Virtual Design Master Season 2

Challenge 2 Submission: Anacreon Lunar Base Deployment

Challenge 2: Adaption and Constraints

Byron Schaller
7-22-2014

Contents

1 Overview

1.1 Executive Summary

We have made it off of Earth, well most of us. Now that we have arrived at the Anacreon Lunar Base the Ship Manufacturing Application Infrastructure must be adapted to the constraints of the Moon.

Lead architect for the terrestrial system, Rob Nelson sacrificed himself in a noble act as we were boarding the final ships. Due to his absence it falls upon the rest of the architects on Team Alpha to carry his vision forward.

Upon arrival at the Anacreon base we were met with unexpected challenges. First, the network will only support IP version 6. The antiquated IP version 4 systems must be adapted to the new standard moving forward.

The second challenge should have been expected. Space, and more importantly power, is scarce. As such the original infrastructure must be scaled to fit in just one half of a standard data center rack, only 21 U of space.

Luckily some hardware from the OEMs originally selected for the terrestrial project survived and was brought with us. We must leverage the Cisco and NetApp gear as we plan the implementation.

The future of mankind depends on the completion of our journey to Mars. That journey depends on the virtual infrastructure described in this document.

Good luck and Godspeed.

1.2 Requirements

Identifier	Requirement
R01	Create a deployment design for Manufacturing Application for the Anacreon lunar site.
R02	The application must remain true to Rob Nelson's original design as much as possible.
R03	Design must by highly available.
R04	Design must include an IPv6 topology.

1.3 Constraints

Identifier	Constraint
C01	Cisco has been selected as the compute and networking platform vendor.
C02	Netapp has been selected as the storage platform vendor.
C03	VMware has been selected as the virtualization software platform vendor.
C04	All components must fit in a 21U rack enclosure.

C05	All IP networking must be IPv6.

1.4 Assumptions

Identifier	Assumption
A01	The world economy has collapsed since the Zed outbreak rendering money useless. Therefore cost is not a factor.
A02	All power on the Anacreon station is regulated and highly available.
A03	The cooling system can accommodate any equipment that can fit in a 21U rack.
A04	All needed software is available and licensing is not a concern.

1.5 Risks

Identifier	Risk
K01	The system is reliant on the Lunar base's power grid.
K02	Small concessions on theoretical availability have been made for large gains in capacity and performance.
K03	The VMware vCloud Automation Center Appliance is not supported when running IPv6.
K04	The power consumption of the NetApp FAS2552 is unknown.

1.6 Contributors

Byron Schaller – Author

Dan Barber – Reviewer

Daryl Atkinson – Technical Advisor

Samwell Tarly – Brother of the Night's Watch

2 Physical Design

2.1 Physical Rack Design

Pursuant to Constraint C04 the entire infrastructure of the Anacreon system must not exceed 21 U.

The complete system for the Anacreon site occupies 20U. The following table details the breakdown by component.

Component	Height (U)
Cisco UCS 5108 Chassis A	6
Cisco UCS 5108 Chassis B	6
Cisco 6248 Fabric Interconnect A	1
Cisco 6248 Fabric Interconnect B	1
Cisco Nexus 5548UP A	1
Cisco Nexus 5548UP B	1
NetApp FAS2552 A	2
NetApp FAS2552 B	2
Total	**20**

The following diagram details the configuration:

2.2 Physical Compute Design

In adherence with vendor choice in Constraint C01, Cisco UCS will be the server platform deployed. Due to the space limitation specified in Constraint C04 the UCS B-series blade system will be used. This allows for the maximum amount of processing power with the smallest physical and environmental footprint.

Using converged network adapters to service both the IP and FCOE networks reduces cabling and eliminates the need for dedicated FC switches and HBAs. This cuts down on overall power and cooling needs. Both of which are limited.

This system was designed to have the maximum compute capacity possible. Per Assumption A01, cost was not taken into consideration.

The 21U rack will contain 2 6U Cisco UCS B 5108 8-slot Bade Chassis for a total of 12U. Each 5108 will be configured with 8 fans, 4 power supplies, 8 B230 M2 half-width blades, and 2 2208 Fabric Extenders (FEX).

Maximum number of hosts in this design is 16.

A dual chassis configuration was chosen to eliminate the chassis its self as a single point of failure. Although loss of half the available systems would greatly impact overall system performance it would still not be a total loss.

The B230 M2 was chosen as it provides the most per-U compute power as well as the smallest possible per-server failure domain.

The Cisco UCS VIC 1280 Mezzanine card was chosen because it is the ideal I/O interface for VMware ESXi 5.5. It offers 40 GB of throughput (twice that of the first generation VIC), as well as the ability to create virtual NICS and HBAs. The VIC 1280 also supports hardware fail-over in the event FEX A or B failing.

The 2208 Fabric Extender was chosen because it has twice the throughput of the first generation FEX modules. It is also required by Cisco UCS VIC 1280 mezzanine card.

The 300 GB Enterprise Performance SSDs will be configured in a RAID 1 array. The resulting volume will we used by vSphere as a Virtual Flash Read Cache as well as the local install of ESXi.

The following table details the configuration of each B230 M2 blade.

Type	Component
CPU	2 x Intel® Xeon® Processor E7-2870 (30M Cache, 2.40 GHz, 10-core)
Memory	32 x 16 GB DIMMs (512GB @1066 Mhz)
SSD	2x 300 GB Enterprise Performance SSD RAID 1
Mezzanine Card	Cisco UCS VIC 1280
Virtual NIC	2x Virtual NICs
Virtual HBA	2x Virtual FC HBAs

The following table details the total compute capacity for the system.

Metric	Capacity
CPU Cores	320
CPU Ghz	768 Ghz
Logical CPUs (w/ Hyper-Threading Enabled)	640
Memory	8 TB

2.2.1 Operating System

All hosts will be installed with VMware vSphere 5.5 Enterprise Plus.

2.2.2 Naming Convention

Hosts will follow this naming convention:

Lunar-ESXi-C*XX*B*XX*

Where C*XX* is either C01 or C02 depending on the chassis the server resides in and B*XX* is between B01 and B08 depending on the slot the server resides in.

2.2.3 UCS Manager

All blades will be deployed via UCS Manager with service profiles. As of this writing, version 2.2 of UCS Manager is known to have issues, therefore version 2.1 will be deployed.

2.3 Physical Storage Design

Pursuant to constraint C02, The Foundation as chosen NetApp as the storage vendor.

The storage subsystem in the Anacreon 21U configuration will consist of 2 NetApp FAS2552s. Each 2552 will be installed with dual controllers and 24 1.2 TB SAS drives.

The FAS2552 was chosen for its 2U form factor, FCoE support, connectivity options, capacity, and availability options.

Each controller has 4 10GB Ethernet ports. 2 ports on each controller will be connected to each Cisco Nexus 5548UP switch. Fiber Channel over Ethernet (FCoE) will be the storage networking protocol that is used. The FCoE network will be mapped to VLAN 10.

The Cisco Nexus 5548UPs are connected to the redundant pair of 6248 Fabric Interconnects (FI). The FIs are connected to the 2208 Fabric Extenders (FEX) in each Cisco UCS 5108 Chassis. The FEX are wired to the VIC 1280 Mezzanine cards in each blade server which presents two virtual HBA devices to the vSphere ESXi kernel. (Full diagram is located in the physical networking design section).

The controllers will all participate in a four node DataOnTap Cluster consisting of 2 active/active HA pairs. The controllers are ALUA capable.

Each FAS2552 will be configured for 1 RAID-DP group of 24 drives. Each group will contain one spare and two parity drives. The total usable capacity is 16.758 TB per node, 33.516 TB total.

LUNS will be created in increments of 2TB.

The following table details the capacity breakdown.

Total Disks per Node	24
Capacity per Disk	1.2 TB
Total Raw Capacity per Node	25.2 TB
Overhead per Node	8.442 TB
Total Usable per Node	16.758

2.3.1 Change from Original Design

Rob Nelson's original design contained a Synology RS2414+ for storing surveillance video, crew video journals and a copy of Hari Seldon's *Encyclopedia Galactica*.

The RS2414+ has been excluded from this design due to space constraints. The NetApp FAS2552 Cluster has adequate room to accommodate this data.

2.4 Physical Networking Design

Pursuant to Constraint C01 Cisco has been chosen as the network vendor by The Foundation.

The core of the Anacreon 21U configuration network is a pair of Cisco Nexus 5548UP switches. The switches are configured as peers in a Virtual Port Channel domain.

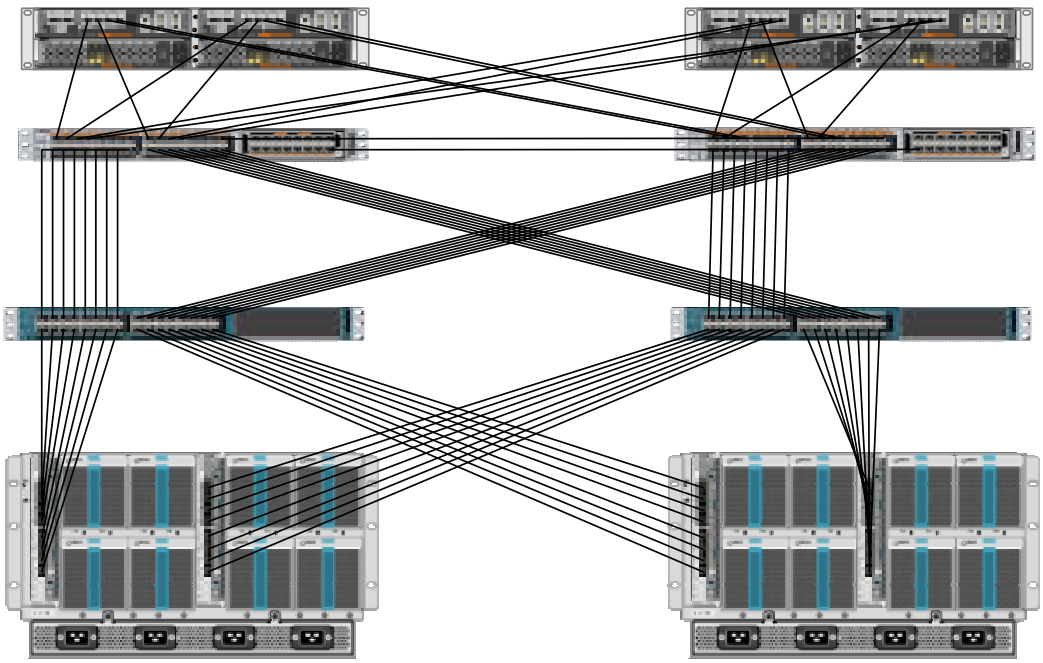
The Nexus 5548UP was chosen because of its 1U form factor, support of FCoE and IPv6, support for Virtual Port Channels, and unified ports.

Each 5548UP has 8 10GB connections to each of the 6248 Fabric Interconnects, 2 10GB connections to each of the NetApp FAS 2552 controllers, and 2 10GB connections to the peer 5548UP. This leaves 18 ports available for expansion and external connectivity on each switch.

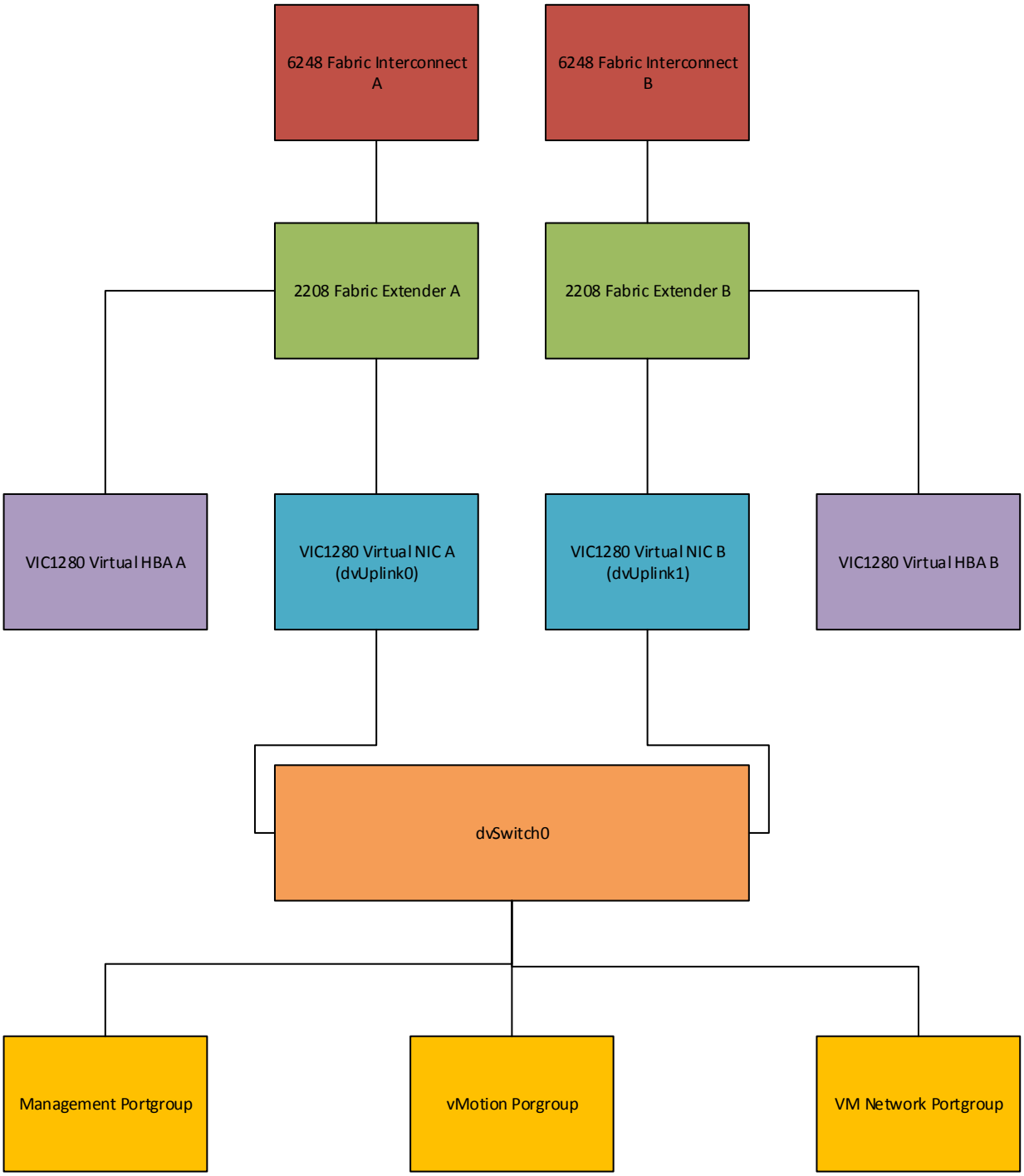
The Virtual Port Channel from the Nexus 5548UPs to the 6248 Fabric Interconnects are configured with 8 active and 8 hot standby connections each as that is the maximum supported.

The Cisco UCS 6248 Fabric Interconnects have 8 10GB connections to either FEX A or FEX B on both 5108 chassis. Each server has 2 connections internally to each FEX via the VIC 1280 Mezzanine cards. The VIC 1280 presents two virtual 10 GB NIC devices to the vSphere ESXi kernel.

The following diagram details the physical connectivity. Each line represents a 10 GB Ethernet connection:



The following diagram details the logical networking connections from the 6248 Fabric Interconnects all the way to the vSphere Distributed Switch Port Groups.



2.5 IPV6 Address Space

2.5.1 Lunar Networking

For scientifically unexplained reasons IPv4 does not work on the moon. Due to this phenomenon all IP networking must be accomplished with IPv6.

This has been documented as Constraint C5.

2.5.2 Usable IP Range

The IETF has granted use rights to The Foundation for the 2222:1:1::/48 network. The Foundation has agreed to implement IEFT Best Current Practice 157 (based on RFC 6177). BCP 157 recommends the use of /64 subnets.

Using /64 subnets on a /48 network allows for the creation of 65536 subnets. The Foundation feels this number will meet current needs and allow for future expansion as needed.

2.5.3 Virtual Local Area Networks (VLANs)

VLANs will be instituted to segment broadcast domains as well for an initial level of security.

The following table lists the VLANs, their intended use, the associated subnet, and whether it will be routed or not.

VLAN	Use	Subnet	Routed?
1	VM Network (Native)	2222:1:1:1::/64	Yes
2	Management Interfaces	2222:1:1:2::/64	Yes
3	vMotion	2222:1:1:3::/64	No
10	FCOE Fabric A	2222:1:1:a::/64	No
11	FCOE Fabric B	2222:1:1:b::/64	No

2.6 Environmental Factors

2.6.1 Power

Electrical power is the lifeblood of the Anacreon lunar outpost. Without clean, reliable, regulated power life would not be possible on the moon. This is the basis of Assumption A2.

With highly available, regulated power available on tap and due to the 21U space constraints a UPS device has not be included in this design.

This also introduces Risk K01, the system is reliant on the lunar base's power.

The two Cisco UCS 5108 Chassis as configured will draw 10,114W at 44.8A at maximum load.

The two 6248 Fabric Interconnects will draw 600W at 2.5A at maximum load.

The two Cisco Nexus 5548UP switches will draw 600W at maximum load.

The power needs of the NetApp FAS2552 is not known at this time. This is documented as Risk K04.

2.6.2 Cooling

It is assumed in Assumption A3 that the cooling system can accommodate any equipment that fits in one 21U rack.

The two Cisco UCS 5108 Chassis as configured will require 25369.3 BTU/hr of cooling.

The two 6248 Fabric Interconnects will require 1876.7 BTU/hr of cooling.

The two Cisco Nexus 5548UP switches will require 1331 BTU/hr of cooling.

The two NetApp FAS2552s will require 3474 BTU/hr of cooling.

The total required for all equipment is 32051 BTU/hr.

2.7 Design Consideration: Capacity vs. Availability

In every design there are considerations. They take the form of concessions and trade-offs. The largest consideration when creating this design was that of capacity vs. availability.

As designed the system is highly available to the order of 99.9%. To achieve 99.99% uptime both the compute and storage capacity would be halved. This was deemed too high a cost to save a theoretical 7 hours of unplanned downtime per year.

This consideration has been documented in Risk K02.

3 Virtual Machine Design

3.1 Original Design

In Rob Nelson's original design for the terrestrial shuttle assembly plants he specified 4 kinds of VMs: management, VMware systems and appliances, Red Hat Enterprise Linux (RHEL) management VMs, and the manufacturing system. The following tables are derived from the configuration specifications listed in his design.

3.1.1 Management Systems

Service	vCPUs	Memory (GB)	System Disk	Data Disk
Domain Controller 1	2	8	60	0
Domain Controller 2	2	8	60	0
RDP Session Host 1	2	32	60	300
RDP Session Host 2	2	32	60	300
RDP Session Host 3	2	32	60	300
RDP Session Host 4	2	32	60	300
RDP Licensing 1	1	4	60	0
RDP Licensing 1	1	4	60	0
vCenter	4	32	100	1000
Total	**18**	**184**	**580**	**2200**

3.1.2 VMware Systems and Appliances

Service	vCPUs	Memory (GB)	System Disk	Data Disk
ID Appliance	1	2	10	0
vCAC Appliance	2	8	30	0
IaaS Components	2	8	30	0
VDPA 1	4	4	3100	0
VDPA 2	4	4	3100	0
vShield Manager	2	8	60	0
vShield Edge	2	1	.5	0
vShield Endpoint (6)	6	3	3000	0
Total	**23**	**38**	**9330.5**	**0**

3.1.3 RHEL Management

Service	vCPUs	Memory (GB)	System Disk	Data Disk
Kickstart	1	.5	100	0
Puppet master	4	8	100	0
Gitolite	2	8	500	0

Jenkins CI	2	8	500	0
Total	**9**	**24.5**	**1200**	**0**

3.1.4 Manufacturing System

Service	vCPUs	Memory (GB)	System Disk	Data Disk
nginx	1	1	50	0
RestMQ	1	2	50	0
MongoDB	2	4	50	200
Watchdog	1	.5	50	0
Load Balancer	1	4	50	0
Total	**6**	**11.5**	**250**	**200**

Note: The original design called for 1 Load Balancer for every 2 servers of each of the other 3 tiers. This is not needed. A Load Balancer is only needed at the presentation tier. Also, 2 Load Balancers should be able to handle the load of hundreds of web nodes.

3.1.5 Totals

Grouping	vCPUs	Memory (GB)	System Disk	Data Disk
All Management	50	246.5	11110.5	2200
Manufacturing (1 instance, no LB)	5	7.5	200	200
Load Balancer (2 instances)	2	8	100	0
Total Available	640	8000	See Note	See Note
Available minus All Management and LBs	588	7745.5	See Note	See Note

Based on these findings the constraining factor on the number of instances of the three-tier manufacturing application is the number of vCPUs.

Note: Storage Capacity is not being considered because all management VMDKs will be thin provisioned and all workloads provisioned by vCloud Automation Center will be deployed via NetApp FlexClone.

With 588 available vCPUs, assuming hyper-threading, and five vCPUs needed per instance (four VMs) of the manufacturing application, 117 parallel instances can be deployed.

At maximum capacity a total of 496 VMs will be deployed. One vCenter Standard server can handle 10,000 powered on VMs, therefore one vCenter server is adequate for this deployment.

3.2 Network Addresses for VMs

3.2.1.1 Management VMs

All management VMs as well as the Load Balancers will have statically assigned IPv6 addresses in the 2222:1:1:1::/64 subnet. A full listing of all machines and addresses can be found in Appendix A.

3.2.1.2 Application VMs

All IP addresses for the Manufacturing Application VMs will be assigned dynamically from a DHCPv6 service that resides on the Domain Controller 1 server.

The scope will include the range of addresses from 2222:1:1:1::2:1 to 2222:1:1:1::2:ffff. All addresses will be dynamically registered with the forward (using AAAA records) and reverse zones in DNS.

3.3 Provisioning Storage Capacity

3.3.1.1 Management VMs

In the original terrestrial design greatly over provisions system drive capacity on nearly all VMs. This was due to an over-abundance of capacity and no concerns for space, power or cooling.

Due to the restrictions placed on the Anacreon base configuration (see Constraint C4) and a desire to preserve as much of the original design as possible, all Management VMs will be thin provisioned.

3.3.1.2 Application VMs

All Application VMs will be provisioned via vCloud Automation Center (vCAC).

vCAC natively supports Fast Provisioning using NetApp's FlexClone technology. FlexClone dramatically reduces time to provision as well as storage capacity utilization. For these reasons, all Blueprints created in vCAC will deploy machines with the NetApp FlexClone option.

4 VMware vSphere Design

4.1 Original Design

The original terrestrial design calls for a vCenter managed cluster of vSphere Enterprise Plus hosts provisioned via Auto Deploy.

All Hosts will be configured for Lock-Down mode.

The vCenter server is to be installed on a Microsoft Windows server. Single Sign-On and Auto Deploy components are to be installed on the same server.

4.1.1 Cluster Settings

The vSphere Cluster will be configured with EVC, HA, and DRS enabled. HA admission policy is to be set for a percent of cluster resources equal to one host.

Host monitoring is enabled with default restart priority (Medium). VM Monitoring is set to disabled.

DRS is configured for Fully Automated with the default threshold (3).

The following table contains the HA and DRS rules from the original design:

Rule Type	VMs
DRS VM-VM Anti-Affinity	DC1, DC2
DRS VM-VM Anti-Affinity	RDPLicense01, RDPLicense02
DRS VM-VM Anti-Affinity	RDPSH01, RDPSH02
DRS VM-VM Anti-Affinity	RDPSH01, RDPSH04
DRS VM-VM Anti-Affinity	RDPSH03, RDPSH04
DRS VM-VM Anti-Affinity	VDPA01, VDPA02
DRS VM-VM Anti-Affinity	DevWatchdog01, DevWatchdog02
DRS VM-VM Anti-Affinity	QAWatchdog01, QAWatchdog02
DRS VM-VM Anti-Affinity	ProdWatchdog01, ProdWatchdog02
VM Override VM Restart Policy - High	Management - vCenter, DCs
VM Override VM Restart Policy - High	Automation - Identity App, vCAC App, Gitolite VM
VM Override VM Restart Policy - High	Manufacturing - Database and Watchdog VMs
VM Override VM Restart Policy - Low	Web Front End VMs

4.1.2 Backup

Backup of virtual machines is handled by VADP per the original design.

4.2 Required Design Modifications

Constraints imposed by the Anacreon base (namely C04 and C05) require that the following modifications be made to the original design.

4.2.1 Auto Deploy

Auto Deploy does not work in an all IPv6 environment due to the fact that the PXE boot infrastructure required by Auto Deploy only works with IPv4.

As a result, all hosts will have ESXi locally installed. Host Profiles will still be used to ensure consistency across hosts.

Auto Deploy is also no longer available for patching. Because of this, vSphere Update Manager will be deployed on the same server as vCenter to apply host updates.

4.2.2 Chassis Based DRS Host Groups

To limit the impact of a possible failure of a 5108 blade chassis, new DRS Host Groups are required. A group will be created for the hosts in each chassis and the original DRS VM-VM Anti-Affinity rules will be changed to Host Group "should" rules with one of each of the VM pairs residing on each chassis.

4.2.3 Single-Sign On

The SSO service has known issues documented in VMware KB 2035454. The service must be configured using FQDNs and not IPs. Also the IPv4 interface must be installed on the network adapter even though it is not used.

4.2.4 HA Admission Control Policy

Admission Control Policy will be disabled. As noted in the above section on availability vs. capacity it has been decided that it is better to run VMs with contention that to disallow startup of a workload.

4.3 Firewall

The original design calls for the use of the Fortigate-300C. Due to the unknown nature of supportability of IPv6 on the Fortigate-300C, it has been replaced with a Cisco ASA 1000V Cloud Firewall.

4.4 Datastore Design

The fundamental change to the underlying storage architecture requires a new datastore design.

Each NetApp FAS2552 will present 8 2TB LUNs to the hosts for a total of 16 2TB LUNs. Storage DRS will be enabled and these LUNs will be aggregated as a single DataStore Cluster.

The dual virtual FC HBAs presented by the VIC1280 are technically over-subscribed by a factor of 2 to 1. To gracefully handle contention caused by bursting Storage IO Control will be enabled.

In alignment with NetApp recommend settings the Path Selection Policy will be set to Round Robin and IO Operation Limit set to 1.

The NetApp FAS2552 supports VAAI so block storage primitives such as Atomic Test and Set and Block Copy will be usable.

4.5 Virtual Flash Read Cache

To improve storage performance and maximize capacity utilization each host has been configured with 2 300GB Enterprise Performance Flash Drives in a RAID 1 array. Space not used by the ESXi install or the host Swap file will be allocated as vFlash Read Cache.

4.6 Host Networking Design

The VIC1280 Mezzanine card presents 2 virtual 10GB NICs to each host. These NICs are both members of an uplink group for a single distributed switch (DVS). The DVS contains 3 port groups: VM Network, Management, and vMotion.

Like the virtual HBAs presented by the VIC1280, the virtual NICs are also over-subscribed by a factor of 2 to 1. To gracefully handle contention caused by bursting Network IO Control (NIOC) is enabled.

The following table details the port group settings:

Port Group	Fail Over	Allocated Bandwidth	NIOC Shares
VM Network	Load Based Teaming	10GB	High
Management	Route Based on Virtual Port	1GB	Normal
vMotion	Route Based on Virtual Port	10GB	Normal

Refer to the Physical Network Design section for a logical diagram of port groups.

4.7 Noteworthy Considerations and Risks

4.7.1 Server FQDNs

As a general principal to avoid issues, all servers referenced in a service configuration need to be configured as FQDNs not IPs.

4.7.2 vCloud Automation Center Appliance

The 6.0.1 version of the vCAC Appliance is not supported when configured with IPv6. However a proof of concept was deployed in a lab and it does in fact work.

The lack of support for this configuration has been documented as Risk K03.

4.7.3 vCenter Web Client

The vSphere Web Client does not work in IPv6 environments. Therefore the vSphere C# client must be used.

4.7.4 vCenter Syslog and Dump Collector

Although not specified in the original design, it is worth noting that the vCenter Syslog Collector and the vCenter Dump Collector do not support IPv6.

5 Application Design

Pursuant to Requirement R02 this design attempts to remain as true as faithful as possible to Rob Nelson's original design. The following changes and risks have been identified in the previous sections:

1. VMs will be provisioned by vCAC natively using NetApp FlexClone.
2. Only 2 Load Balancers will be used.
3. Running the vCAC Appliance in an IPv6 environment is unsupported.

5.1 Puppet Enterprise and IPv6

Puppet Enterprise fully supports IPv6.

5.2 Kickstart and IPv6

Kickstart fully supports IPv6 with the noipv4 kernel option.

5.3 Gitolite and IPv6

Gitolite fully supports IPv6.

5.4 Jenkins and IPv6

Jenkins fully supports IPv6.

For the complete original application design see Appendix B.

6 Appendix A: Static IPv6 Assignments

6.1 VLAN 1: VM Network: 2222:1:1:1::1:0/64

Gateway	:1
Domain Controller 2 1	:2
Domain Controller 2	:3
RDP Session Host 1	:4
RDP Session Host 2	:5
RDP Session Host 3	:6
RDP Session Host 4	:7
RDP Licensing 1	:8
RDP Licensing 1	:9
ID Appliance	:a
vCAC Appliance	:b
IaaS Components	:c
VDPA 1	:d
VDPA 2	:e
vShield Manager	:f
vShield Edge	:10
Kickstart	:11
Puppet master	:12
Gitolite	:13
Jenkins CI	:14

6.2 VLAN 2: Management Network: 2222:1:1:2::/64

Device	Last Double Octect
Gateway	:1
Lunar-ESXi-C01-B01	:2
Lunar-ESXi-C01-B02	:3
Lunar-ESXi-C01-B03	:4
Lunar-ESXi-C01-B04	:5
Lunar-ESXi-C01-B05	:6
Lunar-ESXi-C01-B06	:7
Lunar-ESXi-C01-B07	:8
Lunar-ESXi-C01-B08	:9
Lunar-ESXi-C02-B01	:a
Lunar-ESXi-C02-B02	:b
Lunar-ESXi-C02-B03	:c
Lunar-ESXi-C02-B04	:d
Lunar-ESXi-C02-B05	:e
Lunar-ESXi-C02-B06	:f
Lunar-ESXi-C02-B07	:10
Lunar-ESXi-C02-B08	:11

6248 F1 A	:12
6248 FI B	:13
6248 Cluster VIP	:14
5548UP A	:15
5548UP B	:16
FAS2552 Controller 1	:17
FAS2552 Controller 2	:18
FAS2552 Controller 3	:19
FAS2552 Controller 4	:1a

6.3 VLAN 3: vMotion: 2222:1:1:3::/64

vMotion Interface	Last Double Octect
Lunar-ESXi-C01-B01	:1
Lunar-ESXi-C01-B02	:2
Lunar-ESXi-C01-B03	:3
Lunar-ESXi-C01-B04	:4
Lunar-ESXi-C01-B05	:5
Lunar-ESXi-C01-B06	:6
Lunar-ESXi-C01-B07	:7
Lunar-ESXi-C01-B08	:8
Lunar-ESXi-C02-B01	:9
Lunar-ESXi-C02-B02	:a
Lunar-ESXi-C02-B03	:b
Lunar-ESXi-C02-B04	:c
Lunar-ESXi-C02-B05	:d
Lunar-ESXi-C02-B06	:e
Lunar-ESXi-C02-B07	:f
Lunar-ESXi-C02-B08	:10

7 Appendix B: Original Application Design by Rob Nelson

All of Appendix B originally appeared in Rob Nelson's original design. All content in this section was written by him. No changes, other than formatting, have been made.

7.1 vCAC Portal

The vCAC system components are installed according to the VM Design section. The vCAC portal allows authorized users (vSphere admins, manufacturing application developers, and certain high level Foundation officers) to provision numerous objects types ("X as a Service"). Day One blueprints in the Service catalog will include all of the standardized templates for the existing VMs (see VM Design), Active Directory objects and actions (new users, reset password, etc.). In addition, vCAC offers a REST API that allows the Foundation manufacturing system to interact with vCAC in an automated fashion.

vSphere admins will have the ability to manage the vCAC catalog, adding and updating entries as appropriate. Developers will NOT have the ability to update the catalog, as Puppet will be used to determine the correct application load out for each VM (see Puppet System), and will program the manufacturing system to use the REST API to provision objects as needed.

Management VMs will be restricted to vSphere admins only and can be deployed in any environment/network. Manufacturing VMs are restricted to developers and can be deployed in any of the manufacturing environments/networks (initially Development, QA, Production).

vCAC will rely upon vCenter templates (Windows), kickstart processes (RHEL) and Puppet services to provision new VMs. Other objects will rely on various service specific to the object type, such as ADS for user services.

7.2 Puppet System

The Puppet Configuration Management (CM) product is at the heart of the automation and orchestration engines and the manufacturing continuous delivery/continuous integration processes. The Foundation has chosen Puppet Open Source based on its popularity and familiarity to our personnel, and because of the well documented features and modules that allow it to manage both Windows and Linux VMs (or nodes). Puppet allows the developers to define the desired state of the node and ensure that state is achieved without worrying about how to get there.

Because Puppet applies a desired state to a system, it can start with a bare-bones template and add the required configuration during provisioning. This allows the use of a single Windows VM template per OS level (2012 and 2012R2) and the kickstart of all RHEL VMs. Node definitions can be adjusted and very rapidly, Puppet will bring all of the desired nodes into compliance with the new desired state. This "infrastructure as code" process allows for rapid code development and rapid application deployment.

All system provisioning requests (Windows or Linux) will be made through vCAC. When the user selects a VM via the catalog (see vCAC Portal for more information), vCAC will initiate a workflow to provision the VM.

1. User interactively selects or automated systems submit a REST API query for an item in the catalog and submits the request.

2. vCAC creates the kickstart file for provisioning (Linux only) and deposits it on the Kickstart VM.

3. vCAC initiates the VM provisioning workflow

4. The workflow builds the VM from template (Windows) or creates an new VM of the correct size (Linux)

a. (Linux) The VM receives a kickstart file and downloads/installs the appropriate files.

5. The workflow forces the VM to register with puppet and signs the certificate (ref).

6. The vCAC workflow forces the new VM to check in to Puppet and receive its initial setup.

7. vCAC repeats steps 2-6 for any additional VMs required by the selected catalog entry.

8. vCAC notifies the requestor the catalog request has been fulfilled.

The vSphere administrators are responsible for configuring vCAC, the Kickstart server, the vCAC workflows, and ensuring communication between all components shown above.

The Kickstart server provides RHEL system contents for newly provisioned VMs, plus the kickstart file for the new VM. Each RHEL VM receives nearly the same basic kickstart file, only the node elements such as networking vary. Any two provisioned VMs will start with the exact same minimum level of software required to run Puppet.

The vCAC workflows communicate between vCAC, Kickstart, and newly provisioned nodes. The vSphere administrators and the developers work together to create the proper workflows.

The vSphere administrators will manage the puppet master instance(s) and the puppet code for management VMs. The developers will manage the puppet code for all remaining VMs. Both groups will provide oversight and assistance to each other, ensuring basic sanity checks and adherence to process, which ensures the manufacturing SLA (99.99%) is maintained.

Developers are responsible for aligning the puppet code with manufacturing releases. The puppet code determines the system configuration performed during provisioning step #6. The vCAC catalog will allow selection of a number of environments, including Development, QA, and Production, and developers are responsible for ensuring the Watchdog services provision in the correct environment.

All code changes also involve the Gitolite and Jenkins CI VMs. Gitolite is the authoritative version control source for all Git repositories. Developers will have their own Git workspaces, but only Gitolite and the RDP servers are backed up by VDPA. Jenkins CI is a Continuous Integration tool that is used to test changes to the manufacturing application code. The development workflow defines that all code must be tested by Jenkins CI and pass all tests to be merged upstream. This includes both puppet and manufacturing code, further validating code before deploying to any environment and assisting in maintaining the 99.99% SLA.

vSphere administrators will also follow the Gitolite/Jenkins CI and Dev->Production workflow when modifying puppet code for the management systems. Puppet will also be used to ensure a consistent state and make changes to vCenter. Developers will provide assistance as needed.

The threat from Zeds to humanity's very survival is real and the manufacturing capabilities of the Foundation are vital to reducing that threat. The use of Dev->Production promotion ensures that all changes, even those rapidly required to adapt to future manufacturing requirements, are thoroughly

vetted for software bugs AND material output quality before being introduced to Production. The potential for outages can never be eliminated but combined with other availability measures, the possibility continues to reduce to a near-0 value.