Documenting: Virtual Design Master – Challenge 1

Virtual Design
Master

Presented to: Messrs. Virtual Design Master Judges

Ref: AA-vDM-03-01

Prepared by: **Abdullah Abdullah**

Lebanon, Beirut

Twitter: @do0dzZZ
Blog: http://notes.doodzzz.net

E-mail: abdullah@lowerlayer.net

[Synopsis]

We are now settled on Mars, and ready to build a more permanent infrastructure. Keep in mind that power, cooling, and space are extremely expensive resources on Mars. In order to save space, we have decided not to use a traditional Fibre Channel infrastructure, meaning there will be no dedicated Fibre Channel Switches. We do however have plenty of 10G Ethernet switches, with some 40G Ethernet switches. We have three data centers on the planet, in order to provide high availability for our most critical applications. Our most critical system is our Environmental system, which is responsible for production of water and oxygen, as well as reclamation of waste. Should the environmental systems fail, the pods we live in work in can be sustained for only 20 minutes with the existing oxygen in the pod. We rely on this environmental system to control these resources, as well as to warn us when mechanical components throughout the system are failing or have failed. Our second most critical system is the system which controls our greenhouses. Any failure in this system will likely lead to major issues with our food supply. While we have the ability to communicate via radio if needed, many of the residents on Mars are used to e-mail and collaborative applications and prefer to use them when possible, since it makes them feel more at home. Your infrastructure should also be able to support the deployment of an unknown business critical application in the future.

Table of Contents

1. Executive Summary

1.1. Project Overview:

After jumping from earth to the moon and then finally to Mars the high council of intergalactic datacenter foundation has put their mindset to establishing a reliable and scalable computing environment.

1.2. Intended audience:

The document is to be used by the party responsible for implementing this solution following this design guide (smart monkeys are not allowed and you can't take selfies while moving across the datacenters)

1.3. Project Insights:

To utilize the three datacenters to support the systems that are required for the resident's survival on the planet.

The implementation should support a high available infrastructure spread across the three datacenters, hereunder is a breakdown of the systems along with their priority:

Priority	System	Description
1	Environmental system	Responsible for production of water and oxygen, as well as reclamation of waste. We rely on this environmental system to control these resources, as well as to warn us when mechanical components throughout the system are failing or have failed.
2	Greenhouse Systems Control	Responsible for food supplies
3	E-mail and collaborative applications	Residents utilize these services to communicate with each other.

1.3.1. Project Requirement:

- R001: High availability for the most critical applications.
- R002: Deployment of an unknown business critical application in the future.

1.3.2. Project Constrains:

- C001: Power, cooling, and space are extremely expensive resources on Mars
- C002: Save space (10G Ethernet switches, with some 40G Ethernet switches), No FC Switches.
- C003: Environmental systems fail, pods we live/work in can be sustained for only 20 minutes.
- C004: Greenhouse fail leads to major issues with our food supply.

1.3.3. Project Assumptions:

- A001: The environment will be utilizing VMware as a core technology.
- A002: All software packages and licenses were brought from earth.

- A003: The administration team is fully qualified to administer a VMware infrastructure.
- A004: External components are already available (Active Directory, DNS,
- A003: Server hardware is procured and has C001 taken into consideration.
- A004: Storage hardware is procured and qualifies for usage of iSCSI storage protocols.
- A005: Storage capacity is suitable for the production workloads.
- A006: Connectivity between the datacenters is covered taking into consideration C002.
- A007: Connectivity within each datacenter is covered taking into consideration C002.

2. Architecture Overview

Based on the study please find hereunder a conceptual diagram representing the infrastructure:
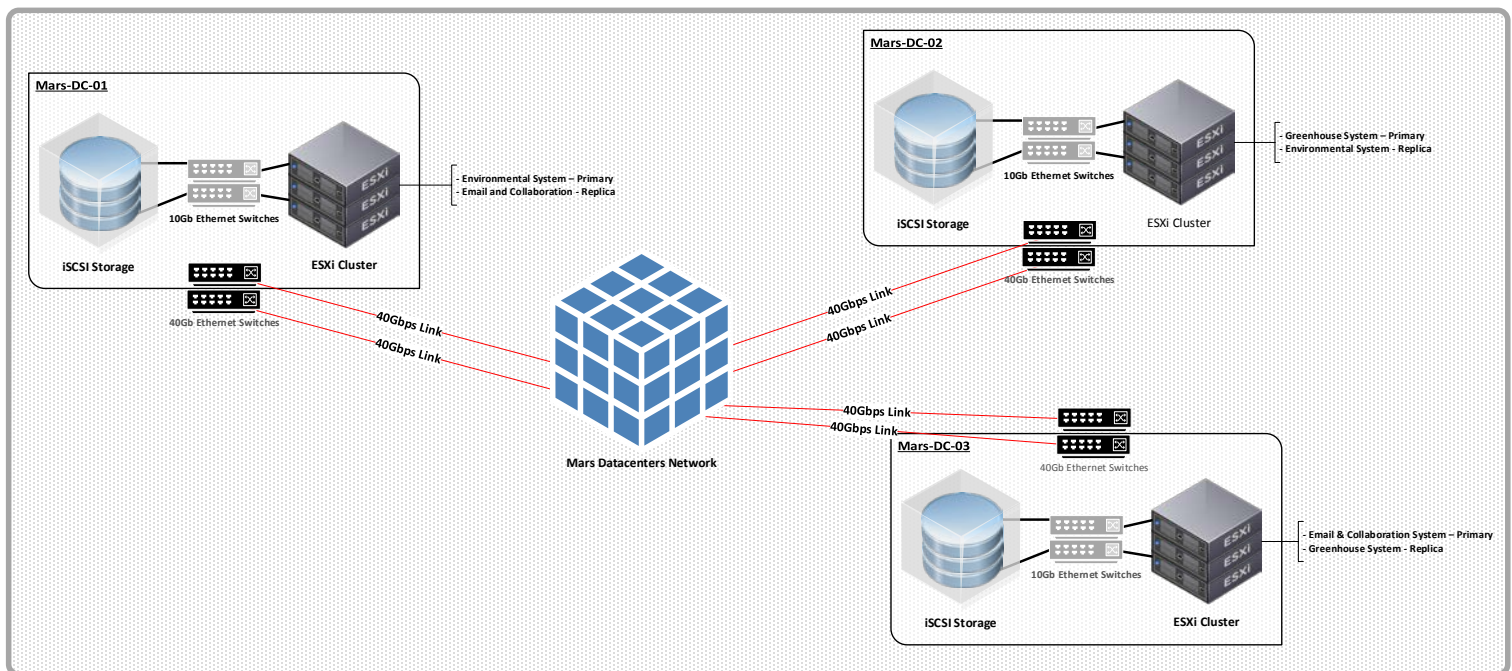
3. Design Summary

3.1. Physical Design:

Since the datacenters are in the process of construction and the hardware implementation is not in the scope of this design summary, after the completion we will add the physical design to this document in appendix A.

3.2. Logical Design:

Hereunder we will be going through the main pillars of the infrastructure and here we will be binding the components required to build the datacenters and achieve the needed requirement.

3.2.1. vSphere Datacenter:

In VMware vSphere, a datacenter is the highest level logical boundary. The datacenter may delineate separate physical sites/locations, or vSphere infrastructures with completely independent purposes.

Within vSphere datacenters, VMware ESXi™ hosts are typically organized into clusters. Clusters group similar hosts into a logical unit of virtual resources, enabling such technologies as:

- VMware vSphere® vMotion®
- VMware High Availability (HA)
- VMware vSphere Distributed Resource Scheduler (DRS)
- VMware vSphere Fault Tolerance (FT)

To address the requirement R001 we have two options to define the best availability:

3.2.1.1. Option 1 Continuous Availability:

- Application level high availability (i.e clustering/loadbalancing).
- All hardware/software single point of failures are pinpointed and thought of.

3.2.1.2. Option 2 High Availability:

- No application high availability required.
- VMs are high available through the hypervisor technology

Hereunder a table summarizing the comparison along with a reflection to the different design qualities:

Design Quality	Option 1	Option 2	Comments
Availability	↑	↑	Both options improve availability, though Option 1 would guarantee a higher level.
Manageability	↓	o	Option 1 would be harder to maintain due to increased complexity.
Performance	o	o	Both design options have no impact on performance.
Recoverability	↑	↑	Both options improve recoverability.
Security	o	o	Both design options have no impact on security.

Legend: ↑ = positive impact on quality; ↓ = negative impact on quality; o = no impact on quality

Design decision: Since option 1 suggests, a drawback in manageability and requires additional intervention and expertise from the staff, our preference is to go with **Option 2**.

3.2.2. vSphere Clusters:

3.2.2.1. Each datacenter will host its own cluster, hereunder you will find a table that reflects our recommendations:

Attribute	Specification
Number of clusters	1 per datacenter
vSphere cluster size per datacenter	3 hosts
Capacity for host failures per cluster	1 host
Dedicated hosts for maintenance capacity per cluster	1 host
Number of "usable" hosts per cluster	3 hosts
Total usable capacity in hosts	9 hosts

3.2.2.2. Each cluster will be configured for:

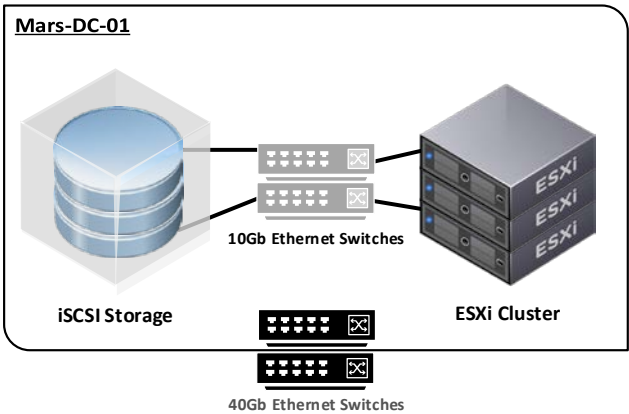| Feature	Status
High Availability	Enabled
Distributed Resource Scheduler	Enabled
Distributed Power Management	Enabled

Figure 01: A single datacenter logical design.

3.2.3. vSphere Hosts:

3.2.3.1. Each ESXi host that will exist in all datacenters will have the specifications reflected in the hereunder table:

Attribute	Specification
Host type and version	ESXi 6.0 (this might change upon the time of implementation)
Number of CPUs Number of cores Processor speed	Sized to support the required workloads.
Memory	Sized to support the required workloads.
Number of NIC ports (10Gb)	4

3.2.4. vSphere Networking:

3.2.4.1. The network design section defines how the vSphere virtual networking will be configured, the network architecture will be defined as follows:

- Separate networks for vSphere management, VM connectivity, iSCSI Storage, vMotion and Replication traffic.
- A distributed switch will be used least 2 active physical (or vNIC) adapter uplink ports.
- Redundancy across different physical adapters to protect against NIC or PCI slot failure.
- Redundancy at the physical switch level.

3.2.4.2. The following tables reflects the networks that will be configured on each vDS in **each cluster:**

Switch	Function	pNICs	Uplinks
VDS1 (vSphere Distributed Switch)	Management, Virtual machine, Storage over IP, VMotion and Replication networks.	12	4

Switch	Port Group	VLAN ID
VDS1 (vSphere Distributed Switch)	Management	100
VDS1 (vSphere Distributed Switch)	Storage over IP (iSCSI)	200
VDS1 (vSphere Distributed Switch)	vMotion	300
VDS1 (vSphere Distributed Switch)	Virtual Machines	400
VDS1 (vSphere Distributed Switch)	Replication	500

A distributed switch was chosen to take advantage of network I/O control, Load-Based Teaming, and Network vMotion. This VDS is configured to use 12 active Ethernet adapters distributed on 4 uplinks. All physical network switch ports connected to these adapters are configured as trunk ports that will pass traffic for all VLANs used by the virtual switch, also the physical NIC ports are connected to redundant physical switches.

Load-based teaming is configured for improved network traffic distribution between the pNICs and Network I/O Control enabled.

Virtual machine network connectivity uses virtual switch port groups and 802.1q VLAN tagging to segment traffic into four VLANs
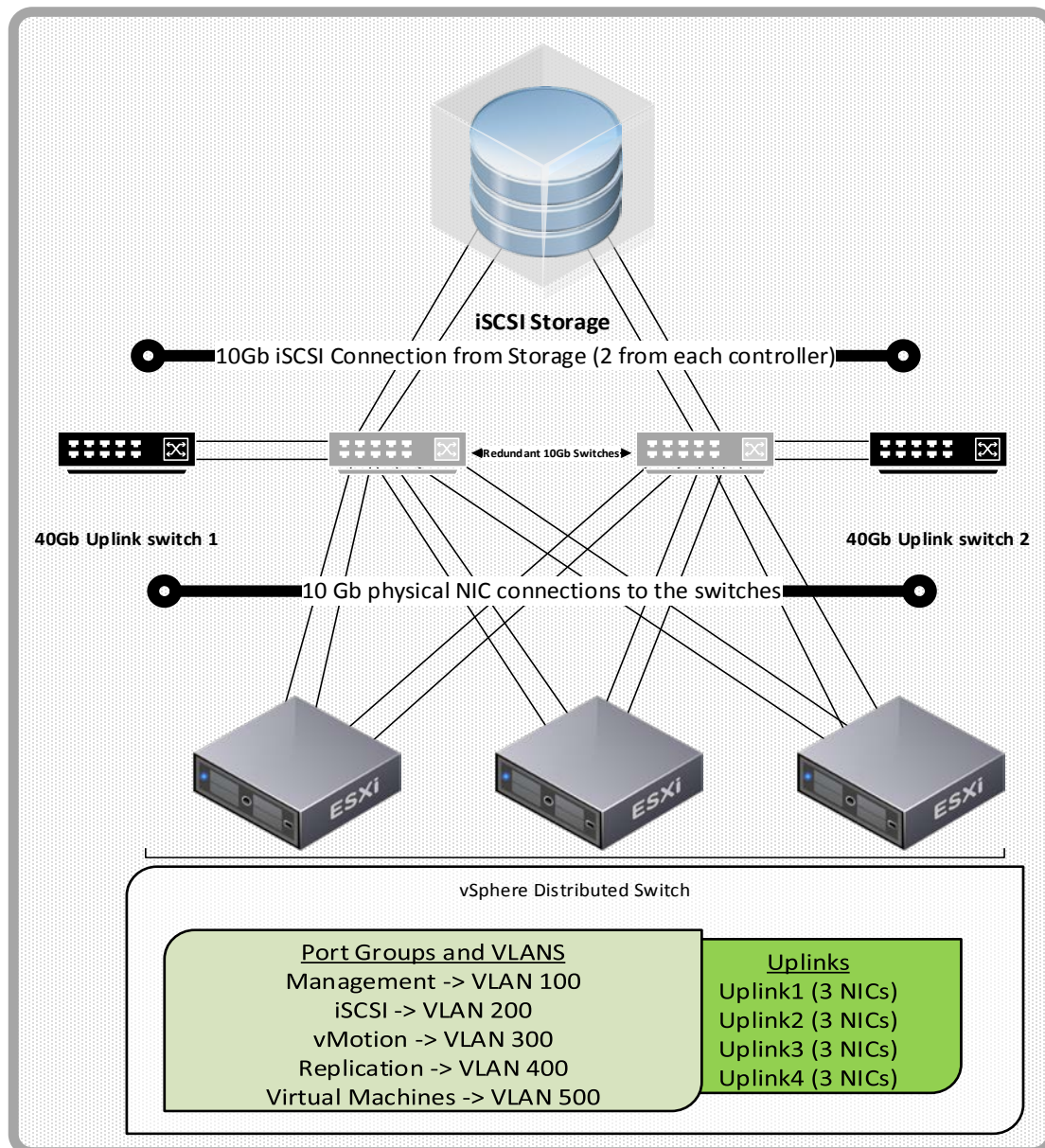

Figure 02: A single datacenter networking physical to logical design.

VDS Configuration Settings

Parameter	Setting
Load balancing	Route based on physical NIC load
Failover detection	Beacon probing
Notify switches	Enabled
Failback	No
Failover order	All active

3.2.5.   vSphere Shared Storage

The shared storage design section defines how the vSphere datastores will be configured. The same storage will be used in all datacenters.

Since we lack the FC switches thus we cannot use an FC storage which eventually lead to the usage of an iSCSI storage and it breaks down to this:

3.2.5.1. Shared Storage Platform

Attribute	Specification
Storage type	iSCSI (IP storage)
Number of storage processors	2 (redundant)
Number of switches Number of ports per host per switch	2 (redundant) 2
LUN size	Sized to support the required workloads.
Total LUNs	Sized to support the required workloads.
VMFS datastores per LUN	1
VMFS version	5

3.2.6.   vSphere Management Platform

3.2.6.1. vCenter Server

This is the heart of any vSphere infrastructure in terms of manageability, since the main intention of having 3 datacenter is to have a high available and robust infrastructure we have chosen to go with vSphere Replication and this suggests having a vCenter Server in each datacenter.

Attribute	Specification
vCenter Server version	6.0
Physical or virtual system	Virtual
Flavor	Appliance
Number of CPUs Processor type Processor speed	4 VMware vCPU N/A
Memory	16GB
Number of NIC and ports	1/1
Authentication	LDAP
High Availability	vSphere HA
Recoverability	Restore from backup

3.2.6.2. vSphere Replication

Since we're entitled to utilize 3 datacenters, it is recommended to split the workload of the systems amongst those datacenters.

In addition each datacenter will host a replica of the one of the systems found in the second and third datacenter and operation goes vice versa, hereunder is a table reflects this assignment:

Datacenter	Workload	Replica	RPO
DC1	Environmental System	DC2	15 minutes
DC2	Greenhouse System	DC3	15 minutes
DC3	e-mail and collaboration	DC1	15 minutes

- In case DC1 fails the environmental system will be restored in DC2.
- In case DC2 failed the greenhouse system will be restored in DC3.
- In case DC3 failed the email and collaboration systems will be restored in DC1.

vSphere replication will be utilizing its own port group and the replication traffic will be isolated completely from production traffic.

3.2.7. vSphere Environment Monitoring

Because the uptime and health of the entire technology infrastructure is critical, we will be leveraging the monitoring and alert features found in each product respectively:

- Hardware (Server, Network, and SAN Infrastructure) will be monitored via their respective management consoles and will be utilizing the e-mail system for notifications.
- vSphere monitoring will leverage the event monitoring and alarm system in it, the vCenter Server will be configured to monitor the health and performance of all critical virtual infrastructure components, including the ESXi hosts, the clusters, vSphere HA, virtual machine operations such as vMotion, and the health of the vCenter Server itself and will be sending notifications through the email system.

3.2.8. vSphere Backup/Restore Considerations

The environments high availability is not only achieved through replication, backup is always a must and it is crucial, we will be utilizing VMware VDP Advanced virtual appliances to:

- Backup the production virtual machines in each datacenter.
- Replicate the backup data between the datacenters.

Datacenter	Workload	VDP Replica
DC1	Environmental System	DC2, DC3
DC2	Greenhouse System	DC1, DC3
DC3	e-mail and collaboration	DC1, DC2

4. Future Workload Deployments

In terms of future workloads the design is very modular and can scale-out/up, it is very imperative to have a stable power source to support expanding the workloads in each datacenter.

That said, the environment is very much flexible and open to growth.

5. Appendices
5.1. Appendix A – Hardware physical design.
5.2. Appendix B – Networking specifications (IPs, Hostnames).
5.3. Appendix C – References
5.3.1. https://pubs.vmware.com/srm-51/index.jsp?topic=%2Fcom.vmware.srm.install_config.doc%2FGUID-128D4414-BA34-4159-8682-FF75A8DFF137.html
5.3.2. http://pubs.vmware.com/vsphere-replication-60/index.jsp?topic=%2Fcom.vmware.vsphere.replication-admin.doc%2FGUID-16677363-4265-4815-9C1C-DAAA3AE500CD.html
5.3.3. http://cormachogan.com/2015/02/12/vsphere-6-0-storage-features-part-4-vmfs-voma-and-vaai/
5.3.4. https://www.vmware.com/support/vsphere-replication/doc/vsphere-replication-60-release-notes.html
5.3.5. http://blogs.vmware.com/vsphere/2014/04/licensing-vdpa-replication.html
5.3.6. https://www.vmware.com/uk/products/vsphere/features/drs-dpm

5.3.7. VCDX Boot Camp - Preparing for the VCDX Panel Defense

5.3.8. VMware vSphere 5.x Datacenter Design Cookbook

5.3.9. VMware vCloud Architecture Toolkit