written per second. The device latency per command is 7.22, which is low.

The second `esxtop` screen, however, shows low throughput (3.95 MB written per second) on vmhba1 but very high device latency (252.04). In this configuration, the SAN cache was disabled. The disk cache is very important, especially for writes.

The use of caches applies not only to SAN storage arrays but also to RAID controllers. A RAID controller typically has a cache with a battery backup. If the battery dies, the cache is disabled to prevent any potential loss of data. Therefore, a problem might occur if you are experiencing very good performance and the performance suddenly degrades. In this case, check the battery status on the RAID controller to see if the battery is still good.

https://t.me/learningnets

7-30 Activity: Hardware Disk Latency (1)

What `esxtop` fields reflect the physical storage device's latency? Select all that apply.

- ❑ GAVG
- ❑ KAVG
- ❑ QAVG
- ❑ DAVG

https://t.me/learningnets

7-31 Activity: Hardware Disk Latency (2)

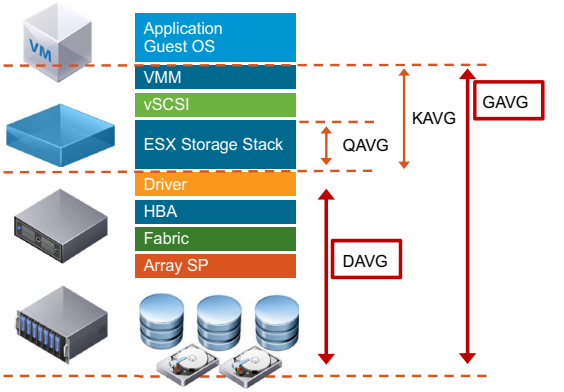
The following `esxtop` fields reflect a physical storage device's latency:

- ✓ GAVG
- ❑ KAVG
- ❑ QAVG
- ✓ DAVG

GAVG = KAVG + DAVG

DAVG is the latency seen at the device driver level.

The Array SP is directly affected by the latency of a physical storage device.



https://t.me/learningnets

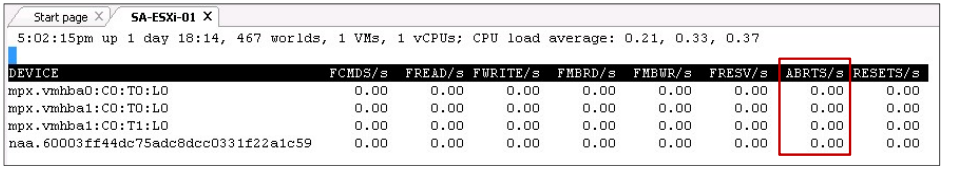
7-32 Overloaded Storage and Command Aborts

If a storage device is experiencing command aborts, the cause of these aborts must be identified and corrected.

To monitor command aborts in `esxtop`:

1. Enter **`u`** to display disks (LUNs).
2. Enter **`f`** and **`L`** to add error statistics to the display.

If ABRTS/s is more than 0 for any LUN, then storage is overloaded on that LUN.



Severely overloaded storage can be the result of several problems in the underlying storage layout or infrastructure. In turn, overloaded storage can manifest itself in many ways, depending on the applications running on the virtual machines. When storage is severely overloaded, operation time-outs can cause commands that are already issued to disk to be terminated (or aborted).

https://t.me/learningnets

7-33 Unexpected Increase in I/O Latency on Shared Storage

When consolidated, applications can share expensive physical resources such as storage.

Situations might occur when high I/O activity on shared storage affects the performance of latency-sensitive applications:

- Very high number of I/O requests are issued concurrently.
- Operations, such as a backup operation in a virtual machine, use the I/O bandwidth of a shared storage device.

To resolve these situations, use Storage I/O Control to control each virtual machine's access to I/O resources of a shared datastore.

Sharing storage resources offers such advantages as ease of management, power savings, and higher resource use. However, situations might occur when high I/O activity on the shared storage might affect the performance of certain latency-sensitive applications:

- A higher-than-expected number of I/O-intensive applications in virtual machines sharing a storage device become active at the same time. As a result, these virtual machines issue a very high number of I/O requests concurrently. Increased I/O requests increases the load on the storage device, and the response time of I/O requests also increases. The performance of virtual machines running critical workloads might be affected because these workloads now must compete for shared storage resources.
- Operations such as vSphere Storage vMotion migration or a backup operation running in a virtual machine use the I/O bandwidth of a shared storage device. These operations cause other virtual machines sharing the device to suffer from resource starvation.

Storage I/O Control can be used to overcome these problems. It continuously monitors the aggregate normalized I/O latency across the shared storage devices (VMFS and NFS datastores). Storage I/O Control can detect the existence of an I/O congestion based on a preset congestion threshold. After the condition is detected, Storage I/O Control uses the resource controls set on the virtual machines to control each virtual machine's access to the shared datastore.

In addition, vSphere DRS provides automatic or manual datastore cluster I/O capacity and performance balancing. vSphere Storage DRS can be configured to perform automatic vSphere Storage vMotion migrations of virtual machines. These migrations can occur when either space use or I/O response time thresholds of a datastore have been exceeded.

https://t.me/learningnets

7-34 Storage Performance Best Practices

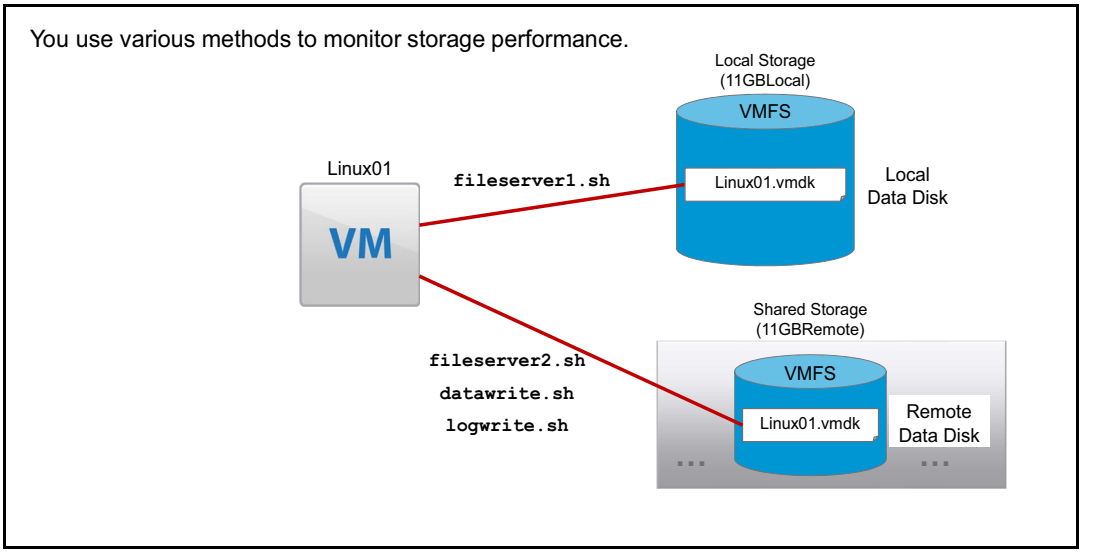
To get the best performance from storage, follow these best practices:

- Configure each LUN with the storage characteristics for applications and virtual machines that use the LUN.
- Avoid oversubscribing paths (SAN) and links (iSCSI and NFS).
- Use Storage DRS and Storage I/O Control whenever applicable.
- Isolate iSCSI and NFS traffic.
- Applications that write a lot of data to storage should not share Ethernet links to a storage device.
- Postpone major storage maintenance until off-peak hours.
- Eliminate all possible swapping to reduce the burden on the storage subsystem.
- In SAN configurations, spread I/O loads over the available paths to the storage devices.
- Strive for complementary workloads.

Balance best practices with the objectives of the organization.

https://t.me/learningnets

7-35 Introduction to Lab 10: Monitoring Storage Performance



You generate various types of disk I/O and compare performance. Your virtual machine is configured with a system disk and two data disks. The system disk and one of the data disks are on the local datastore. The other data disk is on the remote datastore.

You run the following scripts:

- `fileserver1.sh`: This script generates random reads to the local data disk.
- `fileserver2.sh`: This script generates random reads to the remote data disk.
- `datawrite.sh`: This script generates random writes to the remote data disk.
- `logwrite.sh`: This script generates sequential writes to the remote data disk.

Each script starts a program called `aio-stress`, which is a simple command-line program that measures the performance of a disk subsystem. For more information about the `aio-stress` command, see the `readme.txt` file on the test virtual machine in the same directory as the scripts.

https://t.me/learningnets

7-36 Lab 10: Monitoring Storage Performance

Use the `esxtop` command to monitor disk performance across a series of tests

1. Prepare to Run Tests
2. Measure Continuous Sequential Write Activity to a Virtual Disk on a Remote Datastore
3. Measure Continuous Random Write Activity to a Virtual Disk on a Remote Datastore
4. Measure Continuous Random Read Activity to a Virtual Disk on a Remote Datastore
5. Measure Continuous Random Read Activity to a Virtual Disk on a Local Datastore
6. Analyze the Test Results

https://t.me/learningnets

7-37 Review of Lab 10: Monitoring Storage Performance

Storage Counters for Linux01	Sequential Writes / Remote Datastore (Task 2)	Random Writes / Remote Datastore (Task 3)	Random Reads / Remote Datastore (Task 4)	Random Reads / Local Datastore (Task 5)
READS/s				
WRITES/s				

https://t.me/learningnets

7-38 Review of Learner Objectives

You should be able to meet the following objectives:

- Determine which disk metrics to monitor
- Identify metrics in `esxtop`
- Demonstrate how to monitor disk throughput

https://t.me/learningnets

7-39 Key Points

- Storage protocols, storage configuration, RAID levels, queuing, and VMFS configuration are factors that affect storage performance.
- Disk throughput and latency are key metrics when monitoring storage performance.
- If a storage device is experiencing command aborts, the cause of these aborts must be identified and corrected.

Questions?

https://t.me/learningnets

https://t.me/learningnets

MODULE 8

Network Optimization



https://t.me/learningnets

8-2 You Are Here

1. Course Introduction
2. Network Scalability
3. Storage Scalability
4. Host and Management Scalability
5. CPU Optimization
6. Memory Optimization
7. Storage Optimization
8. **Network Optimization**
9. vCenter Server Performance Optimization
10. vSphere Security

https://t.me/learningnets

8-3 Importance

Network performance can be measured in terms of how many packets were dropped when transmitting or receiving data. You should know how to monitor packet drops.

https://t.me/learningnets

8-4 Module Lessons

Lesson 1:	Networking Virtualization Concepts
Lesson 2:	Monitoring Network I/O Activity

https://t.me/learningnets



https://t.me/learningnets

8-6 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Describe network virtualization overhead
- Describe network adapter features that affect performance
- Describe vSphere networking features that affect performance

https://t.me/learningnets

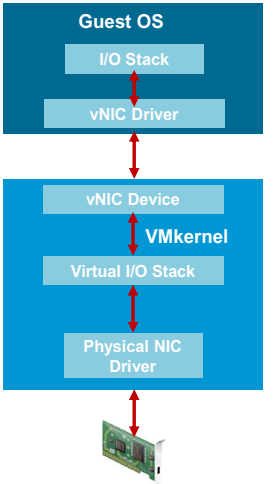
8-7 Network I/O Virtualization Overhead

Packets sent to or received from a virtual machine experience some overhead when compared to a bare-metal (native) environment.

Network I/O virtualization overhead can come from several sources:

- Emulation overhead
- Packet processing
- Scheduling
- Virtual interrupt coalescing
- Halting and waking up the physical CPU
- Halting and waking up the virtual CPU

Network I/O latency can increase due to this virtualization overhead.



The overhead that virtual machines experience is due to packets traversing an extra layer of virtualization stack. The network I/O latency due to the virtualization stack can come from several sources, including the following most common sources:

- Emulation overhead: Certain privileged instructions and some operations to access I/O devices that are executed by the virtual machine are intercepted by the hypervisor. This activity adds some overhead and contributes to network I/O latency.
- Packet processing: The network virtualization stack forwards a network packet from the physical NIC to the virtual machine, and the reverse. This activity requires some computation and processing, such as switching decisions at the virtual switch, inserting and stripping the VLAN tag, and copying packets if necessary. This processing adds some latency on both the transmit and the receive paths.
- Scheduling: Packet transmission and reception involves multiple hypervisor threads and virtual CPUs. The VMkernel scheduler has to schedule and then execute these threads, if they are not already running, on receipt and transmission of a packet. On an idle system, this activity takes a couple of microseconds. On a busy system, if high CPU contention exists, this activity can take tens of microseconds and occasionally milliseconds.
- Virtual interrupt coalescing: Similar to physical NIC interrupt coalescing, the virtual machine does virtual interrupt coalescing. That is, a virtual machine might not be interrupted immediately after receiving or transmitting a packet. Instead, the VMkernel might wait to receive or transmit more than one packet before an interrupt is posted. Also, on the transmit path, the virtual machine might wait until a few packets are queued up before sending the packets down to the hypervisor. Sometimes interrupt coalescing might have a noticeable effect on average latency.

https://t.me/learningnets

- Halting and waking up the physical CPU: Modern hardware and operating systems put the physical CPU in sleep mode (halt) when the CPU is idle. Waking up the CPU from halt mode adds some latency. The amount of latency varies from CPU to CPU. Depending on the load of the system, the CPU might go into deeper sleep states from which it takes even longer to wake up. This overhead is generally common between nonvirtualized and virtualized platforms. Sometimes the side effects of this can be worse in virtualized environments than in physical environments. For example, a CPU might flush its cache on entering halt mode. In this case, the halt operation might affect a virtualized system more than a physical system because the cache footprint is larger in a virtual environment.
- Halting and waking up the virtual CPU: When the guest operating system issues a halt instruction, the VMkernel scheduler might deschedule the virtual CPU. Rescheduling the virtual CPU, and putting the virtual CPU back in the run queue, can take a couple of microseconds.

  Latency introduced due to the virtual or physical CPU halt function is more prominent when the CPU is lightly loaded, that is, when the CPU halts frequently.

https://t.me/learningnets

8-8 VMXNET Network Adapters

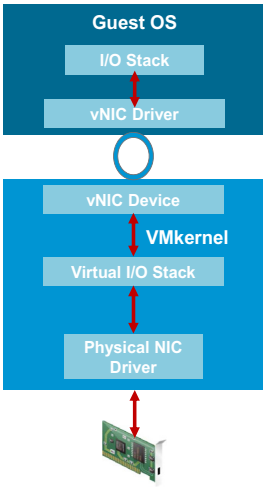
VMXNET is the VMware paravirtualized device driver for virtual networking.

The VMXNET driver has the following performance optimizations:

- Shares a ring buffer between the virtual machine and the VMkernel
- Supports transmission packet coalescing
- Supports interrupt coalescing
- Offloads TCP checksum calculation to the hardware

VMXNET comes in the following versions:

- VMXNET
- VMXNET2 (Enhanced VMXNET)
- VMXNET3



The VMXNET driver implements an idealized network interface that passes through network traffic from the virtual machine to the physical cards with minimal overhead.

The driver improves performance through several optimizations:

- It shares a ring buffer between the virtual machine and the VMkernel. And it uses zero-copy, which in turn saves CPU cycles. Traditional networking uses a series of buffers to process incoming network data and deliver it efficiently to users. However, higher-speed modern networks are turning this approach into a performance bottleneck as the amount of data received from the network often exceeds the size of the kernel buffers. Zero-copy improves performance by having the virtual machines and the VMkernel share a buffer, reducing the internal copy operations between buffers to free up CPU cycles.
- It uses transmission packet coalescing to reduce address space switching.
- It batches packets and issues a single interrupt, rather than issuing multiple interrupts.
- It offloads TCP checksum calculation to the network hardware instead of using the CPU resources of the virtual machine monitor.

https://t.me/learningnets

8-9 Virtual Network Adapters

When you configure a virtual machine, you can add vNICs and specify the adapter type.

ESXi supports several virtual network adapters:

- Vlance
- VMXNET
- E1000 and E1000E
- VMXNET2 (Enhanced VMXNET)
- VMXNET3
- PVRDMA
- SR-IOV passthrough

You can use the Flexible adapter type to configure Vlance and VMXNET.

The following virtual NIC (vNIC) types are supported:

- Vlance: This vNIC is an emulated version of the AMD 79C970 PCnet32 LANCE NIC. Vlance is an older 10 Mbps NIC, with drivers available in most 32-bit guest operating systems. This type cannot be selected directly. You can configure it only by using the Flexible adapter type.
- VMXNET: This vNIC has no physical counterpart. VMXNET is optimized for performance in a virtual machine. This type cannot be selected directly. You can configure it only by using the Flexible adapter type.
- Flexible: This vNIC identifies itself as a Vlance adapter when a virtual machine boots, but it initializes itself and functions as either a Vlance or a VMXNET adapter.
- E1000 and E1000E: The E1000 NIC is an emulated version of the Intel 82545EM Gigabit Ethernet NIC. The E1000E is an emulated version of the Intel 82574L Gigabit Ethernet NIC.
- Enhanced VMXNET (VMXNET2): The Enhanced VMXNET adapter is based on the VMXNET adapter, but it provides some high-performance features, such as jumbo frames and hardware offloads.
- VMXNET3: The VMXNET3 adapter is a paravirtualized NIC designed for performance. It is not related to VMXNET or VMXNET2. The VMXNET3 adapter offers all the features available in VMXNET2 and adds several features, like multiqueue support (called receive-side scaling in Windows), IPv6 offloads, and MSI/MSI-X interrupt delivery.
- PVRDMA: The PVRDMA adapter is a paravirtualized NIC that supports remote direct memory access between virtual machines through the OFED verbs API.

https://t.me/learningnets

- SR-IOV passthrough: SR-IOV passthrough is a representation of a virtual function on a physical NIC with an SR-IOV support.

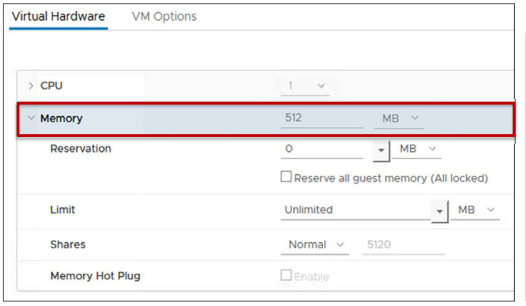
For more information about network adapter types, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

https://t.me/learningnets

8-10 About VMXNET3 Optimization

The VMXNET3 driver is optimized to remove unnecessary metadata from smaller packets.

Optimized usage of pinned (reserved) memory greatly increases the VM's ability to move packets through the OS more efficiently.



VM Virtual Hardware Properties

```
sched.cpu.shares = "normal"
memsize = "512"
sched.mem.min = "512"
sched.mem.minSize = "0"
sched.mem.prealloc.pinnedMainMem = "TRUE"
sched.mem.prealloc = "TRUE"
sched.swap.vmxSwapEnabled = "FALSE"
sched.mem.shares = "normal"
sched.mem.affinity = "all"
powerType.powerOff = "default"
powerType.suspend = "default"
powerType.reset = "default"
toolScripts.afterPowerOn = "TRUE"
```

VM .vmx File

In order to assist the virtual machine to process packets internally as quickly as possible, the VMXNET3 driver is optimized to remove unnecessary metadata from smaller packets (using the Copy Tx function) for small message size ≤ 256 bytes), improving processing time by up to 10 percent.

The use of pinned memory can greatly increase the virtual machine's ability to move packets through the operating system more efficiently. Normally, virtual machine memory addresses must be translated to physical memory addresses, which adds an overhead to the memory access time. Also, if memory is not reserved, some memory pages might be put in swap or moved around to accommodate the demands of other virtual machines. By reserving and pinning the entire virtual machine memory, you address translation as the virtual machine's pages are permanently mapped to specific physical addresses. Thus, the address translation never needs repeating over and over again. Memory can be pinned by configuring VMX options or, more simply, by setting the latency sensitivity of the virtual machine to high.

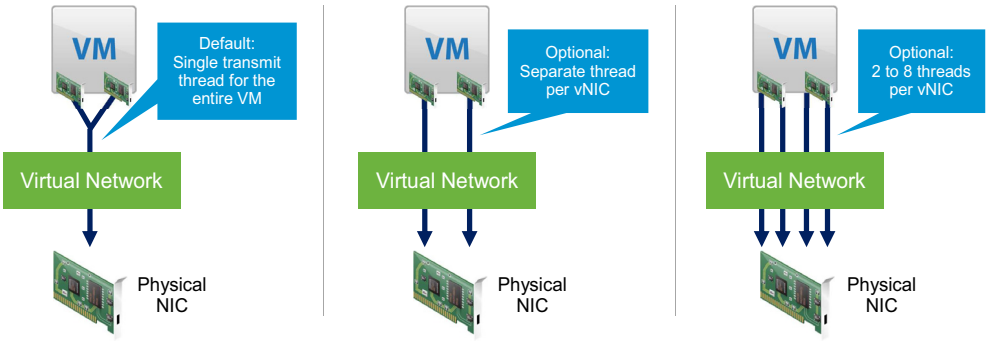
As shown in the screenshots, you add lines in the virtual machine's VMX file to pin all the virtual machine memory. These options ensure that the virtual machine can never use swap, which requires that the virtual machine memory always be backed by real physical memory. All the memory is configured to be preallocated and pinned, which results in the appearance that all the virtual machine memory is permanently active.

https://t.me/learningnets

8-11 Configuring a Single CPU Thread per vNIC

By default, a single VMkernel networking thread is used to process all transmit traffic coming from a VM, regardless of how many virtual NICs (vNICs) the VM has.

To improve a VM's transmit rate when multiple vNICs are available, you can edit the VMX file to allow one or more VMkernel networking threads per vNIC.



CPU scheduling is optimized for network packet processing in the ESXi.

Typically, when virtual machine packets are being processed, a single CPU thread is used to process all transmit traffic coming from a virtual machine, regardless of how many vNICs the virtual machine has. If multiple vNICs are configured for a virtual machine, it is advantageous to allow one CPU thread per vNIC to help speed up the processing time. This arrangement helps boost transmit rates for the virtual machine when more than one vNIC is configured.

You can edit the VMX file to allow one CPU thread per device. Assume, for example, that the VMX file is named `VM1.vmx` and that the virtual machine has two NICs: Ethernet 0 and Ethernet 1. You must add the line `ethernetX.ctxPerDev = "1"` for each device:

```
ethernet0.ctxPerDev = "1"

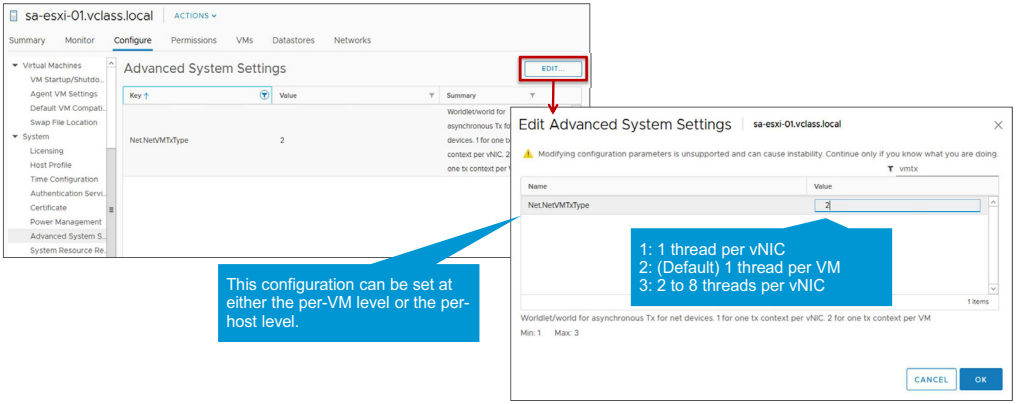
ethernet1.ctxPerDev = "1"
```


https://t.me/learningnets

8-12 Configuring Multiple CPU Threads per vNIC

For VMs that require a high transmit rate and multiple simultaneous TCP streams, you can enable one or more CPU threads per vNIC.

Up to eight separate threads can be created per vNIC.



For VMs that require a high transmit rate and multiple simultaneous TCP streams, multiple CPU threads can be used per vNIC. Up to eight threads can be used by the vNIC, depending on the number of streams.

To enable this feature per VM, you add the line `EthernetX.ctxPerDev="3"` to the VMX file. The value 1 allows a single thread per vNIC. The value 3 allows multiple threads per NIC, up to a maximum of eight threads, created by the host. Using the value 3 does not signify that three threads will be created. It only allows the host to create as many threads as it needs, up to a maximum of eight. To enable this feature per VM, you add `ethernetX.ctxPerDev="3"` to the VMX file.

To set this feature at the host level, change the `Net.NetVMTxType` value to 3 in the ESXi host's advanced settings in vSphere Client.

Acceptable values for `Net.NetVMTxType` are 1, 2, or 3:

- 1: 1 transmit context (CPU thread) per vNIC
- 2 (Default): 1 transmit context (CPU thread) per VM
- 3: 2 through 8 transmit contexts (CPU threads) per vNIC

Keep in mind that many of the CPU scheduling improvements require available CPU threads on a host, so if your CPU is already overprovisioned, implementing some of these features might make matters worse. Adding CPU threads to process traffic flows requires that your CPU be underprovisioned to ensure that network processing does not encounter contention on the CPUs that it is trying to use.

https://t.me/learningnets

8-13 Virtual Switch and Physical NIC Optimization

The virtual switch code is optimized for faster switching using a specific caching technique. This caching technique is designed to keep session information cached. As a result, the switching process takes place faster.

vSphere includes native driver support for Intel NIC cards which removes the overhead of translating from VMkernel to VMKLinux data structures.

Currently-supported native NIC drivers for physical NICs include Mellanox, Emulex, and Intel.

On the physical NIC side, vSphere has broadened the range of native drivers to include Intel cards, in addition to the existing Mellanox and Emulex native drivers. By using native drivers, you can more efficiently structure the data to be sent using the relevant cards and remove the overhead when trying to translate from the VMkernel to the VMKLinux data structures.

https://t.me/learningnets

8-14 About TCP Segmentation Offload

TCP segmentation offload (TSO) improves networking performance by reducing the CPU overhead that is involved with sending large amounts of TCP traffic.

TSO performs the following functions:

- Large TCP packets are offloaded to the adapter for further segmentation.
- The adapter divides the packet into MTU-sized frames.

If the NIC hardware supports this feature, TSO is enabled on VMkernel interfaces by default.

TSO can be manually enabled at the virtual machine level.

TCP segmentation offload (TSO) improves performance for TCP network traffic coming from a virtual machine and for network traffic, such as vSphere vMotion traffic, sent out from the server.

TSO is used in the guest when the VMXNET2 (or later) network adapter is installed. To enable TSO at the virtual machine level, you must replace the existing VMXNET or Flexible virtual network adapter with a VMXNET2 (or later) adapter. This replacement might result in a change in the MAC address of the virtual network adapter.

If TSO becomes disabled for a particular VMkernel interface, the only way to enable TSO is to delete that VMkernel interface and recreate it with TSO enabled.

When the physical NICs provide TSO functionality, the ESXi host can use the specialized NIC hardware to improve performance. However, performance improvements related to TSO do not require NIC hardware support for TSO.

Virtual machines that use TSO on an ESXi host show lower CPU utilization than virtual machines that lack TSO support, when performing the same network activities.

For information about enabling TSO support for a virtual machine or checking whether TSO is enabled on a VMkernel interface, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

For more information about TSO, see VMware knowledge base article 2055140 at http://kb.vmware.com/kb/2055140 and VMware knowledge base article 1009548 at http://kb.vmware.com/kb/1009548.

https://t.me/learningnets

8-15 About Jumbo Frames

Before transmitting packets, the IP layer fragments data into MTU-sized frames:

- The default Ethernet MTU is 1,500 bytes.
- The receive side reassembles the data.

A jumbo frame is an Ethernet frame with a bigger MTU of up to 9,000 bytes:

- It reduces the number of frames transmitted.
- It reduces the CPU utilization on the transmit and receive sides.

Virtual machines must be configured with E1000, E1000E, VMXNET2, or VMXNET3 adapters.

The network must support jumbo frames end to end.

For each packet, the system has to perform a nontrivial amount of work to package and transmit the packet. As Ethernet speed increases, so does the amount of work necessary which results in a greater burden on the system.

Jumbo frames decrease the number of packets requiring packaging, compared to previously-sized packets. That decrease results in less work for network transactions, which frees up resources for other activities.

The physical NICs at both ends, as well as all the intermediate hops, routers, and switches, must support jumbo frames. Jumbo frames must be enabled at the virtual switch level, at the virtual machine, and at the VMkernel interface. ESXi hosts using jumbo frames realize a decrease in load due to network processing. Before enabling jumbo frames, check with your hardware vendor to ensure that your network adapter supports jumbo frames.

For information about enabling jumbo frames, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-networking-guide.pdf.

https://t.me/learningnets

8-16 About SplitRx Mode

SplitRx mode uses multiple physical CPUs to process network packets received in a single network queue.

Enable SplitRx mode in situations where you have multiple virtual machines on an ESXi host receiving multicast traffic from the same source.

SplitRx mode is supported only for VMXNET3 network adapters.

SplitRx mode is enabled when ESXi detects that a single network queue on a physical NIC meets the following conditions:

- The NIC is heavily used.
- The NIC is receiving more than 10,000 broadcast or multicast packets per second.

Multicasting is an efficient way of disseminating information and communicating over the network. A single sender can connect to multiple receivers and exchange information while conserving network bandwidth. Financial stock exchanges, multimedia content delivery networks, and commercial enterprises often use multicasting as a communication mechanism. Multiple receivers can be enabled on a single ESXi host. Because the receivers are on the same host, the physical network does not have to transfer multiple copies of the same packet. Packet replication is carried out in the hypervisor instead.

SplitRx mode is an ESXi feature that provides a scalable and efficient platform for multicast receivers. SplitRx mode typically improves throughput and CPU efficiency for multicast traffic workloads.

SplitRx mode is enabled only if the network traffic is arriving on a physical NIC, not when the traffic is entirely internal to the ESXi host. If the traffic is entirely internal, SplitRx mode can be manually enabled.

SplitRx mode is individually configured for each vNIC. For information about how to enable this feature, see *Performance Best Practices for VMware vSphere 6.0* at http://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/vmware-perfbest-practices-vsphere6-0-white-paper.pdf.

https://t.me/learningnets

8-17 About SplitTx Mode

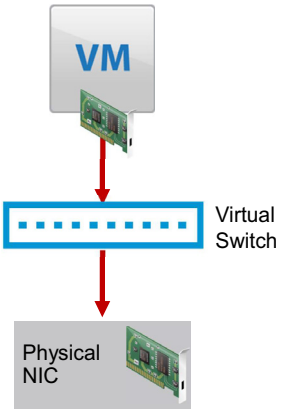
SplitTx mode allows two separate threads to be created for a single flow to parallelize vNIC and physical NIC processing:

- vNIC-to-virtual switch processing
- Virtual switch to physical NIC processing

SplitTx mode must be enabled per host with the following host command:

```
vsish -e set /net/pNics/vmnicX/sched/txMode 1
```

Using two threads significantly speeds up the rate at which the traffic passes through the kernel layers.



https://t.me/learningnets

8-18 Activity: vSphere Networking Features (1)

What networking features help to reduce CPU overhead used for network packet processing on the ESXi host?

- ❑ SplitTx mode
- ❑ SplitRx mode
- ❑ Jumbo frames
- ❑ TCP segmentation offload
- ❑ Multiple CPU threads per vNIC

https://t.me/learningnets

8-19 Activity: vSphere Networking Features (2)

What networking features help to reduce CPU overhead used for network packet processing on the ESXi host?

- ❑ SplitTx mode
- ❑ SplitRx mode
- ✓ Jumbo frames
- ✓ TCP segmentation offload
- ❑ Multiple CPU threads per vNIC

https://t.me/learningnets

8-20 Review of Learner Objectives

You should be able to meet the following objectives:

- Describe network virtualization overhead
- Describe network adapter features that affect performance
- Describe vSphere networking features that affect performance

https://t.me/learningnets



https://t.me/learningnets

8-22 Learner Objectives

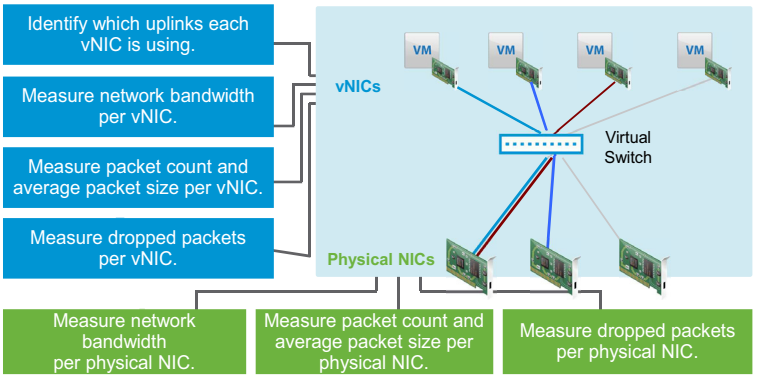
By the end of this lesson, you should be able to meet the following objectives:

- Determine which network metrics to monitor
- View metrics in `esxtop`
- Monitor key network performance metrics

https://t.me/learningnets

8-23 Monitoring for Performance in the Network I/O Stack

Several key points in the network I/O stack should be monitored for performance.



https://t.me/learningnets

8-24 Network Capacity Metrics

Network performance depends on application workload and network configuration.

Dropped network packets indicate a bottleneck in the network.

Identify network problems to determine available bandwidth and compare with expectations:

- Check key metrics.
- A vSphere environment has the following significant network statistics:
  - Network usage
  - Network packets received
  - Network packets transmitted
  - Received packets dropped
  - Transmitted packets dropped

To determine whether packets are being dropped, you can use the advanced performance charts in vSphere Client or use the `esxtop` command.

If received packets are being dropped, adjust the virtual machine CPU shares. If packets are not being dropped, check the size of the network packets, the data received rate, and the data transmitted rate. In general, the larger the network packets, the faster the network speed. When the packet size is large, fewer packets are transferred, which reduces the amount of CPU required to process the data. When network packets are small, more packets are transferred, but the network speed is slower because more CPU is required to process the data. In some instances, large packets can result in high latency. To rule out this problem, check network latency.

If packets are not being dropped and the data receive rate is slow, the host probably lacks the CPU resources required to handle the load. Check the number of virtual machines assigned to each physical NIC. If necessary, perform load balancing by moving virtual machines to different virtual switches or by adding more NICs to the host. You can also move virtual machines to another host or increase the CPU resources of the host or virtual machines.

https://t.me/learningnets

8-25 esxtop Networking Statistics

You can use `esxtop` to track network statistics per virtual machine or per NIC (virtual or physical).

Enter **n** to change to the networking view.

`esxtop` provides the following useful network counters:

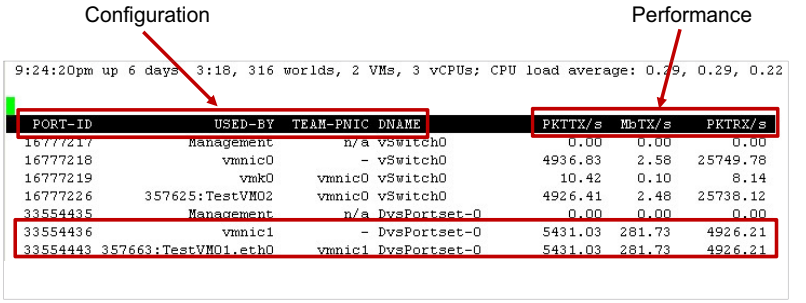
- MbTX/s: Amount of data transmitted (in megabits) per second
- MbRX/s: Amount of data received (in megabits) per second
- PKTTX/s: Average number of packets transmitted per second in the sampling interval
- PKTRX/s: Average number of packets received per second in the sampling interval
- %DRPTX: Percentage of outbound packets dropped in the sampling interval
- %DRPRX: Percentage of inbound packets dropped in the sampling interval

https://t.me/learningnets

8-26 esxtop Network Output

To display network statistics in `esxtop`, enter **n**.

By default, the command shows the columns PKTTX/s, MbTX/s, PKTRX/s, MbRX/s, %DRPTX, and %DRPRX. If you do not see all of the columns, adjust the width of the window.



In the `esxtop` window, configuration information about the objects is listed in the leftmost columns, followed by the performance metrics.

In this example, the TestVM01 virtual machine is connected to a distributed switch, and vmnic1 is connected to the same distributed switch.

The USED-BY column identifies network connections by physical adapter or vSphere network object. vmnic0 is an example of a physical adapter listing. A VMkernel port, such as vmk0, is an example of a vSphere network object. Another example of a vSphere network object is a virtual machine's NIC, identified as 357663:TestVM01.eth0. The value 357663 is an internal virtual machine ID, TestVM01 is the virtual machine name, and eth0 identifies the network interface.

https://t.me/learningnets

8-27 Dropped Network Packets

Packet data might need to be buffered before being passed to the next step in the delivery process.

Network packets are buffered in queues in the following cases:

- The destination is not ready to receive the packets.
- The network is too busy to send the packets.

The queues are finite in size:

- vNIC devices buffer packets when they cannot be handled immediately.
- If the queue in the vNIC fills, packets are buffered by the virtual switch port.

When these queues fill up, no more packets can be received which causes additional arriving packets to be dropped.

Network packets might get stored (buffered) in queues at multiple points along their route from the source to the destination. Network switches, physical NICs, device drivers, and network stacks contain queues where packet data or headers might get buffered.

TCP/IP networks use congestion-control algorithms that limit, but do not eliminate, dropped packets. When a packet is dropped, the TCP/IP recovery mechanisms work to maintain in-order delivery of packets to applications. However, these mechanisms operate at a cost to both networking performance and CPU overhead, a penalty that becomes more severe as the physical network speed increases.

vSphere presents vNICs, such as VMXNET or virtual E1000 devices, to the guest operating system running in a virtual machine. For received packets, the vNIC buffers packet data coming from a virtual switch until it is retrieved by the device driver running in the guest operating system. The virtual switch contains queues for packets sent to the vNIC.

If the guest operating system does not retrieve packets from the vNIC quickly enough, the queues in the vNIC device can fill up. This condition can in turn cause the queues in the corresponding virtual switch port to fill up. If a virtual switch port receives a packet bound for a virtual machine when its packet queue is full, the port must drop the packet.

https://t.me/learningnets

8-28 Dropped Packets in vNICs

If a guest operating system fails to retrieve packets quickly enough from the vNIC, received packets are dropped and a network throughput issue might exist.

Cause	Solution
High CPU utilization	Increase CPU resources provided to the virtual machine.
	Increase the efficiency with which the virtual machine uses CPU resources.
Improper guest operating system driver configuration	Tune the network stack in the guest operating system. Try using VMXNET3.
	Add vNICs to the virtual machine and spread the network load across them.

When the applications and the guest operating system are driving the virtual machine to high CPU use, extended delays might occur. These delays occur from the time the guest operating system receives notification that packets are available until those packets are retrieved from the vNIC. Sometimes, the high CPU use might be caused by high network traffic. The reason is because the processing of network packets can place a significant demand on CPU resources.

Device drivers for networking devices often have parameters that are tunable from within the guest operating system. These parameters control such behavior as whether to use interrupts or perform polling. Improper configuration of these parameters can cause poor network performance and dropped packets in the networking infrastructure.

When the virtual machine is dropping receive packets due to high CPU use, adding vCPUs might be necessary in order to provide sufficient CPU resources. If the high CPU use is due to network processing, ensure that the guest operating system can use multiple CPUs when processing network traffic. See the operating system documentation for the appropriate CPU requirements.

Applications that have high CPU use can often be tuned to improve their use of CPU resources. You might be able to tune the networking stack within the guest operating system to improve the speed and efficiency with which it handles network packets. See the documentation for the operating system.

In some guest operating systems, all of the interrupts for each NIC are directed to a single processor core. As a result, the single processor can become a bottleneck, leading to dropped receive packets. Adding more vNICs to these virtual machines enables the processing of network interrupts to be spread across multiple processor cores.

https://t.me/learningnets

8-29 Dropped Received Packets in Physical NICs

If an ESXi host fails to retrieve packets quickly enough from the physical NIC, received packets are dropped and a network throughput issue might exist.

Cause	Solution
High CPU utilization	Consider migrating virtual machines to other ESXi hosts to reduce CPU utilization.

https://t.me/learningnets

8-30 Dropped Transmit Packets in Physical NICs

Virtual machine traffic might exceed the physical capabilities of the uplink NICs or the networking infrastructure.

Virtual switch buffers might become full, causing additional transmitted packets arriving from the virtual machine to be dropped. Dropped packets are a network throughput issue.

Cause	Solution
Traffic from a virtual machine or a set of virtual machines sharing a virtual switch exceeds the physical capabilities of the uplink NICs or the networking infrastructure.	Add uplink capacity to the virtual switch.
	Migrate some virtual machines with high network demand to other ESXi hosts that have network capacity.
	Enhance the networking infrastructure.
	Reduce network traffic.

When a virtual machine transmits packets on a vNIC, those packets are buffered in the associated virtual switch port until they are transmitted on the physical uplink devices.

Adding more physical uplink NICs to the virtual switch might alleviate the conditions that are causing transmit packets to be dropped. However, traffic should be monitored to ensure that the NIC teaming policies selected for the virtual switch lead to proper load distribution over the available uplinks.

If the virtual switch uplinks are overloaded, moving some of the virtual machines to different virtual switches can help rebalance the load.

Sometimes the bottleneck might be in the networking infrastructure, for example, in the network switches or interswitch links. You might have to add capacity in the network to handle the load.

Reducing the network traffic generated by a virtual machine can help alleviate bottlenecks in the networking infrastructure. The implementation of this solution depends on the application and guest operating system being used. Techniques such as using caches for network data or tuning the network stack to use larger packets (for example, jumbo frames) might reduce the load on the network.

https://t.me/learningnets

8-31 Unexpected Increases in Data Transfer Rate

High-speed Ethernet solutions, such as 10/40 Gigabit Ethernet, support different traffic flows through a single physical link.

Each application can use the full bandwidth when no contention exists for the shared network link.

However, when traffic flows contend for the shared network bandwidth, then the performance of the applications might be affected.

Cause	Solution
Resource contention	Use Network I/O Control (shares, reservations, and limits) to distribute the network bandwidth among the different types of network traffic flows.
Few users dominating the resource usage	

Examples of traffic flows are flows from applications running in a virtual machine, vSphere vMotion, and vSphere Fault Tolerance. These traffic flows can coexist and share a single link.

The total demand from all the users of the network can exceed the total capacity of the network link. In this case, users of the network resources might experience an unexpected impact on performance. These instances, though infrequent, can still cause fluctuating performance behavior, which might be frustrating.

If triggered, vSphere vMotion or vSphere Storage vMotion migrations can consume extra network bandwidth when storage is IP/Ethernet-based. In this case, the performance of virtual machines that share network resources with these traffic flows can be affected. The performance impact might be more significant if the applications running in these virtual machines were latency-sensitive or business-critical.

To overcome the problems listed on the slide, use a resource-control mechanism such as Network I/O Control. Network I/O Control allows applications to freely use shared network resources when resources are underused. When the network is congested, Network I/O Control restricts the traffic flows of applications according to their shares, reservations, and limits.

https://t.me/learningnets

8-32 Networking Best Practices

VMware recommends these best practices for configuring your network:

- Use the VMXNET3 network adapter wherever possible. If using VMXNET3 is not possible, then use VMXNET or VMXNET2.
- Use a physical network adapter that supports high-performance features.
- Team NICs for load balancing.
- Use separate virtual switches, each connected to separate physical NICs, to avoid contention between the VMkernel and virtual machines. This guideline applies especially to virtual machines running heavy networking workloads.
- For networks that support different types of traffic flow, use Network I/O Control to allocate and control network bandwidth.
- Ensure that sufficient CPU resources exist for workloads that have high network throughput.

When you install VMware Tools, a VMXNET adapter replaces the default Vlance adapter. VMXNET adapters should be used for optimal performance in any guest operating system for which they are available.

For the best networking performance, use network adapters that support high-performance hardware features, such as TCP checksum offload, jumbo frames, and the ability to handle high-memory DMA.

NIC teams can provide passive failover in the event of hardware failure or network outage. In some configurations, NIC teaming can increase performance by distributing the traffic across those physical network adapters.

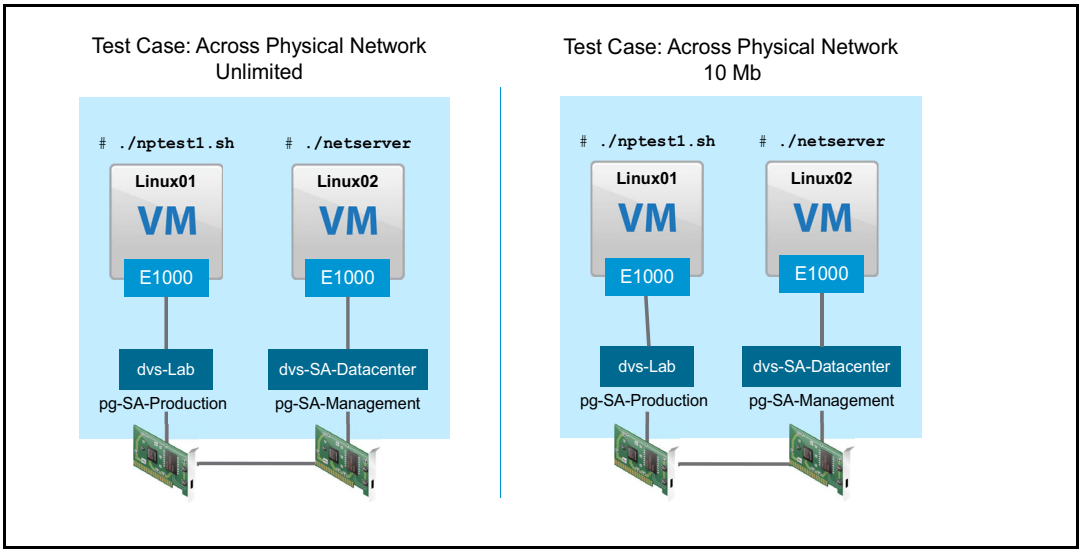
To avoid unnecessary CPU and network overhead when two virtual machines reside on the same ESXi host, connect both virtual machines to the same virtual switch. Virtual machines that communicate with each other on the same ESXi host can also use the VMware Machine Communication Interface device.

Network I/O Control can allocate user-defined proportions of bandwidth for specific needs and can prevent any one resource pool from affecting others.

In a native environment, CPU use plays a significant role in network throughput. To process higher levels of throughput, more CPU resources are needed. Because insufficient CPU resources limit maximum throughput, monitoring the CPU use of high-throughput workloads is essential.

https://t.me/learningnets

8-33 Introduction to Lab 11: Monitoring Network Performance (1)



In this lab, you generate network traffic and compare the network performance of different network adapters connected to different networks or the same network. Both test virtual machines are configured with E1000 network adapters.

You use a client-side script named `nptest1.sh` to generate a lot of network traffic. You are instructed to run this script on the network client virtual machine, Linux01.

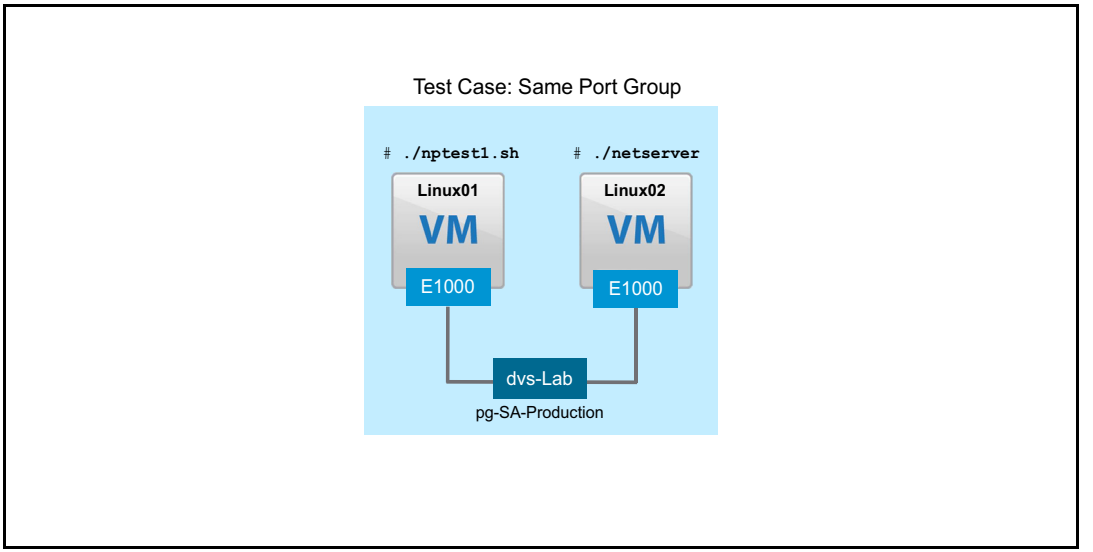
You also use a server-side script called `netserver`. This script runs on the network server virtual machine, Linux02. `netserver` receives the data sent by `nptest1.sh`.

You perform three test cases, two of which are described here:

- Test case 1: The first test virtual machine is connected to the pg-SA-Production port group on the distributed switch named dvs-Lab. The second test virtual machine is connected to the pg-SA-Management port group on the distributed switch named dvs-SA-Datacenter. Network traffic flows between two networks.
- Test case 2: This test case is similar to test case 1, except that the network speed is constrained.

https://t.me/learningnets

8-34 Introduction to Lab 11: Monitoring Network Performance (2)



For test case 3, both the test virtual machines are connected to the pg-SA-Production port group on the dvs-Lab distributed switch.

https://t.me/learningnets

8-35 Lab 11: Monitoring Network Performance

Use the `esxtop` command to monitor network performance

1. Prepare to Monitor Network Performance
2. Prepare the Client and the Server Virtual Machines
3. Measure Network Activity on an ESXi Physical Network Interface
4. Use Traffic Shaping to Simulate Network Congestion
5. Position the Client and the Server on the Same Port Group
6. Restart the Test and Measure Network Activity
7. Stop the Test and Analyze Results
8. Clean Up for the Next Lab

https://t.me/learningnets

8-36 Review of Lab 11: Monitoring Network Performance

Your instructor will review the lab with you.

Network Counter	vmnic2 (Task 3)	vmnic2 10 Mb/s (Task 4)	Linux01.eth0 (Task 6)
MbTX/s			
MbRX/s			

https://t.me/learningnets

8-37 Review of Learner Objectives

You should be able to meet the following objectives:

- Determine which network metrics to monitor
- View metrics in `esxtop`
- Monitor key network performance metrics

https://t.me/learningnets

8-38 Key Points

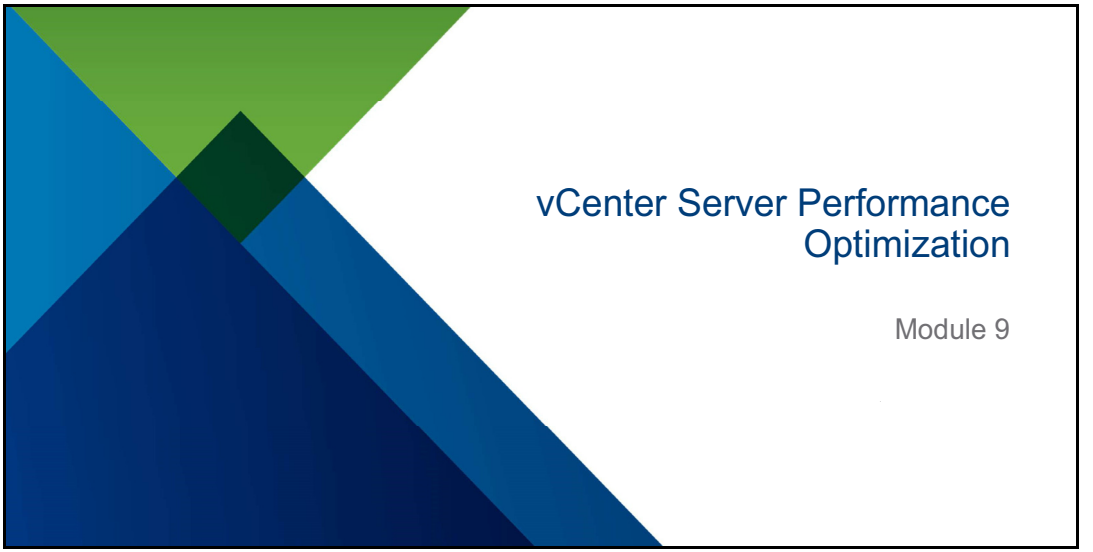
- vSphere uses many performance features of modern network adapters, such as TSO and jumbo frames.
- `esxtop` displays physical and virtual network data on the same screen.
- For network throughput problems, the best indicator is dropped packets.

Questions?

https://t.me/learningnets

MODULE 9

vCenter Server Performance Optimization



https://t.me/learningnets

9-2 You Are Here

1. Course Introduction
2. Network Scalability
3. Storage Scalability
4. Host and Management Scalability
5. CPU Optimization
6. Memory Optimization
7. Storage Optimization
8. Network Optimization
9. **vCenter Server Performance Optimization**
10. vSphere Security

https://t.me/learningnets

9-3 Importance

vCenter Server is a critical component in the vSphere environment. The data center relies on vCenter Server for healthy operation.

An administrator can make the data center run more efficiently and smoothly by using the latest improvements in performance and scalability in vCenter Server.

https://t.me/learningnets

9-4 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

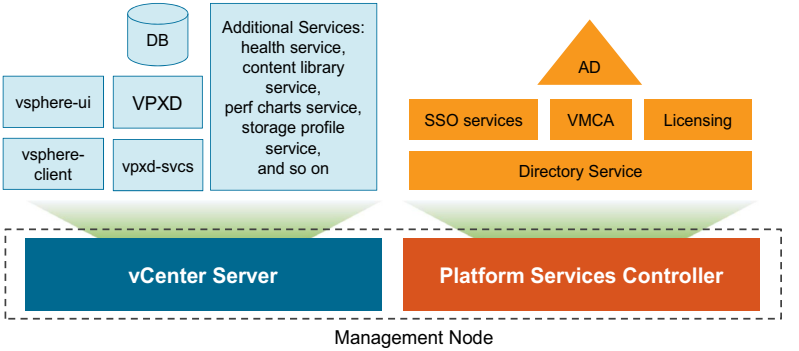
- Review vCenter Server components and services
- Describe the performance considerations for VMware Platform Services Controller™
- Describe the factors that influence vCenter Server performance
- Use VMware vCenter® Server Appliance™ tools to monitor resource usage

https://t.me/learningnets

9-5 vCenter Server Services

The vCenter Server management node consists of many services that are responsible for authenticating users and performing operations requested by users.

In order for vCenter Server to perform efficiently, you must ensure that enough resources are available for these services to run.



The slide shows only a subset of the services found on a management node. vCenter Server services include the following:

- vsphere-ui: Handles vSphere Client (HTML5) login requests to vCenter Server.
- vsphere-client: Handles vSphere Web Client login requests to vCenter Server.
- VPXD: Performs the main business logic of vCenter Server. Responsibilities include sending tasks to appropriate ESXi hosts, retrieving configuration changes from hosts, pushing configuration updates to the database, inserting statistics into the database, and satisfying queries from API clients.
- vpxd-svcs: Performs authorization for different types of inventory elements that vCenter Server is not authoritative on, such as content library-related authorization.
- health service: Monitors host and service health
- content library service: Manages content libraries.
- perf charts service: Manages the performance overview charts that you display with vSphere Client or vSphere Web Client.
- Storage profile service: Provides information to help in VM placement and compliance checks.

The following services are found on Platform Services Controller:

- SSO services: Consist of the security token service, administration server, and the identity management service.

https://t.me/learningnets

- Directory Service: A single sign-on service, the directory service contains the set of roles and privileges that each user is allowed to see. This service also stores licenses and tags.
- VMCA: Uses the VMware Endpoint Certificate Store to serve as a local repository for certificates.
- Licensing service: Provides centralized license management to vSphere and products that integrate with vSphere.

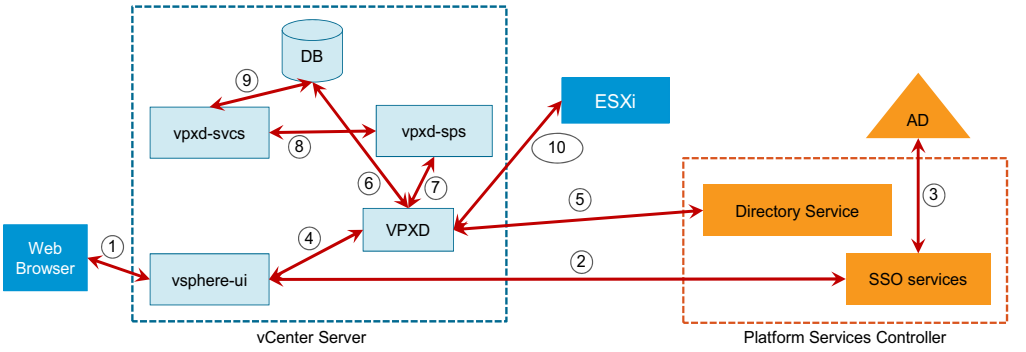
For more information and a list of all the Platform Services Controller services, see *Platform Services Controller Administration* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-platform-services-controller-administration-guide.pdf.

https://t.me/learningnets

9-6 Communication Between Management Services

The services work together and communicate with each other to accomplish all aspects of management, from authenticating user logins to performing tasks on ESXi hosts.

For example, when you power on a VM, services communicate to authenticate the user, send the power-on request to the ESXi host, record changes in the database, and respond to the user.



The services on the management node must communicate to perform the task requested by the user. For example, the user logs in to vCenter Server with vSphere Client and powers on a VM. The following communication must happen between the vCenter Server and Platform Services Controller services:

1. The web browser communicates with the vsphere-ui service.
2. vsphere-ui communicates with the SSO services to authenticate the user account.
3. Assuming Active Directory is used, the SSO service communicates with Active Directory to verify user identity. The SSO service also keeps track of what vCenter Server instance that the user has logged in to.
4. When the user powers on the VM from vSphere Client, the power on request is sent to VPXD.
5. VPXD communicates with the Directory Service to verify that the user is allowed to perform this operation.
6. VPXD might also need to access the vCenter Server database for information, such as DRS-related information.
7. VPXD also communicates with the storage policy service (vmware-sps) to check for compliance and VM storage placement information, for example on the Gold, Silver, or Bronze storage.
8. vmware-sps communicates with vpxd-svcs to access the storage profile associated with the VM.
9. vpxd-svcs accesses the database to retrieve storage profile information.

https://t.me/learningnets

10. The power on request is sent from VPXD to the ESXi host, and the command is executed.

    The ESXi host communicates with vCenter Server to notify vCenter Server that the operation is complete. Information is exchanged between the ESXi host and VPXD. VPXD persists this information in the database, then notifies the user that the operation is done.

https://t.me/learningnets

9-7 vCenter Server Performance Considerations

Consider the following factors when monitoring and maintaining acceptable levels of performance for vCenter Server and its services:

- Platform Services Controller performance
- Number of concurrent vCenter Server operations
- CPU and memory performance
- Network performance
- Database performance
- User interface performance

https://t.me/learningnets

9-8 Performance Considerations for Platform Services Controllers

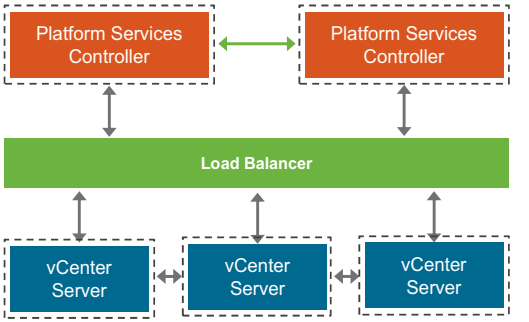
The default size for a Platform Services Controller instance (2 vCPUs, 4 GB memory) is sufficient for most cases.

When vCenter Server and Platform Services Controller instances are on separate nodes, search and login operations can be affected by slow network connections:

- Between vCenter Server and Platform Services Controller
- Between vCenter Server instances

One slow Platform Services Controller node or vCenter Server node can cause search or login operations to slow down.

As a general guideline, the latency between vCenter Server and Platform Services Controller nodes should be less than 30 milliseconds.



In most cases, you do not need to resize the Platform Services Controller virtual appliance because the default size is sufficient.

However, search and login operations can be affected by the vCenter Server-to-Platform Services Controller configuration or the vCenter Server instance-to-vCenter Server instance configuration, especially in a multisite deployment. Latency between these nodes can potentially affect performance.

https://t.me/learningnets

9-9 Concurrent vCenter Server Operations

vCenter Server has the following hard limits:

- 640 concurrent operations are allowed before incoming requests are queued.
- 2000 concurrent sessions are allowed before requests are rejected:
  - Sessions include login sessions and remote console sessions.

You can perform up to 640 concurrent vCenter Server operations before vCenter Server starts queuing the rest of the operations. Also, vCenter Server supports 2000 concurrent sessions (which include login and remote console sessions). Once this session limit is reached, vCenter Server blocks any further login and remote console requests.

https://t.me/learningnets

9-10 vCenter Server CPU Performance

As a general guideline, for CPU performance, vCenter Server CPU usage should not exceed 70 percent.

If vCenter Server CPU usage consistently exceeds 70 percent, perform the following steps:

- Identify which processes are using the most CPU.
  - If the vpxd process is consuming most of the CPU, then consider the following solutions:
    - Add more CPUs to vCenter Server.
    - Add another vCenter Server instance and balance the load across vCenter Server instances.
  - If a Java service is consuming most of the CPU, you might need to increase the heap memory size for that service.
- Examine plug-ins, extensions, or the custom API code that communicate with vCenter Server:
  - You might need to increase resources to accommodate these extensions in order for them to operate optimally.

If high CPU usage is consistent, performance is not normal, and you should check which processes (such as vpxd and vsphere-client) are using the most CPU. Consistent CPU usage of 70 percent or more might mean that a vCenter Server process is performing far more queries than it needs to and thereby increasing its CPU usage.

When you add extensions to vCenter Server, you add code to vSphere Client or vSphere Web Client, you add another extension point from where to get data, and you also invoke more services. For example, for NSX, invoking the vSphere ESX Agent Manager (EAM) service might require more CPUs and memory. These are reasons for which you might need to increase the amount of resources used by the vCenter Server instance beyond the guidance provided in VMware documentation.

https://t.me/learningnets

9-11 vCenter Server Memory Performance

As a general guideline, like CPU usage, vCenter Server memory usage should not exceed 70 percent.

If vCenter Server memory usage consistently exceeds 70 percent, perform the following steps:

- Check if vCenter Server is swapping.
- Check which services have high CPU usage. High CPU usage might indicate excessive swapping because CPU is needed to perform the swapping operation.

More memory might need to be added if high usage is consistent.

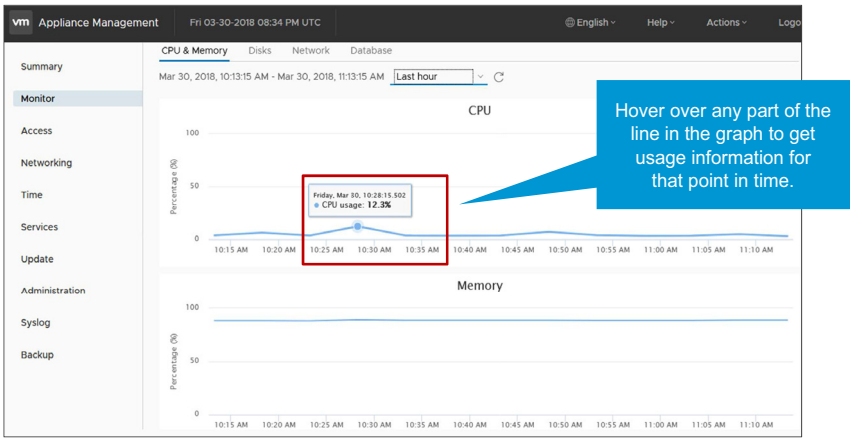
If vCenter Server is running on a VM, set the reservation value to the VM memory size to avoid host level swapping.

You definitely want to avoid swapping out vCenter Server since swapping significantly affects performance.

https://t.me/learningnets

9-12 Monitoring CPU and Memory Usage with VAMI

Use the virtual appliance management interface (VAMI) to monitor CPU and memory usage.



To access the virtual appliance management interface (VAMI), open a web browser and go to https://vCenter_Server_Name:5480. Click **Monitor** in the left pane. The **CPU & Memory** tab displays by default.

https://t.me/learningnets

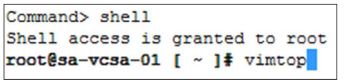
9-13 Using vimtop to Monitor vCenter Server Appliance Resources

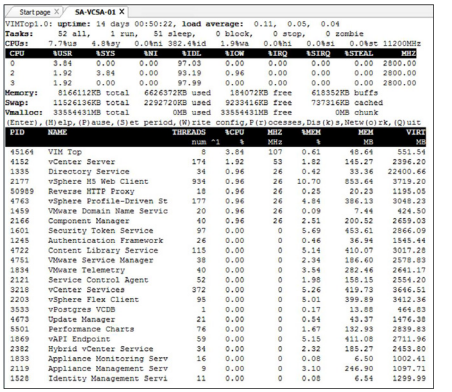
`vimtop` is a text-based tool that provides a simple way to monitor the resource usage of vCenter Server Appliance.

`vimtop` has the following characteristics:

- Similar to the `esxtop` command
- Shows overall resource information about vCenter Server Appliance
- Shows individual services and their resource usage

To start `vimtop`, you must first start the Bash shell from the vCenter Server Appliance command line.





To access `vimtop`, enter **`vimtop`** at the Bash shell command prompt.

To see the `vimtop` help menu, enter **`h`**.

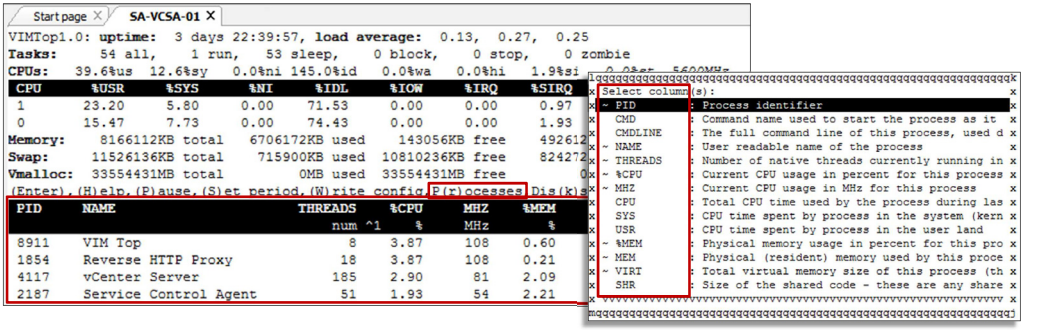
To quit `vimtop`, enter **`q`**.

https://t.me/learningnets

9-14 vimtop Commands and Columns

By default, `vimtop` displays the Processes window.

You can add additional statistics to the display by entering **c** and selecting the desired counters.



To add a metric (or column) to the `vimtop` display, perform the following steps:

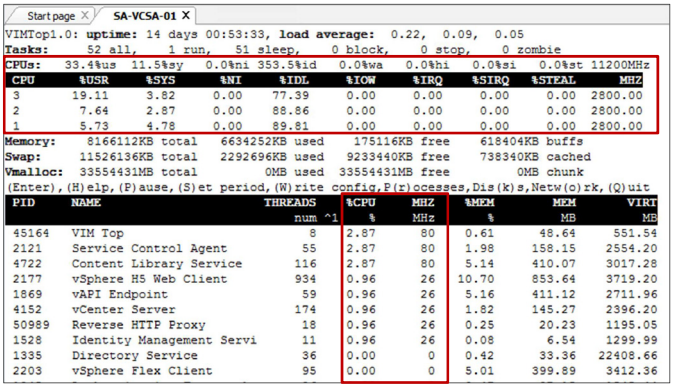
1. Enter **c** to display the column selection screen.
2. Use the up and down arrow keys to highlight a metric in the column selection screen.
3. Press the space bar to select the metric. A tilde (~) appears to the left of the metric name.
4. Press **Esc** to exit the column selection screen.

https://t.me/learningnets

9-15 Monitoring CPU Usage with vimtop

The Processes window shows overall CPU metrics and individual process CPU metrics.

To return to this window, enter **r**.



Aggregate CPU usage can be monitored from the overview pane or the CPU usage of individual process can be examined in the task pane.

Some of the columns displayed include the following:

- Name: User friendly name of the process (service).
- %CPU: Current CPU usage (in percentage) for this process. If the value is 100 percent, then the process is using an entire core.
- MHZ: Current CPU usage (in megahertz) for this process.

https://t.me/learningnets

9-16 Activity: CPU and Memory Usage (1)

As a general guideline, you should keep CPU and memory usage by a vCenter Server instance below what percentage?

- ❑ 60%
- ❑ 70%
- ❑ 80%
- ❑ 90%

https://t.me/learningnets

9-17 Activity: CPU and Memory Usage (2)

As a general guideline, you should keep CPU and memory usage by a vCenter Server instance below what percentage?

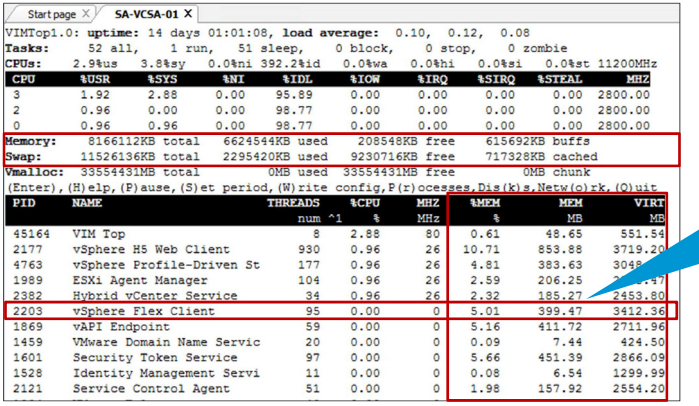
- ❑ 60%
- ✓ 70%
- ❑ 80%
- ❑ 90%

https://t.me/learningnets

9-18 Monitoring Memory Usage with vimtop

The Processes window also displays overall memory and swap metrics.

This window displays memory usage for individual vCenter Server services.



For example, the vSphere Web Client service (referred to here as vSphere Flex Client) is currently using ~400 MB of memory.

Some of the columns displayed include the following:

- %MEM: Physical memory usage (in percentage) for this process.
- MEM: Physical, or resident, memory (in MB) used by this process.
- VIRT: Total virtual memory size (in MB). This value represents the complete working set, which includes resident and swapped memory.

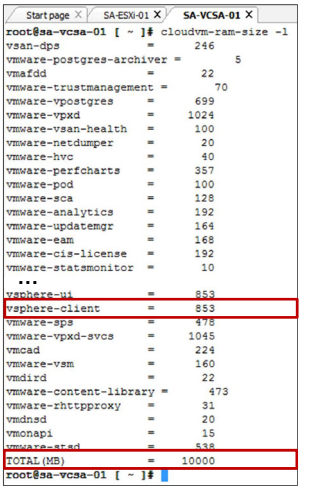
https://t.me/learningnets

9-19 Viewing Service Memory Allocation with cloudvm-ram-size

You can view the amount of memory allocated to the vCenter Server services with the `cloudvm-ram-size` command.

`cloudvm-ram-size -l` displays the memory allocated (in MB) for each service.

For example, the vSphere Web Client service (referred to here as vsphere-client) is allocated 853 MB of memory out of a total of 10,000 MB of memory (allocated to vCenter Server).



9-20 Viewing Service Heap Size

vCenter Server consists of several Java services, such as the vSphere Web Client service (vsphere-client). A Java service uses a heap (pre-reserved area of memory) to store data objects.

Proper heap size is important for Java services to perform optimally.

For example, vSphere Web Client typically needs more heap size for the following scenarios:

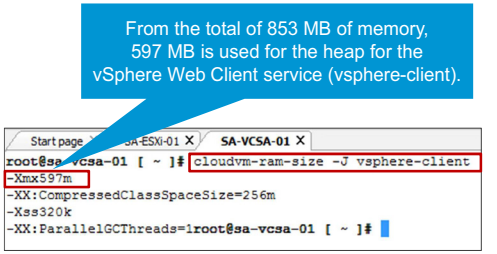
- Multiple vCenter Server instances in enhanced linked mode
- Many extensions and plug-ins
- Large inventories

The following command displays the heap size for a specific service:

- `cloudvm-ram-size -J service_name`

For example:

- `cloudvm-ram-size -J vsphere-client`



https://t.me/learningnets

9-21 Changing the Service Heap Size

You can increase the heap size of a vCenter Server service to resolve performance issues.

To change the heap size, use one of the following options:

- Resize the vCenter Server instance's memory and reboot. Heaps are automatically resized for all services.
- Resize the individual service and restart the service. No reboot is required:
  - To resize an individual service, use the following command:
    - `cloudvm-ram-size -C heap_size service_name`

  For example:
    - `cloudvm-ram-size -C 700 vsphere-client`

  This command changes the heap size of the vSphere Web Client service to 700 MB.

If you do not want to take the downtime involved with rebooting vCenter Server after adding memory, increase the heap size of each individual service and restart that service. The heap is automatically resized.

https://t.me/learningnets

9-22 vCenter Server Network Performance

Many vCenter Server operations involve communication between vCenter Server and the ESXi hosts:

- Host-to-host network bandwidth is typically more critical than vCenter Server to host bandwidth.
- In most cases, operation latency is bound by host latency.

The vCenter Server database plays a critical role in vCenter Server operations:

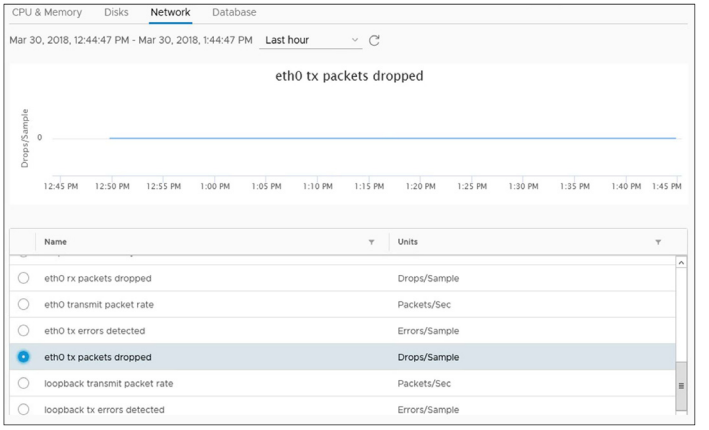
- Minimizing the latency between vCenter Server and its database is very important:

  This guideline applies mainly to a Windows vCenter Server system, since vCenter Server Appliance uses an embedded PostgreSQL database.

https://t.me/learningnets

9-23 Monitoring Network Activity with VAMI

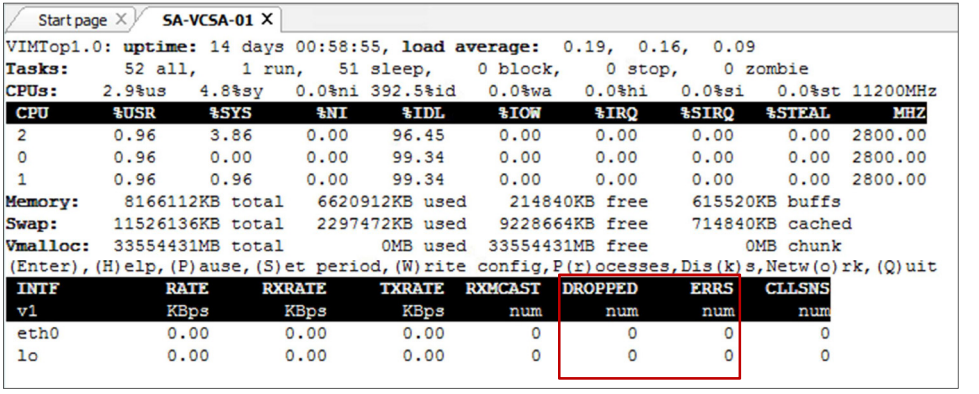
Use VAMI to monitor network activity, which includes byte rate, packets dropped, errors detected, and packet rate



https://t.me/learningnets

9-24 Monitoring Network Activity with vimtop

In the `vimtop` window, enter **o** to display network metrics.



You might monitor the following columns for network activity:

- DROPPED: Number of dropped packets through this network interface because the system ran out of buffers during the last measurement cycle.
- ERRS: Total number of transmit faults and receive faults on this interface.

https://t.me/learningnets

9-25 vCenter Server Database Performance (1)

Several factors should be considered for database performance:

- Under normal vCenter Server usage, the database resources (CPU, memory, disk, and network) should not be saturated.
- Transient high CPU usage (for example, during a rollup) is not a problem.
- Consistently high CPU usage is not normal.

The following symptoms occur with high CPU usage:

- Rollups or Top-N calculations of a large inventory are slow.
- Historical queries of a large inventory are slow.

CPU usage at 100 percent for long periods is not normal. When you see consistently high CPU usage, verify that nothing basic is malfunctioning.

If you have a large inventory, you might have a lot of statistics. If you do not have a good enough disk subsystem, sufficient memory, or sufficient CPU, then rollups or Top_N calculations might be slow.

A large inventory does not mean only a high number of hosts and VMs. A large inventory might mean a lot of datastores, networks, or devices per VM. Some of the queries require a lot of joins or full scans on tables. This is not ideal, but it does happen. In such cases, the database statistics might need to be recomputed or reindexed to make the query engine operate more efficiently.

https://t.me/learningnets

9-26 vCenter Server Database Performance (2)

If you see slow historical queries and consistently high database CPU usage:

- Check the vPostgres database:
  - Recompute database statistics on database tables.
  - Purge database tables, lower the retention policy, or lower the statistics level.
  - Check the vpxd log, database log, database partition sizes, and SEAT tables.

If you are using a third-party tool (which uses the VMware vSphere® Web Services SDK) to obtain historical data:

- Verify queries.
- Check events, tasks, and statistics tables.

If you experience slow rollups or Top-N calculations:

- Add CPUs.
- Check memory.
- Check rollup status.
- Check alarms.

You can check the vpxd log for diagnosis. You might see the `[VdbStatement] Execution elapsed time: …` and `SQL execution took too long` statements in the log. You should expect to see a few of these events (often for Top_N, statistics, and events), and each typically lasts three to four seconds. If these events are extremely frequent and if each event lasts 10 or more seconds, this behavior often indicates that problems are occurring in your database. You should check the resources (I/O, CPU, and memory) and verify that the database is correctly provisioned.

https://t.me/learningnets

9-27 Effects of Changing Statistics Level on Database Traffic

vCenter Server uses a collection level, or statistics level, to determine the amount of data gathered and stored in the database.

When you increase the statistics level, the storage and system requirements might change.

The configuration of VMs and hosts, as well as the number of datastores, directly affects the number of statistics that are generated.

Hardware	VM 1	VM 2	ESXi Host
CPUs	2	2	48 (logical)
Virtual disks	11	1	13
Datastores	1	1	9
NICs	1	1	3

When changing from level 1 to level 2, the number of storage and network counters increases significantly.

Statistics Level	VM 1	VM 2	ESXi Host
1	67	34	223
2	231	148	858
3	263	184	1,779
4	348	196	1,967

The tables show what happens when different statistics levels are used. The configuration of two VMs and a host is shown in the first table. The configuration of your VMs and hosts is critical to how many statistics are generated. In the example, the VM configurations are the same except for the number of virtual disks.

The second table shows what happens when you increase the statistics level. Of the four statistics levels, 1 is the least detailed and 4 is the most detailed. The specific numbers are not as important as the significant increase in statistics from level 1 to level 2. Level 1 reports on aggregated statistics, such as aggregated CPU, aggregated disk latencies, and so on. Level 2 reports more detailed statistics, such as per-NIC and per-datastore statistics.

In the second table, you can see that VM1 has far more statistics than VM2, which is because VM1 has many more disks than VM2. Several datastore and network counters (such as read/write requests per second and packets sent/received) are introduced at level 2. As a result, the larger the number of disks and networks in your configuration, the more statistics that are generated.

In your environment, ensure that you size the infrastructure and database optimally to accommodate the load of a level.

https://t.me/learningnets

9-28 vCenter Server Appliance Disk Usage (1)

The vCenter Server database is the main disk bandwidth consumer and is also write-intensive.

Place write-intensive disk partitions, such as `/storage/db`, on high-speed storage. `/storage/db` is the location of the vCenter Server database.

Disk (VMDK)	Minimum Size	Mount Point	Purpose
VMDK1	12 GB	`/` (10 GB), `/boot` (123 MB), `SWAP` (1 GB)	Kernel images and boot loader configurations
VMDK2	1.8 GB	`/tmp`	Temporary files
VMDK3	25 GB	`SWAP`	Used when system is out of memory to swap to disk
VMDK4	25 GB	`/storage/core`	Core dumps from vpxd process
VMDK5	10 GB	`/storage/log`	vCenter Server and Platform Services Controller logs
VMDK6	10 GB	`/storage/db`	VMware Postgres database storage location

The substantial source of disk I/O tends to be the database. The main disk bandwidth consumers are the following partitions:

- `/storage/db`
- `/storage/dblog`
- `/storage/seat`

If you are using vCenter Server Appliance, ensure that these partitions are on high-speed storage and have sufficient space.

If you need to increase the disk space for vCenter Server Appliance, see VMware knowledge base article 2145603 at http://kb.vmware.com/kb/2145603.

https://t.me/learningnets

9-29 vCenter Server Appliance Disk Usage (2)

Place other database partitions, such as `/storage/dblog` and `/storage/seat`, on high-speed storage as well.

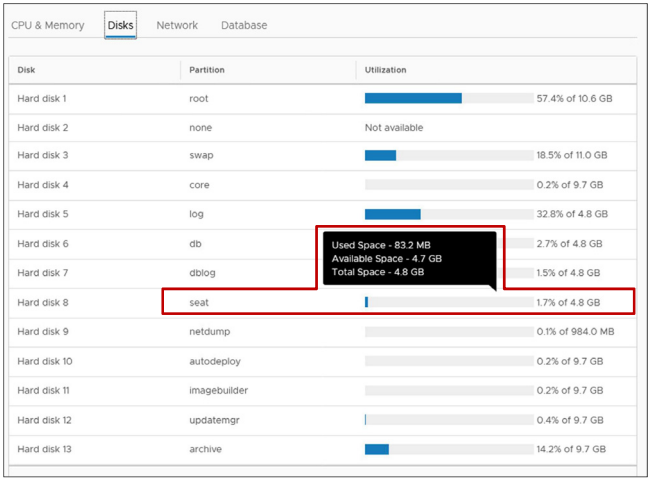
Disk (VMDK)	Minimum Size	Mount Point	Purpose
VMDK7	15 GB	`/storage/dblog`	VMware Postgres database logging location
VMDK8	10 GB	`/storage/seat`	Statistics, events, alarms, and tasks for VMware Postgres
VMDK9	1 GB	`/storage/netdump`	vSphere Netdump collector repository for ESXi dumps
VMDK10	10 GB	`/storage/autodeploy`	vSphere Auto Deploy repository
VMDK11	10 GB	`/storage/imagebuilder`	vSphere ESXi Image Builder repository
VMDK12	100 GB	`/storage/updatemgr`	vSphere Update Manager repository
VMDK13	50 GB	`/storage/archive`	Repository used for archiving files

https://t.me/learningnets

9-30 Monitoring Disk Usage with VAMI

Use VAMI to monitor disk usage by partition.

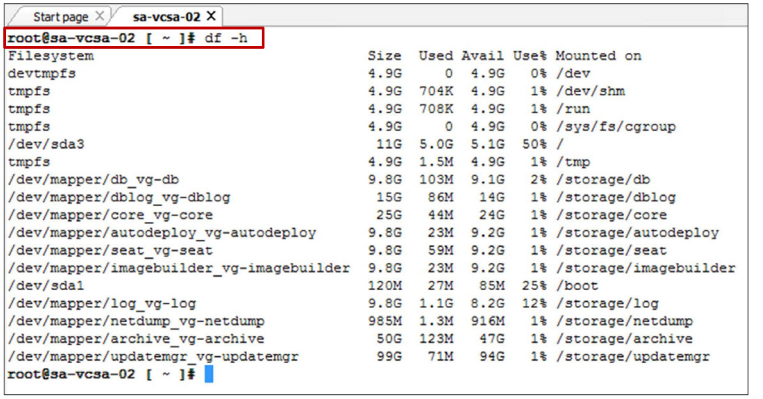
Hover over a partition to view information on used space, available space, and total space.



https://t.me/learningnets

9-31 Monitoring Disk Usage with the df Command

You can also monitor disk usage from the vCenter Server Appliance command line by using the `df -h` command.

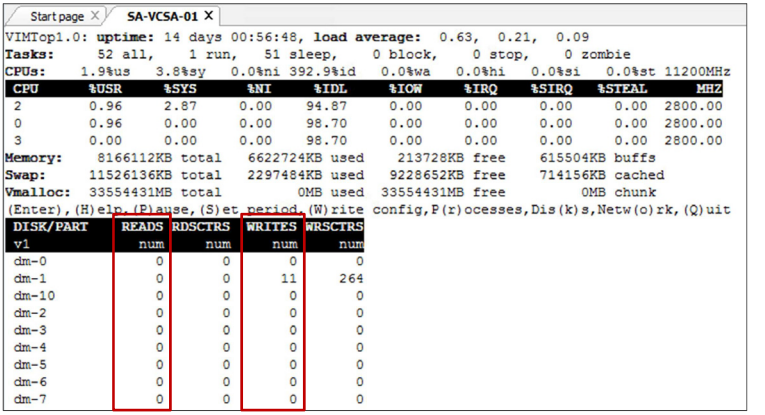


The `df` command displays file system usage information, such as total size, used space, and available space. The `-h` option shows sizes in powers of 1024, for example, in megabytes, gigabytes, and so on.

https://t.me/learningnets

9-32 Monitoring Disk Activity with vimtop

In the `vimtop` window, enter **k** to display disk metrics.



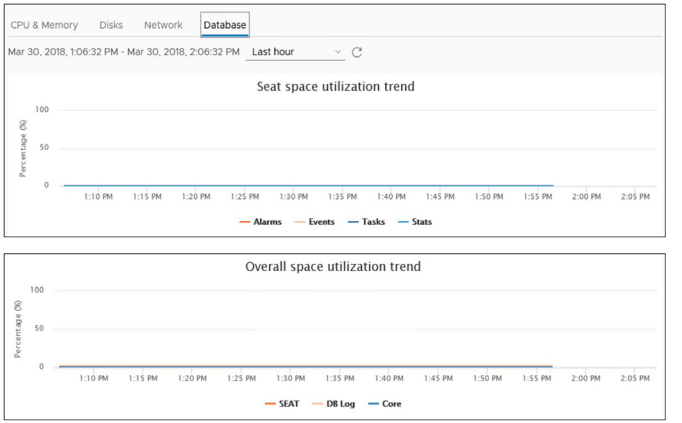
The following columns are displayed:

- READS: Number of reads issued to this partition and completed successfully during the last measurement interval.
- WRITES: Number of writes issued to this partition and completed successfully during the last measurement interval.

https://t.me/learningnets

9-33 Monitoring Database Activity with VAMI

Use VAMI to monitor space utilization trends for SEAT (stats, events, alarms, and tasks) activity, database log activity, and inventory data (core) activity.



9-34 Activity: Network Performance (1)

Host-to-host network bandwidth is typically more critical than vCenter Server-to-host bandwidth.

- ❑ True
- ❑ False

https://t.me/learningnets

9-35 Activity: Network Performance (2)

Host-to-host network bandwidth is typically more critical than vCenter Server-to-host bandwidth.

- ✓ True
- ❑ False

https://t.me/learningnets

9-36 Activity: vCenter Server Tools (1)

Which of the following tools can you use to monitor the free space in the vCenter Server database? Select all that apply.

- ❑ df
- ❑ vimtop
- ❑ cloudvm-ram-size
- ❑ VAMI

https://t.me/learningnets

9-37 Activity: vCenter Server Tools (2)

Which of the following tools can you use to monitor the free space in the vCenter Server database? Select all that apply.

- ✓ df
- ❑ vimtop
- ❑ cloudvm-ram-size
- ✓ VAMI

https://t.me/learningnets

9-38 Review of Learner Objectives

You should be able to meet the following objectives:

- Review vCenter Server components and services
- Describe the performance considerations for Platform Services Controller
- Describe the factors that influence vCenter Server performance
- Use vCenter Server Appliance tools to monitor resource usage

https://t.me/learningnets

9-39 Key Points

- vCenter Server comprises many services.
- You can monitor vCenter Server Appliance using tools such as VAMI, `vimtop`, `df`, and `cloudvm-ram-size`.
- As a general guideline, vCenter Server CPU and memory usage should not exceed 70 percent.
- Low-latency connections both between vCenter Server and the database and between vCenter Server and Platform Services Controller nodes are recommended.
- Because the vCenter Server database is the main disk bandwidth consumer, it is important to ensure that the database partitions have enough space.
- In certain circumstances, increasing service heap size can improve service performance.

Questions?

https://t.me/learningnets

https://t.me/learningnets

MODULE 10

vSphere Security



https://t.me/learningnets

10-2 You Are Here

1. Course Introduction
2. Network Scalability
3. Storage Scalability
4. Host and Management Scalability
5. CPU Optimization
6. Memory Optimization
7. Storage Optimization
8. Network Optimization
9. vCenter Server Performance Optimization
10. **vSphere Security**

https://t.me/learningnets

10-3 Importance

By default, vSphere components are secured by certificates, authorization, limited access, and compliance to security standards.

You harden your vSphere environment against security threats by controlling settings for vCenter Server systems, ESXi hosts, virtual machines, and the vSphere network.

Virtual machine encryption can be used to encrypt virtual machine files and disks to protect confidential customer data.

https://t.me/learningnets

10-4 Module Lessons

Lesson 1:	Configuring ESXi Host Access and Authentication
Lesson 2:	Securing vSphere
Lesson 3:	VMware Certificate Management
Lesson 4:	Securing Virtual Machines

https://t.me/learningnets

10-5 Lesson 1: Configuring ESXi Host Access and Authentication



https://t.me/learningnets

10-6 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

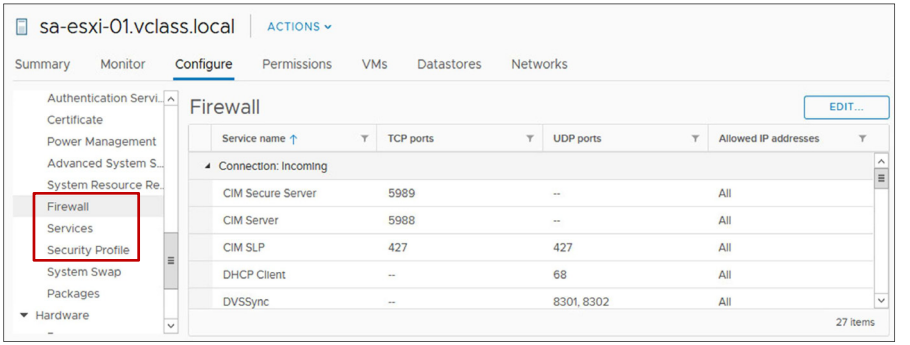
- Configure the ESXi firewall by enabling and disabling services
- Enable and disable lockdown mode on an ESXi host
- Configure user logins to authenticate with directory services

https://t.me/learningnets

10-7 About ESXi Host Security Properties

The ESXi hypervisor architecture has many built-in security features, such as CPU isolation, memory isolation, and device isolation

You can configure additional features such as a firewall, disabling services such as SSH, and configuring lockdown mode.

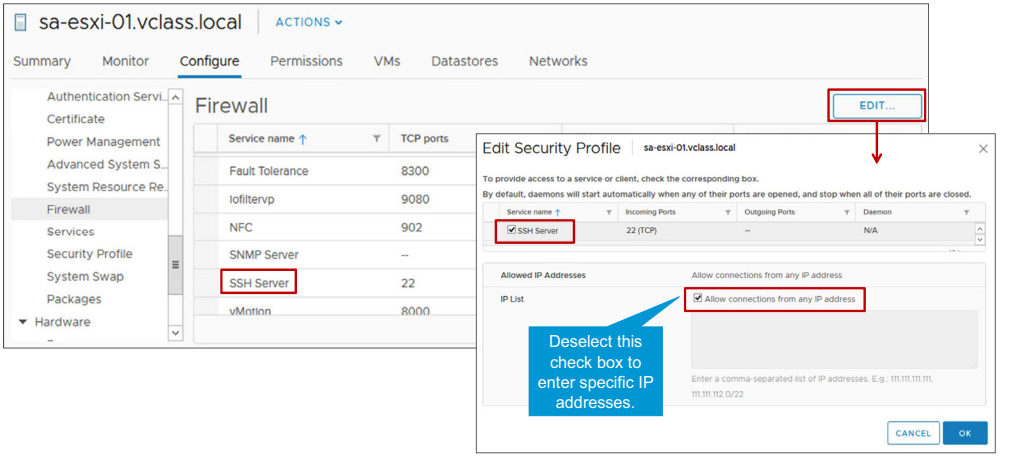


An ESXi host is protected with a firewall. You can open ports for incoming and outgoing traffic, as needed, but should restrict access to services and ports. Using the ESXi lockdown mode and limiting access to the ESXi Shell can further contribute to a more secure environment.

https://t.me/learningnets

10-8 Configuring the ESXi Firewall

To minimize the risk of an attack through the management interface, ESXi includes a service-oriented, stateless firewall.



Firewalls control access to devices in their perimeter by closing all communication pathways, except for the pathways that the administrator explicitly or implicitly designates as authorized. The pathways, or ports, that administrators open in the firewall enable traffic between devices on different sides of the firewall.

With the ESXi firewall engine, rule sets define port rules for each service. For remote hosts, you can specify the IP addresses or range of IP addresses that are allowed to access each service.

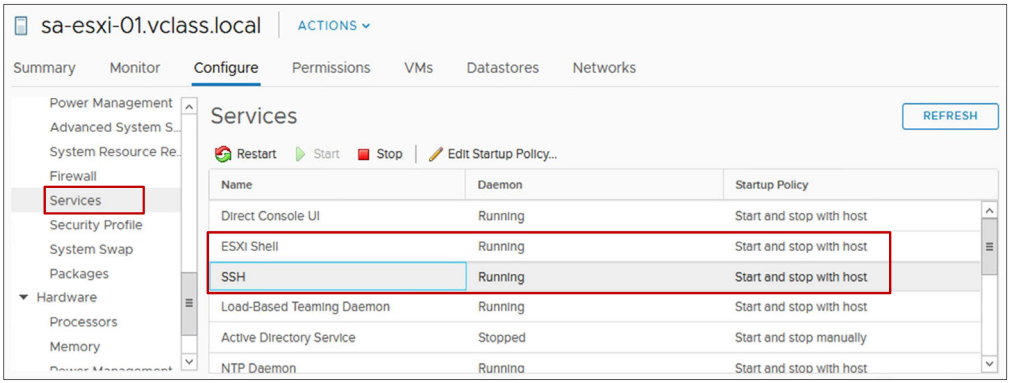
The ESXi firewall can be configured using vSphere Client or VMware vSphere® Command-Line Interface. Firewall management is configured at the command line by using the `esxcli network firewall` namespace.

https://t.me/learningnets

10-9 Configuring ESXi Services

You configure services to start, stop, or restart on an ESXi host. You also control whether the services start and stop with the host or must be started manually.

In this example, the ESXi Shell and SSH services have been started. These services are stopped by default.



You configure services to start, stop, or restart them on an ESXi host. You also control whether the services start and stop with the host or must be started manually.

https://t.me/learningnets

10-10 About Lockdown Mode (1)

To increase the security of your ESXi hosts, you can put your hosts in lockdown mode.

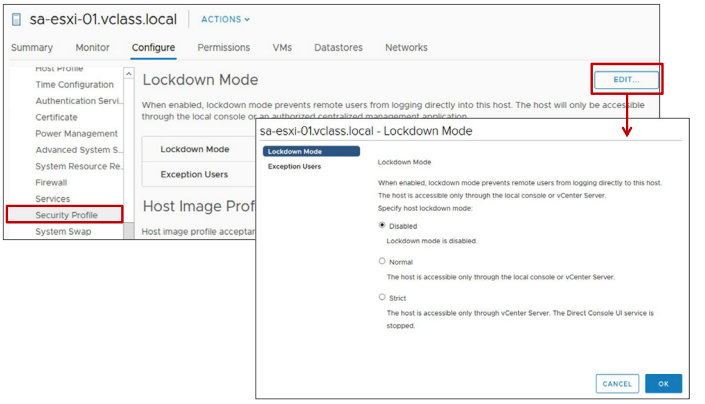
In lockdown mode, some services are disabled and some services are accessible only to certain users.

Two modes are available:

- Normal
- Strict

By default, lockdown mode is disabled.

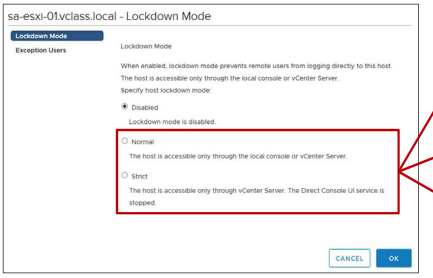
You can enable lockdown mode with vSphere Client, when you create a new ESXi host or modify an existing host's security profile.

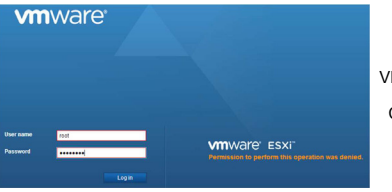


https://t.me/learningnets

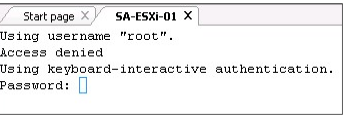
10-11 About Lockdown Mode (2)

When you enable lockdown mode (normal or strict), no users but vpxuser have authentication permissions. Also, users cannot perform operations against the host directly.

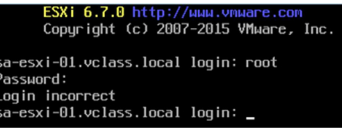




VMware Host Client



SSH Session



Local ESXi Shell

Normal lockdown mode forces all operations to be performed through vCenter Server. The host cannot be accessed directly by using VMware Host Client.

When a host is in lockdown mode, you cannot run commands from vSphere CLI, from a local or remote ESXi Shell, or from a script. External software or management tools might not be able to retrieve or modify information from the ESXi host.

https://t.me/learningnets

10-12 Normal Lockdown Mode Behavior

When normal lockdown mode is enabled on an ESXi host, the following access is allowed:

- The host can be accessed only by vCenter Server, for example, using vSphere Client.
- Access to the Direct Console User Interface (DCUI) is restricted to certain users:
  - Users on the Exception Users list, but they must have administrator privileges
  - Users defined in the DCUI.Access advanced system setting:
    - By default, the root user is defined in this setting.
    - Non-administrative users can also be defined in this setting.
- If the ESXi Shell and SSH services are enabled, then users with administrator privileges who are on the Exception Users list can log in to the host directly using SSH or the local ESXi Shell.

In normal lockdown mode, users listed in the DCUI.Access advanced system setting can log in to the Direct Console User Interface (DCUI). By being on this list, users have emergency access to the DCUI in case the connection to vCenter Server is lost. These users do not require administrative privileges on the host.

https://t.me/learningnets

10-13 Strict Lockdown Mode Behavior

When strict lockdown mode is enabled on an ESXi host, the following access is allowed:

- The host can be accessed only by vCenter Server, for example, using vSphere Client.
- The DCUI service is disabled and cannot be started. Therefore, no users can log in to the DCUI.
- If the ESXi Shell and SSH services are enabled, then users with administrator privileges who are on the Exception Users list can log in to the host directly using SSH or the local ESXi Shell.

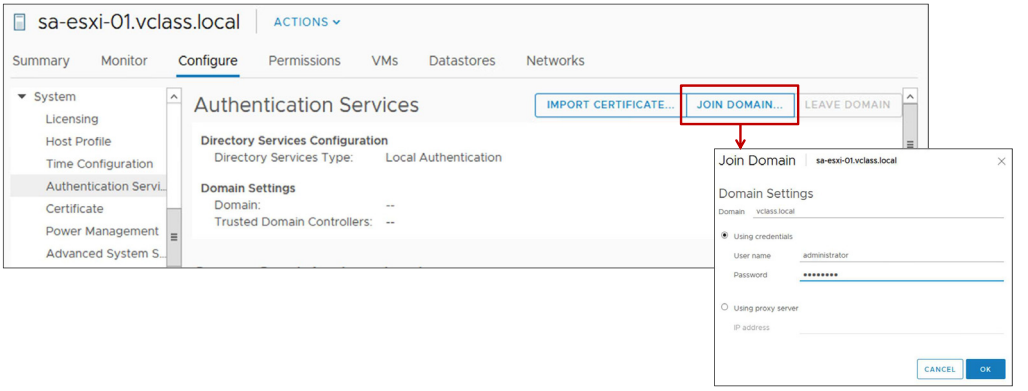
For strict or normal lockdown mode to be an effective security measure, ensure the the ESXi Shell and SSH services are also disabled.

If vCenter Server is unavailable, then vSphere Client is also unavailable. Therefore, hosts that are in strict lockdown mode are not manageable.

https://t.me/learningnets

10-14 Integrating ESXi with Active Directory

You can configure an ESXi host to join an Active Directory domain to manage users and groups. When the ESXi host is added to Active Directory, the ESX Admins domain group is assigned full administrative access to the host, if this group exists.



Although day-to-day vSphere management operations are usually done while logged in to vCenter Server using vSphere Client, the user sometimes requires direct access to the ESXi host. Examples include when accessing local log files and when configuring backups.

vCenter Single Sign-On is the recommended way to manage user access to hosts. The Active Directory (AD) domain to which users belong is added to vCenter Single Sign-On as an identity source. In addition, you can still have local users defined and managed host-by-host, and configured by using vSphere Client. You can use this approach either in place of or in addition to the vCenter Single Sign-On and AD integration.

Whenever you are asked to provide credentials (for example, when using VMware Host Client to log in directly to the ESXi host), you can enter the user name and password of a user in the domain to which the host is joined. The advantage of this user account management model is that you can use AD to manage your ESXi host user accounts. This model is easier and more secure than trying to manage accounts independently per host.

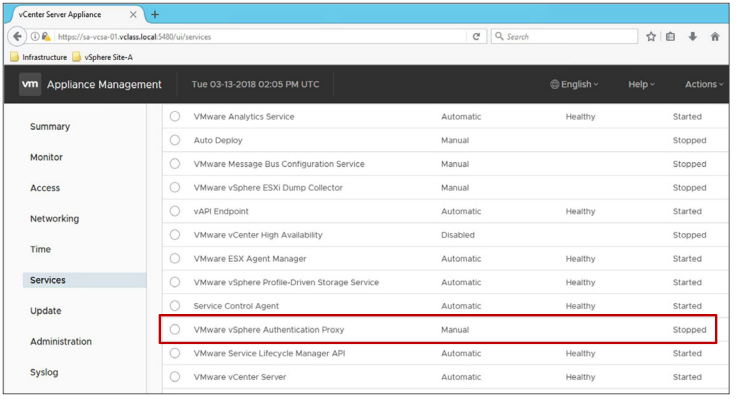
The only user that is defined on the ESXi host is root. The root password is not mapped to an AD account. The initial root password is typically set during ESXi installation but it can be changed later through VMware Host Client or the DCUI.

If the host is integrated with AD, local roles can also be granted to AD users and groups. For example, an AD group can be created to include users who should have an administrator role on a subset of ESXi hosts. On those hosts, the administrator role can be granted to that AD group. For all other servers, those users would not have an administrator role. The ESXi host's subset grants administrator access to the AD group named ESX Admins, which allows the creation of a global administrators group.

https://t.me/learningnets

10-15 VMware vSphere Authentication Proxy Service

You can add ESXi hosts to an Active Directory domain by using the VMware vSphere Authentication Proxy service. This service provides support for joining unattended ESXi hosts to an Active Directory domain by using an account with delegated privileges.



You can use VMware vSphere® Authentication Proxy to join the host to an AD domain, instead of adding the host directly to the AD domain. A chain of trust exists between vSphere Authentication Proxy and the host, which allows vSphere Authentication Proxy to join the host to the AD domain.

vSphere Authentication Proxy is especially useful when used with vSphere Auto Deploy. You can set up a reference host that points to vSphere Authentication Proxy and set up a rule that applies the reference host profile to any ESXi host provisioned with vSphere Auto Deploy. Even if you use vSphere Authentication Proxy in an environment that uses certificates that are provisioned by VMware Certificate Authority or third-party certificates, the process works seamlessly as long as you follow the instructions for using custom certificates with vSphere Auto Deploy.

The VMware vSphere Authentication Proxy service is available on each vCenter Server system. By default, the service is not running. If you want to use vSphere Authentication Proxy in your environment, you can start the service from the virtual appliance management interface (VAMI) or from the command line.

https://t.me/learningnets

10-16 Lab 12: Configuring Lockdown Mode

Configure and test lockdown mode

1. Start the vSphere ESXi Shell and SSH Services
2. Test the SSH Connection
3. Enable and Test Lockdown Mode
4. Disable Lockdown Mode
5. Examine the DCUI.Access List

https://t.me/learningnets

10-17 Review of Learner Objectives

You should be able to meet the following objectives:

- Configure the ESXi firewall by enabling and disabling services
- Enable and disable lockdown mode on an ESXi host
- Configure user logins to authenticate with directory services

https://t.me/learningnets



https://t.me/learningnets

10-19 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Explain the importance of the vSphere Security Configuration Guide
- Discuss recommendations for vCenter Server security
- Summarize strategies to secure the vSphere management network
- Discuss recommendations for ESXi host security
- Plan for secure boot and TPM 2.0 support for an ESXi host
- Discuss general virtual machine and guest operating system protection

https://t.me/learningnets

10-20 About Securing vSphere

Securing vSphere involves aspects of security for the vCenter Server system, ESXi hosts, virtual machines, and the vSphere network. The VMware website provides several security resources.

Topic	Resource
vSphere Security	https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-security-guide.pdf
VMware security policy	http://www.vmware.com/security/
Corporate security response	http://www.vmware.com/support/policies/security_response.html
VMware security advisories	http://www.vmware.com/security/advisories.html

VMware security advisories provide information about security vulnerabilities that are reported in VMware products. You can sign up to receive new and updated advisories by email on the VMware Security Advisories webpage: http://www.vmware.com/security/advisories.html.

For more information about security resources and security best practices and methods, see *vSphere Security* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-security-guide.pdf.

https://t.me/learningnets

10-21 About the vSphere Security Configuration Guide (1)

The vSphere Security Configuration Guide helps you to configure your vSphere environment according to operations security best practices:

- vSphere has, over time, become more secure by default.
- As a result this guide has become less about hardening and more about ensuring that best practices are followed.

For the latest version of the vSphere Security Configuration Guide, see VMware Security Hardening Guides at https://www.vmware.com/security/hardening-guides.html.

Perform a daily audit against the vSphere Security Configuration Guide to identify security vulnerabilities in your vSphere environment.

https://t.me/learningnets

10-22 About the vSphere Security Configuration Guide (2)

The vSphere Security Configuration Guide is in the form of an easy-to-use spreadsheet. Script examples are available to help you automate security checking.

This spreadsheet categorizes each setting based on one or more characteristics.

Characteristic	Description	Example of Setting
Hardening	If set to true, then this setting is an actual hardening setting.	`VM.disable-disk-shrinking-shrink`: Explicitly disable the ability to shrink virtual disks.
Site specific	If set to true, this setting is one that VMware cannot provide default values for.	`ESXi.set-password-policies`: Establish a password policy for password complexity.
Audit	If set to true, this setting should be checked on a regular basis to see if the default value changed.	`vNetwork.reject-mac-changes-dvportgroup`: Ensure that the MAC Address Changes policy is set to reject.

https://t.me/learningnets

10-23 About the vSphere Security Configuration Guide (3)

From the vSphere Security Configuration Guide, you can choose the risk profile to apply to your vSphere environment:

- Assess each guideline in the risk profile for risk management and the impact that the risk poses to the business and operations.

Risk Profile	Target Production Environment	Examples
1	Highest security	Top-secret government or military environments
2	Sensitive	Environments that handle sensitive data, subject to strict compliance rules
3	All	All production environments

The vSphere Security Configuration Guide includes guidelines on how to address certain security vulnerabilities in vSphere components, such as ESXi hosts, virtual machines, and virtual networks. This guide contains metadata to allow for guideline classification and risk assessment.

https://t.me/learningnets

10-24 Controlling Access to vCenter Server Systems

Strictly control access to vCenter Server components to increase security for the system.

vCenter Server access control includes the following best practices:

- Use named accounts and grant the Administrator role only to accounts that are required to have it.
- Restrict datastore browser access to users who really need that privilege.
- Instruct users of vSphere Client to never ignore certificate verification warnings.
- Set the vCenter Server password policy.
- Limit vCenter Server network connectivity by putting the vCenter Server system only on the management network.

Starting with vSphere 6.0, the local administrator no longer has full administrative rights to vCenter Server by default. Instead, assign the vCenter Server Administrator role to one or more named vCenter Server administrator accounts.

Assign the Browse Datastore privilege only to users or groups who really need the privilege. Users with the privilege can view, upload, or download files on datastores associated with the vSphere deployment through the web browser or vSphere Client.

Verify vSphere Client certificates. Without certificate verification, the user might be subject to man-in-the-middle attacks.

By default, vCenter Server automatically changes the vpxuser password every 30 days. You can change this value in vSphere Client.

Ensure that vSphere management traffic is on a restricted network.

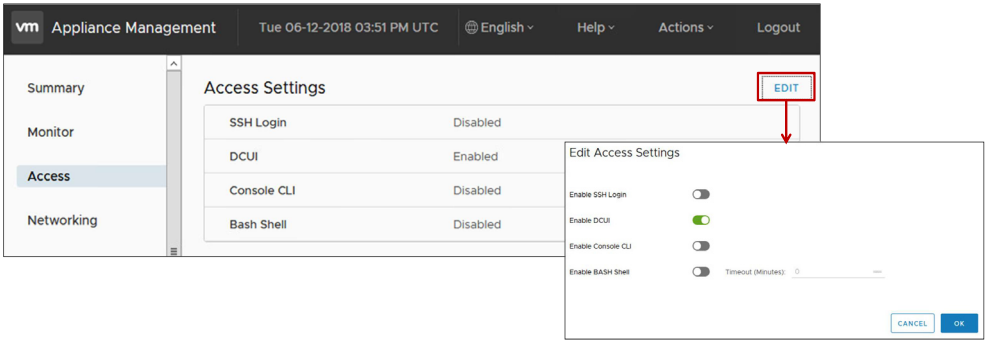
For the complete list of vCenter Server best practices, see *vSphere Security* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-security-guide.pdf.

https://t.me/learningnets

10-25 vCenter Server Access Control

To secure access to vCenter Server, consider the following best practices:

- Do not allow users to log directly in to the vCenter Server system.
- Allow access only to users that have legitimate tasks to perform on vCenter Server.
- Control access to SSH and the command line.



Users who are directly logged in to the vCenter Server instance can cause harm, either intentionally or unintentionally, by altering settings and modifying processes. Those users also have potential access to vCenter Server credentials, such as the SSL certificate.

Consider auditing login events on a regular basis to ensure that only authorized users are accessing vCenter Server.

Use VAMI to control access to SSH, the DCUI, the console command line, and the Bash shell on a vCenter Server Appliance instance.

https://t.me/learningnets

10-26 Securing vCenter Server with TLS 1.2

Transport Layer Security (TLS) is a cryptographic protocol that provides endpoint authentication and secure communications over any transport. The predecessor to TLS is Secure Sockets Layer (SSL).

In vCenter Server 6.7, TLS 1.2 is enabled by default:

- TLS 1.0 and TLS 1.1. protocols are disabled.

If you need to use the TLS 1.0 and TLS 1.1 protocols to support products or services that do not support TLS 1.2, you can enable and disable these protocols by using the TLS Configurator Utility.

For more information about the TLS Configuration Utility, see VMware knowledge base article 2147469 at https://kb.vmware.com/kb/2147469.

For a list of VMware products that support disabling TLS 1.0 and TLS 1.1, see VMware knowledge base article 2145796 at http://kb.vmware.com/kb/2145796.

https://t.me/learningnets

10-27 Securing vCenter Server Systems (1)

For Windows vCenter Server systems, ensure that the operating system on the host machine is secure:

- Maintain supported operating system, database, and hardware versions on the vCenter Server system.
- Keep the vCenter Server system properly patched.
- Provide the operating system with antivirus and antimalware software.
- Ensure that Remote Desktop Protocol host configuration settings are set to their highest encryption level.

Communication between the Windows vCenter Server system and an external MS SQL Server database is encrypted through the TLS 1.2 protocol.

Follow security best practices for Windows vCenter Server systems:

- If vCenter Server is not running on a supported Windows operating system, it might not run properly, making vCenter Server vulnerable to attacks.
- Stay up to date with operating system patches so that the server is less vulnerable to attack.
- Provide operating system protection (such as antivirus and antimalware software) on the Windows vCenter Server system.
- Ensure that Remote Desktop (RDP) Host Configuration settings are set to ensure the highest level of encryption according to industry-standard guidelines or internal guidelines on each Windows computer in the infrastructure.

https://t.me/learningnets

10-28 Securing vCenter Server Systems (2)

For vCenter Server Appliance instances, ensure that vCenter Server networking is secure:

- Configure Network Time Protocol to ensure that all systems use the same relative time source.
- Restrict access to components that are required to communicate with vCenter Server Appliance.

Ensure that your vCenter Server Appliance instance is installed with the latest security patches:

- VMware regularly releases patches on a monthly basis for vCenter Server Appliance.
- Download patch ISO images from https://my.vmware.com/group/vmware/patch.

Follow security best practices for vCenter Server Appliance instances:

- Ensure that all systems use the same relative time source. Synchronized systems are essential for certificate validation. Network Time Protocol also makes it easier to track an intruder in log files.
- Restrict access to components that are required to communicate with the vCenter Server Appliance instance. Blocking access from unnecessary systems reduces the potential for attacks on the operating system.

VMware makes patches available on a monthly basis. These patches might be related to third-party products in the platform, core product functionality, or both. These patches can only be applied in between major releases of vCenter Server Appliance. For example, patches released for the initial release of vCenter Server Appliance 6.7 are not applicable to vCenter Server Appliance 6.7 Update 1 because any patches previously made available will be included with the Update 1 release.

For details on how to patch vCenter Server Appliance, see *vCenter Server Upgrade* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-vcenter-server-67-upgrade-guide.pdf.

https://t.me/learningnets

10-29 Securing vSphere Management Networks

Securing vSphere management networks resembles securing physical networks, though with some special characteristics:

- Balance firewall usage against virtual machine performance.
- Secure the physical switch on each ESXi host to prevent attackers from gaining access to the host and its virtual machines.
- Secure standard switch ports and vSphere distributed switches with security policies.
- Use physically isolated networks. If physical isolation is not possible, then use VLANs.

Create and install the vCenter Server system on a management network whenever possible. Avoid putting the vCenter Server system on production or storage networks, or on a network with access to the Internet.

vCenter Server systems require network connectivity to only the following systems:

- ESXi hosts
- The vCenter Server database
- Other vCenter Server systems in the same vCenter Single Sign-On domain
- Systems that are authorized to run management clients
- Systems that run add-on components, such as vSphere Update Manager
- Infrastructure services such as DNS, AD, and Network Time Protocol
- Other systems that run components that are essential to the functionality of the vCenter Server system

For the best protection of your hosts, ensure that physical switch ports are configured with spanning tree disabled and that the nonnegotiate option is configured for trunk links between external physical switches and virtual switches in Virtual Switch Tagging mode.

https://t.me/learningnets

10-30 Securing ESXi Hosts

VMware constrains vSphere parameters and settings in order to secure ESXi hosts.

To maintain ESXi host security, consider these guidelines:

- Manage access to the ESXi host:
  - Avoid using the root user account.
  - Use AD to manage users instead of maintaining local user accounts.
  - Always use vSphere Client to administer ESXi hosts managed by a vCenter Server instance.
  - Use VMware Host Client to administer ESXi hosts only if vCenter Server is unavailable.
- Manage the use of security services and functions:
  - Keep the ESXi host firewall enabled.
  - Leave the SSH and ESXi shell services disabled.
  - Enable lockdown mode to further increase the security of ESXi hosts.
- Use VMware Certificate Authority to provision each ESXi host with a signed certificate. VMware CA is the root certificate authority (CA).

Create at least one named user account and assign it full administrative privileges. This user account can be either a local account or a domain account that belongs to the ESX Admins domain group. Use this account instead of the root account. Set a very complex password for the root account and limit its use, but do not remove the root user.

You can configure ESXi hosts to use AD to manage users. Using AD alleviates maintaining local user accounts on each ESXi host.

ESXi includes a firewall that is enabled by default. You can manage firewall settings by using security profiles. Many essential security settings for ESXi hosts can be customized using security profiles. Using security profiles is especially useful for single host management.

You can use SSH keys to restrict, control, and secure access to an ESXi host. SSH access and vSphere ESXi Shell access are disabled on ESXi hosts by default.

Lockdown mode is disabled by default. When you use lockdown mode, operations on ESXi hosts must be performed through vCenter Server.

You manage certificates from vSphere Client. ESXi hosts can use VMware CA, a custom certificate authority, or thumbprint mode. By default, a certificate signed by VMware CA is provisioned when you boot an ESXi host from installation media.

For the complete list of ESXi hypervisor and ESXi host security best practices, see *vSphere Security* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-security-guide.pdf.

https://t.me/learningnets

10-31 UEFI Secure Boot for ESXi Hosts

Secure Boot is part of the Unified Extensible Firmware Interface (UEFI) firmware standard.

With Secure Boot enabled, a machine refuses to load any UEFI driver unless the operating system bootloader is cryptographically signed.

ESXi supports secure boot if it is enabled in the hardware.



Machine Enabled with UEFI Secure Boot

To ensure that boot code is executed without modification, Unified Extensible Firmware Interface (UEFI) uses a digital signature provided by a trusted code creator. The digital signature is embedded in every executable code section. Using public-private key pairs, the code creator signs the code with a private key, which can be checked against the public key in a preinstalled signature before the code is executed. If the executable code is marked as modified or invalid, then the code is not executed in the boot path. The system might take an alternate boot path, or the user might be notified to take remedial actions.

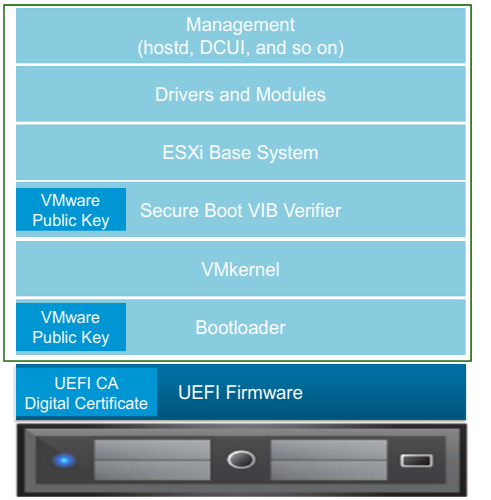
https://t.me/learningnets

10-32 Secure Boot Sequence for ESXi Hosts

ESXi supports UEFI Secure Boot at each level of the boot stack.

The boot sequence proceeds as follows:

- The UEFI firmware validates the bootloader and the VMkernel.
- The secure boot VIB verifier verifies every vSphere installation bundle package that is installed on the system.



If the security verifications pass during the boot sequence, the entire system is booted, with the root of trust in certificates that are part of the Unified Extensible Firmware Interface (UEFI) firmware.

If UEFI Secure Boot does not succeed at any level of the boot sequence, an error results. If you attempt to boot with a bootloader that is unsigned, an error results. The exact error message depends on the hardware vendor. If a VIB or driver has been tampered with, a purple screen appears on the ESXi host stating that the UEFI Secure Boot failed to verify signatures of the invalid VIBs.

https://t.me/learningnets

10-33 TPM 2.0 on ESXi Hosts

ESXi can use Trusted Platform Module (TPM) chips to enhance host security.

TPM protects users from software-based attacks that attempt to steal sensitive information by corrupting system and BIOS code, or by modifying the platform's configuration.

TPM is an industry-wide standard for secure cryptoprocessors. The Trusted Computing Group (TCG) is responsible for TPM technical specifications.

The dedicated microprocessor is designed to secure hardware by integrating cryptographic keys into devices.

vSphere 6.7 introduces support for TPM 2.0.

TPM 1.2 and TPM 2.0 are two vastly different implementations:

- Servers are shipped with either the TPM 1.2 or the TPM 2.0 chip.



Trusted Platform Module (TPM) chips are found in most of today's computers, from laptops, to desktops, to servers. The TPM chip usually is part of the system board and therefore the user may not be able to change it after purchase. It is important for users to select the correct TPM hardware at the time of purchase.

Since the initial publication, the Trusted Computing Group (TCG) has released two major revisions: 1.2 and 2.0. TPM hardware is designed to be compliant with either 1.2 or 2.0 specifications. TPM 2.0 is not backward compatible.

https://t.me/learningnets

10-34 About Remote Attestation

TPM hardware attests to an ESXi host's identity. This is called remote attestation.

Remote attestation is the process of authenticating and attesting to the state of the host's software at a given point in time.

TPM hardware records and securely stores measurements of the hypervisor image:

- The measurements are stored in Platform Configuration Registers (PCRs).
- The recorded state of these measurements is called a quote.
- These measurements can be used to detect changes for anything that can be loaded into memory.

vSphere uses TPM to provide remote attestation of the hypervisor image, based on hardware root of trust. The hypervisor image consists of the following elements:

- ESXi software (hypervisor) in VIB (package) format
- Third-party VIBs
- Third-party drivers

When TPM is enabled, ESXi measures the entire hypervisor stack when the system boots. The measurements include the VMkernel, kernel modules, drivers, native management applications that run on ESXi, and any boot-time configuration options. All VIBs that are installed on the system are measured.

By comparing this image to an image of the expected known good values, third-party solutions can leverage this feature to detect tampering of the hypervisor image.

https://t.me/learningnets

10-35 Remote Attestation Process

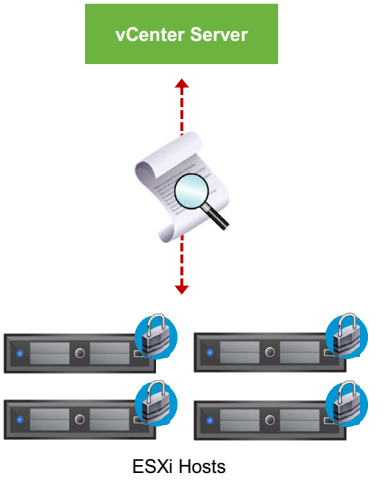
vCenter Server remotely verifies the measurements taken by the TPM 2.0 chip on the ESXi host.

The remote attestation process has the following high-level steps:

1. vCenter Server establishes the trustworthiness of the remote TPM on the ESXi host and creates an attestation key on it.
2. vCenter Server retrieves the attestation report from the ESXi host:

   The report contains the quote of PCRs which is signed by the attestation key.
3. vCenter Server verifies the authenticity of the signed quote from the ESXi host:

   If the signed quote is invalid, then remote attestation fails and the host is not trusted.



When an ESXi host is added to, rebooted from, or reconnected to vCenter Server, vCenter Server requests an attestation key from the host. Part of the attestation key creation process also involves the verification of the TPM hardware itself, to ensure that a known (and trusted) vendor has produced it.

vCenter Server requests that the host sends an attestation report. By checking that the information corresponds to a configuration it deems trusted, vCenter Server identifies the platform on a previously untrusted host.

vCenter Server verifies the authenticity of the signed quote, infers the software versions, and determines the trustworthiness of said software versions.

https://t.me/learningnets

10-36 Requirements for Using TPM 2.0

To use a TPM 2.0 chip, your vSphere environment must meet these requirements:

- vCenter Server 6.7
- ESXi 6.7 with TPM 2.0 chip installed and enabled in UEFI
- UEFI Secure Boot enabled

Review the TPM 2.0 chips certified by VMware by using the VMware Compatibility Guide at http://www.vmware.com/go/hcl.

https://t.me/learningnets

10-37 Activity: ESXi Secure Boot (1)

If UEFI Secure Boot is enabled on an ESXi host and an invalid VIB signature is found during the boot sequence, a purple screen appears stating that the UEFI Secure Boot failed.

- ❑ True
- ❑ False

https://t.me/learningnets

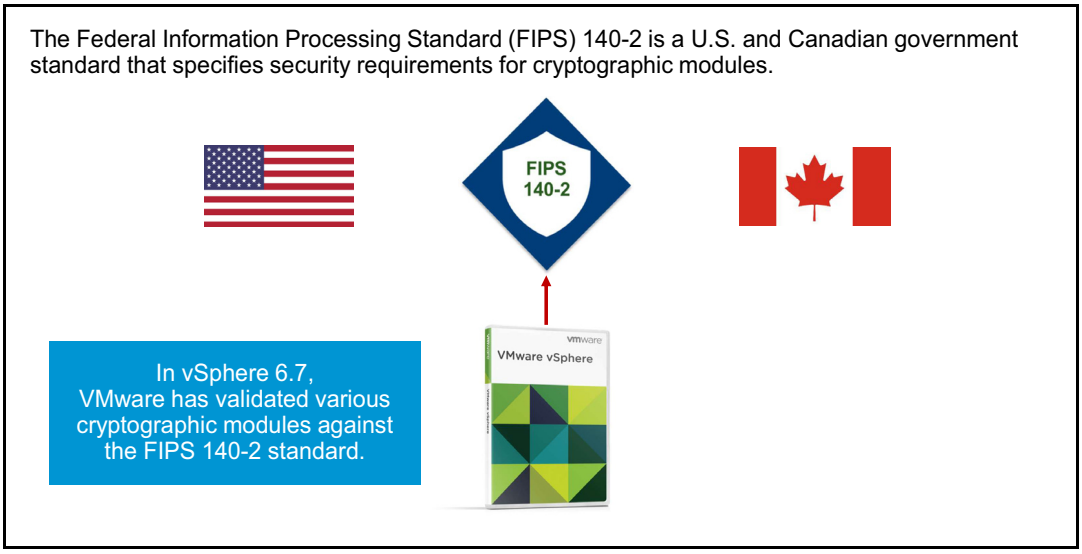
10-38 Activity: ESXi Secure Boot (2)

If UEFI Secure Boot is enabled on an ESXi host and an invalid VIB signature is found during the boot sequence, a purple screen appears stating that the UEFI Secure Boot failed.

- ✓ True
- ❑ False

https://t.me/learningnets

10-39 vSphere Compliance with FIPS 140-2 (1)



The Federal Information Processing Standards (FIPS) 140-2 standard specifies and validates the cryptographic and operational requirements for the modules within security systems that protect sensitive information. These modules employ NIST-Approved security functions such as cryptographic algorithms, key sizes, key management, and authentication techniques.

For more information about the FIPS security certification and validation for vSphere, see VMware Certifications at https://www.vmware.com/support/support-resources/certifications.html.

https://t.me/learningnets

10-40 vSphere Compliance with FIPS 140-2 (2)

By default, vSphere 6.7 performs cryptographic operations in FIPS 140-2 mode.

During the vSphere 6.7 installation, FIPS 140-2 validated cryptographic modules are installed and the FIPS 140-2 mode is enabled.

When FIPS 140-2 mode is enabled, the following functionality is present:

- vSphere uses the VMware VMkernel Cryptographic Module, which is a library that provides FIPS 140-2 approved cryptographic services.
- vSphere does not use cryptographic modules that are not compliant with FIPS 140-2.

FIPS 140-2 mode is not compatible with third-party tools that use old cryptographic modules that are disallowed.

https://t.me/learningnets

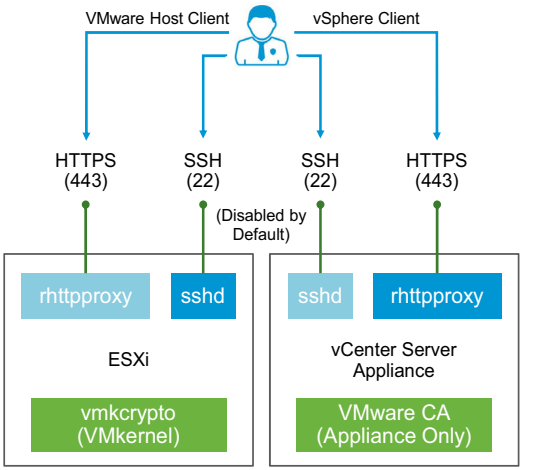
10-41 FIPS 140-2 Architecture

In vSphere 6.7, the management interfaces use FIPS-certified cryptography modules to secure communication:

- Any HTTPS connection, such as VMware Host Client and vSphere Client
- SSH to ESXi or vCenter Server

The following components also use FIPS-certified cryptography modules:

- The VMkernel vmkcrypto module
- VMware CA

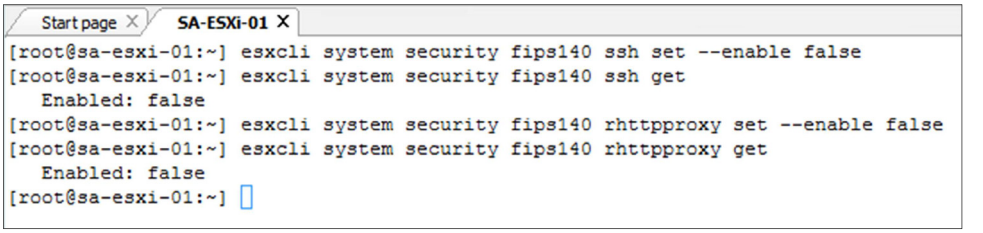


The vmkcrypto module contains the encryption algorithms used by ESXi.

https://t.me/learningnets

10-42 Enabling and Disabling FIPS 140-2 Mode on ESXi Hosts

You can enable and disable FIPS 140-2 mode on an ESXi host by using the `ESXCLI` command.

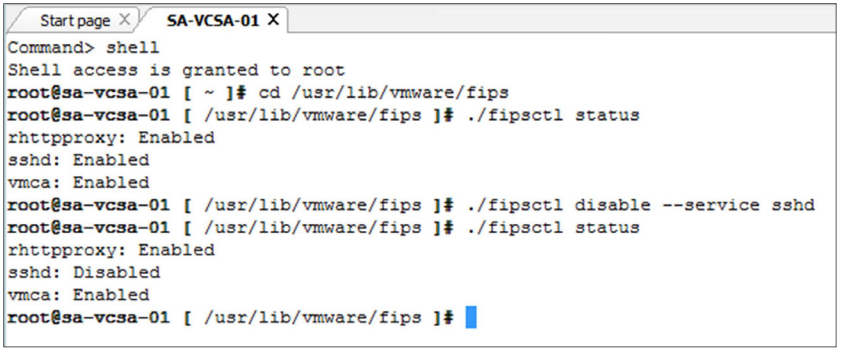


With the `esxcli system security fips140` command, you can view and modify the FIPS-140 settings for ssh and rhttpproxy. To enable FIPS-140 on ssh and rhttpproxy, use `set --enable true`. To disable FIPS-140, use `set --enable false`.

https://t.me/learningnets

10-43 Enabling and Disabling FIPS 140-2 Mode on vCenter Server Appliance

You can enable and disable FIPS 140-2 mode using the `fipsctl` command from the vCenter Server Appliance command line.



With the `fipsctl` command, you can view and modify the FIPS-140 settings for the rhttpproxy, sshd, and vmca services on vCenter Server Appliance.

To enable FIPS-140 on rhttpproxy, ssh, and vmca, use `enable --service service_name`.

To disable FIPS-140, use `disable --service service_name`.

https://t.me/learningnets

10-44 Virtual Machine Protection Overview

Employ the same security measures for a virtual machine as you would for an equivalent physical server:

- Keep all security measures current and patched.
- Install antivirus and antimalware software. Consider using this software with VMware vShield Endpoint™ for hypervisor-level protection.
- Disable unnecessary functions in virtual machines.
- Ensure that the copy and paste operations between guest operating systems and the remote console are disabled.
- Restrict users from running commands on a virtual machine.
- Use templates to deploy virtual machines.
- Keep the operating system patches up to date.

Using templates to deploy virtual machines removes the risk of misconfiguration. All virtual machines are created with a known baseline level of security.

https://t.me/learningnets

10-45 Activity: Security Technologies (1)

Match each description on the left with the corresponding security technology on the right.

1. Government standard that specifies security requirements for cryptographic modules	FIPS
2. Industry-wide standard for secure cryptoprocessors	TLS
3. Cryptographic protocol that provides secure communications over any transport	TPM

https://t.me/learningnets

10-46 Activity: Security Technologies (2)

Match each description on the left with the corresponding security technology on the right.



https://t.me/learningnets

10-47 Review of Learner Objectives

You should be able to meet the following objectives:

- Explain the importance of the vSphere Security Configuration Guide
- Discuss recommendations for vCenter Server security
- Summarize strategies to secure the vSphere management network
- Discuss recommendations for ESXi host security
- Plan for secure boot and TPM 2.0 support for an ESXi host
- Discuss general virtual machine and guest operating system protection

https://t.me/learningnets



https://t.me/learningnets

10-49 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Use VMware Certificate Authority and VMware Endpoint Certificate Store to configure vSphere security certificate management
- Describe solution users and their certificate requirements
- Describe common VMware CA modes
- Replace VMware CA certificates with enterprise or third-party commercial certificates
- Use vSphere Certificate Manager to manage vSphere certificates

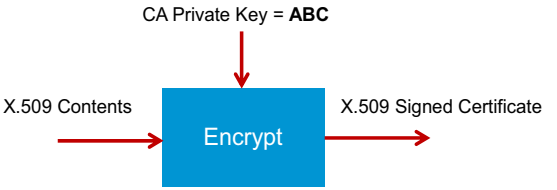
https://t.me/learningnets

10-50 Reviewing Public Key Certificates

Public key or digital certificates are electronic documents that are digitally signed by a trusted certificate source, such as a CA.

A certificate can be signed by a CA after you submit a certificate signing request (CSR), or it can be self-signed.

A CA uses its own private key to sign a digital certificate that validates the end user's public key in the X.509 contents.



Digital certificates that you use with entities that are external to your organization are typically signed by a certificate authority (CA).

You can use self-signed certificates as well. However, other parties are unlikely to trust your certificates because your signing certificate is not embedded in their browsers or systems.

You can use self-signed certificates for internal use, where you can add your public key to all of your internal systems so that they trust the self-signed certificates. You can also use self-signed certificates because you decide to rely on a public key infrastructure that uses a web of trust, such as Pretty Good Privacy.

https://t.me/learningnets

10-51 About CAs

CAs are entities that issue public-key certificates:

- Commercial CAs that charge to issue certificates
- Community CAs, such as CAcert.org, that provide free certificates

Some organizations act as their own CA and issue self-signed certificates.

To trust a certificate signed by a CA, clients such as web browsers must have the CA's root certificate installed:

- The root certificates of large commercial CAs are preinstalled in modern browsers.
- If you use self-signed certificates, you must install and manage your own root certificates.

CAs play an important role in public key infrastructure systems. When a Secure Sockets Layer (SSL) or Transport Layer Security (TLS) client connects to a server, the server sends its public key to the client to authenticate the server. However, the server cannot merely send its public key in plain text because it would be easily corruptible by an attacker who can inject themselves into the middle of the communication.

Instead, the server sends an X.509 certificate to the client. The X.509 certificate contains the server's name (the subject), its public key, and other information. X.509 certificates are signed by a trusted CA, which means that they are encrypted with the CA's private key.

The client trusts the CA because the client already has the CA's public key, which has either been preinstalled or manually installed by an administrator. For example, browsers such as Safari, Firefox, Internet Explorer, and Chrome have public keys of the most common CAs preinstalled.

SSL or TLS usage is not limited to websites and browsers, although these are common and well-known uses. Any client or server software solution that requires secure authentication of the parties can use SSL or TLS in the manner described here.

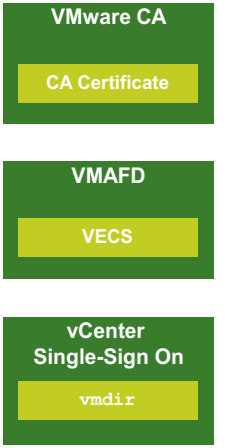
https://t.me/learningnets

10-52 vSphere Certificate Management

vSphere 6.x has a certificate management architecture that simplifies managing your vSphere certificates.

Several certificate management components exist:

- VMware CA issues certificates for vSphere components to use.
- VMware Authentication Framework Daemon (VMAFD):
  - VMware Endpoint Certificate Store (VECS) is a local (client-side) repository for certificates, private keys, and other typical keystore data.
- VMware vCenter® Single Sign-On and VMware Directory Service use certificates.



In vSphere 6.x, VMware CA provisions each ESXi host and each vCenter Server service with certificates that are signed by VMware CA by default. The VMware Authentication Framework Daemon implements the VMware Endpoint Certificate Store (VECS) and other authentication functions. The VMware Directory Service (vmdir) handles SAML certificate management for authentication in conjunction with vCenter Single Sign- On.

This simple analogy shows how these components relate to one another: A certificate is like a driver's license in that it is a digital card that authenticates your identity. In this analogy, VECS is the wallet in which you store your driver's license. VMware CA is the Department of Motor Vehicles that issues you a driver's license. Finally, vCenter Single Sign-On is the police officer or other authority who checks your driver's license when required to verify your identity. The police officer, exactly like vCenter Single Sign-On and vmdir, also has a driver's license for those times when the officer's identity must be authenticated.

https://t.me/learningnets

10-53 About VMware CA

VMware CA issues and validates certificates.

The main purpose of VMware CA is to issue certificates so that vSphere components can establish trust relationships.

The main consumers of certificates from VMware CA are services that run as part of vCenter Server.

Services request certificates to establish their own identity and to establish the identity of other vSphere entities.

By default, VMware CA acts as a root CA. It issues and validates certificates for vSphere components such as ESXi hosts and solutions. VMware CA can handle all certificate management for you. VMware CA provisions vCenter Server components and ESXi hosts with certificates that use VMware CA as the root certificate authority. If you upgrade to vSphere 6.x from an earlier version of vSphere, all self-signed certificates are replaced with certificates that are signed by VMware CA.

https://t.me/learningnets

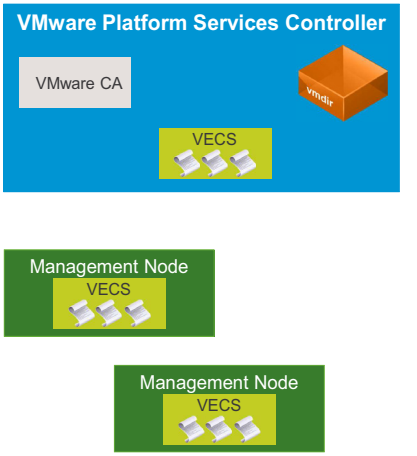
10-54 About the VMware Endpoint Certificate Store

VECS serves as a local, client-side repository for the following items:

- Certificates
- Private keys
- Other certificate information that can be stored in a keystore

VECS contains multiple keystores.

Even if you decide not to use VMware CA, you must use VECS to store all vCenter Server certificates, keys, and other certificate information.



ESXi certificates are stored locally on each host, not in VECS.

VECS runs as part of the VMware Authentication Framework Daemon. VECS runs on every embedded deployment, Platform Services Controller node, and management node. It also holds the keystores that contain the certificates and keys.

The vCenter Single Sign-On service stores the token-signing certificate and its SSL certificate on disk. You can change the token-signing certificate from vSphere Client.

https://t.me/learningnets

10-55 vSphere Certificate Types

vSphere uses the following types of certificates:

- CA certificates
- Solution user certificates
- Machine certificates (SSL):
  - ESXi
  - vCenter Server (embedded)
  - vCenter Server (Management node, with an external Platform Services Controller instance)
  - Platform Services Controller

vSphere components use SSL to communicate securely with one another and with ESXi hosts. SSL communications ensure data confidentiality and integrity. Data is protected and cannot be modified in transit without detection.

vCenter Server services, such as vSphere Client, also use their certificates for their initial authentication to vCenter Single Sign-On. vCenter Single Sign-On provisions each component with a SAML token that the component uses for authentication going forward.

VMware products use standard X.509 version 3 (X.509v3) certificates to encrypt session information that is sent over SSL between components.

https://t.me/learningnets

10-56 vSphere CA Certificates

vSphere CA certificates are trusted root certificates that VMware CA uses to issue other certificates:

- vCenter Single Sign-On certificate
- vmdir certificate

vSphere CA certificates are stored in the trusted root store.

vCenter Single Sign-On certificates include the vCenter Single Sign-On signing certificate and the vmdir certificate. As a rule, you do not need to make any changes to these certificates, but you can replace them in special situations.

The vCenter Single Sign-On signing certificate is used to sign the SAML tokens that vCenter Single Sign-On issues so that clients of vCenter Single Sign-On can verify that the SAML token comes from a trusted source. You can replace this certificate from vSphere Client.

You typically do not need to change or replace the vmdir's certificate. However, if you decide to use a new VMware CA root certificate, and you unpublish the VMware CA root certificate that was used when you provisioned your environment, you must replace the machine SSL certificates, solution user certificates, and certificates for some internal services, including the two vCenter Single Sign-On certificates.

See the vSphere security documentation for details of how to use the command-line utilities to replace the vmdir certificate.

https://t.me/learningnets

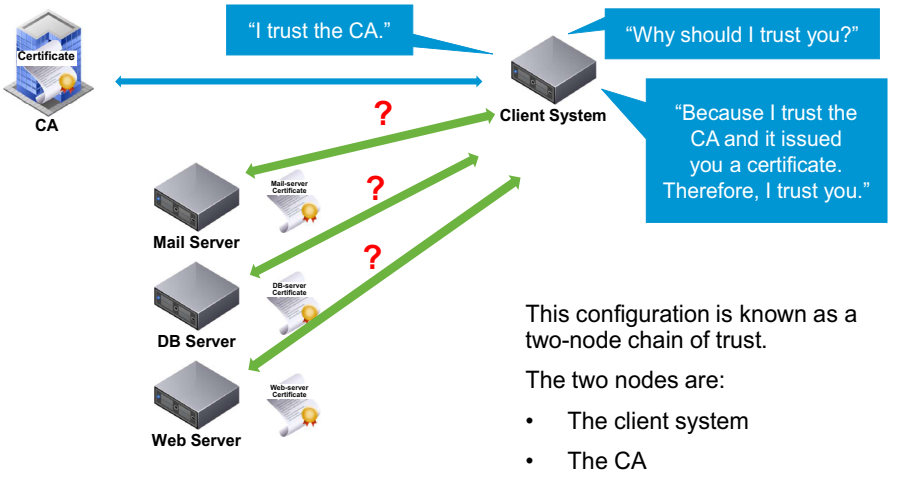
10-57 Chain of Trust (1)

If you trust the CA, then you implicitly trust all of the certificates issued by that CA.

- In order to trust a certificate, you must trust some part of the chain of trust. One of the following must be true:
  - You must say that you explicitly trust the certificate itself.
  - You must say that you explicitly trust the CA that issued it.
- In a self-signed certificate, the issuer and the user are the same system.
  - To use self-signed certificates, every user (client) system must install and explicitly trust every self-signed certificate that is in use in the entire network.
  - Every time a new service is brought on line, all clients must individually install and trust each and every self-signed certificate in the network.
- An in-house or commercial CA eliminates the requirement of each client system installing each and every self-signed certificate as long as:
  - All client systems trust the CA.
  - All certificates come from that trusted CA.

https://t.me/learningnets

10-58 Chain of Trust (2)



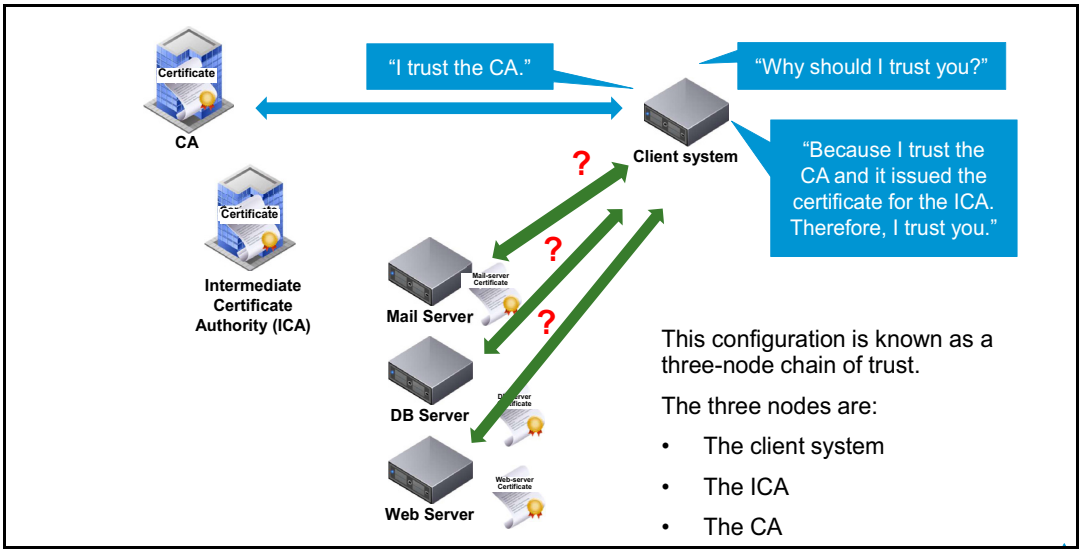
This configuration is known as a two-node chain of trust.

The two nodes are:

- The client system
- The CA

https://t.me/learningnets

10-59 Chain of Trust (3)



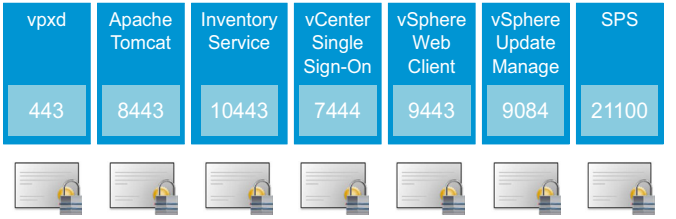
https://t.me/learningnets

10-60 Solution Endpoints Before vSphere 6.0

In vSphere 5.x and earlier, each service listened on a well-defined port.

Each service required its own certificate because authentication was based on SSL thumbprint trust and the thumbprint had to be unique.

The graphic shows the handful of services that existed prior to vSphere 6.0.



In vSphere 5.x and earlier, each vSphere service required its own service certificate. Unlike in vSphere 6.x, only a handful of services existed, including vCenter Server, vCenter Single Sign-On, vSphere Web Client, and the profile-driven storage service. Each certificate had to be unique so that clients could be sure that they were communicating with the desired service. This mechanism relied on a unique certificate thumbprint, which is the computed hash value of the entire certificate.

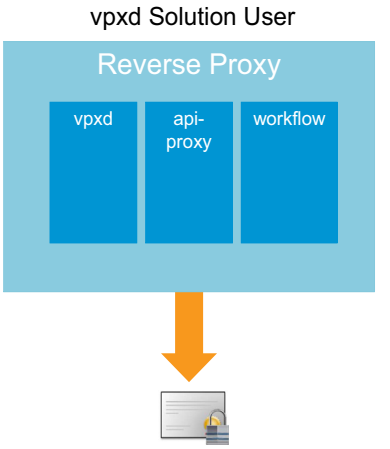
https://t.me/learningnets

10-61 vSphere 6.x Reverse Proxy

In vSphere 6.x, over 25 services require authentication with certificates.

Individual service endpoints are combined. They are serviced by a single endpoint with a single certificate.

The endpoint certificate is assigned to a vSphere solution user. The endpoint certificate is reached through a reverse proxy.



The individual service endpoints that were used in vSphere 5.x are replaced in vSphere 6.x by a reverse proxy that routes traffic to the appropriate service, based on the type of the incoming request. The type is determined by the namespace. Because only a single endpoint exists, only a single certificate is needed for the server to authenticate to the client.

On the slide, the vpxd solution user manages three vSphere services, including vCenter Server (vpxd).

https://t.me/learningnets

10-62 vSphere Solution User Certificates

Solution user certificates are used by the vSphere solution users to authenticate to vCenter Server.

The following default solution users are available:

- vpxd: vCenter Server
- vpxd-extensions: For example, vSphere Auto Deploy.
- vsphere-webclient: vSphere Web Client, performance charts, and so on.
- Machine: Component manager, logging service, license server, and so on.

Instead of 25 or more individual service certificates, vSphere 6.x uses four solution user certificates.

A solution user encapsulates one or more vCenter Server services and uses its solution user certificate to authenticate to vCenter Single Sign-On through a SAML token exchange.

A solution user presents the certificate to vCenter Single Sign-On when it first has to authenticate, after a reboot, and after a timeout has elapsed. The default timeout is 2592000 seconds (30 days) and is configurable from the Single Sign-On Configuration panel in vSphere Client.

For example, the vpxd solution user presents its certificate to vCenter Single Sign-On when it connects to vCenter Single Sign-On. The vpxd solution user receives a SAML token from vCenter Single Sign-On and can then use that token to authenticate to other solution users and services.

The following solution user certificate stores are included in VECS on each management node and each embedded deployment:

- vpxd: This is the vCenter Server service daemon (vpxd) store. vpxd uses the solution user certificate that is stored in this store to authenticate to vCenter Single Sign-On.
- vpxd-extensions: This is the vCenter Server extensions store. It includes the vSphere Auto Deploy service and other services that are not part of other solution users.
- vsphere-webclient: This is the vSphere Web Client store. It also includes some additional services such as the performance chart service.
- machine: Tis is used by component manager, license server, and the logging service.

  The machine solution user certificate has nothing to do with the machine SSL certificate. The machine solution user certificate is used for the SAML token exchange. The machine SSL certificate is used for secure SSL connections for a machine.

https://t.me/learningnets

10-63 vSphere Machine Certificates

ESXi certificates are generated by VMware CA and stored locally on each ESXi host.

Other machine certificates are stored in the MACHINE_SSL_CERT keystore in VECS.

Machine SSL certificates are used by the following entities:

- The reverse proxy service on every vSphere node
- The vmdir service
- The vCenter Server daemon (vpxd)

ESXi certificates are stored locally on each host in the `/etc/vmware/ssl` directory. ESXi certificates are provisioned by VMware CA by default, but you can use custom certificates instead. ESXi certificates are provisioned when the host is first added to vCenter Server and when the host reconnects to vCenter Server.

The machine SSL certificate for each node is used to create an SSL socket on the server side to which SSL clients connect. The certificate is used for server verification and for secure communication such as HTTPS or LDAP over SSL. Every node (embedded deployment, management node, or Platform Services Controller) has its own machine SSL certificate. All services running on that node use this machine SSL certificate to expose their SSL endpoints.

https://t.me/learningnets

10-64 vSphere Certificate Usage Summary

vSphere certificates are provisioned by VMware CA or during a product installation. They are stored in various certificate stores, depending on the certificate's use.

Certificate	Provisioned By	Stored
ESXi certificates	VMware CA (default)	Locally on ESXi
Machine SSL certificates	VMware CA (default)	VECS
Solution user certificates	VMware CA (default)	VECS
vCenter Single Sign-On SSL signing certificate	Provisioned during installation	Locally on node
vmdir SSL certificate	Provisioned during installation	Locally on node

https://t.me/learningnets

10-65 Activity: VMware Endpoint Certificate Store (1)

The VECS is a repository that exists on every management node and that stores vCenter Server certificates and ESXi host certificates.

- ❑ True
- ❑ False

https://t.me/learningnets

10-66 Activity: VMware Endpoint Certificate Store (2)

The VECS is a repository that exists on every management node and that stores vCenter Server certificates and ESXi host certificates.

- ❑ True
- ✓ False

The VECS is a repository that exists on every management node that stores vCenter Server certificates, but not ESXi host certificates. ESXi certificates are stored locally on each host.

https://t.me/learningnets

10-67 VMware CA Modes

VMware CA has the following modes:

- Root CA (default): VMware CA acts as a root CA. VMware CA uses a self-signed certificate to create and issue certificates to vCenter Server, ESXi, and other components that need them.
- Subordinate CA: VMware CA is subordinate to an existing enterprise or external third-party CA.
- Custom: You can choose not to use VMware CA and replace the VMware CA-signed certificates with custom certificates.

By default, the VMware CA is configured to act as a root CA that can use its self-signed root certificate to issue and sign other certificates. However, you can also configure VMware CA as a subordinate CA. In subordinate CA mode, VMware CA is an intermediate CA in the chain of trust that continues up to an enterprise CA or an external third-party CA.

If you do not want to use VMware CA, you can bypass VMware CA altogether and install custom certificates on all your systems. Use subordinate CA mode if you want to issue and install your own certificates. You must issue a certificate for every component, in this situation. All certificate management is your responsibility. All of your custom certificates (except for host certificates) must be installed into VECS.

https://t.me/learningnets

10-68 VMware CA Mode: Root CA

When VMware CA is in root CA mode:

- This mode is the default mode.
- This mode uses a self-signed root certificate.
- The root certificate is a CA certificate that is capable of issuing other certificates.
- All solution user certificates and machine certificates are created from this root CA certificate.

As a normal practice, a CA has a self-signed certificate at the root, especially if the certificate is the first one created in a new domain.

VMware CA uses a self-signed root certificate. It issues certificates to vCenter Server, ESXi, and so on, and manages these certificates. These certificates have a chain of trust that stops at the VMware CA root certificate. VMware CA is not a general-purpose CA and its use is limited to VMware components.

https://t.me/learningnets

10-69 VMware CA Mode: Subordinate CA

When VMware CA is in subordinate CA mode:

- You replace the default self-signed CA certificate that was created during installation.
- You typically replace all issued default certificates after implementation.

This mode requires a CSR issued from VMware CA to be used in an enterprise or commercial CA to generate a new issuing certificate.

An enterprise CA signs the CSR that VMware CA generates. You then configure VMware CA to use this certificate and keys.

In subordinate CA mode, you replace the VMware CA root certificate with a third-party CA-signed certificate that includes VMware CA in the certificate chain. The replacement certificate can be signed by your enterprise CA or by a third-party, external CA.

Going forward, all certificates that VMware CA generates include the full chain. You can replace existing certificates with newly-generated certificates, although such action is not required. The original certificate is still valid if it has not expired or been revoked, and so you can continue to use the certificates that were signed with the original certificate.

The subordinate CA mode combines the security of a third-party, CA-signed certificate with the convenience of the automated certificate management in VMware CA.

https://t.me/learningnets

10-70 Certificate Replacement Options

For vCenter Server solutions, you must replace certificates in the following situations:

- In root CA mode, when you want to replace the VMware CA certificate with a new certificate
- In subordinate CA mode, when you replace the VMware CA certificate with a CA certificate issued by your enterprise or a third-party CA
- In custom mode, when you replace all of your certificates with enterprise or third-party certificates

With the introduction of VMware CA in vSphere 6.0, managing your vSphere certificates is much easier. The options listed on the slide are the three most common situations.

https://t.me/learningnets

10-71 Certificate Manager

You can use vSphere Certificate Manager to perform certificate tasks.

Option	Usage
1	Replace machine SSL certificate with custom CA certificate.
2	Replace VMware CA root certificate with custom CA certificate and replace all certificates.
3	Replace machine SSL certificate with VMware CA certificate.
4	Regenerate a new VMware CA root certificate and replace all certificates.
5	Replace solution user certificates with custom certificates.
6	Replace solution user certificates with VMware CA certificates.
7	Revert last-performed operation by republishing old certificates.
8	Reset all certificates.

You can use vSphere Certificate Manager to replace the certificates in your vSphere environment and to reset all of your vSphere certificates.

To start vSphere Certificate Manager, log in to your vCenter Server system and run the following commands at the command prompt:

- Windows vCenter Server: `C:\Program Files\VMware\vCenter Server\vmcad\certificate-manager`
- VMware vCenter Server Appliance: `/usr/lib/vmware-vmca/bin/certificate-manager`

For more information on the Certificate Manager options and workflows, see *Platform Services Controller Administration* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-platform-services-controller-administration-guide.pdf.

https://t.me/learningnets

10-72 Replacing the VMware CA Certificate in Root CA Mode

To replace the VMware CA certificate in root CA mode, follow these high-level steps.

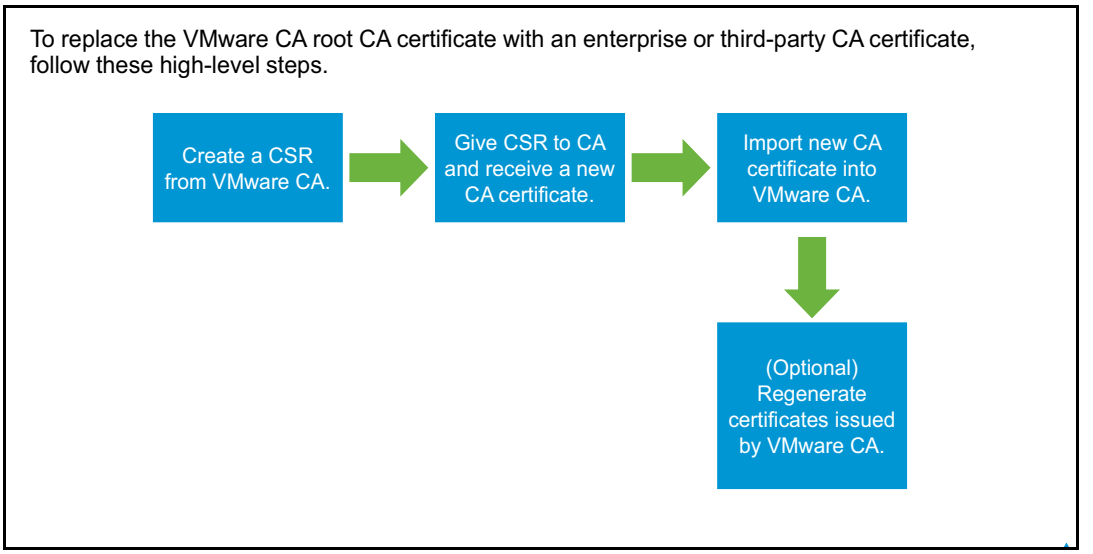


You are not required to replace the certificates that were issued and signed by the previous VMware CA root CA certificate. When you replace the VMware CA root CA certificate, it does not automatically invalidate the previous certificate. It also does not remove the old certificate, so the old certificate continues to be valid in your vSphere environment.

However, if you have a reason to replace the root CA certificate, then you typically also want to replace all of your solution user certificates, your ESXi certificates, and so on, so that they are signed by the current VMware CA root CA certificate. You can replace all of these certificates at once by using option 4 in the vSphere Certificate Manager tool.

https://t.me/learningnets

10-73 Replacing the VMware CA Certificate with an Enterprise CA Certificate



To configure your VMware CA as a subordinate CA, you must import an enterprise or third-party CA certificate into VMware CA. You must generate a certificate signing request that you can present to your enterprise or third-party CA so that the CA can create an appropriate certificate for you.

See *Platform Services Controller Administration* at https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-platform-services-controller-administration-guide.pdf.

https://t.me/learningnets

10-74 Replacing All VMware CA Certificates with Custom Certificates

You can replace all your certificates issued by VMware CA with custom certificates if your site's policies require it.

When you replace VMware CA with custom certificates:

- VMware CA is not in your certificate chain.
- You are responsible for replacing all certificates yourself and for keeping track of certificate expiration.
- VECS is always used, even in this case.

To use custom certificates, send CSRs to your certificate provider and request the following items:

- A root certificate.
- A machine SSL certificate for each machine. The CSR for machine certificates must include the host name and the IP address for the machine.
- Four solution user certificates for each embedded system or management node. Solution user certificates should not include the IP address, host name, or email address. Each certificate should have the name of the solution user.

VMware does not generally recommend that you replace certificates issued by VMware CA with custom enterprise or third-party certificates. However, your own site security policies might dictate that you use only a specific CA for all of your site's systems.

When you replace the VMware CA certificates with custom certificates, you increase your management tasks, which can consequently increase the risk of errors that can lead to security vulnerabilities. You are responsible for all certificate management tasks. And, because you do not use VMware CA, you cannot rely on VMware CA to manage your certificates for you.

https://t.me/learningnets

10-75 ESXi Certificate Replacement Options

Several options exist for ESXi certificate replacement:

- Replace with certificates issued by VMware CA:
  - VMware CA provisions the host certificates.
  - Host certificates include the full chain.
- Use enterprise or third-party certificates:
  - This certificate replacement option enables you to use your own custom certificates.
  - This certificate replacement option requires you to have an enterprise CA.
- Use thumbprint mode:
  - Thumbprint mode is also called legacy mode.
  - Thumbprint mode can be used to retain ESXi 5.5 certificates during an upgrade.

In a vSphere 6.x environment, you should use thumbprint mode sparingly and only temporarily during your upgrade period from older ESXi versions to ESXi 6.7.

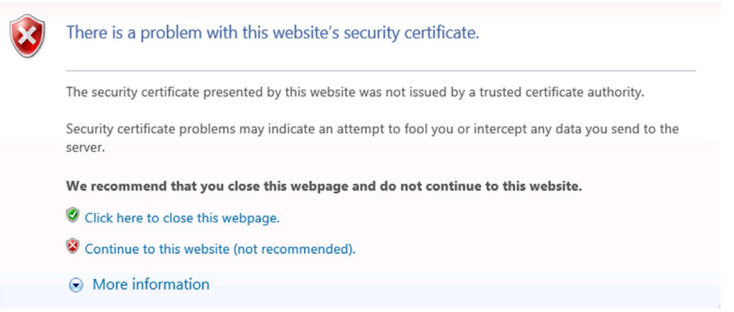
https://t.me/learningnets

10-76 Installing the VMware CA Root Certificate in Your Browser

Your browser does not recognize or trust the VMware CA root certificate.

To eliminate certificate warning messages every time you connect to vCenter Server, you must add the VMware CA root certificate to your browser's list of trusted certificates.

See your browser's vendor documentation.



https://t.me/learningnets

10-77 VMware CA Availability

If VMware CA is not available, the following operations do not work:

- Adding new ESXi hosts
- Refreshing certificates on any ESXi host

In some cases, certificate refreshes on certificate expiry might fail.

If VMware CA is unavailable for short periods of time, operations are not usually disrupted.

VMware CA is included in each Platform Services Controller, whether the Platform Services Controller is a separate node or embedded with vCenter Server. When VMware CA is not available, you lose the ability to manage your certificates. However, issued certificates are still valid. Your ongoing vSphere services or functions should not suffer disruption.

https://t.me/learningnets

10-78 Lab 13: Working with Certificates

Generate and replace a vCenter Server certificate

1. Examine vSphere Certificates
2. Create a Windows 2012 Certificate Authority Template for vSphere
3. Create a Certificate Signing Request
4. Download the CSR to the Student Desktop
5. Request a Signed Custom Certificate
6. Replace a Machine Certificate with the New Custom Certificate

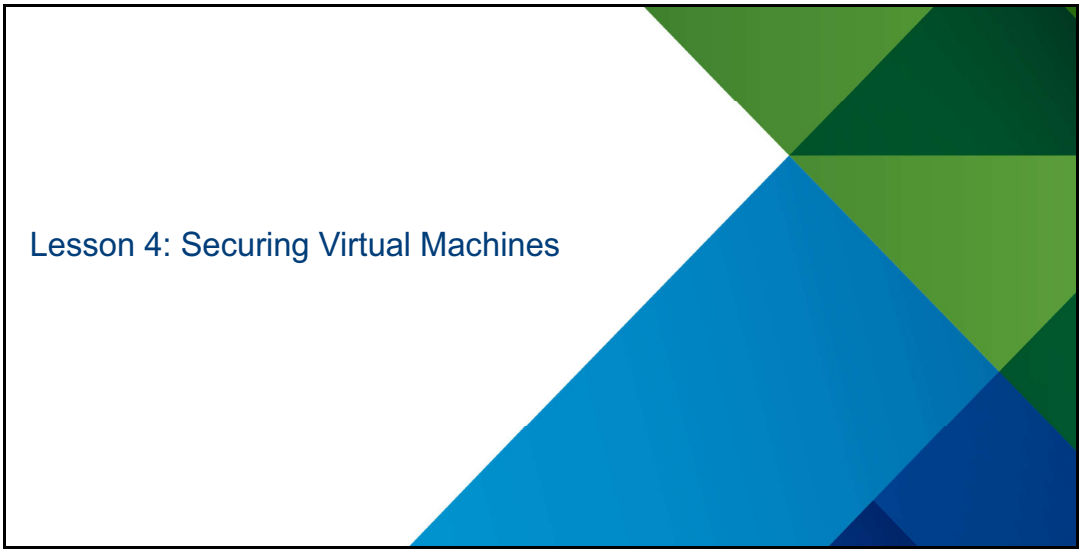
https://t.me/learningnets

10-79 Review of Learner Objectives

You should be able to meet the following objectives:

- Use VMware Certificate Authority and VMware Endpoint Certificate Store to configure vSphere security certificate management
- Describe solution users and their certificate requirements
- Describe common VMware CA modes
- Replace VMware CA certificates with enterprise or third-party commercial certificates
- Use vSphere Certificate Manager to manage vSphere certificates

https://t.me/learningnets



https://t.me/learningnets

10-81 Learner Objectives

By the end of this lesson, you should be able to meet the following objectives:

- Set up encryption in your vSphere environment
- Encrypt virtual machines
- Manage encrypted virtual machines
- Encrypt core dumps
- Enable encrypted vSphere vMotion
- Describe support for virtual machine security features:
  - UEFI secure boot
  - vTPM
  - Virtualization-based security

https://t.me/learningnets

10-82 Business Use Case: Securing Virtual Machines

A common requirement in a company's security policy is to protect virtual machine data from administrative users:

- Problem:
  - A large company has several vSphere and storage administrators.
  - The company must protect its confidential data.
  - The company must reduce the risk of someone easily downloading a virtual machine disk (VMDK) file or the entire virtual machine to a removable storage device and leaving the company with this data.
- Solution:
  - With virtual machine encryption, the company can secure confidential data on a VMDK file so that this data is unreadable without the digital key that was used to encrypt the disk.
  - The key is not readable in any file. It is secured by an additional layer of encryption.
  - The company grants only a limited number of people access to the key.

The digital key used to encrypt a virtual machine is secured in an additional layer of encryption, much like taking the key needed to open a safe and placing it in another safe.

You allow regular vSphere administrators to continue with their daily activities without change, but you grant cryptographic privileges to only a subset of these administrators. Thus, access to virtual machine files does not imply that storage administrators or vSphere administrators can walk away with the virtual machine (VMDK) file and read its contents offsite.

https://t.me/learningnets

10-83 About VM Encryption

VM encryption provides the following functionality.

Function	Description
Encryption	• Protection of VM disks and metadata files, such as NVRAM and VSWP • Multilayer key protection

Orchestration	• Simplified deployment using storage policies
	• Agnostic:
	  • Storage-agnostic
	  • Guest operating system-agnostic

Key control	• Key management provided by key servers • Use of Key Management Interoperability Protocol (KMIP) standard • Nonpersistence of keys for added security
Access control	• Role for administrators without cryptography permissions • Cryptographic tasks authorized only to administrators with appropriate permissions

Introduced in vSphere 6.5, virtual machine encryption protects not only the data on disk but also the swap file, as well as any guest-specific information that might be contained in the VMX or NVRAM files. Virtual machine encryption also provides a multilayer key protection mechanism that makes breaking the security of the virtual machine almost impossible.

Virtual machine encryption makes it simple to configure encryption by using the orchestration capabilities of vCenter Server storage policies. By using storage policies, you can apply encryption to any type of virtual machine, regardless of the guest operating system or the storage on which it resides.

By using external, enterprise-grade key servers, exposure to risk is limited by ensuring that vSphere does not need to manage keys or keep key data on disk anywhere in the vSphere environment. Using the industry-standard Key Management Interoperability Protocol (KMIP), vSphere can be integrated with various third-party key management server (KMS) vendors and products.

You can limit access to cryptographic operations by enforcing roles and permissions for users. Only users with the required permissions can perform cryptographic tasks such as encrypting or decrypting a virtual machine.

https://t.me/learningnets

10-84 Advantages of VM Encryption

VM encryption has several advantages over similar offerings on the market:

- It does not require in-guest agents.
- Its methodology is uniform across all guest operating systems.
- It protects all virtual machine data, including swap files, not only virtual disk files.
- Fine-grained key control: Virtual machines and disks can use different keys.
- It offers easy orchestration through virtual machine storage policies.

https://t.me/learningnets

10-85 Virtual Machine Encryption Architecture

The VM encryption architecture consists of the following components:

- Key management server (KMS):
  - Is an external server that provides keys to services.
  - Is provided by a third-party security vendor.
- vCenter Server:
  - Communicates with a KMS and requests keys on behalf of the ESXi hosts.
- ESXi hosts:
  - Run encrypted VMs.

https://t.me/learningnets

10-86 About the KMS

To prepare the environment for virtual machine encryption, you must set up the KMS.

The KMS has the following characteristics:

- Must be compatible with KMIP 1.1:
  - KMIP is an industry-standard language for the management of security keys.
- Provides key management service for KMIP clients, such as vCenter Server
- Can be configured with a KMIP proxy server
- Is accessed over IP using IPv4 only, IPv6 only, or mixed mode (IPv6 and IPv4)

For vSphere, the KMS must be a server that communicates using the KMIP protocol. vCenter Server implements a KMIP client that can issue KMIP-formatted commands to request keys using specific identifiers. A KMIP server can then return the key values to vCenter Server using a secure channel.

Having a KMS means that all of the management tasks relating to keys can be centrally managed on a single server or server cluster. Multiple vCenter Server instances can access the same KMS cluster and can be granted access to the same keys. Having a centrally-managed key store reduces the attack surface for your key management. Instead of keys existing in multiple locations, which you then have to secure, securing a single or limited number of servers is far easier and safer.

KMIP proxy servers can be established to provide the key management capability of the KMS to a remote location. KMIP proxy servers enable you to keep your critical key server safe in your central data center while still allowing remote offices to use its services without compromising security.

Finally, using the KMIP requires only an IP-capable network, which is readily available in most data centers.

https://t.me/learningnets

10-87 Key Management Clusters

The KMS has a critical role in protecting virtual machines.

Protect the KMS by creating a key management cluster:

- A key management cluster is a group of KMIP-based key management servers that replicate keys to one another.
- A cluster must contain key management servers from one vendor.
- A cluster prevents single points of failure.
- One default cluster is allowed for each vCenter Server instance.
- The cluster is identified in vCenter Server with a Cluster Name value.

To protect against the potential failure of a KMS (also called a KMIP server), KMIP enables the formation of KMIP clusters, which provide some failover protection. vCenter Server keeps a list of key management servers in the cluster. If the first KMS on the list is not available in the cluster, then vCenter Server tries to communicate with the next KMS on the list, and so on.

To form a KMIP cluster, the KMIP servers must be from a single vendor. See the vendor documentation for steps on how to create a KMIP cluster.

After the keys are replicated, a client can request the key from any of the servers participating in the cluster. So if one server goes down, another server can provide the key.

While multiple KMIP clusters can be added to vCenter Server, one cluster must be identified as the default cluster. The default cluster is the cluster from which vCenter Server requests new keys. Keys can be retrieved from nondefault clusters if the cluster is specifically identified. Clusters of KMIP servers are identified using a user-defined friendly name, populated in vSphere Client.

https://t.me/learningnets

10-88 Role of vCenter Server in Virtual Machine Encryption

vCenter Server plays an important role in virtual machine encryption:

- Stores KMS credential information
- Manages keys used for encryption:
  - Identifies keys by UUID
  - Retrieves keys from the KMS
  - Pushes keys to ESXi hosts when required
- Manages permissions
- Implements and manages storage policies
- Records events for auditing purposes

vCenter Server itself cannot be encrypted.

vCenter Server stores the KMS login information, and it is the only vSphere component that can communicate with the KMS and retrieve keys. An ESXi host itself cannot request keys from the KMS. vCenter Server performs key management tasks on behalf of the ESXi hosts.

vCenter Server provides a mechanism whereby keys can be identified by a unique identifier, which can be stored in a virtual machine's VMX file or VMDK file. vCenter Server can use these identifiers to request keys from the KMS server when required and then push these keys to the ESXi hosts.

vCenter Server provides the framework to manage permissions for cryptographic tasks. You can restrict access to critical operations to a subset of your administrators, while ensuring that your other administrators can continue with their daily tasks.

vCenter Server manages storage policies that define whether or not a virtual machine is to be encrypted.

vCenter Server records events for auditing purposes. Event information includes the identity of the user who initiated the event.

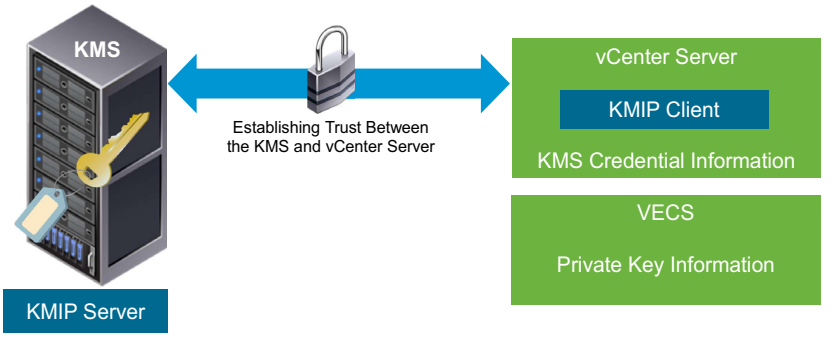
https://t.me/learningnets

10-89 KMS and vCenter Server Communication

To secure communication between the KMIP server (KMS) and the KMIP client (vCenter Server), trust must be established.

vCenter Server stores the private key information in the VECS.

vCenter Server 6.7 uses TLS 1.2 when communicating with the KMS.



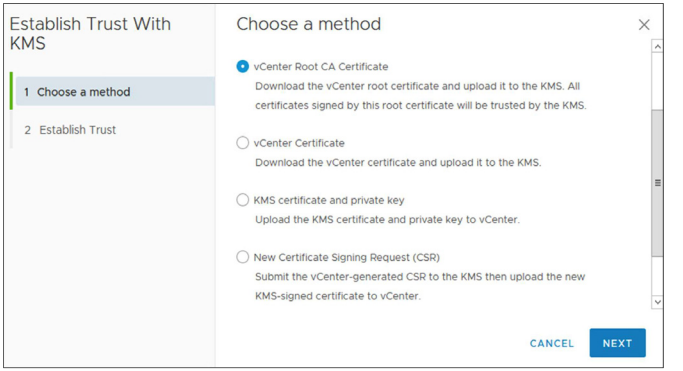
After trust is established between the KMS (KMIP server) and the vCenter Server (KMIP client), vCenter Server stores the KMS credential information. The VECS stores the SSL private key information.

https://t.me/learningnets

10-90 KMIP Client Certificate

The type of certificate used by the KMIP client (vCenter Server) depends on the KMS vendor:

- Some vendors accept self-generated certificates.
- Some vendors require that the KMS generate a trusted certificate for the client, which needs to be downloaded by the client.
- Other vendors accept only a client certificate provided by the KMS itself. Check with the KMS vendor for requirements.



Different KMS vendors have varying requirements for the format of the client certificate that is used by vCenter Server.

Some vendors allow the use of the client's own certificate, such as a self-signed or CA-signed certificate. Other vendors require that the certificate used by the client be signed by the KMS server.

vSphere Client enables you to provide certificate details based on the security requirements of the different types of implementations.

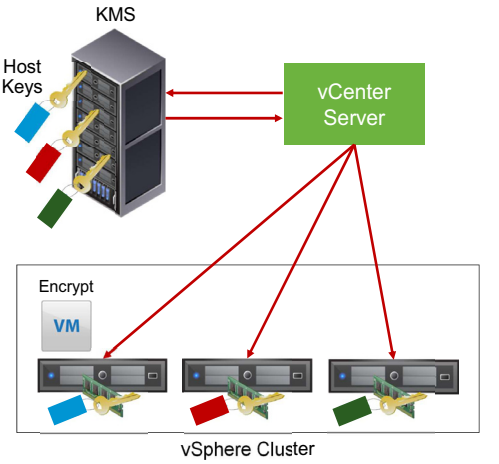
https://t.me/learningnets

10-91 Making an ESXi Host Cryptographically Safe

An ESXi host must be cryptographically safe before it can manage encrypted VMs.

The following sequence of events puts an ESXi host in a crypto-safe state:

1. The user performs an encryption task, for example, create an encrypted VM.
2. vCenter Server requests a separate key from the KMS for each host in the cluster. These keys are known as host keys.
3. vCenter Server retrieves the host keys from the KMS and pushes a host key to each host in the cluster.
4. The host installs its host key in its memory key cache and enables core dump encryption. The host key is used to encrypt core dumps.



Users with the required privileges can perform cryptographic operations. These operations include creating encrypted virtual machines, encrypting existing virtual machines, and decrypting encrypted virtual machines.

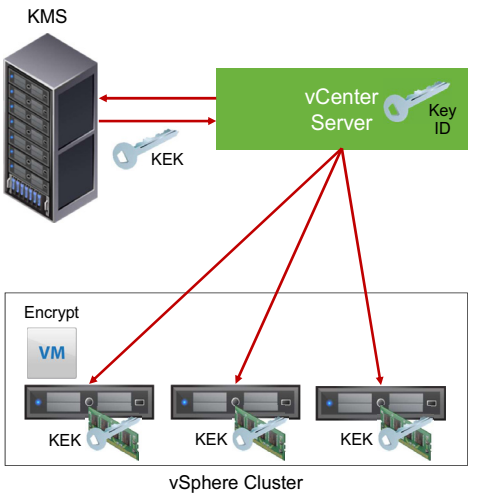
When host encryption mode is enabled on one host in a cluster, it is also enabled for all other hosts in the cluster. Host encryption mode cannot be disabled while any hosts in the cluster are managing encrypted virtual machines.

https://t.me/learningnets

10-92 VM Encryption Process Flow (1)

Now that the host is cryptographically safe, it can receive keys for the VM:

- vCenter Server requests a VM key from the KMS.
- The KMS gives a VM key to vCenter Server. This key is called the key encryption key (KEK).
- vCenter Server passes the KEK directly to the ESXi host. If the host belongs to a cluster, vCenter Server sends the KEK to each host in the cluster. The KEK is stored in the host's memory.
- The KEK itself is never stored on the vCenter Server system, only the key ID is.



When a cryptographic operation takes place, vCenter Server must determine whether or not a VM key needs to be pushed to the ESXi hosts.

vCenter Server retrieves the key encryption key (KEK) from the KMS but the KEK is never stored anywhere on vCenter Server, not even in its memory. The KEK remains in vCenter Server memory only while the key is in transit to the ESXi host. vCenter Server passes the KEK to the ESXi host that is performing the cryptographic operation.

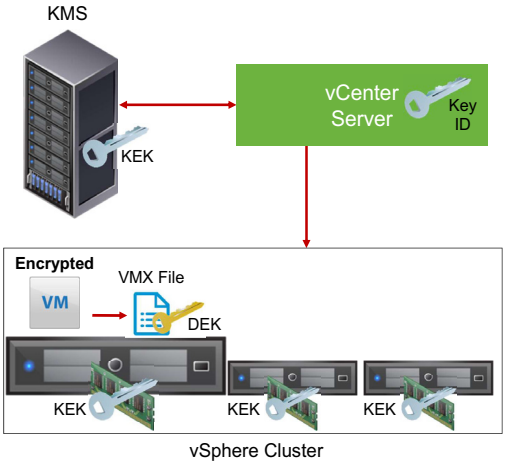
If the host belongs to a cluster, the cluster is deemed the security boundary. In this case, vCenter Server gives the KEK to all hosts in the cluster so that any cluster operations, such as vSphere vMotion migrations and vSphere HA operations, can take place.

https://t.me/learningnets

10-93 VM Encryption Process Flow (2)

- The ESXi host generates a key called the data encryption key (DEK).
- The DEK is used to encrypt the virtual machine and its disks.
- The ESXi host uses the KEK to encrypt the DEK.
- The DEK is stored in the VM's configuration file in an encrypted format.

KEKs remain in ESXi memory until the next reboot. For this reason, core dumps are encrypted.



If an ESXi host reboots, then a VM's KEK is no longer in the host's memory. vCenter Server requests the KEK with the corresponding key ID from the KMS and makes it available to the ESXi host once the host is running. The ESXi host can then decrypt the VM's DEK, as needed.

Because KEKs are stored in the host's memory and might be discovered if a core dump is analyzed, vSphere encrypts all core dumps to provide an additional level of security.

https://t.me/learningnets

10-94 Summary of Keys

The KMS provides keys to secure ESXi hosts and VMs.

Key Type	Description	Key Location
Host key	Used to encrypt core dumps	Stored in the ESXi host's memory
Data encryption key (DEK)	Used to encrypt the VM	Stored in the VM's configuration file (stored in an encrypted format)
Key encryption key (KEK)	Used to encrypt the DEK	Stored in memory of each ESXi host in the cluster

https://t.me/learningnets

10-95 Activity: VM Keys (1)

An ESXi host can pull a VM's key from the Key Management Server when required.

- ❑ True
- ❑ False

https://t.me/learningnets

10-96 Activity: VM Keys (2)

An ESXi host can pull a VM's key from the Key Management Server when required.

- ❑ True
- ✓ False

vCenter Server retrieves a VM's key from the KMS, then pushes this key to ESXi hosts when required. ESXi hosts cannot pull a VM's key from either the KMS or vCenter Server.

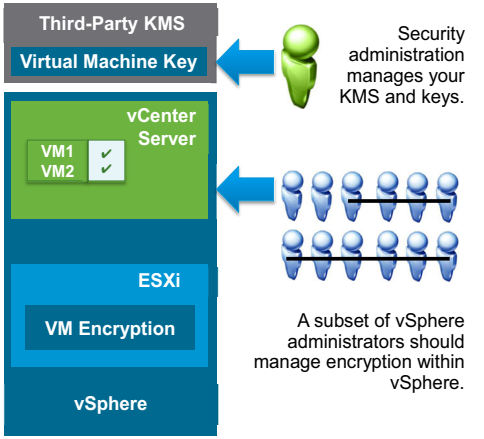
https://t.me/learningnets

10-97 Managing Virtual Machine Encryption

By default, the vCenter Server Administrator role has cryptographic privileges.

However, not all administrators should be able to control encryption operations or have access to keys.

vCenter Server provides a role called No Cryptography Administrator, which enables you to control which administrators have encryption privileges.



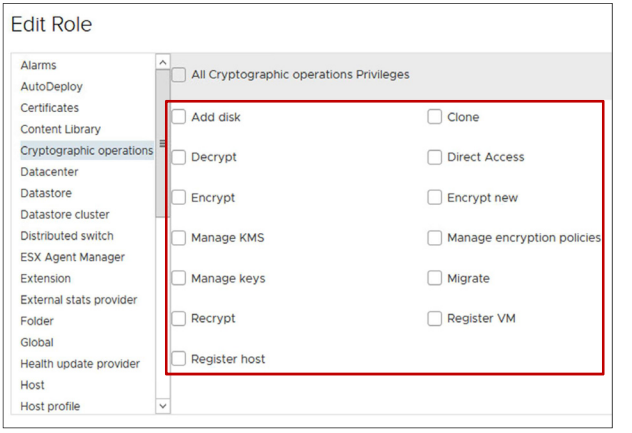
https://t.me/learningnets

10-98 vCenter Server Role: No Cryptography Administrator

The No Cryptography Administrator role has most of the same virtual machine privileges as the Administrator role does.

This role does not include the following privileges:

- Cryptographic Operations
- Global > Diagnostics
- Host > Inventory > Add host to cluster
- Host > Inventory > Add standalone host
- Host > Local operations > Manage user groups



A predefined role called No Cryptography Administrator is provided. You can also create custom roles that include cryptographic privileges.

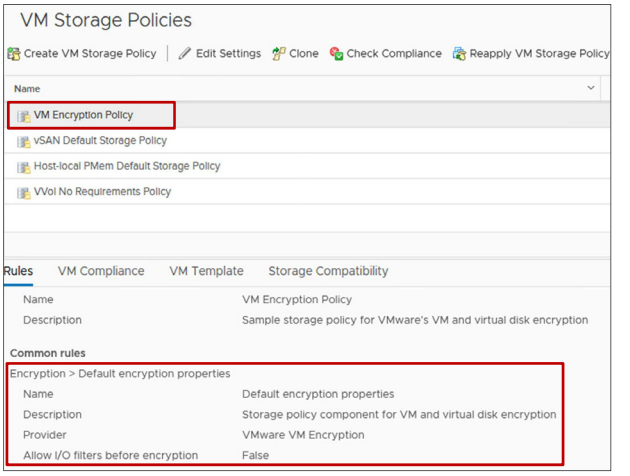
You do not want a user with the No Cryptography Administrator role to be able to add hosts to the vCenter Server inventory, because this task might require the installation of a host key, thus making the task a cryptographic operation.

https://t.me/learningnets

10-99 VM Encryption Prerequisites

To encrypt a VM, you need the following items:

- vCenter Server, version 6.5 or later.
- An ESXi host, version 6.5 or later.
- A registered KMS that can provide vCenter Server with keys.
- An encryption storage policy. For example, you can use the built-in VM encryption policy.



A key part of allowing encrypted virtual machines is to create a storage policy with the encryption filter enabled.

You create a common rule in the storage policy that defines the encryption properties. The default encryption properties are appropriate in most cases. Create a custom policy only if you want to combine encryption with other features, such as caching or replication. You must set the Allow I/O filters before encryption value to True to enable the use of a replication filter before the encryption filter. The replication filter would see plain text before it is encrypted.

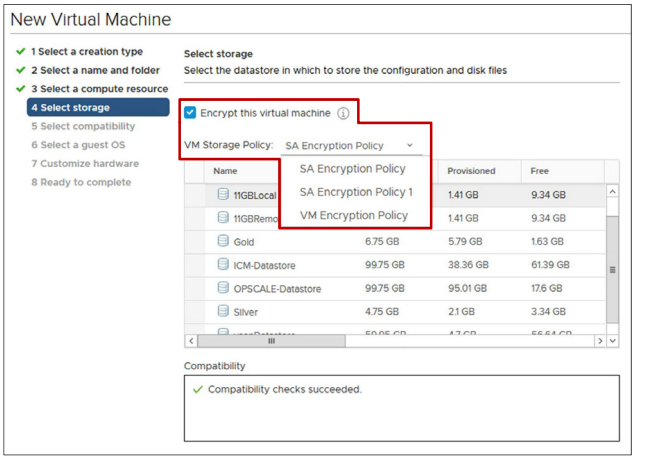
https://t.me/learningnets

10-100 Encrypting New Virtual Machines

A new VM with an encryption storage policy applied is encrypted immediately upon creation.

No additional steps need to be taken to encrypt the new VM.

The encryption storage policy can be applied to any VM, regardless of both the operating system type and the storage on which the VM resides.



To encrypt a new VM, perform the following steps:

1. In the New Virtual Machine wizard, select the **Encrypt this virtual machine** check box to show only the encryption policies.
2. Select an encryption policy from the **VM Storage Policy** drop-down menu.

https://t.me/learningnets

10-101 Encrypting Existing Virtual Machines

Encrypting an existing VM is slightly different from encrypting a new VM.

When a VM encryption storage policy is applied to an existing VM, vSphere automatically performs the following tasks:

- A duplicate disk is created with the encryption I/O filter applied.
- Data is copied from the unencrypted disk to the new disk and the data is encrypted.
- After all the data is copied, the new encrypted VMDK file is attached to the VM.
- The old unencrypted disk is deleted.

This operation is performed on one disk at a time. You need at least the capacity of the largest disk available in the datastore to complete this operation.

This operation is allowed only when the VM is powered off.

Because the data of an existing virtual machine already exists, a duplicate disk is created with the encryption I/O filter applied. Data is then copied from the unencrypted disk to the new encrypted disk, applying the encryption during the copy process. After all the data is copied, the new encrypted VMDK file is attached to the virtual machine, and the old unencrypted disk is deleted.

The virtual machine must be powered off so that no mirror driver is employed.

https://t.me/learningnets

10-102 Backing Up Encrypted Virtual Machines (1)

You can back up encrypted virtual machines with a backup and recovery solution that uses VMware vSphere® Storage APIs - Data Protection.

The backup and recovery solution might also provide its own encryption mechanism.

The ability to back up an encrypted virtual machine depends on the transport mode used by the backup solution.

Transport Mode	Description
SAN mode	• Leveraged by backup servers that have direct access to the physical storage where the virtual machine files reside.
Hot-add mode	• A proxy virtual machine residing on a host with direct access to the virtual machine's disks being backed up can perform the reads from the virtual disk directly and perform a disk-to-disk copy of the backup data.
NBD/NBDSSL mode	• When no other transport is available, this mode is used. • The ESXi host performs the reads and then sends the data over a LAN, using the NFC protocol.

A transport mode defines the method used to transfer data to the backup server.

For more information about the transport modes, see *Virtual Disk Development Kit Programming Guide* at https://code.vmware.com/web/sdk/67/vddk.

https://t.me/learningnets

10-103 Backing Up Encrypted Virtual Machines (2)

The backup server must first determine whether a virtual machine is encrypted.

If the virtual machine is encrypted:

- SAN mode is not supported.
- NBD/NBDSSL is available and can be used.

If hot-add mode is used and the proxy machine is a physical machine:

- Hot-add must be ruled out, and NBD/NBDSSL is used.

If hot-add mode is used and the proxy machine is a virtual machine:

- Hot-add mode is available but the proxy machine must also be encrypted.

The user with which the backup appliance user was registered with vCenter Server must have cryptographic privileges.

The backup data is stored on the backup server in plain text:

- Grant restore privileges only to trusted individuals.
- Have a policy in place to reencrypt a restored virtual machine.

If the virtual machine is encrypted, then SAN mode is immediately ruled out as an available transport mode. Thus, the backup products that rely on this mode cannot back up encrypted virtual machines. The blocks are encrypted, and without vSphere assistance, the backup server cannot subsequently use the data.

If the backup proxy is a physical machine, then hot-add mode is ruled out, and NBD/NBDSSL mode is used. You rely on the ESXi host to open the disk, decrypt the data, and send the backup data in plain text to the backup server.

If the user with which you register the backup appliance to vCenter Server does not have cryptographic privileges, then the disk cannot be decrypted in order to send the data to the backup server.

For any of the transport modes, the backup data is stored on the backup server in plain text. Thus, a virtual machine could be restored as a plain-text machine. Privileges to restore virtual machines should be granted only to trusted individuals, most likely users with cryptographic privileges, so that the virtual machine can be immediately reencrypted after it is restored.

https://t.me/learningnets

10-104 Encrypted VM Operations

vSphere 6.7 introduces support for the following operations on encrypted VMs:

- Suspending and resuming an encrypted VM
- Taking a snapshot of the memory of an encrypted VM
- Unlocking an encrypted VM

https://t.me/learningnets

10-105 Unlocking an Encrypted VM (1)

A VM goes into a locked state when its KEK is unavailable.

A VM's KEK might be unavailable due to the following reasons:

- The KEK is missing on the ESXi host.
- The KEK is missing on the KMS cluster.
- The KMS cluster encountered an error.

vCenter Server triggers an alarm when the VM goes into a locked state.

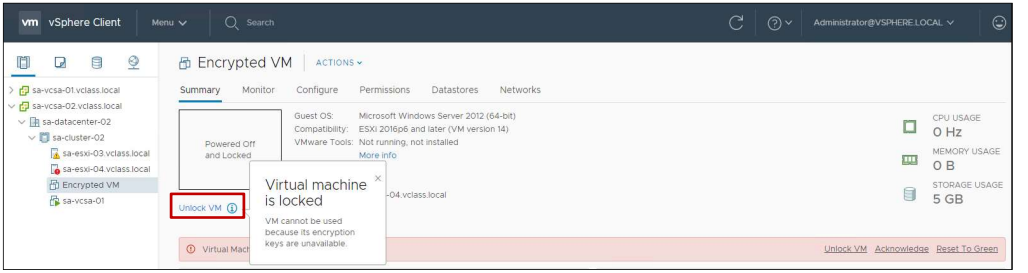
https://t.me/learningnets

10-106 Unlocking an Encrypted VM (2)

To unlock the VM, first ensure that the required key is available on the KMS cluster.

Click **Unlock VM** in vSphere Client to unlock the encrypted VM:

- The Unlock Virtual Machine API sends VM encryption key to the ESXi host.



https://t.me/learningnets

10-107 Activity: Cryptographic Privileges (1)

Which one of the following predefined vCenter Server roles must a user have to perform cryptographic operations?

- ❑ Administrator
- ❑ Virtual Machine Power User
- ❑ Virtual Machine Console User
- ❑ No Cryptography Administrator

https://t.me/learningnets

10-108 Activity: Cryptographic Privileges (2)

Which one of the following predefined vCenter Server roles must a user have to perform cryptographic operations?

- ✓ Administrator
- ❑ Virtual Machine Power User
- ❑ Virtual Machine Console User
- ❑ No Cryptography Administrator

You can also create a custom role that has certain cryptographic privileges, but the only predefined role with cryptographic privileges is the Administrator role.

https://t.me/learningnets

10-109 About Encrypted Core Dumps

A core dump is a saved state of memory at the time of a system crash:

- It is used in debugging various types of system crashes.
- It is used by VMware technical support when debugging crashes such as purple screen crashes.
- It can contain sensitive data, including keys used for virtual machine encryption.

VMkernel core dumps are stored in `/var/core` on the ESXi host.

It is important to protect core dumps, even from VMware technical support personnel who are trying to help you debug your system crash.

If you use virtual machine encryption and an error occurs on the ESXi host, then the resulting core dump is encrypted to protect customer data.

A core dump is very useful when debugging system crashes, especially purple screen crashes. If the VMkernel crashes, this crash usually results in a purple screen. It is the most common type of core dump analyzed by VMware technical support.

If any ESXi hosts in a cluster are running encrypted VMs and an ESXi host crashes, the host creates an encrypted core dump which is stored in the `/var/core` folder on the ESXi host. Core dumps that are included in the `vm-support` package are also encrypted.

https://t.me/learningnets

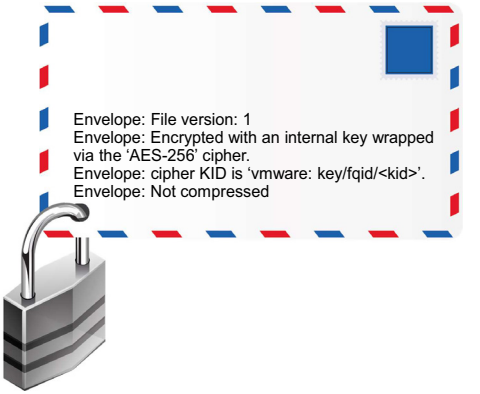
10-110 How Core Dumps Are Encrypted

Core dumps are encrypted as follows:

- An internal one-time key (AES-256) is generated by the host.
- The core dump is placed in a digital envelope.
- The envelope and core dump are encrypted using the internal key.
- The internal key is encrypted using the host key and then stored as part of the envelope.

To view the contents of the core dump file, the host key is required to open the envelope.

Since the host key is available only on the host, VMware technical support cannot access the core dump to help you debug a system crash.



Normally, the host key should be available only on the host or cluster where the original host resides. Thus, an encrypted core dump is not useful to a VMware Technical Support person who is trying to help you debug a system crash.

The `vm-support` command collects core dumps as is. That is, the host key is required to read them.

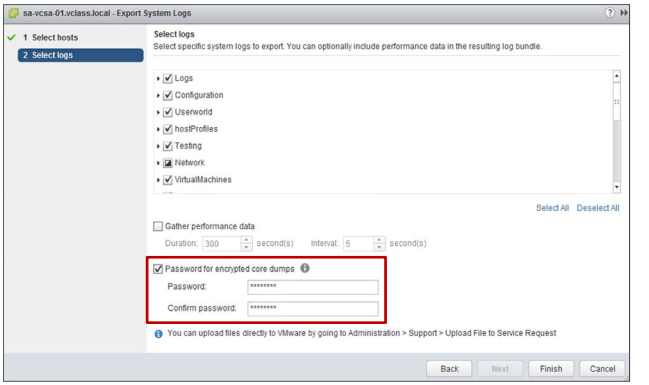
https://t.me/learningnets

10-111 Providing Passwords for Encrypted Core Dumps

With vSphere Web Client, you can generate a log bundle and provide a password for accessing encrypted core dumps.

The password is required to decrypt encrypted core dumps. Without the password, core dump data is indecipherable.

When working with VMware support personnel, provide them with the password so that they can access the core dump to help troubleshoot your issues.



By providing a password, the core dumps are reencrypted with a password-protected envelope. During the support log bundle collection, a new envelope file is created with a new incident key, and the envelope is protected by a password.

Using a secure mechanism, you provide the password to VMware Support, or appropriate support staff, who can view the core dump contents by using the password.

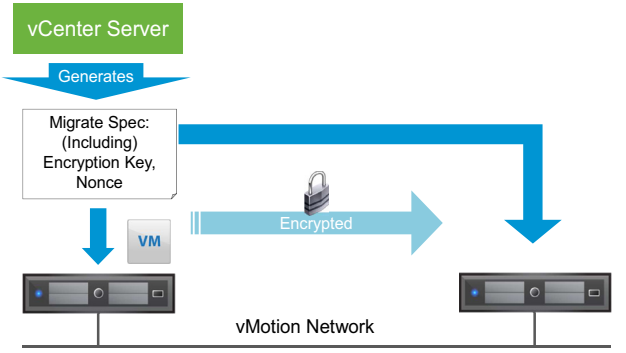
You can set a password for encrypted core dumps when you export the system logs.

https://t.me/learningnets

10-112 About Encrypted vSphere vMotion

Encrypted vSphere vMotion secures the confidentiality, integrity, and authenticity of data that is transferred with vSphere vMotion.

Encrypted vSphere vMotion supports all variants of vSphere vMotion for unencrypted virtual machines, including migration across different vCenter Server instances and versions.



Encrypted vSphere vMotion secures communication on the vMotion network so that sensitive data cannot be discovered during memory migration.

vCenter Server generates a one-use (nonce) encryption key and includes this key in a communication (or migration specification) to both the source and the destination ESXi hosts. This key is used to encrypt and decrypt the data sent over the vMotion network.

In a vSphere vMotion migration across vCenter Server instances, the local vCenter Server instance generates the encryption key and includes the key in a communication over an SSL secured channel to the remote vCenter Server instance. The two vCenter Server instances then pass the required keys to their local ESXi hosts, and these hosts communicate directly using these keys.

For encrypted virtual machines, the virtual machine's memory is encrypted over the vMotion network. The disks are already encrypted and transmitted over the vMotion network as is (encrypted). For nonencrypted virtual machines, the virtual machine's memory is encrypted over the vMotion network, but the disks are transmitted as is (nonencrypted).

For virtual machines that are encrypted, migration with vSphere vMotion always uses encrypted vSphere vMotion. You cannot turn off encrypted vSphere vMotion for encrypted virtual machines.

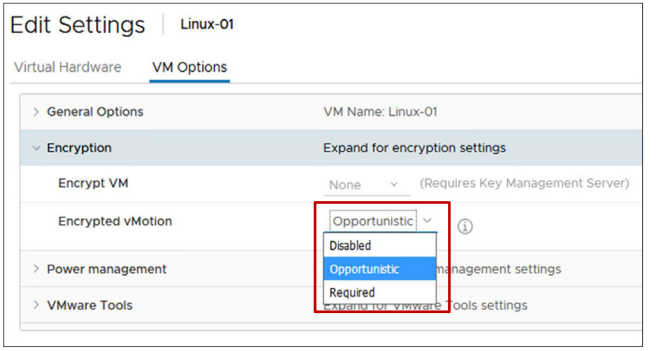
https://t.me/learningnets

10-113 Configuring Encrypted vSphere vMotion

vSphere vMotion always uses encryption when migrating encrypted virtual machines.

For virtual machines that are not encrypted, you can edit the virtual machine's settings to use one of the following states:

- Disabled: Do not use encrypted vSphere vMotion.
- Opportunistic (default): Use encrypted vSphere vMotion if the source and destination hosts support it.
- Required: Use only encrypted vSphere vMotion. If the source or destination host does not support encrypted vSphere vMotion, then the migration is not allowed.



You might want to ensure that vSphere vMotion migrations of certain virtual machines are always encrypted. In such a case, change the encrypted vSphere vMotion setting for a virtual machine to Required.

All virtual machines default to the Opportunistic encryption setting.

When you encrypt a virtual machine, it keeps a record of the current encrypted vSphere vMotion setting. If you later disable encryption for that virtual machine, the encrypted vMotion setting remains at Required until you change the setting explicitly.

The encryption setting is not stored in the VMX file. The setting is stored only in the vCenter Server database.

https://t.me/learningnets

10-114 Encrypted vSphere vMotion Requirements

Encrypted vSphere vMotion requires vSphere 6.5 or later.

Consider the following points when choosing encrypted vSphere vMotion settings:

- If set to Required, the vSphere vMotion migration proceeds with encryption only if vCenter Server and both ESXi hosts are installed with version 6.5 or later.
- If set to Opportunistic, one of the following actions occurs:
  - The vSphere vMotion migration proceeds with encryption if vCenter Server and both hosts are installed with version 6.5 or later.
  - The vSphere vMotion migration proceeds without encryption if any host or vCenter Server is installed with version 6.0 or earlier.
- If set to Disabled, the vSphere vMotion migration proceeds without encryption regardless of the vSphere version.

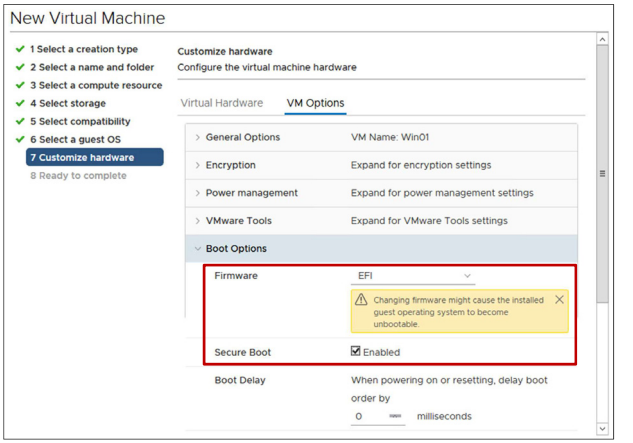
https://t.me/learningnets

10-115 Virtual Machine Protection with UEFI Secure Boot

You can enable secure boot in a virtual machine if the following prerequisites are met:

- The virtual machine uses EFI firmware.
- The virtual hardware is version 13 or later.
- The guest operating system supports UEFI Secure Boot.

To properly support UEFI Secure Boot in a virtual machine, see the guest OS vendor documentation for the proper way to configure and use EFI firmware.



For certain virtual machine hardware versions and operating systems, you can enable UEFI Secure Boot exactly as you can for a physical machine.

In an operating system that supports UEFI Secure Boot, each piece of boot software is signed, including the bootloader, the operating system kernel, and operating system drivers. The virtual machine's default configuration includes several code-signing certificates:

- A Microsoft certificate that is used only for booting Windows
- A Microsoft certificate that is used for third-party code that is signed by Microsoft, such as Linux bootloaders
- A VMware certificate that is used only for booting ESXi inside a virtual machine

The virtual machine's default configuration includes one certificate for authenticating requests to modify the Secure Boot configuration, including the Secure Boot revocation list, from inside the virtual machine, which is a Microsoft KEK certificate. In almost all cases, replacing the existing certificates is not necessary.

For a list of operating systems that support UEFI Secure Boot, see the VMware Compatibility Guide at http://www.vmware.com/resources/compatibility.

https://t.me/learningnets

10-116 Using vTPM in a VM (1)

vSphere 6.7 introduces support for the Virtual Trusted Platform Module (vTPM) device, which lets you add a TPM 2.0 virtual cryptoprocessor to a VM.

A vTPM device is a software emulation of the TPM functionality:

- It enables the guest operating system to create and store private keys in such a way that they are never exposed to this guest operating system.
- It enables the guest operating system to use the private key for encryption or signing.

With a vTPM device, a third party can remotely attest to (validate) the identity of the firmware and the guest operating system.

vTPM has the following use cases:

- An operating system can verify that the firmware loaded was not compromised since the last run.
- An application can verify that the operating system did not load any malicious components.

Compromising the guest operating system usually compromises its secrets, but enabling a vTPM in the guest greatly reduces this risk.

You can add a vTPM device to either a new virtual machine or an existing virtual machine.

https://t.me/learningnets

10-117 Using vTPM in a VM (2)

vTPM depends on VM encryption to secure virtual TPM data.

When you configure vTPM, VM encryption automatically encrypts the VM's home directory (which contains nvram, *.vmx, *.vmsn, snapshots, core files, and so on) but not the disks.

You can choose to add encryption explicitly for the VM and its disks.

You can back up a VM enabled with a vTPM:

- The backup must include all VM data, including the nvram file.
- If your backup does not include the nvram file, then you cannot restore a VM with a vTPM.
- Since the VM home files are encrypted, ensure that the encryption keys are available at the time of a restore.

You can remove a vTPM from a VM. However, removing vTPM causes all encrypted information on a VM to become unrecoverable.

Before removing vTPM from a VM, disable any applications in the guest operating system that use vTPM.

vTPM can be enabled on a virtual machine by editing the virtual machine's settings with vSphere Client.

vTPM can be removed from a virtual machine but the virtual machine must be powered off first. Before you remove vTPM, disable any applications that use vTPM. If you do not disable these applications, the virtual machine might fail to boot when powered on.

https://t.me/learningnets

10-118 Prerequisites to Add vTPM

To use vTPM, your vSphere environment must meet specific requirements:

- Component requirements:
  - ESXi 6.7
  - vCenter Server 6.7
  - KMS configured in vCenter Server to encrypt a VM
- Virtual machine requirements:
  - EFI firmware
  - Virtual machine hardware version 14 or later
  - Windows 10 (64-bit) or Windows Server 2016 (64-bit) guest operating system

vTPM does not require a physical TPM 2.0 chip to be present on the ESXi host. However, if you want to perform host attestation, an external entity, such as a TPM 2.0 physical chip, is required.

https://t.me/learningnets

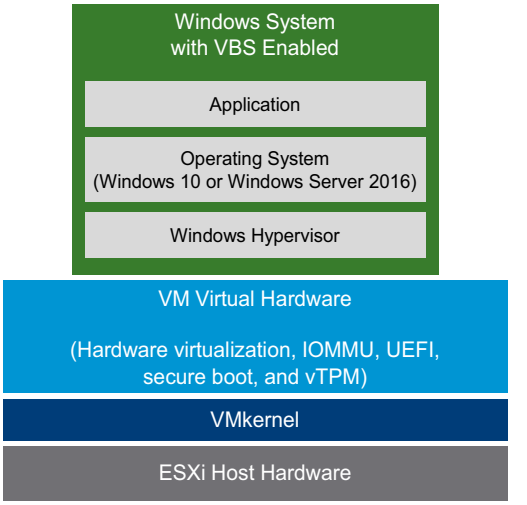
10-119 Virtualization-Based Security for Windows Guest Operating Systems

Windows 10 and Windows Server 2016 introduced a new security technology called virtualization-based security (VBS).

VBS uses Microsoft's Hyper-V virtualization technology to isolate core Windows operating system services in a segregated, virtualized environment.

This virtualized environment provides the following security benefits:

- An additional level of protection to the key services in Windows
- An isolated execution environment for user activity
- Enhanced host operating system protection from attacks



Microsoft virtualization-based security (VBS), a feature of Windows 10 and Windows Server 2016 operating systems, uses hardware and software virtualization to enhance system security by creating an isolated, hypervisor-restricted, specialized subsystem.

VBS allows you to use the following Windows security features to harden your system as well as isolate key system and user secrets from being compromised:

- Credential Guard: Aims to isolate and harden key system and user secrets against compromise.
- Device Guard: Provides a set of features designed to work together to prevent and eliminate malware from running on a Windows system.
- Configurable Code Integrity: Ensures that only trusted code runs from the boot loader onwards.

See the topic on virtualization-based security in the Microsoft documentation for more information.

https://t.me/learningnets

10-120 VBS Requirements and Limitations

To use VBS in a VM, the following prerequisites must be met:

- The virtual machine must be of virtual hardware version 14 or later.
- The guest operating system must be either Windows 10 or Windows Server 2016, or later.
- The VM must be located on an ESXi host version 6.7 or later.

VMs that have VBS enabled use UEFI and secure boot by default.

The following features are not supported with VBS:

- vSphere Fault Tolerance
- PCI Passthrough

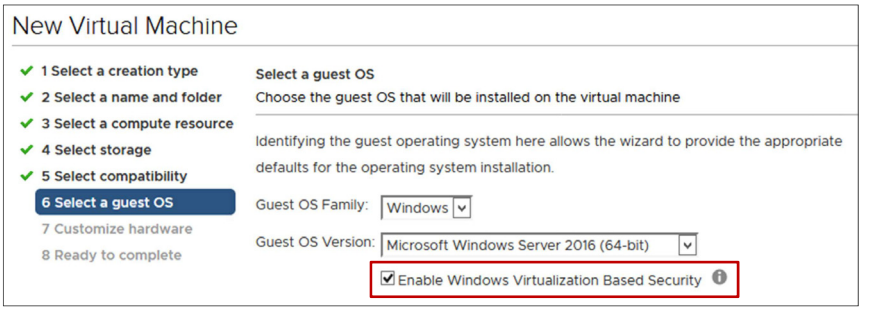
https://t.me/learningnets

10-121 Enabling VBS

To enable VBS when you create a new VM, select Windows 10 or Windows Server 2016 and click **Enable Windows Virtualization Based Security**.

Enabling VBS exposes hardware-assisted virtualization, the input-output memory management unit (IOMMU), EFI, and Secure Boot to the guest operating system.

After you enable VBS for a VM, you must enable VBS within the guest operating system.



https://t.me/learningnets

10-122 What's Next

Consider attending additional courses and pursuing VMware certifications as your next steps.

Course Name	Learning Path
vSphere: Troubleshooting Workshop	Data Center Virtualization
vSphere: Design Workshop	Data Center Virtualization
vSAN: Deploy and Manage	vSAN
vSAN: Troubleshooting Workshop	vSAN

Certification Name
VMware Certified Professional 6.5 – Data Center Virtualization (VCP6.5-DCV)
VMware Certified Advanced Professional 6 – Data Center Virtualization Deployment (VCAP6-DCV)
VMware Certified Advanced Professional 6.5 – Data Center Virtualization Design (VCAP6.5-DCV)
VMware vSAN 2017 Specialist badge

The vSphere: Optimize and Scale course helps prepare you for the VCP6.5-DCV certification and also satisfies the training requirement for this certification.

The vSphere: Troubleshooting Workshop course helps prepare you for the VCAP6-DCV certification.

The vSphere: Design Workshop course helps prepare you for the VCAP6.5-DCV certification.

The vSAN: Deploy and Manage course prepares you for the VMware vSAN 2017 Specialist badge.

https://t.me/learningnets

10-123 Lab 14: Virtual Machine Encryption

Register a key management server with vCenter Server and encrypt a virtual machine

1. Verify Access to the Key Management Server
2. Register the KMS with vCenter Server
3. Create an Encryption Storage Policy
4. Encrypt a Virtual Machine
5. Use Encrypted vSphere vMotion to Migrate Virtual Machines

https://t.me/learningnets

10-124 Review of Learner Objectives

You should be able to meet the following objectives:

- Set up encryption in your vSphere environment
- Encrypt virtual machines
- Manage encrypted virtual machines
- Encrypt core dumps
- Enable encrypted vSphere vMotion
- Describe support for virtual machine security features:
  - UEFI secure boot
  - vTPM
  - Virtualization-based security

https://t.me/learningnets

10-125 Key Points

- A best practice is to define a role using the smallest number of privileges possible for better security and added control.
- Hardening a vSphere environment to be more secure involves configuring the vCenter Server system, ESXi hosts, virtual machines, and the vSphere network.
- VMware CA provisions each ESXi host and each vCenter Server service with certificates that are signed by VMware CA.
- Virtual machine encryption cannot operate without vCenter Server because vCenter Server communicates with the KMS and manages keys used for encryption.
- If you use virtual machine encryption and the ESXi host crashes, then the resulting core dump is encrypted to protect customer data.
- Encrypted vSphere vMotion is available for all virtual machines, whether or not the virtual machines are encrypted.

Questions?

https://t.me/learningnets

https://t.me/learningnets