#### Go to: Before You Buy or Deploy - Considerations for Design and Procurement

#### **Contents**

- 1 General Guidelines for SAN/NAS
- 2 SAN/NAS Link Provisioning and High Availability
- 3 Best Practices for Storage Array LUNs for Unified Communications
  Applications
- <u>4 IOPS and other Performance Requirements</u>
  - ♦ 4.1 Unified Communications Manager
  - ♦ <u>4.2 Cisco Emergency Responder</u>
  - ♦ 4.3 Cisco Intercompany Media Engine
  - ♦ <u>4.4 Cisco Unity Connection</u>
    - ♦ 4.4.1 Summary of IOPS for Unity Connection
  - ♦ <u>4.5 TelePresence Applications</u>
    - ♦ 4.5.1 Cisco TelePresence Video Communication Server (Cisco VCS)
    - ♦ 4.5.2 Cisco TelePresence Conductor (TelePresence Conductor)
  - ♦ 4.6 Unified Presence
  - ♦ 4.7 Cisco Unified Attendant Consoles
  - ◆ 4.8 Cisco Customer Collaboration/Contact Center
    - ♦ 4.8.1 Cisco Contact Center Enterprise
    - ♦ 4.8.2 Cisco Contact Center Express/IPIVR IOPS
    - ♦ 4.8.3 Workforce Optimization (WFO) IOPS
    - ♦ 4.8.4 Cisco MediaSense
    - ♦ 4.8.5 Cisco Finesse
    - ♦ 4.8.6 Cisco Unfied Email Interaction Manager and Web Interaction Manager
    - ♦ 4.8.7 Cisco Unified Intelligence Center
    - ♦ 4.8.8 Cisco Customer Voice Portal (Windows 2003 R2 Server)
    - ♦ 4.8.9 Cisco Customer Voice Portal (Windows 2008 R2 Server)

#### **General Guidelines for SAN/NAS**

- Adapters for storage access must follow supported hardware rules (click here).
- Cisco UC apps use a 4 kilobyte block size to determine bandwidth needs.
- Design your deployment in accordance with the UCS High Availability guidelines (see <a href="http://www.cisco.com/en/US/solutions/collateral/ns340/ns517/ns224/ns944/white-paper-c11-553711.html">http://www.cisco.com/en/US/solutions/collateral/ns340/ns517/ns224/ns944/white-paper-c11-553711.html</a>).
- 10GbE networks for NFS, FCoE or iSCSI storage access should be configured using <u>Cisco Platinum Class QOS</u> for the storage traffic.
- Ethernet ports for LAN access and ethernet ports for storage access may be separate or shared. Separate ports may be desired for redundancy purposes. It is the customer's responsibility to ensure external LAN and storage access networks meet <a href="UC app latency">UC app latency</a>, performance and capacity requirements.
- In absence of UCS 6100/6200, normal QoS (L3 and L2 marking) can be used starting from the first upstream switch to the storage array.

Contents 1

- With UCS 6100/6200
  - ◆ FC or FCoE: no additional requirements. Automatically handled by Fabric Interconnect switch.
  - ♦ iSCSI or NFS: Follow these best practices:
    - ♦ Use a L2 CoS between the chassis and the upstream switch.
    - ♦ For the storage traffic, recommend a <u>Platinum class QoS, CoS=5</u>, no drop (Fiber Channel Equivalent)
    - ♦ L3 DSCP is optional between the chassis and the first upstream switch.
    - ♦ From the first upstream switch to the storage array, use the normal QoS (L3 and L2 marking). Note that iSCSI or NFS traffic is typically assigned a separate VLAN.
    - ♦ iSCSI or NFS: Ensure that the traffic is prioritized to provide the right IOPS. For a configuration example, see the FlexPod Secure Multi-Tenant (SMT) documentation (<a href="http://www.imaginevirtuallyanything.com/us/">http://www.imaginevirtuallyanything.com/us/</a>).
- The storage array vendor may have additional best practices as well.
- If disk oversubscription or storage thin provisioning are used, note that UC apps are designed to use 100% of their allocated vDisk, either for UC features (such as Unity Connection message store or Contact Center reporting databases) or critical operations (such as spikes during upgrades, backups or statistics writes). While thin provisioning does not introduce a performance penalty, not having physical disk space available when the app needs it can have the following harmful effects
  - degrade UC app performance, crash the UC app and/or corrupt the vDisk contents.
  - ♦ lock up all UC VMs on the same LUN in a SAN

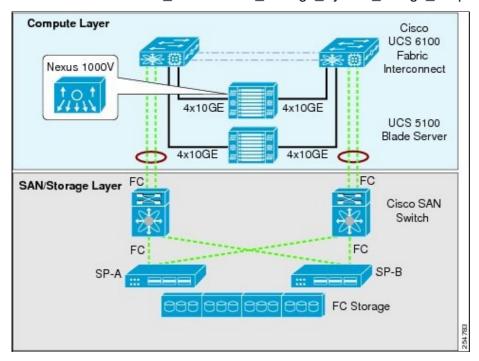
# SAN/NAS Link Provisioning and High Availability

Consider the following example to determine the number of physical Fiber Channel (FC) or 10Gig Ethernet links required between your storage array (such as the EMC Clariion CX4 series or NetApp FAS 3000 Series) and SAN switch for example, Nexus or MDS Series SAN Switches), and between your SAN switch and the UCS Fabric Interconnect Switch. This example is presented to give a general idea of the design considerations involved. You should contact your storage vendor to determine the exact requirement.

Assume that the storage array has a total capacity of 28,000 Input/output Operations Per Second (IOPS). Enterprise grade SAN Storage Arrays have at least two service processors (SPs) or controllers for redundancy and load balancing. That means 14,000 IOPS per controller or service processor. With the capacity of 28,000 IOPS, and assuming a 4 KByte block size, we can calculate the throughput per storage array controller as follows:

- 14,000 I/O per second \* (4000 Byte block size \* 8) bits = 448,000,000 bits per second
- 448,000,000/1024 = 437,500 Kbits per second
- 437,500/1024 = ~428 Mbits per second

Adding more overhead, one controller can support a throughput rate of roughly 600 Mbps. Based on this calculation, it is clear that a 4 Gbps FC interface is enough to handle the entire capacity of one Storage Array. Therefore, Cisco recommends putting four FC interfaces between the storage array and storage switch, as shown in the following image, to provide high availability.



Note: Cisco provides storage networking and switching products that are based on industry standards and that work with storage array providers such as EMC, NetApp, and so forth. Virtualized Unified Communications is supported on any storage access and storage array products that are supported by Cisco UCS and VMware. For more details on storage networking, see <a href="http://www.cisco.com/en/US/netsol/ns747/networking">http://www.cisco.com/en/US/netsol/ns747/networking</a> solutions sub program home.html.

# **Best Practices for Storage Array LUNs for Unified Communications Applications**

There are various ways to create partitions or Logical Unit Numbers (LUNs) in the storage array to meet the IOPS requirement for Cisco Unified Communications applications (see <u>IO Operations Per Second (IOPS)</u>}.

The best practices mentioned below are meant only to provide guidelines. Data Center storage administrators should carefully consider these best practices and adjust them based on their specific data center network, latency, and high availability requirements.

The storage array Hard Disk Drive (HDD) must be a Fibre Channel (FC) class HDD. These hard drives could vary in size. The current most popular HDD (spindle) sizes are:

- 450 GB, 15K revolutions per minute (RPM) FC HDD
- 300 GB, 15K RPM FC HDD

Both types of HDD provide approximately 180 IOPS. Regardless of the hard drive size used, it is important to try to balance IOPS load and disk space usage.

LUN size must be less than 2 terabytes (TB) for the virtual machine file system to recognize it. For Cisco Unified Communications virtual applications, the recommendation is to create a LUN size of between 500 GB and 1.5 TB, depending on the size of the disk and RAID group type used. Also as a best practice, select the LUN size so that the number of Unified Communications virtual machines per LUN is between 4 and 8. Do not allocate more than eight virtual machines (VMs) per LUN or datastore. The total size of all Virtual Machines (where total size = VM disk + RAM copy) must not exceed 90% of the capacity of a datastore.

LUN filesystem type must be VMFS. Raw Device Mapping (RDM) is not supported.

The following example illustrates an example of these best practices for UC:

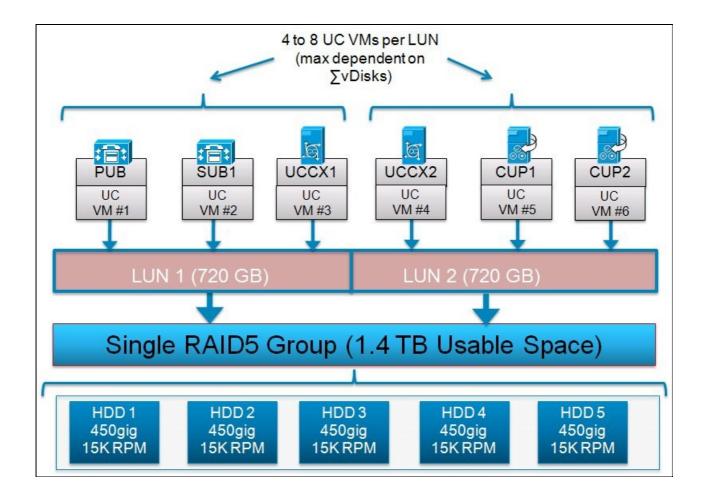
For example, assume RAID5 (4+1) is selected for a storage array containing five 450 GB, 15K RPM drives (HDDs) in a single RAID group. This creates a total RAID5 array size of approximately 1.4 TB usable space. This is lower than the total aggregate disk drive storage space provided by the five 450 GB drives (2.25 TB). This is to be expected because some of the drive space will be used for array creation and almost an entire drive of data will be used for RAID5 striping.

Next, assume two LUNs of approximately 720 GB each are created to store Unified Communications application virtual machines. For this example, between one and three LUNs per RAID group could be created based on need. Creating more than three LUNs per RAID group would violate the previously mentioned recommendation of a LUN size of between 500 GB and 1.5 TB.

A RAID group with RAID 1+0 scheme would also be valid for this example and in fact in some cases could provide better IOPS performance and high availability when compared to a RAID 5 scheme.

The above example of storage array design should be altered based on your specific Unified Communications application IOPS requirements.

Below is a graphic of an example configuration following these best practices guidelines, note there are other designs possible.



# **IOPS** and other Performance Requirements

This page illustrates IOPS under various conditions for Unified Communications applications. This area is under construction. Check back frequently for updates.

Storage performance must support the sum of UC VM OVA IOPS. Note that addressing IOPS requirements may require higher disk/spindle counts, which may result in excess storage capacity

IOPS utilization should be monitored for each application to ensure that the aggregate IOPS is not exceeding the capacity of the array. Prolonged buffering of IOPS against an array may result in degraded system performance and delayed reporting data availability.

# **Unified Communications Manager**

This section provides the IOPS data for a Cisco Unified Communications Manager system under load. These values are per active VM OVA. Which VM OVAs are active, and how many are active simultaneously, depends on how the CUCM cluster nodes are setup with respect to service activation, redundancy groups, etc. (see www.cisco.com/go/ucsrnd for details).

CUCM storage access is on average 93-98% sequential writes with an IO block size of 4 kilobytes.

- Active call processing: As a reference, the following steady state IOPS were observed at various loads (expressed in Busy Hour Call Attempts):
  - 10K BHCA produces ~ 35 IOPS
  - 25K BHCA produces ~ 50 IOPS
  - 50K BHCA produces ~ 100 IOPS
  - 100K BHCA produces ~ 150 IOPS
- Software upgrades during business hours generate 800 to 1200 IOPS in addition to steady state IOPS.
- CDR/CMR via CDR Analysis and Reporting (CAR)
  - ◆ CUCM sending CDR/CMR to the external billing server does not incur any additional IOPS.
  - Enabling CAR continuous loading results in around 300 IOPS average on the system.
  - Scheduled uploads are around 250 IOPS for Publisher's VM only.
- Trace collection is 100 IOPS (occurs on all VMs for which tracing is enabled).
- Nightly backup (usually Publisher's VM only) is 50 IOPS.

# Cisco Emergency Responder

This section provides the IOPS Data for a Cisco Emergency Responder System under Load. These tests use published OVA template for Cisco Emergency Responder.

- Major Discovery: IOPS value of 66.
- Export/Import: IOPS value of 325.
- DRS Backup/Restore: During a backup, the value is 29; during a restore, it reaches 1000 IOPS.

• **Upgrade**: IOPS Value of 4.5.

# Cisco Intercompany Media Engine

Use values for Cisco Unified Communications Manager above (they will be conservatively high).

# **Cisco Unity Connection**

## **Summary of IOPS for Unity Connection**

<b>IOPS Type</b>	1vCPU	2vCPU	4vCPU	7vCPU
Avg Total	78.66	129.01	217.89	201.73
Peak Total	760.99	714.06	867.88	773.70
Avg Read	5.90	19.60	45.67	9.53
Peak Read	601.84	469.46	693.21	526.22
Avg Write	72.77	109.40	172.22	192.20
Peak Write	247.14	293.51	371.97	413.50

# **TelePresence Applications**

For Cisco TelePresence Manager and Cisco TelePresence Multipoint Switch: 100 IOPS is a good typical value to plan around. Otherwise use values for Cisco Unified Communications Manager above (they will be conservatively high).

## Cisco TelePresence Video Communication Server (Cisco VCS)

This section provides the IOPS data for a Cisco VCS system under load.

The recommended hardware on which to run a Cisco VCS VM is Cisco UCS C200 ? M2, UCS C210 ? M2, UCS C220 M3, or UCS B200 ? M2 with:

- Processor supporting AESNI feature
- 6GB of RAM per VM
- 132GB disk space per VM (for a 4GB virtual disk 1 and a 128GB virtual disk 2)
- R2XX-LBBU (Raid disk battery backup to enable cache)
- Four hard disks (450GB SAS 15K RPM 3.5in HDD/hot plug/C200 drive sled)
- PCI card Intel Quad port GbE Controller (E1G44ETG1P20)

#### Test specifications:

- Four individual Cisco VCS VMs running X7.2
- Single C220 M3 TRC running ESXi 5.0.0 (update 1)

#### Test results:

Activity	Average IOPS	Average r/w latency	Peak read latency	Peak write latency
Performance test*	5	0.5ms	15ms	2ms
Software upgrade	35	5ms	300ms	160ms

## UC\_Virtualization\_Storage\_System\_Design\_Requirements

Core dump	40	10ms	100ms	10ms
Burnin test suite	35	50ms	400ms	60ms

<sup>\*</sup> Performance test comprised 100 simultaneous interworked and encrypted calls at 768kbp/s

Note that Cisco VCS software X7.2 and later takes advantage of AES acceleration on Intel Advanced Encryption Standard New Instruction set available on the Cisco C2xx family of servers. This provides enhanced performance for encrypted calls.

## **Cisco TelePresence Conductor (TelePresence Conductor)**

This section provides the IOPS data for a TelePresence Conductor system under load.

The recommended hardware on which to run a TelePresence Conductor VM is Cisco UCS C200? M2, UCS C210? M2, UCS C220 M3, or UCS B200? M2 with:

- Processor supporting AESNI feature
- 6GB of RAM per VM
- 132GB disk space per VM (for a 4GB virtual disk 1 and a 128GB virtual disk 2)
- R2XX-LBBU (Raid disk battery backup to enable cache)
- Four hard disks (450GB SAS 15K RPM 3.5in HDD/hot plug/C200 drive sled)
- PCI card Intel Quad port GbE Controller (E1G44ETG1P20)

#### Test specifications:

- Four individual TelePresence Conductor VMs running XC1.2
- Single C220 M3 TRC running ESXi 5.0.0 (update 1)

#### Test results:

Activity	Average IOPS	Average r/w latency	Peak read latency	Peak write latency
Performance test*	5	0.5ms	15ms	2ms
Software upgrade	35	5ms	300ms	160ms
Core dump	40	10ms	100ms	10ms
Burnin test suite	35	50ms	400ms	60ms

<sup>\*</sup> Performance test comprised 2400 endpoints, 15 calls per second, 30 MCUs, 2 person conference

#### **Unified Presence**

The CUP 1000 user OVA generates about 60 IOPS during steady-state. Expect similar peaks as CUCM during backups and upgrades.

#### **Cisco Unified Attendant Consoles**

This section provides the IOPS data for Cisco Unified Attendant Consoles (using Enterprise Attendant Console to produce heaviest load). Load test and average operations are represented. These tests use the published OVA template for Cisco Unified Attendant Consoles.

L	OAD TESTING	
IOPS (Read)	IOPS (Write)	

## UC\_Virtualization\_Storage\_System\_Design\_Requirements

	Min	Max	Average	Min	Max	Average	Operating Conditions
Cisco Unified Attendant Console Server	0	1013.605	0.442	0	1112.87	12.121	5,000 calls per hour, 25 concurrent operators, continuous LDAP sync, full logging
Upgrade	0	320.006	3.691	0	105.001	10.341	Major upgrade
			AVERA	GE O	PERATI	ONS	
	IOPS (Read)			IOPS (Write)			
		IOPS (Re	ead)		IOPS (W	rite)	
	Min	IOPS (Re	ad) Average	Min	· ` `	T	Operating Conditions
Cisco Unified Attendant Console Server		Ì		1	· ` `	T	Operating Conditions  Synchronized directory, cached BLF, minimum online changes

# **Cisco Customer Collaboration/Contact Center**

# **Cisco Contact Center Enterprise**

The SAN must be able to handle the following Unified CCE application disk I/O characteristics.

The data below is based on CCE 8.5(3) running on Windows Server 2008 R2 Enterprise

Unified CCE	IOPS			Disk R	ead KB	ytes / sec	Disk W	rite KI	Bytes / sec	Onoroting
Component	Peak	Avg.	95th Pct	Peak	Avg.	95th Pct	Peak	Avg.	95th Pct.	Operating Conditions
Router	27	14	19	619	11	12	436	246	313	12,000
Logger	3662	917	1793	12821	735	11571	38043	2777	11832	agent;100
HDS	1017	321	693	2748	250	2368	19351	1441	3879	CPS, ECC: 5 scalars @ 40 bytes each; 200 Reporting users at max query load
Router	118	11	13	1681	14	5	971	66	172	8,000
Logger	1614	502	1065	6137	341	4993	30379	558	5911	agents;60
HDS	1395	165	541	1995	186	1776	12525	1299	2965	CPS; ECC:5 scalars @ 40 bytes each; 200 Reporting users at max query load

The data below is based on CCE 8.0(2) running on Windows Server 2003 R2

		IOPS	3	Disk R	ead KB	ytes / sec	Disk V	Vrite K sec	Bytes /	
Unified CCE Component	Peak	Avg.	95th Pct.	Peak	Avg.	95th Pct.	Peak	Avg.	95th Pct.	Operating Conditions
Router Logger	20 1,000	8 600	10 700	520 4,000	30 600	180 2,500	400 12,000	60 3,000	150 7,000	8,000 agents; 60 cps; ECC: 5 scalars @ 35 bytes each; No
HDS	1,600	1,000	1,100	600	70	400	6,000	2,000	3,800	reporting.
Agent PG	125	40	70	300	5	20	2,000	1,200	1,500	2,000 agents, 15 cps
HDS	3,900	2,500	3,800	75,000	30,000	50,000	9,500	2,200	5,800	8,000 agents; 60 cps; 200 reporting users at max query load; ECC: 5 scalars @ 35 bytes each.
ROGGER	610	360	425	2,700	400	1,600	7,500	2,150	4,300	4,000 agents; 30 cps; ECC: 5 scalars @ 35 bytes each

# Cisco Contact Center Express/IPIVR IOPS

The following table lists the IOPS numbers for Unified Contact Center Express/IPIVR.

IOPS Type	2vCPU
IOPS Avg per VM	150
IOPS Max spike per VM	1500

# **Workforce Optimization (WFO) IOPS**

The following table lists the IOPS numbers for Cisco Unified Workforce Optimization 8.5(2) onward.

Workforce Optimization Component	IOPS (Read)	IOPS (Write)	Operating Conditions
QM 8.5(2+) Base Server	4	2	Idle
	4	12	Recording 300 concurrent calls
	11	38	During recording upload
QM 8.5(2+) Recording Server or WFM 8.5(2+)	1	2	Idle
	74	187	Recording 300 concurrent calls
	102	71	During recording upload

## Cisco MediaSense

The following table lists the IOPS numbers for Cisco MediaSense partitions.

Media Partitions	<b>IOPS Specifications</b>
Average IOPS for non-media partitions	1700 ops/sec
Average IOPS for media partitions	50 ops/sec
Average bandwidth for non-media partitions	7300 kbytes/sec
Average bandwidth for media partitions	1000 kbytes/sec

## **Cisco Finesse**

The following table lists the IOPS numbers for Cisco Finesse.

IOPS Type	IOPS Specification			
Max IOPS	160			
95th percentile	150			

## Cisco Unfied Email Interaction Manager and Web Interaction Manager

EIM-WIM component	IOPS		Disk Read KBytes / sec		Disk Write KBytes / sec				
Envi- w hvi component	Peak	95th Pct.	Avg.	Peak	95th Pct.	Avg.	Peak	95th Pct.	Avg.
Application Server	11.7	3.21	1.55	51	1	0.726	103	19	8.32
File Server	43.65	25.23	14.36	189	2	2	1769	1025.9	446
Database Server	736.9	552	263.31	35450	23184	3903	5737	2872	1625
Messaging Server	9.15	2.71	1.47	57	1	0.7	88	15.1	8
Services Server	11.15	3.155	1.53	55	1	0.73	1.89	17	9
Web Server	11.2	3.36	1.43	55	1	0.8	173	21	10.24
Cisco Media Blender	12.25	7.45	5.51	47	0	0.516	384	289	177.47

#### **Cisco Unified Intelligence Center**

- CPU usage (average) should not exceed 60% for the ESXi Server and for each of the individual processors, and for each VM.
- Memory usage (average) should not exceed 80% for the ESXi Server and for each of the VMs.
- VM snapshots are not supported in production since they have significant impact on system performance.
- The SAN must be able to handle the following Cisco Unified Intelligence Center application disk I/O characteristics.
- Enable hyper threading on all ESXi servers.

Cisco MediaSense

The following table lists the IOPS numbers for Cisco Unified Intelligence Center.

Details	Write (Kilobytes/Sec)	Read (Kilobytes/Sec)	IOPS
Maximum	7758.00	466.00	781.40
95 Percentile	6446.30	433.10	628.34
Average	5727.44	74.32	460.17
Minimum	4897.00	0.00	288.75

## Cisco Customer Voice Portal (Windows 2003 R2 Server)

The following table lists the IOPS numbers for Cisco Unified Customer Voice Portal running on Windows 2003 R2 Server.

#### **IOPS Summary: (Unit in Number, Data are Blade-based)**

<b>CVP Server component</b>	Average	MAX	95th Percentile
Call Server/VXML Servers	380	1536	797
Reporting Server	692	3403	2024
Operation Server	NA (Insignificant)	NA	NA

## Cisco Customer Voice Portal (Windows 2008 R2 Server)

#### • IOPS Numbers:

The following section lists the IOPS numbers for Cisco Unified Customer Voice Portal running on Windows 2008 R2 Server.

## The following IOPS metrics are measured in peak load conditions with failover:

<b>CVP Server</b>	Average	MAX	95th Percentile
CS/VXML/Media Servers	50	1300	124
Reporting Server	329	1250	984
Operations Console	22	100	80

#### The following IOPS metrics are measured in peak load conditions with no failover:

CVP Server	Average	MAX	95th Percentile
CS/VXML/Media Servers	21	450	30
Reporting Server	40	930	56
Operations Console	22	100	80

## • Read/Write Ratio:

#### The following table provides Read/Write Ratio:

# UC\_Virtualization\_Storage\_System\_Design\_Requirements

CVP Server	Average	MAX	95th Percentile
CS/VXML/Media Servers	199/327	2249/1654	411/928
Reporting Server	3563/3000	14375/15390	10010/11506
Operations Console	22/21	107/94	67/79

**Back to: Unified Communications in a Virtualized Environment**