VMware vSphere: Optimize and Scale

Lecture Manual
ESXi 7 and vCenter Server 7

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

https://t.me/learningnets

VMware vSphere:
Optimize and Scale
Lecture Manual
ESXi 7 and vCenter Server 7
Part Number EDU-EN-VSOS7-LECT (4/2020)

Copyright © 2020 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, Project Photon OS™, vCenter Linked Mode, VMware Certificate Authority, VMware ESX®, VMware ESXi™, VMware Horizon®, VMware Horizon® 7, VMware Horizon® 7 on VMware Cloud™ on AWS, VMware Horizon® View™, VMware Host Client™, VMware NSX®, VMware NSX® Manager™, VMware NSX-T™ Data Center, VMware Photon™, VMware Platform Services Controller™, VMware PowerCLI™, VMware Tanzu™, VMware Tanzu™ Kubernetes Grid™, VMware Tanzu™ Kubernetes Grid™ service for vSphere®, VMware Tools™, VMware vCenter Server®, VMware vCenter® Lifecycle Manager™, VMware vCenter® Server Appliance™, VMware vCenter® Single Sign-On, VMware vCloud Director®, VMware vCloud Director® for Service Providers, VMware vCloud®, VMware Verify™, VMware View®, VMware vRealize® Suite Lifecycle Manager™, VMware vSAN™, VMware vShield Endpoint™, VMware vShield™, VMware vSphere®, VMware vSphere® API, VMware vSphere® API for Storage Awareness™, VMware vSphere® Authentication Proxy, VMware vSphere® Auto Deploy™, VMware vSphere® Client™, VMware vSphere® Command-Line Interface, VMware vSphere® Distributed Resource Scheduler™, VMware vSphere® Distributed Switch™, VMware vSphere® ESXi™ Dump Collector, VMware vSphere® ESXi™ Image Builder CLI, VMware vSphere® ESXi™ Shell, VMware vSphere® Fault Tolerance, VMware vSphere® High Availability, VMware vSphere® Network I/O Control, VMware vSphere® PowerCLI™, VMware vSphere® Storage APIs - Array Integration, VMware vSphere® Storage APIs - Data Protection, VMware vSphere® Storage DRS™, VMware vSphere® Storage I/O Control, VMware vSphere® Storage vMotion®, VMware vSphere® Trust Authority™, VMware vSphere® Update Manager™, VMware vSphere® Virtual Volumes™, VMware vSphere® VMFS, VMware vSphere® vMotion®, VMware vSphere® Web Client, VMware vSphere® with Kubernetes, and VMware Workstation™ are registered trademarks or trademarks 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 Course Introduction....2
1-3 Importance....3
1-4 Learner Objectives....4
1-5 Course Outline....5
1-6 Typographical Conventions....6
1-7 References (1)....7
1-8 References (2)....8
1-9 VMware Online Resources....9
1-10 Review of vSphere User Interfaces....10
1-11 About Scalability....11
1-12 About Optimization....12
1-13 About Security....13
1-14 Data Center Layout....14
1-15 VMware Education Overview....16
1-16 VMware Certification Overview....17
1-17 VMware Digital Badge Overview....18
1-18 About VMware Skyline....19

Module 2 Network Scalability 21

2-2 Importance....22
2-3 Module Lessons....23
2-4 Lesson 1: Creating vSphere Distributed Switches....24
2-5 Learner Objectives....25
2-6 About Distributed Switches....26
2-7 Distributed Switch Architecture....27
2-8 Standard and Distributed Switches: Shared Features....29
2-9 Distributed Switch Features....30

https://t.me/learningnets

2-10 Viewing VDS in vSphere Client 31
2-11 Load-Based NIC Teaming 32
2-12 Discovery Protocols 33
2-13 Configuring CDP or LLDP 34
2-14 About Port Binding 36
2-15 About the Traffic Filtering and Marking Policy 38
2-16 Creating Network Traffic Rules 39
2-17 Network Traffic Rule: Example 1 40
2-18 Network Traffic Rule: Example 2 41
2-19 Marking Network Traffic (1) 42
2-20 Marking Network Traffic (2) 44
2-21 Access to NSX Virtual Distributed Switch (1) 45
2-22 Access to NSX Virtual Distributed Switch (2) 46
2-23 Access to NSX Virtual Distributed Switch (3) 47
2-24 Activity: Distributed Switch Features (1) 48
2-25 Activity: Distributed Switch Features (2) 49
2-26 Labs 50
2-27 Lab 1: Accessing the Lab Environment 51
2-28 Lab 2: Configuring vSphere Distributed Switches (1) 52
2-29 Lab 3: Configuring vSphere Distributed Switches (2) 53
2-30 Review of Learner Objectives 54
2-31 Lesson 2: Managing vSphere Distributed Switches 55
2-32 Learner Objectives 56
2-33 About Health Check for Distributed Switches 57
2-34 Health Check Example 58
2-35 Enabling Health Check 60
2-36 Monitoring Health Check Results 61
2-37 Backing Up and Restoring Configurations 62
2-38 Exporting Distributed Switch Configurations 63
2-39 Restoring and Importing Distributed Switch Configurations 65
2-40 Rollback and Recovery of the Management Network 66

https://t.me/learningnets

2-41 Rollback Details 67
2-42 Recovery Through the DCUI 68
2-43 Lab 4: Managing vSphere Distributed Switches 69
2-44 Review of Learner Objectives 70
2-45 Lesson 3: Using Network I/O Control in Distributed Switches 71
2-46 Learner Objectives 72
2-47 About Network I/O Control (1) 73
2-48 About Network I/O Control (2) 74
2-49 About the Bandwidth Allocation Model for System Traffic 75
2-50 Configuring Bandwidth Allocations for System Traffic 76
2-51 Reserving Bandwidth for System Traffic (1) 78
2-52 Reserving Bandwidth for System Traffic (2) 79
2-53 Bandwidth Aggregation for Network Resource Pools 80
2-54 Creating Network Resource Pools 81
2-55 Bandwidth Allocation for VM Traffic 82
2-56 Bandwidth Allocation for Individual VMs 83
2-57 Bandwidth Admission Control in vSphere DRS (1) 84
2-58 Bandwidth Admission Control in vSphere DRS (2) 85
2-59 Bandwidth Admission Control in vSphere HA 86
2-60 Activity: Network I/O Control (1) 87
2-61 Activity: Network I/O Control (2) 88
2-62 Review of Learner Objectives 89
2-63 Lesson 4: NetFlow and Port Mirroring in Distributed Switches 90
2-64 Learner Objectives 91
2-65 About NetFlow 92
2-66 About Network Flows 93
2-67 Network Flow Analysis (1) 95
2-68 Network Flow Analysis (2) 96
2-69 Configuring NetFlow on Distributed Switches 97
2-70 Enabling NetFlow on a Distributed Port Group 99
2-71 About Port Mirroring 100

https://t.me/learningnets

2-72 Port Mirroring Session Types 101
2-73 Port Mirroring Session Properties 103
2-74 Source and Destination in a Port Mirroring Session 104
2-75 Lab 5: Using Port Mirroring 105
2-76 Review of Learner Objectives 106
2-77 Key Points 107

Module 3 Storage Scalability 109

3-2 Importance 110
3-3 Module Lessons 111
3-4 Lesson 1: VMFS Datastores 112
3-5 Learner Objectives 113
3-6 About VMFS Datastores 114
3-7 VMFS5 and VMFS6 Datastores 115
3-8 VMFS6 Datastores 116
3-9 VMFS 6 New Features in vSphere 7 117
3-10 ESXi Support for VMFS5 and VMFS6 118
3-11 Storage Device Formats 119
3-12 Storage Devices Supported by VMFS Datastores 121
3-13 Limitations of Using 4Kn Drives 122
3-14 Snapshot Formats for VMFS 123
3-15 Migrating Your Data from VMFS5 to VMFS6 124
3-16 Activity: VMFS6 Features (1) 125
3-17 Activity: VMFS6 Features (2) 126
3-18 Review of Learner Objectives 127
3-19 Lesson 2: vSphere Storage APIs - Array Integration 128
3-20 Learner Objectives 129
3-21 About vSphere Storage APIs - Array Integration 130
3-22 Hardware Acceleration APIs for Block Storage 131
3-23 Hardware Acceleration for NAS 133
3-24 Array Thin Provisioning APIs 135

https://t.me/learningnets

3-25 About Space Reclamation (1) 136
3-26 About Space Reclamation (2) 137
3-27 Space Reclamation with VMFS 138
3-28 Performing Manual Space Reclamation on a VMFS5 Datastore 139
3-29 Enabling Automatic Space Reclamation on a VMFS6 Datastore 141
3-30 Configuring the Automatic Unmap Rate 142
3-31 Space Reclamation Support for Guest Operating Systems (1) 143
3-32 Space Reclamation Support for Guest Operating Systems (2) 144
3-33 Space Reclamation Support for Guest Operating Systems (3) 145
3-34 Activity: vSphere Storage APIs – Array Integration (1) 146
3-35 Activity: vSphere Storage APIs – Array Integration (2) 147
3-36 Review of Learner Objectives 148
3-37 Lesson 3: vSphere Storage APIs for Storage Awareness and vSphere APIs for I/O Filtering 149
3-38 Learner Objectives 150
3-39 About vSphere API for Storage Awareness 151
3-40 Benefits of Storage Providers 153
3-41 Comparison of Storage Provider Versions 154
3-42 Storage Provider Categories 155
3-43 Registering a Storage Provider 156
3-44 About vSphere APIs for I/O Filtering 157
3-45 Types of I/O Filters 158
3-46 How I/O Filters Work (1) 159
3-47 How I/O Filters Work (2) 160
3-48 I/O Filter Providers 161
3-49 Displaying I/O Filter Provider Information 163
3-50 Activity: I/O Filters (1) 164
3-51 Activity: I/O Filters (2) 165
3-52 Review of Learner Objectives 166
3-53 Lesson 4: Storage Policy Based Management Overview 167
3-54 Learner Objectives 168

https://t.me/learningnets

3-55 About Storage Policy Based Management (1) 169
3-56 About Storage Policy Based Management (2) 170
3-57 About VM Storage Policies 171
3-58 About Storage Policy Rules 172
3-59 Storage Policy Example (1) 173
3-60 Storage Policy Example (2) 174
3-61 About Storage Policy Components 175
3-62 Storage Policy Component Example 176
3-63 Activity: Storage Policy Rules (1) 177
3-64 Activity: Storage Policy Rules (2) 178
3-65 Review of Learner Objectives 179
3-66 Lesson 5: Using VM Storage Policies 180
3-67 Learner Objectives 181
3-68 Creating and Managing Storage Policies 182
3-69 Step 1: Configuring Storage to Be Used for Storage Policies 183
3-70 Step 2: Creating Storage Policy Components 184
3-71 Step 3: Creating VM Storage Policies 185
3-72 Example: Creating a Silver Tier Storage Policy (1) 186
3-73 Example: Creating a Silver Tier Storage Policy (2) 187
3-74 Example: Creating a Silver Tier Storage Policy (3) 188
3-75 Example: Creating a Silver Tier Storage Policy (4) 189
3-76 Step 4: Applying the Storage Policy to the VM 190
3-77 Step 5: Checking Compliance for the VM Storage Policy 191
3-78 Lab 6: Using Policy Based Storage 192
3-79 Review of Learner Objectives 193
3-80 Lesson 6: Storage Policies for vSAN 194
3-81 Learner Objectives 195
3-82 About vSAN (1) 196
3-83 About vSAN (2) 197
3-84 vSAN Disk Architecture 198
3-85 Object-Based Storage 199

https://t.me/learningnets

3-86 vSAN Storage Providers 200
3-87 vSAN Storage Policies 201
3-88 Creating vSAN Storage Policies 202
3-89 vSAN Capabilities (1) 203
3-90 vSAN Capabilities (2) 204
3-91 Lab 7: Creating vSAN Storage Policies 205
3-92 Review of Learner Objectives 206
3-93 Lesson 7: Storage Policies for vSphere Virtual Volumes 207
3-94 Learner Objectives 208
3-95 About vSphere Virtual Volumes 209
3-96 vSphere Virtual Volumes Architecture 210
3-97 vSphere Virtual Volumes Storage Providers 211
3-98 Protocol Endpoints 213
3-99 Storage Containers and vSphere Virtual Volumes Datastores 214
3-100 Virtual Volumes 215
3-101 Mapping VM Files to Virtual Volumes 216
3-102 vSphere Virtual Volumes Replication 217
3-103 Virtual Volume Storage Policies 219
3-104 Virtual Volume Storage Policy: Examples 220
3-105 Creating Virtual Volumes Storage Policies (1) 221
3-106 Creating Virtual Volumes Storage Policies (2) 222
3-107 Activity: Displaying Storage Capabilities (1) 223
3-108 Activity: Displaying Storage Capabilities (2) 224
3-109 Review of Learner Objectives 225
3-110 Lesson 8: Storage I/O Control 226
3-111 Learner Objectives 227
3-112 About Storage I/O Control 228
3-113 Enabling Storage I/O Control (1) 229
3-114 Enabling Storage I/O Control (2) 230
3-115 Monitoring Latency 231
3-116 Automatic Threshold Detection 232

https://t.me/learningnets

3-117 Storage Policy Components for Storage I/O Control 233
3-118 Configuring Storage I/O Control 234
3-119 Storage I/O Control Requirements 235
3-120 Activity: Storage I/O Control (1) 236
3-121 Activity: Storage I/O Control (2) 237
3-122 Review of Learner Objectives 238
3-123 Lesson 9: Datastore Clusters and vSphere Storage DRS 239
3-124 Learner Objectives 240
3-125 About Datastore Clusters 241
3-126 Datastore Cluster Requirements 242
3-127 Host Cluster to Datastore Cluster Relationships 243
3-128 vSphere Storage DRS Overview 244
3-129 Initial Disk Placement 245
3-130 vSphere Storage DRS Affinity Rules 246
3-131 Storage Migration Recommendations 247
3-132 Datastore Correlation Detector 248
3-133 Configuring vSphere Storage DRS Runtime Settings 249
3-134 vSphere Storage DRS Maintenance Mode 251
3-135 Backups and vSphere Storage DRS 252
3-136 Activity: vSphere Storage DRS (1) 253
3-137 Activity: vSphere Storage DRS (2) 254
3-138 Review of Learner Objectives 255
3-139 Key Points 256

Module 4 Host and Management Scalability 257

4-2 Importance 258
4-3 Module Lessons 259
4-4 Lesson 1: VMware Certificate Management 260
4-5 Learner Objectives 261
4-6 Certificate Use Case 262
4-7 About Certificate Authorities 263

https://t.me/learningnets

4-8 About VMware CA 265
4-9 About the VMware Endpoint Certificate Store 266
4-10 vSphere Certificate Types 267
4-11 Chain of Trust (1) 268
4-12 Chain of Trust (2) 269
4-13 Chain of Trust (3) 270
4-14 Activity: ESXi Certificates (1) 271
4-15 Activity: ESXi Certificates (2) 272
4-16 VMware CA Modes 273
4-17 VMware CA: Root CA Mode 274
4-18 Replacing the VMware CA Certificate in Root CA Mode 275
4-19 VMware CA: Subordinate CA Mode 276
4-20 Replacing the VMware CA Certificate with a Subordinate CA Certificate 277
4-21 VMware CA: Hybrid Mode 278
4-22 Hybrid Mode: Replacing a vCenter Server Machine Server Certificate with a Custom Certificate 279
4-23 Certificate Replacement Options 280
4-24 Certificate Manager 281
4-25 Replacing All VMware CA Certificates with Custom Certificates 282
4-26 ESXi Certificate Replacement Options 283
4-27 Renewing or Refreshing ESXi Certificates 284
4-28 VMware CA Availability 285
4-29 Lab 8: Working with Certificates 286
4-30 Review of Learner Objectives 287
4-31 Lesson 2: vCenter Server Identity Federation 288
4-32 Learner Objectives 289
4-33 About vCenter Server IDP Federation for ADFS 290
4-34 vCenter Server IDP Federation Use Cases 291
4-35 vCenter Server IDP Federation Components 292
4-36 Prerequisites for vCenter Server IDP Federation 293
4-37 User Login with the vCenter Server Built-In Identity Provider 294

https://t.me/learningnets

4-38 Configuring vCenter Server IDP Federation in vSphere (1) 295
4-39 Configuring vCenter Server IDP Federation in vSphere (2) 296
4-40 Configuring vCenter Server IDP Federation in vSphere (3) 297
4-41 Configuring vCenter Server IDP Federation in vSphere (4) 298
4-42 Login Workflow of the vCenter Server IDP Federation for ADFS 299
4-43 Enabling ADFS on an Existing Enhanced Linked Mode Configuration 300
4-44 Lab 9: Configuring Identity Federation to Use Microsoft ADFS 301
4-45 Review of Learner Objectives 302
4-46 Lesson 3: vSphere Trust Authority 303
4-47 Learner Objectives 304
4-48 Overview of VM Encryption 305
4-49 UEFI Secure Boot for ESXi Hosts 306
4-50 UEFI Secure Boot Sequence for ESXi Hosts 307
4-51 TPM 2.0 on ESXi Hosts 308
4-52 About Remote Attestation 309
4-53 Remote Attestation Process 310
4-54 Requirements for Using TPM 2.0 311
4-55 About vSphere Trust Authority 312
4-56 About vSphere Trust Authority Administrator 313
4-57 Use Case for vSphere Trust Authority 314
4-58 vSphere Trust Authority Components (1) 315
4-59 vSphere Trust Authority Components (2) 316
4-60 vSphere Trust Authority Workflow 317
4-61 ESXi Components 318
4-62 ESXi Components 319
4-63 Attestation Service 320
4-64 Attestation Service Trust 321
4-65 Key Provider Service 322
4-66 Trusted Infrastructure Agent 323
4-67 vCenter Server Topologies (1) 324
4-68 vCenter Server Topologies (2) 325

https://t.me/learningnets

4-69 About Enabling vSphere Trust Authority 326
4-70 About VM Encryption with vSphere Trust Authority 327
4-71 VM Encryption with a Trusted Key Provider 328
4-72 Identifying VM Encryption Key Providers 329
4-73 Migrating a VM Encrypted with a Trusted Key Provider (1) 330
4-74 Migrating a VM Encrypted with a Trusted Key Provider (2) 331
4-75 Labs 332
4-76 Lab 10: Assigning a vSphere Trust Authority Administrator 333
4-77 Lab 11: Enabling and Configuring vSphere Trust Authority 334
4-78 Lab 12: Encrypting a VM with a Trusted Key Provider 335
4-79 Review of Learner Objectives 336
4-80 Lesson 4: Host Profiles 337
4-81 Learner Objectives 338
4-82 About Host Profiles 339
4-83 Implementing Host Profiles 341
4-84 Monitoring for Compliance 343
4-85 Applying Host Profiles 344
4-86 Host Customization 345
4-87 Managing Host Customizations (1) 347
4-88 Managing Host Customizations (2) 348
4-89 Standardization Across Multiple vCenter Server Instances 349
4-90 Lab 13: Using Host Profiles 350
4-91 Review of Learner Objectives 351
4-92 Lesson 5: Content Libraries 352
4-93 Learner Objectives 353
4-94 About Content Libraries 354
4-95 Content Library Benefits 355
4-96 Content Library Types 356
4-97 Storing Content in Content Libraries 358
4-98 Creating Content Libraries 359
4-99 Advanced Configuration 360

https://t.me/learningnets

4-100 Populating Content Libraries 361
4-101 Using the Content Library to Deploy VMs (1) 362
4-102 Using the Content Library to Deploy VMs (2) 363
4-103 Updating Content Library Items 364
4-104 Publishing Content Libraries for External Use 365
4-105 Content Library Subscriptions 366
4-106 Published and Subscribed Content Libraries 367
4-107 Synchronizing and Versioning Content Libraries 369
4-108 Content Library Maximums 371
4-109 Lab 14: Creating a Content Library 372
4-110 Review of Learner Objectives 373
4-111 Lesson 6: Resource Pools 374
4-112 Learner Objectives 375
4-113 About Resource Pools 376
4-114 Resource Pool Attributes 377
4-115 Reasons to Use Resource Pools 378
4-116 Resource Pool Example 379
4-117 Resource Pool Example: CPU Shares 380
4-118 Resource Pool Example: CPU Contention 381
4-119 About Expandable Reservations 382
4-120 Example of Expandable Reservation (1) 383
4-121 Example of Expandable Reservation (2) 384
4-122 Resource Pool Summary Tab 385
4-123 Resource Reservation Details 386
4-124 Resource Pool Shares 387
4-125 Diluted Shares 389
4-126 About Scalable Shares (1) 390
4-127 About Scalable Shares (2) 391
4-128 Enabling Scalable Shares 392
4-129 Lab 15: Managing Resource Pools 393
4-130 Review of Learner Objectives 394

https://t.me/learningnets

4-131 Key Points 395

Module 5 CPU Optimization 397

5-2 Importance 398
5-3 Module Lessons 399
5-4 Lesson 1: CPU Virtualization Concepts 400
5-5 Learner Objectives 401
5-6 About the CPU Scheduler 402
5-7 About VM Worlds 403
5-8 CPU Scheduler Features 404
5-9 CPU Scheduler Support for SMP VMs 405
5-10 Relaxed Co-Scheduling 406
5-11 CPU Performance Factors 407
5-12 Warning Sign: Ready Time 408
5-13 Transitioning Between World States 409
5-14 Activity: CPU Scheduler (1) 410
5-15 Activity: CPU Scheduler (2) 411
5-16 Security Boundaries 412
5-17 Review of Learner Objectives 413
5-18 Lesson 2: NUMA Architecture in vSphere 414
5-19 Learner Objectives 415
5-20 How the ESXi NUMA Scheduler Works 416
5-21 NUMA-Aware ESXi 417
5-22 Terminology: vCPUs, pCPUs, Cores, and More 418
5-23 About the NUMA Client 419
5-24 Wide VM NUMA Support 420
5-25 Virtual NUMA Topology 421
5-26 Optimizing the Number of Cores per Socket for VMs (1) 422
5-27 Optimizing the Number of Cores per Socket for VMs (2) 423
5-28 Optimizing the Number of Cores per Socket for VMs (3) 424
5-29 Activity: VM Cores Per Socket (1) 425

https://t.me/learningnets

5-30 Activity: VM Cores Per Socket (2) 426
5-31 Memory Considerations with vNUMA (1) 427
5-32 Memory Considerations with vNUMA (2) 428
5-33 Memory Considerations with vNUMA (3) 429
5-34 Memory Considerations with vNUMA (4) 430
5-35 vNUMA Considerations 431
5-36 Review of Learner Objectives 432
5-37 Lesson 3: Using esxtop 433
5-38 Learner Objectives 434
5-39 About esxtop (1) 435
5-40 About esxtop (2) 436
5-41 Getting Help in esxtop 437
5-42 Customizing Fields in esxtop 438
5-43 Review of Learner Objectives 439
5-44 Lesson 4: Monitoring CPU Activity 440
5-45 Learner Objectives 441
5-46 CPU Key Performance Indicators for ESXi Hosts 442
5-47 CPU Key Performance Indicators for VMs 443
5-48 Using esxtop to Monitor ESXi Host CPU 445
5-49 Using esxtop to View CPU Metrics per VM 446
5-50 Using esxtop to View Single CPU Statistics 448
5-51 Using esxtop to Monitor VM vCPU 449
5-52 Important Metrics to Monitor 450
5-53 Example: Identifying CPU Overcommitment 451
5-54 Host CPU Saturation 452
5-55 Resolving Host CPU Saturation 453
5-56 Activity: CPU Metrics (1) 454
5-57 Activity: CPU Metrics (2) 455
5-58 Guest CPU Saturation 456
5-59 Using One vCPU in an SMP VM 457
5-60 Low Guest CPU Use 459

https://t.me/learningnets

5-61 Introduction to Lab 8: Monitoring CPU Performance 460
5-62 Lab 16: Monitoring CPU Performance 461
5-63 Review of Lab: Monitoring CPU Performance 462
5-64 Review of Learner Objectives 463
5-65 Key Points 464

Module 6 Memory Optimization 465

6-2 Importance 466
6-3 Module Lessons 467
6-4 Lesson 1: Memory Virtualization Concepts 468
6-5 Learner Objectives 469
6-6 Virtual Memory Overview 470
6-7 Application and Guest OS Memory Management 471
6-8 VM Memory Management 472
6-9 Memory Reclamation 473
6-10 Reason for Reclaiming Memory 474
6-11 Guest OS Memory Terminology 475
6-12 VM Memory Reclamation Techniques 477
6-13 Transparent Page Sharing (1) 479
6-14 Transparent Page Sharing (2) 481
6-15 Impact of Hardware MMU on Transparent Page Sharing 482
6-16 Impact of NUMA Architecture on Transparent Page Sharing 484
6-17 Transparent Page Sharing and Intra-VM Memory Sharing 485
6-18 Using Salting with Transparent Page Sharing (1) 486
6-19 Using Salting with Transparent Page Sharing (2) 487
6-20 Using Salting with Transparent Page Sharing (3) 488
6-21 Using Salting with Transparent Page Sharing (4) 489
6-22 Using Salting with Transparent Page Sharing (5) 490
6-23 Memory Ballooning in the Guest OS 491
6-24 Reclaiming Memory with Ballooning 493
6-25 Memory Compression 494

https://t.me/learningnets

6-26 Reclaiming Memory with Host Swapping 496
6-27 Activity: Identifying Memory Reclamation Features (1) 497
6-28 Activity: Identifying Memory Reclamation Features (2) 498
6-29 Sliding Scale Mem.MemMinFreePct 499
6-30 Criteria for Reclaiming Host Memory 500
6-31 Transitioning from One Memory State to Another (1) 502
6-32 Transitioning from One Memory State to Another (2) 504
6-33 Review of Learner Objectives 506
6-34 Lesson 2: Monitoring Memory Activity 507
6-35 Learner Objectives 508
6-36 Memory Usage Metrics in the Guest OS 509
6-37 About Consumed Host Memory and Active Guest Memory 510
6-38 Using esxtop to Monitor Memory Usage 511
6-39 Host Ballooning Activity in esxtop (1) 513
6-40 Host Ballooning Activity in esxtop (2) 515
6-41 Host Memory Compression Activity in esxtop 516
6-42 Monitoring Host Cache Swapping 518
6-43 Host Swapping Activity in esxtop: Memory Screen 519
6-44 Host Swapping Activity in esxtop: CPU Screen 520
6-45 Causes of Active Host-Level Swapping 521
6-46 Activity: Identifying esxtop Fields (1) 523
6-47 Activity: Identifying esxtop Fields (2) 524
6-48 Resolving Host-Level Swapping 525
6-49 Reducing Memory Overcommitment 526
6-50 Enabling the Balloon Driver in VMs 528
6-51 Reducing VM Memory Reservation 529
6-52 Dedicating Memory to Critical VMs 530
6-53 Introduction to Lab: Monitoring Memory Performance 531
6-54 Lab 17: Monitoring Memory Performance 532
6-55 Review of Lab: Monitoring Memory Performance 533
6-56 Review of Learner Objectives 534

https://t.me/learningnets

6-57 Key Points 535

Module 7 Storage Optimization 537

7-2 Importance 538
7-3 Module Lessons 539
7-4 Lesson 1: Storage Virtualization Concepts 540
7-5 Learner Objectives 541
7-6 Storage Performance Overview 542
7-7 Storage Protocol Performance 543
7-8 About SAN Configuration 544
7-9 About Storage Queues 546
7-10 Network Storage: iSCSI and NFS 547
7-11 Review of Learner Objectives 548
7-12 Lesson 2: Enhancements to NVMe Storage Systems 549
7-13 Learner Objectives 550
7-14 Enhanced Support for NVMe 551
7-15 Supported NVMe Topologies 552
7-16 FC-NVMe and NVMe over RDMA (RoCE v2) 553
7-17 Support for FC-NVMe 554
7-18 Requirements for FC-NVMe 555
7-19 Support for NVMe over RDMA (RoCE v2) 556
7-20 Requirements for NVMe over RDMA (RoCE v2) 557
7-21 Configuring ESXi for NVMe over RDMA (RoCE v2) (1) 558
7-22 Configuring ESXi for NVMe over RDMA (RoCE v2) (2) 559
7-23 Configuring ESXi for NVMe over RDMA (RoCE v2) (3) 560
7-24 Configuring ESXi for NVMe over RDMA (RoCE v2) (4) 561
7-25 Pluggable Storage Architecture 562
7-26 High-Performance Plug-In for NVMe 563
7-27 Support for High-Performance Plug-In Multipathing 564
7-28 High-Performance Plug-In Best Practices 565
7-29 Review of Learner Objectives 566

https://t.me/learningnets

7-30 Lesson 3: Support for iSER 567
7-31 Learner Objectives 568
7-32 Support for iSER-Enabled vSphere Virtual Volumes Arrays 569
7-33 Configuring iSER 570
7-34 Requirements for iSER 571
7-35 Running iSER 572
7-36 Review of Learner Objectives 573
7-37 Lesson 4: Monitoring Storage Activity 574
7-38 Learner Objectives 575
7-39 Where Storage Problems Can Occur 576
7-40 Storage Key Performance Indicators 577
7-41 Monitoring Disk Throughput by Storage Adapter (1) 578
7-42 Monitoring Disk Throughput by Storage Adapter (2) 580
7-43 Correlating Storage Devices to Datastores (1) 581
7-44 Correlating Storage Devices to Datastores (2) 582
7-45 Monitoring Path Statistics for a Specific Storage Device (1) 583
7-46 Monitoring Path Statistics for a Specific Storage Device (2) 584
7-47 Monitoring Individual VMDK Statistics for a VM (1) 585
7-48 Monitoring Individual VMDK Statistics for a VM (2) 586
7-49 Disk Latency Metrics 587
7-50 Monitoring Disk Latency with esxtop 589
7-51 Monitoring Commands and Command Queuing 590
7-52 Example: Disk Latency and Queuing 591
7-53 Example: Bad Disk Throughput 593
7-54 Activity: Hardware Disk Latency (1) 594
7-55 Activity: Hardware Disk Latency (2) 595
7-56 Overloaded Storage and Command Aborts 596
7-57 Unexpected Increase in I/O Latency on Shared Storage 597
7-58 Storage Performance Best Practices 599
7-59 Introduction to Lab: Monitoring Storage Performance 600
7-60 Lab 18: Monitoring Storage Performance 601

https://t.me/learningnets

7-61 Review of Lab: Monitoring Storage Performance 602
7-62 Review of Learner Objectives 603
7-63 Key Points 604

Module 8 Network Optimization 605

8-2 Importance 606
8-3 Module Lessons 607
8-4 Lesson 1: Network Virtualization Concepts 608
8-5 Learner Objectives 609
8-6 Network I/O Virtualization Overhead 610
8-7 VMXNET Network Adapters 612
8-8 Virtual Network Adapters 614
8-9 About VMXNET3 Optimization 616
8-10 Configuring a Single CPU Thread per vNIC 618
8-11 Configuring Multiple CPU Threads per vNIC 619
8-12 About TCP Segmentation Offload 621
8-13 About Jumbo Frames 623
8-14 About SplitRx Mode 624
8-15 About SplitTx Mode 626
8-16 Activity: vSphere Networking Features (1) 627
8-17 Activity: vSphere Networking Features (2) 628
8-18 Review of Learner Objectives 629
8-19 Lesson 2: Monitoring Network I/O Activity 630
8-20 Learner Objectives 631
8-21 Monitoring for Performance in the Network I/O Stack 632
8-22 Network Capacity Metrics 633
8-23 About esxtop Network Statistics 634
8-24 About esxtop Network Output 635
8-25 Dropped Network Packets 637
8-26 Dropped Packets in vNICs 639
8-27 Dropped Received Packets in Physical NICs 641

https://t.me/learningnets

8-28 Dropped Transmit Packets in Physical NICs 642
8-29 Unexpected Increases in Data Transfer Rate 643
8-30 Networking Best Practices 644
8-31 Introduction to Lab: Monitoring Network Performance (1) 646
8-32 Introduction to Lab: Monitoring Network Performance (2) 648
8-33 Lab 19: Monitoring Network Performance 649
8-34 Review of Lab: Monitoring Network Performance 650
8-35 Review of Learner Objectives 651
8-36 Key Points 652

Module 9 vCenter Server Performance Optimization 653

9-2 Importance 654
9-3 Lesson 1: vCenter Server Performance Optimization 655
9-4 Learner Objectives 656
9-5 vCenter Server Performance Considerations 657
9-6 vCenter Server CPU Performance 658
9-7 vCenter Server Memory Performance 659
9-8 Monitoring CPU and Memory Use with the VAMI 660
9-9 Using vimtop to Monitor vCenter Server Appliance Resources 661
9-10 vimtop Commands and Columns 662
9-11 Monitoring CPU Use with vimtop 663
9-12 Activity: CPU and Memory Use (1) 664
9-13 Activity: CPU and Memory Use (2) 665
9-14 Monitoring Memory Use with vimtop 666
9-15 Viewing Service Memory Allocation with cloudvm-ram-size 667
9-16 Viewing Service Heap Size 668
9-17 Changing Service Heap Size 669
9-18 vCenter Server Network Performance 670
9-19 Monitoring Network Activity with the VAMI 671
9-20 Monitoring Network Activity with vimtop 672
9-21 vCenter Server Database Performance (1) 673

https://t.me/learningnets

9-22 vCenter Server Database Performance (2) 674
9-23 Effects of Changing Statistics Level on Database Traffic 675
9-24 vCenter Server Appliance Disk Use (1) 677
9-25 vCenter Server Appliance Disk Use (2) 678
9-26 vCenter Server Appliance Disk Use (3) 679
9-27 Monitoring Disk Use with the VAMI 680
9-28 Monitoring Disk Usage with the df Command 681
9-29 Monitoring Disk Activity with vimtop 682
9-30 Monitoring Database Activity with the VAMI 683
9-31 Activity: Network Performance (1) 684
9-32 Activity: Network Performance (2) 685
9-33 Activity: vCenter Server Tools (1) 686
9-34 Activity: vCenter Server Tools (2) 687
9-35 Review of Learner Objectives 688
9-36 Key Points 689

Module 10 Appendix: vSphere Auto Deploy 691

10-2 Importance 692
10-3 Appendix Lessons 693
10-4 Lesson 1: vSphere ESXi Image Builder 694
10-5 Learner Objectives 695
10-6 ESXi Images 696
10-7 vSphere Installation Bundles 697
10-8 ESXi Image Deployment 698
10-9 vSphere ESXi Image Builder Overview 699
10-10 vSphere ESXi Image Builder Architecture 700
10-11 Building ESXi Images 701
10-12 Step 1: Starting the Image Builder Service 702
10-13 Step 2: Connecting to a Software Depot 703
10-14 Step 3: Creating an Image Profile 704
10-15 Step 4: Generating the New ESXi Image 705

https://t.me/learningnets

10-16 Review of Learner Objectives 706
10-17 Lesson 2: Overview of vSphere Auto Deploy 707
10-18 Learner Objectives 708
10-19 About vSphere Auto Deploy 709
10-20 vSphere Auto Deploy Modes 711
10-21 vSphere Auto Deploy Stateless Mode 712
10-22 vSphere Auto Deploy Mode: Stateless Caching 713
10-23 vSphere Auto Deploy Mode: Stateful Installation 714
10-24 Locations for Configuration and State Information 715
10-25 vSphere Auto Deploy Architecture 717
10-26 Rules Engine Basics 719
10-27 PXE Boot Infrastructure Setup 721
10-28 Initial Boot of an Autodeployed ESXi Host (1) 722
10-29 Initial Boot of an Autodeployed ESXi Host (2) 723
10-30 Initial Boot of an Autodeployed ESXi Host (3) 724
10-31 Initial Boot of an Autodeployed ESXi Host (4) 725
10-32 Initial Boot of an Autodeployed ESXi Host (5) 726
10-33 Subsequent Boot of a Stateless ESXi Host (1) 727
10-34 Subsequent Boot of a Stateless ESXi Host (2) 728
10-35 Subsequent Boot of a Stateless ESXi Host (3) 729
10-36 Subsequent Boot of a Stateless ESXi Host (4) 730
10-37 Activity: vSphere Auto Deploy Modes (1) 731
10-38 Activity: vSphere Auto Deploy Modes (2) 732
10-39 Running Custom Scripts on Stateless Hosts 733
10-40 Booting Many Stateless Hosts 734
10-41 Stateless Caching Host Profile Configuration 735
10-42 Subsequent Boot of a Stateless Caching Host 736
10-43 Stateful Installation Host Profile Configuration 738
10-44 Subsequent Boot of a Stateful Installation Host 739
10-45 Stateful Installation Using vSphere Lifecycle Manager 740
10-46 Review of Learner Objectives 741

https://t.me/learningnets

10-47 Lesson 3: Configuring vSphere Auto Deploy 742
10-48 Learner Objectives 743
10-49 Configuring a vSphere Auto Deploy Environment 744
10-50 Step 1: Preparing the DHCP Server 745
10-51 Step 2: Starting the vSphere Auto Deploy Service 746
10-52 Step 3: Preparing the TFTP Server 747
10-53 Step 4: Creating Deployment Rules 748
10-54 Step 5: Activating Deployment Rules 749
10-55 Managing the vSphere Auto Deploy Environment 750
10-56 Using vSphere Auto Deploy with vSphere Update Manager 751
10-57 Review of Learner Objectives 752
10-58 Key Points 753

Module 11 Appendix: vSphere Security 755

11-2 Importance 756
11-3 Appendix Lessons 757
11-4 Lesson 1: Configuring ESXi Host Access and Authentication 758
11-5 Learner Objectives 759
11-6 About ESXi Host Security Properties 760
11-7 Configuring the ESXi Firewall 761
11-8 Configuring ESXi Services 762
11-9 About Lockdown Mode (1) 763
11-10 About Lockdown Mode (2) 764
11-11 Normal Lockdown Mode Behavior 765
11-12 Strict Lockdown Mode Behavior 766
11-13 Integrating ESXi with Active Directory 767
11-14 VMware vSphere Authentication Proxy Service 769
11-15 Review of Learner Objectives 770
11-16 Lesson 2: Securing vSphere 771
11-17 Learner Objectives 772
11-18 About Securing vSphere 773

https://t.me/learningnets

11-19 About the vSphere Security Configuration Guide 774
11-20 Controlling Access to vCenter Server Systems 775
11-21 vCenter Server Access Control 776
11-22 Securing vCenter Server Appliance 777
11-23 Securing vSphere Management Networks 779
11-24 Securing ESXi Hosts 781
11-25 VM Protection Overview 783
11-26 Review of Learner Objectives 784
11-27 Lesson 3: Security Standards and Protocols 785
11-28 Learner Objectives 786
11-29 Securing vCenter Server with TLS 1.2 787
11-30 Activity: ESXi Secure Boot (1) 788
11-31 Activity: ESXi Secure Boot (2) 789
11-32 vSphere Compliance with FIPS 140-2 (1) 790
11-33 vSphere Compliance with FIPS 140-2 (2) 791
11-34 FIPS 140-2 Architecture 792
11-35 Enabling and Disabling FIPS 140-2 Mode on ESXi Hosts 793
11-36 Enabling and Disabling FIPS 140-2 Mode on vCenter Server Appliance 794
11-37 Activity: Security Technologies (1) 795
11-38 Activity: Security Technologies (2) 796
11-39 Review of Learner Objectives 797
11-40 Lesson 4: Securing VMs 798
11-41 Learner Objectives 799
11-42 VM Protection with UEFI Secure Boot 800
11-43 Using vTPM in a VM (1) 802
11-44 Using vTPM in a VM (2) 803
11-45 Prerequisites for Adding vTPM 804
11-46 Virtualization-Based Security for Windows Guest Operating Systems 805
11-47 VBS Requirements and Limitations 807
11-48 Enabling VBS 808
11-49 About SGX 809

https://t.me/learningnets

11-50 vSGX Use Cases 810
11-51 SGX Terminology 811
11-52 Intel SGX Architecture 812
11-53 vSGX Requirements 813
11-54 vSGX Restrictions 814
11-55 Enabling and Disabling vSGX for a VM 815
11-56 Review of Learner Objectives 816
11-57 Lesson 5: VM Encryption 817
11-58 Learner Objectives 818
11-59 About VM Encryption 819
11-60 Overview of VM Encryption 820
11-61 Advantages of VM Encryption 821
11-62 Business Use Case: Securing VMs 822
11-63 VM Encryption Interoperability (1) 823
11-64 VM Encryption Interoperability (2) 825
11-65 VM Encryption Architecture 827
11-66 VM Encryption Prerequisites 828
11-67 Encrypting New VMs 829
11-68 Encrypting Existing VMs 830
11-69 Unlocking Encrypted VMs (1) 831
11-70 Unlocking Encrypted VMs (2) 832
11-71 About the KMS 833
11-72 KMIP Client Certificate 834
11-73 Key Management Clusters 835
11-74 Summary of Keys 836
11-75 Role of vCenter Server in VM Encryption 837
11-76 KMS and vCenter Server Communication 838
11-77 Making an ESXi Host Cryptographically Safe 839
11-78 VM Encryption Process Flow (1) 840
11-79 VM Encryption Process Flow (2) 841
11-80 Activity: VM Keys (1) 842

https://t.me/learningnets

11-81 Activity: VM Keys (2) 843
11-82 Managing VM Encryption 844
11-83 vCenter Server Role: No Cryptography Administrator 845
11-84 Activity: Cryptographic Privileges (1) 846
11-85 Activity: Cryptographic Privileges (2) 847
11-86 Pushing Keys to Hosts 848
11-87 Cross vCenter Migration and Cloning of an Encrypted VM 849
11-88 Architecture for Cross vCenter Server Migrations and Clones (1) 850
11-89 Architecture for Cross vCenter Server Migrations and Clones (2) 851
11-90 Architecture for Cross vCenter Server Migrations and Clones (3) 852
11-91 Performing a Shallow Recrypt of VMs with Snapshots 853
11-92 VM Encryption Operations When Cloning a VM 854
11-93 Cloning an Encrypted VM with Instant Clone Technology 855
11-94 Cloning a VM: Encrypting or Decrypting the Destination VM 856
11-95 System Requirements When Cloning a VM 857
11-96 Backing Up Encrypted VMs (1) 858
11-97 Backing Up Encrypted VMs (2) 859
11-98 About Encrypted Core Dumps 860
11-99 Core Dump Encryption 861
11-100 Providing Passwords for Encrypted Core Dumps 862
11-101 VM Encryption Events and Alarms 863
11-102 Warning: Encrypted Core Dump Exists (1) 864
11-103 Warning: Encrypted Core Dump Exists (2) 865
11-104 Critical Event: Insufficient Disk Space 866
11-105 Critical Event: Key Generation Failed 867
11-106 Critical Event: KMS Unreachable 868
11-107 About Encrypted vSphere vMotion 869
11-108 Configuring Encrypted vSphere vMotion 871
11-109 Encrypted vSphere vMotion Requirements 872
11-110 Review of Learner Objectives 873
11-111 Key Points 874

https://t.me/learningnets

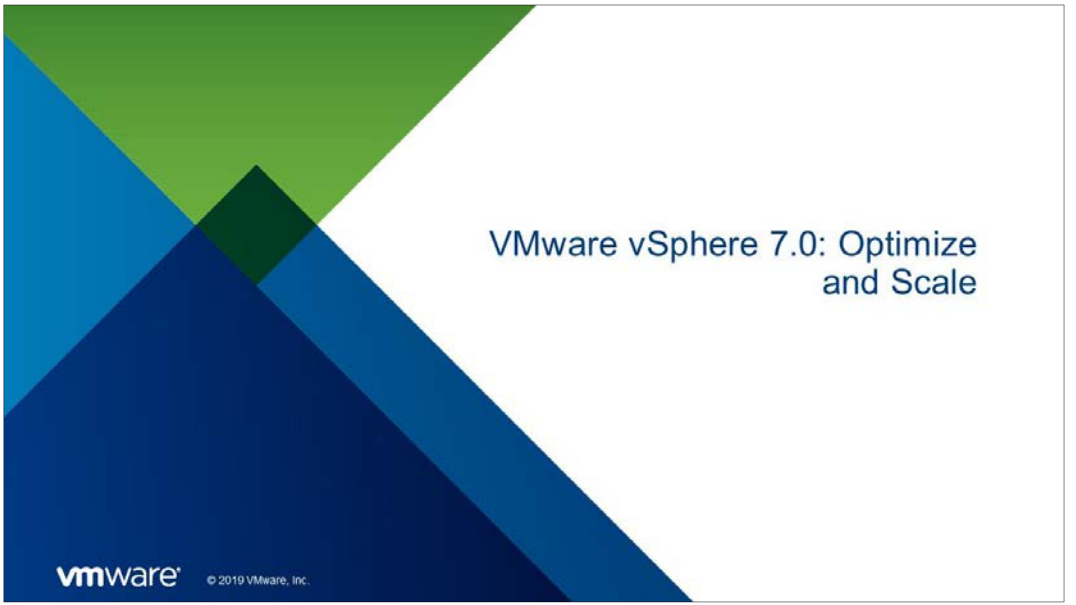
11-112 Auto Deploy 1 878
11-113 Auto Deploy 2 879
11-114 vSphere Security 1 882
11-115 vSphere Security 2 883

https://t.me/learningnets

https://t.me/learningnets

Module 1

Course Introduction



https://t.me/learningnets

https://t.me/learningnets

1-3 Importance

To configure and maintain a highly available and scalable vSphere environment, you must optimize and scale ESXi hosts and vCenter Server instances.

https://t.me/learningnets

1-4 Learner Objectives

After completing this course, you should be able to meet the following objectives:

- Configure and manage vSphere networking and storage for a large enterprise
- Use the vSphere Client and vSphere ESXi Shell to manage vSphere
- Harden the vSphere environment against security threats
- Create a content library for deploying virtual machines
- Monitor and analyze key performance indicators for compute, storage, and networking resources for ESXi hosts
- Optimize the performance of ESXi and vCenter Server

https://t.me/learningnets

1-5 Course Outline

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. Appendix: vSphere Auto Deploy
11. Appendix: vSphere Security

Those who participate in the VMware vSphere: Optimize and Scale instructor-led training receive additional recorded lecture material as an exclusive benefit from VMware Learning Zone. This material is available to participants who sign up at the VMware Learning Zone for a free trial access period at https://vmwarelearningzone.vmware.com/oltpublish/site/program.do?dispatch=showCourseSession&id=0dd4d914-733a-11ea-9f48-0cc47adeb5f8

https://t.me/learningnets

1-6 Typographical Conventions

The following typographical conventions are used in this course.

Conventions	Use and Examples
`Monospace`	Identifies command names, command options, parameters, code fragments, error messages, filenames, folder names, directory names, and path names: • Run the `esxtop` command. • ... found in the `var/log/messages` file.
`Monospace Bold`	Identifies user inputs: • Enter **`ipconfig/release`**.
Boldface	Identifies user interface controls: • Click the **Configuration** tab.
Italic	Identifies book titles: • *vSphere Virtual Machine Administration*
< >	Indicates placeholder variables: • <ESXi_host_name> • ... the `Settings/<Your_Name>.txt` file

https://t.me/learningnets

1-7 References (1)

Title	Location
vSphere ESXi Installation and Setup	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-70-installation-setup-guide.pdf
vCenter Server Installation and Setup	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-vcenter-server-70-installation-guide.pdf
vCenter Server and Host Management	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-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

1-8 References (2)

Title	Location
vSphere Virtual Machine Administration	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-virtual-machine-admin-guide.pdf
vSphere Networking	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-networking-guide.pdf
vSphere Storage	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-storage-guide.pdf
vSphere Resource Management	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-resource-management-guide.pdf
vSphere Host Profiles	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-host-profiles-guide.pdf
vSphere Security	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-security-guide.pdf
vSphere Monitoring and Performance	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-monitoring-performance-guide.pdf

https://t.me/learningnets

1-9 VMware Online Resources

VMware vSphere Blog: http://blogs.vmware.com/vsphere
VMware Communities: http://communities.vmware.com
- Start a discussion.
- Access the knowledge base.
- Access documentation, technical papers, and compatibility guides.
- Access communities.
- Access user groups.

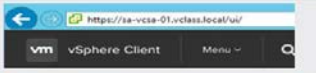
VMware Support: http://www.vmware.com/support
VMware Hands-on Labs: http://hol.vmware.com
VMware Education: http://www.vmware.com/education
- Access course catalog and worldwide course schedule.

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

https://t.me/learningnets

1-10 Review of vSphere User Interfaces

UI	Description
vSphere Client	• Connects to vCenter Server • HTML5-based client • https://*vCenter Server*/ui
VMware Host Client	• Connects directly to an ESXi host • HTML5-based client • https://*ESXi host*/ui
Data Center CLI (DCLI)	• A simplified command-line interface (CLI) that you can use to automate tasks in your VMware Cloud on AWS, NSX-T, and vCenter Server environments
esxcli	• A CLI framework in vSphere that provides a modular architecture for various components called namespaces (nmp, corestorage, and swiscsi) running in the VMkernel
VMware PowerCLI	• Command-line and scripting tool built on Windows PowerShell



The Flash-based client, known as the vSphere Web Client, is no longer available.

https://t.me/learningnets

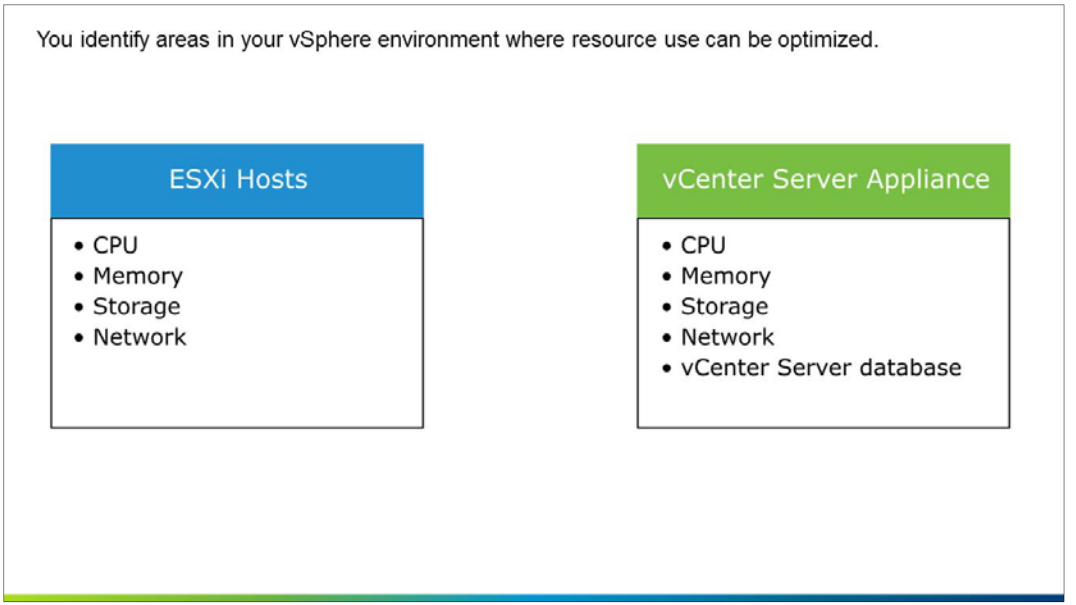
1-11 About Scalability

With vSphere technology, you can build a scalable infrastructure.

Networking	• vSphere distributed switches
Storage	• VMFS5 and VMFS6 • Storage APIs • vSAN and vSphere Virtual Volumes • Storage policies • Datastore clusters and vSphere Storage DRS
Hosts and VMs	• Content libraries • Host profiles • Resource pools

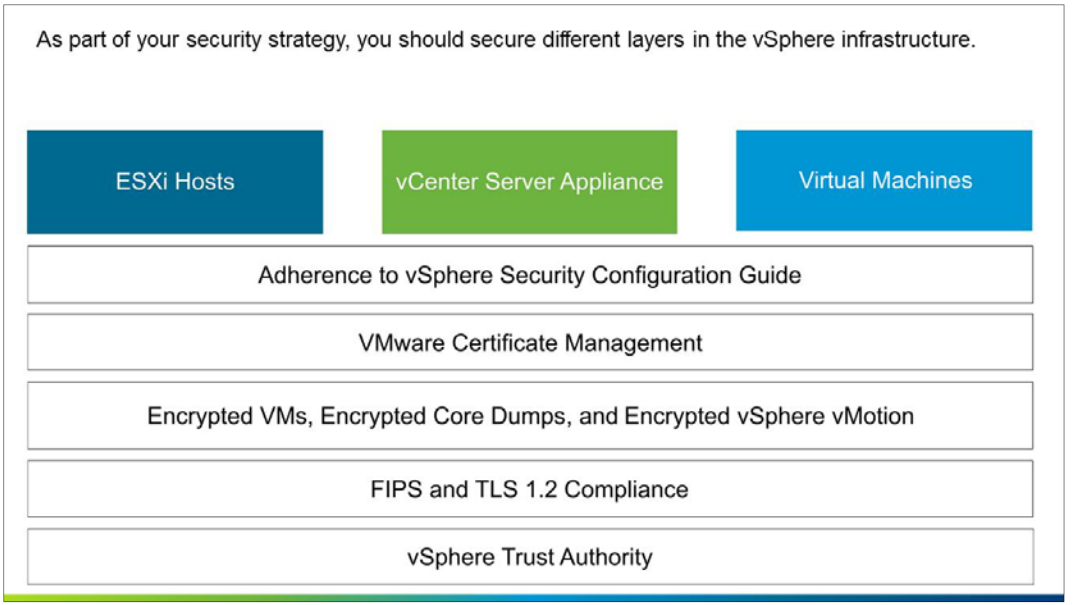
https://t.me/learningnets

1-12 About Optimization



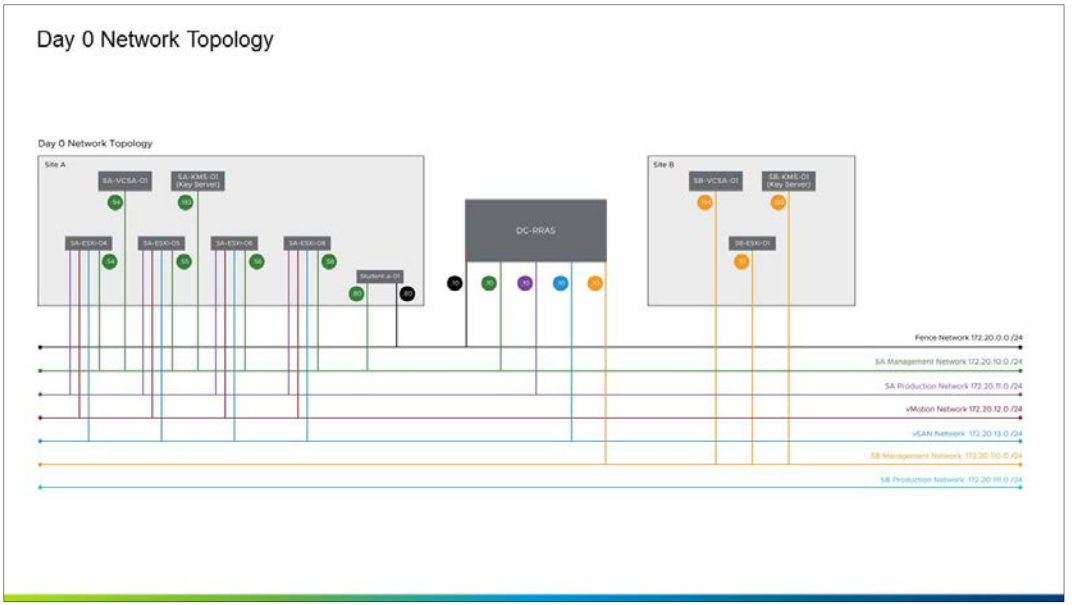
https://t.me/learningnets

1-13 About Security



https://t.me/learningnets

1-14 Data Center Layout



The data center used in the labs consists of Site A and Site B networks that share a domain controller (DC-RRAS).

Site A has three ESXi 7.0 hosts (SA-ESXi-04, SA-ESXi-05, and SA-ESXi-06), a vCenter Server Appliance (SA-VCSA-01), a Key Server (SA-KMS-01), and an administrator system (Student-a-01). These servers are managed through the SA Management Network (172.20.10.0/24).

Site B has one ESXi 7.0 host (SB-ESXi-01), a vCenter Server Appliance (SB-VCSA-01), and a Key Server (SB-KMS-01). These servers are managed through the SB Management Network (172.20.110.0/24).

These servers are connected through the following networks:

- SA Production Network (172.20.11.0/24)
- vMotion Network (172.20.12.0/24)

https://t.me/learningnets

- vSAN Network (172.20.13.0/24)
- SB Production Network (172.20.111.0/24)

https://t.me/learningnets

1-15 VMware Education Overview

Your instructor will introduce other Education Services offerings available to you:

- VMware Learning Paths:
  - Help you find the course that you need based on the product, your role, and your level of experience
  - Can be accessed at https://www.vmware.com/education
- VMware Learning Zone, which is the official source of digital training, includes the following options:
  - On Demand Courses: Self-paced learning that combines lecture modules with hands-on practice labs
  - VMware Lab Connect: Self-paced, technical lab environment that lets you practice skills learned during instructor-led training
  - Certification Exam Prep: Comprehensive video-based reviews of exam topics and objectives to help you take your certification exam
- For more information, see https://vmwarelearningzone.vmware.com.

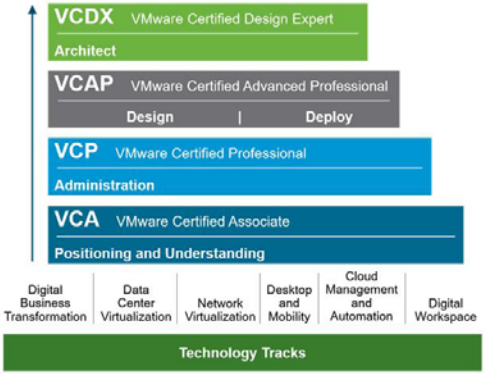
https://t.me/learningnets

1-16 VMware Certification Overview

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

Certifications are grouped into technology tracks. Each track offers one or more levels of certification (up to four levels).

For the complete list of certifications and details about how to attain these certifications, see https://vmware.com/certification.



https://t.me/learningnets

1-17 VMware Digital Badge Overview

VMware digital badges are emerging as the IT industry standard to recognize achievements, and to nurture and progress talent.
Digital badges have the following features:

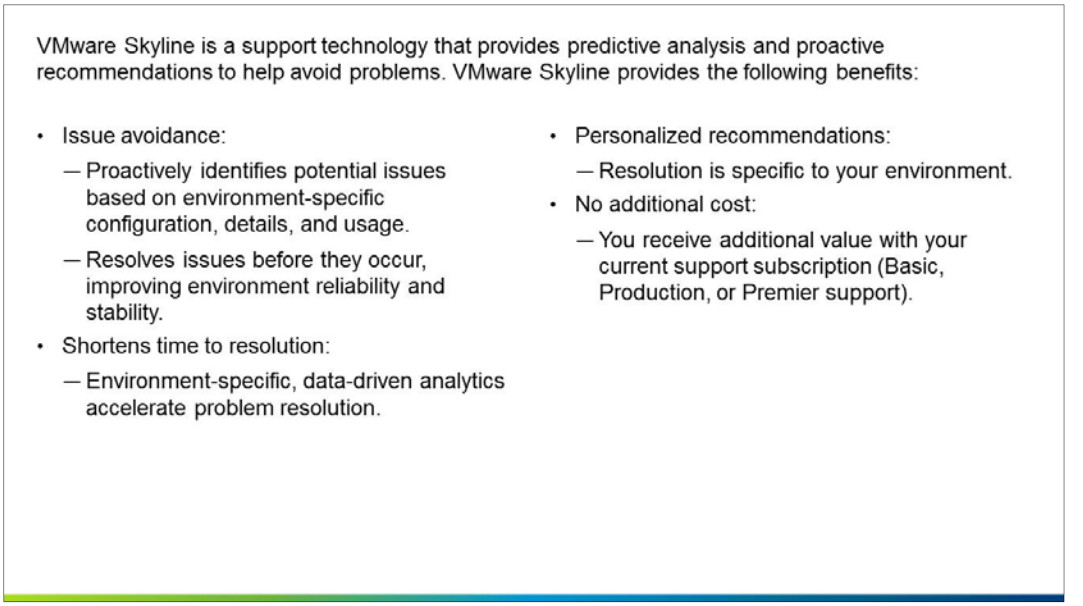
- Digital emblem that symbolizes skills and achievements
- Easy to share in social media (LinkedIn, Twitter, Facebook, blogs, and so on)
- Tethered to VMware to validate and verify achievement

For the complete list of digital badges, see http://www.pearsonvue.com/vmware/badging.

Digital badges contain metadata with skill tags and accomplishments, and are based on Mozilla's Open Badges standard.

https://t.me/learningnets

1-18 About VMware Skyline



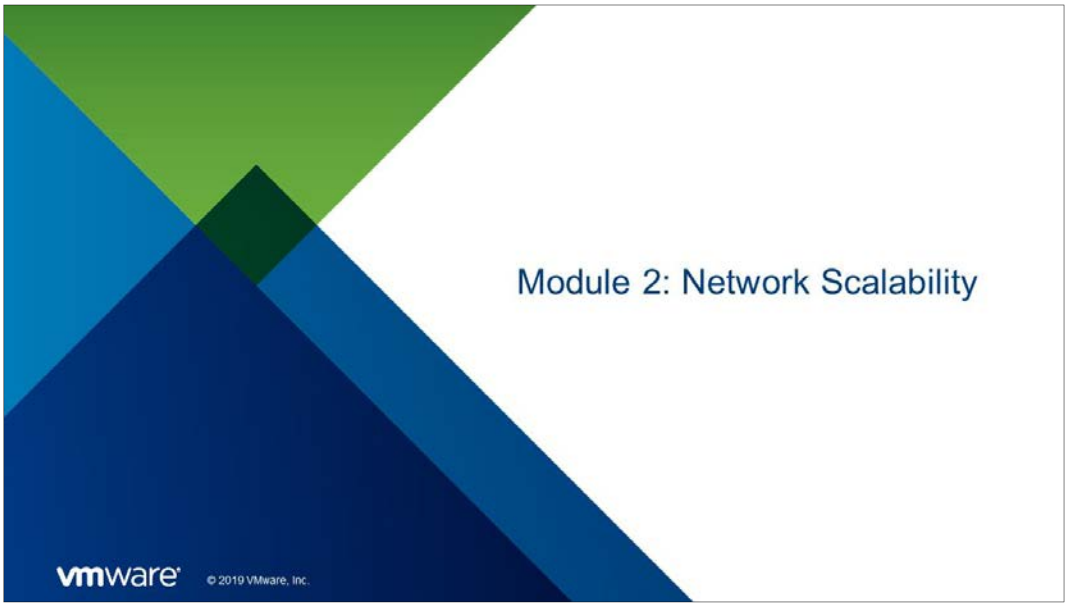
VMware Skyline shortens the time it takes to resolve a problem so that you can get back to business quickly. The VMware Technical Support engineers can use VMware Skyline to view your environment's configuration and the specific, data-driven analytics to help speed up problem resolution.

https://t.me/learningnets

https://t.me/learningnets

Module 2

Network Scalability



https://t.me/learningnets

2-2 Importance

As you scale your vSphere environment, the features and functions of vSphere Distributed Switch can help you manage networking.

https://t.me/learningnets

2-3 Module Lessons

1. Creating vSphere Distributed Switches
2. Managing vSphere Distributed Switches
3. Using Network I/O Control in Distributed Switches
4. NetFlow and Port Mirroring in Distributed Switches

https://t.me/learningnets

https://t.me/learningnets

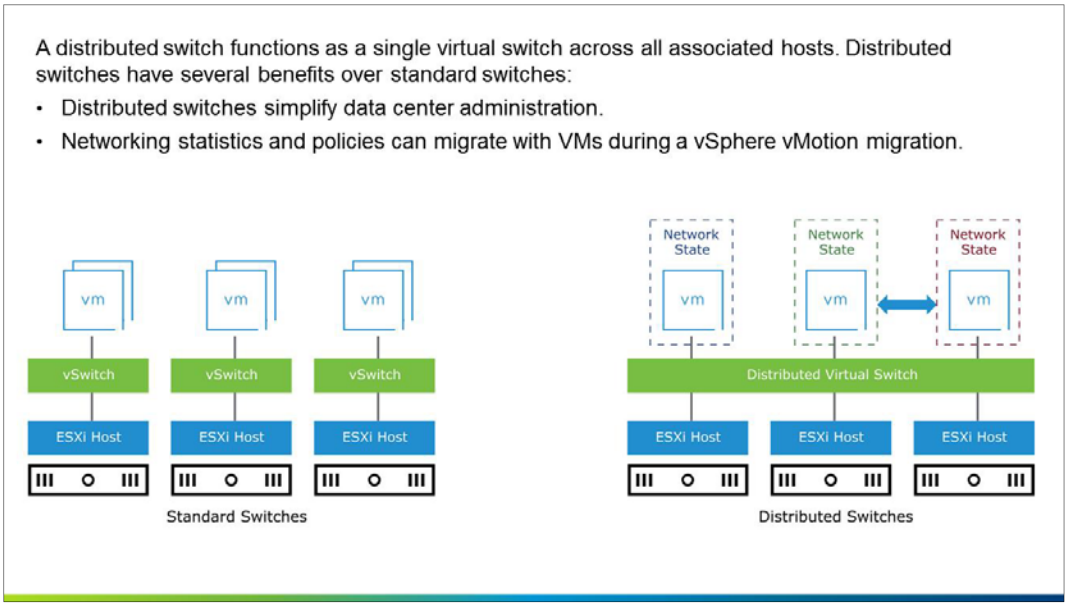
2-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe distributed switches
- Recognize how the NSX virtual distributed switch is supported in vSphere 7
- Identify the benefits and features of distributed switches
- Create a distributed switch

https://t.me/learningnets

2-6 About Distributed Switches

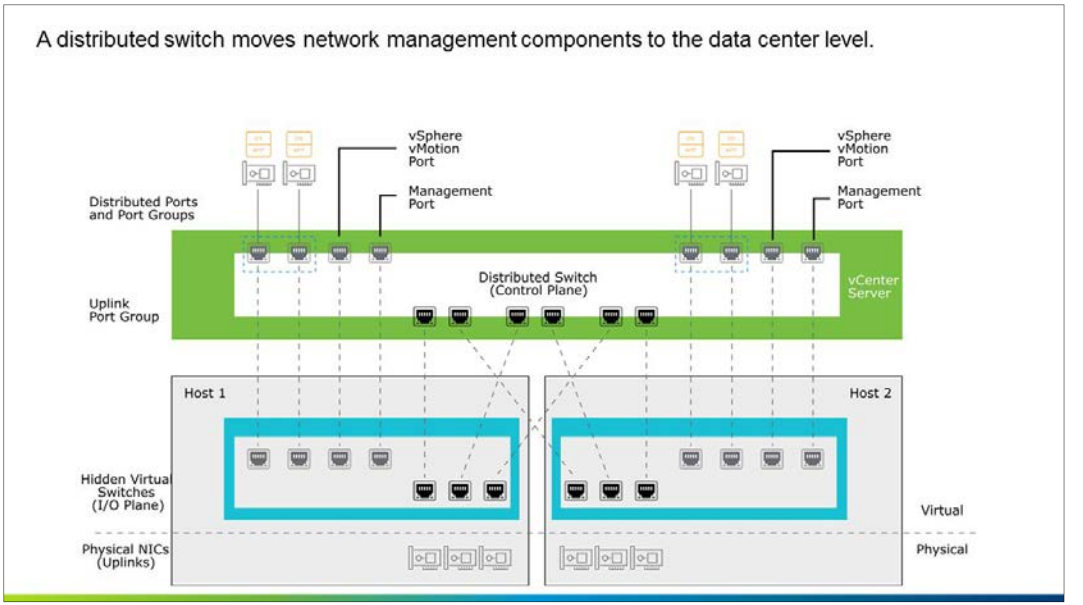


Having the network configuration at the data center level (vSphere Distributed Switch), not at the host level (standard switch), gives distributed switches the following advantages:

- Data center setup and administration are simplified through this centralized network configuration. For example, adding a host to a cluster and making it compatible with vSphere vMotion is much easier than with a standard switch.
- Distributed ports migrate with their clients. For example, when you migrate a VM with vSphere vMotion, the distributed port statistics and policies move with the VM, which simplifies debugging and troubleshooting.

https://t.me/learningnets

2-7 Distributed Switch Architecture



Managed by vCenter Server, a distributed switch is a logical entity that you can use to create and maintain a consistent virtual networking configuration throughout all your 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 configures 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 VM or a VMkernel interface, to a distributed port. vCenter Server stores the state of distributed ports in the vCenter Server database.

https://t.me/learningnets

With a distributed port group, you can 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 VMs on different hosts can communicate with each other only if both VMs have uplinks in the same broadcast domain.

https://t.me/learningnets

2-8 Standard and Distributed Switches: Shared Features

Standard and distributed switches share some features.

Feature	Standard Switch	Distributed Switch
Layer 2 switch	✓	✓
VLAN segmentation	✓	✓
IPv6 support	✓	✓
802.1Q tagging	✓	✓
NIC teaming	✓	✓
Outbound traffic shaping	✓	✓

https://t.me/learningnets

2-9 Distributed Switch Features

Distributed switches have several features that standard switches do not have.

Feature	Standard Switch	Distributed Switch
Inbound traffic shaping		✓
Configuration backup and restore		✓
Private VLANs		✓
Link Aggregation Control Protocol		✓
Data center level management		✓
vSphere vMotion migration of network state		✓
Network I/O Control		✓
Per-port policy settings		✓
Port state monitoring		✓
NetFlow		✓
Port mirroring		✓
Support for NSX-T Data Center		✓

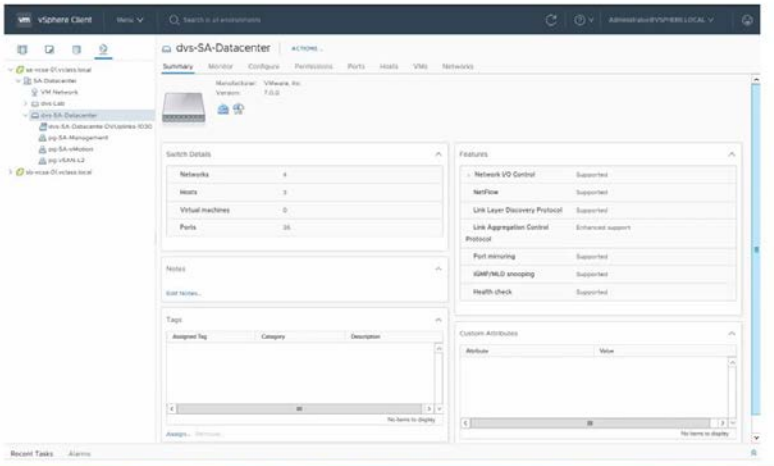
During a vSphere vMotion migration, a distributed switch tracks the virtual networking state (for example, counters and port statistics) as the VM moves from host to host. This tracking provides a consistent view of a virtual network interface, regardless of the VM location or vSphere vMotion migration history.

Tracking simplifies network monitoring and troubleshooting activities where vSphere vMotion is used to migrate VMs between hosts.

https://t.me/learningnets

2-10 Viewing VDS in vSphere Client

In the vSphere Client, you can access VDS objects.



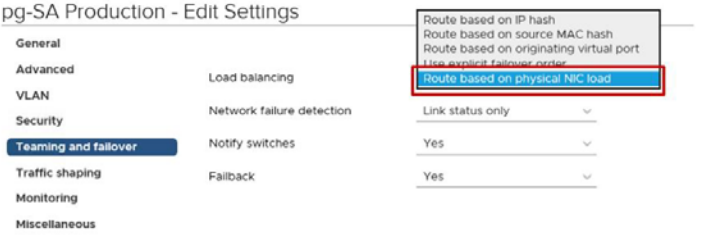
https://t.me/learningnets

2-11 Load-Based NIC Teaming

Load-based NIC teaming ensures that physical NIC capacity in a NIC team is optimized. This feature works in the following ways:

- It 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**.



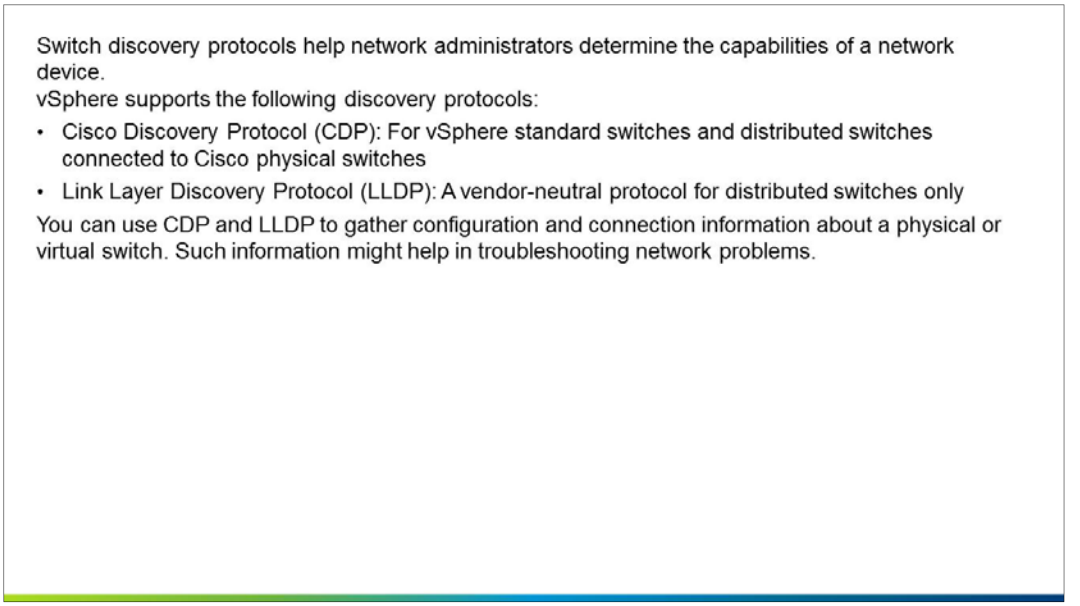
Load-based NIC teaming checks the real load of the uplinks and reduces the load on overloaded uplinks. No changes on the physical switch are required.

The distributed switch calculates uplinks for VMs by checking the VM 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 VM with the highest I/O is moved to a different uplink.

Load-based NIC teaming is not the default teaming policy. You must configure the policy to use it.

https://t.me/learningnets

2-12 Discovery Protocols



Discovery protocols are available in vSphere virtual switches.

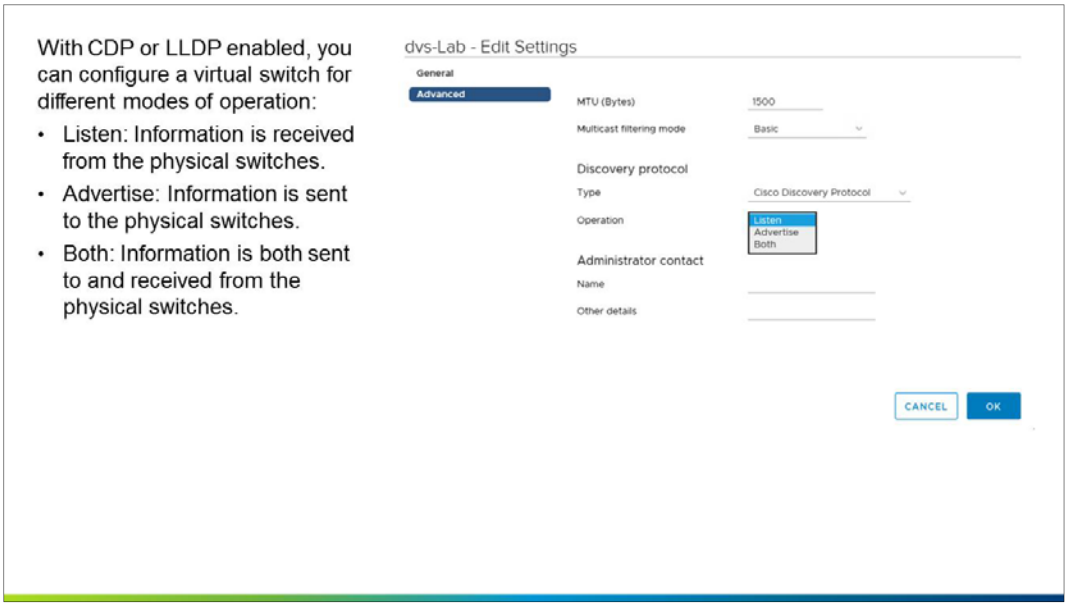
Cisco Discovery Protocol (CDP) was developed by Cisco Systems to broadcast connected devices information at network layer 2. CDP is supported in vSphere since version 4.0.

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

Network devices use CDP or LLDP to advertise their identity, capabilities, and neighbors on a network.

https://t.me/learningnets

2-13 Configuring CDP or LLDP



With CDP and LLDP, the vSphere Client can identify properties of a physical switch, such as switch name, port number, and port speed or 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:

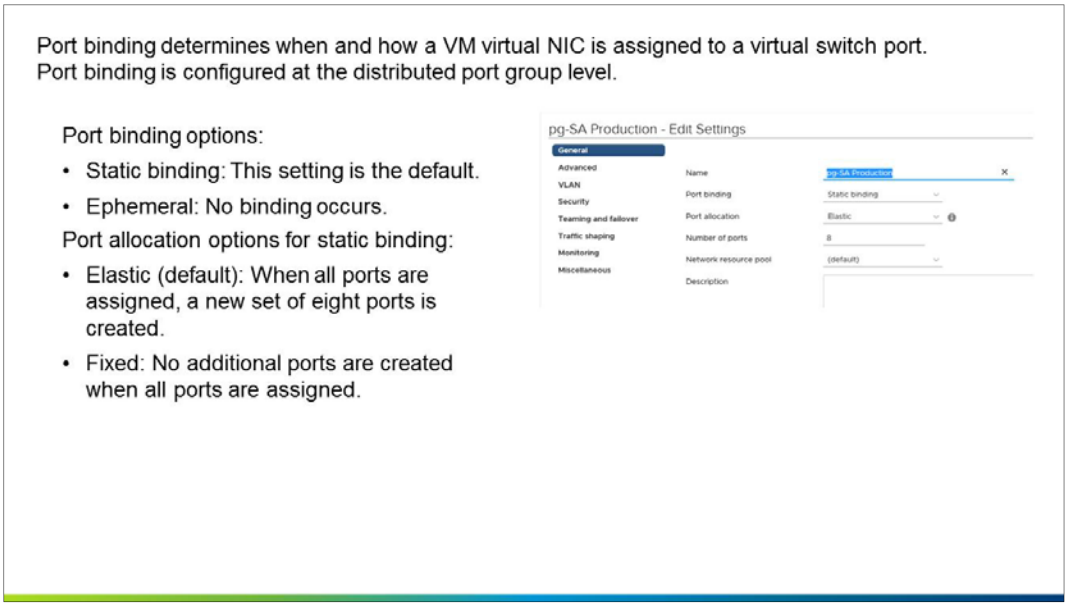
- Listen (default): The ESXi host detects and displays information about the associated physical switch port, but information about the virtual switch is not available to the physical switch administrator.
- Advertise: The ESXi host provides information about the virtual switch to the physical switch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

https://t.me/learningnets

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

https://t.me/learningnets

2-14 About Port Binding



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

If you select static binding, 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 VM when the VM is powered on and its NIC is in a connected state. The port is deleted when the VM is powered off or the VM NIC is disconnected.

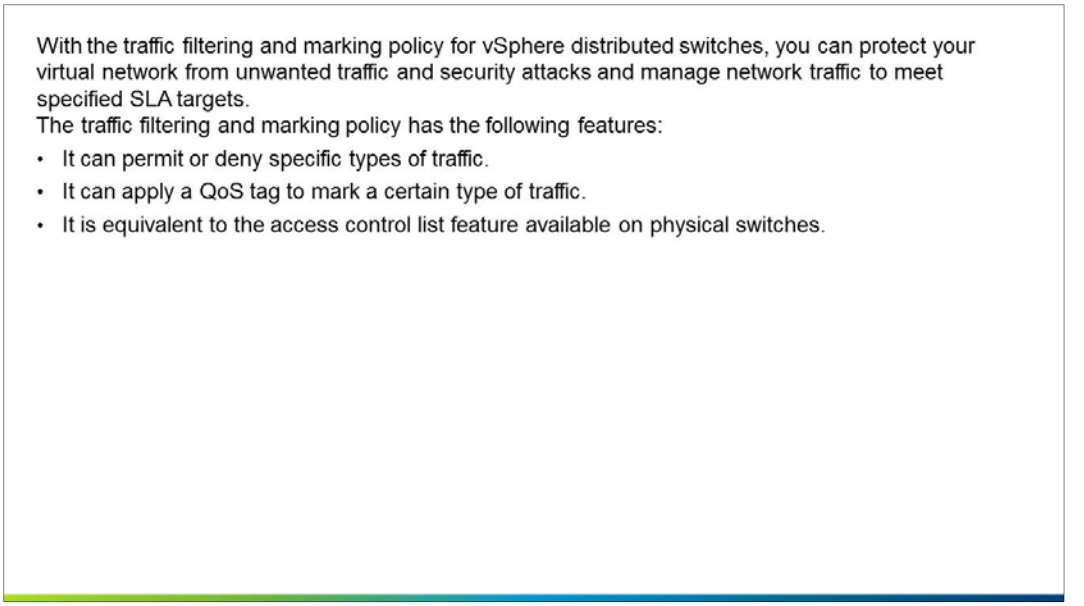
You can make ephemeral port assignments through the ESXi host and vCenter Server, providing flexibility in managing VM connections through the host when vCenter Server is down. Only an ephemeral binding allows you to modify VM network connections when vCenter Server is down. However, network traffic is unaffected by a vCenter Server failure, regardless of the port binding type.

https://t.me/learningnets

Use ephemeral port groups only for recovery purposes when you want to provision ports directly on an ESXi host, bypassing vCenter Server, and not in any other case.

https://t.me/learningnets

2-15 About the Traffic Filtering and Marking Policy



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.

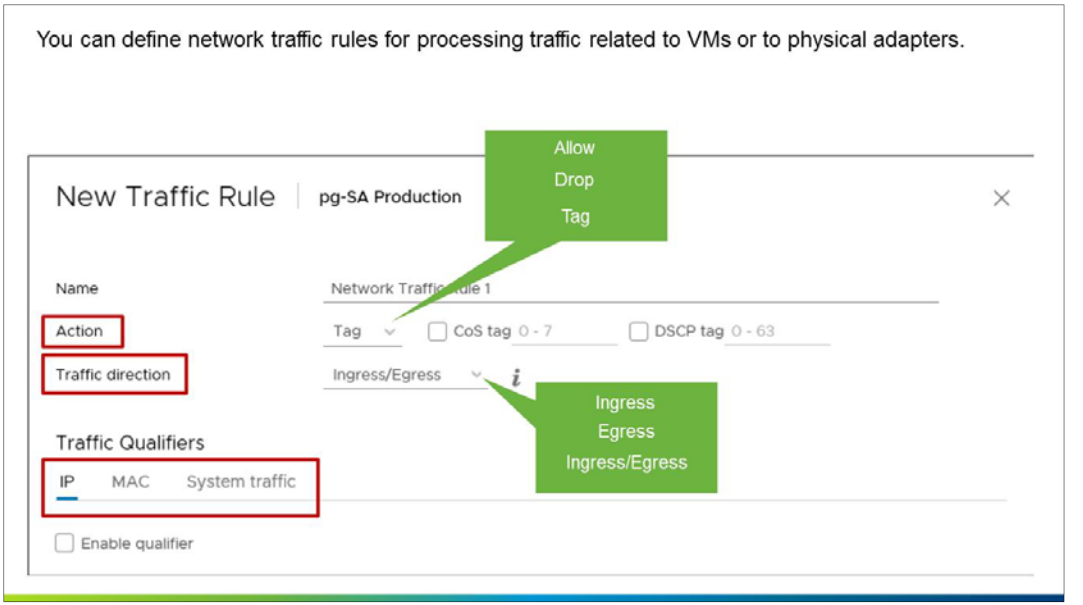
Use the traffic filtering and marking policy to create a set of rules for security and QoS tagging of packets flowing through distributed switch ports.

The distributed switch applies rules on traffic at different places in the data stream. It can apply network traffic rules on the data path between the VM network adapter and the distributed port. It 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



In general, a network traffic rule consists of a qualifier for traffic and 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.

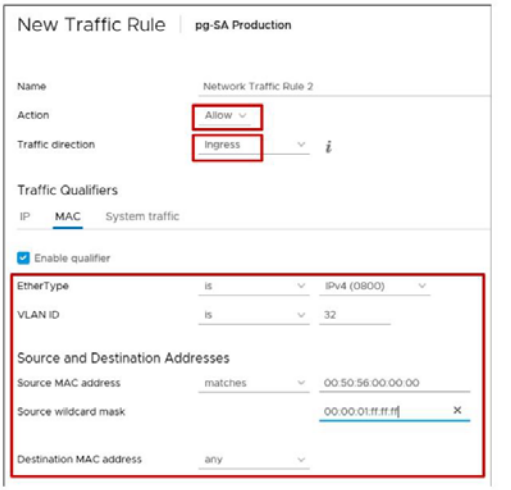
Traffic direction is for the distributed switch. The direction can be ingress (traffic entering the distributed switch), egress (traffic leaving the distributed switch), or both. The direction influences how you identify the traffic source and destination.

A qualifier represents a set of matching criteria related to a networking layer. You can match traffic based on the 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

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

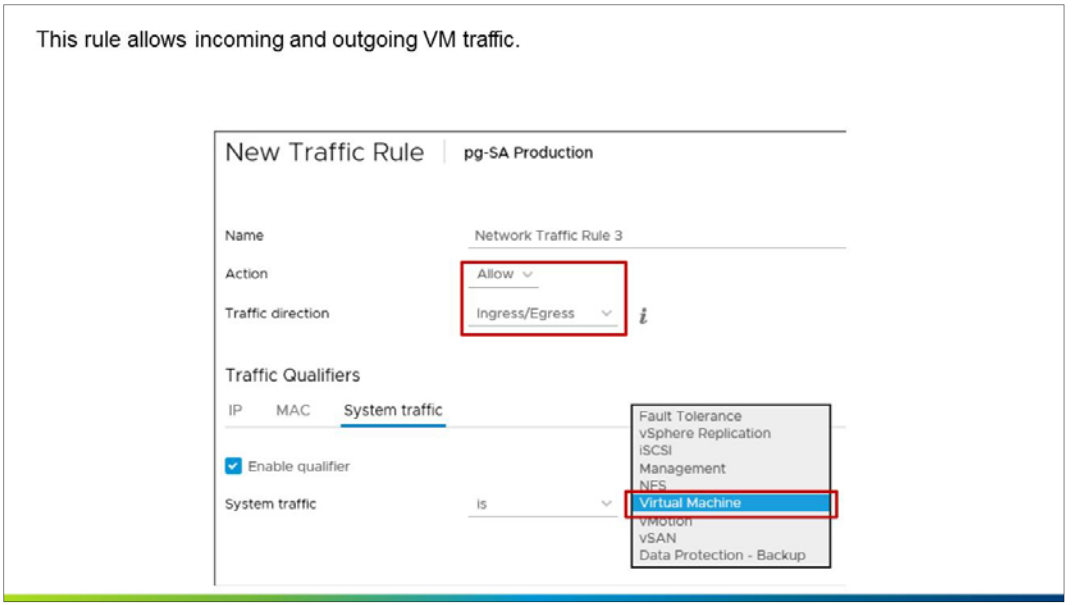


Use a MAC traffic qualifier to define matching criteria for the layer 2 (data link layer) properties of packets. These properties include MAC address, VLAN ID, and a protocol that consumes the frame payload (IPv4, IPv6, or Address Resolution Protocol).

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



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.

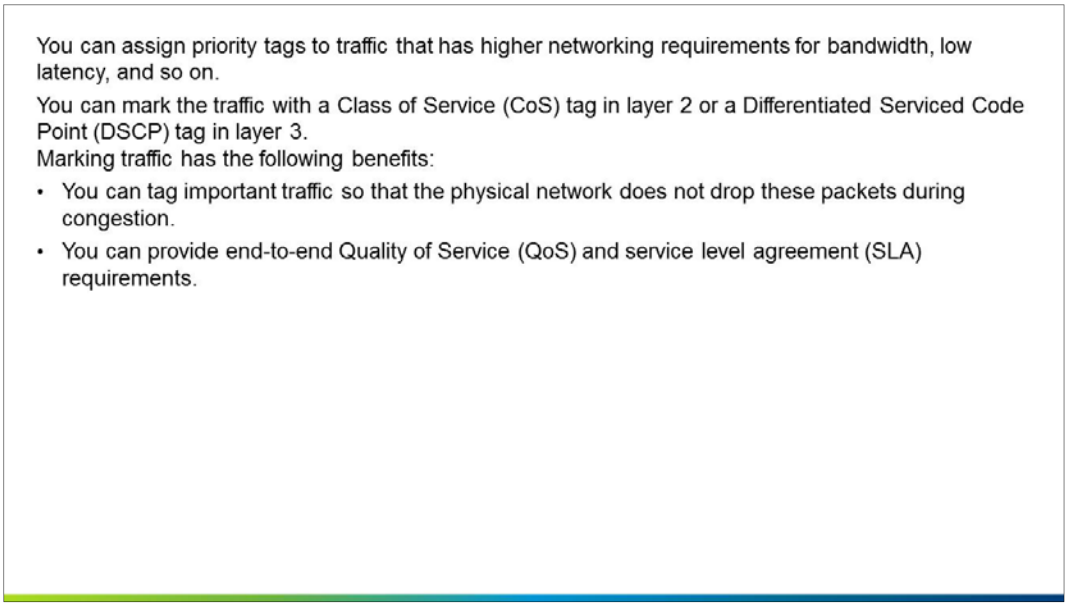
In this example, the rule enforces that the traffic through the pg-SA Production port group must be VM traffic. The distributed switch drops any other types of traffic that are sent to this port group.

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.

When you use 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 (1)



Marking, or priority tagging, is a way to mark traffic that has higher Quality of Service (QoS) demands. With priority tagging, the network can recognize different classes of traffic. Network devices can handle the traffic from each class according to its priority and requirements.

Similar to traffic filtering, marking also requires that you classify 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 Class of Service (CoS) priority tag, a Differentiated Serviced Code Point (DSCP) tag, or both.

You can mark traffic with a CoS priority tag in the layer 2 packet header. Accepted values are 0 through 7.

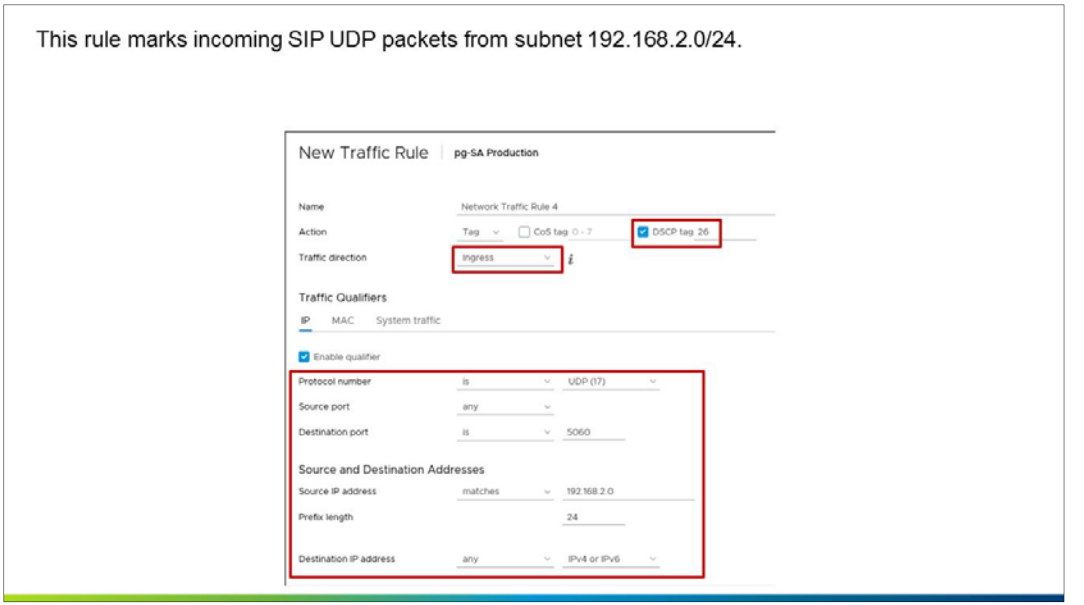
You can mark traffic with a DSCP tag in the layer 3 packet header. Accepted values are 0 through 63.

https://t.me/learningnets

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.

https://t.me/learningnets

2-20 Marking Network Traffic (2)



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 (SIP) for VoIP usually has a DSCP tag equal to 26.

https://t.me/learningnets

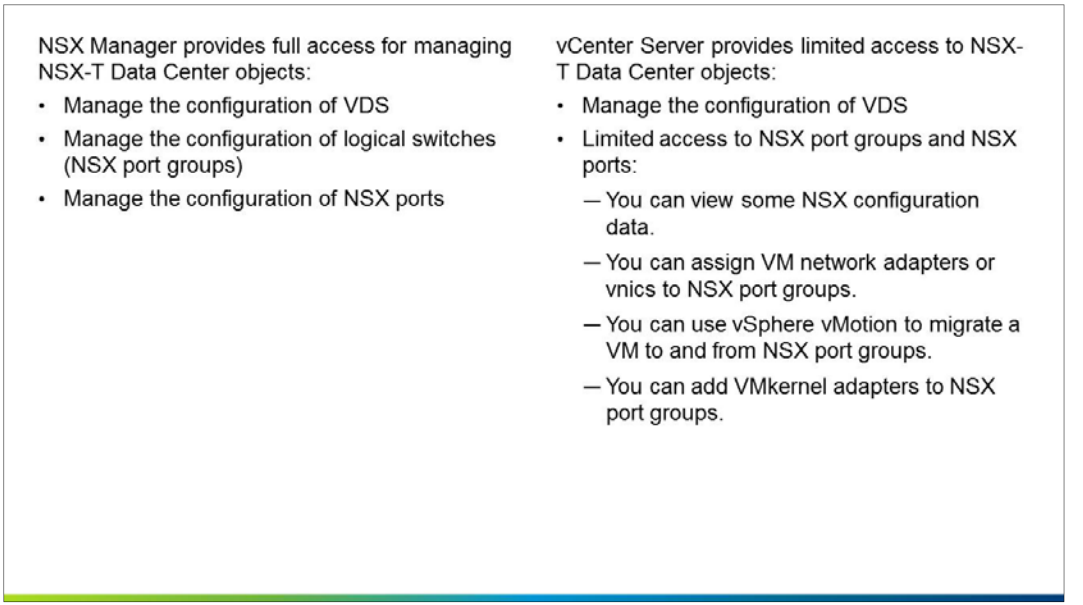
2-21 Access to NSX Virtual Distributed Switch (1)

vSphere 7 supports both VDS and the NSX virtual distributed switch (N-VDS).

- NSX logical switches can be projected in vCenter as a form of virtual port groups.
- vSphere and NSX must be configured with the following minimum versions:
  - vCenter Server 7.0
  - ESXi 7.0
  - VDS 7.0
  - NSX-T Data Center 3.0
- vCenter Server must be registered to the NSX Manager as a compute manager.
- For logical switches to be visible in the vSphere Client, you must configure VDS with the NSX port group from the transport zone.

https://t.me/learningnets

2-22 Access to NSX Virtual Distributed Switch (2)



Starting with vSphere 7, NSX port groups behave as regular distributed port groups so VMkernel adapters can be added to NSX port groups from the vSphere Client.

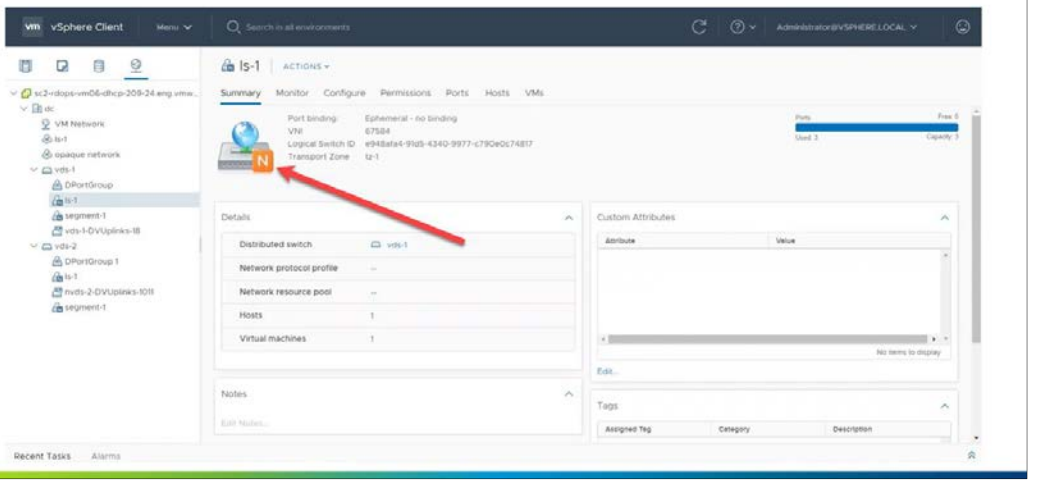
The vSphere Client does not support the following NSX port group actions:

- Delete
- Modify settings
- Use in port mirroring session
- Include configuration in VDS exports

https://t.me/learningnets

2-23 Access to NSX Virtual Distributed Switch (3)

In the vSphere Client, you can visually distinguish NSX virtual distributed switch objects from VDS objects.



https://t.me/learningnets

2-24 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, the VM with the highest I/O is moved to a different uplink.

- True
- False

https://t.me/learningnets

2-25 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, 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 reduces the load if needed.

https://t.me/learningnets

2-26 Labs

Lab: Accessing the Lab Environment
Lab: Configuring vSphere Distributed Switches

https://t.me/learningnets

2-27 Lab 1: Accessing the Lab Environment

Access the lab environment and verify that vSphere licenses are valid:

1. Access Your Lab Environment
2. Verify That the vSphere Licenses Are Valid
3. (Optional) Assign Valid vSphere Licenses
4. Add ESXi Hosts to vCenter Server

https://t.me/learningnets

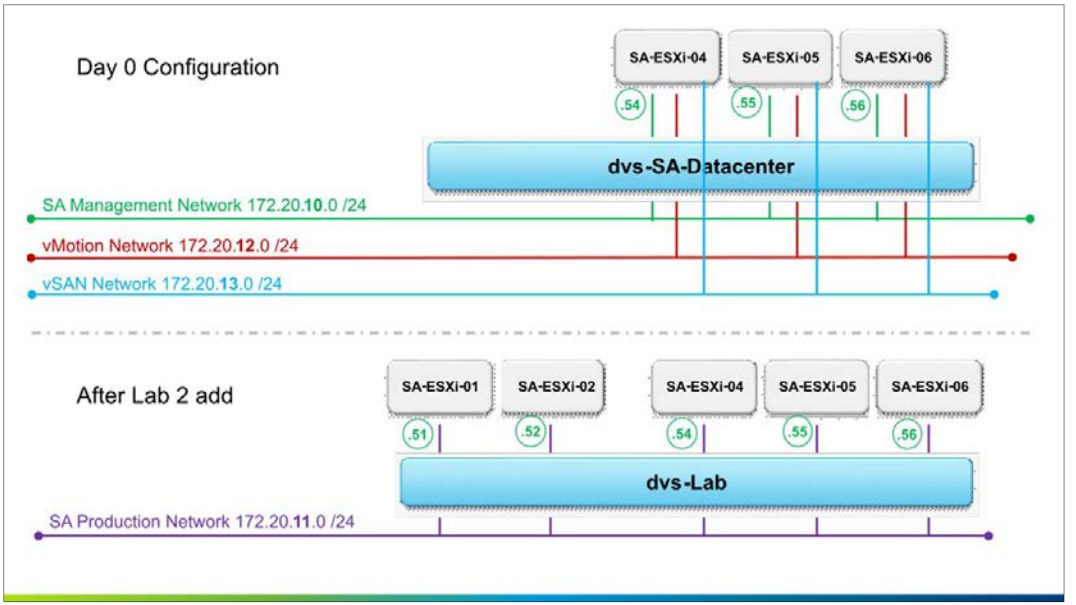
2-28 Lab 2: Configuring vSphere Distributed Switches (1)

Create and configure a distributed switch:

1. Create a Distributed Switch
2. Add ESXi Hosts to the Distributed Switch
3. Examine Your Distributed Switch Configuration
4. Migrate VMs to Another Distributed Switch Port Group

https://t.me/learningnets

2-29 Lab 3: Configuring vSphere Distributed Switches (2)



https://t.me/learningnets

2-30 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe distributed switches
- Recognize how the NSX virtual distributed switch is supported in vSphere 7
- Identify the benefits and features of distributed switches
- Create a distributed switch

https://t.me/learningnets

https://t.me/learningnets

2-32 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Perform a distributed switch health check
- Back up and restore a distributed switch configuration
- Describe the automatic rollback and recovery of networking configurations

https://t.me/learningnets

2-33 About Health Check for Distributed Switches

With the health check support feature, you can identify and troubleshoot configuration errors in a vSphere distributed switch.
By regularly examining settings on distributed and physical switches, health check identifies 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

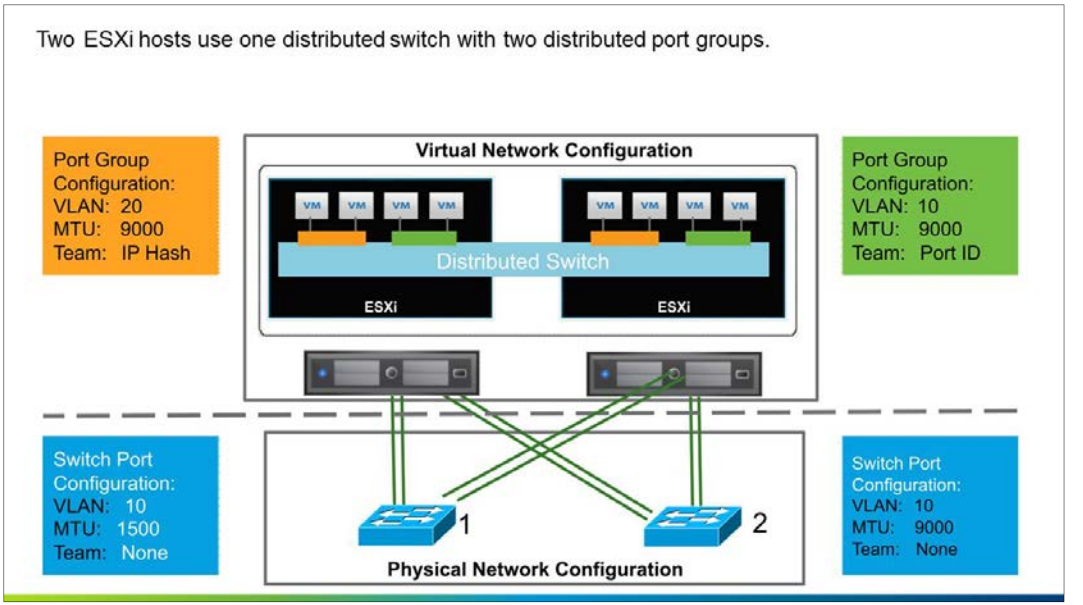
Health check is a feature that detects certain inconsistencies between the physical and the virtual networks. Key parameters, such as VLAN tags, MTUs, and the NIC teaming configuration, must be configured consistently on the physical and the virtual switches. An inconsistent configuration can lead to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

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

As a best practice, enable health check for brief time periods because this feature creates many MAC addresses and can affect production traffic.

https://t.me/learningnets

2-34 Health Check Example



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

Compare the virtual port group (VLAN 10) 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 not experience connectivity issues.

Compare the virtual port group (VLAN 20) to the settings on physical switch 1. The VLAN IDs and the MTU settings are different. The virtual port group 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.

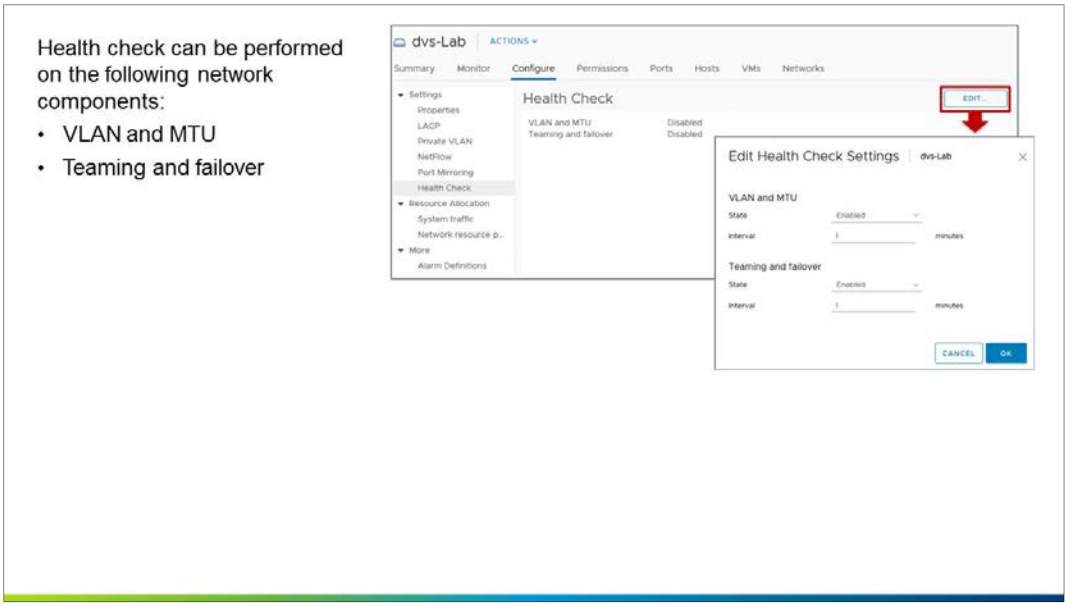
Health check can detect and report the configuration differences between the port group and the switch port configuration by using layer 2 Ethernet packets. At one-minute intervals (by default),

https://t.me/learningnets

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 the vSphere Client.

https://t.me/learningnets

2-35 Enabling Health Check



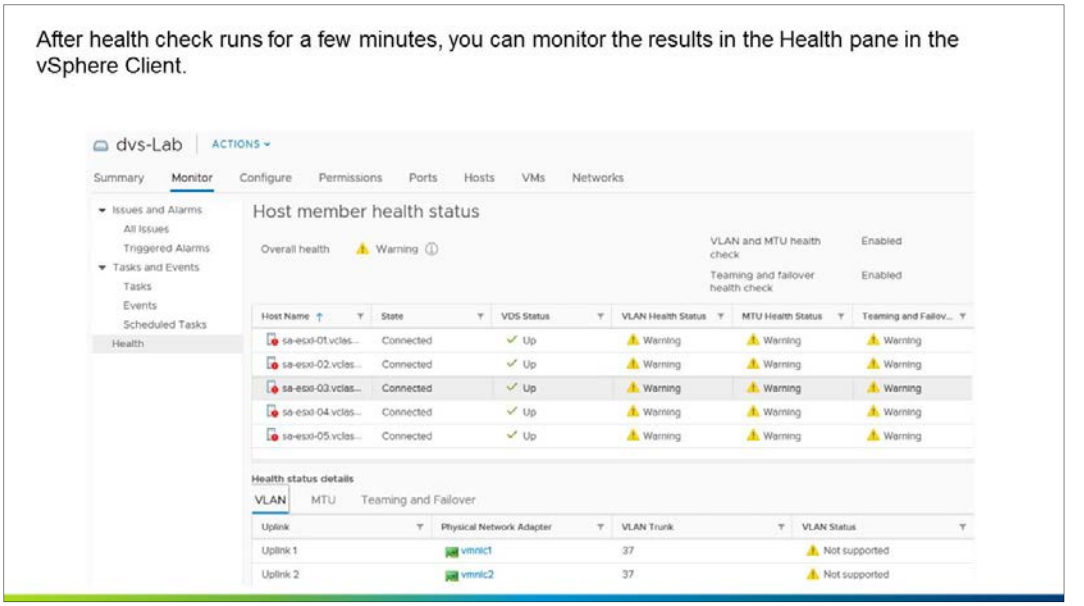
Health check has the following requirements:

- 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/7.0/vsphere-esxi-vcenter-server-70-networking-guide.pdf.

https://t.me/learningnets

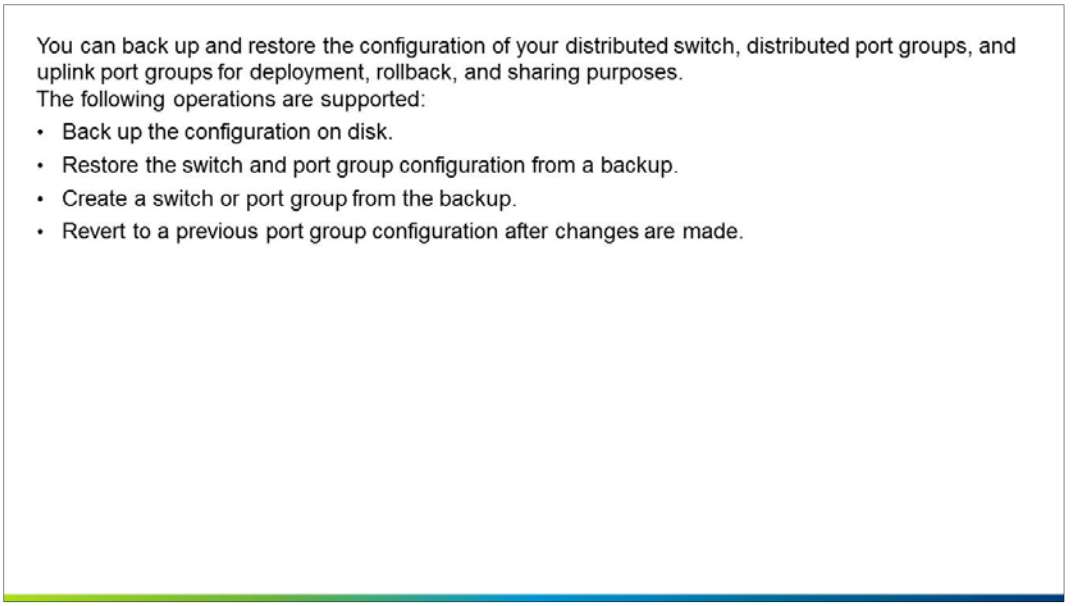
2-36 Monitoring Health Check Results



The health status details are organized into three categories: VLAN, MTU, and Teaming and Failover.

https://t.me/learningnets

2-37 Backing Up and Restoring Configurations



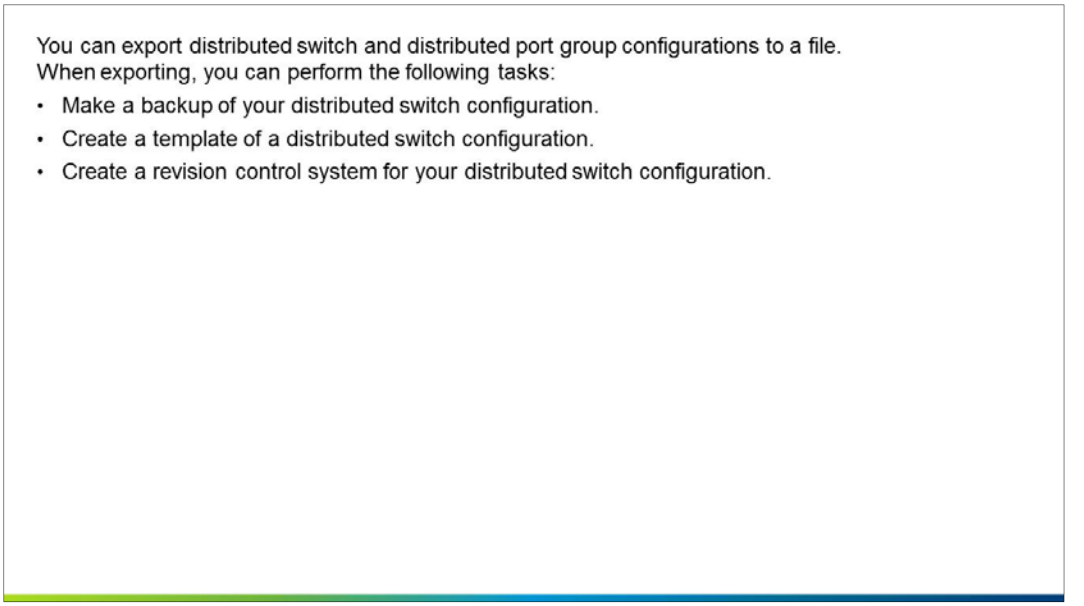
You perform these operations by using the export, import, and restore functions available for distributed switches.

You have other available restore options if a switch configuration is lost, for example, when vCenter Server database corruption occurs, or if the virtual switch or port group settings are misconfigured.

Ways to restore the virtual switch include restoring the database completely or rebuilding the switch. Although both measures restore the switch, they can be time-consuming.

https://t.me/learningnets

2-38 Exporting Distributed Switch Configurations



You export the distributed switch and distributed port group configuration to a file on the system that is running the vSphere Client. The file preserves valid network configurations, which can then be distributed to other deployments.

By making periodic backups with the export function, you can preserve your distributed switch configuration if a failure occurs, 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.

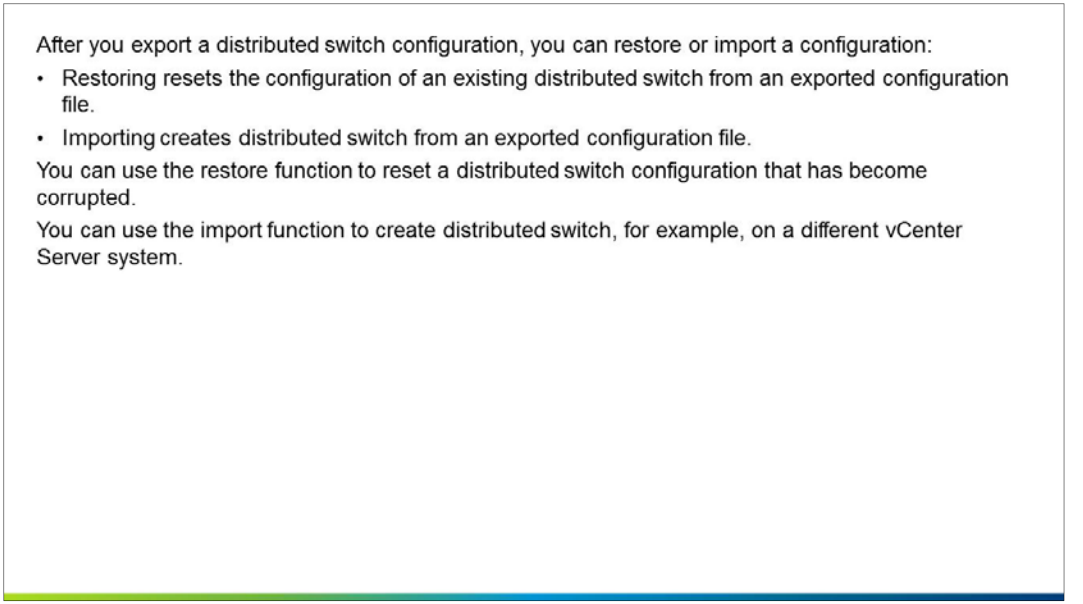
You can automate this task with the following PowerCLI cmdlets:

https://t.me/learningnets

- `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

2-39 Restoring and Importing Distributed Switch Configurations



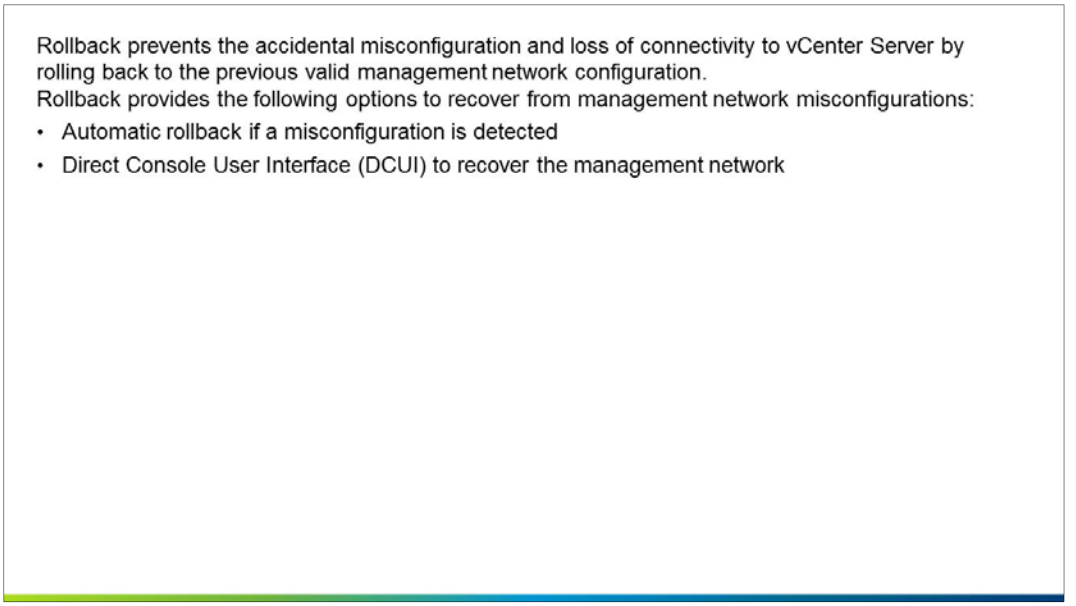
With the restore and import functions, you can recreate a distributed switch configuration.

Restoring a distributed switch configuration overwrites the current settings of the distributed switch and its port groups. Full functionality is restored if network settings fail or the vCenter Server database is corrupted. 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

2-40 Rollback and Recovery of the Management Network

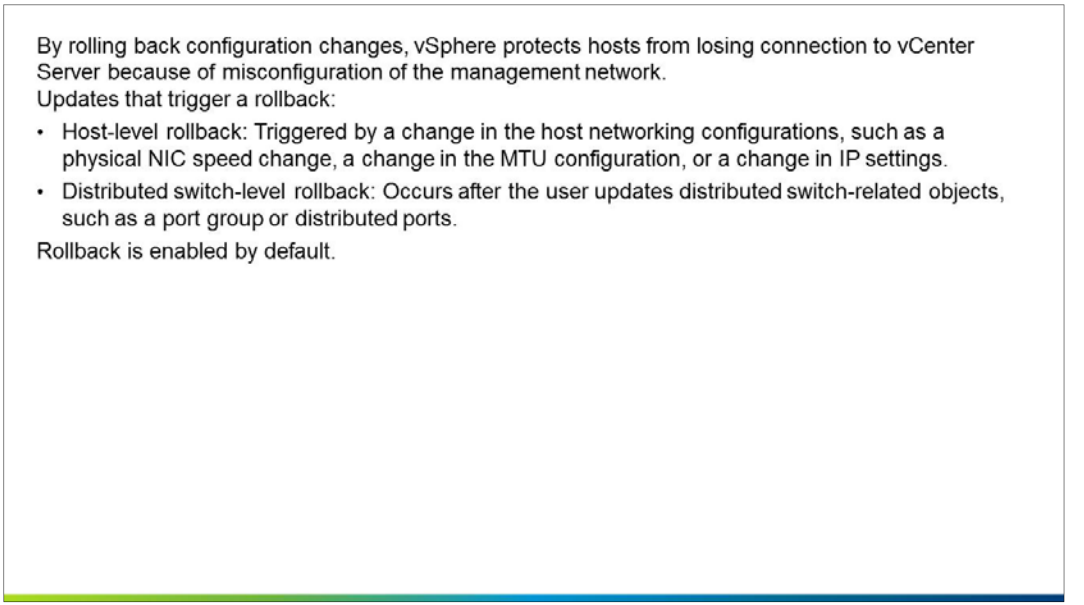


Automatic rollback and recovery is a vSphere feature that helps prevent management network outages. You protect the management network in the following ways:

- The automatic rollback feature detects any configuration changes on the management network. If the host cannot reach vCenter Server, the changes do not 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 those parameters for all hosts connected to the switch.

https://t.me/learningnets

2-41 Rollback Details



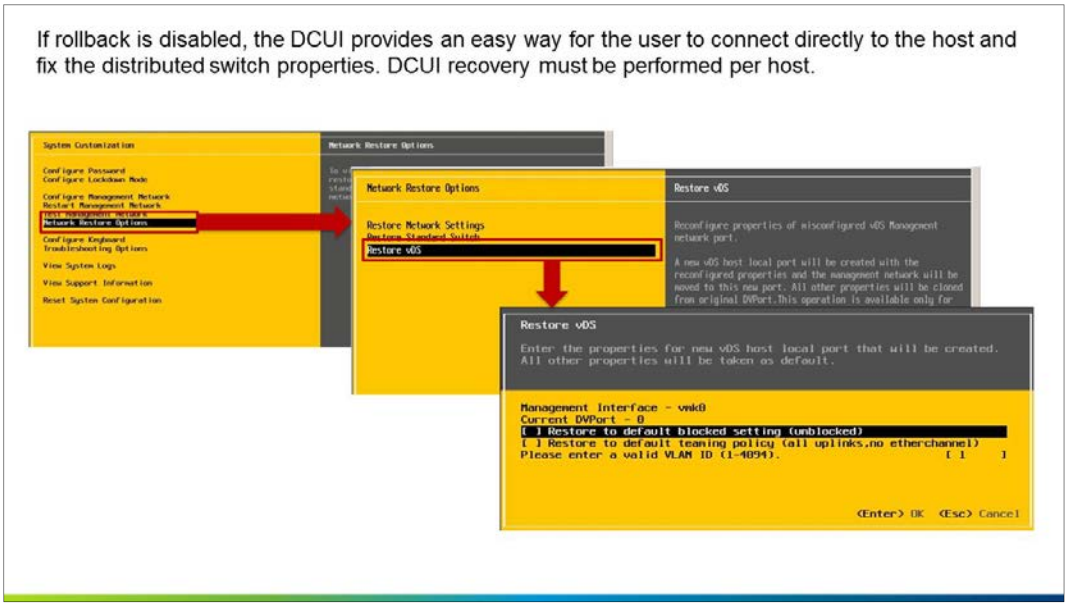
The following types of updates can trigger a rollback:

- 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.
- If misconfiguration of the distributed virtual switch causes connectivity issues between the hosts and the vCenter Server system, a rollback can occur. 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

2-42 Recovery Through the DCUI



To roll back the configuration manually from the DCUI, take these steps:

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.
4. Under Network Restore Options, select **Restore vDS** to roll back the distributed network settings.

   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

2-43 Lab 4: Managing vSphere Distributed Switches

Perform a health check, remediate the vSphere Distributed Switch (VDS) issue and back up a distributed switch:

1. Add a New Port Group to VDS
2. Enable the VDS Health Check
3. Investigate the VDS Health Check Status
4. Remediate the VDS Issue
5. Disable the VDS Health Check Service
6. Back Up the VDS Configuration

https://t.me/learningnets

2-44 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Perform a distributed switch health check
- Back up and restore a distributed switch configuration
- Describe the automatic rollback and recovery of networking configurations

https://t.me/learningnets

Lesson 3: Using Network I/O Control in Distributed Switches

https://t.me/learningnets

2-46 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

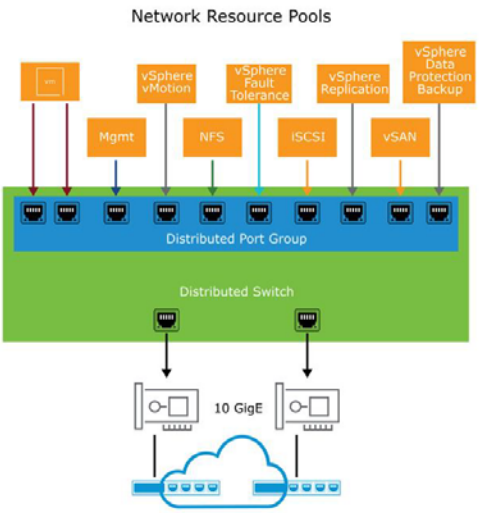
- Describe how Network I/O Control enhances performance
- Describe how to use Network I/O Control to allocate bandwidth to different types of system traffic
- Describe how to use Network I/O Control to allocate bandwidth to VMs

https://t.me/learningnets

2-47 About Network I/O Control (1)

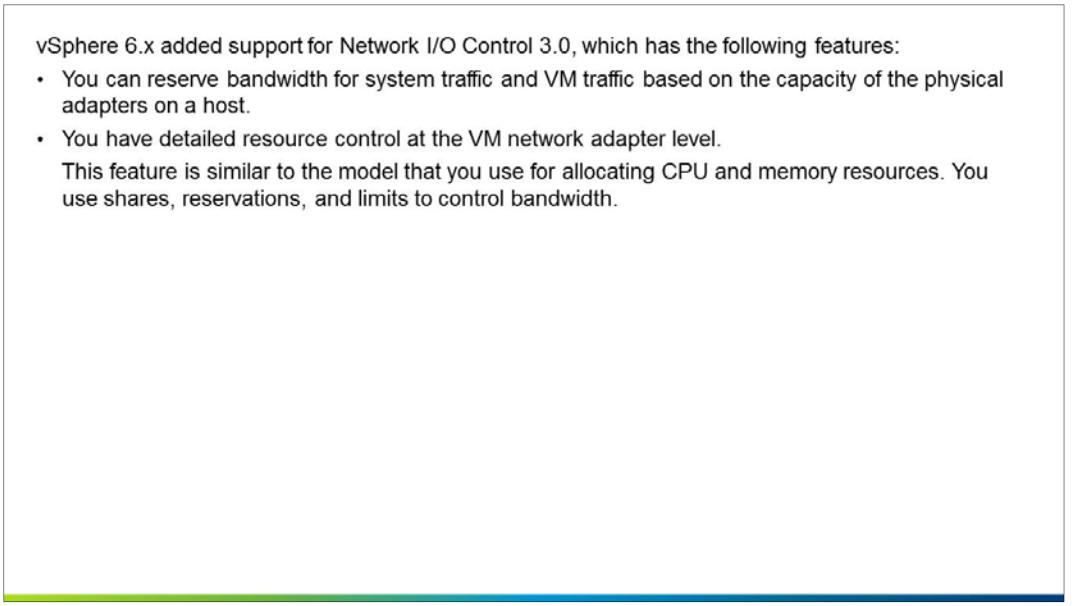
You can use Network I/O Control 3.0 to allocate network bandwidth to business-critical applications and to resolve situations where several types of traffic compete for common resources.

Network I/O Control 3.0 allocates network bandwidth over distributed switches by using network resource pools for VM and system traffic.



https://t.me/learningnets

2-48 About Network I/O Control (2)



For more information about the procedure for upgrading Network I/O Control, see vSphere Networking at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-networking-guide.pdf.

https://t.me/learningnets

2-49 About the Bandwidth Allocation Model for System Traffic

Network I/O Control 3.0 uses several configuration parameters to allocate bandwidth to system traffic (such as management, iSCSI, vSAN, and VM).

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 to 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.

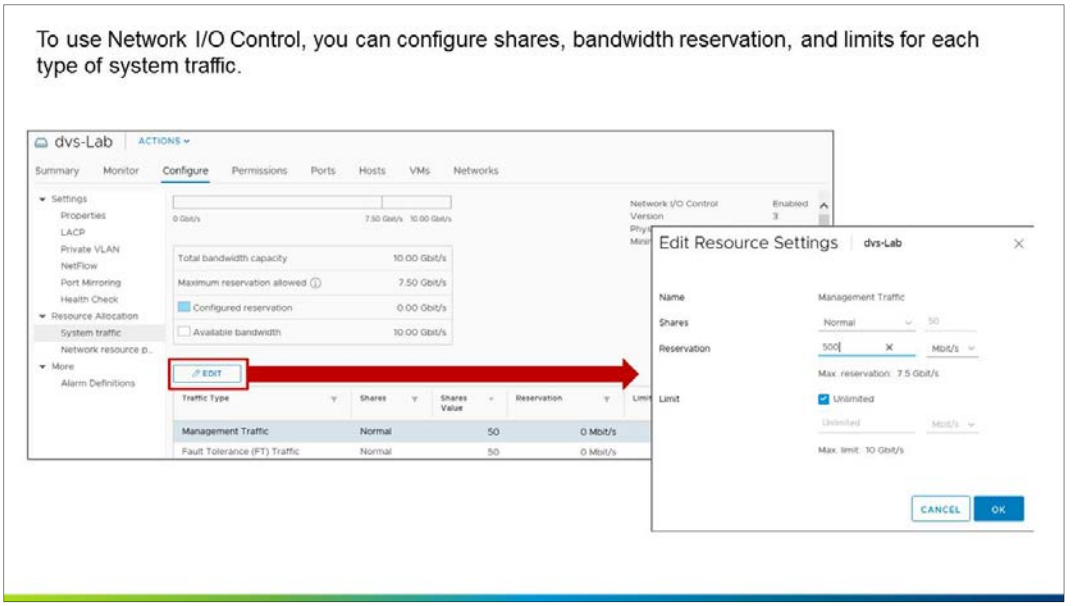
The amount of bandwidth that is available to a system traffic type is determined by its relative shares and the amount of data that the other system features transmit.

Reserved bandwidth that is unused becomes available to other types of system traffic. However, Network I/O Control 3.0 does not redistribute unused system traffic capacity to VM 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. The vacant bandwidth is not allocated to VM system traffic.

Network I/O Control 3.0 cannot safely meet a potential need for bandwidth for system traffic. This failure to reallocate unused reserved bandwidth to VM system traffic is especially true of a new iSCSI path where you must provide bandwidth to a new VMkernel adapter.

https://t.me/learningnets



You can use Network I/O Control 3.0 to configure bandwidth allocation for the traffic that is related to the main features of vSphere:

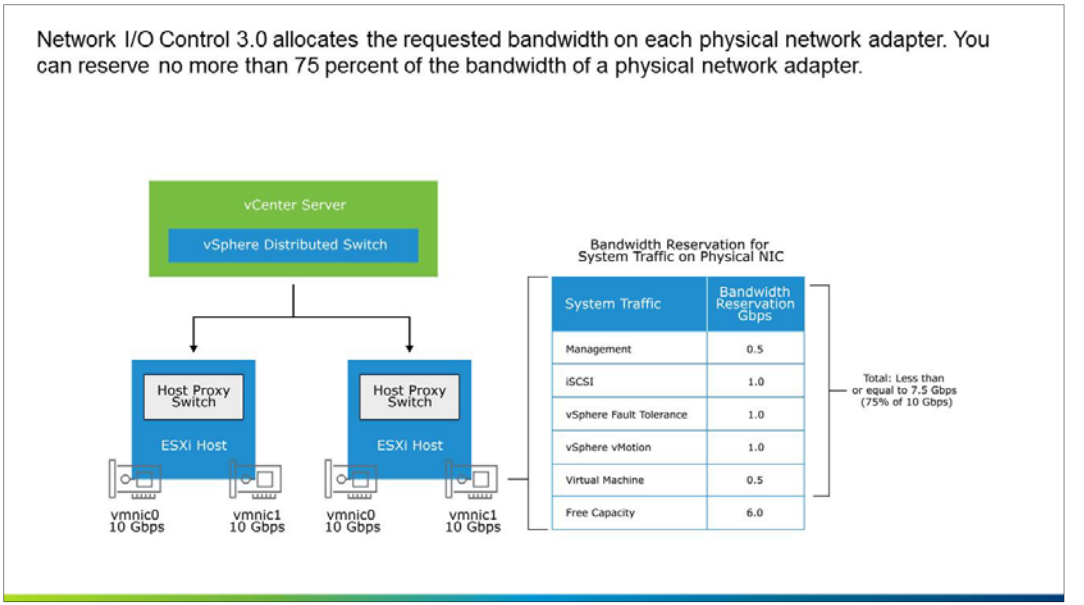
- VMs
- Management
- vSphere vMotion
- NFS
- vSphere Fault Tolerance
- iSCSI
- vSphere Replication

https://t.me/learningnets

- vSAN
- vSphere Data Protection

https://t.me/learningnets

2-51 Reserving Bandwidth for System Traffic (1)



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

For 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 VM traffic, and so on.

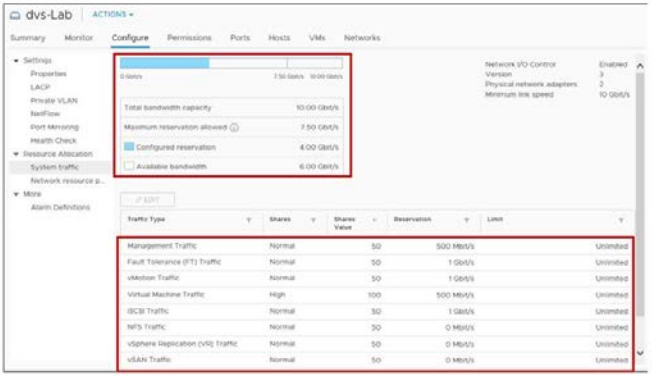
You might leave more capacity unreserved so that the host can dynamically allocate bandwidth according to shares, limits, and use, and to reserve only enough bandwidth for the operation of a system feature.

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, 30 Gbps (.75 x 40) is reserved.

https://t.me/learningnets

2-52 Reserving Bandwidth for System Traffic (2)

The vSphere Client shows bandwidth use information and the share, reservation, and limit values for each traffic type.

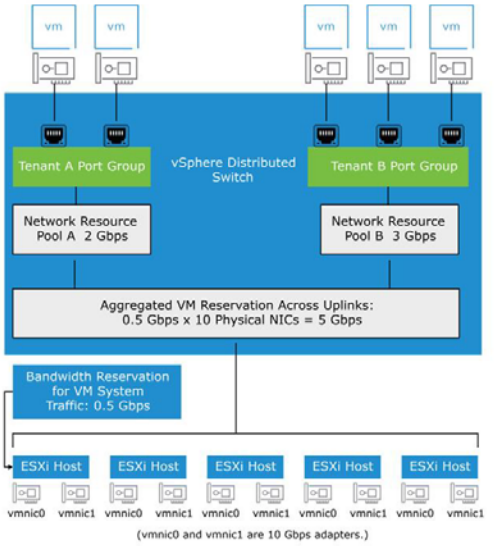


https://t.me/learningnets

2-53 Bandwidth Aggregation for Network Resource Pools

A network resource pool represents a part of the bandwidth that is 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.

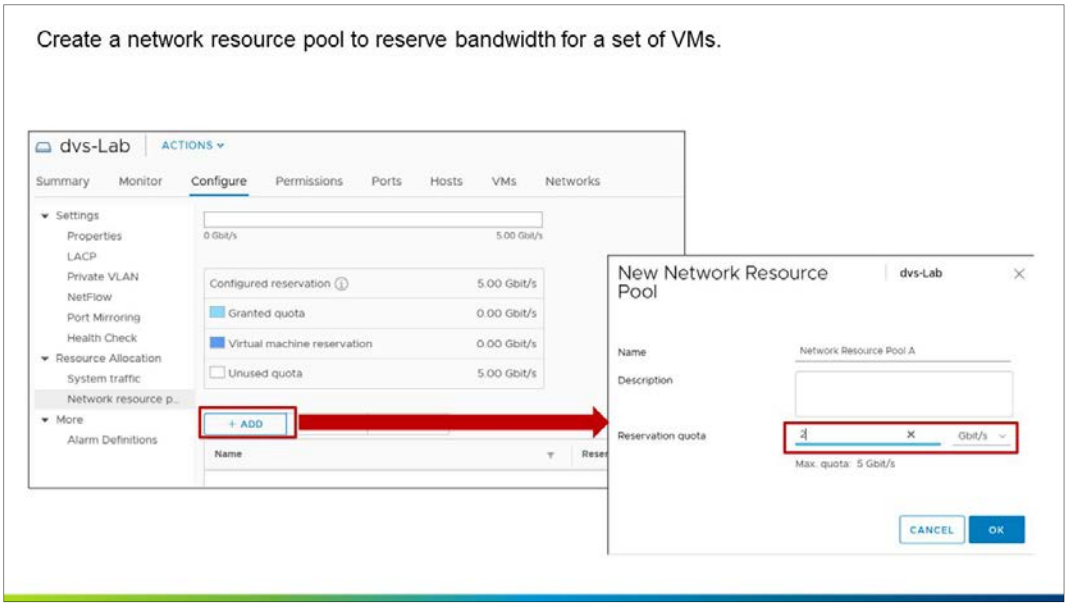


Network I/O Control 3.0 allocates bandwidth for VMs across the entire distributed switch and on the physical adapter carrying the VM traffic.

To use Network I/O Control 3.0 to enable bandwidth allocation for VMs, configure the VM system traffic. The bandwidth reservation for VM traffic is also used in admission control. When you power on a VM, admission control verifies that enough bandwidth is available.

https://t.me/learningnets

2-54 Creating Network Resource Pools



For example, if the VM system traffic has 0.5 Gbps reserved for each 10 GbE uplink on a distributed switch that has 10 uplinks, the total aggregated bandwidth that is available for VM 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 VM receives bandwidth from the pool through the distributed port group that the VM is connected to.

https://t.me/learningnets

2-55 Bandwidth Allocation for VM Traffic

Bandwidth is allocated to individual VMs based on configured shares, reservations, and limits for the VM network adapters.

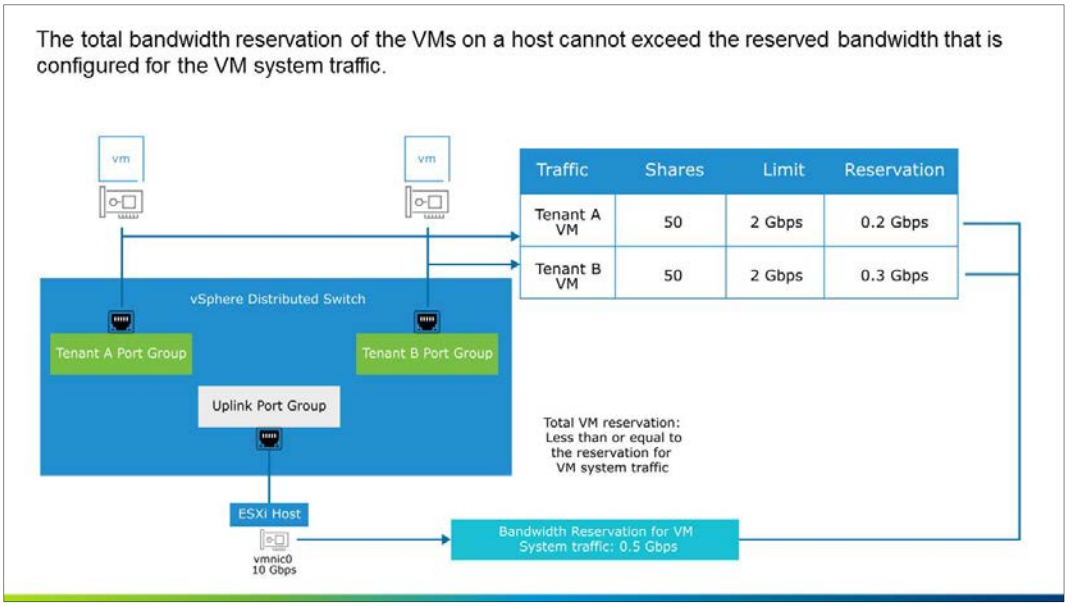
Bandwidth Parameter	Description
Shares	The relative priority of the traffic type through the VM network adapter against the capacity of the physical adapter that is carrying the VM traffic to the network. Use the following values to define the number of shares: • Low: 25 • Normal: 50 • High: 100 • Custom: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

A network resource pool provides a reservation quota to VMs. The quota represents a portion of the bandwidth that is reserved for VM system traffic on the physical adapters connected to the distributed switch.

You can set aside bandwidth from the quota for the VMs 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

2-56 Bandwidth Allocation for Individual VMs

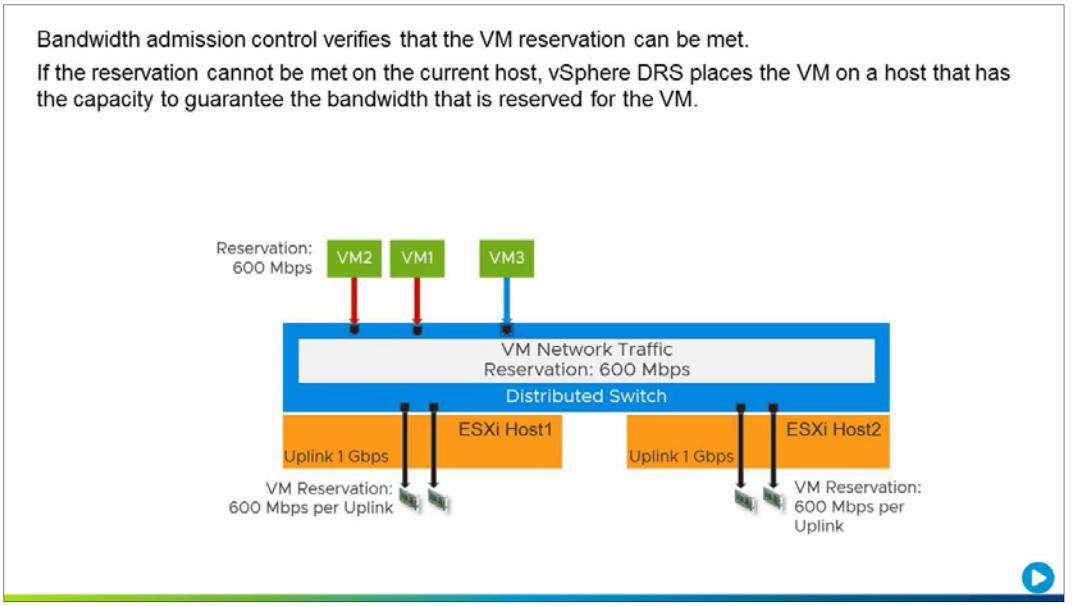


To guarantee bandwidth, Network I/O Control 3.0 implements a traffic placement engine that becomes active if a VM has bandwidth reservation configured. The distributed switch tries to direct the traffic from a VM network adapter to a physical adapter that can supply the required bandwidth and is in the 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 VM network adapter asks for a limit of 200 Mbps and the average bandwidth configured in the traffic-shaping policy is 100 Mbps, the effective limit becomes 100 Mbps.

https://t.me/learningnets

2-57 Bandwidth Admission Control in vSphere DRS (1)



If you power on a VM that is in a cluster, vSphere DRS places that VM on a host with the capacity to guarantee the bandwidth that is reserved for the VM, according to the active teaming policy.

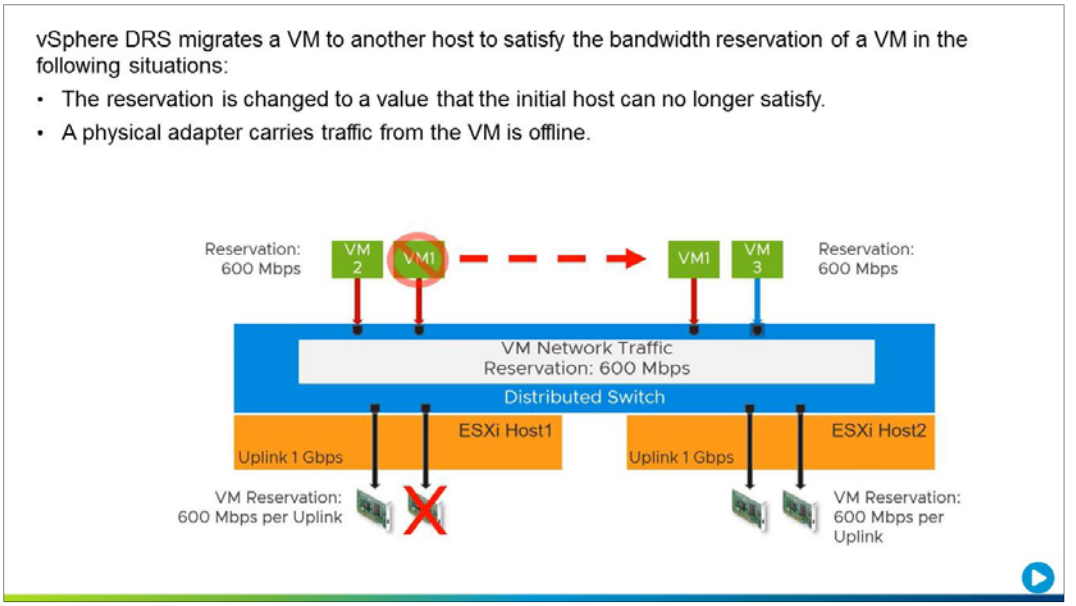
To use admission control in vSphere DRS, you must perform the following tasks:

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

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

https://t.me/learningnets

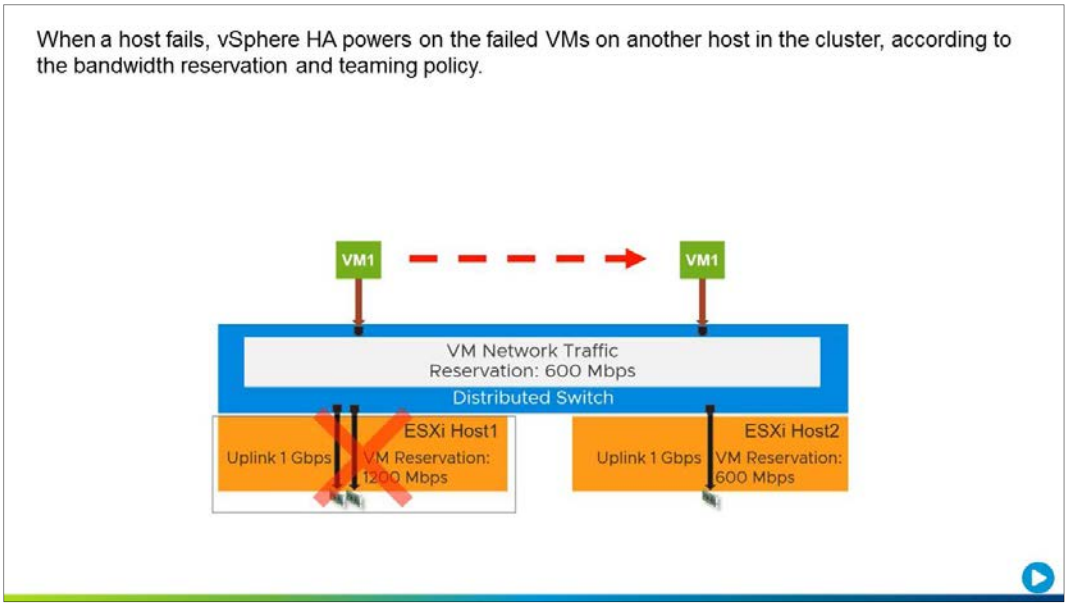
2-58 Bandwidth Admission Control in vSphere DRS (2)



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

https://t.me/learningnets

2-59 Bandwidth Admission Control in vSphere HA



Bandwidth admission control prevents a VM from starting if the bandwidth reservation for that VM cannot be met.

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

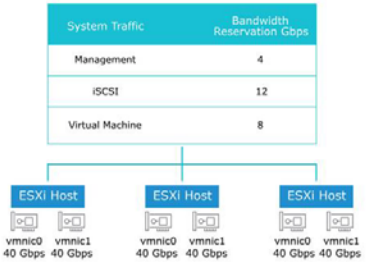
- Allocate bandwidth for VM system traffic.
- Configure the bandwidth requirements of a VM that is connected to the distributed switch.

https://t.me/learningnets

2-60 Activity: Network I/O Control (1)

You want to guarantee that the VMs connected to a distributed port group called Prod 1 can access 25 percent of the available reserved bandwidth for VM traffic.
Use the bandwidth reservation information in the example to answer the following question.

What must you configure to provide the appropriate bandwidth to the VMs on Prod 1?



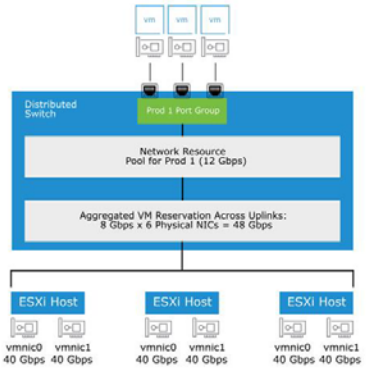
https://t.me/learningnets

2-61 Activity: Network I/O Control (2)

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

Twenty-five percent of the aggregated amount is 12 Gbps.

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



https://t.me/learningnets

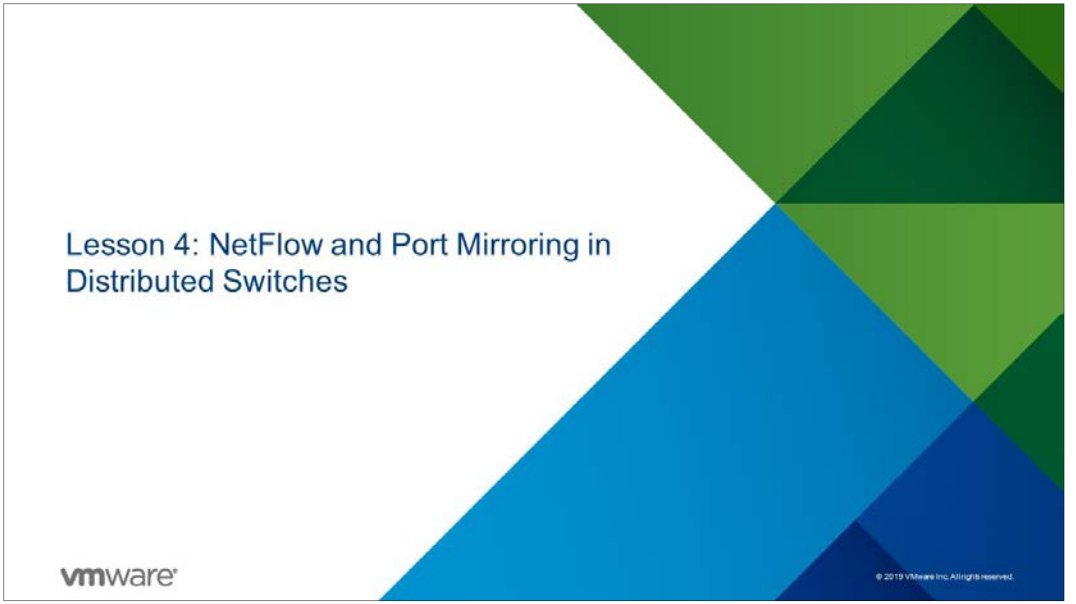
2-62 Review of Learner Objectives

After completing this lesson, 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 VMs

https://t.me/learningnets

2-63 Lesson 4: NetFlow and Port Mirroring in Distributed Switches



https://t.me/learningnets

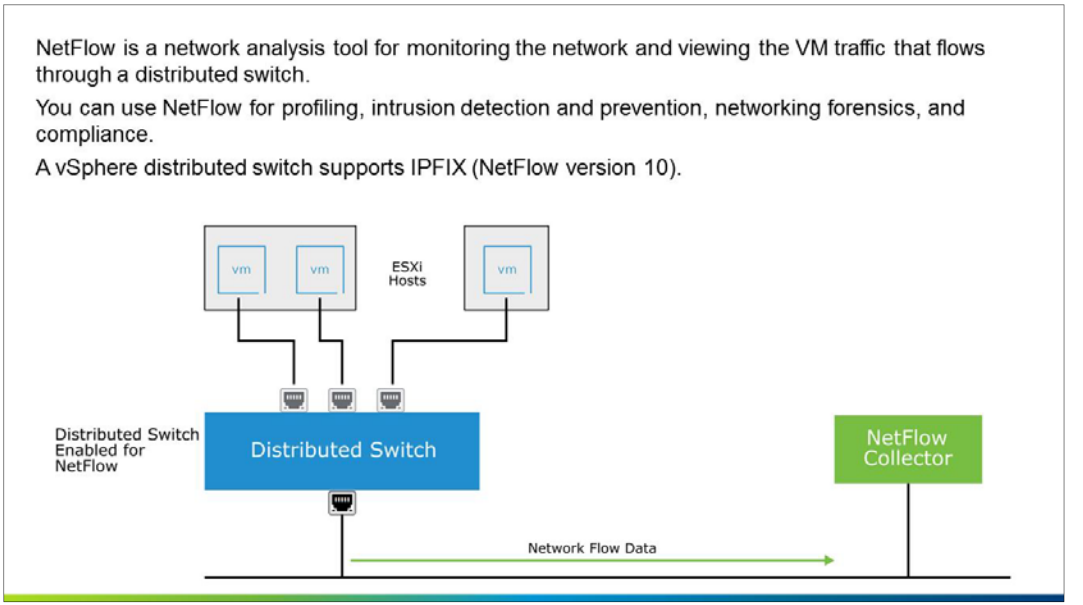
2-64 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe how to use NetFlow for monitoring network security and performance
- Configure port mirroring on a distributed switch

https://t.me/learningnets

2-65 About NetFlow



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

Standard switches do not support NetFlow.

NetFlow collectors are available from third-party providers.

https://t.me/learningnets

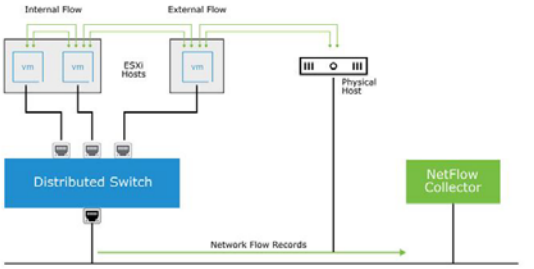
2-66 About Network Flows

A network flow is a unidirectional sequence of packets. Each packet shares a common set of properties.

NetFlow captures the following types of flows:

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

Flow records are sent to a NetFlow collector for analysis.



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

Internal flows are generated from intrahost VM traffic, that is, traffic between VMs on the same hosts.

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

A flow is a sequence of packets that share properties such as source and destination IP addresses, source and destination ports, input and output interface IDs, and protocol.

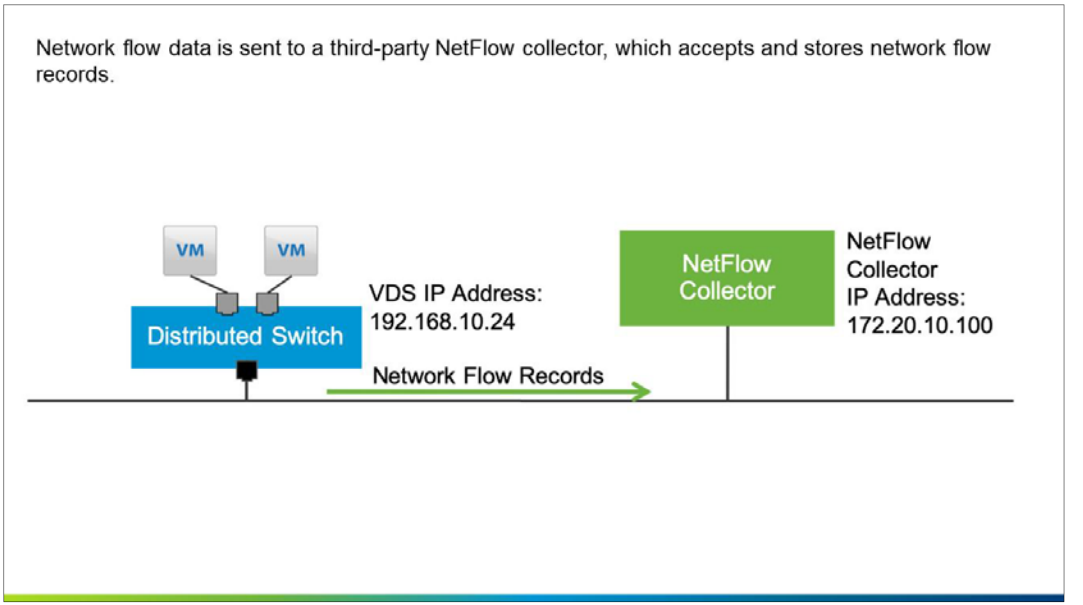
A flow is unidirectional. Flows are processed and stored as flow records by supported network devices, such as a distributed switch. The flow records are sent to a NetFlow collector for additional analysis.

https://t.me/learningnets

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

2-67 Network Flow Analysis (1)



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 for 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 VM sent and received in the past 24 hours

https://t.me/learningnets

2-68 Network Flow Analysis (2)

A NetFlow collector stores flow-based data and analyzes the collected data:

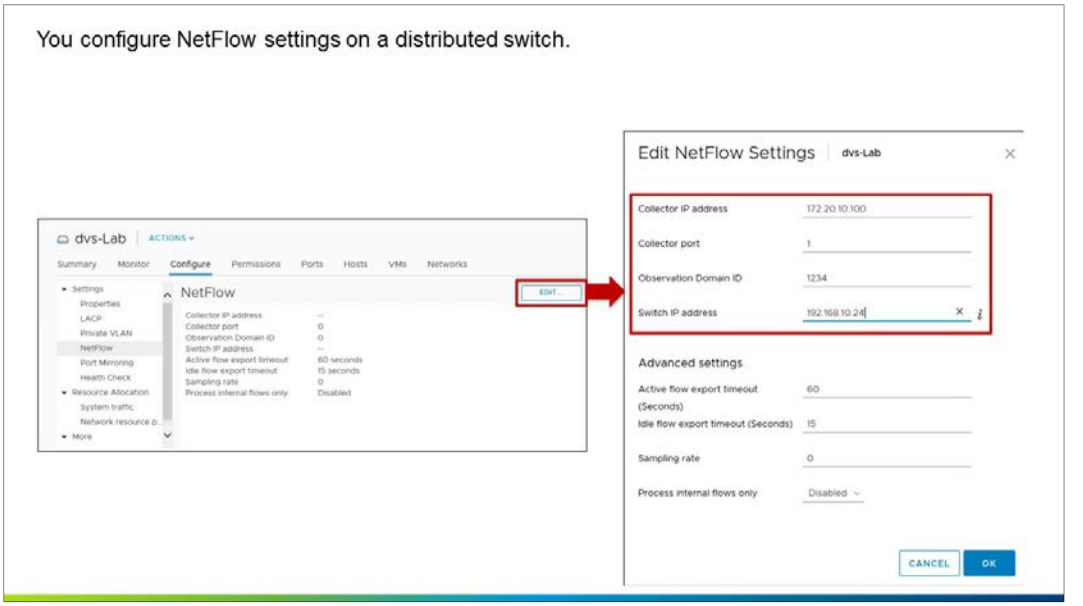
- Includes a storage system for long-term storage of flow-based data:
  - You can investigate and isolate excessive network bandwidth use, bottlenecks, and unexpected application traffic.
  - You can view historical records to diagnose the cause outages or breaches.
- Mines, aggregates, and reports on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

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 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 required. 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

2-69 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 that are 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. With the address, the NetFlow collector can interact with the distributed switch as a single switch, rather than as a separate, unrelated switch for each associated host. If this IP address is not configured, the host's management IP address is used instead.

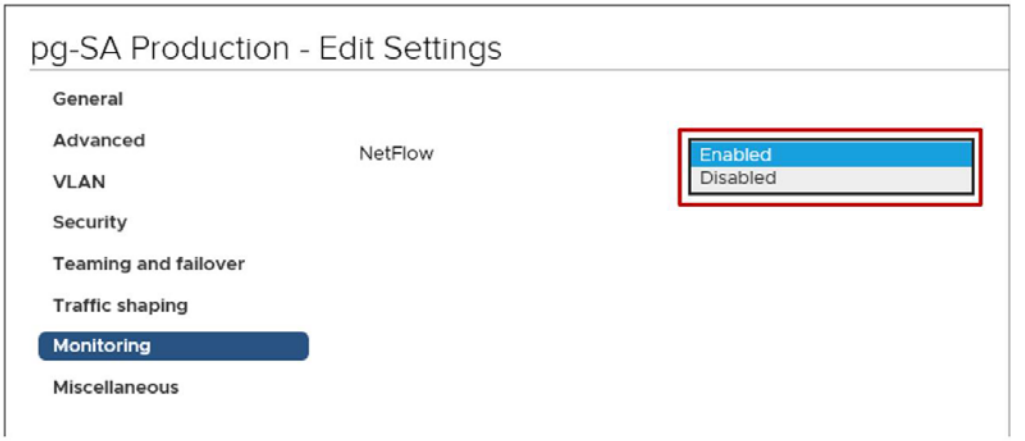
https://t.me/learningnets

For a description of the advanced NetFlow settings, see vSphere Networking at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-networking-guide.pdf.

https://t.me/learningnets

2-70 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

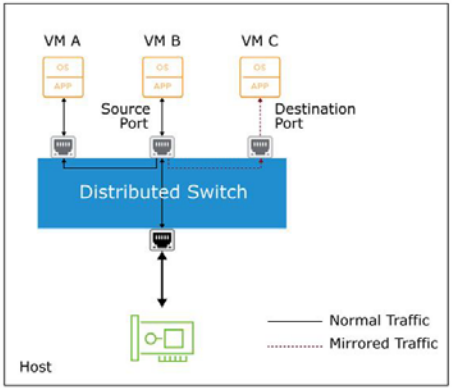
2-71 About Port Mirroring

Port mirroring duplicates network packets from a source to a destination. It has the following uses:

- Helps with troubleshooting
- Provides input for network analysis appliances

Many network switch vendors implement port mirroring in their products.
vSphere port mirroring supports the following switch functionalities:

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



On physical switches, administrators often must mirror traffic to special ports when troubleshooting network-related problems. Port mirroring is commonly used for network appliances that require the monitoring of network traffic, for example, intrusion detection systems.

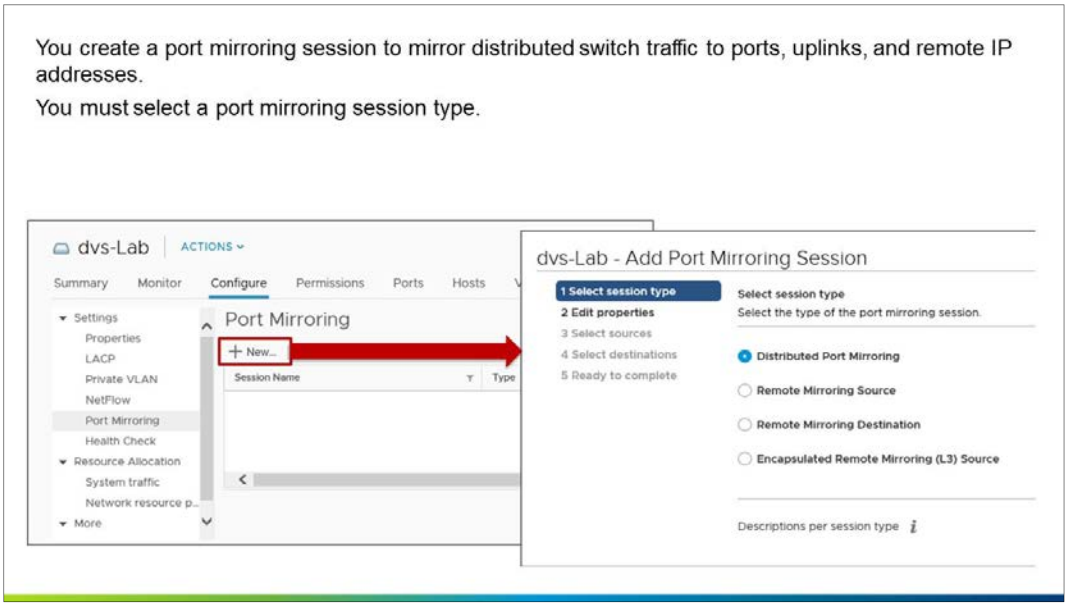
Many network switch vendors implement port mirroring in their products. For example, port mirroring on a Cisco Systems switch is called Switched Port Analyzer (SPAN).

With port mirroring, you do not need to enable promiscuous mode on a distributed switch 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

2-72 Port Mirroring Session Types



You can select from the following port mirroring session options:

- **Distributed Port Mirroring**: Port mirroring between distributed ports can only be local. If the source and the destination ports are on different hosts because of 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. A new mirroring path from the source port to B's destination uplink is created on B. The uplink that is used is determined by the uplink name specified in the session.

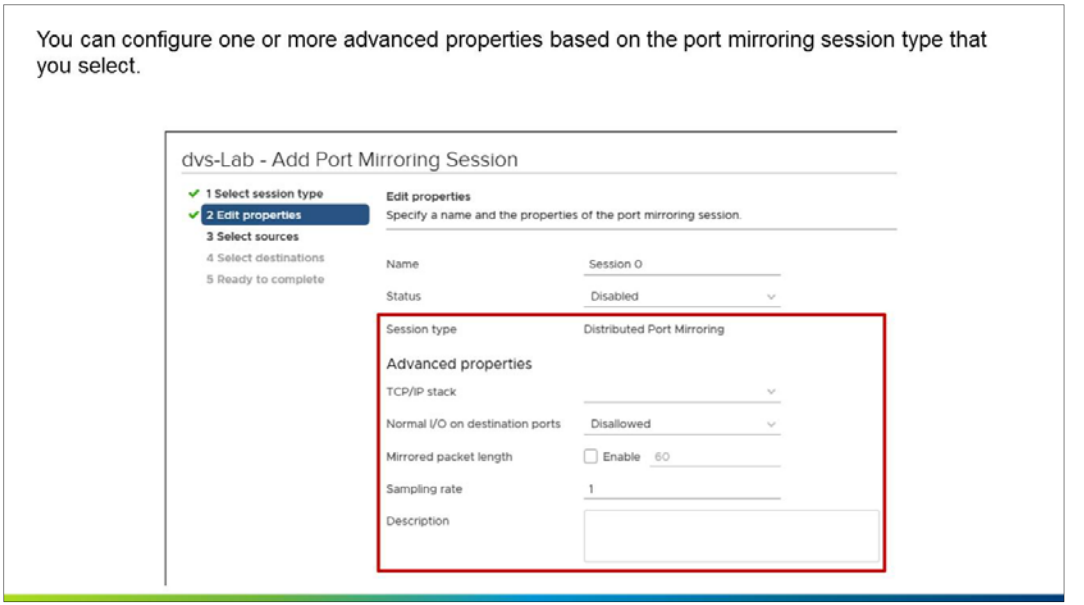
https://t.me/learningnets

- **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

2-73 Port Mirroring Session Properties



A distributed port mirroring type session can have the following properties:

- TCP/IP stack type (default or mirror)
- Normal I/O on destination ports (can be allowed or disallowed)
- 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/7.0/vsphere-esxi-vcenter-server-70-networking-guide.pdf.

https://t.me/learningnets

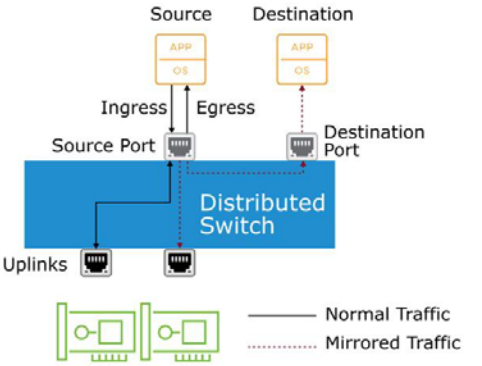
2-74 Source and Destination in a Port Mirroring Session

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 VM into the distributed switch. Egress traffic direction is traffic flowing from the distributed switch into the source VM.

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

2-75 Lab 5: 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

2-76 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe how to use NetFlow for monitoring network security and performance
- Configure port mirroring on a distributed switch

https://t.me/learningnets

2-77 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.
- Distributed switches support the use of network analysis and troubleshooting tools, specifically NetFlow and port mirroring.
- In vSphere 7, you can access VDS and NSX virtual distributed switch.

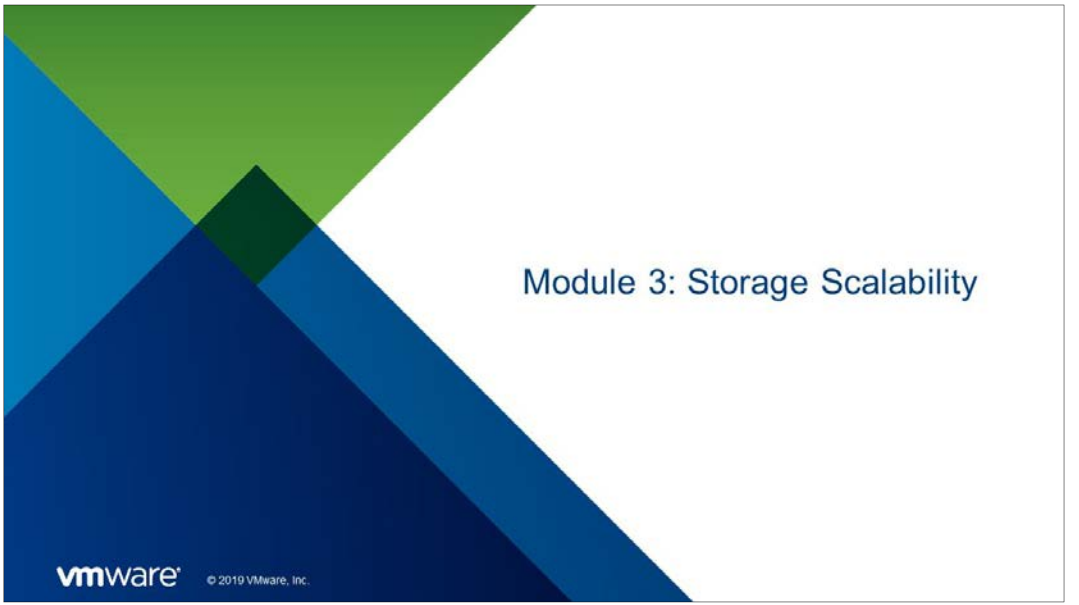
Questions?

https://t.me/learningnets

https://t.me/learningnets

Module 3

Storage Scalability



https://t.me/learningnets

3-2 Importance

The scalability features and software-defined storage configurations in vSphere can help you handle growth efficiently in your environment. For datastore growth and balancing issues, you can use vSphere Storage DRS. And with knowledge and skills in storage policy based management, you can align storage with the application demands of your VMs.

https://t.me/learningnets

3-3 Module Lessons

1. VMFS Datastores
2. vSphere Storage APIs - Array Integration
3. vSphere Storage API for Storage Awareness and vSphere APIs for I/O Filtering
4. Storage Policy Based Management Overview
5. Using VM Storage Policies
6. Storage Policies for vSAN
7. Storage Policies for vSphere Virtual Volumes
8. Storage I/O Control
9. Datastore Clusters and vSphere Storage DRS

https://t.me/learningnets

https://t.me/learningnets

3-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

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

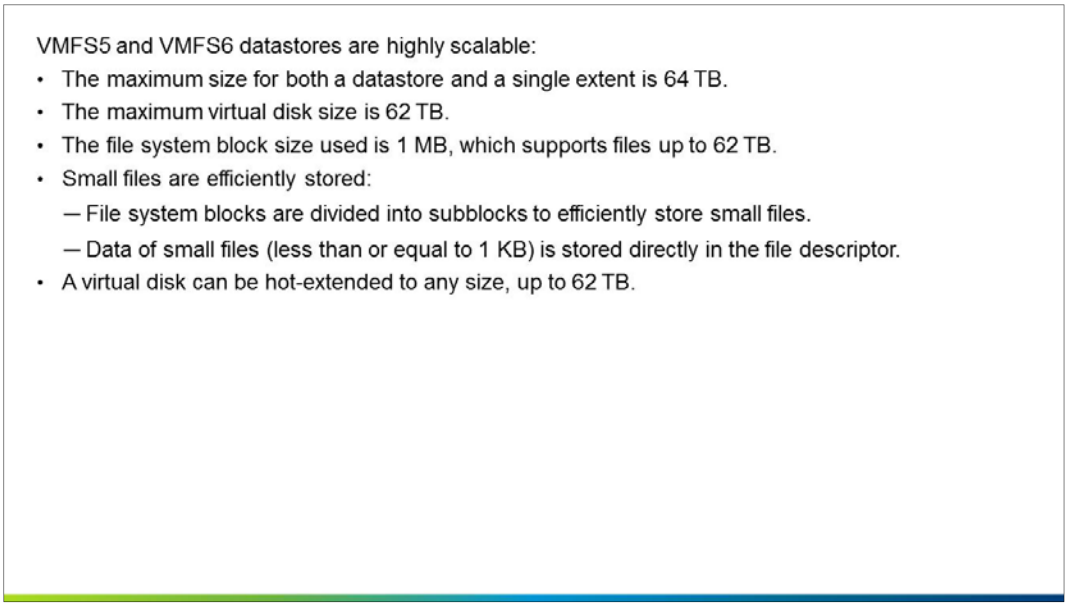
https://t.me/learningnets

3-6 About VMFS Datastores

VMFS is a high-performance, cluster file system that is optimized for VMs.
vSphere 6.5 and later versions support VMFS5 and VMFS6 datastores.
In vSphere 6.7, VMFS3 datastores are upgraded to VMFS5 when mounted.

https://t.me/learningnets

3-7 VMFS5 and VMFS6 Datastores



For more information about the hot extend feature, see vSphere Storage at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-storage-guide.pdf.

https://t.me/learningnets

3-8 VMFS6 Datastores

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 VMFS-6 at https://storagehub.vmware.com/t/vsphere-6-5-storage-1/vmfs-6-5/.

https://t.me/learningnets

3-9 VMFS 6 New Features in vSphere 7

Shared VMDK on VMFS6 supports SCSI-3 Persistent Reservations (SCSI-3 PR) at the virtual disk (VMDK) level:

- A Windows Server Failover Cluster (WSFC) can use shared disks on VMFS.
- It can support clustered virtual disks on specific datastores (requires supported hardware).
- You can use the vSphere Client to migrate from RDMs to VMFS.

Affinity 2.0 Manager is designed to be smarter and more efficient in the allocation of resources for thin disks:

- It minimizes overhead in the I/O path.
- Region Maps allows the Affinity Manager to quickly supply information about available RCs to the Resource Manager.

Affinity 2.0 Manager is a smarter and more efficient way to allocate resources and minimize overhead in the I/O path:

- Creates a cluster-wide view of the resource allocation status of the resource clusters (RC), dividing disk space into “regions.”
- Ranks regions into Region Maps based on the number of blocks allocated by the current host as well as other hosts.

https://t.me/learningnets

3-10 ESXi Support for VMFS5 and VMFS6

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 and later	Automatically upgraded to VMFS5 on first mount.	Yes	Yes	1024	4096

https://t.me/learningnets

3-11 Storage Device Formats

ESXi supports storage devices with traditional and advanced sector formats:

- 512n: A 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

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

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 leads to 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 improved performance is true only when the guest operating system is configured to issue I/O that are aligned to the 4K sector size.

https://t.me/learningnets

Because legacy applications and operating systems do not support 4K native drives, the drive industry introduced 4K sector size drives in 512-byte emulation mode to support legacy applications. The 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 help drive adoption of 4K drives. The goal is to move to 4K native drives.

https://t.me/learningnets

3-12 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 OS 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

3-13 Limitations of Using 4Kn Drives

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

- Only local SAS and 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.
- Only the Native Multipathing Plug-In (NMP) can claim the 4Kn devices:
  - The High-Performance Plug-In (HPP) cannot be used to claim these devices.
  - Third-party multipathing plug-ins are not supported.
- For vSAN, 4Kn drives can only be used as capacity HDDS in Hybrid mode.

https://t.me/learningnets

3-14 Snapshot Formats for VMFS

On the VMFS datastore, the delta disk of a VM snapshot is a sparse disk. Delta disks use different sparse formats depending on the type of VMFS datastore.
The SEsparse format is similar to VMFSsparse, with some enhancements:

- SEsparse is space efficient and supports space reclamation.
- This disk format is recommended for virtual desktop infrastructure workloads.

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.

When you take a snapshot, the state of the virtual disk is preserved, preventing the guest operating system from writing to the disk. 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.

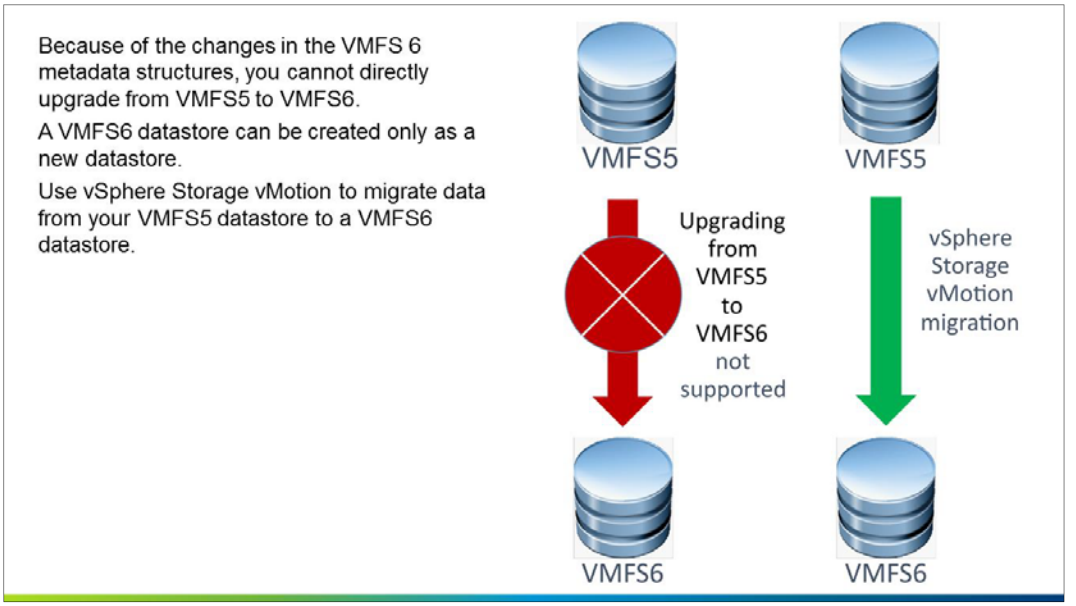
VMFSsparse is implemented on top of VMFS. The VMFSsparse layer processes I/O issued to a snapshot VM. Technically, VMFSsparse is a redo-log that starts empty, immediately after a VM snapshot is taken. The redo-log expands to the size of its base VM disk (VMDK), when the entire VMDK is rewritten with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

With SEsparse space reclamation, blocks that are deleted by the guest OS 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 OS deletes 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

3-15 Migrating Your Data from VMFS5 to VMFS6



For more information, see VMware knowledge base article 2147824 at https://kb.vmware.com/s/article/2147824.

https://t.me/learningnets

3-16 Activity: VMFS6 Features (1)

Which 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-extends of a virtual disk to any size up to 62 TB
- ❒ Parallel disk operations and concurrent transactions

https://t.me/learningnets

3-17 Activity: VMFS6 Features (2)

Which 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-extends of a virtual disk to any size up to 62 TB
- ✓ Parallel disk operations and concurrent transactions

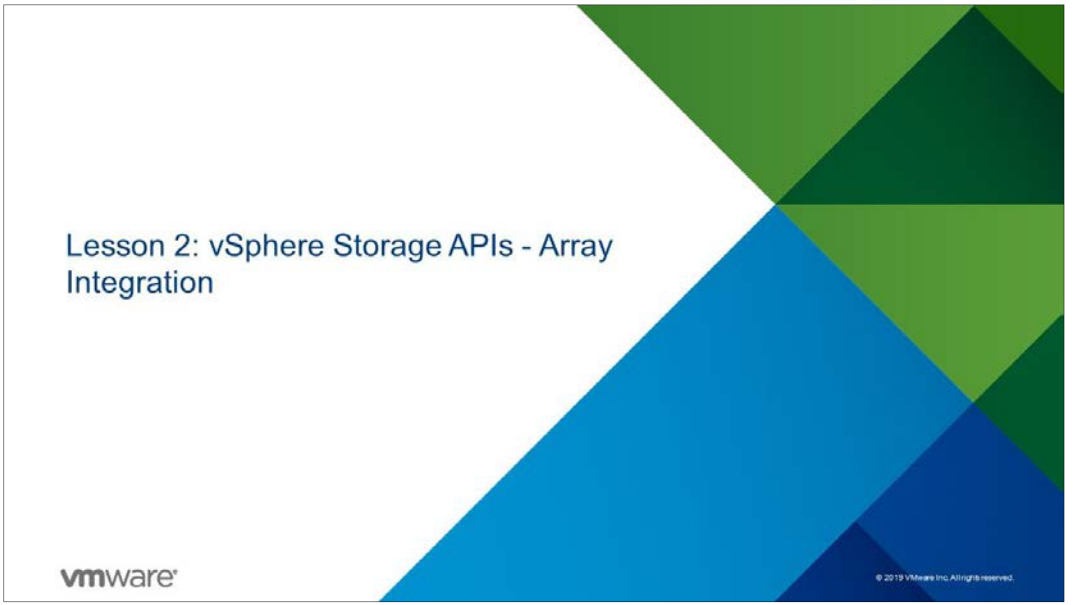
https://t.me/learningnets

3-18 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

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

https://t.me/learningnets



https://t.me/learningnets

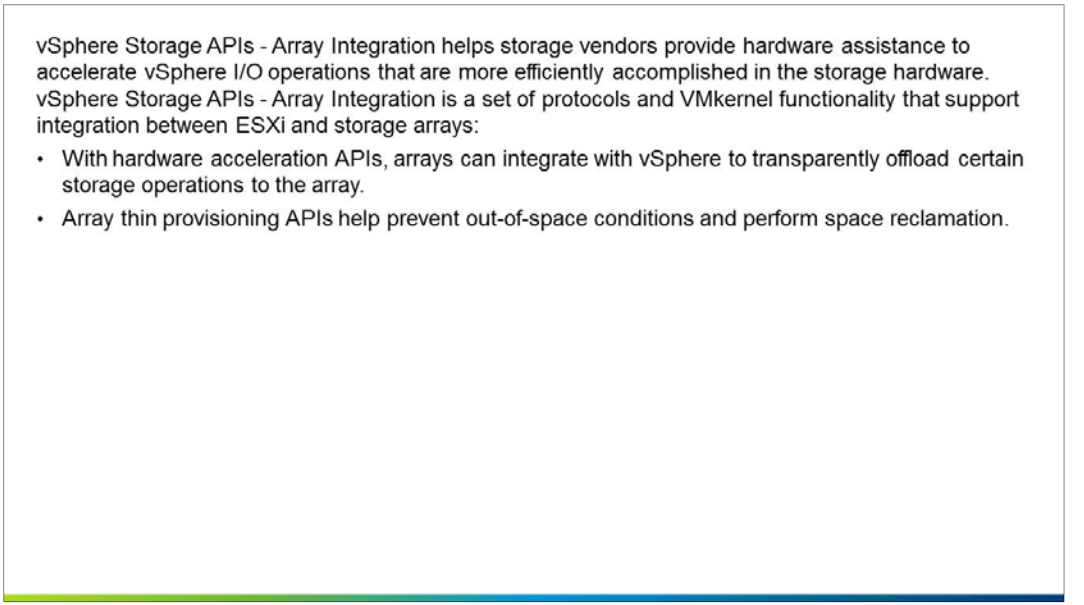
3-20 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe how vSphere Storage APIs - Array Integration helps storage arrays integrate with vSphere
- Explain how space reclamation works with VMFS datastores

https://t.me/learningnets

3-21 About vSphere Storage APIs - Array Integration



vSphere Storage APIs is a family of APIs that are 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 VM or per virtual disk file.

vSphere Storage APIs - Array Integration plug-ins can improve data transfer performance, and they are transparent to the end user.

https://t.me/learningnets

3-22 Hardware Acceleration APIs for Block Storage

Using hardware acceleration APIs, the ESXi host can offload specific VM and storage management operations to storage hardware:

- Use of these APIs significantly reduces the CPU overhead on the host.
- Block storage devices, such as Fibre Channel and iSCSI devices, support hardware acceleration.
- Hardware acceleration for block storage devices supports the several 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 template
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 storage hardware assistance, your host performs array 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: Storage arrays can make full copies of data in the array. The host does not need to read and write the data. This operation reduces the time and network load when cloning VMs, provisioning from a template, or migrating with vSphere Storage vMotion.
- Block zeroing: Storage arrays can zero out several blocks to provide newly allocated storage, free of previously written data. This operation reduces the time and network load when creating VMs and formatting virtual disks. This operation is activated when the Thick Provision Eager Zero option is used for a disk.
- Hardware-assisted locking: This option supports discrete VM locking without use of SCSI reservations. Disk locking is per sector, unlike SCSI reservations, which locks the entire logical unit number (LUN).

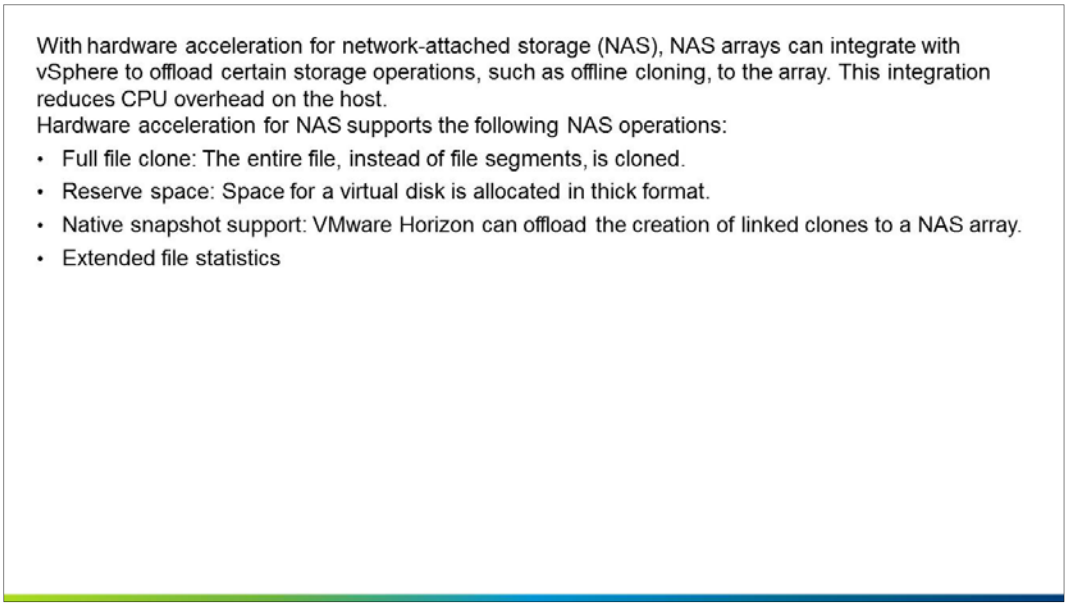
https://t.me/learningnets

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

3-23 Hardware Acceleration for NAS



With hardware acceleration, your host can 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: Storage arrays 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.

https://t.me/learningnets

- Native snapshot support: This operation enables View to offload the creation of linked clones to a NAS array.
- Extended file statistics: Storage arrays can accurately report space utilization.

Hardware acceleration for NAS is implemented through vendor-specific NAS plug-ins. Vendors typically create and support these plug-ins and distribute them as vSphere installation bundles (VIBs) through a webpage.

https://t.me/learningnets

3-24 Array Thin Provisioning APIs

With array thin provisioning APIs, a host can integrate with physical storage and detect space use 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 occur:
    - Virtual disks and VM files are deleted from the datastore.
    - Virtual disks and VM files are migrated off 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 that is required 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.

When the datastore expands, or when vSphere Storage vMotion is used to migrate VMs to a thin-provisioned LUN, the array thin provisioning APIs enable the host to communicate with the LUN. The host can warn about breaches in physical space or about out-of-space conditions

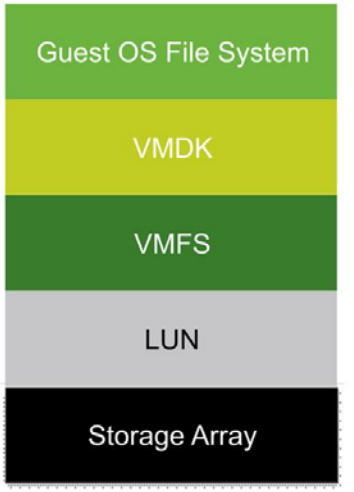
No installation steps are required for the array thin provisioning extensions. Array thin provisioning works on all VMFS5 and VMFS6 volumes. Device firmware that is enabled for this API must use the array thin provisioning features. ESXi continuously checks for firmware that is compatible with array thin provisioning. After firmware is upgraded, ESXi starts using the array thin provisioning features.

https://t.me/learningnets

3-25 About Space Reclamation (1)

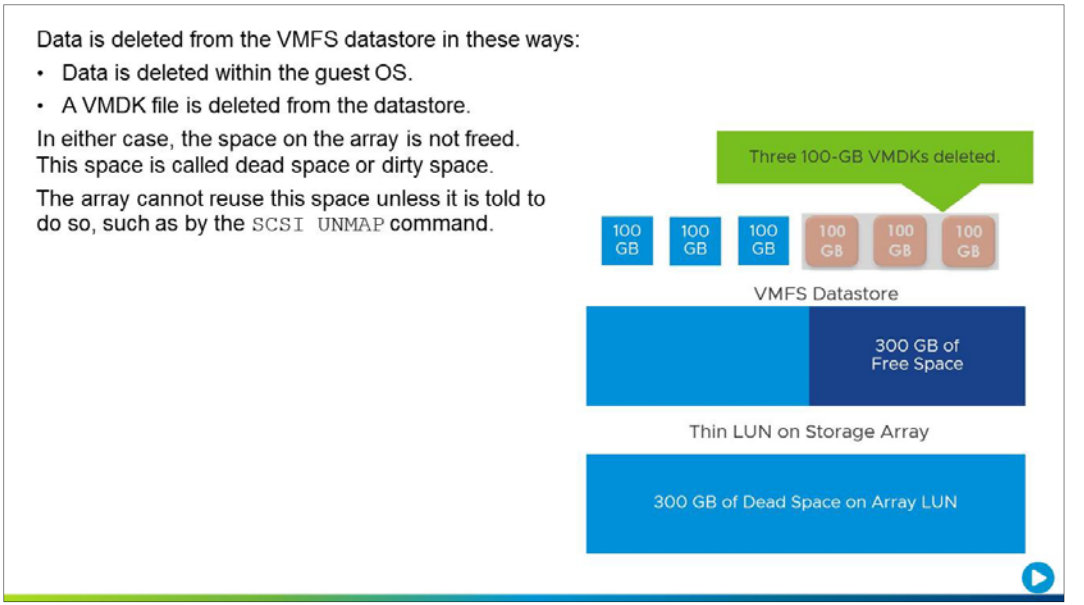
In a virtualized environment, different levels of abstraction exist at the storage layer:

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



https://t.me/learningnets

3-26 About Space Reclamation (2)



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

3-27 Space Reclamation with VMFS

ESXi supports the reclamation of free space. When you run the SCSI unmap command, an ESXi host informs the storage array to reclaim free space:

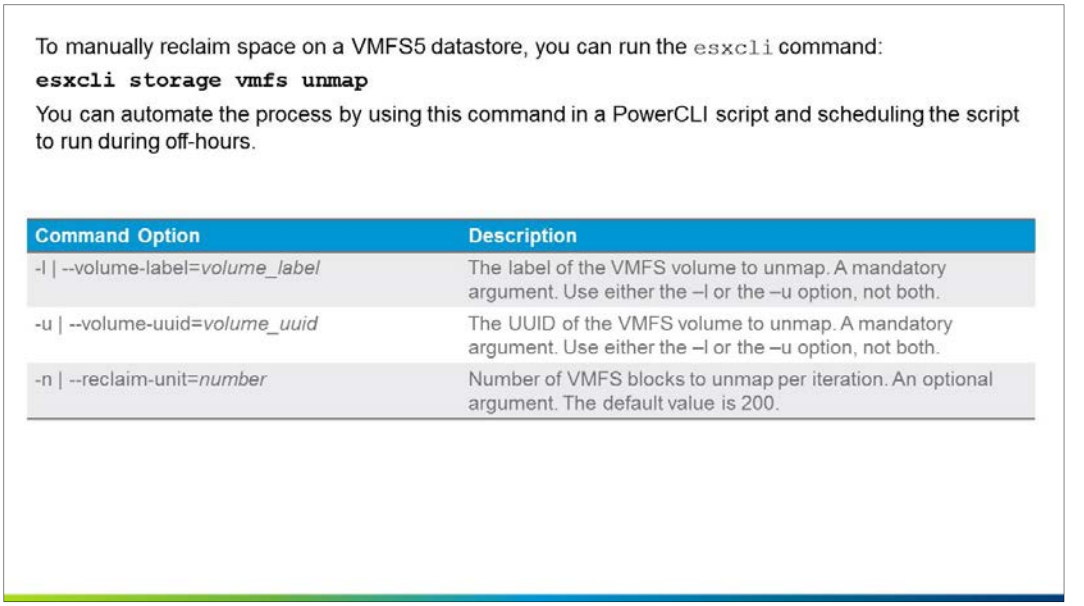
- With a VMFS5 datastore, you can manually send the unmap command with the `esxcli` command.
- With a VMFS6 datastore, the ESXi host automatically sends the unmap command to the storage array.

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

When you run the unmap command, an ESXi host informs the storage array that the VM files are removed or deleted from a thin-provisioned VMFS datastore. At this notification, the array can reclaim the freed blocks.

https://t.me/learningnets

3-28 Performing Manual Space Reclamation on a VMFS5 Datastore



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

This command can be run without a maintenance window:

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

When you run the command, it might send many `unmap` requests at a time, which can lock some of the resources during this operation.

https://t.me/learningnets

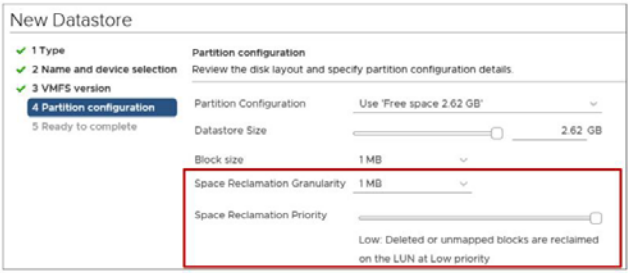
For more information about this command, see vSphere Storage at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-storage-guide.pdf.

https://t.me/learningnets

3-29 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 is 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.

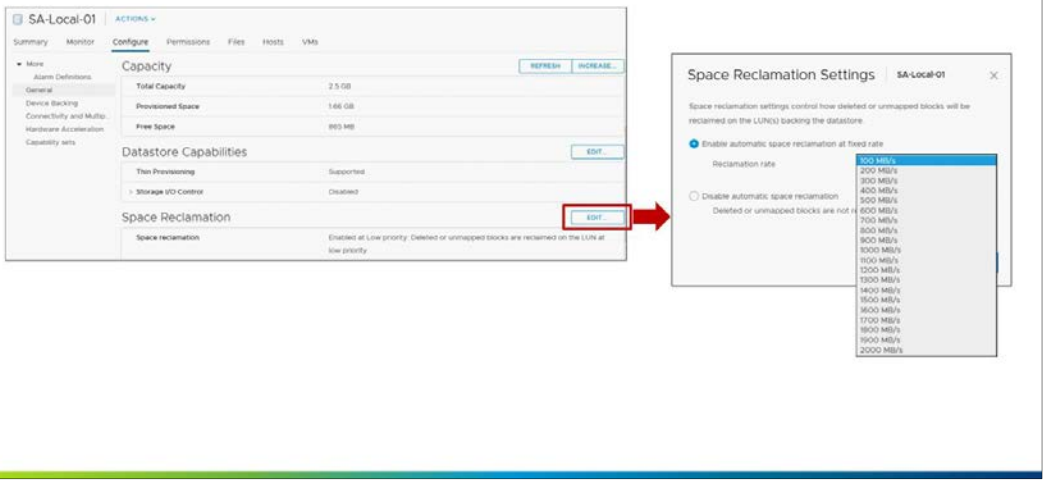
By default, the LUN performs the space reclamation operation at a low rate. You can also set the space reclamation priority to None to disable the operation for the datastore.

https://t.me/learningnets

3-30 Configuring the Automatic Unmap Rate

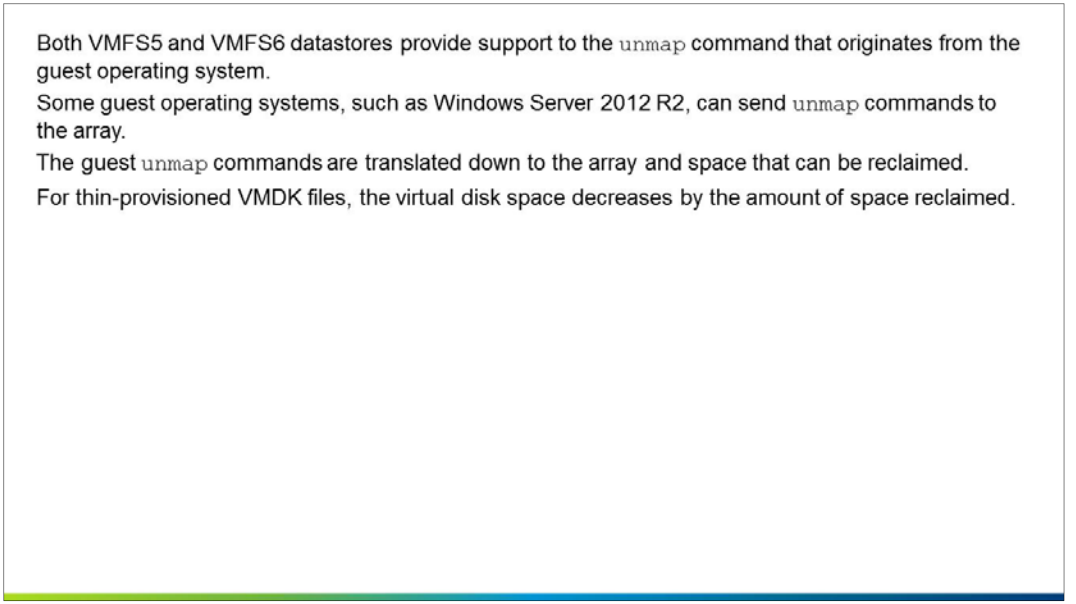
With vSphere 6.7 and later, you can set a fixed unmap rate for the datastore. The unmap rate is the rate at which automatic `unmap` commands are processed.

You can change the unmap rate to match the storage array's capabilities.



https://t.me/learningnets

3-31 Space Reclamation Support for Guest Operating Systems (1)



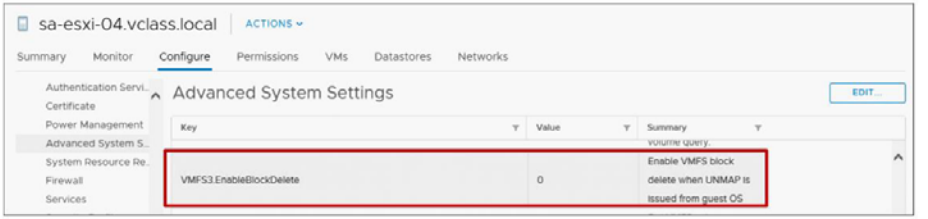
ESXi supports the `unmap` commands that are issued directly from a guest operating system to reclaim storage space.

https://t.me/learningnets

3-32 Space Reclamation Support for Guest Operating Systems (2)

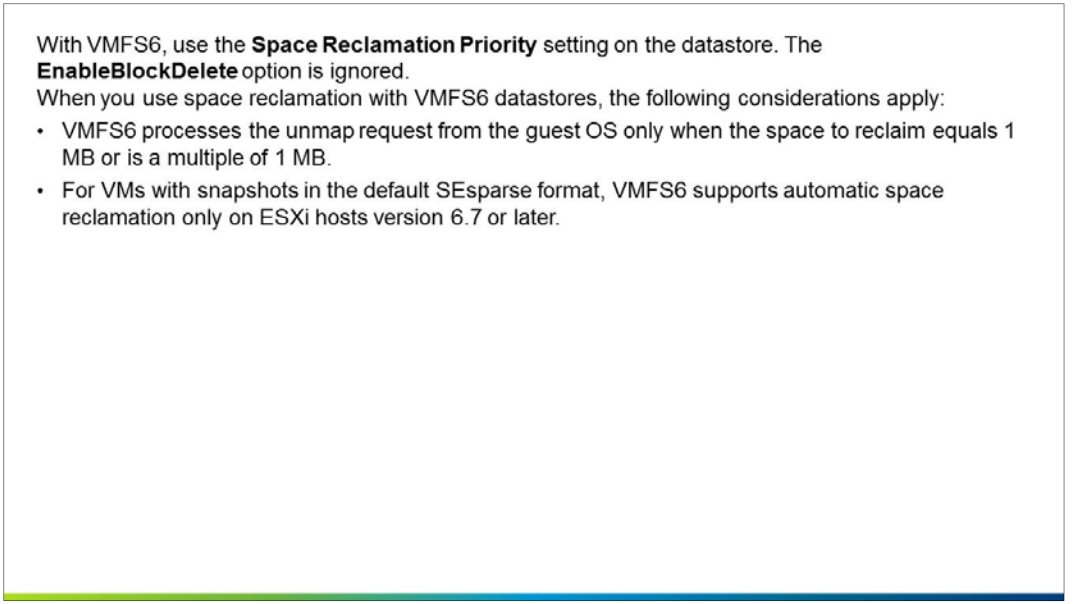
VMFS5 supports the automatic space reclamation requests for a limited number of guest operating systems.
To send unmap requests from the guest OS to the array, the VM must meet the following prerequisites:

- The virtual disk must be thin-provisioned.
- The guest OS must be able to identify the virtual disk as thin-provisioned.
- VM 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

3-33 Space Reclamation Support for Guest Operating Systems (3)



VMFS6 generally supports automatic space reclamation requests 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

3-34 Activity: vSphere Storage APIs – Array Integration (1)

Which actions apply to an ESXi host that uses storage devices that support vSphere Storage APIs – Array Integration? Select all that apply.

- ❒ Inform the storage array that it can reclaim free space in a VMFS datastore.
- ❒ Offload certain storage operations, such as offline cloning, to a NAS array.
- ❒ Retrieve information about a storage array, such as its storage topology and capabilities.
- ❒ Offload specific storage operations to a block storage device, such as Fibre Channel and iSCSI.

https://t.me/learningnets

3-35 Activity: vSphere Storage APIs – Array Integration (2)

Which actions apply to an ESXi host that uses storage devices that support vSphere Storage APIs – Array Integration? Select all that apply.

- ✓ Inform the storage array that it can reclaim free space in a VMFS datastore.
- ✓ Offload certain storage operations, such as offline cloning, to a NAS array.
- ❒ Retrieve information about a storage array, such as its storage topology and capabilities.
- ✓ Offload specific storage operations to a block storage device, such as Fibre Channel and iSCSI.

https://t.me/learningnets

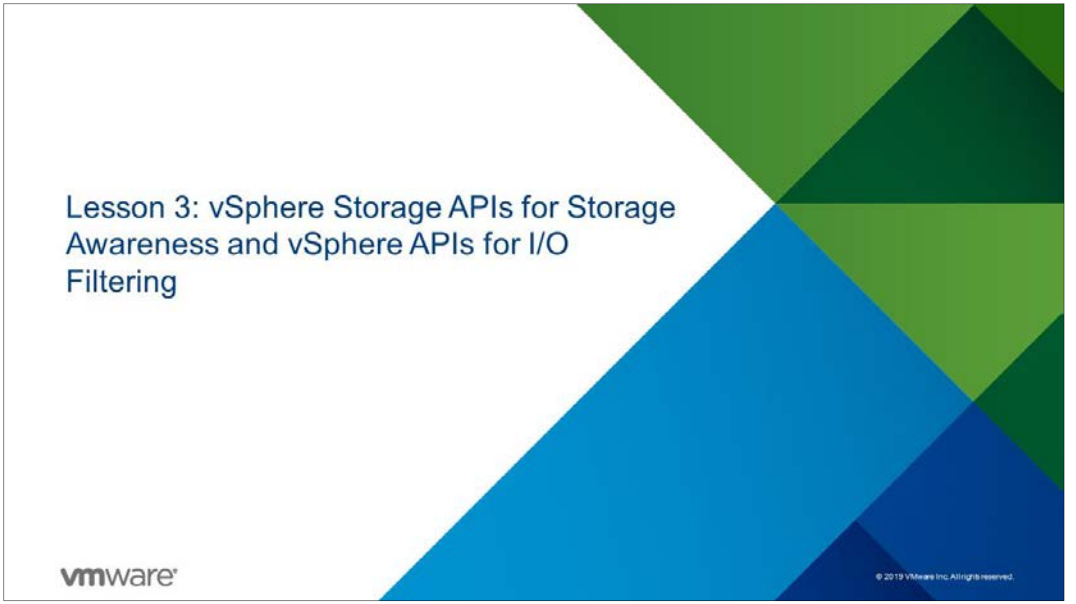
3-36 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe how vSphere Storage APIs - Array Integration helps storage arrays integrate with vSphere
- Explain how space reclamation works with VMFS datastores

https://t.me/learningnets

Lesson 3: vSphere Storage APIs for Storage Awareness and vSphere APIs for I/O Filtering



https://t.me/learningnets

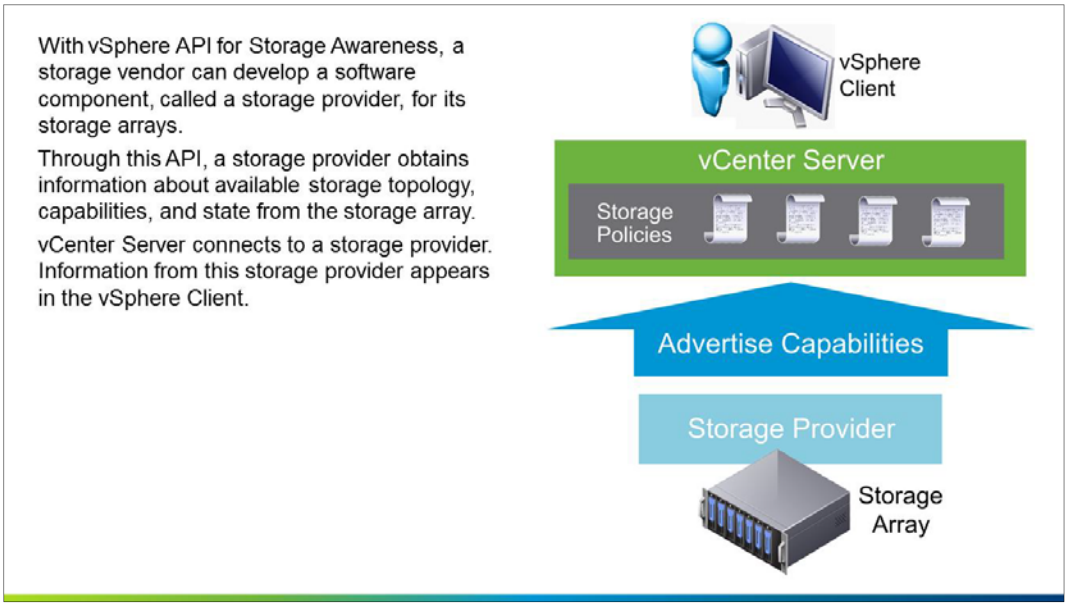
3-38 Learner Objectives

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

- Describe how vSphere API for Storage Awareness ensures that a VM's storage requirements are met
- Explain how vendors can create data services for VMs by using vSphere APIs for I/O Filtering

https://t.me/learningnets

3-39 About vSphere API for Storage Awareness



In vCenter Server, vSphere administrators cannot access the storage capabilities of the storage array on which their VMs are stored. VMs are provisioned to a storage black box. vSphere administrators see only a LUN identifier, such as a Network Address Authority ID (NAA ID) or a T10 identifier.

A storage vendor can use 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 storage vendor decides.

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 the vSphere Client.

A storage provider can report information about one or more storage devices. It can support connections to a single vCenter Server instance or to multiple vCenter Server instances.

https://t.me/learningnets

For information about the vSphere API for Storage Awareness program, see APIs for Storage Awareness (VASA) Package at https://developercenter.vmware.com/web/dp/other-programs/storage/vasa.

https://t.me/learningnets

3-40 Benefits of Storage Providers

Storage providers give vSphere administrators information about the topology, capabilities, and state of the physical storage devices for their VMs. Administrators can use this information in the following ways:

- To monitor the health and use of their physical storage devices
- To configure VM storage policies with the correct storage in terms of space, performance, and service-level agreement requirements

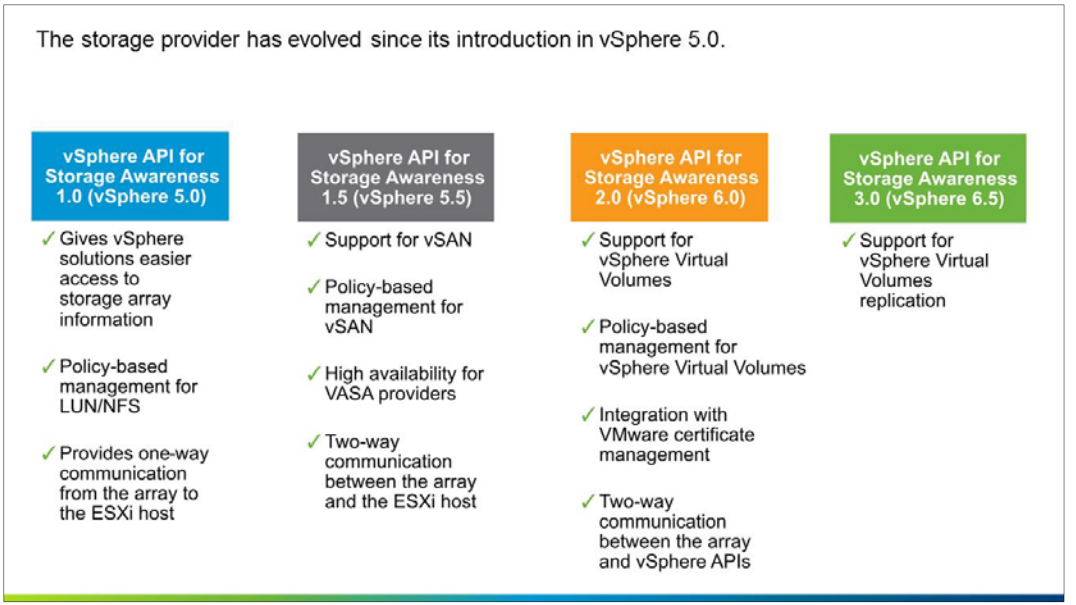
A storage provider provides capability information about the storage configuration, status, and storage data services in your environment. vSphere Storage DRS uses the information supplied by a storage provider to make VM placement and migration decisions that are compatible with the storage system.

A storage provider provides capability information:

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

https://t.me/learningnets

3-41 Comparison of Storage Provider Versions



https://t.me/learningnets

3-42 Storage Provider Categories

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

Persistence storage providers:

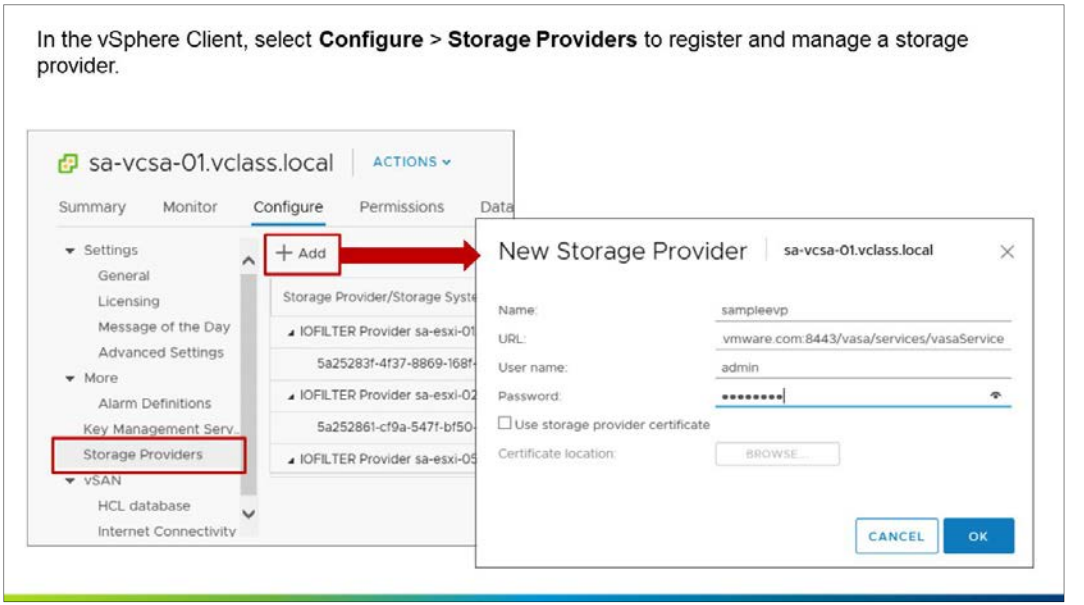
- Manage arrays and storage abstractions
- Can provide data services, such as replication
- Examples:
  - vSAN storage provider
  - vSphere Virtual Volumes storage provider

Third-party storage providers:

- Offered by a third party and typically require registration
- Examples:
  - vSphere Virtual Volumes storage provider
  - vSphere APIs for Storage Awareness storage providers

https://t.me/learningnets

3-43 Registering a Storage Provider



From the Storage Providers pane, you can register a storage provider. All system storage capabilities that are presented by the storage provider appear in the vSphere Client.

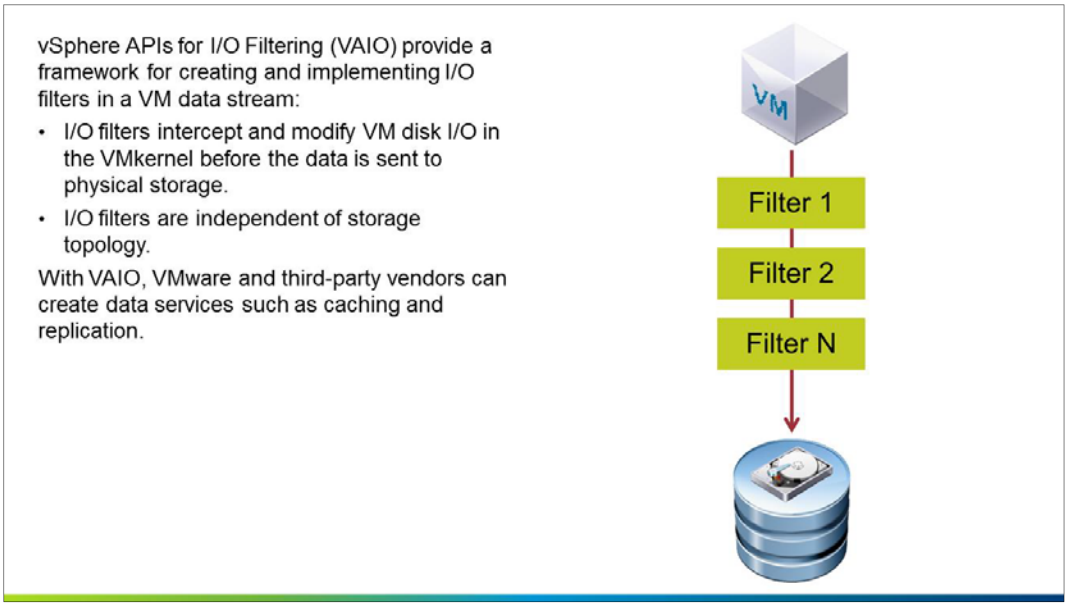
To register a storage provider, the storage vendor provides a URL, a login account, and a password.

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 how to *Register Storage Providers* see https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-B50E19B8-C1AD-4FE1-AD90-4DD6D29831FE.html.

https://t.me/learningnets

3-44 About vSphere APIs for I/O Filtering



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

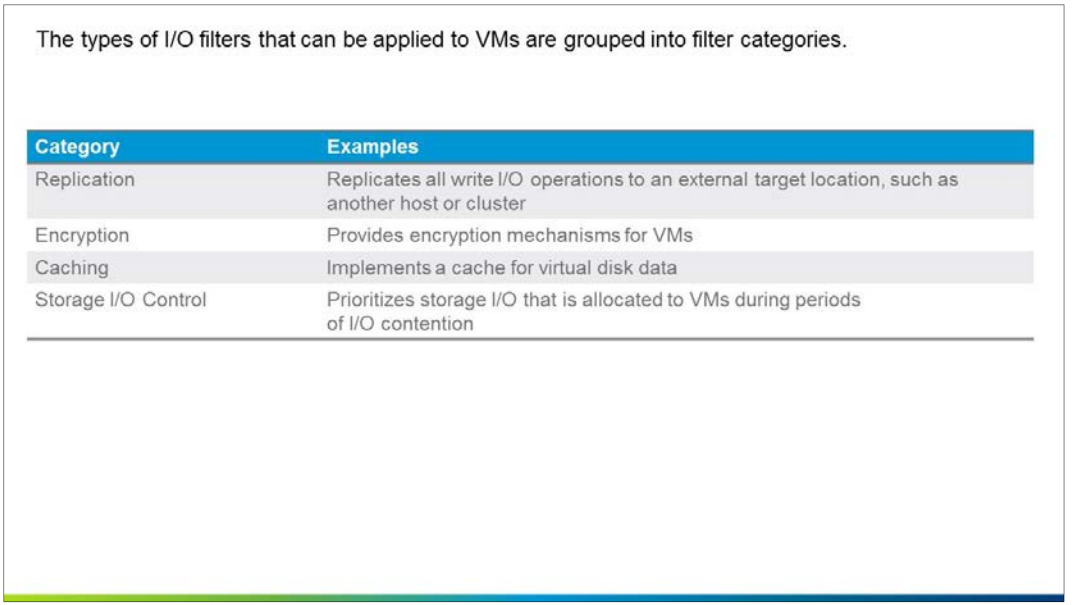
VMware offers certain categories of I/O filters. In addition, third-party vendors can create the I/O filters. Typically, these I/O filters are distributed as packages that provide an installer for deploying the filter components on vCenter Server and ESXi host clusters.

I/O filters can support all datastore types:

- VMFS
- NFS 3
- NFS 4.1
- vSphere Virtual Volumes
- vSAN

https://t.me/learningnets

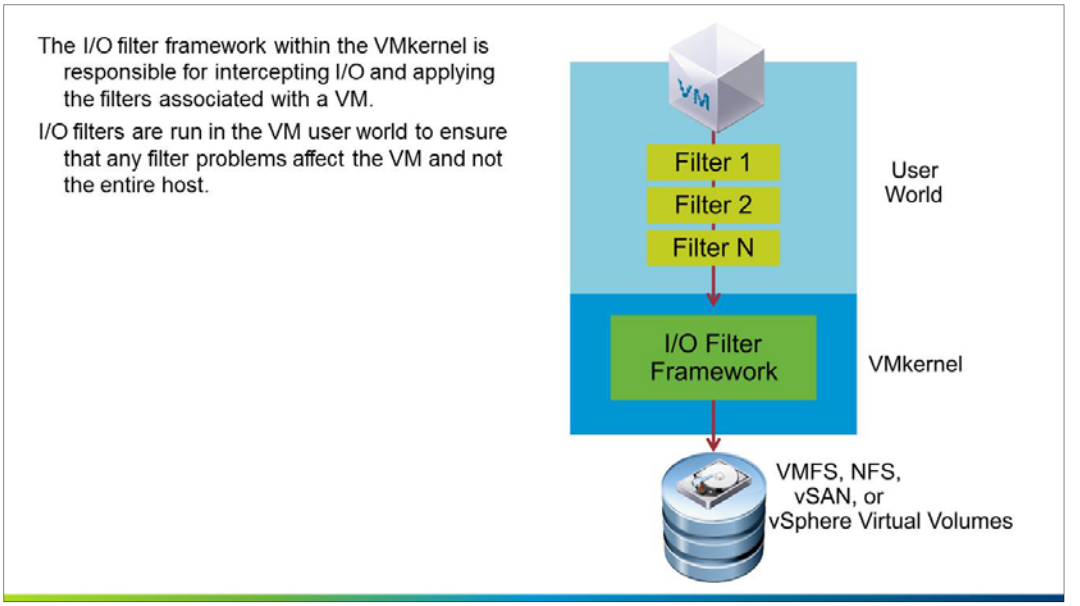
3-45 Types of I/O Filters



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

3-46 How I/O Filters Work (1)



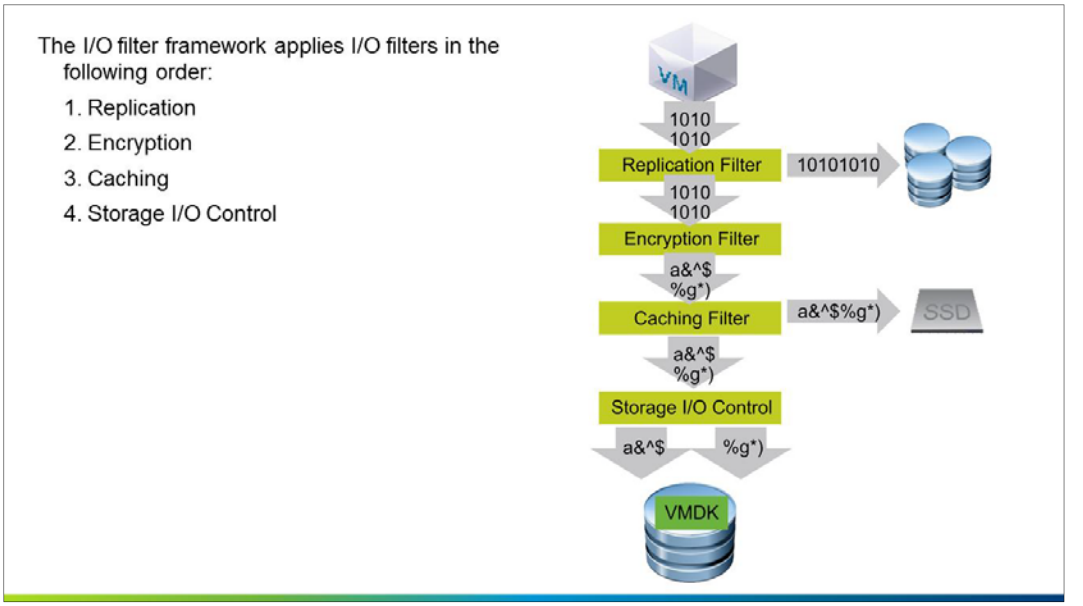
Filters operate by intercepting data in the VMkernel before the data is sent out to physical storage.

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 VM 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 VM'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

3-47 How I/O Filters Work (2)



The order in which I/O passes through these filter categories is predefined. This order, as shown on the slide, is important, especially if you plan to implement a replication filter and an encryption filter.

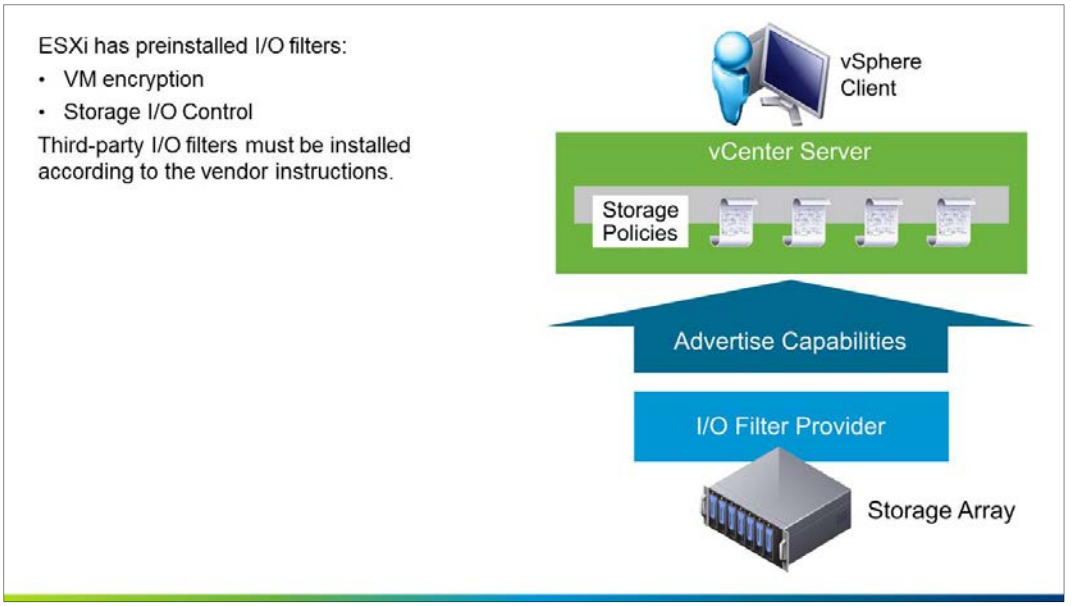
Data passes from the filters in this predefined order before being forwarded to the storage on which the VM files or objects reside.

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.

Because the replication filter comes before the encryption filter, the data that is replicated is in plain text. Additional security measures must be taken to protect these transmissions and the data on the replication target.

https://t.me/learningnets

3-48 I/O Filter Providers



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. You can use these capabilities to build storage policies.

Storage providers for I/O filtering are software components in vSphere. These providers integrate with I/O filters and report data service capabilities that are supported by I/O filters 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.

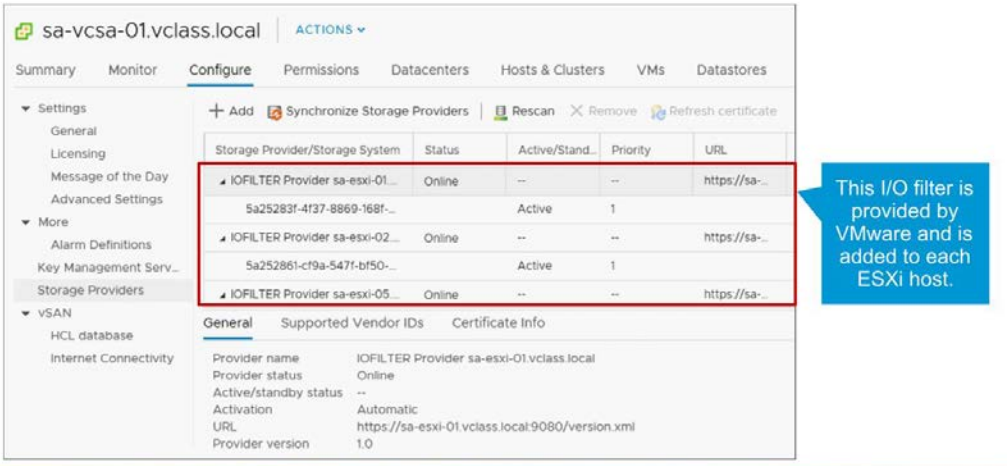
https://t.me/learningnets

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

https://t.me/learningnets

3-49 Displaying I/O Filter Provider Information

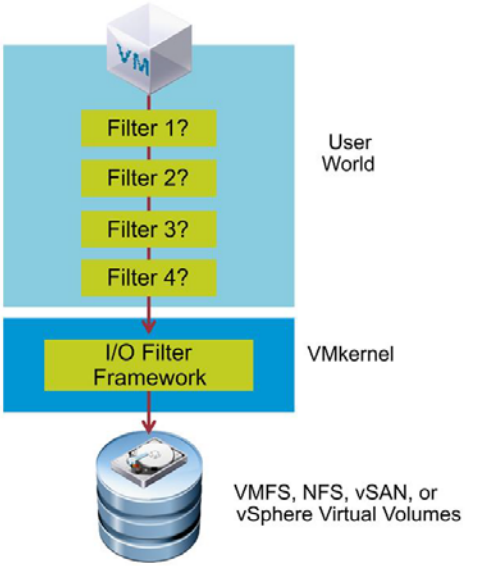
To view the I/O filter providers for each host, select the vCenter Server object in the inventory and select **Configure > Storage Providers**.



https://t.me/learningnets

3-50 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.

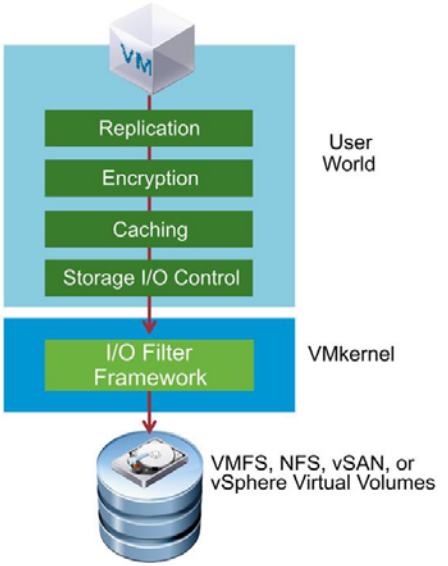


https://t.me/learningnets

3-51 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

3-52 Review of Learner Objectives

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

- Describe how vSphere API for Storage Awareness ensures that a VM's storage requirements are met
- Explain how vendors can create data services for VMs by using vSphere APIs for I/O Filtering

https://t.me/learningnets

https://t.me/learningnets

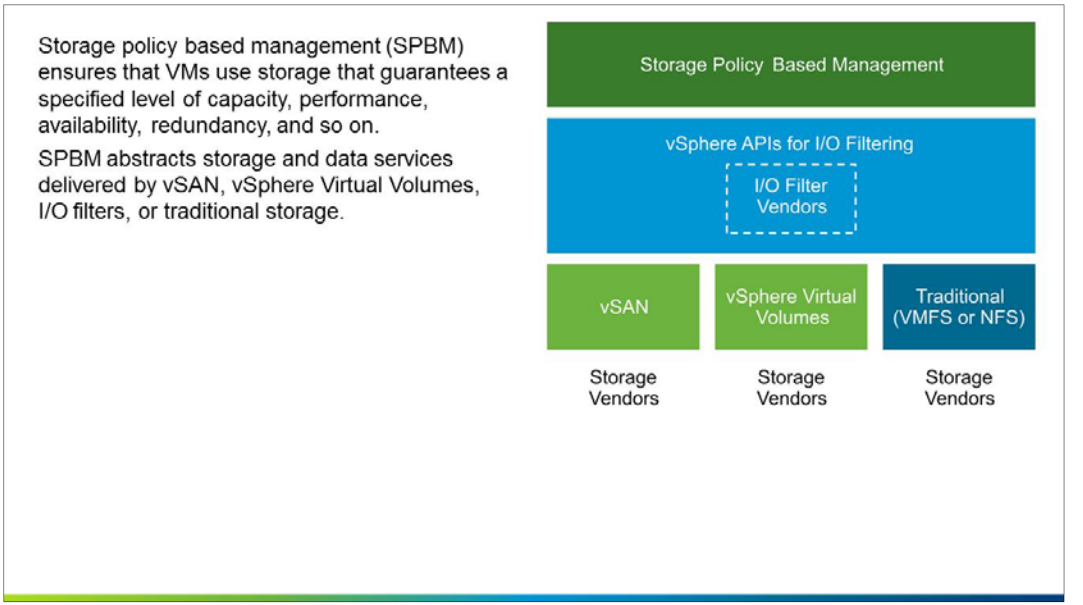
3-54 Learner Objectives

At the end of this lesson, you should be able to meet the following objective:

- Describe storage policy based management
- Recognize the categories of storage policy rules

https://t.me/learningnets

3-55 About Storage Policy Based Management (1)



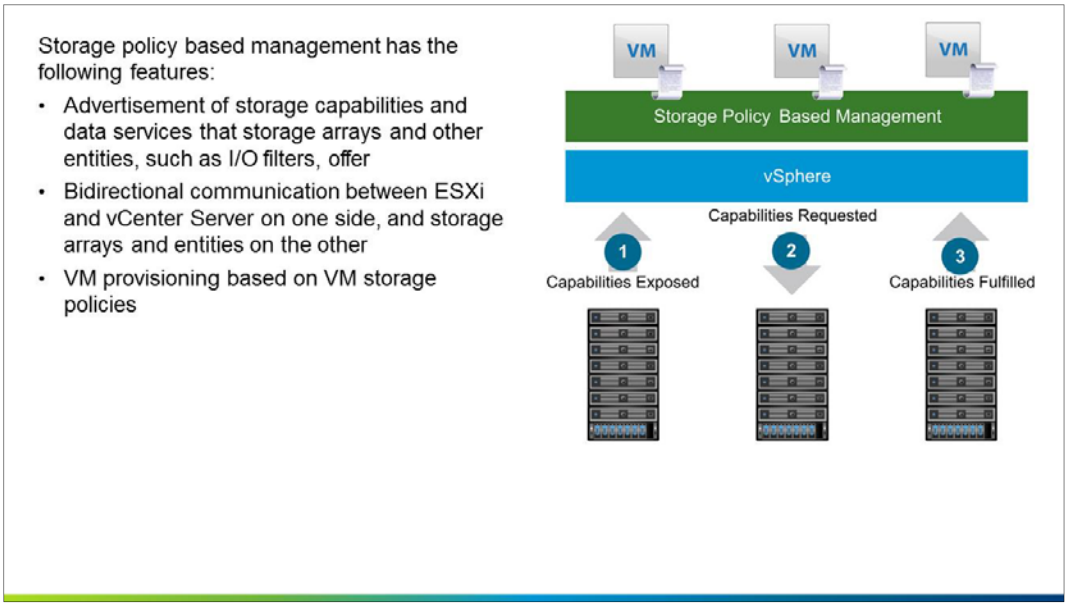
VMware partners and storage vendors can support these storage and data services.

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

When an application is to be provisioned, vSphere uses the Storage Policy Based Management module to match the policy requirements against configured storage container capabilities. In this way, vSphere determines where the VMDK files should exist.

https://t.me/learningnets

3-56 About Storage Policy Based Management (2)



VM storage policies minimize the amount of storage planning that you must do for each VM.

For example, you can use VM 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.

The capabilities exposed are performance, capacity, protection, and so on.

The capabilities requested are aligned on VM boundaries.

The capabilities fulfilled are datastore provisioned, and the policy is kept with the datastore.

https://t.me/learningnets

3-57 About VM Storage Policies

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.

VM storage policies are essential to VM provisioning.

You use VM storage policies during the provisioning of a VM to ensure that a VM's disks are placed on the storage that is best for its situation.

For example, with VM storage policies, you can ensure that the VM running a critical I/O intensive database is placed in the gold tier.

Ideally, you want to create the best match of predefined VM storage requirements with available physical storage properties.

https://t.me/learningnets

3-58 About Storage Policy Rules

Rules are the basic elements of a storage policy. Each rule is a statement that describes a single requirement for VM 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 VM storage objects are allocated within a datastore to achieve 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 VM.

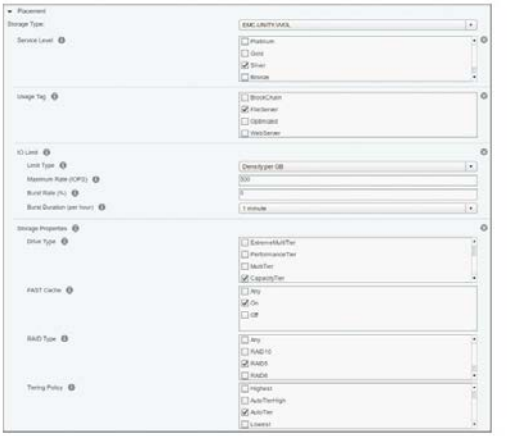
Data service rules activate specific data services. Storage systems or other entities can provide these services. They can also be installed on your hosts and vCenter Server. You include the data service rules in the storage policy components.

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

https://t.me/learningnets

3-59 Storage Policy Example (1)

When you create a storage policy, you can select one or more rules that represent service levels demanded by VMs.
These rules define storage characteristics.



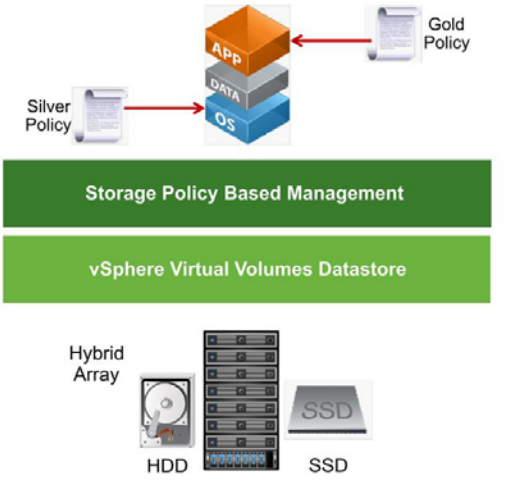
Each array vendor has an identified namespace. The namespace contains the storage container identifiers and assigned capabilities. The capabilities appear as options in the client interface. In this example, the vSphere Client is used.

https://t.me/learningnets

3-60 Storage Policy Example (2)

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



https://t.me/learningnets

3-61 About Storage Policy Components

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

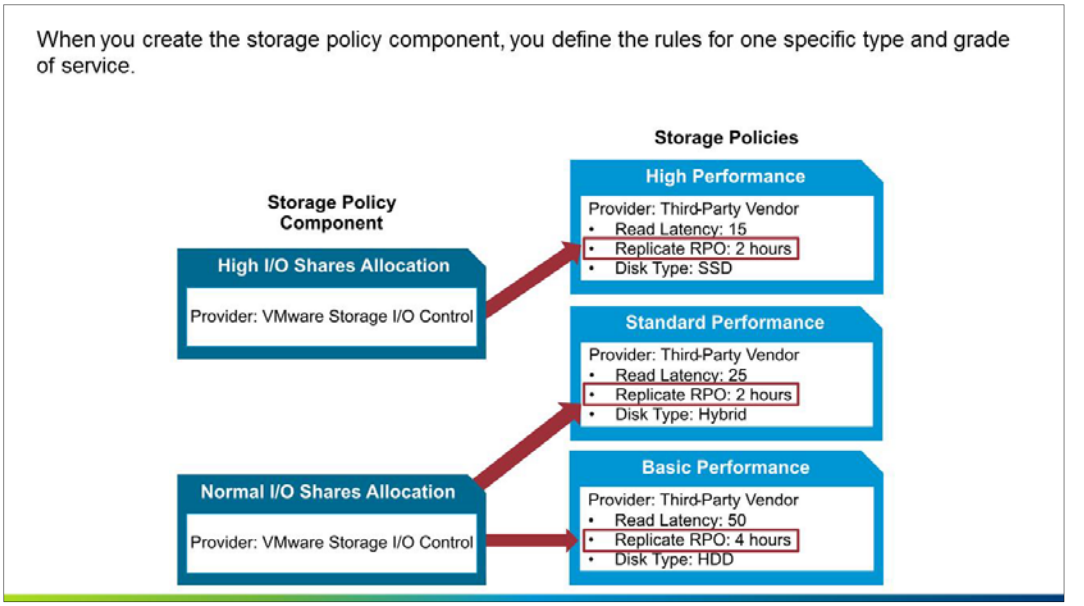
Each component describes one type of service from a storage provider.
The services vary depending on the storage provider that you use, but they generally belong to one of the following categories:

- 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

3-62 Storage Policy Component Example



You must add the component to the storage policy and 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

3-63 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

3-64 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, not data service rules, guide decisions about VM placement. Data service rules activate specific data services (such as caching or replication) for the VM.

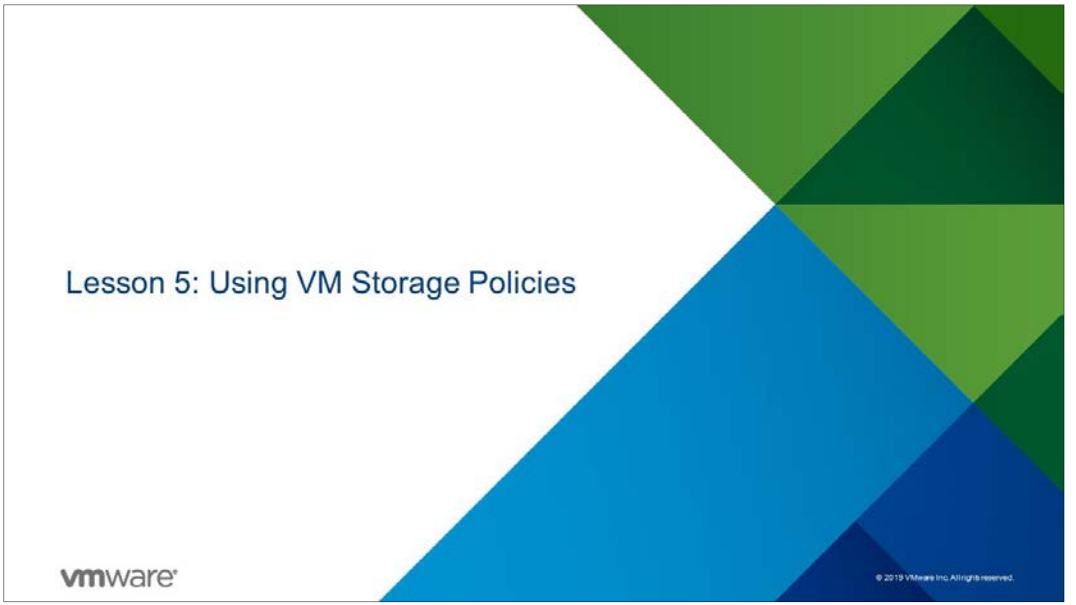
https://t.me/learningnets

3-65 Review of Learner Objectives

At the end of this lesson, you should be able to meet the following objective:

- Describe storage policy based management
- Recognize the categories of storage policy rules

https://t.me/learningnets



https://t.me/learningnets

3-67 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Create VM storage policies
- Use VM storage policies to create tiered storage

https://t.me/learningnets

Creating and Managing Storage Policies

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

1. Configure storage to 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

3-69 Step 1: Configuring Storage to Be Used for Storage Policies

When you create storage policies, the vSphere Client shows information from 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, the appropriate provider should be 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, create datastore tags:

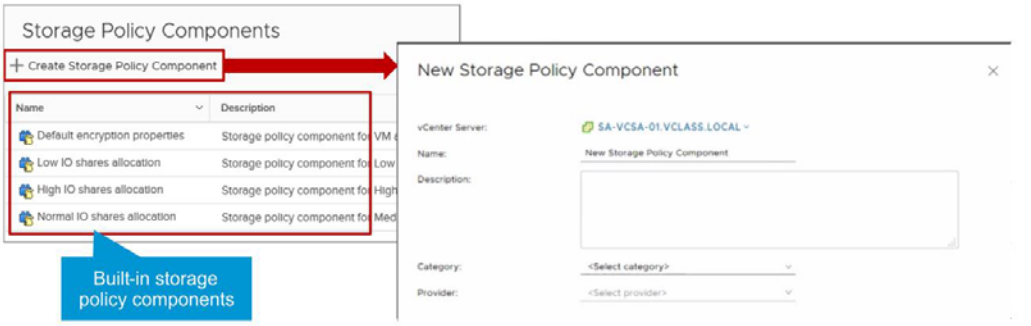
- Use tags to indicate a property that is not communicated through the storage providers, such as geographical location or administrative group.

https://t.me/learningnets

3-70 Step 2: Creating Storage Policy Components

You 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 Storage I/O Control.

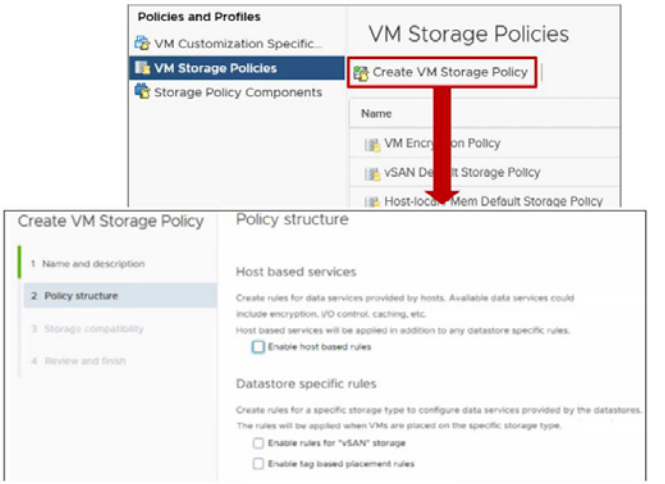


https://t.me/learningnets

3-71 Step 3: Creating VM Storage Policies

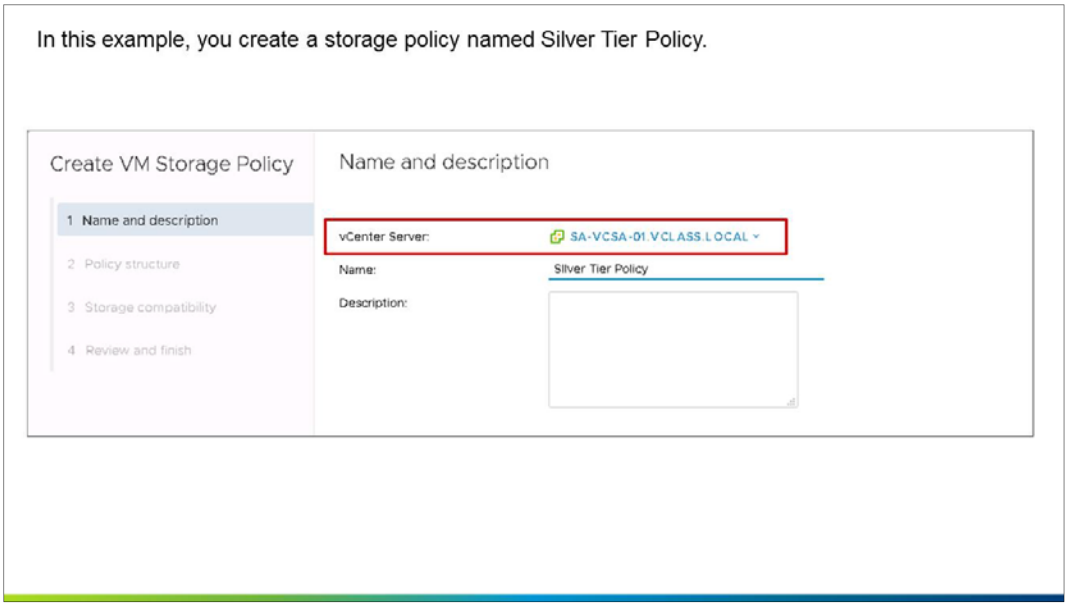
When you create a storage policy, you must select the policy structure:

- **Host based services** define rules to enable and configure data services provided by your host.
- **DataStore specific rules** describe requirements or services for a storage type.



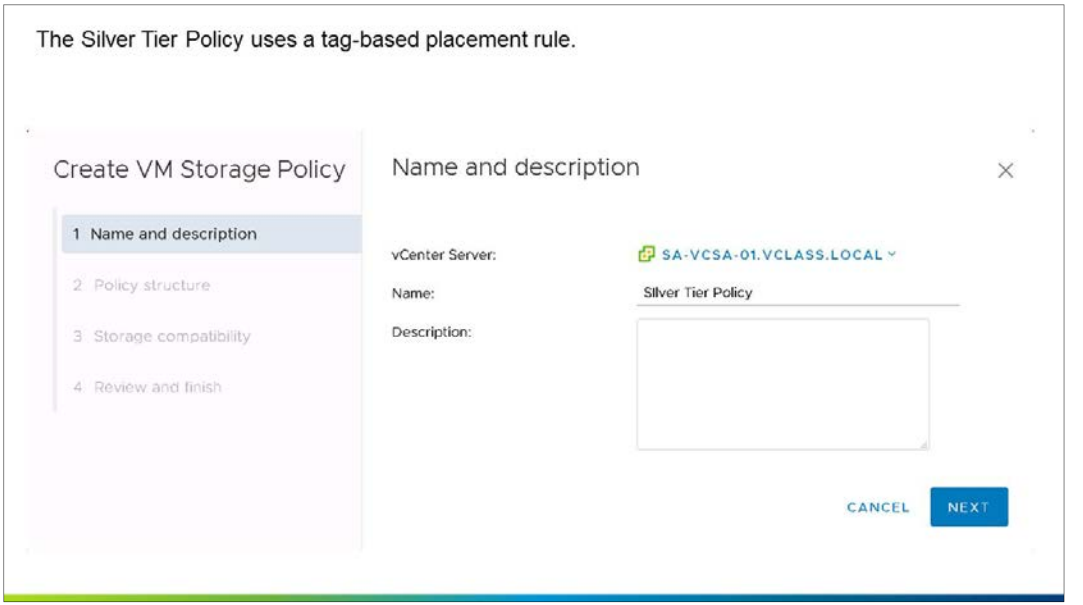
https://t.me/learningnets

3-72 Example: Creating a Silver Tier Storage Policy (1)



https://t.me/learningnets

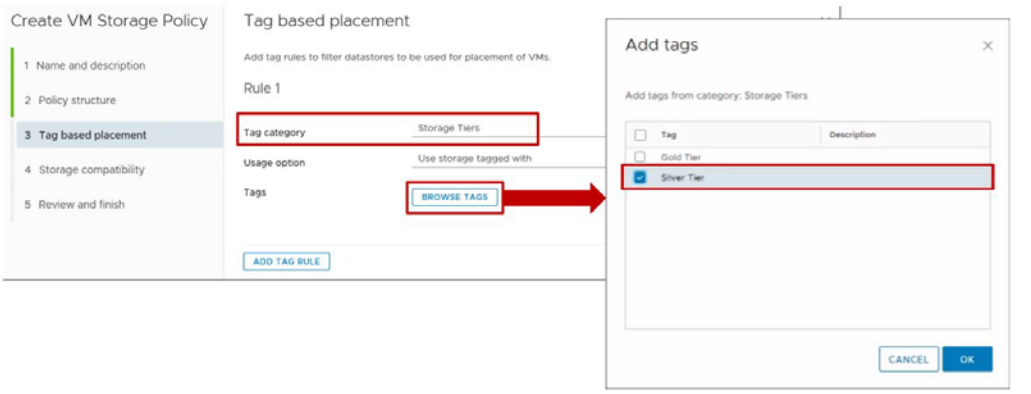
3-73 Example: Creating a Silver Tier Storage Policy (2)



https://t.me/learningnets

3-74 Example: Creating a Silver Tier Storage Policy (3)

You select the **Storage Tiers** tag category. Then, you select the tag called **Silver Tier**.
The Silver Tier Policy uses a datastore that is associated with the tag Silver Tier.



https://t.me/learningnets

3-75 Example: Creating a Silver Tier Storage Policy (4)

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

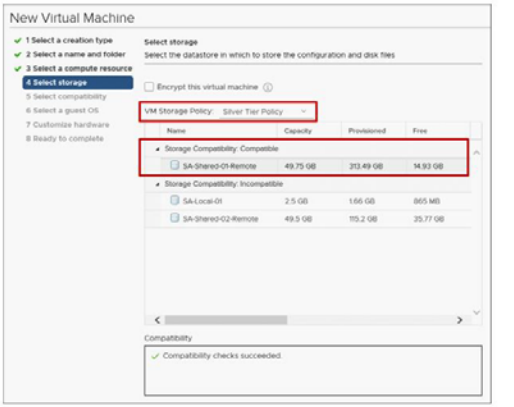
https://t.me/learningnets

3-76 Step 4: Applying 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.

When you assign the Silver Tier Policy to a VM, a list of compatible datastores is shown.

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



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

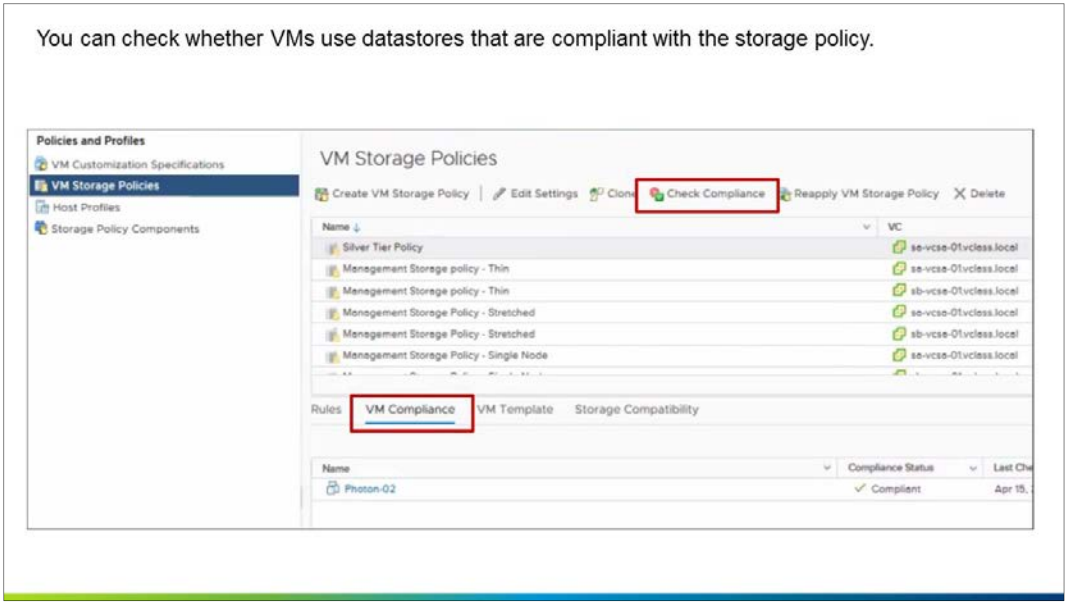
When a VM storage policy is selected, the datastores are divided into two categories: compatible and incompatible.

You can select datastores outside of the VM storage policy, but these datastores put the VM into a noncompliant state.

By using VM storage policies, you can easily see whether storage is compatible or incompatible. You do not need to ask the vSAN administrator, or refer to a spreadsheet of NAA IDs, every time you deploy a VM.

https://t.me/learningnets

3-77 Step 5: Checking Compliance for the VM Storage Policy



You can use VM storage policies during the ongoing management of VMs.

You can periodically check whether a VM is migrated to or created on inappropriate storage, potentially making it noncompliant.

You can also use storage information to monitor the health and use of storage and report when the VM's storage is noncompliant.

https://t.me/learningnets

3-78 Lab 6: Using 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 VM's Storage
3. Configure Storage Tags
4. Create VM Storage Policies
5. Assign Storage Policies to VMs

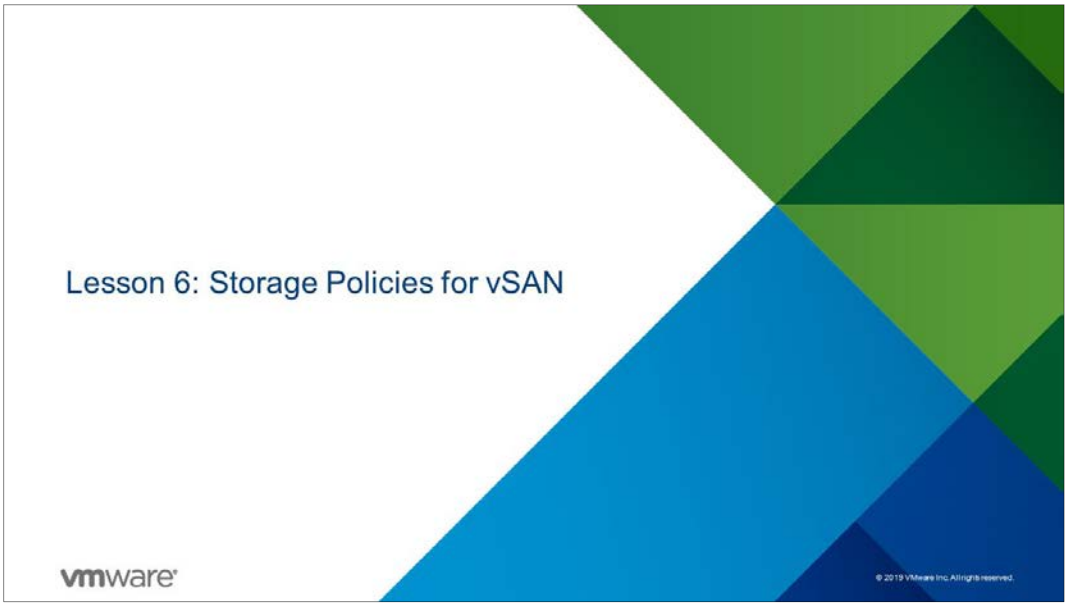
https://t.me/learningnets

3-79 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Create VM storage policies
- Use VM storage policies to create tiered storage

https://t.me/learningnets



https://t.me/learningnets

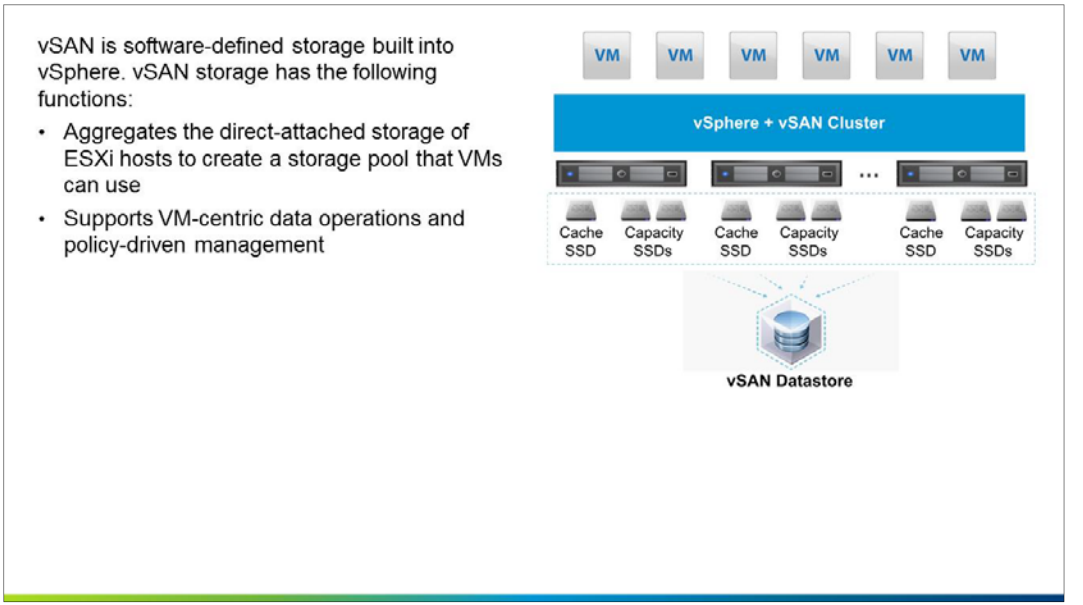
3-81 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Identify components in the vSAN architecture
- Create vSAN storage policies

https://t.me/learningnets

3-82 About vSAN (1)



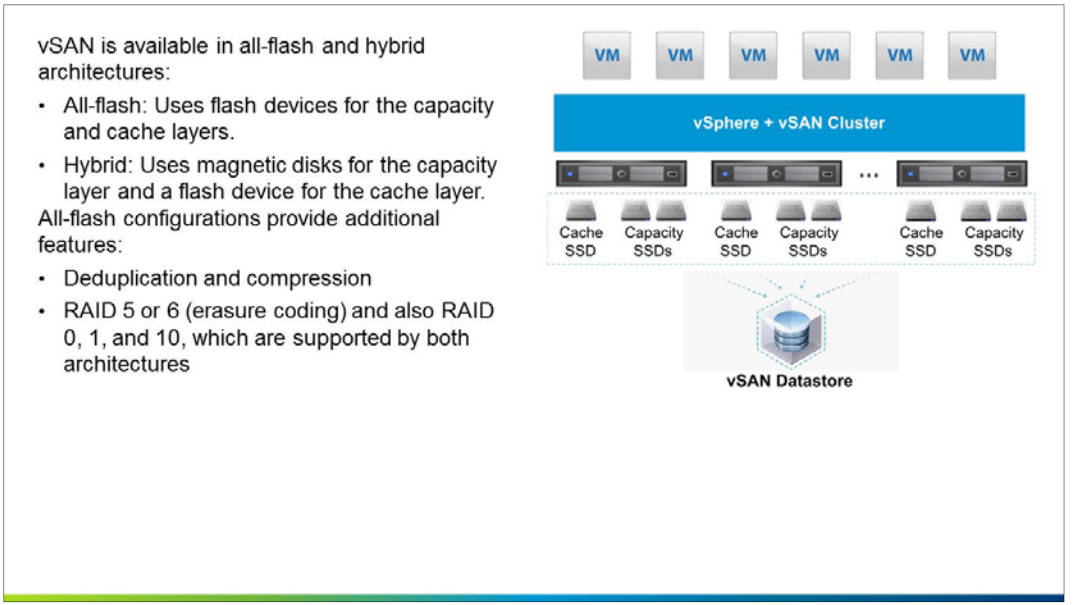
In a vSAN environment, ESXi hosts are configured to form a vSAN cluster. All ESXi hosts communicate through a dedicated network. At least three hosts in the cluster must have local storage devices. Although it is not a best practice, hosts without local storage can share their compute resources and use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

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

- Storage policy per VM architecture: With multiple policies per datastore, each VM can have different storage configurations in terms of availability, performance, and capacity.
- vSphere and vCenter Server integration: vSAN capability is built into the VMkernel and requires no appliance.
- 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 VMs that are configured for vSAN.

https://t.me/learningnets

3-83 About vSAN (2)



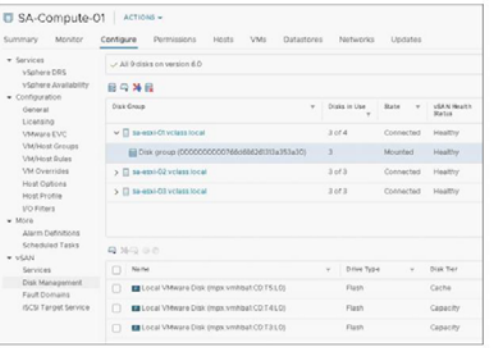
All-flash configurations can use features that are not available in hybrid configurations:

- 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

3-84 vSAN Disk Architecture

- Disks from up to 64 ESXi hosts are grouped to form a vSAN datastore.
- Each host that contributes storage in a vSAN cluster contains at least one disk group.
- Each host can support a maximum of five disk groups.
- A disk group contains at most one cache device and between one and seven capacity devices.



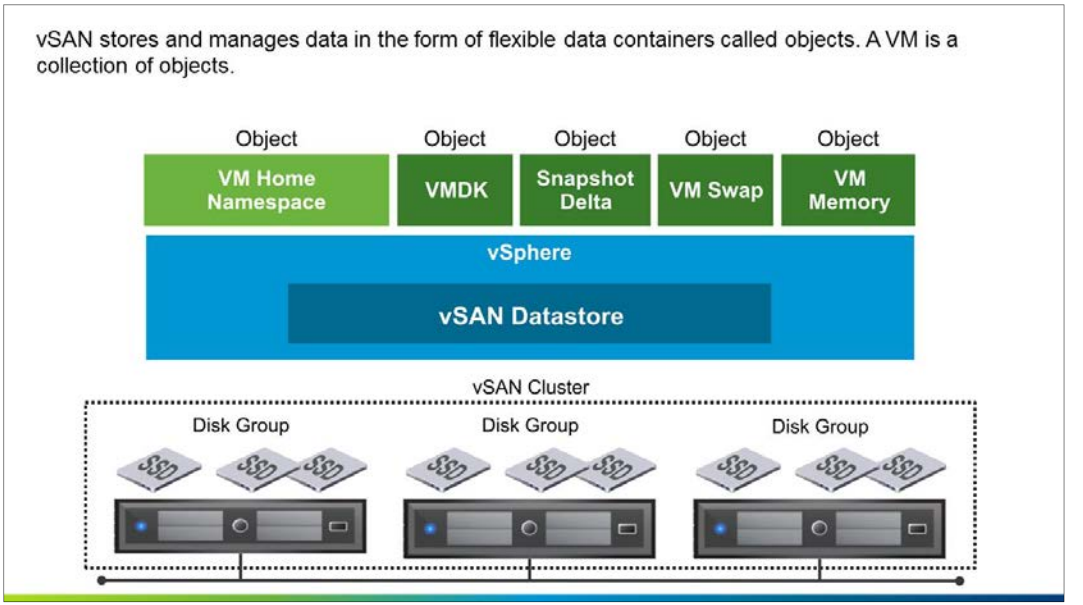
In the vSAN cluster, the local storage components of the ESXi hosts are combined to create a vSAN datastore. Only one datastore can be created for each vSAN cluster.

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.

At most, a vSAN host can have 5 disk groups, each containing up to 7 capacity devices, resulting in a maximum of 35 capacity devices for each host.

https://t.me/learningnets

3-85 Object-Based Storage



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

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

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

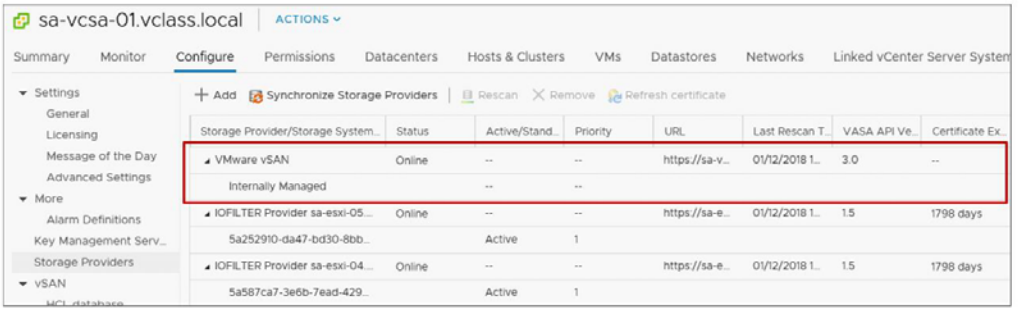
https://t.me/learningnets

3-86 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 the vSphere Client.

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



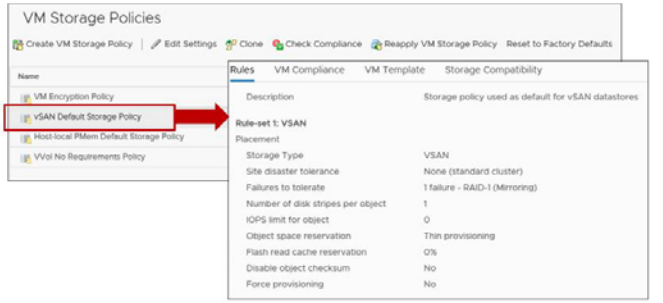
vSAN storage providers report a set of underlying storage capabilities to vCenter Server. They also communicate with the vSAN layer to report the storage requirements of the VMs.

https://t.me/learningnets

3-87 vSAN Storage Policies

You can create multiple VM storage policies for a single vSAN datastore:

- vSAN has a default VM storage policy.
- You can create custom storage policies that use vSAN capabilities.
- The default vSAN storage policy is used, unless you select a different storage policy.
- You use the vSphere Client to view, create, and modify policies.

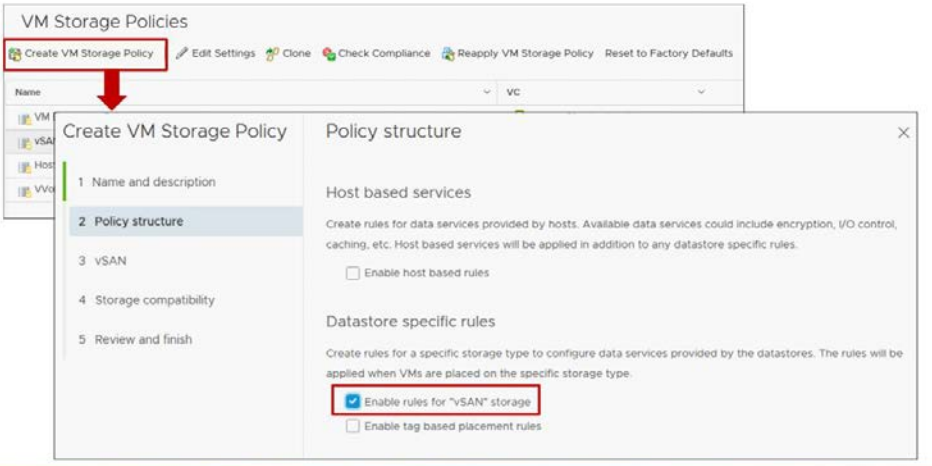


VMs that are deployed to vSAN datastores are assigned at least one VM storage policy. If a storage policy is not assigned to the VM that is provisioned, a default storage policy is applied to the VM from the datastore. If a custom policy is not applied to the vSAN datastore, the vSAN default storage policy is used.

https://t.me/learningnets

3-88 Creating vSAN Storage Policies

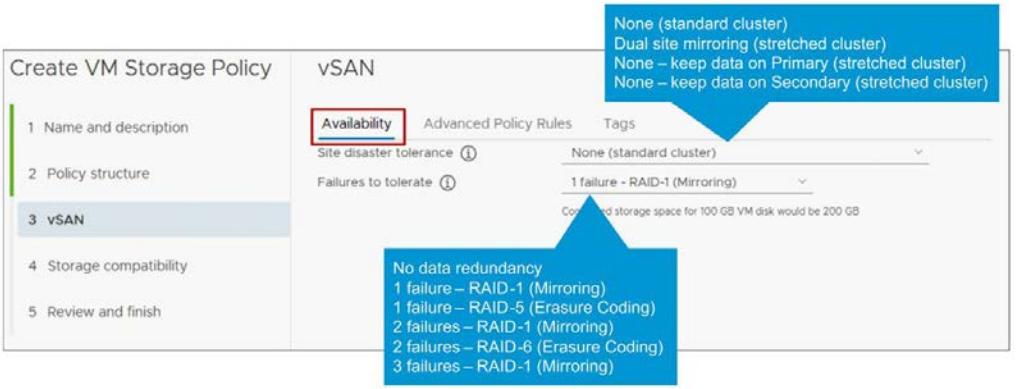
To create a vSAN storage policy, you must use rules for the vSAN storage type.



https://t.me/learningnets

3-89 vSAN Capabilities (1)

You define the rule set or set of capabilities. On the **Availability** tab, you can define site disaster tolerance and number of failures to tolerate.

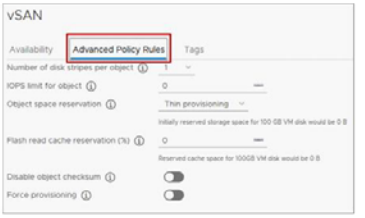


https://t.me/learningnets

3-90 vSAN Capabilities (2)

On the **Advanced Policy Rules** tab, you can configure one or more advanced vSAN options.

On the **Tags** tab, 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/7.0/vsan-70-administration-guide.pdf.

https://t.me/learningnets

3-91 Lab 7: 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 VM
4. Make the VM Compliant
5. Create an Invalid Storage Policy

https://t.me/learningnets

3-92 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Identify components in the vSAN architecture
- Create vSAN storage policies

https://t.me/learningnets

https://t.me/learningnets

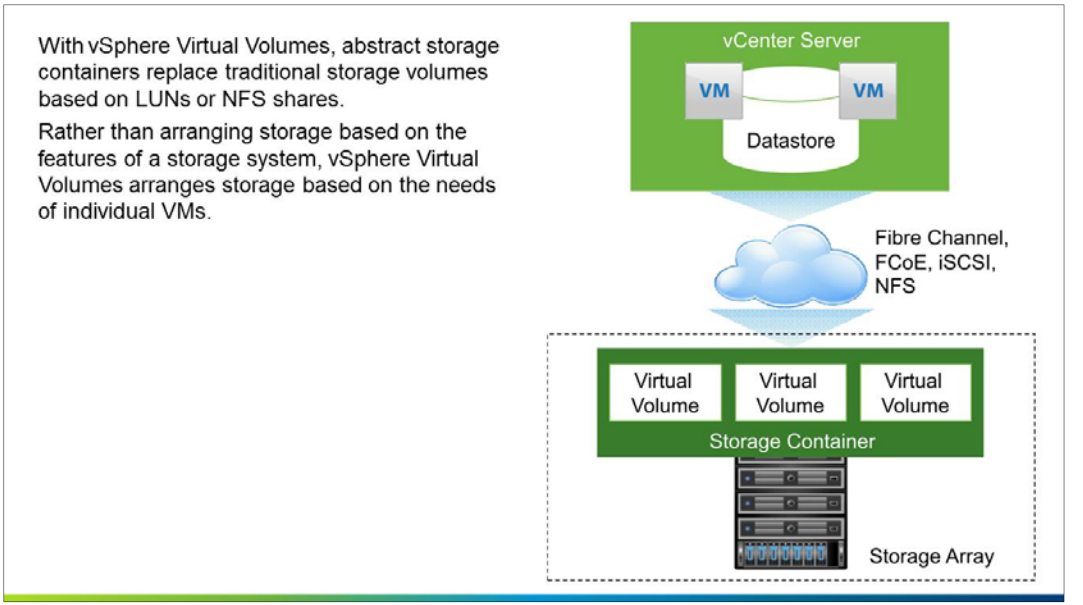
3-94 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Recognize components in the vSphere Virtual Volumes architecture
- Explain the rule set for creating vSphere Virtual Volumes storage policies

https://t.me/learningnets

3-95 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.

Traditionally, vSphere storage management uses a datastore-centric approach. With this approach, storage administrators and vSphere administrators discuss in advance the underlying storage requirements for VMs. The storage administrator 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 VM storage. Typically, the datastore is the lowest level at which data management occurs, from a storage perspective. However, a single datastore contains multiple VMs that might have different requirements.

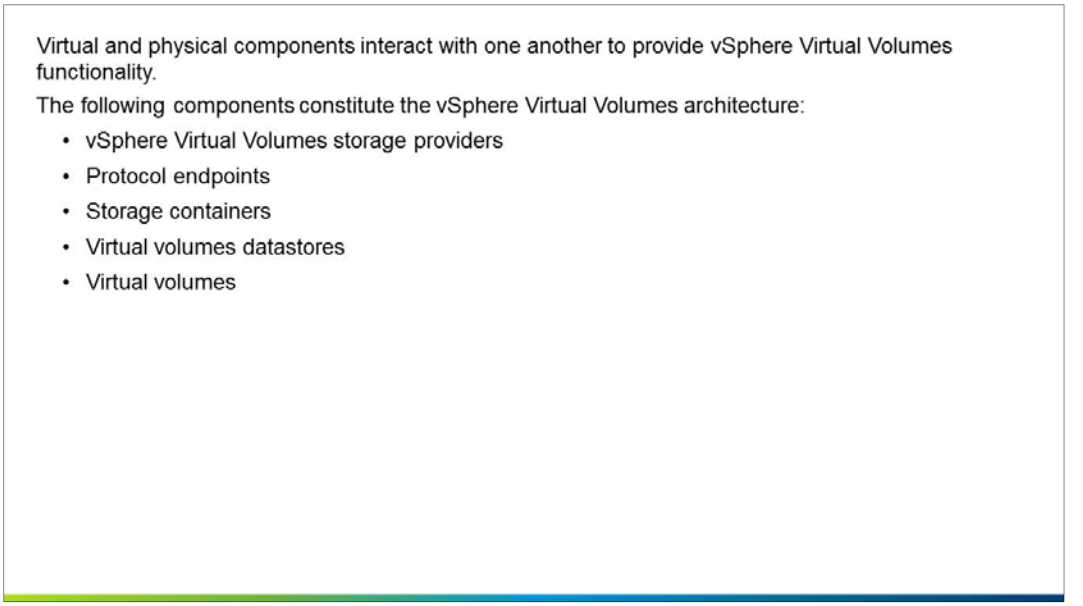
With the traditional approach, differentiation per VM is difficult.

With vSphere Virtual Volumes, you can differentiate VM services per application by offering a new approach to storage management.

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

https://t.me/learningnets

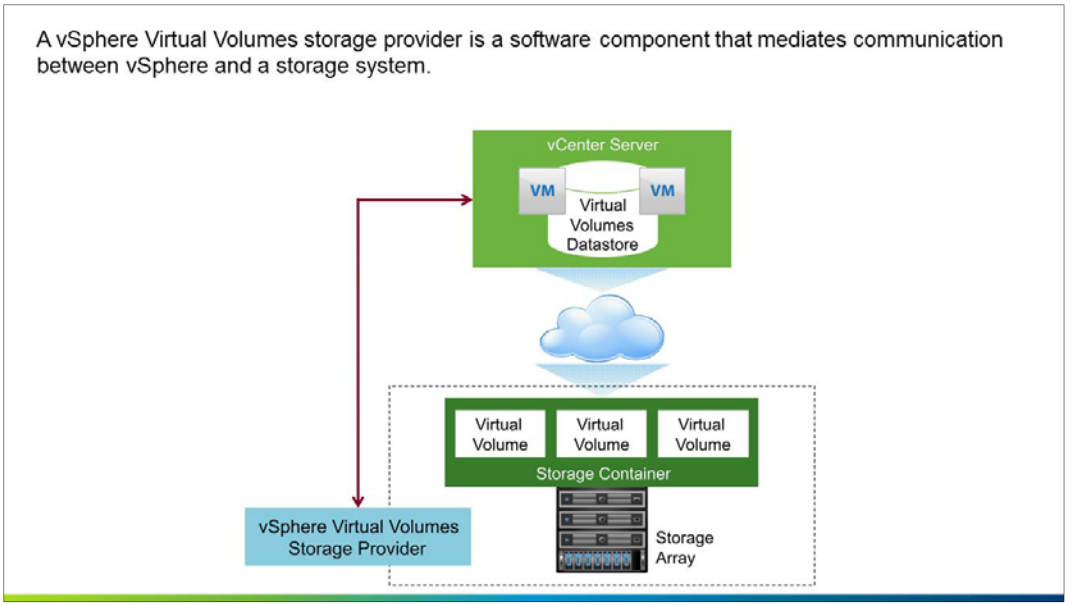
3-96 vSphere Virtual Volumes Architecture



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

https://t.me/learningnets

3-97 vSphere Virtual Volumes Storage Providers



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 VM storage requirements, which you define in a storage policy, to the storage layer. In this way, a virtual volume created in the storage layer meets the requirements outlined in the policy.

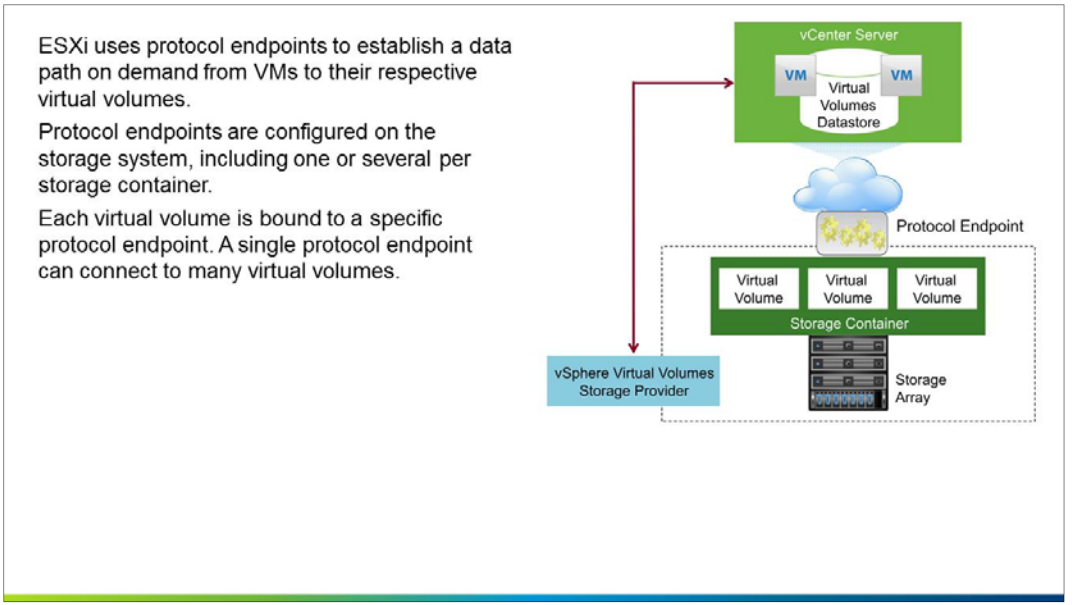
Typically, vendors supply storage providers that can integrate with vSphere and provide support to virtual volumes. Every storage provider must be certified by VMware and properly deployed. For information about deploying the vSphere Virtual Volumes storage provider, contact your storage vendor.

https://t.me/learningnets

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

3-98 Protocol Endpoints



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 VM 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 few protocol endpoints.

A storage administrator configures protocol endpoints. Protocol endpoints are part of physical storage and are exported, 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 the vSphere Client. Protocol endpoints can also be discovered during a storage rescan.

https://t.me/learningnets

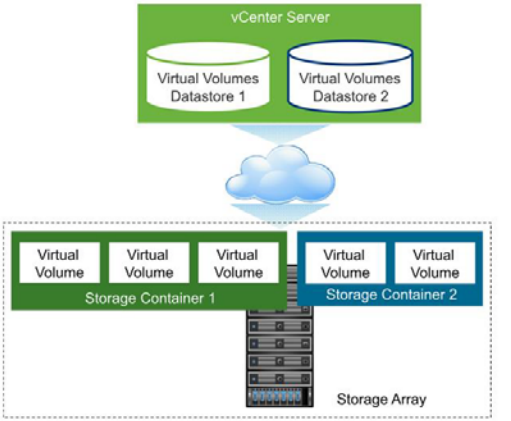
3-99 Storage Containers and vSphere Virtual Volumes Datastores

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 the vendor-specific implementation, but each storage system requires at least one container.

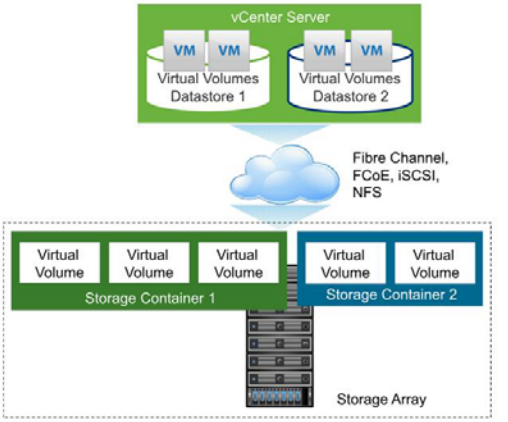
After vCenter Server discovers storage containers exported by storage systems, you must map a storage container to a virtual volumes datastore. The virtual volumes datastore that you create corresponds directly to the specific storage container and becomes the container's representation in vCenter Server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

https://t.me/learningnets

3-100 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 VM files, virtual disks, and their derivatives.

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

https://t.me/learningnets

3-101 Mapping VM Files to Virtual Volumes

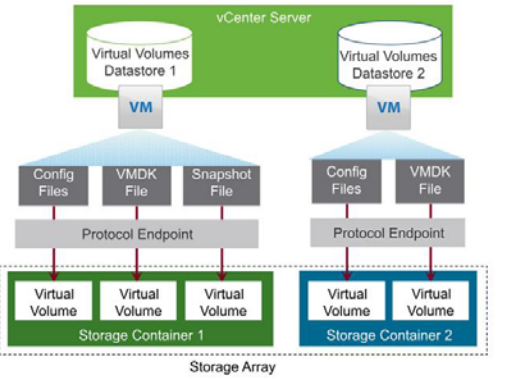
The storage system creates virtual volumes for the core elements that make up the VM:

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. vSphere can then offload intensive storage operations, such as taking snapshots, cloning, and replication, to the storage system.



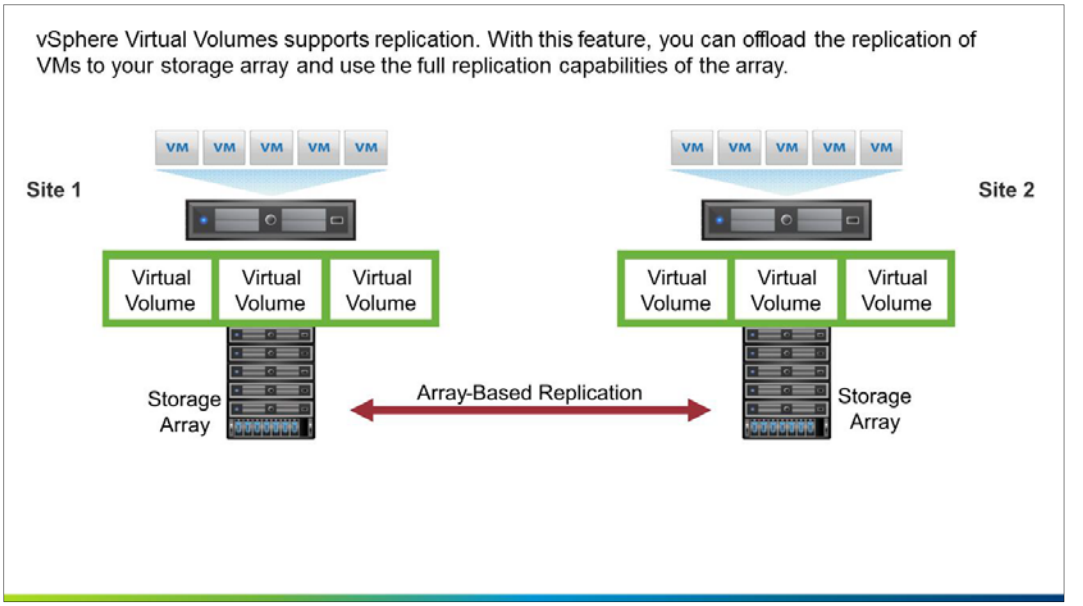
A data virtual volume corresponds directly to each virtual disk (.vmdk) file. Like virtual disk files on traditional datastores, virtual volumes are presented to VMs as SCSI disks.

A configuration virtual volume, or a home directory, represents a small directory that contains metadata files for a VM. 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 VM components and virtual disk derivatives, such as clones, snapshots, and replicas. These virtual volumes include a swap virtual volume to hold VM swap files and a virtual memory volume to hold the contents of VM memory for a snapshot.

https://t.me/learningnets

3-102 vSphere Virtual Volumes Replication



vSphere Virtual Volumes replication has the following properties:

- Available with vSphere API for Storage Awareness version 3.0
- 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 vary with different storage vendors. The following requirements generally apply to all vendors:

- The storage arrays that you use to implement replication must be compatible with vSphere Virtual Volumes.
- The arrays must integrate with the version of the storage (VASA) provider that is compatible with vSphere Virtual Volumes replication.

https://t.me/learningnets

- 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 by the storage system.

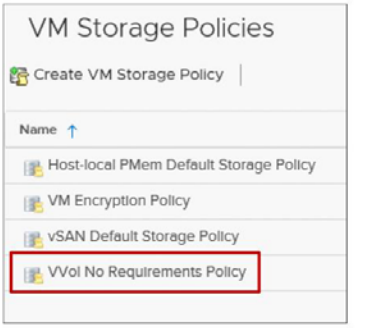
- When applicable, replication groups and fault domains for virtual volumes must be preconfigured by the storage system.

https://t.me/learningnets

3-103 Virtual Volume Storage Policies

A VM that runs on a virtual volumes datastore requires a VM storage policy.
If you do not create a VM storage policy compatible with the virtual volumes datastore, the system uses the No Requirements default policy:

- This generic virtual volumes policy contains no rules or storage specifications.
- Storage arrays can determine the optimum placement for the VM objects.



https://t.me/learningnets

3-104 Virtual Volume Storage Policy: Examples

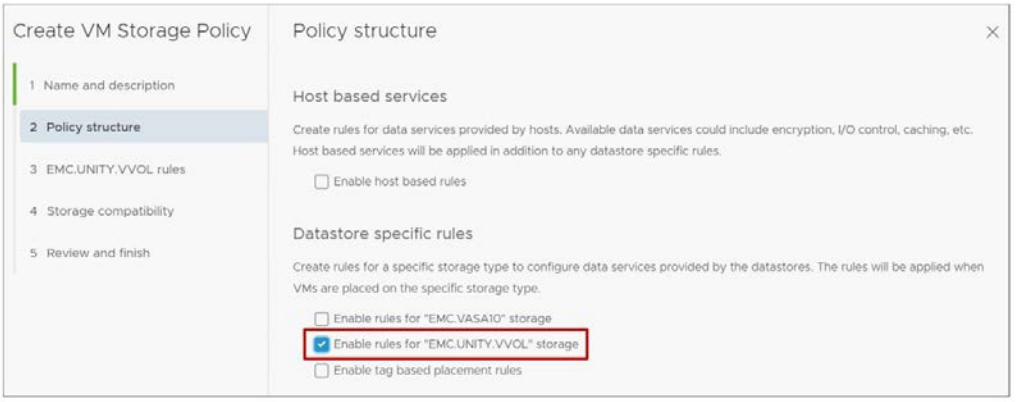
By using different virtual volumes for different VM components, you can apply storage policies at detailed levels, for example:

- 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.
- 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 business-critical VMs.

https://t.me/learningnets

3-105 Creating Virtual Volumes Storage Policies (1)

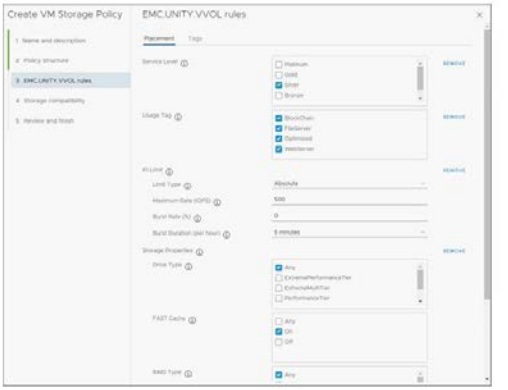
A virtual volumes storage policy can include host-based service rules and datastore-specific rules.



https://t.me/learningnets

3-106 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

3-107 Activity: Displaying Storage Capabilities (1)

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

- ❒ That tags are created for the virtual volumes storage array
- ❒ That virtual volumes storage policy components are preconfigured
- ❒ That the virtual volumes storage provider is registered with vCenter Server
- ❒ That the VVol No Requirements storage policy is not deleted

https://t.me/learningnets

3-108 Activity: Displaying Storage Capabilities (2)

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

- ❒ That tags are created for the virtual volumes storage array
- ❒ That the virtual volumes storage policy components are preconfigured
- ✓ That the virtual volumes storage provider is registered with vCenter Server
- ❒ That the VVol No Requirements storage policy is not deleted

https://t.me/learningnets

3-109 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Recognize components in the vSphere Virtual Volumes architecture
- Explain the rule set for creating vSphere Virtual Volumes storage policies

https://t.me/learningnets

https://t.me/learningnets

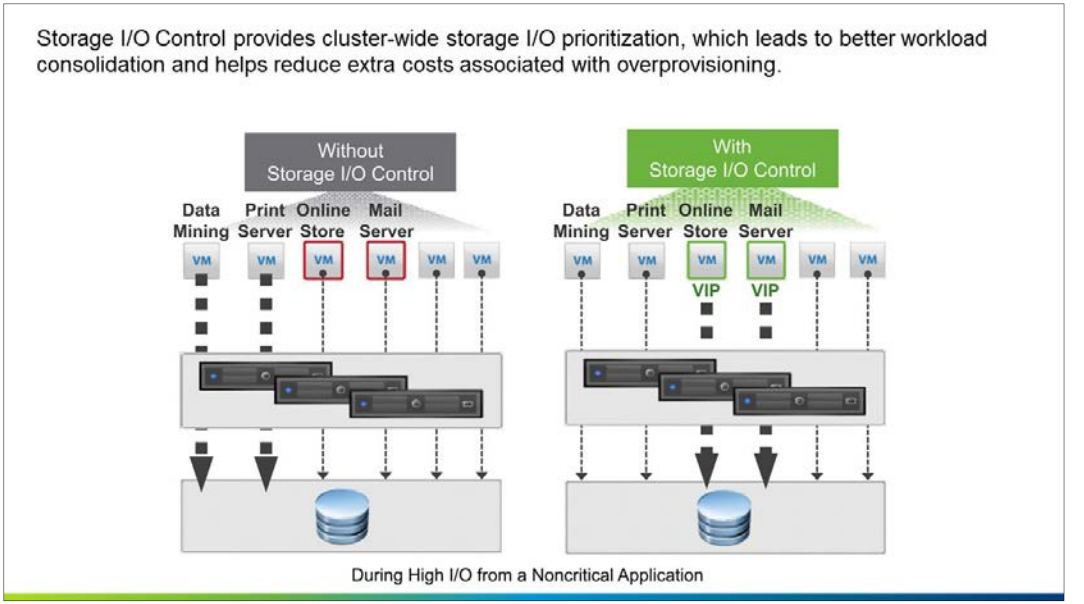
3-111 Learner Objectives

After completing 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

3-112 About Storage I/O Control



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

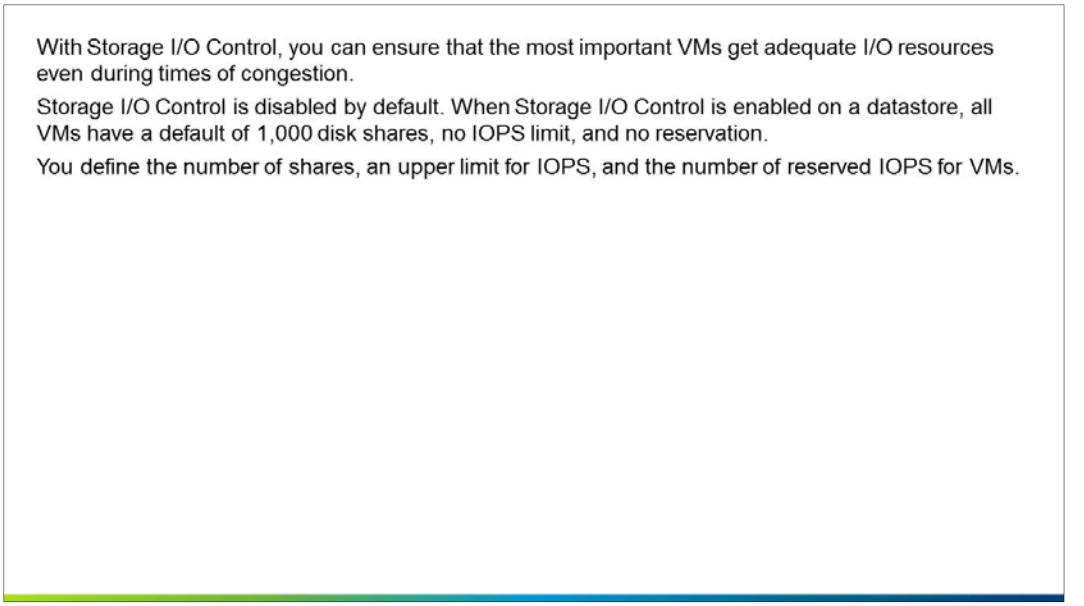
Storage I/O Control extends the constructs of shares, limits, and reservations to handle storage I/O resources. Storage I/O Control is a proportional-share, I/O operations per second (IOPS) scheduler that, under contention, throttles IOPS. You can control the amount of storage I/O that is allocated to VMs during periods of I/O congestion. Controlling storage I/O ensures that more important VMs get preference over less important VMs 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, Storage I/O Control gathers latency statistics for an ESXi host, sent to vCenter Server, and stored in the vCenter Server database. With these statistics, vSphere Storage DRS can decide whether a VM should be migrated to another datastore.

For more information, see *Managing Storage I/O Resources* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.resmgmt.doc/GUID-7686FEC3-1FAC-4DA7-B698-B808C44E5E96.html.

https://t.me/learningnets

3-113 Enabling Storage I/O Control (1)

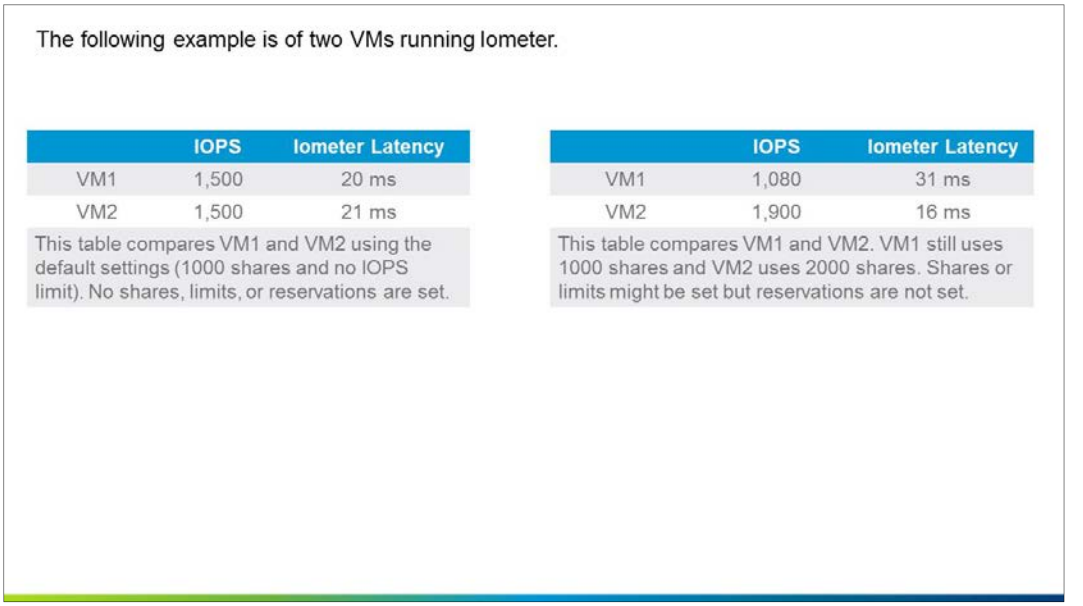


Storage I/O Control provides quality-of-service capabilities for storage I/O in the form of I/O shares, limits, and reservations that are enforced across all VMs accessing a datastore, regardless of which host they are running on.

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 VM that accesses that datastore is allocated I/O resources in proportion to its shares.

https://t.me/learningnets

3-114 Enabling Storage I/O Control (2)



When you allocate storage I/O resources, you can limit the IOPS that are allowed for a VM. By default, the number of IOPS that are allowed for a VM is unlimited. If the limit that you want to set for a VM 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 VM. 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.

VMs VM1 and VM2 are running an I/O load generator called Iometer.

Each VM 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 VM achieves are identical. However, with Storage I/O Control enabled, the IOPS achieved by the VM with more shares (VM2) is greater than the IOPS of VM1. The example assumes that each VM is running enough load to cause a bottleneck on the datastore.

https://t.me/learningnets

3-115 Monitoring Latency

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.

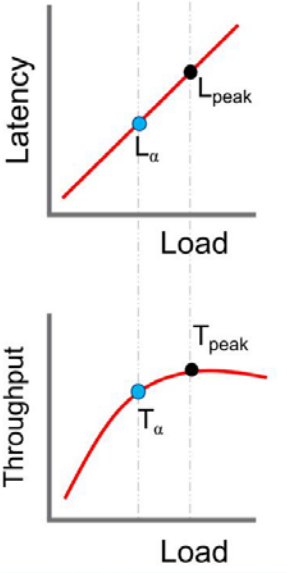
https://t.me/learningnets

3-116 Automatic Threshold Detection

With the injector-based model, Storage I/O Control determines the peak throughput of the storage 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.



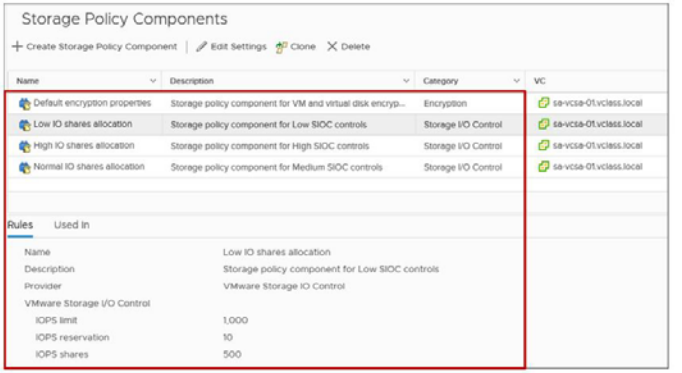
The I/O injector model determines the peak throughput of a datastore. The resulting peak throughput measurement can be used to determine the peak latency of a datastore.

https://t.me/learningnets

3-117 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 define Storage I/O Control capabilities.

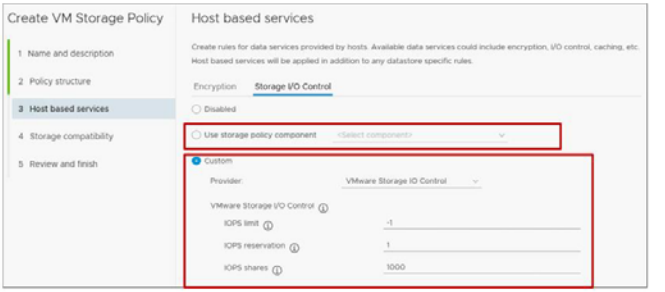


https://t.me/learningnets

3-118 Configuring Storage I/O Control

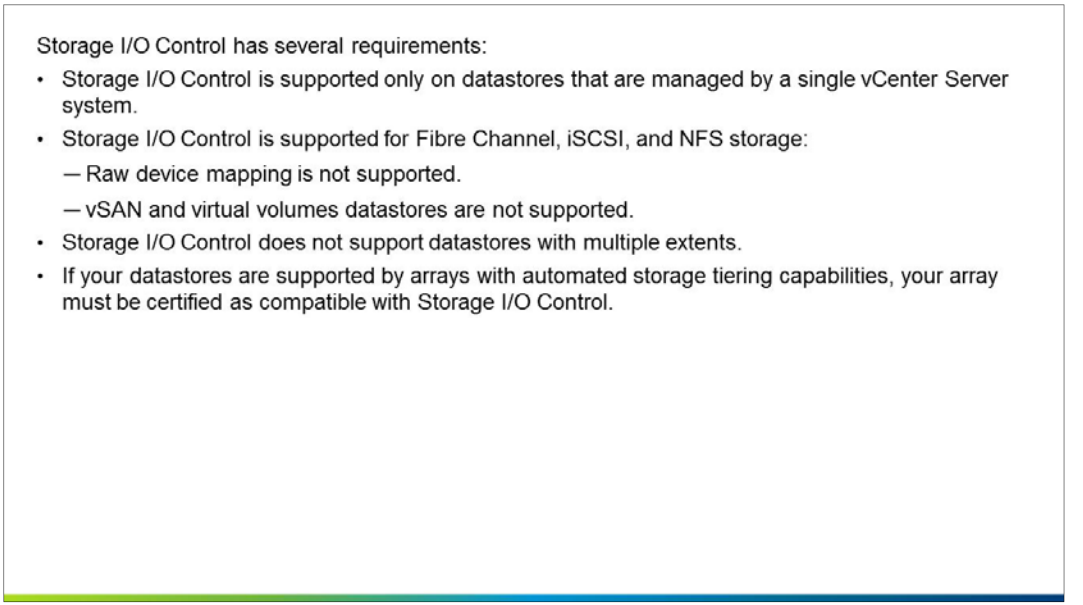
You can use a storage policy component for Storage I/O Control.

Or, you can create a custom storage policy component, which defines the IOPS limit, IOPS reservation, and IOPS shares. You use the component in the storage policy and assign the policy to a VM.



https://t.me/learningnets

3-119 Storage I/O Control Requirements



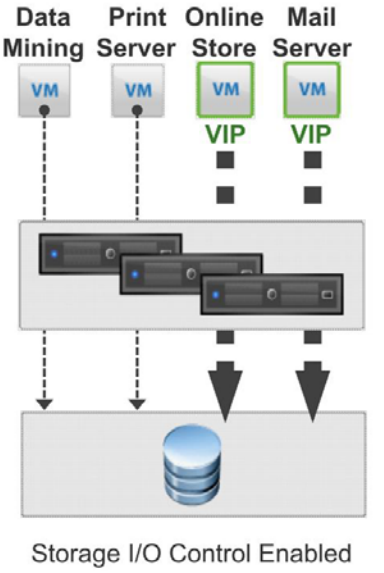
Automated storage tiering is the ability of an array (or a group of arrays) to migrate LUNs, virtual volumes, or parts of LUNs or virtual volumes to different types of storage media (SSD, Fibre Channel, SAS, SATA, and so on) based on user-set policies and current I/O patterns. No special certification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

3-120 Activity: Storage I/O Control (1)

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

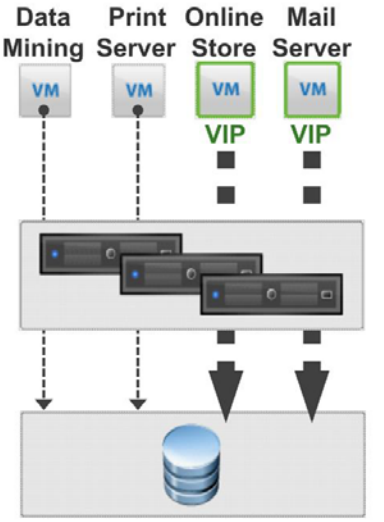


https://t.me/learningnets

3-121 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 2,000.
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 1,000 disk shares.



Storage I/O Control Enabled

https://t.me/learningnets

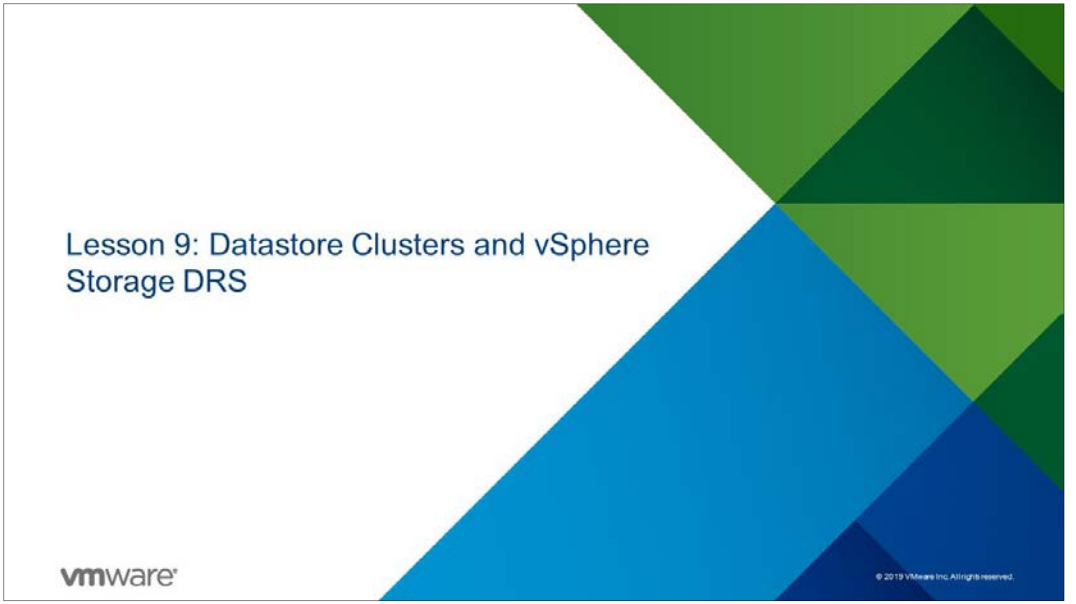
3-122 Review of Learner Objectives

After completing 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

3-123 Lesson 9: Datastore Clusters and vSphere Storage DRS



https://t.me/learningnets

3-124 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

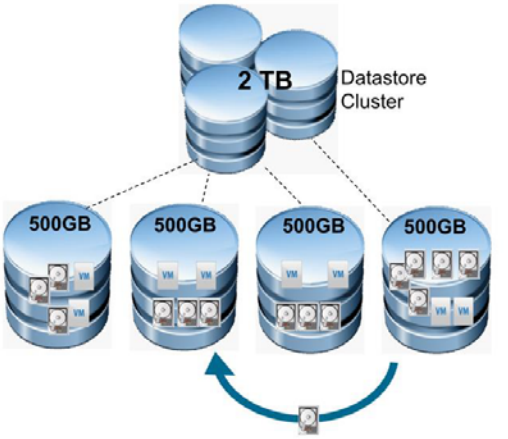
- Create a datastore cluster
- Configure vSphere Storage DRS
- Explain how Storage I/O Control and vSphere Storage DRS complement each other

https://t.me/learningnets

3-125 About Datastore Clusters

A datastore cluster is a collection of datastores that are grouped together but that work as separate entities.

When you enable a datastore cluster with vSphere Storage DRS, the datastores on the cluster work 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 that work as a single unit.

https://t.me/learningnets

3-126 Datastore Cluster Requirements

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

- 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 that is enabled for vSphere Storage DRS.

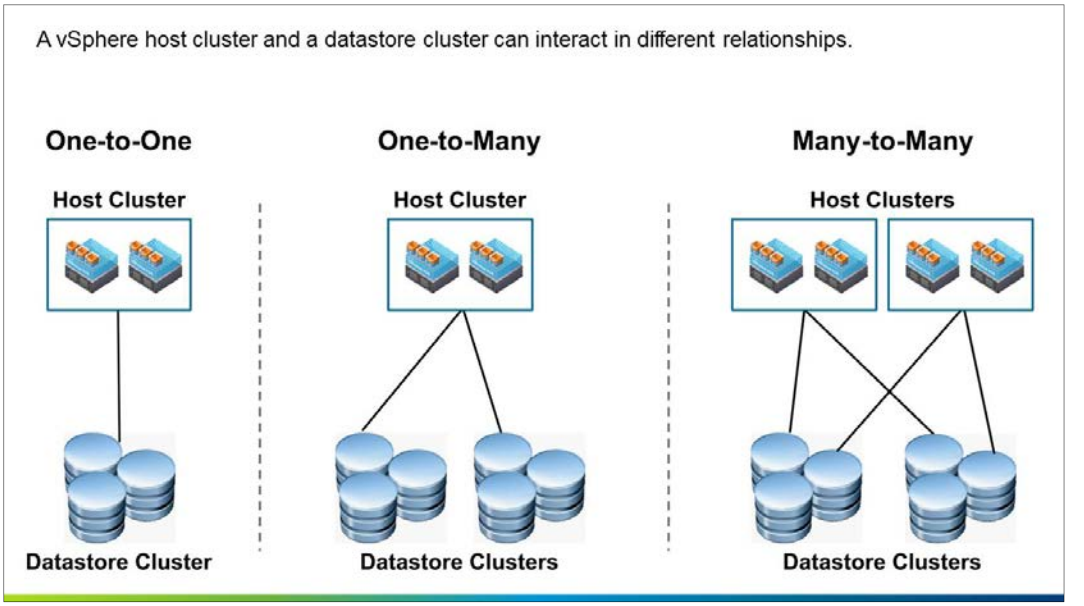
    This exception applies only to storage arrays that do not support vSphere API for Storage Awareness.
- Datastores shared across multiple data centers cannot be included in a datastore cluster.
- Datastores with hardware acceleration enabled should not be included in the same datastore cluster as datastores 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 VMs between NFS and VMFS datastores.

https://t.me/learningnets

3-127 Host Cluster to Datastore Cluster Relationships



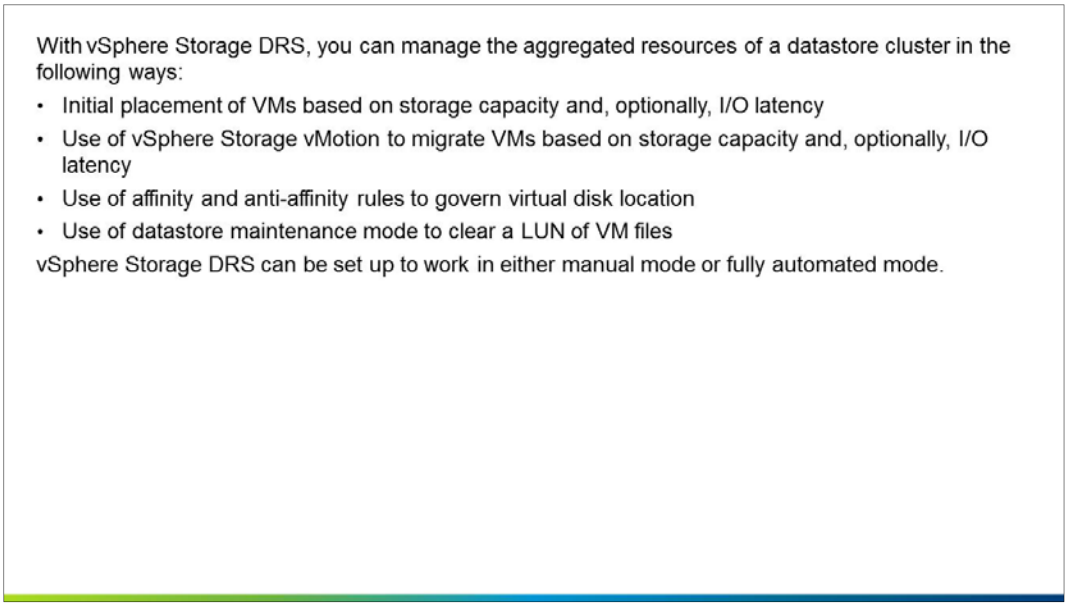
Host clusters and datastore clusters can coexist in the virtual infrastructure. A host cluster is a vSphere DRS or a vSphere HA cluster.

Load balancing by vSphere DRS and vSphere Storage DRS can occur at the same time. vSphere DRS balances VMs 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 VMs 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

3-128 vSphere Storage DRS Overview



vSphere Storage DRS manages the placement of VMs 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 VMs 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.

https://t.me/learningnets

3-129 Initial Disk Placement

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

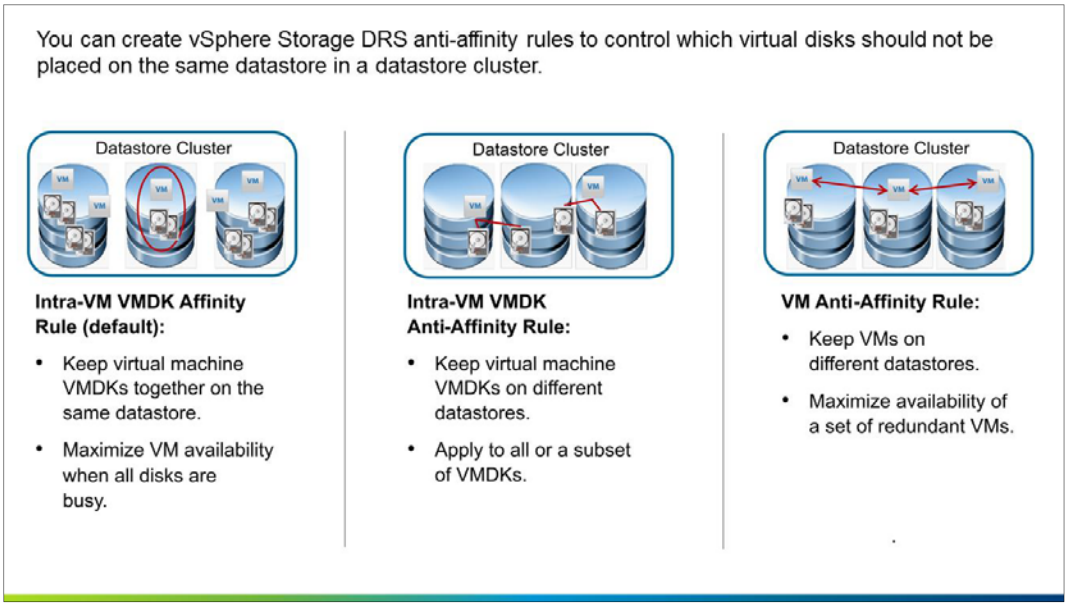
When VMs are created, cloned, or migrated, you select a datastore cluster rather than a single datastore. vSphere Storage DRS selects a member datastore based on capacity and, optionally, IOPS load.

By default, a VM'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 VM is created, cloned, or migrated, and the user selects a datastore cluster, vSphere Storage DRS selects 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

3-130 vSphere Storage DRS Affinity Rules



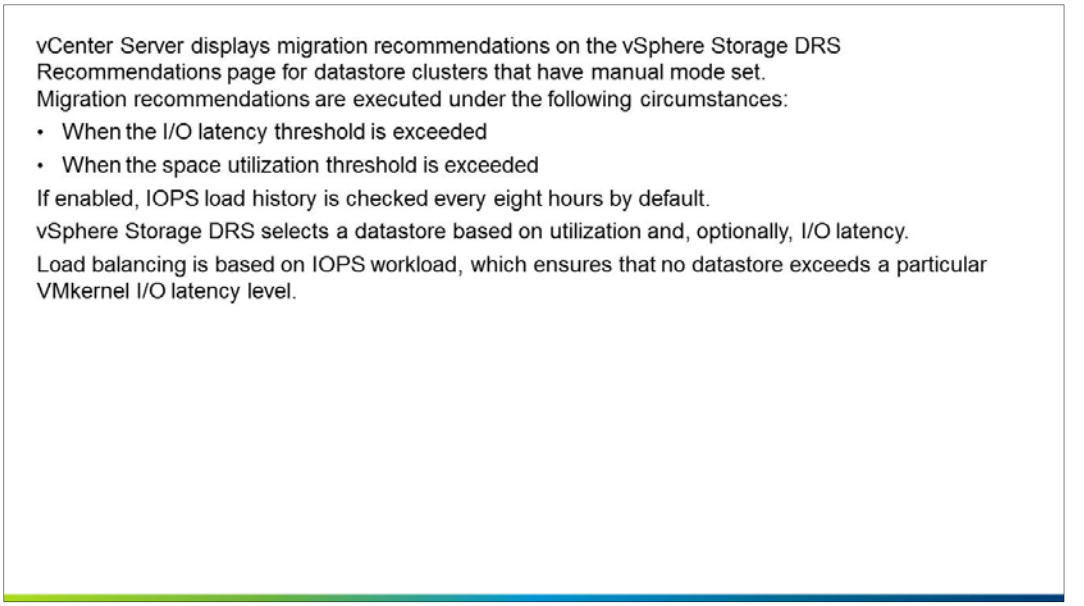
By default, all of a VM's disks can be on the same datastore.

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 VM’s virtual disks on separate datastores.

VM anti-affinity rules keep VMs on separate datastores. This rule is useful when redundant VMs must always be available.

https://t.me/learningnets

3-131 Storage Migration Recommendations



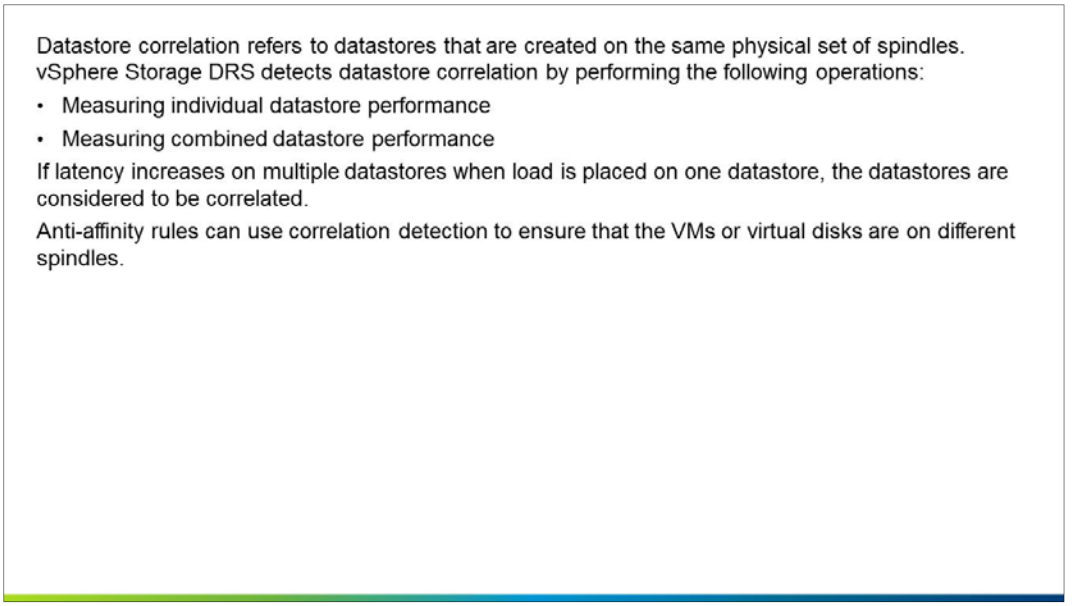
Storage migration recommendations work as follows:

- 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.

To achieve optimal balance of VMs across datastores in the datastore cluster, vSphere Storage DRS considers moving powered-off VMs to other datastores.

https://t.me/learningnets

3-132 Datastore Correlation Detector



The purpose of datastore correlation is to help vSphere Storage DRS in determining where to move a VM. For example, you gain little advantage by moving a VM from one datastore to another if both datastores are supported 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.

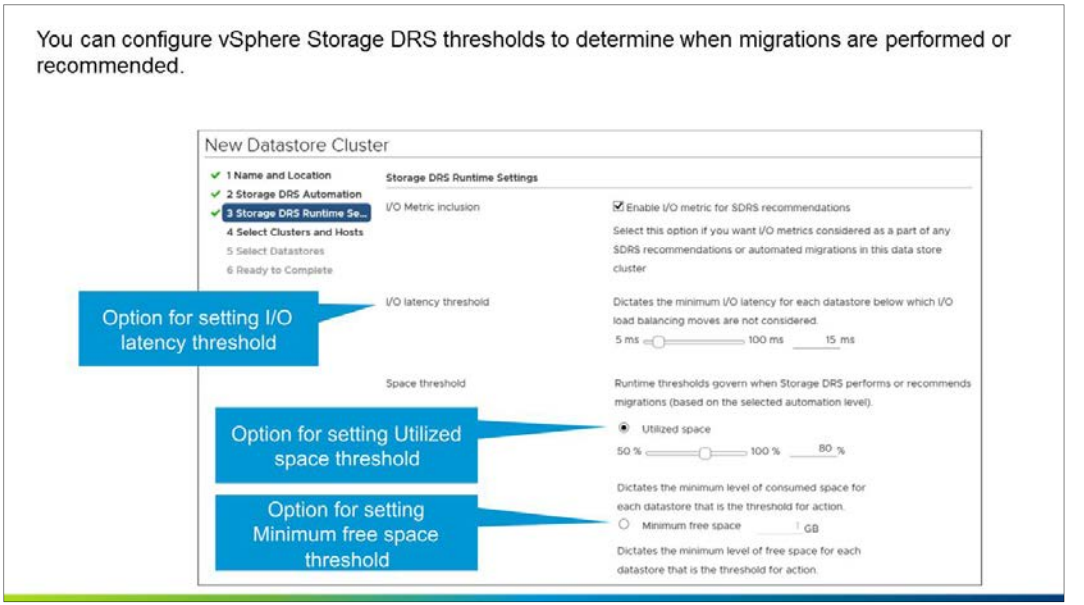
The datastore correlation detector can also be used for anti-affinity rules, ensuring that VMs and virtual disks are not only kept on separate datastores but also kept on different spindles on the back end.

Correlation is determined by a long-running background process.

Datastore correlation is enabled by default.

https://t.me/learningnets

3-133 Configuring vSphere Storage DRS Runtime Settings



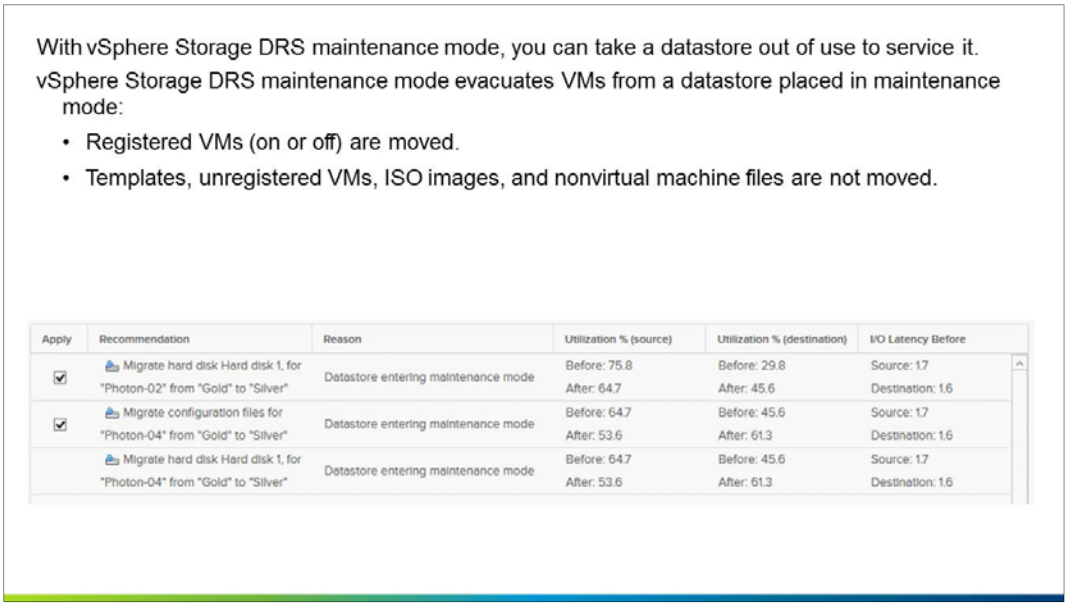
- When you create datastore cluster, you select or deselect the **Enable I/O metric for SDRS recommendations** check box to enable or disable IOPS metrics use in recommendations or automated migrations.
- When you enable I/O load balancing, Storage I/O Control is enabled for all the datastores in the datastore cluster if it is not already enabled for the cluster.

https://t.me/learningnets

- 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

3-134 vSphere Storage DRS Maintenance Mode



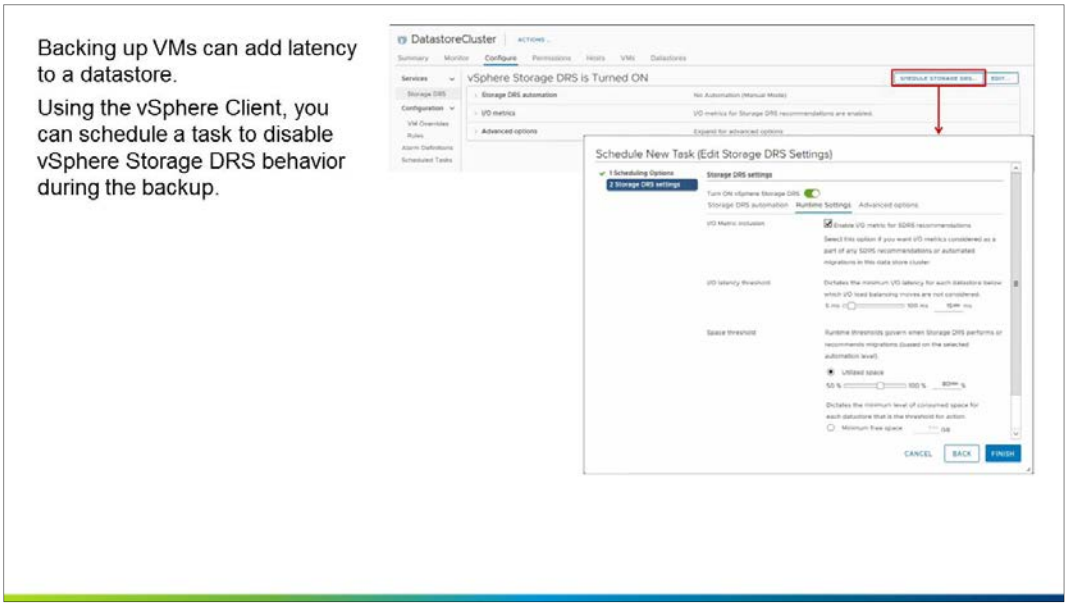
With vSphere Storage DRS, you can place a datastore in maintenance mode. vSphere Storage DRS maintenance mode is available to datastores in a datastore cluster that is enabled for vSphere Storage DRS. Standalone datastores cannot be placed in maintenance mode.

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, VMs are automatically migrated to other datastores.

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

https://t.me/learningnets



You can configure scheduled tasks to change vSphere Storage DRS behavior.

You can use scheduled tasks to change the vSphere Storage DRS configuration of the datastore cluster to match enterprise activity.

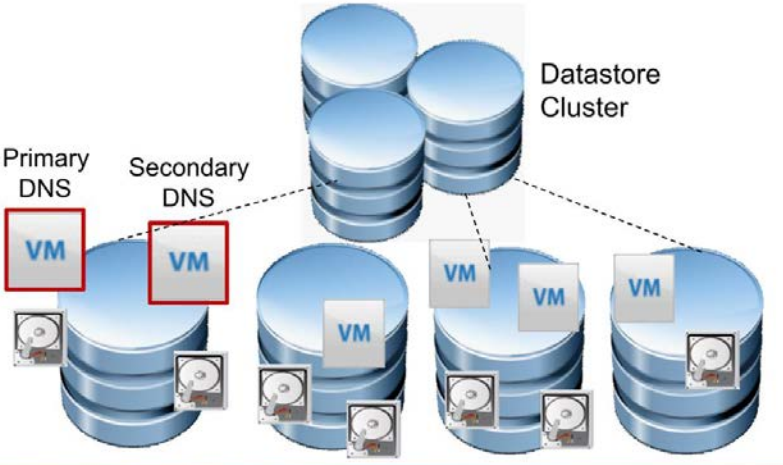
For example, if the datastore cluster is configured to perform migrations based on I/O latency, you might disable the use of I/O metrics by vSphere Storage DRS during the backup window.

You can reenable I/O metrics use after the backup window closes.

https://t.me/learningnets

3-136 Activity: vSphere Storage DRS (1)

Consider the datastore cluster example. How do you configure vSphere Storage DRS to eliminate the single point of failure for the primary DNS and secondary DNS servers?

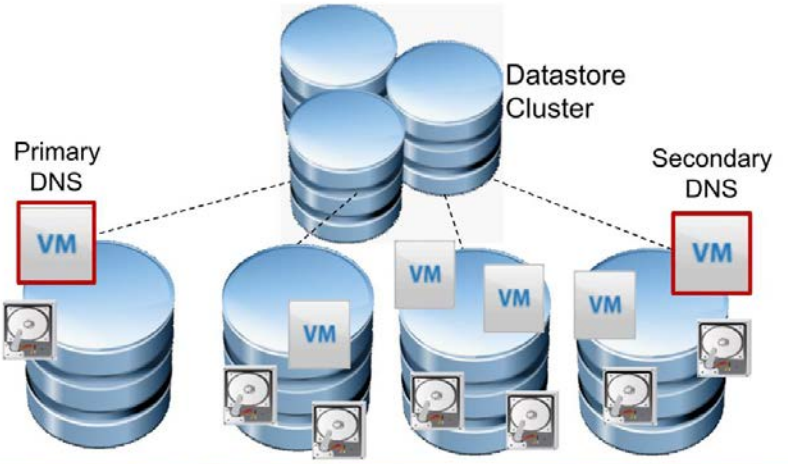


https://t.me/learningnets

3-137 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

3-138 Review of Learner Objectives

After completing 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

3-139 Key Points

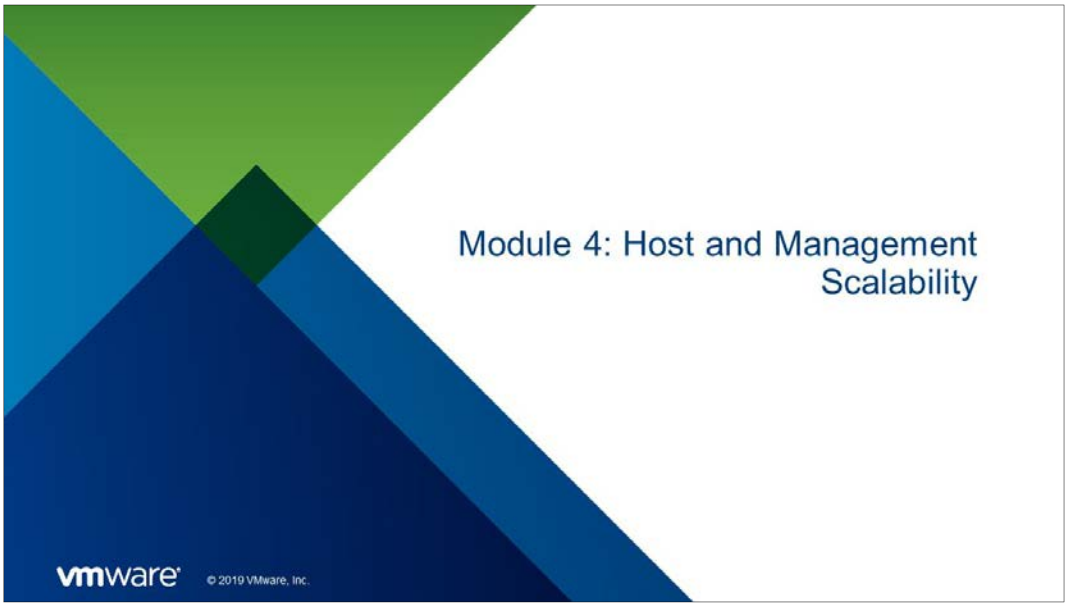
- With a VMFS6 datastore, the ESXi host automatically sends the SCSI `unmap` command to the storage array, and the unmap rate is configurable.
- With vSphere API for Storage Awareness, storage vendors can provide information about the capabilities of their storage arrays to vCenter Server.
- With vSphere APIs for I/O Filtering, VMware and third-party vendors can 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 provides cluster-wide storage I/O prioritization.
- A datastore cluster that is 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

Module 4

Host and Management Scalability



https://t.me/learningnets

4-2 Importance

To scale your vSphere environment effectively, you must understand the vSphere features and functions that help you manage hosts in your environment.

https://t.me/learningnets

4-3 Module Lessons

1. VMware Certificate Management
2. vCenter Server Identity Federation
3. vSphere Trust Authority
4. Host Profiles
5. Content Libraries
6. Resource Pools

https://t.me/learningnets

https://t.me/learningnets

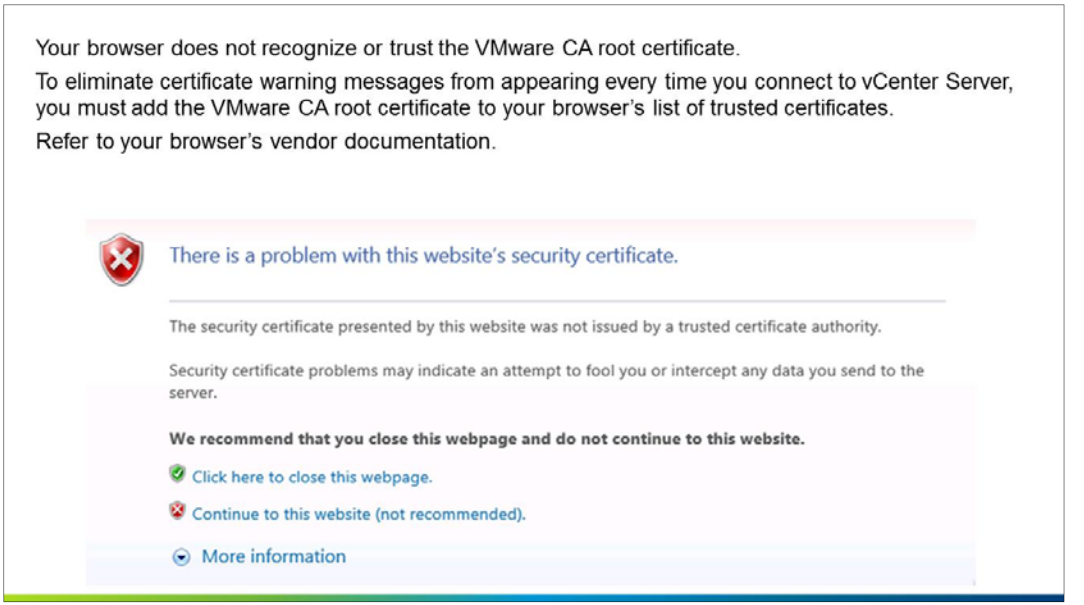
4-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Use VMware CA 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 the vSphere Client to manage vSphere certificates

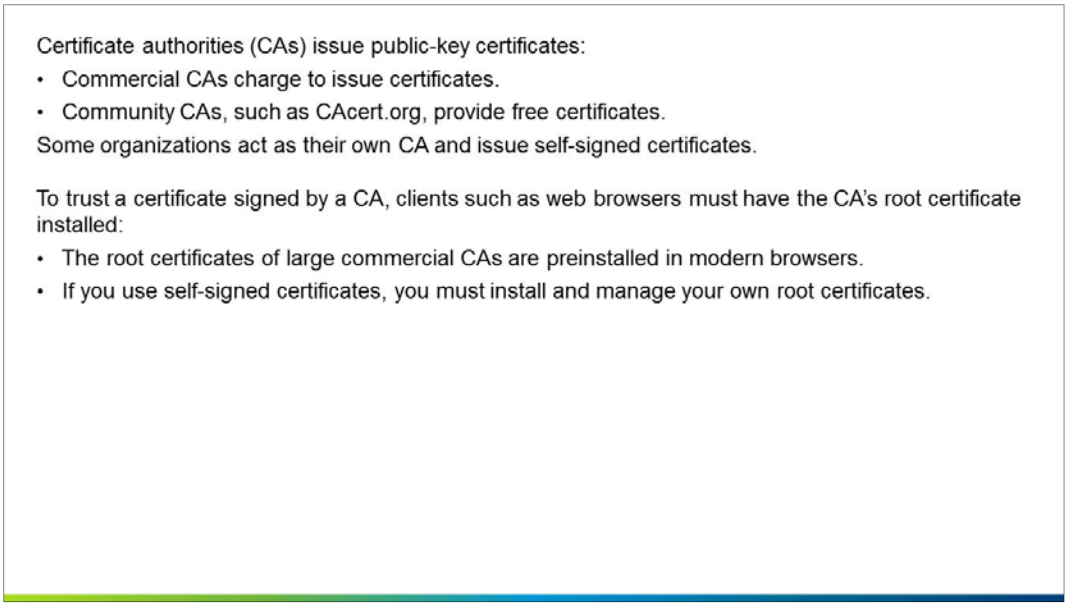
https://t.me/learningnets

4-6 Certificate Use Case



https://t.me/learningnets

4-7 About Certificate Authorities



Public key cryptographic systems use pairs of keys. The data is encrypted using a private key, which is known only to the owner. The data can be decrypted only by using a public key, which might be widely disseminated.

CAs are important 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 send its public key in plain text because such text is easily corruptible by attackers who can inject themselves into 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.

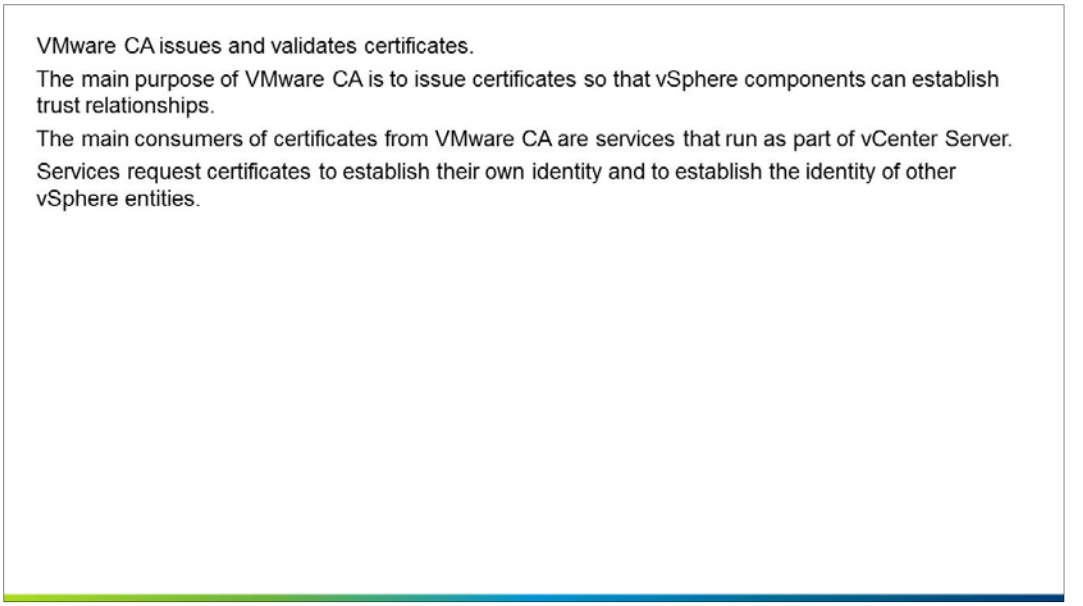
The client trusts the CA because the client already has the CA's public key, which is either preinstalled or manually installed by an administrator. For example, browsers such as Safari, Firefox, Internet Explorer, and Chrome have public keys of the most common CAs preinstalled.

https://t.me/learningnets

SSL or TLS use is not limited to websites and browsers, although these uses are common and well known. 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

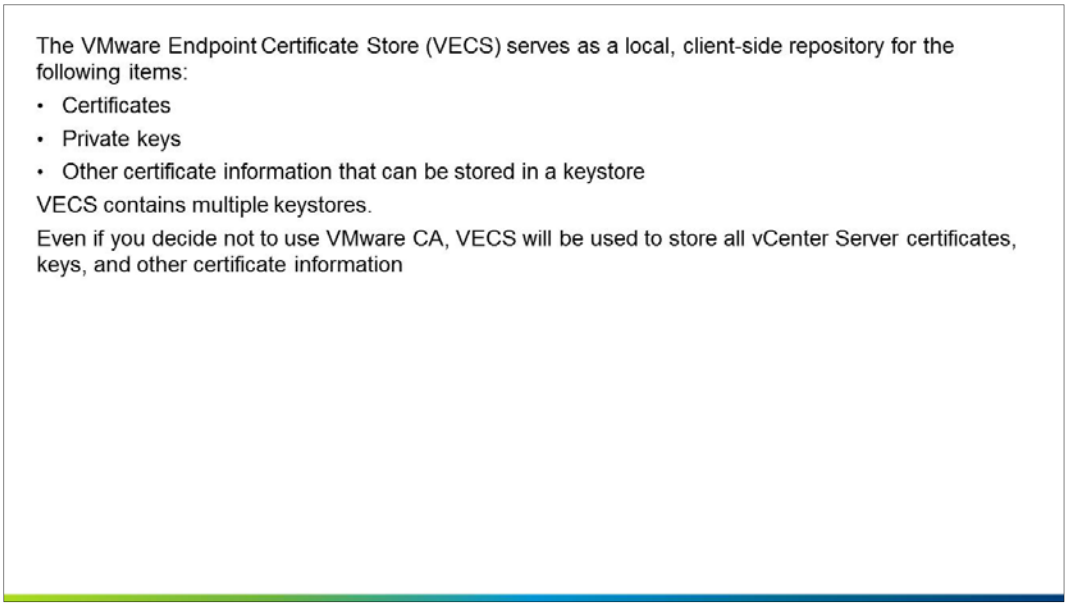
4-8 About VMware CA



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

4-9 About the VMware Endpoint Certificate Store



ESXi certificates are stored locally on each host, not in the VMware Endpoint Certificate Store (VECS).

VECS runs as part of the VMware Authentication Framework Daemon on vCenter Server. It holds the keystores that contain the certificates and keys. With support for multiple keystores, the key manager can manage multiple hosts that require different access credentials. Different certificates for different host names can be served from the same IP address. Therefore, multiple secure (HTTPS) websites (or any other service over TLS) can be served by the same IP address without requiring all those sites to use the same certificate.

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 the vSphere Client.

https://t.me/learningnets

4-10 vSphere Certificate Types

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 the vSphere Client, also use 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 from that point on.

VMware products use standard X.509 version 3 (X.509v3) certificates to encrypt session information that is sent over SSL between components.

Solution user certificates continue to be used, but they might not be displayed in the Certificate Manager client.

https://t.me/learningnets

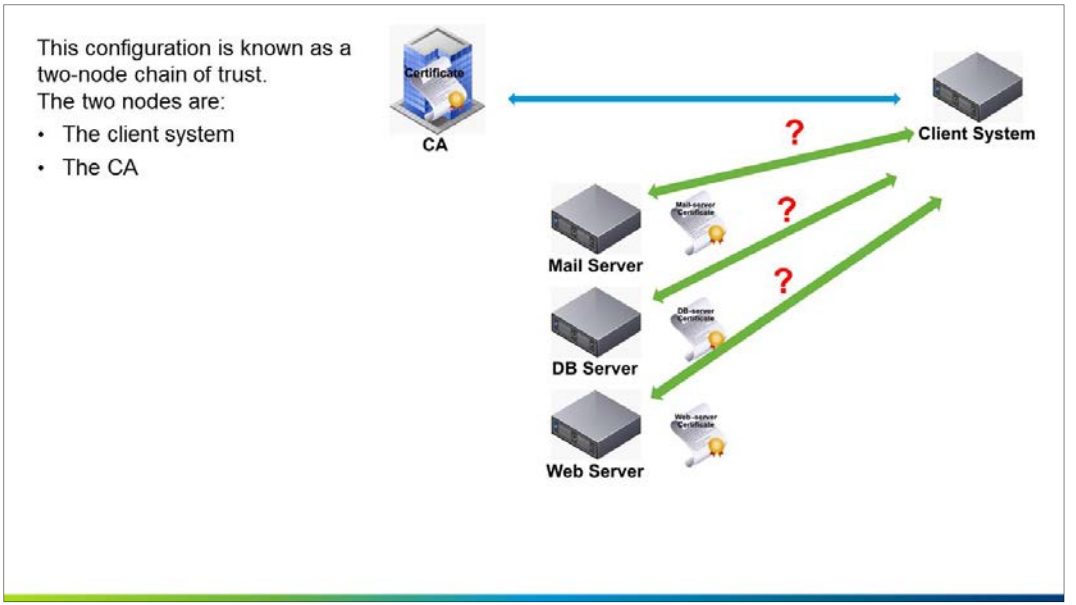
4-11 Chain of Trust (1)

If you trust the CA, you implicitly trust all the certificates issued by that CA.

- To trust a certificate, you must trust some part of the chain of trust. One of the following conditions 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 used in the network.
  - Every time a new service is brought online, all clients must individually install and trust every self-signed certificate in the network.
- An in-house or commercial CA eliminates the requirement that each client system installs every self-signed certificate as long as:
  - All client systems trust the CA.
  - All certificates come from that trusted CA.

https://t.me/learningnets

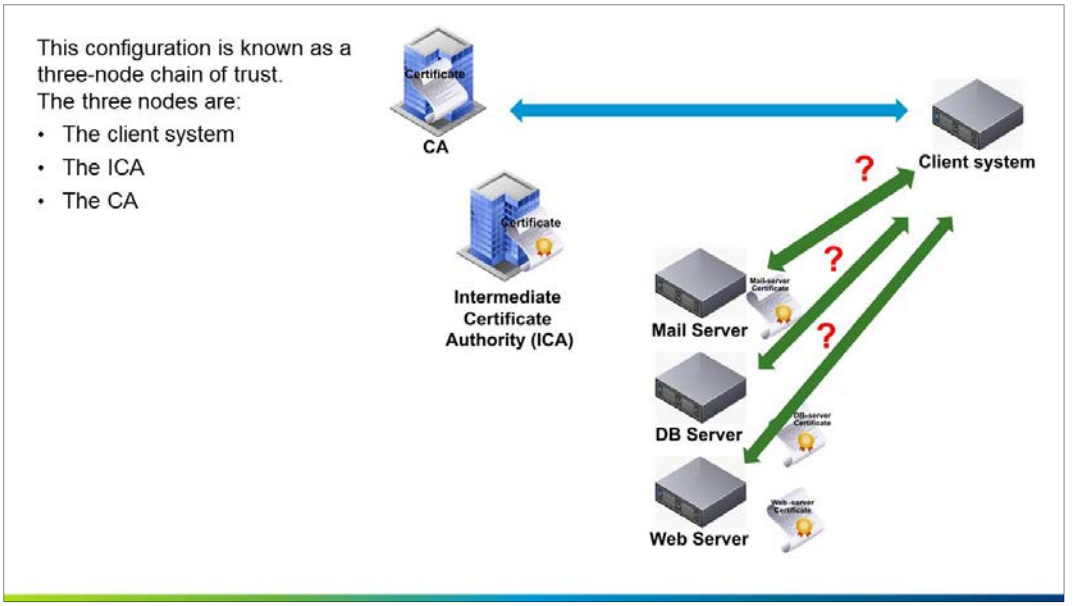
4-12 Chain of Trust (2)



The client trusts the server because the Certificate Authority (CA) issued the server its certificate and the client trusts the CA.

https://t.me/learningnets

4-13 Chain of Trust (3)



The client trusts the server because the Intermediate Certificate Authority (ICA) issued the server its certificate, the CA issued the ICA its certificate, and the client trusts the CA.

https://t.me/learningnets

4-14 Activity: ESXi Certificates (1)

Which of the following is a unique aspect of an ESXi certificate generated by VMware CA?

- ❒ It is stored locally in the MACHINE_SSL_CERT keystore in VECS.
- ❒ It is stored locally in the vCenter Server daemon (vpxd).
- ❒ It is stored locally on the relevant ESXi host.
- ❒ It is stored locally on every vSphere node.

https://t.me/learningnets

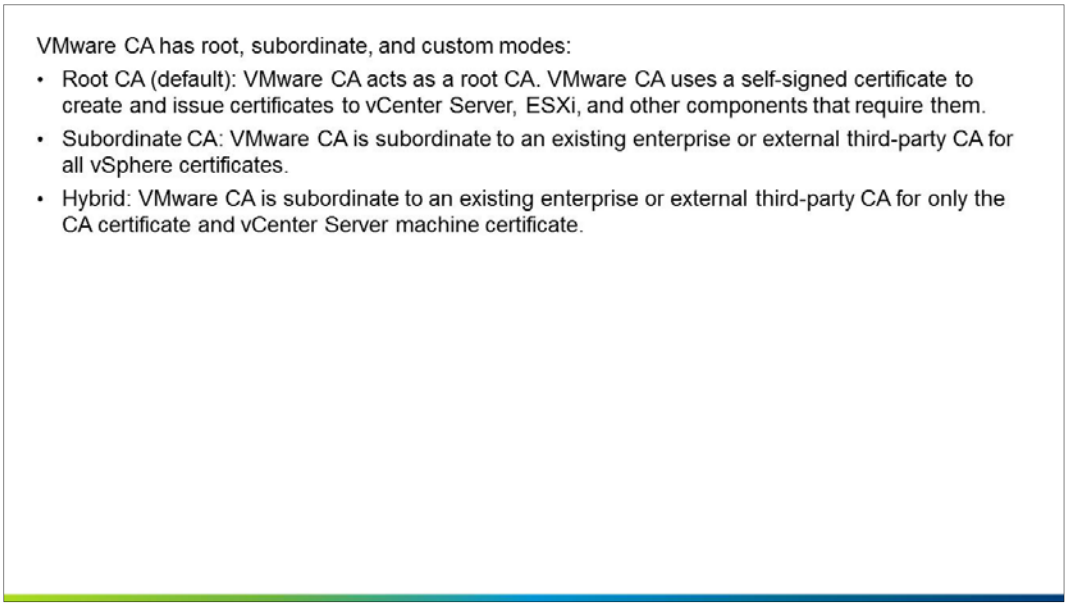
4-15 Activity: ESXi Certificates (2)

Which of the following is a unique aspect of an ESXi certificate generated by VMware CA?

- ❒ It is stored locally in the MACHINE_SSL_CERT keystore in VECS.
- ❒ It is stored locally in the vCenter Server daemon (vpxd).
- ✔ It is stored locally on the relevant ESXi host.
- ❒ It is stored locally on every vSphere node.

https://t.me/learningnets

4-16 VMware CA Modes



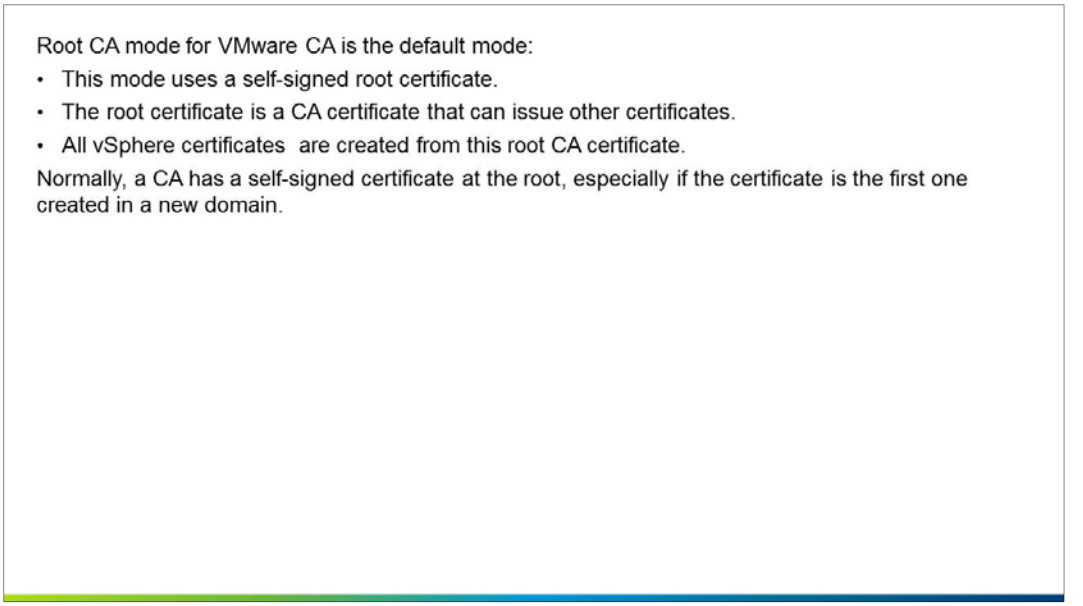
By default, VMware CA is configured to act as a root CA that can use its self-signed root certificate to issue and sign other certificates. However, you can also configure VMware CA as a subordinate CA. In subordinate CA mode, VMware CA is an intermediate CA in the chain of trust that continues up to an enterprise CA or an external third-party CA.

If you do not want to use VMware CA, you can bypass VMware CA and install custom certificates on all your systems.

Use subordinate CA mode if you want to issue and install your own certificates. In this situation, you must issue a certificate for every component. Certificate management is your responsibility. All your custom certificates (except for host certificates) must be installed in VECS.

https://t.me/learningnets

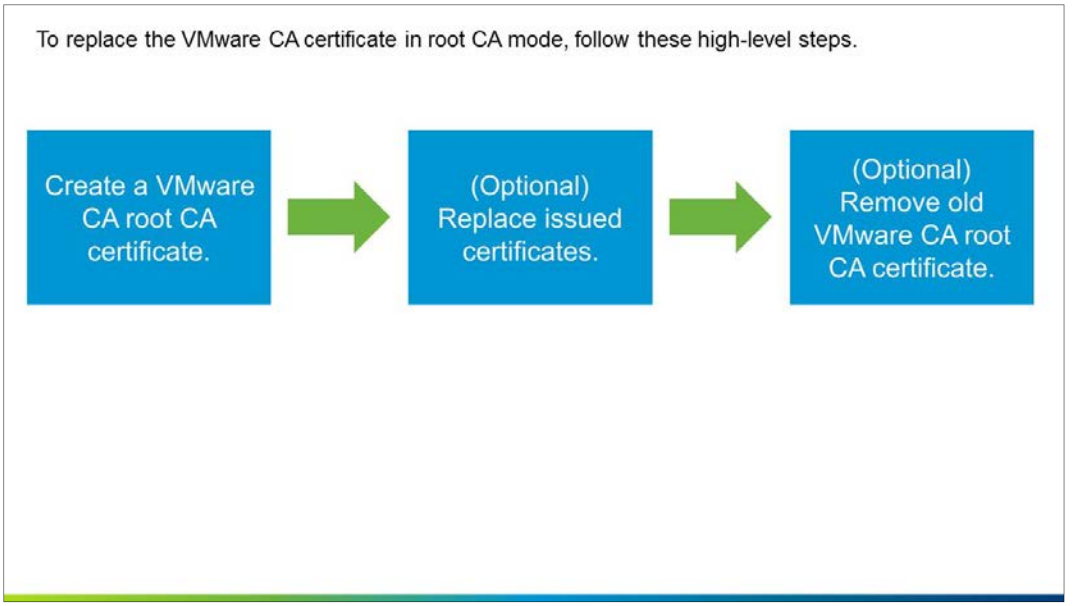
4-17 VMware CA: Root CA Mode



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

4-18 Replacing the VMware CA Certificate in Root CA Mode

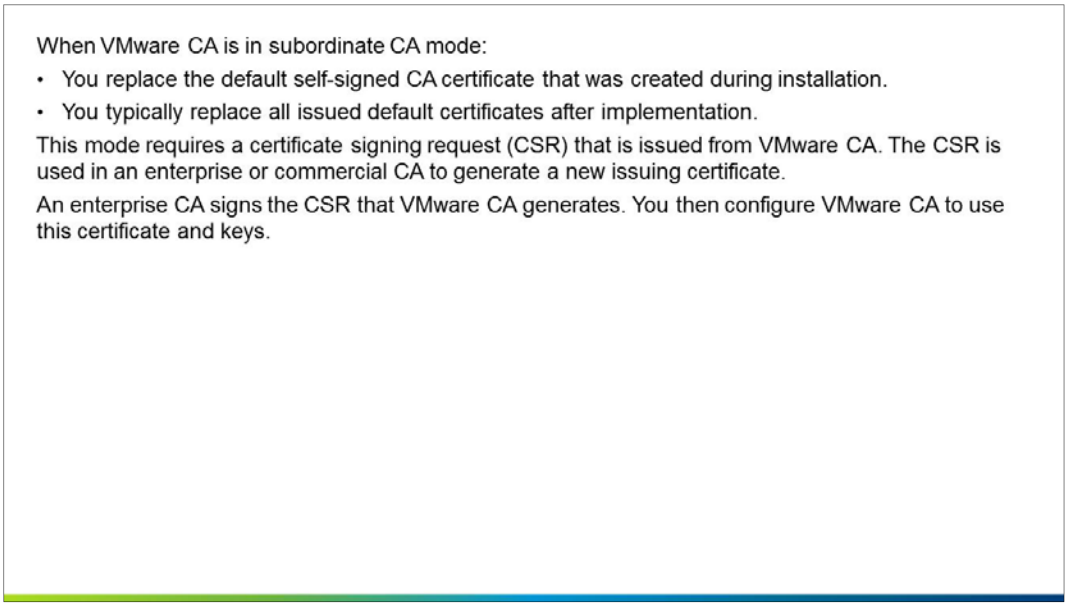


You are not required to replace the certificates that were issued and signed by the previous VMware CA root CA certificate. When you replace the VMware CA root CA certificate, it does not automatically invalidate the previous certificate. It also does not remove the old certificate, so the old certificate remains valid in your vSphere environment.

However, when you want to replace the root CA certificate, you typically also want to replace all 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 these certificates at once by using the vSphere Certificate Manager tool.

https://t.me/learningnets

4-19 VMware CA: Subordinate CA Mode



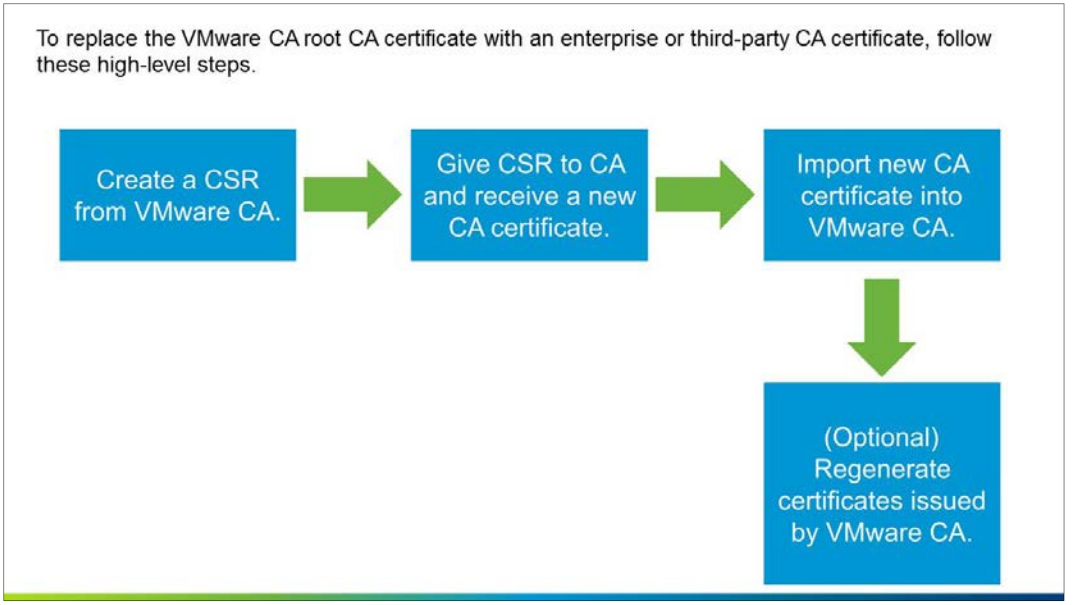
In subordinate CA mode, you replace the VMware CA root certificate with a third-party CA-signed certificate that includes VMware CA in the certificate chain. The replacement certificate can be signed by your enterprise CA or by a third-party, external CA.

All certificates that VMware CA generates include the full chain. You can replace existing certificates with newly generated certificates, although this step is not required. The original certificate is still valid if it has not expired or been revoked, and 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

4-20 Replacing the VMware CA Certificate with a Subordinate 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 (CSR) that you can present to your enterprise or third-party CA so that the CA can create an appropriate certificate for you.

https://t.me/learningnets

4-21 VMware CA: Hybrid Mode

In Hybrid Mode, VMware CA supplies some certificates but uses custom certificates for other parts of your infrastructure:

- Company policy may not allow intermediate CAs.
- Replace the vCenter Server Machine SSL Certificate with a custom certificate that was signed by the enterprise CA and saved in VECS.
- All other vSphere traffic is secured by the default self-signed certificates issued by VMware CA.

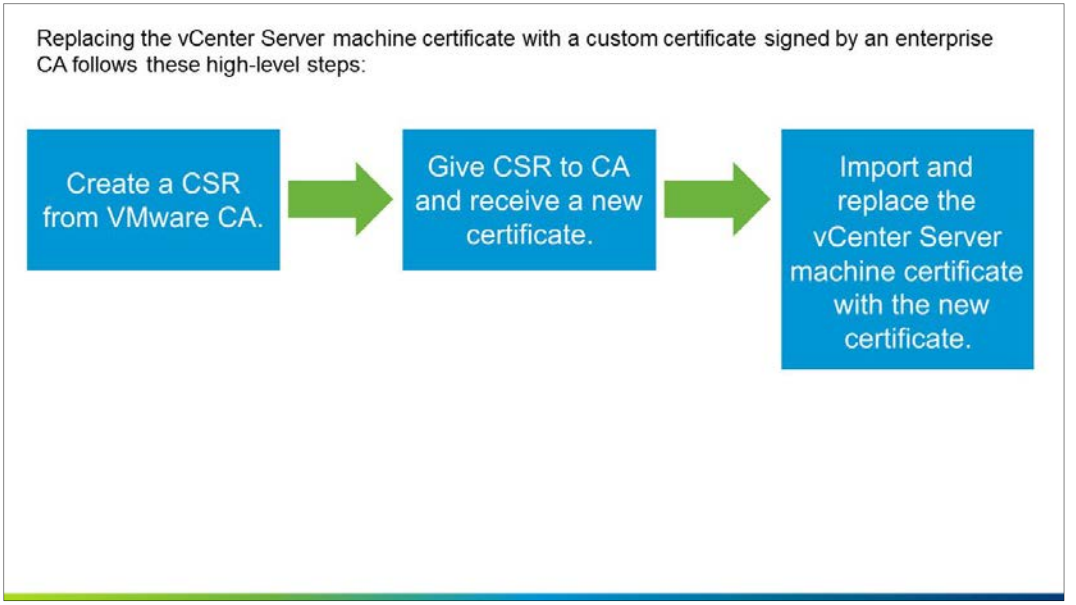
This mode requires a certificate signing request (CSR) that is issued by VMware CA.

An enterprise CA uses the CSR to sign a new certificate that will be used to replace the vCenter Server machine certificate.

For more information about replacing certificates, see *Certificate Replacement Overview* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.authentication.doc/GUID-4469A6D3-048A-471C-9CB4-518A15EA2AC0.html

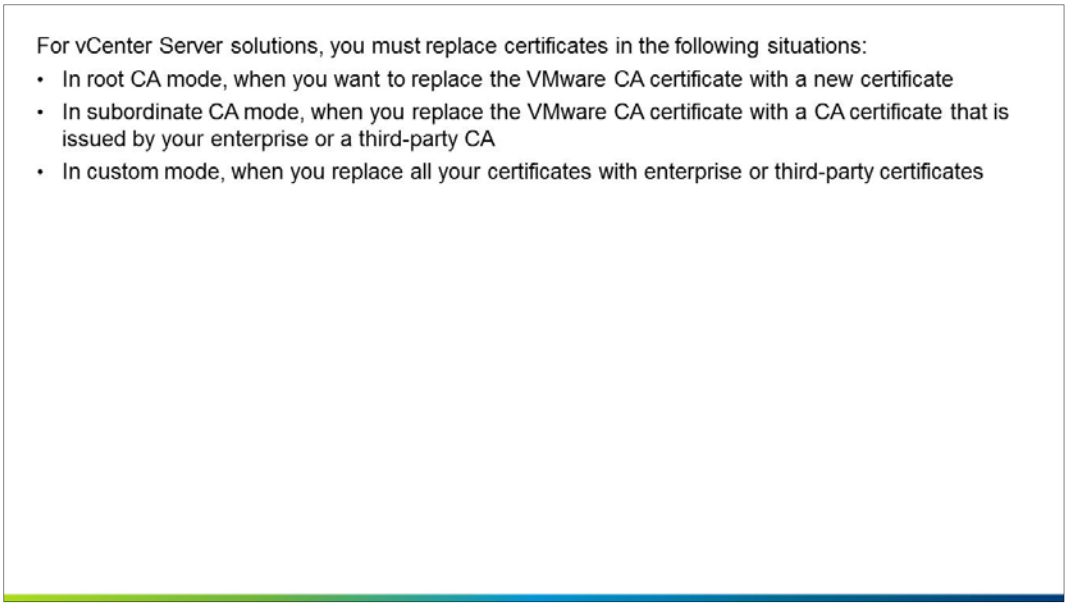
https://t.me/learningnets

4-22 Hybrid Mode: Replacing a vCenter Server Machine Server Certificate with a Custom Certificate



https://t.me/learningnets

4-23 Certificate Replacement Options



With the introduction of VMware CA in vSphere 6.0, managing your vSphere certificates is much easier. The situations described on the slide are the three most common ones.

https://t.me/learningnets

4-24 Certificate Manager

With vSphere 7, you can use the vSphere Client to replace the certificates in your vSphere environment and to reset all vSphere certificates. You can also use the vSphere Certificate Manager.

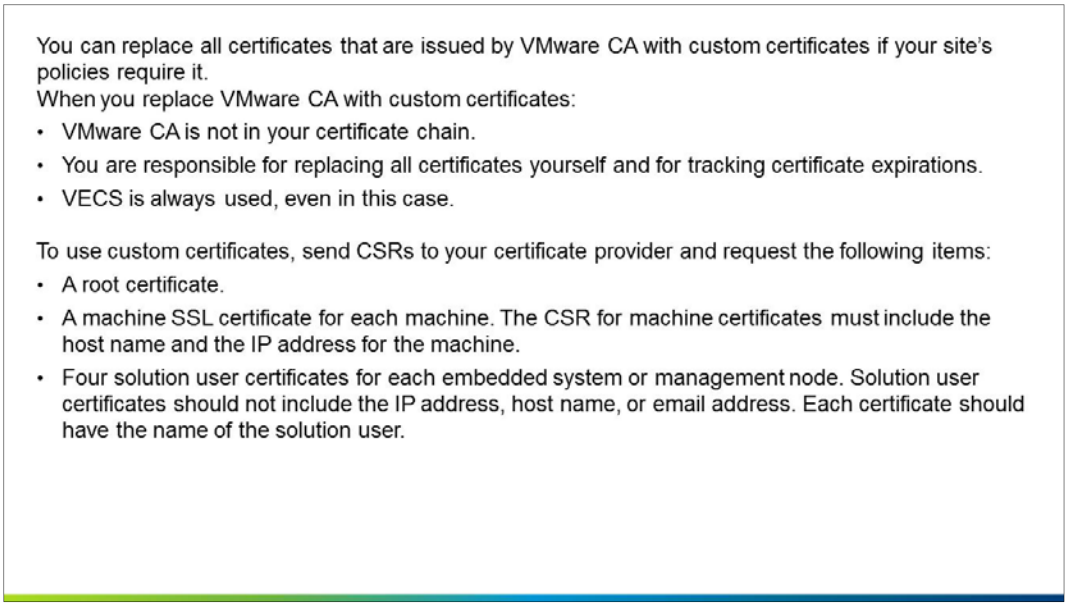
To start the vSphere Certificate Manager, log in to your vCenter Server system and run the commands at the command prompt:

```
/usr/lib/vmware-vmca/bin/certificate-manager
```

For more information about certificate management, see *Certificate Management Overview* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.authentication.doc/GUID-3D0DE463-D0EC-442E-B524-64759D063E25.html.

https://t.me/learningnets

4-25 Replacing All VMware CA Certificates with Custom Certificates



Replacing VMware CA issued certificates with custom enterprise or third-party certificates is not recommended. However, your own site security policies might dictate that you use only a specific CA for all your site's systems.

When you replace the VMware CA certificates with custom certificates, you increase your management tasks, which can increase the risk of errors that lead to security vulnerabilities. You are responsible for all certificate management tasks. Because you do not use VMware CA, you cannot rely on VMware CA to manage your certificates for you.

https://t.me/learningnets

4-26 ESXi Certificate Replacement Options

Several options for ESXi certificate replacement are available:

- 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:
  - You can use your own custom certificates.
  - You must have an enterprise CA.

https://t.me/learningnets

4-27 Renewing or Refreshing ESXi Certificates

If VMCA assigned certificates to your ESXi hosts (6.0 and later), you can renew or refresh those certificates from the vSphere Web Client.

- Renewing certificates:
  - Typically, you renew certificates when they are about to expire, but you can provision a host with a new certificate for other reasons.
  - Renewing retrieves a fresh signed certificate for the host from VMCA.
  - If the certificate is already expired, you must disconnect the host and reconnect it.
  - By default, vCenter Server renews the certificates of a host with the Expired, Expiring immediately, or Expiring status each time the host is added to the inventory or reconnected.
- Refreshing CA certificates pushes all certificates in the TRUSTED_ROOTS store in the vCenter Server VECS store to the host.

To renew or refresh CA certificates from the vSphere Client:

1 Browse to the host in the vSphere Client inventory.

2 Click the **Manage** tab and click **Settings**.

3 Under System, click **Certificate**.
You can view detailed information about the selected host's certificate.

4 Choose to **Renew** or **Refresh**.

5 Click Yes to confirm.

https://t.me/learningnets

4-28 VMware CA Availability

In the unlikely event that VMware CA is not available, some operations do not work:

- Adding new ESXi hosts
- Refreshing or renewing certificates on any ESXi host

Sometimes, certificate refreshing on certificate expiry might fail.

If VMware CA is unavailable for short periods of time, operations are not usually disrupted.

VMware CA is 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 be disrupted.

https://t.me/learningnets

4-29 Lab 8: Working with Certificates

Generate and replace a vCenter Server certificate using the vSphere Client:

1. Examine the Machine SSL Certificate
2. Generate a CSR for the Custom Certificate
3. Request a Custom Certificate
4. Replace the Current Certificate with a Custom Certificate

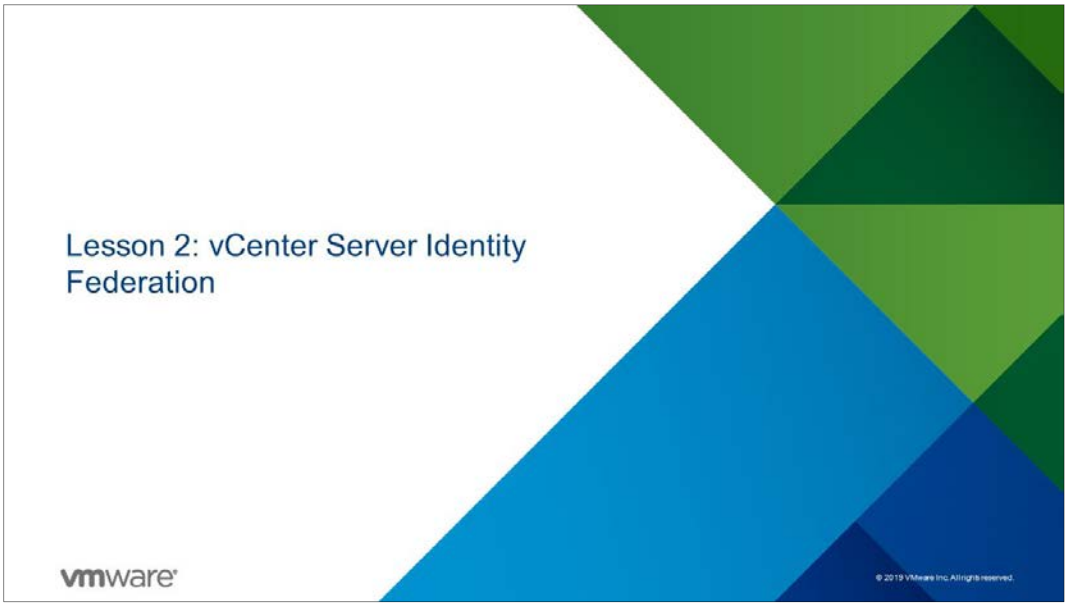
https://t.me/learningnets

4-30 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Use VMware CA 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 the vSphere Client to manage vSphere certificates

https://t.me/learningnets



https://t.me/learningnets

4-32 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe identity federation
- Recognize use cases for identity federation
- Describe the architecture of identity federation
- Configure identity federation

https://t.me/learningnets

4-33 About vCenter Server IDP Federation for ADFS

vCenter Server 7.0 supports identity provider (IDP) federation for Active Directory Federation Services (ADFS).

In a vCenter Server identity provider federation configuration, the external IDP interacts with the identity source on behalf of vCenter Server.

In this federation authentication setup, vCenter Server does not perform the following functions:

- Handling the user's credentials
- Creating the initial identity token
- Managing or updating user and group information in the identity source

https://t.me/learningnets

4-34 vCenter Server IDP Federation Use Cases

You can use an external IDP to replace vCenter Server as the IDP and enable single sign-on (SSO) with federated infrastructure and applications.

You can use the vCenter Server IDP federation in the following scenarios:

- When using SSO with the existing ADFS-integrated infrastructure and applications
- For improving data center security because vCenter Server never handles the user's credentials
- When using the authentication mechanisms that are supported by the external IDP

https://t.me/learningnets

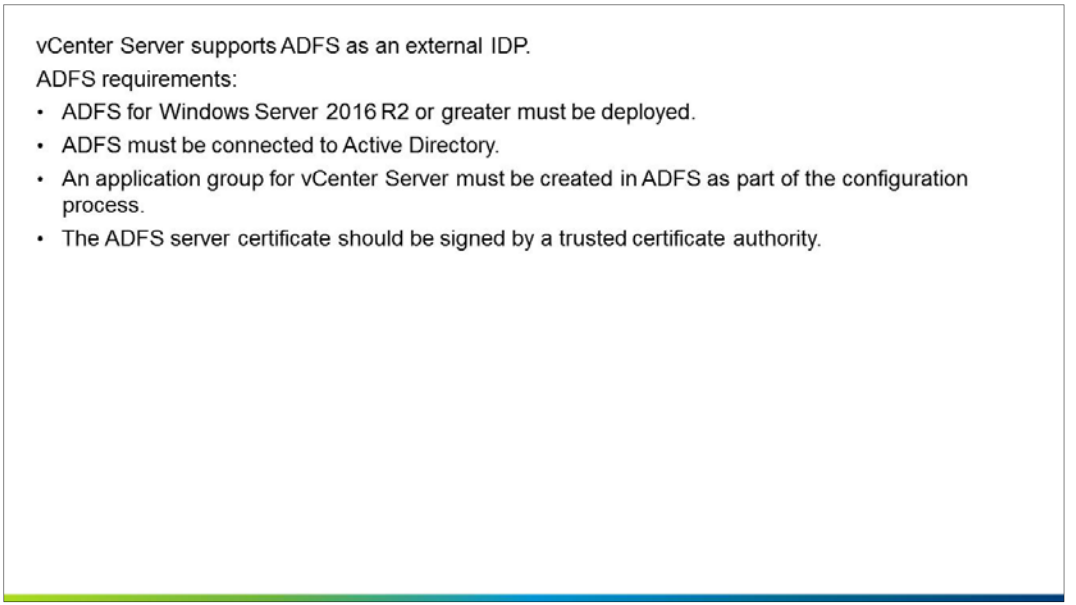
4-35 vCenter Server IDP Federation Components

The vCenter Server 7.0 IDP federation configuration includes the following vSphere and third-party components:

- vCenter Server
- An IDP service that is configured on vCenter Server
- An ADFS server and associated Microsoft AD domain
- An ADFS application group
- AD groups and users that map to vCenter Server groups and users

https://t.me/learningnets

4-36 Prerequisites for vCenter Server IDP Federation



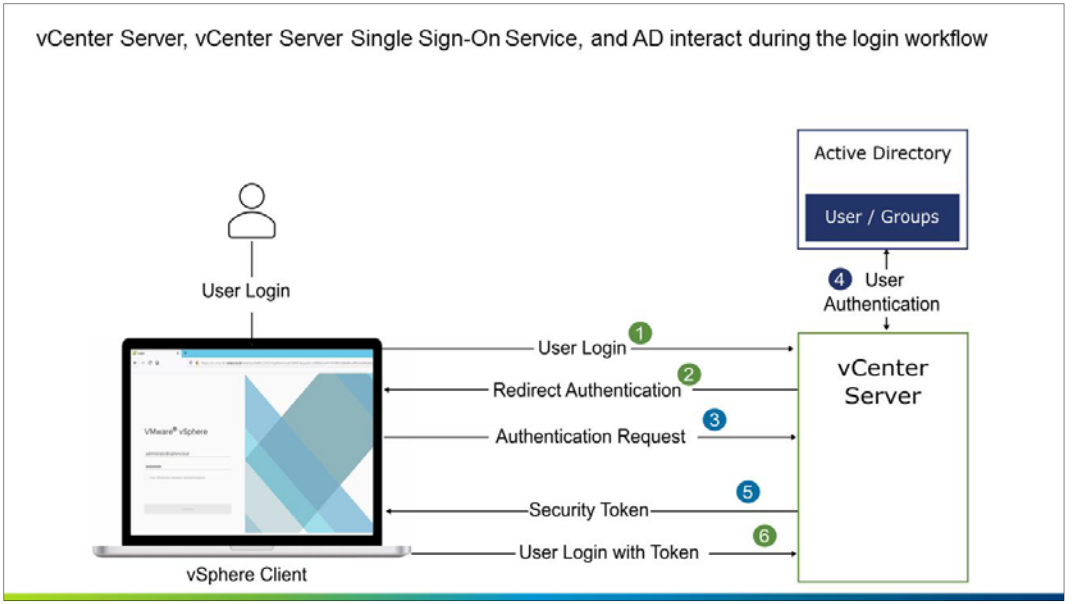
vCenter Server supports only one external IDP (one ADFS source) and the local OS or vsphere.local identity source.

You cannot use multiple external IDPs.

The vCenter Server IDP federation uses OpenID Connect (OIDC) for user login to vCenter Server.

https://t.me/learningnets

4-37 User Login with the vCenter Server Built-In Identity Provider



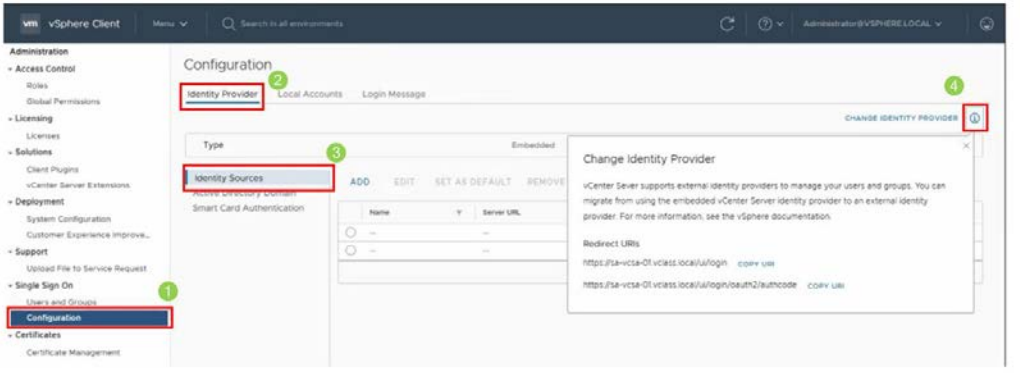
Before using vCenter Server IDP Federation for ADFS, this is how the vCenter Server, vCenter Server Single Sign-On Service, and AD would interact during the login workflow:

1. The user logs in to vCenter Server on the usual vCenter Server landing page.
2. vCenter Server redirects the authentication request to the vCenter Single Sign on Service
3. If needed, the vCenter Single Sign on Service prompts the user to log in with Active Directory credentials
4. The Single Sign On service authenticates the user with Active Directory
5. The vCenter Single Sign-On service signs all tokens with a signing certificate, and stores the token signing certificate on disk. The certificate for the service itself is also stored on disk.
6. vCenter Server uses the token to log in the user.

https://t.me/learningnets

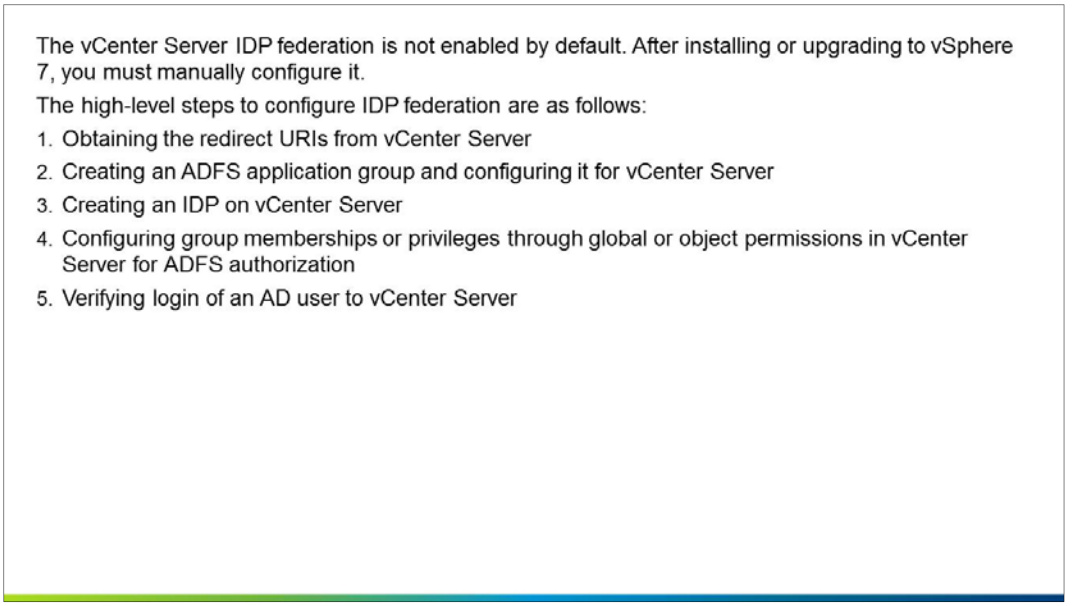
4-38 Configuring vCenter Server IDP Federation in vSphere (1)

In the vSphere Client, you obtain the redirect URIs from vCenter Server.



https://t.me/learningnets

4-39 Configuring vCenter Server IDP Federation in vSphere (2)

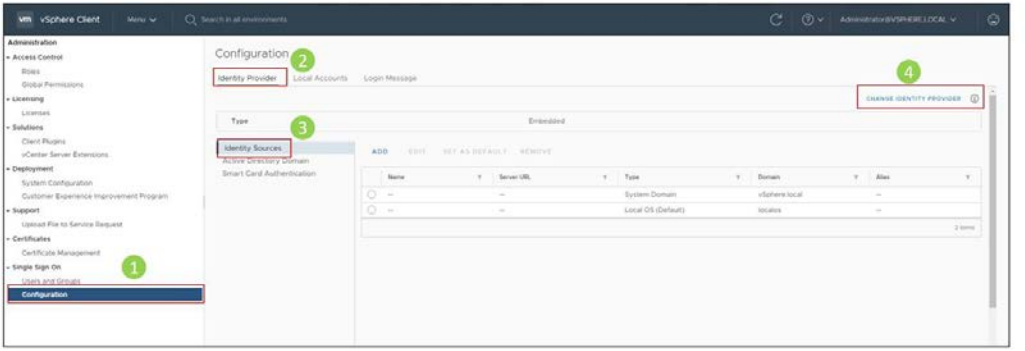


For more information about configuring vCenter Server IDP federation for vCenter Server, see the VMware knowledge base article at <link TBD.>

https://t.me/learningnets

4-40 Configuring vCenter Server IDP Federation in vSphere (3)

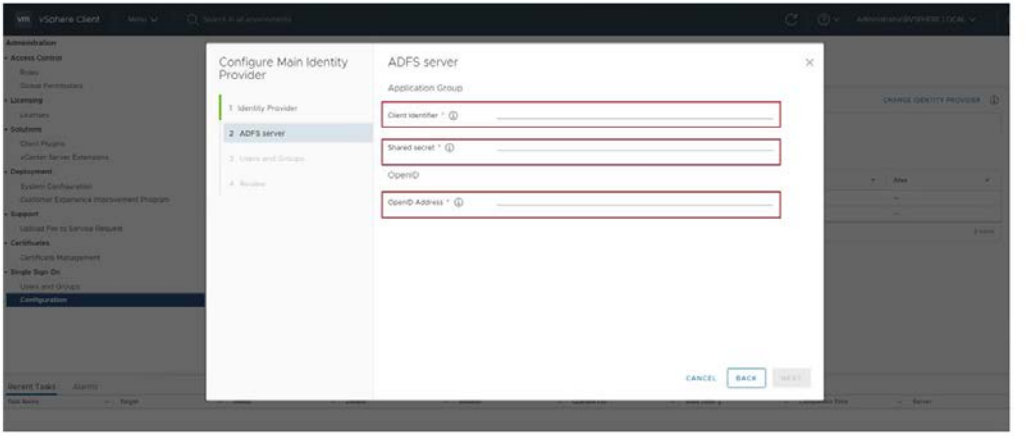
In the vSphere Client, you can change the vCenter Server identity provider federation.



https://t.me/learningnets

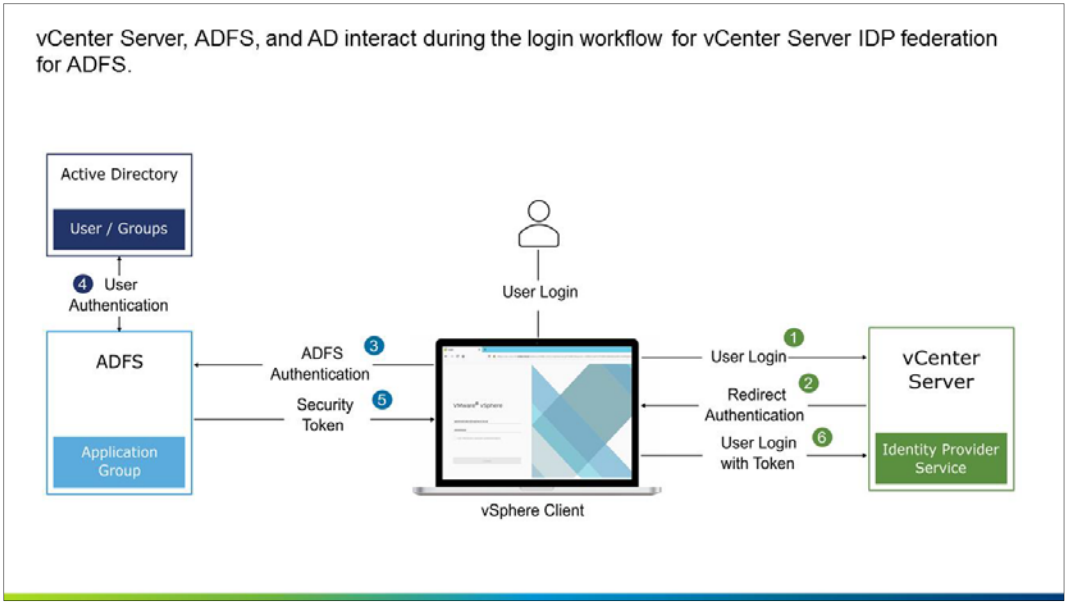
4-41 Configuring vCenter Server IDP Federation in vSphere (4)

In the vSphere Client, you provide application group information.



https://t.me/learningnets

4-42 Login Workflow of the vCenter Server IDP Federation for ADFS



vCenter Server, ADFS, and AD interact during the login workflow as follows:

1. The user logs in to vCenter Server on the usual vCenter Server landing page.
2. vCenter Server redirects the authentication request to ADFS.
3. If needed, ADFS prompts the user to log in with Active Directory credentials.
4. ADFS authenticates the user with Active Directory.
5. ADFS issues a security token with group information from Active Directory.
6. vCenter Server uses the token to log in the user.

If the external IDP is unreachable, the login process falls back to the vCenter Server landing page, showing an appropriate information message. Administrative users can still log in using their local accounts (local OS or vsphere.local identity sources).

https://t.me/learningnets

4-43 Enabling ADFS on an Existing Enhanced Linked Mode Configuration

To enable ADFS on an existing Enhanced Linked Mode configuration:

1. Deploy a minimum of two vCenter Server nodes in an Enhanced Linked Mode configuration.
2. Configure ADFS on one of the linked vCenter Server nodes. The ADFS configuration is replicated to the other vCenter Server node.
3. Add all redirect URIs for both vCenter Server nodes to the configured OAuth Application Group in ADFS.

https://t.me/learningnets

4-44 Lab 9: Configuring Identity Federation to Use Microsoft ADFS

Configure the vCenter Server identity provider federation with Microsoft ADFS:

1. Configure vCenter Server Identity Provider Federation
2. Add Permissions to vCenter Server for an AD Account
3. Log In to vCenter Server Using an AD Account

https://t.me/learningnets

4-45 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe identity federation
- Recognize use cases for identity federation
- Describe the architecture of identity federation
- Configure identity federation

https://t.me/learningnets

https://t.me/learningnets

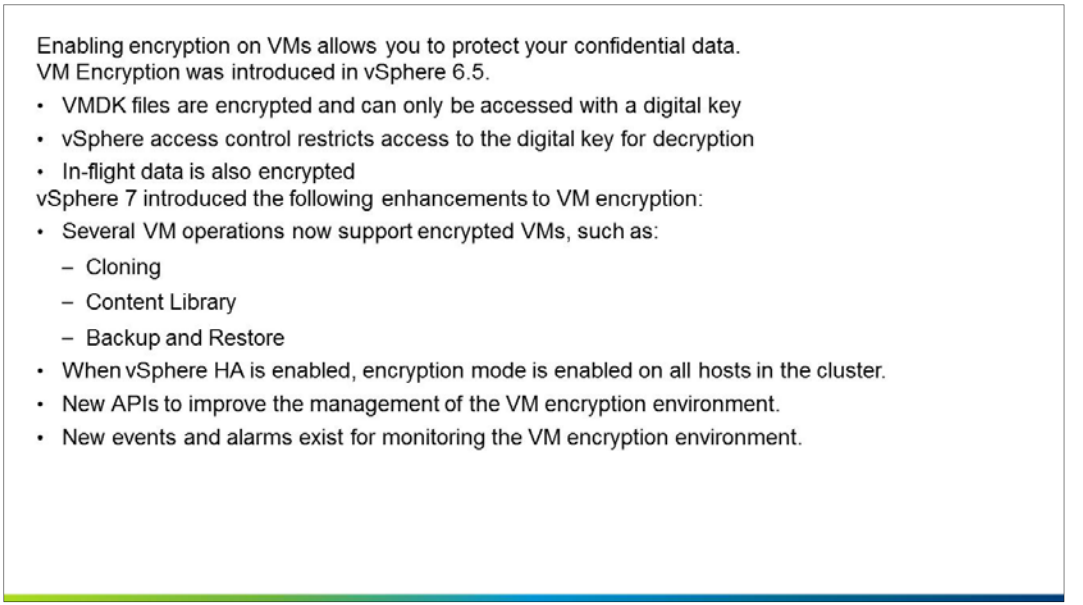
4-47 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe vSphere Trust Authority
- List the use cases of vSphere Trust Authority
- Explain the general architecture of vSphere Trust Authority
- Configure vSphere Trust Authority

https://t.me/learningnets

4-48 Overview of VM Encryption

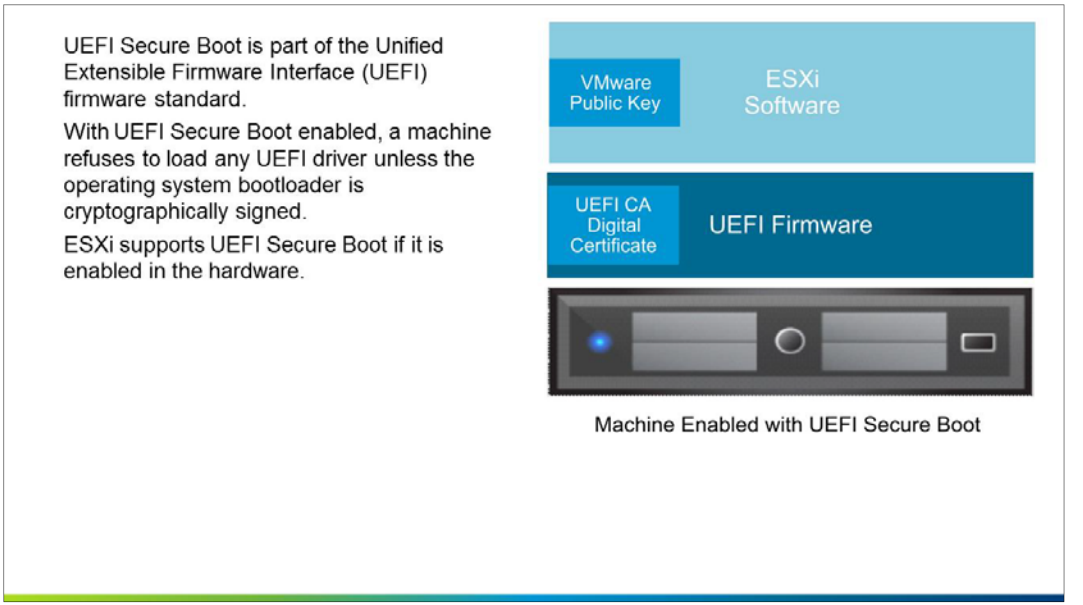


The enhancements introduced in vSphere 7 improved the general functionality of VM encryption and are based on competitive analysis and feedback from VMware customers and partners.

To mitigate the risk of an administrator accessing data with authorization, VMs can be encrypted and access to encryption/decryption functionality can be provided to a subset of administrators. This is achieved using the inbuilt cryptographic privileges in vSphere.

https://t.me/learningnets

4-49 UEFI Secure Boot for ESXi Hosts



Starting with vSphere 7, a vSphere Trust Authority Trusted Cluster must have UEFI Secure Boot enabled.

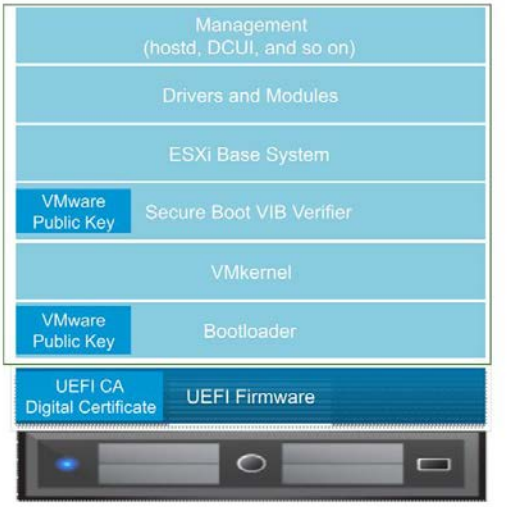
To ensure that boot code is executed without modification, Unified Extensible Firmware Interface (UEFI) uses a digital signature provided by a trusted code creator. The digital signature is embedded in every executable code section. Using public-private key pairs, the code creator signs the code with 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, 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

4-50 UEFI Secure Boot Sequence for ESXi Hosts

ESXi supports UEFI Secure Boot at each level of the boot stack.
The boot sequence proceeds as follows:

1. The UEFI firmware validates the bootloader and the VMkernel.
2. The secure boot VIB verifier verifies every vSphere installation bundle package that is installed on the system.



https://t.me/learningnets

4-51 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 secures hardware by integrating cryptographic keys into devices.

vSphere 6.7 introduced 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.

https://t.me/learningnets

4-52 About Remote Attestation

TPM hardware attests to an ESXi host's identity. This process is called remote attestation.

Remote attestation authenticates and attests 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.

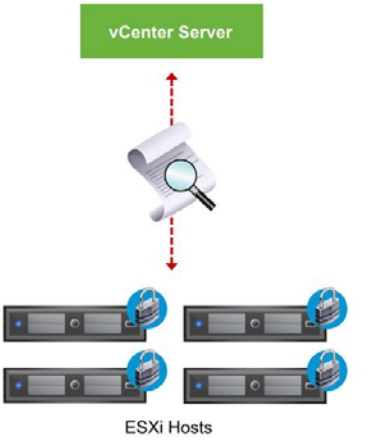
https://t.me/learningnets

4-53 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 Platform Configuration Registers (PCR), which is signed by the attestation key.
   - vCenter Server verifies the authenticity of the signed quote from the ESXi host. If the signed quote is invalid, 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 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

4-54 Requirements for Using TPM 2.0

To use a TPM 2.0 chip, your vSphere environment must meet these requirements:

- vCenter Server 6.7 or later
- ESXi 6.7 or later with TPM 2.0 chip installed and enabled in UEFI
- UEFI Secure Boot enabled

Review the TPM 2.0 chips that are certified by VMware by using the VMware Compatibility Guide at http://www.vmware.com/go/hcl.

https://t.me/learningnets

4-55 About vSphere Trust Authority

vSphere Trust Authority secures a vSphere environment and its workloads:

- Uses a hardware root of trust to secure the environment
- Reports on the software running on a host
- Measures or attests software using a Trusted Platform Module (TPM)

A Trust Authority Administrator must configure the following parameters:

- Permitted versions of ESXi
- Trusted TPM manufacturers or devices

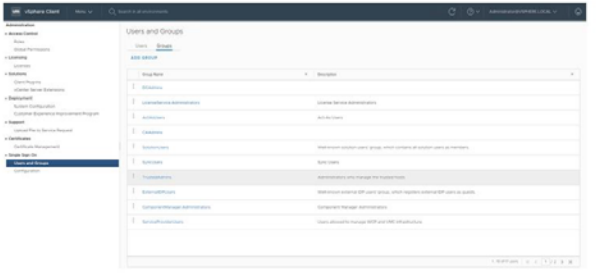
A trusted key provider is required. A valid attestation report is needed before access is granted to the key provider.

https://t.me/learningnets

4-56 About vSphere Trust Authority Administrator

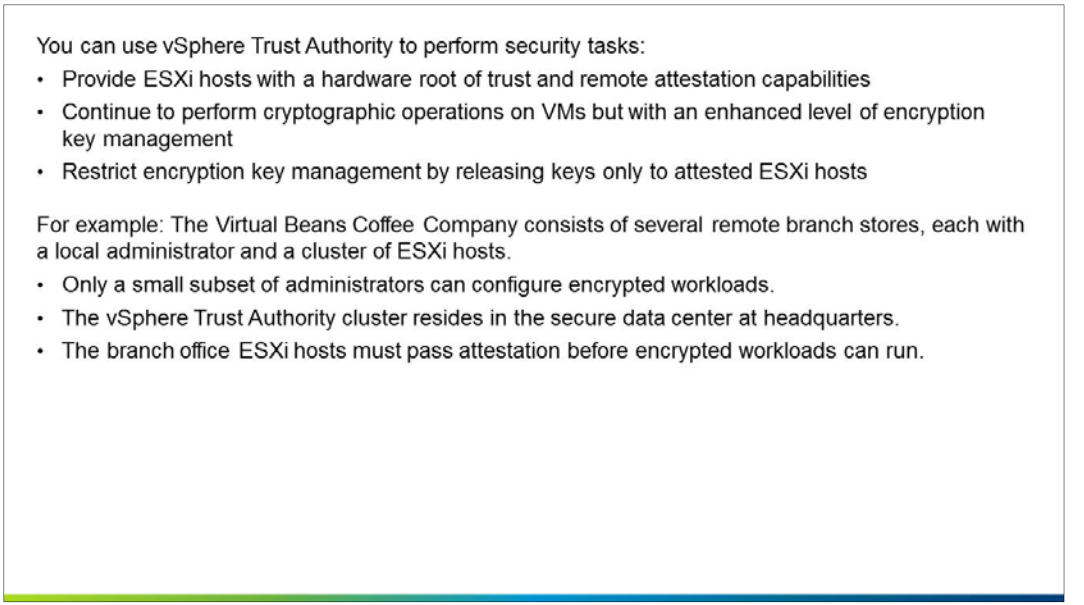
To enable vSphere Trust Authority, you must add a user to the vSphere TrustedAdmins group.

- This user becomes the Trust Authority Administrator.
- Use a separate user from the vCenter Server Administrator as your Trust Authority Administrator.
- You must add a Trust Authority Administrator for both the Trust Authority Cluster and the Trusted Cluster.
- Group membership is validated by ESXi using SAML token authentication.



https://t.me/learningnets

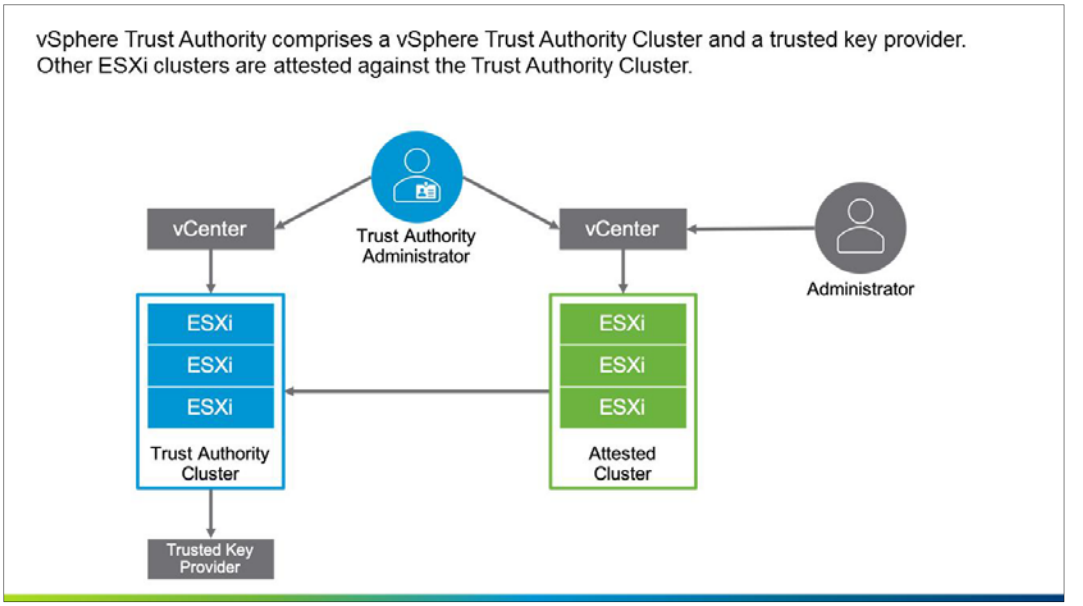
4-57 Use Case for vSphere Trust Authority



In vSphere 7, you can create a trusted computing base, which consists of a secure, manageable set of ESXi hosts. vSphere Trust Authority implements a remote attestation service for the ESXi hosts you want to trust. Furthermore, vSphere Trust Authority improves upon TPM 2.0 attestation support (added to vSphere in the 6.7 release), to implement access restrictions on encryption keys and so better protect VM workload secrets. In addition, vSphere Trust Authority allows only authorized Trust Authority administrators to configure vSphere Trust Authority services, and configure Trust Authority hosts. The Trust Authority Administrator can be the same user as the vSphere administrator user, or a separate user.

https://t.me/learningnets

4-58 vSphere Trust Authority Components (1)

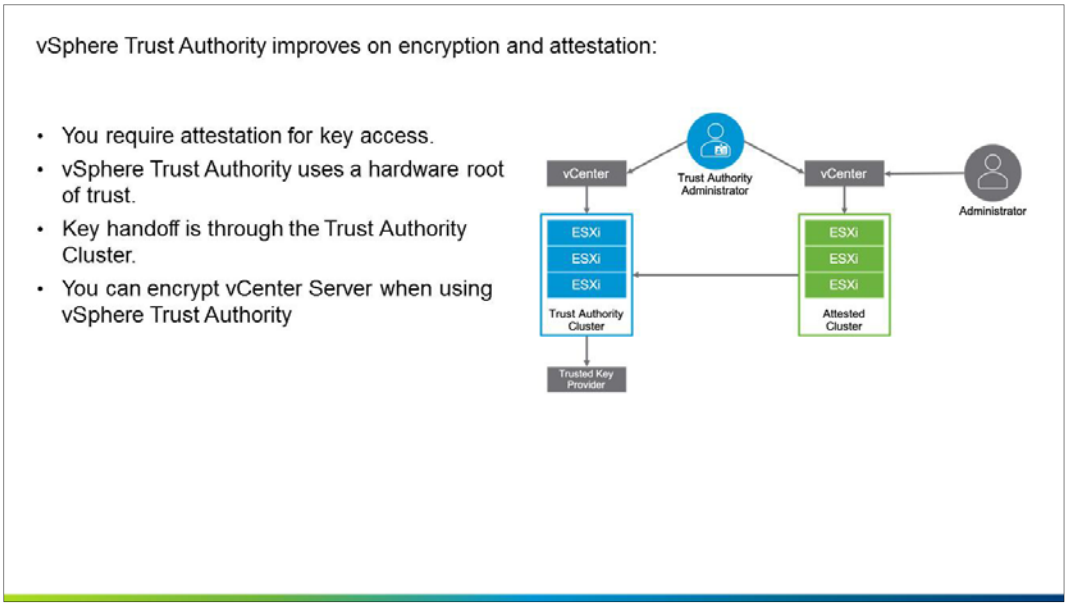


Several concepts apply to vSphere Trust Authority:

- Trust Authority Cluster: A set of restricted ESXi hosts with a known good configuration.
- Administered by a smaller number of trusted administrators.
- Trusted Key Provider (also called key management server [KMS]): A key provider that only the Trust Authority Cluster should know.
- Attested cluster: Also known as a Trusted Cluster. The attestation report for a cluster must be validated by the Trust Authority Cluster.

https://t.me/learningnets

4-59 vSphere Trust Authority Components (2)



Before vSphere 7, in vSphere 6.x, vSphere Trust Authority worked in the following ways:

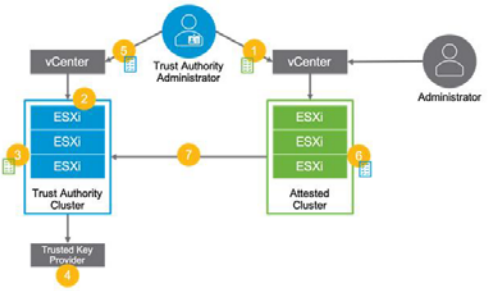
- Attestation is a read-only view, and no repercussions occur when it fails.
- The attester is vCenter Server.
- Key handoff occurs through vCenter.
- You cannot safely encrypt vCenter Server because it forms a dependency on key access.

https://t.me/learningnets

4-60 vSphere Trust Authority Workflow

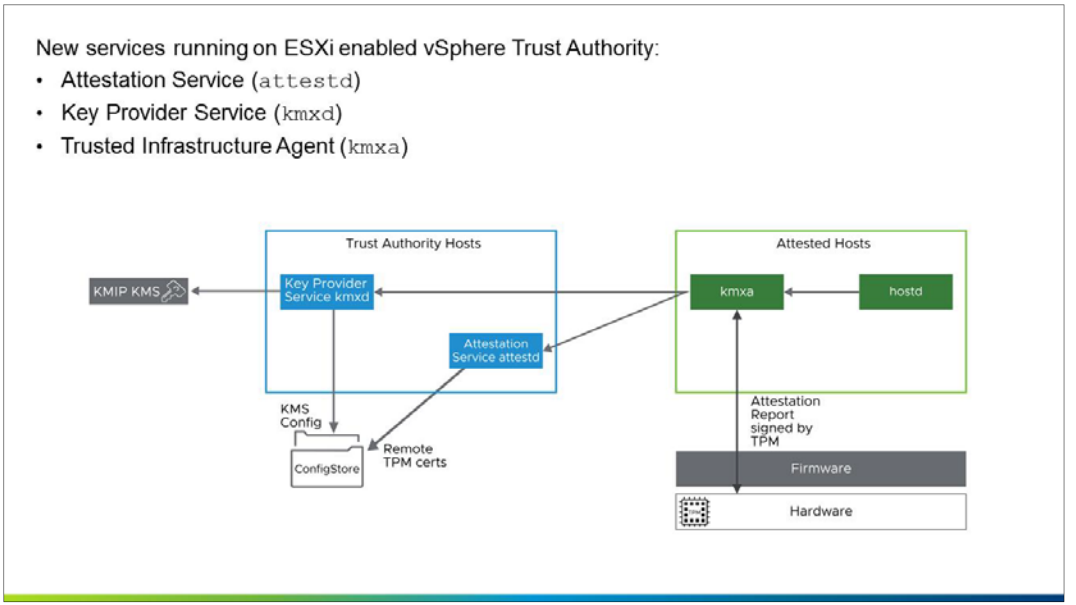
In the vSphere Trust Authority workflow, a Trust Authority Administrator performs the following tasks:

1. Exports the identity of the hosts to be attested
2. Enables Trusted Host Services on a Trust Authority Cluster
3. Imports the identity of the attesting hosts into the Trust Authority Cluster
4. Configures a trusted key provider
5. Exports the identity of the Trust Authority Cluster
6. Imports the Trust Authority Cluster identity into the attesting cluster
7. Configures the Trust Authority Cluster as the key provider for the attesting cluster



https://t.me/learningnets

4-61 ESXi Components



All vSphere Trust Authority services run behind the Reverse Proxy and API Forwarder. The Trusted Infrastructure Agent communicates over port 443 to the Trust Authority Hosts.

The following internal ports are used:

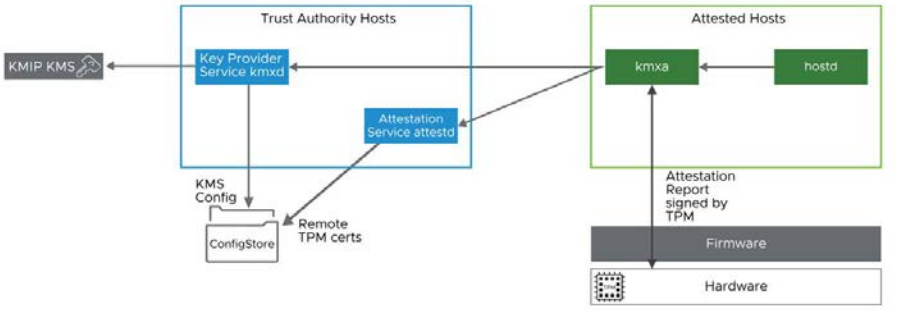
- Attestation Service : Port 7889
- Key Provider Service : Port 7888
- Trusted Infrastructure Agent : Port 7890

https://t.me/learningnets

4-62 ESXi Components

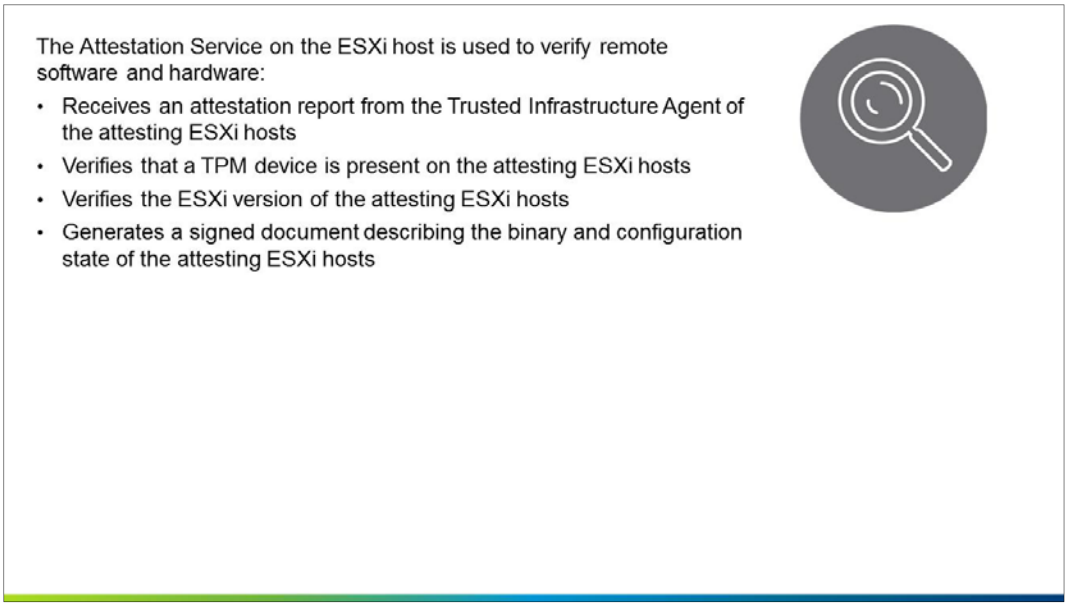
Services running on ESXi enabled vSphere Trust Authority:

- Attestation Service (`attestd`)
- Key Provider Service (`kmxd`)
- Trusted Infrastructure Agent (`kmxa`)



https://t.me/learningnets

4-63 Attestation Service



The Attestation Service performs the following functions:

- Generates a signed document that contains assertions describing the binary and configuration state of the remote ESXi hosts in the Trusted Cluster.
- Attests the state of the ESXi hosts using a Trusted Platform Module (TPM) 2.0 chip for software measurement and reporting. The TPM measures the software stack and configuration data and reports it to the Attestation Service.
- Verifies that the software measurement signature can be attributed to a previously configured trusted TPM endorsement key (EK).
- Ensures that the software measurement matches one of a set of previously blessed ESXi images.
- Signs a SAML token that it issues to the ESXi host, providing the assertions about the identity, validity, and configuration of the ESXi host.

https://t.me/learningnets

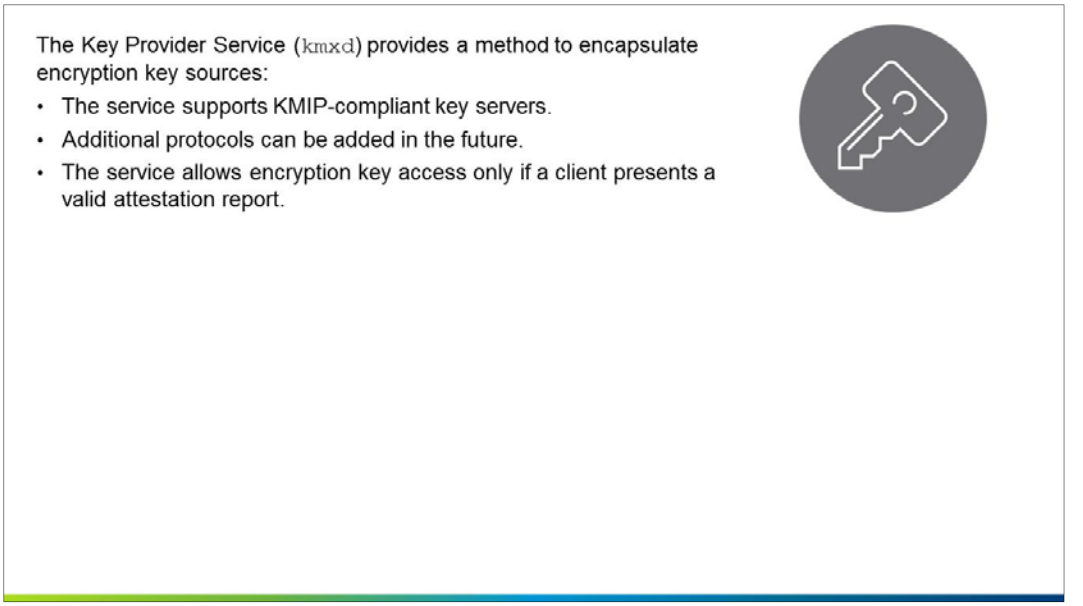
4-64 Attestation Service Trust

The Attestation Service does not know what to trust.
A Trust Authority Administrator must configure the following items:

- Which versions of ESXi to trust
- Which TPM device manufacturers to trust
- Optionally, which specific TPM devices to trust

https://t.me/learningnets

4-65 Key Provider Service



With the Key Provider Service, vCenter Server and ESXi hosts do not require direct key management server (KMS) credentials.

In vSphere Trust Authority, an ESXi host can access an encryption key by authenticating with the Key Provider Service rather than with vCenter Server (as in vSphere 6.7).

For the Key Provider Service to connect to a KMS, the Trust Authority Administrator must configure a trust setup. For most servers that are compliant with the Key Management Interoperability Protocol (KMIP), you configure a trust setup by creating client and server certificates.

https://t.me/learningnets

4-66 Trusted Infrastructure Agent

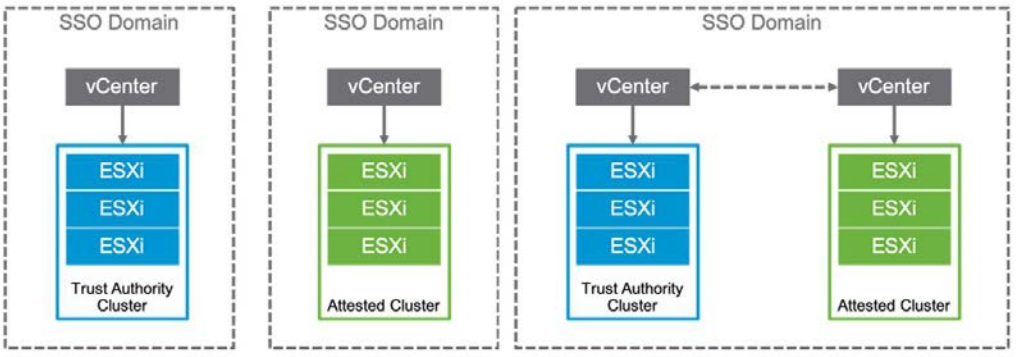
Trusted Infrastructure Agent communicates with a Trusted Key Provider Service and Attestation Service :

- The Trusted Infrastructure Agent retrieves encryption keys from the Trusted Key Provider Service.
- The agent can be issued with encryption keys only if the ESXi host passes attestation.
- The ESXi host does not receive encryption keys from vCenter Server when using a Trusted Key Provider.

https://t.me/learningnets

4-67 vCenter Server Topologies (1)

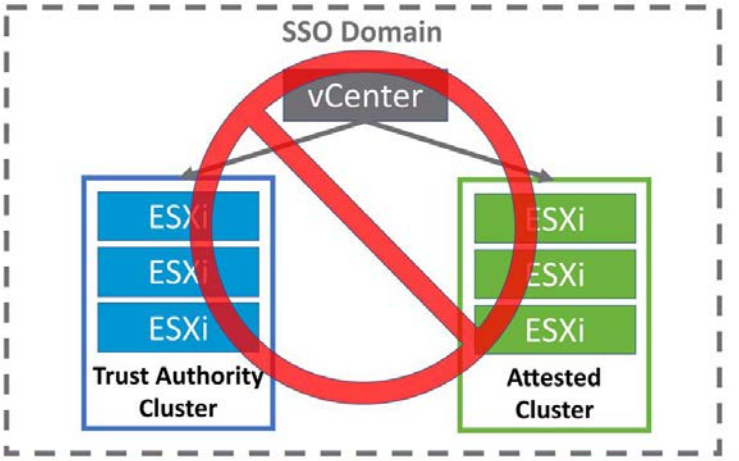
vSphere Trust Authority requires two or more vCenter Server systems. The Trust Authority cluster must reside in a different vCenter Server inventory than the Trusted (Attested) cluster but can participate in Enhanced Linked Mode



https://t.me/learningnets

4-68 vCenter Server Topologies (2)

A single vCenter Server system managing both the Trust Authority cluster and Trusted (attested) cluster is not supported.



https://t.me/learningnets

4-69 About Enabling vSphere Trust Authority

In vSphere 7, you enable vSphere Trust Authority primarily through vSphere PowerCLI:

- vSphere PowerCLI 12.0 provides a new set of vSphere PowerCLI cmdlets.
- A user who runs the cmdlets must be a member of the TrustedAdmins single sign-on group.
- A user interface is planned for a future release.

https://t.me/learningnets

4-70 About VM Encryption with vSphere Trust Authority

The user experience of VM encryption does not change when it is used with vSphere Trust Authority in vSphere 7:

- You can encrypt a VM by applying an encryption storage policy.
- An attested cluster uses a trusted key provider.
- You can mix a trusted key provider and a standard key provider in the same vSphere environment..
- When you add a virtual TPM device to a VM, the VM is encrypted.

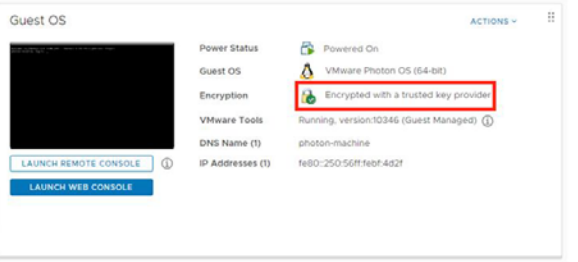
https://t.me/learningnets

4-71 VM Encryption with a Trusted Key Provider

A VM that is encrypted with a trusted key provider has a slightly different icon from a VM encrypted with a standard key provider. It also has the following characteristics:

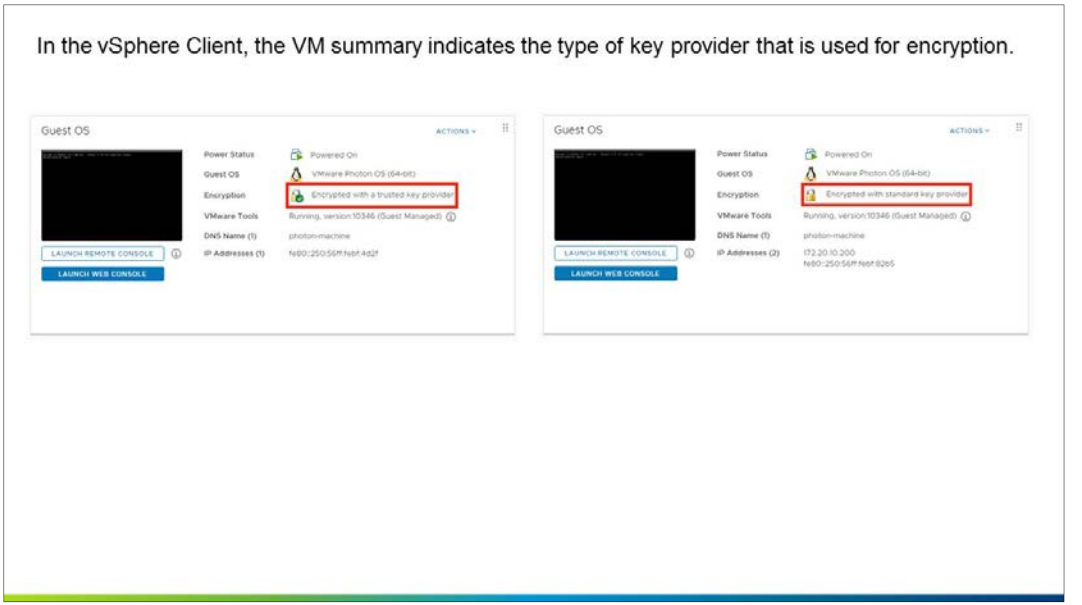
- Can run only on ESXi hosts attested against the vSphere Trust Authority Cluster
- Cannot be migrated to an ESXi host that is not attested against the same vSphere Trust Authority Cluster

A VM encrypted with a Standard Key Provider can be migrated to an attested ESXi host, but it continues to use the standard key provider.



https://t.me/learningnets

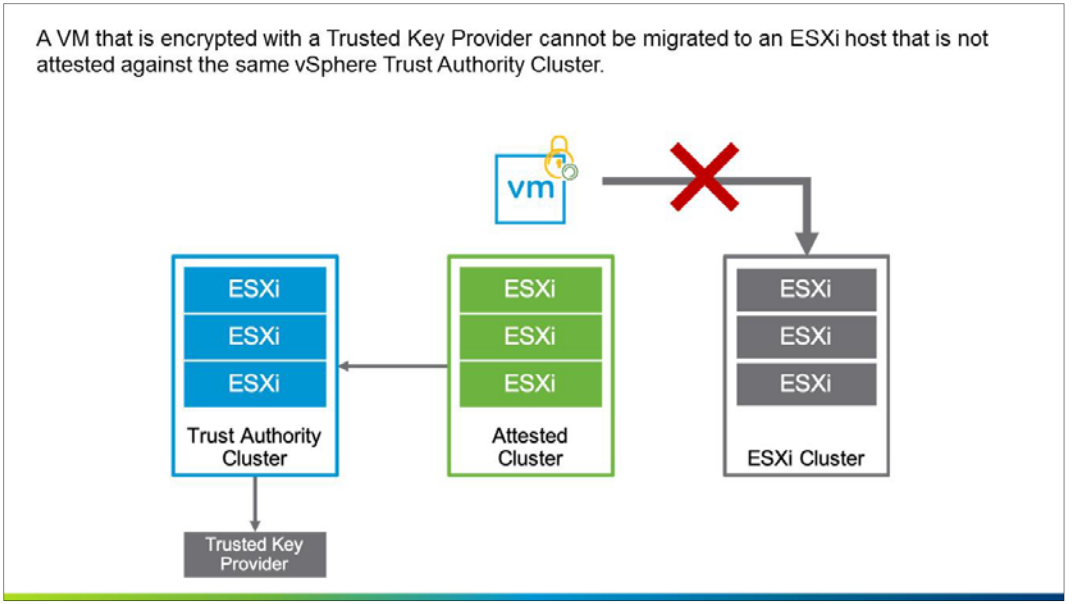
4-72 Identifying VM Encryption Key Providers



The `crypto-util dictionary describe /path/vm.vmx` command does not indicate whether the VM is encrypted with a standard or trusted key provider. However, the key provider name is present in the output so that a correlation can be made.

https://t.me/learningnets

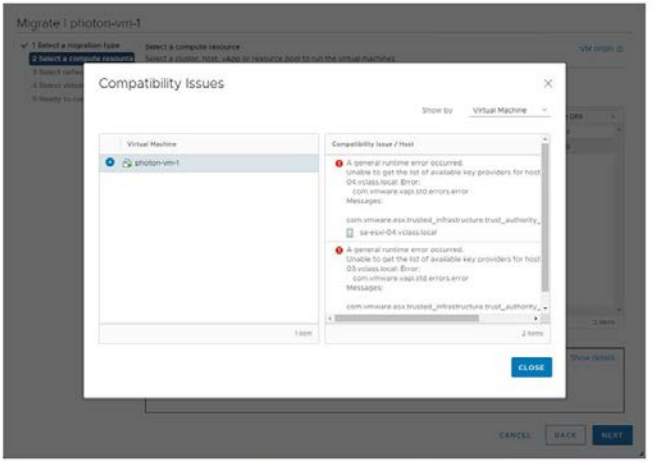
4-73 Migrating a VM Encrypted with a Trusted Key Provider (1)



https://t.me/learningnets

4-74 Migrating a VM Encrypted with a Trusted Key Provider (2)

The vSphere Client displays a compatibility issue indicating that the destination host is not trusted and attested.



https://t.me/learningnets

Labs

- Lab: Assigning a vSphere Trust Authority Administrator
- Lab: Enabling and Configuring vSphere Trust Authority
- Lab: Encrypting a VM with a Trusted Key Provider

https://t.me/learningnets

4-76 Lab 10: Assigning a vSphere Trust Authority Administrator

Assign a vSphere Trust Authority Administrator role to allow a user to configure and manage vSphere Trust Authority:

1. Assign a vSphere Trust Authority Administrator

https://t.me/learningnets

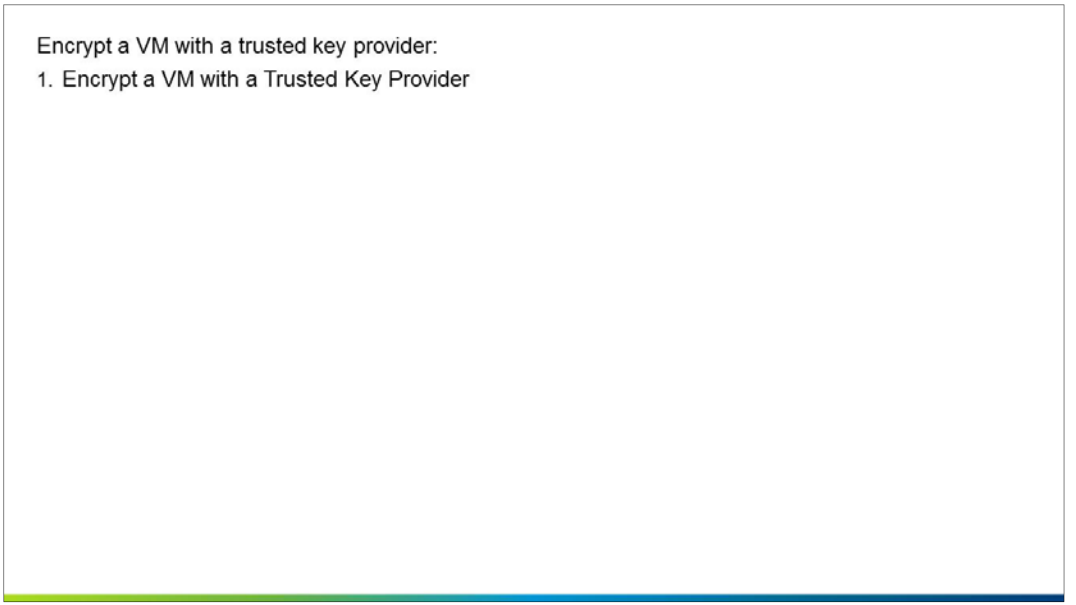
4-77 Lab 11: Enabling and Configuring vSphere Trust Authority

Enable and configure vSphere Trust Authority:

1. Preconfigure the Environment
2. Export the TPM Certificate and ESXi Image Metadata
3. Export the Trusted User Principal
4. Enable vSphere Trust Authority Services
5. Import the Trusted Host Information to the Trust Authority Cluster
6. Create a Trusted Key Provider on the Trust Authority Cluster
7. Export the Trust Authority Cluster Settings
8. Import the Trust Authority Cluster Settings into the Trusted Hosts Cluster
9. Configure the Trusted Key Provider for the Trusted Hosts Cluster

https://t.me/learningnets

4-78 Lab 12: Encrypting a VM with a Trusted Key Provider



https://t.me/learningnets

4-79 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe vSphere Trust Authority
- List the use cases of vSphere Trust Authority
- Explain the general architecture of vSphere Trust Authority
- Configure vSphere Trust Authority

https://t.me/learningnets

https://t.me/learningnets

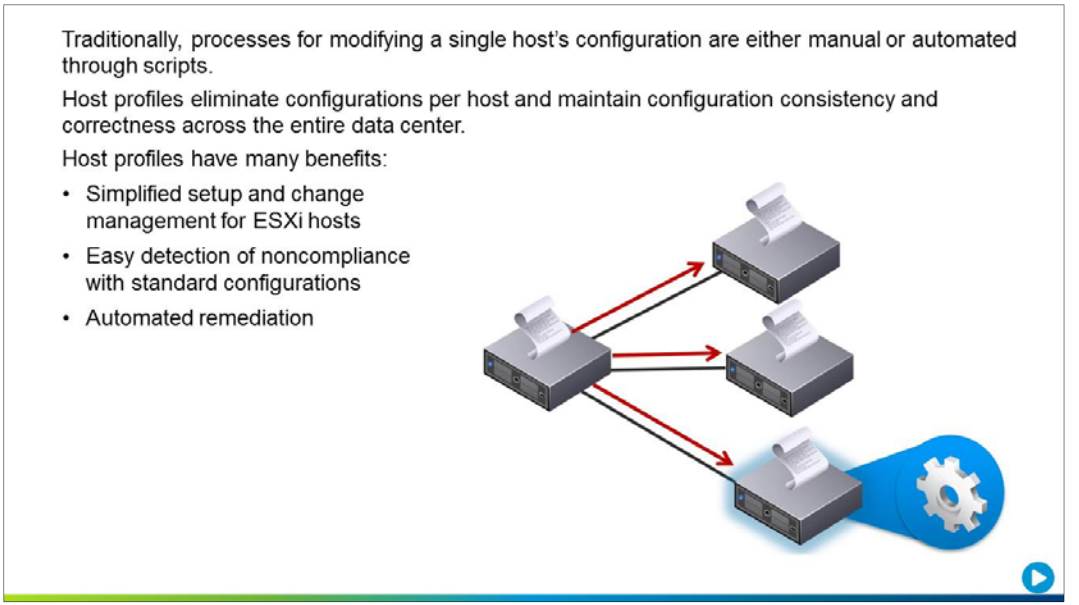
4-81 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe the benefits of implementing host profiles
- Explain how the Host Profiles feature works
- Use a host profile to manage ESXi configuration compliance
- Use host profiles to standardize ESXi configurations across multiple vCenter Server instances

https://t.me/learningnets

4-82 About Host Profiles



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, for example:

- 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 (NTP) server
- Security settings: Firewall settings, Active Directory configuration, ESXi services, and so on.
- DNS and routing: DNS server, IP default gateway, and so on.

https://t.me/learningnets

Traditionally you would modify the configuration of a single host:

- Manually, using the vSphere Client or the command prompt
- Automatically, using scripts

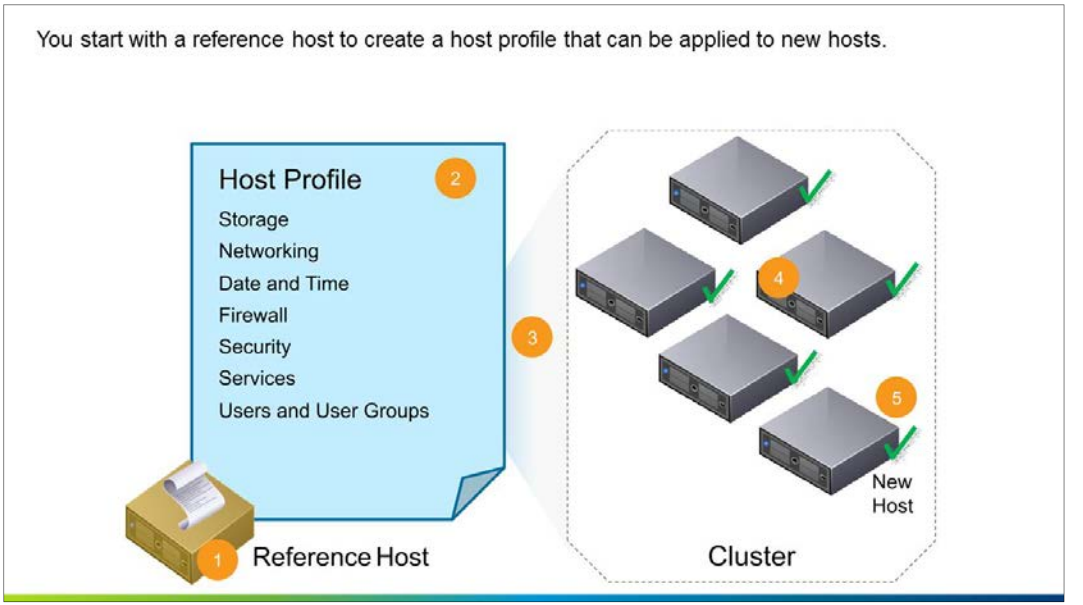
Using the Host Profiles feature, you can export configuration settings from a master reference host and save them as a portable set of policies, called a host profile. You can use this profile to quickly configure other hosts in the data center. Configuring hosts with this method reduces the setup time 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

4-83 Implementing 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.

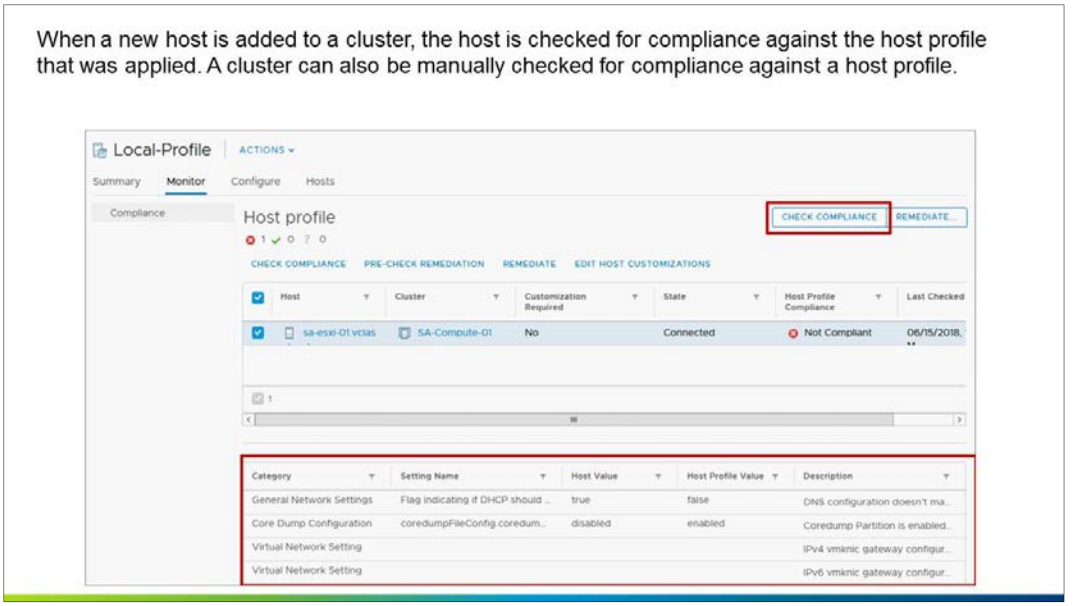
https://t.me/learningnets

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

4-84 Monitoring for Compliance



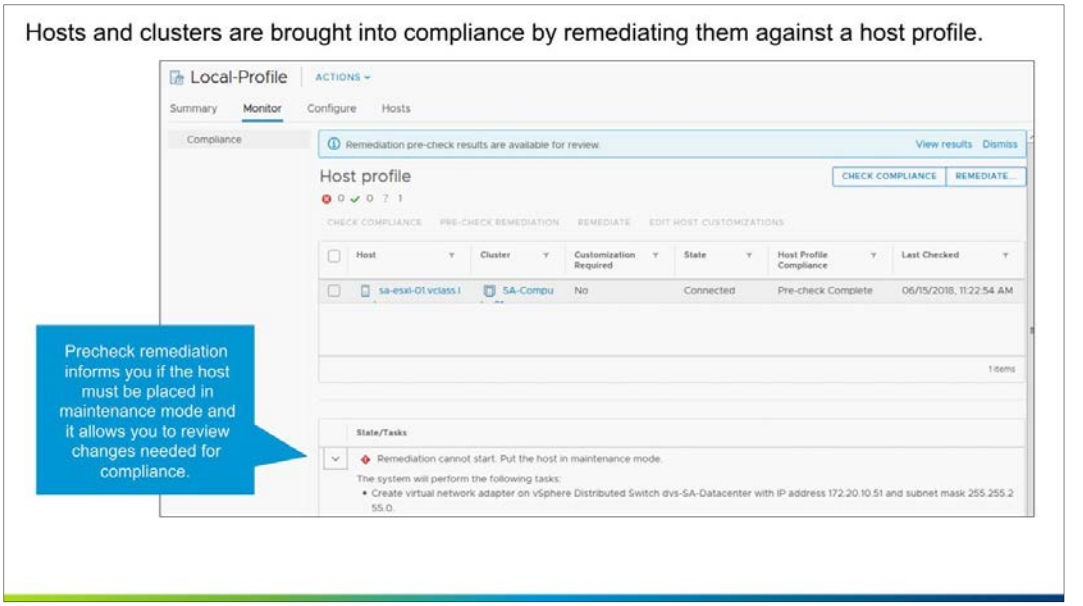
After the host profile is created and associated with a set of hosts or clusters, you can verify the compliance status in various places in the vSphere Client:

- 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).

You can also check compliance for individual components or settings. These built-in cluster compliance checks are useful for vSphere DRS and vSphere DPM.

https://t.me/learningnets

4-85 Applying Host Profiles



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. If this list of actions shows parameters that you do not want, you can cancel applying the host profile.

Certain host profile policy configurations require that the host is 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 into maintenance mode at remediation. For other cases, the remediation process stops if the host is not placed into maintenance mode before remediation.

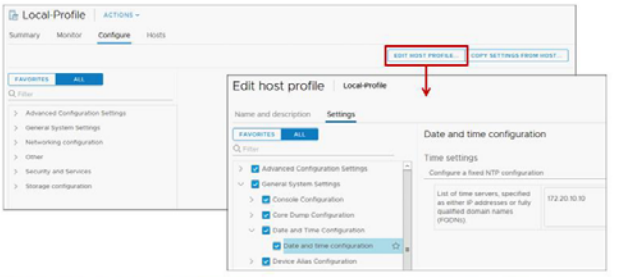
https://t.me/learningnets

4-86 Host Customization

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 that are provisioned with 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. An answer file is equivalent to host customization.

You can check the status of a host customization. If the host customization has all the required user input answers, the status is Complete. If the host customization is missing some of the

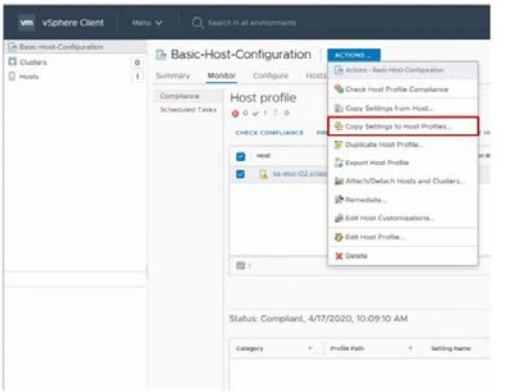
https://t.me/learningnets

required user input answers, 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

4-87 Managing Host Customizations (1)

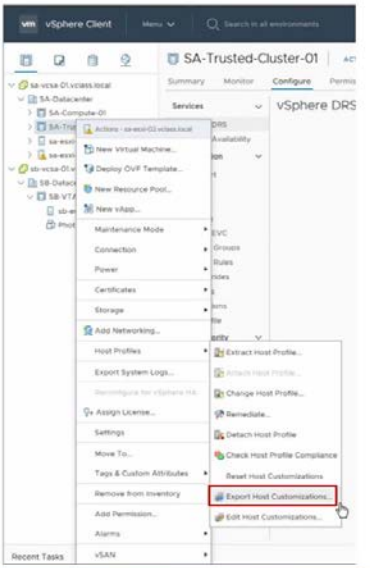
You can copy settings across host profiles. Using this functionality, you can maintain a golden master host profile for global settings that are common across all host profiles.



https://t.me/learningnets

4-88 Managing Host Customizations (2)

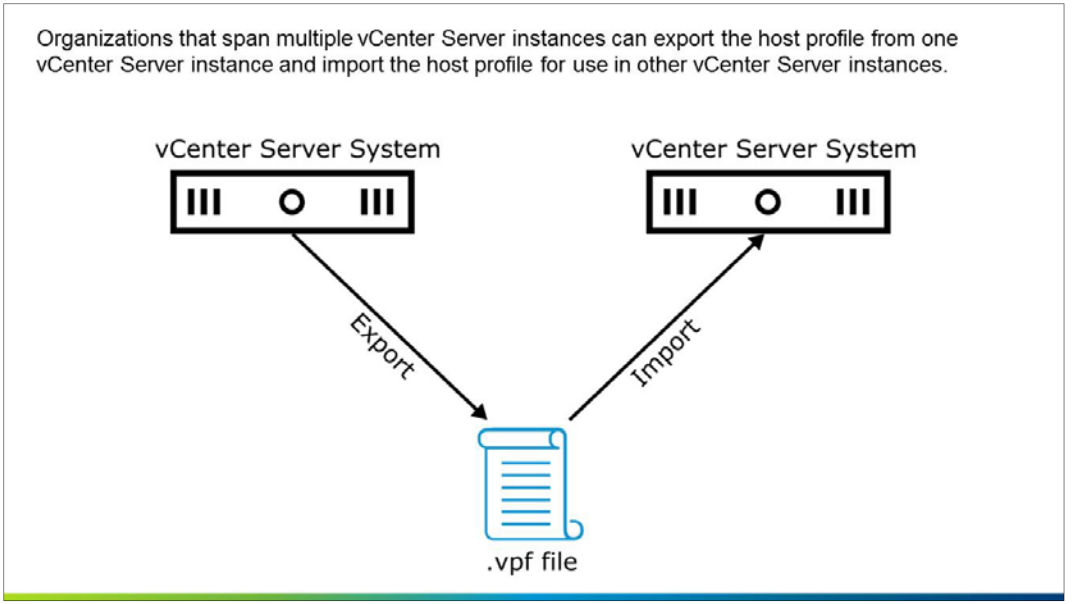
You can edit host customizations by exporting the data to a CSV file and then editing the CSV file. With this functionality, you can create host customizations offline.



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

4-89 Standardization Across Multiple 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 i a security measure that stops passwords from being exported in plain text. After the profile is im are prompted to reenter values for the password and the password is applied to a host.

https://t.me/learningnets

4-90 Lab 13: Using Host Profiles

Use host profiles to manage host configuration compliance:

1. Preconfigure ESXi Hosts
2. Create and Export a Host Profile
3. Import a Host Profile
4. Duplicate and Edit a Host Profile
5. Attach an ESXi Host to a Host Profile
6. Run an Initial Compliance Check
7. Introduce a Configuration Drift
8. Run a Compliance Check and Remediate the Configuration Drift
9. Detach the Host Profile

https://t.me/learningnets

4-91 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe the benefits of implementing host profiles
- Explain how the Host Profiles feature works
- Use a host profile to manage ESXi configuration compliance
- Use host profiles to standardize ESXi configurations across multiple vCenter Server instances

https://t.me/learningnets

https://t.me/learningnets

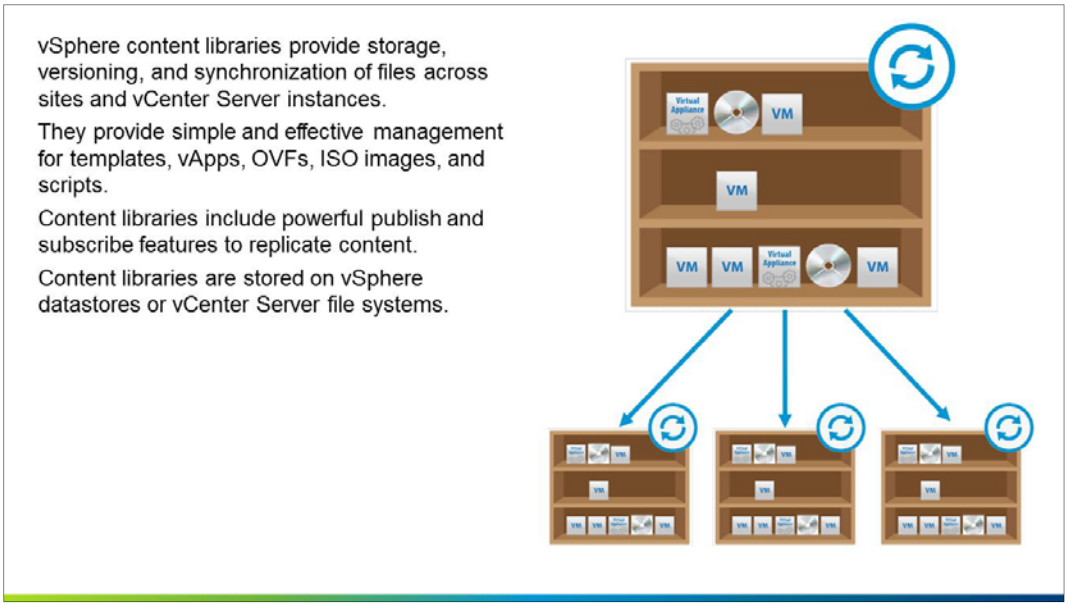
4-93 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the purpose of content libraries in a vSphere environment
- Identify vSphere requirements for content libraries
- Create a local content library
- Subscribe to a published content library
- Deploy VMs from a content library

https://t.me/learningnets

About Content Libraries



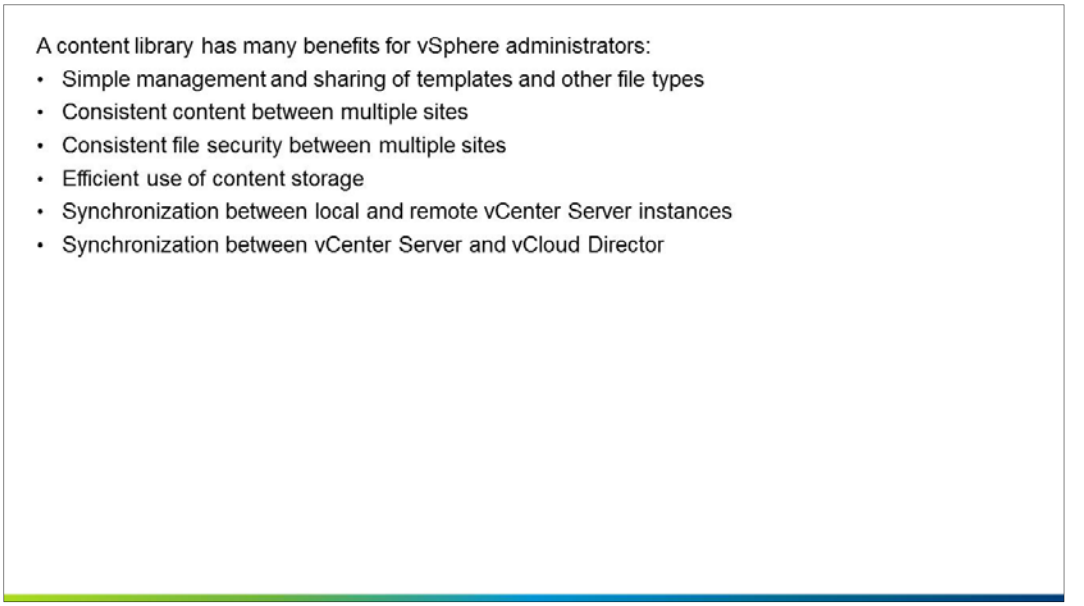
Content libraries provide a simple and effective way to manage multiple content types in a small vSphere environment or across several vCenter Server instances. A content library can be synchronized across remotely located sites and against remotely located vCenter Server instances.

Content libraries allow subscriptions to published data sets, resulting in consistent data across multiple sites. A vSphere administrator controls publication and maintains consistent data and security on that data. Data can be synchronized automatically across sites or on demand when needed.

The type of storage that is used for content libraries depends on the data use in the library.

https://t.me/learningnets

4-95 Content Library Benefits



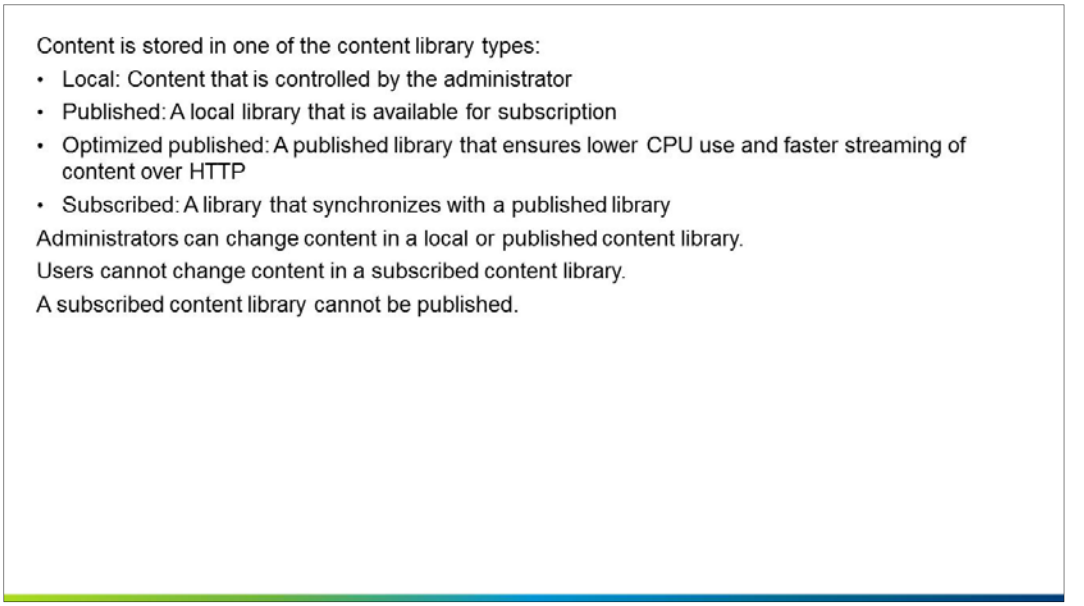
Storage and consistency is a key reason to install and use a content library. Sharing content and ensuring that this content is kept up to date is a major task. For example, you might create a central content library to store the master copies of OVF templates, ISO images, or other file types for a main vCenter Server instance.

You can publish this content library so that other libraries that are located across the world can 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 vCenter Single Sign-On or Active Directory in the current release of vCenter Server.

https://t.me/learningnets

4-96 Content Library Types



Use an optimized published content library when the subscribed libraries reside on a remote vCenter Server system and Enhanced Linked Mode is not used.

You cannot deploy VMs from templates in an optimized library. You must first create a subscribed library and then deploy VMs from the subscribed library.

A local library is the simplest form of library. A local library is available for use in data centers that are objects to the local vCenter. A published library is a local library that is available for subscription. Using the vCenter Server database, the content library tracks version changes. 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.

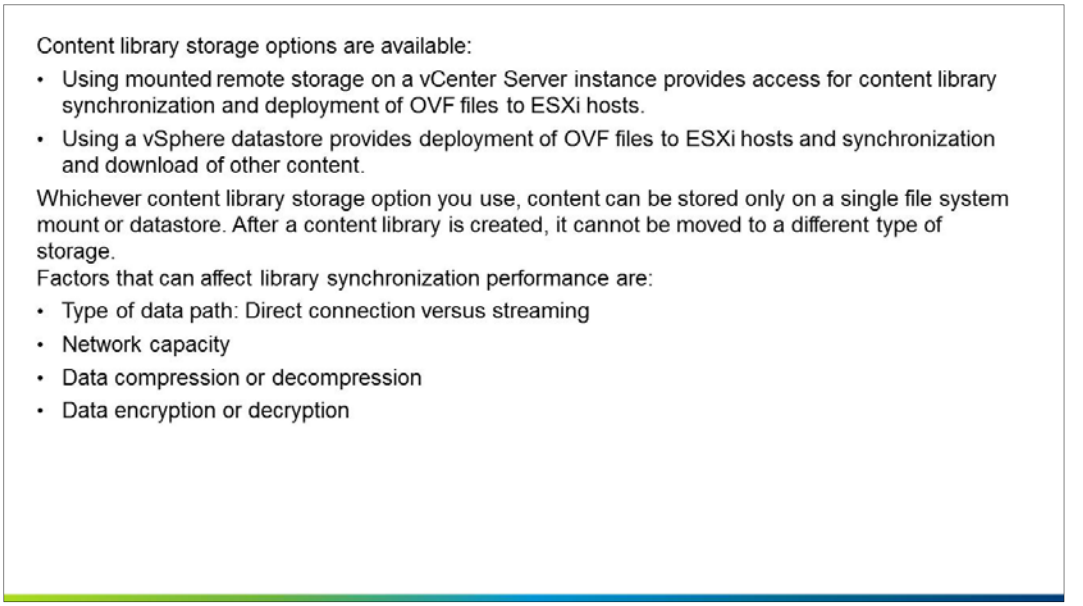
An optimized published library is a library that is optimized to ensure lower CPU use and faster streaming of the content over HTTP. Use this library as a main content depot for your subscribed libraries. You cannot deploy VMs from an optimized library. You must first configure a

https://t.me/learningnets

subscribed library to subscribe to the optimized published library and then deploy VM from the subscribed library.

https://t.me/learningnets

4-97 Storing Content in Content Libraries



Content library storage can be a datastore in your vSphere inventory or it can be a local file system on a vCenter Server system.

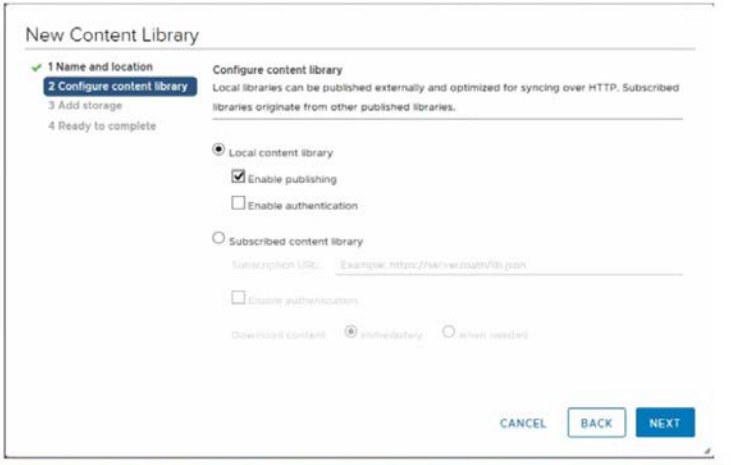
Content library storage can also be a remote storage location. If you use a Windows-based vCenter Server system, the storage can be an SMB server. If you use vCenter Server Appliance, the storage can be an NFS server.

For more information about how to increase the performance of content libraries, see the VMware blog article "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

Creating Content Libraries

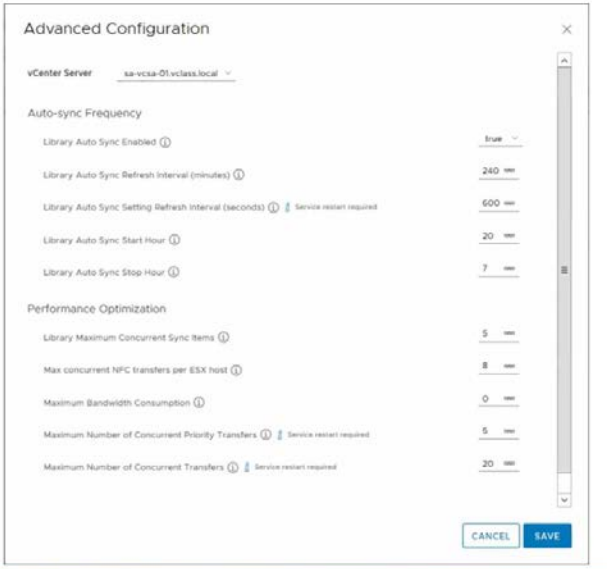
When you create a content library, you select the content library type.



https://t.me/learningnets

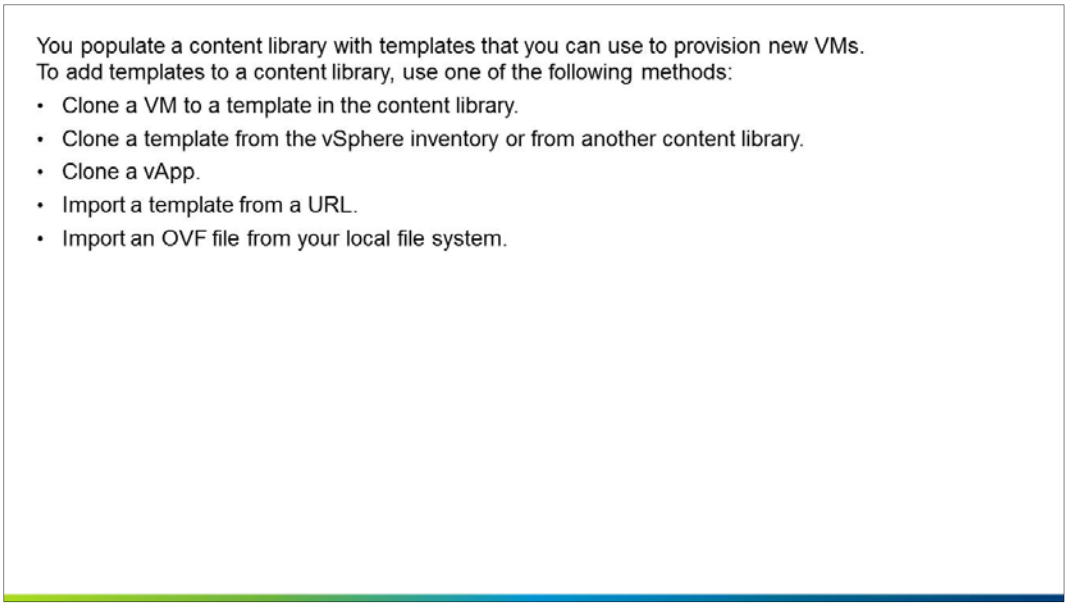
4-99 Advanced Configuration

Use the Advanced Configuration page to control and optimize how a published library stores and synchronizes content.



https://t.me/learningnets

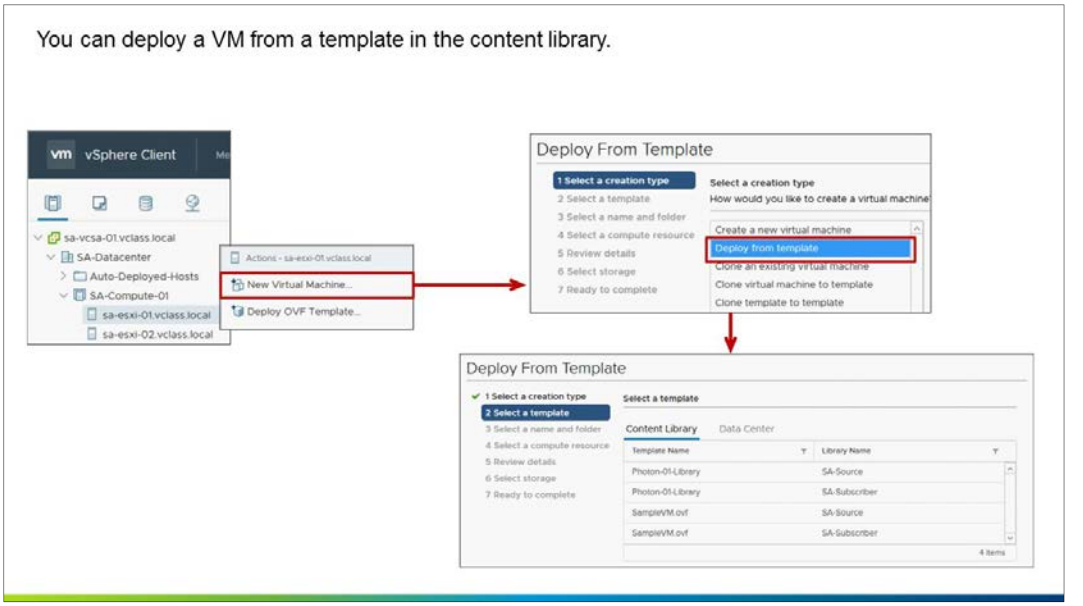
4-100 Populating Content Libraries



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

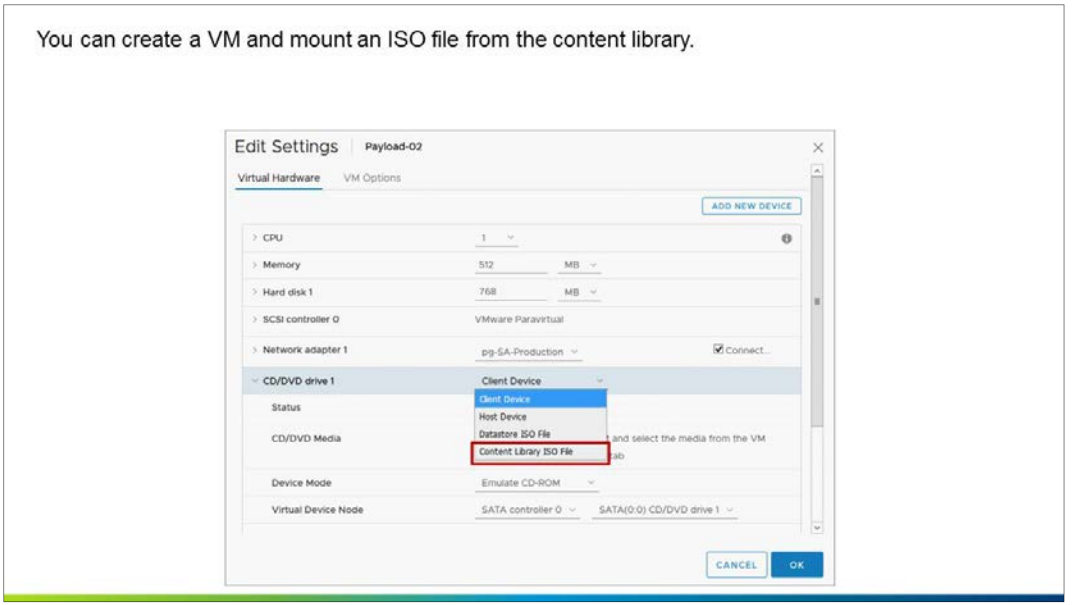
4-101 Using the Content Library to Deploy VMs (1)



You can use a VM template from a content library to deploy a VM to a host or a cluster in your vSphere inventory. You can also apply a guest OS customization specification to the VM.

https://t.me/learningnets

4-102 Using the Content Library to Deploy VMs (2)

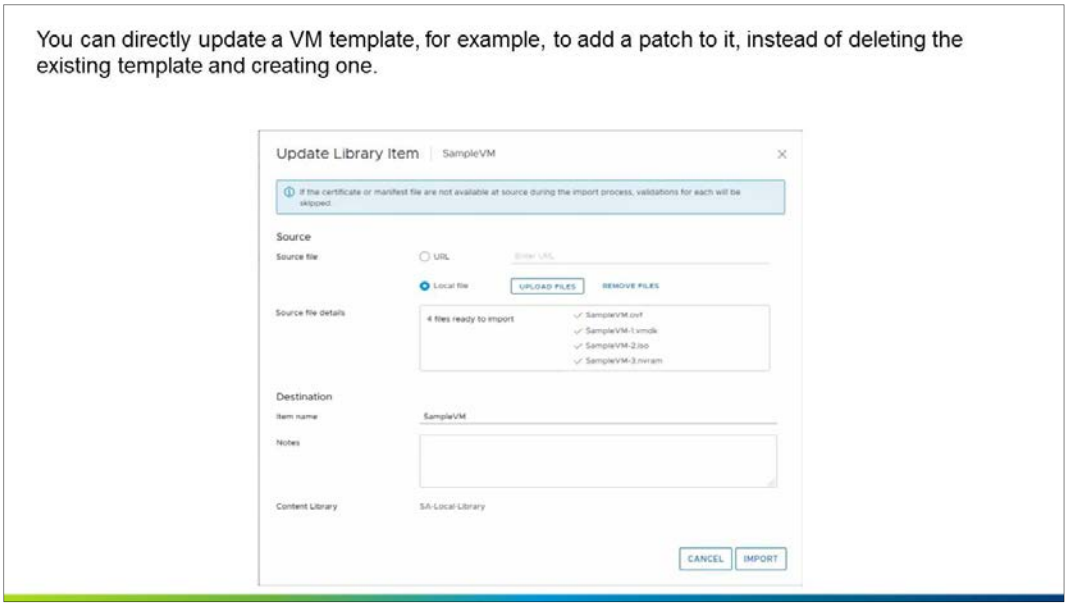


To install a guest OS and its applications on a new VM, you can connect the CD/DVD device to an ISO file that is stored in a content library.

A VM can only use an ISO image that is stored in a content library if its ESXi host can access the datastore that originally stored the ISO image.

https://t.me/learningnets

4-103 Updating Content Library Items

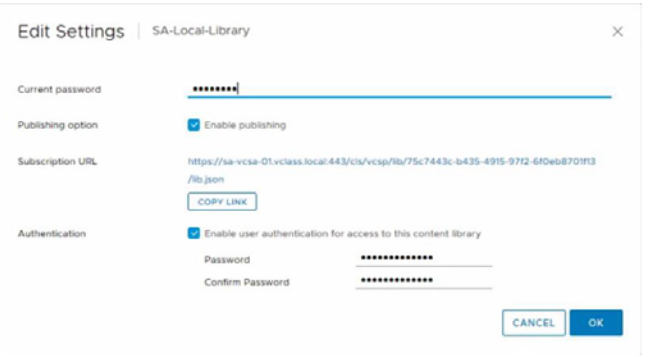


Managing and keeping your virtual environment up to date might require that you update the content of a library item.

https://t.me/learningnets

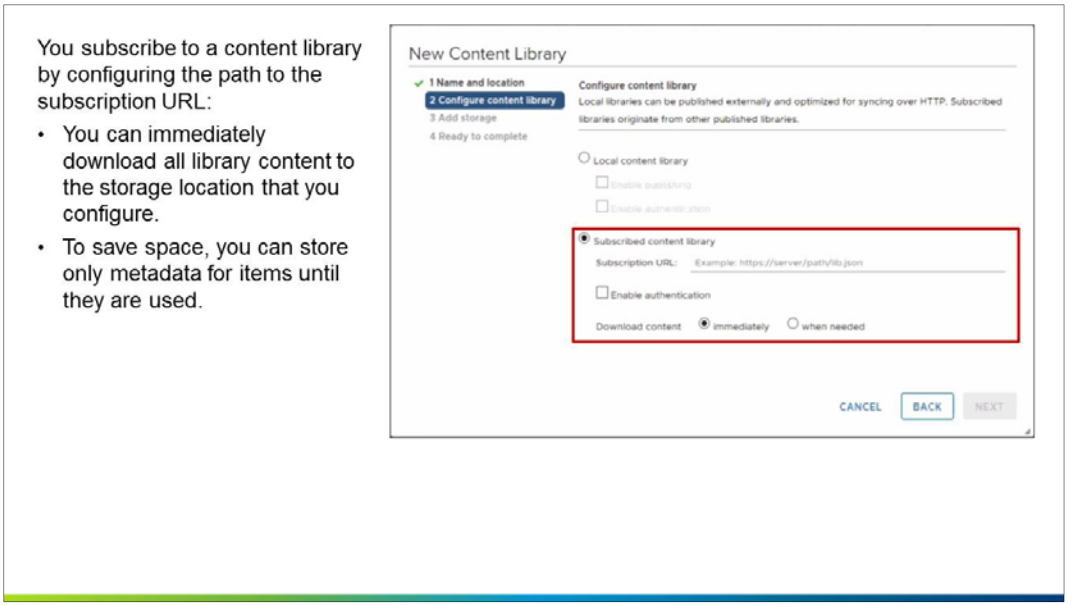
4-104 Publishing Content Libraries for External Use

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

4-105 Content Library Subscriptions

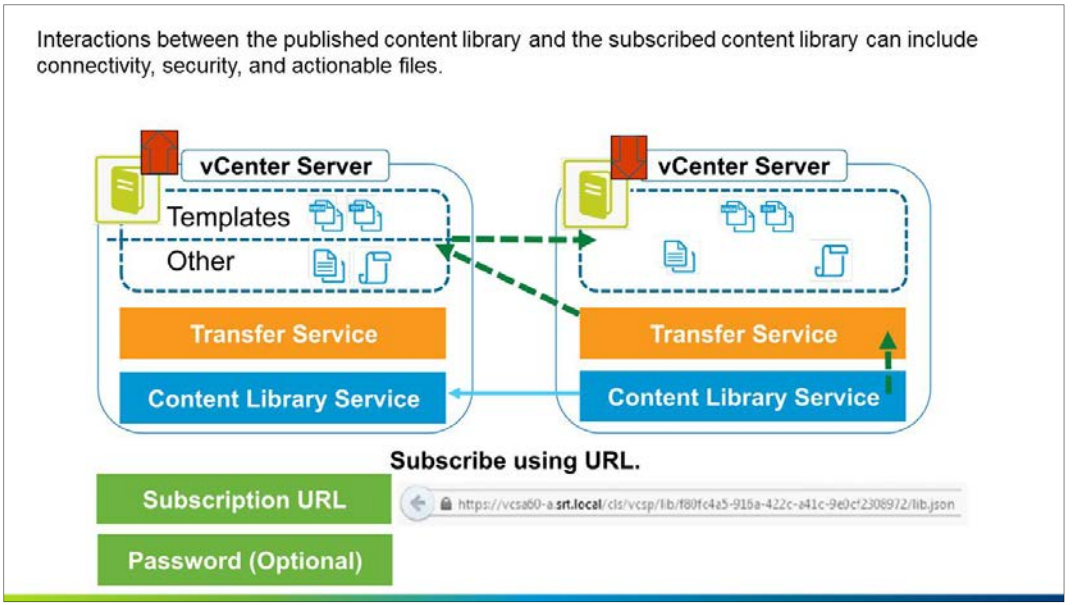


When a content library subscription is created, the administrator selects how the content synchronizes with the published content library. Content can be downloaded immediately if space for the content is not a concern.

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

4-106 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 VMs 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 OS on a new VM.

The administrator can publish the library. Publishing generates a URL that points to the `lib.json` file of that library. The administrator can enable authentication and assign a password.

https://t.me/learningnets

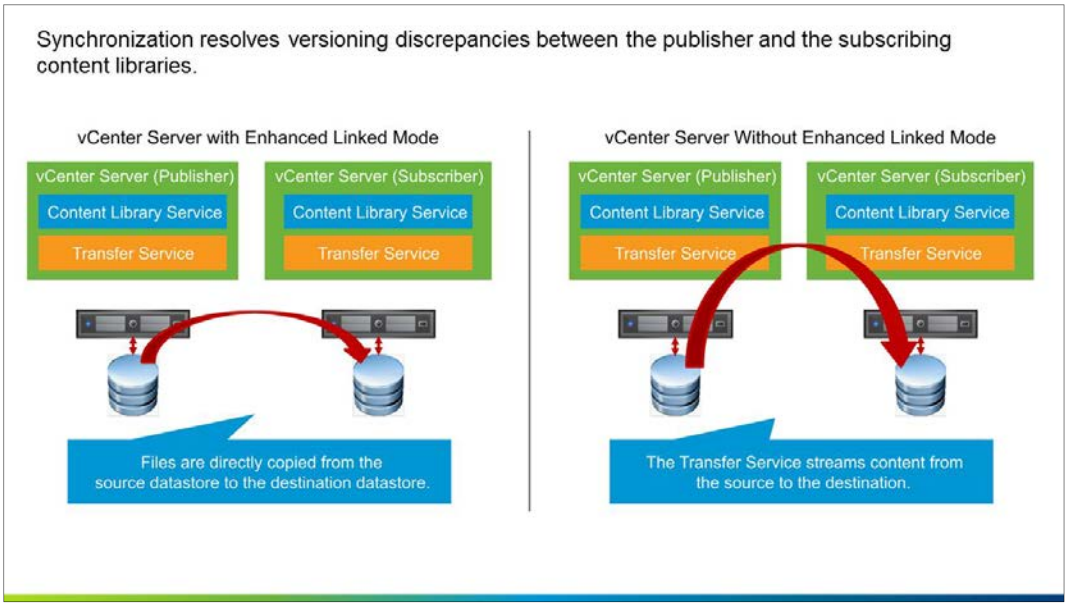
By default, the user name is `vcsp`. You cannot set a user name during the creation of the content library, but you can change the user name in the content library properties after creation.

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.

For more information, see *vSphere Virtual Machine Administration* at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-virtual-machine-admin-guide.pdf.

https://t.me/learningnets

4-107 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 that is 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 required. If a discrepancy is found, the version number of each item in the library is checked. If 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

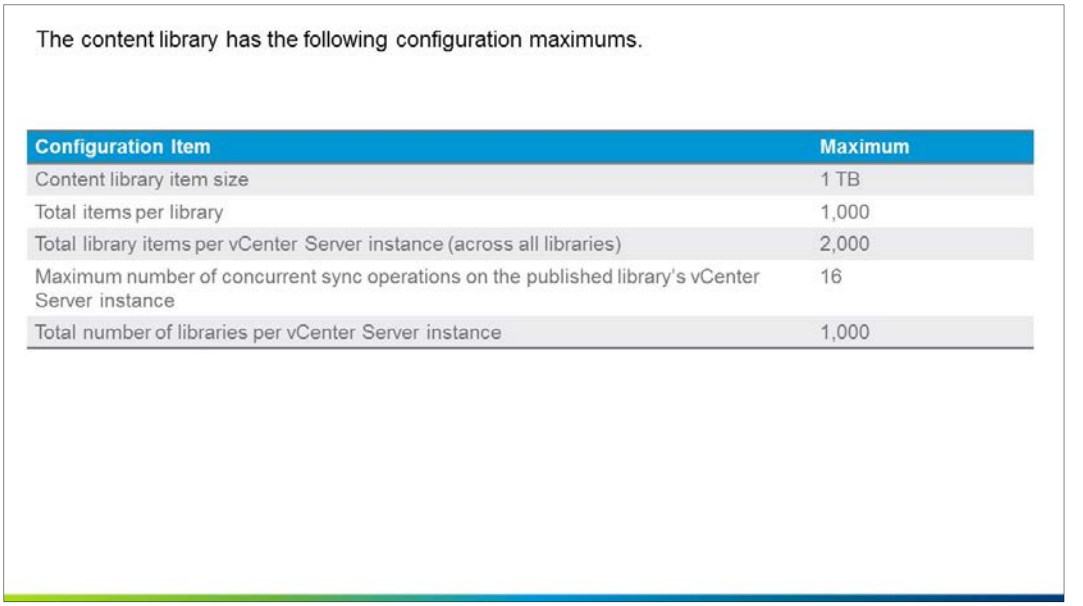
https://t.me/learningnets

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 scenarios are possible, based on the storage configuration:

- Both published and subscribed libraries reside on a datastore.
- Both published and subscribed libraries reside on an NFS file system mounted on the vCenter Server instances.
- The published library resides on an NFS file system, whereas the subscribed library resides on a datastore.

https://t.me/learningnets

4-108 Content Library Maximums



The content library can be supported by a datastore or stored to available storage on vCenter Server.

Regardless of the option selected, the content library can be supported 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 the vSphere Client.

https://t.me/learningnets

4-109 Lab 14: Creating a Content Library

Create a multisite content library:

1. Create a Local Content Library
2. Upload Data to the Content Library
3. Create a Subscribed Content Library
4. Create a Subscription for VM Templates
5. Clone a Template to the Local Library
6. Synchronize the Content Libraries
7. Deploy a VM from the Subscribed Content Library
8. Clean Up for the Next Lab

https://t.me/learningnets

4-110 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the purpose of content libraries in a vSphere environment
- Identify vSphere requirements for content libraries
- Create a local content library
- Subscribe to a published content library
- Deploy VMs from a content library

https://t.me/learningnets

https://t.me/learningnets

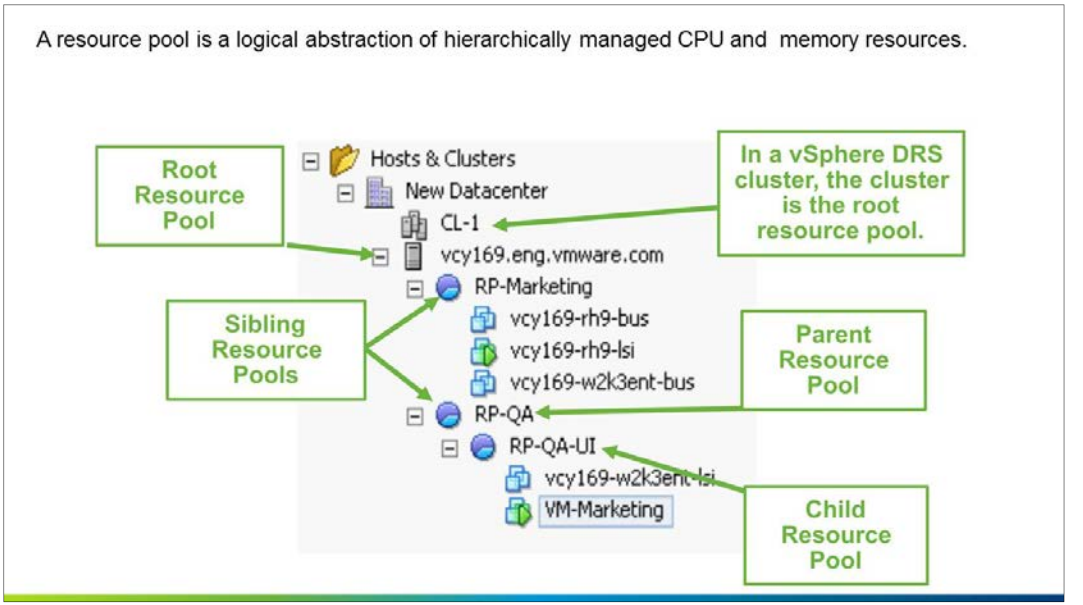
4-112 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Create a resource pool
- Set resource pool attributes
- Describe expandable reservations
- Describe scalable shares

https://t.me/learningnets

4-113 About Resource Pools



A resource pool is a logical abstraction for managing resources. Resource pools can be grouped into hierarchies and used to hierarchically partition available CPU and memory resources. In the example on the slide, RP-QA is the parent resource pool for RP-QA-UI. RP-Marketing and RP-QA are siblings.

With a resource pool, you can divide and allocate resources to VMs and other resource pools. You can control the aggregate CPU and memory resources of the compute resource. The compute resource can be either a standalone host or a vSphere DRS cluster. Resource pools are also used to delegate privileges to other users and groups.

The topmost resource pool is called the root resource pool. Each standalone host and each vSphere DRS cluster has an (invisible) root resource pool that groups the resources of that host or cluster. The root resource pool does not appear because the resources of the host (or cluster) and the root resource pool are always the same.

https://t.me/learningnets

4-114 Resource Pool Attributes

You can create a child resource pool of any ESXi host, resource pool, or vSphere DRS cluster:

- Shares: Low, Normal, High, Custom
- Reservation: In MHz, GHz, MB, or GB
- Reservation Type:
  - **Expandable** check box selected: VMs and subpools can draw from this pool's parent.
  - **Expandable** check box deselected: VMs and subpools can draw only from this pool, even if its parent has free resources.
- Limit:
  - In MHz, GHz, MB, or GB
  - Unlimited access, by default, up to maximum amount of accessible resources

Like VMs, resource pools have reservation, limit, and share values for CPU and memory resources.

Resource pools also have a setting that enables reservations to be expanded. With this expandable reservation attribute, a resource pool that cannot satisfy a reservation request can search through its hierarchy to find unreserved capacity to satisfy the reservation request.

https://t.me/learningnets

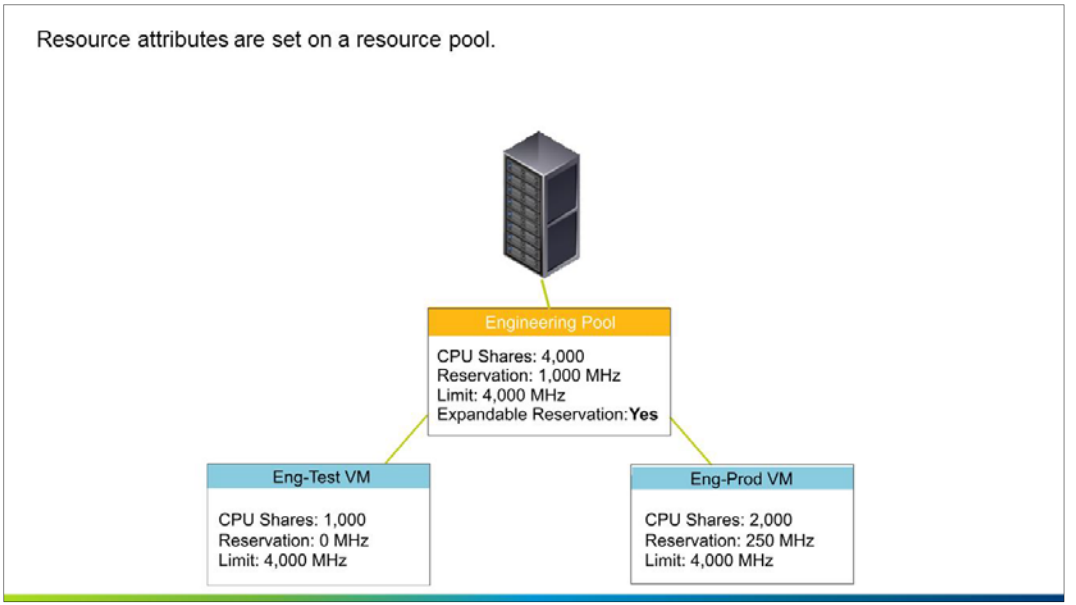
4-115 Reasons to Use Resource Pools

With resource pools, you can delegate control over resources of a standalone host or a vSphere DRS cluster. Using resource pools can result in the following benefits:

- Flexible hierarchical organization: Add, remove, or reorganize resource pools or change resource allocations as needed.
- Isolation between pools, sharing in pools: Top-level administrators can make a pool of resources available to a department-level administrator.
- Access control and delegation: VM creation and management are performed within the boundaries of the resources to which the resource pool is entitled. Delegation is usually done with permission settings.
- Separation of resources from hardware: If you use vSphere DRS clusters, the resources of all hosts are always assigned to the cluster.
- Management of sets of VMs running a multitier service: Group VMs for a multitier service in a resource pool.

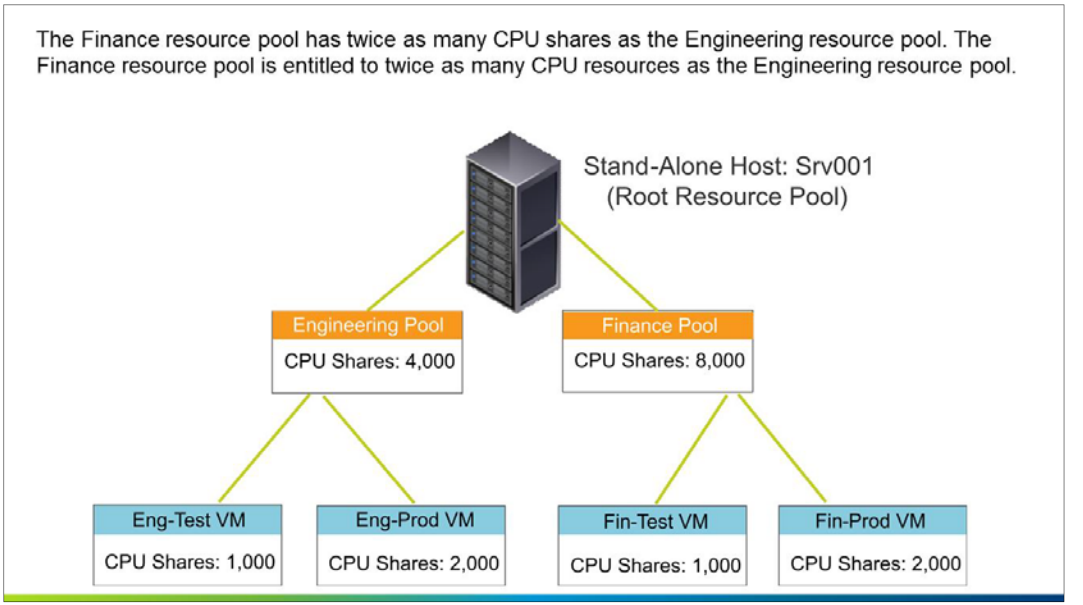
https://t.me/learningnets

4-116 Resource Pool Example



https://t.me/learningnets

4-117 Resource Pool Example: CPU Shares



Resource pools can be organized hierarchically. The root resource pool is the topmost resource pool. The root resource pool includes the sum for all CPUs (in megahertz) and the sum of all the installed RAM (in megabytes) available in the compute environment (standalone host or cluster).

On the slide, the root resource pool is a standalone host named Svr001. It has 12,000 MHz of CPU and 4 GB of RAM, available for use by other resource pools or VMs.

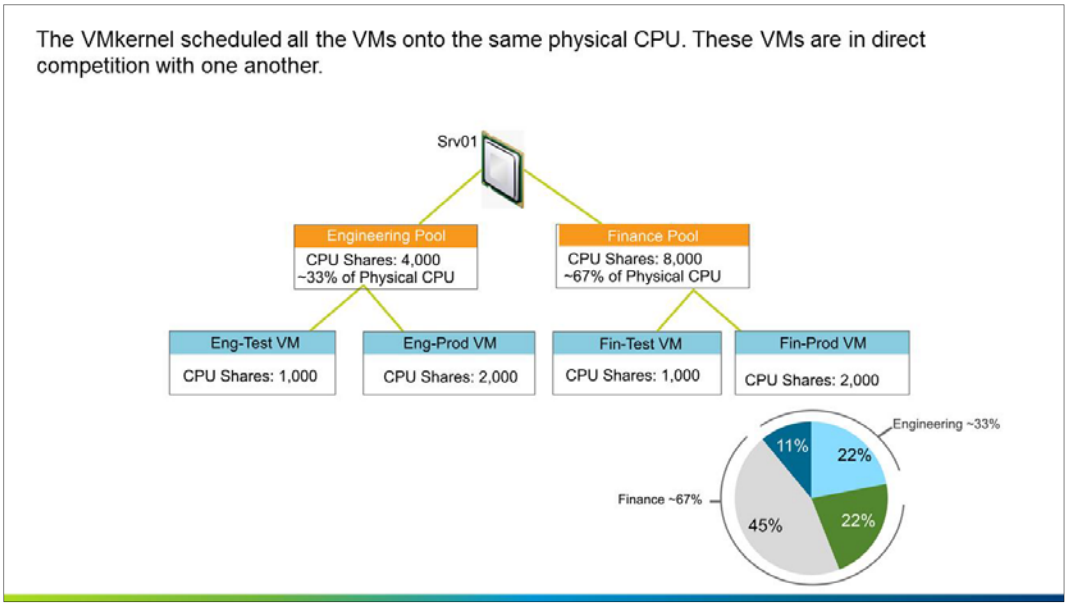
Except for the root resource pool, every resource pool has a parent resource pool. A resource pool might contain child resource pools or only VMs that are powered on in the pool.

A child resource pool is used to allocate resources from the parent resource pool for the child's consumers. Administrative control can also be delegated to individuals or organizations. A child resource pool cannot exceed the capacity of the parent resource pool. By creating a child pool, you reserve resources from the parent pool, whether or not VMs in the child pool are powered on.

Shares specify the relative priority or importance of either a resource pool or VM. If a resource pool has twice as many shares of a resource as another resource pool, it is entitled to consume twice as much of that resource. The same logic applies to VMs.

https://t.me/learningnets

4-118 Resource Pool Example: CPU Contention



The example on the slide uses approximations to explain how the number of shares affects the amount of CPU allocated to a VM.

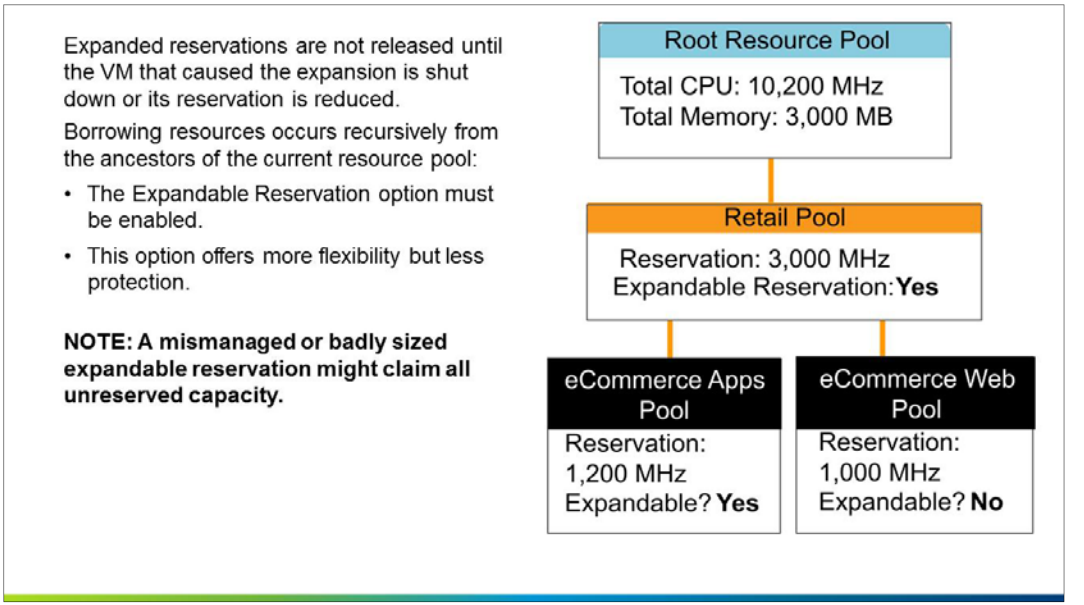
All VMs are running on the same physical CPU, Srv01.

The Engineering pool gets 33 percent of the CPU and splits its allotment between VMs Eng-Test and Eng-Prod. Likewise, the Finance pool gets 67 percent of the CPU and splits its allotment between VMs Fin-Test and Fin-Prod. A VM's resource settings are constrained by the resources of the resource pool to which the VM belongs.

The VM Eng-Test gets approximately 33 percent of the CPU allocation of the Engineering resource pool [1,000/(1,000 + 2,000)]. This figure is equal to about 11 percent of the physical CPU (33 percent of 33 percent equals approximately 11 percent). Each of the VMs gets a percentage of the physical CPU allocated to its resource pool based on its individual share allocation.

https://t.me/learningnets

4-119 About Expandable Reservations



In the example, eCommerce Apps is an expandable child resource pool.

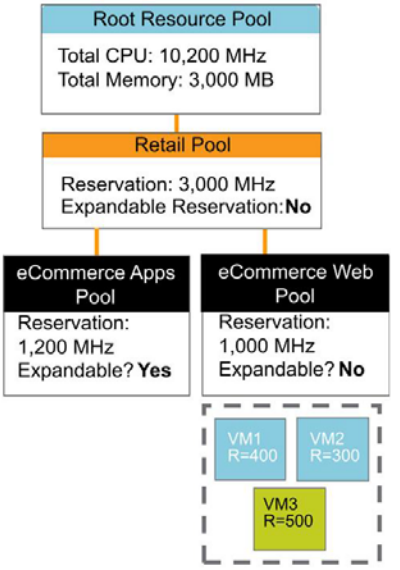
The reservation of a child resource pool cannot exceed that of its parent. The search for unused resources goes through the ancestry of the root resource pool or to the first resource pool where the Expandable Reservation Type is not enabled.

Use expandable reservation carefully. A single child resource pool can use all its parent's available resources, leaving nothing directly available for other child resource pools. This issue can occur if VMs are improperly moved around to get them to start based on resource reservations.

You might want to disable the **Expandable** Reservation Type when you configure a fixed amount of resources to a group. For example, an IT administrator whose customers are different organizations in the company that have paid for a fixed amount of CPU and memory resources might want to disable the **Expandable** Reservation Type.

https://t.me/learningnets

4-120 Example of Expandable Reservation (1)



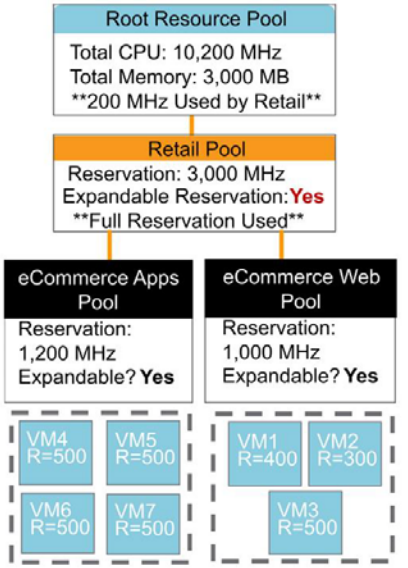
eCommerce resource pools reserve 2,200 MHz of the 3,000 MHz that the Retail pool has reserved.

You power on VMs in the eCommerce Web pool.
With the Expandable Reservation option disabled on the eCommerce Web pool, VM3 cannot start with a reservation of 500 MHz:

- Lower the VM reservation.
- Enable **Expandable** as the Reservation Type.
- Increase the eCommerce Web pool's reservation.

https://t.me/learningnets

4-121 Example of Expandable Reservation (2)

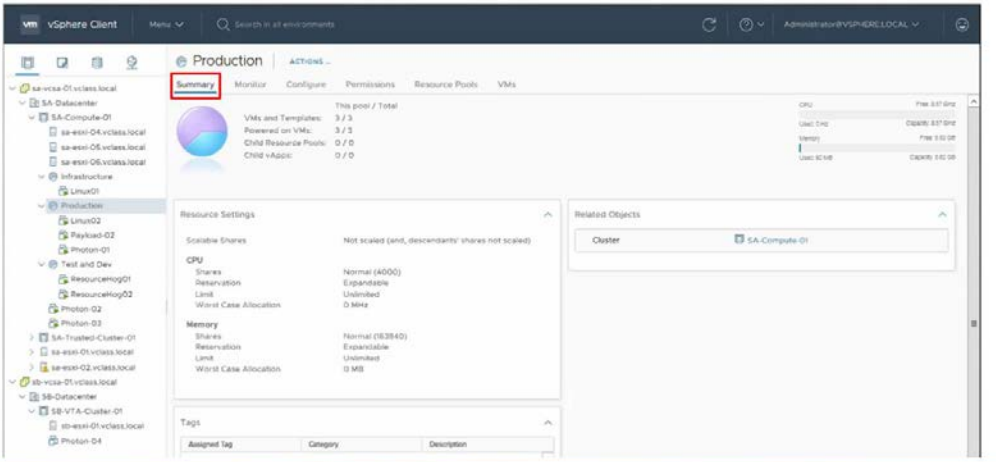


The **Expandable** Reservation Type is enabled on the eCommerce Web pool.
The system considers the resources available in the child resource pool and its direct parent resource pool.
The VM's reservation is charged against the reservation for eCommerce Web pool.
eCommerce Web pool's reservation is charged against the reservation for Retail.

https://t.me/learningnets

4-122 Resource Pool Summary Tab

The resource pool's **Summary** tab displays information that applies to the host machine and its resources.



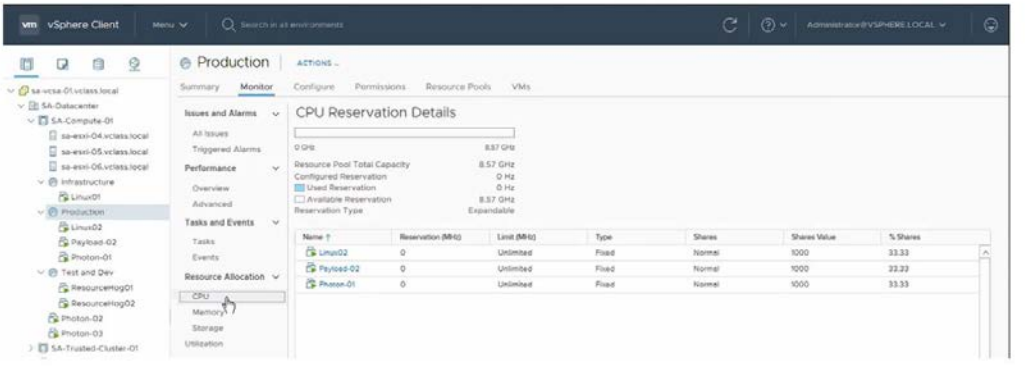
The **Summary** tab provides important information:

- The Resource Settings pane displays CPU and Memory share settings.
- The Resource Consumers pane displays the number of VMs, number of powered-on VMs, and child resource pools that are in the selected resource pool.
- The Tags pane displays tags assigned to objects that reside in the resource pool.

https://t.me/learningnets

4-123 Resource Reservation Details

In the vSphere Client, you can view information about a resource pool's CPU, memory, and storage resources.

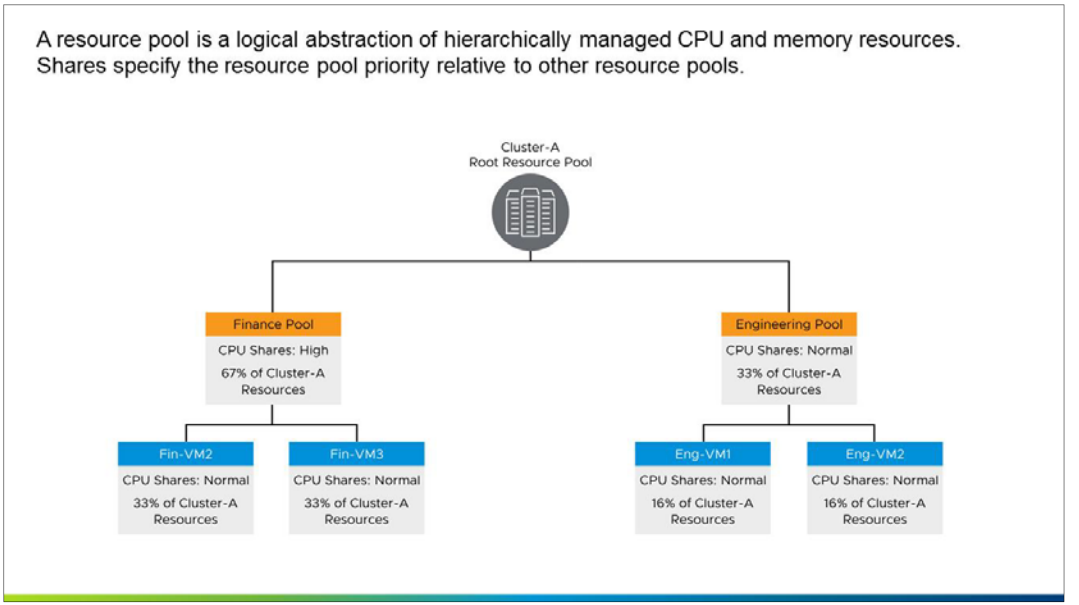


For the resource pool's CPU and memory resources, the following information is provided:

- Amount of CPU and memory reservation that is configured
- Type of reservation used (expandable or fixed)
- Amount of reservation in use by VMs and child pools
- Amount of CPU and memory that is available to be reserved

https://t.me/learningnets

4-124 Resource Pool Shares



Resource pools can be organized hierarchically. The root resource pool is the topmost resource pool. The root resource pool includes the sum for all CPUs (in megahertz) and the sum of all the installed RAM (in megabytes) available in the compute environment (standalone host or cluster).

A child resource pool is used to allocate resources from the parent resource pool for the child's consumers. A child resource pool cannot exceed the capacity of the parent resource pool. By creating a child pool, you reserve resources from the parent pool. In the example, the root resource pool is Cluster-A. Cluster-A has two child resource pools: the Engineering pool and the Finance pool.

Shares specify the relative priority or importance of a resource pool. In the example, you want the finance VMs to have twice as much processor time as the engineering VMs. You can configure the Finance pool with twice as many CPU shares as the Engineering pool.

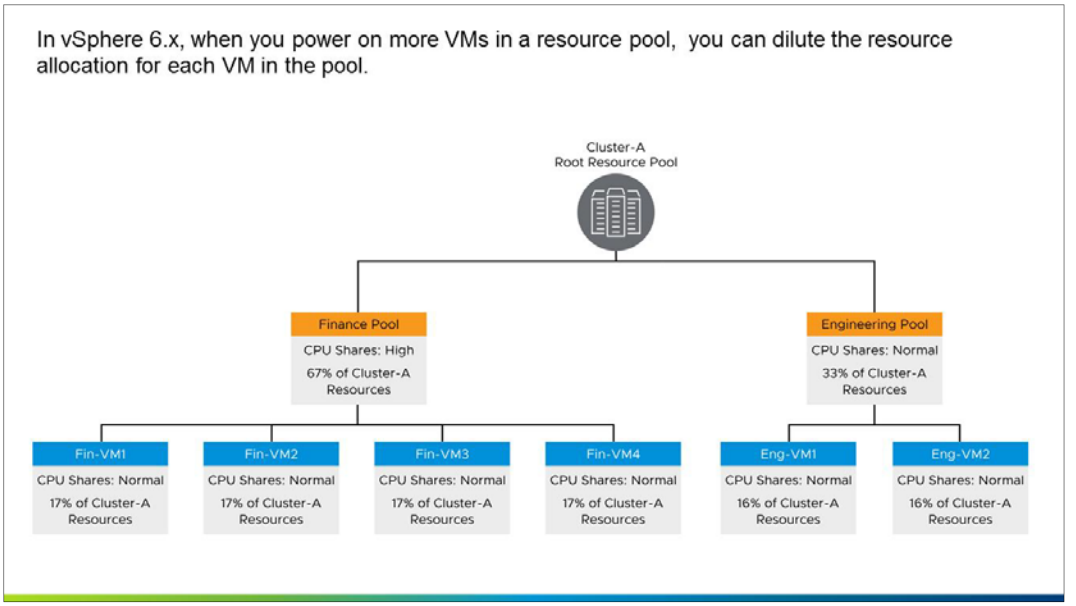
You can also specify shares per VM and distribute the resource allocation within the resource pool. VM shares are relative only to other powered-on VMs within the same resource pool. In the example, all VMs have an equal number of shares. However, the Finance VMs are entitled to

https://t.me/learningnets

about 67% of the cluster processor time. Each Finance VM gets about 33% of the entire cluster CPU resources, or twice as much as for engineering VMs.

https://t.me/learningnets

4-125 Diluted Shares

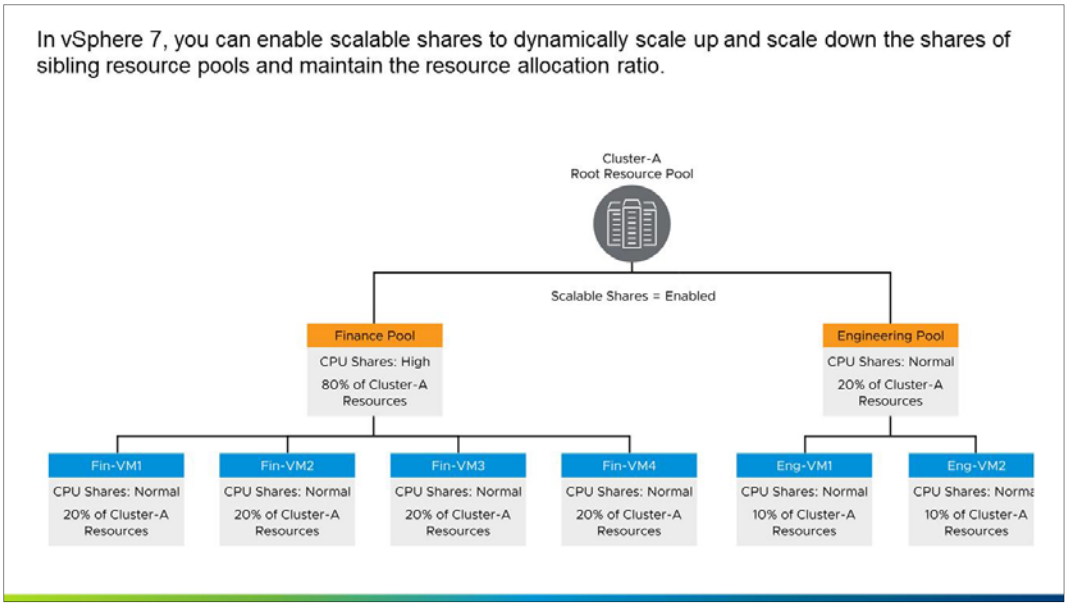


When you power on additional VMs in a resource pool, the resource allocation per VM is recalculated based on the value configured for shares. In the example, two more VMs are powered on in the Finance pool. All VMs in the cluster have an equal number of shares. Therefore, the resources are divided by four.

Because the finance VMs are entitled to about 68% of the cluster processor time, each finance VM gets about 17% of the entire cluster CPU resources. The value of shares within the Finance pool is diluted and is no longer twice as much as the engineering VMs.

https://t.me/learningnets

4-126 About Scalable Shares (1)



When you enable scalable shares on a resource pool, all child resource pools maintain the original ratio of shares configured in the parent pool.

In the example, you enable scalable shares on Cluster-A to dynamically adjust the resource allocation of child resource pools. All VMs have an equal number of shares, so the shares configuration of the child resource pools adjusts to give every finance VM twice as much as the engineering VMs. The shares configuration of the child resource pools adjusts to 80% for the Finance pool, and 20% for the Engineering pool. Each finance VM gets 20% of the Cluster-A processor time, and the engineering VMs get 10%.

https://t.me/learningnets

4-127 About Scalable Shares (2)

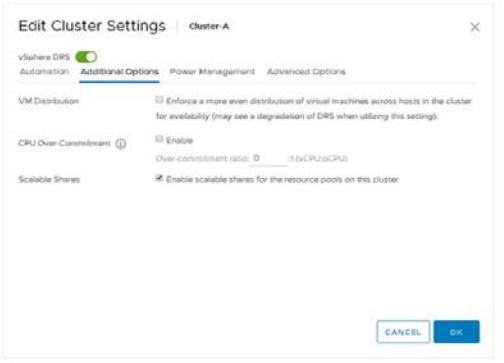
The scalable shares feature makes the child resource pools scalable:

- DRS dynamically scales up and scales down the shares of the resource pools to maintain a resource allocation ratio.
- The scalable shares algorithm maintains the resource allocation ratio by summing the shares of all entities in the resource pool and multiplying by the user-specified initial shares value.
- Scalable shares make the child resource pool scalable. It does not make the parent resource pool scalable.

https://t.me/learningnets

4-128 Enabling Scalable Shares

You can enable scalable shares at the cluster or resource pool level. The resource allocation of child resource pools becomes scalable.



https://t.me/learningnets

4-129 Lab 15: Managing Resource Pools

Create and use resource pools:

1. Maintain VMs
2. Create CPU Contention
3. Create Resource Pools
4. Verify Resource Pool Functionality

https://t.me/learningnets

4-130 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Create a resource pool
- Set resource pool attributes
- Describe expandable reservations
- Describe scalable shares

https://t.me/learningnets

4-131 Key Points

- vSphere certificate management simplifies the process of certificate replacement.
- VMware CA provisions each ESXi host and each vCenter Server service with certificates that are signed by VMware CA.
- vCenter Server identity federation extends authorization to Microsoft ADFS.
- vSphere Trust Authority secures a vSphere environment and its workloads.
- Using the Host Profiles feature, you can 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.
- A resource pool is a logical abstraction for flexible management of resources, such as CPU and memory.

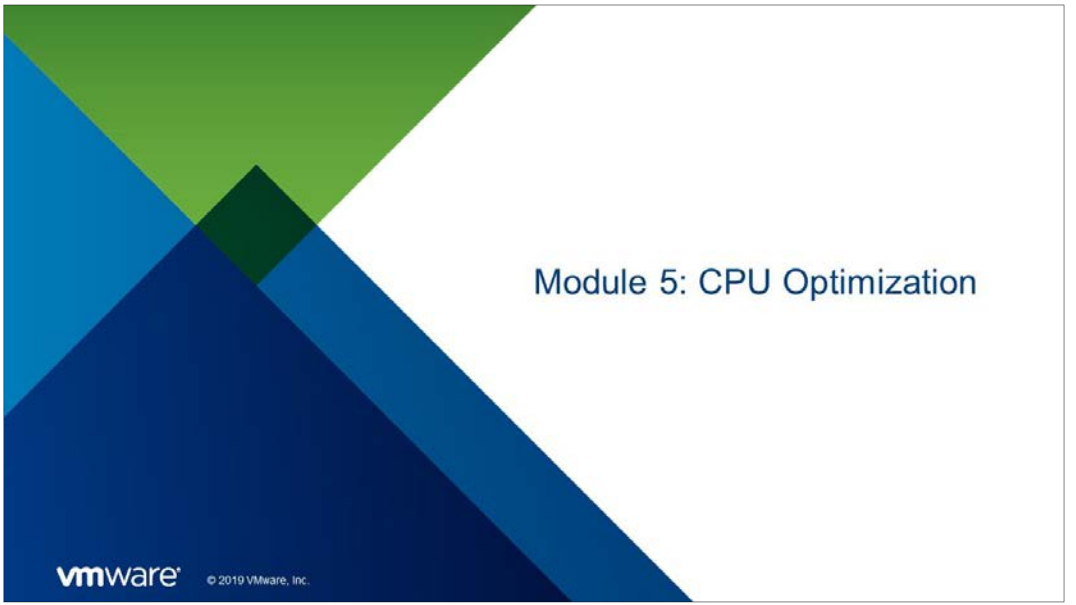
Questions?

https://t.me/learningnets

https://t.me/learningnets

Module 5

CPU Optimization



https://t.me/learningnets

5-2 Importance

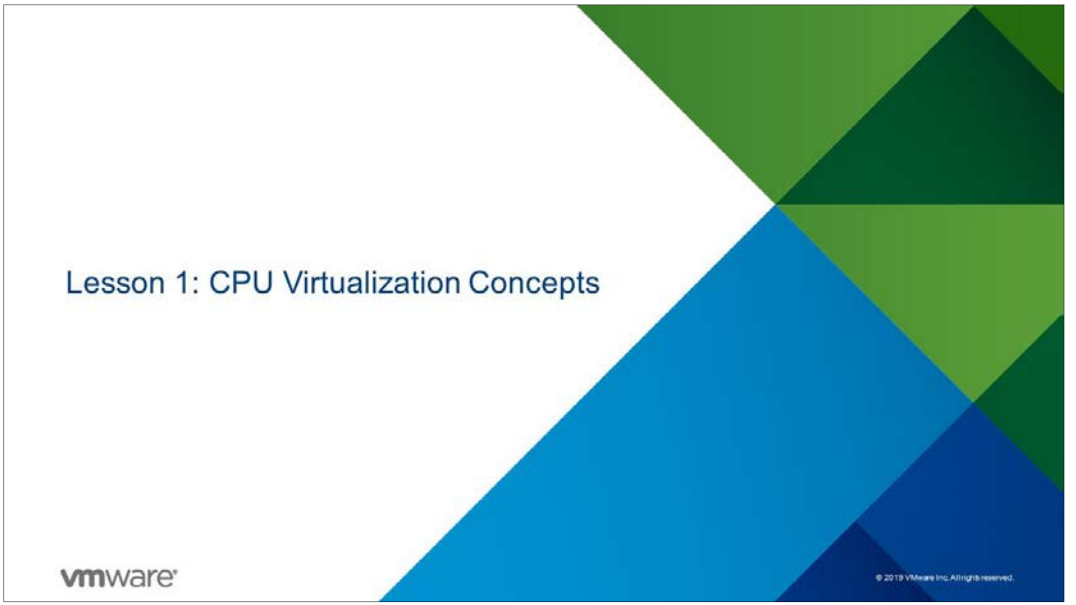
In a vSphere environment, multiple VMs run on the same host. To prevent a situation in which VMs are allocated insufficient CPU resources, you must know how to monitor both host and VM CPU use.

https://t.me/learningnets

5-3 Module Lessons

Lesson: CPU Virtualization Concepts
Lesson: NUMA Architecture in vSphere
Lesson: Using esxtop
Lesson: Monitoring CPU Activity

https://t.me/learningnets



https://t.me/learningnets

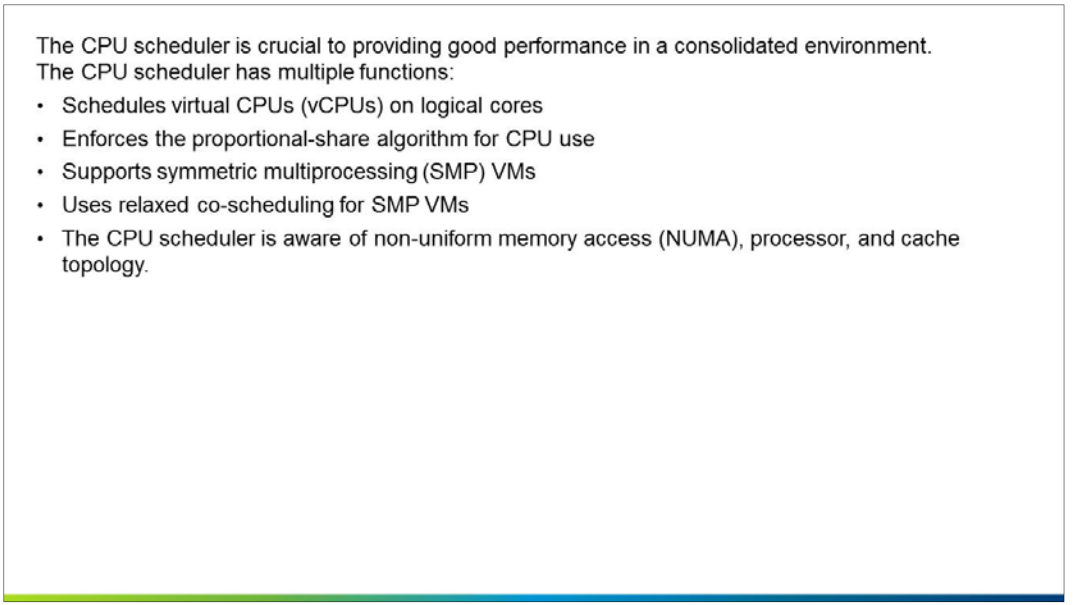
5-5 Learner Objectives

After completing 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

5-6 About the CPU Scheduler



A CPU capable of simultaneous multithreading (SMT) has two logical cores per physical core.

Because most modern processors are equipped with multiple cores per processor, building a system with tens of cores running hundreds of VMs is easy. In such a large system, allocating CPU resources efficiently and fairly is critical.

https://t.me/learningnets

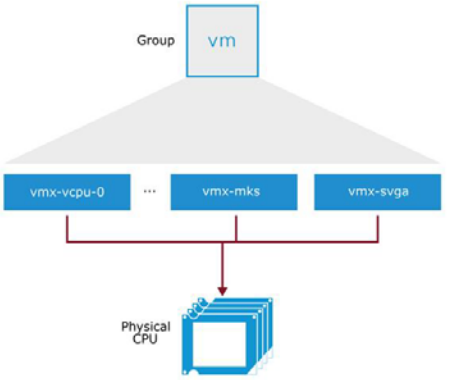
5-7 About VM Worlds

A world is an execution context that is scheduled on a logical core. A world is like a process in conventional operating systems.

A VM is a collection, or group, of worlds.

Different types of VM worlds exist. Some examples of VM worlds:

- One for each vCPU (`vmx-vcpu-0`, `vmx-vcpu-1`, and so on)
- One for the VM's mouse, keyboard, and screen (`vmx-mks`)
- One for the SVGA device (`vmx-svga`)
- The CPU scheduler chooses which world to schedule on a logical core.



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 use. 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-VM worlds exist as well. These non-VM worlds are VMkernel worlds, which are used to perform various system tasks. Examples of these non-VM worlds include the `idle` and `vmotionServer` worlds.

https://t.me/learningnets

5-8 CPU Scheduler Features

Features of the CPU scheduler include:

- The CPU scheduler allocates CPU resources and coordinates CPU use.
- The CPU scheduler uses dynamic and transparent CPU resource allocation:
  - It checks physical CPU use every 2 to 30 milliseconds and migrates vCPUs as necessary.
- The CPU scheduler allows a vCPU to run on a logical core for 50 milliseconds (the default time slice) before another vCPU of the same priority gets scheduled.

  While a vCPU is running on a core, other vCPUs must wait their turn to be scheduled on that core, which can introduce queuing.
- The CPU scheduler enforces the proportional-share algorithm for CPU use:
  - It hosts time-slice physical cores across all VMs when CPU resources are overcommitted.
  - It 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 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 VMs so that each VM 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 consumed less than its entitlement, the world is considered high priority and is likely to be chosen to run next.

https://t.me/learningnets

5-9 CPU Scheduler Support for SMP VMs

ESXi uses a form of co-scheduling that is optimized to run SMP VMs 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 considers skew when scheduling vCPUs in an SMP VM:

- Skew is the difference in execution rates between vCPUs in an SMP VM.
- A vCPU's skew increases when it is not making progress but one of its sibling vCPUs is.
- A vCPU is considered to be in a skewed state if the execution rates between vCPUs in the VM exceed a threshold (typically more than a few milliseconds).

A symmetric multiprocessing (SMP) VM presents the guest OS 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 VM.

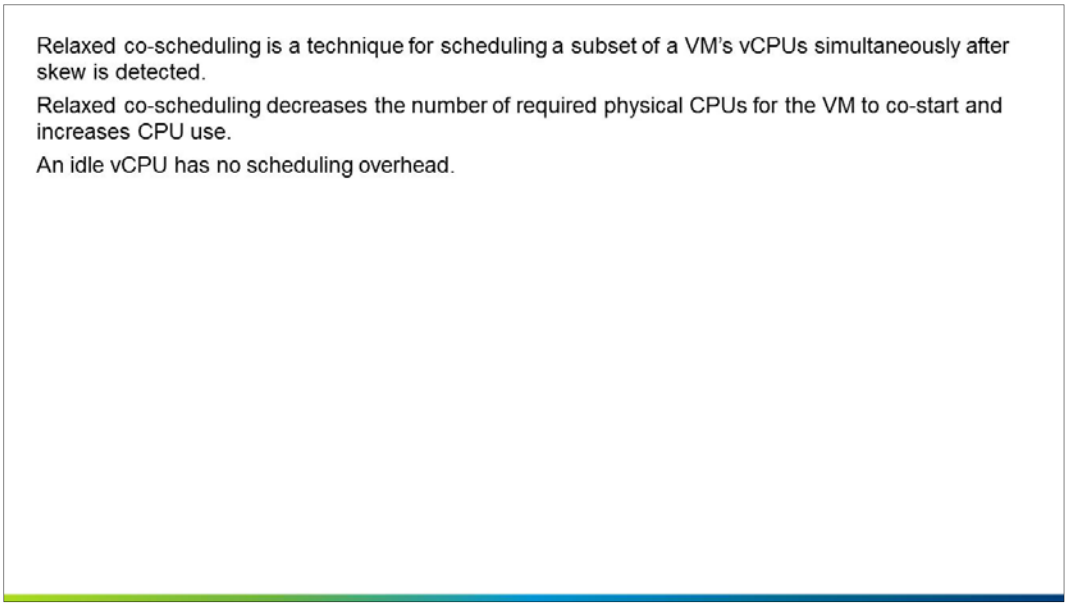
Without co-scheduling, the vCPUs associated with an SMP VM would be scheduled independently, breaking the guest's assumptions regarding uniform progress.

The progress of each vCPU in an SMP VM 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 VM. 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

5-10 Relaxed Co-Scheduling



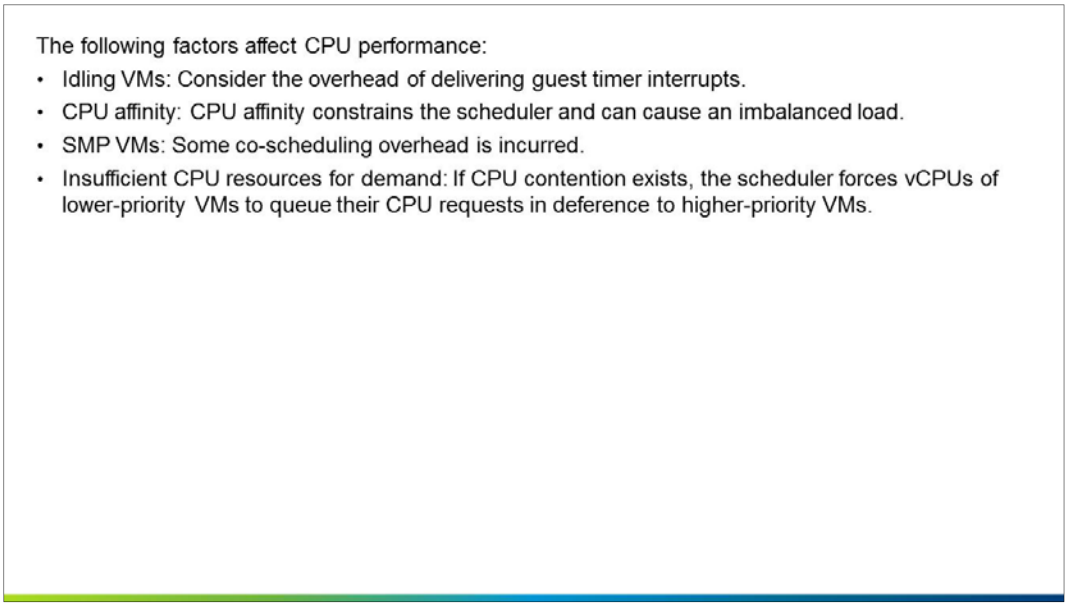
With relaxed co-scheduling, only the vCPUs that are skewed must be co-started. Relaxed co-scheduling ensures that, when any vCPU is scheduled, all other vCPUs that are behind are also scheduled, reducing skew. This approach is called relaxed co-scheduling because only a subset of a VM's vCPUs must be 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 VMs. For example, an ESXi host with two physical cores might be running one vCPU each from two different VMs, if their sibling vCPUs are idling, without incurring co-scheduling overhead.

https://t.me/learningnets

5-11 CPU Performance Factors



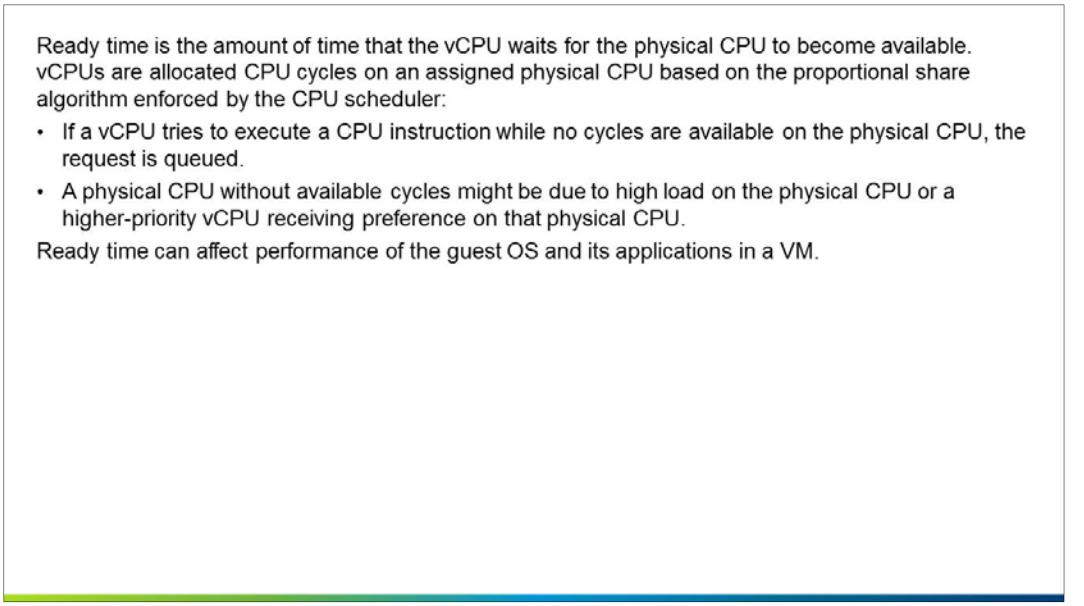
You might think that idle VMs do not cost anything in terms of performance. However, timer interrupts still must be delivered to these VMs. Lower timer interrupt rates can help guest OS performance.

You should use as few vCPUs in your VM as possible to reduce the amount of timer interrupts necessary, and to reduce any co-scheduling overhead that might be incurred. Also, use SMP kernels, and not uniprocessor kernels, in SMP VMs.

CPU capacity is a finite resource. Even on a server that allows additional processors to be configured, a maximum always exists as to the number of processors that can be installed. As a result, performance problems might occur when insufficient CPU resources are available to satisfy demand.

https://t.me/learningnets

5-12 Warning Sign: Ready Time

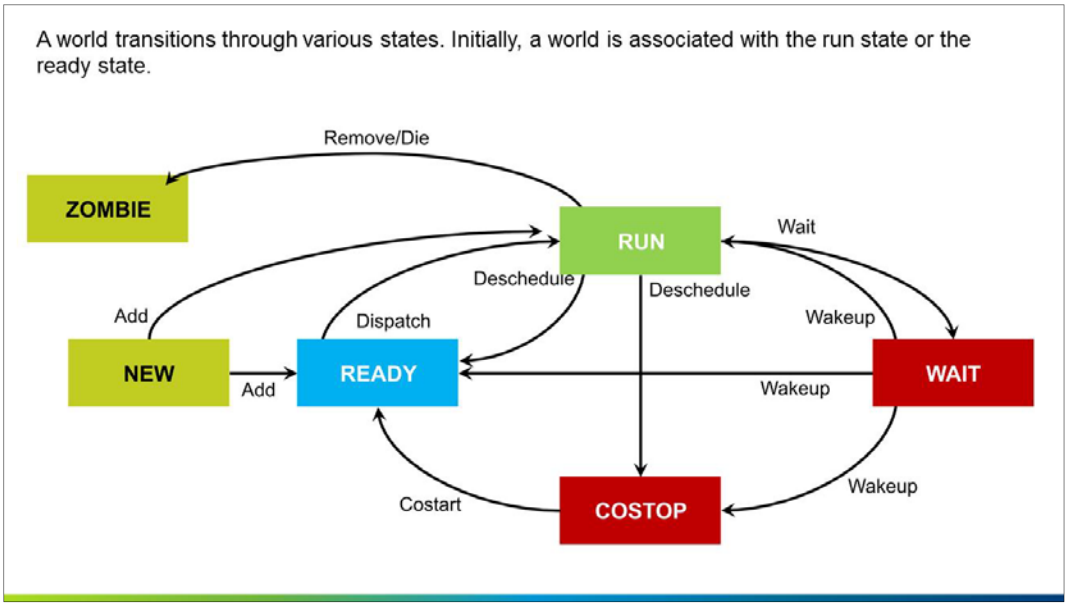


To achieve best performance in a consolidated environment, you must consider ready time. Ready time is the time that a VM must wait in the queue in a ready-to-run state before it can be scheduled on a CPU.

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

5-13 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 co-stop 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

5-14 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 VM to run in
- ❒ Determines which worlds are entitled to CPU time
- ❒ Migrates worlds from busy processors to idle processors

https://t.me/learningnets

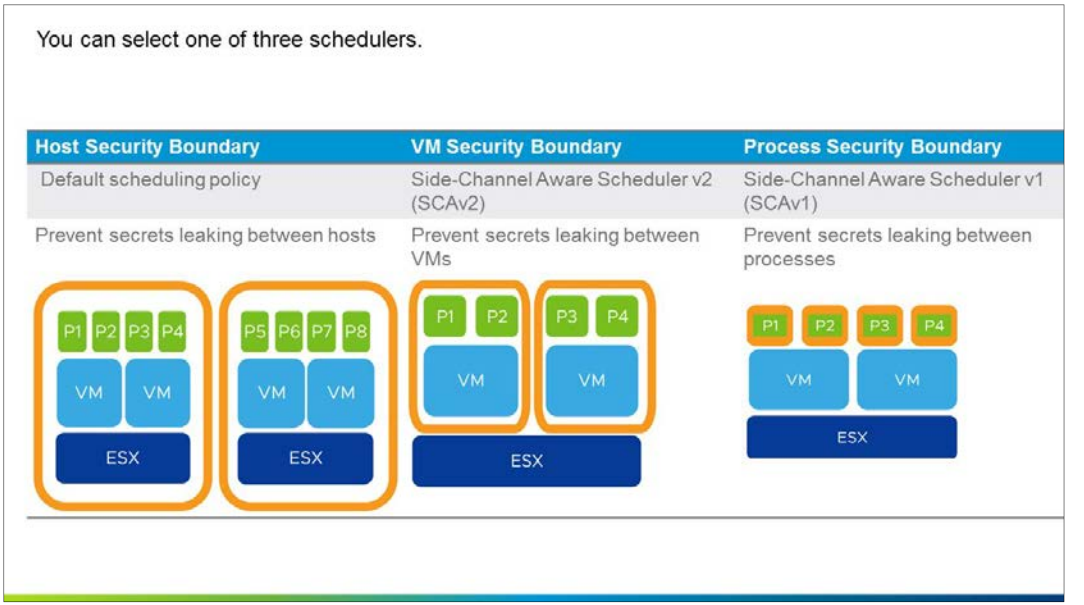
5-15 Activity: CPU Scheduler (2)

What tasks does the CPU scheduler perform? Select all that apply.

- ✓ Schedules vCPUs on physical CPUs
- ❒ Creates a world for each VM 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 Security Boundaries



For an overview of the security issues, descriptions of each scheduler option, and the results of performance testing with different scenarios, see *Performance of vSphere 6.7*

Scheduling Options at https://blogs.vmware.com/vsphere/2019/05/which-vsphere-cpu-scheduler-to-choose.html

https://t.me/learningnets

5-17 Review of Learner Objectives

After completing 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

https://t.me/learningnets

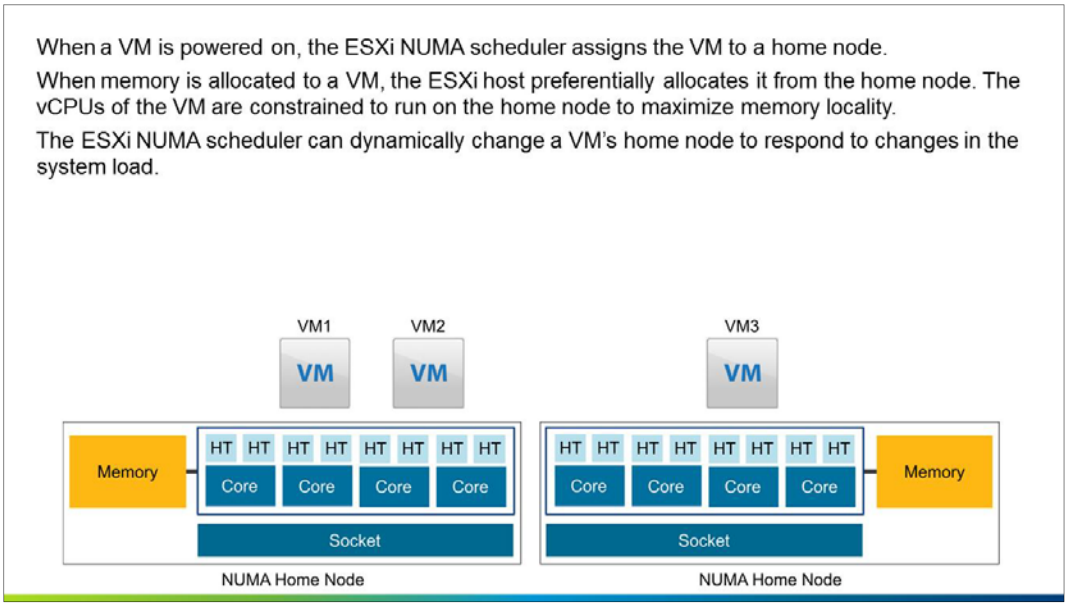
5-19 Learner Objectives

After completing 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 How the ESXi NUMA Scheduler Works



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 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 doing so improves overall memory locality.

https://t.me/learningnets

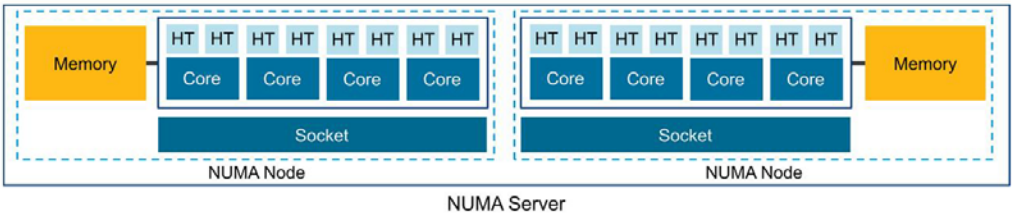
5-21 NUMA-Aware ESXi

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.

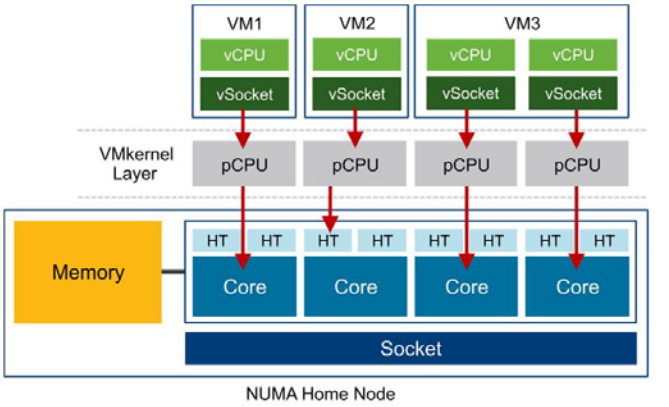
In a NUMA host, the delay incurred when accessing memory varies for different memory locations.

When poor NUMA locality occurs, the VM's performance might be less than if its memory were all local.

https://t.me/learningnets

5-22 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.



To understand how NUMA functions in vSphere, a review of CPU components at the virtual and physical layers is helpful.

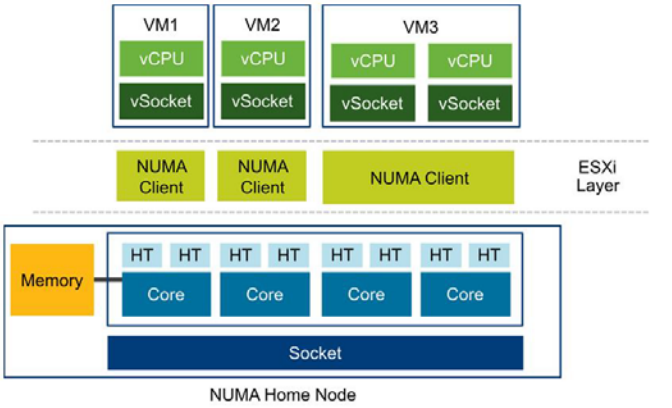
https://t.me/learningnets

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.



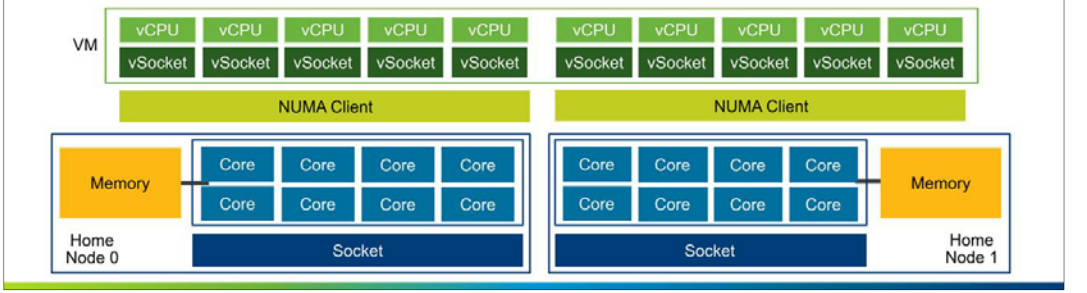
When a VM is powered on, the number of vCPUs in a VM are compared with the number of physical cores in a NUMA node. If enough physical cores exist in the NUMA node to satisfy the vCPU count, a single NUMA client is created. A uniform memory address space is presented to the VM.

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 several NUMA clients and is called 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, hyperthreads 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

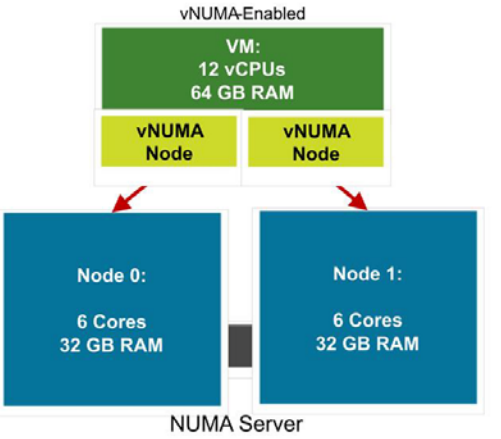
5-25 Virtual NUMA Topology

When a wide VM is created, a virtual NUMA (vNUMA) topology is created and presented to the guest OS.

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 OS 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 OS 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. 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* on the Frank Denneman website at https://frankdenneman.nl/.

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 use.

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 OS 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.

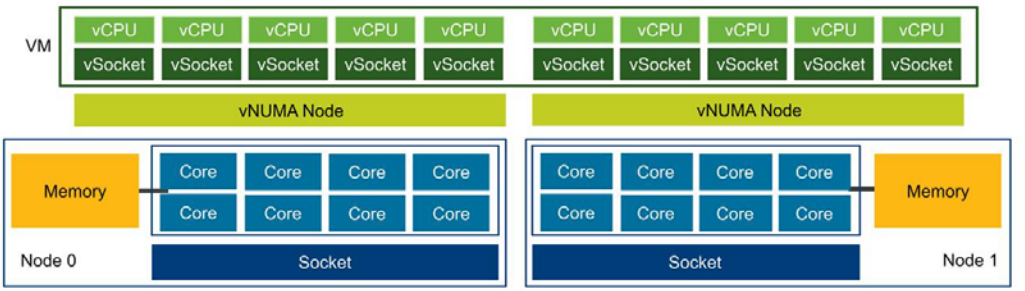
When you create VM, the number of vCPUs that you specify is divided by the cores per socket value to give you the number of sockets. The default Cores per Socket value is 1.

To ensure both an optimal vNUMA topology and performance, regardless of what vSphere version you use, 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)

In the example, you have a VM that is configured with 10 vCPUs and 1 core per socket (10 sockets). This VM runs on a dual-socket, eight cores per socket system.



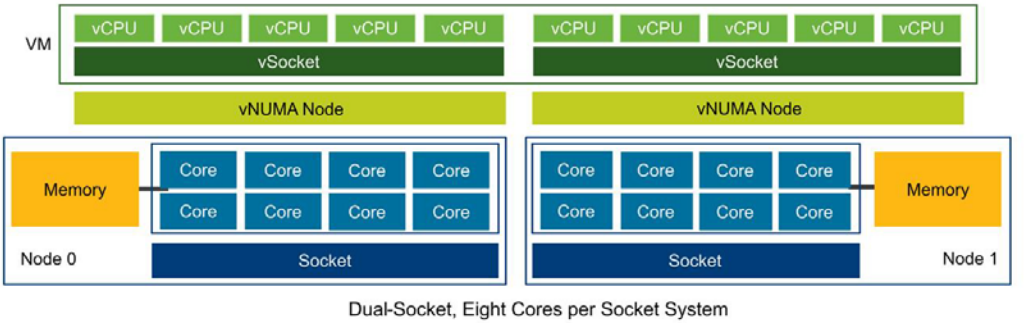
Dual-Socket, Eight Cores per Socket System

https://t.me/learningnets

5-28 Optimizing the Number of Cores per Socket for VMs (3)

To optimize guest OS 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.

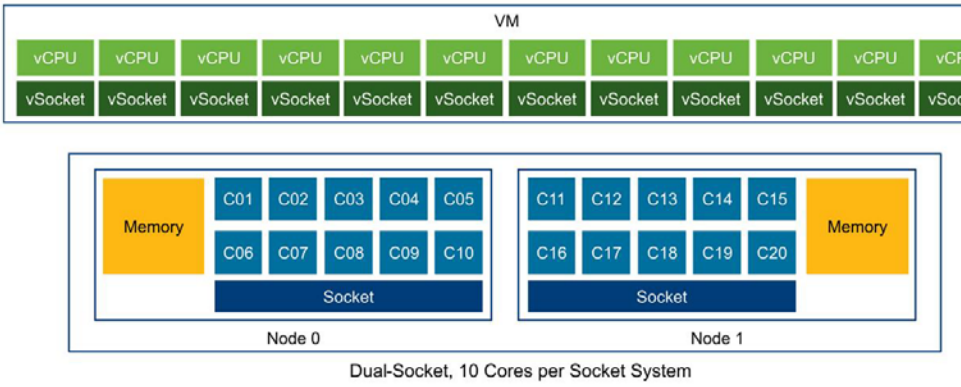


Dual-Socket, Eight Cores per Socket System

https://t.me/learningnets

5-29 Activity: VM Cores Per Socket (1)

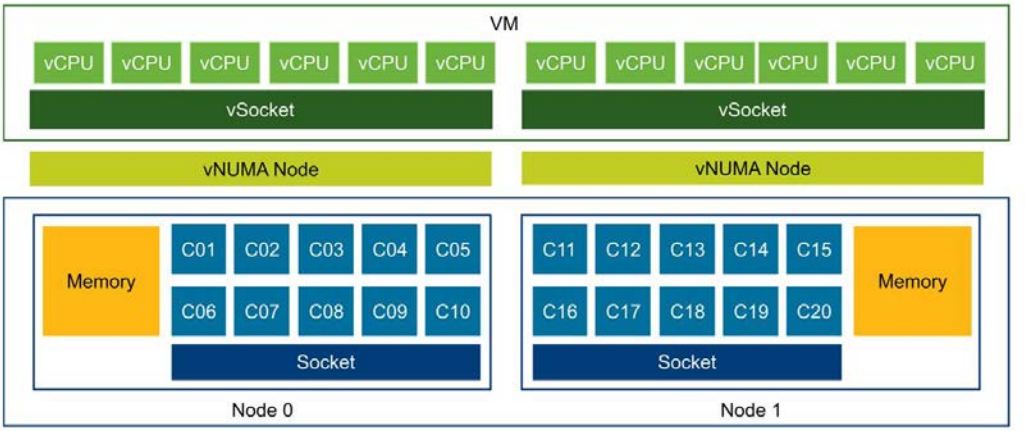
For optimal guest OS performance, what should the number of cores per socket be for a 12-vCPU VM running on a dual-socket, 10 cores per socket system?



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.



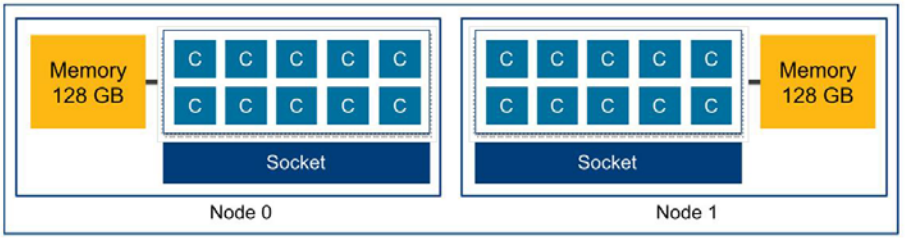
https://t.me/learningnets

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.

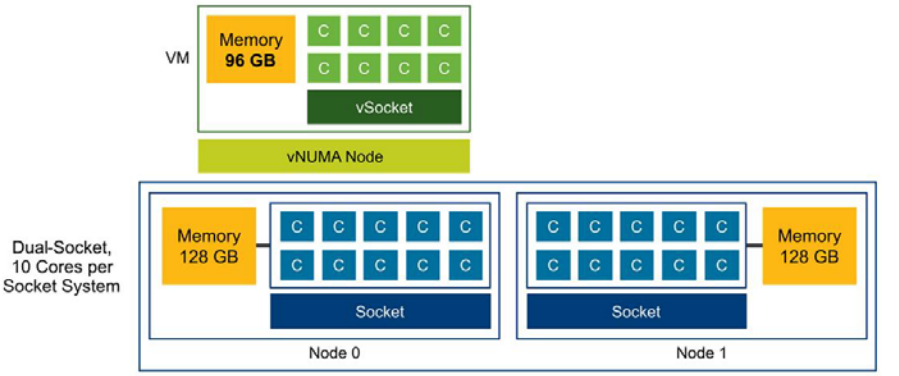


Dual-Socket, 10 Cores per Socket System

https://t.me/learningnets

5-32 Memory Considerations with vNUMA (2)

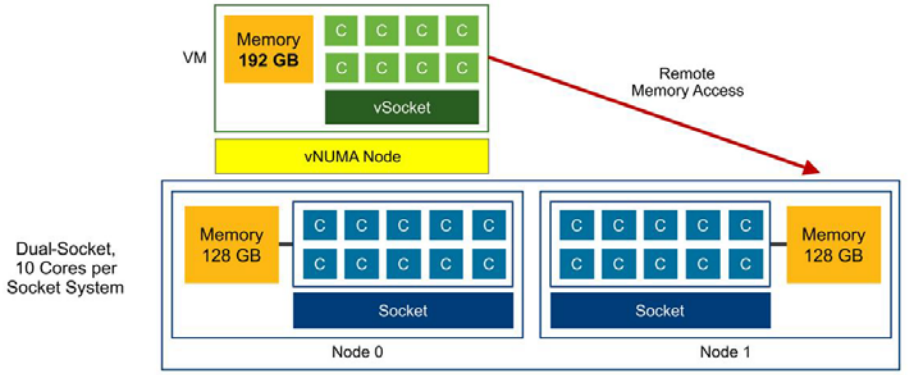
Continuing with the example, you create a VM with 96 GB of RAM, 1 socket, and 8 cores per socket. For this VM, ESXi creates a single vNUMA node. The VM fits into a single physical NUMA node.



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's memory will likely span 2 physical NUMA nodes, resulting in remote memory access.

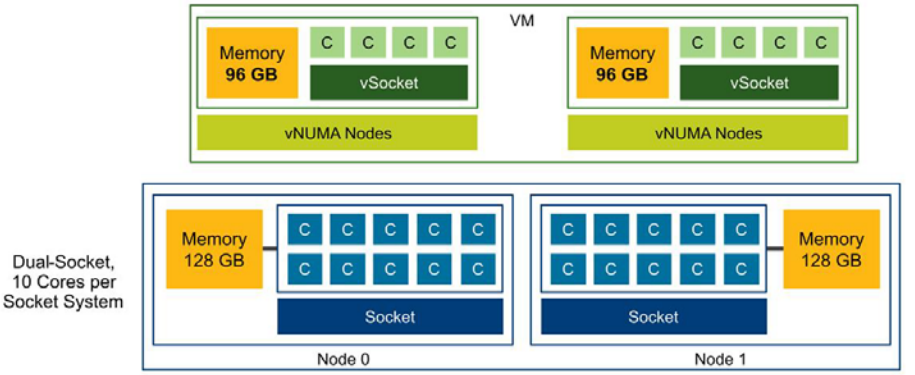


https://t.me/learningnets

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. To complete this configuration, you must also set **numa.vcpu.maxPerMachineNode** 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 an optimal vNUMA configuration:

- Always configure the VM vCPU count to be reflected as cores per a single socket, until you exceed either the physical core count or the total memory available on a single physical NUMA node.
- When you must configure more vCPUs than there are physical cores in the NUMA node, or if you assign more memory than a NUMA node contains, you must 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.
- Though 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 and *Decoupling of Cores per Socket from Virtual NUMA Topology in vSphere 6.5* on the Frank Denneman website at https://frankdenneman.nl/.

https://t.me/learningnets

5-36 Review of Learner Objectives

After completing 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

https://t.me/learningnets

5-38 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe how to use the `esxtop` tool in interactive mode

https://t.me/learningnets

5-39 About esxtop (1)

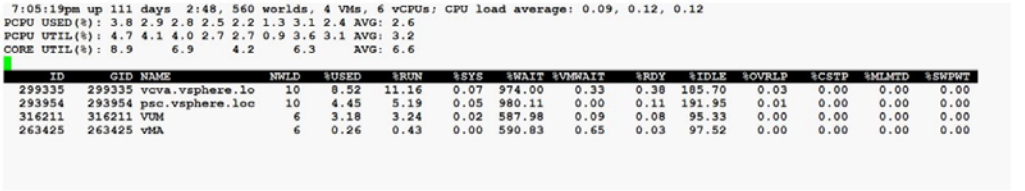
The primary real-time performance monitoring tool for vSphere is esxtop:

- It can be run from a host's vSphere ESXi Shell as `esxtop`.
- It can also be run remotely as `resxtop`.

`esxtop` provides much of the same information as the vSphere Client but supplements the output with more detailed information.

`esxtop` works like the `top` performance utility found in Linux operating systems.

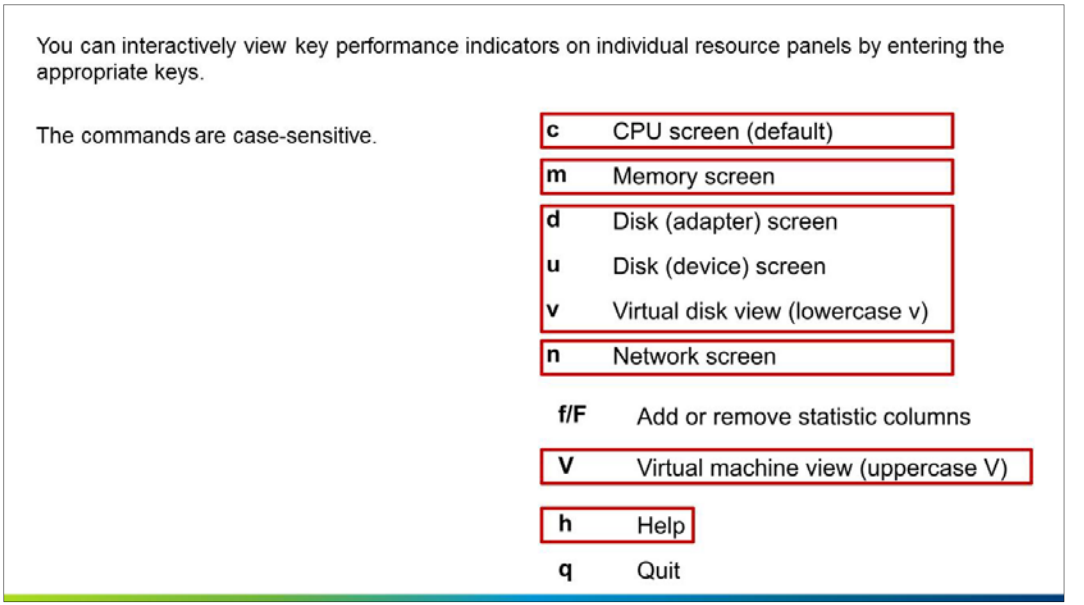
`esxtop` displays information in percentages of sample time (5 seconds by default).



You can run the `esxtop` utility using vSphere ESXi Shell to communicate with the management interface of the ESXi host. To do this, you must have root user privileges.

https://t.me/learningnets

5-40 About esxtop (2)



You use lowercase and uppercase letters to specify which columns appear in which order on the CPU, memory, storage adapter, storage device, VM storage, and network panels. You can add or remove columns containing data in the respective `esxtop` panels.

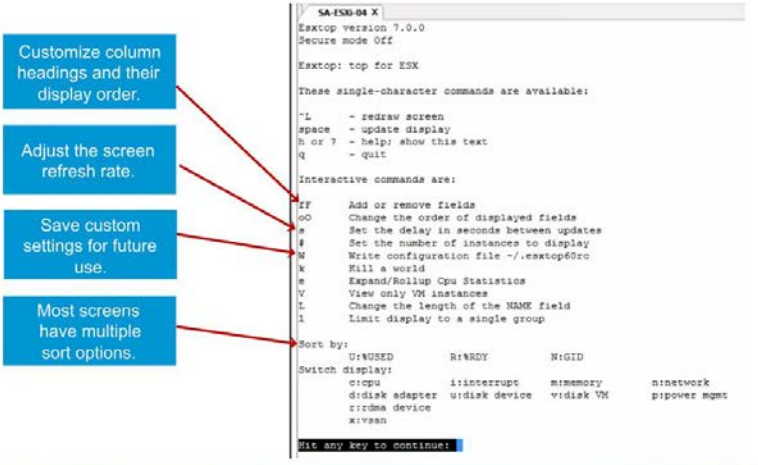
For more information about using `esxtop` or `resxtop` in interactive mode, see Using esxtop or resxtop in Interactive Mode

at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.monitoring.doc/GUID-49D45588-AE12-4143-9CCD-3E77368C506E.html.

https://t.me/learningnets

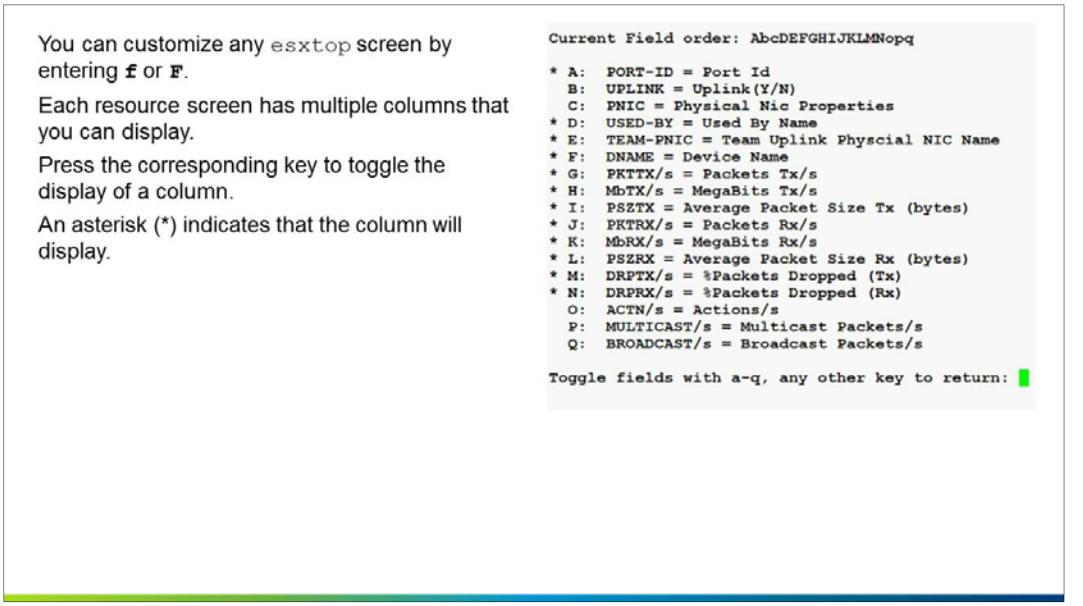
5-41 Getting Help in esxtop

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



For more information about the `esxtop` utility, see vSphere Monitoring and Performance at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.monitoring.doc/GUID-A31249BF-B5DC-455B-AFC7-7D0BBD6E37B6.html.

https://t.me/learningnets

5-43 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe how to use the `esxtop` tool in interactive mode

https://t.me/learningnets

https://t.me/learningnets

5-45 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Identify key CPU metrics to monitor
- Use metrics in esxtop
- Monitor CPU use

https://t.me/learningnets

5-46 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

5-47 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 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 VM:

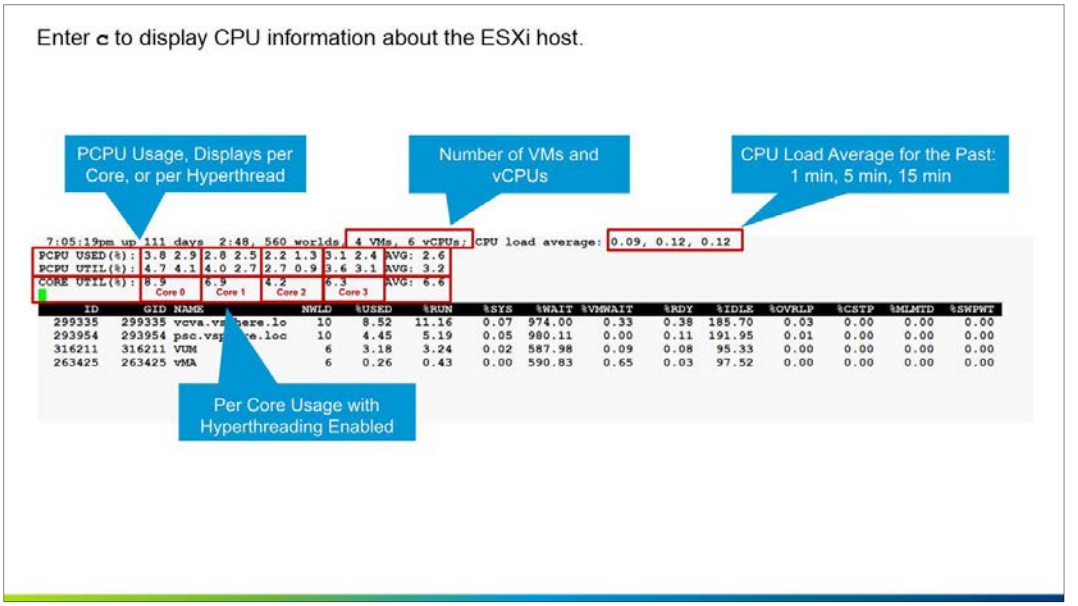
- %USED: CPU use. The percentage of physical CPU core cycles used by a group of worlds running VMs, resource pools, or other worlds.
- %RDY: Percentage of time that the VM was ready to run but was not provided CPU resources on which to execute.
- %CSTP: Percentage of time that the vCPUs of a VM spent in the co-stop state, waiting to be co-started. This value gives an indication of the co-scheduling overhead incurred by the VM. If this value is low, any performance problems should be attributed to other issues and not to the co-scheduling of the VM's vCPUs.
- %MLMTD: Percentage of time that the VMkernel did not run the VM because that would violate the VM's limit setting.
- %VMWAIT: Percentage of time that the VM spent in a blocked state waiting for events.

https://t.me/learningnets

- %SWPWT: Percentage of time that the VM spends waiting for the VMkernel to swap memory.
- NWLD: Number of members in the VM of the world that is running.

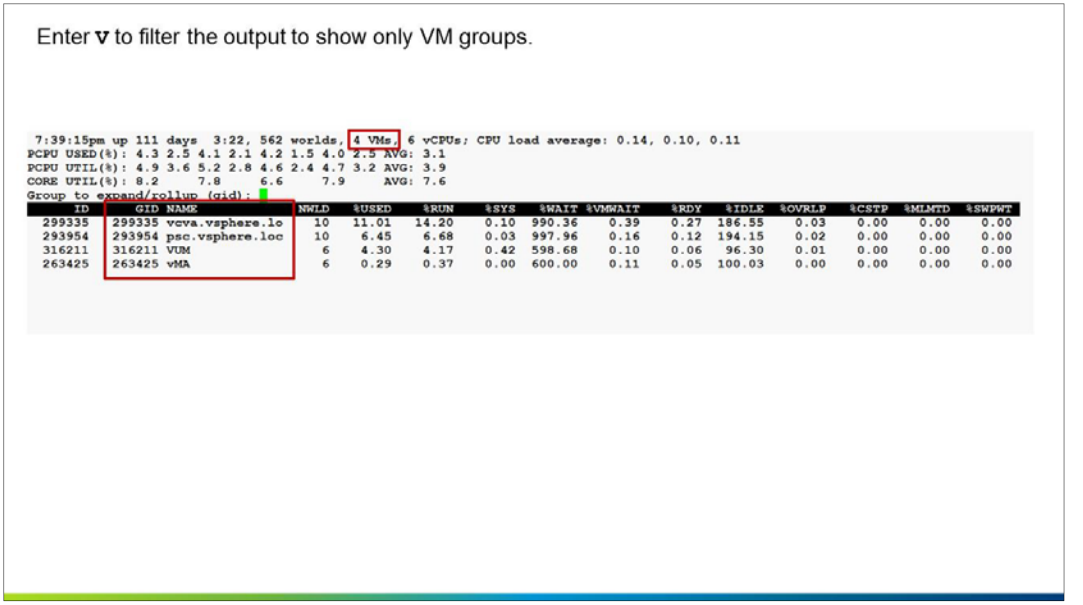
https://t.me/learningnets

5-48 Using esxtop to Monitor ESXi Host CPU



https://t.me/learningnets

5-49 Using esxtop to View CPU Metrics per VM



The `esxtop` tool 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:

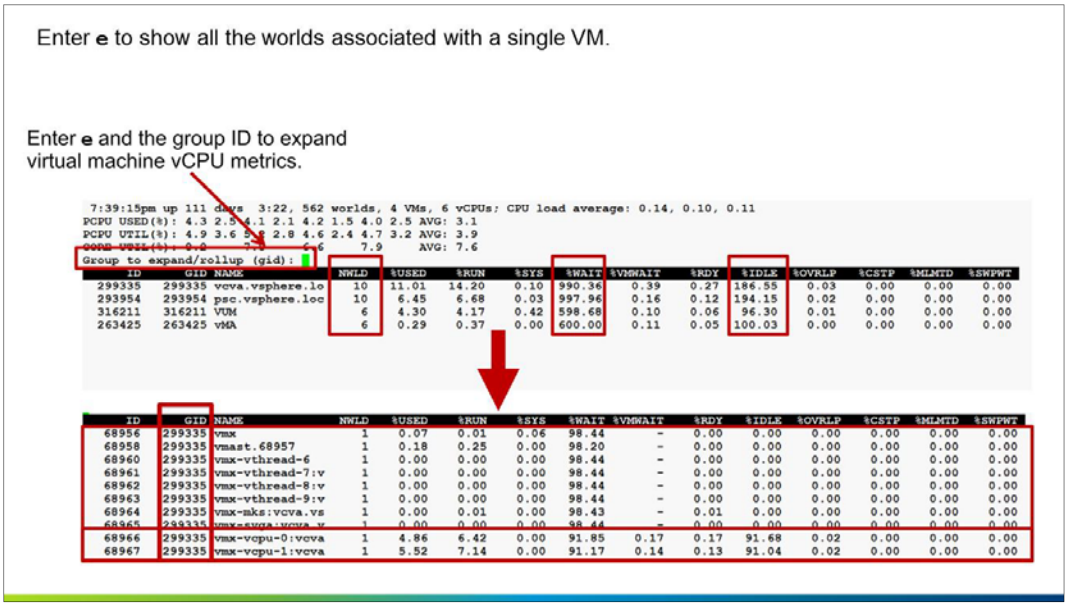
- Press the spacebar: Immediately updates the current screen.
- Enter **s**: Prompts you for the delay between updates, in seconds. The default value is 5 seconds and the minimum value is 2 seconds.
- Enter **V**: Displays VM instances only. Do not confuse this command with lowercase v which shifts the view to the storage VM resource use screen. If you accidentally use lowercase, enter `c` to return to the CPU resource use panel.
- Enter **e**: Toggles between CPU statistics being displayed, expanded, or unexpanded. The expanded display includes CPU resource use statistics broken down by individual worlds

https://t.me/learningnets

belonging to a VM. All percentages of the individual worlds are percentages of a single physical CPU.

https://t.me/learningnets

5-50 Using esxtop to View Single CPU Statistics

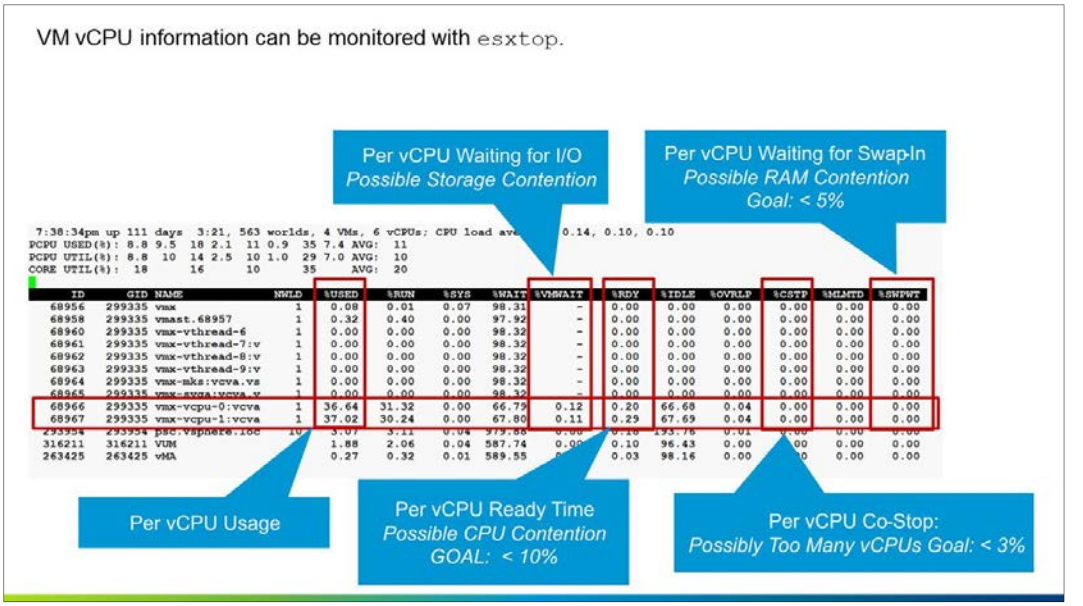


The default `esxtop` window shows aggregate information for all vCPUs in a VM. Enter **e** and enter the group ID to expand the group and display all the worlds associated with that group.

In the example, the VM group with group ID 299335 is expanded to show all the worlds associated with that group.

https://t.me/learningnets

5-51 Using esxtop to Monitor VM vCPU



CPU use, ready time, wait time, and co-stop information for each vCPU can be tracked in this `esxtop` panel.

https://t.me/learningnets

5-52 Important Metrics to Monitor

Monitoring both high-use values and ready time is important to accurately depict CPU performance.
High-use values alone do not always indicate poor performance:

- This value is often an indicator of high use.
- 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 VMs than can be scheduled on the given physical CPUs.

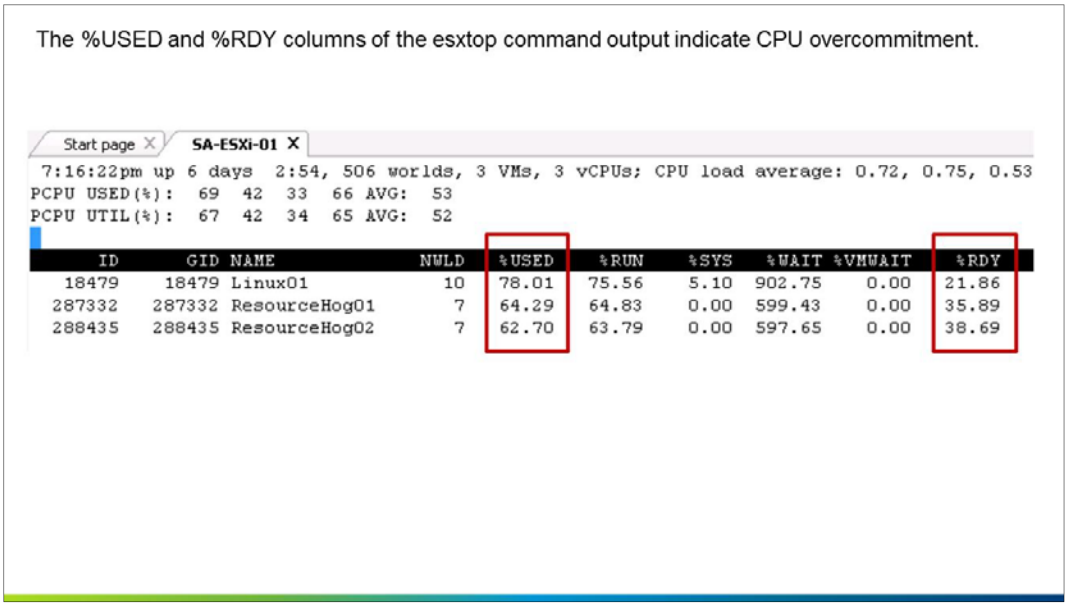
Though high use is not necessarily an indication of poor CPU performance, it is often mistakenly understood as such. Administrators in the physical space view high use as a warning sign because of the limitations of physical architectures. However, one of the goals of virtualization is to efficiently use the available resources. Detecting high CPU use, and queuing, might indicate a performance problem.

As an analogy, consider a freeway. If every lane of 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 it was designed to do, moving many 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

5-53 Example: Identifying CPU Overcommitment

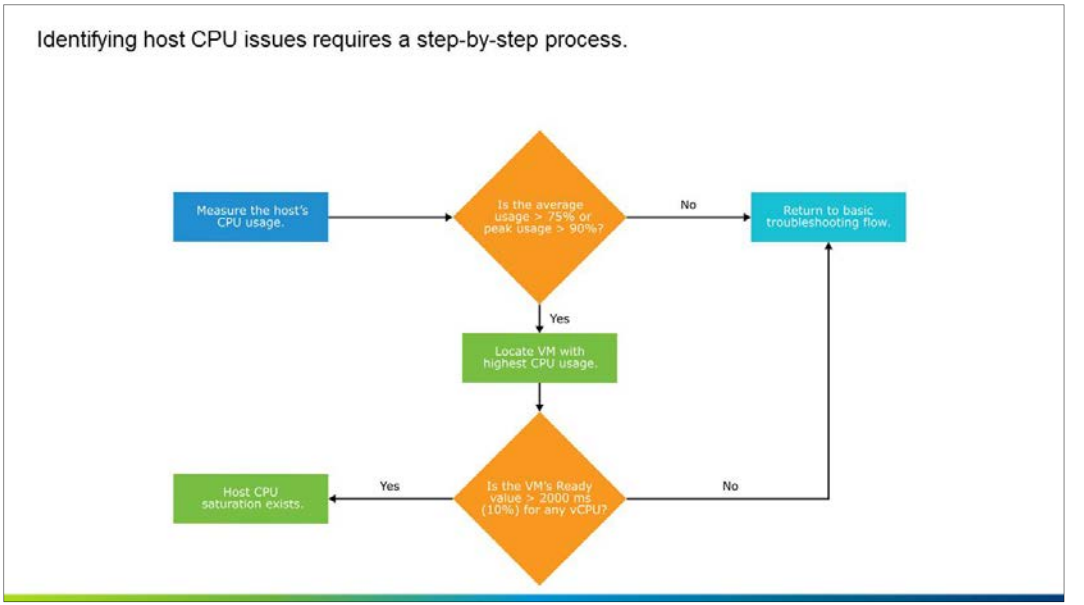


The example uses `esxtop` to detect CPU overcommitment. Looking at the pCPU line near the top of the window, you can determine that the host has four CPUs, three of which are moderately used.

Three active VMs are shown: Linux01, ResourceHog01, and ResourceHog02. These VMs 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 VMs 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

5-54 Host CPU Saturation



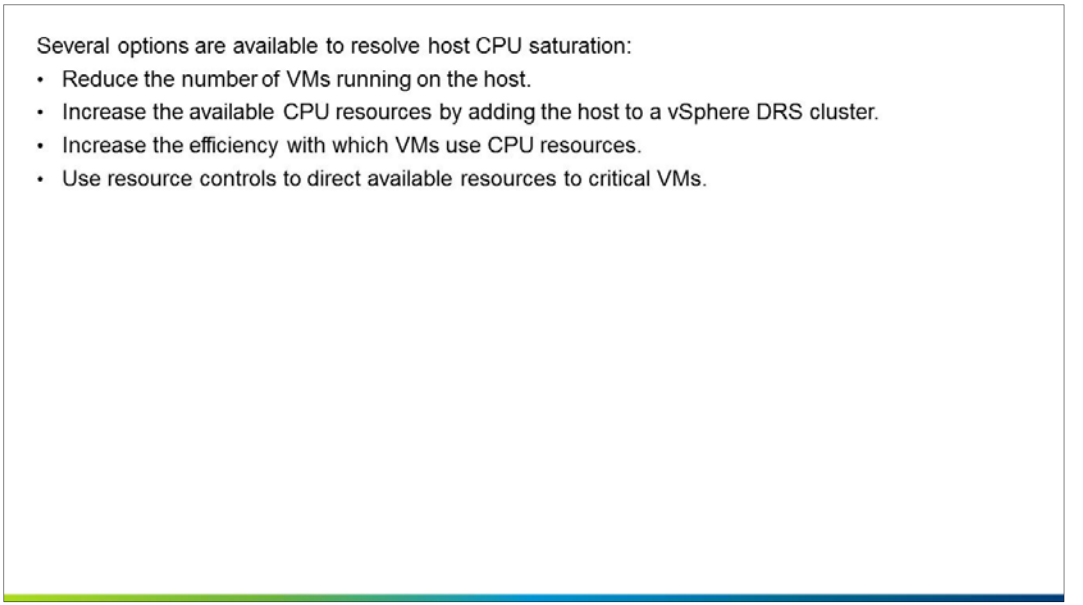
In the esxtop tool, ready time is reported in a percentage format. A figure of 5 percent means that the VM spent 5 percent of its last sample period waiting for available CPU resources.

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

5-55 Resolving Host CPU Saturation



Populating an ESXi host with too many VMs running compute-intensive applications can make supplying sufficient resources to any individual VM impossible.

When the host has a few VMs with high CPU demand, you might solve the problem by first increasing the efficiency of CPU use in the VMs. VMs with high CPU demands are those most likely to benefit from simple performance tuning. In addition, unless care is taken with VM placement, moving VMs with high CPU demands to a new host might move the problem to that host.

When the host has many VMs with moderate CPU demands, reducing the number of VMs on the host by rebalancing the load is the simplest solution. Performance tuning on VMs with low CPU demands yields smaller benefits. It is also time-consuming if the VMs are running different applications.

When the host has a mix of VMs 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 VMs, increasing efficiency might be the best approach.

https://t.me/learningnets

5-56 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

5-57 Activity: CPU Metrics (2)

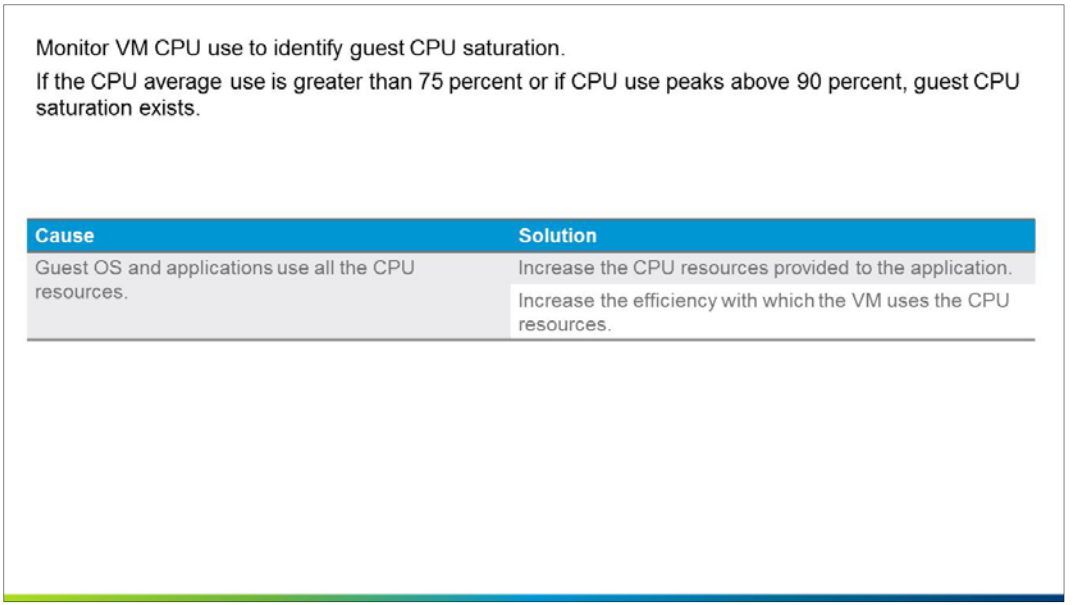
Which of the following metrics in `esxtop` can give you an indication of memory contention?

- ❒ %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

5-58 Guest CPU Saturation



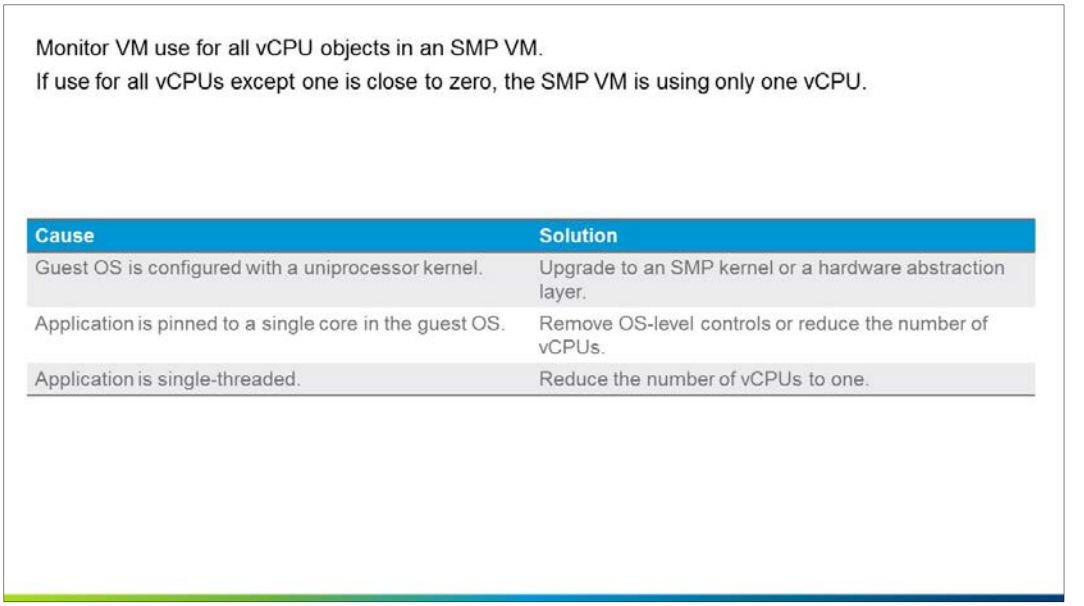
Guest CPU saturation occurs when the application and OS running in a VM use all the CPU resources that the ESXi host is providing to that VM. The occurrence of guest CPU saturation does not necessarily indicate that a performance problem exists.

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 OS 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 VM continues to experience CPU saturation despite adding CPU resources, the tuning and behavior of the application and OS should be investigated.

https://t.me/learningnets

5-59 Using One vCPU in an SMP VM



When a VM 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 OS might be configured with a uniprocessor kernel (with Linux) or hardware abstraction layer (HAL) (with Windows). For a VM to take advantage of multiple vCPUs, the guest OS running in the VM must be able to recognize and use multiprocessor cores. Follow the documentation provided by your OS vendor to verify for the type of kernel or HAL that is being used by the guest OS.

The application might be pinned to a single core in the guest OS. 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 OS, the application might run on only vCPU0. To determine whether this is the case, inspect the commands used to start the application or other OS-level tunings applied to the application. The inspection of these values must be accompanied by an understanding of the OS vendor's documentation regarding restricting CPU resources.

https://t.me/learningnets

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 VM might lead to low guest CPU use, 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 VM might help in determining whether the application is single-threaded.

https://t.me/learningnets

5-60 Low Guest CPU Use

Monitor a VM to identify low CPU use.

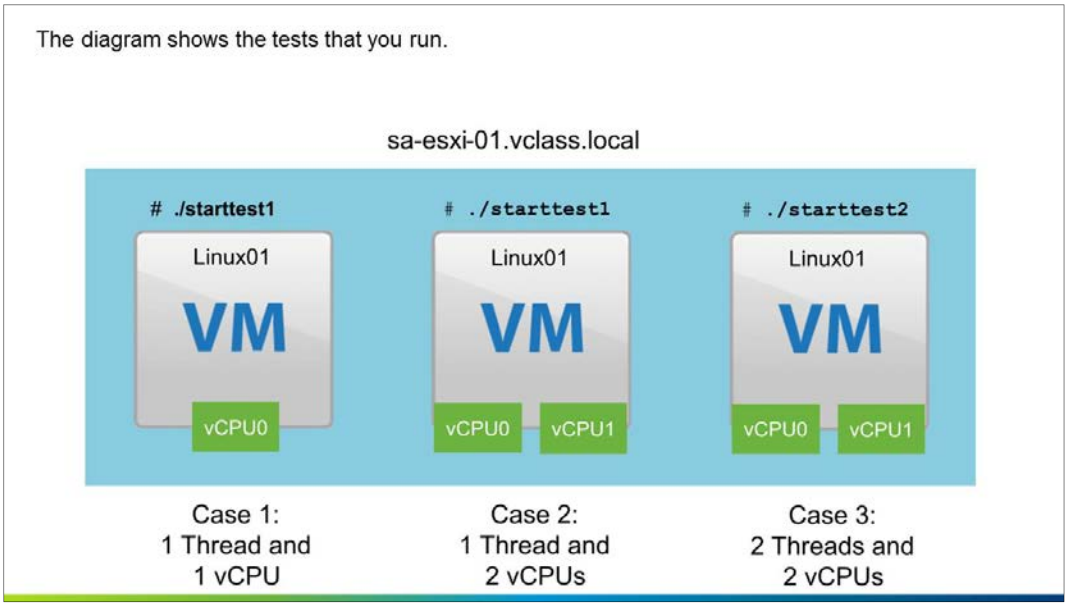
If the average use is less than 75 percent, guest CPU use is low.

Cause	Solution
High storage response times.	Troubleshoot slow or overloaded storage.
High response times from external systems.	Solution is dependent on application and environment.
Poor application or OS tuning.	
Application pinned to cores in guest OS.	Remove OS-level controls or reduce the number of vCPUs.
Too many configured vCPUs.	Reduce the number of vCPUs.
Restrictive resource allocations.	Modify the VM resource settings.

To determine whether the problem is low guest CPU use, use vSphere Web Client to verify the CPU use of the VMs on the host.

https://t.me/learningnets

5-61 Introduction to Lab 8: Monitoring CPU Performance



In this lab, you compare the performance of a MySQL database running on a single-vCPU VM with a dual-vCPU VM. You run three tests. Each test measures CPU use, ready time, idle time, and operations per minute.

You record the data after each test:

- Case 1: On a single-vCPU VM, you run the starttest1 program. starttest1 simulates a single-threaded application accessing the database.
- Case 2: On a dual-vCPU VM, you run the starttest1 program.
- Case 3: On a dual-vCPU VM, you run the starttest2 program. starttest2 simulates a dual-threaded application accessing the database.

https://t.me/learningnets

5-62 Lab 16: Monitoring CPU Performance

Use the `esxtop` command to monitor CPU performance:

1. Run a Single-Threaded Program in a Single-vCPU VM
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 VM
5. Record Statistics for Case 2: One Thread and Two vCPUs
6. Run a Dual-Threaded Program in a Dual-vCPU VM
7. Record Statistics for Case 3: Two Threads and Two vCPUs
8. Analyze the Test Results

https://t.me/learningnets

5-63 Review of Lab: Monitoring CPU Performance

Organize your results for this lab.

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

5-64 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Identify key CPU metrics to monitor
- Use metrics in esxtop
- Monitor CPU use

https://t.me/learningnets

5-65 Key Points

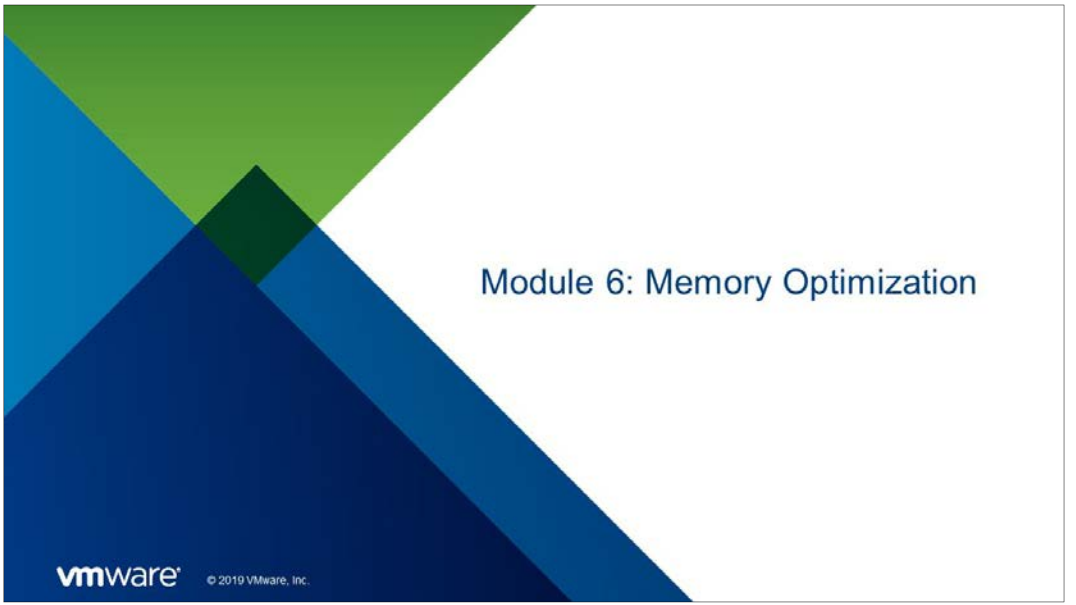
- CPU problems are usually caused by host CPU resources that are insufficient to satisfy vCPU demand.
- High CPU use values with a high ready time often indicate CPU performance problems.
- For CPU-related problems, you should verify host CPU saturation and guest CPU saturation.

Questions?

https://t.me/learningnets

Module 6

Memory Optimization



https://t.me/learningnets

6-2 Importance

Although vSphere employs various mechanisms to efficiently allocate memory, you might still encounter a situation in which VMs are allocated insufficient physical memory.

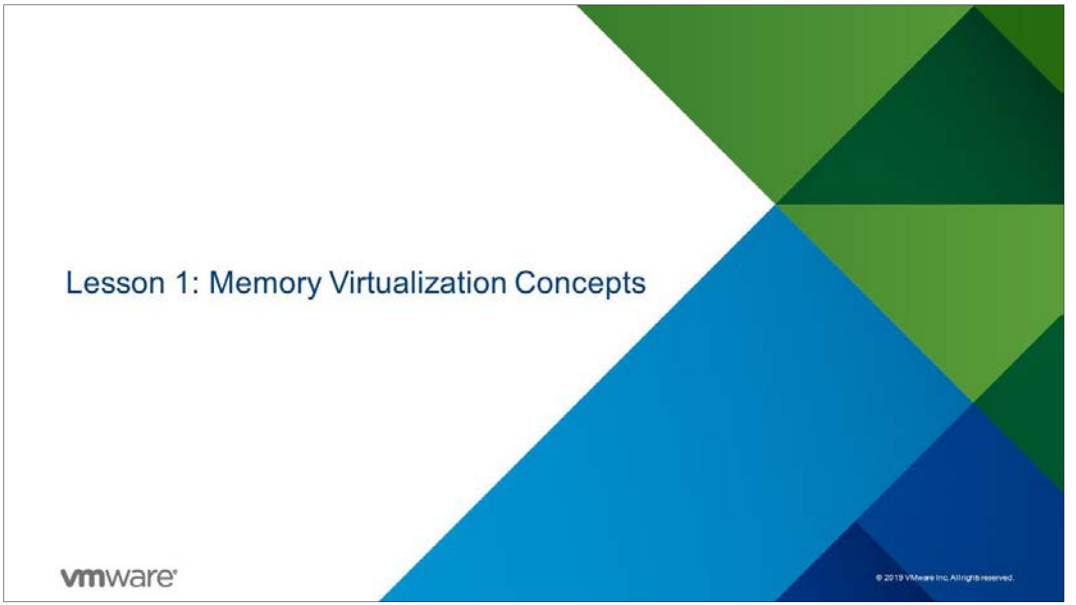
You should know how to monitor memory usage of the host and the VMs.

https://t.me/learningnets

6-3 Module Lessons

Lesson: Memory Virtualization Concepts
Lesson: Monitoring Memory Activity

https://t.me/learningnets



https://t.me/learningnets

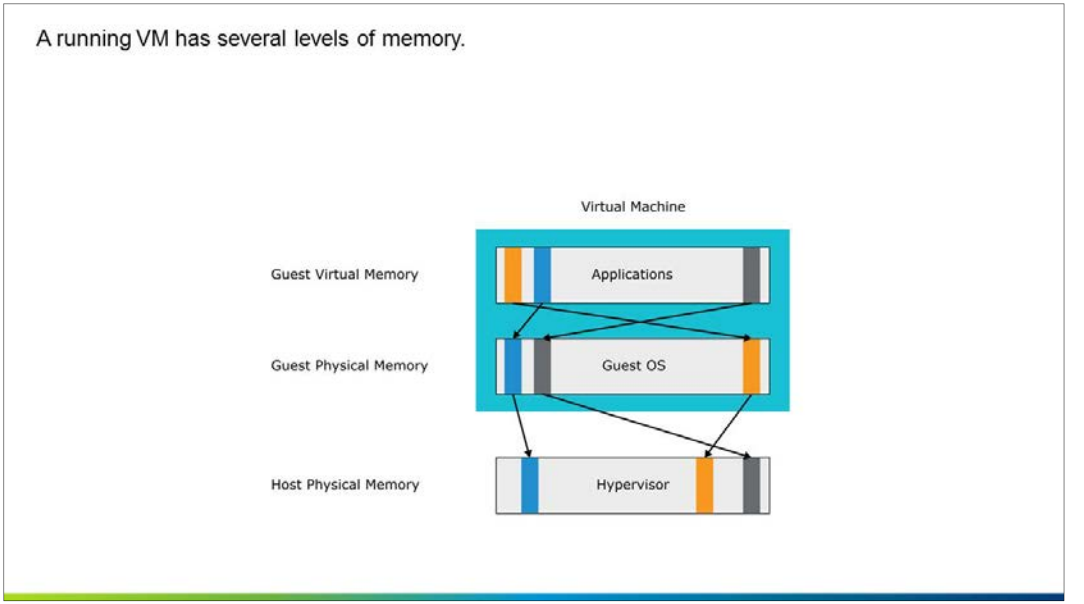
6-5 Learner Objectives

After completing 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

6-6 Virtual Memory Overview

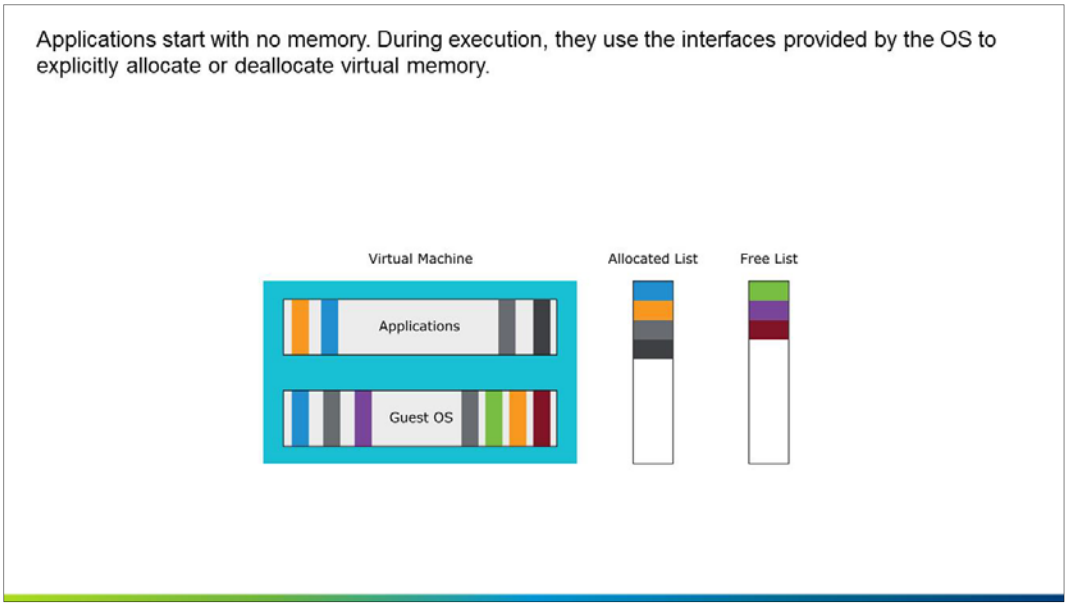


The following levels of memory exist when running a VM:

- Guest virtual memory: Memory visible to the applications running in the VM. This memory is a continuous virtual address space presented by the guest OS to applications.
- Guest physical memory: Memory visible to the guest OS running in the VM. The guest OS 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

6-7 Application and Guest OS Memory Management



In a nonvirtual environment, the OS owns all physical memory in the system. The hardware does not provide interfaces for the OS to explicitly allocate or free physical memory. The OS establishes the definitions of allocated or free physical memory.

Operating systems have different implementations to realize this abstraction. Whether a physical page is free depends on which list the page resides in.

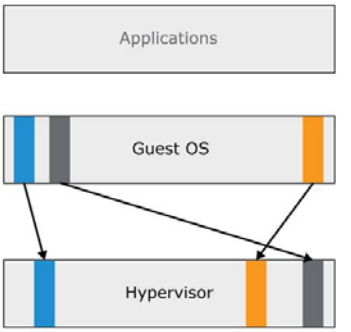
https://t.me/learningnets

6-8 VM Memory Management

VMs start with no host physical memory allocated to them.

Host physical memory is allocated on demand:

- The guest OS does not explicitly allocate memory.
- Memory is allocated on the VM's first access to memory (read or write).



Because a VM runs a guest OS and one or more applications, the VM memory management properties combine both application and OS memory management properties. Like an application, when a VM first starts, it has no preallocated host physical memory. Like an OS, the VM cannot explicitly allocate host physical memory through standard interfaces.

The hypervisor intercepts the VM's memory accesses and allocates host physical memory for the VM on its first access to the memory. The hypervisor always writes zeros to the host physical memory before assigning it to a VM.

https://t.me/learningnets

6-9 Memory Reclamation

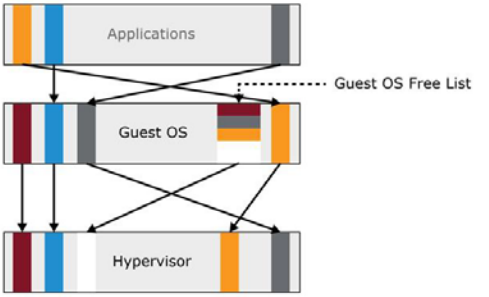
Guest physical memory is not freed in the typical sense:

- 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 VM can accrue lots of host physical memory.

The hypervisor cannot directly reclaim host physical memory that is in the guest OS free list.



VM memory deallocation acts like an OS, such that the guest OS 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.

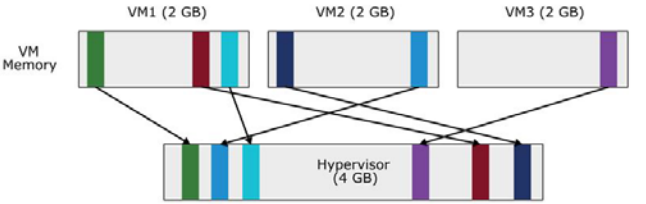
It is difficult for the hypervisor to know when to free host physical memory when guest physical memory is deallocated, or freed, because the guest OS free list is not accessible to the hypervisor. The hypervisor is completely unaware of which pages are free or allocated in the guest OS. As a result, the hypervisor cannot reclaim host physical memory when the guest OS frees guest physical memory.

https://t.me/learningnets

6-10 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.



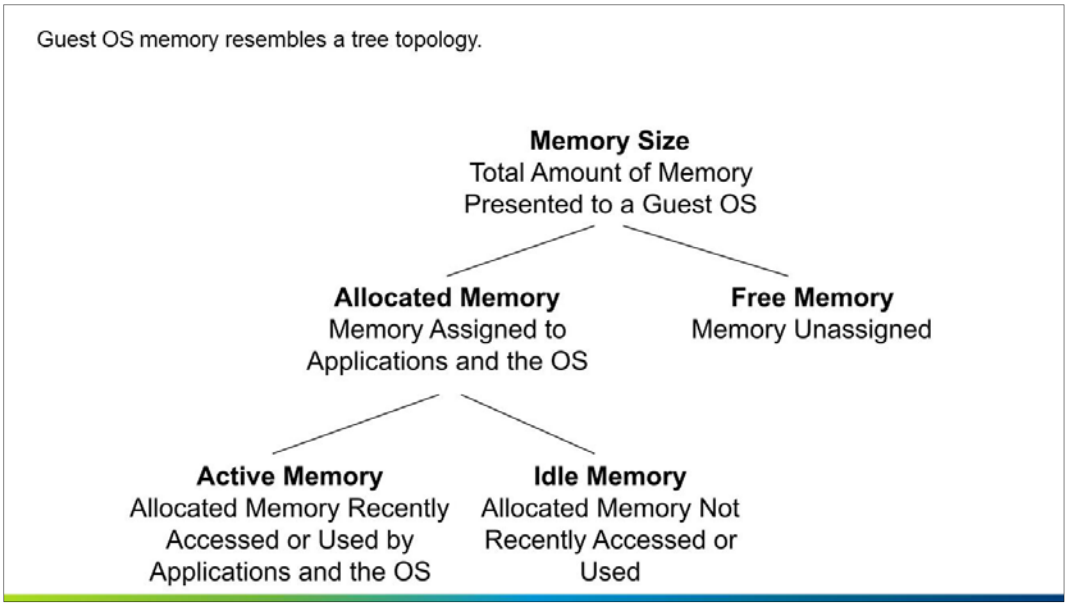
With memory overcommitment, ESXi ensures that the host physical memory is consumed by active guest memory as much as possible. Typically, some VMs might be lightly loaded compared to others, and so, for much of the time, their memory sits idle. Memory overcommitment allows the hypervisor to use memory-reclamation techniques to take the inactive or unused host physical memory away from the idle VMs and give it to other VMs that will actively use it.

With memory overcommitment, each VM has a smaller footprint in host physical memory, making it possible to fit more VMs on the host while still achieving good performance for all VMs. In the example, you can enable a host with 4 GB of physical memory to run three VMs with 2 GB of VM memory each.

This action assumes that all VMs are using the default setting for memory reservation, which is 0. In such a case, all the VMs would power on. If all the VMs had a memory reservation of 2 GB each, without memory overcommitment, then only one VM can be run. The hypervisor cannot reserve host physical memory for more than one VM because each VM has overhead memory as well.

https://t.me/learningnets

6-11 Guest OS Memory Terminology



https://t.me/learningnets

The memory used by the guest OS in the VM can be described in the following way:

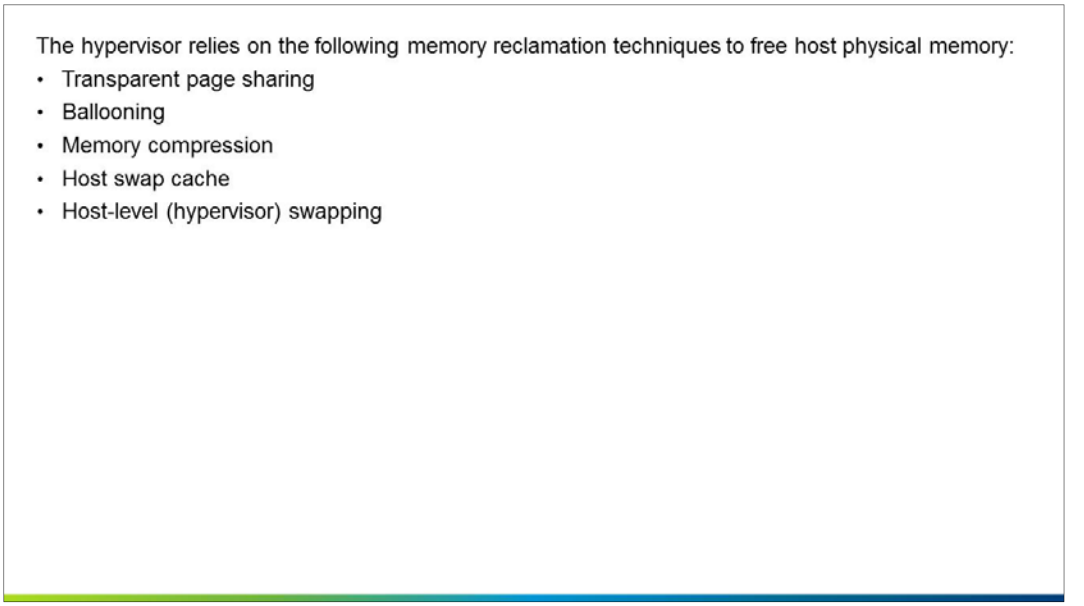
- Memory size is the total amount of memory that is presented to the guest. By default, the memory size corresponds to the memory configured when the VM is created.
- The total amount of memory (memory size) can be divided into two parts:
    - Free memory: Memory that is not assigned to the guest OS or to applications.
    - Allocated memory: Memory that is assigned to the guest OS 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.

Keeping a VM's active memory in host physical memory is vital for maximizing VM performance.

These terms are used to describe the guest OS's memory. The ESXi host does not know about a guest's free, active, or idle memory.

https://t.me/learningnets

6-12 VM Memory Reclamation Techniques



When multiple VMs are running, some VMs might have identical sets of memory content. Opportunities exist for sharing memory across VMs and in a single VM. 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 VMs. When a VM writes to the page, a copy of the page is created for the VM 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 insufficient to reclaim memory. If needed, the hypervisor will swap out guest physical memory to a virtual swap (VSWP) file.

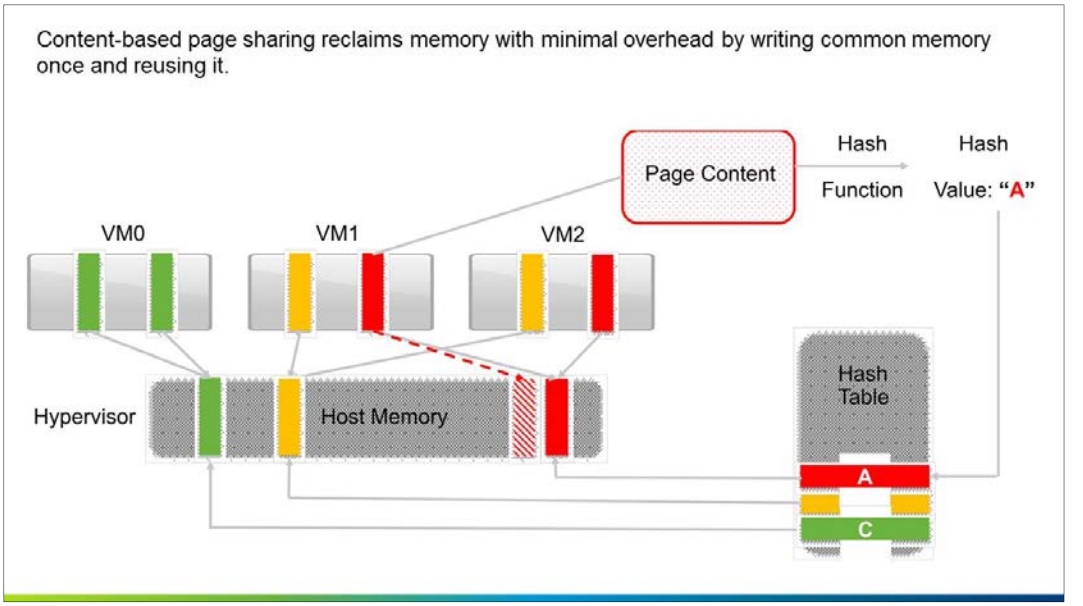
Host swap cache is an optional memory reclamation technique that uses local flash storage to cache a VM's memory pages. By using local flash storage, the VM avoids the latency associated with a storage network that would be used if it swapped memory pages to VSWP files.

https://t.me/learningnets

When severe memory pressure occurs and the hypervisor must swap memory pages to disk, the hypervisor swaps to a host swap cache rather than to a VSWP file. When a host runs out of space on the host cache, a VM's cached memory is migrated to a VM's regular VSWP file. Each host must have its own host swap cache configured.

https://t.me/learningnets

6-13 Transparent Page Sharing (1)



Many physical memory pages on a host often store identical contents. For example, if many VMs are running Windows 2012R2, each VM 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 VMs.

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 VM and inaccessible to the guest OS. Because of this invisibility, sensitive information cannot be leaked from one VM to another.

https://t.me/learningnets

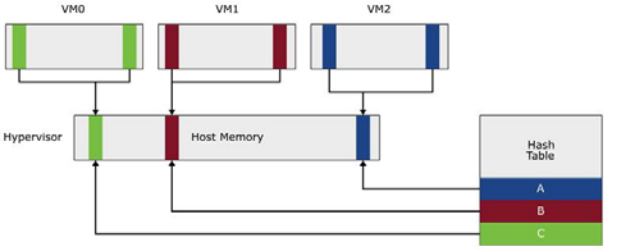
A standard copy-on-write technique is used to handle writes to the shared host physical pages. An attempt to write to the shared pages generates a minor page fault. In the page fault handler, the hypervisor transparently creates a private copy of the page for the VM and remaps the affected guest physical page to this private copy. In this way, VMs can safely modify the shared pages without disrupting other VMs sharing that memory. Writing to a shared page incurs overhead compared to writing to nonshared pages because of the extra work performed in the page fault handler.

https://t.me/learningnets

6-14 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 each VM (intra-VM) is enabled by default.



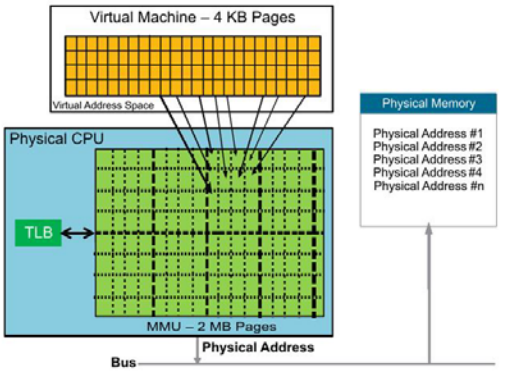
For information about the security risks associated with inter-VM TPS, see VMware knowledge base article 2080735 at http://kb.vmware.com/kb/2080735.

https://t.me/learningnets

6-15 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 scans and shares small pages (4 KB).
- The MMU uses large pages (2 MB).
- 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 lookaside misses.

https://t.me/learningnets

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.

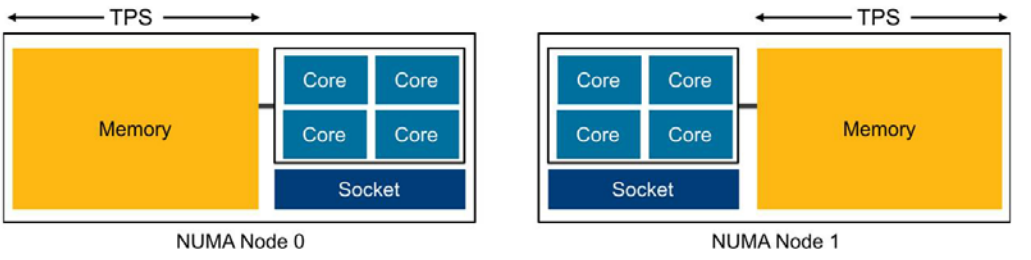
When using EPT or RVI, the esxtop tool might show zero or few shared pages, because TPS uses small pages (4 KB) and EPT and RVI use large pages (2 MB). An ESXi host tracks what pages can be shared. If memory resources become overcommitted, ESXi breaks the large pages into small pages and begins sharing memory between VMs.

https://t.me/learningnets

6-16 Impact of NUMA Architecture on Transparent Page Sharing

TPS for ESXi systems is 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 VMs use shared pages, they do not access memory remotely and therefore do not incur the latency penalty associated with remote memory access.

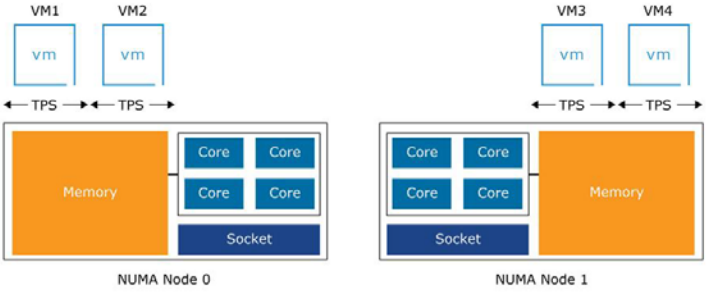


https://t.me/learningnets

6-17 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

6-18 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 security concerns related to 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 Mem.ShareForceSalting host configuration option is used to force the concept of salting. Each VM can have a unique salt value set in the VM'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.

For more information about TPS capabilities, see VMware knowledge base article 2097593 at http://kb.vmware.com/kb/2097593.

https://t.me/learningnets

6-19 Using Salting with Transparent Page Sharing (2)

The host setting is used to achieve a particular goal.

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

https://t.me/learningnets

6-20 Using Salting with Transparent Page Sharing (3)

The 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				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

When Mem.ShareForceSalting is set to 0, salting is not enforced and all VMs have inter-VM TPS enabled in the host.

https://t.me/learningnets

6-21 Using Salting with Transparent Page Sharing (4)

The 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**			Yes
1		**Goal: Inter-VM TPS for All VMs as Default**			Yes
1		Permits Some VMs with No Inter-VM TPS			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

When Mem.ShareForceSalting is 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.

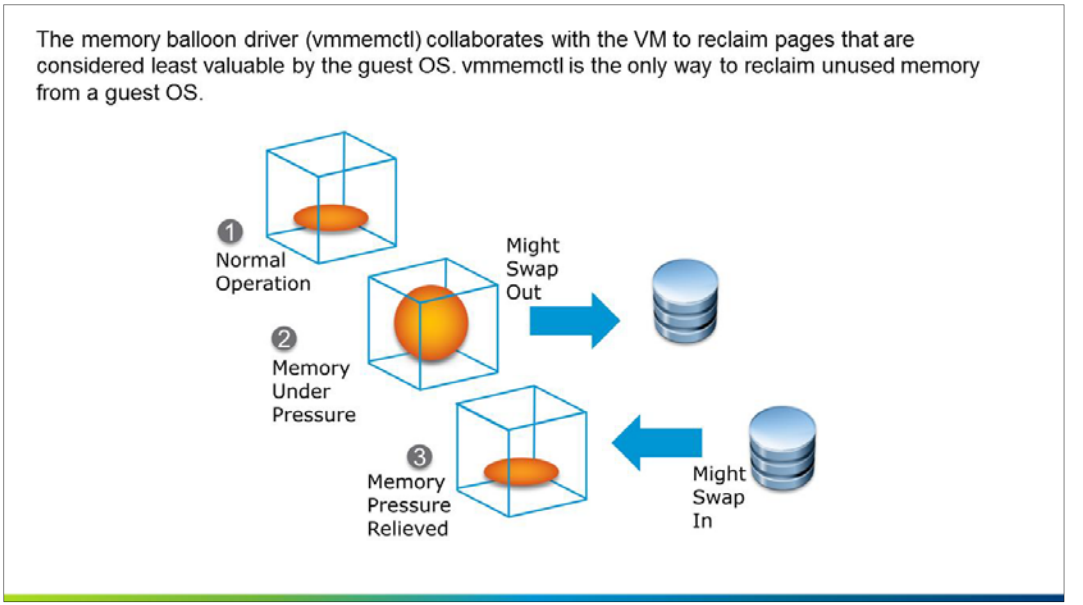
https://t.me/learningnets

6-22 Using Salting with Transparent Page Sharing (5)

The default setting is when Mem.ShareForceSalting is set to 2. 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

6-23 Memory Ballooning in the Guest OS



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 OS, causing the guest to use its own native memory management algorithms. When memory is tight, the guest OS 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:

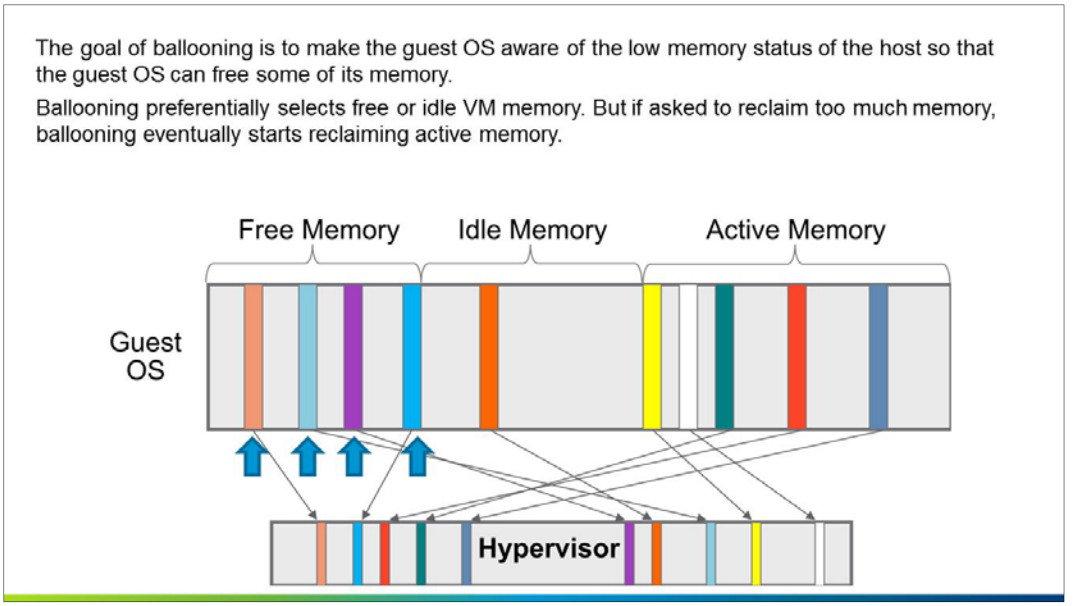
1. Under normal operation, the VMkernel does not need to reclaim memory from the VM. Memory pressure does not exist and ballooning is not active.
2. The balloon driver artificially creates memory pressure on the VM. The guest OS responds to the memory pressure by swapping out the optimal pages according to its analysis.
3. Memory pressure is relieved, ballooning is inactive, and the guest OS swaps pages into memory as needed.

https://t.me/learningnets

If necessary, you can limit the amount of memory that vmmemctl reclaims by setting the `sched.mem.maxmemctl` parameter for a specific VM. This option specifies the maximum amount of memory that can be reclaimed from a VM in megabytes.

https://t.me/learningnets

6-24 Reclaiming Memory with Ballooning

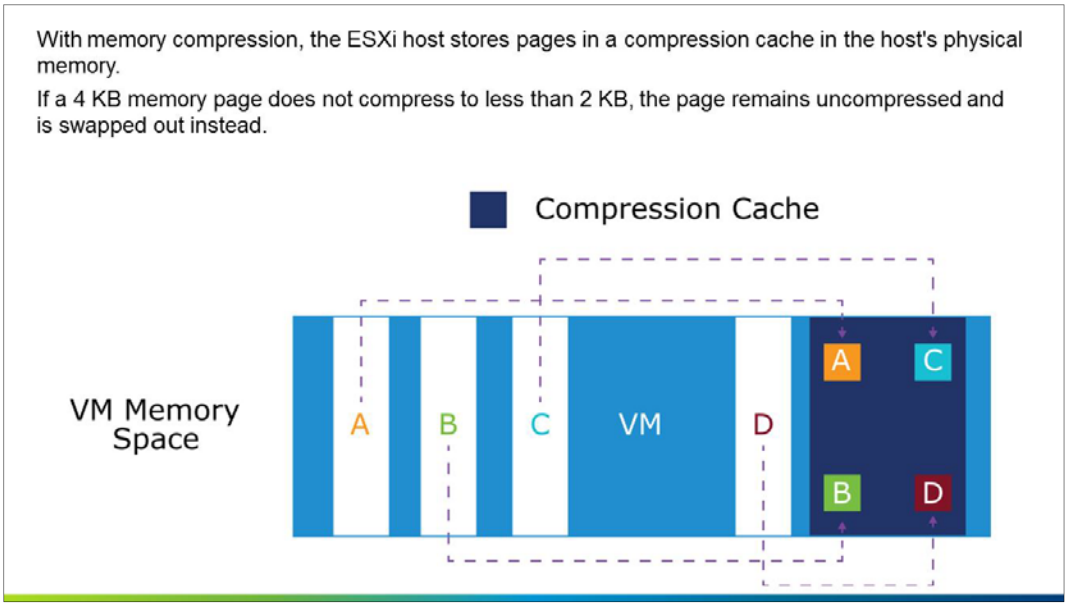


Due to the VM's isolation, the guest OS is unaware that it is running in a VM and unaware of the states of other VMs on the same host. When the hypervisor runs multiple VMs, the total amount of free host physical memory might become low. Because a guest OS cannot detect the host physical memory shortage, none of the VMs free guest physical memory.

The guest OS determines whether it must page out guest physical memory to satisfy the balloon driver's allocation requests. If the VM 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 OS decides which guest physical pages are to be paged to satisfy the balloon driver's allocation requests.

https://t.me/learningnets

6-25 Memory Compression



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 significantly faster than the disk access. ESXi determines whether a page can be compressed by verifying 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.

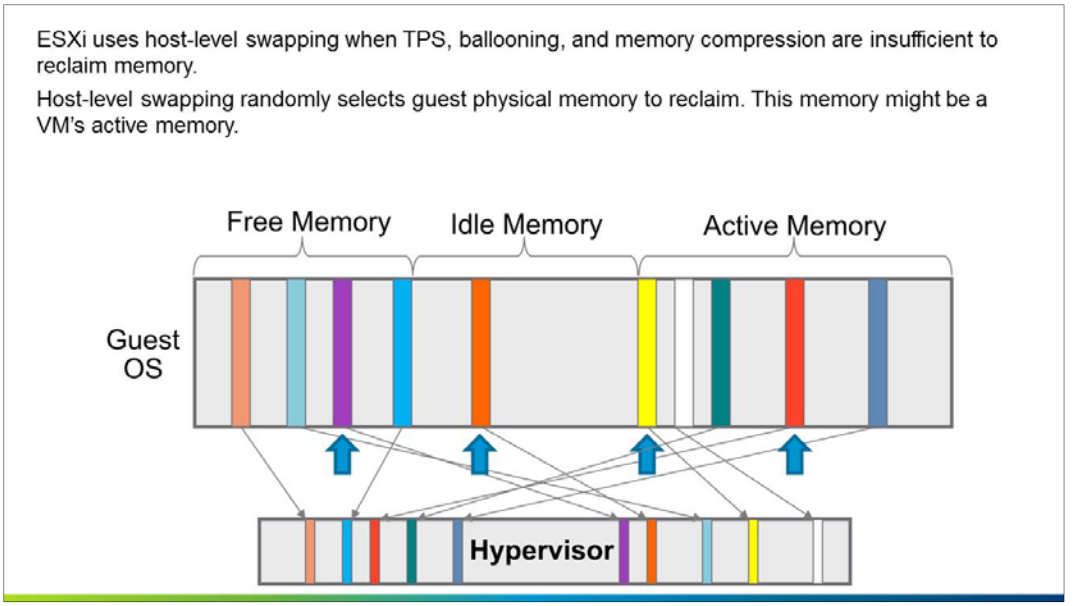
For this example, assume that ESXi must reclaim 8 KB physical memory (that is, two 4 KB pages) from a VM. With memory compression, a swap candidate page is compressed and stored using 2 KB of space in a per-VM compression cache. Each compressed page yields 2 KB memory space for ESXi to reclaim. To reclaim 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.

https://t.me/learningnets

The compression cache size starts at 0 and, by default, can increase 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. Because this setting is applied at the host level, it affects all VMs running on that host.

https://t.me/learningnets

6-26 Reclaiming Memory with Host Swapping



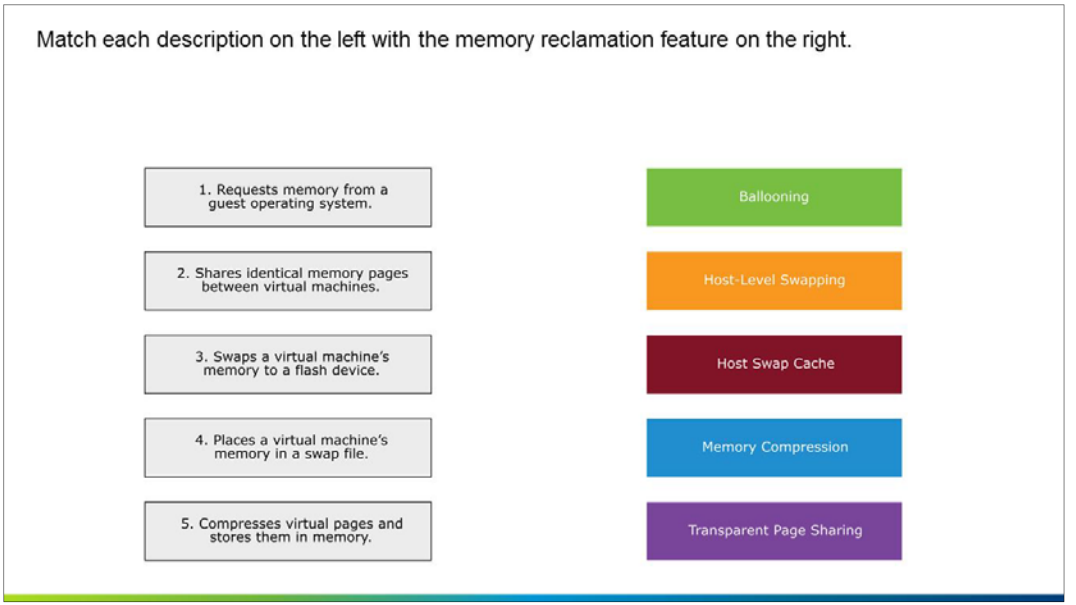
When starting a VM, the hypervisor creates a separate swap VSWP file for the VM to support host-level swapping. Then, if necessary, the hypervisor can directly swap out guest physical memory to the swap file, which frees host physical memory for other VMs.

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 OS. For example, the guest OS will never page out its kernel pages because those pages are critical to ensure guest kernel performance. However, the hypervisor cannot identify those guest kernel pages, so it might swap out those physical pages.

ESXi swapping is distinct from the swapping performed by a guest OS due to memory pressure in a VM. Guest OS level swapping might occur even when the ESXi host has ample resources.

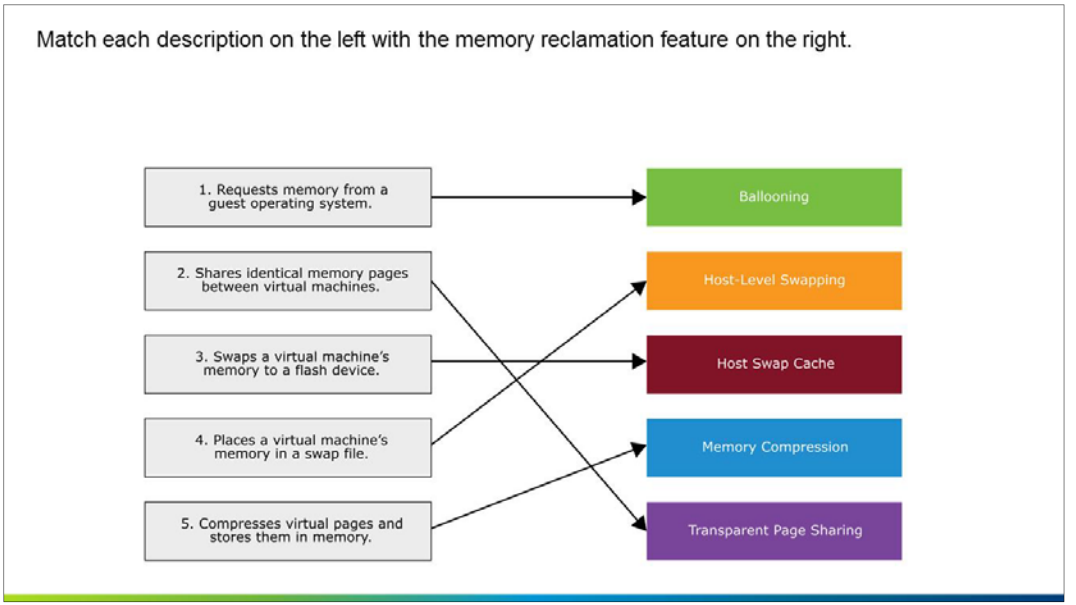
https://t.me/learningnets

6-27 Activity: Identifying Memory Reclamation Features (1)



https://t.me/learningnets

6-28 Activity: Identifying Memory Reclamation Features (2)



https://t.me/learningnets

6-29 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 that 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

6-30 Criteria for Reclaiming Host Memory

ESXi host physical memory is reclaimed based on the high, clear, soft, hard, and low states.

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 VM guest memory use 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 insufficient 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.

https://t.me/learningnets

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 VMs that consume more memory than their target memory allocations.

https://t.me/learningnets

6-31 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 in the following way:

- 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.

https://t.me/learningnets

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* on the Frank Denneman website at https://frankdenneman.nl/.

https://t.me/learningnets

6-32 Transitioning from One Memory State to Another (2)

When memory pressure decreases (that is, when free memory is increasing), a different set of thresholds is 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 the 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 the 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.

https://t.me/learningnets

ESXi transitions from Clear state to the 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* on the Frank Denneman website at https://frankdenneman.nl/.

https://t.me/learningnets

6-33 Review of Learner Objectives

After completing 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

https://t.me/learningnets

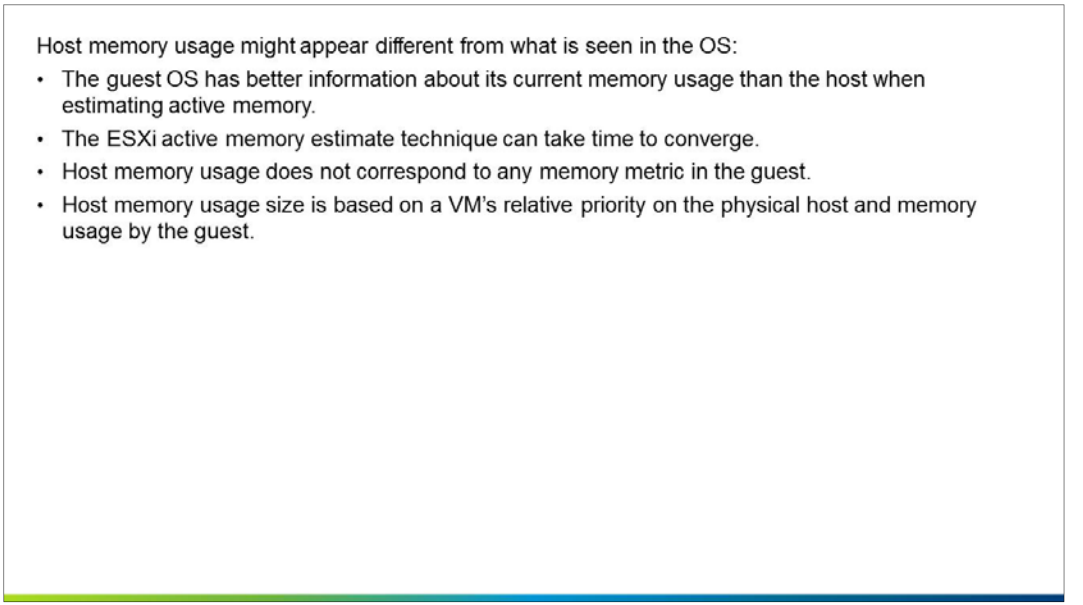
6-35 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Identify key memory metrics to monitor
- Use metrics in the esxtop tool
- Monitor host memory usage and virtual memory usage
- Monitor host swapping activity
- Monitor host ballooning activity

https://t.me/learningnets

6-36 Memory Usage Metrics in the Guest OS

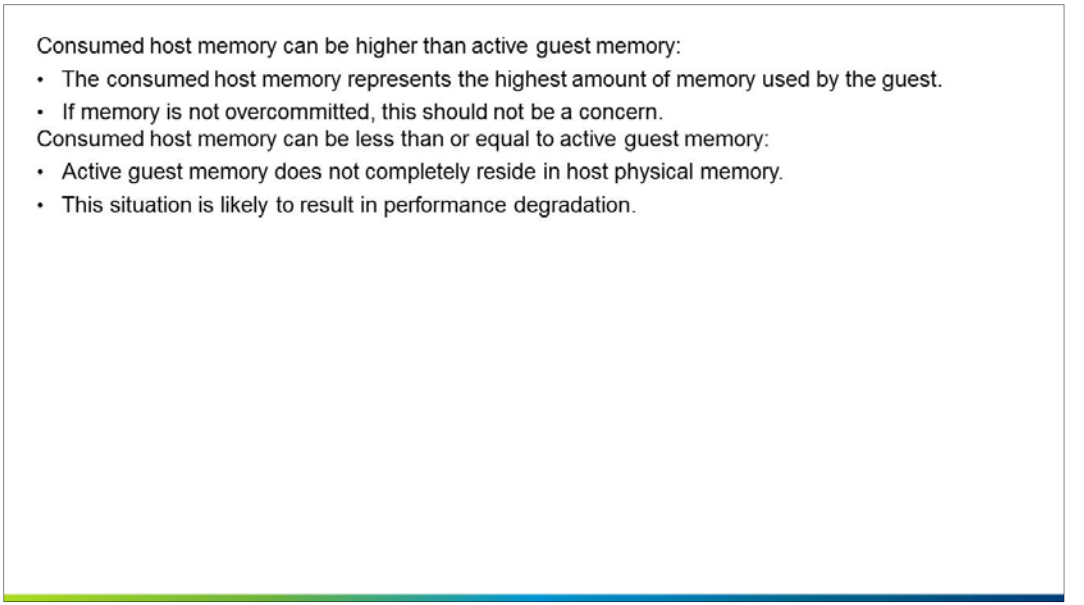


For guest memory usage, the values reported in vSphere Web Client might be different from the active memory usage reported by the guest OS. The first reason for this difference is that the guest OS generally has a better idea than the hypervisor of what memory is active. The guest OS knows what applications 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 OS'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 OS. It has no meaning because the guest OS does not know that it is running in a VM or that other VMs exist on the same physical host.

https://t.me/learningnets

6-37 About Consumed Host Memory and Active Guest Memory



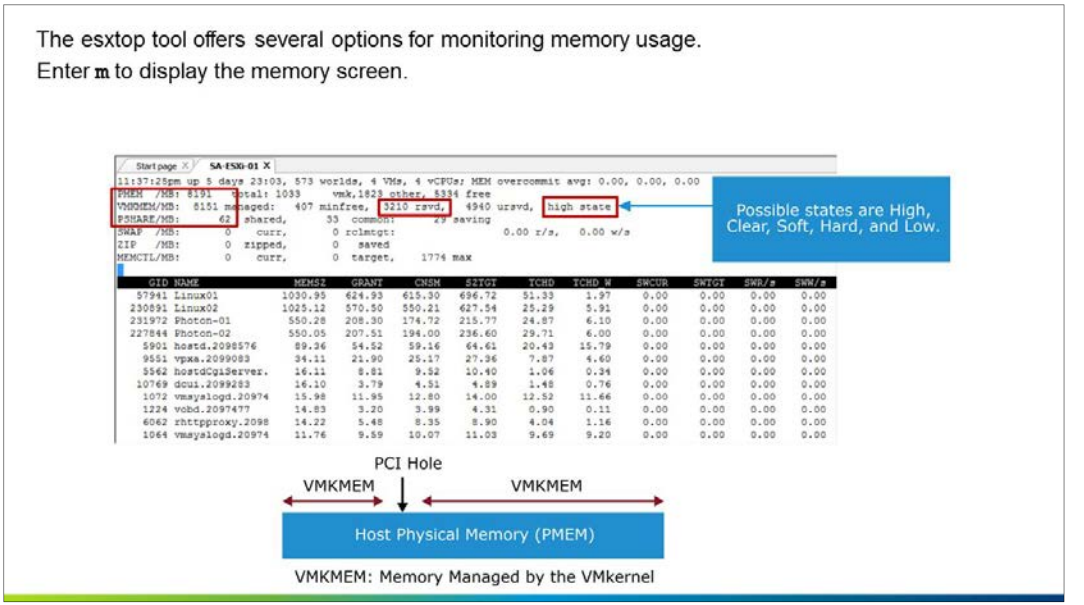
Consumed host memory is greater than active guest memory because, for physical hosts that are not overcommitted on memory, consumed host memory represents the highest amount of memory usage by a VM. In the past, this VM might have actively used a large amount of memory.

Because the host physical memory is not overcommitted, the hypervisor has no reason to invoke ballooning or host-level swapping to reclaim memory. Therefore, 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 VM 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 VM is swapped out by the hypervisor. Either case of lower consumed host memory is probably due to high memory overcommitment.

https://t.me/learningnets

6-38 Using esxtop to Monitor Memory Usage



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 VMs 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 indicates 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.

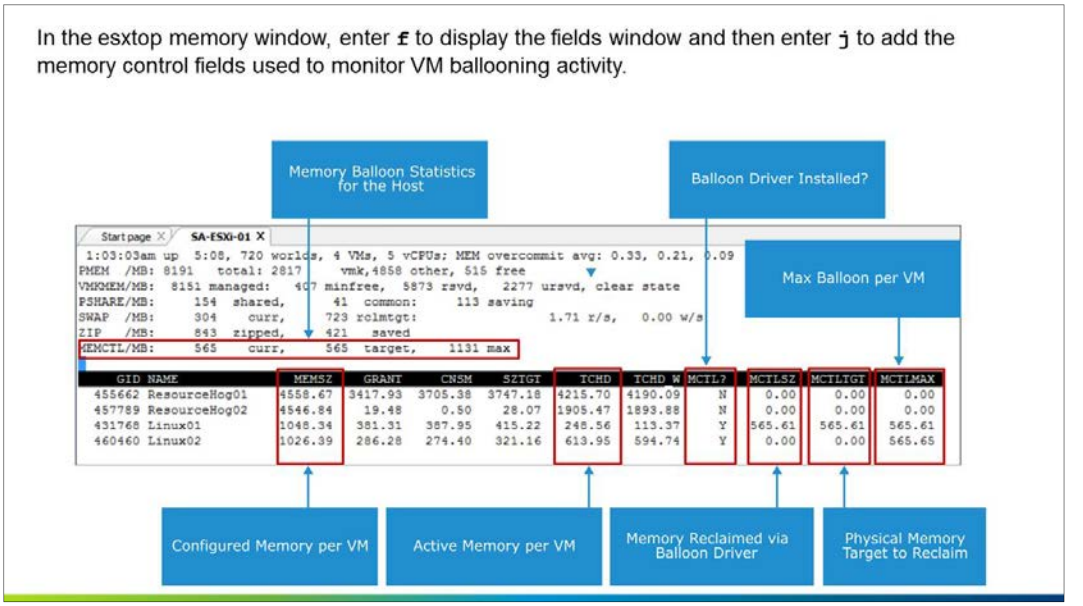
https://t.me/learningnets

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

6-39 Host Ballooning Activity in esxtop (1)



The following metrics are useful in monitoring ballooning activity:

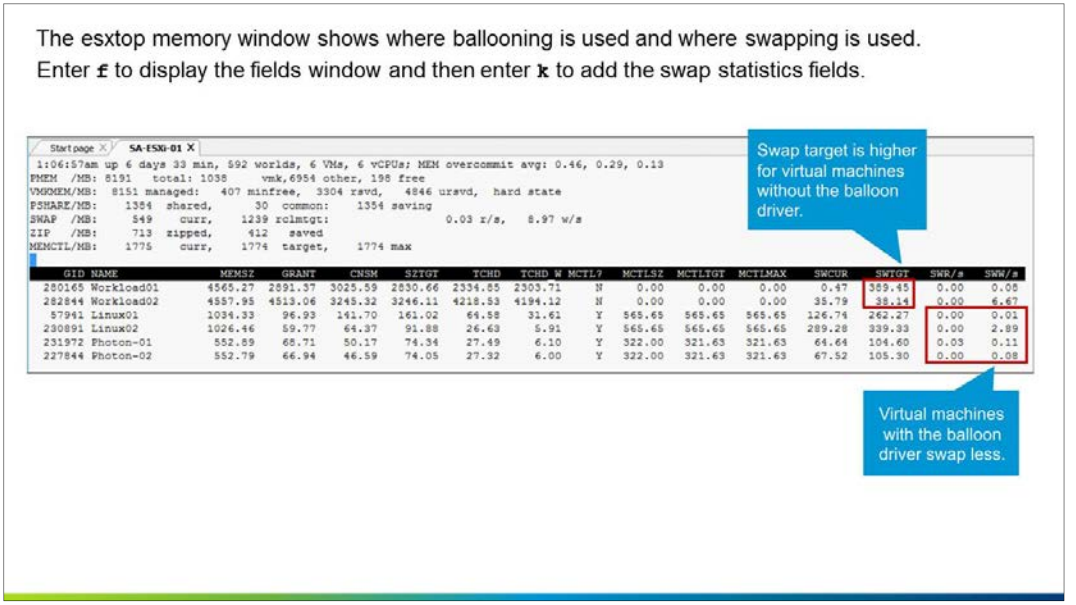
- MEMCTL/MB: This line displays the memory balloon statistics for the entire host. All numbers are in megabytes:
  - Curr: The total amount of physical memory reclaimed using the balloon driver (vmmemctl).
  - Target: The total amount of physical memory that ESXi wants to reclaim with the balloon driver.
  - Max: 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 VM.

https://t.me/learningnets

- MCTLSZ: This value is reported for each VM and represents the amount of physical memory that the balloon driver is holding for use by other VMs.
- MCTLTGT: This value is the amount of physical memory that the host wants to reclaim from the VM through ballooning.

https://t.me/learningnets

6-40 Host Ballooning Activity in esxtop (2)

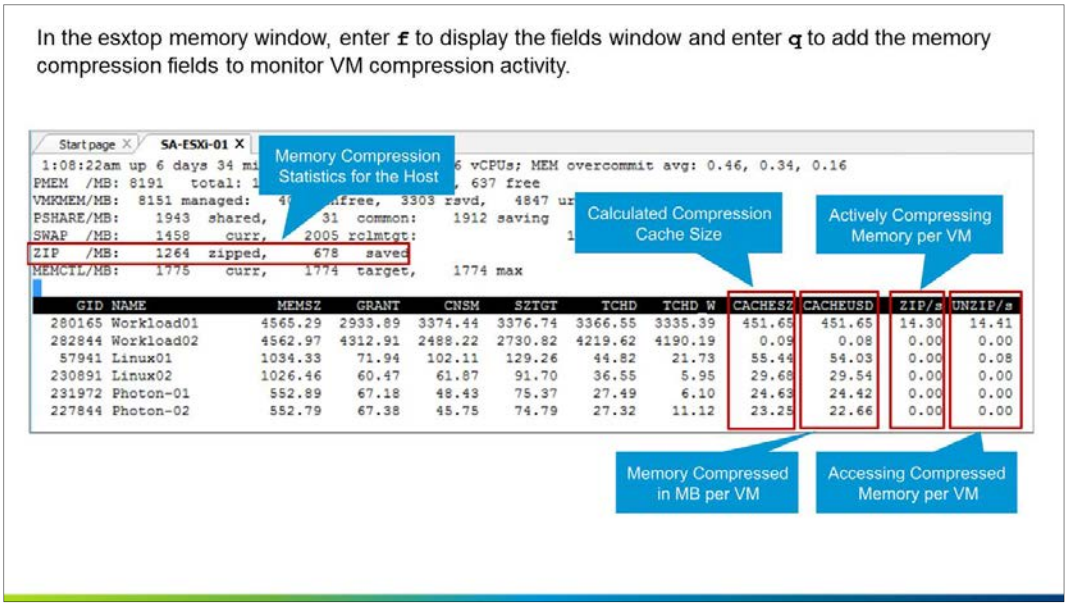


The example shows why ballooning is preferable to swapping. Multiple VMs are running, and all the VMs are running a memory-intensive application. The MCTL? field shows that VMs named `Workload01` and `Workload02` do not have the balloon driver installed. The balloon driver is installed on all the other VMs. If the balloon driver is not installed, typically VMware Tools was not installed on the VM.

The VMs with no balloon driver installed usually have a higher swap target (SWTGT) than the VMs that have the balloon driver installed. VMs that have the balloon driver installed can reclaim memory by using the balloon driver. If ESXi cannot reclaim memory through ballooning, it must resort to other methods, such as host-level swapping.

https://t.me/learningnets

6-41 Host Memory Compression Activity in esxtop



ESXi provides a memory compression cache to improve VM 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 must be swapped, ESXi first tries to compress the page. Pages that can be compressed to 2 KB or smaller are stored in the VM'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

https://t.me/learningnets

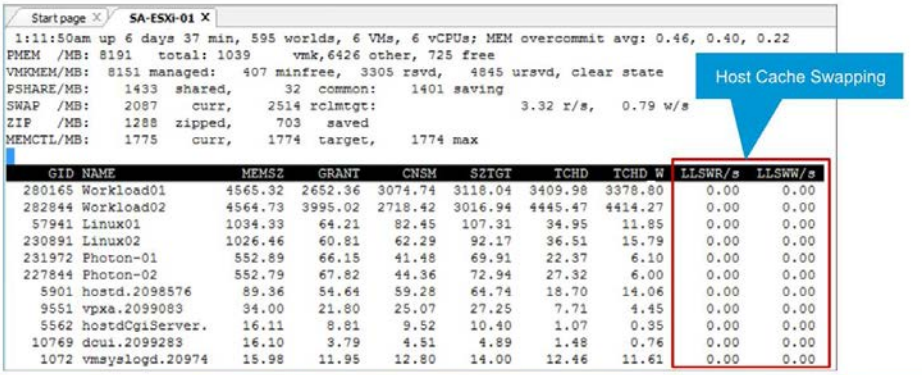
- ZIP/s (MB/s): Compressed memory per second
- UNZIP/s (MB/s): Decompressed memory per second

https://t.me/learningnets

6-42 Monitoring Host Cache Swapping

In the esxtop memory window, enter **f** to display the fields window and enter **1** to add the swap statistics fields to monitor host cache swapping activity.

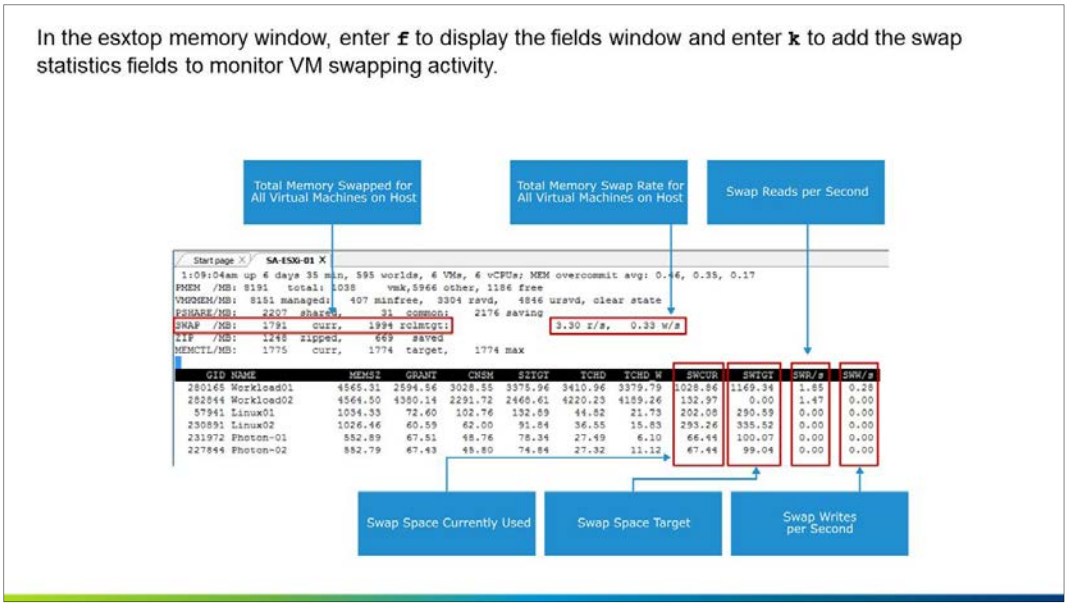
esxtop reports on host cache swapping in the LLSWR/s and LLSWW/s fields.



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

6-43 Host Swapping Activity in esxtop: Memory Screen



A useful metric in the memory screen is SWAP/MB. This metric represents total swapping for all the VMs 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.

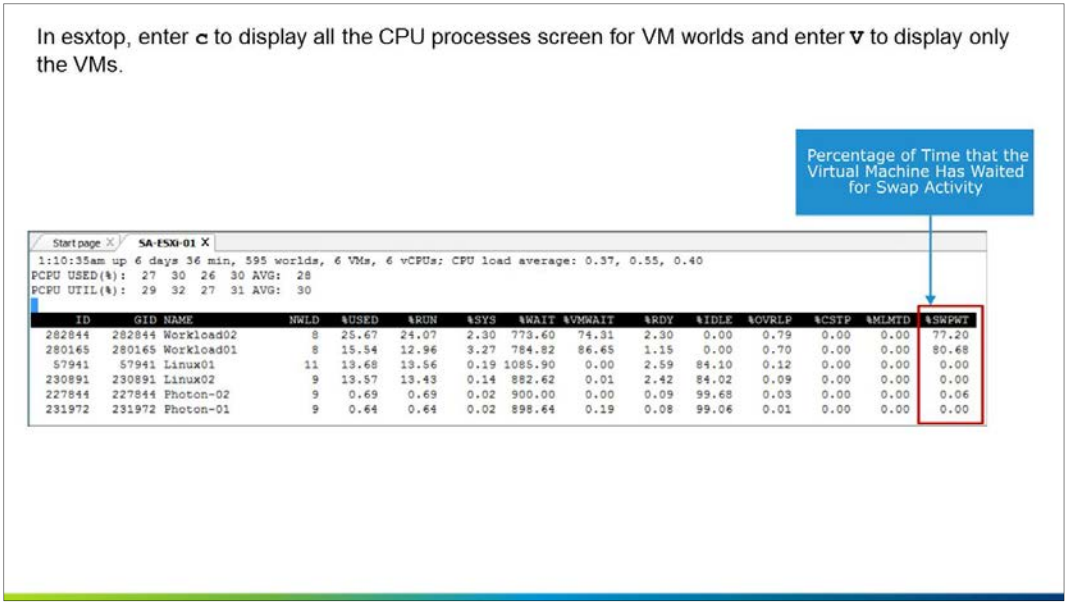
To monitor host-level swapping activity per VM, enter uppercase **V** in the esxtop window to display only the VMs.

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 VM.
- SWTGT: The amount of swap space that the host expects the VM to use.

https://t.me/learningnets

6-44 Host Swapping Activity in esxtop: CPU Screen



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 VM. This metric represents the percentage of time that the VM 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

6-45 Causes of Active Host-Level Swapping

The basic cause of host-level swapping is memory overcommitment from using memory-intensive VMs 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 VMs 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, VM allocations, and VM workload. In addition, the amount of overhead memory must be considered 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 VMs, the host might be forced to swap the memory of other VMs. Swapping memory with other VMs might occur even though VMs are not actively using all their reserved memory.
- Balloon drivers in VMs not running or disabled: In this condition, if the balloon driver is not running or was deliberately disabled in some VMs, the amount of balloonable memory on the

https://t.me/learningnets

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

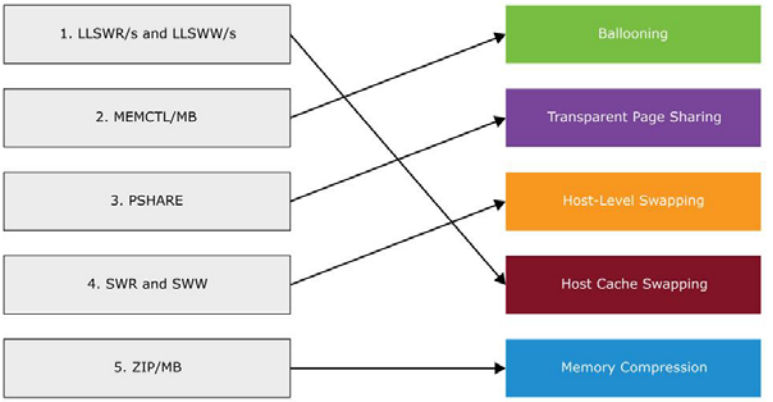
6-46 Activity: Identifying esxtop Fields (1)

Match the esxtop fields on the left with the memory reclamation technique that each monitors on the right.

https://t.me/learningnets

6-47 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

6-48 Resolving Host-Level Swapping

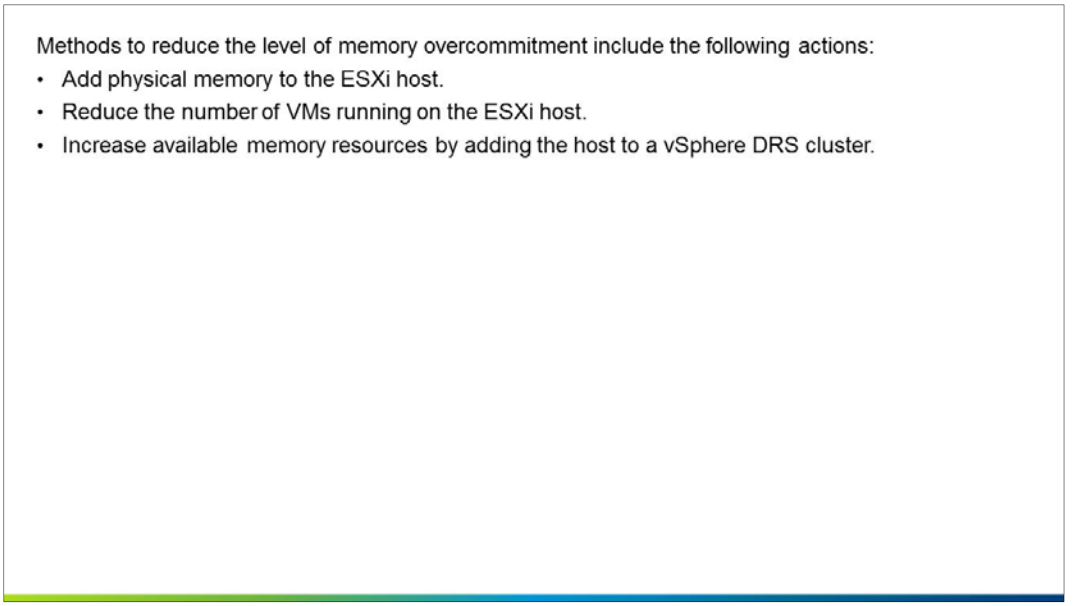
You can resolve performance problems that are caused by active host-level swapping in several ways:

- Reduce the level of memory overcommitment: Use vSphere vMotion.
- Enable the balloon driver in all VMs: Install VMware Tools.
- Add a host to the DRS cluster.
- Add memory to the host.
- Reduce memory reservations: Configured VM memory can be overcommitted, not reserved VM memory.
- Use resource controls to dedicate memory to critical VMs: Use these controls only as a last resort.

In most situations, reducing memory overcommitment levels is the correct 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

6-49 Reducing Memory Overcommitment



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 VMs running on the ESXi host: One way to reduce the number of VMs is to use vSphere vMotion to migrate VMs 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 VM on the host and the available memory resources on the target hosts. You must ensure that the migrated VMs will not cause swapping to occur on the target hosts.

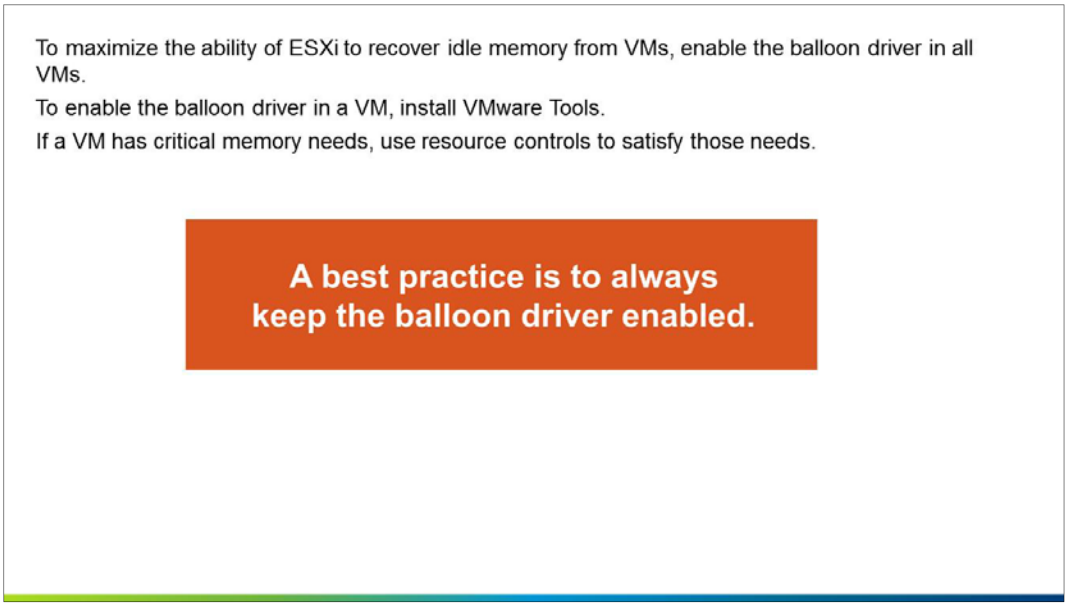
https://t.me/learningnets

If additional ESXi hosts are not available, you can reduce the number of VMs by powering off noncritical VMs. Reducing the number of running VMs makes additional memory resources available for critical applications. Even idle VMs 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 VMs and hosts or account for peak usage periods.

https://t.me/learningnets

6-50 Enabling the Balloon Driver in VMs



If a VM has critical memory needs, 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 VM.

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 VM. 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 VM.

https://t.me/learningnets

6-51 Reducing VM Memory Reservation

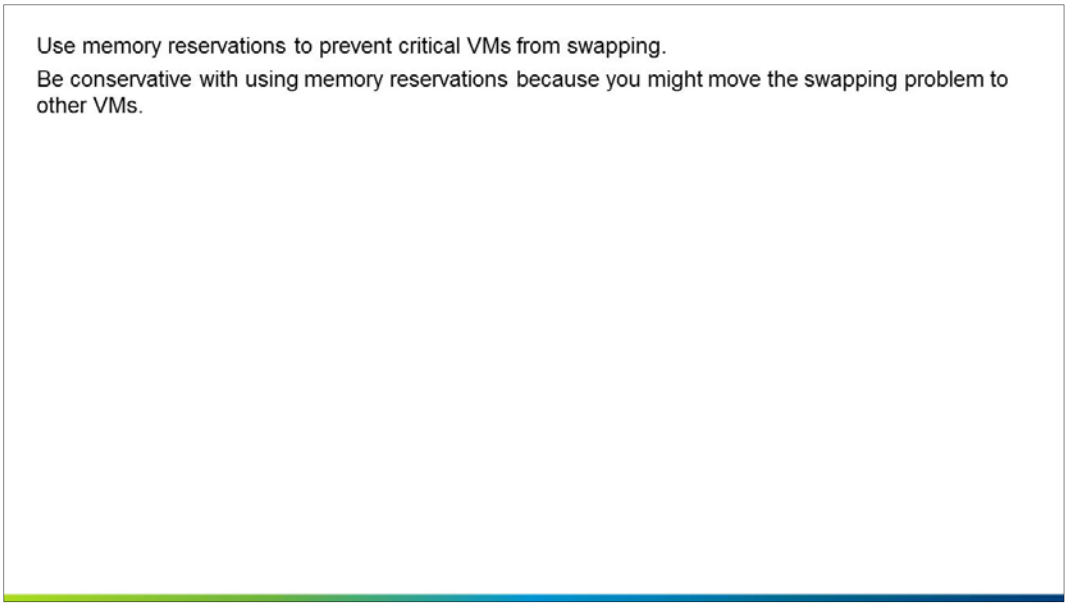
Reevaluate the memory reservation of a VM if this reservation causes the host to swap VMs without reservations.

Reduce the memory reservation of a VM if the VM is not using its full reservation.

If the reservation cannot be reduced, memory overcommitment must be reduced.

https://t.me/learningnets

6-52 Dedicating Memory to Critical VMs

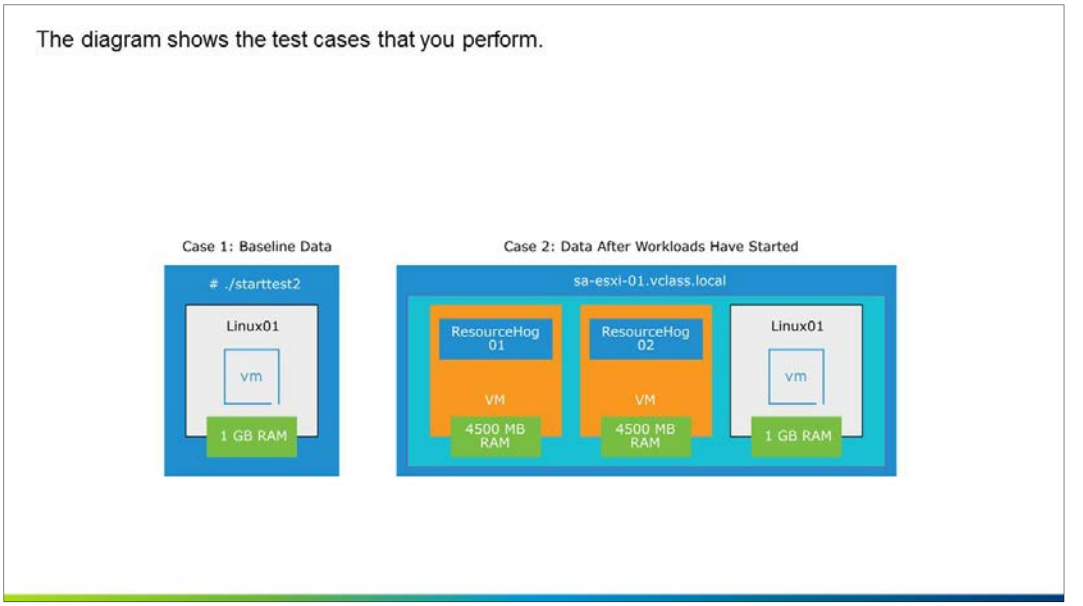


When reducing the level of memory overcommitment is impossible, configure your performance critical VMs with sufficient memory reservations to prevent them from swapping.

However, using memory reservations only moves the swapping problem to other VMs whose performance will be severely degraded. In addition, the swapping of other VMs might still affect the performance of VMs with reservations, because of the added disk traffic and memory management overhead. This approach should be used only with caution when no other options exist.

https://t.me/learningnets

6-53 Introduction to Lab: Monitoring Memory Performance



In this lab, you perform two tests:

- Test case 1: Run starttest2 on your test VM 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

Lab 17: Monitoring Memory Performance

Use the `esxtop` command to monitor memory performance under load:

1. Generate Database Activity in the Test VM
2. Check for Overcommitment of VM Memory
3. Configure esxtop to Report VM 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

6-55 Review of Lab: Monitoring Memory Performance

Your instructor will review the following questions with you:

- For Linux01, did the values for MCTLSZ and MCTLTGT converge?
- For Linux01, did the values for SWCUR and SWTGT converge?
- Which VMs do not have the balloon driver installed?
- Which VMs experienced swapping activity? Did the %SWPWT value exceed 5 percent for any of these VMs?
- 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

6-56 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Identify key memory metrics to monitor
- Use metrics in the esxtop tool
- Monitor host memory usage and virtual memory usage
- Monitor host swapping activity
- Monitor host ballooning activity

https://t.me/learningnets

6-57 Key Points

- The hypervisor uses memory reclamation techniques, such as TPS, ballooning, memory compression, and host-level swapping, are used to reclaim host physical 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 VMs.

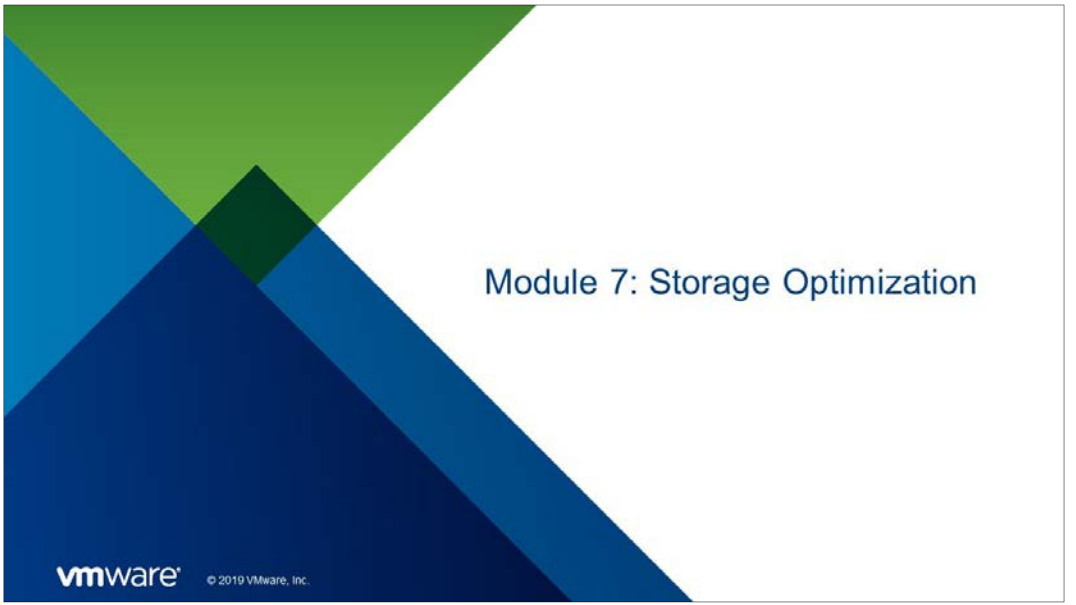
Questions?

https://t.me/learningnets

https://t.me/learningnets

Module 7

Storage Optimization



https://t.me/learningnets

7-2 Importance

Storage performance is one of the most critical aspects of overall workload performance. Monitoring and diagnosing storage performance and throughput, at the ESXi host and VM level, enables you to optimize storage performance and minimize bottlenecks.

https://t.me/learningnets

7-3 Module Lessons

Lesson: Storage Virtualization Concepts
Lesson: Enhancements to NVMe Storage Systems
Lesson: Support for iSER
Lesson: Monitoring Storage Activity

https://t.me/learningnets

https://t.me/learningnets

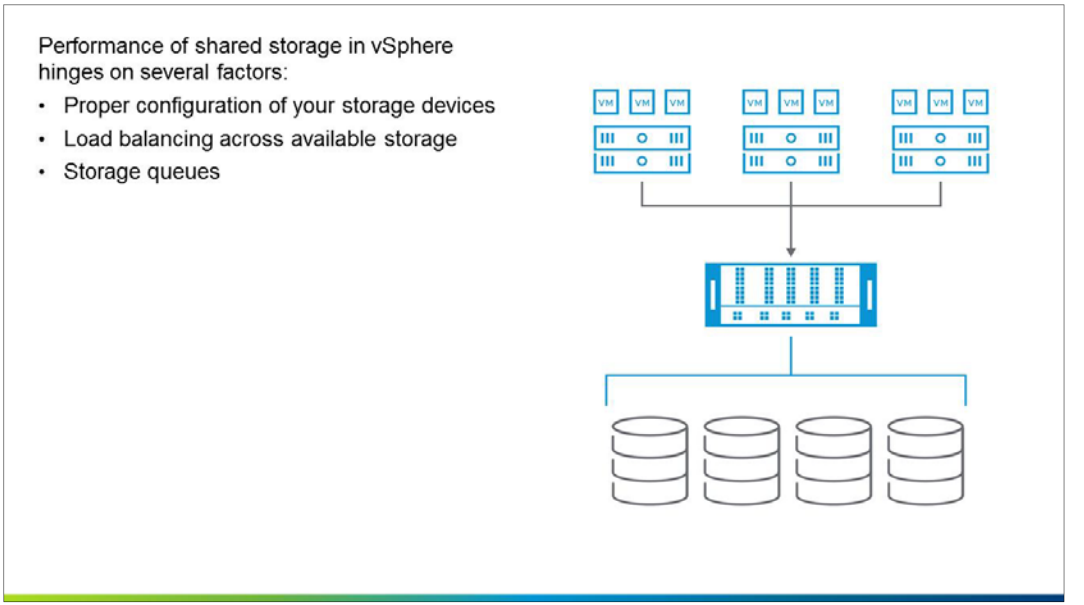
7-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe factors that affect storage performance

https://t.me/learningnets

7-6 Storage Performance Overview

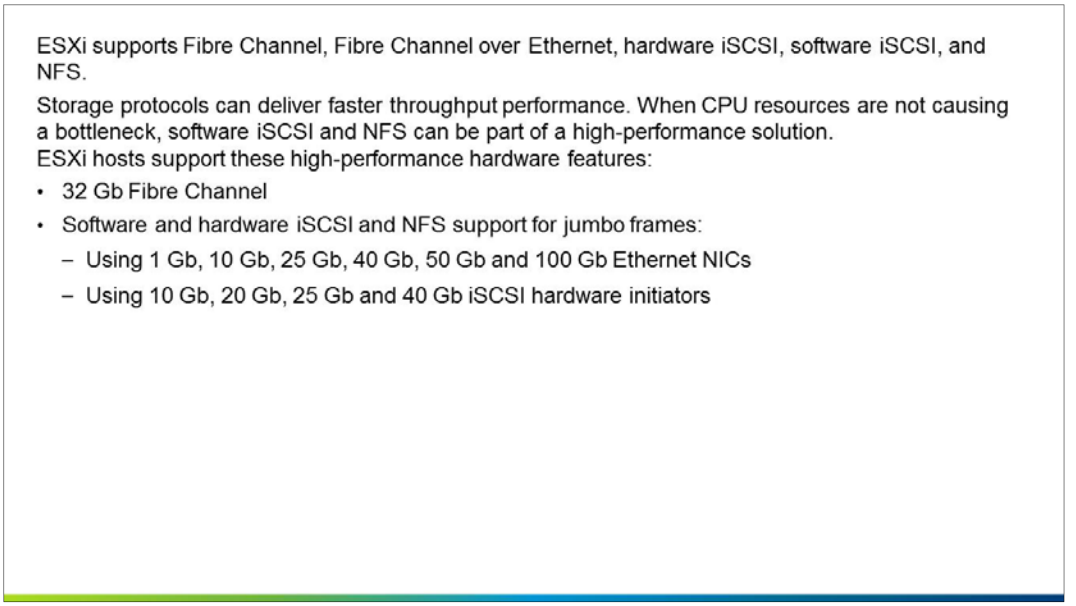


ESXi enables multiple hosts to share physical storage using its optimized storage stack. Shared storage of VMs uses VMFS, NFS, vSAN, and vSphere Virtual Volumes. Shared storage enables vSphere vMotion, vSphere DRS, and vSphere HA.

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

7-7 Storage Protocol Performance

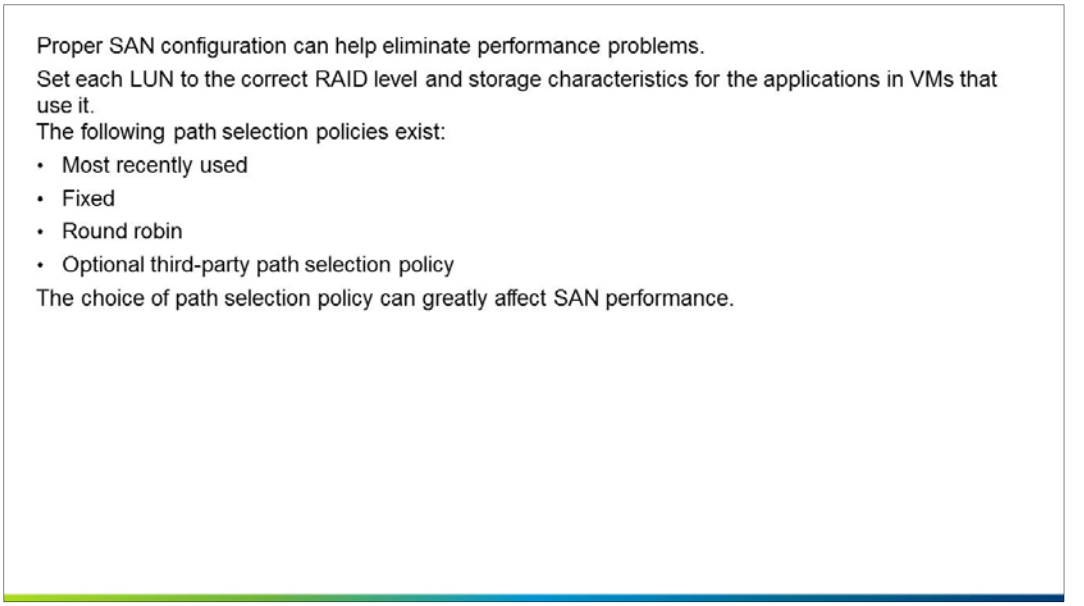


When using Fibre Channel, Fibre Channel over Ethernet (FCoE), and hardware iSCSI, a major part of the protocol processing is offloaded to the host bus adapter (HBA). Therefore, the cost of each I/O is very low.

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 higher cost is because of 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 can deliver high throughput performance when CPU resource is not causing a bottleneck.

https://t.me/learningnets

7-8 About SAN Configuration



Storage performance depends on workload, hardware, vendor, RAID level, cache size, stripe size, and so on. Consult VMware documentation and storage vendor documentation for more information about configuring your storage devices.

Because each application running in your vSphere environment has different requirements, you can improve throughput and minimal latency by selecting the appropriate RAID level and path selection policy.

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.

https://t.me/learningnets

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:

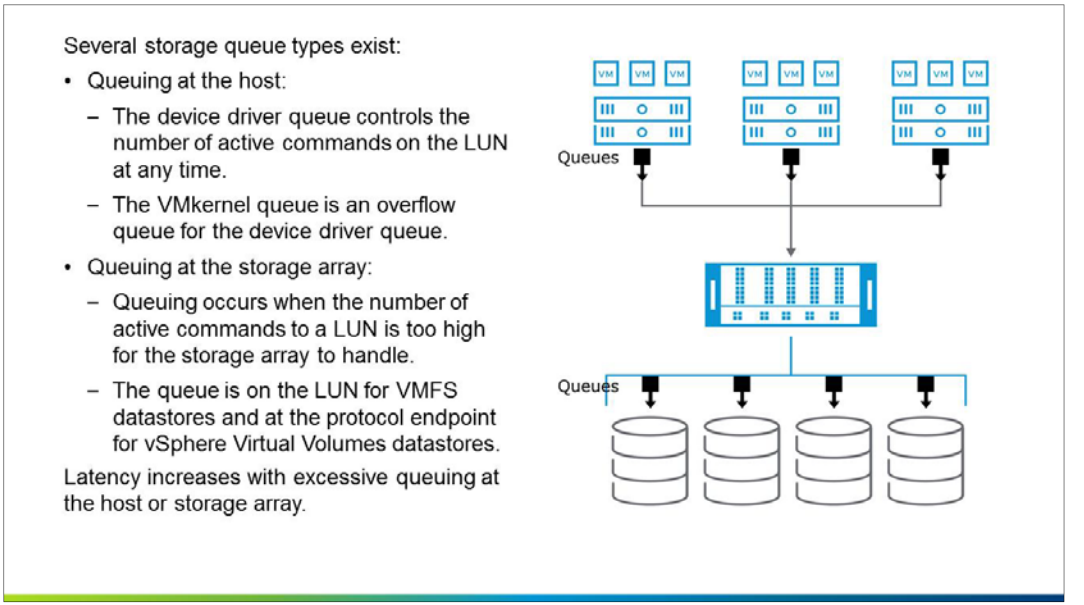
- For active-passive storage arrays, only the paths to the active controller are used in the Round Robin policy.
- For active-active storage arrays, all paths are used in the Round Robin policy.

For more information, see vSphere Storage at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-storage-guide.pdf.

For more information on path selection policies, see *Viewing and Managing Paths* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-98F0851E-12F2-4FB2-A794-7643DD2E6F9B.html.

https://t.me/learningnets

7-9 About Storage Queues



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 command starts.

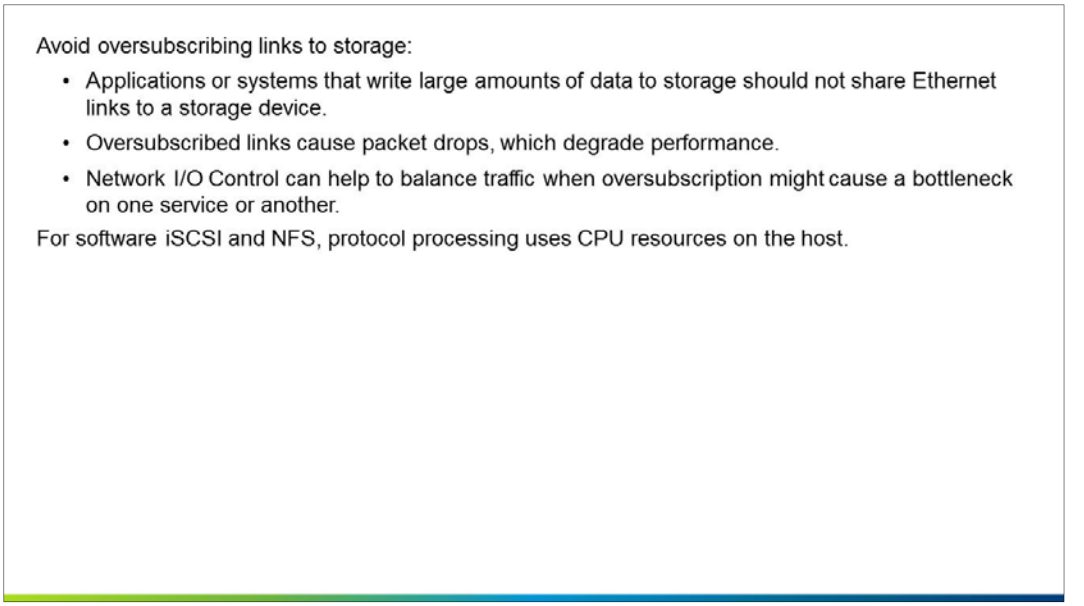
The kernel queue can be thought of as an overflow queue for the device driver queues. A kernel queue includes features that optimize storage. These features include multipathing for failover and load balancing, prioritization of storage activities based on VM and cluster shares, and optimizations to 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 VMs exceeds the LUN queue depth, the excess commands are queued in the ESXi kernel, which increases latency.

In addition to LUN queue depth, command queuing can also occur at the storage array.

https://t.me/learningnets

7-10 Network Storage: iSCSI and NFS



Applications and systems, such as data acquisition or transaction logging systems, perform best with multiple connections to storage devices.

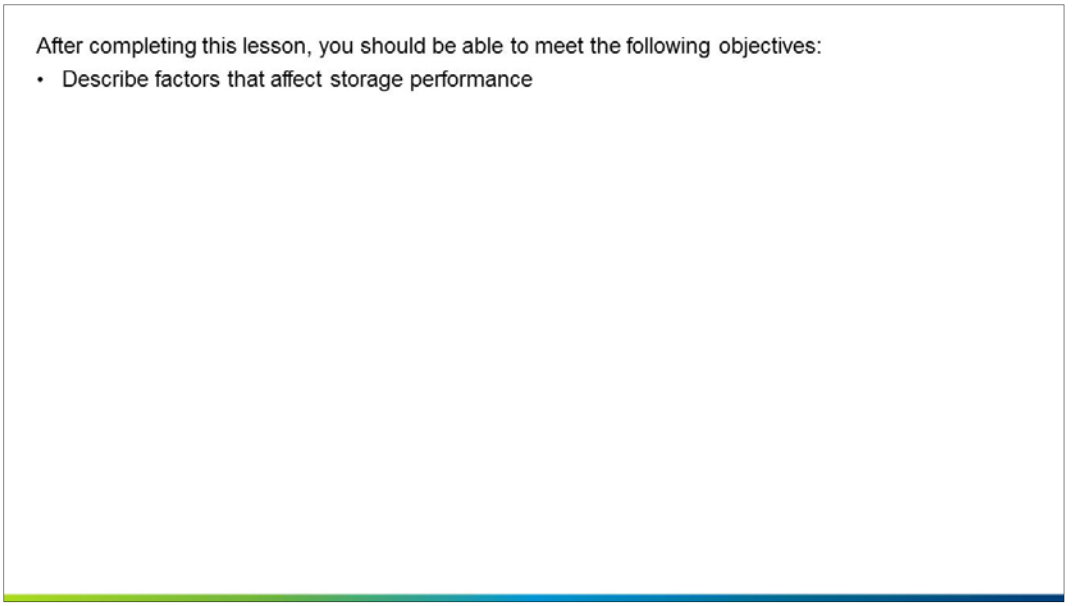
For iSCSI and NFS, design your network topology to prevent Ethernet bottlenecks. Bottlenecks where multiple links are routed through fewer links can result in oversubscription and dropped network packets. When the number of links transmitting near capacity get switched to a smaller number of links, such oversubscription becomes possible. Recovering from dropped network packets results in large performance degradation.

In shared configurations, using VLANs or VPNs is not a suitable solution to the problem of link oversubscription. However, separating VLANs for NFS and iSCSI minimizes network interference from other packet sources.

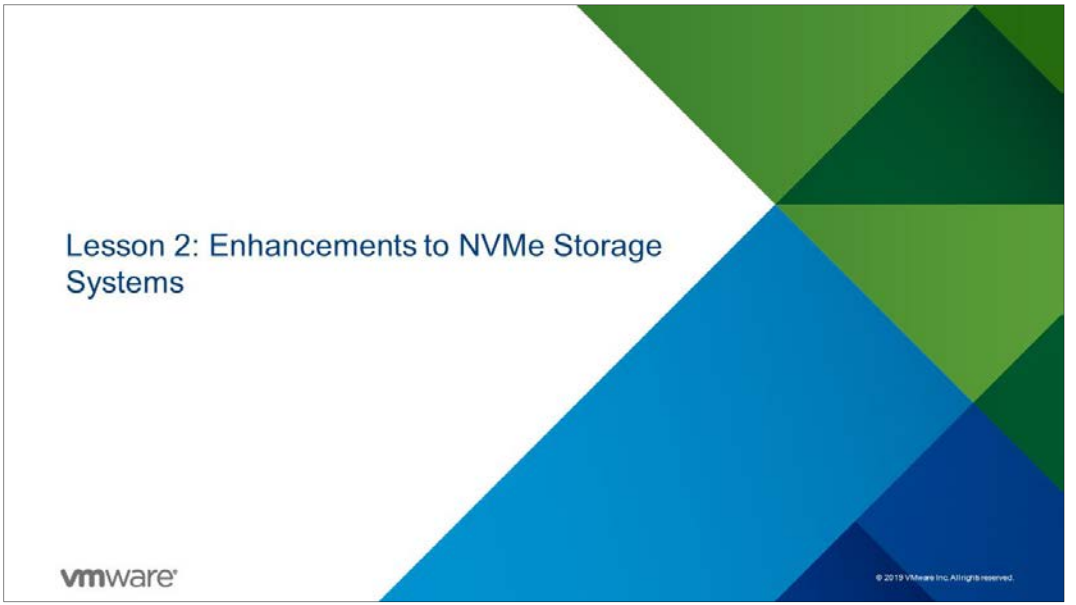
Finally, with software-initiated iSCSI and NFS, the network protocol processing takes place on the host system and therefore requires more CPU resources than other storage options.

https://t.me/learningnets

7-11 Review of Learner Objectives



https://t.me/learningnets



https://t.me/learningnets

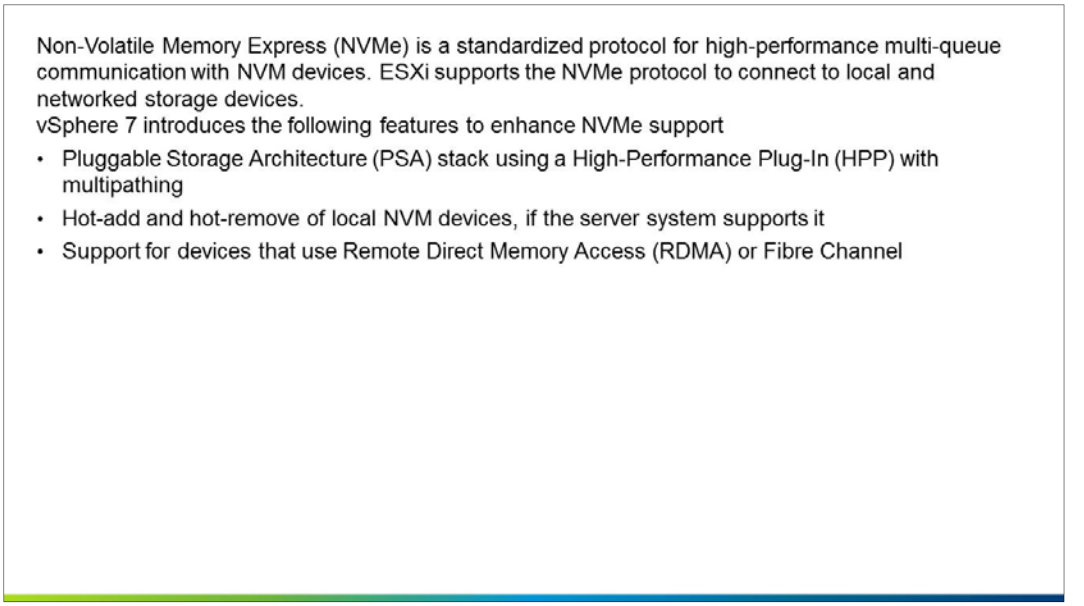
7-13 Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Describe the enhancements to NVMe support
- Configure ESXi for NVMe over RDMA (RoCE v2)
- Describe the enhancements to VMware Pluggable Storage Architecture
- Describe the purpose of the High-Performance Plug-in

https://t.me/learningnets

7-14 Enhanced Support for NVMe



ESXi uses Pluggable Storage Architecture (PSA) as a VMkernel layer for managing storage multipathing.

The PSA is an open and modular framework. It coordinates software modules that are responsible for multipathing operations.

Modules include generic multipathing modules (MPPs) that VMware provides, Native Multipathing Plug-In (NMP) and High-Performance Plug-In (HPP), and third-party MPPs.

https://t.me/learningnets

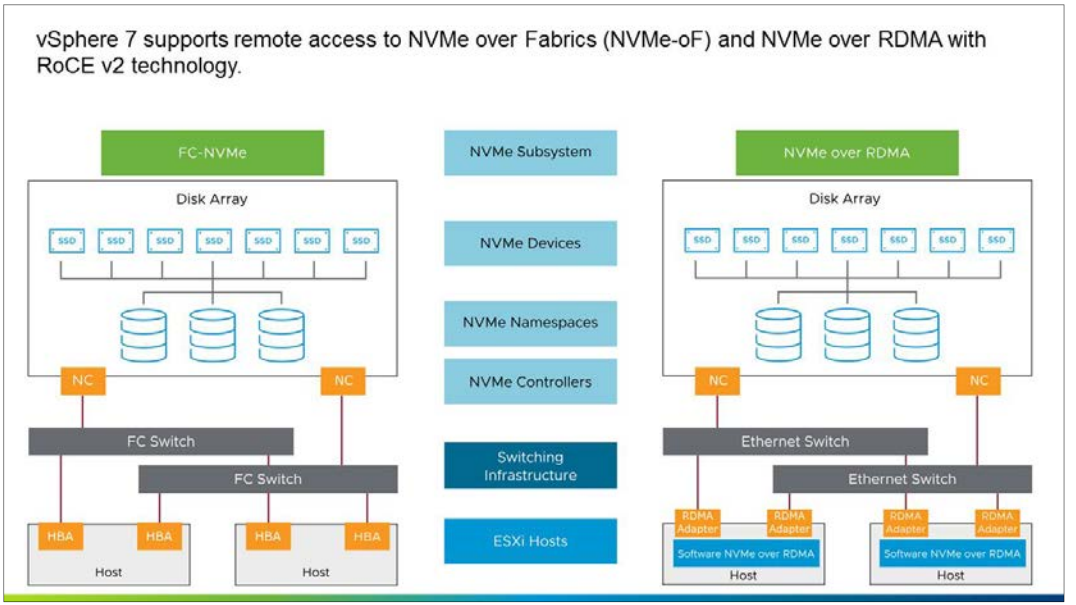
7-15 Supported NVMe Topologies

In vSphere 7, the NVMe storage can be directly attached to a host using a Peripheral Component Interconnect Express (PCIe) interface or indirectly through different fabric transports.

NVMe Transport	ESXi 7 Support
NVMe over PCIe	Local storage
NVMe over RDMA	Shared NVMe over Fabrics (NVMe-oF) storage with RDMA over Converged Ethernet (RoCE) v2 technology
NVMe over Fibre Channel (FC-NVMe)	Shared NVMe-oF storage with Fibre Channel

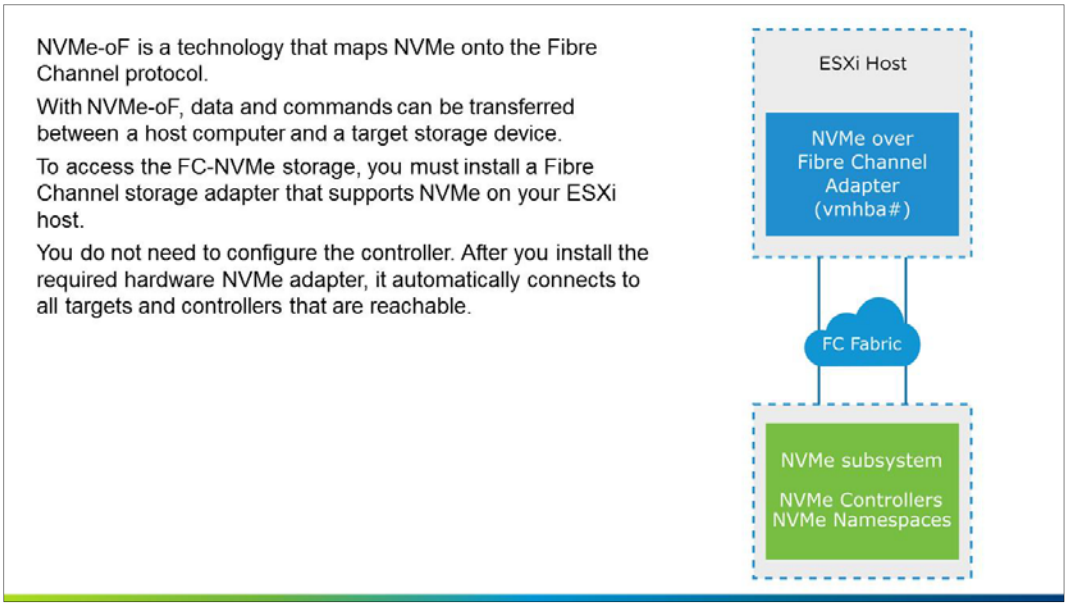
https://t.me/learningnets

7-16 FC-NVMe and NVMe over RDMA (RoCE v2)



https://t.me/learningnets

7-17 Support for FC-NVMe



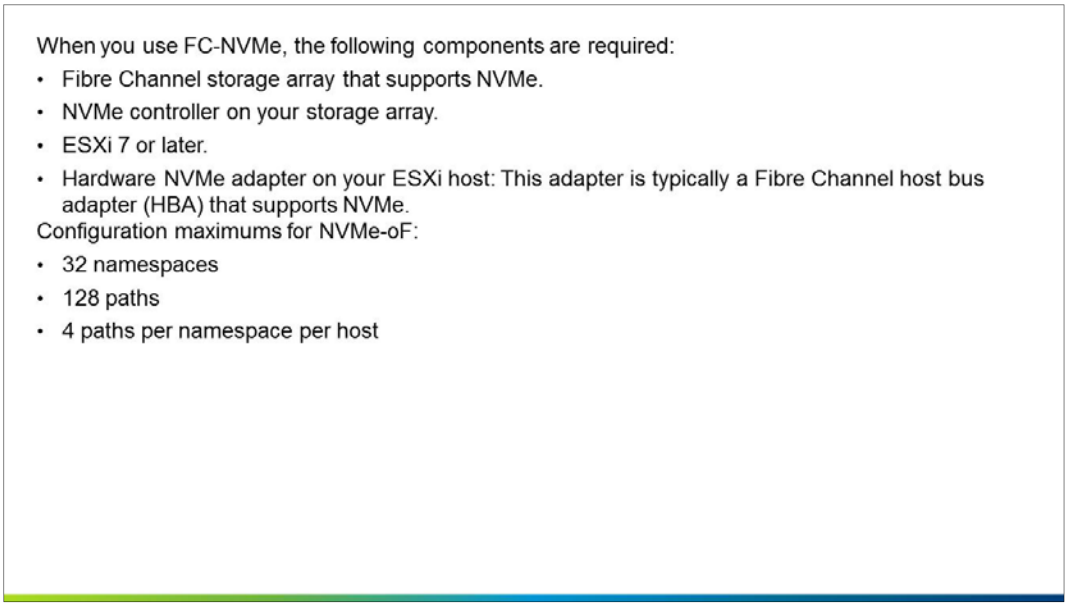
Your ESXi host detects the device and displays it in the vSphere Client as a standard Fibre Channel adapter (vmhba) with the storage protocol indicated as NVMe.

You do not need to configure the controller. After you install the required hardware NVMe adapter, it automatically connects to all targets and controllers that it can reach.

You can later reconfigure the adapter and disconnect its controllers or connect other controllers that were not available during the host boot.

https://t.me/learningnets

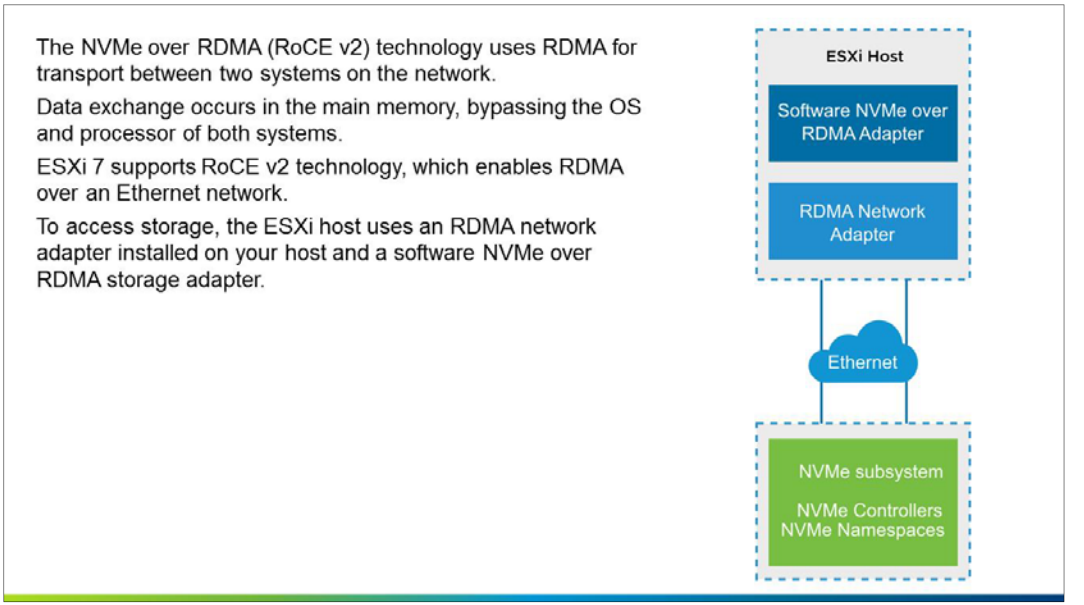
7-18 Requirements for FC-NVMe



For more information about NVMe concepts, see *About VMware NVMe Storage* chapter in the *vSphere Storage* documentation at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-storage-guide.pdf.

https://t.me/learningnets

7-19 Support for NVMe over RDMA (RoCE v2)



NVMe over RDMA with RoCE v2 technology provides ultra-efficient storage to optimize performance-sensitive workloads over your existing Ethernet fabric.

You must configure both adapters to use them for storage discovery. The adapter configuration process involves setting up VMkernel binding for an RDMA network adapter and enabling a software NVMe over RDMA adapter.

For more information about configuring adapters for NVMe over RDMA (RoCE v2) storage, see *Configure Adapters for NVMe over RDMA (RoCE v2) Storage* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-293BC371-B260-4CDD-9BB3-94E834851689.html.

https://t.me/learningnets

7-20 Requirements for NVMe over RDMA (RoCE v2)

NVMe over RDMA (RoCE v2) requires the following components:

- NVMe storage array with NVMe over RDMA (RoCE v2) transport support
- NVMe controller on your storage array
- ESXi 7 or later
- Ethernet switches supporting a lossless network
- Network adapter on your ESXi host that supports RoCE v2
- Software NVMe over RDMA adapter enabled on your ESXi hosts

Configuration maximums for NVMe-oF:

- 32 namespaces
- 128 paths
- 4 paths per namespace per host

https://t.me/learningnets

7-21 Configuring ESXi for NVMe over RDMA (RoCE v2) (1)

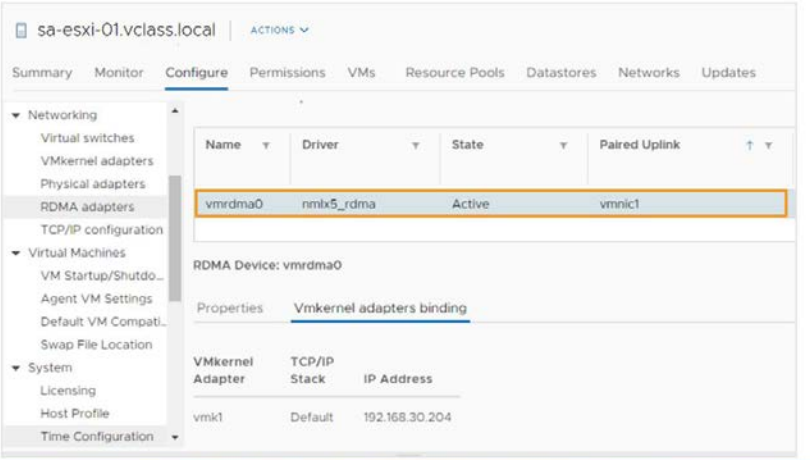
To configure your ESXi host to use NVMe over RDMA (RoCE v2):

1. On your ESXi host, install an adapter that supports RDMA (RoCE v2).
2. In the vSphere Client, verify that the RDMA adapter is discovered by your host.
3. Configure VMkernel binding for the RDMA adapter.
4. Verify the VMkernel binding configuration for the RDMA adapter.
5. Enable the software NVMe over RDMA adapter.
6. Discover NVMe controllers on your array (automatic or manual).
7. Rescan storage to discover NVMe namespaces.
8. Create your datastores.

https://t.me/learningnets

7-22 Configuring ESXi for NVMe over RDMA (RoCE v2) (2)

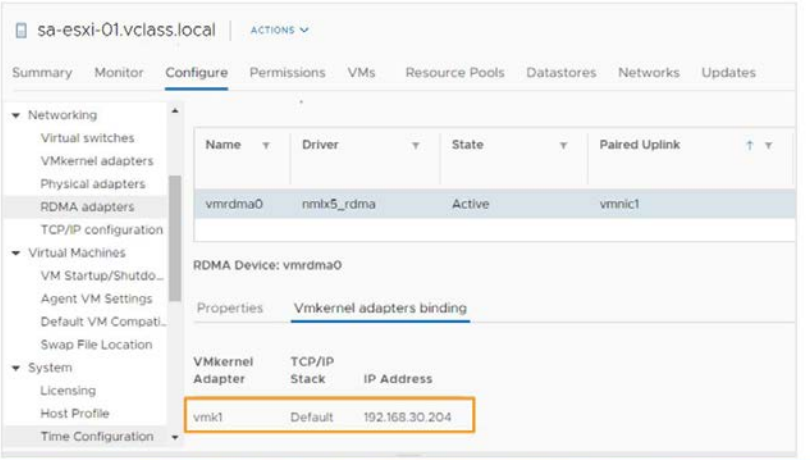
Identify the physical network adapter that corresponds to the RDMA adapter.



https://t.me/learningnets

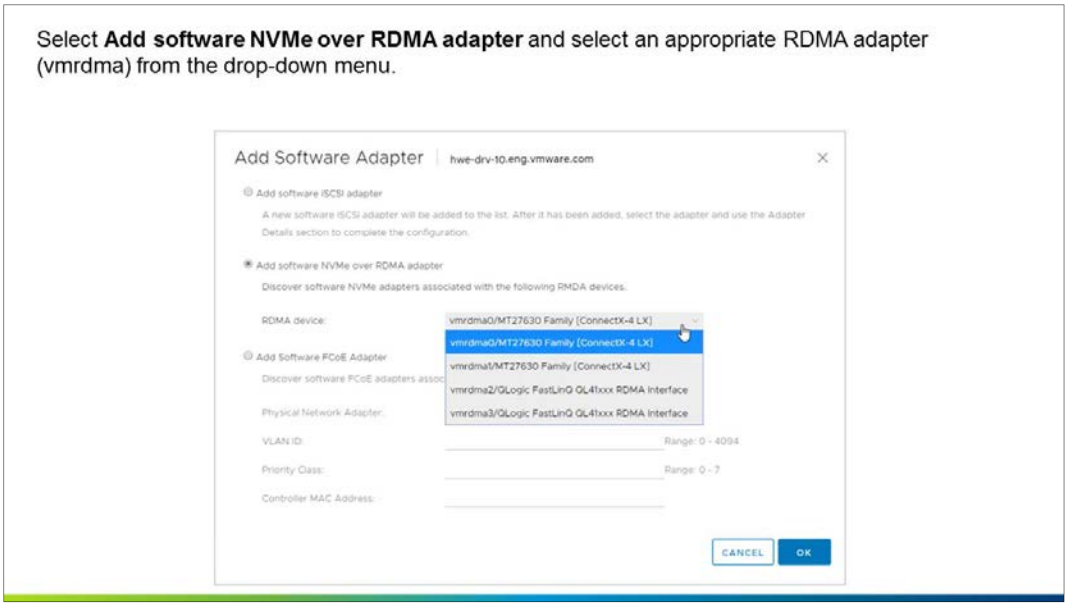
7-23 Configuring ESXi for NVMe over RDMA (RoCE v2) (3)

Configure the VMkernel binding for the RDMA adapter.



https://t.me/learningnets

7-24 Configuring ESXi for NVMe over RDMA (RoCE v2) (4)



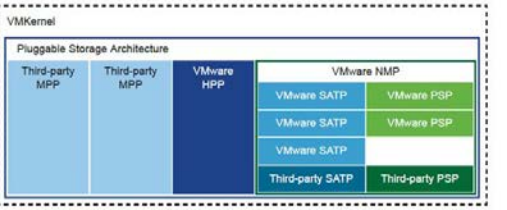
The software NVMe over RDMA adapter appears on the list as a vmhba storage adapter. You can remove the adapter if you must free the underlying RDMA network adapter for other purposes.

For more information about configuring adapters for NVMe over RDMA (RoCE v2) storage, see *Configure Adapters for NVMe over RDMA (RoCE v2) Storage* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-293BC371-B260-4CDD-9BB3-94E834851689.html.

https://t.me/learningnets

7-25 Pluggable Storage Architecture

Pluggable Storage Architecture is an open and modular framework that coordinates various software modules responsible for multipathing operations.



Acronym	Definition
PSA	Pluggable Storage Architecture
NMP	Native Multipathing Plug-in: Generic VMware multipathing module that is used with SCSI storage devices
PSP	Path Selection Plug-in: Handles path selection for a SCSI storage device
SATP	Storage Array Type Plug-in: Handles path failover for a given SCSI storage array

Acronym	Definition
MPP (third party)	Multipathing Plug-in: A multipathing module developed and provided by a third party
HPP	High-Performance Plug-in provided by VMware: Used with ultra-fast local and networked flash devices
PSS	Path Selection Scheme: Handles multipathing for NVMe storage devices

VMware provides generic native multipathing modules, called VMware NMP and VMware HPP.

In addition, the PSA offers a collection of VMkernel APIs that third-party developers can use.

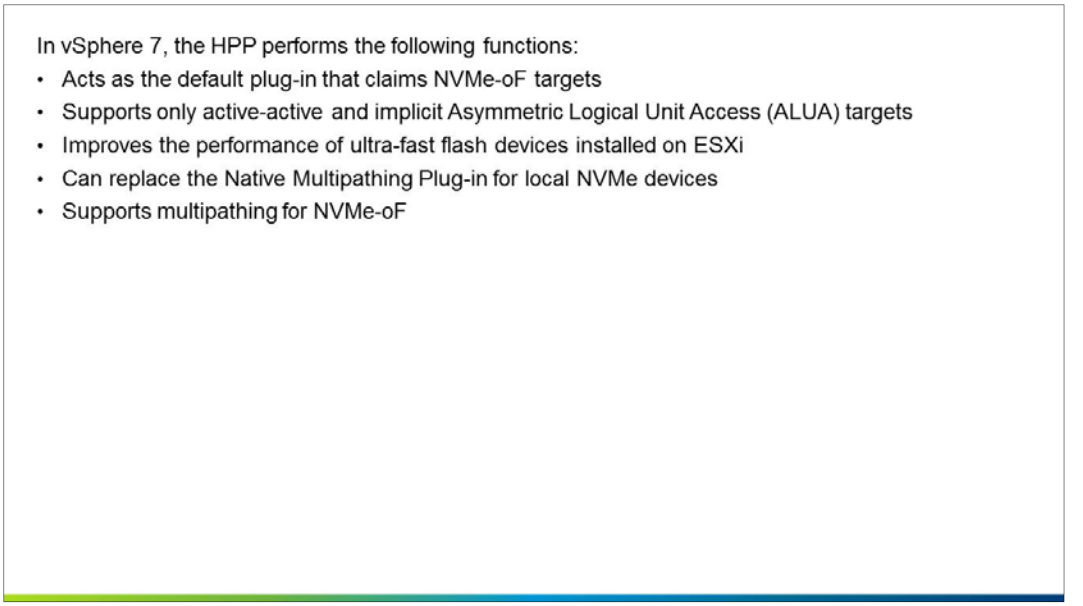
Software developers can create their own load balancing and failover modules for a particular storage array. These third-party multipathing modules (MPPs) can be installed on the ESXi host and run in addition to the VMware native modules, or as their replacement.

As the diagram on the slide shows, multiple third-party MPPs can run in parallel with the VMware NMP or HPP.

When installed, the third-party MPPs can replace the behavior of the native modules. The MPPs can take control of the path failover and the load-balancing operations for the specified storage devices.

https://t.me/learningnets

7-26 High-Performance Plug-In for NVMe



In vSphere 7, HPP integrates path selection and failover logic within a single module, reducing cross-module calls.

To support multipathing, HPP uses a Path Selection Scheme (PSS) when selecting physical paths for I/O requests. You can use vSphere ESXi Shell or the vSphere CLI to configure the HPP and PSS.

For local NVMe devices, NMP remains the default plug-in, but you can replace it with HPP.

For more information about configuring and using HPP, see *VMware High Performance Plug-In and Path Selection Schemes* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-F7B60A5A-D077-4E37-8CA7-8CB912173D24.html.

https://t.me/learningnets

7-27 Support for High-Performance Plug-In Multipathing

The HPP improves the performance of NVMe devices on your ESXi host.

HPP Support	vSphere 7
Storage devices	Local NVMe PCIe Shared NVMe-oF (active-active and implicit ALUA targets only)
Multipathing	Yes
Second-level plug-ins	No Path Selection Scheme (PSS)
SCSI-3 persistent reservations	No
4Kn devices with software emulation	No
vSAN	No

https://t.me/learningnets

7-28 High-Performance Plug-In Best Practices

To achieve the fastest throughput from a high-speed storage device, follow these recommendations:

- Use the HPP for NVMe local or networked devices.
- Do not activate the HPP for HDDs or slower flash devices.
- When using FC-NVMe, follow general recommendations for Fibre Channel storage.
- When using NVMe-oF, do not mix transport types to access the same namespace.
- Before registering NVMe-oF namespaces, ensure that active paths are presented to the host.
- Configure your VMs to use VMware Paravirtual controllers.
- Configure your VMs to use high-latency sensitivity.
- If a single VM drives a significant share of the device's I/O workload, consider spreading the I/O across multiple virtual disks. Attach the disks to separate virtual controllers in the VM.

Do not activate the HPP for HDDs or slower flash devices. The HPP does not provide any performance benefits with devices incapable of at least 200,000 IOPS.

If you do not attach the disks to separate virtual controllers in the VM, I/O throughput might be limited. This limitation is due to saturation of the CPU core that is responsible for processing I/O operations on a virtual storage controller.

For information about device identifiers for NVMe devices that support only NGUID ID format, see *NVMe Devices with NGUID Device Identifiers* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-E38066BA-8AC1-4A5D-8E2C-EB24FCC8E9E1.html.

https://t.me/learningnets

7-29 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Describe the enhancements to NVMe support
- Configure ESXi for NVMe over RDMA (RoCE v2)
- Describe the enhancements to VMware Pluggable Storage Architecture
- Describe the purpose of the High-Performance Plug-in

https://t.me/learningnets

https://t.me/learningnets

7-31 Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Describe support for iSER in vSphere 7
- Configure iSER on ESXi hosts

https://t.me/learningnets

7-32 Support for iSER-Enabled vSphere Virtual Volumes Arrays

iSCSI Extension for RDMA (iSER) provides high-performance iSCSI connectivity for standard RDMA NICs.

In addition to supporting traditional iSCSI, ESXi supports the iSER protocol. When the iSER protocol is enabled, the iSCSI framework on the ESXi host can use the RDMA transport instead of TCP/IP.

vSphere 7 also adds iSER support for vSphere Virtual Volumes.
In vSphere 7, support for iSER requires the following:

- RoCE-enabled devices
- A target array that supports iSER on RoCE

https://t.me/learningnets

7-33 Configuring iSER

Configure iSER on your ESXi host so that the iSCSI framework on the host can use Remote Direct Memory Access (RDMA) transport instead of TCP/IP.

Configuration Step	Description
Enable the VMware iSER adapter.	Use the `esxcli` command to enable the VMware iSER adapter.
Modify general properties for iSCSI or iSER adapters.	If needed, change the default name and alias assigned to your adapter.
Configure port binding for iSCSI or iSER.	You must create network connections to bind the iSER engine and the RDMA-capable network adapter.
Configure dynamic or static discovery for iSCSI and iSER.	With the dynamic discovery, each time the initiator contacts a specified iSER storage system, it sends the `SendTargets` request to the system. With static discovery, manually enter information for the targets.
Set up CHAP for iSCSI or iSER.	This step is necessary only if your environment uses the Challenge Handshake Authentication Protocol (CHAP).
Enable jumbo frames for networking.	If your environment supports jumbo frames, enable them for the adapter.

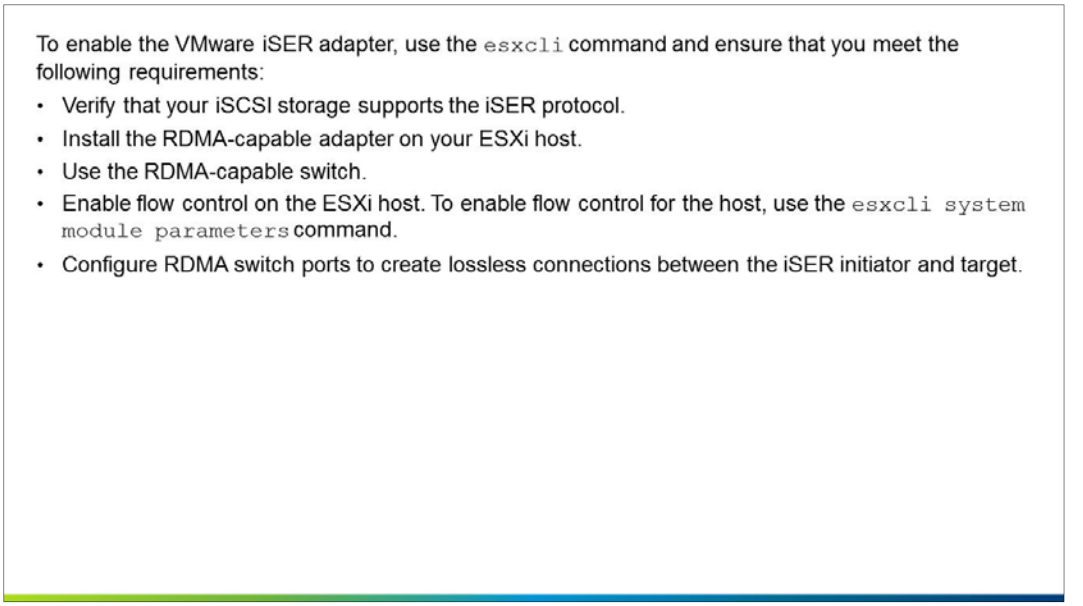
When installed on the host, the RDMA-capable adapter appears in vCenter Server as a network adapter (vmnic).

To make the adapter functional, you must enable the VMware iSER component and connect the iSER adapter to the RDMA-capable vmnic. You can then configure properties, such as targets and CHAP, for the iSER adapter.

NOTE: iSER does not support NIC teaming. When configuring port binding, use only one RDMA adapter per virtual switch.

https://t.me/learningnets

7-34 Requirements for iSER



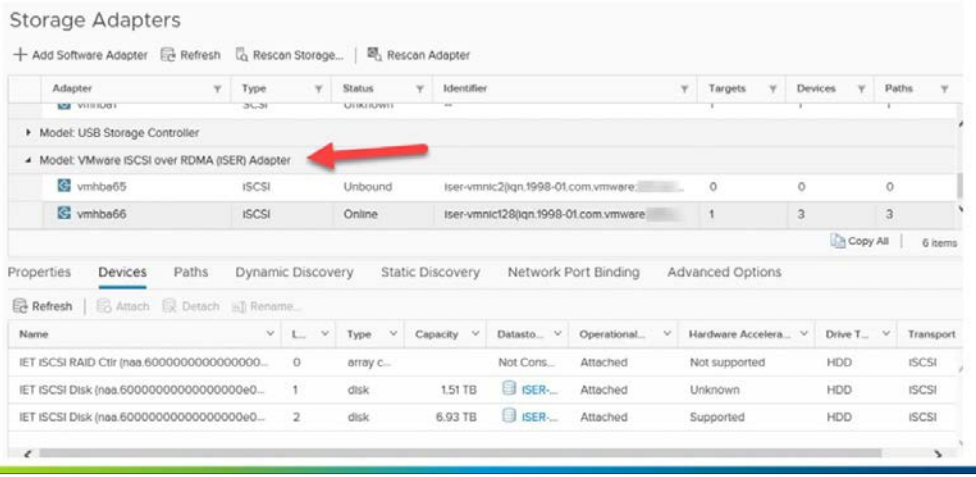
To enable flow control, see the VMware knowledge base article 1013413 at http://kb.vmware.com/kb/1013413.

https://t.me/learningnets

7-35 Running iSER

Run the following command from the ESXi host to enable the iSER adapter.

```
esxcli rdma iser add
```



https://t.me/learningnets

7-36 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Describe support for iSER in vSphere 7
- Configure iSER on ESXi hosts

https://t.me/learningnets

https://t.me/learningnets

7-38 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

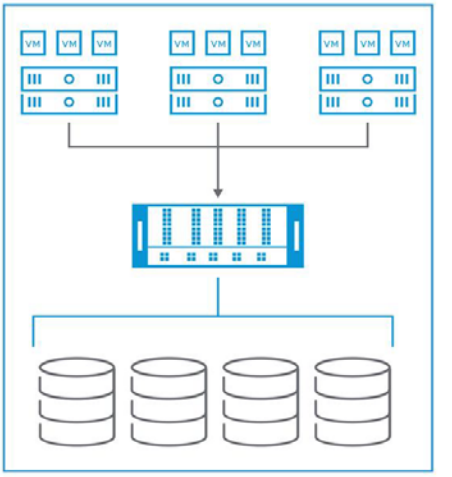
https://t.me/learningnets

7-39 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:

- VM: VM 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

7-40 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:

- 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 problems:

- Are you getting the expected I/O per second (IOPS)?
- Are you getting the expected bandwidth (read/write)?

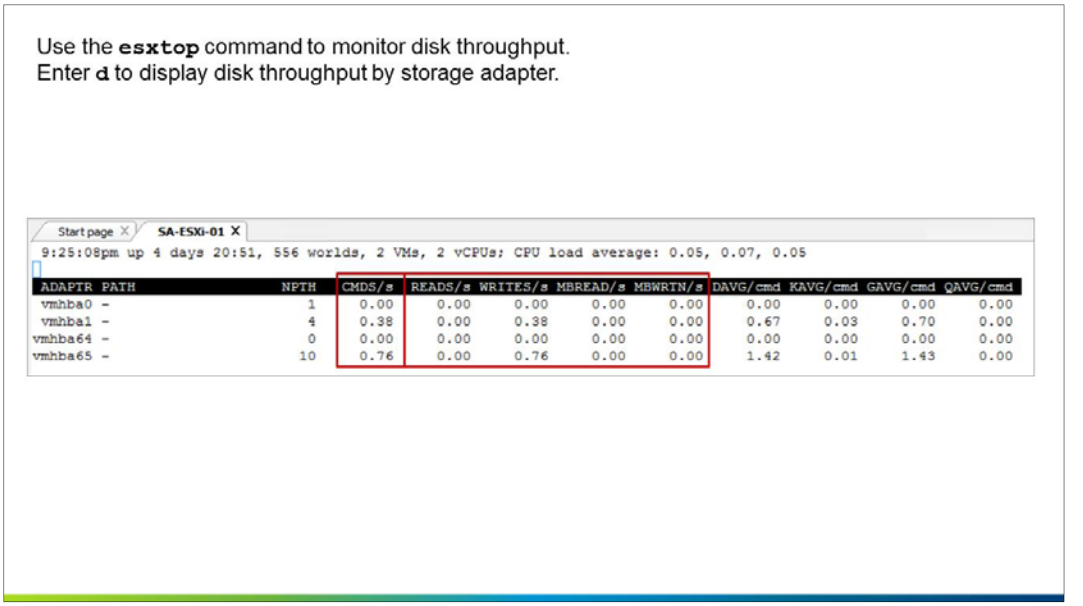
The same questions apply to disk latency. Verify disk latency and compare it with your expectations.

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

7-41 Monitoring Disk Throughput by Storage Adapter (1)



The esxtop window displays disk activity per disk adapter. To display the set of columns, enter **f**. Ensure that the following columns are selected: A (adapter name), B (path name), C (number of paths), E (I/O statistics), and G (overall latency statistics). If necessary, select columns to display by entering **a**, **b**, **c**, **e**, or **g**.

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 such as Iometer.

Disk throughput can be monitored using esxtop:

- READS/s: Number of disk reads per second
- WRITES/s: Number of disk writes per second

https://t.me/learningnets

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

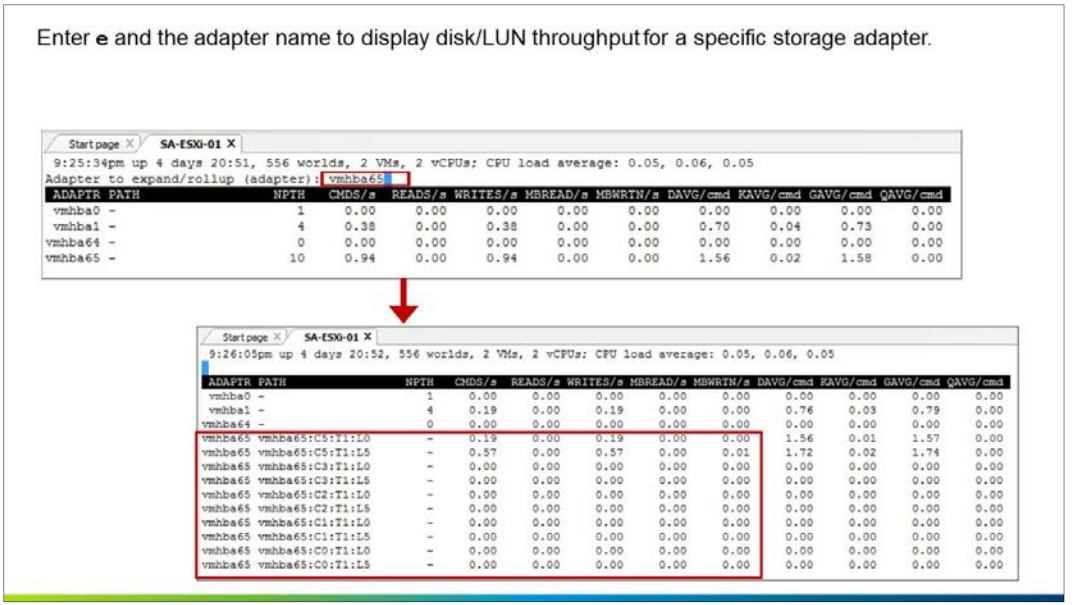
All these metrics can be monitored per HBA (vmhba#).

CMDS/s shows the total amount of commands per second and includes IOPS (Input/Output Operations Per Second). It also shows other SCSI commands such as SCSI reservations, locks, vendor string requests, unit attention commands, and so on, being sent to or coming from the device or VM being monitored.

In most cases, CMDS/s = IOPS, unless there are many metadata operations (such as SCSI reservations).

https://t.me/learningnets

7-42 Monitoring Disk Throughput by Storage Adapter (2)



In this example, vmhba65 is entered in the esxtop window. The esxtop window displays disk/LUN throughput information for each of the adapter's paths.

https://t.me/learningnets

7-43 Correlating Storage Devices to Datastores (1)

Use the 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.

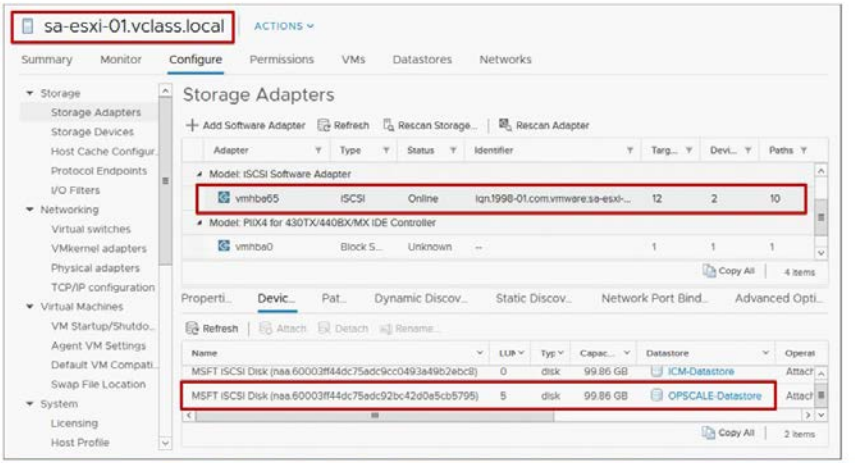
Storage administrators usually keep a record (for example, in a spreadsheet) of each storage device, its unique ID, and its use.

You can use the 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

7-44 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.

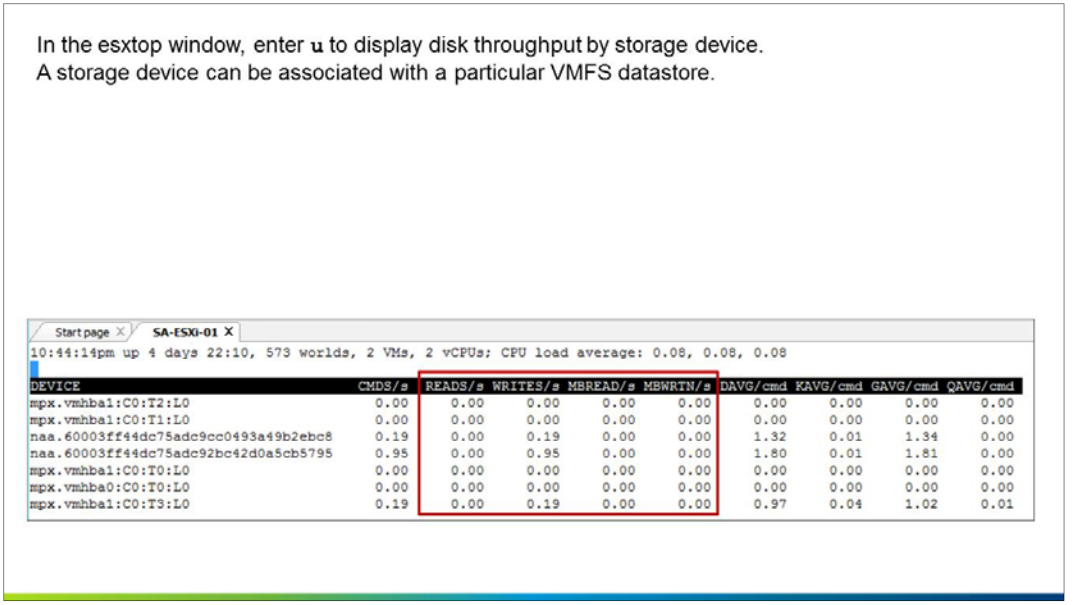


You can also identify what ESXi hosts can access that particular LUN. In this example, the naa.60003ff44dc75adc92bc42d0a5cb5795 LUN is accessed by the sa-esxi-01.vclass.local host.

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

7-45 Monitoring Path Statistics for a Specific Storage Device (1)



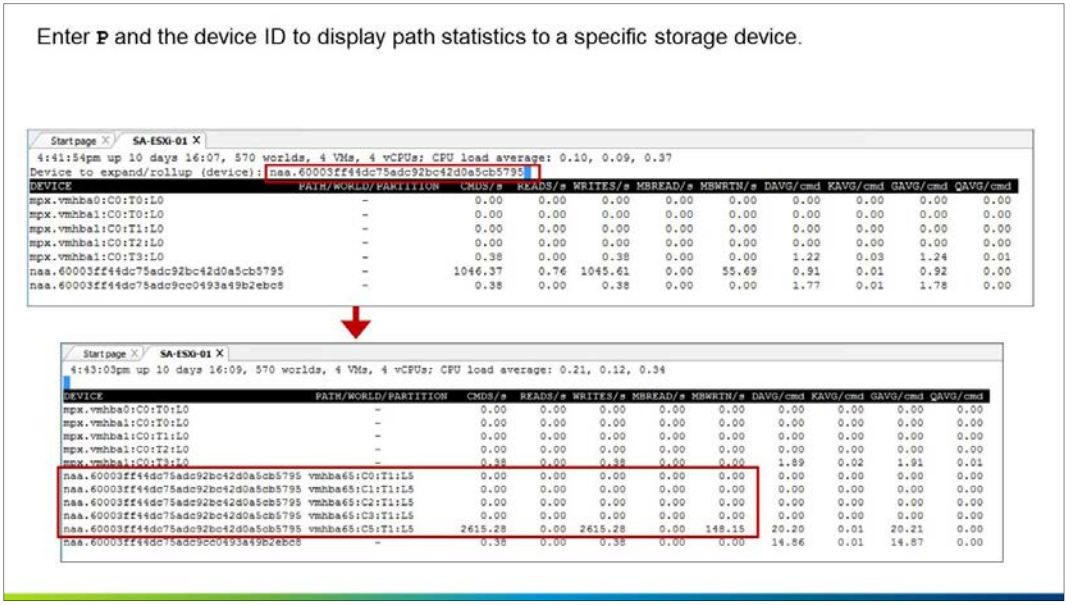
Storage devices (or LUNs) are identified by a unique identifier, such as a Network Address Authority ID (NAA ID) or a T10 identifier.

This esxtop window shows disk activity by storage device. To display the set of columns, enter **f**.

Ensure that the following columns 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 columns to display by entering **a**, **b**, **f**, **g**, or **i**.

https://t.me/learningnets

7-46 Monitoring Path Statistics for a Specific Storage Device (2)

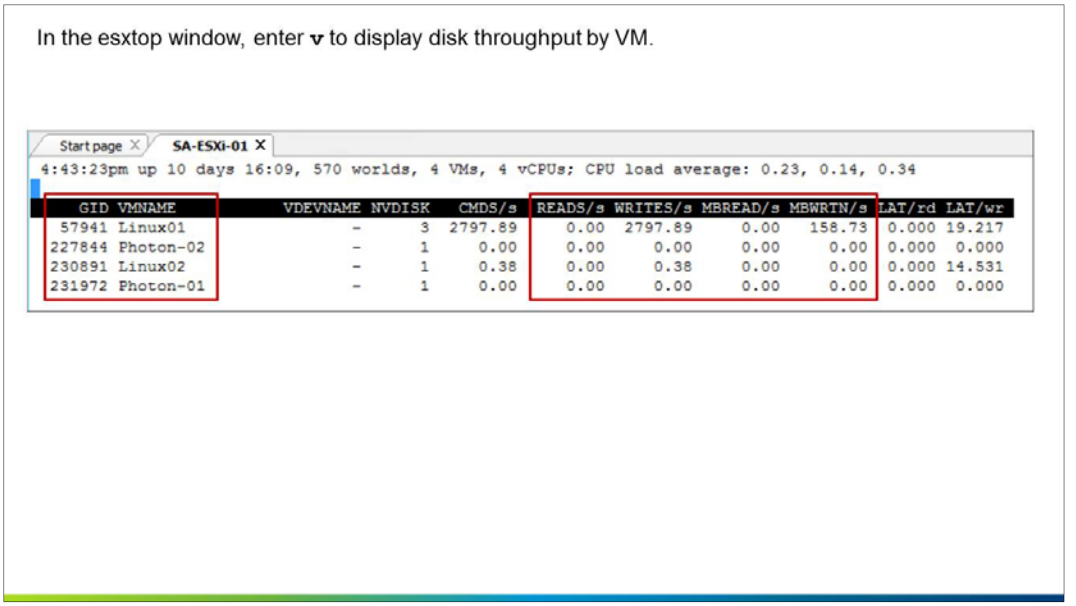


In this example, naa.60003ff44dc75adc92bc42d0a5cb5795 is entered in the esxtop window.

`esxtop` displays path statistics for this particular device.

https://t.me/learningnets

7-47 Monitoring Individual VMDK Statistics for a VM (1)

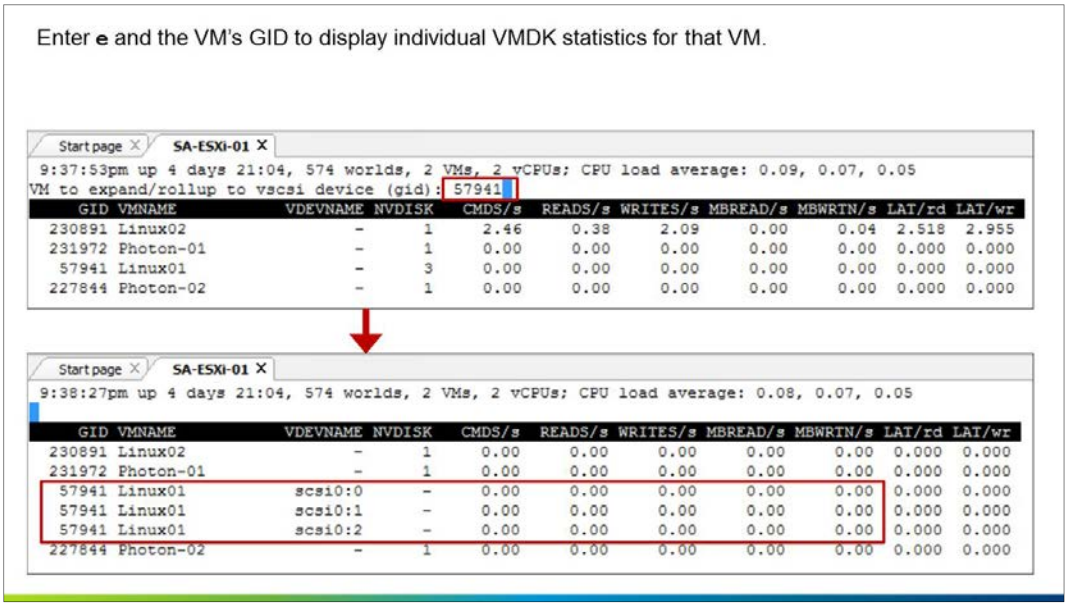


This esxtop window shows disk activity per VM. To display the set of columns, enter **f**.

Ensure that the following columns 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 columns to display by entering **b**, **c**, **d**, **e**, **i**, **j**, or **k**.

https://t.me/learningnets

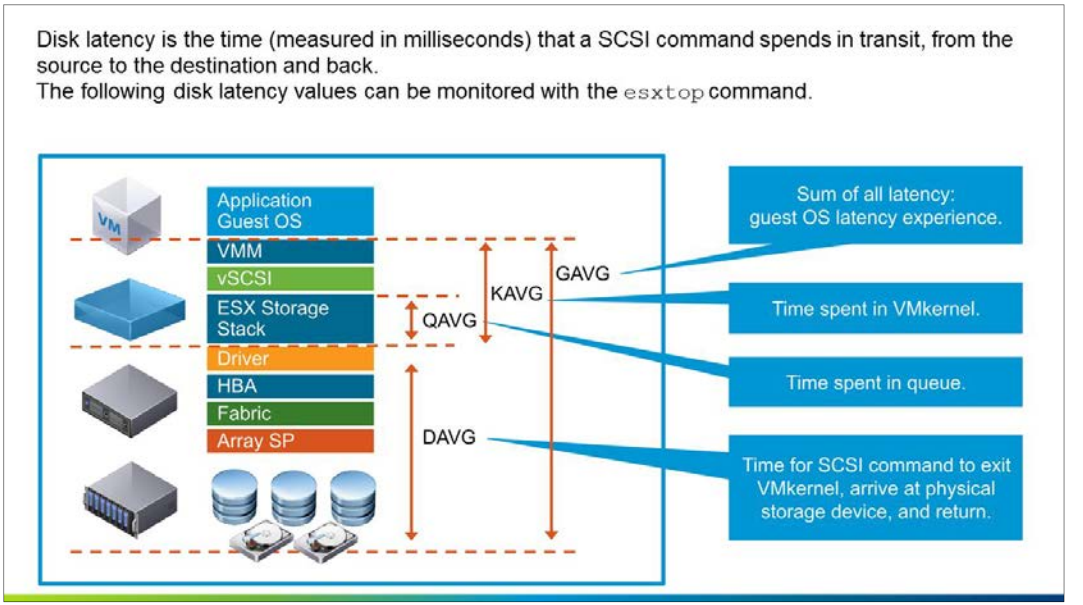
7-48 Monitoring Individual VMDK Statistics for a VM (2)



In this example, the GID for the Linux01 VM (57941) is entered in the esxtop window. The esxtop window displays disk throughput statistics for each VMDK connected to the VM.

https://t.me/learningnets

7-49 Disk Latency Metrics



Latency statistics are measured at the various layers of the ESXi storage stack.

GAVG is the round-trip latency that the guest OS 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 much queuing occurs in ESXi, KAVG can be as high as or higher than DAVG. If this situation occurs, verify 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, verify 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

https://t.me/learningnets

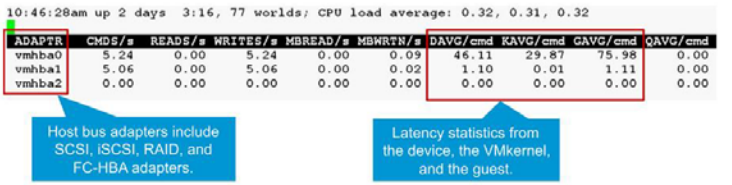
intermediate data also matches DAVG values, the storage is likely underconfigured for the application. Adding disk spindles or changing the RAID level might help in such cases.

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

7-50 Monitoring Disk Latency with esxtop

In addition to disk throughput, you can monitor disk latency in the esxtop disk adapter window (enter **d**).



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 that 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.

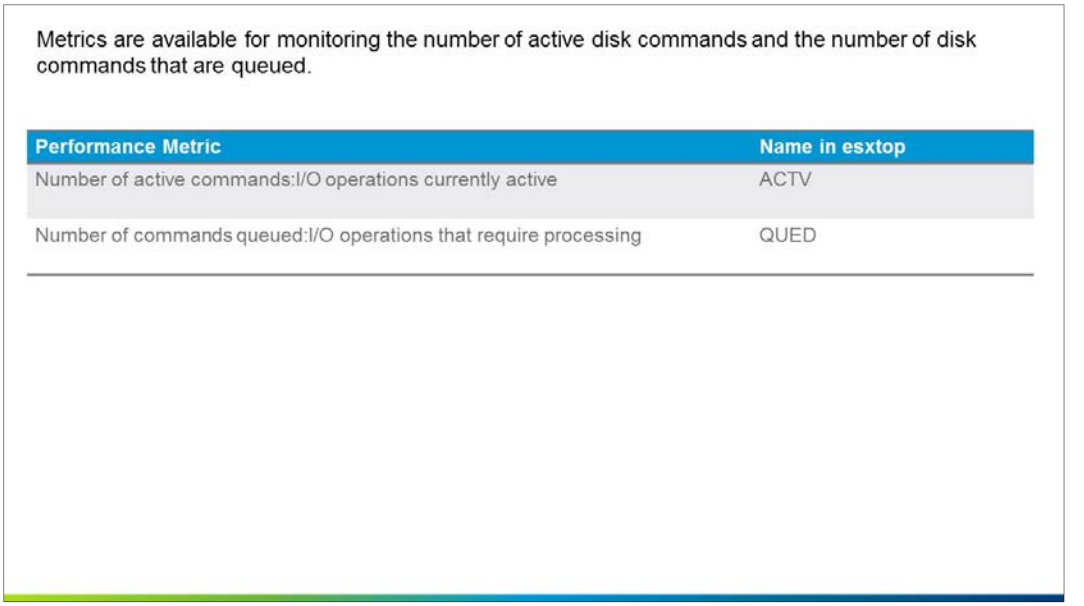
KAVG/cmd is the average amount of time that it takes the VMkernel to service a disk operation. This number represents the time taken by the CPU to manage I/O. Because processors are much faster than disks, this value should be close to zero. A value of 1 or 2 is considered high for this metric.

GAVG/cmd is the total latency seen from the VM when performing an I/O request.

GAVG is the sum of DAVG plus KAVG.

https://t.me/learningnets

7-51 Monitoring Commands and Command Queuing

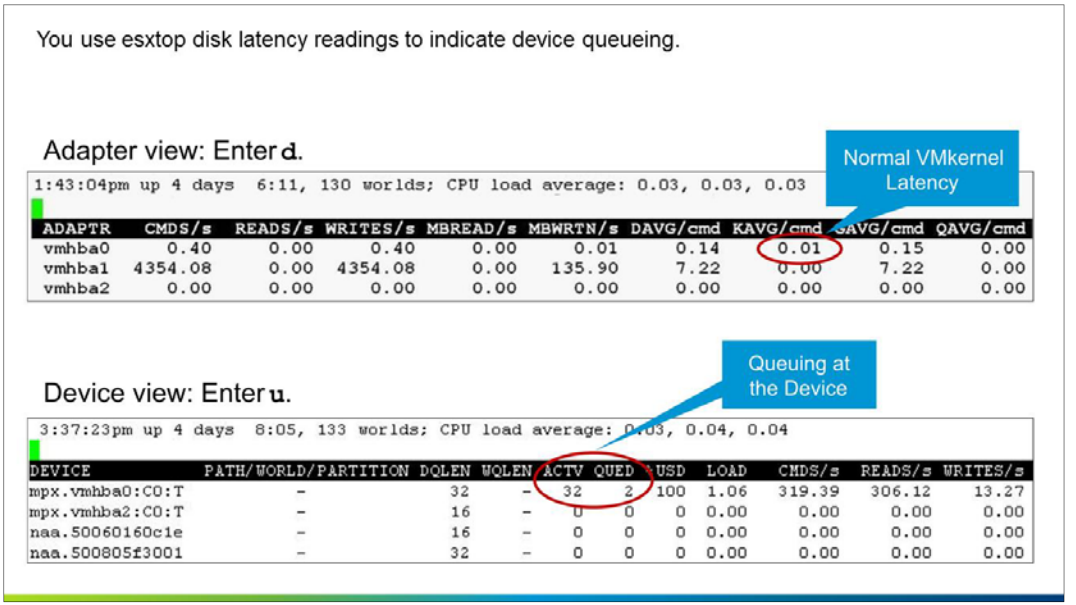


The metrics in the table provide information about your disk performance. They are often used to further interpret the latency values that you might be observing:

- 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

7-52 Example: Disk Latency and Queuing



This example shows monitoring the kernel latency value, KAVG/cmd. This value is being monitored for the vmhba0 device. In the first **esxtop** window (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 window (enter **u** in the window), 32 active I/O operations (ACTV) and 2 I/O operations 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/O operations (that is, more than 32) exist simultaneously, the device gets bottlenecked to only 32 outstanding I/O operations at a time. To resolve this problem, you might change the queue depth of the device driver.

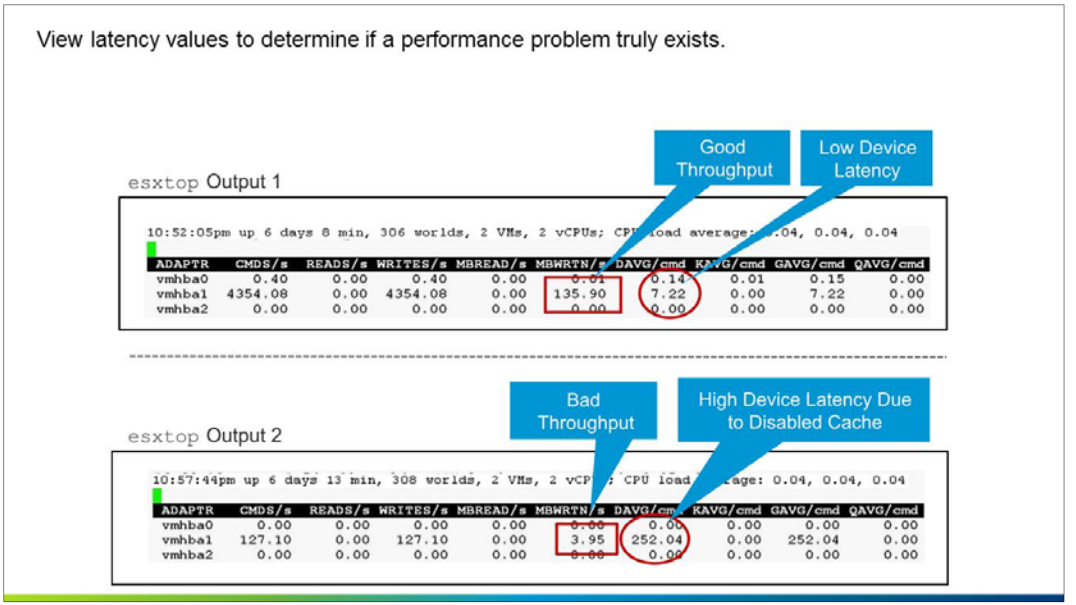
For more information about changing the queue depth of the device driver, see *About vSphere Monitoring and Performance* at https://docs.vmware.com/en/VMware-

https://t.me/learningnets

vSphere/7.0/com.vmware.vsphere.monitoring.doc/GUID-A8B06BE0-E5FC-435C-B12F-A31618B21E2C.html.

https://t.me/learningnets

7-53 Example: Bad Disk Throughput



In the example, the first esxtop window (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.

However, the second esxtop window 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 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, verify the battery status on the RAID controller to see if the battery is still good.

https://t.me/learningnets

7-54 Activity: Hardware Disk Latency (1)

What **esxtop** columns reflect the physical storage device's latency? Select all that apply.

- ❒ GAVG
- ❒ KAVG
- ❒ QAVG
- ❒ DAVG

https://t.me/learningnets

7-55 Activity: Hardware Disk Latency (2)

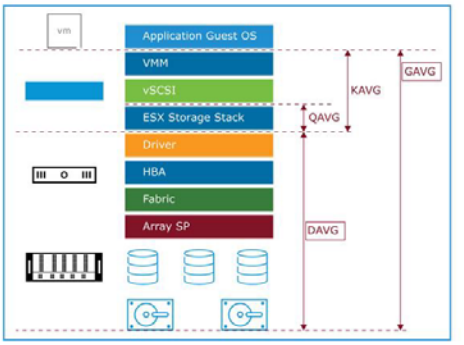
The following **esxtop** columns reflect a physical storage device's latency:

- ✓ GAVG
- ❒ KAVG
- ❒ QAVG
- ✓ DAVG

GAVG = KAVG + DAVG

DAVG is the latency seen at the device driver level.

The latency of a physical storage device directly affects the Array SP.



https://t.me/learningnets

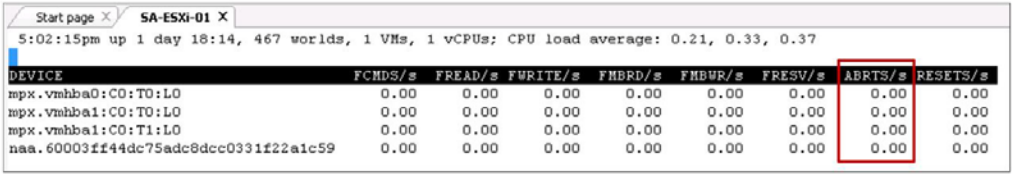
7-56 Overloaded Storage and Command Aborts

If a storage device is experiencing command aborts, the cause of these aborts must be identified and corrected.

To monitor command aborts in the esxtop window:

- Enter `u` to display disks (LUNs).
- Enter `f` and `L` to add error statistics columns to the display.

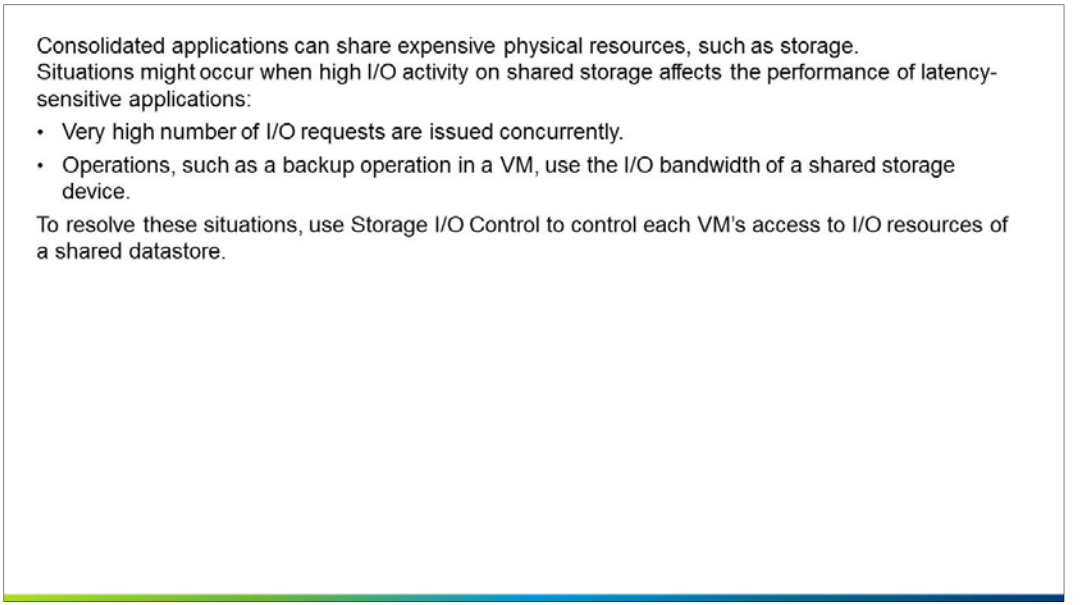
If ABRTS/s is more than 0 for any LUN, storage is overloaded on that LUN.



Severely overloaded storage can result from several problems in the underlying storage layout or infrastructure. Overloaded storage can manifest itself in many ways, depending on the applications running on the VMs. When storage is severely overloaded, operation timeouts can cause commands that are already issued to disk to be terminated (or aborted).

https://t.me/learningnets

7-57 Unexpected Increase in I/O Latency on Shared Storage



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 VMs sharing a storage device become active at the same time. As a result, these VMs 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 VMs 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 VM use the I/O bandwidth of a shared storage device. These operations cause other VMs sharing the device to suffer from resource starvation.

https://t.me/learningnets

Storage I/O Control can 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 VMs to control each VM’s access to the shared datastore.

Also, vSphere DRS can automatically or manually balance datastore cluster I/O capacity and performance. vSphere Storage DRS can perform automatic vSphere Storage vMotion migrations of VMs. These migrations can occur when a datastore's threshold for space use or I/O response time are exceeded.

https://t.me/learningnets

7-58 Storage Performance Best Practices

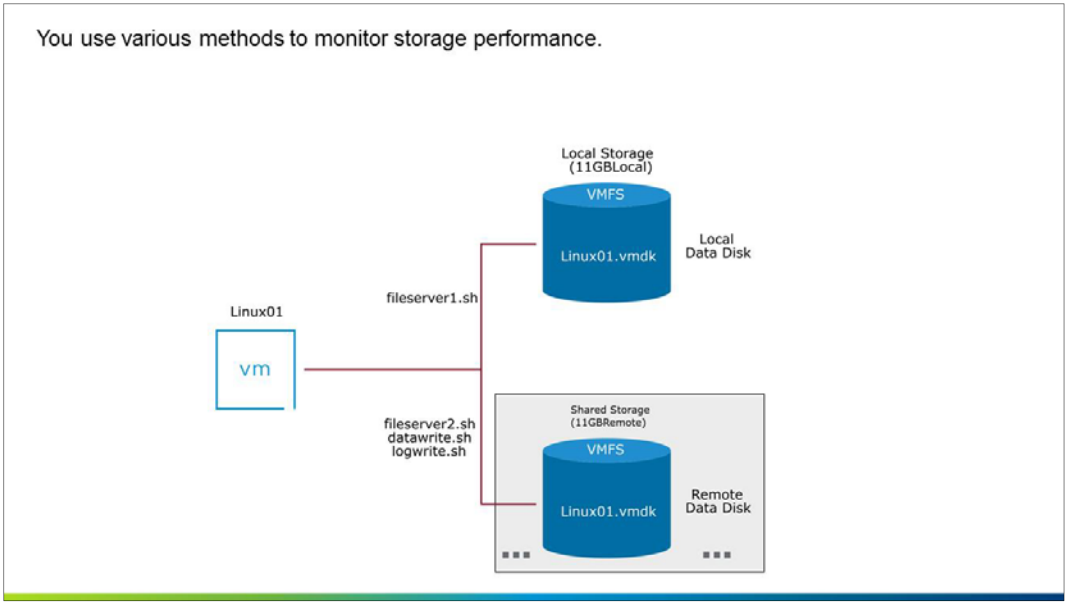
To get the best performance from storage, follow these best practices:

- Configure each LUN with the storage characteristics for applications and VMs that use the LUN.
- Avoid oversubscribing storage paths and storage devices. Use multipathing to avoid single points of failure and increase bandwidth.
- Use Storage DRS and Storage I/O Control whenever applicable.
- Isolate iSCSI and NFS traffic.
- Whenever possible, physically separate VM traffic from storage traffic, otherwise use vLANs.
- 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

7-59 Introduction to Lab: Monitoring Storage Performance



You generate various types of disk I/O and compare performance. Your VM 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 an application called `aio-stress`, which is a simple command-line application that measures the performance of a disk subsystem. For more information about the `aio-stress` command, see the `readme.txt` file on the test VM in the same directory as the scripts.

https://t.me/learningnets

7-60 Lab 18: 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

7-61 Review of Lab: 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/ LocalDatastore (Task 5)
READS/s				
WRITES/s				

https://t.me/learningnets

7-62 Review of 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

7-63 Key Points

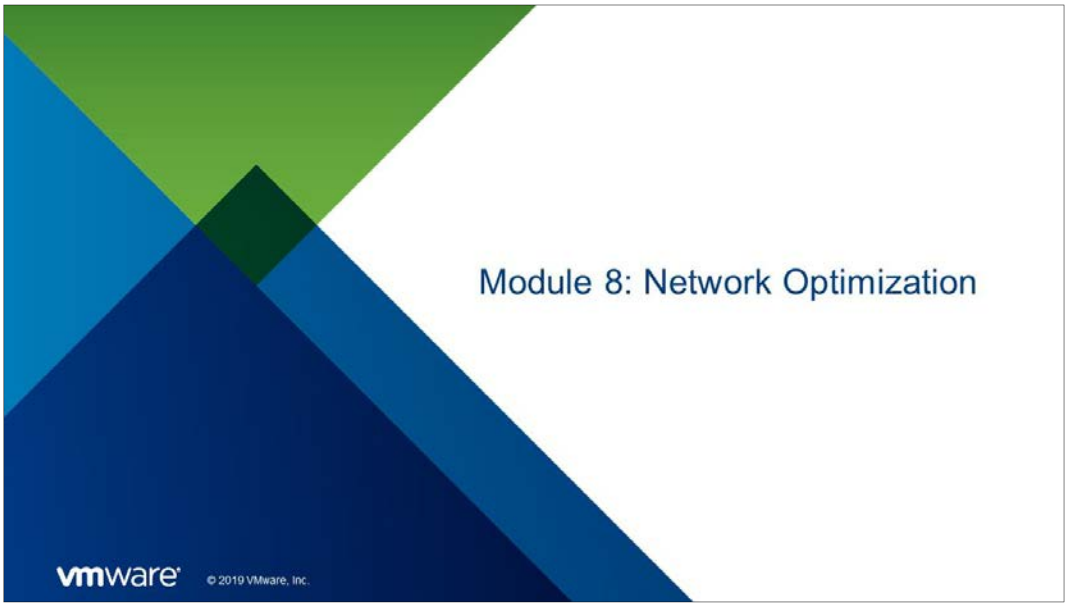
- Storage protocols, storage configuration, RAID levels, queuing, and VMFS configuration are factors that affect storage performance.
- Disk throughput and latency are key metrics when monitoring storage performance.
- If a storage device is experiencing command aborts, the cause of these aborts must be identified and corrected.

Questions?

https://t.me/learningnets

Module 8

Network Optimization



https://t.me/learningnets

8-2 Importance

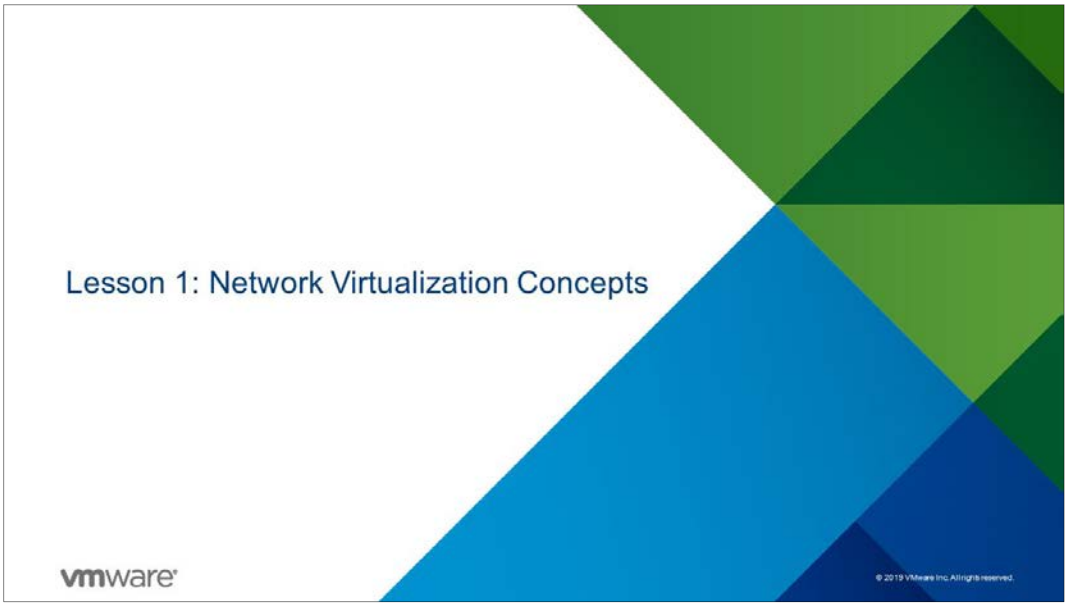
Network performance can be measured in terms of how many packets were dropped when transmitting or receiving data. Therefore, you should know how to monitor packet drops.

https://t.me/learningnets

8-3 Module Lessons

Lesson: Network Virtualization Concepts
Lesson: Monitoring Network I/O Activity

https://t.me/learningnets



https://t.me/learningnets

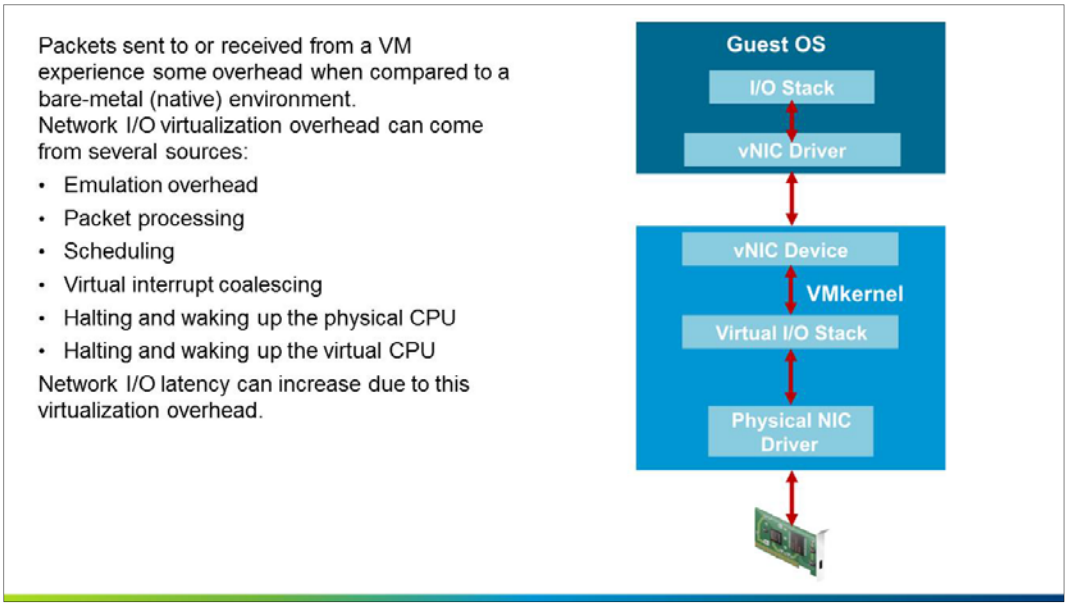
8-5 Learner Objectives

After completing 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

8-6 Network I/O Virtualization Overhead



The overhead that VMs 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 the VM executes 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 VM, 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

https://t.me/learningnets

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 VM does virtual interrupt coalescing. That is, a VM 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 VM 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.

- 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.

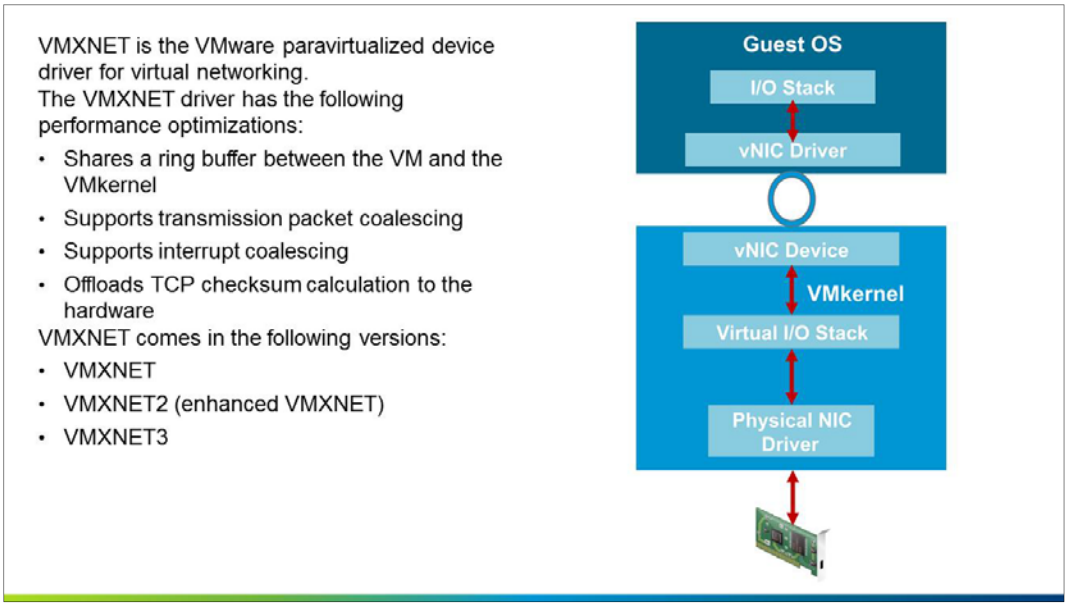
  Sometimes the side effects of this halting and waking can be worse in virtualized environments than in physical environments. For example, a CPU might flush its cache on entering halt mode. In this case, the halt operation might affect a virtualized system more than a physical system because the cache footprint is larger in a virtual environment.

- Halting and waking up the virtual CPU: When the guest OS 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

8-7 VMXNET Network Adapters



The VMXNET driver implements an idealized network interface that passes through network traffic from the VM to the physical cards with minimal overhead.

The driver improves performance through several optimizations:

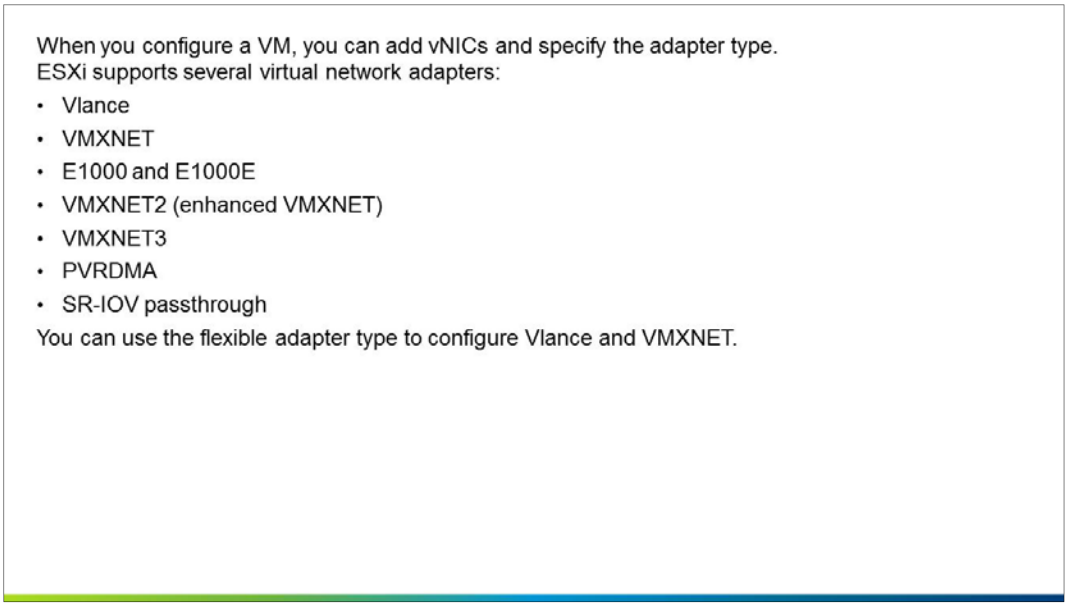
- It shares a ring buffer between the VM and the VMkernel. And it uses zero-copy which, in turn, saves CPU cycles. Traditional networking uses a series of buffers to process incoming network data and deliver it efficiently to users. However, higher-speed modern networks are turning this approach into 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 VMs 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.

https://t.me/learningnets

- 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 VM monitor.

https://t.me/learningnets

8-8 Virtual Network Adapters



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 VM. 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 VM 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.

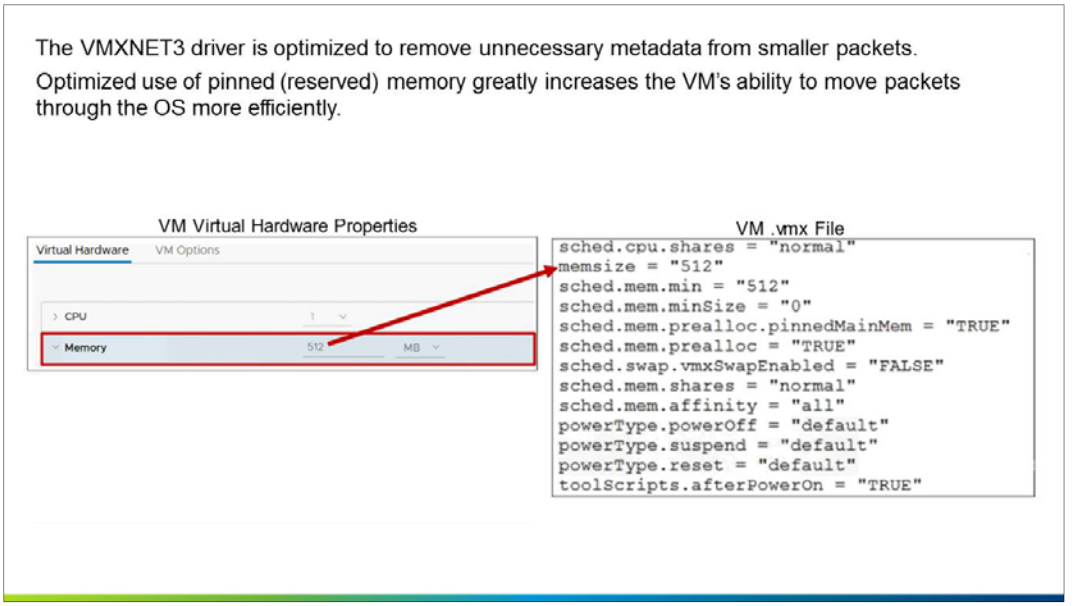
https://t.me/learningnets

- Enhanced VMXNET (VMXNET2): The enhanced VMXNET adapter is based on the VMXNET adapter but it provides 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 VMs through the OFED verbs API.
- 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 *Network Adapter Basics* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.vm_admin.doc/GUID-AF9E24A8-2CFA-447B-AC83-35D563119667.html.

https://t.me/learningnets

8-9 About VMXNET3 Optimization



To assist the VM in processing packets internally as quickly as possible, the VMXNET3 driver is optimized to remove unnecessary metadata from smaller packets. It is optimized 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 VM's ability to move packets through the OS more efficiently. Normally, VM 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 memory or moved to accommodate the demands of other VMs. By reserving and pinning the entire VM memory, you address translation as the VM's pages are permanently mapped to specific physical addresses. Therefore, the address translation never needs repeating.

Memory can be pinned by configuring VMX options or, more simply, by setting the latency sensitivity of the VM to high.

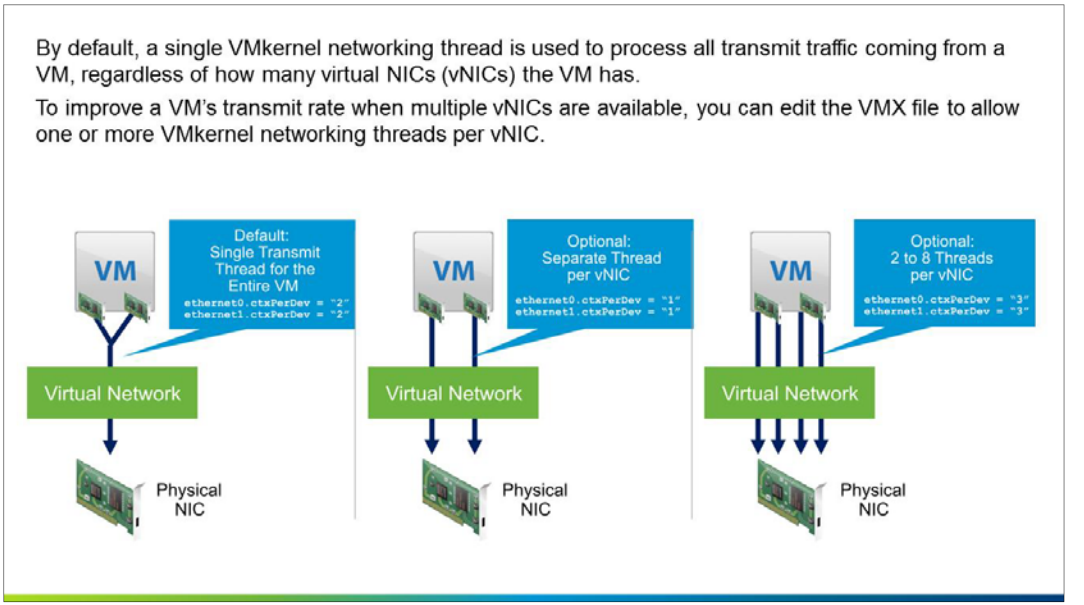
As shown, you add lines in the VM's VMX file to pin all VM memory. These options ensure that the VM can never use swap memory, which requires that the VM memory always be backed by

https://t.me/learningnets

real physical memory. All the memory is configured to be preallocated and pinned, which results in the appearance that all VM memory is permanently active.

https://t.me/learningnets

8-10 Configuring a Single CPU Thread per vNIC



CPU scheduling is optimized for network packet processing in the ESXi.

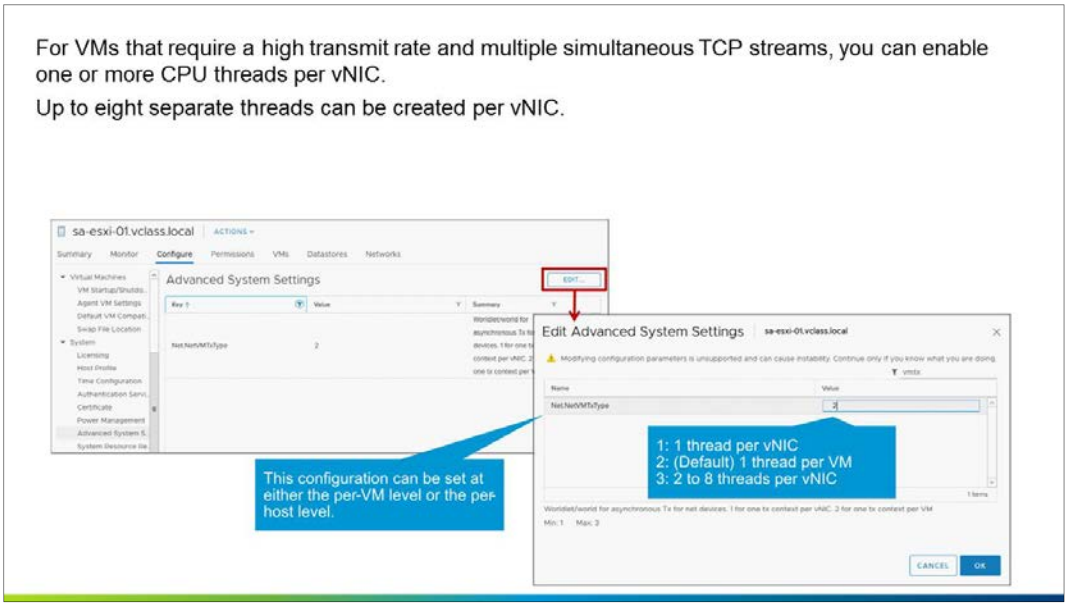
Typically, when VM packets are being processed, a single CPU thread is used to process all transmit traffic coming from a VM, regardless of how many vNICs the VM has. If multiple vNICs are configured for a VM, it is advantageous to allow one CPU thread per vNIC to help speed up the processing time. This arrangement helps boost transmit rates for the VM 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 VM 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

8-11 Configuring Multiple CPU Threads 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, 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 the vSphere Client.

https://t.me/learningnets

Acceptable values for Net.NetVMTxType are 1, 2, or 3:

1 1 transmit context (CPU thread) per vNIC

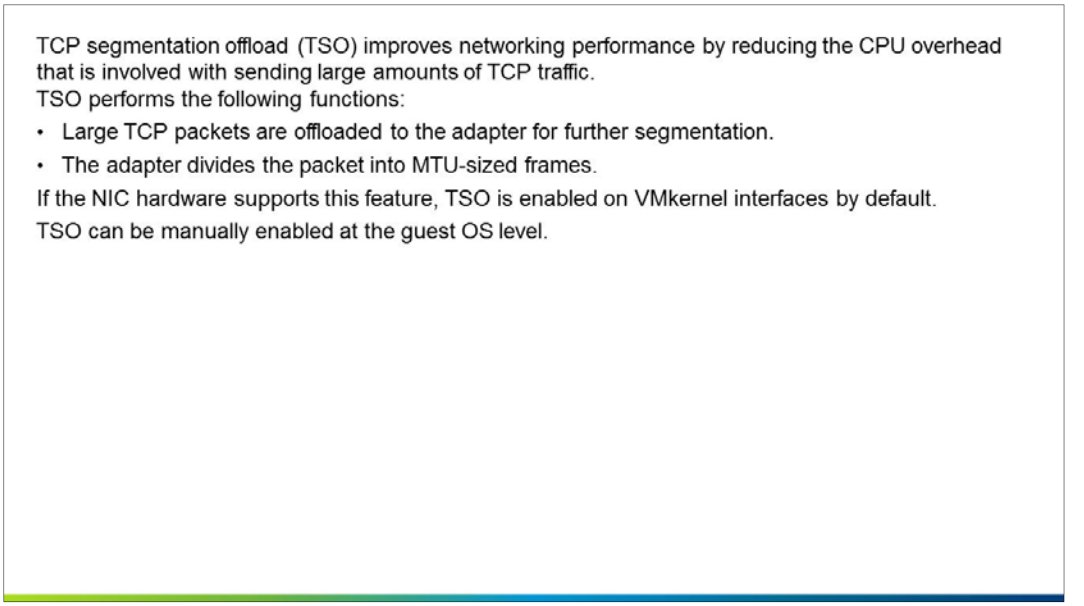
2 (Default): 1 transmit context (CPU thread) per VM

3 2 through 8 transmit contexts (CPU threads) per vNIC

Many of the CPU scheduling improvements require available CPU threads on a host. 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

8-12 About TCP Segmentation Offload



TCP segmentation offload (TSO) improves performance for TCP network traffic coming from a VM 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 VM 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.

VMs that use TSO on an ESXi host show lower CPU utilization than VMs that lack TSO support, when performing the same network activities.

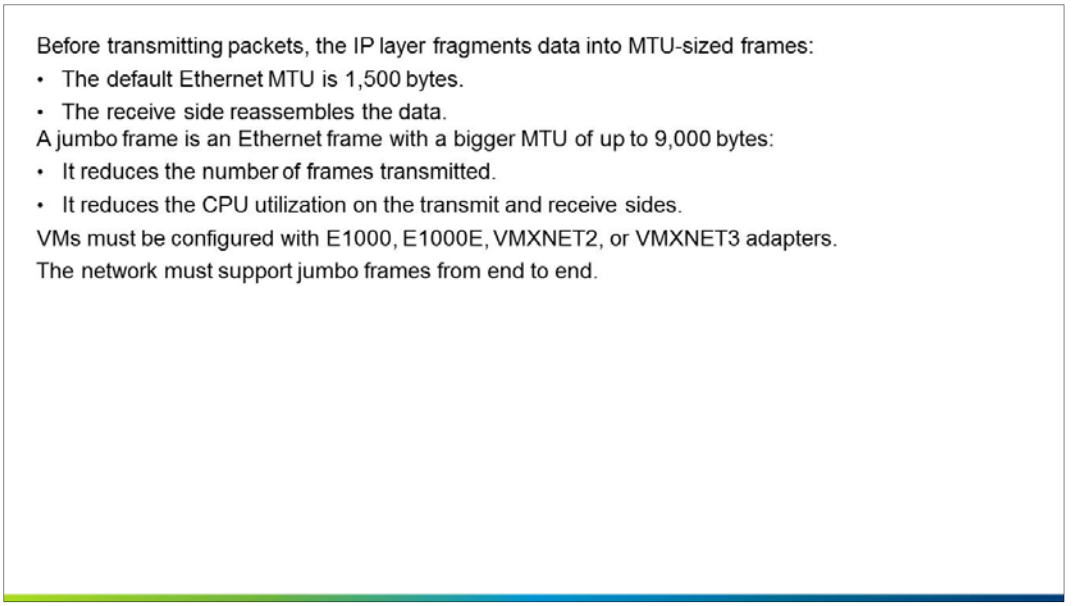
https://t.me/learningnets

For information about enabling TSO support for a VM or checking whether TSO is enabled on a VMkernel interface, see *TCP Segmentation Offload* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.networking.doc/GUID-E105A601-9331-496C-A213-F76EA3863E31.html.

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

8-13 About Jumbo Frames



For each packet, the system must 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 VM, and at the VMkernel interface. ESXi hosts using jumbo frames realize a decrease in load due to network processing. Before enabling jumbo frames, verify with your hardware vendor to ensure that your network adapter supports jumbo frames.

For information about jumbo frames, see *Jumbo Frames* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.networking.doc/GUID-53F968D9-2F91-41DA-B7B2-48394D997F2A.html.

https://t.me/learningnets

8-14 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 VMs 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.

https://t.me/learningnets

SplitRx mode is individually configured for each vNIC. For more information about how to enable this feature, see the *ESXi Networking Considerations* section at https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/performance/vsphere-esxi-vcenter-server-67-performance-best-practices.pdf.

https://t.me/learningnets

8-15 About SplitTx Mode

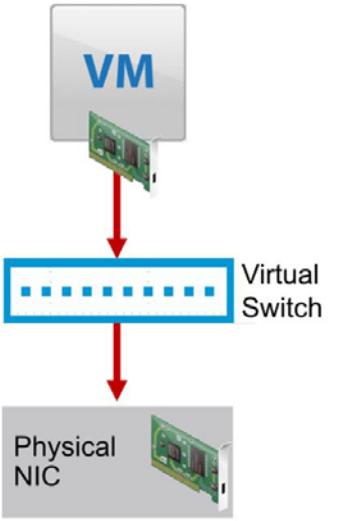
With SplitTx mode, two separate threads can be created for a single flow to parallelize vNIC and physical NIC processing:

- 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

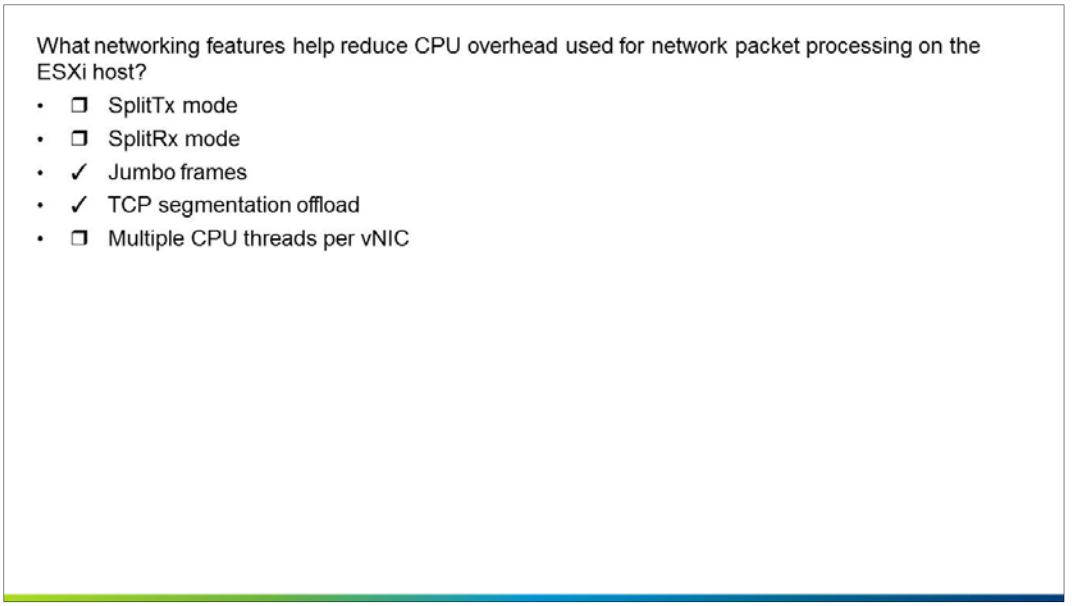
8-16 Activity: vSphere Networking Features (1)

What networking features help 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

8-17 Activity: vSphere Networking Features (2)



https://t.me/learningnets

8-18 Review of Learner Objectives

After completing 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

https://t.me/learningnets

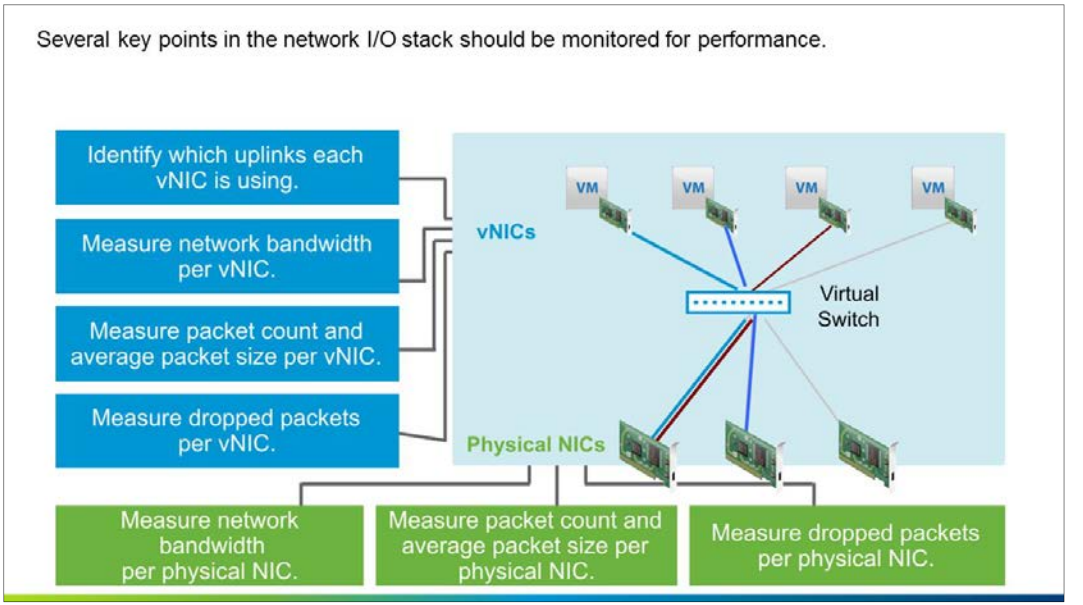
8-20 Learner Objectives

After completing 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

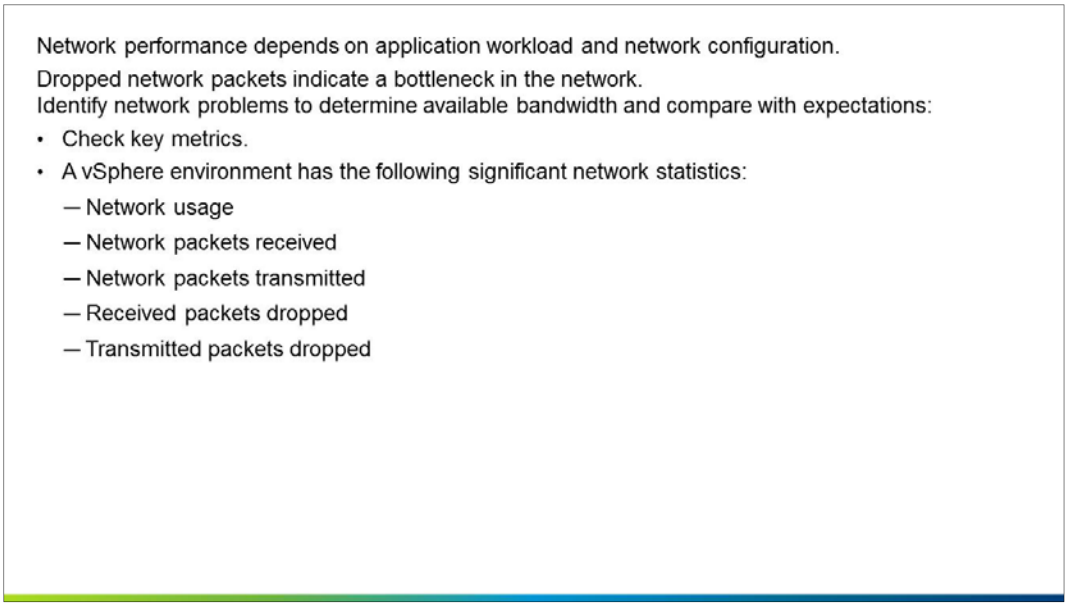
https://t.me/learningnets

8-21 Monitoring for Performance in the Network I/O Stack



https://t.me/learningnets

8-22 Network Capacity Metrics



To determine whether packets are being dropped, you can use the advanced performance charts in the vSphere Client or use the **`esxtop`** command.

If received packets are being dropped, adjust the VM 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 VMs assigned to each physical NIC. If necessary, perform load balancing by moving VMs to different virtual switches or by adding more NICs to the host. You can also move VMs to another host or increase the CPU resources of the host or VMs.

https://t.me/learningnets

8-23 About esxtop Network Statistics

You can use **esxtop** to track network statistics per VM or per NIC (virtual or physical).

Enter **n** to change to the networking view.

esxtop provides the following useful network counters:

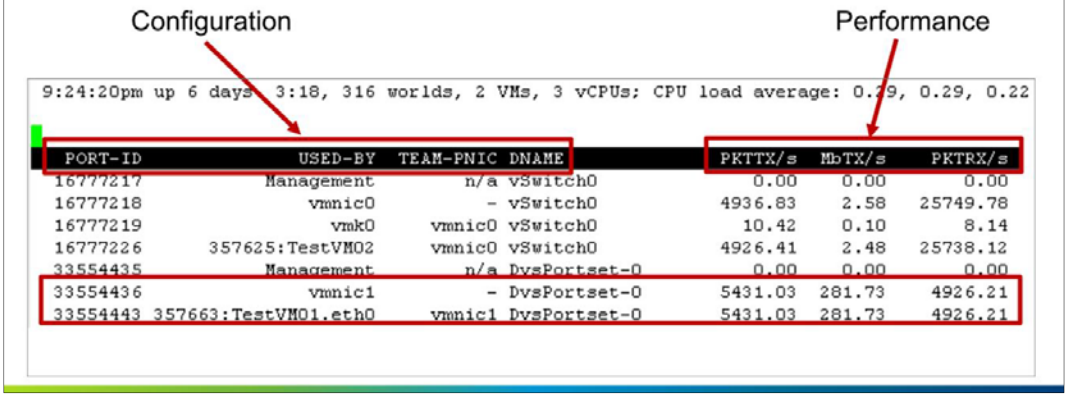
- MbTX/s: Amount of data transmitted (in megabits) per second
- MbRX/s: Amount of data received (in megabits) per second
- PKTTX/s: Average number of packets transmitted per second in the sampling interval
- PKTRX/s: Average number of packets received per second in the sampling interval
- %DRPTX: Percentage of outbound packets dropped in the sampling interval
- %DRPRX: Percentage of inbound packets dropped in the sampling interval

https://t.me/learningnets

8-24 About 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 VM 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 VM's NIC, identified as 357663:TestVM01.eth0. The value 357663 is an internal VM ID, TestVM01 is the VM name, and eth0 identifies the network interface.

- PKTTX/s – Average number of packets transmitted per second in the sampling interval
- MbTX/s – Amount of data transmitted in Mbps
- PKTRX/s – Average number of packets received per second in the sampling interval

https://t.me/learningnets

- MbRX/s – Amount of data received in Mbps
- %DRPTX – Percentage of transmit packets dropped.
- %DRPRX – Percentage of receive packets dropped

You might see `Shadow of vmnic` in some lines under the USED-BY column in the Network view of ESXTOP. Shadow ports are used to transmit and receive health check data for each uplink. You use exstop to monitor these shadow ports to debug and troubleshoot the network health check feature.

For more information about shadow ports, see *What Are the Shadow of vmnic in ESXTOP?* at https://www.virtuallyghetto.com/2013/01/what-are-shadow-of-vmnic-in-esxtop.html

https://t.me/learningnets

8-25 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 OS running in a VM. 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 OS. The virtual switch contains queues for packets sent to the vNIC.

https://t.me/learningnets

If the guest OS 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 VM when its packet queue is full, the port must drop the packet.

https://t.me/learningnets

8-26 Dropped Packets in vNICs

If a guest OS 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 VM.
	Increase the efficiency with which the VM uses CPU resources.
Improper guest OS driver configuration	Tune the network stack in the guest OS. Try using VMXNET3.
	Add vNICs to the VM and spread the network load across them.

When the applications and the guest OS are driving the VM to high CPU use, extended delays might occur. These delays occur from the time the guest OS 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 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 OS. 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 VM is dropping receive packets due to high CPU use, adding vCPUs might be necessary to provide sufficient CPU resources. If the high CPU use is due to network processing, ensure that the guest OS can use multiple CPUs when processing network traffic. See the OS documentation for the appropriate CPU requirements.

https://t.me/learningnets

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 OS to improve the speed and efficiency with which it handles network packets. For more information, eee the documentation for the OS.

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 VMs enables the processing of network interrupts to be spread across multiple processor cores.

https://t.me/learningnets

8-27 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 VMs to other ESXi hosts to reduce CPU utilization.

https://t.me/learningnets

8-28 Dropped Transmit Packets in Physical NICs

VM 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 VM to be dropped. Dropped packets are a network throughput issue.

Cause	Solution
Traffic from a VM or a set of VMs 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 VMs with high network demand to other ESXi hosts that have network capacity.
	Enhance the networking infrastructure.
	Reduce network traffic.

When a VM 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.

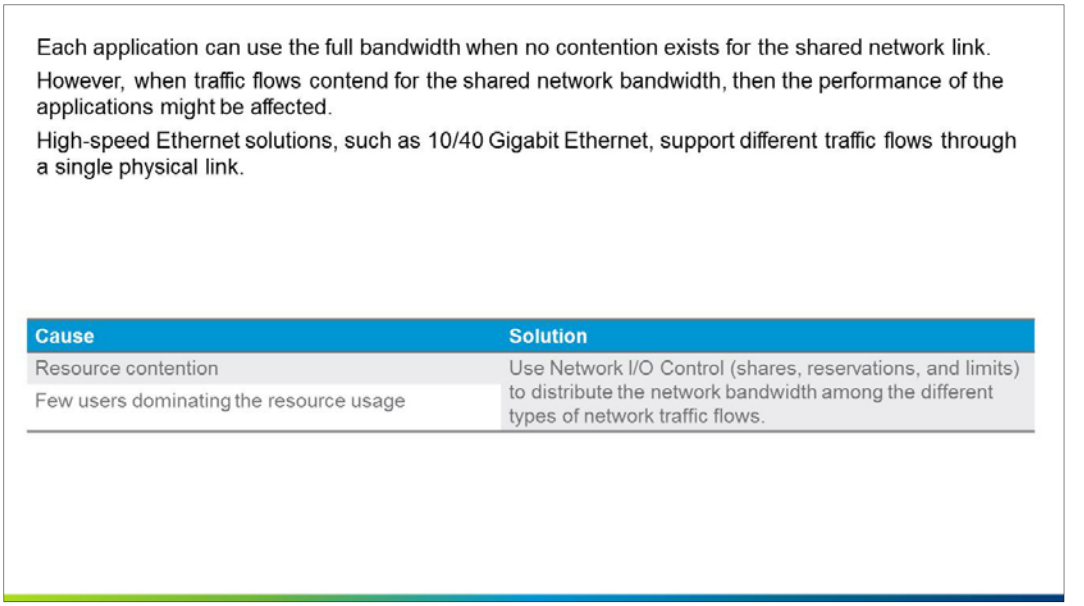
If the virtual switch uplinks are overloaded, moving some of the VMs to different virtual switches can help rebalance the load.

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 VM can help alleviate bottlenecks in the networking infrastructure. The implementation of this solution depends on the application and guest OS 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

8-29 Unexpected Increases in Data Transfer Rate



Examples of traffic flows are flows from applications running in a VM, vSphere vMotion, and vSphere Fault Tolerance. These traffic flows can coexist and share a single link.

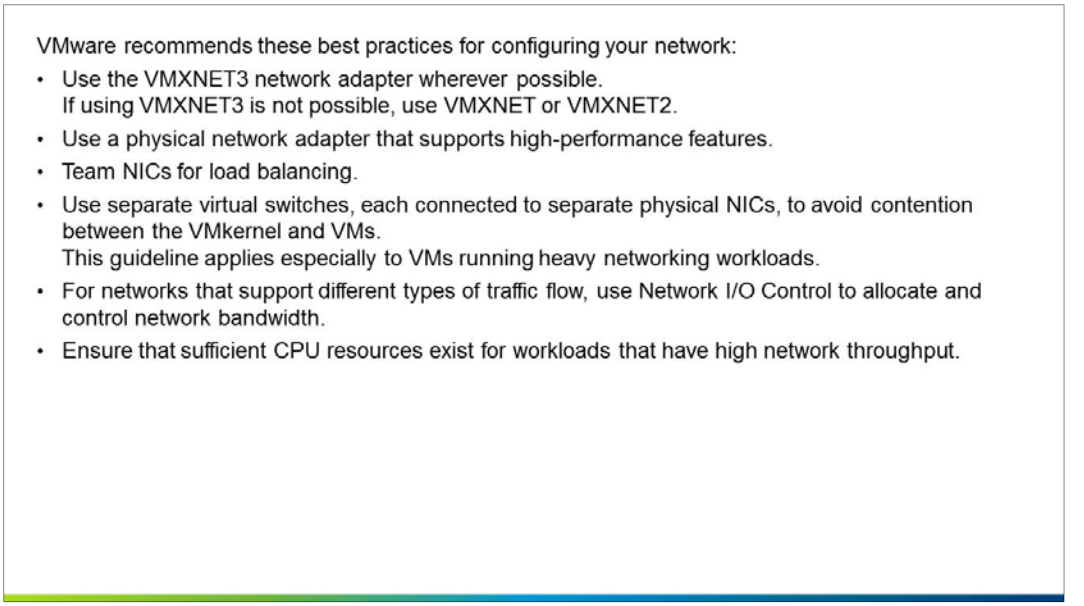
The total demand from all network users can exceed the total capacity of the network link. In this case, users of the network resources might experience an unexpected impact on performance. These instances, though infrequent, can still cause fluctuating performance behavior, which might be frustrating.

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 VMs that share network resources with these traffic flows can be affected. The performance impact might be more significant if the applications running in these VMs 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

8-30 Networking Best Practices



When you install VMware Tools, a VMXNET adapter replaces the default Vlance adapter.

VMXNET adapters should be used for optimal performance in any guest OS 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 if hardware failure or a network outage occurs. 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 VMs reside on the same ESXi host, connect both VMs to the same virtual switch. VMs that communicate with each other on the same ESXi host can also use the VMware Machine Communication Interface device.

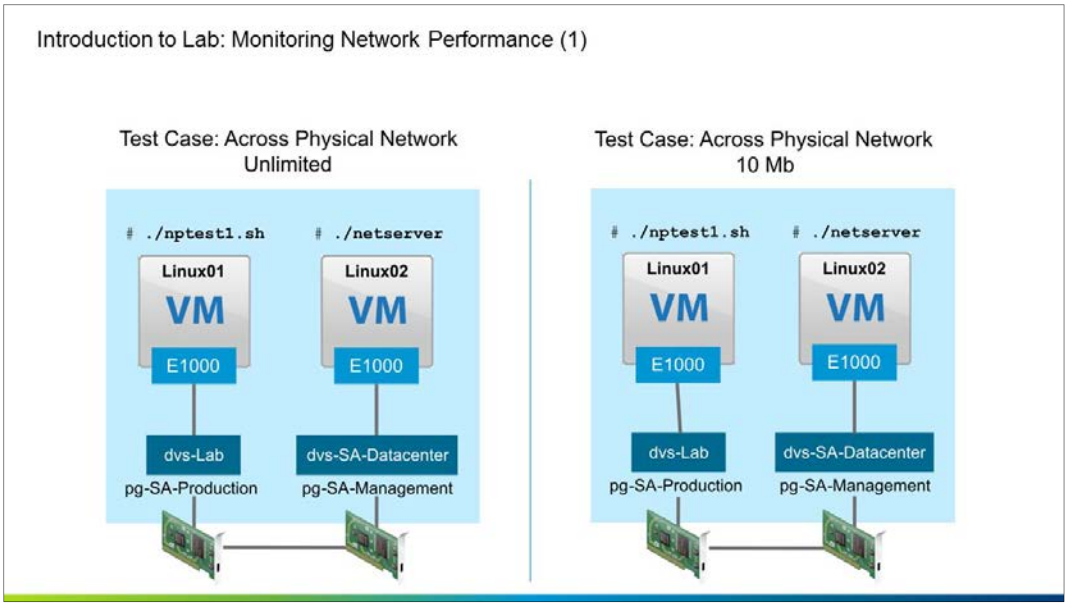
https://t.me/learningnets

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

8-31 Introduction to Lab: 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 VMs 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 VM, Linux01.

You also use a server-side script called netserver. This script runs on the network server VM, Linux02, and receives the data sent by nptest1.sh.

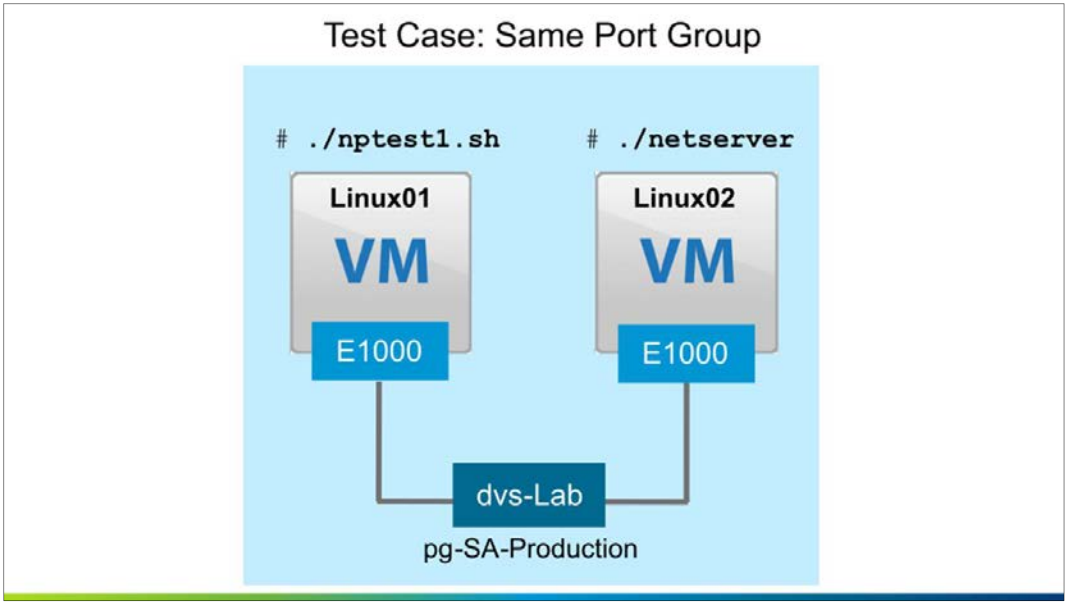
https://t.me/learningnets

You perform three test cases, two of which are described here:

- Test case 1: The first test VM is connected to the pg-SA-Production port group on the distributed switch named dvs-Lab. The second test VM 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

8-32 Introduction to Lab: Monitoring Network Performance (2)



For test case 3, both test VMs are connected to the pg-SA-Production port group on the dvs-Lab distributed switch.

https://t.me/learningnets

8-33 Lab 19: Monitoring Network Performance

Use the `esxtop` command to monitor network performance:

1. Prepare to Monitor Network Performance
2. Prepare the Client and the Server VMs
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

https://t.me/learningnets

8-34 Review of Lab: 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

8-35 Review of Learner Objectives

After completing 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

8-36 Key Points

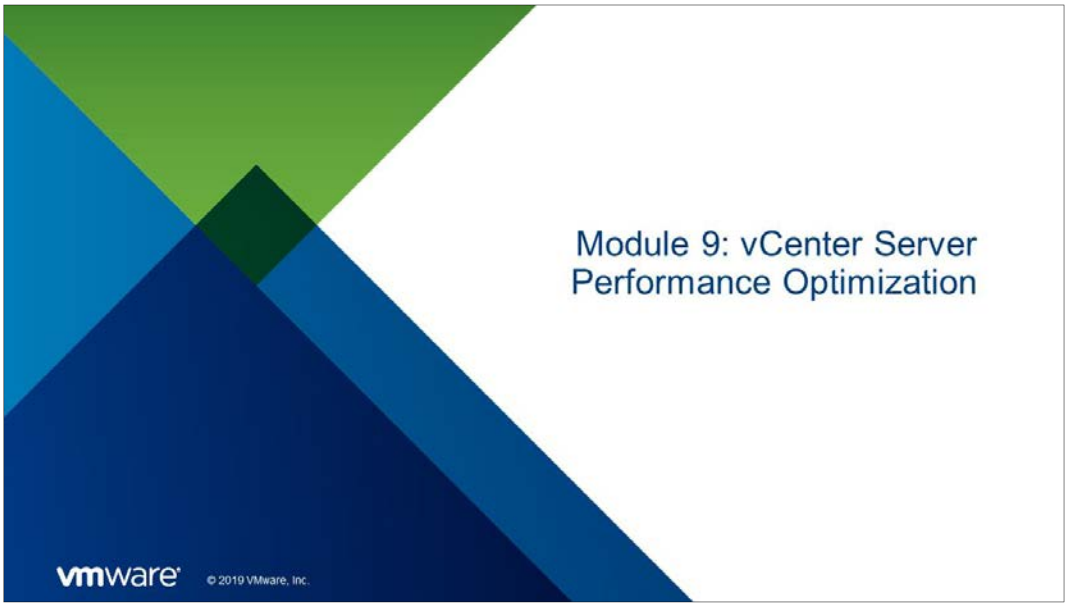
- vSphere uses many performance features of modern network adapters, such as TSO and jumbo frames.
- The **`esxtop`** tool displays physical and virtual network data in the same window.
- The presence of dropped packets is an indicator of network problems.

Questions?

https://t.me/learningnets

Module 9

vCenter Server Performance Optimization



https://t.me/learningnets

9-2 Importance

vCenter Server is a critical component in the vSphere environment. Data centers rely on vCenter Server for healthy operation.

You can make the data center run more efficiently and smoothly by using the latest improvements to vCenter Server performance and scalability.

https://t.me/learningnets

Lesson 1: vCenter Server Performance Optimization

https://t.me/learningnets

9-4 Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Review vCenter Server components and services
- Describe the factors that influence vCenter Server performance
- Use vCenter Server Appliance tools to monitor vCenter Server resource use

https://t.me/learningnets

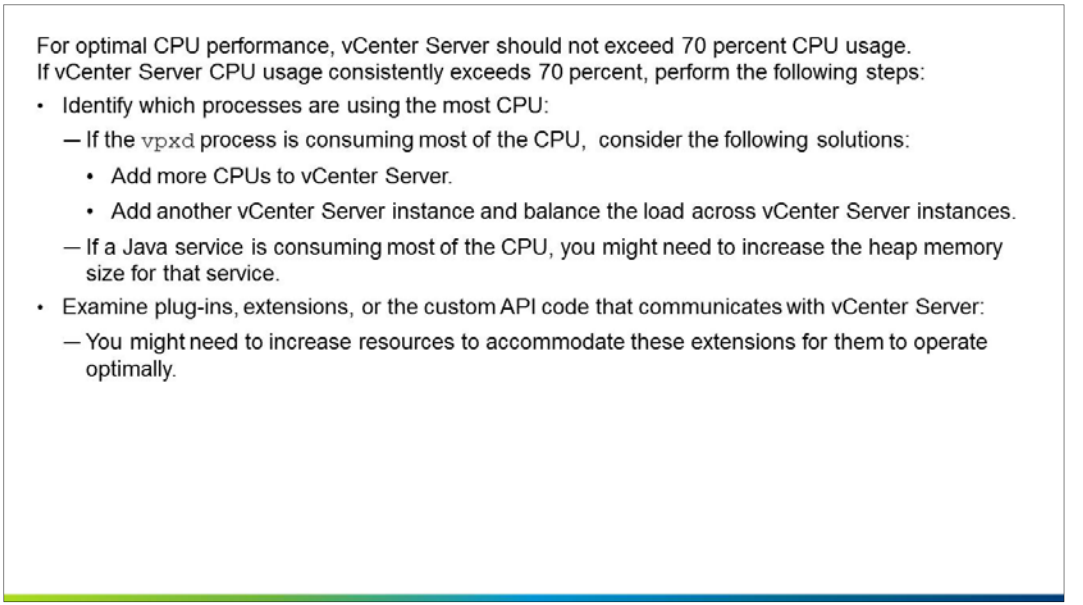
9-5 vCenter Server Performance Considerations

Consider the following factors when monitoring and maintaining acceptable levels of performance for vCenter Server and its services:

- CPU and memory performance
- Network performance
- Database performance
- User interface performance

https://t.me/learningnets

9-6 vCenter Server CPU Performance



Consistent high CPU usage can cause poor performance. Verify which processes, such as `vpxd` and `vsphere-client`, use the most CPU. CPU usage that is consistently higher than 70 percent can cause a vCenter Server process to perform more queries than necessary.

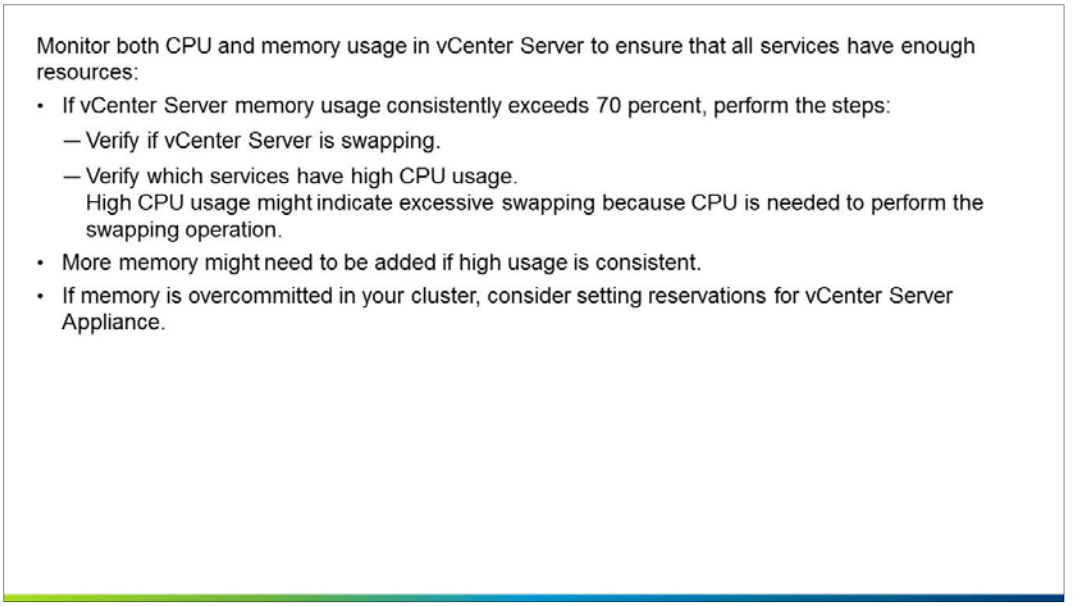
The following actions might require increasing the amount of resources used by the vCenter Server instance beyond the guidance provided in the VMware documentation:

- Adding extensions to vCenter Server
- Adding code to the vSphere Client
- Adding another extension point from where to get data
- Running more services might require more CPUs and memory.

  For example, with NSX, running the vSphere ESX Agent Manager (EAM) service is known to increase demand on resources.

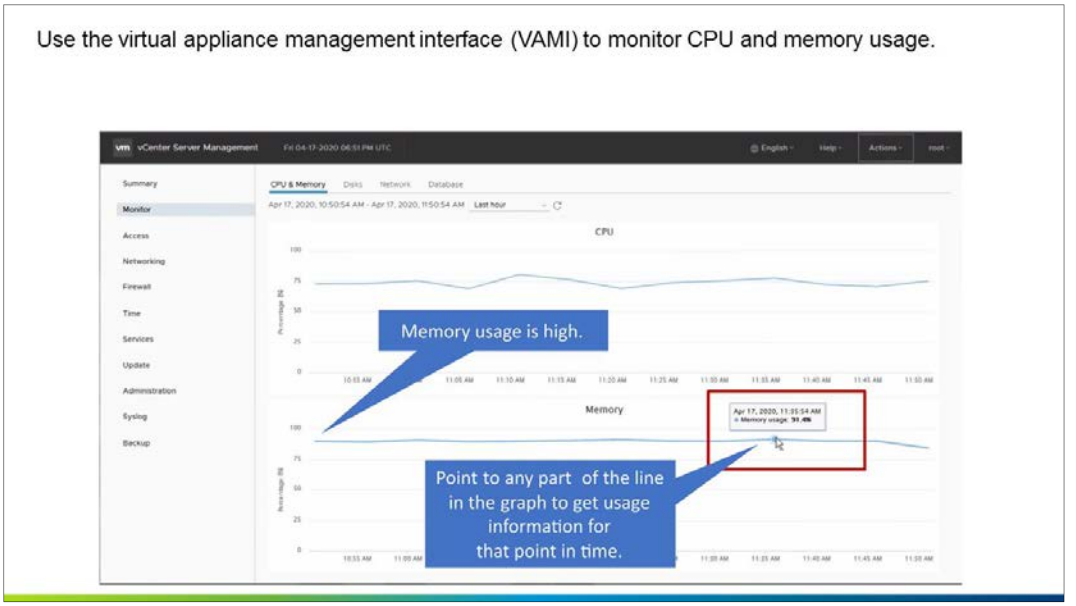
https://t.me/learningnets

9-7 vCenter Server Memory Performance



vCenter Server and vCenter Server Appliance are critical components of your software-defined data center (SDDC). For maximum performance and uptime, ensure that vCenter Server Appliance is provisioned with enough resources to keep its systems running optimally.

https://t.me/learningnets



To access the virtual appliance management interface (VAMI), open a web browser and navigate to `https://vCenter_Server_Name:5480`. Click **Monitor** in the left pane. The **CPU & Memory** tab displays by default.

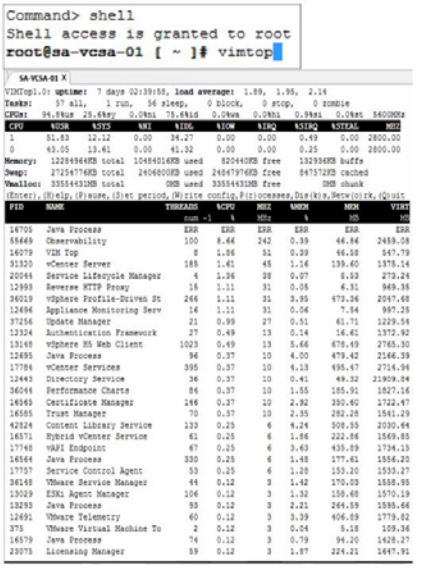
https://t.me/learningnets

9-9 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.
The `vimtop` tool as 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**.

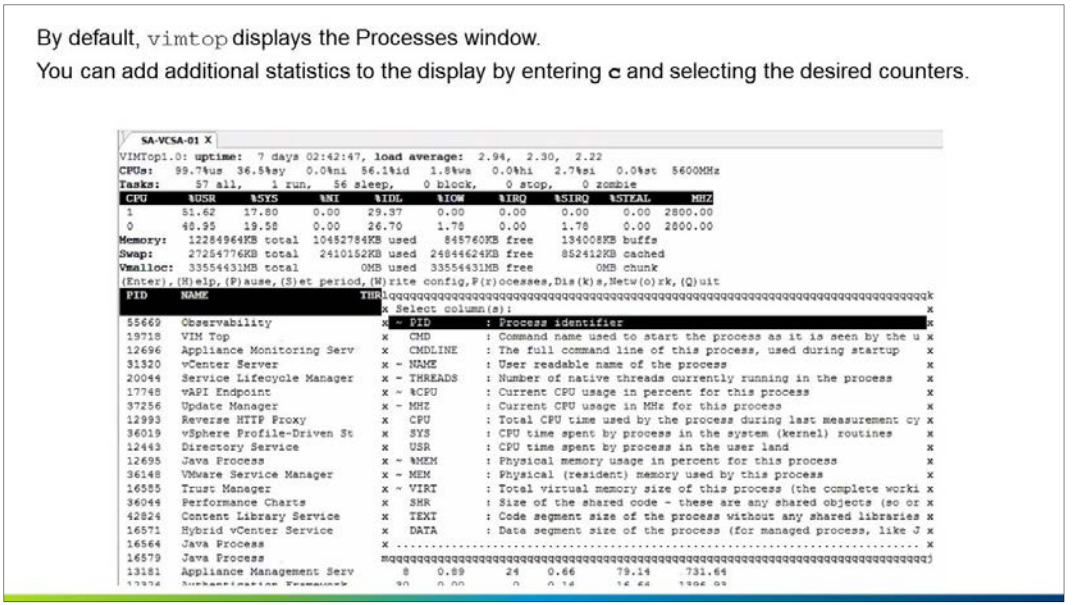
To quit `vimtop`, enter **q**.

https://t.me/learningnets

9-10 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.



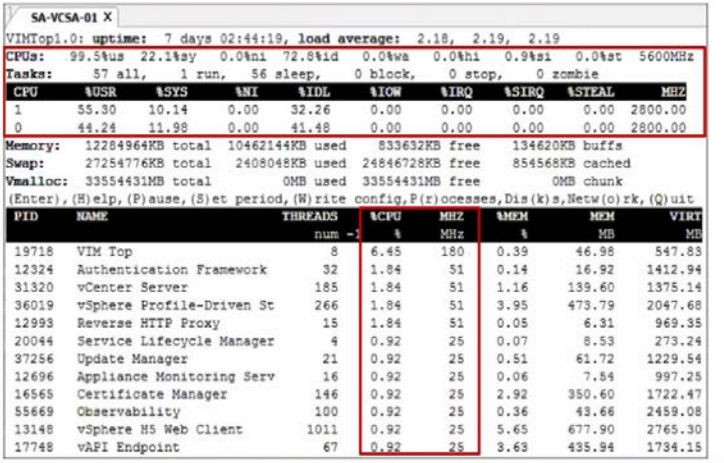
To add a metric (or column) to the `vimtop` display, perform the following steps:

1. Enter **c** to display the column selection pane.
2. Use the up and down arrow keys to highlight a metric in the column selection panel.
3. Press the space bar to select the metric.
   A tilde (**~**) appears to the left of the metric name.
4. Press the Esc key to exit the column selection panel.

https://t.me/learningnets

9-11 Monitoring CPU Use with vimtop

The Processes window shows overall CPU metrics and individual process CPU metrics.
To return to this window, enter **r**.



You can monitor aggregate CPU use from the overview pane or examine CPU use of individual process in the task pane.

Some columns that display include:

- Name: User-friendly name of the process (service).
- %CPU: Current CPU use (in percentage) for this process. If the value is 100 percent, the process is using an entire core.
- MHZ: Current CPU use (in megahertz) for this process.

https://t.me/learningnets

9-12 Activity: CPU and Memory Use (1)

As a general guideline, you should keep CPU and memory use by a vCenter Server instance below what percentage?

- ❒ 60%
- ❒ 70%
- ❒ 80%
- ❒ 90%

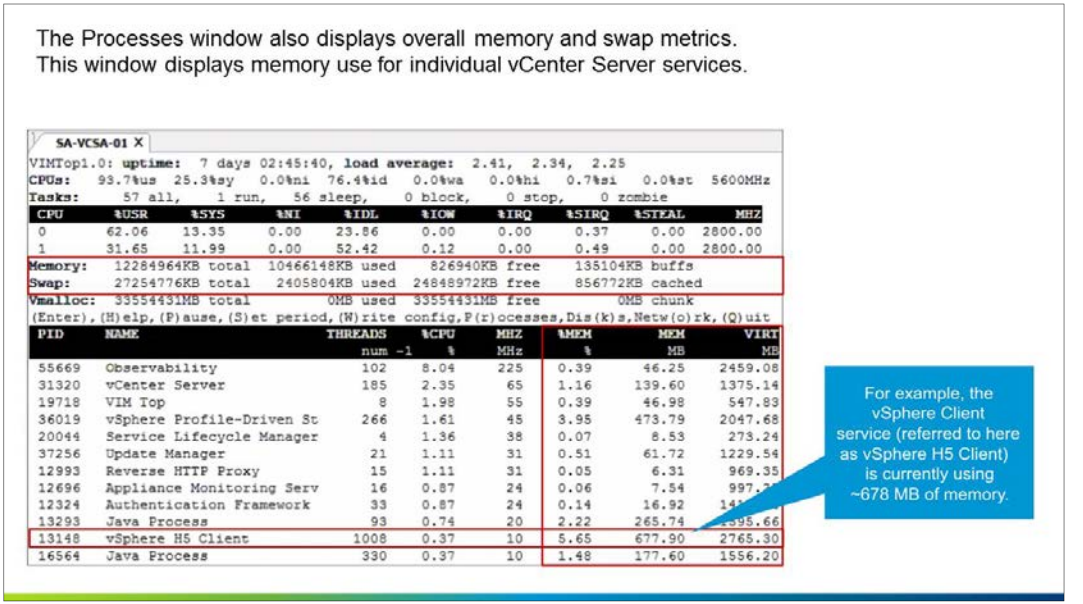
https://t.me/learningnets

9-13 Activity: CPU and Memory Use (2)

As a general guideline, you should keep CPU and memory use by a vCenter Server instance below what percentage?

- ❒ 60%
- ✓ 70%
- ❒ 80%
- ❒ 90%

https://t.me/learningnets



Some columns that display include:

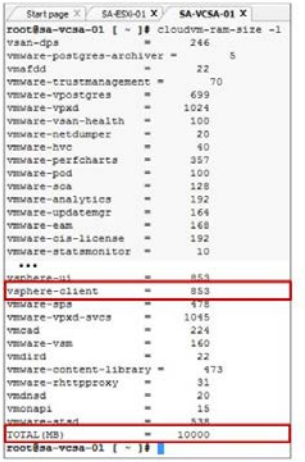
- %MEM: Physical memory use (in percentage) for this process.
- MEM: Physical or resident memory (in MB) used by this process.
- VIRT: Total virtual memory size (in MB). This value represents the complete working set, which includes resident and swapped memory.

https://t.me/learningnets

9-15 Viewing Service Memory Allocation with cloudvm-ram-size

Use the `cloudvm-ram-size` command shows the amount of memory allocated to the vCenter Server services.
The `cloudvm-ram-size -l` command displays the memory allocated (in MB) for each service.

For example, the vSphere 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).



https://t.me/learningnets

9-16 Viewing Service Heap Size

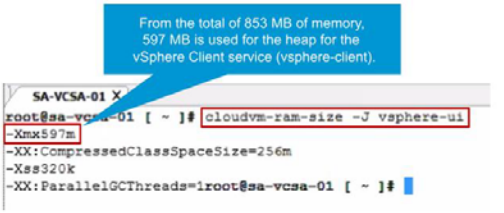
vCenter Server consists of several Java services, such as the vSphere Client service (vsphere-client). A Java service uses a heap (pre-reserved area of memory) to store data objects.

Proper heap size is important for Java services to perform optimally.
For example, vSphere Web Client might need more heap size in these 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`



https://t.me/learningnets

9-17 Changing 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, the following command changes the heap size of the vSphere Web Client service to 700 MB:

  ```
  cloudvm-ram-size -C 700 vsphere-client
  ```

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

9-18 vCenter Server Network Performance

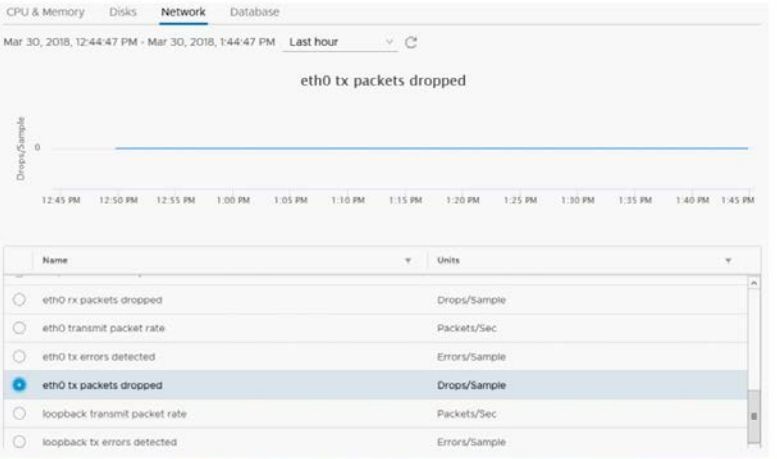
Many vCenter Server operations involve communication between vCenter Server and the ESXi hosts:

- 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.

https://t.me/learningnets

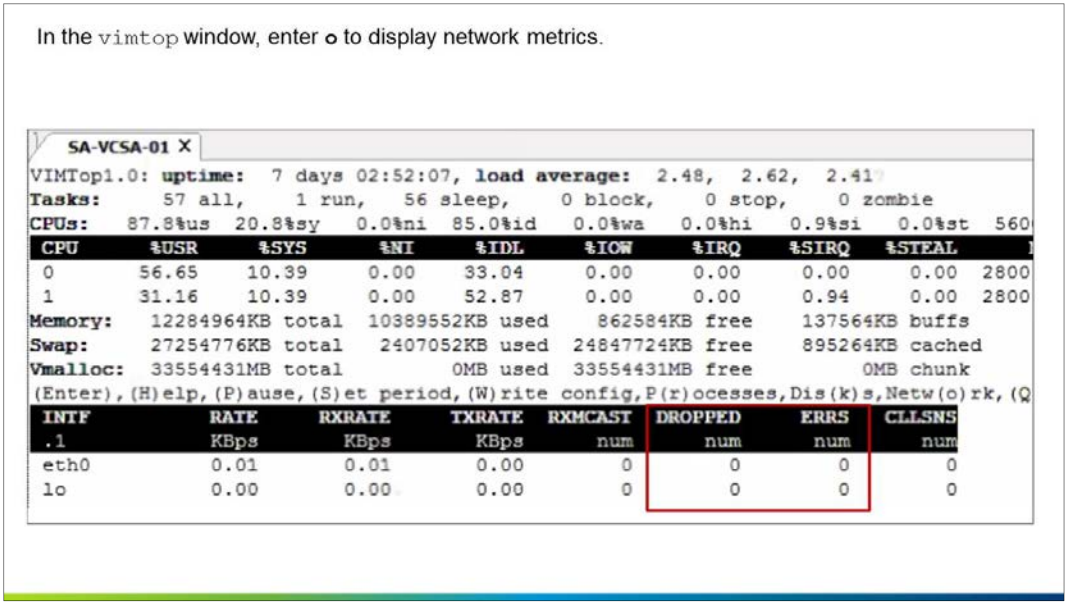
9-19 Monitoring Network Activity with the VAMI

Use the VAMI to monitor network activity, which includes byte rate, packets dropped, errors detected, and packet rate



https://t.me/learningnets

9-20 Monitoring Network Activity with vimtop

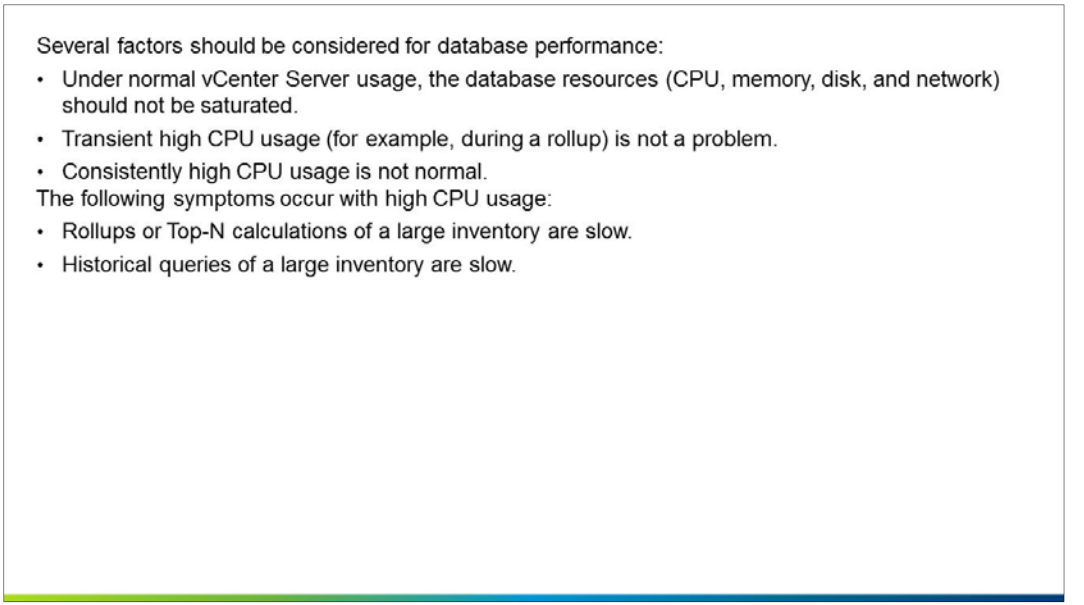


You can monitor the following columns for network activity:

- 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

9-21 vCenter Server Database Performance (1)



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 many 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 only mean a high number of hosts and VMs, it might mean many datastores, networks, or devices per VM. Some queries require many joins or full scans on tables, which is not ideal but does occur. 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

9-22 vCenter Server Database Performance (2)

If you see slow historical queries and consistently high database CPU usage:

- Verify 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.

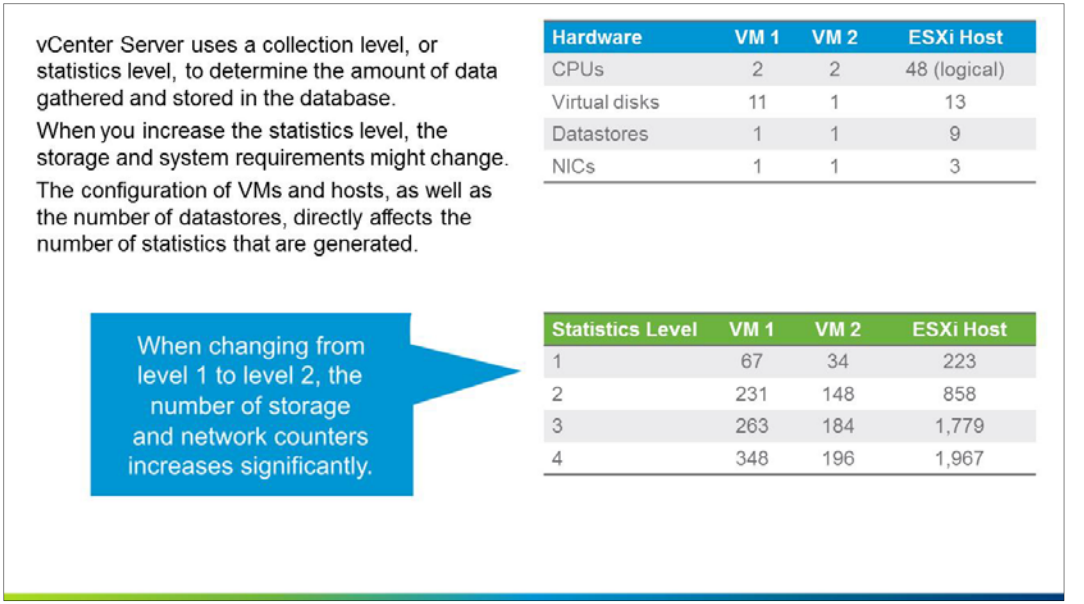
If you experience slow rollups or Top-N calculations:

- Add CPUs.
- Verify memory.
- Verify rollup status.
- Verify alarms.

You can verify 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), each of which typically lasts between three and four seconds. If these events are extremely frequent and each lasts 10 or more seconds, this behavior often indicates problems in your database. Verify the resources (I/O, CPU, and memory) and that the database is correctly provisioned.

https://t.me/learningnets

9-23 Effects of Changing Statistics Level on Database Traffic



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

https://t.me/learningnets

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

9-24 vCenter Server Appliance Disk Use (1)

The vCenter Server database is the main disk bandwidth consumer and is write-intensive. Place write-intensive disk partitions, such as `/storage/db`, on high-speed storage.

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` is the location of the vCenter Server database.
- `/storage/dblog`
- `/storage/seat`

When using vCenter Server Appliance, ensure that these partitions are on high-speed storage and have sufficient space.

If you must 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

9-25 vCenter Server Appliance Disk Use (2)

Place other database partitions, such as `/storage/dblog` and `/storage/seat`, on high-speed storage.

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

https://t.me/learningnets

9-26 vCenter Server Appliance Disk Use (3)

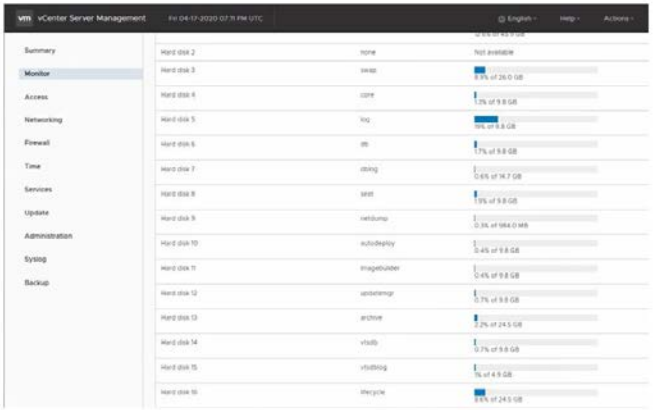
Disk (VMDK)	Minimum Size	Mount Point	Purpose
VMDK13	50 GB	/storage/archive	Repository used for archiving files
VMDK14	10 GB	/storage/vtsdb	VMware vTSDB service repository that stores the stats
VMDK15	5 GB	/storage/vtsdblog	VMware vTSDB service repository that stores the logs of the service
VMDK16	100 GB	/storage/lifecycle	Workload Control Plane service stage directory that stores the binaries for install, update, and upgrade

https://t.me/learningnets

9-27 Monitoring Disk Use with the VAMI

Use the VAMI to monitor disk use by partition.

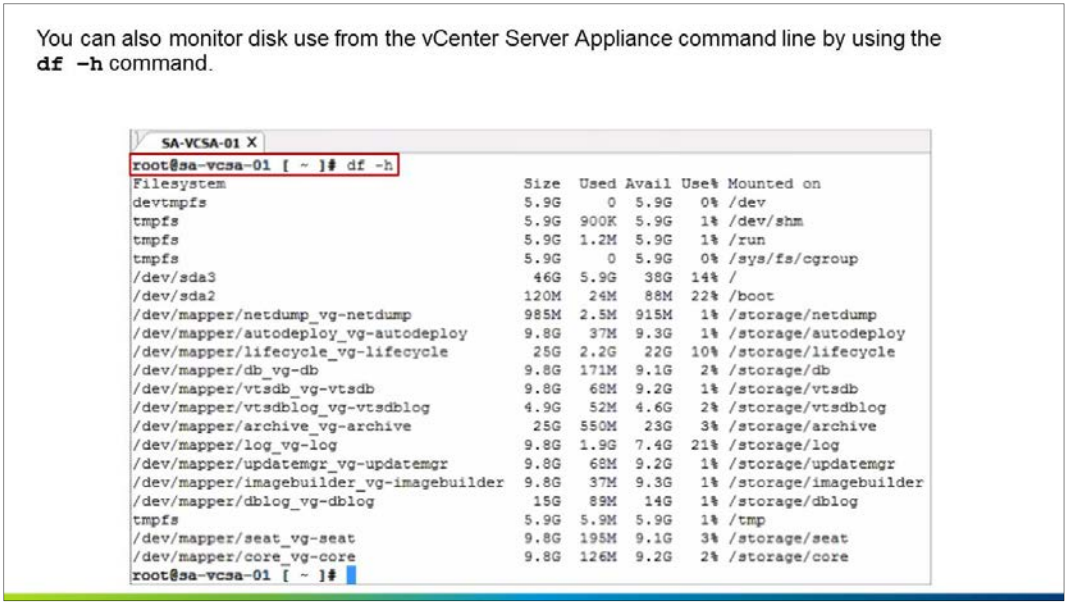
Point to a partition to view information on used space, available space, and total space.



https://t.me/learningnets

9-28 Monitoring Disk Usage with the df Command

You can also monitor disk use from the vCenter Server Appliance command line by using the **df -h** command.

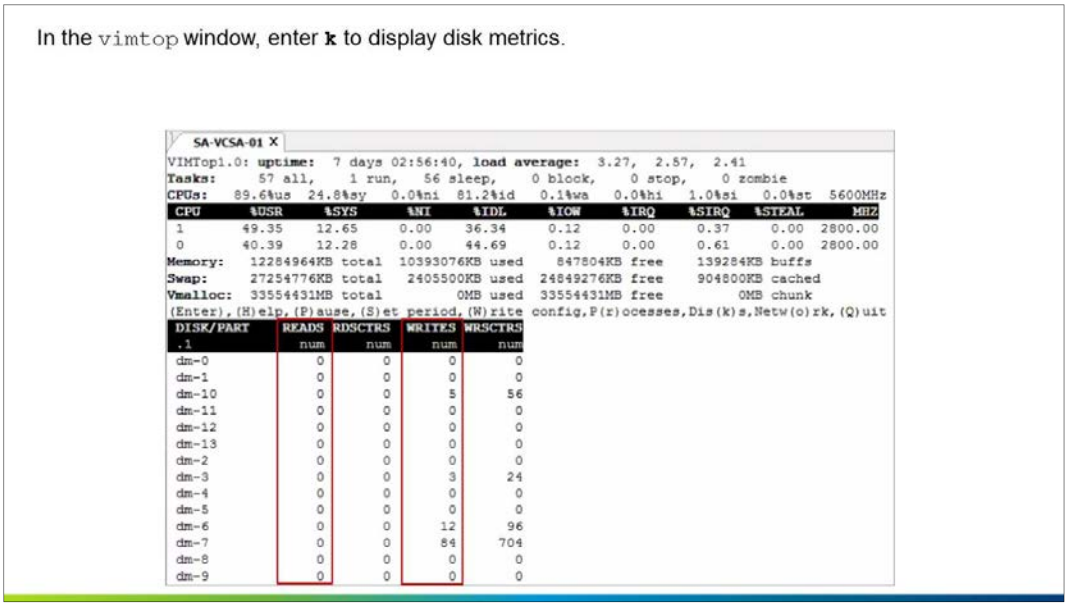


The **df** command displays file system use information, such as total size, used space, and available space. The **-h** option shows sizes in powers of 1024, for example, in megabytes, gigabytes, and so on.

https://t.me/learningnets

9-29 Monitoring Disk Activity with vimtop

In the `vimtop` window, enter **k** to display disk metrics.



The following columns show:

- 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.

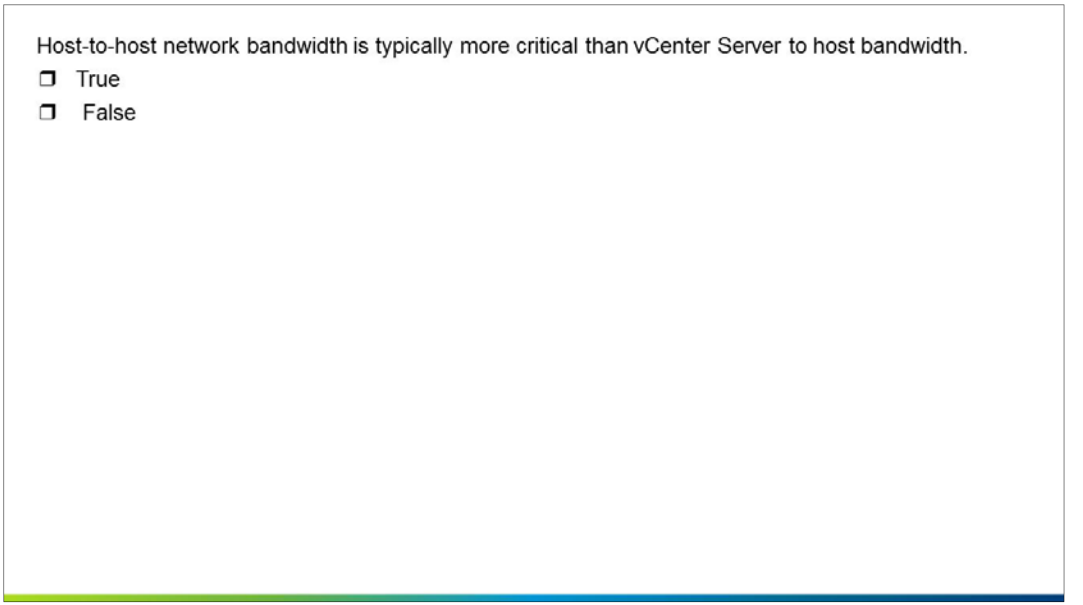
https://t.me/learningnets

9-30 Monitoring Database Activity with the VAMI

Use the VAMI to monitor space usage trends for SEAT (stats, events, alarms, and tasks) activity, database log activity, and inventory data (core) activity.

https://t.me/learningnets

9-31 Activity: Network Performance (1)



https://t.me/learningnets

9-32 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

9-33 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

9-34 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

9-35 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Review vCenter Server components and services
- Describe the factors that influence vCenter Server performance
- Use vCenter Server Appliance tools to monitor vCenter Server resource use

https://t.me/learningnets

9-36 Key Points

- vCenter Server comprises many services.
- You can monitor vCenter Server Appliance using tools such as the VAMI, `vimtop`, `df`, and `cloudvm-ram-size`.
- As a general guideline, vCenter Server CPU and memory usage should not exceed 70 percent.
- 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.

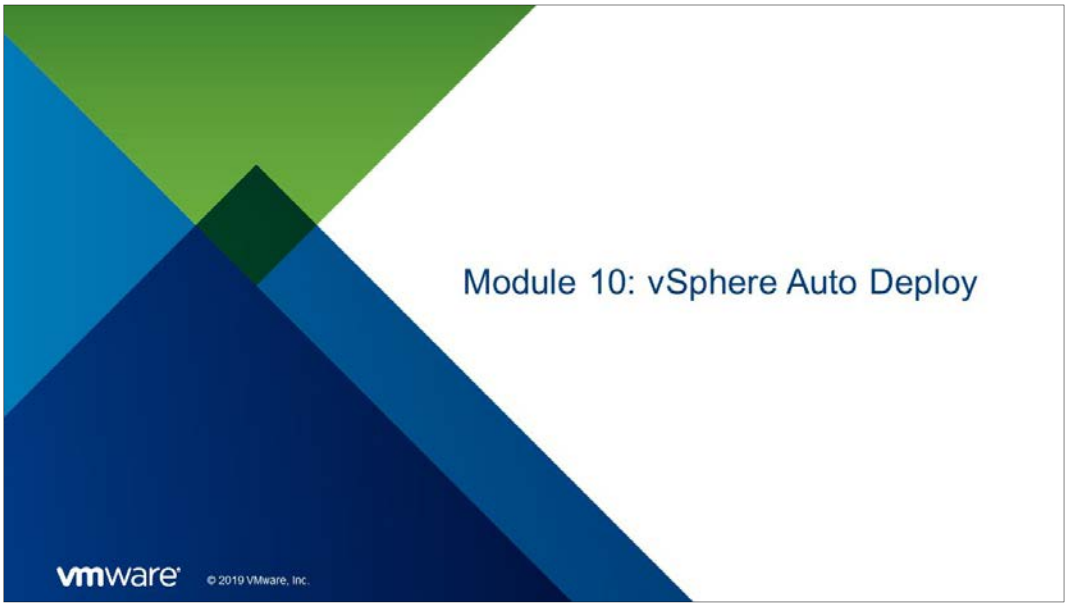
Questions?

https://t.me/learningnets

https://t.me/learningnets

Module 10

Appendix: vSphere Auto Deploy



https://t.me/learningnets

10-2 Importance

You can use the vSphere ESXi Image Builder CLI to manage software depots, image profiles, and software packages (VIBs). With vSphere Auto Deploy, you can deploy an ESXi image that is created with vSphere ESXi Image Builder.

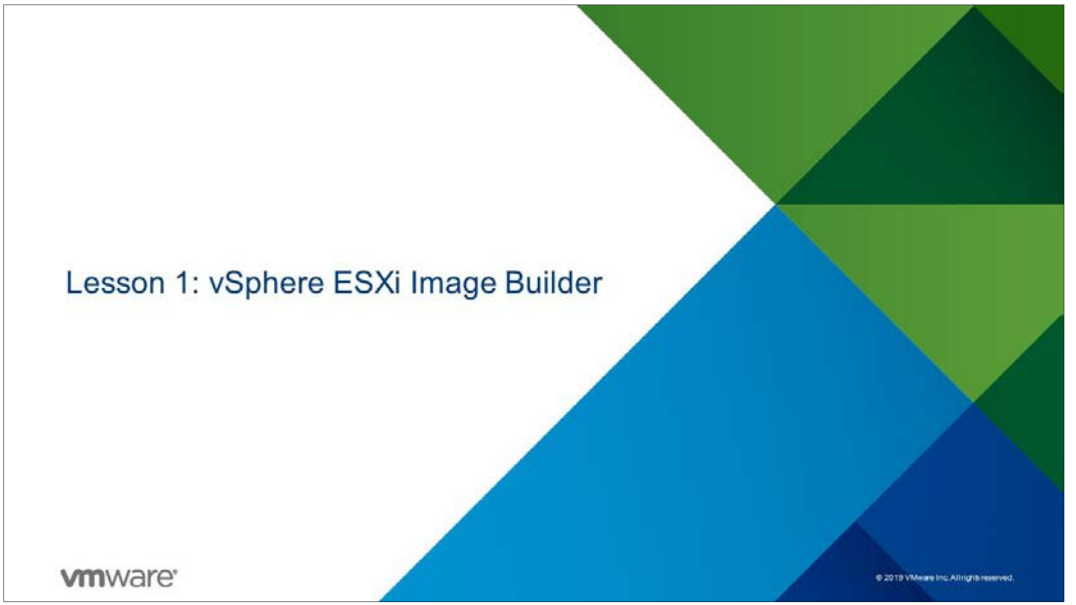
https://t.me/learningnets

10-3 Appendix Lessons

1. vSphere ESXi Image Builder
2. Overview of vSphere Auto Deploy
3. Configuring vSphere Auto Deploy

https://t.me/learningnets

10-4 Lesson 1: vSphere ESXi Image Builder



https://t.me/learningnets

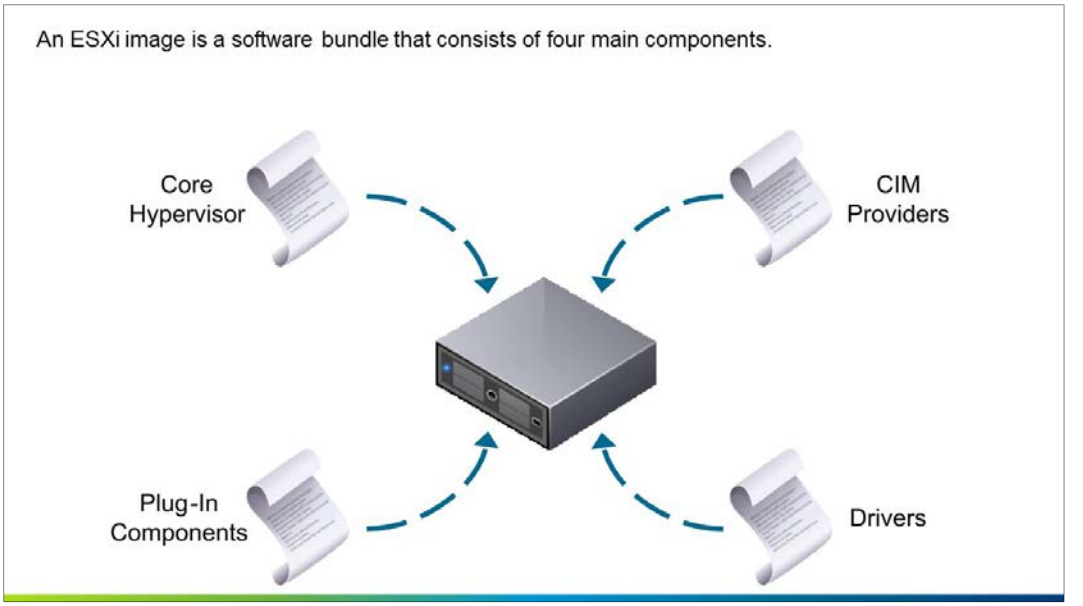
10-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the purpose of vSphere ESXi Image Builder
- Use the vSphere Client to create an ESXi image

https://t.me/learningnets

10-6 ESXi Images



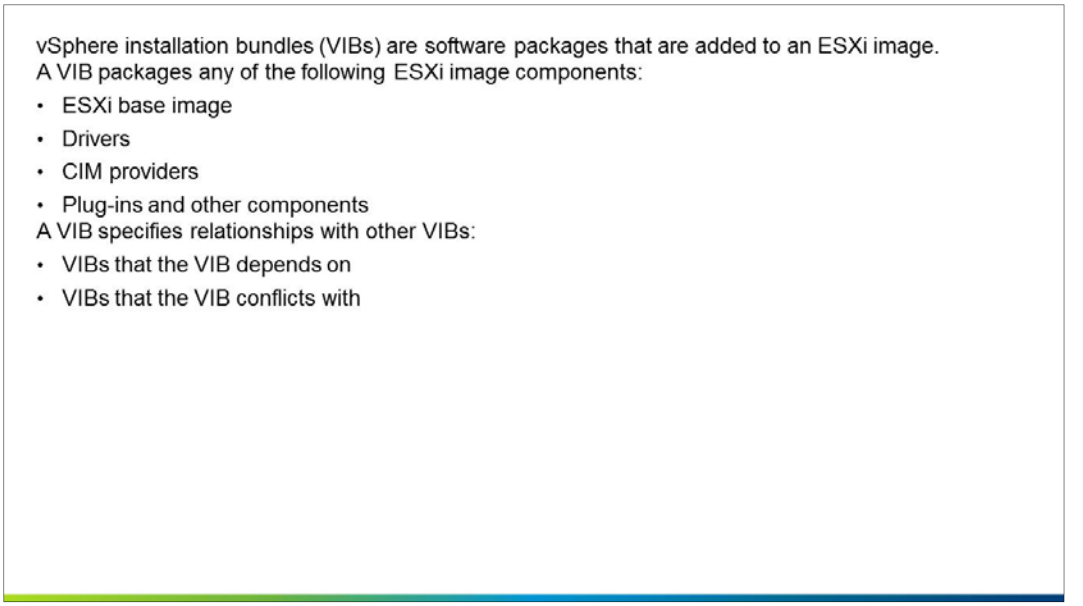
An ESXi image contains all the software necessary to run on an ESXi host and includes the following components:

- 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

10-7 vSphere Installation Bundles

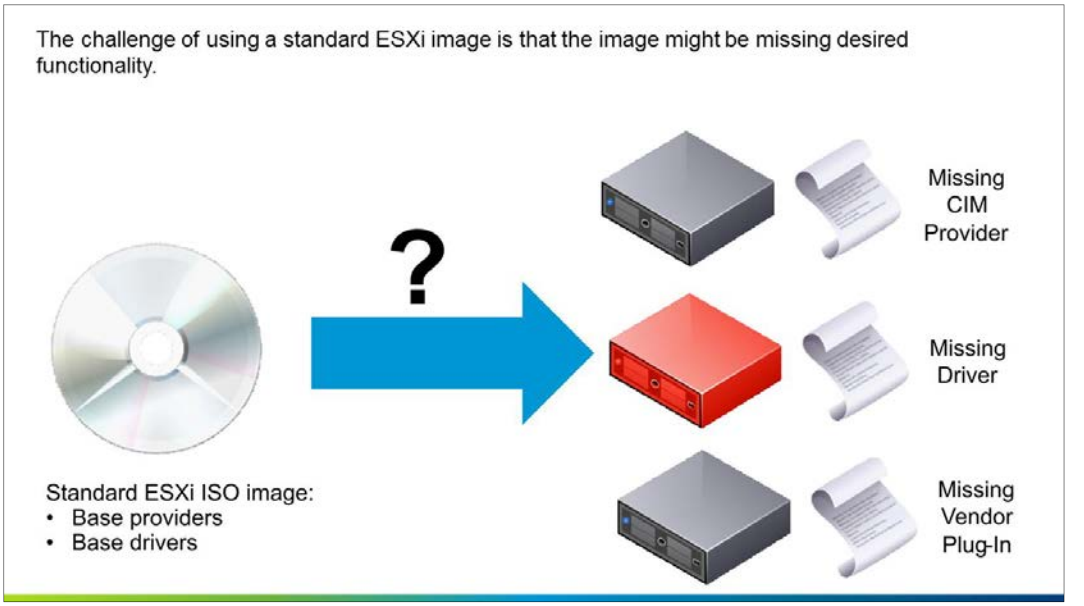


An ESXi image includes one or more vSphere installation bundles (VIBs). A VIB is an ESXi software package. VMware and its partners package solutions, drivers, CIM providers, and applications that extend the ESXi platform in VIBs.

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

10-8 ESXi Image Deployment



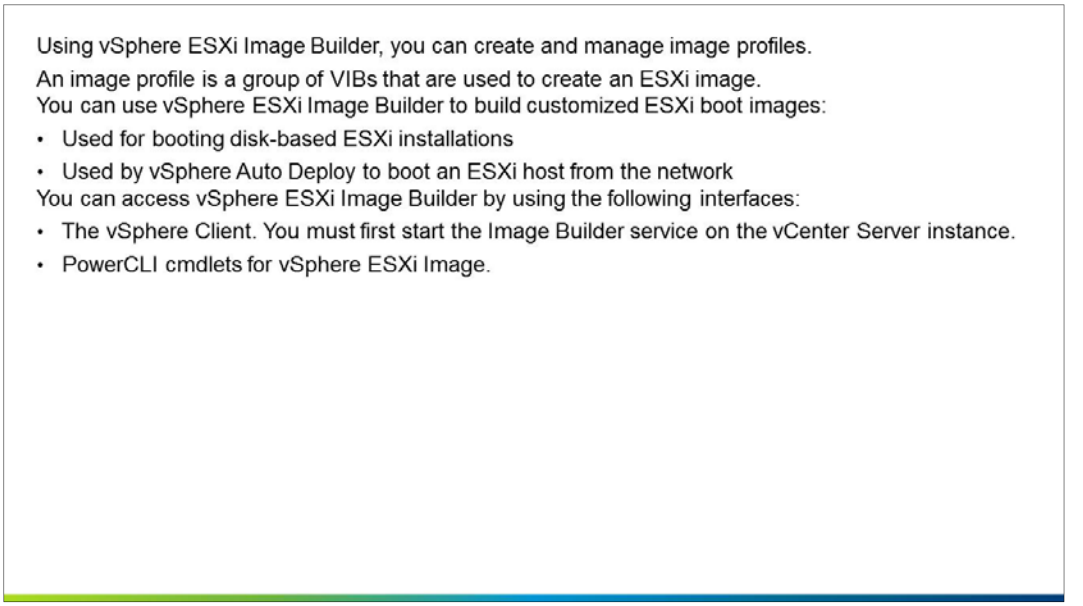
VMware provides standard ESXi images, which are available on the VMware website. ESXi images can also be provided by VMware partners.

The challenge that administrators face when using the standard ESXi image that VMware provides 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 the vSphere Client.

https://t.me/learningnets

10-9 vSphere ESXi Image Builder Overview



vSphere ESXi Image Builder is a utility for customizing ESXi images. You access vSphere ESXi Image Builder through the vSphere Client. You use vSphere ESXi Image Builder to create and manage VIBs, image profiles, software depots, and software channels.

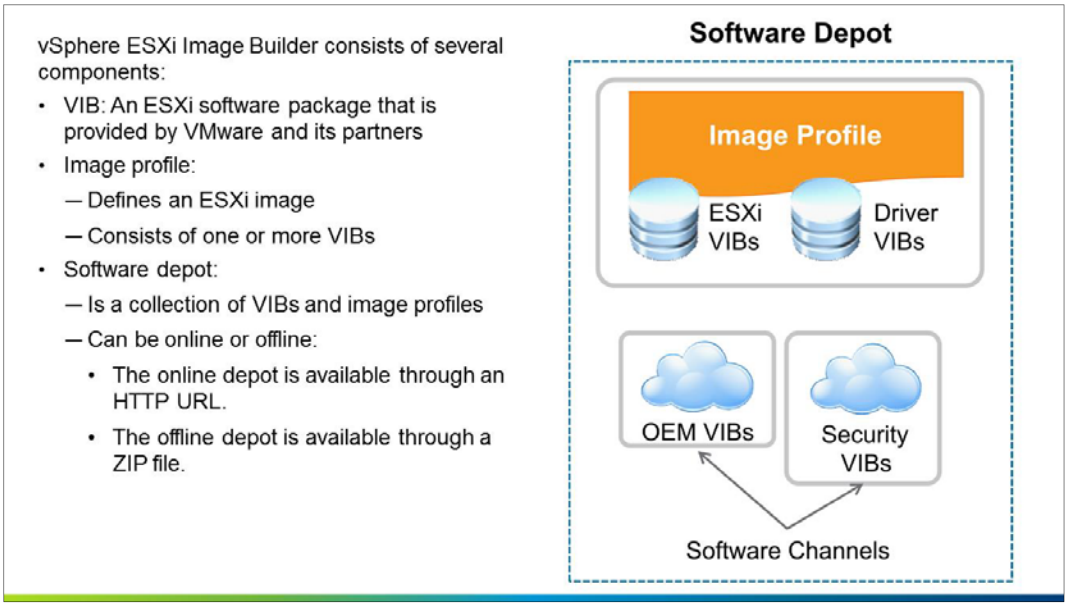
For more information about vSphere ESXi Image Builder and its prerequisite software, see *vSphere ESXi Image Builder Installation and Usage* at

https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.esxi.install.doc/GUID-62B15826-B529-4519-B57A-98DFD0CC5522.html.

For more information about PowerCLI, see the documentation page at https://www.vmware.com/support/developer/PowerCLI.

https://t.me/learningnets

10-10 vSphere ESXi Image Builder Architecture



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 the vSphere 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 that you can 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.

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

10-11 Building ESXi Images

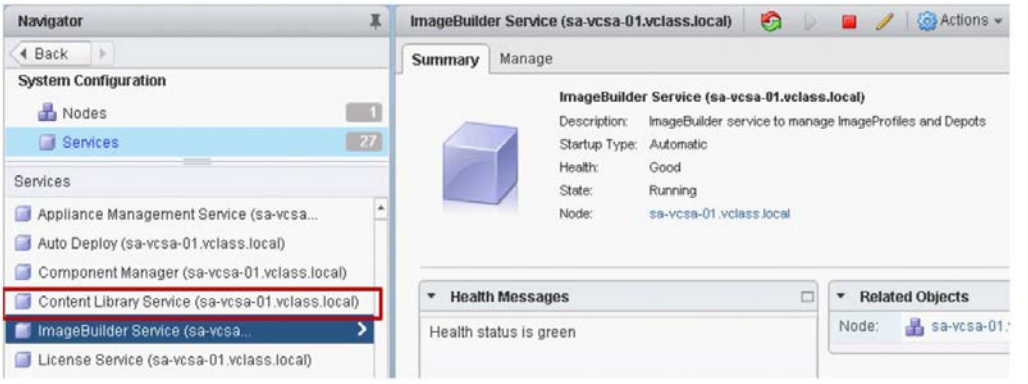
To build an ESXi image:

1. Start the Image Builder service.
2. Connect to a software depot.
3. Create an image profile.
4. Generate the new ESXi image.

https://t.me/learningnets

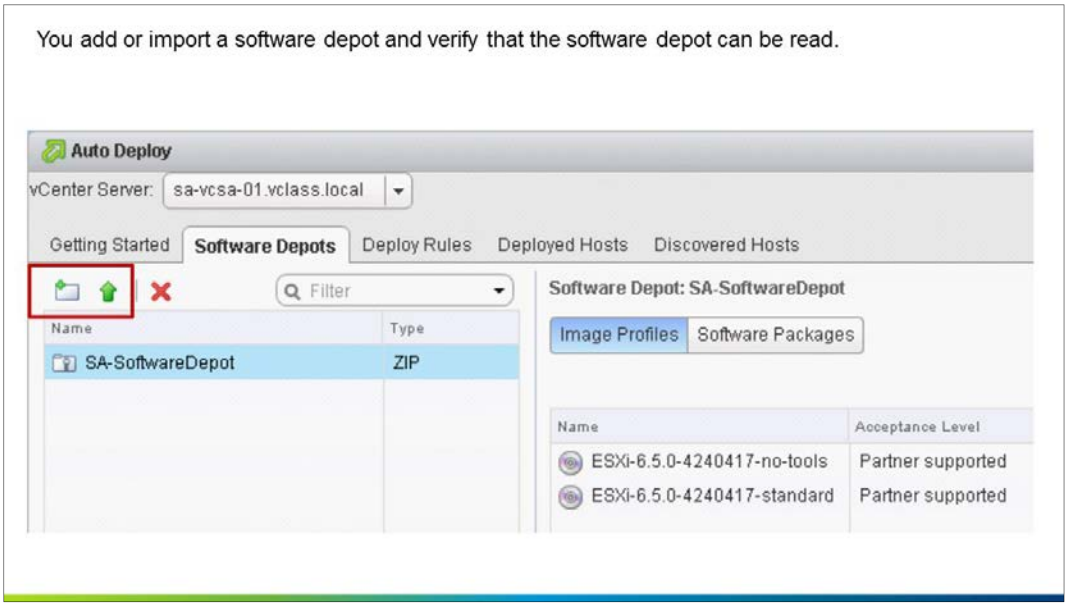
10-12 Step 1: Starting the Image Builder Service

Using the Image Builder service, you can manage and customize image profiles through the vSphere Client.



https://t.me/learningnets

10-13 Step 2: Connecting to a Software Depot



Before you can create or customize an ESXi image, vSphere ESXi Image Builder must be able to access one or more software depots.

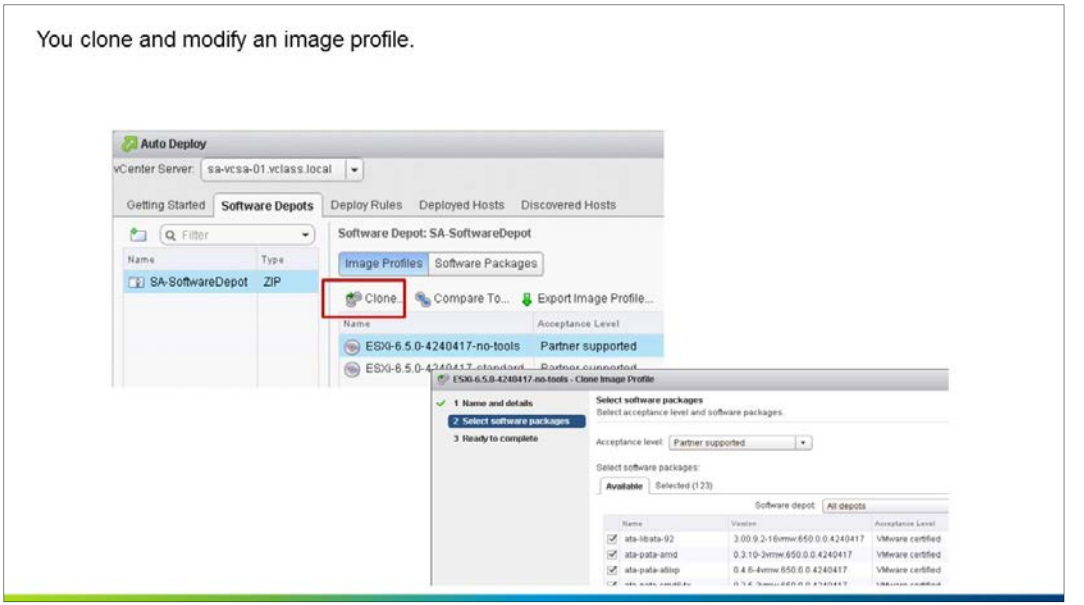
Use the vSphere Client to add an ESXi software depot or offline depot ZIP file to vCenter Server.

You can then create 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 that are found in the depot.

https://t.me/learningnets

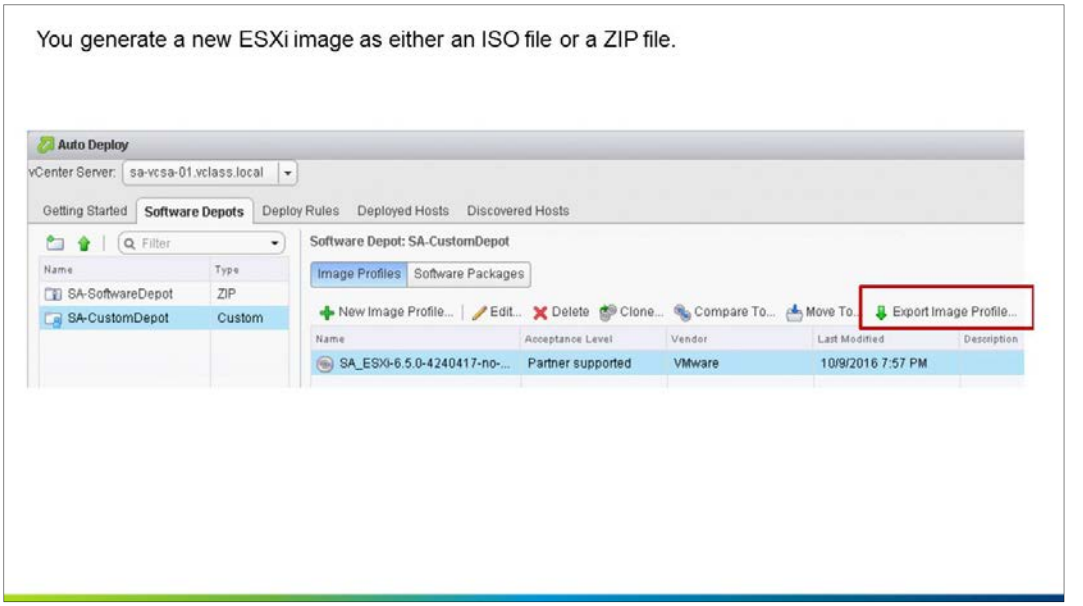
10-14 Step 3: Creating an Image Profile



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.

https://t.me/learningnets

10-15 Step 4: Generating the New ESXi Image



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 vSphere Update Manager for remediating ESXi hosts. The exported image profile can also be used with vSphere Auto Deploy to boot ESXi hosts.

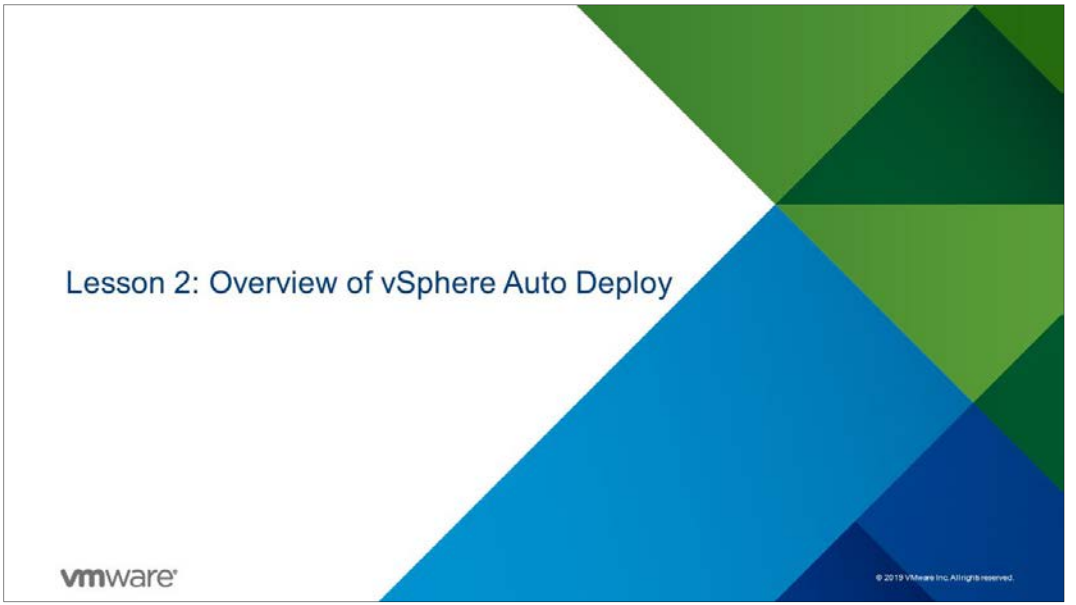
https://t.me/learningnets

10-16 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the purpose of vSphere ESXi Image Builder
- Use the vSphere Client to create an ESXi image

https://t.me/learningnets



https://t.me/learningnets

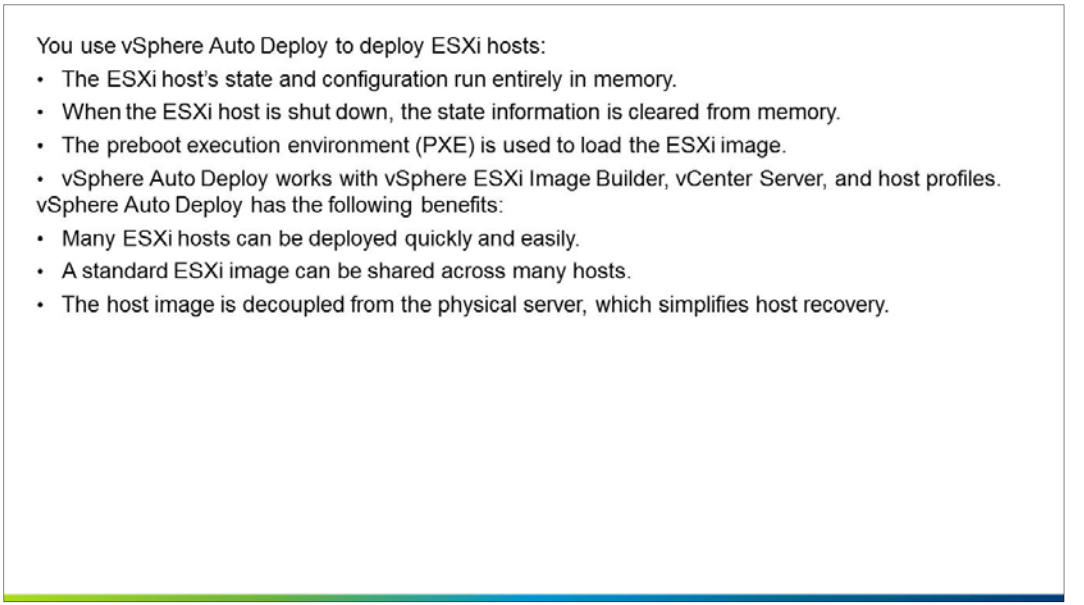
10-18 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the purpose of vSphere Auto Deploy
- Describe how an autodeployed ESXi host boots

https://t.me/learningnets

10-19 About vSphere Auto Deploy



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 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 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.

With vSphere Auto Deploy, many hosts can be rapidly deployed. 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, freeing the boot device for other uses.

https://t.me/learningnets

For more information about vSphere Auto Deploy, see *Using vSphere Auto Deploy to Reprovision Hosts*

at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.esxi.upgrade.doc/GUID-E2A5AB39-3EA7-4D26-8F17-3AB4BCC1EBC4.html.

https://t.me/learningnets

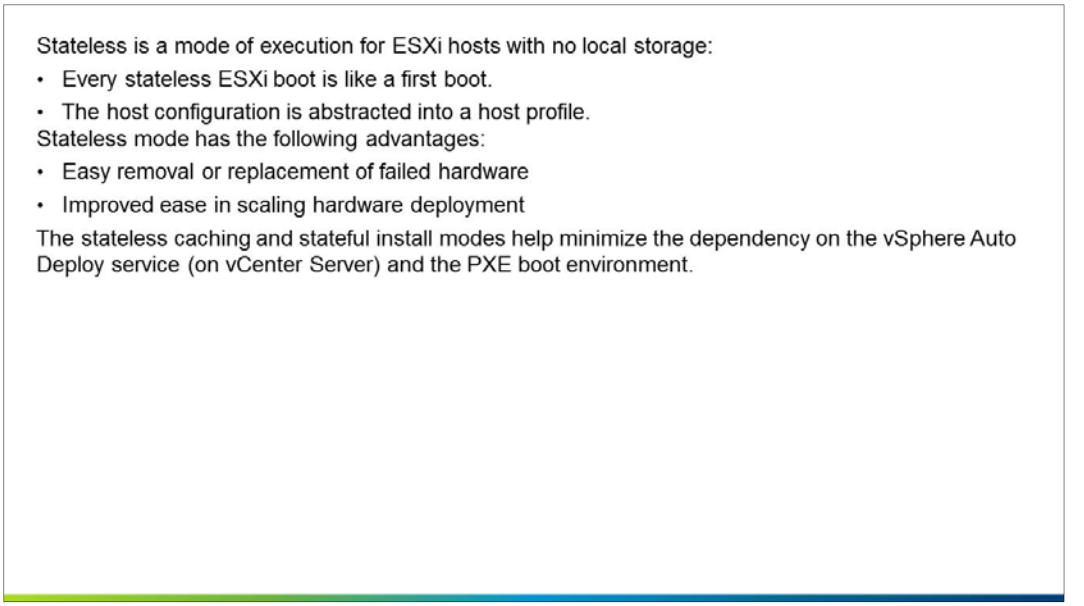
10-20 vSphere Auto Deploy Modes

You can autodeploy hosts by using the following modes:

- Stateless
- Stateless caching
- Stateful install

https://t.me/learningnets

10-21 vSphere Auto Deploy Stateless Mode

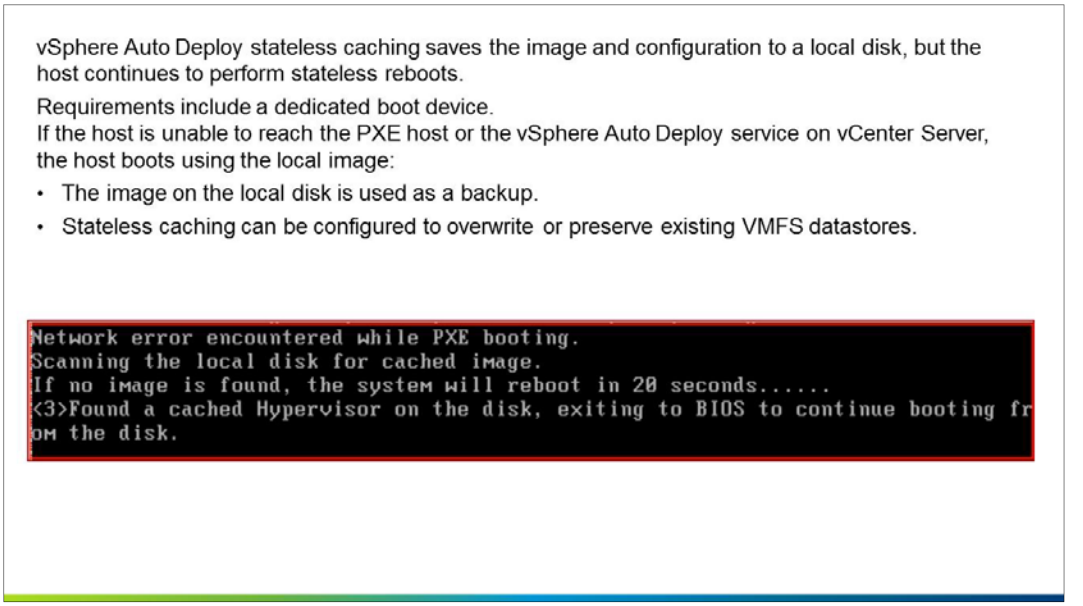


With stateless mode, the ESXi host boots with networking connectivity to vCenter Server through the built-in standard switch. If the host profile specifies distributed switch membership, vCenter Server joins the ESXi host to the distributed switch.

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, vCenter Server, or the Dynamic Host Configuration Protocol (DHCP) and Trivial File Transfer Protocol (TFTP) servers are unavailable.

https://t.me/learningnets

10-22 vSphere Auto Deploy Mode: Stateless Caching



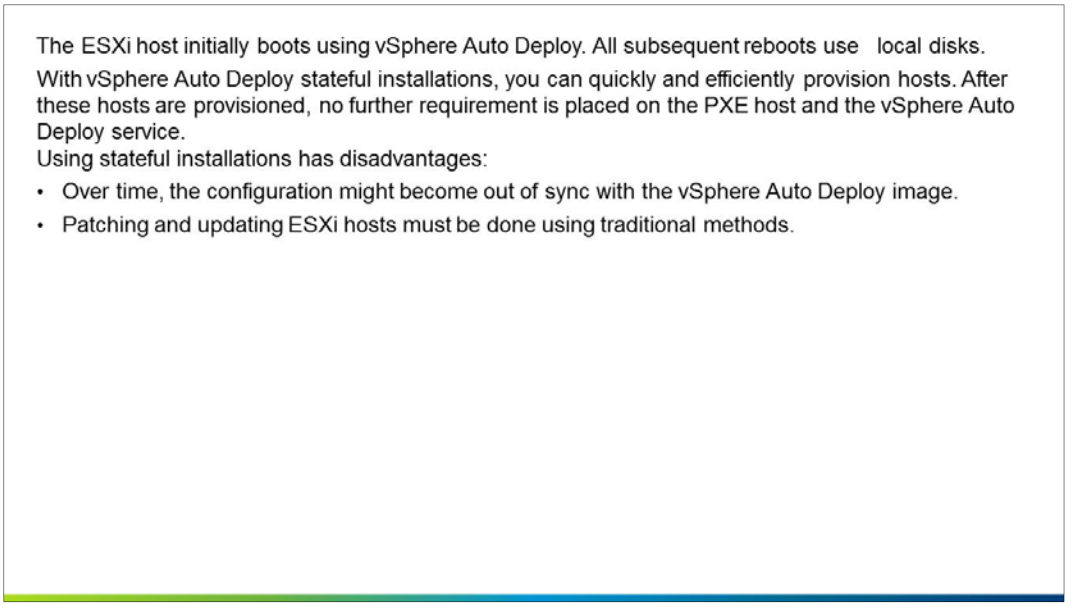
vSphere Auto Deploy stateless caching PXE boots the ESXi host and loads the image in memory which is similar to the process for loading an ESXi host using stateless mode. However, when the host profile is applied to the ESXi host, the image running in memory is copied to a boot device. The saved image acts as a backup in case the PXE infrastructure, vCenter Server, or vSphere Auto Deploy service are unavailable. If the host must reboot and it cannot contact the DHCP or TFTP server, or vSphere Auto Deploy service, the network boot times out and the host reboots using the cached disk image.

Stateless caching might help ensure availability of an ESXi host because the host can boot if an outage occurs that affects the DHCP, TFTP, or vCenter Server, or the vSphere Auto Deploy service. However, 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 you can 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 is assigned to the host.

https://t.me/learningnets

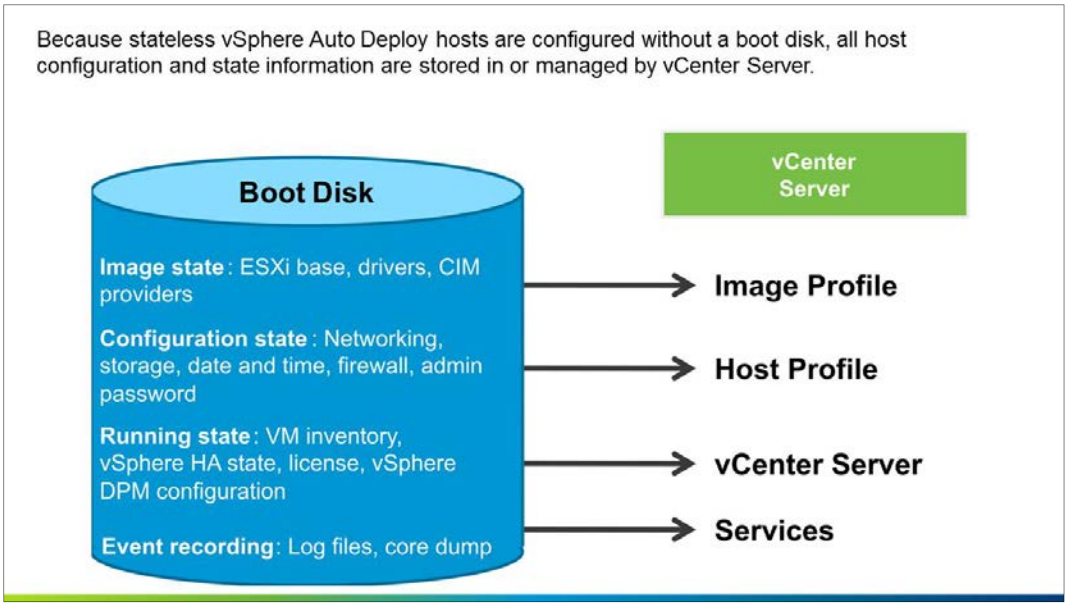
10-23 vSphere Auto Deploy Mode: Stateful Installation



Setting up stateful installation is similar to configuring stateless caching. But instead of attempting to use the vSphere Auto Deploy service on 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

10-24 Locations for Configuration and State Information



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 the host's information. The host information that is stored includes the following items:

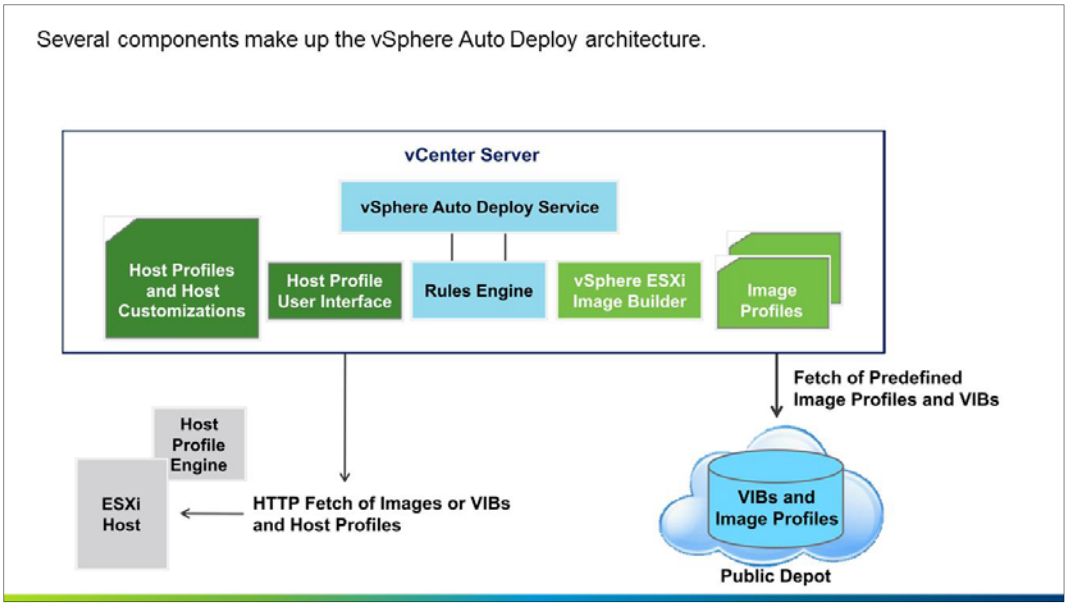
- Image state: Executable software to run on the ESXi host. The information is part of the image profile, which can be created and customized with the vSphere Client or vSphere ESXi Image Builder.
- Configuration state: Configurable settings that determine how the host is configured. Examples include virtual switch settings, boot parameters, and driver settings. Host profiles are created using the host profile user interface in the vSphere Client.

https://t.me/learningnets

- Running state: Settings that apply while the ESXi host is up and running. This state also includes the location of the VM in the inventory and the VM autostart information. The vCenter Server instance manages state information.
- Event recording: Information found in log files and core dumps. This information can be managed by vCenter Server, using services such as vSphere ESXi Dump Collector and vSphere Syslog Collector. vSphere Syslog Collector uses Linux rsyslogd.

https://t.me/learningnets

10-25 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 the 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 the vSphere 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

https://t.me/learningnets

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

10-26 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.

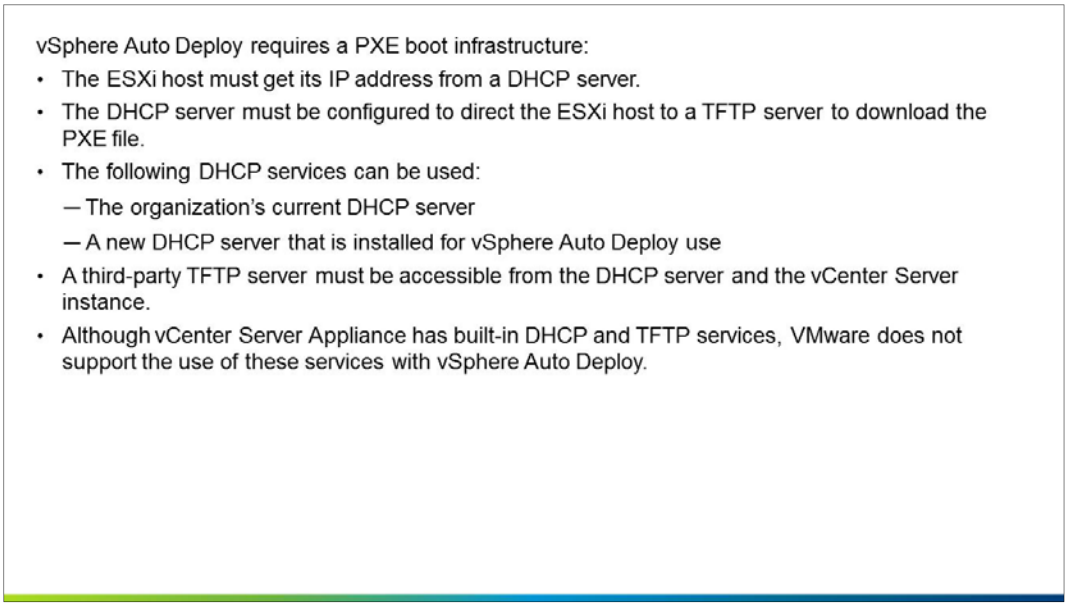
https://t.me/learningnets

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 the vSphere 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 used to boot the host. If more than one item is mapped by the rules, the vSphere Auto Deploy service uses the first item in the rule set.
- Working rule set: You can use the working rule set to test changes to rules before making them active. For example, you can use the vSphere Client to test compliance with the working rule set.

https://t.me/learningnets

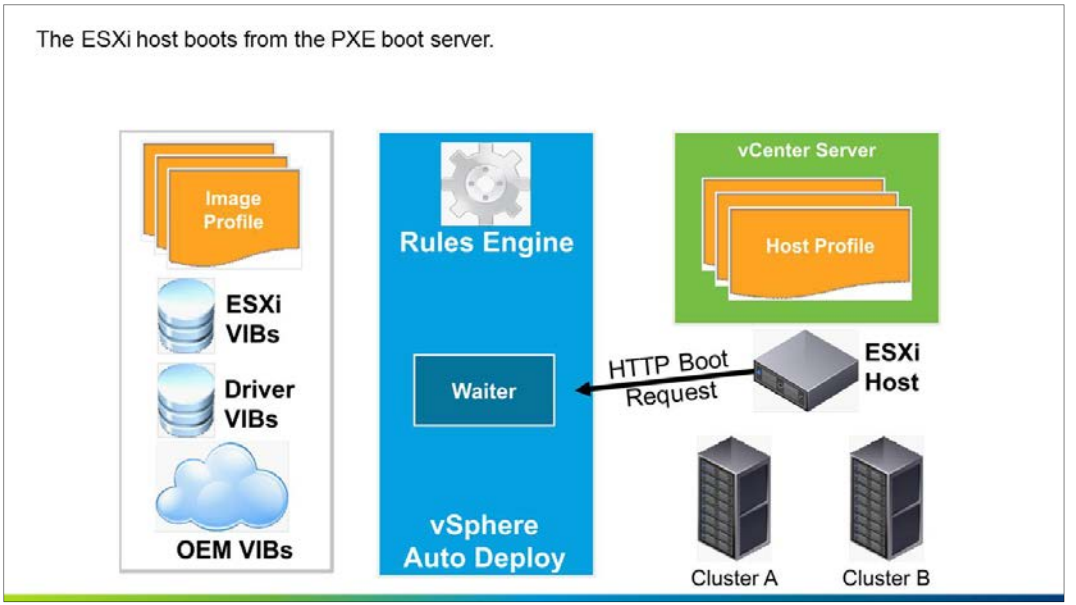
10-27 PXE Boot Infrastructure Setup



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

10-28 Initial Boot of an Autodeployed ESXi Host (1)

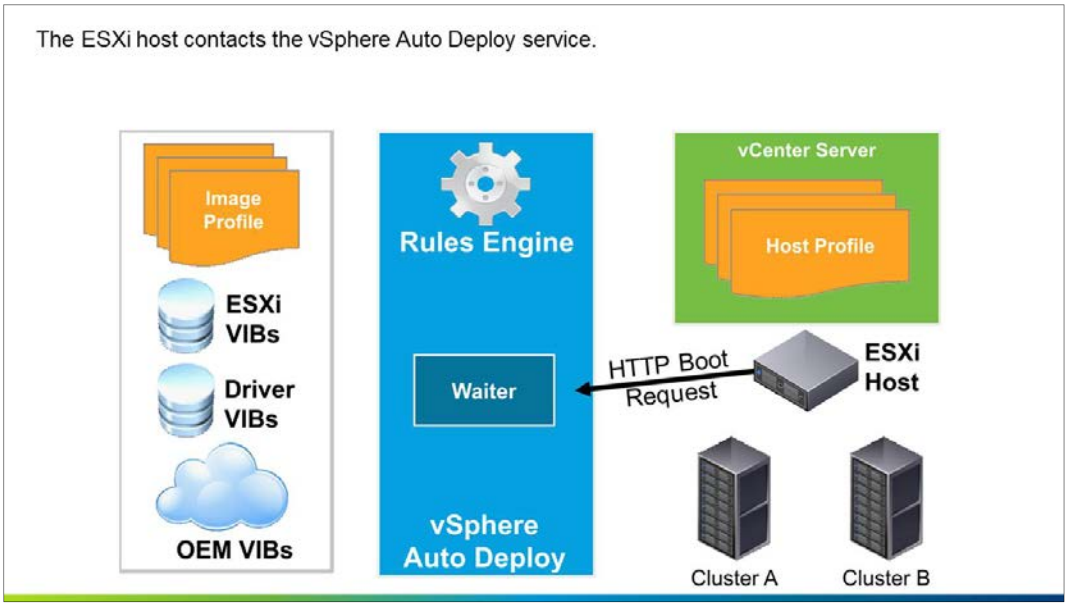


When an autodeployed host boots for the first time, the following events occur:

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

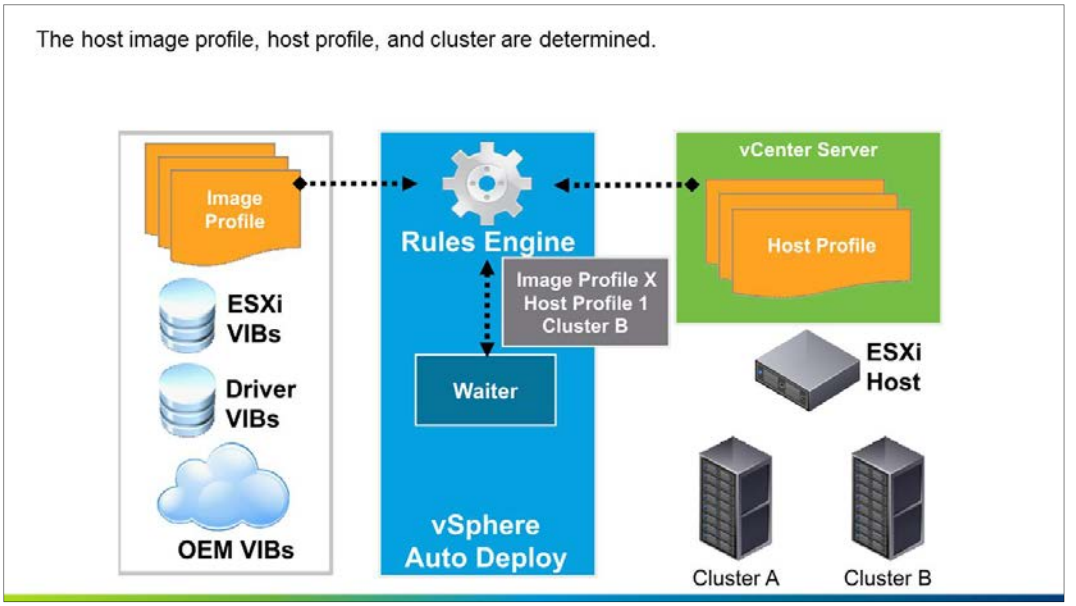
10-29 Initial Boot of an Autodeployed ESXi Host (2)



The gPXE configuration file instructs the host to make an HTTP boot request to the vSphere Auto Deploy service.

https://t.me/learningnets

10-30 Initial Boot of an Autodeployed ESXi Host (3)

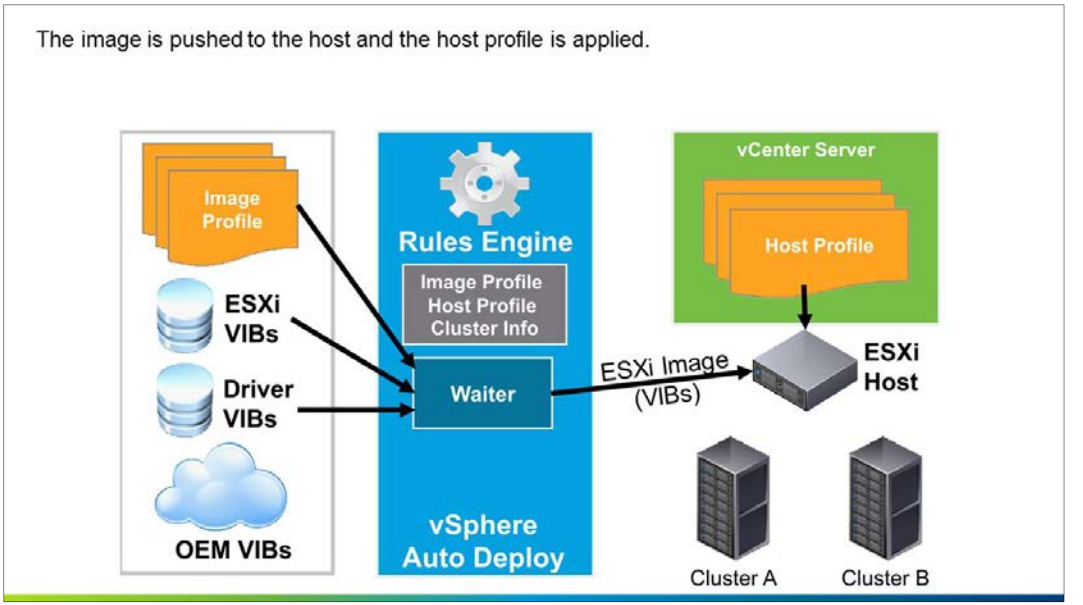


The vSphere Auto Deploy service queries the rules engine for the following information about the host:

- The image profile to use
- The host profile to use
- Which cluster the host belongs to, if any

https://t.me/learningnets

10-31 Initial Boot of an Autodeployed ESXi Host (4)



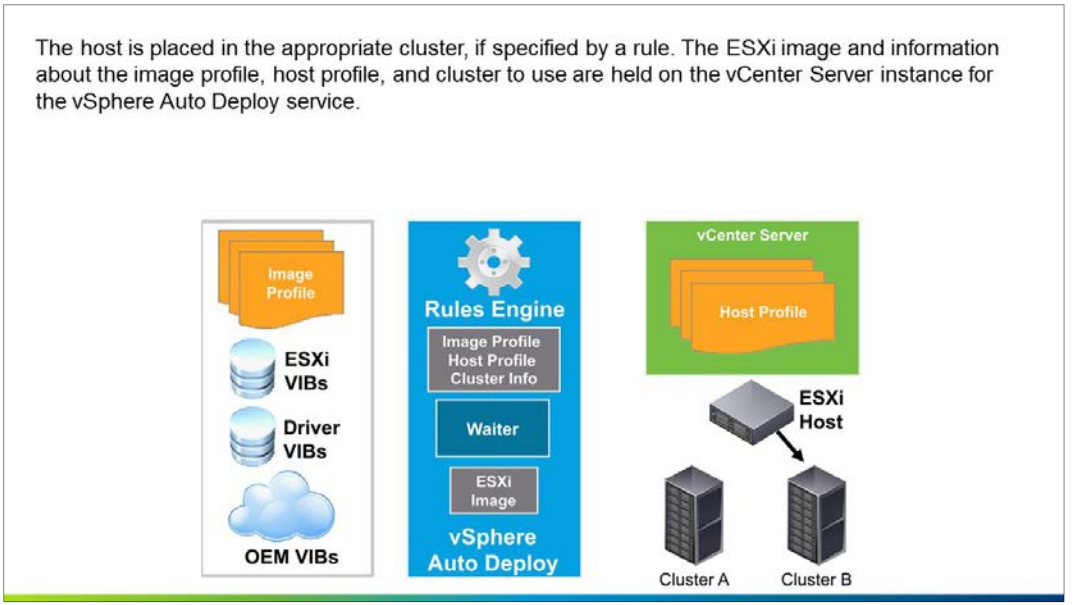
The vSphere Auto Deploy service streams VIBs specified in the image profile to the host.

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

10-32 Initial Boot of an Autodeployed ESXi Host (5)



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 is 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.

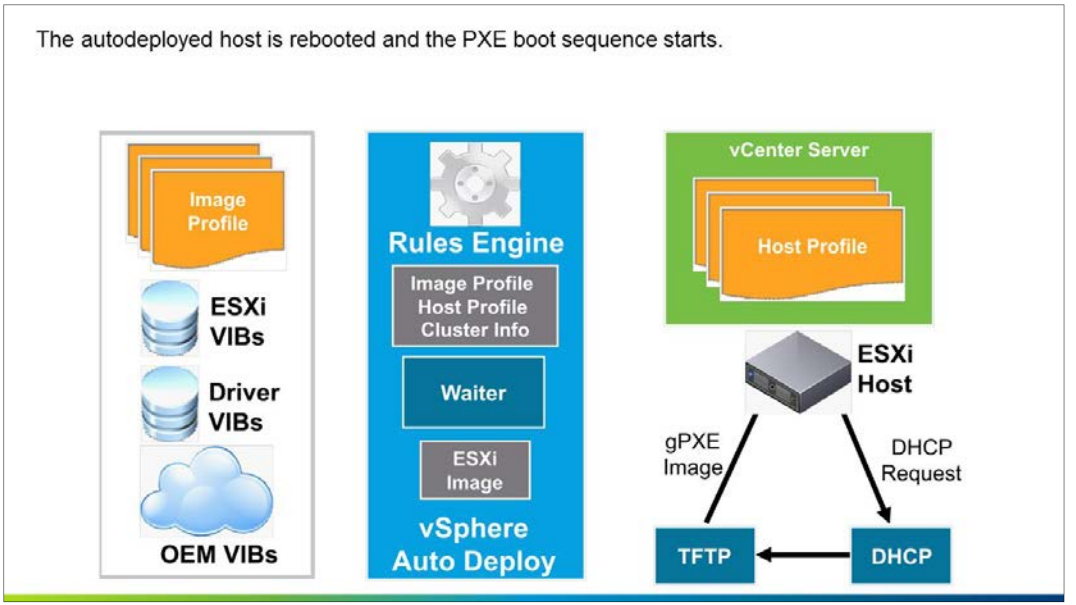
If the host is part of a fully automated vSphere DRS cluster, the cluster rebalances itself based on the new host, moving VMs onto the new host.

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

10-33 Subsequent Boot of a Stateless ESXi Host (1)

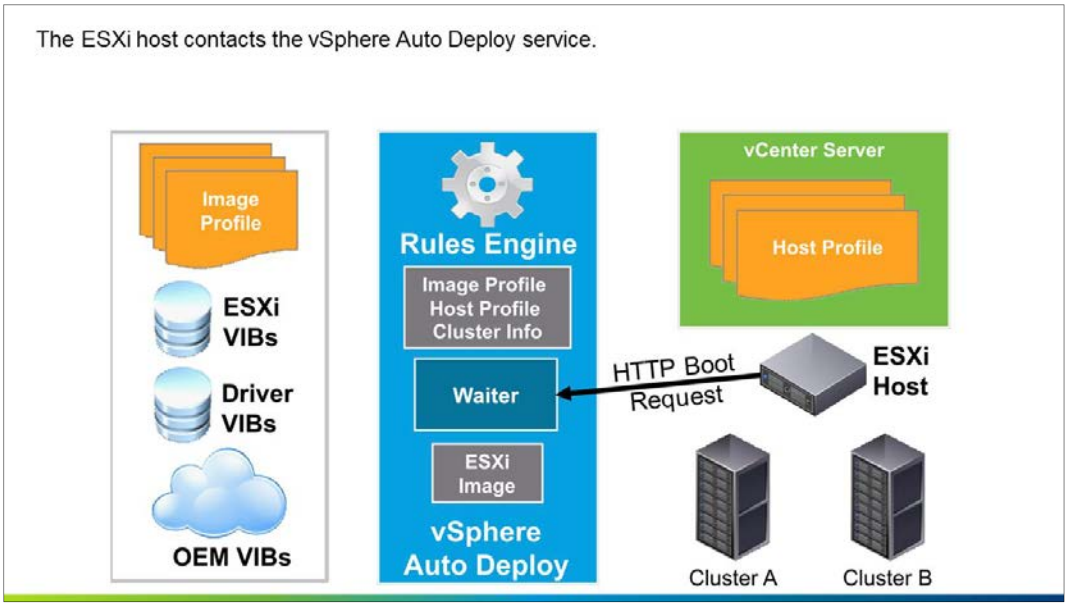


When an autodeployed ESXi host is rebooted, a slightly different sequence of events occurs:

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

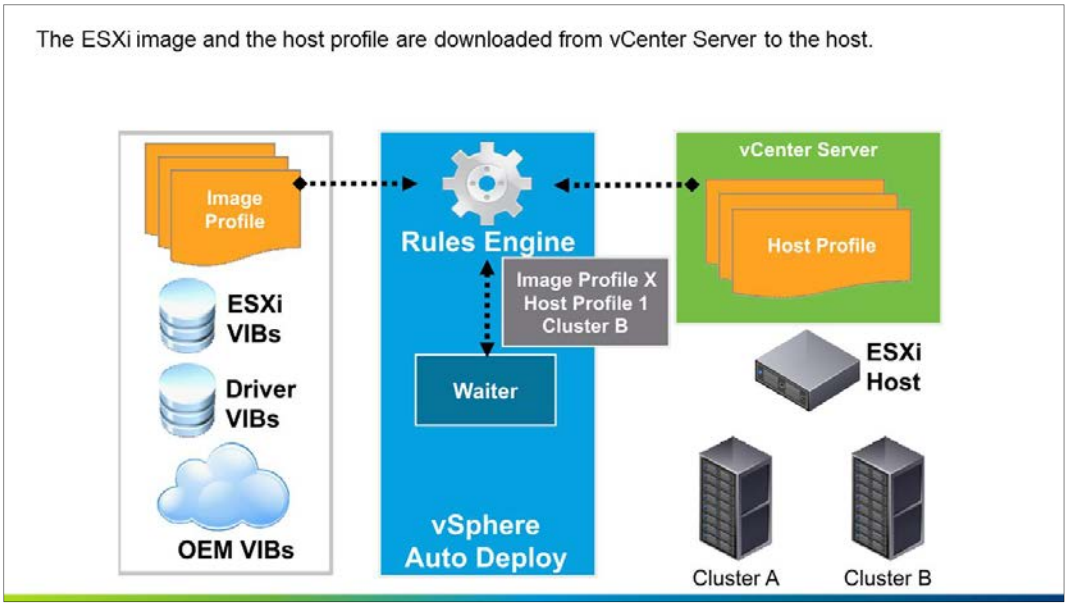
10-34 Subsequent Boot of a Stateless ESXi Host (2)



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

10-35 Subsequent Boot of a Stateless ESXi Host (3)



The subsequent boot sequence differs from the initial boot sequence.

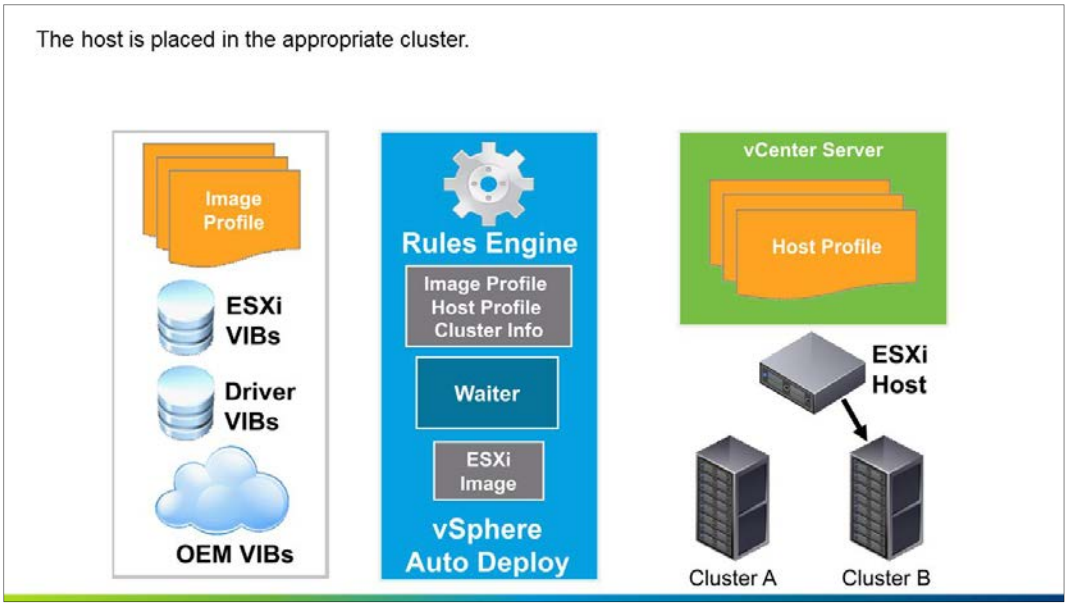
When an ESXi host is booted for the first time, vSphere Auto Deploy queries the rules engine for information about the host. The information about the host's image profile, host profile, and cluster is stored on the vSphere Auto Deploy service.

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

10-36 Subsequent Boot of a Stateless ESXi Host (4)



The ESXi host is placed in its assigned cluster on the vCenter Server instance.

https://t.me/learningnets

10-37 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

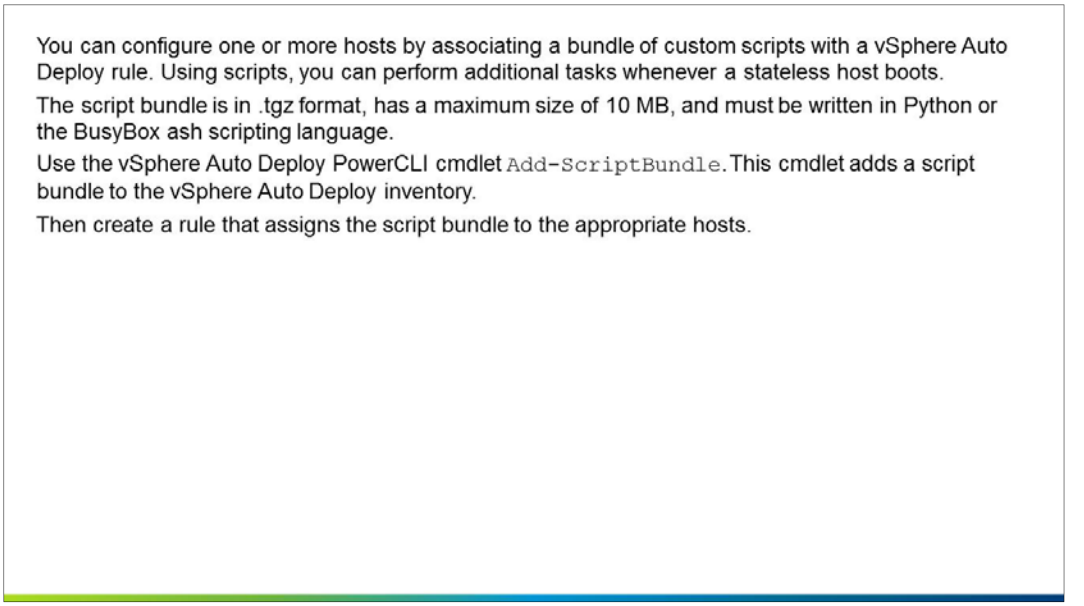
10-38 Activity: vSphere Auto Deploy Modes (2)

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

10-39 Running Custom Scripts on Stateless Hosts

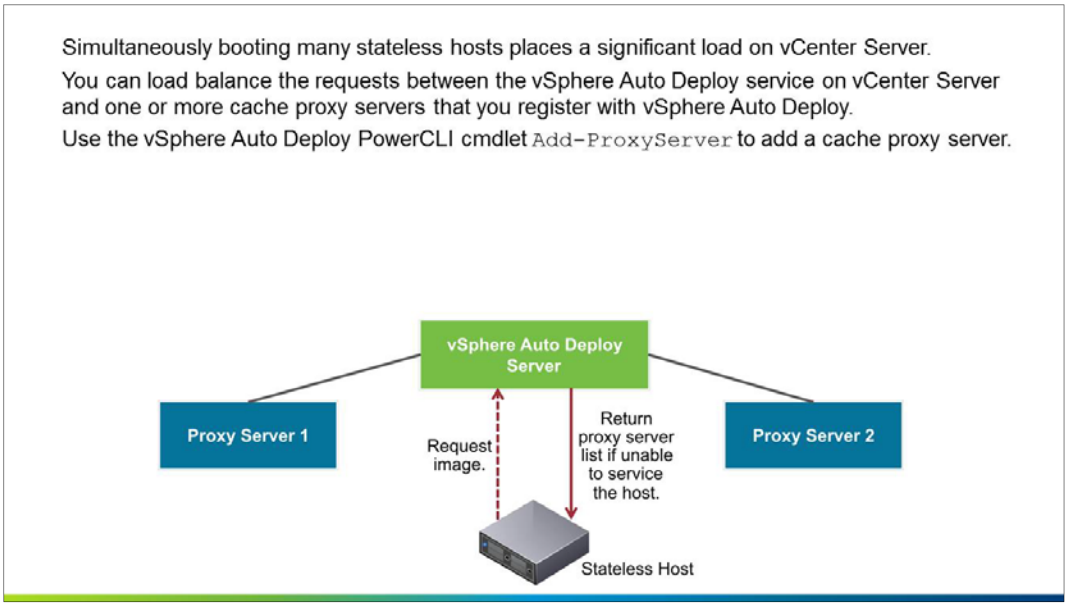


The scripts run in alphabetical order after the initial ESXi boot workflow of the host.

For more information about how to add a script bundle, see *Working with Script Bundles* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.esxi.install.doc/GUID-485DCF6A-4674-47E4-AD44-A1432774D792.html?hWord=N4IghgNiBcIM4GMBOBLADgFwAQCMCuAdgCYQCmIAvkA.

https://t.me/learningnets

10-40 Booting Many Stateless Hosts



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 service might take a long time to respond to a request.

Proxy servers (off-the-shelf proxy servers) help lessen the load on the vSphere Auto Deploy service that is on vCenter Server. If proxy servers front the vSphere Auto Deploy service, a list of proxy servers is sent to the stateless host when, in an extreme situation, vSphere Auto Deploy cannot serve the image. The stateless host tries to boot from the first proxy server on the list. If this proxy server is available, the image is served to the host. But if the first proxy server is overloaded or unavailable, 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 https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.esxi.install.doc/GUID-836883D3-24C8-46A1-BD66-0F4756A64E04.html.

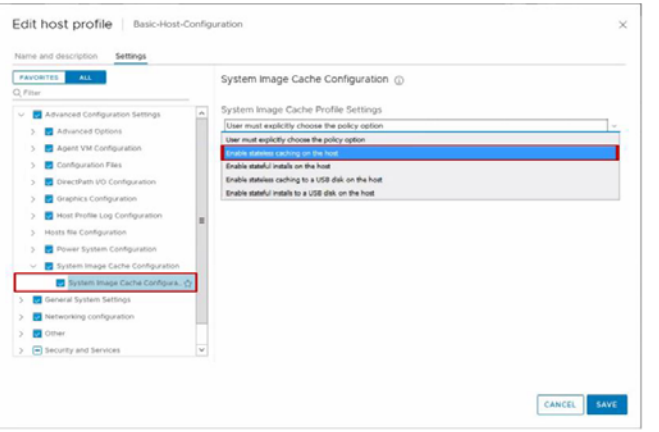
https://t.me/learningnets

10-41 Stateless Caching Host Profile Configuration

You use host profiles to configure autodeployed 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.

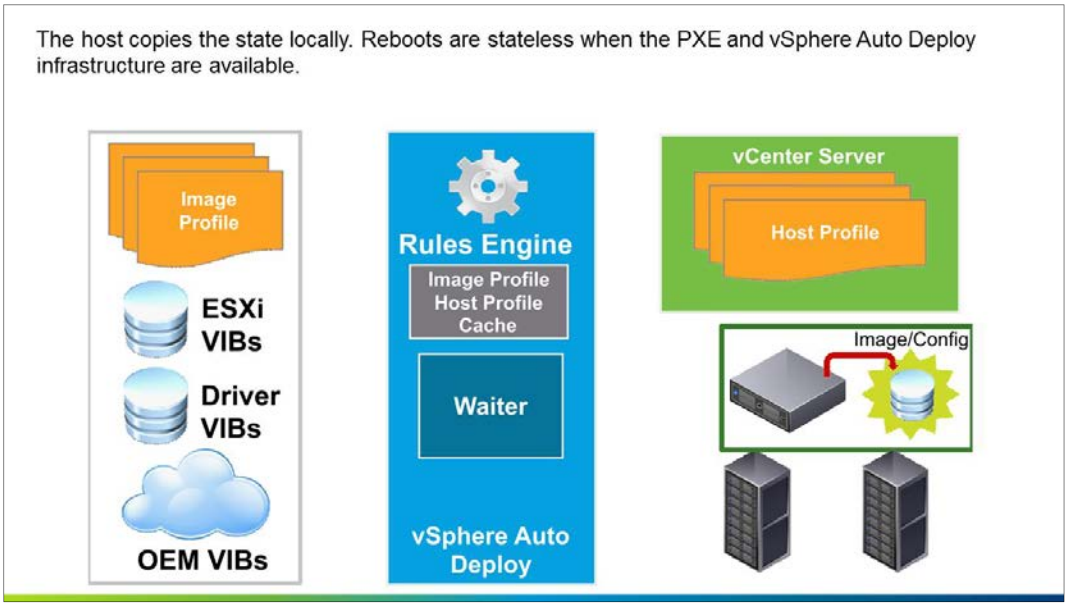
The host runs stateless under normal operations.



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

10-42 Subsequent Boot of a Stateless Caching Host



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 VMs 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 VMs to each host.

https://t.me/learningnets

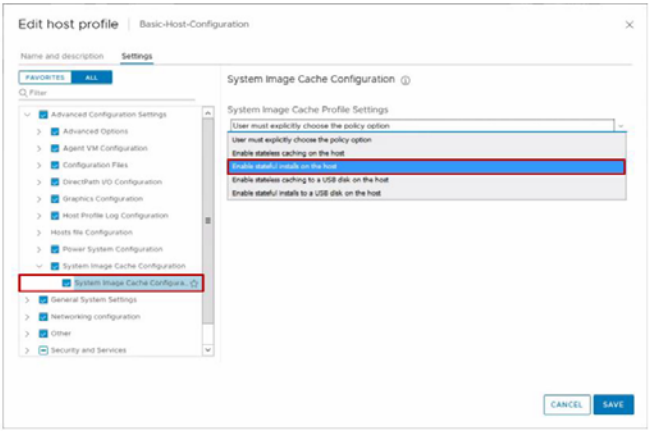
- If the vCenter Server system is not available, vSphere DRS does not work. The vSphere Auto Deploy service cannot add hosts to vCenter Server. You can connect to each ESXi host by using VMware Host Client and add VMs to each host.
- If you change your setup while connectivity is lost, the changes are lost when the connection to the vSphere Auto Deploy service is restored.

https://t.me/learningnets

10-43 Stateful Installation Host Profile Configuration

To configure a stateful installation host profile:

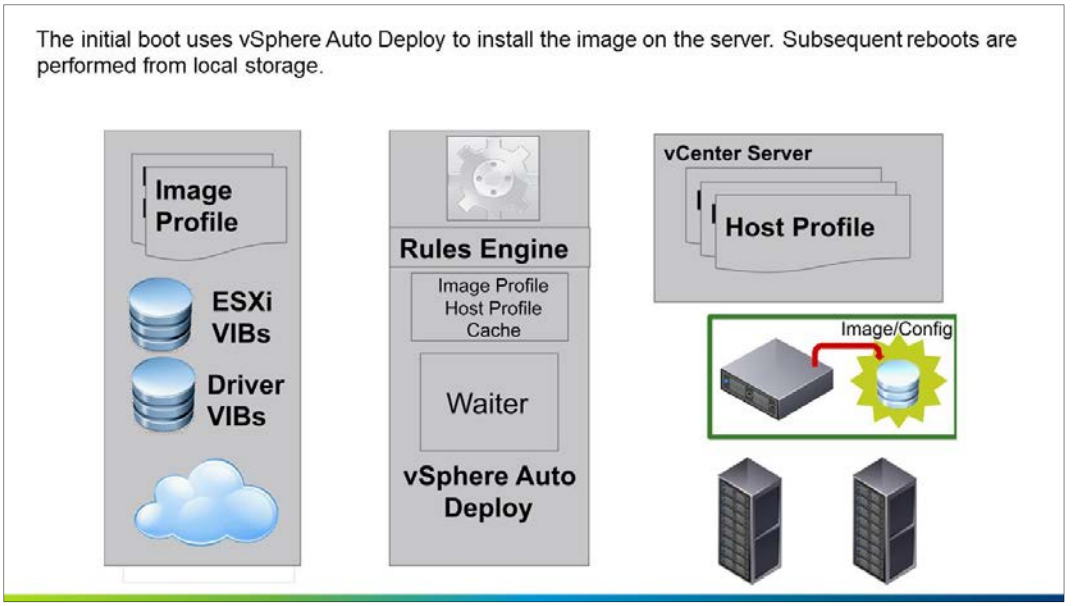
1. Create the host profile with stateful installation configured.
2. Use vSphere Auto Deploy to boot the ESXi host. The ESXi image is saved to disk, and the host profile is applied.
3. Direct future boots to the local disk. No future attempts are made to contact the PXE server or the vSphere Auto Deploy service.



You configure stateful installation by 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

10-44 Subsequent Boot of a Stateful Installation Host

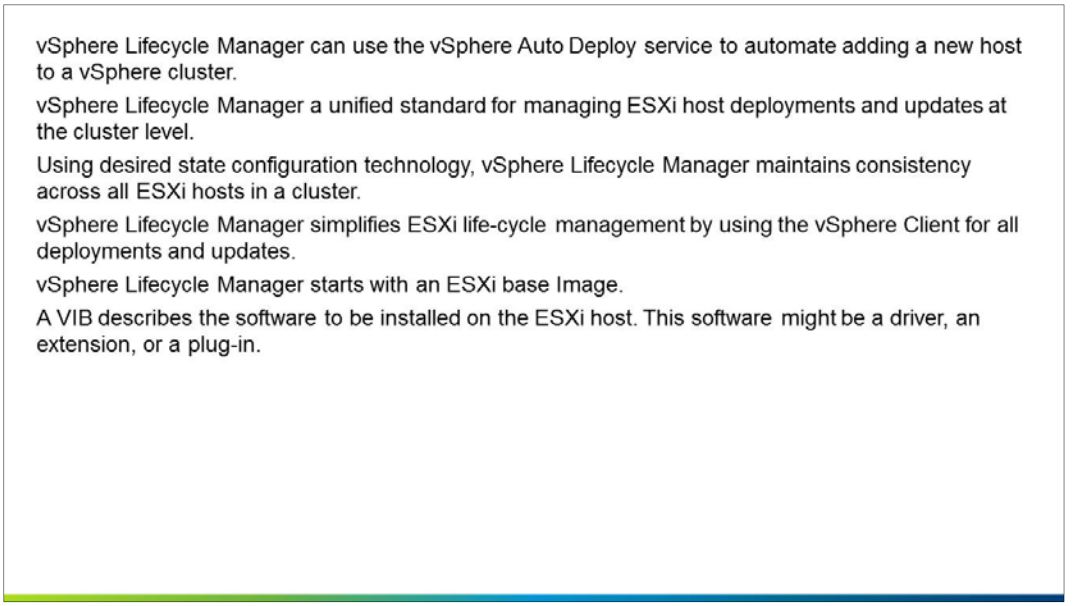


With stateful installation, the vSphere Auto Deploy service is used to provision new ESXi hosts. The first time that the host boots, it uses the PXE host, the DHCP server, and the vSphere Auto Deploy service, like a stateless host does. All subsequent reboots use the image that is saved to the local 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

10-45 Stateful Installation Using vSphere Lifecycle Manager



vSphere Lifecycle Manager introduces a new architecture to coordinate the validation and remediation of vSphere clusters. It consults the Hardware Compatibility Lists to find the latest releases and the hardware that is compatible with each release.

Depots that are compatible with vSphere Update Manager are used to publish ESXi releases. These depots can be hosted online or be provided as local files. You can use VMware predefined depots, or you can add custom depots to access base images, add-ons, and components.

The vSphere Lifecycle Manager components that are in ESXi hosts and vCenter Server Appliance instances are hosted in the vSphere Update Manager vAPI endpoint. During compliance check and remediation, ESXi Health Perspectives queries ESXi health so that vSphere Lifecycle Manager can safely update the hosts.

https://t.me/learningnets

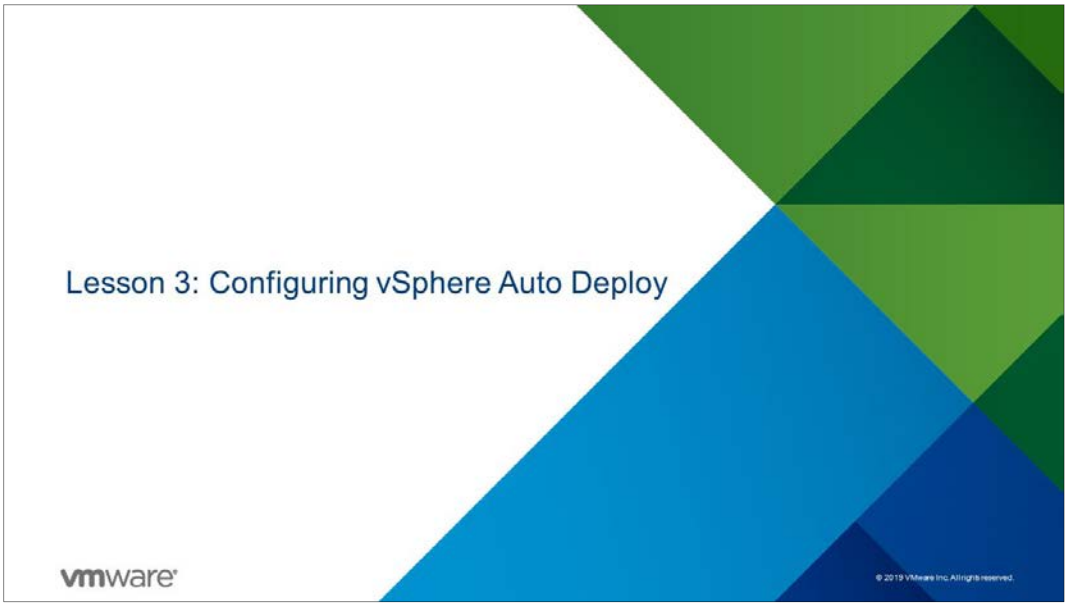
10-46 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the purpose of vSphere Auto Deploy
- Describe how an autodeployed ESXi host boots

https://t.me/learningnets

10-47 Lesson 3: Configuring vSphere Auto Deploy



https://t.me/learningnets

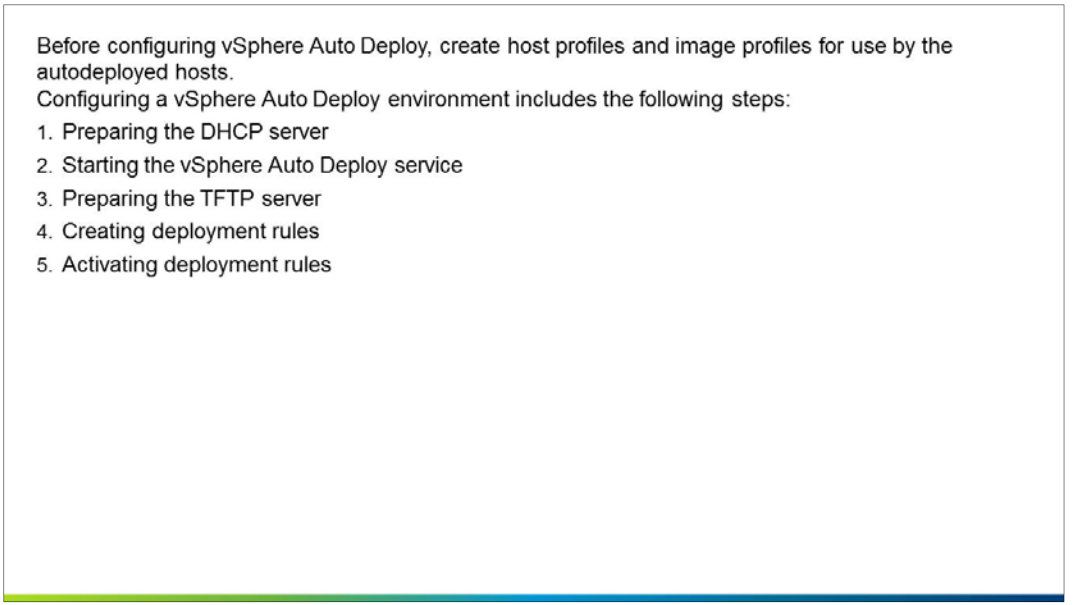
10-48 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Configure a vSphere Auto Deploy environment
- Use vSphere Auto Deploy to deploy ESXi hosts

https://t.me/learningnets

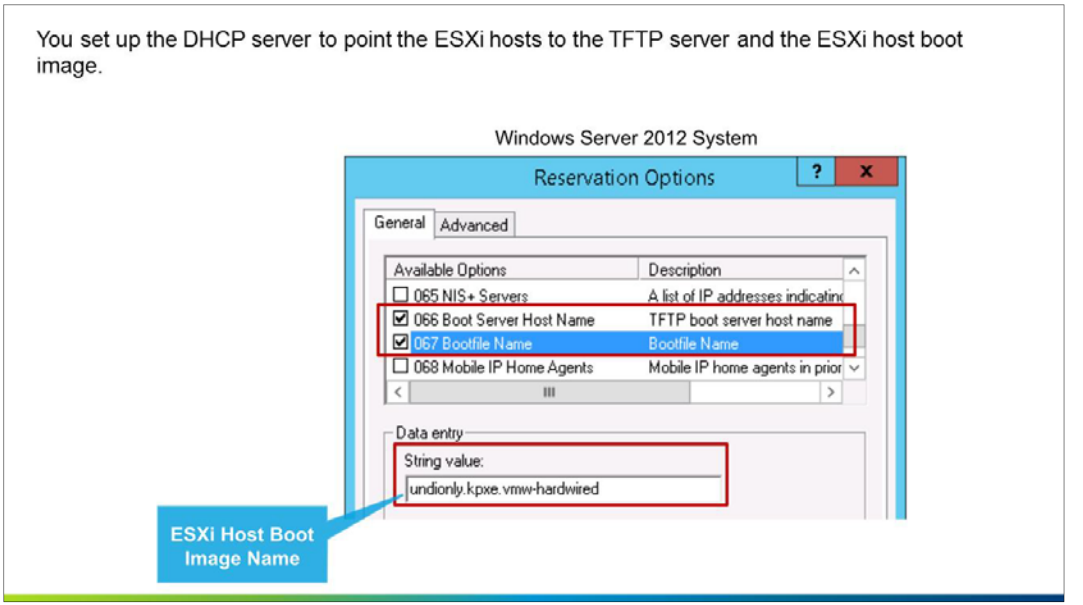
10-49 Configuring a vSphere Auto Deploy Environment



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

10-50 Step 1: Preparing the DHCP Server



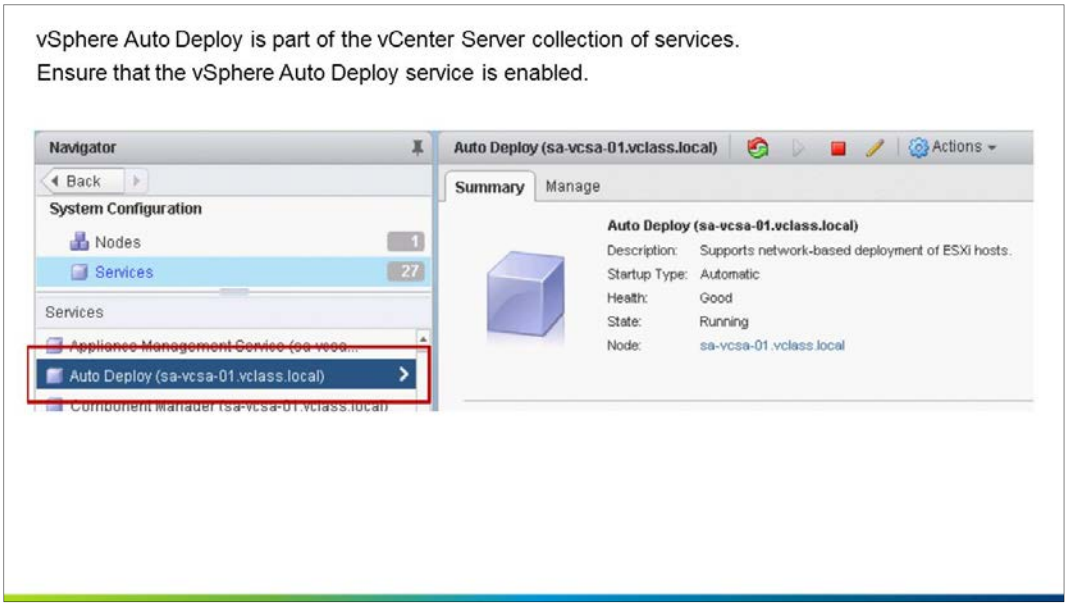
You must set up the DHCP server to serve each target ESXi host with an iPXE binary. In this example, the DHCP server is included with Windows Server 2012.

The iPXE binary file **`undionly.kpxe.vmw-hardwired`** 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

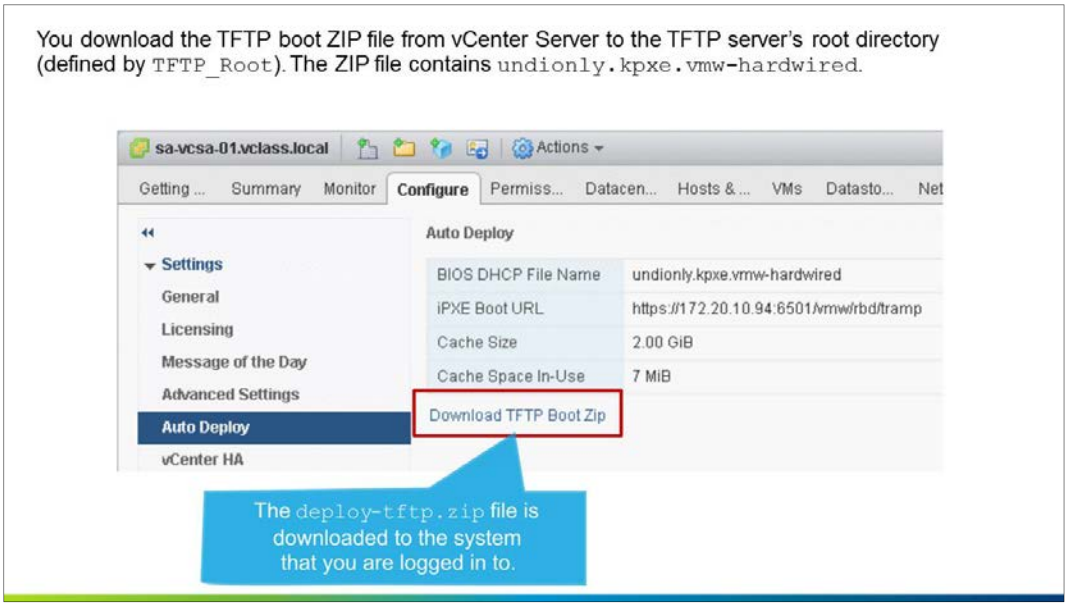
10-51 Step 2: Starting the vSphere Auto Deploy Service



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

10-52 Step 3: Preparing the TFTP Server



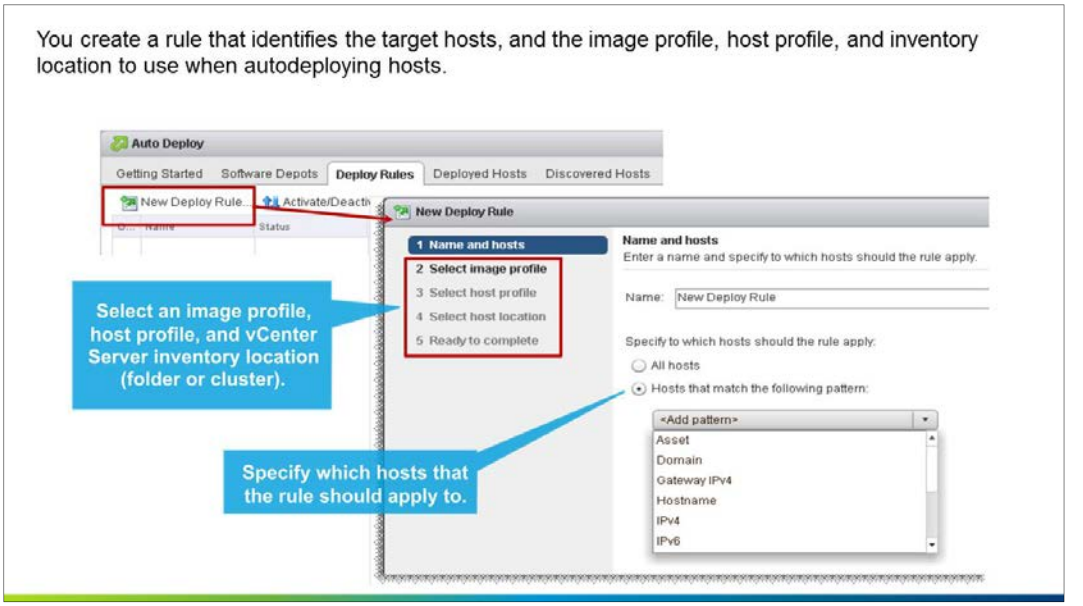
After you prepare the DHCP server, you must start the Auto Deploy service on vCenter Server and configure the TFTP server. You must download a TFTP boot ZIP file from your vSphere Auto Deploy service. The TFTP server serves the boot images that vSphere Auto Deploy provides.

When you click **Download TFTP Boot Zip**, the `deploy-tftp.zip` file is downloaded to the system on which the vSphere 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 ttps://docs.vmware.com/en/VMwarevSphere/6.7/vsphere-esxi-67-installation-setup-guide.pdf.

https://t.me/learningnets

10-53 Step 4: Creating Deployment Rules



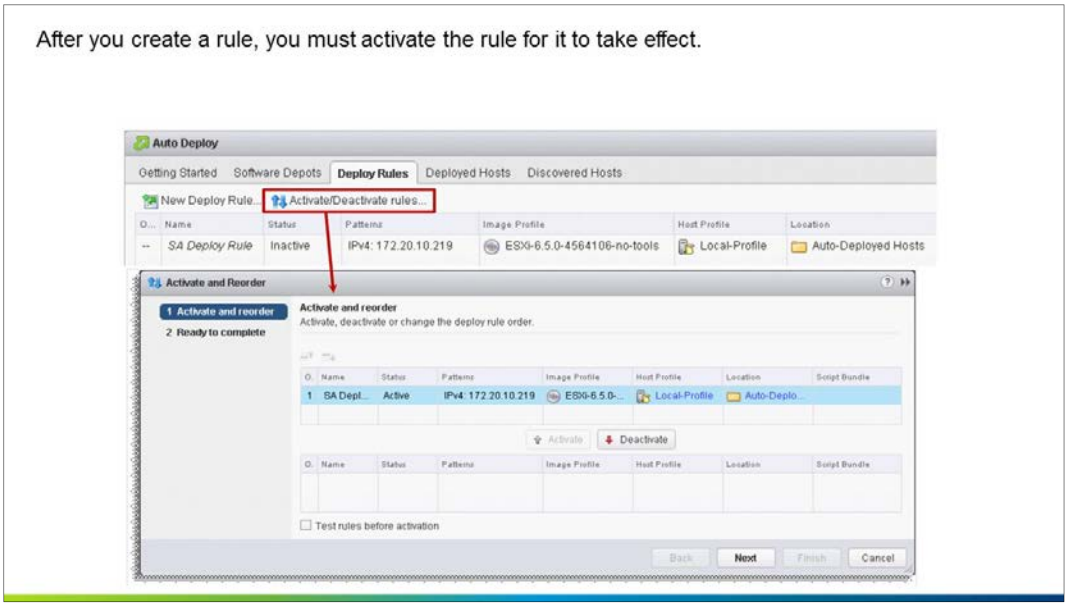
Rules can assign image profiles and host profiles to a set of hosts, or specify the location (folder or cluster) of a host on the target vCenter Server system. A rule can identify target hosts by boot MAC address, SMBIOS information, BIOS UUID, vendor, model, or fixed DHCP IP address. Usually, rules apply to multiple hosts.

After you create a rule, you must add it to a rule set. The only supported rule sets are active and working.

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 the vSphere Client, you can edit a rule if it is in the inactive state.

https://t.me/learningnets

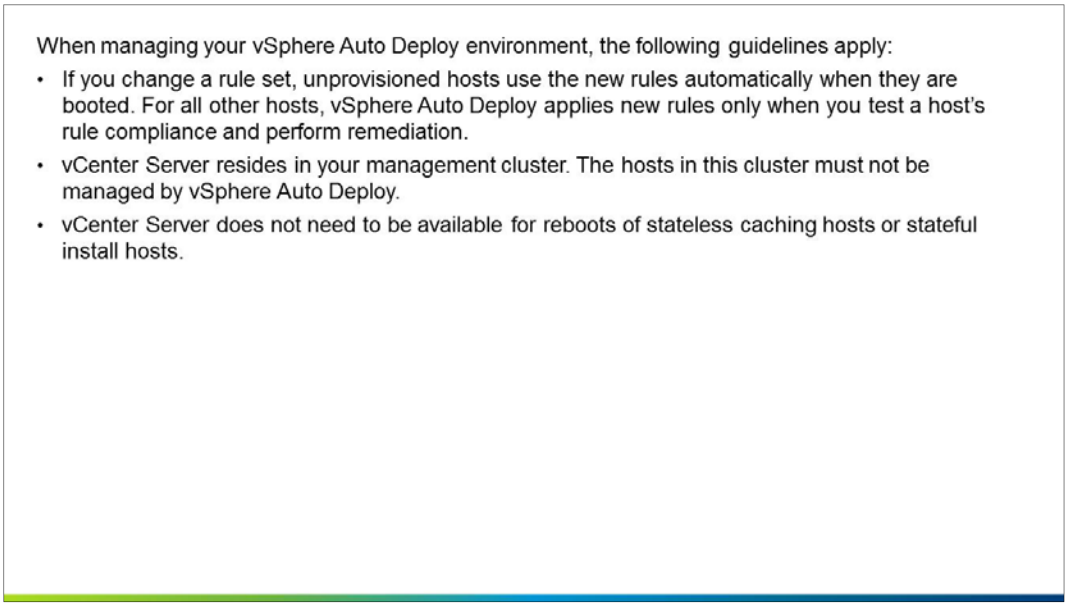
10-54 Step 5: Activating Deployment Rules



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

10-55 Managing the vSphere Auto Deploy Environment



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 cannot boot or reboot hosts. Installing vCenter Server in a VM and placing the VM in a vSphere HA cluster is recommended for keeping the VM available.

For information about how to Test and Repair Rule Compliance see https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.esxi.upgrade.doc/GUID-4915B724-500E-4FB3-BAC2-0EA46CFBD7EE.html.

https://t.me/learningnets

10-56 Using vSphere Auto Deploy with vSphere Update Manager

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:

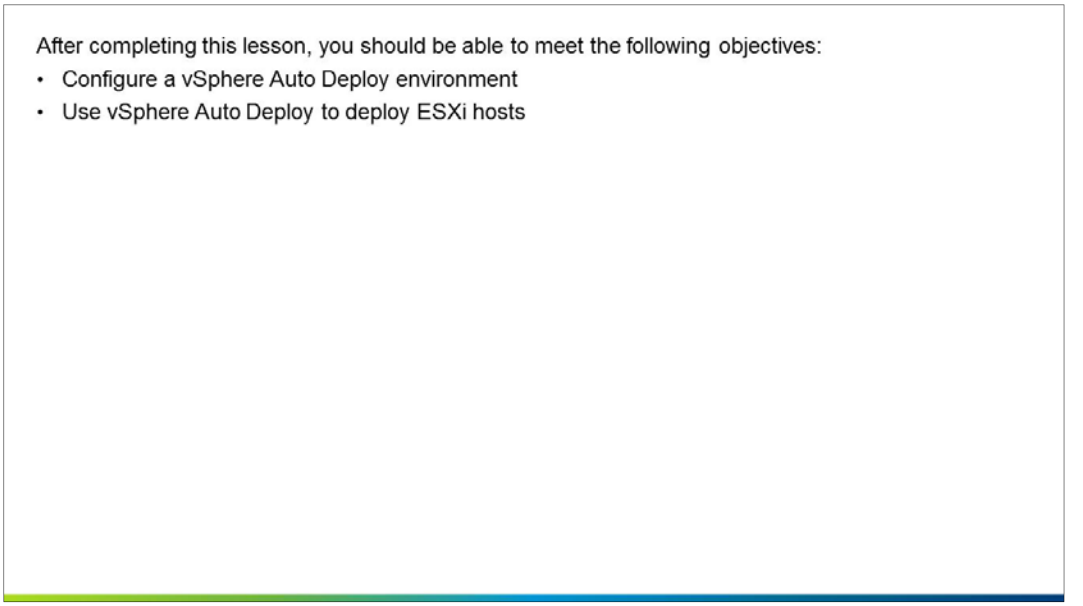
1. Manually update the image that vSphere Auto Deploy uses with patches.
   If rebooting is possible, rebooting is all that is required to update the host.
2. If a reboot cannot be performed, create a baseline in vSphere Update Manager and remediate the host.

vSphere Update Manager can be used to upgrade hosts that boot using vSphere Auto Deploy. Image profiles must be patched and modified using the vSphere 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

10-57 Review of Learner Objectives



https://t.me/learningnets

10-58 Key Points

- vSphere ESXi Image Builder is used to create ESXi installation images that include a set of updates, patches, and drivers.
- You can use vSphere Auto Deploy to deploy large numbers of ESXi hosts quickly and easily.
- With the Host Profiles feature, you can 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.

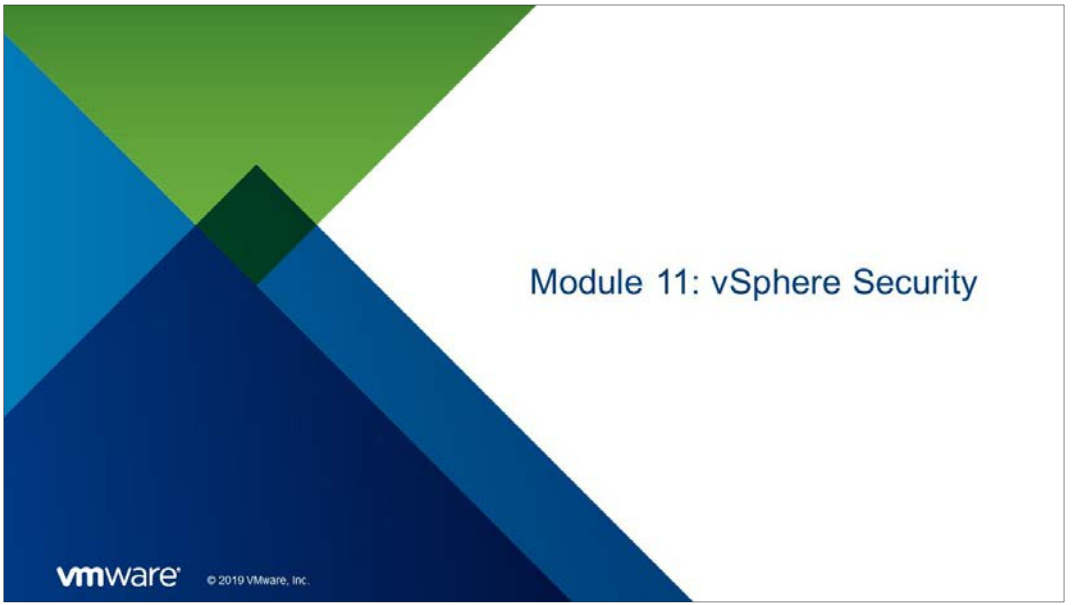
Questions?

https://t.me/learningnets

https://t.me/learningnets

Module 11

Appendix: vSphere Security



https://t.me/learningnets

11-2 Importance

By default, vSphere components are secured by certificates, authorization, limited access, and compliance to security standards.

You can further harden your vSphere environment against security threats by controlling settings for vCenter Server systems, ESXi hosts, VMs, and the vSphere network.

You can encrypt VM files and disks to protect confidential customer data.

https://t.me/learningnets

11-3 Appendix Lessons

1. Configuring ESXi Host Access and Authentication
2. Securing vSphere
3. Security Standards and Protocols
4. Securing VMs
5. VM Encryption

https://t.me/learningnets

11-4 Lesson 1: Configuring ESXi Host Access and Authentication

https://t.me/learningnets

11-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

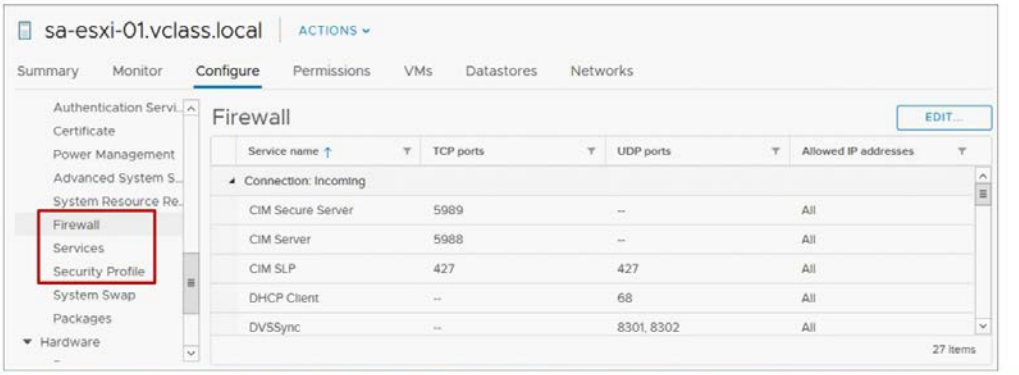
- Configure the ESXi firewall by enabling and disabling services
- Enable and disable lockdown mode on an ESXi host
- Configure user logins to authenticate with directory services

https://t.me/learningnets

11-6 About ESXi Host Security Properties

The ESXi hypervisor architecture has many built-in security features, such as CPU isolation, memory isolation, firewall, and device isolation.

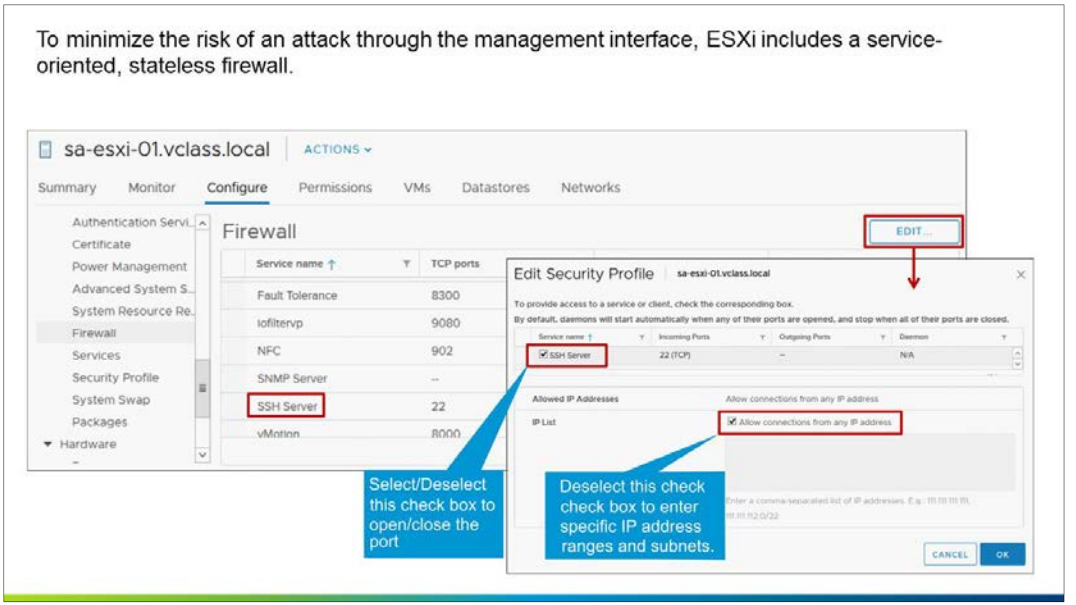
You can configure additional features such as a firewall, disable services such as SSH, and configure lockdown mode.



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 vSphere ESXi Shell can further contribute to a more secure environment.

https://t.me/learningnets

11-7 Configuring the ESXi 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 allow 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 define the IP addresses or range of IP addresses that are allowed to access each service.

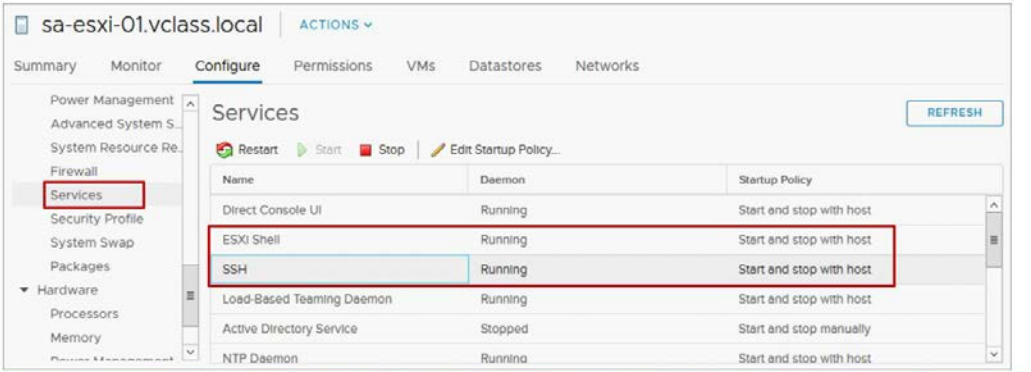
You can configure the ESXi firewall by using the vSphere Client or vSphere CLI. You configure firewall management at the command line by using the **`esxcli network firewall`** namespace.

https://t.me/learningnets

11-8 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, vSphere ESXi Shell and SSH services are started. These services are stopped by default.



https://t.me/learningnets

11-9 About Lockdown Mode (1)

To increase the security of your ESXi hosts, you can put your hosts in lockdown mode.

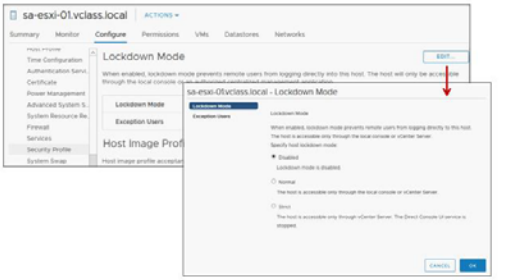
In lockdown mode, some services are disabled and some services are accessible only to certain users.

The available modes are:

- Normal
- Strict

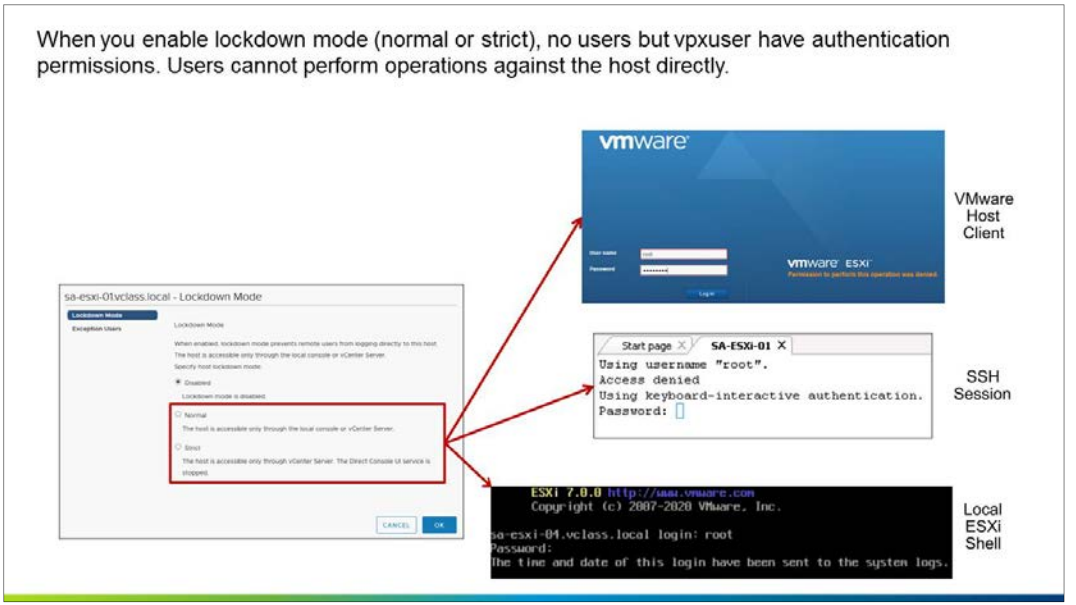
By default, lockdown mode is disabled.

You can enable lockdown mode with the vSphere Client when you create ESXi host or modify an existing host's security profile.



https://t.me/learningnets

11-10 About Lockdown Mode (2)



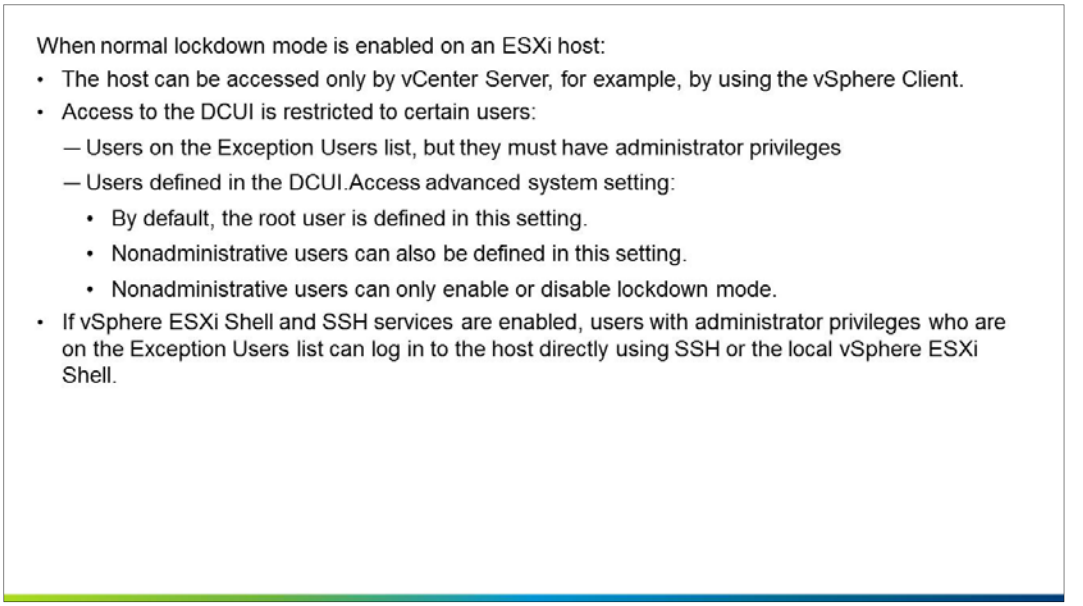
You can enable lockdown mode:

- When adding the host to the vCenter Server
- By using the Direct Console User Interface (DCUI)
- By using a host's security profile in the vSphere Client

In lockdown mode, all host operations must be performed through vCenter Server. The host cannot be accessed directly through the VMware Host Client.

https://t.me/learningnets

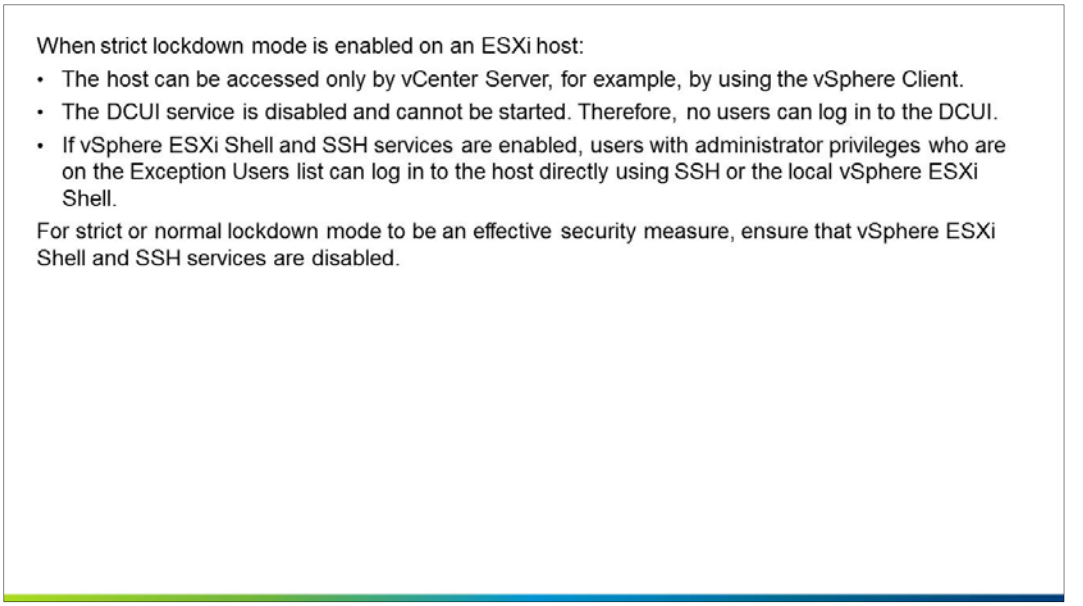
11-11 Normal Lockdown Mode Behavior



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 if the connection to vCenter Server is lost. These users do not require administrative privileges on the host. Nonadministrative DCUI.Access users can only disable or enable lockdown.

https://t.me/learningnets

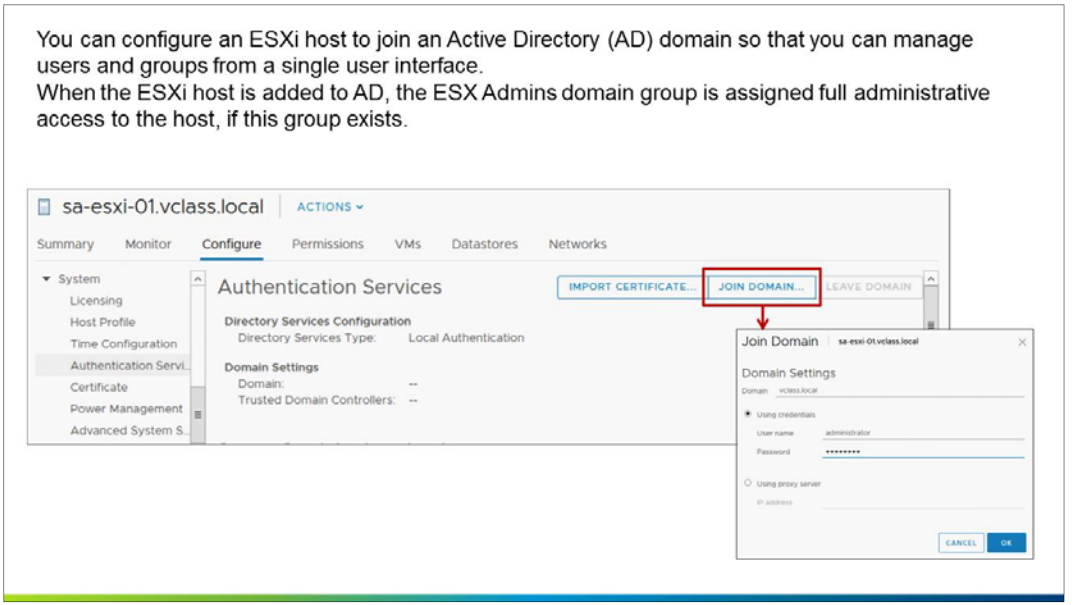
11-12 Strict Lockdown Mode Behavior



If vCenter Server is unavailable, the vSphere Client is also unavailable. Therefore, hosts that are in strict lockdown mode are not manageable.

https://t.me/learningnets

11-13 Integrating ESXi with Active Directory



Although day-to-day vSphere management operations are done while logged in to vCenter Server using the vSphere Client, the user sometimes requires direct access to the ESXi host, for example, when accessing local log files.

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. You can still define and manage local users host-by-host and configure them using the vSphere Client. You can use this approach either in place of or in addition to the vCenter Single Sign-On and AD integration.

Whenever 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.

https://t.me/learningnets

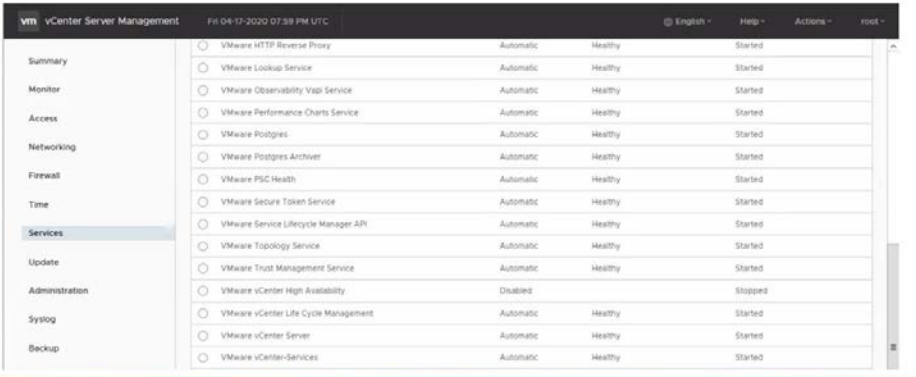
The only user that is defined on the ESXi host is root. The root password is not mapped to an AD account. The initial root password is typically set during ESXi installation, but it can be changed later through VMware Host Client, vSphere ESXi Shell, or the DCUI.

You can create other domain groups if you only want to grant administrative rights to a subset of hosts. You give that domain group administrator rights on the specific hosts and add the domain users to the group.

https://t.me/learningnets

11-14 VMware vSphere Authentication Proxy Service

You can automate the process of adding ESXi hosts to an AD domain by using the vSphere Authentication Proxy service. This service provides support for joining autodeployed ESXi hosts to an AD domain by using an account with delegated privileges.



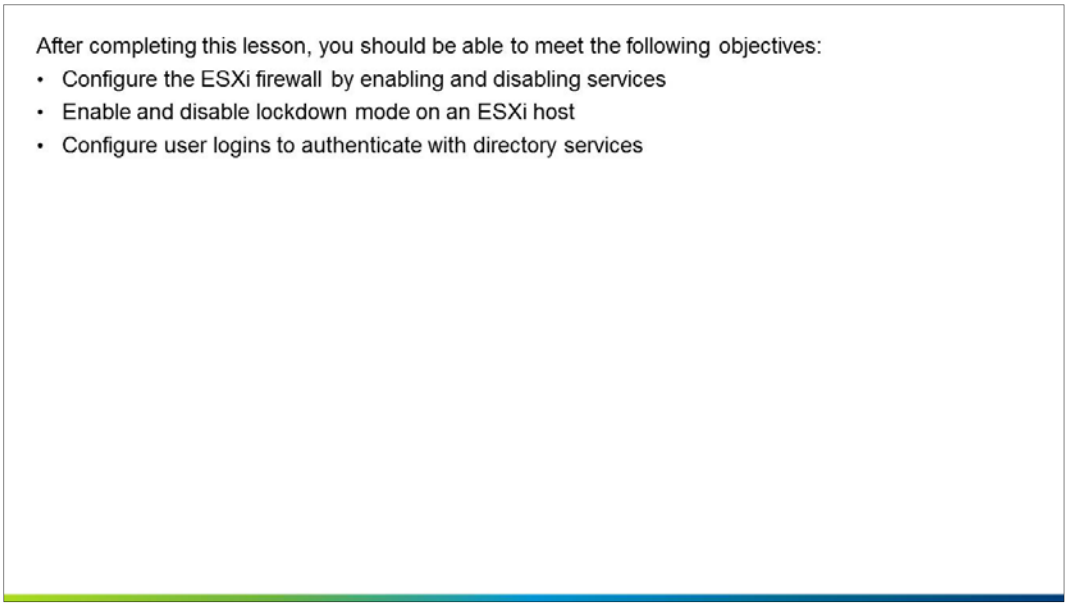
You can use vSphere Authentication Proxy to join the host to an AD domain, instead of adding the host directly to the AD domain. A chain of trust exists between vSphere Authentication Proxy and the host, which allows vSphere Authentication Proxy to join the host to the AD domain.

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, provisioned by VMware CA or third-party certificates, the process works seamlessly if you follow the instructions for using custom certificates with vSphere Auto Deploy.

The 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

11-15 Review of Learner Objectives



https://t.me/learningnets

https://t.me/learningnets

11-17 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the importance of the vSphere Security Configuration Guide
- Identify recommendations for vCenter Server security
- Recognize strategies for securing the vSphere management network
- Identify recommendations for ESXi host security
- Recognize guidelines for general VM and guest operating system protection

https://t.me/learningnets

11-18 About Securing vSphere

Securing vSphere involves aspects of security for the vCenter Server system, ESXi hosts, VMs, and the vSphere network. The VMware website provides several security resources.

Topic	Resource
vSphere security	https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-esxi-vcenter-server-70-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 at http://www.vmware.com/security/advisories.html.

https://t.me/learningnets

11-19 About the vSphere Security Configuration Guide

The vSphere Security Configuration Guide helps you to configure your vSphere environment according to security best practices.

Over time, vSphere has become more secure by default.

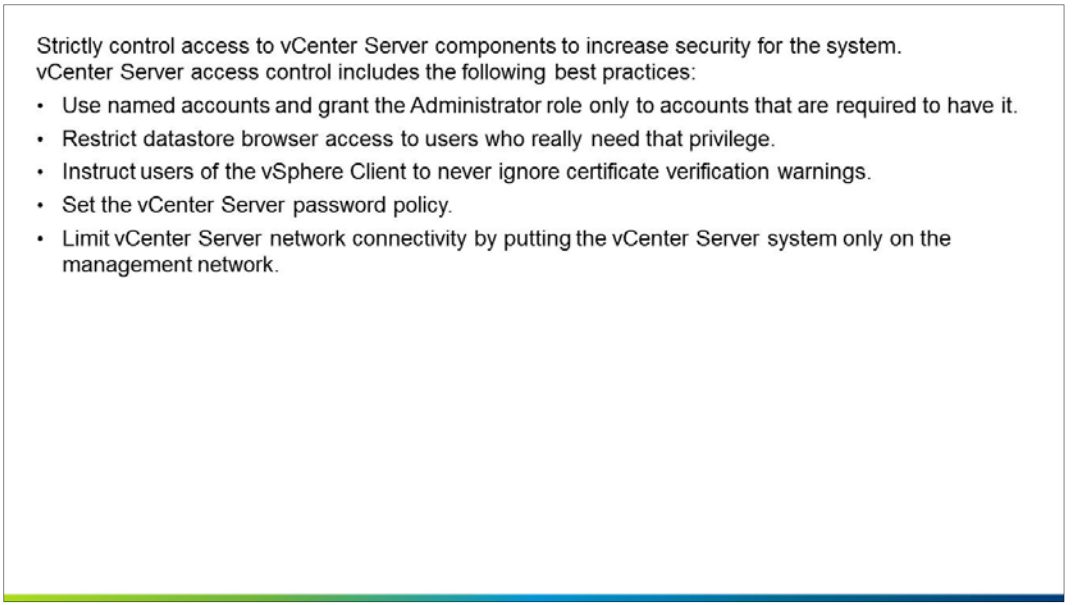
As a result, this guide is less about hardening and more about ensuring that best practices are followed.

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

11-20 Controlling Access to vCenter Server Systems



Starting with vSphere 6.0, the local administrator no longer has full administrative rights to vCenter Server by default. Instead, you 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 vSphere Client.

Verify vSphere Client certificates. Without certificate verification, the user might be subject to man-in-the-middle attacks.

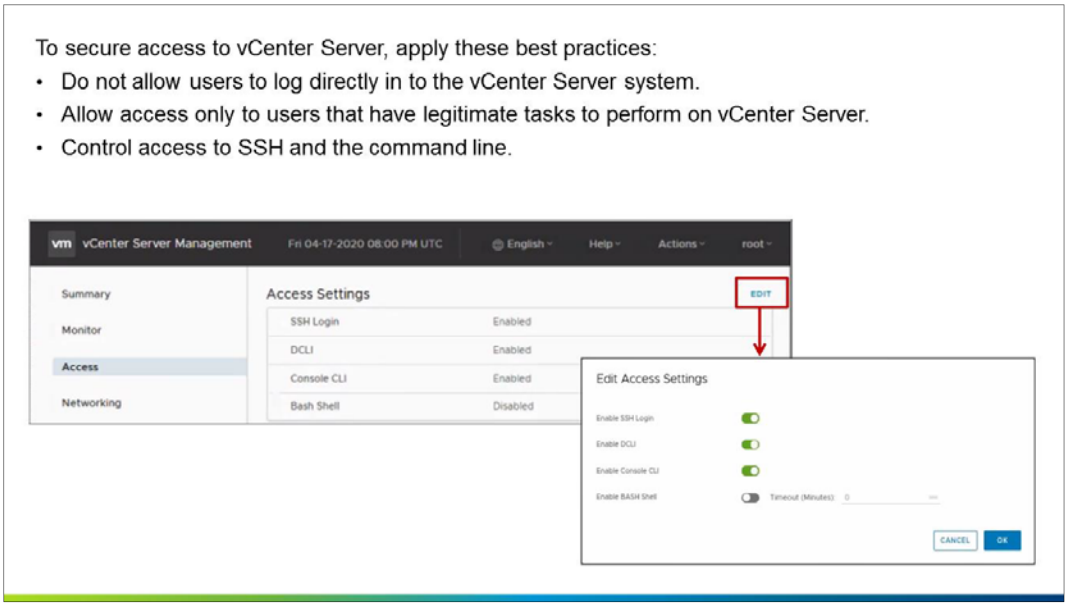
By default, vCenter Server automatically changes the vpxuser password every 30 days. You can change this frequency in the vSphere Client.

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/7.0/vsphere-esxi-vcenter-server-70-security-guide.pdf.

https://t.me/learningnets

11-21 vCenter Server Access Control



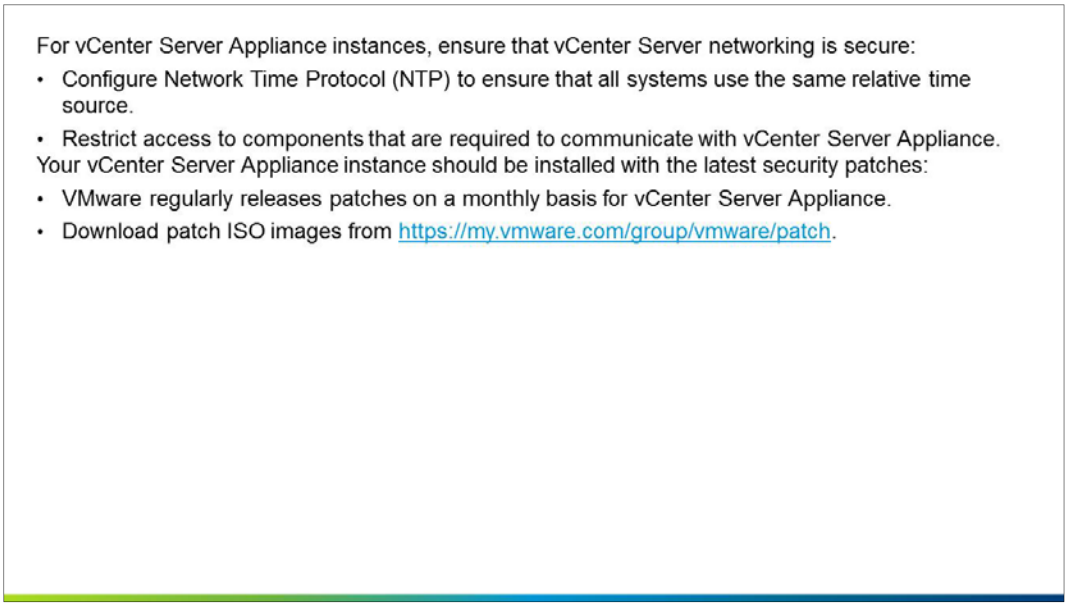
Users who are directly logged in to the vCenter Server instance can cause harm, either intentionally or unintentionally, by altering settings and modifying processes. Those users also have potential access to vCenter Server credentials, such as the SSL certificate.

Consider auditing login events regularly to ensure that only authorized users are accessing vCenter Server.

Use the 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

11-22 Securing vCenter Server Appliance



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 (NTP) 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.

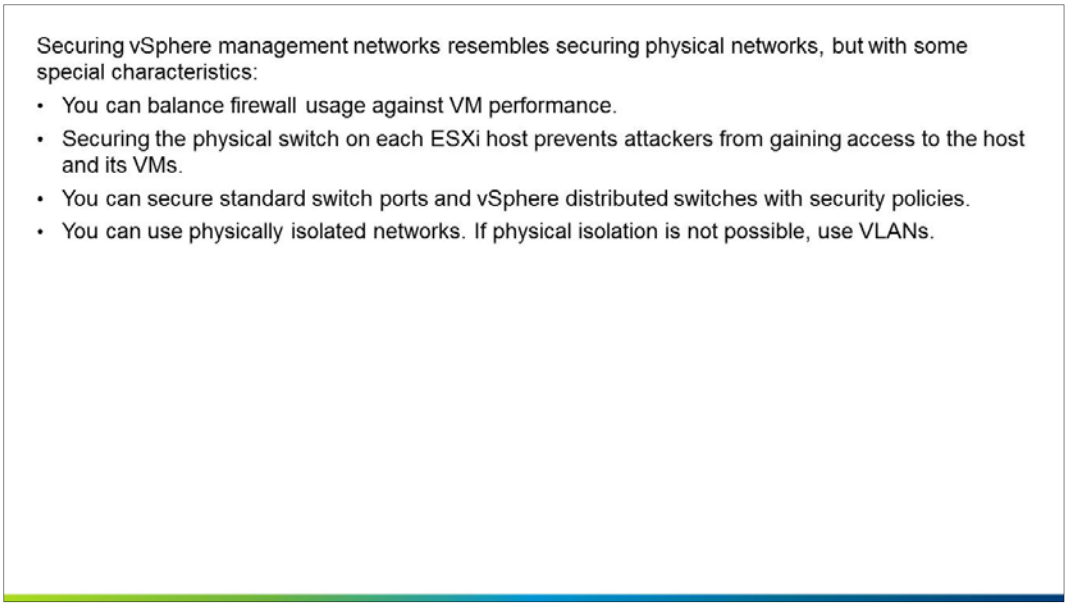
VMware makes patches available on a monthly basis. These patches might be related to third-party products in the platform, core product functionality, or both. The patches can only be applied between major releases of vCenter Server Appliance. For example, patches for the initial release of vCenter Server Appliance 6.7 are not applicable to vCenter Server Appliance 6.7 Update 1 because patches that are previously made available are included with the Update 1 release.

https://t.me/learningnets

For more information about how to patch vCenter Server Appliance, see vCenter Server Upgrade at https://docs.vmware.com/en/VMware-vSphere/7.0/vsphere-vcenter-server-70-upgrade-guide.pdf.

https://t.me/learningnets

11-23 Securing vSphere Management Networks



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 with the following systems:

- ESXi hosts
- Other vCenter Server systems in the same vCenter Single Sign-On domain
- Infrastructure services such as DNS, AD, and NTP
- Other systems that run components that are essential to the functionality of the vCenter Server system

https://t.me/learningnets

For the best protection of your hosts, ensure that physical switch ports are configured with the Spanning Tree Protocol 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

11-24 Securing ESXi Hosts

To secure ESXi hosts, VMware constrains vSphere parameters and settings.
Consider these guidelines for maintaining ESXi host security:

- 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 the 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 vSphere ESXi Shell services disabled.
  - Enable lockdown mode to further increase the security of ESXi hosts.
- Use VMware CA 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 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 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.

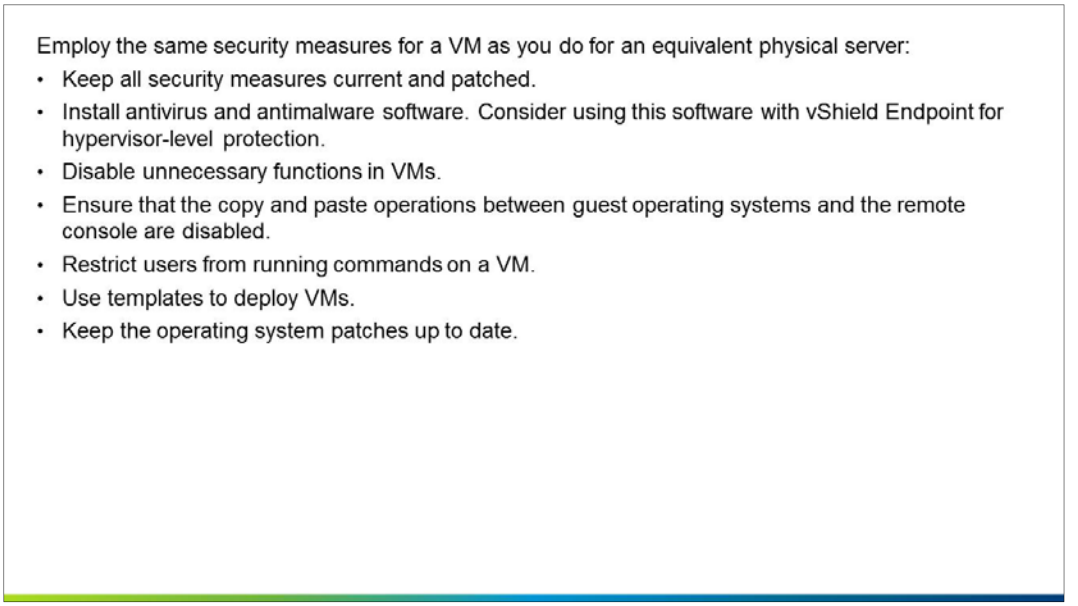
https://t.me/learningnets

You manage certificates from the 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/7.0/vsphere-esxi-vcenter-server-70-security-guide.pdf.

https://t.me/learningnets

11-25 VM Protection Overview



Using templates to deploy VMs removes the risk of misconfiguration. All VMs are created with a known baseline level of security.

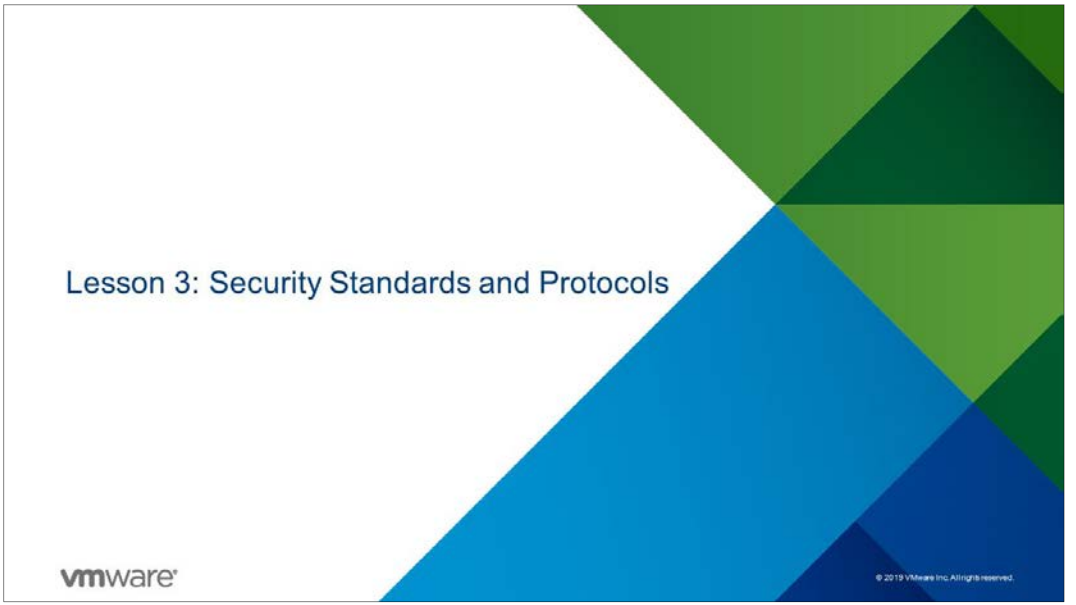
https://t.me/learningnets

11-26 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Explain the importance of the vSphere Security Configuration Guide
- Identify recommendations for vCenter Server security
- Recognize strategies for securing the vSphere management network
- Identify recommendations for ESXi host security
- Recognize guidelines for general VM and guest operating system protection

https://t.me/learningnets



https://t.me/learningnets

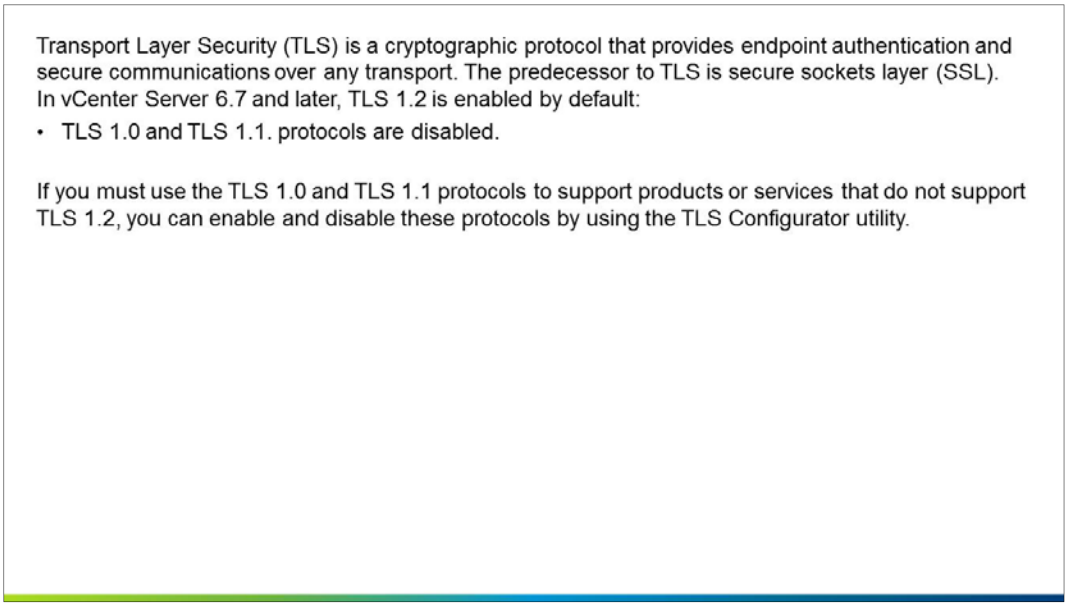
11-28 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Secure vCenter Server with TLS 1.2
- Use UEFI Secure Boot for ESXi hosts
- Plan for secure boot and TPM 2.0 support for an ESXi host
- Enable FIPS 140-2 for ESXi hosts and vCenter Server instances

https://t.me/learningnets

11-29 Securing vCenter Server with TLS 1.2



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.

For more information about the TLS Configurator utility, see *Managing TLS Protocol Configuration with the TLS Configurator Utility* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.security.doc/GUID-82028A21-8AB5-4E2E-90B8-A01D1FAD77B1.html.

https://t.me/learningnets

11-30 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

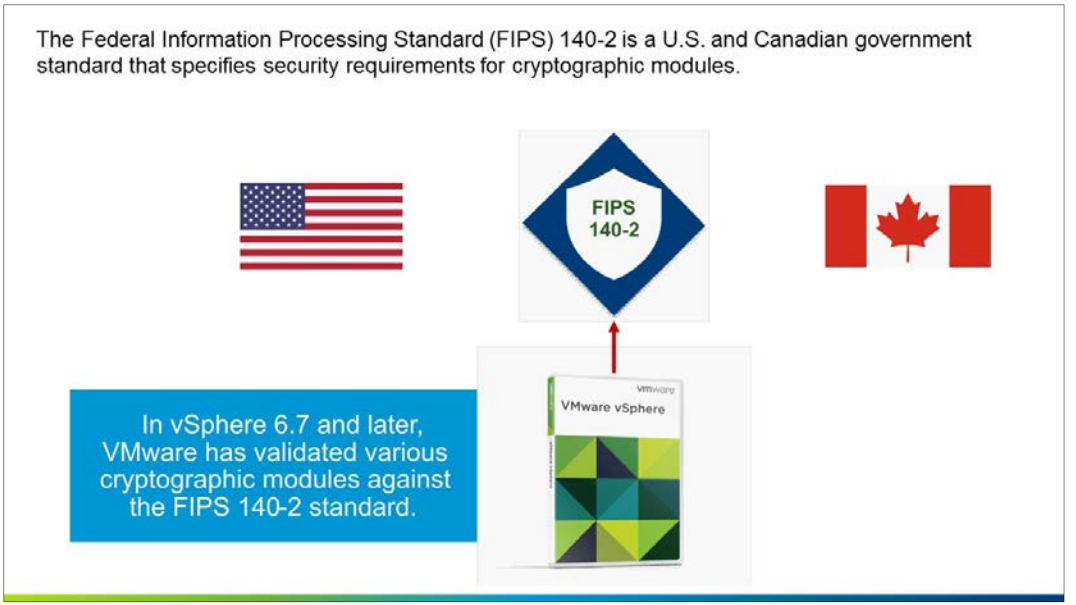
11-31 Activity: ESXi Secure Boot (2)

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

11-32 vSphere Compliance with FIPS 140-2 (1)

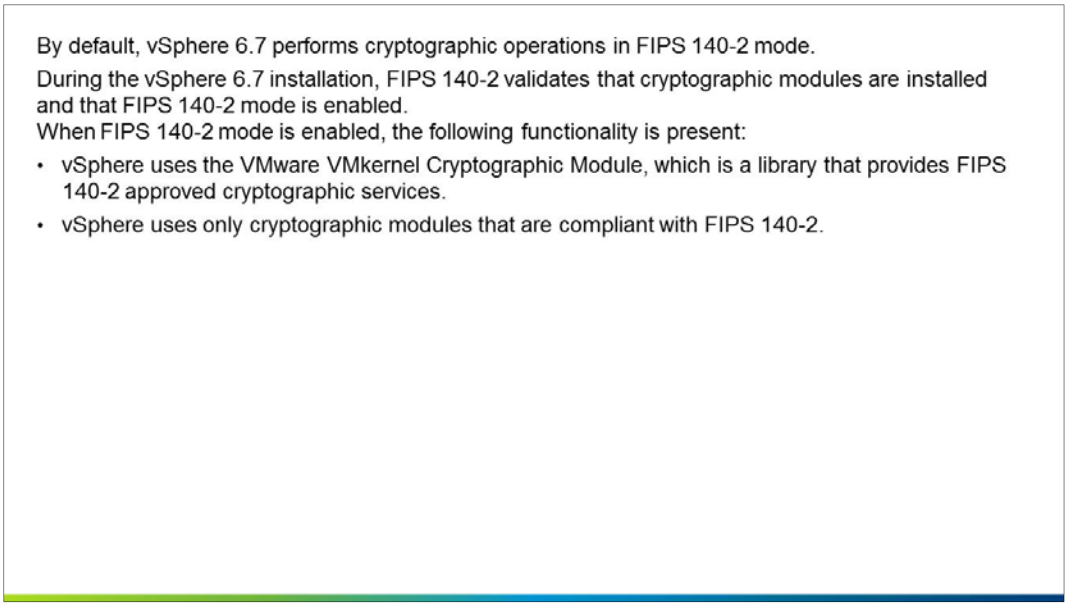


The Federal Information Processing Standards (FIPS) 140-2 standard specifies and validates the cryptographic and operational requirements for the modules within security systems that protect sensitive information. These modules use National Institute of Standards and Technology (NIST) approved security 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

11-33 vSphere Compliance with FIPS 140-2 (2)



FIPS 140-2 mode is not compatible with third-party tools that do not enable cryptographic modules.

https://t.me/learningnets

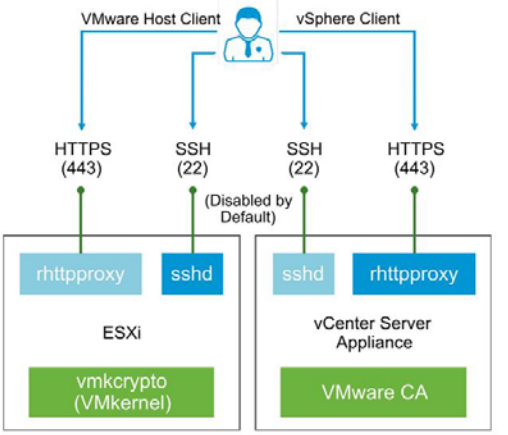
11-34 FIPS 140-2 Architecture

In vSphere 6.7 and later, 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:

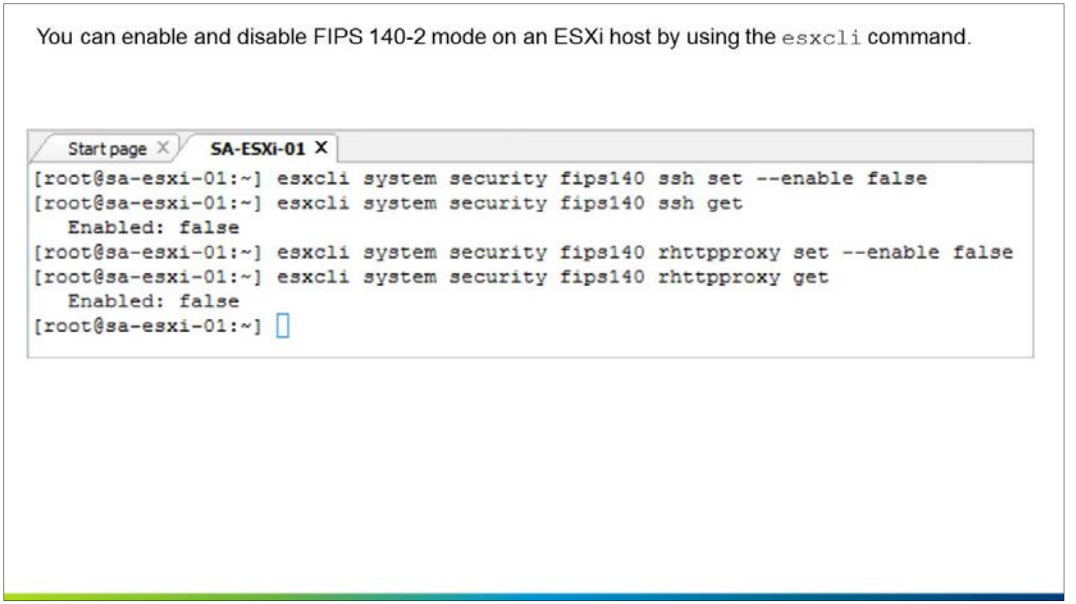
- The VMkernel vmkcrypto module
- VMware CA



The vmkcrypto module contains the encryption algorithms used by ESXi.

https://t.me/learningnets

11-35 Enabling and Disabling FIPS 140-2 Mode on ESXi Hosts



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, run set `--enable true`.

To disable FIPS-140, run `set --enable false`.

https://t.me/learningnets

11-36 Enabling and Disabling FIPS 140-2 Mode on vCenter Server Appliance

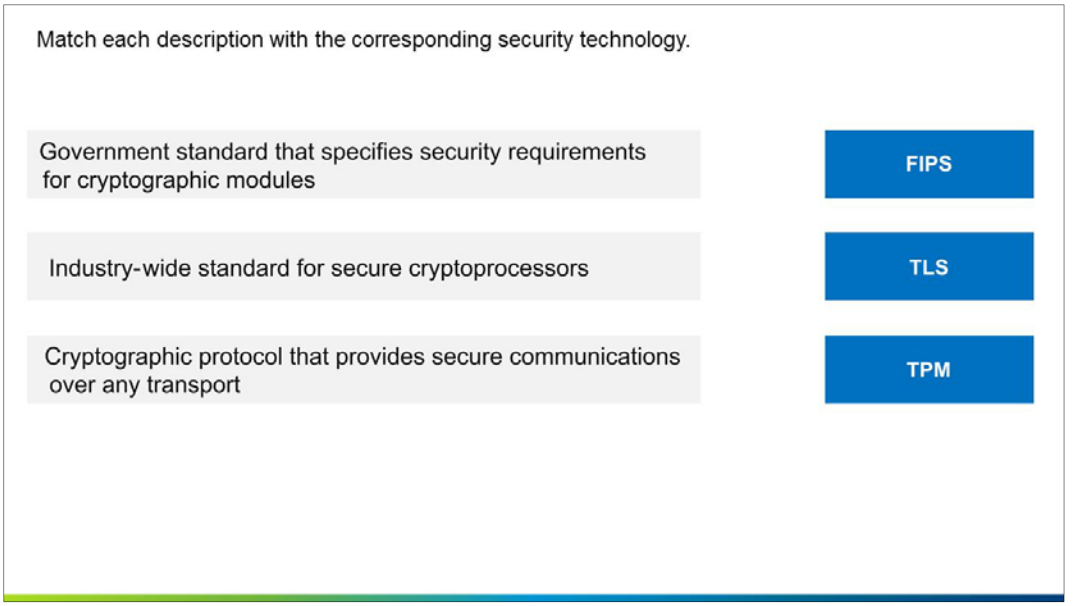
With the `fipsctl` command, you can view and modify the FIPS-140 settings for the rhttpproxy, sshd, and vmca services on vCenter Server Appliance.

To enable FIPS-140 on rhttpproxy, ssh, and vmca, run `enable –-service service_name`.

To disable FIPS-140, run `disable --service service_name`.

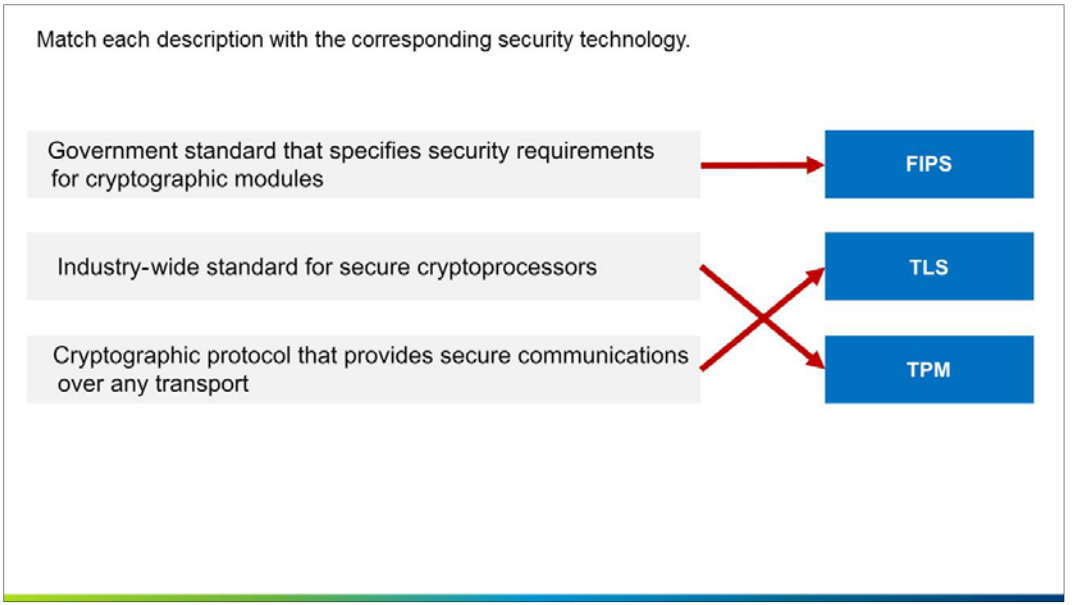
https://t.me/learningnets

11-37 Activity: Security Technologies (1)



https://t.me/learningnets

11-38 Activity: Security Technologies (2)



https://t.me/learningnets

11-39 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Secure vCenter Server with TLS 1.2
- Use UEFI Secure Boot for ESXi hosts
- Plan for secure boot and TPM 2.0 support for an ESXi host
- Enable FIPS 140-2 for ESXi hosts and vCenter Server instances

https://t.me/learningnets

https://t.me/learningnets

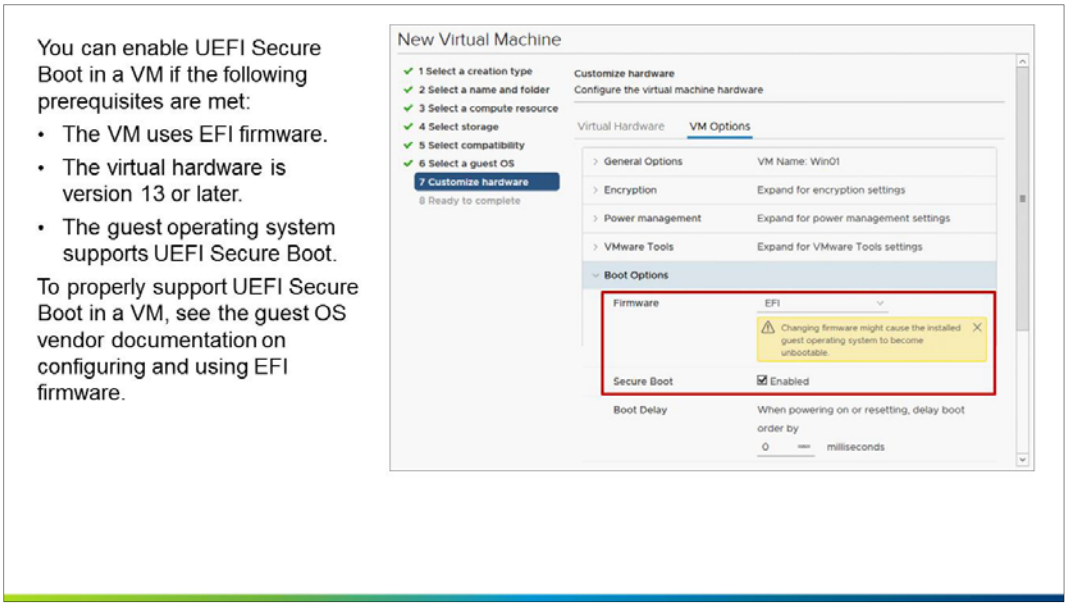
11-41 Learner Objectives

After completing this lesson, you should be able to meet the following objective:

- Describe support for VM security features

https://t.me/learningnets

11-42 VM Protection with UEFI Secure Boot



For certain VM hardware versions and operating systems, you enable UEFI Secure Boot exactly as you enable it 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 VM'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 and signed by Microsoft, such as Linux bootloaders
- A VMware certificate that is used only for booting ESXi inside a VM

The VM's default configuration includes one certificate for authenticating requests to modify the UEFI Secure Boot configuration, including the UEFI Secure Boot revocation list from inside the VM, which is a Microsoft key encryption key (KEK) certificate. In most cases, replacing the existing certificates is not necessary.

https://t.me/learningnets

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

11-43 Using vTPM in a VM (1)

vSphere 6.7 introduced support for the Virtual Trusted Platform Module (vTPM) device. With this device, you can add a TPM 2.0 virtual cryptoprocessor to a VM.
A vTPM device is a software emulation of the TPM functionality:

- The guest operating system can create and store private keys in such a way that they are never exposed to this guest operating system.
- The guest operating system can 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:

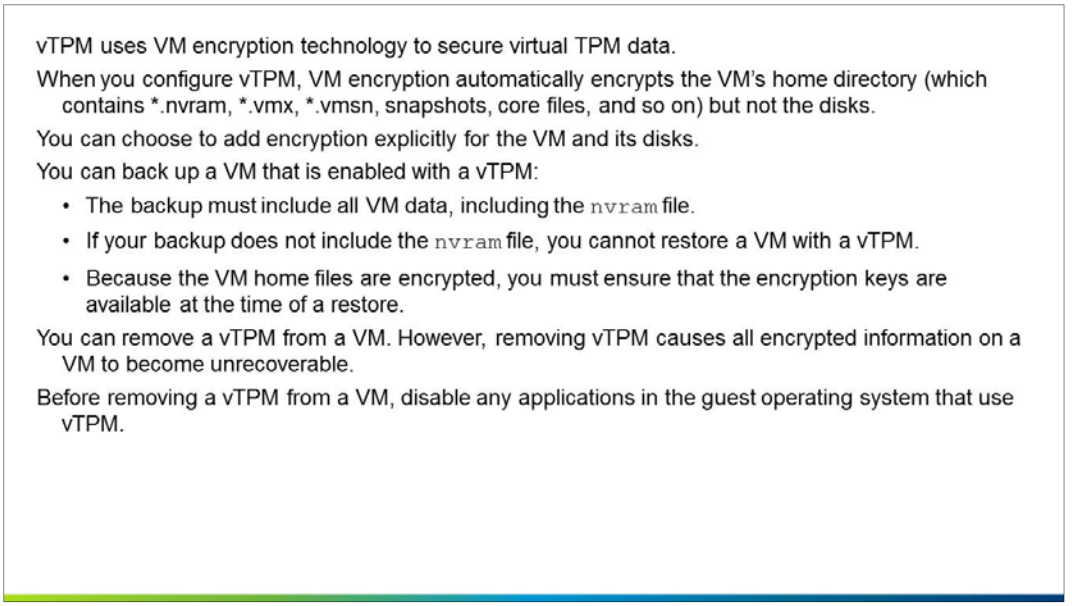
- An operating system can verify that the firmware loaded was not compromised since the last run.
- An application can verify that the operating system did not load any malicious components.

Compromising the guest operating system usually compromises its secrets, but enabling a vTPM in the guest reduces this risk.

You can add a vTPM device to either a new VM or an existing VM.

https://t.me/learningnets

11-44 Using vTPM in a VM (2)

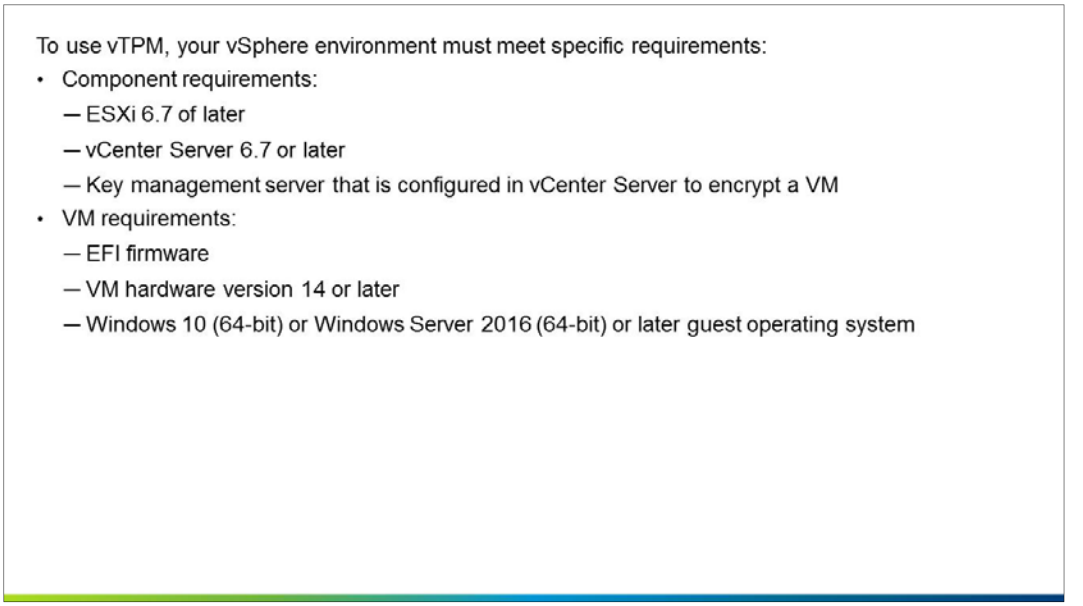


vTPM can be enabled on a VM by editing the VM's settings with the vSphere Client.

vTPM can be removed from a VM but the VM must be powered off first. Before you remove vTPM, disable any applications that use vTPM. If you do not disable these applications, the VM might fail to boot when powered on.

https://t.me/learningnets

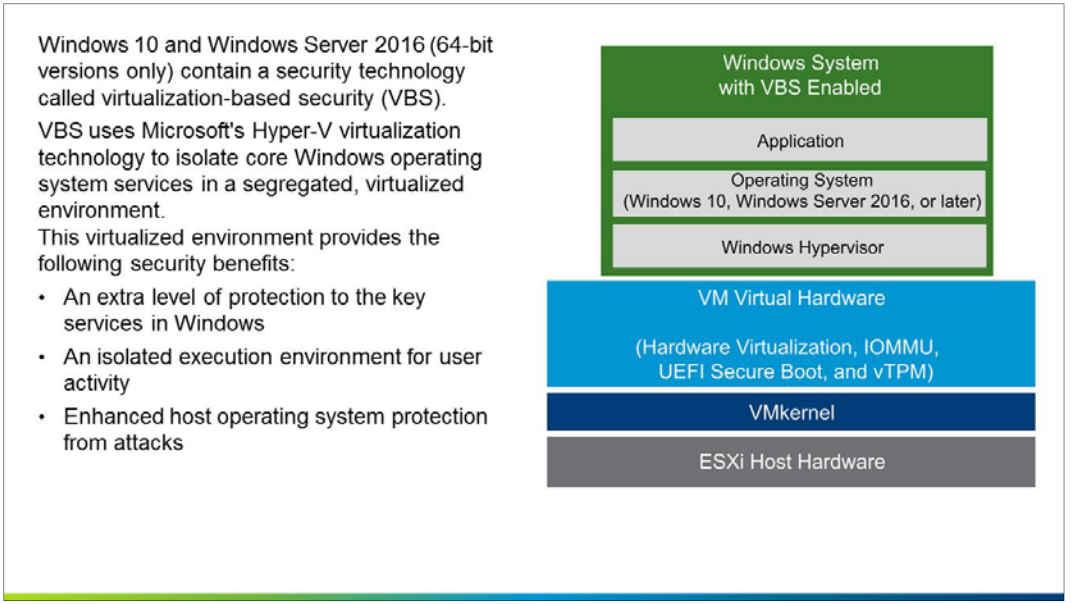
11-45 Prerequisites for Adding vTPM



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

11-46 Virtualization-Based Security for Windows Guest Operating Systems



Microsoft virtualization-based security (VBS), a feature of Windows 10 and Windows Server 2016 and later operating systems, uses hardware and software virtualization to enhance system security by creating an isolated, hypervisor-restricted, specialized subsystem.

With VBS, you can use the following Windows security features to harden your system and isolate key system and user secrets so that they are not compromised:

- Credential Guard: Aims to isolate and harden key system and user secrets against compromise
- Device Guard: Provides a set of features that 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

https://t.me/learningnets

For more information about virtualization-based security, see the Microsoft documentation.

https://t.me/learningnets

11-47 VBS Requirements and Limitations

To use VBS in a VM, you must meet prerequisites:

- The VM must be of virtual hardware version 14 or later.
- The guest operating system must be either Windows 10 or Windows Server 2016, or later (64 bit).
- The VM must be on an ESXi host version 6.7 or later with a compatible TPM device installed and enabled for vTPM support.

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

11-48 Enabling VBS

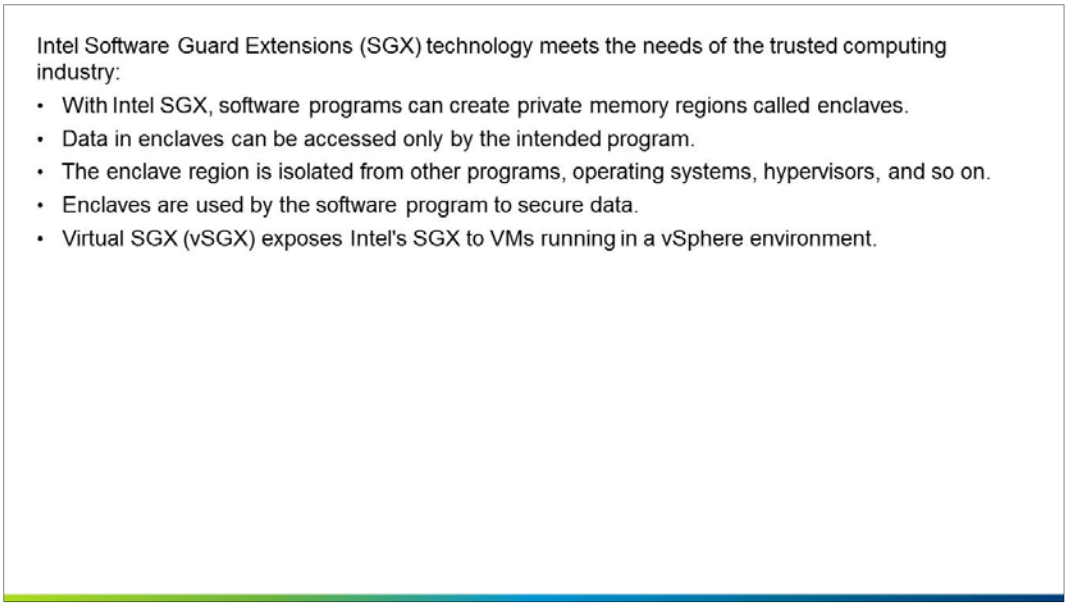
To enable VBS when you create VM, select Windows 10 or Windows Server 2016 or later and select the **Enable Windows Virtualization Based Security** check box.

Enabling VBS exposes hardware-assisted virtualization, the input-output memory management unit (IOMMU), EFI, and UEFI 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

11-49 About SGX



Trusted computing refers to various technologies and proposals for resolving computer security problems. The problems are resolved through hardware enhancements and associated software modifications. Several major hardware manufacturers and software vendors, collectively known as the Trusted Computing Group (TCG), cooperate in this venture and create specific plans.

For more information about the Trusted Computing Group, see https://trustedcomputinggroup.org.

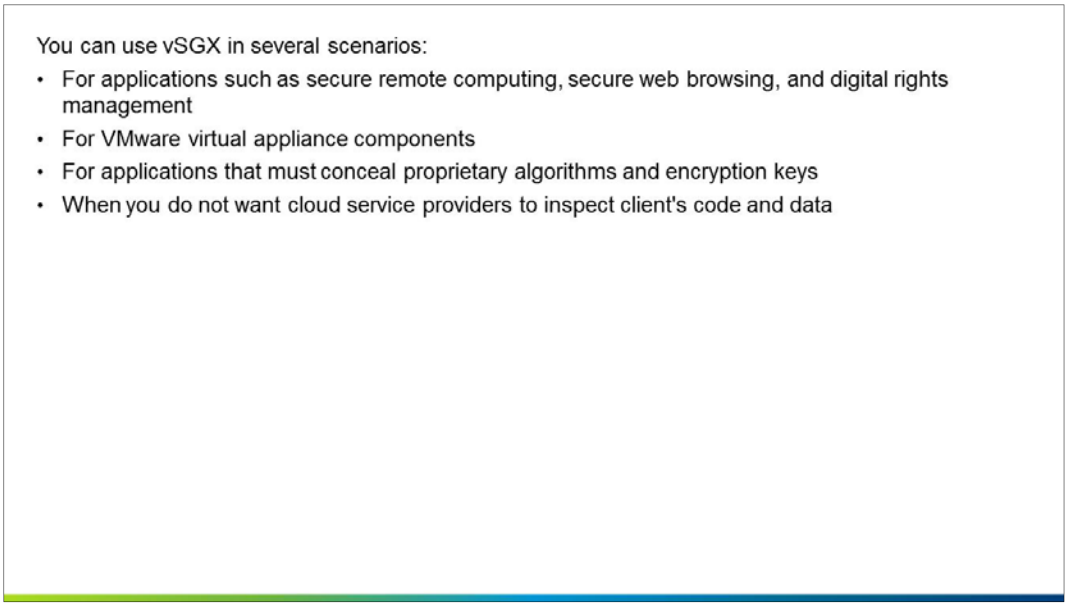
Data inside an application can be compromised by rogue software or a bug in the operating system or in the hypervisor.

Programmers can use Intel SGX to create private memory regions called enclaves. Data in enclaves is protected from unauthorized access and modification by rogue software running at higher privilege levels. Intel SGX provides this protection without disrupting the ability of legitimate software to schedule and manage the use of platform resources.

vSGX is implemented as part of the vSphere core virtualization stack. With vSGX, applications can run on VMs to create their own enclaves.

https://t.me/learningnets

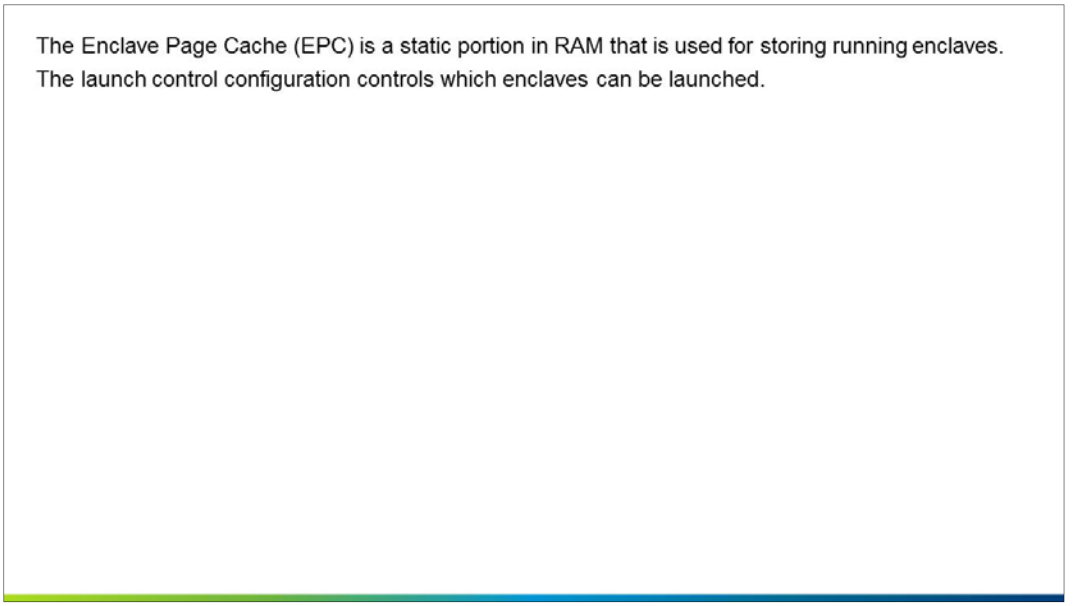
11-50 vSGX Use Cases



Many third-party applications can use vSGX.

https://t.me/learningnets

11-51 SGX Terminology

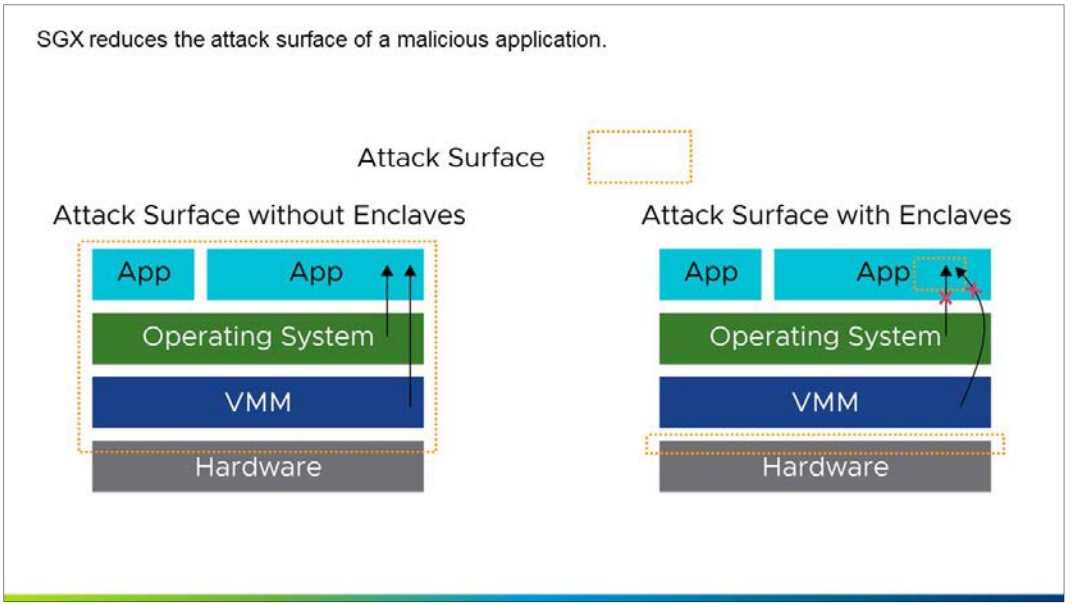


The Enclave Page Cache (EPC) is a static portion of physical RAM. This portion is allocated at boot time by the BIOS. EPC stores running enclaves. It is a limited resource, typically not more than 93 MB.

Launch control configuration is an architectural component from which the platform can control which enclaves can be launched.

https://t.me/learningnets

11-52 Intel SGX Architecture



Applications frequently work with private information such as passwords, encryption keys, bank account numbers, and so on. This private data should be accessed only by the designated recipient. The job of the operating system is to protect such private data from other applications and users.

Operating systems and applications often employ safeguards to protect this private data. Despite these protections, most computer systems are still vulnerable.

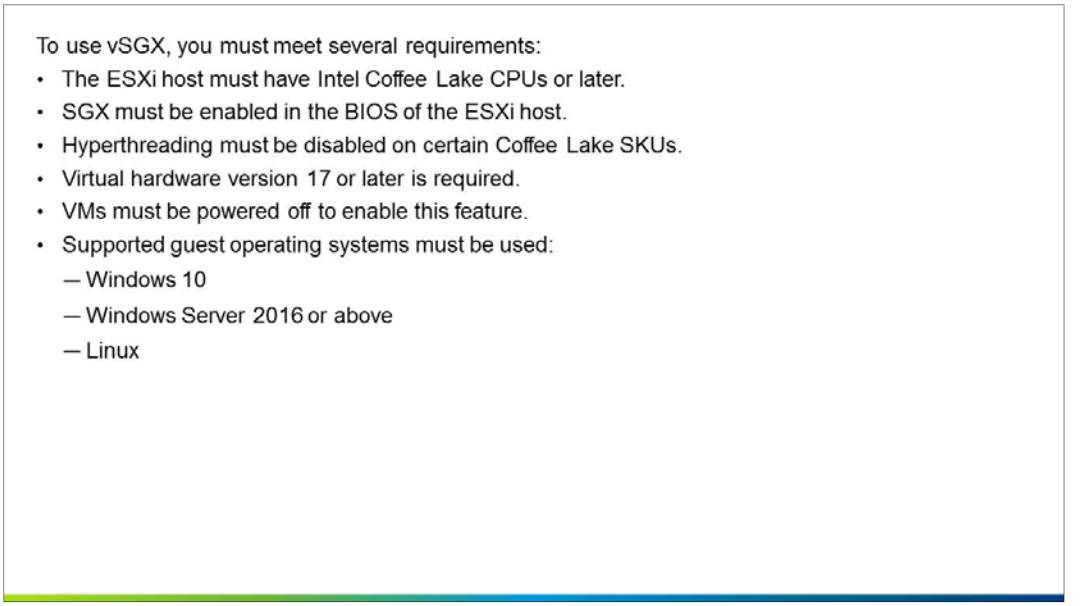
Malware with administrative privileges can access all operating system resources including all applications running on that system. Malware can target an application's protected data to extract encryption keys and secret data directly from memory.

With Intel SGX, applications can create enclaves, which protect the confidentiality and integrity of applications from privileged malware.

Privileged malware is malware that gets privileges by manipulating the underlying system.

https://t.me/learningnets

11-53 vSGX Requirements

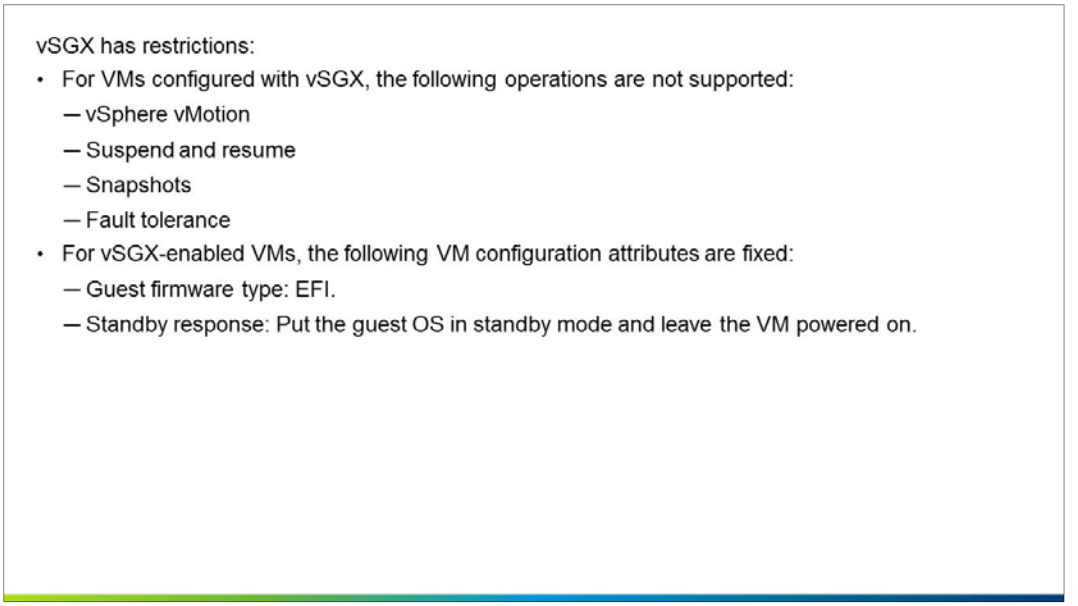


In theory, all versions of Linux are supported. In practice, a special kernel module published by Intel must be loaded. This special kernel module is available for only a few Linux distributions.

For more information about this module, see Intel SGX SDK at https://01.org/intel-softwareguard-extensions and Intel SGX SDK download at https://01.org/intel-software-guard-extensions/downloads.

https://t.me/learningnets

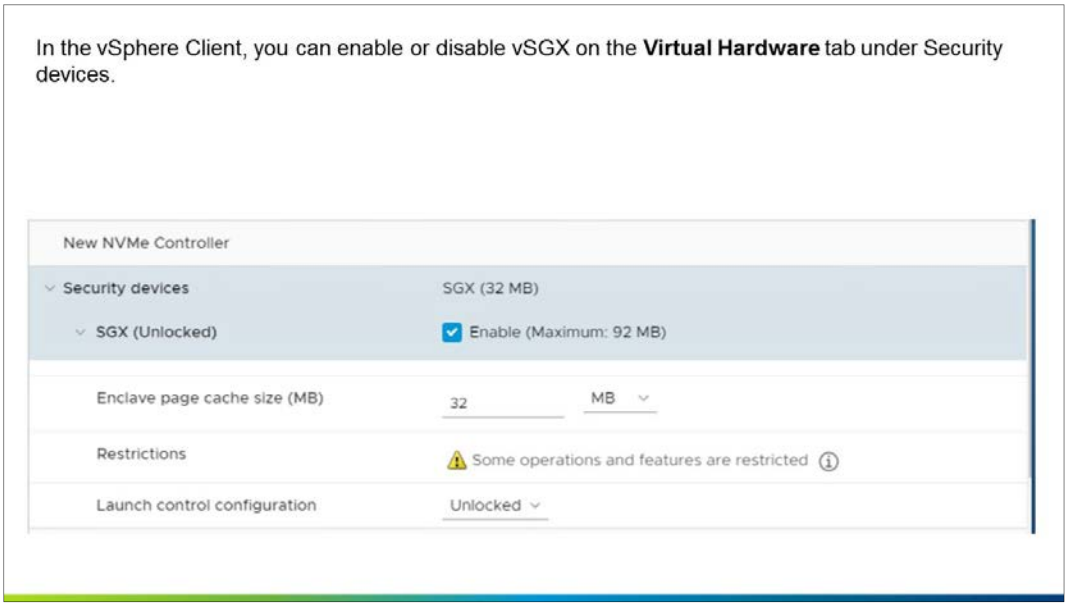
11-54 vSGX Restrictions



Cold migrations can be performed.

https://t.me/learningnets

11-55 Enabling and Disabling vSGX for a VM



vSGX can be enabled or disabled using the **Edit Settings** menu for the VM.

The VM must be powered off to enable or disable the feature.

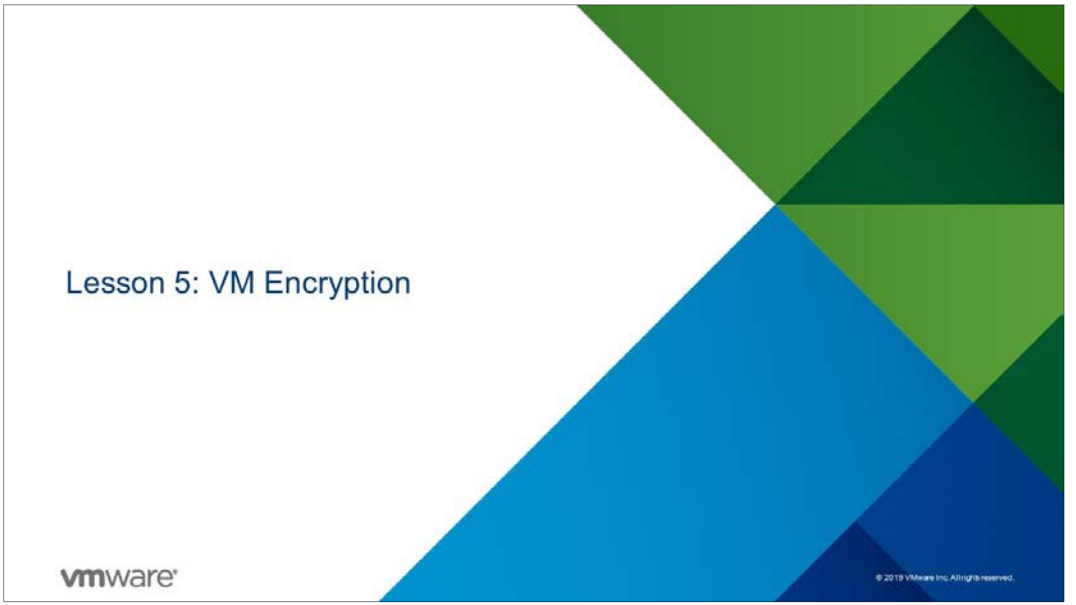
https://t.me/learningnets

11-56 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objective:

- Describe support for VM security features

https://t.me/learningnets



https://t.me/learningnets

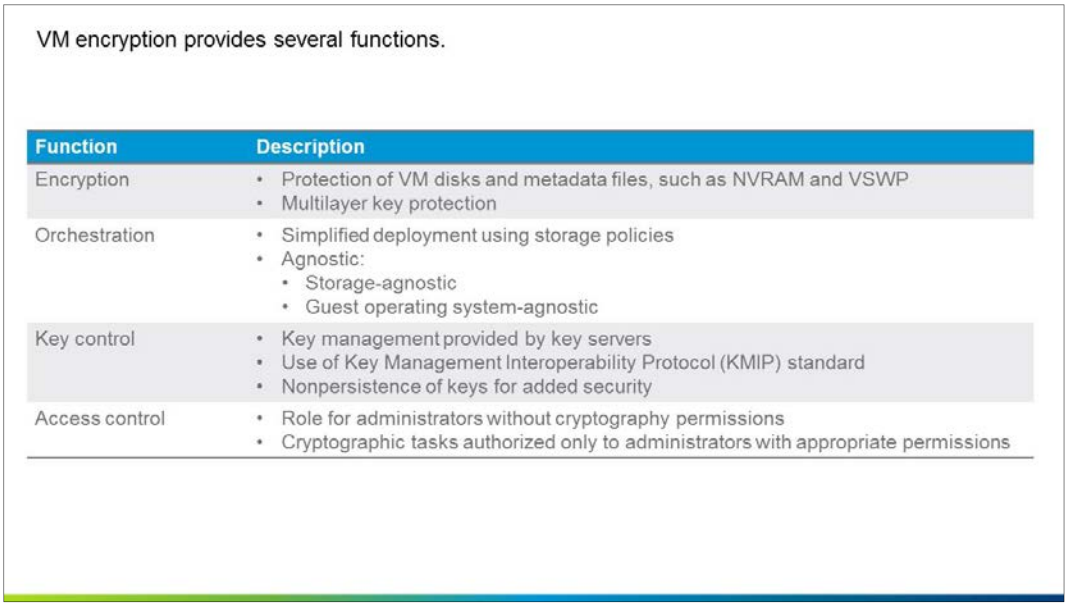
11-58 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe the components of a VM encryption architecture
- Manage encrypted VMs
- Encrypt core dumps
- Use encryption with instant clones
- Migrate Encrypted VMs
- List VM encryption events and alarms
- Enable encrypted vSphere vMotion

https://t.me/learningnets

11-59 About VM Encryption



Introduced in vSphere 6.5, VM encryption protects not only the data on disk but also the swap file and any guest-specific information that might be contained in the `VMX` or `NVRAM` files. VM encryption also provides a multilayer key protection mechanism that makes breaking the security of the VM almost impossible.

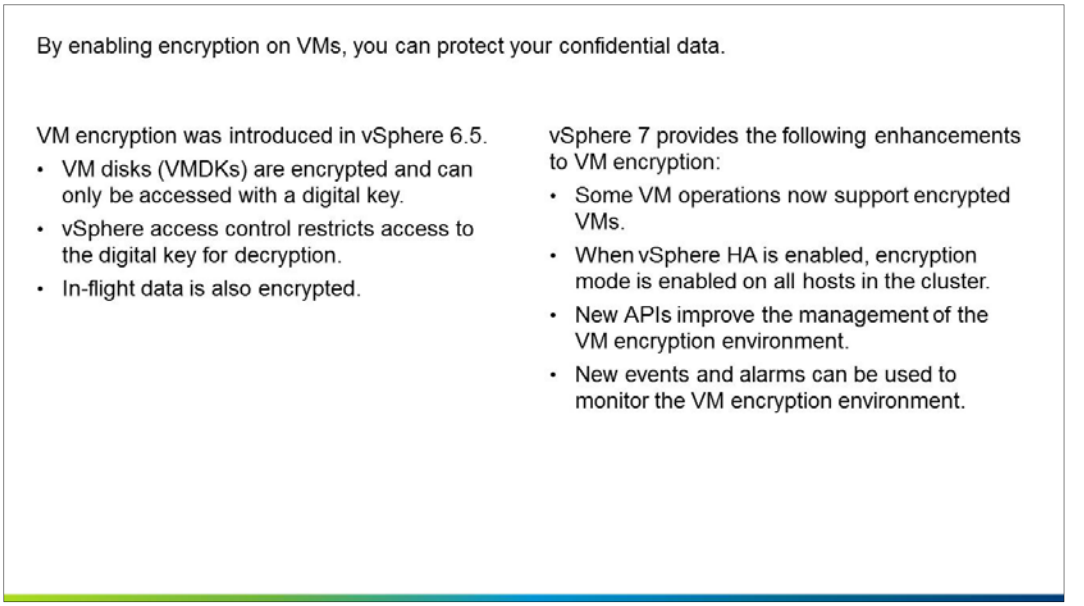
VM encryption makes it simple to configure encryption by using the orchestration capabilities of vCenter Server storage policies. By using storage policies, you can apply encryption to any type of VM, regardless of the guest operating system or the storage on which it resides.

By using external, enterprise-grade key servers, exposure to risk is limited because 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 VM.

https://t.me/learningnets

11-60 Overview of VM Encryption



The enhancements in vSphere 7 improve the general functionality of VM encryption. They are based on competitive analysis and feedback from VMware customers and partners.

To mitigate the risk of an administrator accessing data without authorization, VMs can be encrypted, and access to the encryption and decryption functionality can be provided to a subset of administrators. You can provide this type of encryption by using the inbuilt cryptographic privileges in vSphere.

https://t.me/learningnets

11-61 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 VM data, including swap files, not only virtual disk files.
- It enables detailed key control: VMs and disks can use different keys.
- It offers easy orchestration through VM storage policies.
- It is compatible with KMIP, an industry-standard language for the management of security keys. KMIP allows the interoperable exchange of data between different key management servers and clients.

https://t.me/learningnets

11-62 Business Use Case: Securing VMs

A common requirement in a company's security policy is to protect VM 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 VMDK file or the entire VM to a removable storage device and leaving the company with this data.
- Solution:
  - With VM 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 extra layer of encryption.
  - The company grants only a limited amount of people access to the key.

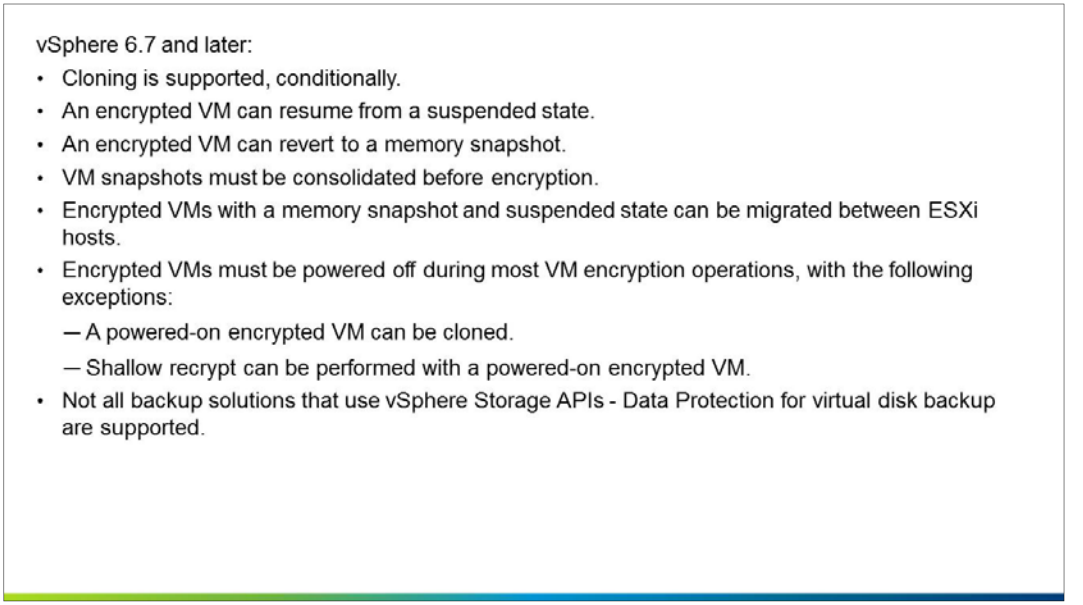
The digital key used to encrypt a VM is secured in an extra layer of encryption, much like taking the key needed to open a safe and placing it in another safe.

You allow regular vSphere administrators to continue with their daily activities without change, but you grant cryptographic privileges to a subset of these administrators.

Access to VM files does not mean that storage administrators or vSphere administrators can walk away with the VM (VMDK) file and read its contents offsite.

https://t.me/learningnets

11-63 VM Encryption Interoperability (1)



vSphere VM encryption has limitations in the devices and features that it can interoperate with in vSphere 6.5 and later releases:

- Full clones are supported. The clone inherits the parent encryption state, including keys. You can re-encrypt a full clone to use new keys or decrypt the full clone.
- Linked clones are supported, and the clone inherits the parent encryption state including keys. You cannot decrypt the linked clone or re-encrypt a linked clone with different keys.
- Starting with vSphere 6.7, you can resume from a suspended state of an encrypted VM or revert to a memory snapshot of an encrypted VM. You can migrate an encrypted VM with memory snapshot and suspended state between ESXi hosts.

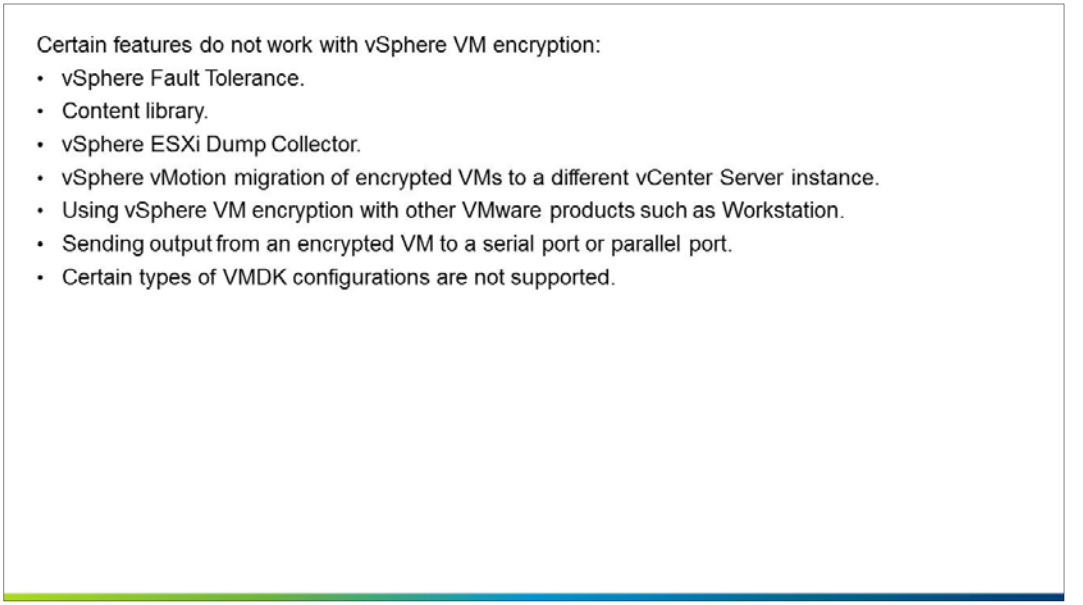
https://t.me/learningnets

- You cannot perform the following tasks on an encrypted VM:
  - For most VM encryption operations, the VM must be powered off. You can clone an encrypted VM, and you can perform a shallow recrypt while the VM is powered on.
  - You cannot encrypt a VM that has existing snapshots. Consolidate all existing snapshots before you perform the encryption.
- You can use vSphere VM encryption with IPv6 mode or in mixed mode. You can configure the KMS with IPv6 addresses. Both vCenter Server and the KMS can be configured with only IPv6 addresses.

For more information about VM encryption interoperability, see https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.security.doc/GUID-C0AF1F3A-67B4-41A6-A933-7E52A3603D9D.html.

https://t.me/learningnets

11-64 VM Encryption Interoperability (2)



Encrypted migration with vSphere vMotion of an unencrypted VM is supported.

You cannot send output from an encrypted VM to a serial port or parallel port. Even if the configuration appears to succeed, output is sent to a file.

Certain types of VMDK configurations are not supported with vSphere VM encryption:

- Flash Read Cache.
- A named virtual disk unassociated with a VM, also called First Class Disk.
- RDM (Raw Device Mapping).
- Multiwriter or shared disks (MSCS, WSFC, or Oracle RAC). If a virtual disk is encrypted, and you select **Multi-writer** on the VM's Edit Settings page, the **OK** button is disabled.

https://t.me/learningnets

For more information, see *Virtual Machine Encryption Interoperability* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.security.doc/GUID-C0AF1F3A-67B4-41A6-A933-7E52A3603D9D.html.

https://t.me/learningnets

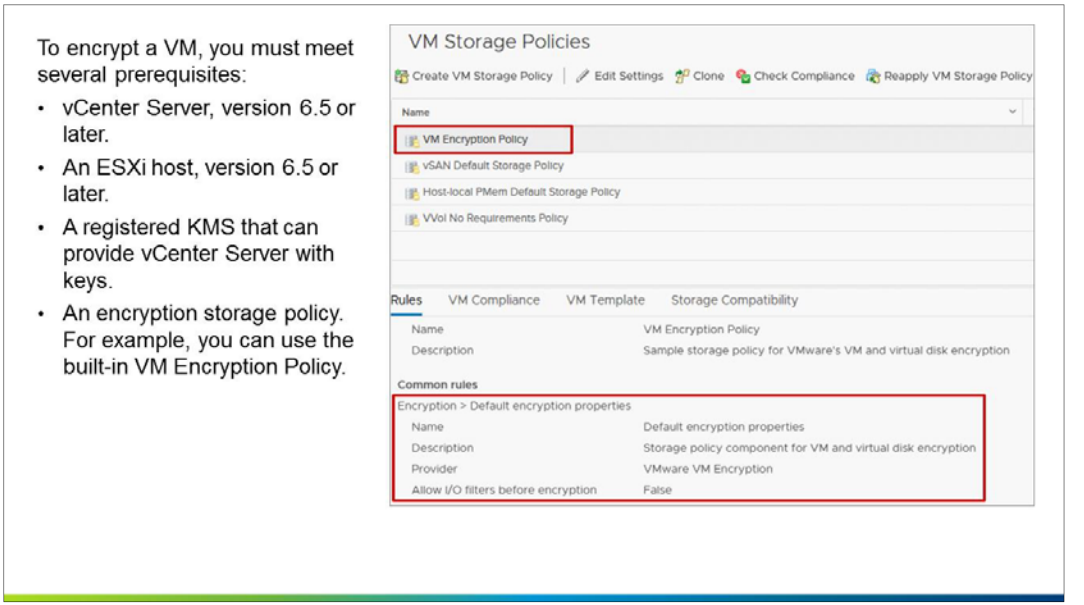
11-65 VM Encryption Architecture

The VM encryption architecture consists of the following components:

- Key management server (KMS):
  - An external server that provides keys to services
  - Provided by a third-party security vendor
- vCenter Server:
  - Communicates with a KMS and requests keys on behalf of ESXi hosts
- ESXi hosts:
  - Run encrypted VMs

https://t.me/learningnets

11-66 VM Encryption Prerequisites



To encrypt VMs, you must 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. You 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 sees plain text before it is encrypted.

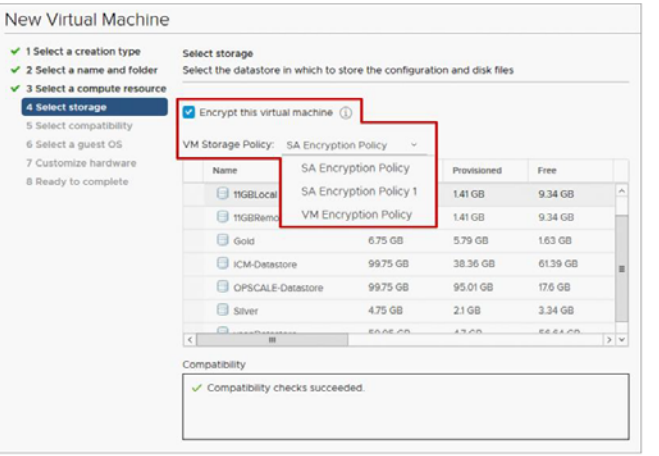
https://t.me/learningnets

11-67 Encrypting New VMs

If a new VM has an encryption storage policy applied to it, the VM is encrypted immediately when it is created.

You do not need to take additional steps 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.

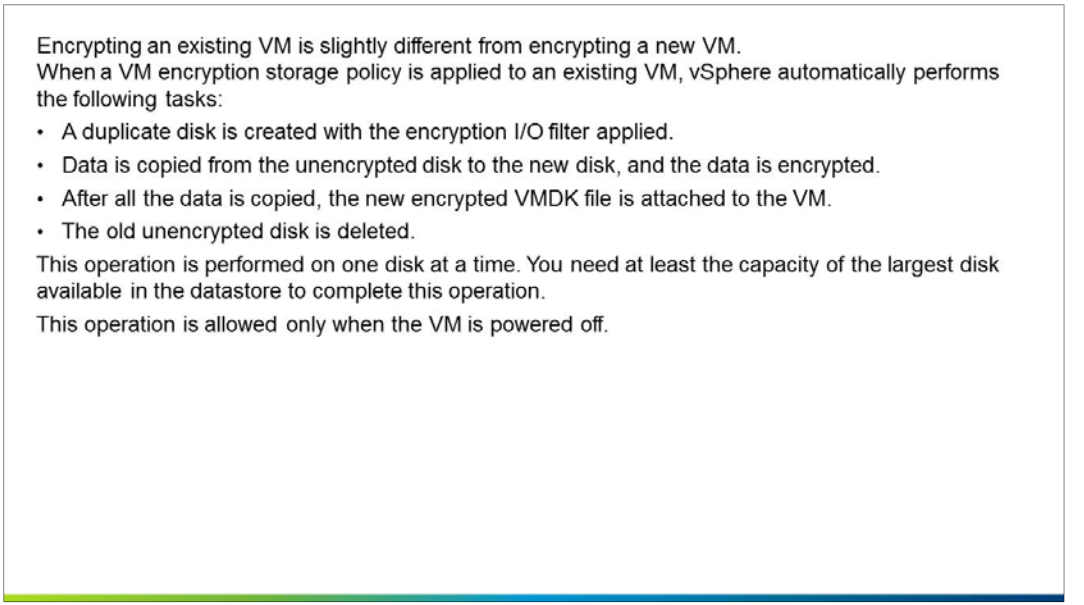


To encrypt a new VM, perform the following steps:

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

11-68 Encrypting Existing VMs



For the data of an existing VM, 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 VM, and the old unencrypted disk is deleted.

The VM must be powered off so that no mirror driver is employed.

https://t.me/learningnets

11-69 Unlocking Encrypted VMs (1)

A VM goes into a locked state when its key encryption key (KEK) is unavailable.
A VM's KEK might be unavailable for 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

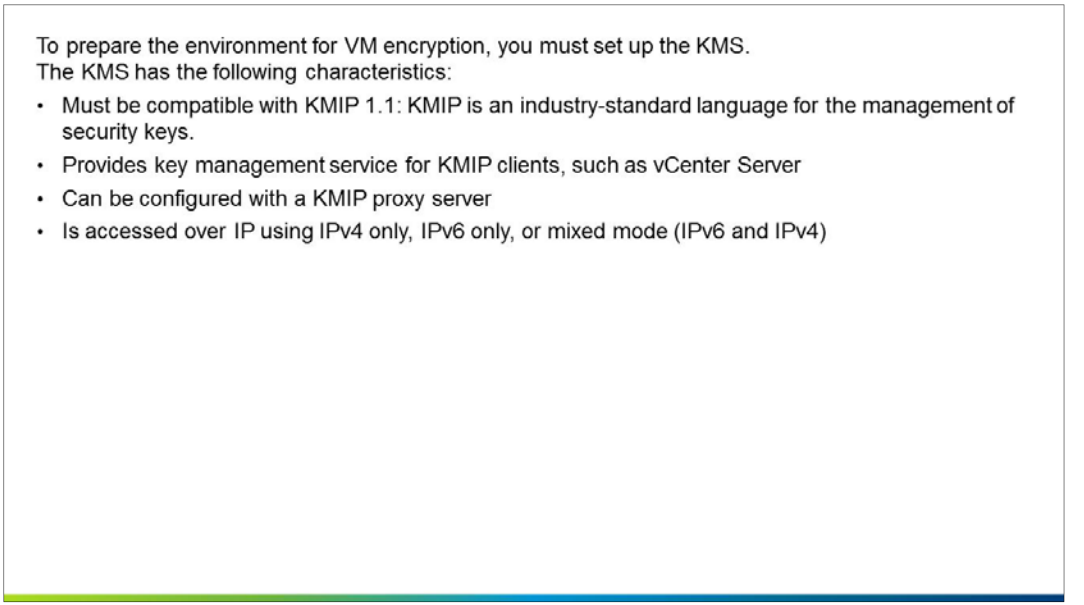
11-70 Unlocking Encrypted VMs (2)

To unlock a VM, first ensure that the required key is available on the KMS cluster.
In the vSphere Client, click **Unlock VM** to unlock the encrypted VM:

- The Unlock Virtual Machine API sends the VM encryption key to the ESXi host.

https://t.me/learningnets

11-71 About the KMS



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 return the key values to vCenter Server using a secure channel.

Having a KMS means that all 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 must secure, a single or limited number of servers are secured, which is easier and safer.

KMIP proxy servers can provide the key management capability of the KMS to a remote location. With KMIP proxy servers, you can keep your critical key server safe in your central data center while still allowing remote offices to use its services without compromising security.

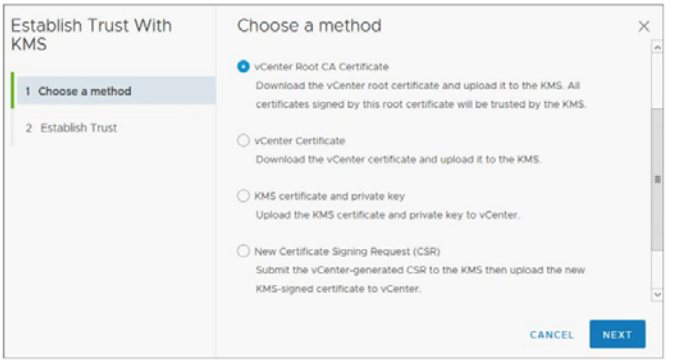
Finally, using the KMIP requires only an IP-capable network, which is readily available in most data centers.

https://t.me/learningnets

11-72 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 must be downloaded by the client.
- Other vendors accept only a client certificate provided by the KMS itself. Check with the KMS vendor for requirements.



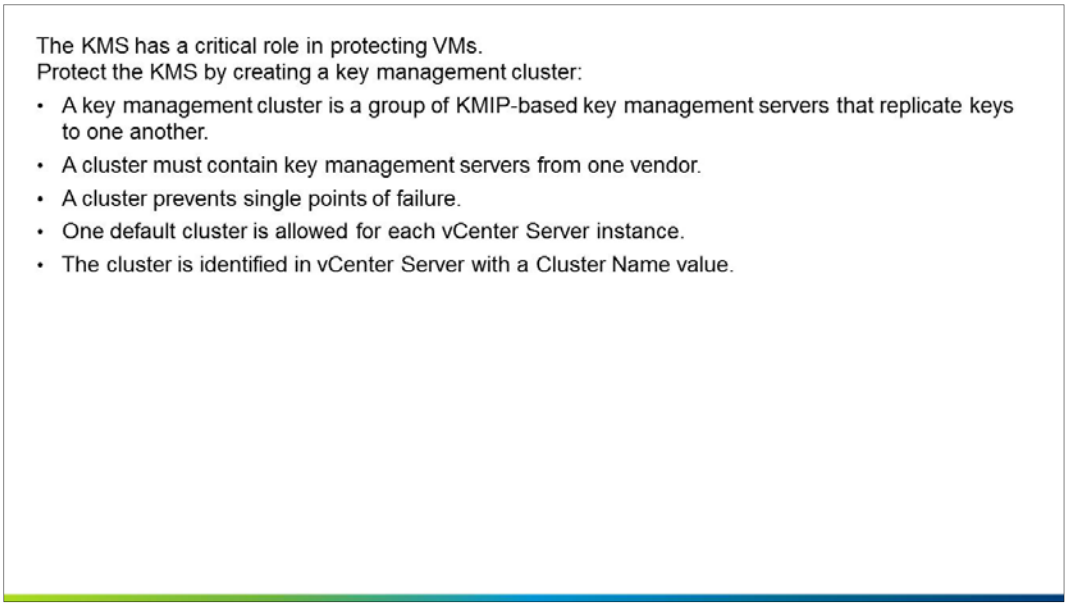
Different KMS vendors have varying requirements for the format of the client certificate that is used by vCenter Server.

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.

You can use the vSphere Client to provide certificate details based on the security requirements of the different types of implementations.

https://t.me/learningnets

11-73 Key Management Clusters



To protect against the potential failure of a KMS (also called a KMIP server), KMIP clusters are formed. These clusters 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, 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.

Although 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 identified. Clusters of KMIP servers are identified using a user-defined friendly name, populated in the vSphere Client.

https://t.me/learningnets

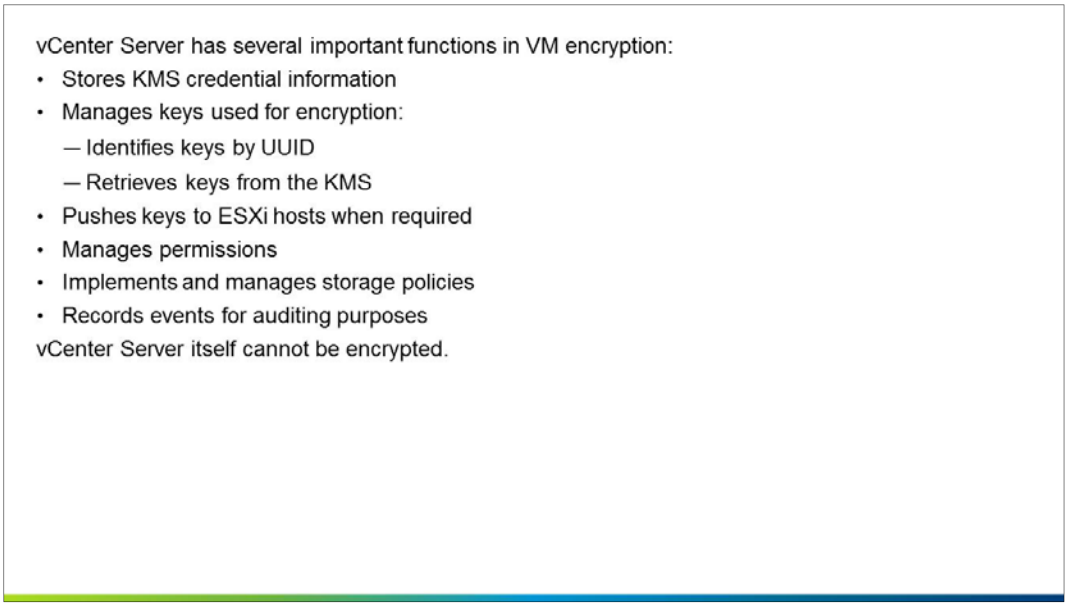
11-74 Summary of Keys

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 (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

11-75 Role of vCenter Server in VM Encryption



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 VM'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.

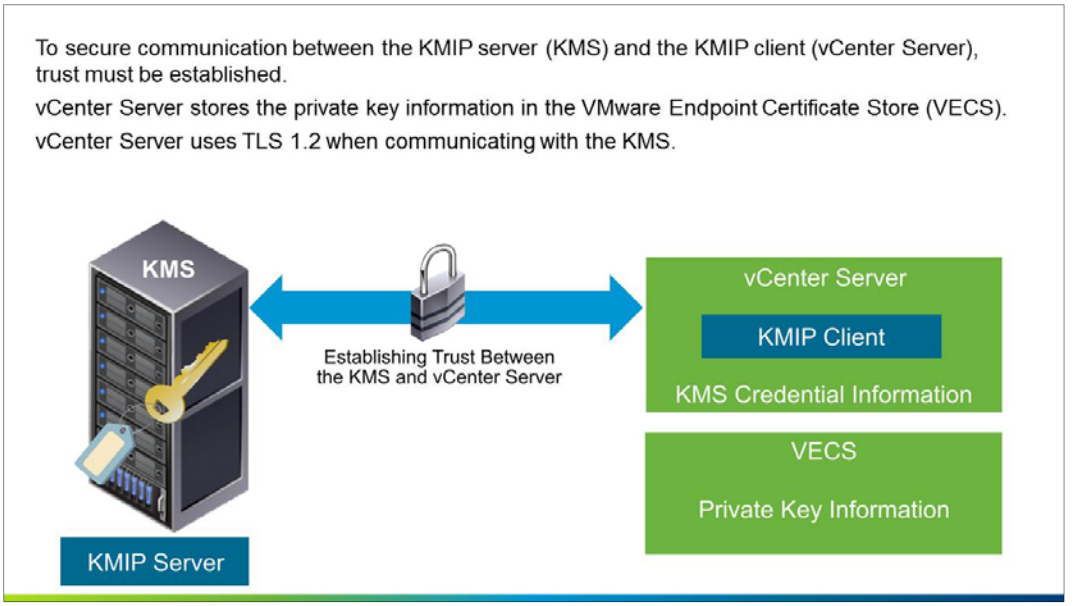
vCenter Server provides the framework for managing permissions for cryptographic tasks. You can restrict access to critical operations to a subset of your administrators, while ensuring that your other administrators can continue with their daily tasks.

vCenter Server manages storage policies that define whether a VM is encrypted.

vCenter Server records events for auditing purposes. Event information includes the identity of the user who initiated the event.

https://t.me/learningnets

11-76 KMS and vCenter Server Communication



After trust is established between the KMS (KMIP server) and the vCenter Server (KMIP client), vCenter Server stores the KMS credential information. The VMware Endpoint Certificate Store (VECS) stores the SSL private key information.

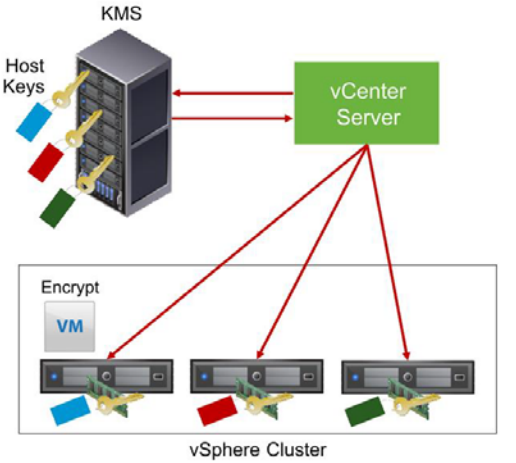
https://t.me/learningnets

11-77 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, creates 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.

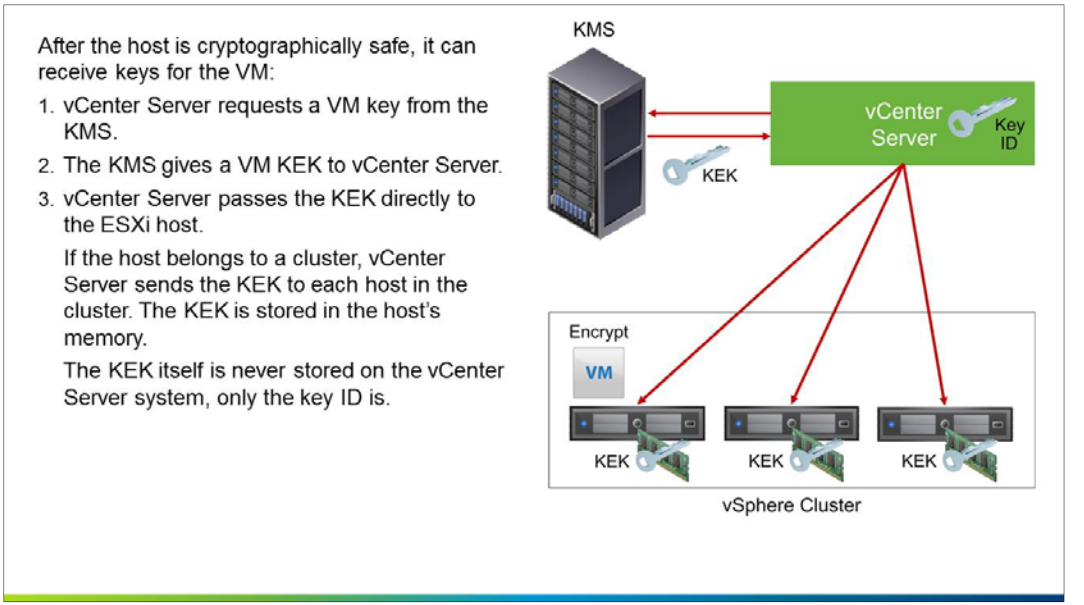


Users with the required privileges can perform cryptographic operations. These operations include creating encrypted VMs, encrypting existing VMs, and decrypting encrypted VMs.

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 hosts in the cluster are managing encrypted VMs.

https://t.me/learningnets

11-78 VM Encryption Process Flow (1)



When a cryptographic operation occurs, vCenter Server must determine whether a VM key must 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.

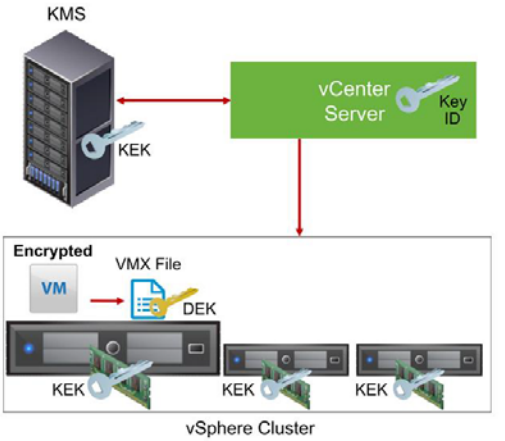
If the host belongs to a cluster, the cluster is deemed the security boundary. In this case, vCenter Server gives the KEK to all hosts in the cluster so that any cluster operations, such as vSphere vMotion migrations and vSphere HA operations, can occur.

https://t.me/learningnets

11-79 VM Encryption Process Flow (2)

KEKs remain in ESXi memory until the next reboot. For this reason, core dumps are encrypted.

4. The ESXi host generates a key called the data encryption key (DEK).
5. The DEK is used to encrypt the VM and its disks.
6. The ESXi host uses the KEK to encrypt the DEK.
7. The DEK is stored in the VM's configuration file in an encrypted format.



If an ESXi host reboots, 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 when 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 extra level of security.

https://t.me/learningnets

11-80 Activity: VM Keys (1)

An ESXi host can pull a VM's key from the KMS when required.

- ❒ True
- ❒ False

https://t.me/learningnets

11-81 Activity: VM Keys (2)

An ESXi host can pull a VM's key from the KMS when required.

❒ True

✔ False

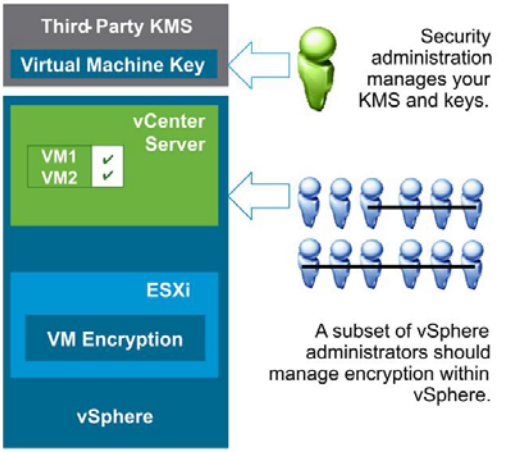
vCenter Server retrieves a VM's key from the KMS and pushes this key to ESXi hosts when required. ESXi hosts cannot pull a VM's key from either the KMS or vCenter Server.

https://t.me/learningnets

By default, the vCenter Server Administrator role has cryptographic privileges.

However, not all administrators should be able to control encryption operations or access keys.

vCenter Server provides a role called No Cryptography Administrator, with which you can control which administrators have encryption privileges.

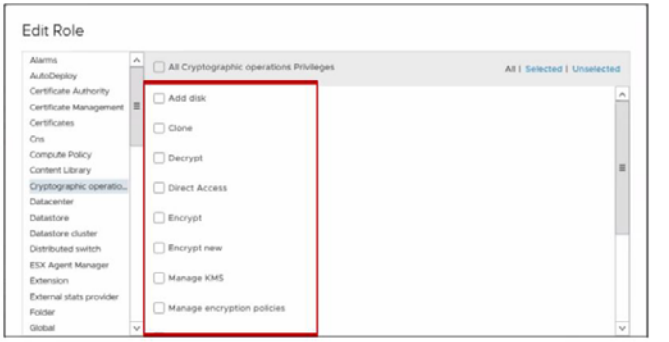


https://t.me/learningnets

11-83 vCenter Server Role: No Cryptography Administrator

This role does not include the following privileges:

- Cryptographic Operations
- Global Diagnostics
- Add host to cluster
- Add standalone host
- 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 add hosts to the vCenter Server inventory because this task might require the installation of a host key, making the task a cryptographic operation.

https://t.me/learningnets

11-84 Activity: Cryptographic Privileges (1)

Which predefined vCenter Server role allows a user to perform cryptographic operations?

- ❒ Administrator
- ❒ Virtual Machine Power User
- ❒ Virtual Machine Console User
- ❒ No Cryptography Administrator

https://t.me/learningnets

11-85 Activity: Cryptographic Privileges (2)

Which predefined vCenter Server role allows a user 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

11-86 Pushing Keys to Hosts

Before vSphere 7, when a cryptographic operation occurs on an ESXi host, these conditions apply:

- The host must enter crypto-safe (encryption) mode.
- When one host in a cluster enters crypto-safe mode, all the other hosts in the cluster must enter crypto-safe mode.
- vCenter Server pushes VM keys to all the hosts in the cluster.

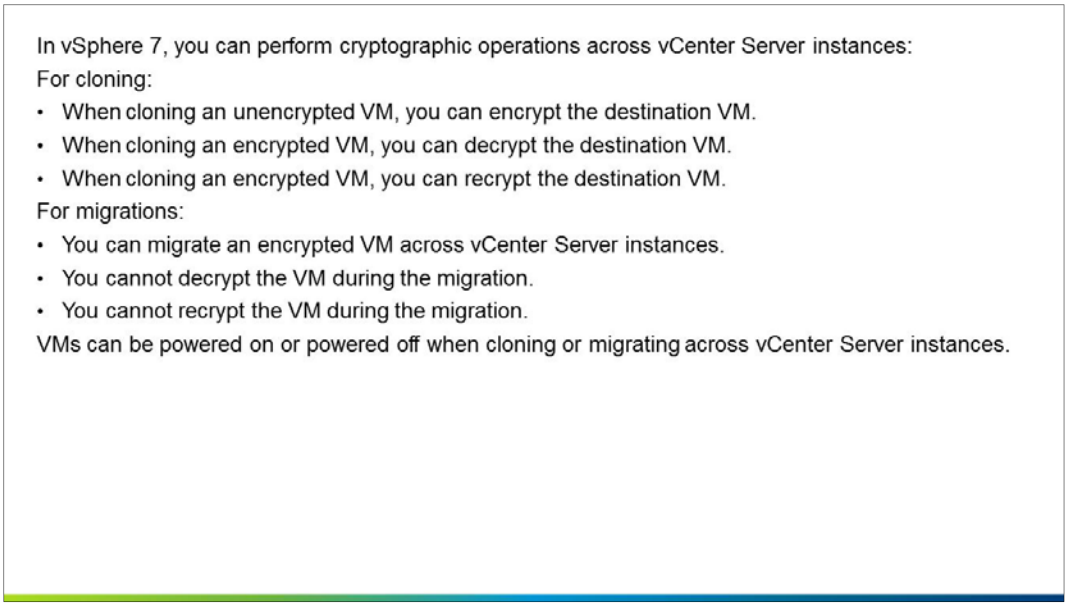
Starting with vSphere 7, the following conditions apply:

- If vSphere HA is enabled on a cluster, and at least one encrypted VM exists on a host in the cluster, vCenter Server pushes the necessary VM keys to all hosts in the cluster.
- If vSphere HA is not enabled, vCenter Server pushes VM keys only to the ESXi hosts on which the VM and its disks are registered.

vSphere DRS can be enabled or not. Only vSphere HA is checked.

https://t.me/learningnets

11-87 Cross vCenter Migration and Cloning of an Encrypted VM



During the migration operation, you cannot change keys and, therefore, recrypting the VM during the migration is not supported. Also, you cannot change the storage policy during migration, and, therefore, decrypting the VM during the migration is not supported.

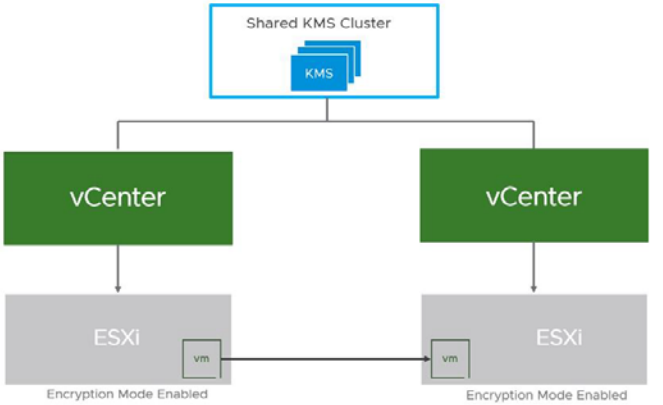
https://t.me/learningnets

11-88 Architecture for Cross vCenter Server Migrations and Clones (1)

The source and destination vCenter Server instances must point to the same KMS cluster, known as the shared KMS cluster.

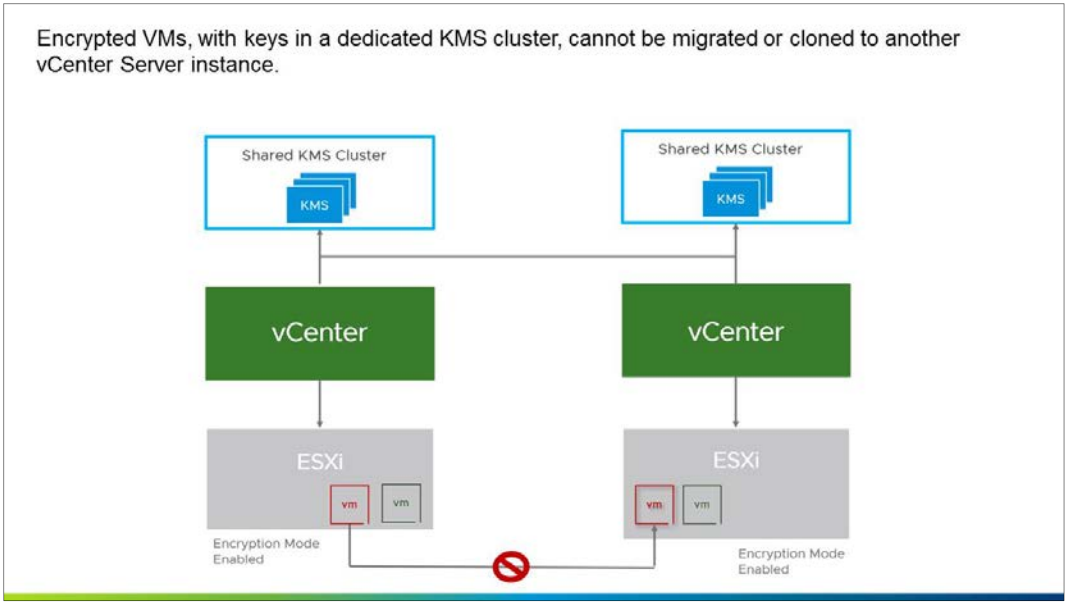
Encrypted VMs, with keys in the shared KMS cluster, can be migrated or cloned between vCenter Server instances.

Both source and destination ESXi hosts must be vSphere 7.



https://t.me/learningnets

11-89 Architecture for Cross vCenter Server Migrations and Clones (2)



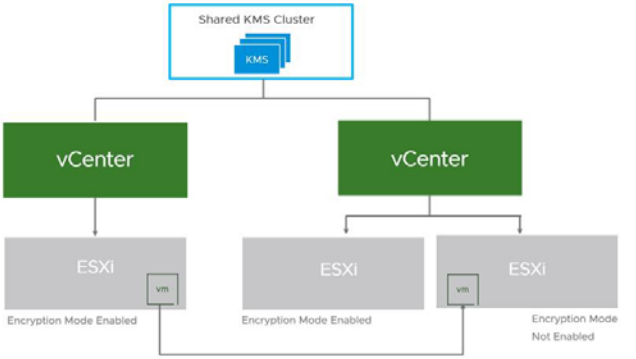
VM keys must reside in the shared KMS cluster for VM migrations and cloning to occur between vCenter Server instances.

https://t.me/learningnets

11-90 Architecture for Cross vCenter Server Migrations and Clones (3)

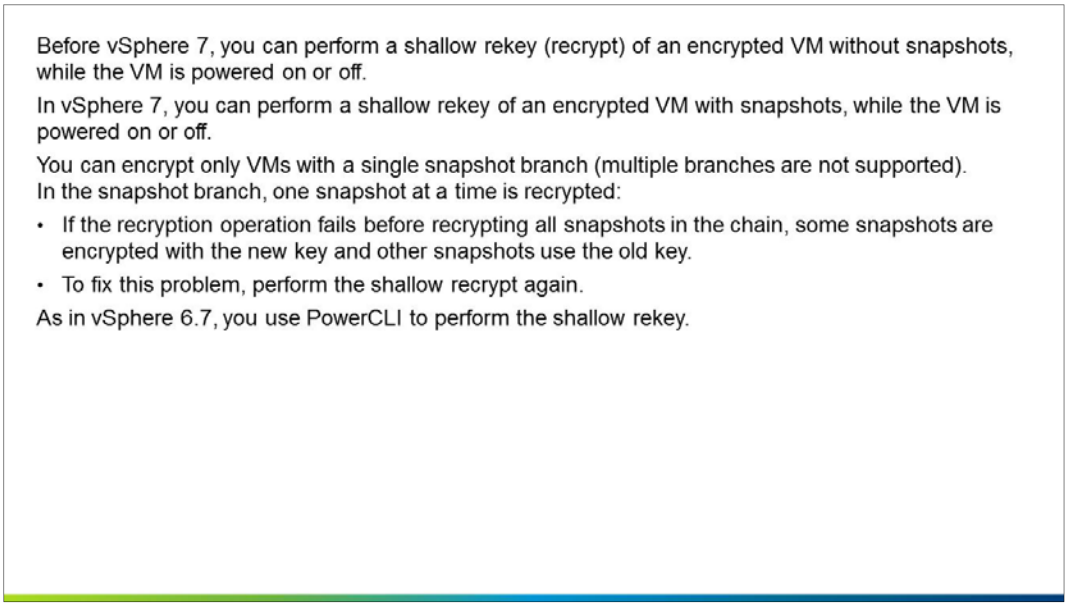
If the destination host is not in encryption mode, verify that the default KMS cluster is configured on the destination vCenter Server instance.

You can set the default KMS cluster to the shared KMS cluster or a different KMS cluster.



https://t.me/learningnets

11-91 Performing a Shallow Recrypt of VMs with Snapshots

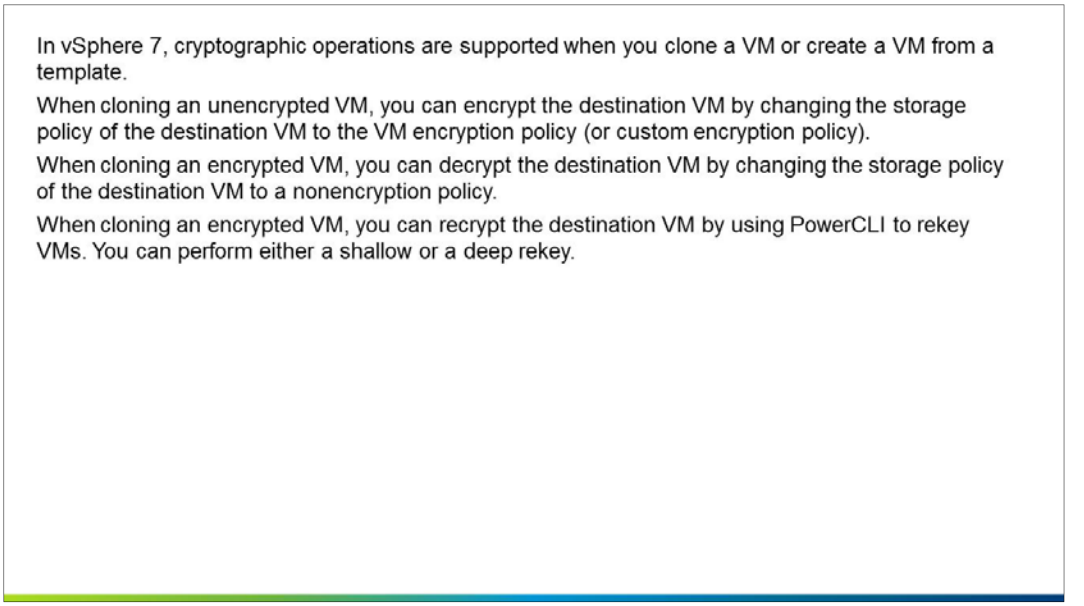


A deep recryption replaces all keys in a VM with new keys, from top to bottom. This operation requires rewriting all encrypted data in the VM.

A shallow recryption replaces only the top-level managed keys in a VM, namely, the VMDiskKey. Only a small amount of encrypted data must be rewritten.

https://t.me/learningnets

11-92 VM Encryption Operations When Cloning a VM

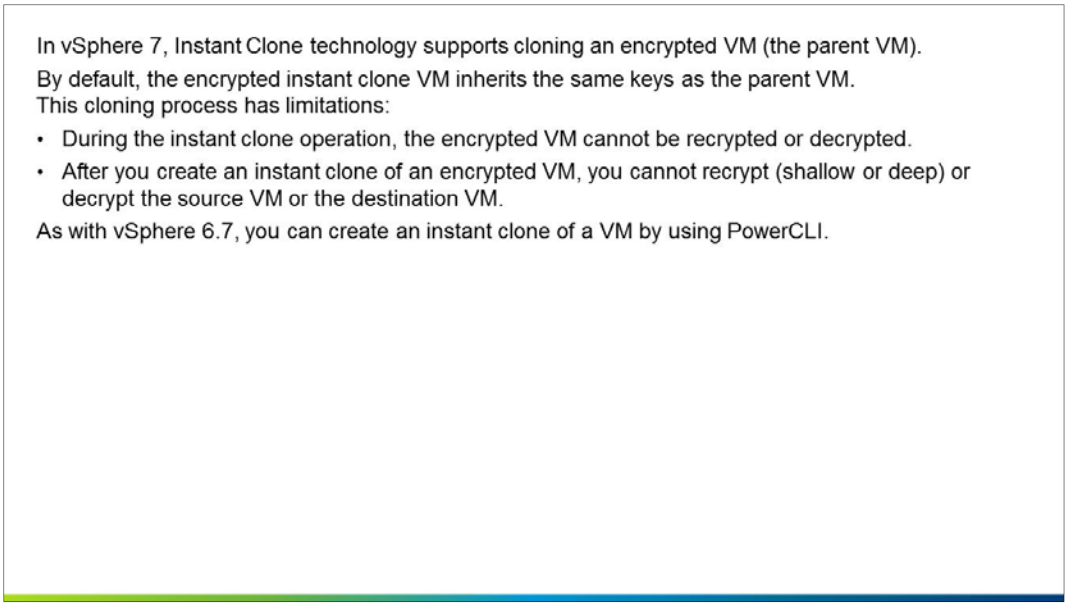


A shallow rekey (or recrypt) changes the KEK associated with the VM. The data encryption key (DEK) is wrapped with the new KEK.

Deep rekey (or recrypt) changes the DEK and recrypts all data using the new DEK. This function requires temporary disk space and is slower than a shallow rekey because all data must be rewritten.

https://t.me/learningnets

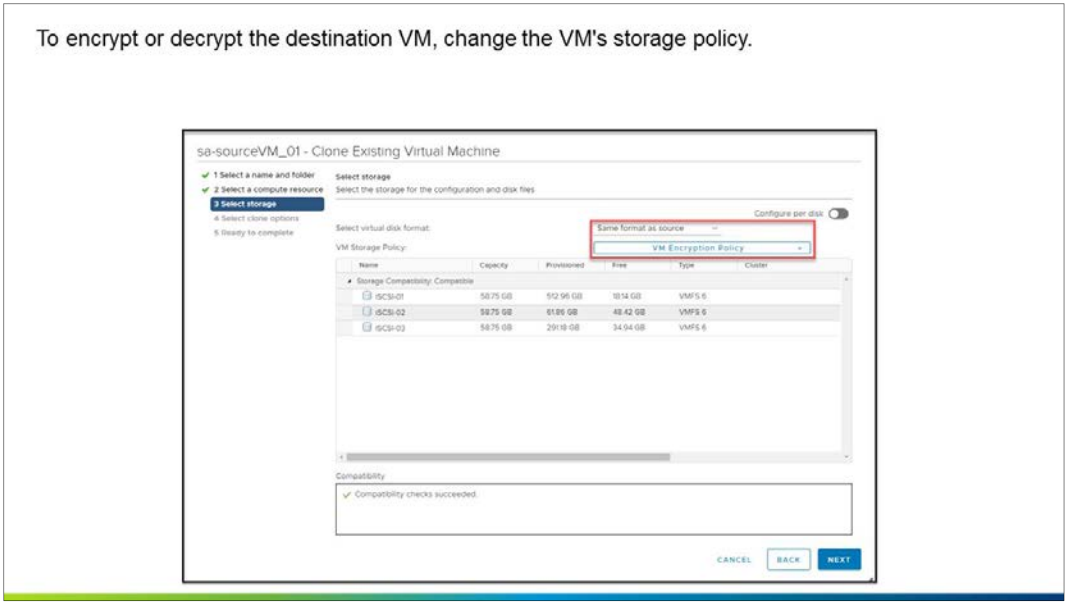
11-93 Cloning an Encrypted VM with Instant Clone Technology



You cannot change the storage policy or change encryption keys during the instant clone operation.

https://t.me/learningnets

11-94 Cloning a VM: Encrypting or Decrypting the Destination VM



Decrypting a VM can be a security risk. Therefore, as a best practice, restrict most users from having the privilege to decrypt VMs.

https://t.me/learningnets

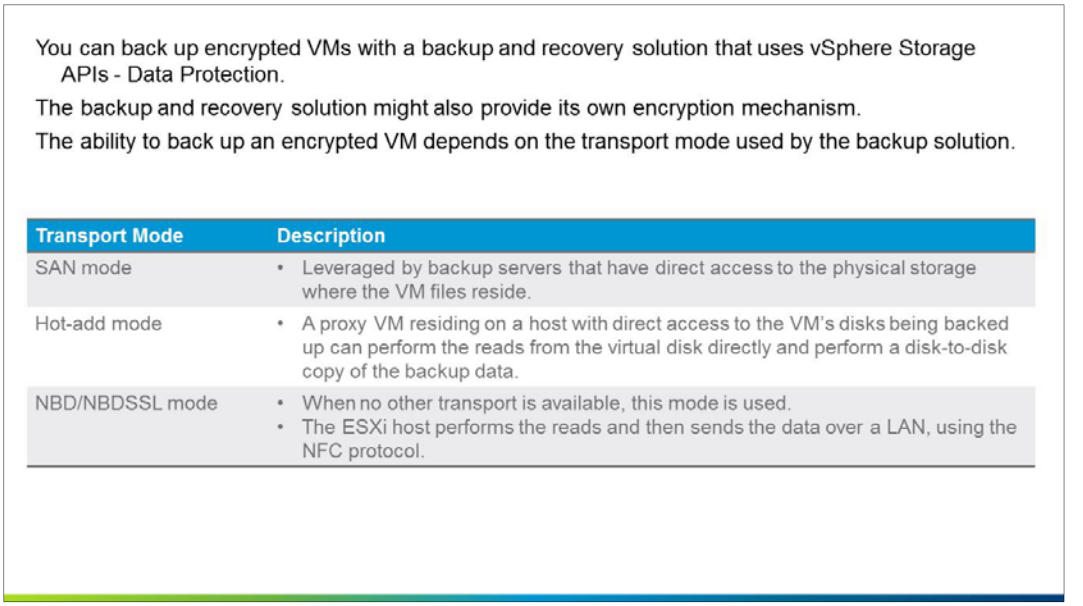
11-95 System Requirements When Cloning a VM

For encrypting, decrypting, and recrypting VMs during cloning, the following requirements must be met:

- A default KMS server must be configured in vCenter Server.
- If cloning occurs across ESXi hosts, both source and destination ESXi hosts must be vSphere 7.

https://t.me/learningnets

11-96 Backing Up Encrypted VMs (1)

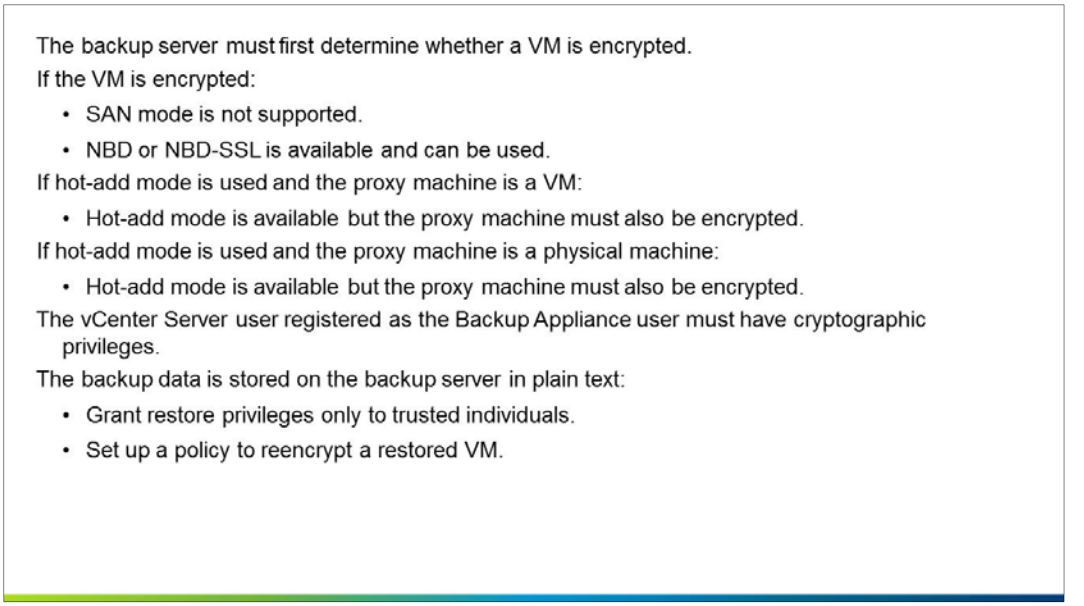


A transport mode defines the method that is used to transfer data to the backup server.

For more information about transport modes, see *Virtual Disk Development Kit (VDDK)* at https://code.vmware.com/web/sdk/7.0/vddk.

https://t.me/learningnets

11-97 Backing Up Encrypted VMs (2)



If the VMis encrypted, SAN mode is immediately ruled out as an available transport mode. The backup products that rely on this mode cannot back up encrypted VMs. The blocks are encrypted, and without vSphere assistance, the backup server cannot use the data.

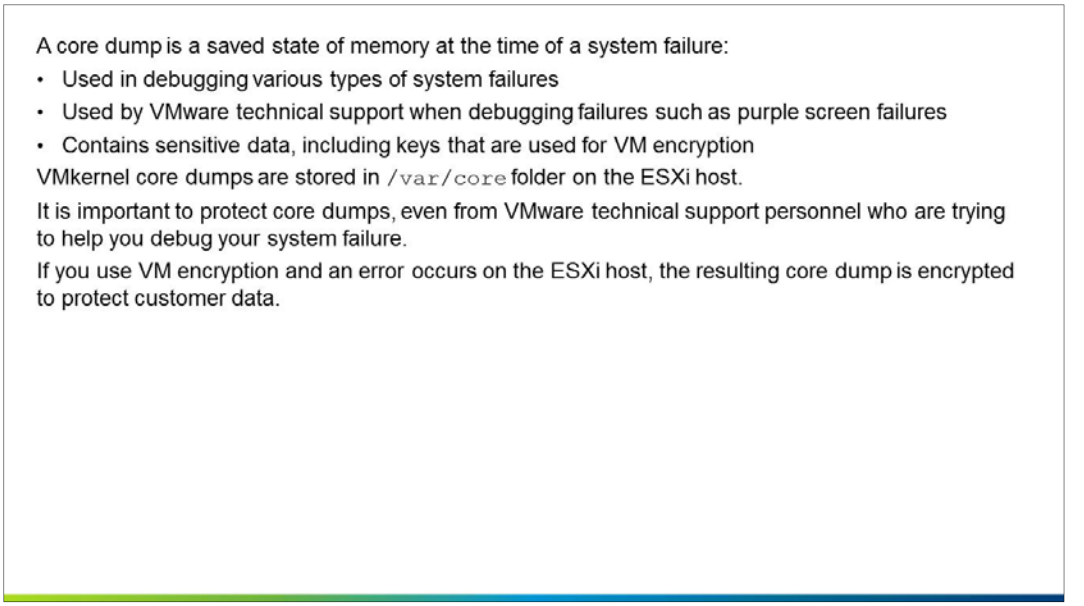
If the backup proxy is a physical machine, hot-add mode is ruled out, and NBD or NBD-SSL mode is used. You rely on the ESXi host to open the disk, decrypt the data, and send the backup data in plain text to the backup server.

If the user with which you register the backup appliance to vCenter Server does not have cryptographic privileges, the disk cannot be decrypted 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. In this way, a VM can be restored as a plain-text machine. Privileges to restore VMs should be granted only to trusted individuals, most likely users with cryptographic privileges, so that the VM can be immediately re-encrypted after it is restored.

https://t.me/learningnets

11-98 About Encrypted Core Dumps



A core dump is useful when you must debug system failures, especially purple screen failures. If the VMkernel fails, this failure 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 fails, the host creates an encrypted core dump that is stored in the **`/var/core`** folder on the ESXi host. Core dumps that are included in the **`vm-support`** package are also encrypted.

The core dump is encrypted using the host key.

https://t.me/learningnets

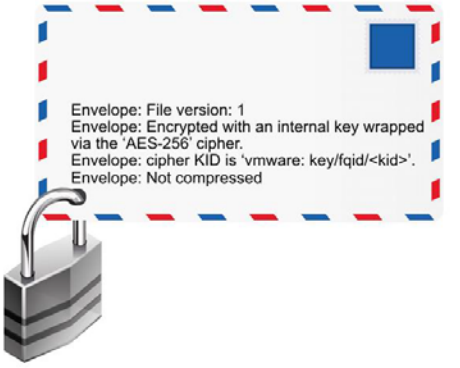
11-99 Core Dump Encryption

Core dumps are encrypted as follows:

1. An internal one-time key (AES-256) is generated by the host.
2. The core dump is placed in a digital envelope.
3. The envelope and core dump are encrypted using the internal key.
4. 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, you must use the host key to open the envelope.

Because the host key is available only on the host, VMware technical support cannot access the core dump to help you debug a system failure.



Normally, the host key should be available only on the host or cluster where the original host resides. Therefore, an encrypted core dump is not useful to a VMware technical support person who is trying to help you debug a system failure.

The **`vm-support`** command collects core dumps as is. That is, the host key is required to read them.

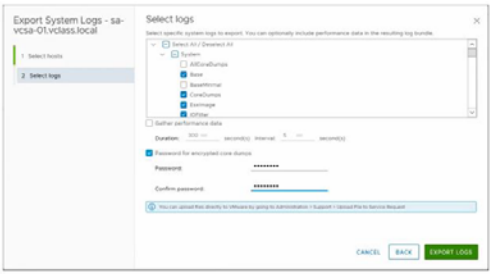
Digital envelope refers to an encryption methodology that uses two keys. vSphere encrypts and decrypts core dumps using an internal one-time key and the ESXi host key.

https://t.me/learningnets

11-100 Providing Passwords for Encrypted Core Dumps

With the vSphere 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.



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 technical 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

11-101 VM Encryption Events and Alarms

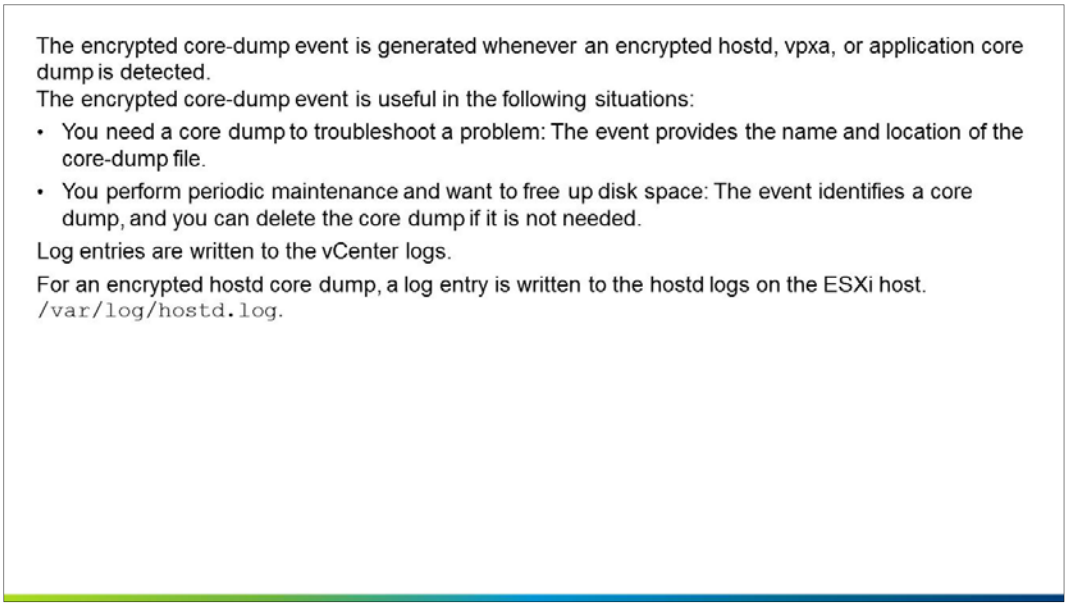
vSphere 7 logs warnings and errors that provide information about the running state of vCenter Server and the VM encryption environment:

- Warning event:
  - Encrypted core dump exists.
- Critical events:
  - Insufficient disk space to perform the cryptographic operation
  - Key creation failed on the KMS cluster.
  - KMS server is unreachable.

These events help you find possible issues.

https://t.me/learningnets

11-102 Warning: Encrypted Core Dump Exists (1)

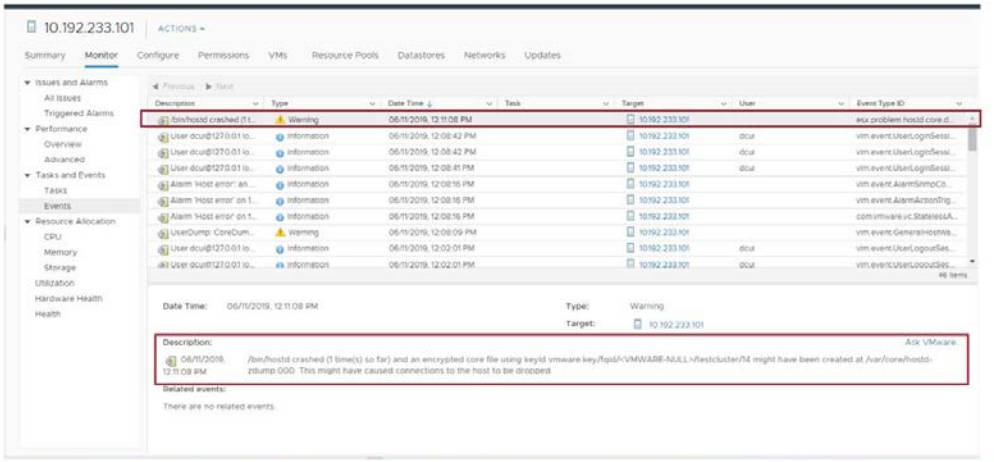


As with vSphere 6.x, you can decrypt or recrypt an encrypted core dump on your ESXi host by using the `crypto-util` CLI.

https://t.me/learningnets

11-103 Warning: Encrypted Core Dump Exists (2)

vCenter Server generates and captures warning events when an encrypted hostd core dump exists. The message is written to the hostd logs.



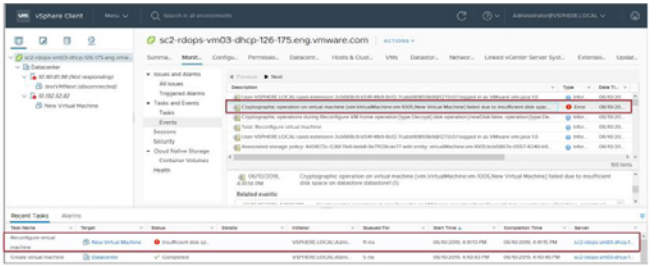
https://t.me/learningnets

11-104 Critical Event: Insufficient Disk Space

A critical event is posted when encryption fails because of insufficient disk space.

If disk space is low, and you try to encrypt or decrypt a powered-off VM, this event might occur. Twice the amount of disk space is required temporarily to complete encrypting or decrypting operations.

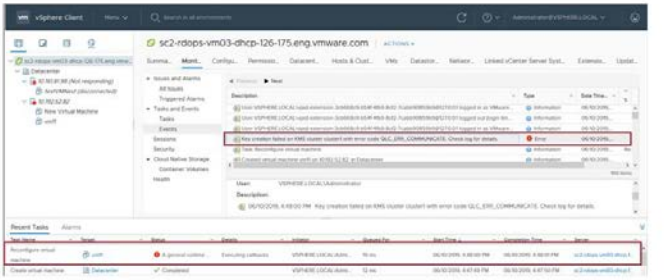
Locate the datastore where the encrypted VM resides and monitor its disk space.



https://t.me/learningnets

11-105 Critical Event: Key Generation Failed

A critical event is posted when the KMS fails to generate a new key for vCenter Server. The event is also captured in the vCenter Server logs.



Search for Key creation failed in the `/var/log/vmware/vpxd/vpxd.log` file.

Verify that the KMS is powered on and operating properly.

Verify that the KMS and vCenter Server can communicate with each other.

https://t.me/learningnets

11-106 Critical Event: KMS Unreachable

The KMS Unreachable critical event is posted when vCenter Server detects that its connection with the KMS is down.

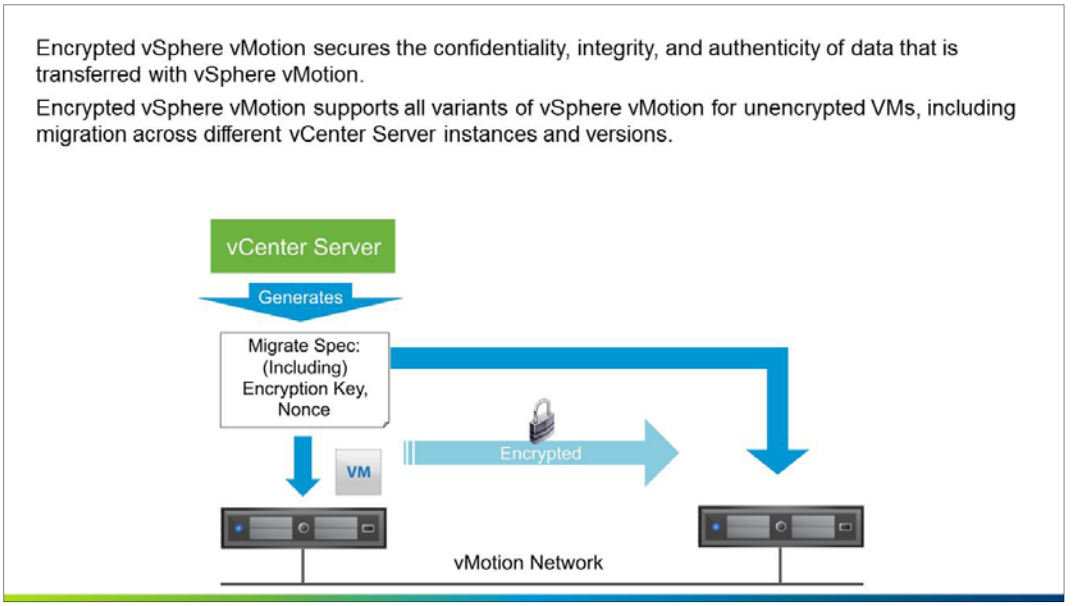
If the connection is down, the KMS health status alarm is triggered.

If the KMS health status alarm is triggered, take the following actions:

- Check vCenter Server logs for related messages.
- Verify that the KMS server is not down or powered off.
- Check the connection between the KMS server and vCenter Server:
  - If a proxy is required to reach your KMS, check that it is still reachable.
  - Verify that you are using the correct IP addresses or FQDNs.

https://t.me/learningnets

11-107 About Encrypted vSphere vMotion



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 pass the required keys to their local ESXi hosts, and these hosts communicate directly using these keys.

For encrypted VMs, the VM's memory is encrypted over the vMotion network. The disks are already encrypted and transmitted over the vMotion network as is (encrypted).

For nonencrypted VMs, the VM's memory is encrypted over the vMotion network, but the disks are transmitted as is (nonencrypted).

https://t.me/learningnets

For VMs that are encrypted, migration with vSphere vMotion always uses encrypted vSphere vMotion. You cannot turn off encrypted vSphere vMotion for encrypted VMs.

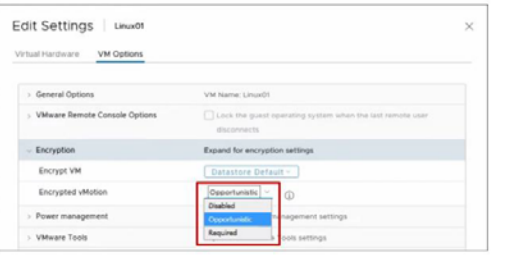
https://t.me/learningnets

11-108 Configuring Encrypted vSphere vMotion

vSphere vMotion always uses encryption when migrating encrypted VMs.

For VMs that are not encrypted, you can edit the VM'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, the migration is not allowed.



https://t.me/learningnets

11-109 Encrypted vSphere vMotion Requirements

Encrypted vSphere vMotion requires vSphere 6.5 or later.
When choosing encrypted vSphere vMotion settings, consider the following descriptions:

- 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 instance 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

11-110 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe the components of a VM encryption architecture
- Manage encrypted VMs
- Encrypt core dumps
- Use encryption with instant clones
- Migrate Encrypted VMs
- List VM encryption events and alarms
- Enable encrypted vSphere vMotion

https://t.me/learningnets

11-111 Key Points

- When defining a role, use 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, VMs, and the vSphere network.
- VMware CA provisions each ESXi host and each vCenter Server service with certificates that are signed by VMware CA.
- VM encryption cannot operate without vCenter Server because vCenter Server communicates with the KMS and manages keys that are used for encryption.
- If you use VM encryption and the ESXi host fails, the resulting core dump is encrypted to protect customer data.
- Encrypted vSphere vMotion is available for all VMs, whether or not the VMs are encrypted.

Questions?

https://t.me/learningnets