Executive Summary

The world has been taken over by zombies and they have ravaged nearly everyone on the planet. A billionaire is providing the investment to create a highly scalable and full orchestrated environment which will be the backbone of a manufacturing facility. This facility is building ships to transport what's left of the human race to the moon before sending them to Mars for colonization.

This system is now being moved to the moon to provide the same functionality to help build ships that will transport the rest of the human race to Mars for colonization.

Business Requirements

Number	Item
R01	Design must accommodate current application in smaller footprint
R02	Failure tolerances must be identified
R03	Maximum sizes must be established
R04	Must fit into 21U

Business Constraints

Number	Item
C01	An IPv6 network is the only network that exists
C02	Can only use previous hardware vendors
C03	Can only use previous software vendors
C04	Hardware must accommodate a 21U form factor

Business Assumptions

Number	Item
A01	No DR on-site
A02	UPS Power is provided
A03	Air-conditioning is provided
A04	Limited professional support services for environment

Business Risks

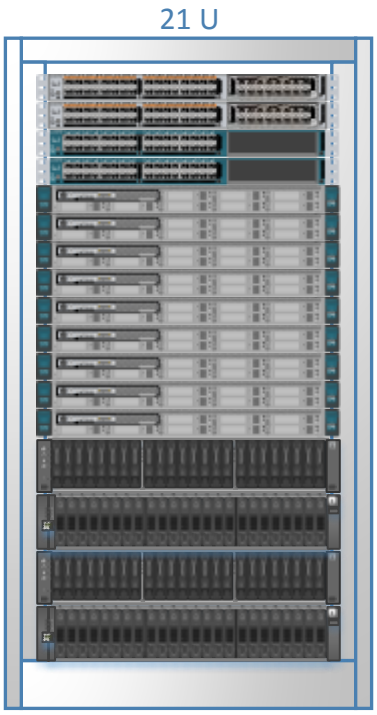
Number	Item
K01	vCAC Appliances have IPv6 support but not officially supported
K02	No off-site DR

Document Purpose

This document serves as the configuration document to layout the scaled-down version of the Earth-based manufacturing depot. While the same application that is running on Earth will run on the moon, there are serious design considerations that will impact what that architecture looks like. This document serves as the template for that configuration.

Physical Datacenter Overviews

- Datacenter rack layout



The rack will be 21U in accordance with R04. The networking, compute and storage will all be housed within the rack. The hardware components are listed as a constraint to use previous vendors in the manufacturing design, C02, however due to scale they were changed to provide the most scale in the smallest footprint while still supporting the application.

    - Networking
        - Cisco 5548UP – 2
        - Cisco 6248 - 2
    - Compute
        - Cisco C220 M3 – 9
    - Storage
        - NetApp E2252 – 2
            - 2246 – 2 – one connected to each E2552
    - Consumed rack space – 21U
- Power configurations

The installed equipment will draw the following power requirements. Some numbers are obtained from maximums from the vendor websites and some are actual tested numbers based on the vendor's documentation. The racks will consist of two power distribution units that will be connected to separate uninterruptible power supplies. There is an assumption, A02, that the facility will have a generator capable of supplying power in case of main grid failure. The rack will have connections to UPS-A and UPS-B. The numbers are a best representation of the data provided:

Rack	Item	UPS-A Watts	UPS-B Watts
1	Cisco 5548UP	193W	193W
1	Cisco 5548UP	193W	193W
1	Cisco 6248	193W	193W
1	Cisco 6248	193W	193W
1	Cisco C220 M3	375W	375W
1	Cisco C220 M3	375W	375W
1	Cisco C220 M3	375W	375W
1	Cisco C220 M3	375W	375W
1	Cisco C220 M3	375W	375W
1	Cisco C220 M3	375W	375W
1	Cisco C220 M3	375W	375W
1	Cisco C220 M3	375W	375W
1	NetApp E2552	150W	150W
1	NetApp 2246	150W	150W
1	NetApp E2552	150W	150W
1	NetApp 2246	150W	150W

Total Consumed Power	4372	4372

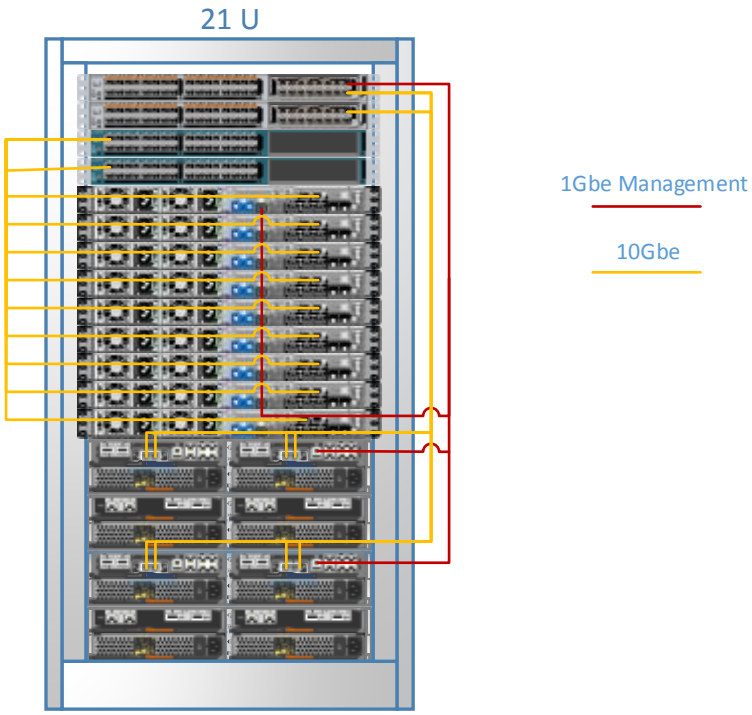
- HVAC configurations

The equipment will need to be cooled during operation. There is an assumption, A03, that the facility will have a room for the datacenter equipment to reside in. Using the Table 1 provided in this document, the amount of AC tonnage required to cool the equipment is as follows:

  - Formula – Total Watts * 0.000283 = Tons
  - Calculation – 8744 * 0.000283 = 2.01112 Tons

Physical Infrastructure Overviews

- Network Infrastructure



The following configuration was created to handle the scale-down infrastructure needed to accommodate the application. The network infrastructure consists of a core switching network composed of a pair of Cisco 5548UP switches. There are Cisco 6248UP Fabric Interconnects that are providing the UCS Manager and Service Profiles for the Cisco C220 M3s. These machines connect into the fabric extenders which then connects back to the 5548UPs. This configuration allows the core server and networking components to fit into the 21U requirement.

- Network/VLAN Design

One of the business constraints, C01, states that the Moon facility only has access to an IPv6 network. IPv6 addresses and networks can have significant numbers of hosts and networks, more so than the entire IPv4 address space. The following VLANs are required to split out traffic for broadcast domain purposes as well as ease of management and troubleshooting. One flat network has inherent limitations in terms of scalability. This configuration will allow maximum scaling and segregation of network traffic.

The following prefix will be used and 5 networks established to split traffic out based on Team Alpha's name and the networks necessary to provide connectivity to the infrastructure.

IPv6 Prefix – 2001:aaaa:aaaa:aaa0::/64 – 2001:aaaa:aaaa:aaa4::/64 = 5 networks with 18446744073709551616 IP addresses per network.

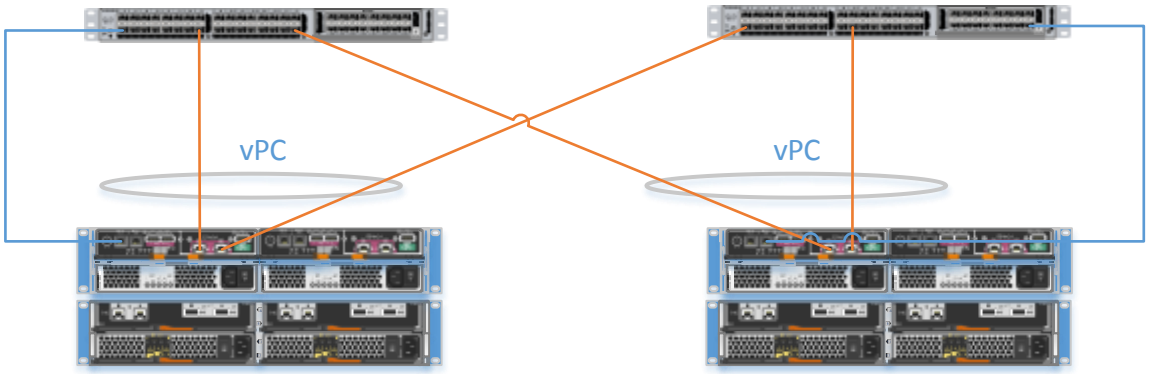
VLAN	Purpose	IPv6 Network
A	Management Network	2001:aaaa:aaaa:aaa0::/64
B	VM Network	2001:aaaa:aaaa:aaa1::/64
C	vMotion	2001:aaaa:aaaa:aaa2::/64
D	NFS	2001:aaaa:aaaa:aaa3::/64
E	Replication	2001:aaaa:aaaa:aaa4::/64

There are two concerns with using IPv6 with the management components. The vCAC appliances have support for IPv6 however do not 'officially' support using its configuration. Given the state of the world being in complete disarray, 'officially' supported is less valid. This configuration is necessary to provide the automation tier of the application. It will be listed as a risk, K01.

- Host Infrastructure
    - Cisco C220 M3 x9

CPU	RAM	NIC	Power	Storage
2x2.8Ghz E5-2680	256GB – 1866Mhz	VIC 1225	Dual 750W	Dual SD Card

- Storage Infrastructure



The storage infrastructure physical layout is shown above. Certain components have been removed for clarity of connectivity. The NetApp E2552's connect to each Cisco Nexus 5548UP via a vPC based on LACP. This provides redundancy in case of network connection loss, controller loss and network switch loss. Dual configurations are being provided as a means to replicate the NFS volumes between arrays to provide against array or shelf failure using Synchronous SnapMirror. The second array also provides the shared storage for the video system. All volumes will be replicated between arrays using Synchronous SnapMirror to ensure all data is protected.

With the constraint of IPv6 only network infrastructure, C01, the only supported IP-based protocol for vSphere that is supported is NFS. This protocol in IPv6 form, is supported across the design.

The NetApp disk layout is as follows:

Type of Disk	Use	Number	Size	Spares	RaidSize	Parity	RaidSets	Total Space
SSD	Flash Pool	4	1.6TB	0	0	0	0	4.8
SAS	Aggregate	39	900GB	3	18	4	2	20.3TB

The Flash Pool is used to provide hybrid characteristics to the arrays and is a 1:1 ratio of Flash Pool to Aggregate. Only one controller will be serving active data from each NetApp. This is done because of the way that NetApp requires splitting the disks across controllers. Each controller will be assigned 3 SAS drives to provide for each controllers root volume, and then all other disks are assigned to controller 1. Controller 2 is a hot-standby in case of failure and will pick up the NFS connections and resume serving storage. This is done to maximize the amount of available disk space within the rack.

Virtualization Infrastructure Overviews

- vSphere Definitions

One of the requirements was to scale down but still use the same software components, R01, with a constraint that the software vendors had to be used, C03.

vSphere Component	Description
VMware vSphere	The core products of the VMware vSphere environment include: • ESXi – 3 instances will compose the management cluster and 6 hosts will comprise the compute cluster for VM consumption • vCenter Server – 1 installed VM instance • vCenter Server Database – 1 VM instance for the single instance of vCenter Server • SSO – Single Sign-on component that is required for connecting to the vSphere Client and vSphere Web Client • vSphere Client – Still needed to manage VMware Update Manager • vSphere Web Client – Used to manager the vSphere environment • VDPA – used to back up the virtual machine instances
vCloud Automation Center	• Identity Appliance • vCAC Appliance • IaaS Components – installed on a Windows machine • vShield Manager • vShield Edge • vShield Endpoint

- vSphere Component Definitions

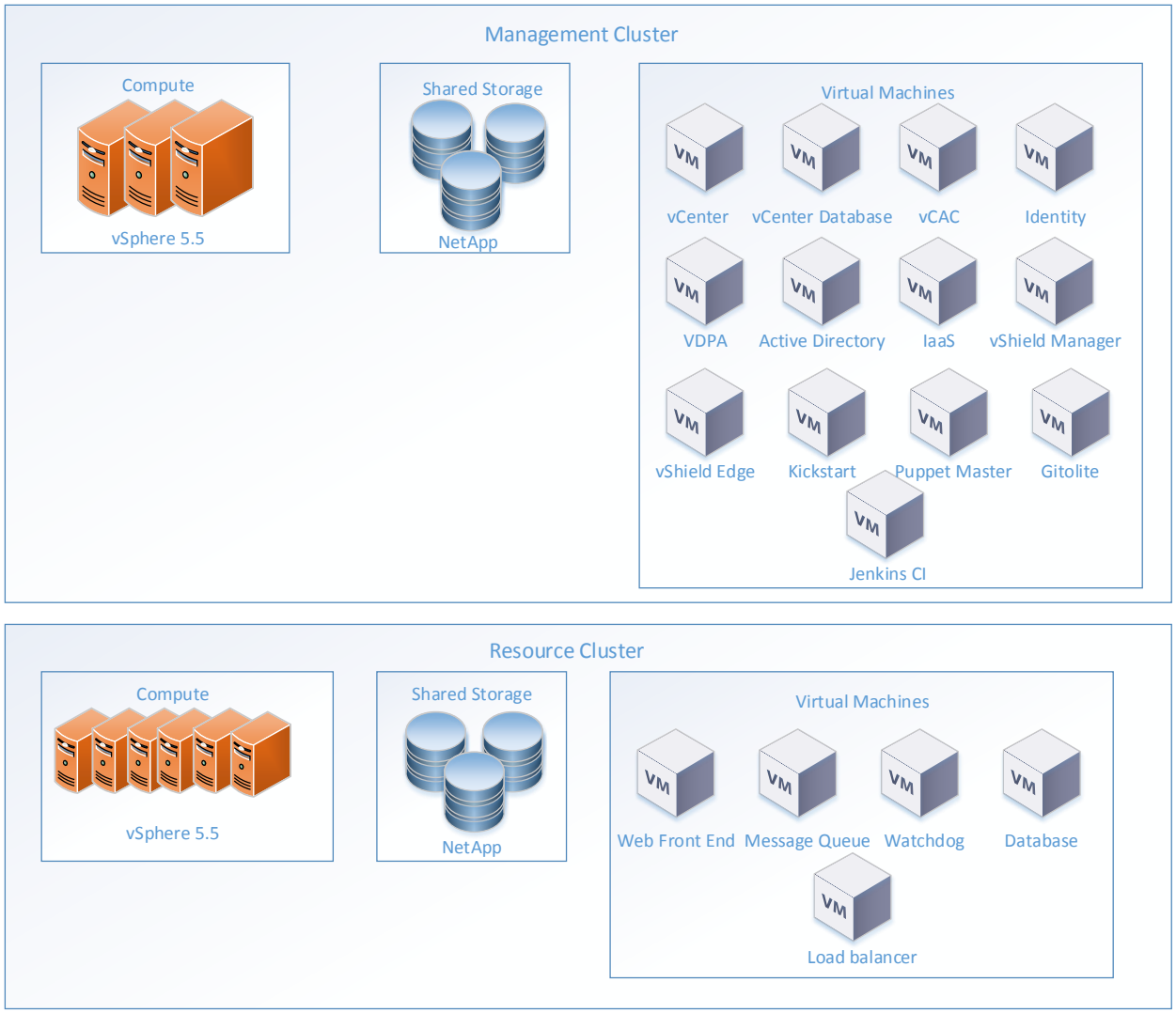
Design Section	vSphere Components
vSphere Architecture – Management Cluster	• vCenter Server and vCenter Database • vCenter Cluster and ESXi hosts • Single Sign-On • vSphere Update Manager
vSphere Architecture – Compute Cluster	• vCenter Cluster and ESXi hosts

- vSphere Architecture Design Overview
- High level Architecture

The vSphere components are being split out to facilitate ease of troubleshooting of the management components without disruption of the resource components. They are split out as follows:

Management cluster that will host the management components of the vSphere infrastructure. They are split out to ensure they have dedicated resources in which to consume.

Compute cluster that will host the application layer of the deployment. Splitting the compute cluster from the management cluster helps facilitate adding more resources later if necessary to scale out the deployment.



- Site Considerations

The vSphere management and compute clusters are both residing within the same facility. This will provide the lowest latency for management as well as a consistent datacenter in which to manage the clusters.

There are no other sites that are in scope for this project.

- Design Specifications
- vSphere Architecture Design – Management Cluster
  - Computer Logical Design
    - Datacenter

One datacenter will be built to house the two clusters for the environment.

- vSphere Cluster

Below is the cluster configuration for the management cluster for the environment.

Attributes	Specification
Number of ESXi Hosts	3
DRS Configuration	Fully Automated
DRS Migration Threshold	Level 3
HA Enable Host Monitoring	Enabled
HA Admission Control Policy	Disabled
VM restart priority	Medium – vCenter and vCenter DB set High priority
Host Isolation response	Leave powered on
VM Monitoring	Disabled

To fulfill the requirement of R02 in that fault tolerances need to be established, the HA admission control policy is set to Disabled to ensure that the core management functions are online at all times regardless of host resources. The currently tolerance in the management cluster would be N+1 for maintenance purposes, however with the amount of resources that each host has the cluster could run on one node in extreme cases.

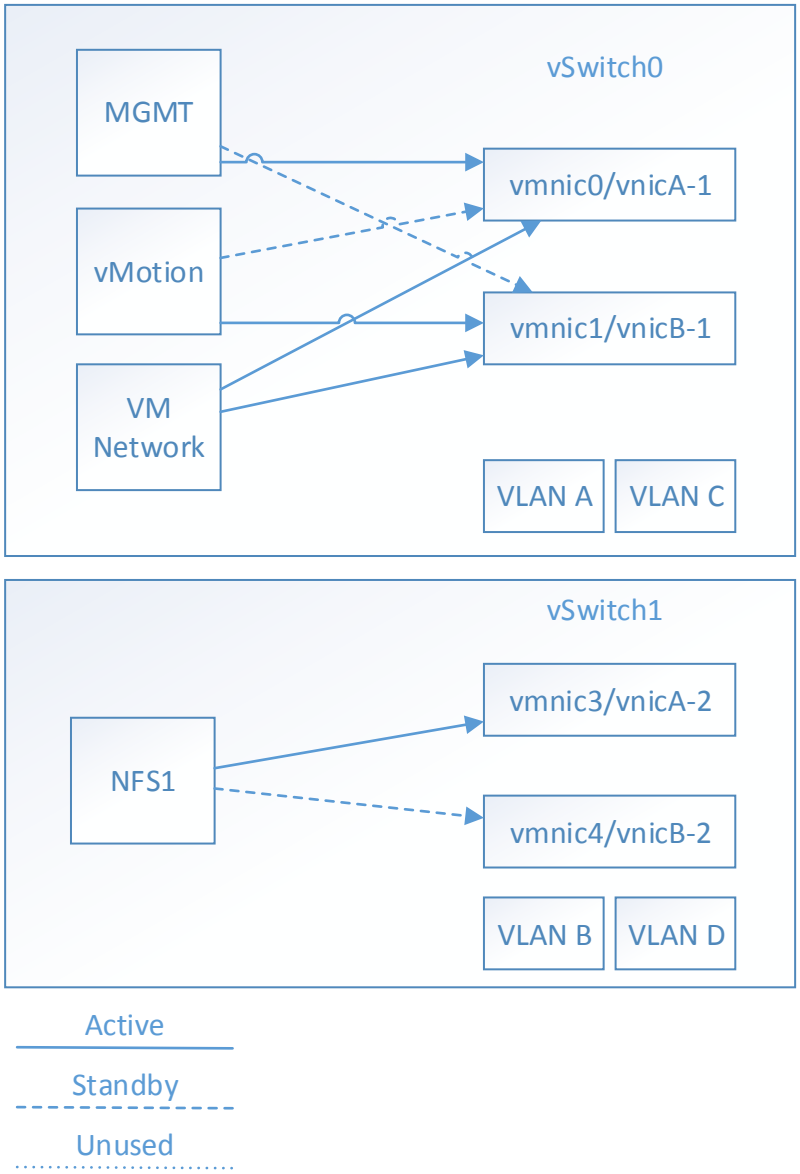
- Host logical Design

Attribute	Specification
Host Type and Version	VMware ESXi Installable
Processors	X86 Compatible
Storage	FlexFlash SD for local ESXi install, shared storage for VMs
Networking	Connectivity to all needed VLANS
Memory	Sized for workloads

- Network Logical Design

Switch Name	Switch Type	Function	# of Physical Ports
vSwitch0	Standard	Management vMotion VM Network	2x10Gbe
vSwitch1	Standard	NFS	2x10Gbe

The VIC 1225 in the Cisco C220 M3 allows the ability to split out 2 10Gbe Network connections into 256 vnics. The configuration is to create 4 vnics through the Service Profile, 2 with bindings to Fabric Interconnect A and Fabric Interconnect B. A pair of vnics, one going to either Fabric Interconnect, will compose the port group uplinks necessary for failover purposes and redundancy within the vSphere environment. This is illustrated below:



Port group configurations

Attribute	Setting
Load balancing	Route based on originating virtual port ID
Failover Detection	Link Status Only
Notify Switches	Yes
Failover Order	MGMT – Active vmnic0/Standby vmnic1 vMotion – Standby vmnic0/Active vmnic1 VM Network – Active vmnic3/Active vmnic4 NFS1 – Active vmnic3/Standby vmnic4

- Shared Storage Logical Design

Attribute	Specification
Number of NFS datastores to start	2
Volume1 Size	3TB
Volume2 Size	4TB

The VDAP appliance will reside on its own data store. All others VMs will reside on the same datastore so NetApp deduplication can reduce the overall footprint.

- Management Components

This is the list of Management components that will be running on the management cluster:

- vCenter Server
- vCenter Database
- vCenter Update Manager
- vCAC Appliance
- Identity Appliance
- IaaS
- Kickstart
- Puppet Master
- Gitolite
- Jenkins CI
- vShield Manager
- vShield Edge
- VDAP

- Management Components Resiliency Considerations

Component	HA Enabled?
vCenter Server	Yes
vCenter Database	Yes
vCAC Appliance	Yes
Identity Appliance	Yes
Active Directory	Yes
IaaS	Yes
Kickstart	Yes
Puppet Master	Yes
Gitolite	Yes
Jenkins CI	Yes
vShield Manager	Yes
vShield Edge	Yes
VDAP	Yes

- Management Server Configurations

VM	vCPUs	RAM	Disk1	Disk2	Disk3	Controller	Quantity
vCenter Server	2	16GB	40GB	200GB	N/A	LSI Logic SAS	1
vCenter Database	2	16GB	40GB	100GB	N/A	LSI Logic SAS	1
vCAC Appliance	2	8GB	30GB	N/A	N/A	LSI Logic SAS	1
Identity Appliance	1	2GB	10GB	N/A	N/A	LSI Logic SAS	1
IaaS	2	8GB	30GB	N/A	N/A	LSI Logic SAS	1
Active Directory	1	4GB	40GB	N/A	N/A	LSI Logic SAS	1
Kickstart	1	512MB	100GB	N/A	N/A	LSI Logic SAS	1
Puppet Master	4	8GB	100GB	N/A	N/A	LSI Logic SAS	1
Gitolite	2	8GB	500GB	N/A	N/A	LSI Logic SAS	1
Jenkins CI	2	8GB	500GB	N/A	N/A	LSI Logic SAS	1
vShield Manager	2	8GB	60GB	N/A	N/A	LSI Logic SAS	1
vShield Edge	2	1GB	5GB	N/A	N/A	LSI Logic SAS	1
VDAP	4	4GB	3100GB	N/A	N/A	LSI Logic SAS	1

- vSphere Architecture Design – Compute Cluster
  - Computer Logical Design
    - Datacenter

One datacenter will be built to house the two clusters for the environment.

- vSphere Cluster

Attributes	Specification
Number of ESXi Hosts	6
DRS Configuration	Fully Automated
DRS Migration Threshold	Level 3
HA Enable Host Monitoring	Enabled
HA Admission Control Policy	Disabled
VM restart priority	Medium -
Host Isolation response	Leave powered on
VM Monitoring	Disabled

To fulfill the requirement of R02 in that fault tolerances need to be established, the HA admission control policy is set to Disabled because the system should stay running regardless of degradation. A running system that runs slower is more important than one that runs at maximum performance. The currently tolerance in the compute cluster would be N+3. The limiting factor is disk space in this configuration. However each host is fully capable of driving 3 groups each. This means that the cluster could sustain a 3 host loss without performance degradation.

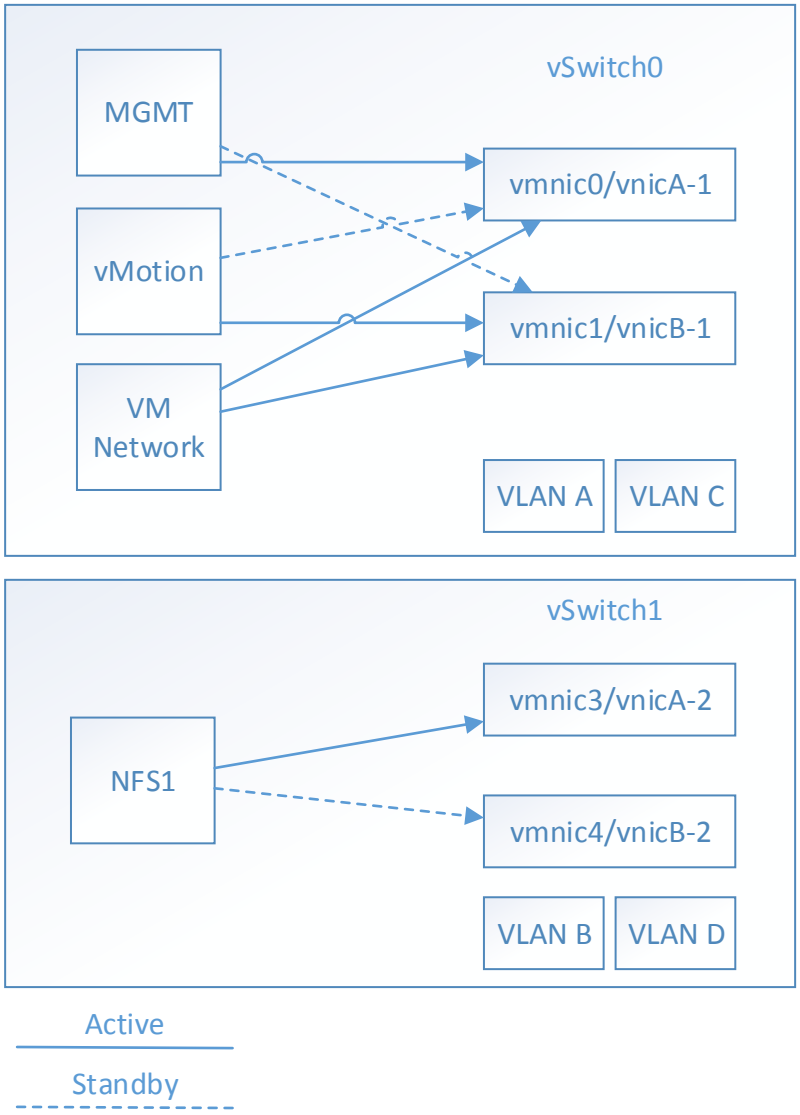
- Host Logical Design

Attribute	Specification
Host Type and Version	VMware ESXi Installable
Processors	X86 Compatible
Storage	FlexFlash SD for local ESXi install, shared storage for VMs
Networking	Connectivity to all needed VLANS
Memory	Sized for workloads

- Network Logical Design

Switch Name	Switch Type	Function	# of Physical Ports
vSwitch0	Standard	Management vMotion VM Network	2x10Gbe
vSwitch1	Standard	NFS	2x10Gbe

The network configuration for the Compute Cluster follows the exact same pattern as that of the Management cluster for simplicity.



Port group configurations

Attribute	Setting
Load balancing	Route based on originating virtual port ID
Failover Detection	Link Status Only
Notify Switches	Yes
Failover Order	MGMT – Active vmnic0/Standby vmnic1 vMotion – Standby vmnic0/Active vmnic1 VM Network – Active vmnic3/Active vmnic4 NFS1 – Active vmnic3/Standby vmnic4

- Shared Storage Logical Design

Attribute	Specification
Number of volumes	1
Volume1 Size	10.4TB

- Compute Components

This is a list of the components that will be running on the compute cluster for the environment:

- Web Front End
- Message Queue
- Database
- Watchdog
- Load Balancer

- Compute Component Resiliency Considerations

Component	HA Enabled?
Web Front End	Yes
Message Queue	Yes
Database	Yes
Watchdog	Yes
Load Balancer	Yes

- Compute Server Configurations

VM	vCPUs	RAM	Disk1	Disk2	Controller	Quantity
Web Front End	1	1GB	50GB	N/A	LSI Logic SAS	6
Message Queue	1	2GB	50GB	N/A	LSI Logic SAS	6
Database	2	4GB	50GB	200GB	VMware Para-virtual Controller	6
Watchdog	1	512MB	50GB	N/A	LSI Logic SAS	6
Load Balancer	1	4GB	50GB	N/A	LSI Logic SAS	9

The current application system consumes the following resources in each grouping:

vCPUs	Ghz	RAM	Disk	VMs Total
6	16.8	81GB	1650GB	33

If the application grows in a similar fashion as shown above, the deduplication savings would be significant. Transparent Page Sharing of the like VMs would result in significant memory savings as well on each host. Each host consists of the following:

CPU	Ghz	RAM	NIC	Power	Storage
2x2.8Ghz E5-2680	44.8	256GB – 1866Mhz	VIC 1225	Dual 750W	Dual SD Card

Giving us a total cluster resource numbers of reserving 80% for overhead:

Ghz	pRAM
215.04	1209GB

The NetApp array consists of the following amount of disk space available for host consumption taking into consideration resources reserved for replicated volumes. No deduplication savings is taken into consideration on space consumed or available, but 80% reserved for overhead.

Total Disk Empty	Space reserved for replicated volumes	Total space left to consume
20.3TB	7TB	10.4TB

Taking into account the following ratios and figures we get an approximate maximum of each host and approximation of maximum storage consumed:

Figure	Description	Results
vCPU to PCPU ratio	4:1	64 vCPUs
vRAM to pRAM ratio	1.2:1	302GB
Possible TPS Savings	40%	56.4GB more RAM
Possible Deduplication Savings	60%	990GB more space available

The theoretical maximums for requirement, R03, that number of groups of that the entire compute cluster would be able to hold would be taking into account the above ratios leaving:

		Max App groups
Max vCPUs	384	64
Max Ghz	860.16	51
Max vRAM	2540GB	31
Max Disk	16.64TB	10
Max number of app groups by least common denominator		**10**

Max number of VMs by least common denominator		**330**
Max ratio of VMs per host by least common denominator		**55:1**

- vSphere Security
    - Host Security

Hosts will be placed into lockdown mode to prevent root access. This would ensure that only access can be done through the DCUI.

    - Network Security

All virtual switches will have the following settings:

Attribute	Setting
Promiscuous Mode	Management Cluster – Reject Compute Cluster - Reject
MAC Address Changes	Management Cluster – Reject Compute Cluster – Reject
Forged Transmits	Management Cluster – Reject Compute Cluster - Reject

    - vCenter Security

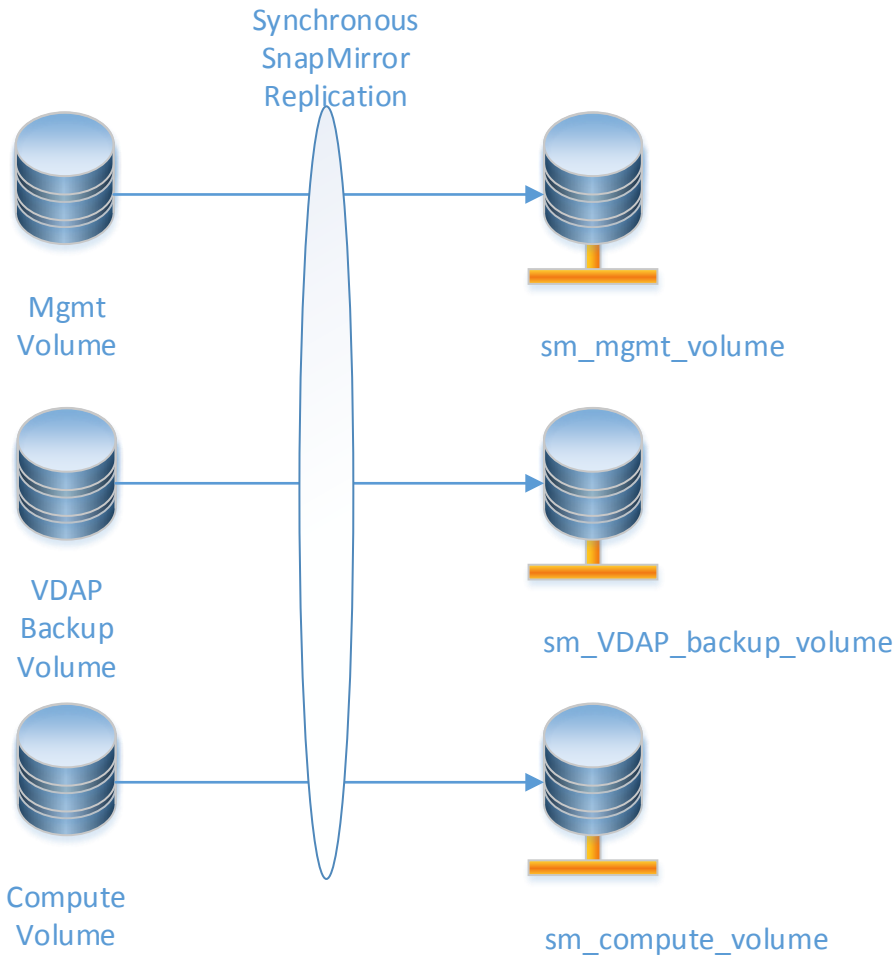
By default when vCenter is added to an Active Directory domain, the Domain Administrators group is granted local administrator permissions to the vCenter Server. A new vCenter Admins group will be created, appropriate users will be added to the group and that group will become the new local administrators on the vCenter Server. The Domain Administrators group will be removed.

Backup, Recovery and Replication Overviews

- VDAP
    - Master Server Configuration
        - Server storage will come from primary NetApp array
        - Attached 3.1TB disk to hold backups
        - Weekly Fulls, Nightly incrementals
- Synchronous SnapMirror Replication

NetApp has support for SnapMirror on IPv6 networks. This includes the Synchronous SnapMirror method. Synchronous SnapMirror requires a very low latency network connection between NetApp arrays as the writes have to be confirmed on the remote array before another can be done on the primary controller. Given that all transactions are going to occur in the same networking fabric, this will provide the lowest latency necessary to facilitate this configuration.

Synchronous SnapMirror will be configuring for the management cluster volume, the VDAP volume and the compute cluster volume. This configuration will help to alleviate the risk of no off-site DR, K02.



Appendix A – Bill of Materials

Equipment	Quantity
Cisco 5548UP	2
Cisco 6248UP	2
Cisco C220 M3	9
NetApp E2552	2
NetApp DS2246	2

Racks	1
PDUs	2