DCUCD

Data Center Unified Computing Design

Volume 3

Version 4.0

Student Guide

Text Part Number: 97-3026-01

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Table of Contents

Volume 3

Module 7

Evaluate Cisco Unified Computing System Solutions

Overview

This module evaluates the designed Cisco Unified Computing solution success criteria and assesses the return on investment (ROI) and total cost of ownership (TCO).

Objectives

Upon completing this module, you will be able to evaluate ROI and TCO using the Cisco UCS ROI tool. This includes the ability to meet these objectives:

- Identify design success criteria
- Describe how to evaluate design and solution ROI and TCO

Lesson 1

Evaluating Solution Design

Overview

This lesson identifies design success criteria.

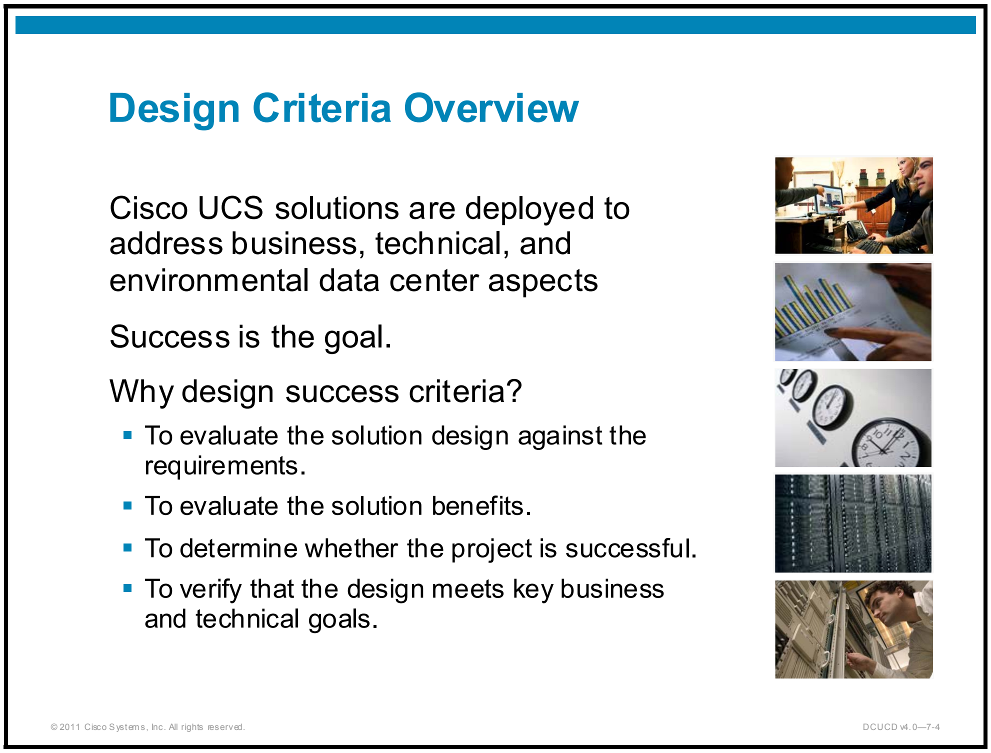
Objectives

Upon completing this lesson, you will be able to identify design success criteria. This ability includes being able to meet this objective:

- Recognize design success criteria

Evaluating Design Success Criteria

This topic describes design success criteria.

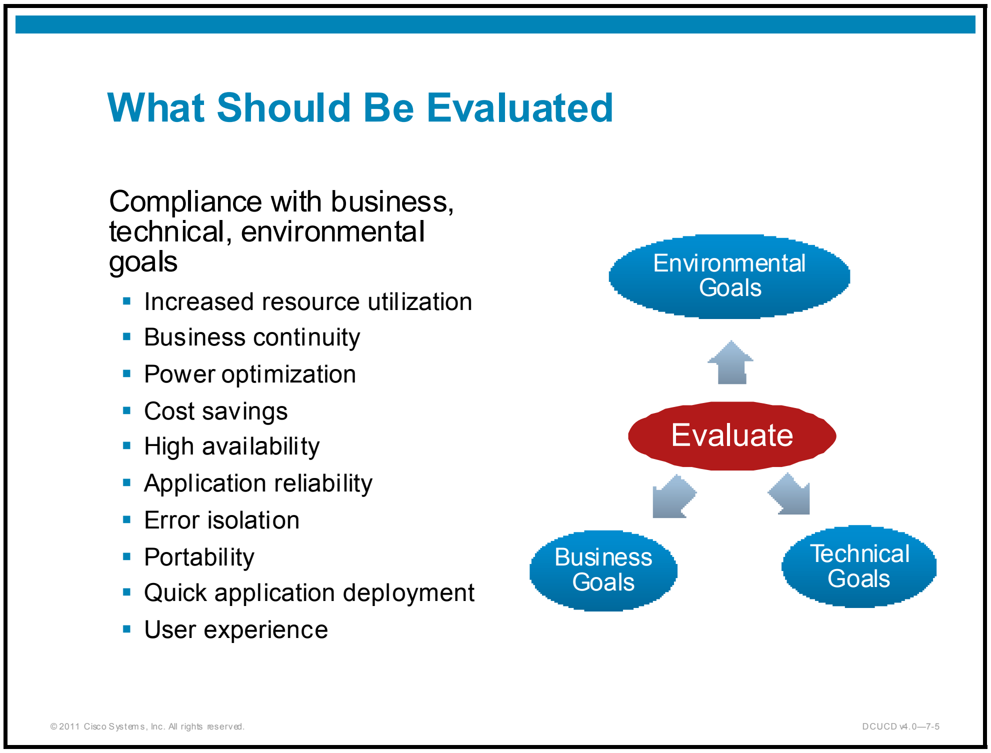


The Cisco Data Center Unified Computing System solution is used to address business, technical, and environmental aspects.

Success is an important goal to strive for. A Cisco Data Center UCS project is no exception. Due to the different perceptions that can exist about what constitutes success, it may sometimes be difficult to tell whether a project is successful. Time, cost, and quality are often used as criteria for evaluating the success of a project.

To confirm that solution expectations have been met, a project design should be evaluated against certain success criteria. The design success criteria are used to answer the following basic questions:

- Does the solution design fulfill the requirements that have been established for the project?
- Does the solution create benefits according to a business, technical, and/or environmental perspective?
- Were the key business and technical goals that were identified at the start of the project fulfilled?



What are the key aspects of the solution? Have the stated project goals been achieved? Three key questions to ask when evaluating a project are listed here:

- How has the solution met the business goals that were set before deploying the solution?
- How has the solution met the technical goals that were set before deployment?
- How has the solution met the environmental goals that were set before deployment?

The goals that the solution should fulfill must create certain benefits, which, in the case of Cisco Data Center UCS solution, are as follows:

- **More efficient resource usage:** The Cisco UCS solution provides the capability for IT organizations to ensure that resources will be available and accessible for applications.
- **Error isolation:** The Cisco UCS solution serves as a safeguard to provide security to the system against different possible disruptions or faults that can be a consequence of unavoidable circumstances, such as operating system failure.
- **Increased overall security:** Although it seems different that Cisco UCS designates users and applications into various VMs, the solution actually increases security levels among interrelated and diverse segments. In other words, using the Cisco UCS solution makes security better even when using VMs that are sharing the same physical hardware.
- **Quick provisioning:** Rather than using troublesome procedures for storage installation and management, the Cisco UCS solution offers the capability to create new VMs instantly without requiring physical servers. If properly designed, the solution also shortens the time that is needed to set up storage and data management systems. What used to take weeks can now be accomplished in several minutes.
- **Portability and mobility:** The use of intangible equipment and virtual resources allows effortless movement of VMs from one physical server to another.

Using Design Criteria

Compare computing infrastructure, not servers

- Illustrate computing platform-level TCO and ROI
- Illustrate how Cisco UCS architecture affects TCO and ROI

Cisco UCS solution simplifies design results in these ways:

- Fewer overall components, less server sprawl
- More reliable, flexible, available, and scalable infrastructure
- Easier to manage and troubleshoot infrastructure

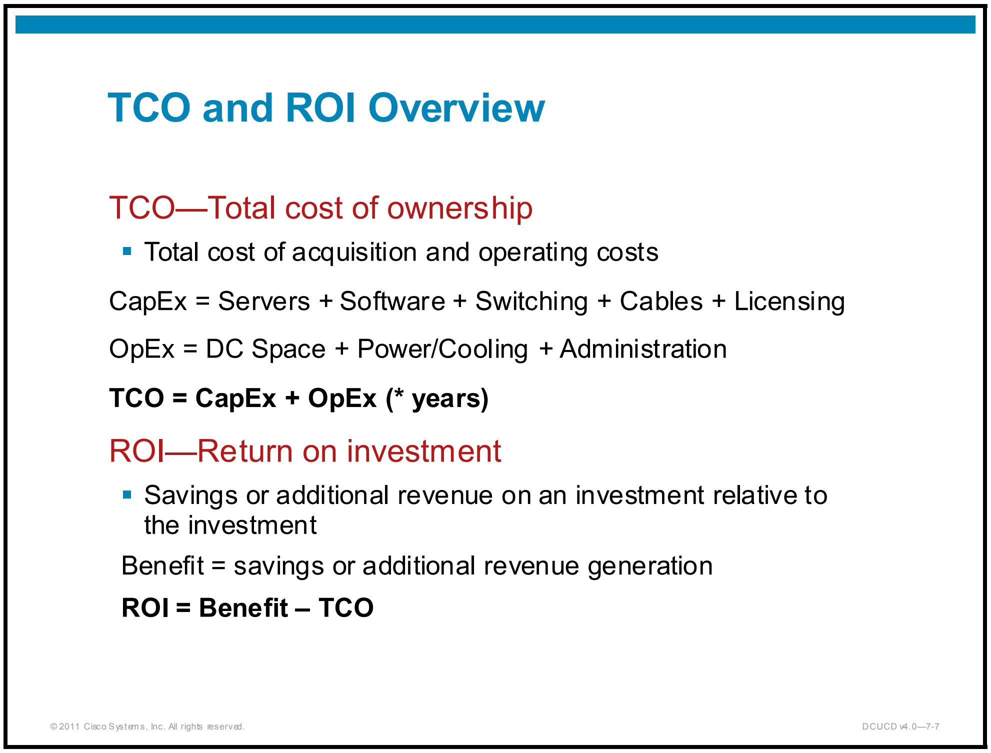
 DCUCD v4.0—7-6

Evaluating the solution and, therefore, the design success criteria requires information about the existing environment as well as the new solution.

In the following example, some aspects and criteria of design success are evaluated manually—that is, without the use of specialized tools—in order to illustrate how many aspects influence that design success. Keep in mind that you should not simply compare servers, but evaluate the complete solution—hardware and software aspects, physical aspects, and management and staff effort.

Although comparing server-to-server infrastructures will typically indicate that the Cisco UCS solution gives financial benefits, comparing all the aspects gives accurate results and further emphasizes the benefits that the customer receives from the proposed solution.

Today, you can use tools (such as the Cisco UCS ROI tool) that help you determine the design success and quantify the solution.



TCO

Total cost of ownership (TCO) analysis includes the total cost of acquisition and operating costs. A TCO analysis is used to gauge the viability of any capital investment. An enterprise may use it as a product or process comparison tool.

Analyzing TCO helps IT to reduce perception of risk and helps to gain a proper understanding of what the solution will cost.

TCO comprises capital expenditures (CapEx) and operating expenditures (OpEx), which are measured for all the following years in which the solution will be used.

CapEx

These are expenditures that create future benefits. A capital expenditure is incurred when a business spends money either to buy fixed assets or to add to the value of an existing fixed asset with a useful life that extends beyond the taxable year. CapEx are used by a company to acquire or upgrade physical assets such as equipment, property, or industrial buildings.

OpEx

Operating expenditures are the ongoing costs of running a product, business, or system.

ROI

Return on investment (ROI) is the ratio of money gained or lost (whether realized or unrealized) on an investment relative to the amount of money invested. It is the difference between the solution benefit and TCO, where the benefit is the savings and/or additional revenue that the solution brings over time.

Example—Existing Solution

	Count
Servers	1000
Storage	40
Network (cable/port)	3000
Racks	200
Power drops	400

	Value
Server deployment time (hours)	40h
Recovery time (hours)	20h
Average CPU utilization (%)	<10%
Average memory utilization (%)	20%
Average storage utilization (%)	25%

	Eth	FC	Total
Adapters	2000	2000	4000
Switches	65	6	71
Cables	3130	2172	5302
Management	1000	0	1000

 DCUCD v4.0—7-8

The example in the figure evaluates success criteria, based on the following numbers:

- 1000 servers are deployed.
- 40 storage devices are deployed.
- 3000 network ports and cables are used.
- 200 cabinet racks are used.

For the LAN and SAN networks, the following numbers apply:

- Servers are equipped with 2 LAN adapters and 2 HBAs totaling 2000 Ethernet and 2000 Fibre Channel adapters.
- 65 LAN switches are used for Ethernet connectivity:
  - 63 switches are used in the access layer; 2 are used as core switches.
- 6 Fibre Channel switches are used for SAN connectivity:
  - 2 SAN islands are used, with 3 switches in each.
  - SAN switches are interconnected with four Inter-Switch Links (ISLs).
- 40 storage devices are connected to the SAN with 4 Fibre Channel links each.
- 3130 cables are used as the Ethernet cabling infrastructure.
- 2172 cables are used as the Fibre Channel cabling infrastructure.

From an operational perspective, the following numbers apply:

- Planned downtime is 3 hours per month.
- Unplanned downtime for the past year was 18 hours.
- The customer needs 20 hours to recover in case of a failure (for example, server failure).
- It takes 40 hours to deploy a new application on a new server.

- Average server CPU utilization is less than 10 percent.
- Average server memory utilization is around 20 percent.
- Average storage utilization is around 25 percent.

Example—Deployed Solution

Network characteristics

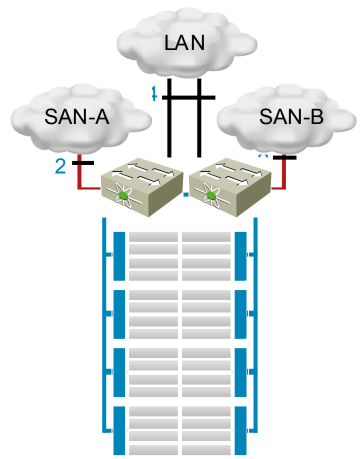
- Server connectivity—Cisco UCS 6100
- Existing SAN—Cisco MDS
- LAN connectivity—Cisco Nexus 7000
- Unified fabric with FCoE

Compute characteristics

- Cisco B-Series
- Server virtualization with VMware

Storage characteristics

- Two disk arrays

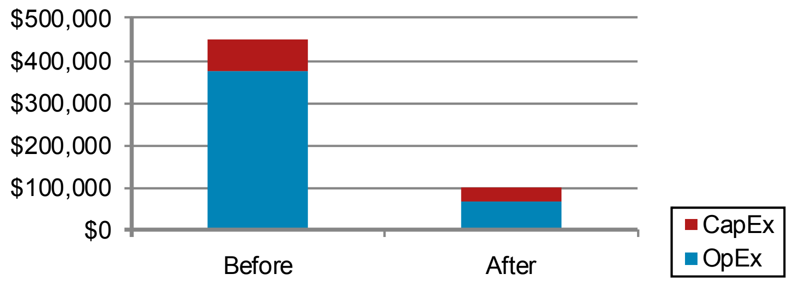


 DCUCD v4.0—7-9

You have redesigned a data center for a customer by deploying Cisco UCS, and have decided to use the following:

- For the network:
  - Cisco UCS 6140XP Fabric Interconnect cluster for server connectivity
  - Cisco Multilayer Director Switches (MDS) in the existing SAN
  - Cisco Nexus 7000 for LAN connectivity to provide high throughput using 10 Gigabit Ethernet
  - Unified fabric with Fibre Channel over Ethernet (FCoE) to simplify server connectivity
- For the server:
  - Cisco UCS 5108 chassis with blade servers
  - VMware to virtualize and consolidate servers
- For the storage component:
  - Two disk arrays

Example—Evaluating Business Goals



	Before	After
Server deployment time (hours)	40 hr	3 hr
Recovery time (hours)	20 hr	0 hr*
Average CPU utilization (%)	<10%	82%
Average memory utilization (%)	20%	85%
Average storage utilization (%)	25%	75%

*VMware High Availability automatically manages physical server failure.

Evaluating the business goals involves comparing the CapEx and OpEx before and after the Cisco UCS solution is used. The main reduction presents the OpEx, since much less power is used and much less cooling is required.

Apart from that, you are also comparing these factors:

- How much time it takes to deploy a new server
- How much time it takes to implement failover in case of a server failure
- What the average resource utilization levels are

The time that is required to deploy a server is minimized from almost two days to three hours. Recovery is instantaneous since VMware High Availability is used. In the event of server failure, VMware High Availability detects the failure and starts the VMs from the failed server on another ESX server.

The server deployment time has been reduced due to a combination of different mechanisms. First, the Cisco UCS solution speeds up several previously lengthy administrator tasks:

- Templates are used for new physical server deployments.
- Server LAN and SAN connectivity is deployed via Cisco UCS Manager, without the need to disturb the network and storage team.

Second, since VMware vSphere is used for server unified computing in case of available physical computing resources, the new server is rapidly deployed by using a server template in the form of a “gold image,” and by using VMware vSphere for server virtualization.

In viewing the average utilization levels, the numbers reveal that the Cisco UCS solution helped customers to raise average CPU utilization from less than 10 to 82 percent, average memory utilization from 20 to 85 percent, and average storage utilization from 25 to 75 percent.

Example—Evaluating Technical Goals

Quantity	Before	After
Servers	1000	80
Storage	40	2
Network (cable/port)	3000	48
Racks	200	4
Power drops	400	42

	Before	After
Planned downtime	3 hr/month	0*
Unplanned downtime	18 hr/year	10 min/year
Recovery	20 hr	0**

* VMware vMotion is used to move the VM.
**VMware High Availability automatically manages physical server failure.

 DCUCD v4.0—7-11

Next, you will evaluate the design against the stated technical goals to see whether it meets the technical design criteria. First, compare the device count before and after the Cisco UCS solution is implemented:

- The physical server count is reduced from 1000 servers to just 80 servers, yielding an average server consolidation ratio of 12.5 to 1.
- The storage devices have also been consolidated from 40 to just 2 larger disk arrays.
- The cable and port count has been reduced by a factor of 62, from 3000 to 48.
- The facility now has more spare room, since 196 racks have been freed for use, which prolongs the data center lifespan by making extra space available for growth.

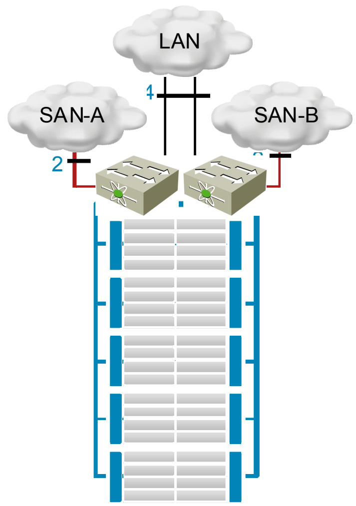
From the perspective of technical responsiveness, the Cisco UCS solution in a virtualized data center creates a reduction in the time that is required for planned, unplanned, and recovery tasks.

The planned downtime is minimized to almost zero hours since VMware vMotion is used to move the virtualized server—the VM—away from the physical server that has been selected for maintenance. Apart from that, the maintenance time that is related to server firmware (adapter, enclosure, BIOS, and so on) is also minimized with the use of central server management—the Cisco UCS Manager.

Unplanned downtime has been reduced to a minimum since multiple high availability mechanisms have been employed in combination with the network—Gateway Load Balancing Protocol (GLBP), redundant links in EtherChannel and virtual port channel (vPC), virtualized server environment where VMware High Availability is used to solve server failure, and an application-based cluster, which also manages application-related failures.

Recovery time has been minimized as a consequence of using a combination of high availability mechanisms.

Example—Evaluating Cabling Benefits



Before	Eth	FC	Total
Adapters	2000	2000	4000
Switches	65	6	71
Cables	3130	2172	5302
Management	1000	0	1000

After	Eth	FC	Total
Adapters	80	0	80
Switches	4	2	6
Cables	44	4	48
Management	2	0	0

Finally, in viewing the benefits from a cabling perspective, the passive infrastructure has significantly been reduced.

- Adapters have been reduced by a factor of 25, from 4000 to only 80. This was made possible with the use of FCoE, thus deploying converged network adapters (CNAs). Each of the 80 servers has one CNA with redundancy support.
- The number of switches has been reduced to six:
  - 2 Cisco UCS 6100 Series Fabric Interconnects used to connect the 10 blade enclosures
  - 2 Cisco Nexus 7000 Series switches used for core switches
  - 2 Cisco MDS switches used to connect the Fibre Channel-based disk arrays

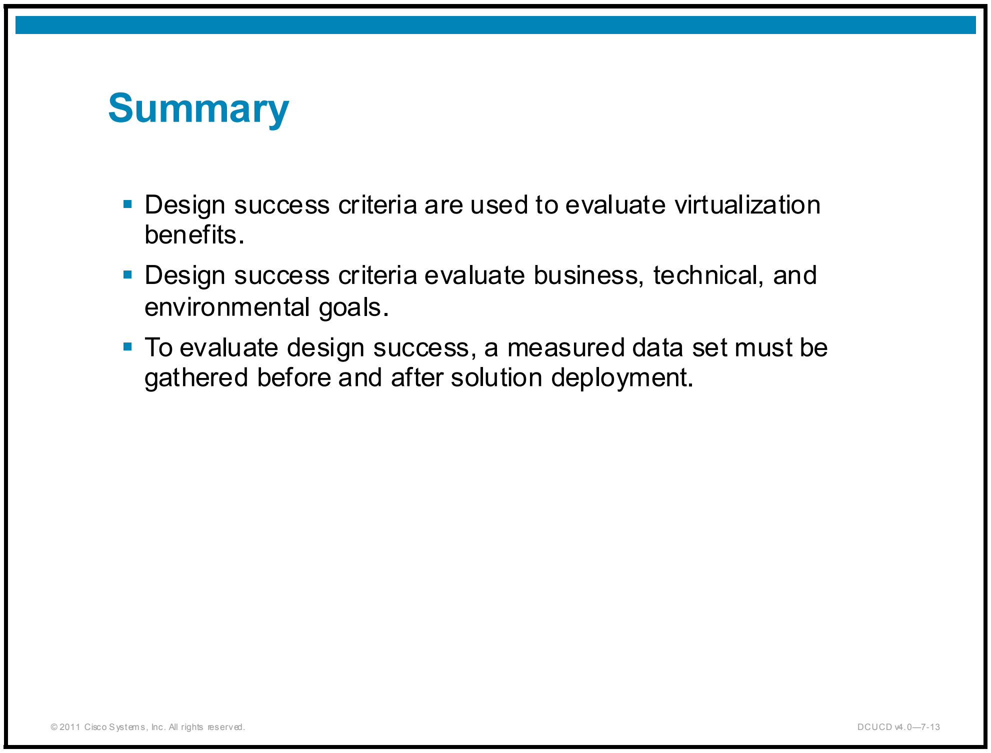
Cabling count has been significantly reduced—from 5302 to only 48 cables.

- Each of the 10 blade enclosures is connected using 2 cables to each of 2 Cisco UCS 6100 Fabric Interconnects in a cluster, totaling 40 cables for server to fabric interconnect connectivity.
- Each Cisco UCS 6100 Fabric Interconnect is connected to Cisco Nexus 7000 Series Switches using two cables. In total, four cables are used.
- The 2 disk arrays are connected to MDS switches with 16 cables—8 per disk array to connect each disk array to both SAN islands.
- The MDS Fibre Channel switches are connected to Cisco UCS 6100 Series Fabric Interconnects with two cables, each totaling four cables.

The management connectivity for the servers is embedded with a Cisco UCS 5100 Blade Enclosure to Cisco UCS 6100 cabling— no special cabling is required to manage servers.

Summary

This topic summarizes the key points that were discussed in this lesson.



Lesson 2

Determining Solution ROI and TCO

Overview

This lesson describes how to evaluate design and solution return on investment (ROI) and total cost of ownership (TCO).

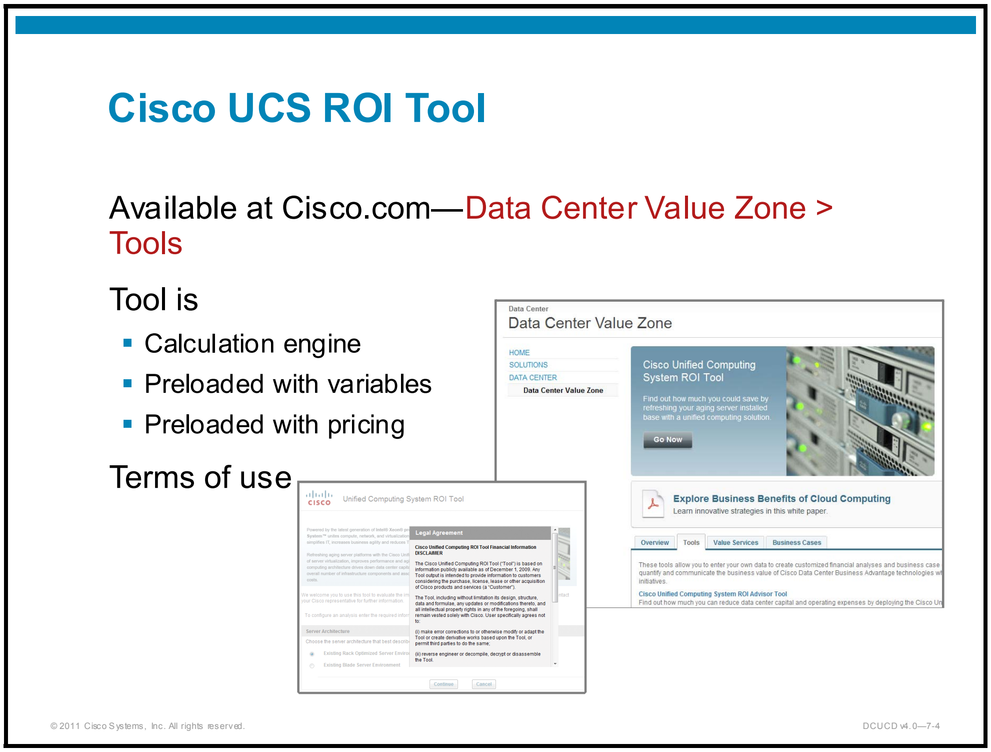
Objectives

Upon completing this lesson, you will be able to evaluate solution ROI and TCO using the Cisco UCS ROI tool. This includes the ability to meet this objective:

- Evaluate the Cisco UCS ROI tool

Using the Cisco UCS ROI Tool

This topic evaluates the Cisco UCS ROI tool.



The Cisco UCS ROI tool compares the capital and operating costs of maintaining an existing computing infrastructure (server and network environment) with the cost of implementing a Cisco UCS solution (Cisco UCS and accompanying devices) over the same time period (typically five years).

The analysis measures the benefit of the Cisco UCS solution by calculating the difference in TCO between these two environments and the ROI in the proposed Cisco UCS solution—in other words, the analysis helps to discover how much data center capital and operating expenses can be reduced.

The tool delivers a comparison between existing rack-mount or blade servers and networks, and the new Cisco UCS environment with Cisco UCS B-Series blade servers or C-Series rack optimized servers and unified fabric networking.

The tool is available on the Cisco.com website on the Data Center Value Zone under the Tools section:

- http://www.cisco.com/en/US/partner/solutions/ns340/ns517/ns224/tools/data_center_value_zone.html
- http://www.cisco.com/survey/exit.html?https://express.salire.com/Go/Cisco/Unified-Computing-ROI-Tool.aspx

There are two options for the Advisor tool—the express version, which is publicly available, and the advanced version that is available to Cisco system engineers and specialized partners.

Tool Purpose

The ROI tool is helpful in the following ways:

- Sophisticated calculation engine
- Preloaded with directional pricing
- Preloaded with directional data center environmental variables

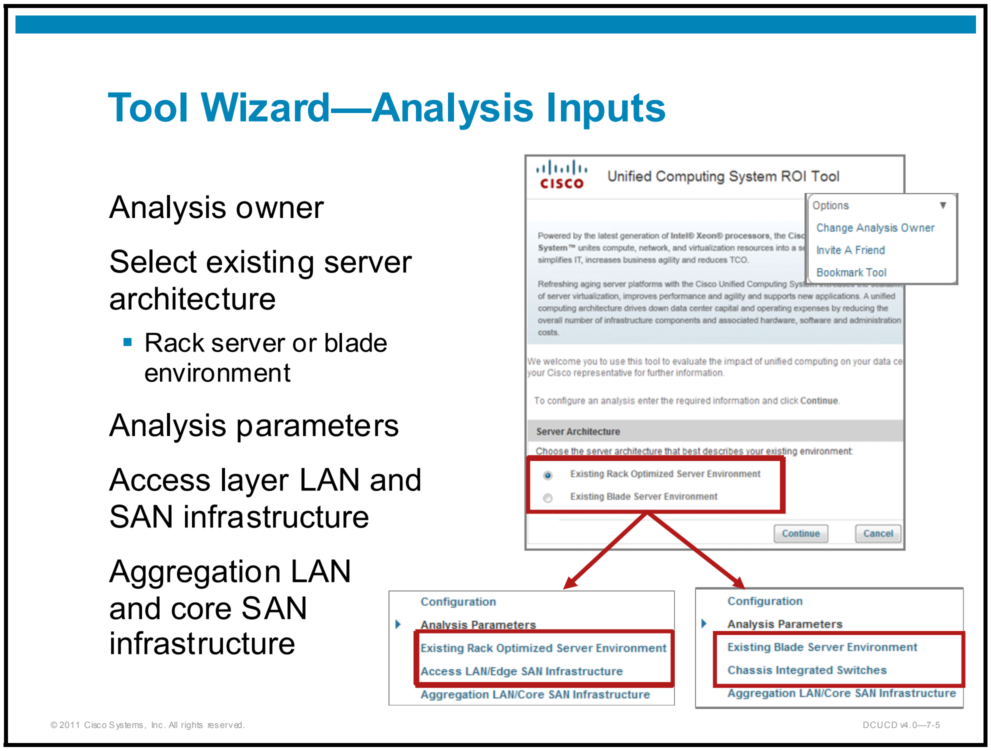
The tool is not designed for the following:

- A guide to UCS Architecture and Data Center design—the solution design requires understanding and planning each and every component of the solution.
- An up-to-the-minute competitive pricing configurator. Accurate scenario modeling requires fresh pricing inputs by user.
- It is not representative of the unique environment of the customer.

Terms of Use

The Cisco Unified Computing ROI Tool is based on information publicly available as of December 1, 2009. Any Tool output is intended to provide information to customers considering the purchase, license, lease, or other acquisition of Cisco products and services.

The Tool is designed to provide general guidance only; it is not intended to provide business, purchasing, legal, accounting, tax or other professional advice. The Tool is not a substitute for the professional or business judgment of the customer.



The tool estimates the financial effect of upgrading an existing server and associated network infrastructure (current state) to a Cisco UCS infrastructure (new state). The financial analysis model projects cost structures for the current and new state over the chosen period of the analysis (between 3 and 10 years).

Depending on the current-state environment, the tool allows a comparison between existing rack-optimized servers or blade servers and the associated network and a new state that is based on either of the following:

- Cisco UCS with B-Series blade servers and integrated unified fabric
- Cisco UCS with C-Series rack-mount servers and Cisco Nexus unified fabric

Current State

IT departments need to support new business applications and growth of existing workloads and data stores, while maintaining service levels to end users and containing overall costs. Today, compute and storage requirements continue to grow at an unprecedented rate, while more than 70 percent of IT budget is spent simply to maintain and manage existing infrastructure.

The current state of infrastructure is amplifying these challenges. In most cases, data center environments are still assembled from individual compute, network, and storage network components. Administrators spend significant amounts of time manually accomplishing basic integration tasks rather than focusing on more strategic, proactive initiatives. One of the main initiatives that data center managers are undertaking to address this situation is consolidation of physical servers through server virtualization. While server virtualization can have immediate impact by enabling the improved utilization of server hardware, current infrastructures are generally not well-suited to easy deployment, expansion, and migration of virtualized workloads to meet changing business needs.

New State

The Cisco UCS solution is a next-generation data center platform that unites compute, network, storage access, and virtualization into a cohesive system designed to reduce TCO and increase business agility. The system integrates a low-latency, lossless 10 Gigabit Ethernet unified network fabric with enterprise-class, x86-architecture servers. The system is an integrated, scalable, multichassis platform in which all resources participate in a unified management domain. IT departments have the choice of unified computing blade or rack form factors.

Primary Assumptions

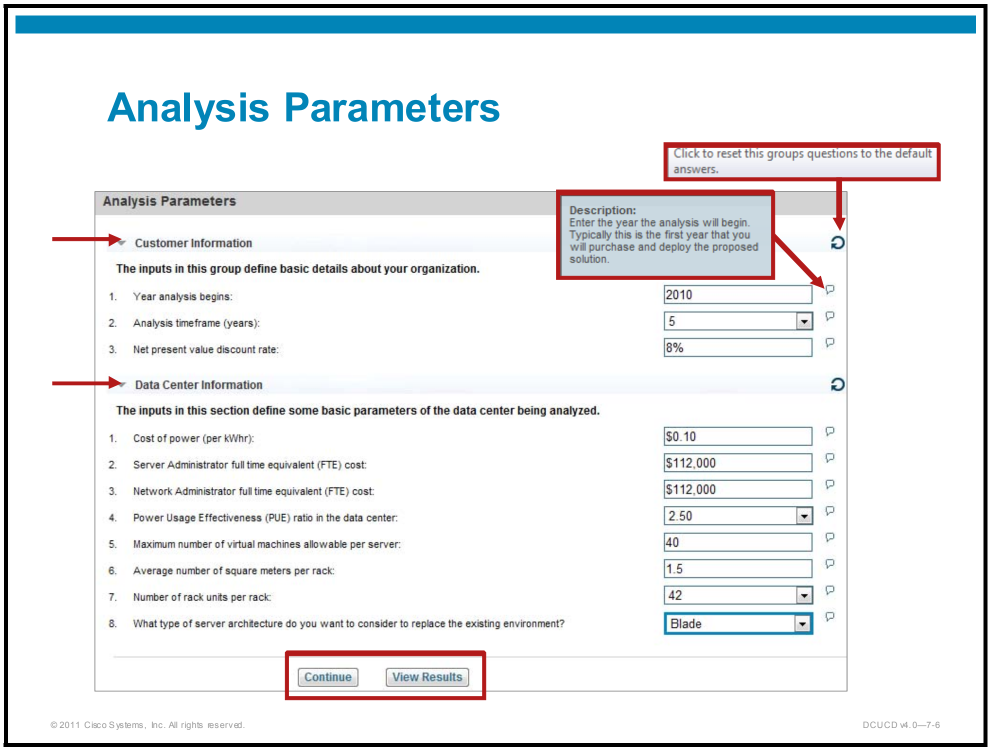
These are the primary assumptions:

- The server consolidation ratio depends on the anticipated compute improvements that will be gained by moving from your current server CPU to the Cisco UCS solution based on the latest Intel Xeon processors. Your consolidation ratio may vary based on workload and may be higher in workloads that are memory-intensive.
- The model assumes that your current-state network infrastructure uses a 1 Gigabit Ethernet LAN and 1/2/4-Gb Fibre Channel SAN connections.
- The model assumes that you will move your server network to a Cisco 10-Gb unified fabric infrastructure for all Ethernet and Fibre Channel traffic. Review the financial results to determine the associated costs of your compute and I/O components to more clearly understand the cost impacts of each.
- The current-state network infrastructure defaults assume the use of Cisco Catalyst 6500 Series Switches and Cisco MDS 9500 Family end-of-row Fibre Channel switches.
- The default analysis assumes that your end-state compute environment will be a chassis-based Cisco UCS B-Series unified computing blade server environment. You can edit the assumptions if you want the model to use a Cisco UCS C-Series unified computing rack-server environment as the end state.
- The financial analysis is a cash-flow-based model and does not take into account metrics such as depreciation, taxes, and amortization.

Analysis Inputs

When using the tool, the analysis is provided immediately upon choosing the type of existing environment, that is, either rack-optimized or blade server. To tailor the analysis to better reflect the customer environment, the following analysis inputs can be changed from the preloaded values:

- Analysis owner
- Existing server architecture type:
  - Rack-optimized server infrastructure
  - Blade server infrastructure
- Analysis parameters
- Access-layer LAN and SAN infrastructure that can be either:
  - Access LAN or edge SAN infrastructure, in the case of rack-optimized existing environment
  - Chassis integrated switches for blade server environment
- Aggregation LAN or core SAN infrastructure in both cases



Once the existing environment type is selected, the Analysis Parameters section can be updated from the default parameters. This requires some input from the customer in order to get more accurate analysis results. For any input that needs to be entered, you can verify the meaning by hovering the mouse cursor over the balloon at the field where the data needs to be entered. The values can also be reset to the default values by clicking the curved arrow above the section or beside the field where the value is entered..

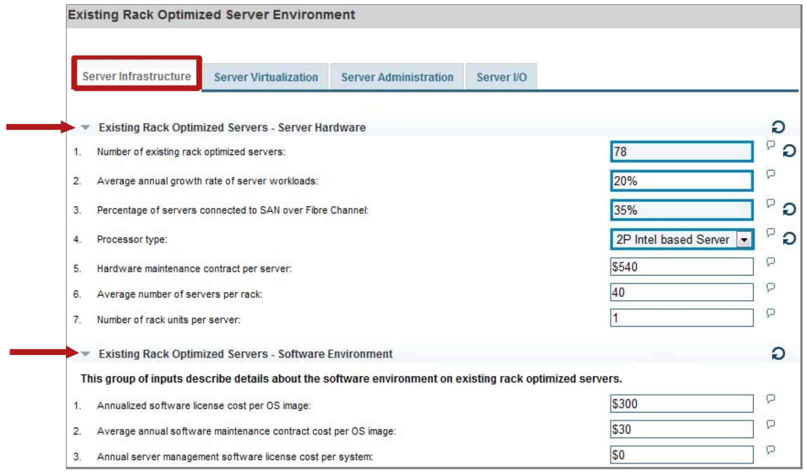
The following parameters can be entered for the existing environment:

- **Customer information:** Information about the customer organization
  - Year of the analysis
  - Analysis timeframe (default is five years)
  - Net present value discount rate
- **Data center information:**
  - Cost of power per kilowatt hour (kWh)
  - Server and network administrator full-time equivalent (FTE) cost
  - Power usage effectiveness ratio in the DC
  - Maximum number of VMs per physical host
  - Average square meters per rack
  - Rack size
  - Type of architecture to replace the existing one—rack-optimized or blade server

Once you are satisfied with the values, click the **Continue** button to enter more detailed information about the existing environment.

Example 1—Rack Environment

Identify existing rack-optimized server environment



 DCUCD v4.0—7-7

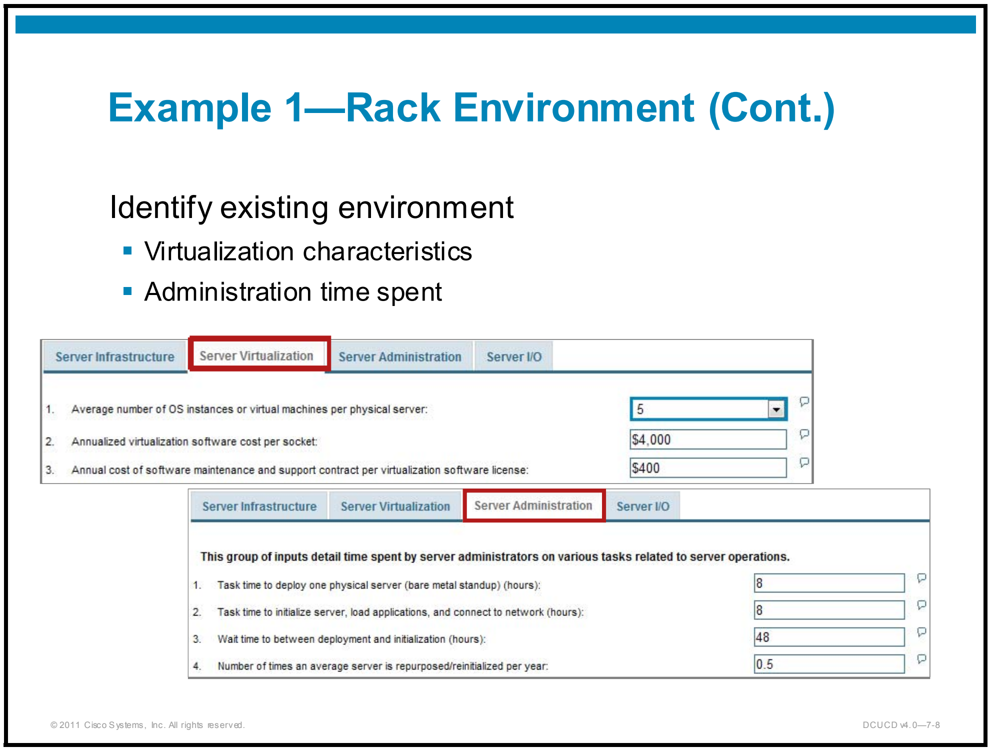
The more detailed existing environment infrastructure information is divided into different sections. Depending on the environment type (rack-optimized or blade server), some sections may be different.

In the example above, the existing environment is of the rack-optimized infrastructure type, and thus the sections that are described here must be entered.

Server Infrastructure

In this section, the designer enters the information that describes the server environment from the hardware and software perspectives. The section is further subdivided into the following subsections:

- Server Hardware:
  - Number of servers
  - Average annual growth rate
  - Percentage of servers that are connected to Fibre Channel SAN
  - CPU type
  - Hardware maintenance cost per server
  - Average quantity of servers per rack
  - Server size in rack units
- Software Environment:
  - Annual cost of software license per operating system image
  - Annual cost of software maintenance contract per operating system image
  - Annual cost of server management software license per system



Server Virtualization

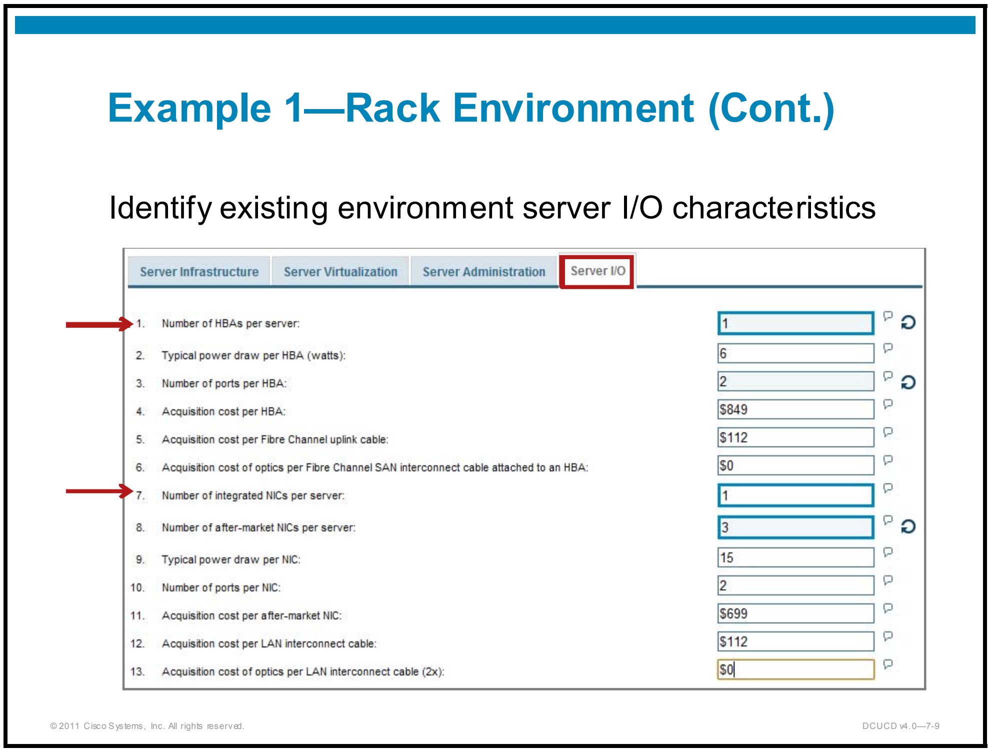
This section is used to enter the existing environment virtualization characteristics if server virtualization is already used:

- Average number of operating system instances or VMs per physical server
- Annual virtualization software licenses cost per socket
- Annual virtualization software license maintenance and support cost

Server Administration

This section is used to enter information about the management and administration effort that is related to the existing server environment:

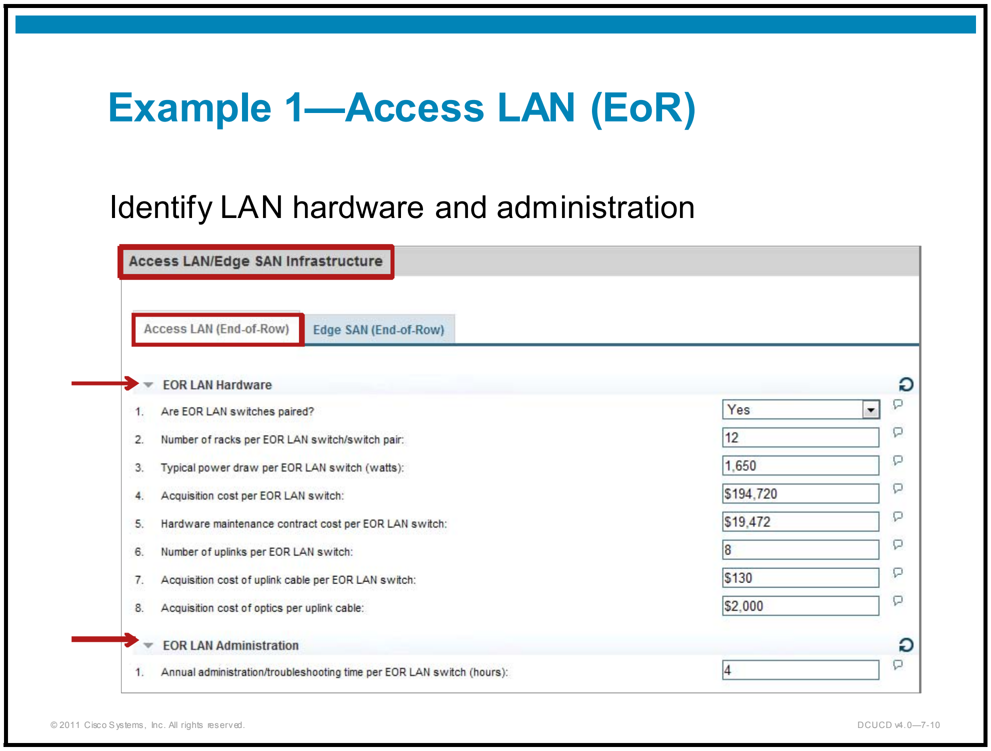
- Time that is required to install bare-metal physical server
- Time that is required to initialize server, load applications, and connect the server to the network
- Idle time between deployment and initialization
- Number of times a server is repurposed or reinitialized per year



Server I/O

This section is used to enter information about the existing server environment LAN and SAN attachment.

- The SAN parameters:
  - HBA quantity per server
  - Port quantity per HBA
  - HBA acquisition cost
  - Fibre Channel uplink cable acquisition cost
  - Acquisition cost of Fibre Channel optics
- The LAN parameters:
  - Integrated NICs quantity
  - Additional NICs quantity
  - Power draw per NIC
  - NIC port quantity
  - Additional NIC acquisition cost
  - LAN interconnect cable and optics cost.

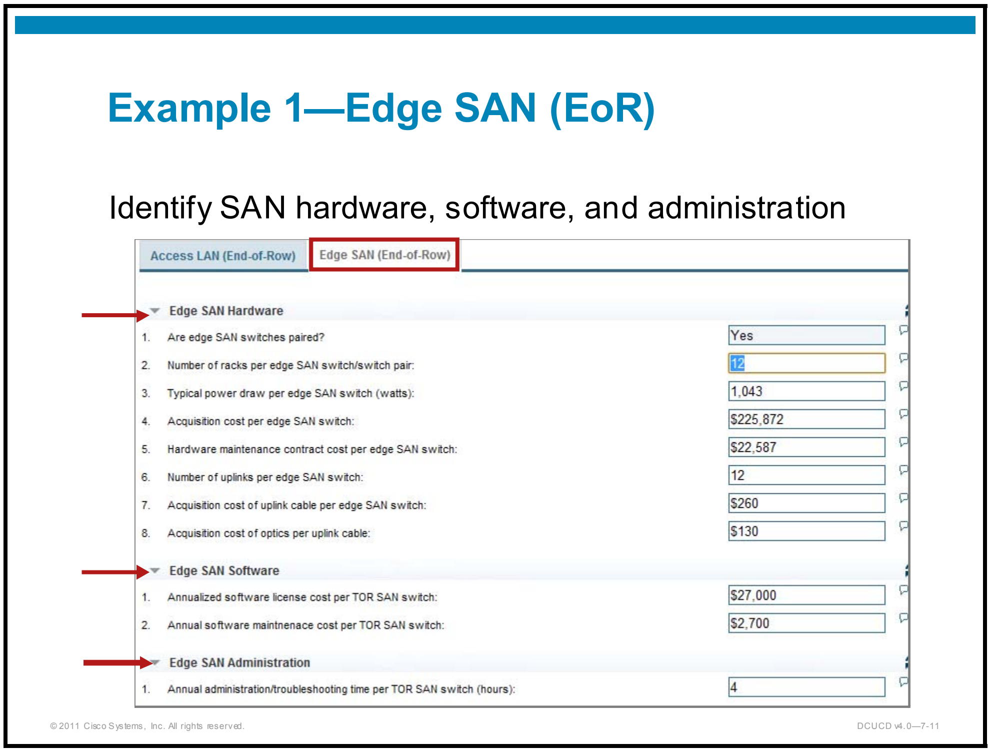


The next section is used to enter information about the existing access LAN and edge SAN infrastructures in the case of rack-optimized server infrastructure.

Access LAN (End-of-Row)

This section presumes that end-of-row (EoR) access LAN equipment is used to attach the rack-optimized servers. The following information can be changed from the defaults:

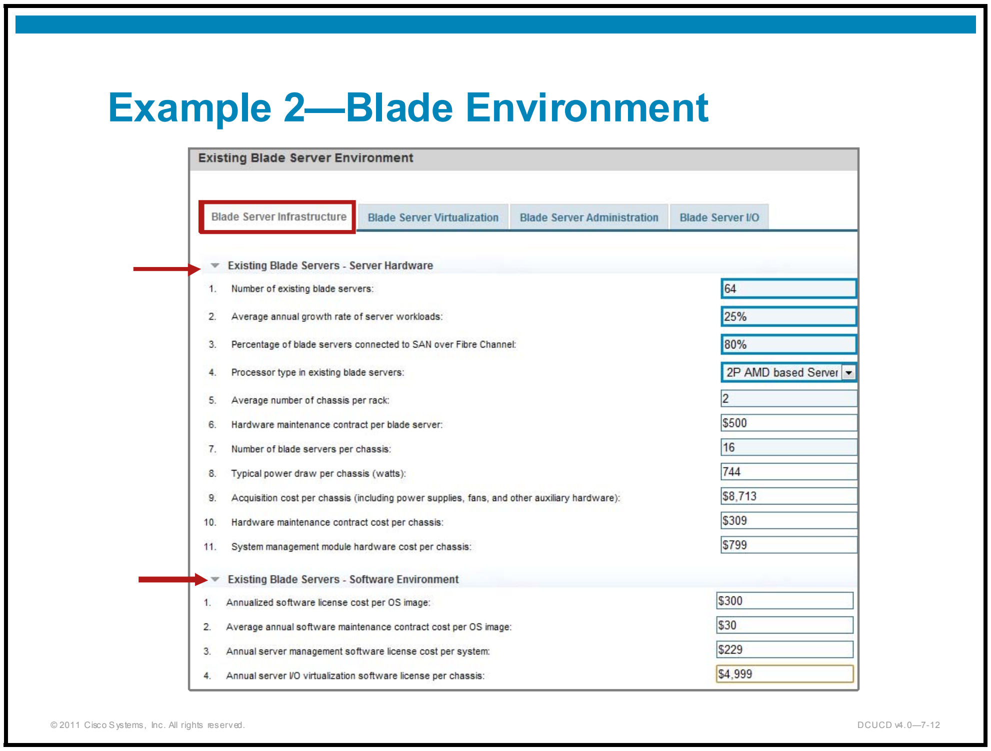
- LAN hardware:
  - Whether the EoR LAN switches are paired in redundant configuration (that is, dual physical fabric design)
  - Number of EoR LAN switches or pairs
  - Typical power draw per switch
  - All the costs that are related to the hardware—acquisition cost and maintenance contract per switch, uplink cable, and optics costs
  - Per LAN switch quantity of uplinks
- LAN administration, which details the annual administrative effort per switch



Edge SAN (End-of-Row)

This section presumes that EoR edge SAN equipment is used to attach the rack-optimized servers. The following information can be changed from the defaults:

- SAN hardware:
  - Whether the EoR SAN switches are paired in redundant configuration (that is, dual physical fabric design)
  - Number of racks per EoR SAN switch or pair
  - Typical power draw per switch
  - All the costs that are related to the hardware—acquisition cost and maintenance contract per switch, uplink cable, and optics costs
  - Per SAN switch quantity of uplinks
- SAN software, which describes the costs of software licenses and maintenance per switch
- SAN administration, which details the annual administrative effort per switch



In the example, when the existing environment is of the server blade infrastructure type, some sections and required information differ from the information that needs to be entered for the rack-optimized environment.

Blade Server Infrastructure

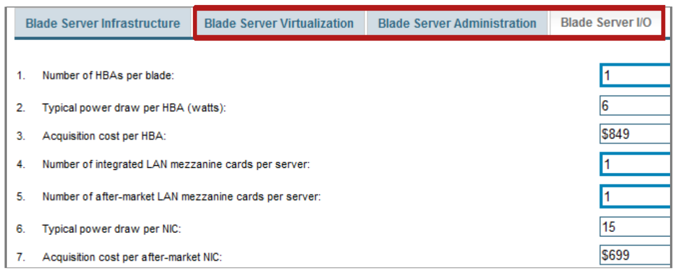
In this section, the designer enters the information that describes the server environment from the hardware and software perspective. The section is further subdivided into the following:

- Server hardware information:
  - Number of blade servers
  - Average annual growth rate of the server workload
  - Percentage of the blade servers that are connected to Fibre Channel SAN
  - CPU type
  - Average number of chassis per rack
  - Hardware maintenance contract cost per blade server
  - Blade server quantity per chassis
  - Chassis acquisition cost, which includes all the chassis components
  - Hardware maintenance contract cost per chassis
  - System management module cost per chassis
- Server software information, consisting of these annual costs:
  - Software license per operating system image
  - Software maintenance contract per operating system image
  - Server management software license per system
  - Server I/O virtualization software license cost per chassis

Example 2—Blade Environment (Cont.)

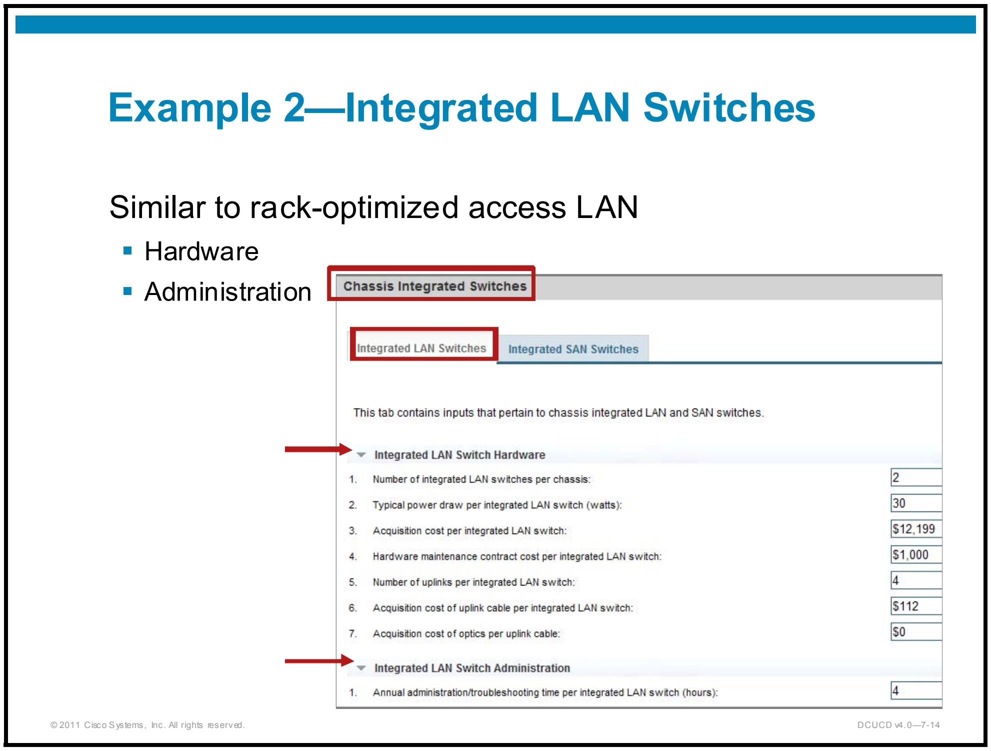
Other existing server environment characteristics

- Similar to rack-optimized
  - Virtualization
  - Administration
  - I/O



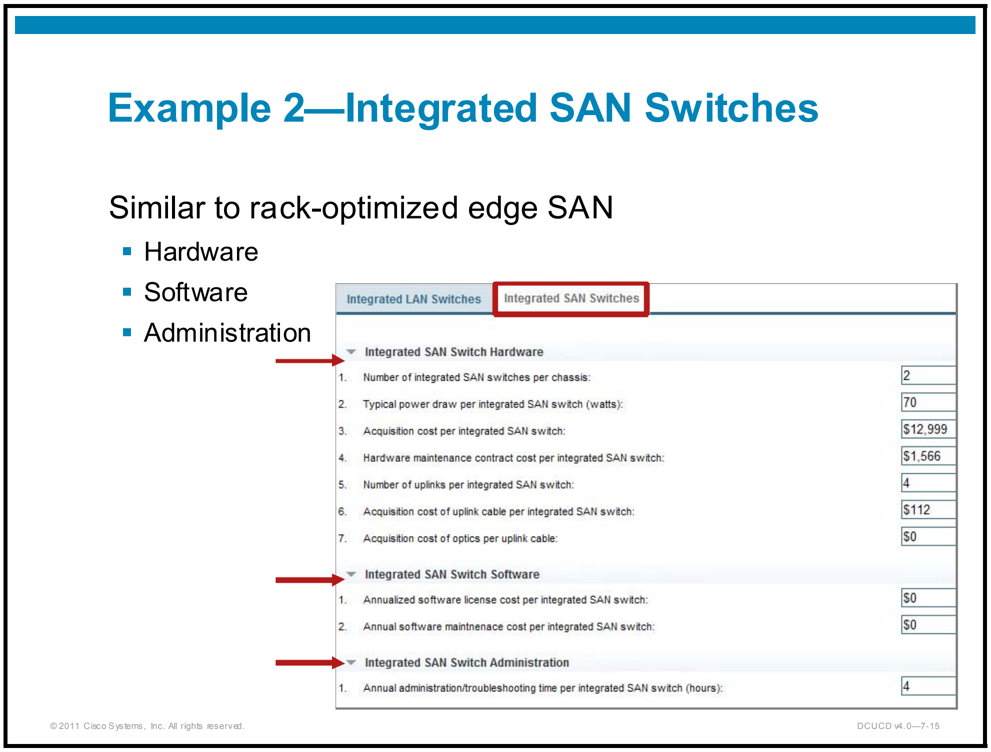
 DCUCD v4.0—7-13

The other sections that are related to blade environment—Blade Server Virtualization, Blade Server Administration, and Blade Serve I/O—are very similar to the ones for the rack-optimized environment with respect to the blade-server specifics.

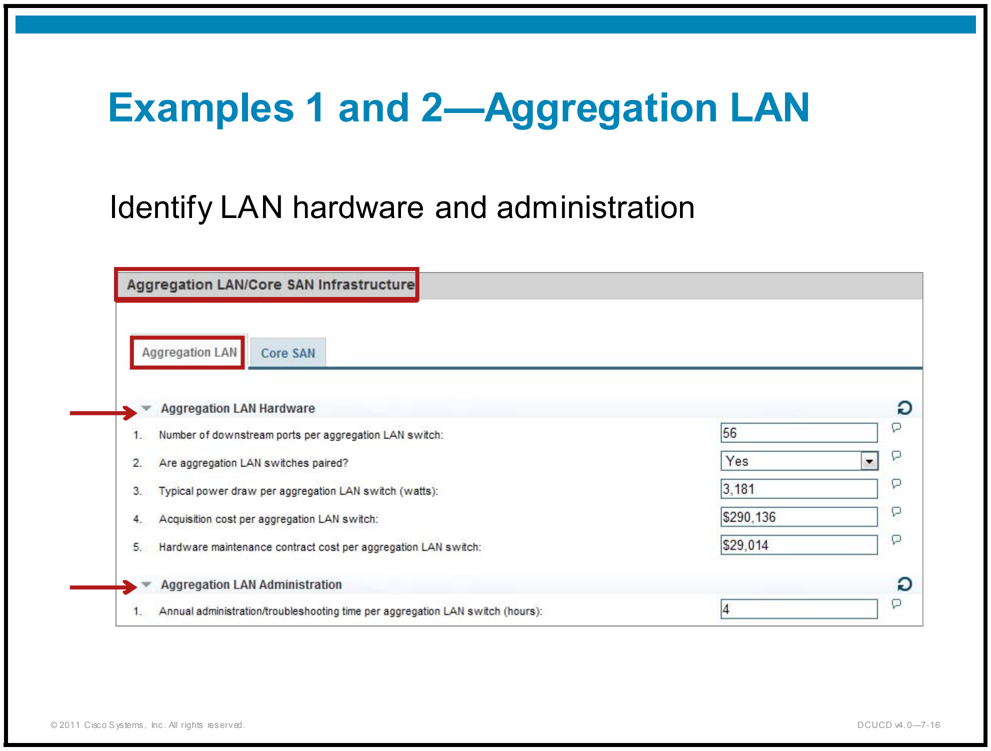


Likewise, the Chassis Integrated Switches section describes the "access layer" of attaching the blade server to the LAN and SAN, similar to the access LAN and edge SAN sections for the rack-optimized servers. The minor differences are in regard to the specifics of the integrated switches.

The Integrated LAN Switches section describes the blade server LAN attachment and related costs.



The Integrated SAN Switches section describes the blade server SAN attachment and related costs.

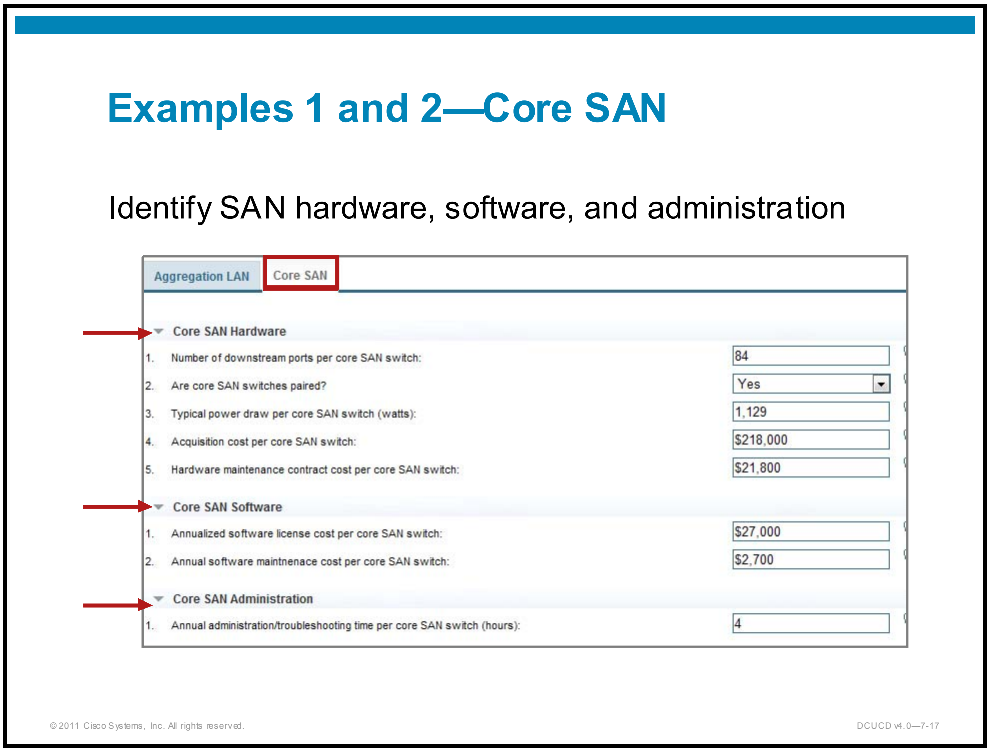


The Aggregation LAN/Core SAN Infrastructure section is the same for both types of environment, and describes how the access layer infrastructure (LAN and SAN, if existing) is attached to the upstream LAN and SAN networks. The section not only describes the hardware characteristics but also the related costs.

Aggregation LAN

This section is used to describe the LAN network to which the server access layer infrastructure is attached.

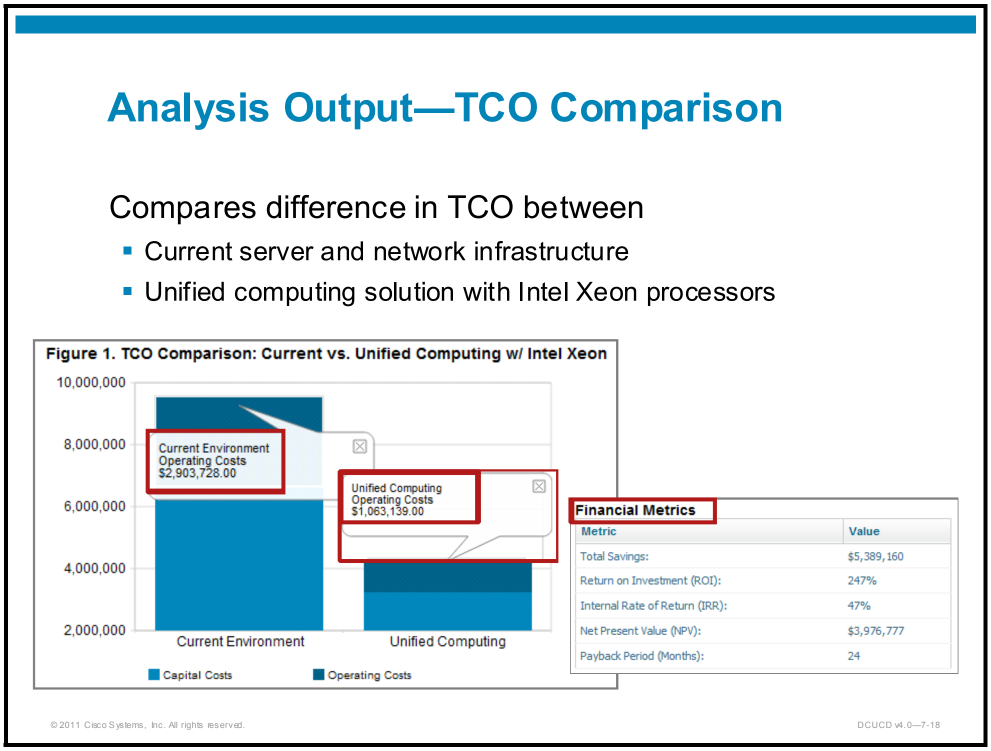
- The Aggregation LAN Hardware subsection requires the following information:
  - Number of ports per aggregation LAN switch toward the access layer
  - Whether the aggregation LAN switches are deployed in the dual physical fabric manner
  - Typical power draw per switch
  - Costs per switch—the acquisition and maintenance
- The Aggregation LAN Administration subsection describes the annual management effort per switch.



Core SAN

This section is used to describe the SAN network to which the server access layer infrastructure is attached.

- The Core SAN Hardware subsection requires the following information:
  - Number of ports per core SAN switch toward the access layer
  - Whether the core SAN switches are deployed in the dual physical fabric manner
  - Typical power draw per switch
  - Costs per switch—the acquisition and maintenance
- The Core SAN Software subsection describes the costs of software licenses and maintenance per switch.
- The Core SAN Administration subsection describes the annual management effort per switch.



Once all the parameters and characteristics are tailored to the specific customer, the analysis can be updated and the results reviewed. Note that the parameters can be changed at any time by editing them and updating the analysis.

The results of the analysis can be examined over the web or downloaded. Both sets of results consist of several sections, as follows.

Figure 1: TCO Comparison

Figure 1 of the analysis illustrates the difference in TCO between the current server and network infrastructure and a Cisco UCS solution that leverages Intel Xeon processors. Hovering the mouse over the graph gives details about the costs that are related to each solution. The table also details financial metrics such as total savings, ROI, and others.

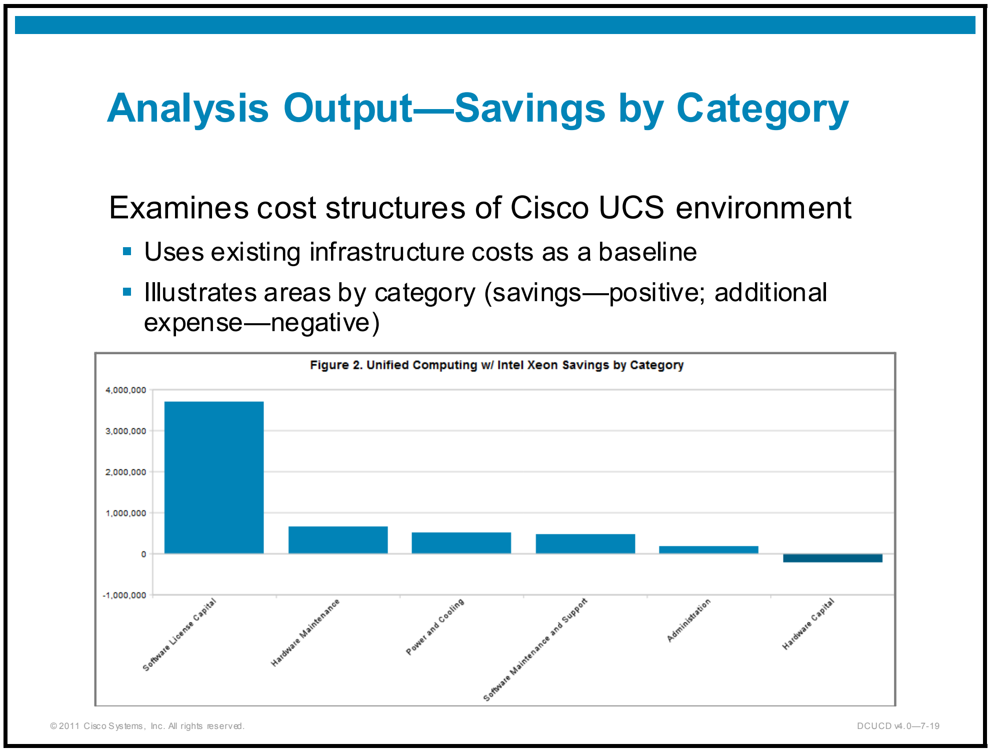


Figure 2: Savings by Category

This section examines the cost structures of a unified computing environment using the existing infrastructure costs as a baseline.

The graph is illustrated with areas of savings that are positive or negative, and organized per category.

Analysis Output—Server Quantity Comparison

Examines difference in server quantity—current vs. Cisco UCS environment

- Consolidates workloads from multiple existing servers onto fewer Cisco UCS servers

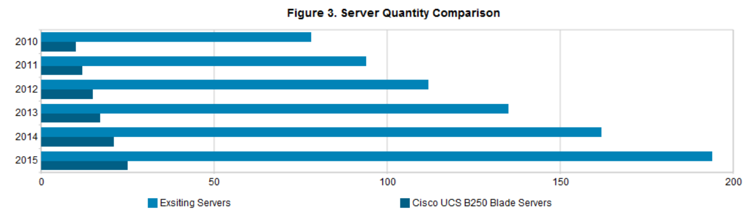


Table 2. Server Profile Comparison

Category	Existing Environment	Cisco UCS Environment
VMs per server	5	39
Server price	$6,000	$30,987
Number of sockets	2	2
Power draw (watts)	385	209

DCUCD v4.0—7-20

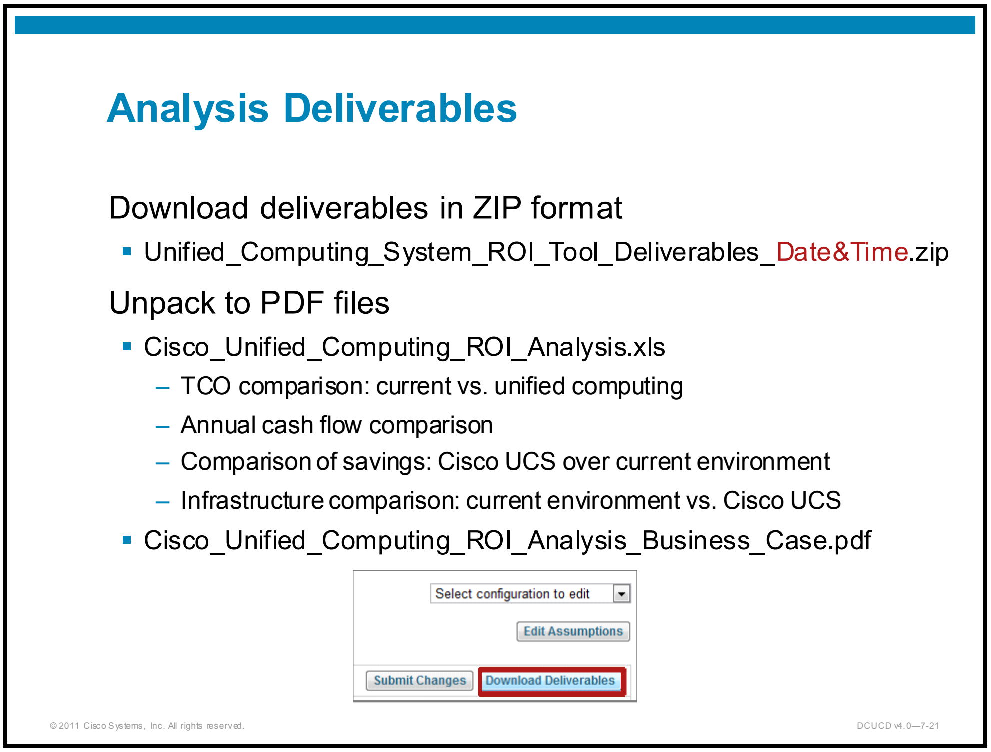
Figure 3: Server Quantity Comparison

This section illustrates the difference between the number of servers in the current environment and the number of Cisco UCS servers with Intel Xeon processors in the proposed environment.

Cisco UCS servers are based on the latest Intel Xeon processors, allowing you to consolidate workloads from multiple existing servers onto a smaller number of Cisco UCS servers.

The information details not only the difference between the solutions for the current environment, but also for the following years as specified by the analysis inputs. The example in the figure specifies the information for the next five years.

The table depicts the comparison between the two server environments, based on the information that was entered for the analysis.



As mentioned previously, the results of the analysis can be downloaded in a zipped format. The following two files are part of the archive:

- Cisco_Unified_Computing_ROI_Analysis.xls.pdf
- Cisco_Unified_Computing_ROI_Analysis_Business_Case.pdf

Analysis Document

This document examines the financial impact of replacing an existing install base of data center servers and the supporting network infrastructure with a unified computing solution from Cisco that is based on the latest Intel Xeon processors.

Implementing unified computing dramatically reduces the amount of hardware that is required to support the server workload in your environment today, while providing additional capacity and more efficient scalability of data center resources.

The results of this analysis are tailored to the organization using data provided through the tool.

This file includes the following sections (with comparison of the current environment to a Cisco UCS solution):

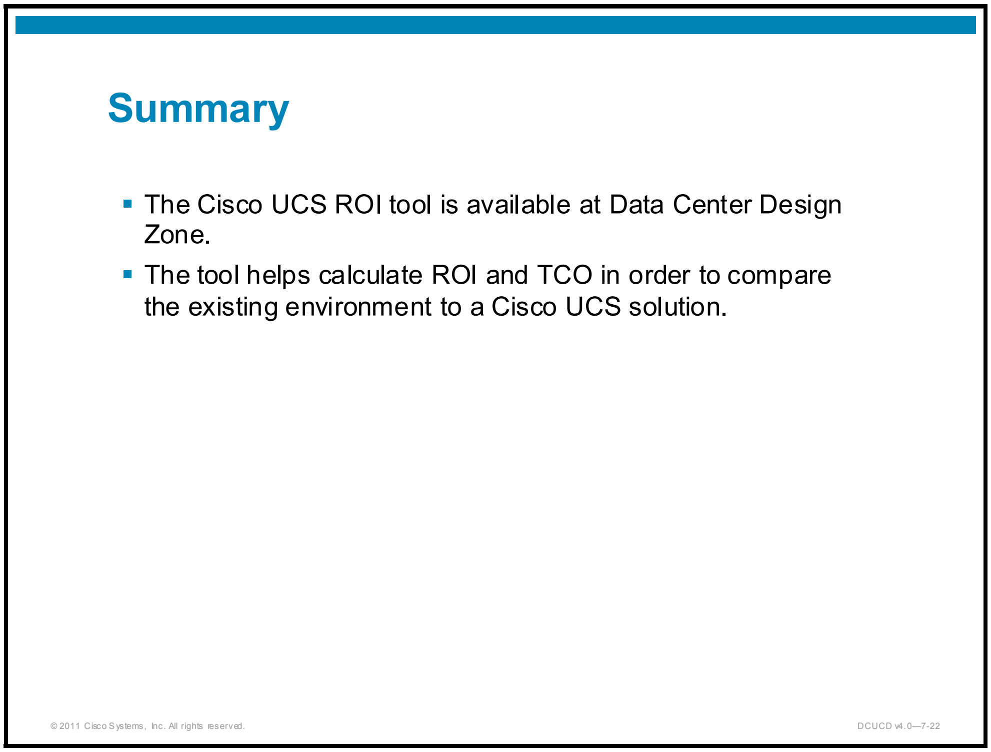
- TCO comparison of current environment and Cisco UCS solution, with financial metrics
- Annual cash flow comparison
- Comparison of savings
- Infrastructure comparison

Analysis Business Case Document

This document describes the rationale behind the Cisco UCS solution.

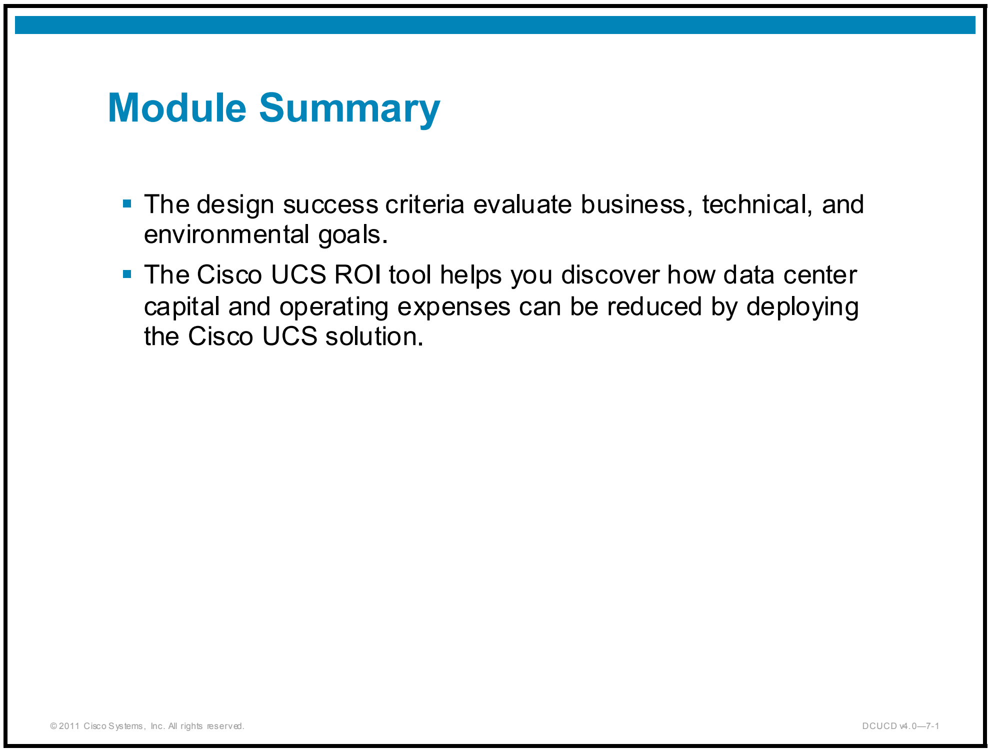
Summary

This topic summarizes the key points that were discussed in this lesson.



Module Summary

This topic summarizes the key points that were discussed in this module.



Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) Why are design success criteria used? (Source: Evaluating Solution Design)

- A) to evaluate the solution against requirements
- B) to calculate CapEx
- C) to calculate OpEx
- D) to determine ROI

Q2) Which factor is typically used to determine the amount of power consumed to cool the equipment in a data center? (Source: Evaluating Solution Design)

- A) BTU
- B) CoP
- C) ROI
- D) RU

Q3) Which two kinds of existing environments can be evaluated in the Cisco UCS ROI tool? (Choose two.) (Source: Determining Solution ROI and TCO)

- A) blade server environment
- B) mainframe environment
- C) LAN environment with integrated switches
- D) rack server environment
- E) HPC environment

Module Self-Check Answer Key

Q1)	A
Q2)	B
Q3)	A, C

Module 8

Plan Migration to Cisco Unified Computing System Solution

Overview

This module evaluates a migration plan for an existing implementation to a Cisco Unified Computing System solution.

Objectives

Upon completing this module, you will be able to design a migration plan for an existing implementation to a Cisco UCS solution. This ability includes being able to meet this objective:

- Evaluate migration plan aspects, actions, deliverables, and dependencies

Lesson 1

Designing a Migration Plan

Overview

This lesson evaluates migration plan aspects, actions, deliverables, and dependencies.

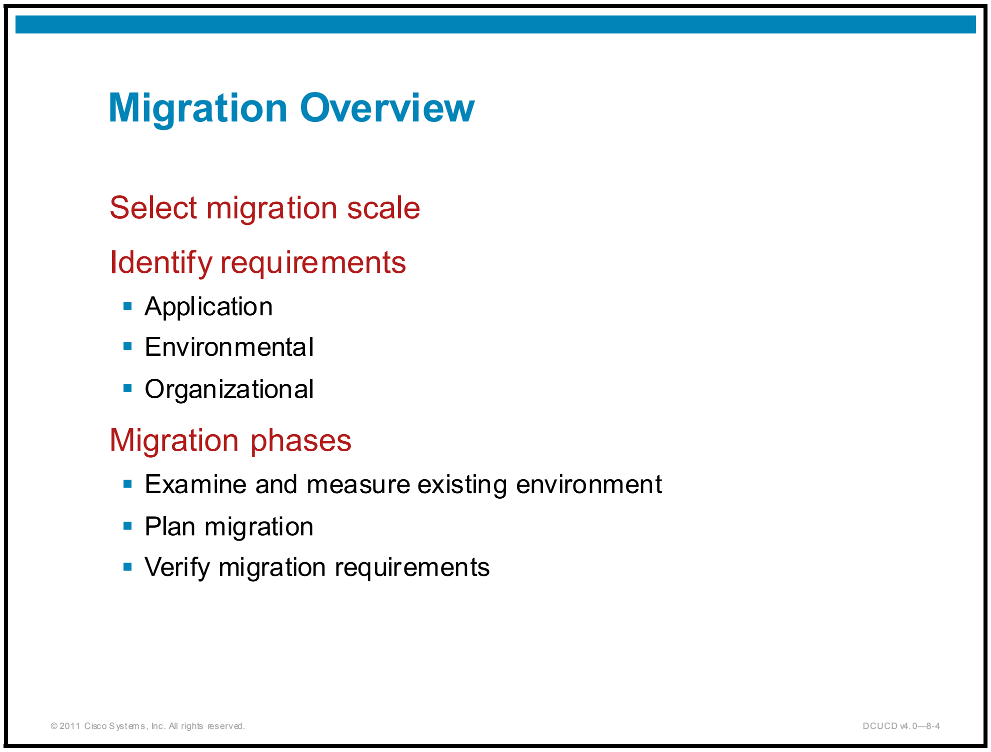
Objectives

Upon completing this lesson, you will be able to identify whether a migration to Cisco UCS should be planned. This includes the ability to meet these objectives:

- Assess migration aspects
- Propose a migration plan

Assessing Migration Aspects

This topic assesses migration aspects.

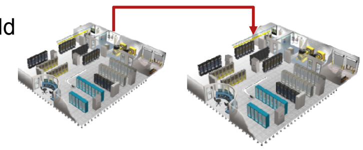


Cisco Data Center Business Advantage architectural framework encompasses technologies that address network, storage, compute, operating system, application services, management, and security aspects of the Unified Computing solution.

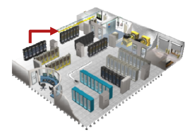
Since the migrations are different, it is important that the solution architect addresses all the aspects and that are part of the migration.

Migration Scale

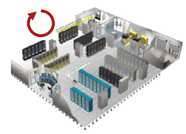
- Full migration to new site
  - Build new physical data center (replace old equipment with Cisco UCS)
  - Migrate operating systems, applications, and data



- Full migration within existing site
  - Build new logical data center within an existing physical data center space (replace old equipment with Cisco UCS)
  - Migrate applications and data



- Partial migration or redesign
  - Redesign an existing data center (add Cisco UCS to the existing deployment)
  - Migrate operating systems, applications, and data



Depending on the type of migration, which can range from simply upgrading the existing data center to include virtualization to building a physically new data center with new equipment, you can determine what migration steps are required to have the least painful migration.

The migration plan should also evaluate these aspects:

- Whether the existing server parameters (MAC, WWN, UUID) can be used and need to be migrated to new servers
- How a new Cisco UCS server infrastructure merges with the existing management tools
- Can the existing management tools for operating system, applications, and virtual infrastructure be used?
- Is any additional configuration like Serial over LAN (SoL) or Intelligent Platform Management Interface (IPMI) required for server management?
- Would adding new servers require a new management application or license since the infrastructure would go beyond the existing management limitations?
- What would happen if the management servers and services were relocated?

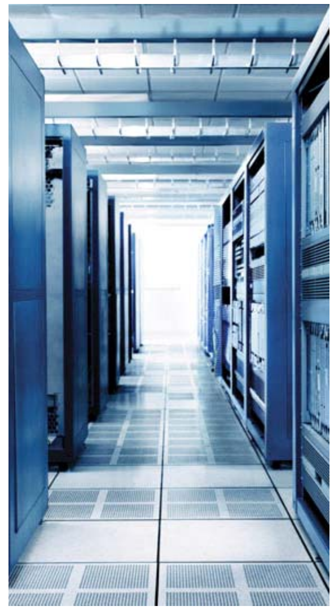
Server Migration Options

Select server migration method

P2P
- Server personality (MAC, WWN, UUID)?
- Operating system and application reinstall?

P2V
- Select the P2V conversion tool (for example, VMware vConvert)
- Prerequisite—Cisco UCS and VMware

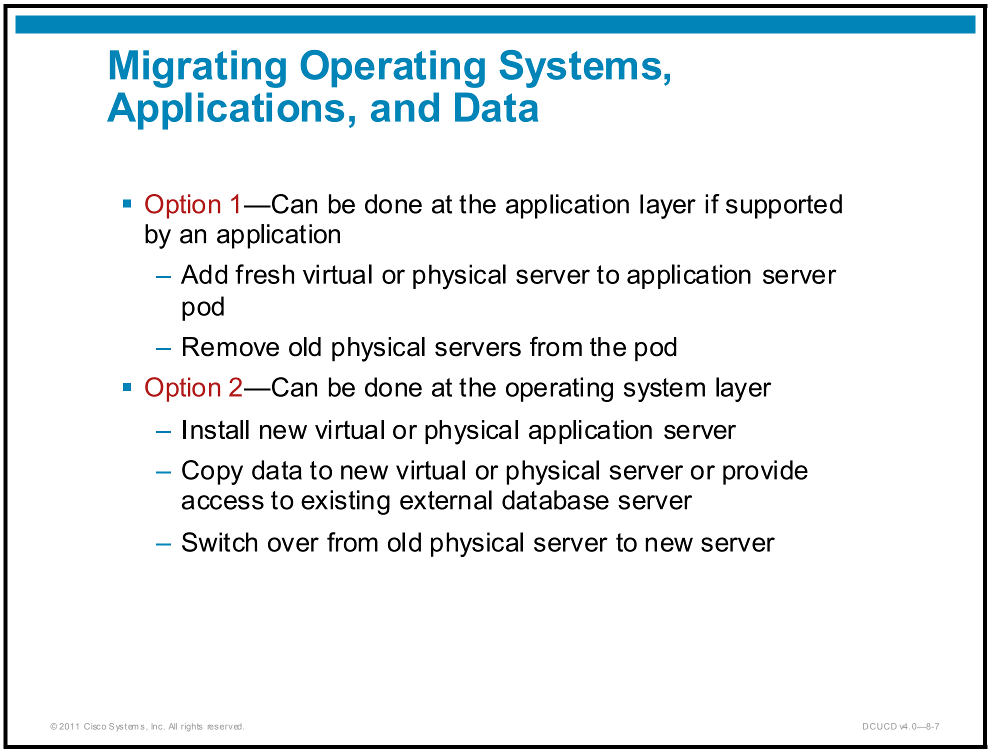
V2V
- Existing VMware infrastructure—migrate VMs using VMotion
- New VMware infrastructure— export and import VMs



 DCUCD v4.0—8-6

The server migration aspects largely depend on the selected migration method.

- **Physical to physical (P2P):** The existing servers are migrated to Cisco UCS server blades in one-to-one fashion. The migration plan must determine whether it is necessary to migrate personality identifiers like MAC, world wide name (WWN), Universally Unique ID (UUID addresses). Next, it has to evaluate whether complete operating system and application reinstall is needed or some cloning tool can be used to migrate the server operating system, applications, and data (if it resides on the local disks).
- **Physical to virtual (P2V):** The existing physical servers are converted to virtual machines (VMs). The migration plan must define the prerequisites—installed Cisco UCS clusters connected to LAN and SAN, implemented VMware infrastructure with proper management. Next, the plan must define which tools will be used for P2V migration (for example, VMware vConvert).
- **Virtual to virtual (V2V):** The existing VMs are migrated to a new VMware virtual infrastructure. The migration plan must define the prerequisites—the installed Cisco UCS clusters connected to LAN and SAN and new ESX hosts that were added to the virtual infrastructure. Second, the migration plan should also consider possibilities in regards to the virtual infrastructure:
  - — Existing virtual infrastructure will be used with new ESX hosts added to the infrastructure. The Cisco UCS clusters should be properly connected to LAN to permit communication between the management infrastructure—VMware vCenter, and to the SAN (should be connected to the same shared storage as other ESX hosts in a cluster) to allow proper operation for VMware services like vMotion, High Availability, Fault Tolerance, Disaster Recovery Solutions, and so on
  - — New virtual infrastructure will be used, thus new management services must be deployed, and ESX hosts must be added to that infrastructure. Next, the migration plan should define how the VMs from the existing virtual infrastructure would be migrated to the new one—using export and import of VMs or using cloning tools.



In general, when migrating from a traditional setup to a virtualized one, you need to migrate a service from a physical server to a virtual server.

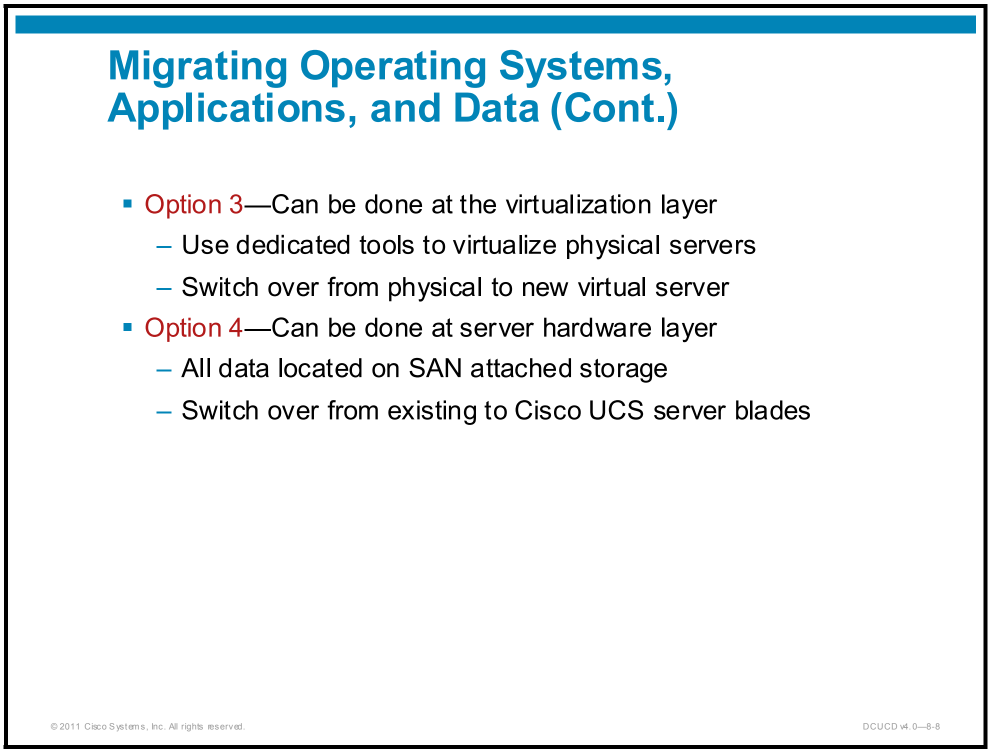
This can be done in four ways.

- **Option 1:** Use application clusters.

  Install a fresh virtual server and an application server. Put the new server into an application cluster with the existing physical server. Repeat this step for as many servers as required. Once there are enough virtual servers in the cluster, start removing the physical servers until only virtual servers remain. Decommission the physical servers.

- **Option 2:** Reinstall the servers.

  Install an operating system and an application server into a virtual server. Migrate data and configuration from the physical server to the virtual server. Switch over from the physical server to the virtual server.



- **Option 3:** Convert physical servers to virtual servers.

  Use dedicated tools that can do a hot or cold conversion of a physical server to a virtual server. Switch over from the physical server to the virtual server.

- **Option 4:** Convert physical servers to physical servers.

  Use the boot from the SAN and store all data on SAN attached storage. Switch over to Cisco UCS by booting the blade server.

Examine Existing Environment

Facility aspects

Rack space

HVAC—cooling capacity

Assess available power

- At facility
- Per rack

Evaluate cabling

- Distances

 DCUCD v4.0—8-9

The assessment should be performed before the migration commences. It is not necessary, but it is strongly advised, to conduct an audit of the existing environment.

For justifying the design, it is of utmost importance to have the relevant information upon which the design and migration are based. This includes utilization of memory and CPU resources, utilization of storage space, inventory details, historical growth report, security policies in place, high-availability mechanisms, dependencies between the data center elements (such as applications, operating system, server, storage, and so on), and limitations of the current infrastructure.

Examine Existing Environment (Cont.)

Analyze performance characteristics

- For servers, LAN, SAN, applications
- CPU utilization level, memory, disk space, bandwidth
- Application responsiveness

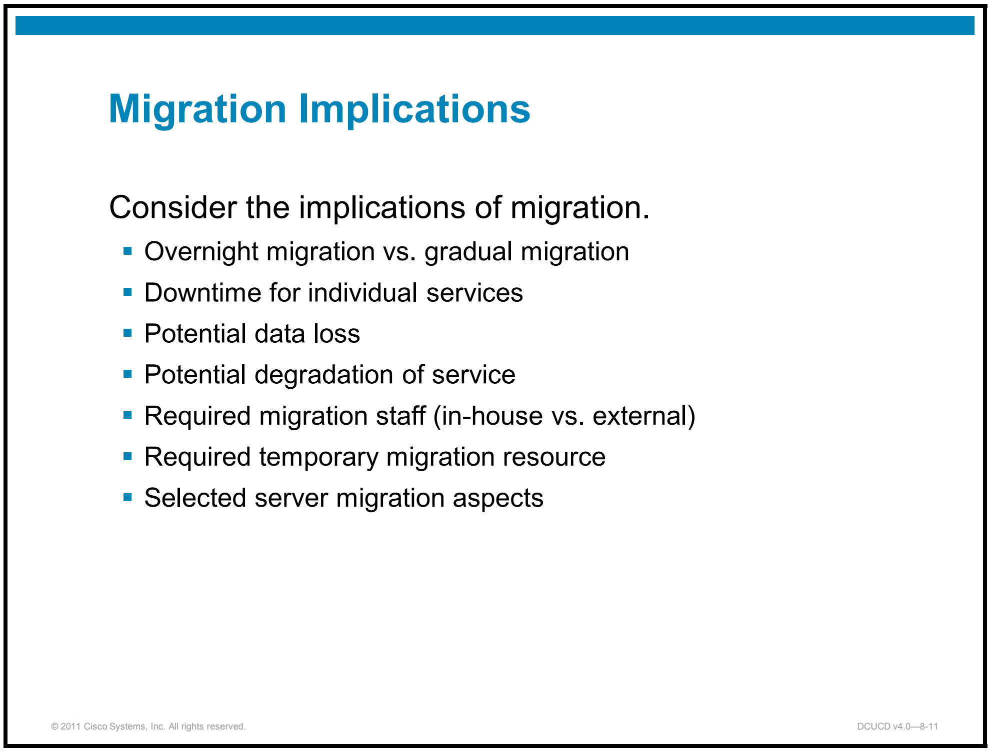
Use reconnaissance and analysis tools

- VMware Capacity Planner, Akorri BalancePoint, Microsoft Assessment and Planning (MAP) Toolkit

 DCUCD v4.0—8-10

It is vital that the assessment is run and information is collected over a long period so that it is relevant.

A lot of information can be collected with the various reconnaissance and analysis tools available from different vendors. For example, if the project involves migration to a VMware vSphere environment from physical servers, the VMware Capacity Planner will help with collecting information about the servers (such as processor power and memory used).



Migration of a large and complex data center can be a very sensitive operation. It is the migration that will reveal any outstanding flaws in the data center design and migration plan. The list in the figure covers the most common considerations that, when properly addressed, help you build a migration plan that can reduce the overall risk.

Designing a Migration Plan

This topic describes migration plans.

Migration Design Overview

Prerequisites

- Cisco UCS pod setup
- LAN and SAN infrastructure
- Connectivity between existing equipment and UCS pod

Migration procedures

- Operating system
- Application
- Data

Verification steps

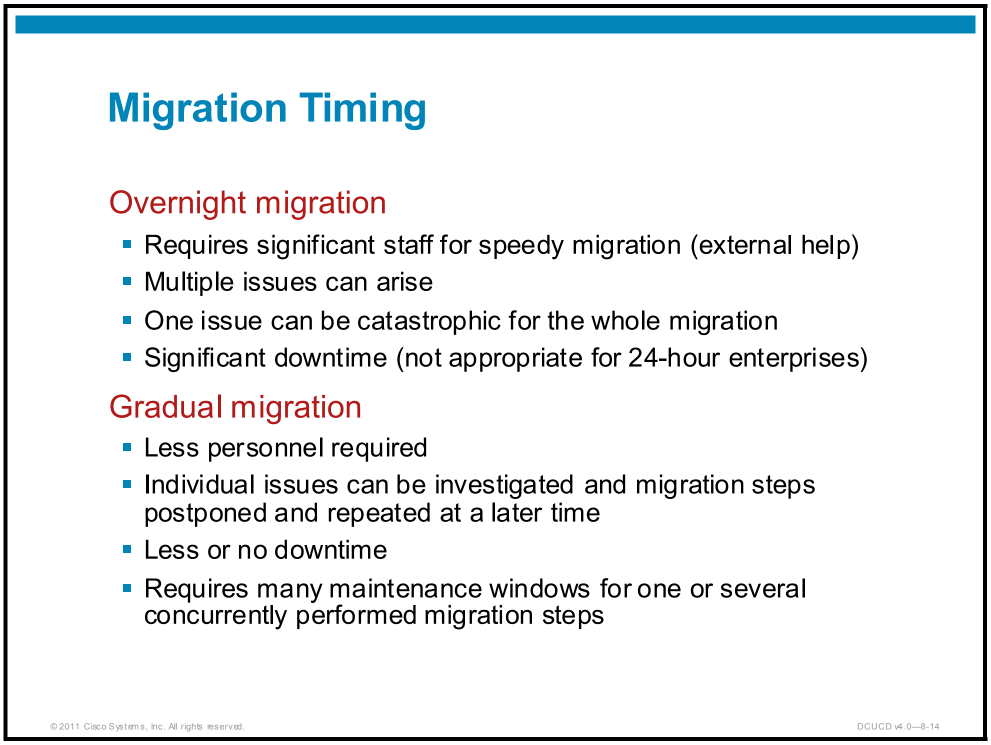
Rollback procedures

- Address migration issues

 DCUCD v4.0—8-13

In general, a migration plan should include a list of migration actions, which can be divided into migration phases. Each migration action should have three main components:

- Detailed task list that is assigned to the appropriate resource
- Verification steps to confirm that the migration was successful
- Rollback procedure to revert to the original setup in case an unforeseen problem is detected that cannot be immediately mitigated



Overnight Migration

Migration of data centers can be performed all at once (the "overnight" migration) wherein the migration plan is designed to migrate all services at once. This is typically feasible only for small data centers that are used only during business hours and not 24 hours a day. This can offer a migration window between two business days (literally overnight) or during the weekend (two entire days). Larger data centers or data centers that are used 24 hours a day typically cannot be migrated in such short times.

Gradual Migration

An alternative to overnight migrations is to devise a migration plan wherein services are being migrated one by one over several maintenance windows (such as nights or weekends). Gradual migration plans should be designed in a way to cause minimum downtime per service (minutes or hours).

Downtime

Cold conversion of servers may require significant downtime (for example, hours)

Hot conversion minimizes required downtime but may not always be available or reliable

- Cannot be used for P2P server migration

Application-layer clustering has zero downtime but

- Is only possible with applications that have clustering functionality
- Requires additional planning for adding new servers to the cluster and decommissioning old servers

 DCUCD v4.0—8-15

When determining how to migrate old servers to new servers, it is recommended that you select the migration option with the minimum downtime, especially for services that require maximum uptime.

Data Loss

Hot server conversion or data copying during busy hours could result in data loss during conversion and switchover.

Replication between existing and new centralized storage is an option, but switchover requires downtime (stop old server, start new server).

 DCUCD v4.0—8-16

Hot conversion means that a conversion is performed while the server is active. This process can potentially lead to problems where data can be lost within moments as a conversion and switchover are being performed. Avoid using this method for services that involve sensitive data where data loss is not acceptable (for example, an online banking transaction server). The preference is a longer downtime versus data loss.

Degradation of Service

Migrated servers should provide the same functionality with equal or better performance.

Consider individual migrated components:

CPU	➔	Guaranteed CPU resources in virtual environment that are equal to or greater than existing CPU power
Memory	➔	Equal or greater than existing memory for virtual machine
Disk	➔	Create virtual disk on the central storage that is equal to or greater than existing disk capacity
Network	➔	Ensure equal or greater bandwidth to virtual machine (on server and network equipment)

Audit the existing environment to get peak and average resource utilization, and provide guarantees for the identified peak periods (do not extend the V2P ratio beyond capacity).

When dimensioning the virtualized data center, it is normal to assign multiple virtual servers per one high-performance server (that is, several CPUs or cores, or both, and large amounts of memory). Care should be taken not to assign too many, which could result in performance degradation after migration. Longer monitoring (for example, over several weeks) should be performed and the data analyzed to determine the average and peak resource utilization of each physical server. The peak utilization should be used to dimension the new data center and the P2V ratio.

Migration Staff

In-house staff may not have any experience with migration.

Use external expertise when lacking in-house expertise for migration.

Consult external experts for creating a migration plan, and use in-house staff to carry out a gradual migration.

 DCUCD v4.0—8-18

A gradual migration plan can be prepared with the help of external experts with experience in data center design and migration. The migration itself, when done gradually, can be performed using in-house staff with optional external oversight or help.

Temporary Resources

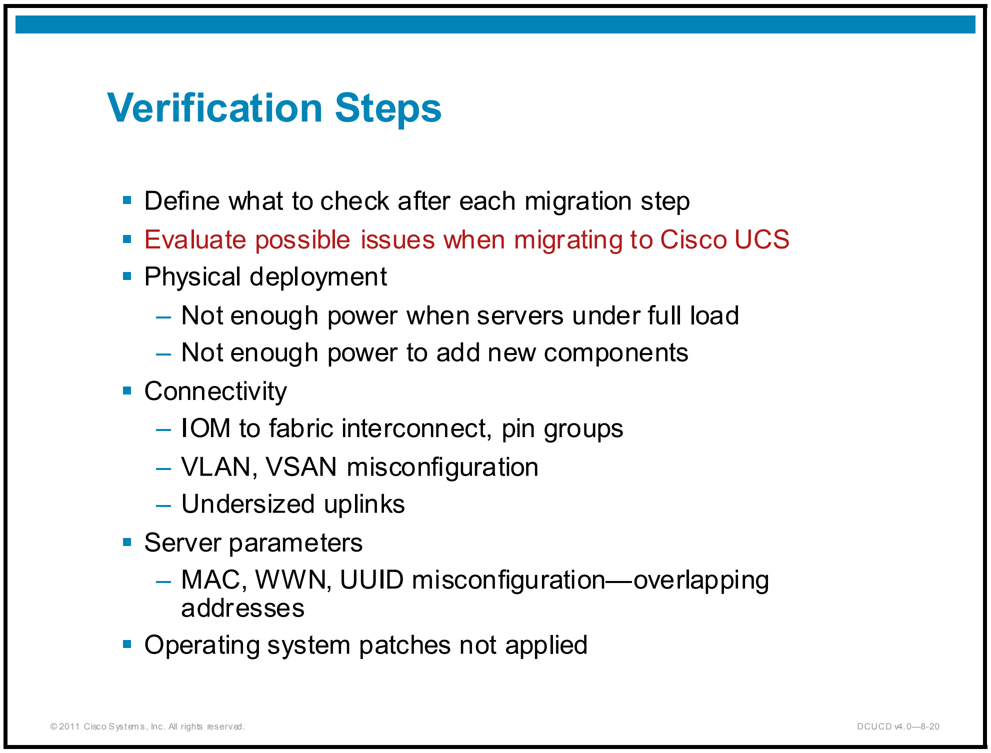
Some resources may be required for the migration only:

- Layer 1 and Layer 2 connectivity between old and new data center (for example, dark fiber or Layer 2 VPN)
- Cisco UCS server blades for migration tools and applications
- Additional storage space

The migration plan should list and detail the temporary resources to ensure a smooth migration.

 DCUCD v4.0—8-19

A migration plan should also list any temporary resources that are needed during the migration process. This would typically include Layer 2 connectivity between the old and the new data centers, and some servers and network devices to help in the migration process.



Verification steps are an important part of the migration plan, since they are used to verify proper operation of the migrated server infrastructure and applications.

To enable you to write proper verification steps, the plan should evaluate the possible problems that could take place during migration.

Physical Deployment-Related Issues

There could be issues related to physical deployment:

- It is possible that, during or after the migration, all or part of the equipment could be shut down due to insufficient power or cooling capacity. The issue could arise from the amount of load that is placed on the servers, which under full load consume more power and produce more heat than when idle or under light load.
- It is possible that adding a new component—such as a Cisco UCS server blade—cannot be completed due to insufficient power. This could be due either to an insufficient power supply installed in the Cisco UCS chassis or to insufficient power at rack level.

Connectivity-Related Issues

There could be issues related to connectivity:

- Misconfigured I/O Module (IOM)-to-fabric interconnect links
  - The IOM-to-fabric interconnect link is not following fabric interconnect A and B design requirements, but instead the uplinks are cross-connected to fabric interconnect A and B in an ad hoc manner—left and right IOM connected to the same fabric interconnect.
  - The wrong configuration of pin groups would result in no connectivity. Pin groups might specify the wrong uplink for a certain server.
- VLAN errors
  - Server network interface card (NIC) adapters might be placed in the wrong VLAN.

  - The upstream switch trunk interface might block required VLANs (due to the allowed VLAN list not being properly configured).
- VSAN errors
  - Server HBAs might be configured with the wrong VSAN.
  - FCoE VLAN ID might be overlapping with existing VLANs.
- Undersized uplinks resulting in poor or no connectivity due to insufficient bandwidth

Server Issues

Overlapping MAC, WWN, or UUID parameters might result in no connectivity.

Operating System Issues

An operating system without proper patches might prevent proper server operation.

Evaluating a Migration Plan

Test implementation and migration steps in advance if possible:

Connectivity
- Interconnect old and new infrastructure (Layer 2)
- Create dummy virtual or physical servers with available IP addresses in existing server VLANs on Cisco UCS
- Test connectivity from Cisco UCS dummy servers to real servers

Conversion to virtual servers
- If possible, hot-convert a real server
- Start the virtual server in isolated environment to verify its standalone operation

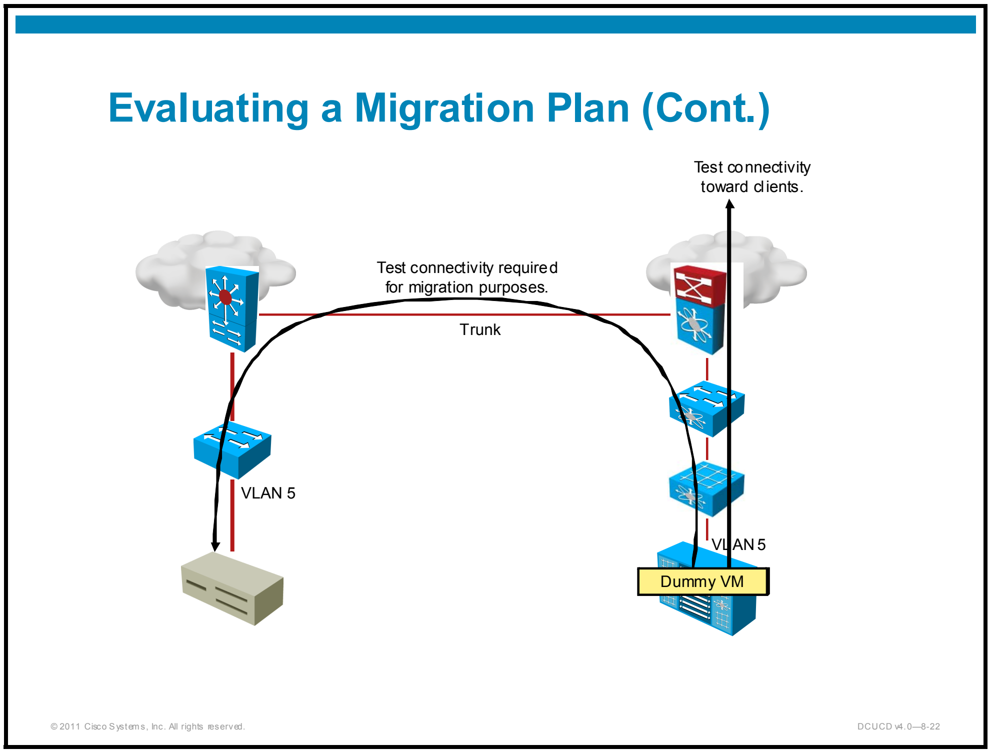
Migration of physical servers
- Use cloning tool to transfer data if possible

Management
- Verify proper operation of existing management tools for operating system and applications

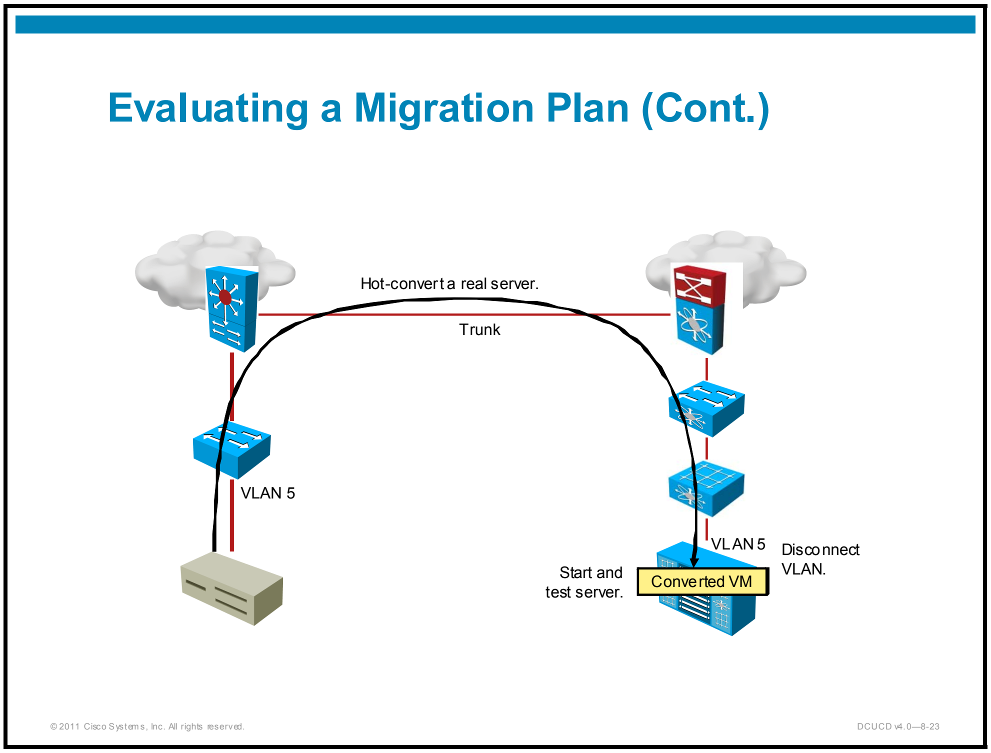
 DCUCD v4.0—8-21

A "dry run" of the migration can be performed—where all actions that do not cause downtime can be performed to ensure that any issues in the migration plan are identified in advance before the actual migration is started.

This evaluation requires a new data center to be built and configured, connectivity to be provided between the old and the new data centers, and some services to be tested in an isolated environment.



The networking and storage infrastructure of a new data center can be built and tested in advance. The operation of real servers can be tested only after they are migrated. To maximize the reliability of the migration process, it is recommended that you create dummy virtual servers and test their connectivity to the existing physical servers (this connectivity is required during the migration period) and their connectivity to clients (this connectivity will be required for normal server functionality after migration).



Additionally, it is recommended that you hot-convert some or all of the existing physical servers and start them in the new virtual environment in an isolated mode, simply to test that the servers were successfully converted and are operational.

The figure illustrates a simplified set of migration steps for a single server using the third migration option—using conversion tools. The old and new data center infrastructures are temporarily connected on Layer 2, and a hot conversion is used to create a new virtual server with the same operating system and applications as the original physical server (only the underlying hardware is different).

Deliverables

Assessment document

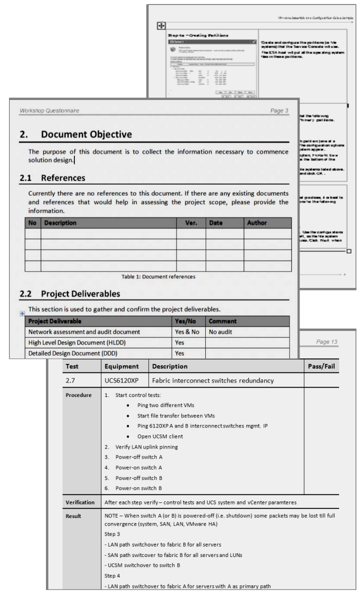
- Describes existing environment with emphasis on performance characteristics

Migration design

- Describes what will be done in migration
- Lists required resources

Migration implementation

- Detailed schedule of migration steps and actions
- Also includes verification steps and rollback procedures



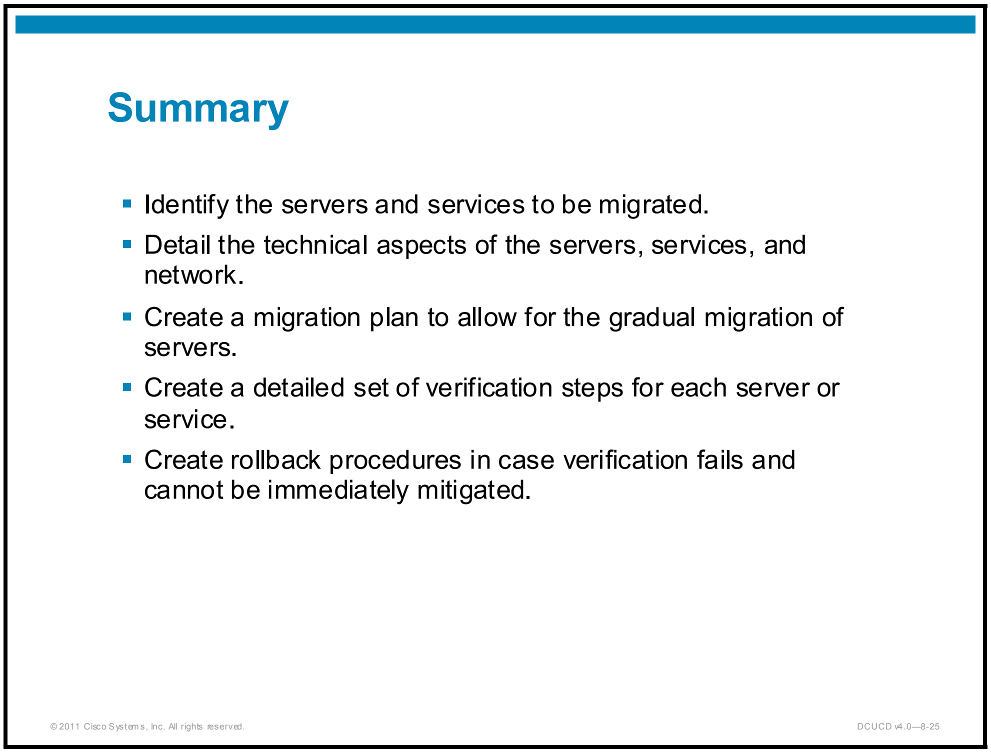
DCUCD v4.0—8-24

The migration should be properly documented with the following documentation:

- **Assessment:** This document describes the existing environment and all the characteristics including performance, with special emphasis on applications and their responsiveness. Migration should not result in degradation of user experience since that would render the project a failure. By measuring the environment before the migration and after the migration, the migration staff can prove that the new environment is working better than the old one.
- **Design:** This document describes the selected migration scale, selected server migration options, required tools, selected operating system, and application migration options and required resources. The migration design document specifies the resources that are necessary to conduct the migration—for example, extra storage space, extra Ethernet ports to connect new equipment before old equipment is decommissioned, extra staff or external specialists, and so on.
- **Implementation plan:** This document lists and describes the sequence of the migration plan in detail, consisting of:
  - **Migration procedures, steps, and actions:** Specifying the actions for conducting the migration (in the correct order) with verification tests and expected results.
  - **Verification procedures:** Specifying the actions that are required to confirm proper behavior.
  - **Rollback procedures:** Specifying the actions that are necessary to revert to a previous state in case of problems during migration.

Summary

This topic summarizes the key points that were discussed in this lesson.



Module Summary

This topic summarizes the key points that were discussed in this module.

Module Summary

- An assessment should be performed over a long period of time before a migration commences.
- The migration plan should include a list of migration actions—detailed task list, verification steps, and rollback procedures.

 DCUCD v4.0—8-1

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) Which migration scenario typically requires additional space in the existing facility? (Source: Designing a Migration Plan)

A) full migration to a new site
B) physical to virtual server migration
C) virtual to virtual server migration
D) partial migration or redesign

Q2) What is the benefit of gradual migration? (Source: Designing a Migration Plan)

A) reduces the potential challenges of the migration
B) enables quick migrations over the weekend
C) no temporary resources required
D) no migration staff required
E) single maintenance window required

Q3) Which conversion method requires the least downtime? (Source: Designing a Migration Plan)

A) cold conversion
B) hot conversion
C) V2V migration
D) overnight migration

Module Self-Check Answer Key

Q1) D

Q2) A

Q3) B

Appendix

Determining a Solution for ROI and TCO Using VMware Chargeback

Overview

This lesson describes how to evaluate a design and solution for the return on investment (ROI) and the total cost of ownership (TCO).

Objectives

Upon completing this lesson, you will be able to meet this objective:

- Evaluate the VMware Chargeback tool

VMware Chargeback

This topic describes the VMware Chargeback tool.

Virtualization Implications

Virtualization versus physical model

- New application = new VM (not hardware)
- IT owns hardware (not business)

What has changed?	IT still needs the ability to:
New virtualization paradigm	Monitor and control server deployments
Ease of VM provisioning makes it simple to deploy new machines	Create accountability, because sprawl can occur
Virtual machines are cheaper than physical machines but are not free	Manage IT priorities by using costs to influence customer behavior

DCUCD v4.0—A1-4

IT enables business users and application owners to perform the following activities:

- Easily request services
- Specify desired service levels
- Consume and pay for those services

These activities are provided with a high degree of reliability, and the application owners and business users do not need to understand the underlying infrastructure.

This approach hides the complexity of the infrastructure from the users. It changes how end users request services and how IT delivers service levels with a dynamic, flexible, and reliable IT infrastructure. This transformation positions IT as a service provider with the ability to interact with end users, clearly understand and manage their expectations, ensure that the infrastructure delivers these needs, and transparently allocate costs for services that are delivered.

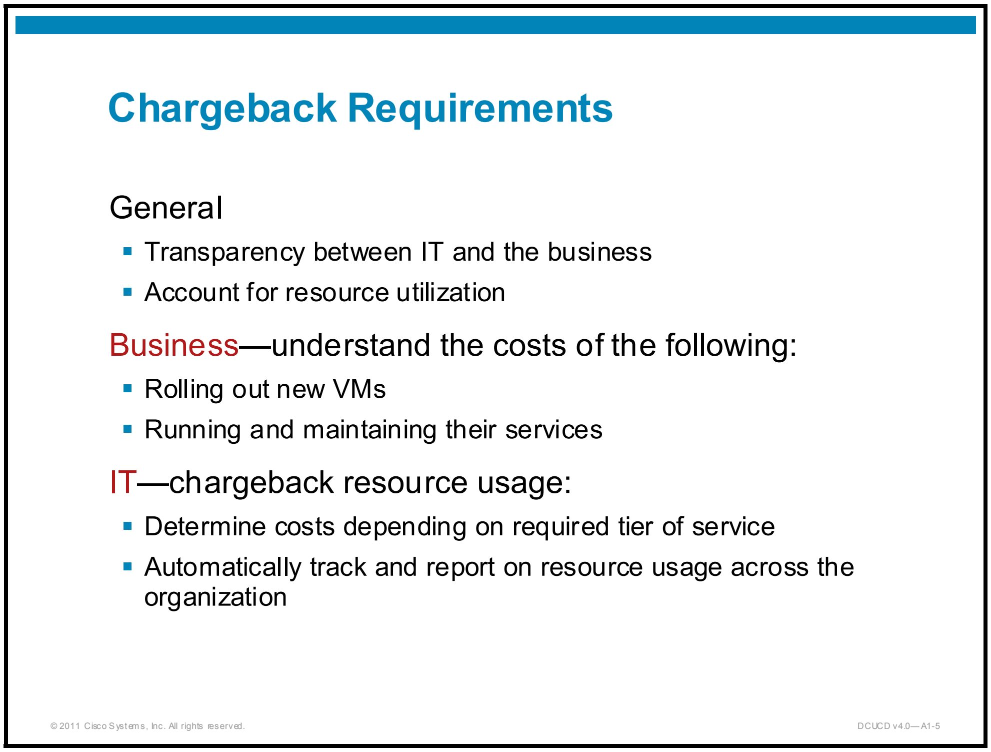
IT is accountable for delivering what the business needs (for example, reliable email service, a responsive customer relationship management (CRM) system, or an e-commerce site that supports peak shopping periods). IT is not held responsible for whether a particular server or other piece of infrastructure is operating correctly.

Virtualization Implications

Server virtualization and the virtual machine (VM) sprawl that it is creating are now making IT organizations re-evaluate their chargeback positions.

The speed and flexibility of virtualization makes some form of chargeback mandatory; otherwise, demand could increase greatly and more low-priority workloads could be deployed that do not justify their costs.

It is not a sustainable approach to focus only on component-level management. IT needs to manage what it is measured on—business outcomes—and this measurement will require a fundamentally different and better approach that raises management to service levels.



IT administrators have a key building block for delivering IT as a service. They need to meter resources managed in virtual infrastructure, which includes server resources, network traffic, public IP addresses, and other services such as DHCP, NAT, and firewalling, and so on.

The chargeback model is required to account for the operational costs that are involved in providing and maintaining an IT infrastructure, including the costs for IT services and applications (for example, the VMware vSphere infrastructure). Measuring resource utilization and calculating the corresponding IT operational cost enable IT to account for the IT resources that are utilized, and to bill for the services that are provided in a virtual infrastructure.

To properly address this problem, the IT administrators can utilize the chargeback tools. The representatives of such tools are VMware Chargeback, which is used in VMware vSphere environments and VKernel, which can be used also in other environments.

VMware vCenter Chargeback

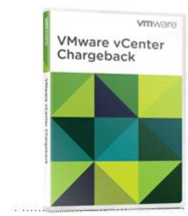
Account, monitor, and report on costs associated with virtual infrastructure

Features

- Fixed, allocation, and utilization-based costing
- Charge different amounts for tiers of infrastructure
- Schedule reports and email results

Benefits

- Improve resource utilization
- Budget optimization



 DCUCD v4.0—A1-6

VMware vCenter Chargeback enables accurate cost measurement, analysis, and reporting of VMs. With vCenter Chargeback, line-of-business owners and IT teams can gain visibility into the actual cost of the virtual infrastructure that is required to support business services.

vCenter Chargeback can map IT costs to business units, cost centers, or external customers. This ability can enable a better understanding of how much resources cost and what can be done to optimize resource utilization.

It also allows organizations to support policy-driven accountability for self-service environments so that business owners can "pay as they go" for IT resources.

VMware Chargeback Aspects

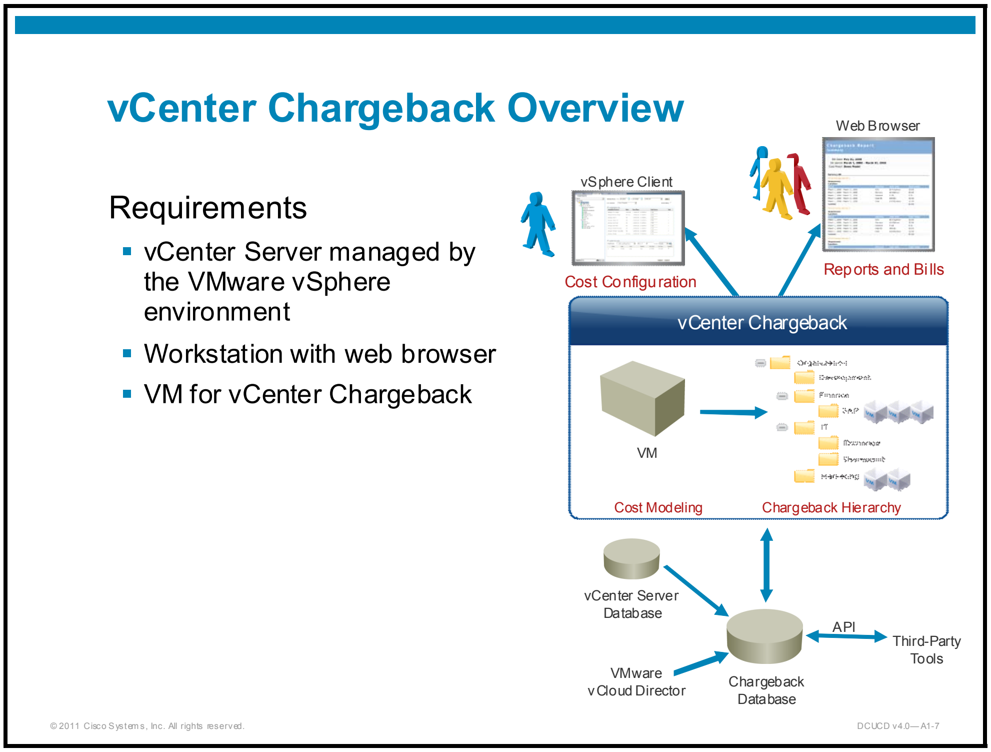
The key benefits of VMware Chargeback are the following:

- Improves resource utilization by gaining accurate visibility into the true costs of virtualized workloads
- Enables line-of-business owners to have complete transparency and accountability for self-service resource requests
- Models infrastructure costs in a flexible way that fits organizational processes and policies

VMware Chargeback allows administrators to perform these functions:

- **Report virtualization costs precisely:** Chargeback takes into account many different factors. These range from hardware costs (CPU, memory, storage, and so on) to additional elements such as power and cooling. Chargeback can incorporate all of these variables to provide IT with a comprehensive virtualized-environment, cost-enabling chargeback to individual business units and the business as a whole.
  - Understand the costs of VMs
  - Properly allocate costs across organization units
  - Comprehensive reporting

- **Customize cost models and metrics:** Chargeback allows for IT administrators to enter their unique cost information and tune chargeback, based on their specific requirements.
  - Support reservation-based costing, utilization-based costing, or a combination of both.
  - Allow for the entry of cost accounting structures, base cost models, fixed costs, and multiple rates that are needed to calculate true costs. Templates can also be used to apply existing models quickly to expanding or new environments.
  - Export of data to third-party systems or importing of existing hierarchies.
- **Simplify billing:** Chargeback automatically creates detailed billing reports that can be submitted to business units within an organization. This solution provides business units with a clear view of the resources that are consumed and their associated costs. Administrators can now transition the IT environment from a cost center to a value center.



VMware Chargeback is tightly integrated with VMware vSphere and vCenter Server and thus it requires both programs for its operation.

Chargeback can be used for these functions:

- Generate cost reports that can provide comparison of cost models and resource usage for specified periods of time
- Allow for customizable report templates that allow reports that contain a unique header, footer, logo, title, and so on, to fit company or business unit brand standards
- Report on the most used and most expensive VMs, allowing organizations to gain visibility into the largest contributors to data center costs, which in turn makes it easier to influence and prioritize efficiency efforts

vCenter Chargeback automatically creates detailed reports that can be submitted to business units within an organization or used by IT for decision making and planning. These reports can provide organizations with a clear view of the resources that are consumed and their associated costs, eliminating the perception that VMs are "free."

Costing Models

- Start simple and move to an advanced model over time
- Compare between models with different reporting options
- Ensure that the model aligns with organizational requirements
- Flexible costing options and mix and match between models

VM

Costing Model	Description
Fixed costing	Fixed cost for a VM instance
Allocation-based costing	Variable costs per VM based on allocated resources
Utilization-based costing	Variable costs per VM based on the actual resources utilized

 DCUCD v4.0—A1-8

Multiple cost models provide flexibility in measuring costs. vCenter Chargeback supports three different cost models:

- **Fixed costing:** This model looks at a specific per VM and includes costs such as floor space, power and cooling, and software or administrative overhead.
- **Allocation-based costing:** This model looks at the variable costs per VM that are based on allocated resources, such as the amount of memory, CPU, or storage allocated or reserved for the VM in the vCenter Server.
- **Utilization-based costing:** This model looks at the variable costs per VM that are based on actual resources that are used, including average memory, disk and CPU usage, network I/O, and disk I/O.

The cost models can be combined in a cost template, making it easy to start with a simple chargeback model and align it with organizational requirements. Administrators can customize the chargeback model with unique cost information and flexible cost models. vCenter Chargeback can allow these functions:

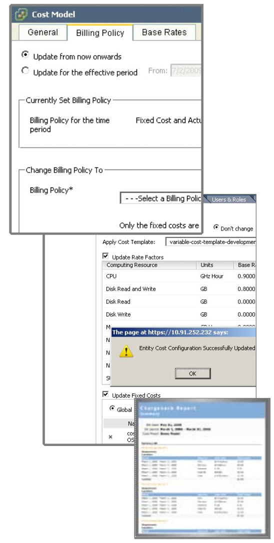
- The ability to generate allocation-based costing, utilization-based costing, or a combination of both to fit the unique costing policies of an organization.
- Reusable cost templates that allow the process of adding cost models to VMs to be standardized
- The entry of base costs, fixed costs, one-time costs, multiple rate factors, and overage fees to model true costs

Accounting Overview

Accounting aspects

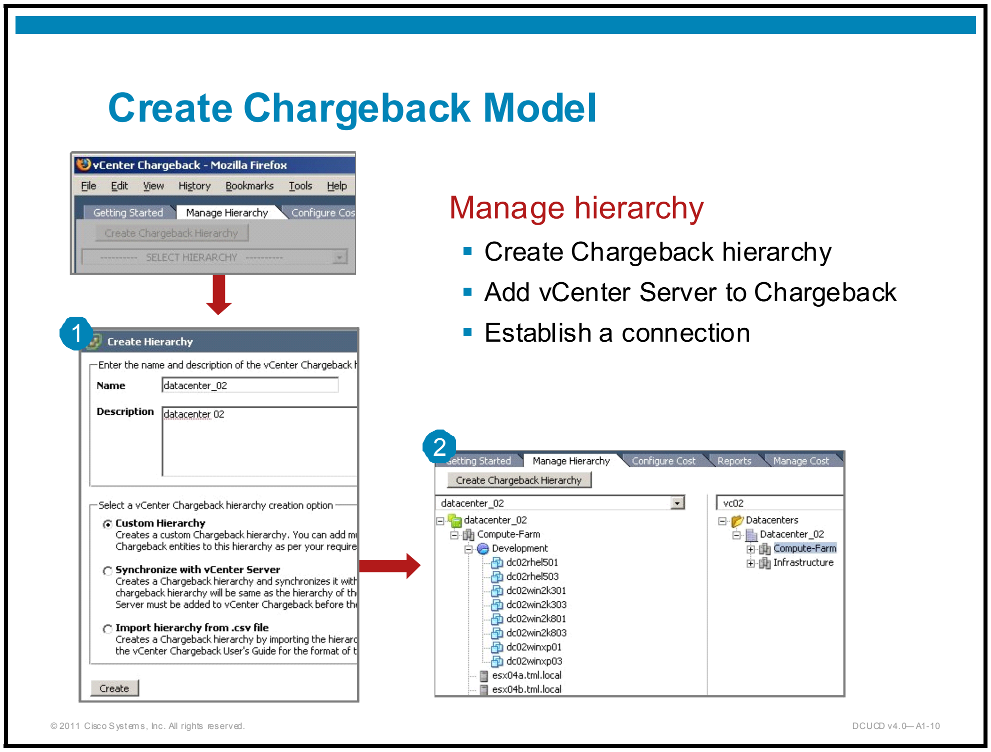
- Cost accounting—base rates associated with tracked metering elements
- Multiple rate factors
  - Added for use of storage, servers, and so on
  - Can distinguish between development versus production system costs
  - Increase rates for additional services (such as backup and disaster recovery)
- Applying fixed costs—to add noncompute attributes (such as licenses, power and cooling, and staff)

Use cost templates to simplify the process of adding costs for VMs.



 DCUCD v4.0—A1-9

Administrators can define the base rates that are associated with metered elements (for example, the CPU, memory, disk, disk I/O, or network I/O). Rate multipliers can be applied to different costing models at run-time, allowing IT administrators to account for the use of higher-end storage or servers, or higher service level agreements (SLAs) (for example, the disaster recovery or backup that is included, production versus development systems). Metered elements that are not needed can also be easily disabled.



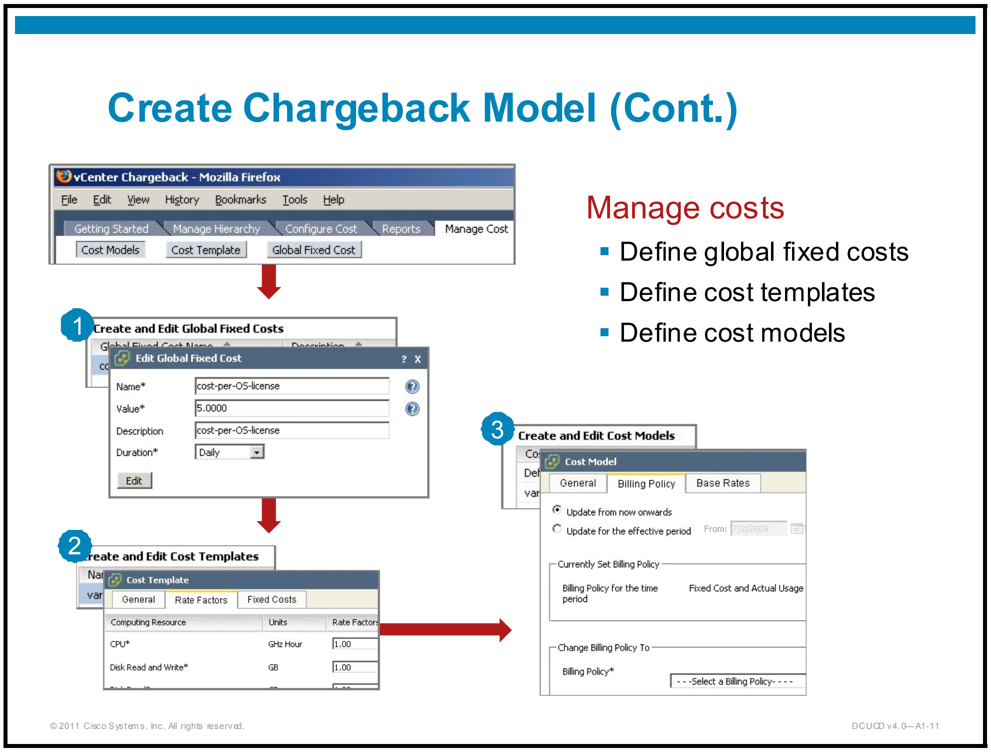
Management Model

Administrators create a hierarchy in Chargeback that provides the basis for determining cost structures and the delivery of reports.

The hierarchy can span multiple vCenter Server instances for organization-wide chargeback. Cost templates simplify and standardize the process of adding costs for virtual machines to the Chargeback hierarchy. Any number of cost templates can be created to support different cost structures across the organization.

Chargeback interacts with the vCenter Server to determine the utilization of the computing resources by various VMs that are created in the vCenter Server hierarchy. Chargeback enables administrators to create more hierarchies that can be different from the vCenter Server hierarchies.

A Chargeback hierarchy defines how the various VMs and ESX Server hosts are assigned to and shared by different departments, cost centers, or business units in an organization.



Manage Costs

The costs have to be managed, which allows for the entry of cost accounting structures, base cost models, fixed costs, and multiple rates that are needed to calculate true costs. Templates can also be used to apply existing models quickly to expanding or new environments.

Global Fixed Costs

Many environments today do not have a current billing or cost-tracking policy for a virtual infrastructure, so this scenario shows how to start with a simple fixed-cost model based on the fixed cost of a VM. With the help of the VMware Chargeback calculator, the administrator can translate physical costs to virtual and determined fixed costs per VM.

The fixed-cost model relies on defining fixed costs that are set up in the current environment, and then these fixed costs are attached to one or more folders in the hierarchy through a cost template.

Cost Template

Chargeback enables the administrator to create cost templates. A cost template consists of rate factors for each of the chargeable resources, global costs, and local fixed costs.

The rate factor is the multiplication factor to be used along with the base rate to calculate the charge per unit of a resource for an entity. Rate factors are useful when you want to charge the entities in a hierarchy differently.

The rate factors enable the administrator to apply a cost that is a multiple or a fraction of the base rate.

A local fixed cost is a cost that is specific to an entity, a group of entities, or to an entire hierarchy. An example of a local fixed cost is the real estate cost for the physical storage of the entities. This cost differs based on the actual geographic location.

The rate factors and local fixed costs are entity-specific. The administrator can set these values in a cost template and use the template to configure the costs at the entity level.

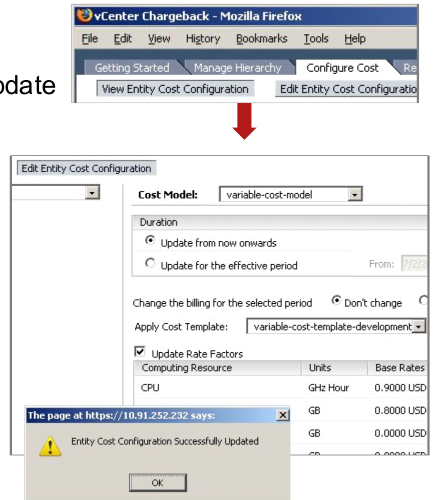
Cost Model

The Chargeback cost model defines the base rate and billing policy. The base rate is the global rate that you want to charge for a unit of the chargeable computing resource, such as the CPU and memory. A billing policy determines the cost types and chargeable units to be considered for calculating the chargeback cost. Chargeback provides several types of billing policies, such as for the fixed cost and for the actual usage.

Use Chargeback Model

Configure costs

- Edit entity cost—choose and update
  - Cost model
  - Cost template
  - Rate factors and fixed costs



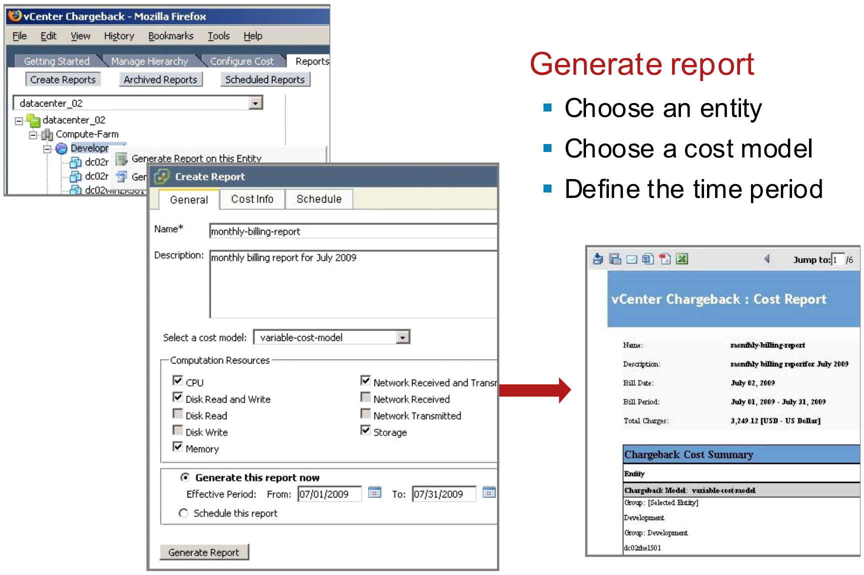
Configure Costs

The base rates that are defined in a cost model are global values and will be applied uniformly to the entire hierarchy or part of the hierarchy for which you generate a report.

In real time, however, the administrator might want to charge each entity or a set of entities differently. That is, the resource usage cost for one VM might differ from the cost for another. This ability can be achieved only if the cost for each entity or a set of entities in the hierarchy is configured separately.

The administrator may want to charge some global fixed costs and entity-specific local fixed costs.

Use Chargeback Model (Cont.)



Generate report

- Choose an entity
- Choose a cost model
- Define the time period

 DCUCD v4.0—A1-13

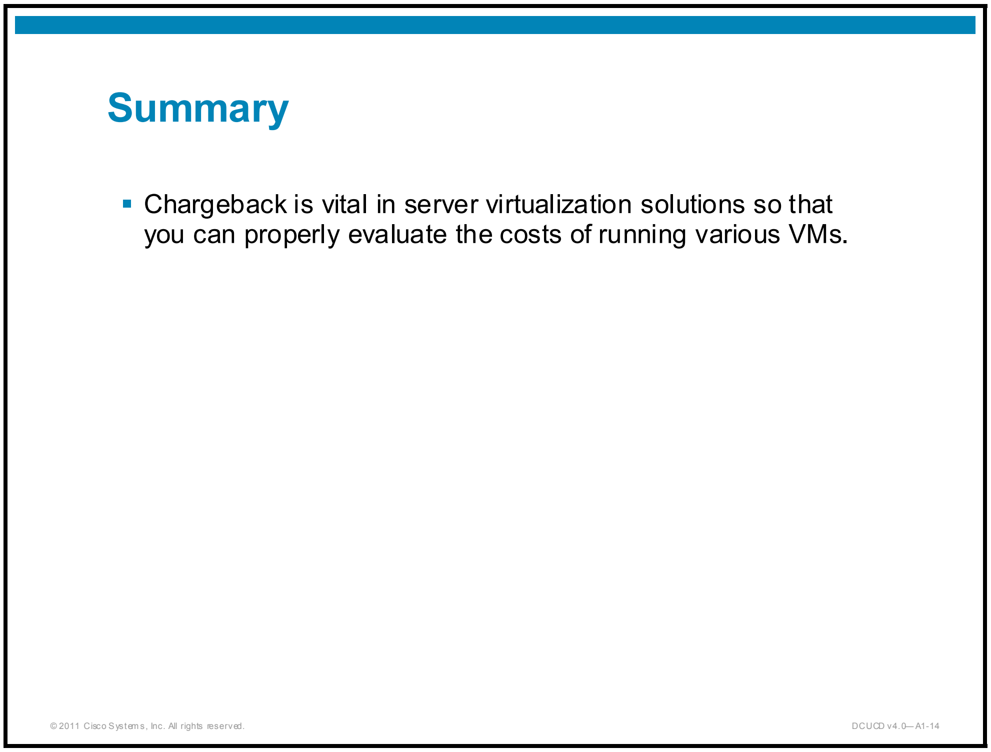
Once the model is defined, it can be used to generate a report about the virtualization infrastructure that is selected. When generating the report, the administrator has to enter the desired cost model and the time period for which the report should be generated.

When the report is generated, the information can be reviewed and the cost per VM can be easily determined (in relation to the cost model used in the report, such as variable, as in the figure, fixed, per utilized resource, and so on).

Using such reports is also vital in understanding the benefits of the complete infrastructure, not only the benefits per part. This ability enables the IT administrators and solution architects to better evaluate the solution.

Summary

This topic summarizes the key point that was discussed in this appendix.



References

For additional information, refer to these resources:

- http://www.cisco.com/en/US/partner/solutions/ns340/ns517/ns224/tools/data_center_value_zone.html
- http://www.cisco.com/survey/exit.html?https://express.salire.com/Go/Cisco/Unified-Computing-ROI-Tool.aspx