



# CHAPTER 3

## Implementation and Configuration

---

Typical InMage deployments require control servers to be deployed in both the Enterprise and SP data centers. Proper sizing of each component, based on change rate, is required to achieve desired recovery point objective (RPO), recovery time objective (RTO), and retention window. This chapter will provide sizing details along with deployment recommendations for each component. The following major topics will be discussed:

- [Master Target—Enterprise and Service Provider, page 3-1](#)
- [InMage Scout Server Details, page 3-6](#)
- [InMage vContinuum, page 3-12](#)
- [InMage Agent Configuration, page 3-13](#)
- [Multi-Tenant Portal—RX Server, page 3-17](#)
- [Summary Tables of Components for All Tenants, page 3-21](#)
- [VMDC 2.3, page 3-24](#)
- [BMC Cloud Lifecycle Management, page 3-49](#)

The above contents are not organized based on deployment order, but are rather general guidelines for each component as a standalone entity. Actual deployment order may vary depending on the method of onboarding the tenant and the SP method of operation.

## Master Target—Enterprise and Service Provider

The InMage disaster recovery solution creates replication pairs between a primary VM and a dedicated VM on a secondary ESX server called "master target" or "MT." The MT may act as the protection target for multiple physical or virtual servers. Each machine (physical or virtual) can only be mapped to a single MT and a single protection plan. A single protection plan can also only span across a single MT. Depending on the number of servers under protection, each tenant will have at least one MT. Based on the type of use case, a tenant may require MTs in both the enterprise and SP:

- Enterprise to SP Recovery Protection: At least one instance of MT required in the SP.
- SP to Enterprise Failback Protection: At least one instance of MT is required in the Enterprise.

MTs can be deployed on Linux or Windows virtual servers, but must be of the same OS family as primary servers being protected. A number of factors, which are described in the following topics, determine deployment and sizing of the MT:

- [Master Target OS-Specific Information, page 3-2](#)
- [Master Target Deployment Limit, page 3-2](#)

- [Volume Requirements, page 3-3](#)

## Master Target OS-Specific Information

The OS that matches that of the primary server needs to be deployed on the MT.

- If Primary VMs are running Windows, the MT needs to be Windows based. Windows 2008 R2 is recommended for Windows-based MTs.
- Otherwise, if the primary server is running Linux, the Linux MT needs to be deployed. The CentOS-based MT is recommended.

Although detailed OS commands vary greatly when setting up a Windows MT from Linux, most of the high level planning steps are very similar. InMage recommends both the Linux and Windows MT to be configured with three volumes:

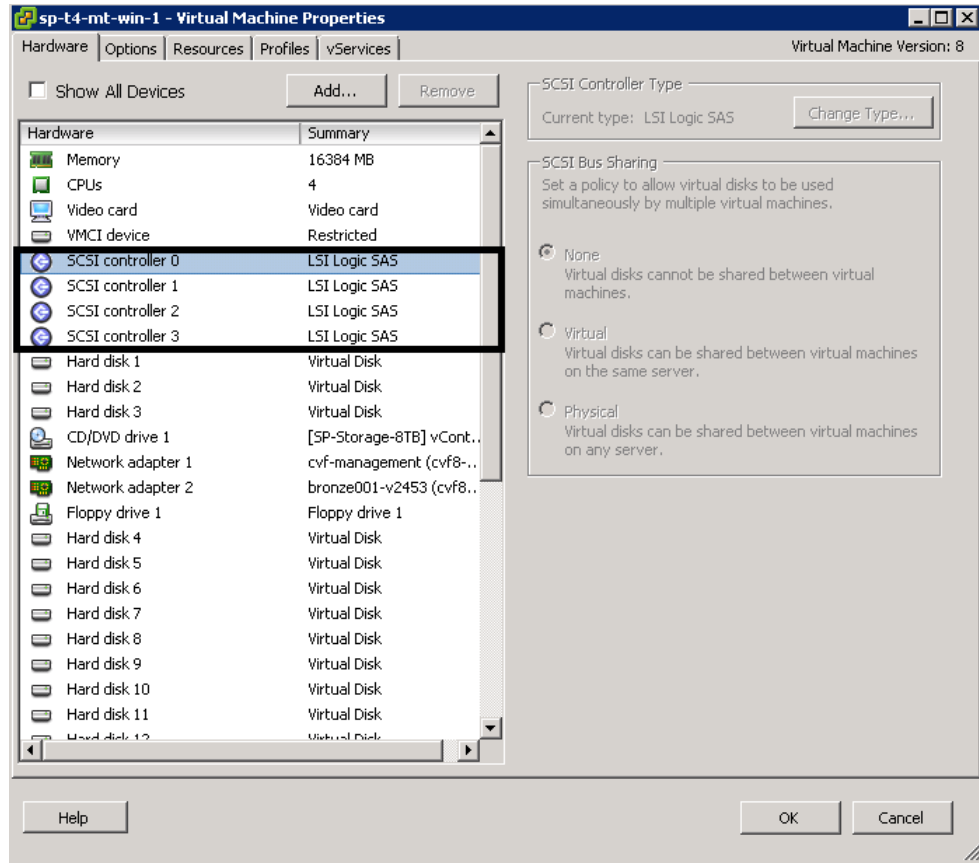
- OS boot volume
- Cache
- Retention

Both Cache and Retention are required for journal. Space required for each volume is based on change rate. Refer to [“Volume Requirements” section on page 3-3](#) for additional details.

## Master Target Deployment Limit

As discussed previously, the vContinuum will create a shadow copy of each disk under protection and dual mount the disk locally on the shadow VM and the MT. Each time a block changes at the source machine, the MT is the target of the changed block. As such, to prevent consistency problems on virtual disks, a lock is placed against all protected volume/VMDKs. Since the MT is a VM running in the vSphere environment, vCenter limits apply to each MT. vCenter limits each host to 4 SCSI controllers; each SCSI controller is capable of supporting 15 VMDKs. Refer to [Figure 3-1](#).

Figure 3-1 Master Target Deployment Limit



Due to the VMware SCSI controller limits, a single MT can support a maximum of 60 VMDKs. Additionally, InMage requires three volumes on the MT for OS, Cache, and Retention. Assuming DR Drill functionality is not required, a MT can support a maximum of 57 VMDKS and therefore protect 57 primary VMs. If DR Drill functionality is required, InMage recommends to not exceed 40 VMDKS per MT. If the number of primary VMs that need protection exceeds 57 (or 40 with DR Drill), then additional MTs must be deployed. Refer to “[Implementation Best Practices](#)” section on page 5-8 for details.

## Volume Requirements

To properly size a deployment, it is important to ensure the InMage MT has sufficient disk space required to support the desired change rate, journal history and intermediate caching. These are described in the following topics:

- [Retention Volume Sizing, page 3-3](#)
- [Cache Volume Sizing, page 3-5](#)

## Retention Volume Sizing

Retention policies effect the sizing of retention volume. The three types of retention policies are:

- **Time based:** The data in the retention window will be overwritten after the specified time period.

- **Space based:** The data in the retention window will be overwritten once the size is exhausted.
- **Time and space based:** The data in the retention window will be overwritten once the time specified or space specified qualifies first.

To ensure the SP meets and maintains SLA, time-based retention is often implemented. In a time-based deployment, two factors determine the size of retention volume:

- Daily Data Change Rate
- Retention Window

Retention window is the time period information about data changes on the primary server is stored. An application can be rolled back to any point in time or application-consistent bookmarks/tags in this window. If the data change rate is expected to be constant over a 24-hour window, the following simple relationship can be used to derive the size of retention drive:

**Retention Drive Size = (Daily Change Rate) \* (Number of hours of restore point or retention window)**

As an example, if an enterprise expects 750GB of data change in an 24-hour window and a retention window/journal history of six hours, the expected retention size will be (750G / 24 hours) \* (6 hours) = 187G (200G rounded up).

If the data change rate follows a traditional bell curve, where lots of change occurs during normal operation and little change occurs after hours, when sizing the retention drive based on time and space, it is important to ensure the retention drive is based on the peak data change window. In a production environment, the daily record of change rate profile is stored in the CX server. This information can be used to fine tune the size of the retention drive as needed. Figure 3-2 is a sample screen capture of the change rate profile screen:

Figure 3-2 Change Rate

The screenshot shows the ScoutCloud CX Monitor interface. At the top, there are navigation tabs for DASHBOARD, PROTECT, MONITOR, RECOVER, and SETTINGS. The current page is 'Monitor >> Statistics/Reports >> Analyzer'. Below the navigation, there are 'Protection Options' and 'Recommended CX Configuration' sections. The 'Recommended CX Configuration' section is highlighted with a red box and shows an 'Average data change rate: 309041.4397 Mbytes/day'. Below this is a table titled 'Pairs Configured' with columns for 'Include', 'Source Host: Source Volume', 'Target Host: Target Volume', 'Bandwidth Required For RPO ~>= 0 Kbits/Sec (PEAK)', 'Bandwidth Required For RPO ~>= 0 Kbits/Sec (AVERAGE)', 'Cumulative data changes(in MBytes)', 'Average data change rate (MBytes/Sec)', 'Retention Storage Required (in MBytes)', and 'Target Storage Required (in MBytes)'. The table contains three rows of data, with the first two rows having their 'Include' checkboxes checked.

Include	Source Host: Source Volume	Target Host: Target Volume	Bandwidth Required For RPO ~>= 0 Kbits/Sec (PEAK)		Bandwidth Required For RPO ~>= 0 Kbits/Sec (AVERAGE)		Cumulative data changes(in MBytes)			Average data change rate (MBytes/Sec)		Retention Storage Required (in MBytes)	Target Storage Required (in MBytes)
			With Compression	Without Compression	Average with Compression	Average without Compression	With Compression	Without Compression	Monitoring Interval(Days)	Compression Enabled	Compression Disabled		
<input checked="" type="checkbox"/>	T4-W2K8-SRC6:C	SP-T4-MT-WIN-1:C ESX 146185F5-CE22-C547-A042A2EB8D384BAC_C	6457.5019	8646.7993	1192.5294	1683.6418	65216.4492	92074.1593	7	0.1078	0.1522	39460.3540	49598.3501
<input checked="" type="checkbox"/>	T4-W2K8-SRC6:C SRV	SP-T4-MT-WIN-1:C ESX 146185F5-CE22-C547-A042A2EB8D384BAC_C_SRV	0.3348	3.4019	0.0441	0.4258	2.4097	23.2841	7	0.0000	0.0000	9.9789	109.9750
<input checked="" type="checkbox"/>	T4-W2K8-SRC20:C	SP-T4-MT-WIN-1:C ESX 146E0D55-40D1-1949-81462F5386A7C38B_C	7047.6979	9464.0244	1895.1813	2548.3528	103642.7276	139363.0420	7	0.1714	0.2304	59727.0180	69865.0141

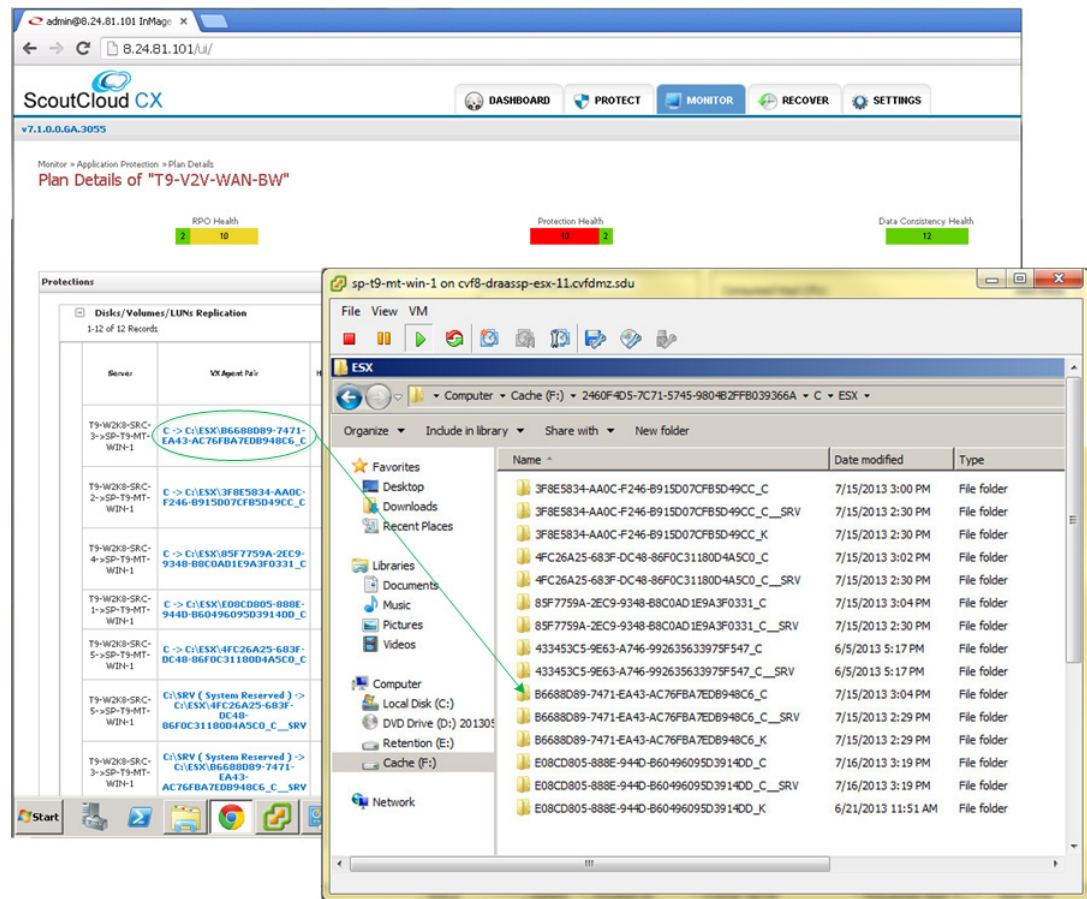
The highlighted box indicates the average change rate per day across all volumes under protection. Average change rate can be removed or added from the data change profile with the enable/disable check box next to "Individual volume."

## Cache Volume Sizing

The cache volume is used by the MT to store new uncompressed incoming data from the primary site prior to be processed and written to journal. For each source disk/volume under protection, a separate and dedicated folder is created on the cache volume on the MT.

Figure 3-3 shows the Cloud CX UI and the vCenter console for the MT named SP-T9-MTWIN-1. In the CX UI, several servers and VX Agent Pairs are shown. The VX Agent Pair associates a local drive on the primary server with a folder on the MT. The naming convention for the folders uses a host ID for the primary server with the volume name appended. Looking at the first row, changes for the C: drive on the T9-W2K8-SRC-3 primary server are cached on the F:\2460F4D5-7C71-5745-9804B2FFB039366A\C\ESX\B6688D89-7471-EA43-AC76FBA7EDB948C6\_C folder on the SP-T9-MT-WIN-1 MT.

Figure 3-3 Cache Volume Sizing



As a design guideline, 500MB of disk space per Virtual Machine Disk (VMDK) under protection should be reserved on the cache volume. The total size of the cache volume is a function of total number of volumes under protection:

**Size of Cache Volume = (Total number of volumes under protection) \* (500MB per volume)**

We expect the cache volume utilization to stay relatively low when no resource bottlenecks exist.

# InMage Scout Server Details

InMage Scout server is a multi-purpose server that, depending on deployment, can be referred to as one of the following:

1. **Process Server (PS):** Caches copies of block changes on primary server(s) residing in enterprise data center and sends them to master targets residing in the SP secondary site. Optionally, compression and encryption of block changes can be performed prior to being transmitted over the WAN link.
2. **Central Configuration Server (CX-CS):** Allows an operator to perform tasks through a web-based UI, such as fault monitoring (RPO violation, agent heartbeat), configuration, accounting, and performance monitoring. This server also generates reports, trend graphs, e-mail, and SNMP trap alerts.
3. **Dual role (PS + CX):** Similar to the CX-CS server, but adds the PS functionality in the server provider server to enable failback protection for the secondary servers back to the primary servers.

This section includes the following topics:

- [Scout Server Storage and Compute Implementation, page 3-6](#)
- [Scout Server Network Implementation, page 3-7](#)
- [Scout Server Network Bandwidth, page 3-8](#)
- [Scout Server Replication Options, page 3-10](#)

## Scout Server Storage and Compute Implementation

To understand Scout Server sizing, it is important to understand the multi-stage processing required to set up a protection plan:

- **Resyncing (Step I):** Baseline copy of primary server's drive under protection is created at the secondary server.
- **Resyncing (Step II):** Data changes during Step I are sent to the secondary server.
- **Differential Sync:** Once Step II completes, only block changes are sent to the secondary server.

To minimize WAN overhead, Scout also supports fast resync, where the replication process starts directly from differential sync instead of replication stages. Fast resync does a block-by-block comparison between storage volume at the primary and secondary sites. Data is sent over the WAN only if the two blocks are different. The primary case for fast resync is if the source went down abruptly causing some changes in memory to be lost.

As shown in [Figure 3-4](#), for each volume protected, disk space is reserved on the PS cache drive for the syncing stages using the following fields:

- **Resync File Threshold (MB):** For the resync steps, which defaults to 2G.
- **Differential Files Threshold (MB):** For differential sync, based on the expected or observed change rate, which defaults to 8G.

**Figure 3-4** Pair Settings

Pair Settings											
Visible	Resync	Profiling Mode	Secure CX-PS to Destination	Secure Source to CX-PS	Resync Mode	RPO Threshold	Replication Pool (1-24)	Resync Files Threshold (MB)	Differential Files Threshold (MB)	Compression Enable	
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Fast	30	2	2048	8192	CX-PS Based	

However, the 8G threshold can also be changed if the change rates for a single volume are higher. A single volume from a DB server, for example, could generate a large number of changes. It is a good practice to monitor cache utilization proactively and adjust thresholds based on application change rate. The main intent behind increasing these thresholds is to cache data on the PS instead of primary servers in the event there is a WAN bottleneck. If no bottlenecks exist, the default settings should work for most application deployment scenarios and provide enough storage to cache changes for a 6-8 hour WAN outage.

Total disk requirements on the processing server, based on default settings, will be the number of volumes \* (8G + 2G), which should not exceed the size of the cache partition. [Table 3-1](#) is the general CPU/memory/disk sizing recommendation by InMage based on change rate:

**Table 3-1** CPU/Memory/Disk Sizing Recommendation

Data Change Rate	CPU	Memory	Boot Volume Disk Type and Memory	Cache Disk
< 300GB	1 Quad Core	8 GB	RAID 1+0 15K disk, 40GB	RAID 1+0 10K/15K disk, 400GB
< 700GB	2 Quad Core	16 GB	RAID 1+0 15K disk, 40GB	RAID 1+0 10K/15K disk, 790GB
< 1 TB	2 Quad Core	32 GB	RAID 1+0 15K disk, 40GB	RAID 1+0 10K/15K disk, 790GB

The above recommendation assumes data encryption and compression as well as potential WAN outage of up to six hours. The CPU requirement could be reduced if data encryption and compression are offloaded to an external appliance. Consult with InMage for alternative deployment considerations.

## Scout Server Network Implementation

“[Component Flow](#)” section on page 2-38 documented data flows between the enterprise and SP among various InMage components that enable CDP and automatic recovery upon disaster declaration.

The primary and secondary server's VX agents communicate configuration and status information to the CX server over the HTTP protocol using standard port 80. The stateful firewall contains predefined rules for both HTTP and FTP/FTPS; i.e., no need to open a range of ports specifically exists. Also, the secondary server VX agents can open connections to port 873 (RSYNC) of the CX server when using resync.

The primary and secondary server agents use FTP/FTPS as the data transfer protocol to send and receive data from the CX server. The FTP protocol comes in two flavors, Active and Passive. InMage defaults to Passive FTP. Active FTP uses fixed ports 20 and 21 on the server side (CX server). In Passive FTP, clients initiate all connections; thus, no client-side firewall settings need to be configured. The clients, however, do need to be able to access port 21 and all ports greater than 1024 on the server side (CX server). It is possible to limit the range of ports to be opened up; this setting is controlled by `/etc/proftpd.conf`. To achieve this, simply create an entry for PassivePorts directive in your `proftpd.conf`:

```
PassivePorts 60000 65535 # allowed ports
```

The FX agents communicate configuration and status information to the CX server over the HTTP protocol. The data transfer protocol by default is a single socket connection to port 874 of the primary or secondary server.

In summary, InMage recommends that network traffic on the following ports not be blocked by hardware- or software-based firewall(s):

**Table 3-2 Network Ports**

Component	Traffic Type (Port)
Source Host	<ul style="list-style-type: none"> <li>• HTTP (80)</li> <li>• HTTPS (443)</li> <li>• Unencrypted data transfer (9080)</li> <li>• Encrypted data transfer (9443)</li> </ul>
CX Configuration Server	<ul style="list-style-type: none"> <li>• SMTP (25)</li> <li>• HTTP (80)</li> <li>• HTTPS (443)</li> <li>• MySQL (3306)</li> <li>• Unencrypted data transfer (9080)</li> <li>• Encrypted data transfer (9443)</li> </ul>
Target Host	<ul style="list-style-type: none"> <li>• HTTP (80)</li> <li>• HTTPS (443)</li> <li>• VX replication data traffic (873)</li> <li>• FX replication (874)</li> <li>• Unencrypted data transfer (9080)</li> <li>• Encrypted data transfer (9443)</li> </ul>
Optional	<ul style="list-style-type: none"> <li>• SNMP (162)</li> </ul>
Process Server	<ul style="list-style-type: none"> <li>• MySQL (3306)</li> <li>• Unencrypted data transfer (9080)</li> <li>• Encrypted data transfer (9443)</li> </ul>

## Scout Server Network Bandwidth

InMage Bandwidth Manager controls and shares bandwidth from the processing server to the master targets. Available WAN bandwidth, along with data change rate, directly impacts target RPO. To determine WAN bandwidth required during peak and off peak hours, InMage offers the ability to profile a single or multiple disk volumes before or after enabling block level replication. To enable profiling before setting up block level replication, data changes can be sent to a local CX-CS server or "InMageProfiler" for analysis of the data rather than sending data changes across the WAN. The objective is to simplify the task of initial sizing of the WAN for disaster recovery. Profiling as part of a regular block level data replication is intended to be regular ongoing capacity management, performed by the cloud admin, to adjust available WAN bandwidth as application usage pattern changes.



To gain valuable insight into data change rate and compressibility on the primary server(s) before setting up block level replication, first install the scout agent on the primary server. Configure a CXCS and CX-PS, aka "InMageProfiler," in the same LAN as the primary server with the following procedure:

- Step 1** From CX UI, navigate to **Protect > Protection Plans > Create Protection Plan > Setup Profiling**.
- Step 2** Provide the protection plan name and choose the type of profiling as **Individual Volume Profiling**.
- Step 3** Select a desired volume and select next to profile. All the hosts are listed under Host Drives. Expand the hosts to select the desired volume and click **Next**.
- Step 4** In the second step, select the target as InMageProfiler. The Replication Options are optional and the CDP retention option will not be available. Click **Next**.

The **Monitor > Volume Protection** page can be used for monitoring:

**Figure 3-5 Monitor/Volume Protection Page**

Monitor >> Volumes >> Reports  
Replication Reports

Statistics Reports Settings

Pair Details							
Primary Server	Primary Volume	Remote Server	Target Volume	Process Server	Replication Pool	Fast Resync Unmatched %	Agent Log
T4-W2K8-SRC2	C	InMageProfiler	P	t4-ps-1.t4.sdu [ 6.126.101.6 ]	24	N/A	N/A

Health Report [ Jul 01, 2013 - Jul 05, 2013 ]											
T4-W2K8-SRC2											
T4-W2K8-SRC2 (C) - PROTECTED <span style="float: right;">Change Rate   RPO   Retention   Health</span>											
Date	Data changes (in MBytes)		Retention Window (Days)		RPO		No. of hours RPO not met	Data Flow Controlled (Hours)	Retention log reset?	Available Consistency Points	Protection Coverage
	With Compression	Without Compression	Policy	Available	Threshold	Max					
Jul 04, 2013	3051.25	4511.41	0	0	30 min	7.23 days	0.03	0	NO	0	100%
Jul 05, 2013	8094.66	11524.09	0	0	30 min	0.26 min	0	0	NO	0	100%
<b>Total:</b>	11145.91	16035.5	N/A	N/A	N/A	N/A	0.03	0	N/A	N/A	100%

Data change profiling as part of a regular block level data replication is enabled by default. Navigate to **Monitor > Analyze Profile Results**.

The network planner can input the following four metrics into the protection options window:

- Proposed available bandwidth
- Desired RPO (minutes)
- Bandwidth adjustment factor
- Retention window (days)

InMage is based on historical data change metrics profile if proposed RPO is achievable during peak and average usage pattern. Based on profile result, the network planner can further adjust the WAN bandwidth or reduce the desired RPO expectations.

**Figure 3-6 Analyze Profiling Results**

Monitor » Statistics/Reports » Analyzer

**Analyzer**

Protection Options	
Cumulative bandwidth available (Kbits/Sec):	<input type="text" value="1000"/>
Desired Worst Case RPO (Min.):	<input type="text" value="1"/>
Bandwidth Adjustment Factor:	<input type="text" value="0.35"/>
Retention Window (Days):	<input type="text" value="3"/>

Recommended CX Configuration	
<input type="button" value="View Configuration"/>	Average data change rate: 497026.2771 MBytes/day

Result			
Data Change	Compression	Bandwidth Required (Kbits/sec)	Rpo Achieved
Peak	Yes	1443.6268	No
	No	1979.2029	No
Average	Yes	746.9172	Yes
	No	1060.3227	No

By default, all volumes under protections are included. Simply unselect a volume(s) to exclude it from the final data change profiling / analysis.

**Figure 3-7 Selecting Volumes for Analysis**

Recommended CX Configuration	
<input type="button" value="View Configuration"/>	Average data change rate: 533310.4646 MBytes/day


Pairs Configured													
Include	Source Host: Source Volume	Target Host: Target Volume	Bandwidth Required For RPO ~ = 0 Kbits/Sec (PEAK)		Bandwidth Required For RPO ~ = 0 Kbits/Sec (AVERAGE)		Cumulative data changes(in MBytes)			Average data change rate (MBytes/Sec)		Retention Storage Required (in MBytes)	Target Storage Required (in MBytes)
			With Compression	Without Compression	Average with Compression	Average without Compression	With Compression	Without Compression	Monitoring Interval(Days)	Compression Enabled	Compression Disabled		
<input checked="" type="checkbox"/>	T4-W2K8-SRC6:C	SP-T4-MT-WIN-1:C:\ESX\1461B5F5-CE22-C547-A042A2E8BD384BAC_C	483116.8089	646908.6852	215926.9877	308570.9205	157836.4816	225556.5593	7	0.2610	0.3729	96667.0968	106805.0929
<input checked="" type="checkbox"/>	T4-W2K8-SRC6:C:\SRV	SP-T4-MT-WIN-1:C:\ESX\1461B5F5-CE22-C547-A042A2E8BD384BAC_C__SRV	25.0516	254.5150	6.9271	66.1408	5.0635	48.3471	7	0.0000	0.0001	20.7202	120.7163

Unselecting the volume does not operationally impact the volume, remove the volume from data protection, or prevent data migration, regular backup, etc. It simply excludes the volume from data profiling.

## Scout Server Replication Options

This section outlines the configurable replication options available on the Scout Server. The settings described below are configured on a per-volume basis. [Figure 3-8](#) is a screen capture taken from a protection deployment.

Figure 3-8 Volume Replication Configuration Options

v7.1.0.0.GA.3055 Switch UI 

Monitor » Volumes » Settings  
**Replication Settings**

Statistics Reports **Settings**

Pair Details							
Primary Server	Primary Volume	Remote Server	Target Volume	Process Server	Replication Pool	Fast Resync Unmatched %	Agent Log
T2-LX-SRC-3	/dev/sda	SP-T2-MT-LX-1	/dev/mapper/36000c290836457ccda08f44077f05bcf	t2-ps-1 [ 6.126.99.201 ]	6	N/A	N/A

Pair Settings										
Visible	Resync	Profiling Mode	Secure CX-PS to Destination	Secure Source to CX-PS	Resync Mode	RPD Threshold	Replication Pool (1-24)	Resync Files Threshold (MB)	Differential Files Threshold (MB)	Compression Enable
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Fast	0	5	2048	192	CX-PS Based

Restart Resync Accept Changes Reset

Select a different Process Server

<div style="border: 1px solid #ccc; padding: 2px;">           SP-T2-PS-1 (8.24.71.101)            t2-ps-1 (6.126.99.201)         </div>	Number of Pair Configured 1
---	--------------------------------

Accept Changes

Automatic Resync

Start between hours  :  and  :  after waiting  minutes

Accept Changes

Retention Settings						
Retention	Retention Log size limit (in MB)	Retention Time limit	Log data directory	Disk Space Threshold (%)	Unused Space (in MB)?	On insufficient disk space
Enabled	0.00	1 day	/mnt/retention/Retention_Logs	80	256.00	Purge older logs

Edit Disable Retention

1. Secure Transport: Refer to “IPsec” section on page 3-33.
2. Batch Resync: Refer to “Scout Server Storage and Compute Implementation” section on page 3-6 regarding Resync. This setting controls number of replication errors that can be resynced simultaneously in a protection plan. Typically, this setting should be set based on WAN bandwidth and storage. If WAN throughput is low, it is recommend to limit the number of concurrent resyncs. Default setting is 3 when setting up a protection plan from vContinuum.
3. Automatic Resync: Used when a replication pair is required to address data inconsistencies automatically. A resync is required if any inconsistency exists between the primary and secondary. For example, the primary server reboots causing data in memory to be lost. Resync is required to ensure data consistency. When the Automatic Resync Option is enabled and data inconsistency occurs, the replication pair waits for a certain period of time (by default, 30 minutes) before performing a forced resync within the Start Between Hours time frame. This ensures data consistency and minimizes manual intervention.
4. Use Compression. The three options are:
  - No Compression: Use this option to replicate data without compression.
  - CX-PS based: Use this option to enable data compression on the CX-PS.
  - Source based: Use this option to compress data on the primary server before sending it to the CX-PS.

5. Resync and Differential Thresholds: Refer to [Scout Server Storage and Compute Implementation, page 3-6](#).
6. RPO Threshold: Worst case RPO before an email alerts are sent and RPO alarms are raised.
7. Bandwidth restriction per MT.

## InMage vContinuum

InMage vContinuum is used to set up application discovery, protection, and failback for VMware vSphere VM as well for physical servers. Together with Scout processing (PS), Scout CX, and MT, vContinuum protects VMs as well as physical servers by replicating them to a secondary ESX/ESXi server and recovering them on the secondary VMware vSphere environment when needed. When integrated with application VSS writers, application backup and recovery can be application consistent instead of crash consistent. Refer to InMage's compatibility matrix for information on supported operating system versions.

[Table 3-3](#) displays the platforms that are supported by vContinuum.

**Table 3-3** Continuum-Supported Platforms

Platform	Version Number
vSphere	ESX 3.5, 3.5 U2, 4.0, 4.1 ESXi 3.5, 3.5 U2, 4.0, 4.1, 5.0, 5.0 U1, 5.1
Guest OS	Windows 2003, 2008, 2008 R2, 2012 SLES 9.x, 10.x, 11.x CentOS 4.x, 5.x, 6.x RHEL 4.x, 5.x, 6.x
vCenter	vCenter 4.0, 4.1, 5.0, 5.0 U1, 5.1



### Note

- To provide failover from enterprise to SP, secondary vSphere (SP) version should be either the same or higher than the source (enterprise) vSphere server. To perform a failback from SP to Enterprise, enterprise vSphere version should be either the same or higher than the SP vSphere. vSphere server may need to be upgraded if failback is required.
- For new installations, InMage recommends:
  - Secondary ESXi Platform: ESXi 5.1
  - MT Platform for Windows: Windows 2008 R2 Enterprise Edition
  - MT Platform for Linux: CentOS 6.2 64-bit
  - CX Scout OS: CentOS 6.2 64-bit
  - vContinuum: Windows 2008 R2

Deploying the vContinuum wizard can occur in two major ways:

- In the case of the Linux-only tenant, the wizard has to be installed on a dedicated VM running Windows 2008 R2, Windows7, or XP desktop.
- In the case of Windows, the vContinuum wizard can be installed on the MT or on a dedicated VM running Windows 2008 R2, Windows7, or XP.

Running the vContinuum wizard on top of the MT server reduces the number of touch points and deployment footprint. This is the deployment model implemented in the Cisco Cloud Validation Facility.

## InMage Agent Configuration

InMage agents comes in various flavors depending on the guest OS. The Unified Windows Agent covers most releases for Windows 2003, 2008, and 2012, as well as some Windows XP, Vista, 7, and 8 releases. The specific Windows edition support is highlighted in [Figure 3-9](#).

**Figure 3-9 InMage Windows Edition Support**

Windows Guest Operating Systems								
GUEST OS			EDITION					
OS Version	Bit	Release	Web	Standard	Enterprise	Data Center	Professional	
Windows 2003	32 bit	Base*	✓	✓	✓			
		SP1*	✓	✓	✓			
		SP2	✓	✓	✓			
		R2 SP1		✓	✓			
		R2 SP2		✓	✓			
	64 bit	Base*			✓	✓		
		SP1*			✓	✓		
		SP2			✓	✓		
		R2 SP1			✓	✓		
		R2 SP2			✓	✓		
Windows 2008	32 bit	SP1	✓	✓	✓	✓		
		SP2	✓	✓	✓	✓		
	64 bit	SP1	✓	✓	✓	✓		
		SP2	✓	✓	✓	✓		
		R2	✓	✓	✓	✓		
Windows 2012#	64 bit	Base		✓		✓		
Windows XP	64	SP2					✓	
Windows Vista	32	Base			✓			
	64	Base			✓			
Windows 7	32	Base					✓	
	64	Base					✓	
Windows 8	64	Base					✓	



### Note

- If Windows 2003 (Base, SP1) source machines have applications that require application quiesce. It is then strongly suggested to upgrade to Windows 2003 SP2 to overcome the VSS-related errors.
- Storage Space is not supported. Protect Windows 2012 VM with ReFS Filesystem, requires matching MT.



### Note

UEFI with Dynamic disk will not work.

- Windows XP 32 bit VM on ESX cannot be protected in V2V workflow, but it can be protected using P2V workflow.

- Windows XP 32 /64 recovered VM will not have network drivers installed automatically; the user needs to install manually.

Unified Agent support for Linux is specific to the distribution and release. RHEL 5/6, CentOS 5/6, and SUSE 10/11 are supported. RHEL 5 U3 / CentOS 5 U3 and older versions require network changes to be manually configured after recovery to secondary data center. For Linux physical-to-virtual protection, GRUB bootloader must be used on the source server. Refer to [InMage\\_vContinuum\\_Compatibility\\_Matrix](#) for complete details.

On the Windows platform, the VSS framework is used by vContinuum and Scout to enforce application consistency. The InMage agent interfaces with the VSS writer to quiesce the application and take application consistent snapshots. Refer to [InMage\\_vContinuum\\_Compatibility\\_Matrix](#) for a complete list of certified applications across various operation systems.

There are two ways to install InMage agents onto the primary servers:

1. **Manual Installation:** Network operator can pre-load the Scout agent on each server:
  - Interactive install
  - Command line install (Silent mode)
  - File-based install (Linux)
2. **Bulk Push:** Network operator can push the Scout agent to each server using a UI:
  - CX UI (Linux)
  - vContinuum (Windows)

The first option works well in a smaller deployment in which there are few primary servers.

InMage\_Scout\_Standard\_User\_Guide has excellent step-by-step installation procedures, which will not be repeated in this document. We do want to point out that for Linux agent install, by default, the installer checks for a free space of 2 GB on /root. This can be suppressed by adding the extra switch -k (space in KBs) or -m (space in MBs) or -g (space in GBs).

Bulk push works better for a medium or large deployment as it automates a great number of repetitive clicks and configuration settings that a cloud operator is required to enter. The choice of UI depends on the primary server OS platform. For Linux bulk install, the CX UI is used and for Windows bulk install the vContinuum is used.

This section includes the following topics:

- [CX UI for Linux Bulk Install and Upgrade, page 3-14](#)
- [vContinuum for Windows Bulk Install and Upgrade, page 3-16](#)

## CX UI for Linux Bulk Install and Upgrade

Depending on distribution of Linux, a distribution-specific agent installer needs to first be loaded into the Software Repository. To upload the desired installers, navigate through **Settings > Agent Installers > Software Repository**. Click **Browse** to upload software.

### *Figure 3-10 Home Installer 1*

Once installers are loaded, navigate through **Settings > Agent Installers > Manage Agent Installation/Upgrade**. Click **Install Agent** to provide install requirements. This is a four-step process:

- Step 1** Select **Push Server**. Select the CX server as the push server.
- Step 2** Select **Remote Server**. Enter the IP address range of the primary server(s). Click **Submit** to enter the primary server username/password. Enter the username/password at the top row if the same username/password will be used for all servers; otherwise, enter the username/password next to individual servers.

**Figure 3-11 Home Installer 2**

Settings » Agent Installers » Software Repository  
Software Repository

- Step 3** Set **Install Options**. This is where installation directory, agent type, and CX IP/port is entered. Installation directory defaults to /usr/local/InMage/, Unified Agent (both file and volume agent), and source Scout Agent.

**Figure 3-12 Home Installer 3**

Settings » Agent Installers » Manage Agent Installation/Upgrade  
Manage Agent Installation/Upgrade

IP Address	Remote Server User Name	Remote Server Password	Message
<input checked="" type="checkbox"/>	root	.....	
<input checked="" type="checkbox"/> 6.126.99.208	root	.....	
<input checked="" type="checkbox"/> 6.126.99.209	root	.....	

- Step 4** Click **Run** after you have reviewed information for accuracy.

**Note**

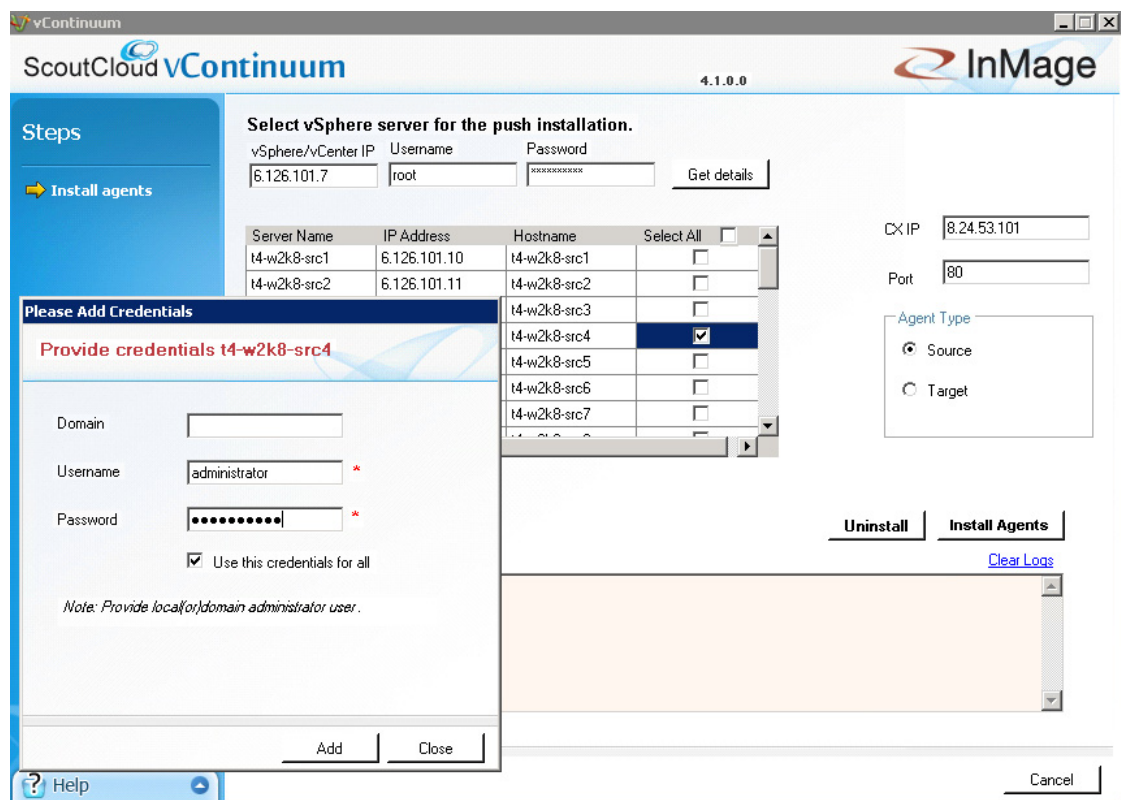
For Linux install, ensure SSH access to the primary server has been enabled.

## vContinuum for Windows Bulk Install and Upgrade

Agent update/install for Windows primary servers is supported from the tenant vContinuum server. vContinuum and CX server pre-bundles compatible windows source agent unlike Linux bulk install. A separate upload into software repository is not required. To install agents, access the vContinuum UI from the vContinuum server and select **Push Agent**. This is a four-step process:

- Step 1** Provide vCenter information.
- Step 2** Select server(s) to install the agent. Currently, the UI does not indicate agent status. It is not possible to determine which server has agent already been installed.
- Step 3** Provide credentials for the server. Default setting is to apply the same credentials to all servers.

**Figure 3-13 vContinuum: Provide Credentials**



- Step 4** vContinuum validates if servers can be reached with provided credentials. Click **Install Agent** to begin the agent install.

A number of outstanding enhancements exist to simplify the user interface. Refer to “[Monitoring, Best Practices, Caveats, and Troubleshooting](#)” section on page 5-1 for details.

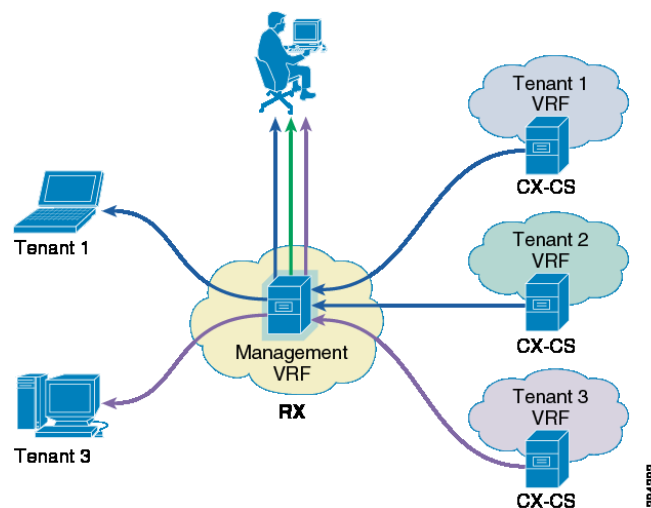


## Multi-Tenant Portal—RX Server

The RX server is a multi-tenant portal for end customers to monitor the protected servers. It also enables the SP to consolidate administration to a single interface rather than a dedicated CX UI for each customer/CX server, greatly simplifying monitoring. The RX server enables:

- Centralized monitoring for all CS servers and source servers for all tenants.
- Central user account management.
- Consolidated reports of protection health of servers, servers details, usage statistics, RPO, and alerts.
- Publish news and announcements by the SP to users.

**Figure 3-14 Multi-Tenant Portal**



As displayed in [Figure 3-14](#), the RX server is an aggregation point for statistics coming from tenant-specific CX-CS servers. Based on role-based access control (RBAC), a tenant user can be limited to monitor statistics from a single CX-CS server while a Managed Services Provider (MSP) user can be assigned to monitor multiple deployments of CX-CS from a single interface.

This section includes the following topics:

- [Multi-Tenant Portal Version Compatibility](#), page 3-17
- [Multi-Tenant Portal User Accounts](#), page 3-18
- [CX Server Registration with Multi-Tenant Portal](#), page 3-20
- [Multi-Tenant Portal Rebranding](#), page 3-21

## Multi-Tenant Portal Version Compatibility

The 7.1 RX server can be either installed on a 64 bit ScoutOS version 5 and Update5, or a ScoutOS version 6 and update 2. When integrating RX with the CX-CS, RX version is always backwards compatible with the CX-CS version. As a rule of thumb, RX should be always higher or equal to the CX-CS version. [Table 3-4](#) is a compatibility table with ScoutOS 6.2 CX.

**Table 3-4** Compatibility Table

	ScoutOS 5.5 RX	ScoutOS 6.2 RX
ScoutOS 5.5 CX	Yes	Yes
ScoutOS 6.2 CX	No	Yes

7.1 RX is compatible with 7.1 GA CX without additional updates/hotfixes. Compatibility with an earlier version of CX release may require additional updates/hotfixes. Refer to InMage Scout Cloud RX release notes for details.

Although the number of users accessing the multi-tenant portal and CX pairing will vary, a minimum of 150G of disk, 1 vCPU, and 4G of memory should be allocated to the VM hosting the RX server.

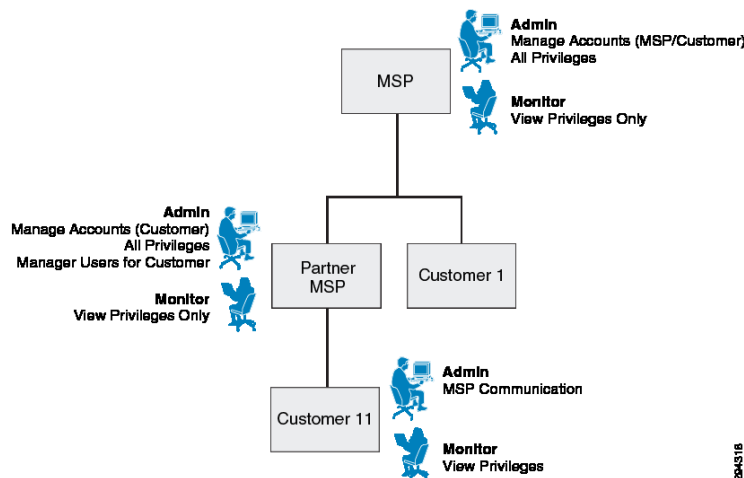
## Multi-Tenant Portal User Accounts

Figure 3-15 lists out various types of user accounts that can be created in the multi-tenant portal. There are two major types of accounts: MSP and Customer:

The root MSP account is created by default with the RX installation; it has full access to all features and data for all customers. Additional sub-MSP accounts can be created in a managed services/white labeling deployment. Account data view would be limited to their own customers.

Customer accounts can be created when new tenants are added.

Both root and sub-MSP accounts can create additional customers and associate the corresponding CXCS server for multi-tenant portal access.

**Figure 3-15** Multi-Tenancy Overview

Users can be set up as monitor user (read only) or administrator (read/write). The ability to perform self-served recovery from the RX portal can be allowed per tenant based on the recovery setting. To enable a user for self-served recovery, first navigate to Accounts > Customer Accounts > Edit customer. Simply select the check box next to Allow to Perform Recovery Options, as shown in [Figure 3-16](#).

Figure 3-16 Self Service Recovery Setting

Accounts » Customer Accounts » Edit Customer

## Edit Customer Account

**Company Details**

Tenant 9	Reference ID: 9
	Contract ID: 9
	Contract Start Date: 05/22/2013
	Contract Expiry Date: 05/01/2014

**User Details**

Full Name	User Name	User Type	Creation Date	Account Status	Action
Tenant 9	tenant9	Administrator	2013-05-22 15:20:17	Active	

**Assigned CS Servers**

IP Address	Host Name
8.24.81.101	sp-t9-ps-1

**Support Contact**

**Recovery Settings**

Allow to Perform Recovery Operations (  Allow Hardware and Network Configuration )

Save Cancel

**Recovery Settings**

Allowed to Perform Recovery Operations with Advanced Options (Hardware and Network Configuration)

The CVD validation focused on SP-managed deployment with direct tenants. [Figure 3-17](#) is a summary view.

Figure 3-17 Summary Tenant View

InMage Systems							
Company	Reference ID	Contract ID	Contract Start Date	Contract Expiry Date	Recovery Allowed?	Account Status	View Dashboard
Tenant 1	1	1	05/22/2013	05/01/2014	✓	✓	
Tenant 10	10	10	05/22/2013	05/01/2014	✓	✓	
Tenant 11	11	11	05/22/2013	05/01/2014	✓	✓	
Tenant 12	12	12	05/22/2013	05/01/2014	✓	✓	
Tenant 2	2	2	05/22/2013	05/01/2014	✓	✓	
Tenant 3	3	3	05/22/2013	05/01/2014	✓	✓	
Tenant 4	4	4	05/22/2013	05/01/2014	✓	✓	
Tenant 5	5	5	05/22/2013	05/01/2014	✓	✓	
Tenant 6	6	6	05/22/2013	05/01/2014	✓	✓	
Tenant 7	7	7	05/22/2013	05/01/2014	✓	✓	
Tenant 8	8	8	05/22/2013	05/01/2014	✓	✓	
Tenant 9	9	9	05/22/2013	05/01/2014	✓	✓	

As displayed in [Figure 3-17](#), in an SP-managed scenario, all tenants are mapped under the default/root MSP account, InMage Systems. CX-CS servers are mapped to each company, and multiple users can exist per company. When a root MSP user logs in, a summary list of tenants along with tenant status, dashboard, and self-recovery status is displayed.

## CX Server Registration with Multi-Tenant Portal

You can point the configuration server to Scout Cloud RX in the following two ways:

- From the RX UI, also known as Pull Method, where Cloud RX pulls the required data from all the registered CX-CS servers OR
- From the CX-CS UI, also known as Pull/Push Method. This is when CX-CS is behind firewall and RX cannot initiate connection to the CX-CS, CX-CS can push data to RX using the Push method.

To register CX-CS with Scout Cloud RX through RX UI:

- 
- Step 1** Navigate to Settings > CS Server.
  - Step 2** Enter the CX IP address or IP Range along with the HTTP port and click Discover.
  - Step 3** CX Servers in the specified IP range are displayed along with the details such as Server IP, Status, and Host Name.
  - Step 4** Enter user name, password, and alias name for CX. This ensures secured CS registration.
  - Step 5** Select the customer account. Select the CS Server and click Register with RX.
- 



### Note

Discover / CX-CS registration is available to MSP admin users only.

To register CX with Scout Cloud RX through CX-CS UI:

- Step 1** Navigate to Settings > RX Server.
- Step 2** Enter the RX IP address and CX alias name. Synchronize data mode defaults to Allow CX to push data to RX.
- Step 3** CX server will be mapped to the "Unassigned CS Servers" pool as shown in [Figure 3-18](#).

**Figure 3-18** Newly Added CX-CS Server In Unassigned Pool

Registered CS Servers										
UnAssigned CS Servers										
	CS Server IP	CS Server Name	CS Alias Name	CS Version	Registration Date	Last Synchronization Time	Synchronization Interval (minutes)	Data Synchronization Method	Communication Type	Action
<input type="checkbox"/>	8.24.85.101	sp-t10-ps-1	sp-t10-ps-1		2013-07-06 15:32:24		5	PUSH	HTTP	

- Step 4** Select the checkbox and click Action to assign the CX-CS server to Customer. Once assigned, the CS server is moved from the unassigned pool to a specific customer as shown in [Figure 3-19](#).

**Figure 3-19** Newly Added CX-CS Server Assigned to Customer

Tenant 10										
	CS Server IP	CS Server Name	CS Alias Name	CS Version	Registration Date	Last Synchronization Time	Synchronization Interval (minutes)	Data Synchronization Method	Communication Type	Action
<input type="checkbox"/>	8.24.85.101	sp-t10-ps-1	sp-t10-ps-1	7.1.0.0	2013-07-06 15:32:24		5	PUSH	HTTP	



**Note**

If a CX-CS server is configured with dual NICs, one is to communicate with remote Scout agents and the other is to communicate with RX. Use the push method instead of pull when registering with the RX server. This is a known limitation and will be addressed in future releases of InMage software.

## Multi-Tenant Portal Rebranding

InMage's Scout Product provides the ability to support user interface customizations as required by its partners. This rebranding allows changing the user interface to carry the partner's company and product name. Also, it allows customizing color and graphics for uniformity across partner products.

Refer to InMage\_Scout\_Cloud\_RX\_Branding document for additional details.

## Summary Tables of Components for All Tenants

This section includes the following topics:

- [InMage Components for Service Provider, page 3-22](#)
- [InMage Component for Enterprise Tenants, page 3-23](#)

## InMage Components for Service Provider

Based on sizing discussion in earlier sections, twelve tenants were configured based on various change rates. [Table 3-5](#) summarizes disk / CPU and memory configuration required for InMage control servers.

**Table 3-5** *Disk/CPU/Memory Configuration (Service Provider)*

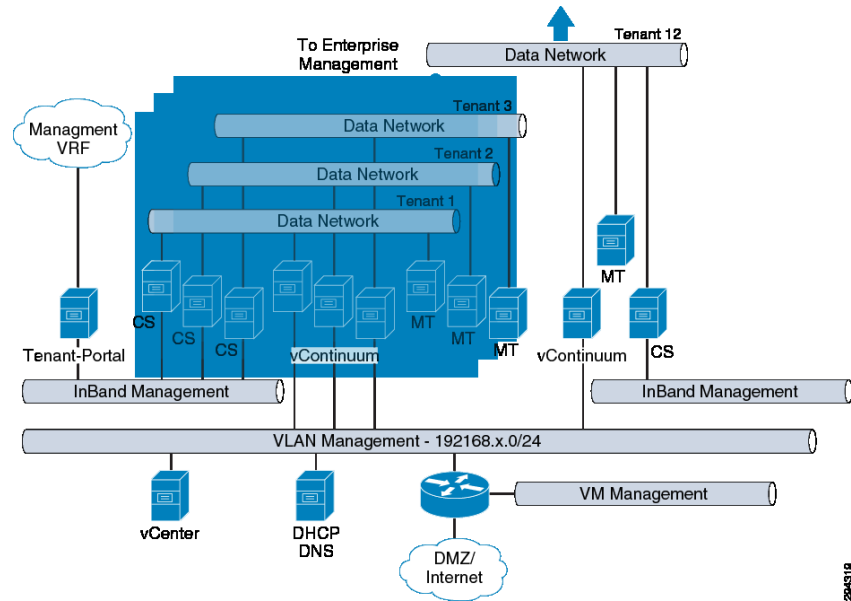
Tenant	CX-PS (GB)	vCPU	RAM (GB)	MT Retention (GB)	MT Cache	vCPU	RAM (GB)
1	800	8	32	2X 300	2X40	2X4	2X16
2	250	4	8	300	40	4	16
3	800	8	32	300	40	4	16
4	800	8	16	300	40	4	16
5	800	8	16	300	40	4	16
6	800	8	16	300	40	4	16
7	800	8	16	300	40	4	16
8	800	8	16	300	40	4	16
9	250	4	8	50	40	2	8
10	250	4	8	50	40	2	8
11	250	4	8	50	40	2	8
12	250	4	8	50	40	2	8

- Tenants 1 and 3 are sized based on 1TB change rate per day.
- Tenants 4, 5, 6, 7 and 8 are sized based on 750GB change rate per day.
- Tenants 2, 9, 10, 11, and 12 are sized based on 350G change rate per day.

MT retention is based on time (9 hours) and space capacity. With the exception of Tenant 2, all tenants are Windows-based tenants running Windows 2008 R2. vContinuum is installed directly on the MT. Tenant 2 is a Linux-only tenant, running CentOS 6.3; as such, a dedicated vContinuum server running windows 2008R2 with 1 vCPU, 3G memory, and 50G disk was utilized. With the exception of Tenants 9 - 12, all MTs are deployed with 4 vCPU, 16G memory, and 300G retention drive.

Deployment topology and logical connectivity between InMage control servers are captured in [Figure 3-20](#).

Figure 3-20 Deployment Topology and Logical Connectivity between InMage Control Servers



From the SP VPC, communication with enterprise management is via inband data NIC part of VMDC server VLAN. A common out-of-band (OOB) management VLAN was extended from the Cloud Validation Facility management POD to the VMDC 2.3 infrastructure. OOB management VLAN is used to communicate with shared resources residing in the management PoD such as vCenter, Active Directory, vCenter Database, and DHCP server.

## InMage Component for Enterprise Tenants

Twelve Enterprise tenants were configured to mirror the SP setup. Table 3-6 summarizes Disk / CPU and memory configuration required for enterprise InMage component. Retention policy, vCPU, and RAM requirements mirror that of the SP site based on 1TB, 750GB, and 350GB change rate/day. MT is deployed in the enterprise for failback protection. Enterprise master target is expected to be idle and not needed until failback after a disaster recovery. It can be pre-provisioned or installed during time of failback. Each tenant will have a dedicated vCenter server managing ESXi resources. Tenant 1 and Tenant 2 have both physical and virtual servers. Both P2V and V2V protection are needed. To generate the desired data change rate, each enterprise is assigned a set of machine machines (Table 3-6).

Table 3-6 Disk/CPU/Memory Configuration (Enterprise)

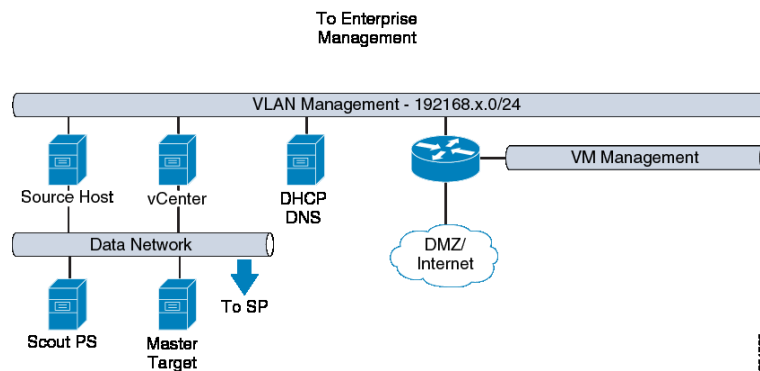
Tenant	PS (GB)	vCPU	RAM (GB)	MT Retention (GB)	MT Cache (GB)	vCPU	RAM (GB)	vCenter disk (GB)	vCPU	RAM (GB)	Total VM
1	800	8	32	2 X 300	2 X 40	2X 4	2 X 16	100	2	4	60
2	250	4	8	300	40	4	16	100	2	4	10
3	800	8	32	300	40	4	16	100	2	4	50
4	800	8	16	300	40	4	16	100	2	4	30
5	800	8	16	300	40	4	16	100	2	4	30
6	800	8	16	300	40	4	16	100	2	4	30
7	800	8	16	300	40	4	16	100	2	4	30

Table 3-6 Disk/CPU/Memory Configuration (Enterprise) (continued)

Tenant	PS (GB)	vCPU	RAM (GB)	MT Retention (GB)	MT Cache (GB)	vCPU	RAM (GB)	vCenter disk (GB)	vCPU	RAM (GB)	Total VM
8	800	8	16	300	40	4	16	100	2	4	30
9	250	4	8	50	40	2	8	100	2	4	7
10	250	4	8	50	40	2	8	100	2	4	7
11	250	4	8	50	40	2	8	100	2	4	8
12	250	4	8	50	40	2	8	100	2	4	8

Deployment topology and logical connectivity between InMage Enterprise control servers are captured in Figure 3-21.

Figure 3-21 Deployment Topology and Logical Connectivity between InMage Enterprise Control Servers



From the Enterprise VPC, communication with the SP management is via the inband data NIC part of VMDC server VLANs. A common OOB management VLAN was extended from the Cloud Validation Facility Management PoD to the VMDC 2.2-based enterprise VPC. OOB management VLAN is used to communicate with shared resources residing in the Management PoD such as vCenter, Active Directory, vCenter Database, and DHCP server.

## VMDC 2.3

This section includes the following topics:

- [VMDC 2.3 Integrated Compute and Storage Stack, page 3-25k](#)
- [Mapping DR Components to VMDC 2.3 Containers, page 3-32](#)
- [Tenant Configuration, page 3-33](#)
- [Connectivity across the WAN, page 3-39](#)
- [Storage Configuration, page 3-42](#)



## VMDC 2.3 Integrated Compute and Storage Stack

The VMDC 2.3 release uses modular blocks for compute and storage, generically referred to as Integrated Compute and Storage (ICS) stacks. Several stacks can be attached to a PoD, providing compute and storage scale. With VMDC 2.3, three ICS stacks can be connected to the Nexus 7004, with 4x 10G links per aggregation switch, for a total of 80 Gbps to the ICS switch layer. Refer to the Cisco VMDC 2.3 Design Guide at <http://www.inwats.cisco.com/publications/viewdoc.php?docid=6637> for more discussion of the scaling factors.

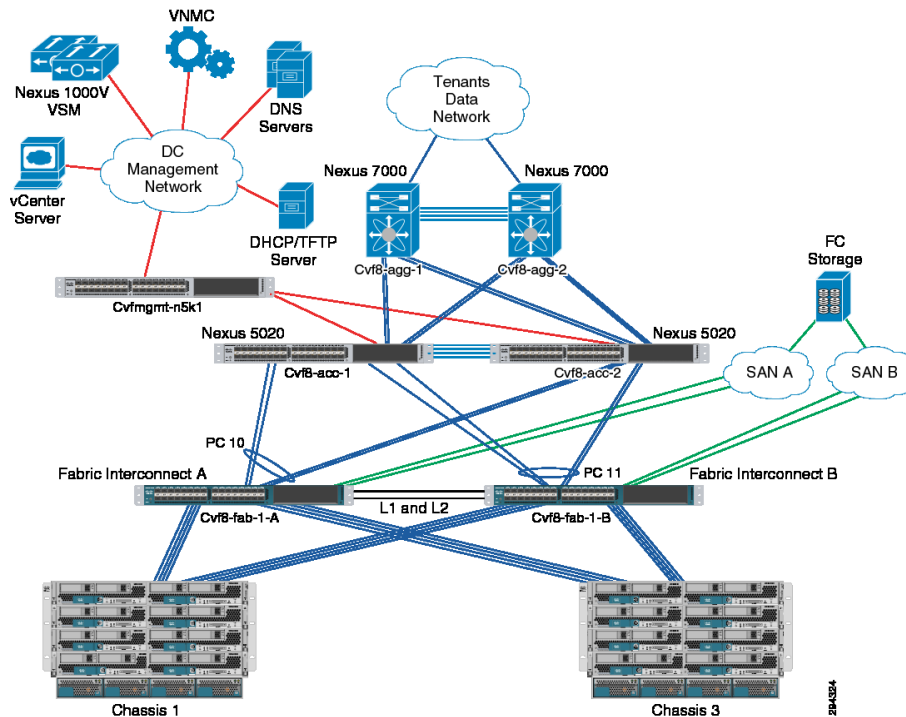
In our implementation, a smaller footprint ICS was built as listed in [Table 3-7](#)

**Table 3-7 ICS Footprint**

Tenant Type	Number of Tenants	Number of VLANs per Tenant	Number of VMs
Gold	4	3	30
Silver	2	3	60
Bronze	6	1	160
Total	12	24	250

The ICS design uses the VNX 5500 as the SAN storage. The details of the ICS buildout are covered in [Figure 3-22](#).

**Figure 3-22 ICS Buildout**



This section includes the following topics:

- [UCS Implementation, page 3-26](#)

- [ESXi Implementation, page 3-28](#)
- [Nexus 1000V, page 3-30](#)
- [VSG Implementation, page 3-32](#)

## UCS Implementation

[Table 3-8](#) shows UCS components for implementation.

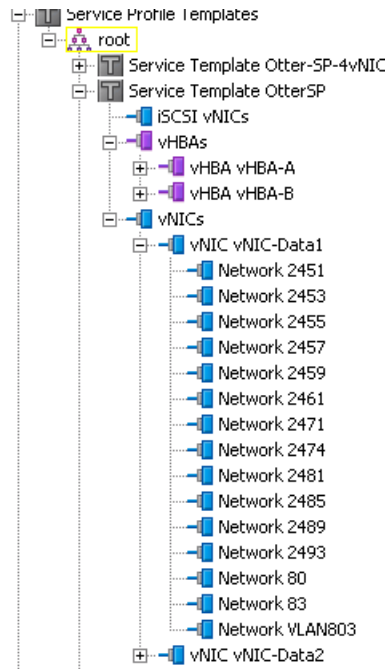
**Table 3-8 UCS Implementation**

Component	Product Name	Quantity
Fabric Interconnect (FI)	Cisco UCS 6120	2
Chassis	Cisco UCS 5108	2
I/O Module	Cisco UCS 2104XP	4
Blade Server	Cisco UCS B200 M3 (2 x 8 cores CPU, 196GB Memory)	3
Blade Server	Cisco UCS B200 M2 (2 x 6 cores CPU, 96GB Memory)	9
Adapter	Cisco UCS VIC 1240	3
Adapter	Cisco UCS M81KR	9

Two UCS 5108 chassis are connected to a pair of UCS 6120 FIs. Each chassis has four server links to each FI. The UCS FIs are configured in End Host (EH) mode into a cluster to provide active/standby management plane redundancy for the UCSM, active/active for data forwarding. The uplinks on the FIs are bundled into port-channels to the upstream Nexus 5000 switch. Both management and data traffic are carried in the same port-channel. Nexus 5000 switches with Fibre Channel (FC) links for access to SAN storage. Each UCS blade is configured with two vHBAs for access to SAN storage via SAN-A and SAN-B for storage multipathing.

The UCSM service-profile in [Figure 3-23](#) is used on the UCSM.

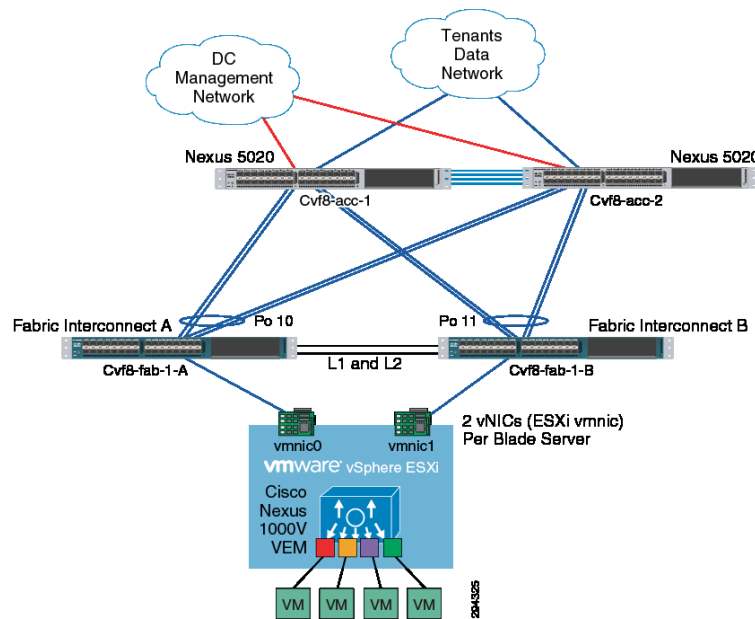
Figure 3-23 UCSM Service-Profile



Each service-profile is associated with a respective server pool and blades from both chassis are made available to the server pool.

The UCS Fabric Interconnects are connected to a pair of Nexus 5000 for redundancy. Both FIs are configured with the same set of VLANs. All VLANs are trunked to all available uplinks and the same uplink port channel is used to carry both management and tenant data traffic.

Figure 3-24 UCS Implementation



## ESXi Implementation

A VMware vSphere Cluster is a grouping of servers with similar characteristics and can be used as one unified compute resource. VMware vSphere Cluster is the foundation used to achieve a pooling of resources, HA, and Distributed Resource Scheduling.

Refer to the following documents for more information on VMware HA and VMware Distributed Resource Scheduler:

- HA Deepdive at <http://www.yellow-bricks.com/vmware-high-availability-deepdiv/>
- Distributed Resource Scheduler Deepdive at <http://www.yellow-bricks.com/drs-deepdive/>

The following set of clusters were created based on the recommendations provided in VMDC 2.3 Design Guide at [http://www.cisco.com/en/US/docs/solutions/Enterprise/Data\\_Center/VMDC/2.3/design\\_guide/VMDC\\_2.3\\_DG\\_1.html#wp1361952](http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VMDC/2.3/design_guide/VMDC_2.3_DG_1.html#wp1361952).

**Table 3-9 Host per Cluster**

Cluster Type	Number of Hosts	Memory	CPU
Bronze	5	900GB	207 GHz
Silver	3	500 GB	129 GHz
Gold	4	700 GB	165 HZ

It is recommended to size the number of hosts per cluster based on capacity and HA requirements of each individual implementation. A minimum of three hosts is recommended to provide non-disruptive host maintenance without loss of VMware HA protection. Cluster size varies from 3 ESXi hosts to 5 ESXi depending on workload type.

The following set of VMware HA parameters is necessary to define capacity requirements and isolation response:

- Admission Control
- Number of Host Failures Permitted
- Restart Priority
- Isolation Response
- Host Monitoring Status

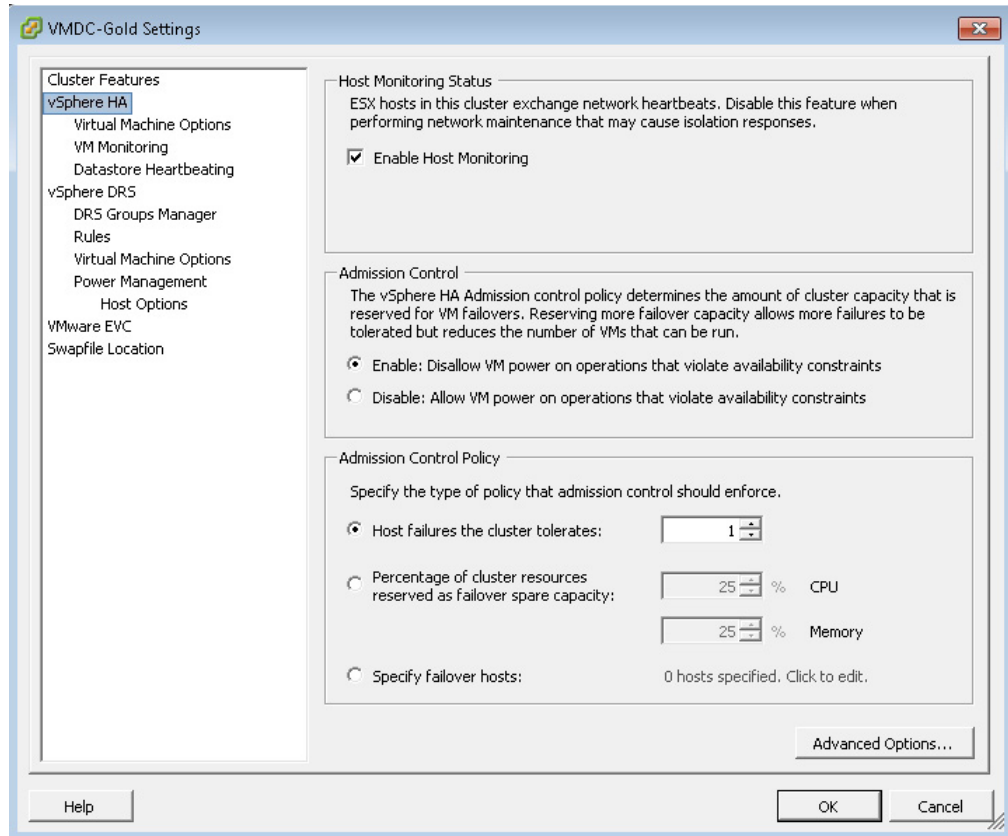
Based on the recommendations provided in VMDC , all three clusters are assigned the same HA parameters. [Table 3-10](#) provides the sample settings for the Gold Cluster.

**Table 3-10 Sample Settings for the Gold Cluster**

Category	Setting
Admission Control	Enabled: Do not power on VMs that violate availability constraints
Number of Host Failures Permitted	1 host failures cluster tolerates
Restart Priority	Medium
Isolation Response	Leave VM powered on
Host Monitoring Status	Enabled

[Figure 3-25](#) shows the Gold Cluster parameters.

Figure 3-25 Gold Cluster HA -1



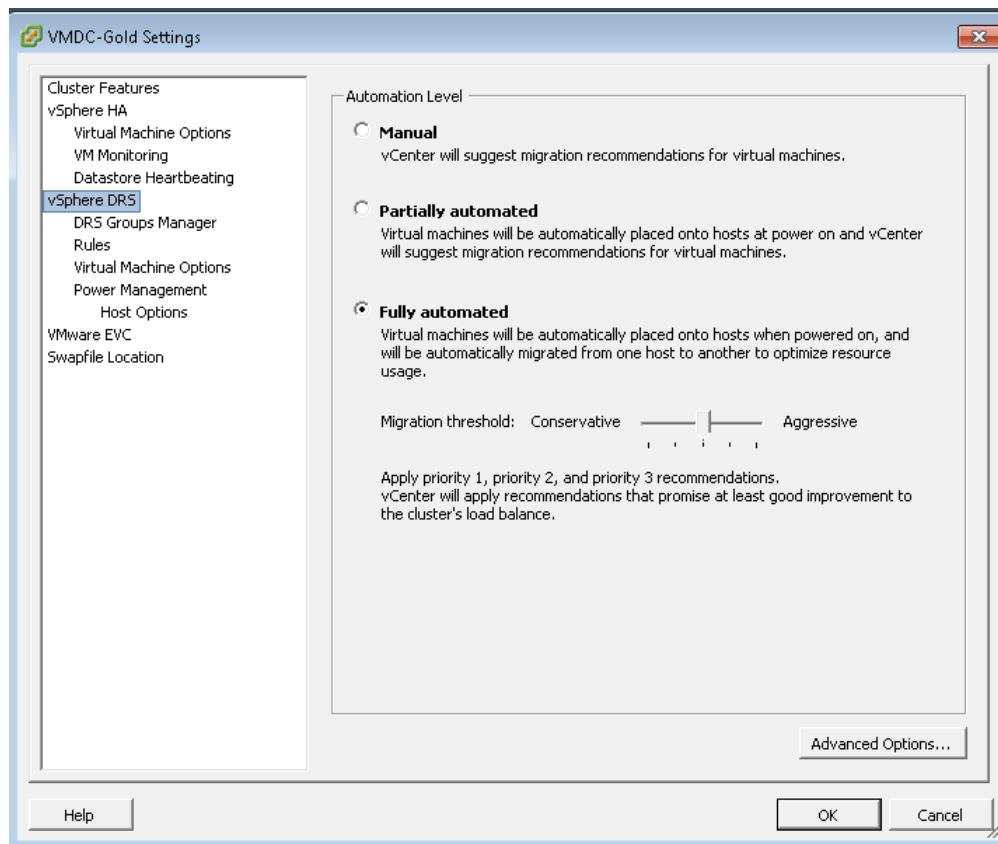
The VMware Distributed Resource Scheduler functions by monitoring the VM (CPU and memory) loads in a virtual computer cluster and, if necessary, moves the VMs from one physical ESX server to another in an attempt to load balance the workload. Distributed Resource Scheduler works in one of three modes: fully automatic, partially automatic, or manual. Based on the recommendations provided in the VMDC 2.3 Design Guide at [http://www.cisco.com/en/US/docs/solutions/Enterprise/Data\\_Center/VMDC/2.3/design\\_guide/VMDC\\_2.3\\_DG\\_1.html#wp1361952](http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VMDC/2.3/design_guide/VMDC_2.3_DG_1.html#wp1361952), all three clusters are assigned the same Distributed Resource Scheduler parameters.

Table 3-11 Distributed Resource Scheduler Parameters

Distributed Resource Scheduler Mode	Fully Automated
Migration Threshold	3 stars

Figure 3-26 shows the Gold Distributed Resource Scheduler setting.

Figure 3-26 Gold Cluster HA -1



vSphere supports cluster sizes of up to 32 servers when HA and/or DRS features are utilized. In general practice, however, the larger the scale of the compute environment and the higher the virtualization (VM, network interface, and port) requirement, the more advisable it is to use smaller cluster sizes to optimize performance and virtual interface port scale. Therefore, in large VMDC deployments, cluster sizes are limited to eight servers; in smaller deployments, cluster sizes of 16 or 32 can be utilized. Gold, Silver, and Bronze compute profiles are created to represent Large, Medium, and Small workload types. Gold has one vCPU/core and 16 GB RAM, Silver has 0.5 vCPU/core and 8 GB RAM, and Bronze has 0.25 vCPU/core and 4 GB of RAM.

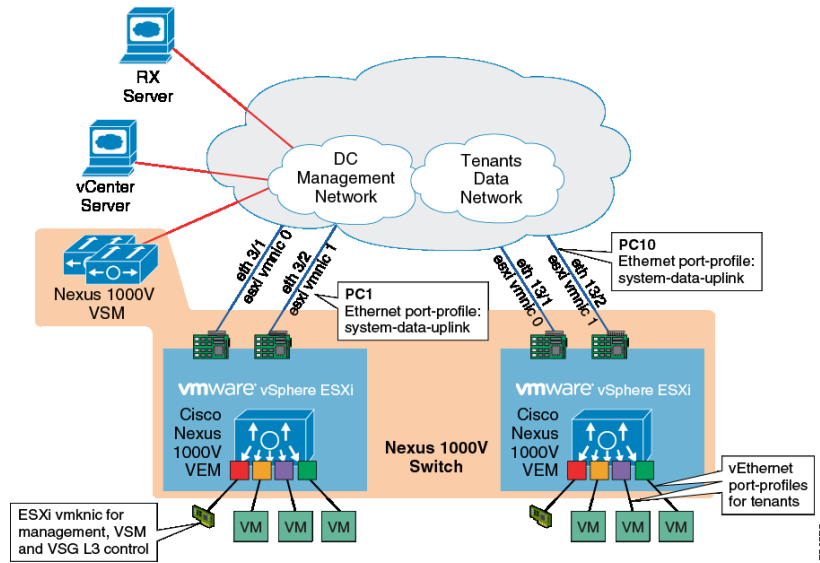
While the VMDC 2.3 architecture works with Vblocks and FlexPods, the system has been validated with VNX 5500.

## Nexus 1000V

The Nexus 1000V Series Switches provide a comprehensive and extensible architectural platform for VM and cloud networking. In our implementation, all networking needs of the VMs are provided by Nexus 1000V Series Switches and it is implemented identically to the VMDC 2.3 specification.

Without reduplicating VMDC 2.3 documentation, [Figure 3-27](#) is a summary description of the implementation.

Figure 3-27 Summary Description of Nexus 1000v Implementation



The Nexus 1000V VSM is configured in L3 SVS mode. In L3 SVS mode, VSM encapsulates the control and packet frames into User Datagram Protocol (UDP) packets. The VSM uses its mgmt0 interface to communicate with the VEMs. The VEMs are located in a different IP subnet from the VSM mgmt0 interface. On each VEM, the vmk0 vmkernel interface is used to communicate with the VSM. The following configuration shows the VSM svcs-domain configuration:

```
svs-domain
domain id 1
control vlan 1
packet vlan 1
svs mode L3 interface mgmt0
```

The UCS is configured with EHV mode and has the upstream L2 switches performing the split between the management and customer production data domains. Each ESXi/VEM host is configured with two NICs (also referred to as the ESXi VM Network Interface Card (VMNIC) or UCS vNIC), carrying both management network and tenants' data network (for UCS Fabric A - fabric B redundancy). On the Nexus 1000V, the following configuration shows the Ethernet port-profile configuration:

```
port-profile type ethernet system-data-uplink
vmware port-group
switchport trunk allowed vlan 80,83,2451,2453,2455,2457,2459,2461,2471
switchport trunk allowed vlan add 2474,2481,2485,2489,2493
switchport mode trunk
switchport trunk native vlan 83
channel-group auto mode on mac-pinning
no shutdown
system vlan 83
max-ports 32
state enabled
```

When the ESXi host is added to the Nexus 1000V DVS, the vmnic0 and vmnic1 interfaces are attached to the system-data-uplink Ethernet uplink port profile. In this implementation, the vmknic ESXi kernel interfaces (vmk0) are also managed by the Nexus 1000V. The following shows the configuration used for the ESXi management.

```
port-profile type vethernet esxi-mgmt-vmknic
capability l3control
vmware port-group
switchport mode access
pinning id 0
```

```

switchport access vlan 83
no shutdown
system vlan 83
max-ports 32
state enabled

```

Refer to VMDC 2.3 (Nexus 1000V Series Switches at [http://www.cisco.com/en/US/partner/docs/solutions/Enterprise/Data\\_Center/VMDC/2.3/implementation\\_guide/VMDC2.3\\_IG2.html#wp2272450](http://www.cisco.com/en/US/partner/docs/solutions/Enterprise/Data_Center/VMDC/2.3/implementation_guide/VMDC2.3_IG2.html#wp2272450)) for additional details.

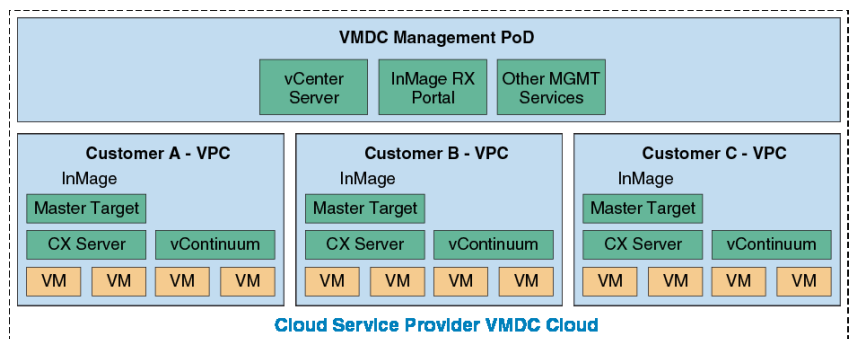
## VSG Implementation

For information about VSG implementation, please see: [http://www.cisco.com/en/US/partner/docs/solutions/Enterprise/Data\\_Center/VMDC/2.3/implementation\\_guide/VMDC2.3\\_IG5.html#wp2277285](http://www.cisco.com/en/US/partner/docs/solutions/Enterprise/Data_Center/VMDC/2.3/implementation_guide/VMDC2.3_IG5.html#wp2277285)

## Mapping DR Components to VMDC 2.3 Containers

As previously discussed in “InMage Components for Service Provider” section on page 3-22, a common management PoD is deployed to host shared services (example: vCenter, DHCP, and DNS). The multi-tenant portal (RX) is part of the management PoD. The remaining InMage specific components are co-located with the production ICS cluster corresponding to service tier (Gold / Silver / Bronze). Refer to [Figure 3-28](#).

**Figure 3-28** VMDC Management PoD



InMage RX server will be deployed with dual NICs:

- Management NIC: Used to communicate with the Configuration Server (CX).
- OOB Management NIC: Server VLAN part of a dedicated VMDC 2.3 container used for OOB tenant access. Refer to “[Out of Band Management Portal Access](#)” section on page 3-34.

InMage CX-CS servers will be deployed with dual NICs:

- Management NIC: Used to communicate with the Multi-Tenant Portal (RX).
- Data NIC: VMDC 2.3 Server VLAN: Used for Scout Agent, MT, and PS server registration and configuration updates.

Master Target: Two deployment scenarios are based on OS:

- InMage Windows MT is deployed with dual NICS because vContinuum is colocated with the MT:
  - Management NIC; used by vContinuum to communicate with shared vCenter in the SP Cloud.
  - Data NIC; used by vContinuum to communicate with tenant vCenter in the enterprise private cloud., and to receive data changes from the enterprise private cloud.



- InMage Linux MT will be deployed with a single NIC: – Data NIC:
  - Used to receive data change from Enterprise Private Cloud - Communication with vCenter is from vContinuum.

Virtual Machines - VM settings are configured during recovery, refer to [Chapter 4, “Recovery Workflows,”](#) for details. Based on the VMDC specification, each VM will have two NICs:

- Management NIC—Protected by the VSG and accessible by a cloud orchestration system, such as BMC or CIAC.
- Data NIC—VMDC 2.3 Server VLAN.

## Tenant Configuration

This section includes the following topics:

- [IPsec, page 3-33](#)
- [Out of Band Management Portal Access, page 3-34](#)
- [VMDC Container Modifications, page 3-35](#)

### IPsec

When sending data across the WAN from a primary site to a secondary site, securing the data transmission is critical. Data may be left vulnerable if encryption between the CX-PS server to a MT in the secondary site is not enabled. Two encryption points are possible in the architecture:

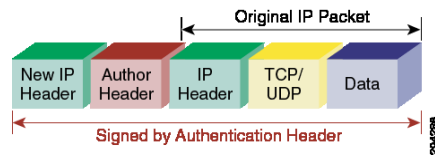
- Primary Server: Encrypt data from the source primary server to local InMage CX-PS.
  - This option requires the InMage agent (DataTap) on the primary server under protection to encrypt the data before sending it to the processing server. Scout PS server does not support protection of a device when it is encrypted outside of InMage. This approach will consume CPU/memory and potentially other resources on the production server. InMage recommends to not use this option unless deployment topology prevents a processing server from being deployed.
- Processing Server: Encrypt data from InMage CX-PS to MT before transmitting the packet out of the enterprise WAN.
  - This option is preferred as it assumes encryption is performed by an external device, not the production server. A number of ways are possible for accomplishing this. InMage supports data encryption within the processing server or external encryption appliance such as the Cisco ASA. Either option will secure the data transmission; however, it is important to remember if the processing server is encrypting traffic across the WAN. Performance penalties can occur on the processing server when compared to unencrypted transmissions.

To simplify capacity management and operational support, external ASA appliances were used in our implementation. ASA pairs deployed in multi-context mode are used to secure the communication between the Enterprise LAN and SP LAN across a L3 MPLS VPN. The benefits of ASA are:

- Multi-context support: Tenant separation and delegation of role / responsibility.
- Capacity monitoring and management: Single pair of ASA VSXs deploying additional vCPU/ MEM to each processing server to support encryption.
- Operational Monitoring: Single tunnel per customer vs. encryption setting per disk/volume under protection

All traffic needing encryption between the enterprise and service are redirected to the IPsec tunnel based on static routing. ASA appliance will encrypt the entire frame and re-encapsulate it with an additional IPsec header:

**Figure 3-29** IPsec Header

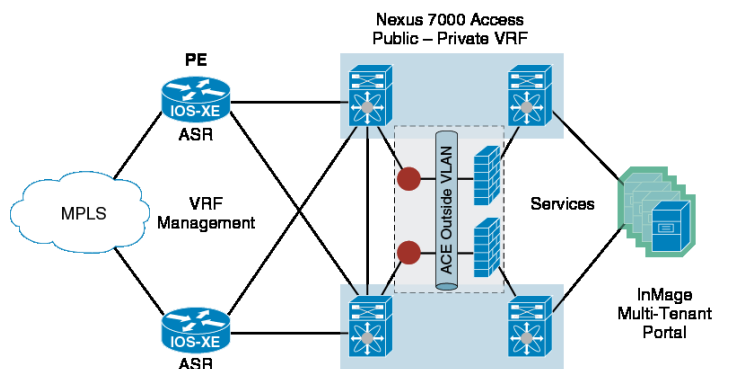


Refer to “[VMDC Container Modifications](#)” section on page 3-35 Updates for details on routing changes.

## Out of Band Management Portal Access

The topology in [Figure 3-30](#) is used to validate the functions and features of the InMage Multi-tenant Portal.

**Figure 3-30** Out of Band Management Portal Access Topology



All DRaaS users share a VMDC 2.3 Gold container for management access. A Gold container is used as it offers the guaranteed resources as well as flexibility in network container resources such as firewall and server load balancing servers.

The Enterprise route targets are imported on the Cisco ASR 1000 for MPLS connectivity across the WAN.

For example:

```
vrf definition mgmt-gold-pub rd 800:800
route-target export 800:800 route-target import 800:800 route-target import 2451:2451
route-target import 3003:3003 route-target import 2471:2471 route-target import 3008:3008
route-target import 2474:2474 route-target import 3015:3015 route-target import 2453:2453
route-target import 3019:3019 route-target import 2455:2455 route-target import 3020:3020
route-target import 2457:2457 route-target import 3023:3023 route-target import 2459:2459
route-target import 3027:3027 route-target import 2461:2461 route-target import 3029:3029
route-target import 3482:3482 route-target import 3010:3010 route-target import 3040:3040
route-target import 3042:3042 route-target import 3034:3034
```

The ASR 1000 peers downstream with the Cisco Nexus 7000 aggregation to learn specific prefixes required for management access.

Traffic flows from the Enterprise to the Cisco ASR 1000 via the MPLS core, and the Cisco ASR 1000 follows the path of the VMDC 2.3 Gold container until it arrives on the Cisco Nexus 7000 aggregation public VRF. From the Cisco Nexus 7000 aggregation, traffic is sent across the firewall to the load balancer for server load balancing (if needed).

## VMDC Container Modifications

This section includes the following topics:

- [VMDC Gold, page 3-35](#)
- [VMDC Silver, page 3-36r](#)
- [VMDC Bronze, page 3-38](#)

### VMDC Gold

Changes to the VMDC Gold container were implicit to the overall architecture. No modifications or changes were made to the baseline VMDC Gold container; server VLANs were neither introduced nor removed. Non-InMage/DRaaS-related traffic streams follow identical network paths as documented in the VMDC 2.3 Design and Implementation guide (<http://www.in-wats.cisco.com/publications/viewdoc.php?docid=6637>). Traffic sourced from the VPC will first be routed to the private VRF default gateway based on default routing. Traffic will be forwarded from the private VRF to the public VRF across the vFW. Once traffic arrives in the public VRF, it will be L3 routed to the ASR1K (PE) towards the L3 VPN.

Traffic to and from the InMage server from the primary enterprise site needs to be forwarded to the IPsec tunnel for encryption or decryption.

- This was accomplished by setting up a static site-to-site tunnel between the SP and enterprise, placing the IPsec-inside interface on Gold server VLAN 1 and the outside interface on Gold server VLAN 2.
- InMage traffic from the enterprise LAN will be sent to the ASA-outside interface residing on server VLAN 2.
- Once received traffic is decrypted by the ASA, it will be forwarded to InMage servers residing on server VLAN 1 via the IPsec-inside interface.
- InMage traffic from the SP to Enterprise is accomplished by adding static routes on the CX and MT server pointing to the inside interface of the ASA.
- Once received traffic is encrypted by the ASA, it will be forwarded out of the ASA-outside interface residing on server VLAN 2 towards the HSRP address of the private VRF interface.
- Once encrypted traffic is received in server VLAN 2, normal VMDC traffic flow occurs.

See [Figure 3-31](#) for details.



- Traffic to and from InMage server from the primary enterprise site needs to be forwarded to the IPsec tunnel for encryption or decryption.
- This was accomplished by setting up a static site-to-site tunnel between the SP and enterprise, placing the IPsec-inside interface on the Silver server VLAN 1 and the outside interface on the Silver server VLAN 2.
- InMage traffic from the enterprise LAN will be sent to the ASA-outside interface residing on server VLAN 2.
- Once received traffic is decrypted by the ASA, it will be forwarded to InMage servers residing on server VLAN 1 via the IPsec-inside interface.
- InMage traffic from SP to Enterprise is accomplished by adding static routes on the CX and MT server pointing to the inside interface of the ASA.
- Once received traffic is encrypted by the ASA, it will be forwarded out of the ASA-outside interface residing on server VLAN 2 towards the HSRP address of the private VRF interface.
- Once encrypted traffic is received in server VLAN 2, normal VMDC traffic flow occurs. Refer to [Figure 3-32](#) for details.

**Figure 3-32 V2V Traffic Flow (Silver Container)**

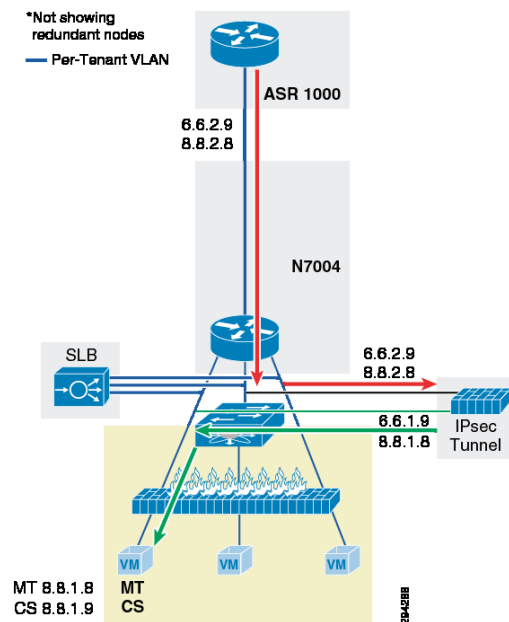


Figure 3-32 is an example of V2V traffic flow between the Enterprise InMage PS server and the SP MT:

- From 6.6.1.9 (Enterprise InMage processing server).
- Destination 8.8.1.8 (SP InMage MT).
- Tunnel Source—Enterprise ASA IPsec-outside interface: 6.6.2.9.
- Tunnel Destination—SPO ASA IPsec-outside interface: 8.8.2.8.

### Enterprise

1. PS receives changes from data tap on the primary server.
2. PS sends traffic (6.6.1.9, 8.8.1.8) to IPsec router.
3. IPsec router encrypts the traffic and sends re-encapsulated packet with (6.6.2.9, 8.8.2.9) to SP.

**Service Provider**

1. IPsec-encrypted traffic with (6.6.2.9, 8.8.2.9) is received by ASR 1000.
2. ASR 1000 forwards (6.6.2.9, 8.8.2.9) to the aggregation Nexus 7000 customer VRF.
3. (6.6.2.9, 8.8.2.9) is sent to ASA IPsec-outside interface for decryption.
4. ASA decrypts the packet, strips header (6.6.2.9, 8.8.2.9) and sends re-encapsulated packet with (6.6.1.9, 8.8.1.8) to InMage MT.

**VMDC Bronze**

Unlike Gold and Silver, we had to make some modifications to the baseline VMDC Bronze container to support encryption with IPsec. This is because unlike Gold and Silver container, Bronze container has only a single server VLAN. It wasn't possible to set up a site-to-site tunnel with only a single server VLAN; an additional IPsec-outside interface is required. This interface could connect directly to the ASR or to the aggregation 7004. Connecting the IPsec interface to the ASR introduced a number of fundamental design changes; it was much simpler to introduce a dedicated SVI interface on the 7004 and extend it to the ASA. The benefits of this approach are the following:

- VMDC Alignment: Minimal changes to VMDC baseline container models.
- Operation Consistency: IPsec configuration is nearly identical among all VMDC container types.
- Single Point of Resource Management: Only VLAN resources on the N7k are required. Does not introduce any changes to ASR or ASA.

Non-InMage/DRaaS-related traffic streams follow identical network paths as documented in the VMDC 2.3 design and implementation guide (<http://www.in-wats.cisco.com/publications/viewdoc.php?docid=6637>). Traffic sourced from the VPC will first be routed to customer VRF default gateway. Once traffic arrives in the customer VRF, it will be L3 routed to the ASR 1000 (PE) toward the L3 VPN.

Traffic to and from InMage server from the primary enterprise site needs to be forwarded to the IPsec tunnel for encryption or decryption.

- This was accomplished by setting up a static site-to-site tunnel between the SP and enterprise, placing the IPsec-inside interface on the Bronze server VLAN 1 and the outside interface on a newly created SVI interface between the ASA and aggregation Nexus 7004.
- InMage traffic from the enterprise LAN will be sent to an ASA-outside interface.
- Once received traffic is decrypted by the ASA, it will be forwarded to InMage servers residing on the server VLAN 1 via the IPsec-inside interface.
- InMage traffic from SP to enterprise is accomplished by adding static routes on the CX and MT server pointing to the inside interface of the ASA.
- Once received traffic is encrypted by the ASA, it will be forwarded out of the ASA-outside interface. Encrypted traffic follows normal VMDC traffic path once received by Nexus 7004.

See [Figure 3-33](#) for details.

Figure 3-33 V2V Traffic Flow (Bronze Container)

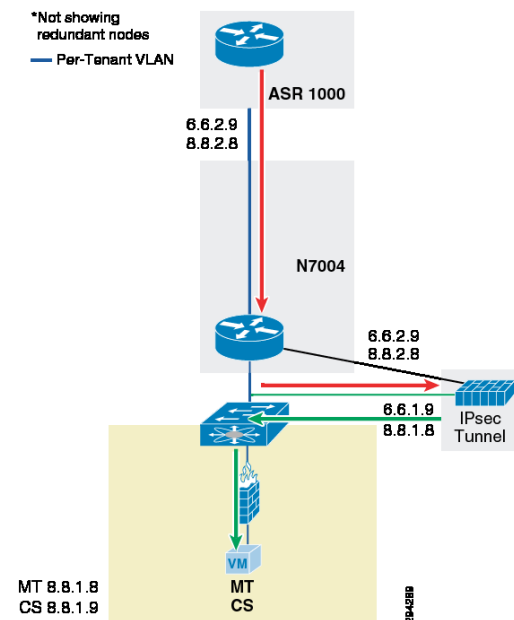


Figure 3-33 is an example of V2V traffic flow between the Enterprise InMage PS server and the SP MT:

- From 6.6.1.9 (Enterprise InMage processing server).
- Destination 8.8.1.8 (Service provider InMage MT).
- Tunnel Source—Enterprise ASA IPsec-outside interface: 6.6.2.9.
- Tunnel Destination—SP ASA IPsec-outside interface: 8.8.2.8.

#### Enterprise

1. PS receives changes from data tap on the primary server.
2. PS sends traffic (6.6.1.9, 8.8.1.8) to IPsec router.
3. IPsec router encrypts the traffic and sends re-encapsulated packet with (6.6.2.9, 8.8.2.9) to SP.

#### Service Provider

1. IPsec encrypted traffic with (6.6.2.9, 8.8.2.9) is received by ASR 1000.
2. ASR1K forwards (6.6.2.9, 8.8.2.9) to the aggregation Nexus 7000 customer VRF.
3. (6.6.2.9, 8.8.2.9) is sent to ASA IPsec-outside interface for decryption.
4. ASA decrypts the packet, strips header (6.6.2.9, 8.8.2.9) and sends re-encapsulated packet with (6.6.1.9, 8.8.1.8) to InMage MT.

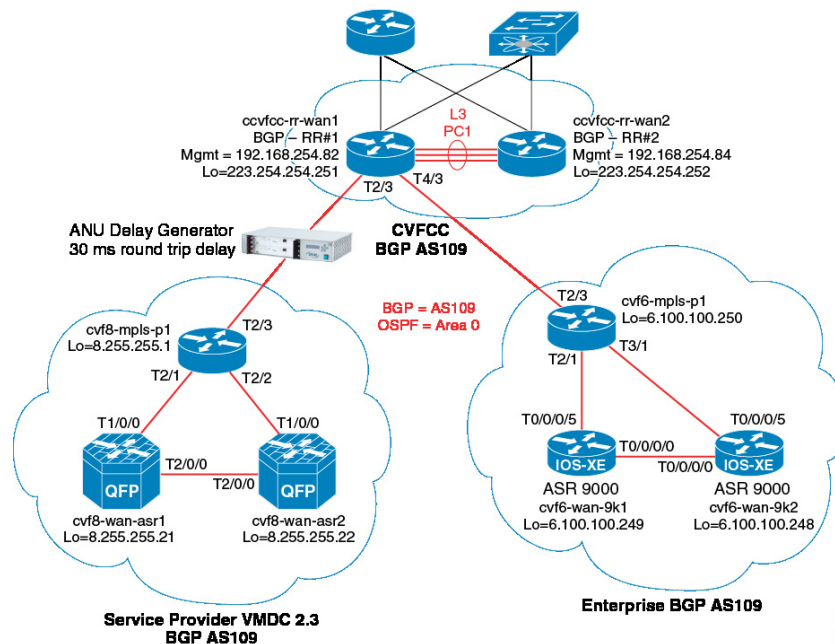
## Connectivity across the WAN

The enterprise is connected to the SP through a MPLS-VPN (L3VPN providing connectivity through an MPLS core) network for data plane connectivity. This VPN network provides connectivity in the data path between the private and public clouds for users and applications in the Enterprise to access applications in the public cloud. The same VPN network is utilized to provide in-band management connectivity between the Enterprise and the SP. Thus, the control plane (management) connectivity

between the InMage processing server on the Enterprise and the CX and MT on the SP side is carried in the same path (in-band) as the data plane connectivity. This model allows the Enterprise and SP management applications (InMage) to have either public or private IP addressing.

Figure 3-34 shows the WAN topology deployed; resources below the PE router are not drawn. Refer to earlier sections for details. Based on VMDC recommendation, a pair of ASR 1006s was deployed as PE routers in the VMDC 2.3 topology, PE routers connected to P routers running MPLS using 10 GE. At the Enterprise site, VMDC 2.2 container was utilized to simulate enterprise tenants. A pair of ASR 9000s was deployed as the PE router.

Figure 3-34 Connectivity across the WAN



MPLS VPN route is a combination of route distinguisher (RD) and actual prefix. RD is a unique identifier used to distinguish the same prefix from different customer. We define it at PE router for particular VRF. Prefix combined with RD and actual IPv4 prefix is called vpnv4 prefix and is carried by MP-BGP. BGP-extended community support is required to carry vpnv4 prefixes and labels. To carry vpnv4 prefixes, we configured iBGP in the core network. It either requires full mesh connections between routers or route reflectors (RR) deployment. We went with the second option using RR. The following is the configuration on the RR:

```
router bgp 109
bgp router-id 223.254.254.251
bgp log-neighbor-changes
neighbor RRCC peer-group
neighbor RRCC remote-as 109
neighbor RRCC description ibgp-to-RR
neighbor RR peer-group
neighbor RR remote-as 109
neighbor RR description CVFCC-WAN-6k1-to-CVFCC-WAN-6k2
neighbor RR update-source Loopback0
neighbor 6.100.100.248 remote-as 109
neighbor 6.100.100.248 peer-group RRCC
neighbor 6.100.100.249 remote-as 109
neighbor 6.100.100.249 peer-group RRCC
neighbor 6.100.100.250 remote-as 109
neighbor 6.100.100.250 peer-group RRCC
```



```

neighbor 8.255.255.1 remote-as 109
neighbor 8.255.255.1 peer-group RRCC
neighbor 8.255.255.21 remote-as 109
neighbor 8.255.255.21 peer-group RRCC
neighbor 8.255.255.22 remote-as 109
neighbor 8.255.255.22 peer-group RRCC
neighbor 223.254.254.252 remote-as 109
neighbor 223.254.254.252 peer-group RR
!
address-family ipv4
neighbor RRCC send-community
neighbor RRCC route-reflector-client
neighbor RR send-community both
neighbor 6.100.100.248 activate
neighbor 6.100.100.249 activate
neighbor 6.100.100.250 activate
neighbor 8.255.255.1 activate
neighbor 8.255.255.21 activate
neighbor 8.255.255.22 activate
neighbor 223.254.254.252 activate no auto-summary
no synchronization
network 223.254.0.0 mask 255.255.0.0 exit-address-family
!
address-family vpnv4
neighbor RRCC send-community both
neighbor RRCC route-reflector-client
neighbor RRCC next-hop-self
neighbor RR send-community both
neighbor 6.100.100.248 activate
neighbor 6.100.100.249 activate
neighbor 6.100.100.250 activate
neighbor 8.255.255.1 activate
neighbor 8.255.255.21 activate
neighbor 8.255.255.22 activate
neighbor 223.254.254.252 activate exit-address-family

```

PE routers are configured to import prefix from remote PE as well as export local prefix. The following is an example of a Gold tenant:

```

vrf definition tenant11-gold-pub rd 3486:3486
route-target export 3486:3486
route-target import 3486:3486
route-target import 3040:3040
!
address-family ipv4
exit-address-family
!

```

A WAN can be distributed over a large geographical region. The data packets need more time to travel through the network. WAN latency, in general, is orders of magnitudes higher than latency in local area networks. TCP throughput and application performance are directly impacted by latency. Many applications are slow or do not work at all for users far away from the data centers. In our implementation, an Anue delay generator was inserted between the P router and the BGP RR to inject round trip delay of 30ms between the primary Enterprise data center and the SP secondary site. [Figure 3-35](#) is a screen capture of Anue configuration.

Figure 3-35 Anue Configuration

The screenshot shows the VMDc configuration interface for Blade 3 (XGEM). The navigation path is 'Configure Classifier → Bandwidth → Delay/Impairments'. The main area displays 'Blade 3 Overview' with a table of profiles. The table has columns for #, Name, Enabled, Delay, Policing, Shaping, Drop, Modify, Corrupt, CRC Corrupt, Reorder, Duplicate, Jitter, Other, and Bandwidth (Mbps). The 'Default' profile is enabled with a delay of 15.000ms. The bandwidth for the 'Default' profile is 462.256 Mbps Tx and 461.183 Mbps Rx.

#	Name	Enabled	Delay	Policing	Shaping	Drop	Modify	Corrupt	CRC Corrupt	Reorder	Duplicate	Jitter	Other	Bandwidth (Mbps)	
														Tx	Rx
0	Default	✓	15.000ms	—	—	—	—	—	—	—	—	—	✓	462.256 79170 Pkts	461.183 78992 Pkts
1	<a href="#">Profile #1</a>														

The following formula can be used to calculate effective TCP throughput based on round trip delay:

$$\text{Throughput} = \text{Window Size} / \text{RTT}$$

All InMage components are running with default window size based on host OS.

Using the standard 64KB TCP window size of a Windows machine:

$$65536 * 8 \text{ bits} / 0.030 \text{ seconds} = 17.4 \text{ Mbps maximum possible throughput}$$

Linux window size is based on:

```
net.core.rmem_max = 131071
net.core.rmem_default = 124928
```

$$131071 * 8 \text{ bits} / 0.030 \text{ seconds} = 34.9 \text{ Mbps maximum possible throughput}$$

Based on the above, the TCP stream between the Enterprise process server and SP MT is limited to 17.4 Mbps.

## Storage Configuration

In the current implementation, the VNX 5500 is used to provide the storage needs of the solution. The VNX 5500 is based on a unified storage architecture and provides Storage Area Network (SAN) and Network-Attached Storage (NAS) capabilities on a single platform. In this solution, only SAN is utilized. The Nexus 5000 is the FC switch that connects server blades and storage to provide SAN capability.

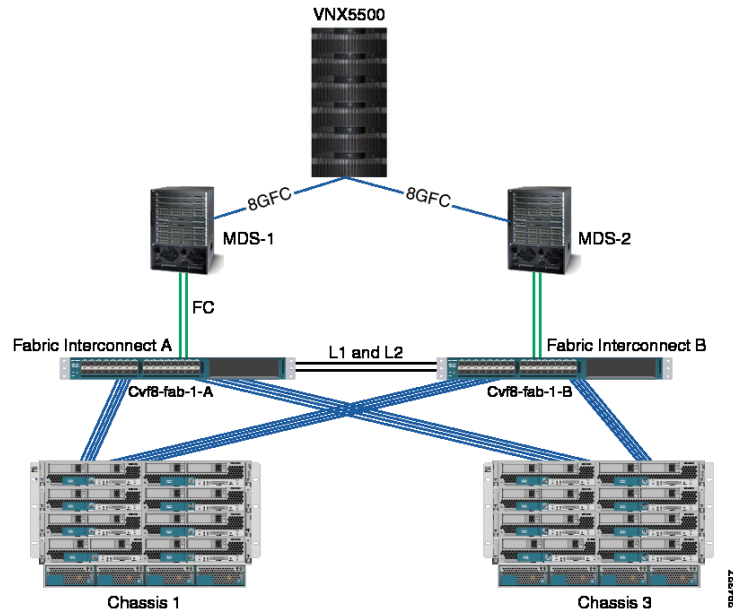
This section presents the following topics:

- [SAN Implementation Overview, page 3-42](#)
- [VNX5500 Configuration Overview, page 3-45](#)

## SAN Implementation Overview

This section explains the Fibre Channel over Ethernet (FCoE) connection from servers to the FI and Fibre Channel (FC) connectivity to carry SAN traffic from the FI to the MDS (storage switch) to VNX 5500 Filers. [Figure 3-36](#) shows an overview of the SAN infrastructure.

**Figure 3-36 Storage Infrastructure Overview**



Features of FC configuration in the data center follow:

- Each blade server has two vHBAs that provide server to storage SAN connectivity. This is to provide server level host bus adapter (HBA) fabric redundancy.
- Storage traffic from server blades to FIs is FCoE. Each virtual SAN (VSAN) is mapped to a unique VLAN that carries storage traffic from server to FI.
- Each FI is mapped to one VSAN. In this case, FI-A (CVF8-FAB-1-A) carries all VSAN88 traffic and FI-B (CVF8-FAB-1-B) carries all VSAN89 traffic.
- FCoE, by default, maps FC traffic to a no-packet drop class using the system QoS policy. This assures that during congestion storage traffic will not be dropped.

Figure 3-37 shows the list of VSANs in the SAN infrastructure: VSAN88, VSAN89. This is to allow multiple SANs and LANs to share a common infrastructure when carried using the same FCoE links between the server and FIs.

**Figure 3-37 Infrastructure VSANs**

Name	ID	Fabric ID	If Type	If Role	Transport	FCoE VLAN ID	Operational State
VSAN default (1)	1	Dual	Virtual	Network	Fc	4048	Ok
VSAN sys8_FCoE_Fab_A (88)	88	A	Virtual	Network	Fc	88	Ok
VSAN sys8_FCoE_Fab_B (89)	89	B	Virtual	Network	Fc	89	Ok

Figure 3-38 shows the vHBA configuration on each server blade. vHBAs are part of a server service profile derived from a server template, consists of two vHBA adapters per server blade. Each vHBA is placed on a unique, isolated SAN network. vHBA0 of all server blades are placed in SAN-A and vHBA1 is placed in SAN-B

**Figure 3-38 Infrastructure vHBAs**

Name	WWPN	Desired Order	Actual Order	Fabric ID	Desired Placement	Actual Placement
vHBA vHBA-A	Derived	1	Unspecified	A	Any	Any
vHBA If sys8_FCoE_Fab_A						
vHBA vHBA-B	Derived	2	Unspecified	B	Any	Any
vHBA If sys8_FCoE_Fab_B						

Figure 3-39 shows the ports used between FI and the MDS switch for SAN traffic. Although 4 ports are configured, only two of the ports are physically cabled. FI-A (fc 2/1, 2/2) connects to MDS-1 (fc 1/19, 1/20) and FI-B (fc 2/1, 2/2) connects to MDS-2 (fc 1/19, 1/20).

**Figure 3-39 Ports Used between FI and MDS Switch**

Name	Fabric ID	If Type	If Role	Transport	Administrative State
FC Interface 2/1	A	Physical	Network	Fc	Enabled
FC Interface 2/2	A	Physical	Network	Fc	Enabled
FC Interface 2/3	A	Physical	Network	Fc	Enabled
FC Interface 2/4	A	Physical	Network	Fc	Enabled

Soft zoning (using World Wide Port Name (WWPN) names) is configured on the MDS to allow servers with specific identity (WWPN) to communicate with VNX filers. Each filer connection has its own WWPN name. The following configuration shows the zoning configuration SAN-A. As mentioned before, vHBA0 of all server blades are placed in the SAN-A and vHBA1 is placed in SAN-B. The WWPN of vHBAs is obtained from the UCSM. The WWPN of VNX filers is fetched using VNX FC port properties.

```

zoneset name cvf8-Fab-a vsan 88
  zone name cvf8-temp-all-vnx5500 vsan 88
    pwnn 50:00:00:25:b5:e1:81:1f
    pwnn 50:00:00:25:b5:e1:81:3e
    pwnn 50:00:00:25:b5:e1:81:3f
    pwnn 50:00:00:25:b5:e1:81:5e
    pwnn 50:00:00:25:b5:e1:81:5f
    pwnn 50:00:00:25:b5:e1:81:7e
    pwnn 50:00:00:25:b5:e1:81:7f
    pwnn 50:00:00:25:b5:e1:81:9e
    pwnn 50:00:00:25:b5:e1:81:9f
    pwnn 50:00:00:25:b5:e1:81:ae
    pwnn 50:00:00:25:b5:e1:81:af
    pwnn 50:00:00:25:b5:e1:81:be
    pwnn 50:00:00:25:b5:e1:81:bf
    pwnn 50:00:00:25:b5:e1:81:de
    pwnn 50:00:00:25:b5:e1:81:df
    pwnn 50:06:01:64:3e:a0:36:1a < VNX

```

```

cvf6-san-mds1# show flogidatabase vsan 88

INTERFACE          VSAN    FCID          PORT NAME          NODE NAME
-----
fc1/19             88      0xb00100     20:41:54:7f:ee:12:5d:40  20:58:54:7f:ee:12:5d:41
fc1/19             88      0xb00101     50:00:00:25:b5:e1:81:7e  20:00:00:25:b5:08:01:2f
fc1/19             88      0xb00102     50:00:00:25:b5:e1:81:ae  20:00:00:25:b5:08:01:4f
fc1/19             88      0xb00104     50:00:00:25:b5:e1:81:1f  20:00:00:25:b5:08:01:9f
fc1/19             88      0xb00108     50:00:00:25:b5:e1:81:5f  20:00:00:25:b5:08:01:af

fc1/19             88      0xb0011e     50:00:00:25:b5:e1:81:9f  20:00:00:25:b5:08:01:bf
fc1/19             88      0xb0011f     50:00:00:25:b5:e1:81:af  20:00:00:25:b5:08:01:ef
fc1/19             88      0xb00123     50:00:00:25:b5:e1:81:5e  20:00:00:25:b5:08:01:3f
fc1/19             88      0xb00126     50:00:00:25:b5:e1:81:df  20:00:00:25:b5:08:01:ff
fc1/20             88      0xb00000     20:42:54:7f:ee:12:5d:40  20:58:54:7f:ee:12:5d:41
fc1/20             88      0xb00001     50:00:00:25:b5:e1:81:7f  20:00:00:25:b5:08:01:cf
fc1/20             88      0xb00002     50:00:00:25:b5:e1:81:3e  20:00:00:25:b5:08:01:0f
fc1/20             88      0xb00004     50:00:00:25:b5:e1:81:de  20:00:00:25:b5:08:01:6f
fc1/20             88      0xb00008     50:00:00:25:b5:e1:81:3f  20:00:00:25:b5:08:01:8f
fc1/20             88      0xb0000f     50:00:00:25:b5:e1:81:bf  20:00:00:25:b5:08:01:df
fc1/20             88      0xb00010     50:00:00:25:b5:e1:81:9e  20:00:00:25:b5:08:01:5f
fc1/20             88      0xb00017     50:00:00:25:b5:e1:81:be  20:00:00:25:b5:08:01:7f
fc1/34             88      0xb00300     50:06:01:64:3e:a0:36:1a  50:06:01:60:be:a0:36:1a

```

We had the choice of implementing a "single initiator zoning" where each zone contains only one host server vHBA and can contain multiple storage array targets in the same zone. Instead, we implemented "multi-initiator zoning" to allow the flexibility of moving hosts between ESXi clusters. Instead of masking at the MDS level, we used VNX storage group to mask specific LUNs to the ESXi Cluster. Refer to [VNX5500 Configuration Overview, page 3-45](#) for details. The FC interface on the MDS switch is used to connect to the VNX 5500 for FC connectivity and is configured below.

```

interface fc1/34
switchport description Connection to VNX5500 port-license acquire
no shutdown

```

## VNX5500 Configuration Overview

VNX has four main configuration elements:

- The physical drives
- The storage pool
- The LUN
- The tiering policy of the LUN

[Figure 3-40](#) shows a high level overview of VNX.

**Figure 3-40 High Level Overview of VNX**

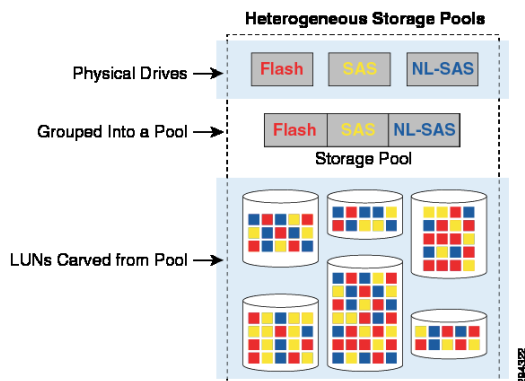


Figure 3-40 was taken from from EMC VNX Virtual Provisioning at <http://www.emc.com/collateral/hardware/white-papers/h8222-vnx-virtual-provisioning-wp.pdf>.

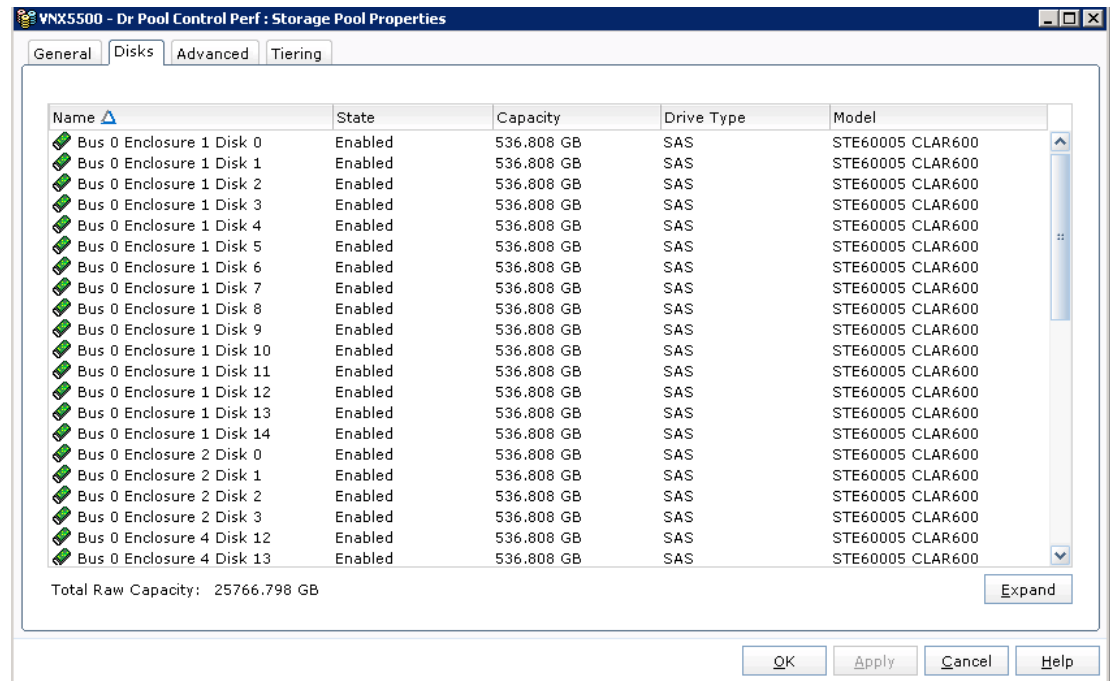
As discussed earlier, FAST VP is a tiering solution that can be utilized with the VNX to reduce total cost of storage ownership. FAST VP operates by continuously collecting performance statistics. Collected data is analyzed once per hour and, based on schedule, data is moved between tiers once every 24 hours during a specified relocation window. The granularity of data is 1GB. Each 1 GB block of data is referred to as a "slice." When FAST VP relocates data, it will move the entire slice to a different storage tier.

To start off, our implementation of FAST VP consists of only SAS disks: this is to provide sufficient performance to on board newly protected VMs and also to allow VNX sufficient time to identify hotter/colder slices of data. Once the desired number of tenants per storage pool is reached, NL-SAS drives can be introduced to move cold data from performance tier to capacity tier. To balance the data split between performance and capacity tier, a manual relocation at the storage pool level can

be initiated by a cloud admin through the Unisphere GUI. Both relocation rate and duration can be specified at the time of manual relocation.

In our implementation, storage pool was sized based on total IOPS to support peak VM transfer from the primary site and change rate of existing VMs under protection. 4800 IOPS was provisioned to support Journal retention and an additional 4800 IOPS to support aggregate workload change rate. In an actual deployment, IOPS requirement will vary considerably. Depending on WAN bandwidth, number of customers on boarding new VMs and change rate of existing VMs under protection, IOPS needs to be sized according to the deployment scenario.

Based on 200 IOPS per SAS disk and RAID 10 configuration, 48 SAS drives were needed to support 4800 IOPS. Figure 3-41 shows a screen capture of disk configuration.

**Figure 3-41** Disk Configuration

Tenant-specific LUN is created on top of the storage pool. Each tenant is assigned a dedicated Journal LUN and workload LUN is shared between tenants. EMC FAST Cache was also utilized to provide read acceleration during the time of recovery. Total of 274GB of usable flash cache was deployed. In a real deployment scenario, the amount of flash cache should be sized based on the overall capacity of recovery workloads.

**Table 3-12** Journal LUN

Journal	LUN Size (GB)
Tenant Control_1	1480
Tenant Control_2	590
Tenant Control_3	1140
Tenant Control_4	1140
Tenant Control_5	1140
Tenant Control_6	1140
Tenant Control_7	1140
Tenant Control_8	1140
Tenant Control_9	340
Tenant Control_10	340
Tenant Control_11	340
Tenant Control_12	340

**Table 3-13 Workload LUN**

Workload	LUN Size (GB)
LUN Gold-1	500
LUN Silver-1	750
LUN Silver-2	750
LUN Bronze-1	900
LUN Bronze-2	900
LUN Bronze-3	900

All of the LUNs are mapped to the corresponding storage group, based on ESXi cluster, as shown in [Table 3-14](#). As discussed in earlier sections, LUN masking is implemented at the storage array level to simplify the ability to move hosts between clusters for various test scenarios.

**Table 3-14 Storage Group DR-Bronze**

Storage Group	ESX Cluster Name	Hosts	LUNs
DR-Bronze	Bronze	cvf8-draassp-esx-4.cvfdmz.sdu	Tenant Control_1
		cvf8-draassp-esx-5.cvfdmz.sdu	Tenant Control_4
		cvf8-draassp-esx-6.cvfdmz.sdu	Tenant Control_5
		cvf8-draassp-esx-7.cvfdmz.sdu	Tenant Control_6
		cvf8-draassp-esx-8.cvfdmz.sdu	Tenant Control_7
			Tenant Control_8
			LUN Bronze-1
			LUN Bronze-2
		LUN Bronze-3	

**Table 3-15 Storage Group DR-Silver**

Storage Group	ESX Cluster Name	Hosts	LUNs
DR-Silver	Silver	cvf8-draassp-esx-2.cvfdmz.sdu	Tenant Control_2
		cvf8-draassp-esx-3.cvfdmz.sdu	Tenant Control_3
		cvf8-draassp-esx-3-1.cvfdmz.sdu	LUN Silver-1
			LUN Silver-2

**Table 3-16 Storage Group DR-Gold**

Storage Group	ESX Cluster Name	Hosts	LUNs
DR-Gold	Gold	cvf8-draassp-esx-9.cvfdmz.sdu	Tenant Control_9
		cvf8-draassp-esx-10.cvfdmz.sdu	Tenant Control_10
		cvf8-draassp-esx-11.cvfdmz.sdu	Tenant Control_11



**Table 3-16 Storage Group DR-Gold (continued)**

Storage Group	ESX Cluster Name	Hosts	LUNs
		cvf8-draassp-esx-12.cvfdmz.sdu	Tenant Control_12
			LUN Gold-1

## BMC Cloud Lifecycle Management

DRaaS 1.0 leverages existing capabilities of Cloud Orchestration for VMDC with BMC Cloud Lifecycle Management 3.1 SP1. Initial workflows of onboarding tenants, network container creation, firewall policy changes, and server load balancer updates align with VMDC 2.3 operational method and procedures. This is well documented by the SDU BMC-CLM team. Refer to [BMC Design and Implementation Guide](#) for additional details.

