

Bachelor's Thesis (UAS)

Degree Program in Information Technology

2015

Antti Hurme

VMWARE PUBLIC CLOUD INFRASTRUCTURE

– Development environment design and
implementation



TURUN AMMATTIKORKEAKOULU
TURKU UNIVERSITY OF APPLIED SCIENCES

BACHELOR'S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Degree Programme in Information Technology

30.8.2015 | 37 pages

Instructor: Patric Granholm

Antti Hurme

VMWARE PUBLIC CLOUD INFRASTRUCTURE – DEVELOPMENT ENVIRONMENT DESIGN AND IMPLEMENTATION

The goal of this thesis was to create a development environment for Lounea's Public Cloud infrastructure for testing and development purposes. Lounea is a telecommunications company founded in 1887 as Oy Salo-Turku-Paimio puhelinyhdistys. Lounea has increased its focus in datacenter operations during the past years and public cloud services are a part of it. The objective was to analyze the current infrastructure and, based on the result, plan and implement a functional development environment with similar configuration. This thesis focuses on the design of a public cloud infrastructure created using VMware products to create a Software Define Datacenter for customer production virtual machines. The development environment had to have similar functionality as the production environment for results to be valid in both environments.

The thesis resulted in a functional VMware public cloud infrastructure built on legacy hardware dedicated for this purpose. The installation has later been used to plan and implement major version upgrades of the production environment with minimal downtime.

KEYWORDS:

VMware, vCloud Director, public cloud, Software Defined Datacenter, virtualization, networking, redundancy, datacenter design

FOREWORD

The main part of this thesis was carried out during the design and construction of the public cloud test environment over several months in 2014. The need for a test environment for Lounea's public cloud infrastructure allowed for an excellent thesis project, albeit it being extremely broad and technically challenging. The technologies required for a fully functional virtualization environment have been a part of my full time job for the past three years and yet there is room for improvement and thus this project has provided me with the tools to strengthen my knowledge.

I would like to thank my supervisor Patric Granholm for the extended support and Turku University of Applied Sciences for extending my study time which allowed me to complete the work at a prolonged time.

CONTENTS

1 INTRODUCTION	7
1.2 Concepts and Definitions	8
2 PRODUCTION INFRASTRUCTURE ANALYSIS	9
2.1 Public Cloud	9
2.2 Physical Infrastructure	9
2.3 VMware Infrastructure	12
2.3.1 ESXi Clusters	12
2.4 Security Analysis	16
2.5 High Availability design	16
2.5.1 Identity management, authentication and domain services	17
2.5.2 VMware vCenter Server	17
2.5.3 vCloud Networking and Security, vCloud Director	18
3 DEVELOPMENT ENVIRONMENT	19
3.1 Hardware	19
3.1.1 Blade Servers	20
3.1.2 Storage	21
3.2 Design Considerations	22
3.2 Implementation	23
3.2.1 Hardware installation and configuration	24
3.2.2 Hypervisor installation and configuration	26
3.2.3 Core services	28
3.2.4 VMware vCenter	29
3.2.5 vShield Network and Security and vCloud Director	31
3.3 Real world usage	34
4 CONCLUSION AND FUTURE WORK	35
REFERENCES	37

FIGURES

Figure 1. Physical iSCSI Ethernet topology.	10
Figure 2. Physical LAN network topology.	11
Figure 3. Physical management LAN network topology.	12
Figure 4. Virtual Network configuration.	15
Figure 5. vCloud Director service relations.	18
Figure 6. Blade server physical configuration.	21
Figure 7. HP Fibre Channel SAN Configuration Diagram.	27

TABLES

Table 1. Blade Server resource configuration	20
--	----

ACRONYMS, ABBREVIATIONS, AND SYMBOLS

IAAS	Infrastructure as a Service
SAAS	Software as a Service
PAAS	Platform as a Service
DNS	Domain Name System
DHCP	Dynamic Host Configuration Protocol
vCD	vCloud Director
vCNS	vCloud Networking and Security
vCS	vCenter Server
OS	Operating System
UPS	Uninterruptible Power Supply
PDU	Power Distribution Unit
iSCSI	Internet Small Computer Interface
TCP	Transport Control Protocol
VLAN	Virtual LAN
SAN	Storage Area Network
LUN	Logical Unit Numbering
DRS	Distributed Resource Scheduler
FT	Fault Tolerance
HA	High Availability

1 INTRODUCTION

Lounea Oy provides, among other services, a cloud infrastructure for clients to either expand or replace their own server needs. Servers are used to enable the company's goal by providing services and applications to their employees. Thus, the nature of a server is to be available at any given time. A cloud service provider rents out its capacity to the customers and one datacenter can provide computing and storage capacity to multiple customers. As such, an uptime of 99,9% is promised to the customers and from a service provider viewpoint, this means that all actions carried out in the datacenter can not cause disruptions to the service.

A development and test environment provides a safe system to plan and execute any changes on a live system without causing downtime on the production service. The goals of this thesis are to plan and implement a working test environment replicating the production environment as identical as possible with the hardware provided. As the hardware was not purchased, but rather consisted of end-of-life hardware decommissioned from production use, this thesis will focus on the theory for a high availability system in which the production environment is and on ways of implementing similar fault-tolerance and high-availability on a smaller scale.

This thesis consists of three parts, which are divided in the following manner:

1. Production environment analysis and public cloud infrastructure design concepts.
2. Development environment design and implementation
3. Future design considerations and use cases

This thesis discusses the different components required to provide a VMware vCloud Director environment which is used in the production environment. As this public cloud infrastructure was built and designed before this thesis, this thesis only focuses on VMware public cloud infrastructure and its components. Physical switch and router configuration and electrical details are left outside this thesis' scope.

1.2 Concepts and Definitions

Infrastructure as a Service (IAAS) refers to a service provider renting out slices of a dedicated computing and storage infrastructure to clients. The client is responsible for the operating systems and software which are purpose-built using the hardware provided by an IAAS service provider.

Platform as a Service (PAAS) focuses on building a standardized framework for the core functionalities which a service would require. PAAS usually refers to the service provider taking care of the Operating Systems and sometimes the underlying infrastructure to allow the software developer to focus on the core application.

Software as a Service (SAAS) is a service provided to the end user as is. The service can be hosted email, DNS, Customer management software or anything the end-user might need. The software is accessed using a client and the user does not have to be concerned with the underlying infrastructure.

Virtualization abstracts the underlying physical hardware and provides the required resources for an operating system through software. This makes it possible to use a mixed environment of different hardware vendors without impact on the guest software.

Hypervisor is the software which provides the necessary components to enable virtualization on a server. Hypervisor handles all input and output requests from a guest operating system and processes them on underlying hardware. Hypervisors make it possible to share resources between multiple guests (VMware, Inc. 2013e).

2 PRODUCTION INFRASTRUCTURE ANALYSIS

Lounea public cloud infrastructure is used by several customers around Finland to support their business operations and they all rely on the service provided to be excellent. The following chapter will describe the production environment with its different components that build up a fully functional IAAS platform.

2.1 Public Cloud

A public cloud infrastructure is purpose-built to provide end-users with a flexible and easily accessible environment for their business critical applications. The As-a-Service concept has been changing the IT-landscape by allowing customers to offload some or all underlying IT infrastructure to different service providers ranging from infrastructure as a service to software as a service. This has allowed IT-professionals to focus on supporting their business while offloading some of the more tedious work to a service provider.

Lounea Oy provides an IAASsolution to its customers with its public cloud infrastructure. As a service provider, Lounea Oy focuses on providing a highly available and fault tolerant system for its customers business critical applications. Lounea has traditionally been a telecommunications company, while lately shifting towards providing Internet connectivity and IT-services which include a VMware public cloud service.

2.2 Physical Infrastructure

Lounea Oy has two datacenters in its cloud infrastructure, one located in Salo and the other one in Forssa. Both locations are almost identical in their design and hardware to provide a balanced infrastructure which can handle a disaster recovery in the second site.

Each site has its own storage, networking and computing capabilities which are built with redundancy and high-availability in mind. Eliminating any single-points of failure from the physical layer to the software layer is as important as the other.

Starting from the power distribution, all physical hardware has been configured to tolerate power supply and power grid failures. Uninterruptable power (UPS) is provided in

minimum to half of all the power supply units which is typically the minimum required to keep any system alive. As UPS is a battery-powered, this will only provide enough power for a few minutes. In case of a longer power outage, generators have been placed to automatically start in case of a major blackout. This ensures that all systems will maintain full operation even if power distribution from the electrical company is severed for longer periods of time. Equipment cooling is connected in similar fashion and is also built redundant with at least two physical units.

All network components from storage to internet connectivity are redundant to any physical components failure. Storage systems have two uplinks for each controller, connected to two different switches in the same network. All servers have a minimum of two network uplinks for each networking purpose. Dividing physical networking to two physical switches results in a redundant networking topology even if a physical switch suffers from failure.

Storage connectivity is configured to use Internet Small Computer System Interface, more commonly known as iSCSI. iSCSI is an IP-based storage networking standard that allows servers to use a common, shared storage area network (SAN). All virtual machines are stored on shared storage to allow any hypervisor in the same storage network to run the virtual machine in question. In the production environment, iSCSI traffic is separated to its own physical infrastructure to provide an additional layer of security as well as to provide dedicated storage switching capacity.

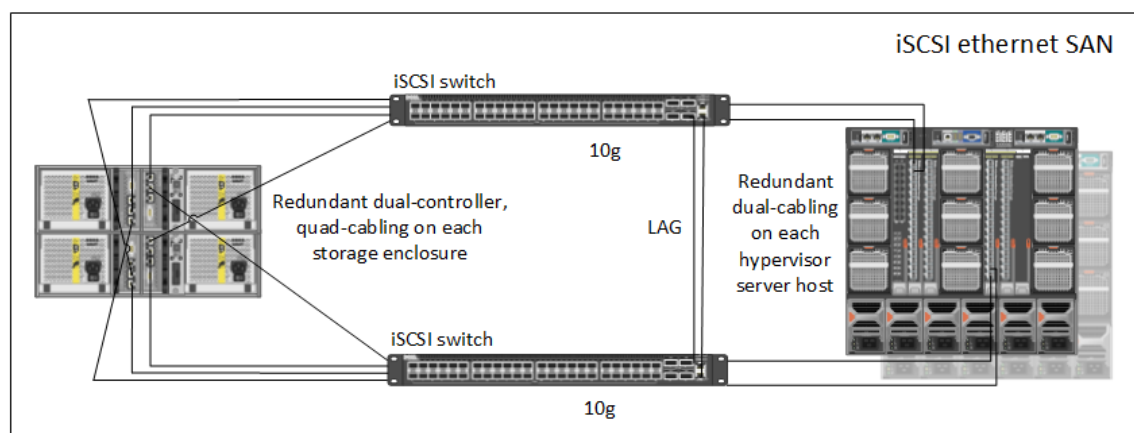


Figure 1. Physical iSCSI Ethernet topology.

In the same manner, all client internet traffic is separated from the management traffic with physical switches. All servers are connected to two switches for client traffic which is eventually routed through two redundant routers, or optionally a virtual LAN (VLAN) is transferred to the clients own site.

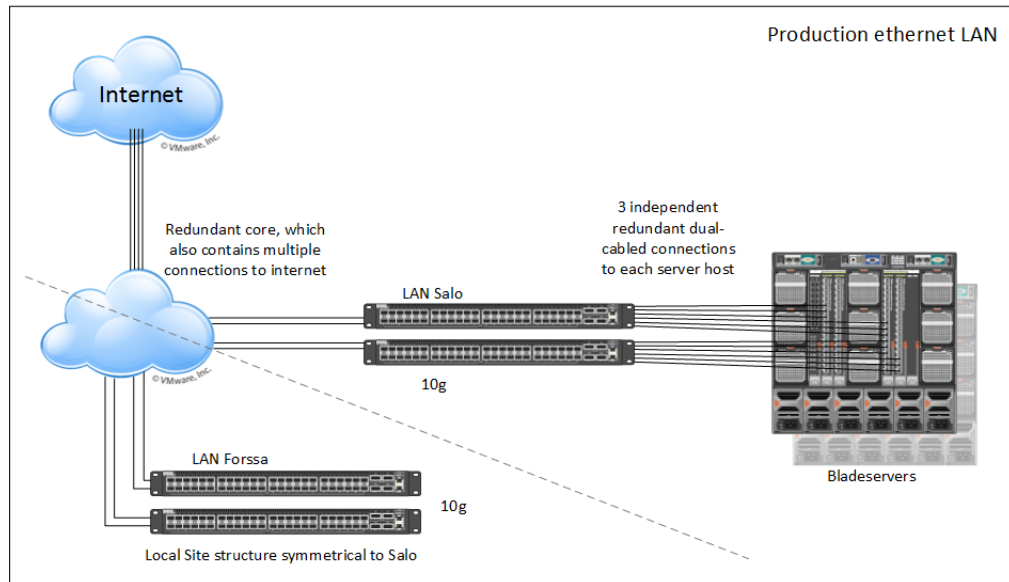


Figure 2. Physical LAN network topology.

IP traffic for administrative purposes is also made redundant with two switches to each server, storage and other equipment with management interfaces. Figure 3 contains a visual representation of this configuration.

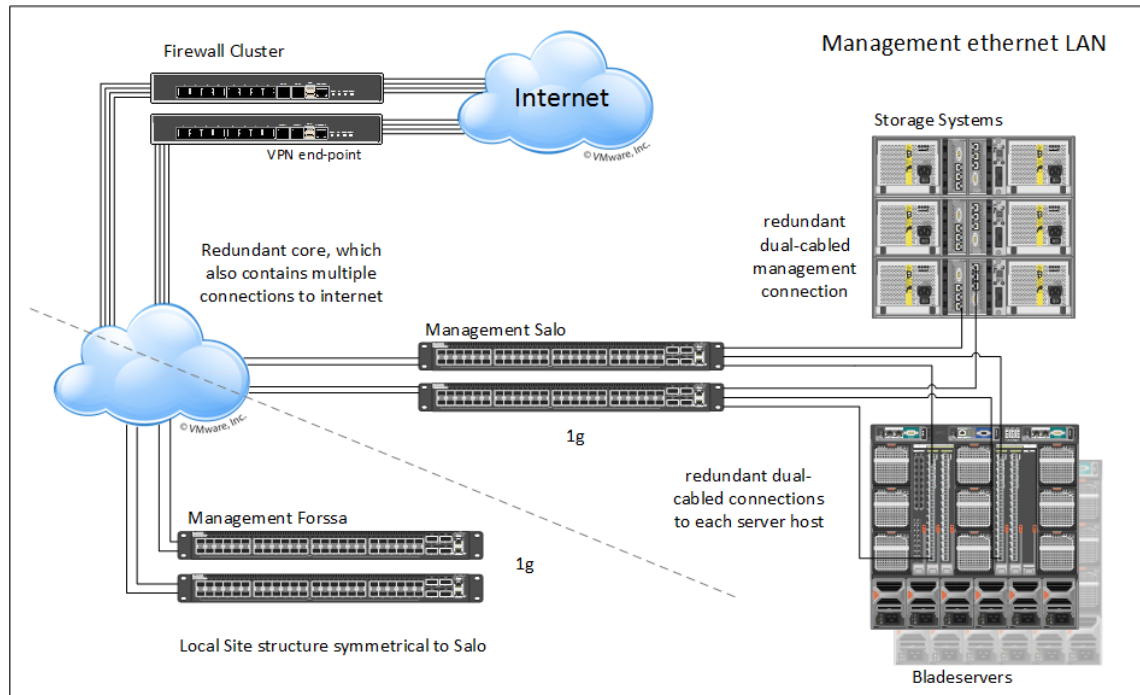


Figure 3. Physical management LAN network topology.

2.3 VMware Infrastructure

The VMware product suite has been chosen to serve as the virtualization platform for Lounea public cloud. VMware provides several products from the most basic hypervisor OS named ESXi to management software like vCenter server and vCloud Director.

This chapter provides information about each component and their core functionality. All of the following components are used in the production environment.

2.3.1 ESXi Clusters

Each physical server has a standalone installation of VMware ESXi installed and is configured identically when they serve a similar purpose. Those ESXi hosts which serve the client VMs share a similar configuration as where those hosts running management VMs have similar configuration. A cluster is formed out of the hosts serving the client VMs and a separate cluster is formed for the management environment. A cluster consists of two or more computing servers, in this case installed with VMware ESXi and provides the administrator with a single administrative unit that has the combined computing power of all hosts attached to the cluster. When a virtual machine is powered on, it can start on any of the hosts but will prefer the host with most resources available.

A cluster will also provide additional capabilities, such as fault tolerance, high availability and distributed resources. When a cluster is formed of two or more hosts, rapid recovery from a host failure is possible with high availability. ESXi hosts are constantly polled by a vCenter Server to ensure they respond, but hosts also communicate among themselves using a heartbeat signal to ensure they are responding. If a host fails to respond to the heartbeat signal for three consecutive times, a fail-over will occur and virtual machines will be restarted on the other hosts in the cluster. The production environment server capacity has been planned to withstand one physical server failure, which means that all computing resources (CPU and RAM) have been calculated using the formula $(\text{number of hosts} - 1) * \text{RAM/CPU}$. This ensures that all running virtual machines will be powered back on in case of one hypervisor failure.

As briefly mentioned before, the client and management clusters have been separated. This design choice ensures that any action taken on either cluster has minimal impact on the other. There are two client clusters, one in Salo and the second one in Forssa. These clusters act independently and have their own vCenter servers for management. The management cluster consists of its own vCenter server and the VMs running in this cluster contain all the necessary servers and services to provide an end-user experience.

The VMware vCenter Server is the fundamental component to create a clustered environment and offers a single point of view to all sites. All three vCenter servers are linked together to allow the administrators single log-in to the management interface to administer any of the three sites without changing administrative programs. The vCenter server also provides, but is not limited to, vMotion, DRS, Distributed Switches, HA, and fail-over capabilities.

2.3.1.1 vMotion

vMotion and storage vMotion allows a virtual machine to migrate between servers in an online state, or optionally with storage vMotion to move between different datastores without disruption to the virtual machine operation. vMotion is the only management traffic that shares the client Virtual Machine LAN infrastructure, due to the fact that those switches provide 10 Gbit switching capability while the management switches are limited to a 1Gbit transfer speed.

2.3.1.2 DRS

VMware Distributed Resource Scheduling (DRS) provides the automatic virtual machine placement within the cluster during initial power-on sequence. DRS chooses a host that meets the virtual machine resource requirements, and chooses a host that has the least amount of consumed resources. DRS also takes into account any affinity rules that might be placed on the virtual machine, for example “Should run on host X or must not run on host Y”. VMware DRS also makes cluster wide resource pools possible, which is required by vCloud Director to create virtual datacenter containers.

As the production environment also is both DRS and HA enabled, the vCenter Server will optimize virtual machine placement in a fail-over situation. This provides fast rebalancing of the remaining host resources between the virtual machines.

2.3.1.3 Virtual Distributed Switches

Virtual Distributed Switches (VDS) provide a single virtual switch that stretches over multiple VMware ESXi hosts in a single cluster and provides an easier method for the IT administrator to manage a complex virtual network. VDS have set uplinks on each ESXi host, and these uplinks are configured in a redundant manner to provide uninterrupted network connectivity in case of a physical switch failure.

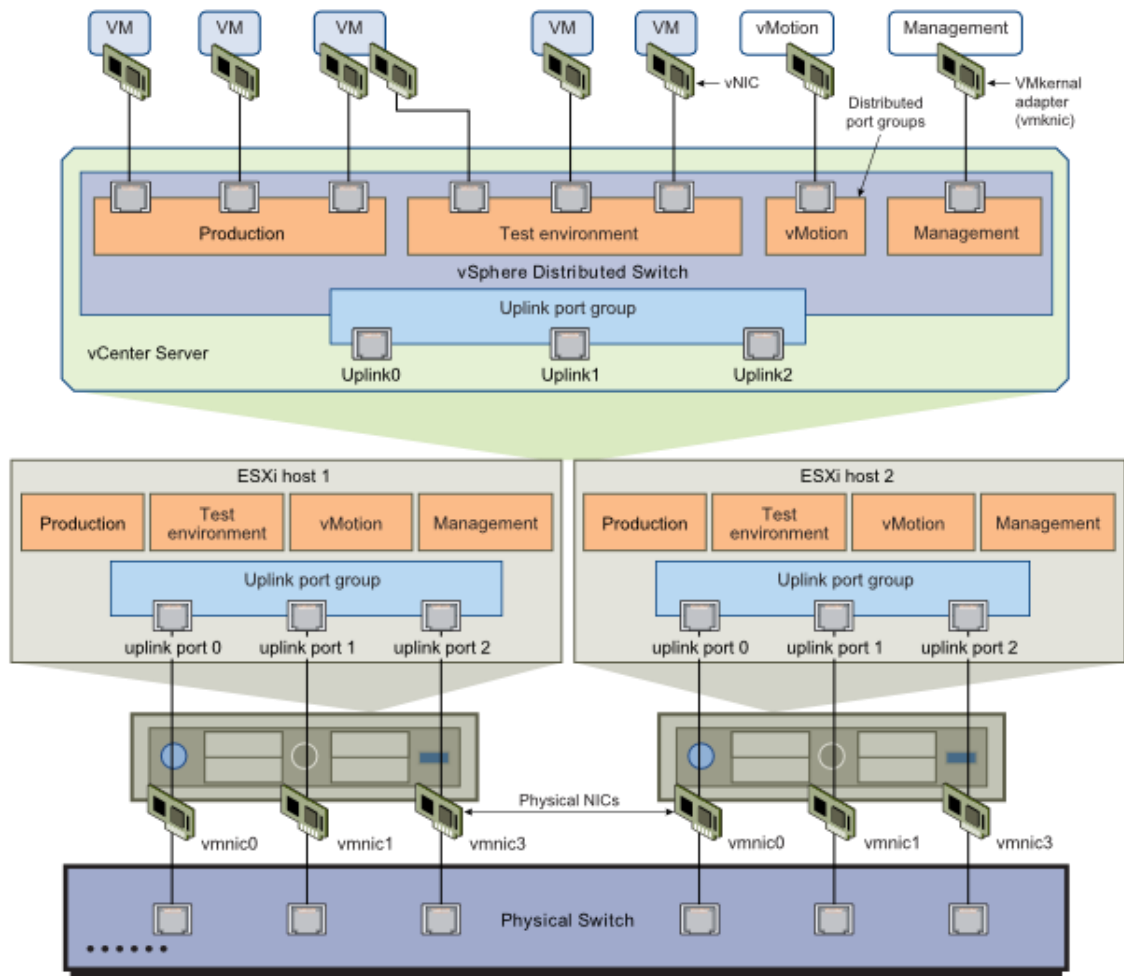


Figure 4. Virtual Network configuration.

Source: [image online] Available through: VMware publications website <<http://pubs.vmware.com/vsphere-51/topic/com.vmware.vsphere.networking.doc/GUID-BE95F7DF-96C2-4A83-B769-3D18B0D6F8C1-high.png>> [Accessed 21 May 2015]

VDS provide port groups which may be configured with different security and networking settings, and virtual machines are connected to different port groups based on the clients networking requirements. Although VDS support VMware management, iSCSI and other traffic to be configured, the production environment does not utilize virtual distributed switches or these networking needs.

2.4 Security Analysis

The production environment contains a mixed operating system infrastructure, consisting of both Windows and different Linux distributions. As such, multiple security base-lines have been agreed upon to provide a safe environment without obstructing day-to-day administration to a great extent. Therefore, some security-hardening guidelines have not been applied and the development environment will not contain all security aspects that are configured in the production infrastructure.

Authentication is handled by Microsoft Active Directory as well as Open LDAP. This provides a good centralized user management that is easily managed. Linux SSH access is only allowed to a specific user with public-key authentication. Any root access remotely is denied.

Any service that requires authentication to any system, has its own username and password. In this way, it is easier to isolate any authentication and or unauthorized access. Firewalls are enabled on each individual host and operating system as well as on the perimeter network providing external networking access. Firewalls deny all traffic that is not explicitly allowed.

VMware provides good security hardening guides which have been used as a reference for the VMware security configuration. These guides can be found on their public website (<http://www.vmware.com/security/hardening-guides>).

2.5 High Availability design

High availability design is the key component in providing a successful public cloud infrastructure. Lounea ensures its customers a 99,9% accessibility for their infrastructure and as such, any single point of failure in hardware has been eliminated where possible.

Building a fault-tolerant system is not limited to hardware, but all additional services should be built as resilient as possible. As there are two physical datacenter locations, a connection issue between the two sites might occur and the logical infrastructure must be resilient to such an event.

2.5.1 Identity management, authentication and domain services

Authentication is one of the core components in any environment that provides a service. If the system can not verify the identity of a user and filter out unauthorized access, then anyone could make changes anywhere. Although all systems deny access to a given resource if the identity can not be confirmed, any disruptions to the authentication services would prohibit the cloud to function as intended. All authentication services have been configured to provide authentication services even if the communication is severed between the two datacenters, and in normal operation to provide load balancing and software failover between the sites. To achieve this, two Microsoft Active Directory (AD) service servers have been deployed, each on its own separate VMware cluster. Active Directory also provides Lightweight Directory Access Protocol (LDAP) user and computer account authentication and Domain Name Services (DNS) resolution to the other components in the datacenter. DNS is configured on each client to query any DNS entry from its local DNS server first, and only if that name resolution fails, query the second DNS server on the second site.

Since VMware vCenter Server 5.1, any VMware related authentication is passed through VMware Single Sign On (SSO) service. The SSO service functions as an arbiter during authentication and two VMware SSO servers have been deployed in a similar manner as the Microsoft AD servers using the Multisite deployment method. The SSO server provides the user or service a token during the first successful authentication and each consecutive authentication is processed using this token, regardless of the VMware service the authentication process originates from.

Microsoft AD, DNS and VMware SSO provide the necessary authentication services and can function independently if necessary during a network loss between the two datacenters.

2.5.2 VMware vCenter Server

VMware vCenter server is currently the only major component which can not be built with fault-tolerance in mind from a software perspective. A VMware cluster can only be linked to one vCenter server and, as such, its possible failure would mean loss of VMware management. In addition, other vCenter server-managed functions, such as DRS would not function. As each VMware ESXi host in a cluster is connected to a vCenter server,

there have been three vCenter servers deployed. The vCenter server responsible for the clusters in Forssa has been placed in Forssa and the vCenter servers managing clusters in Salo has been placed in Salo. This ensures that VMware clusters and hosts remain fully functional even if the network connectivity between the sites is disrupted.

2.5.3 vCloud Networking and Security, vCloud Director

Both vCloud Networking and Security (vCNS) and vCloud Director (vCD) provide the necessary functionality for a safe public cloud virtual machine environment. vCNS provides the necessary functionality to provision and manage vShield firewalls and vCD provides the user interface accessible by the clients. Both of these services have been deployed redundantly on both sites. vCNS only serves the vShield firewalls on its location, and is connected only the vCenter server on the same location. vCD is fully redundant with a shared database, and either instance serves users on either location.

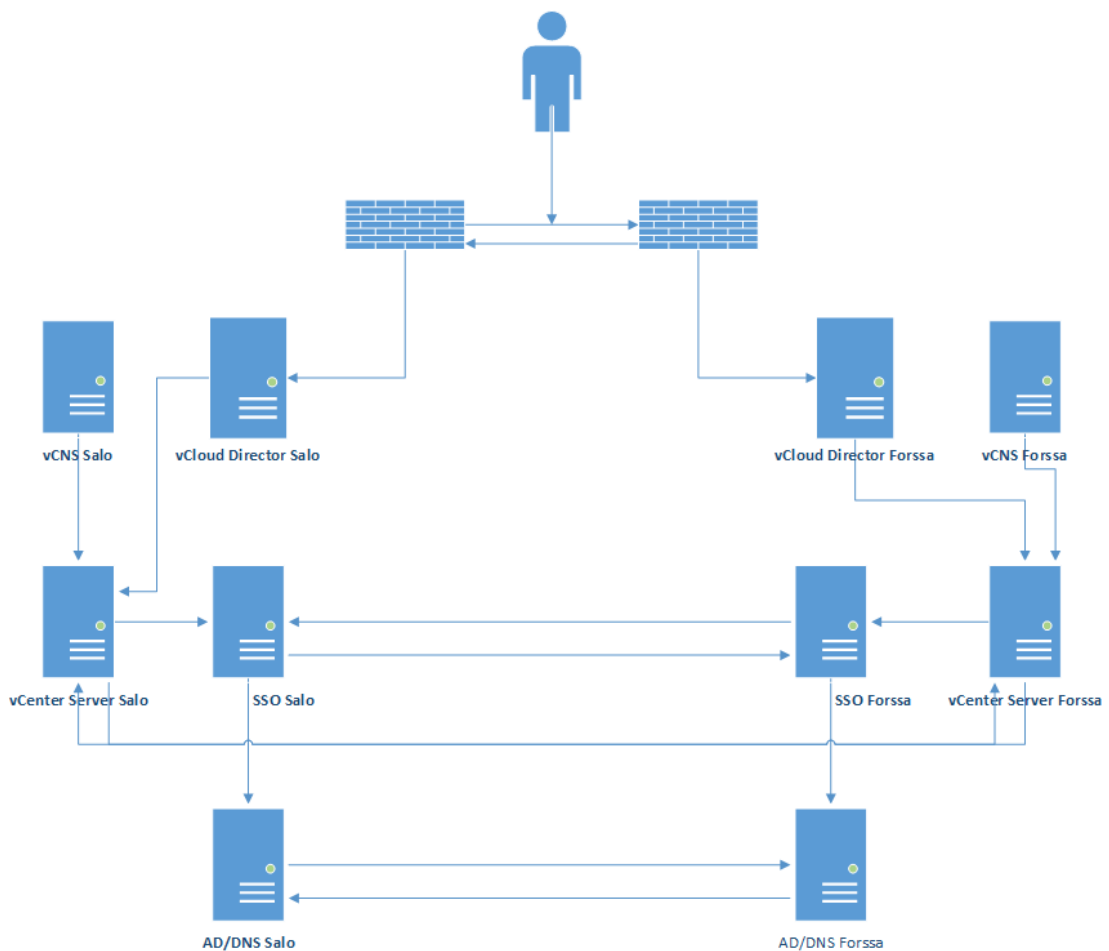


Figure 5. vCloud Director service relations.

3 DEVELOPMENT ENVIRONMENT

The production environment is large and has multiple physical locations and hundreds of virtual machines running. The development environment should focus on the key concepts and replicate them with limited resources. As some services may be consolidated into the same server to save resources, others can not be so because the core concept would differ too much. This chapter introduces the key components described earlier and discusses the design choices made regarding the development infrastructure.

3.1 Hardware

The hardware chosen for the development environment was an old HP P7000 blade chassis. The chassis itself supports a maximum of 12 blade servers, 9 of which were occupied. The blade chassis has two physical switches for LAN traffic as well as two separate Fiber Channel switches for storage embedded for each half-height blade server. These switches are connected to two external Fiber Channel switches and two 24-port 1000BASE-T LAN switches. Lan switches are slower than what can be found in the production environment, but are sufficient for a smaller test environment. The physical LAN topology will be similar to the physical environment, and any redundancy tests may be conducted at the development environment with valid results comparable for the production environment. The main difference in the network infrastructure is the combination of management and client LAN traffic to the same switches to reduce costs and complexity.

3.1.1 Blade Servers

There is a total of 9 blade servers in the blade chassis chosen for this project. The blade servers are out of warranty, and thus are not suitable for production use according to the company policy. The company policy dictates that any critical production hardware must have valid warranty to minimize production downtime.

The blade servers are still operational and thus can be used for any non-critical use, like this development environment.

Table 1. Blade Server resource configuration

Blade Server	Specifications	Notes
Proliant BL460c G1	2 Intel® Xeon® E5440 48 GB DDR2	Management
Proliant BL460c G1	2 Intel® Xeon® E5440 48 GB DDR2	Management
Proliant BL460c G1	2 Intel® Xeon® E5440 48 GB DDR2	Management
Proliant BL460c G1	2 Intel® Xeon® E5430 34 GB DDR2 Memory	Resource1
Proliant BL460c G1	2 Intel® Xeon® E5430 34 GB DDR2 Memory	Resource1
Proliant BL460c G1	2 Intel® Xeon® E5450 34 GB DDR2 Memory	Resource1
Proliant BL460c G1	1 Intel® Xeon® E5430 24 GB DDR2 Memory	Resource2
Proliant BL460c G1	1 Intel® Xeon® E5345 24 GB DDR2 Memory	Resource2
Proliant BL460c G1	1 Intel® Xeon® E5345 24 GB DDR2 Memory	Resource2
Proliant BL460c G1	1 Intel® Xeon® E5345 24 GB DDR2 Memory	Resource2
Proliant BL460c G1	1 Intel® Xeon® E5430 24 GB DDR2 Memory	Resource2

The blade servers have different computing capacities and resources are more scarce compared to the production environment. The limited hardware capabilities have to be taken into account while creating the VMware infrastructure. Specific design changes are described in greater detail in Section 3.2.

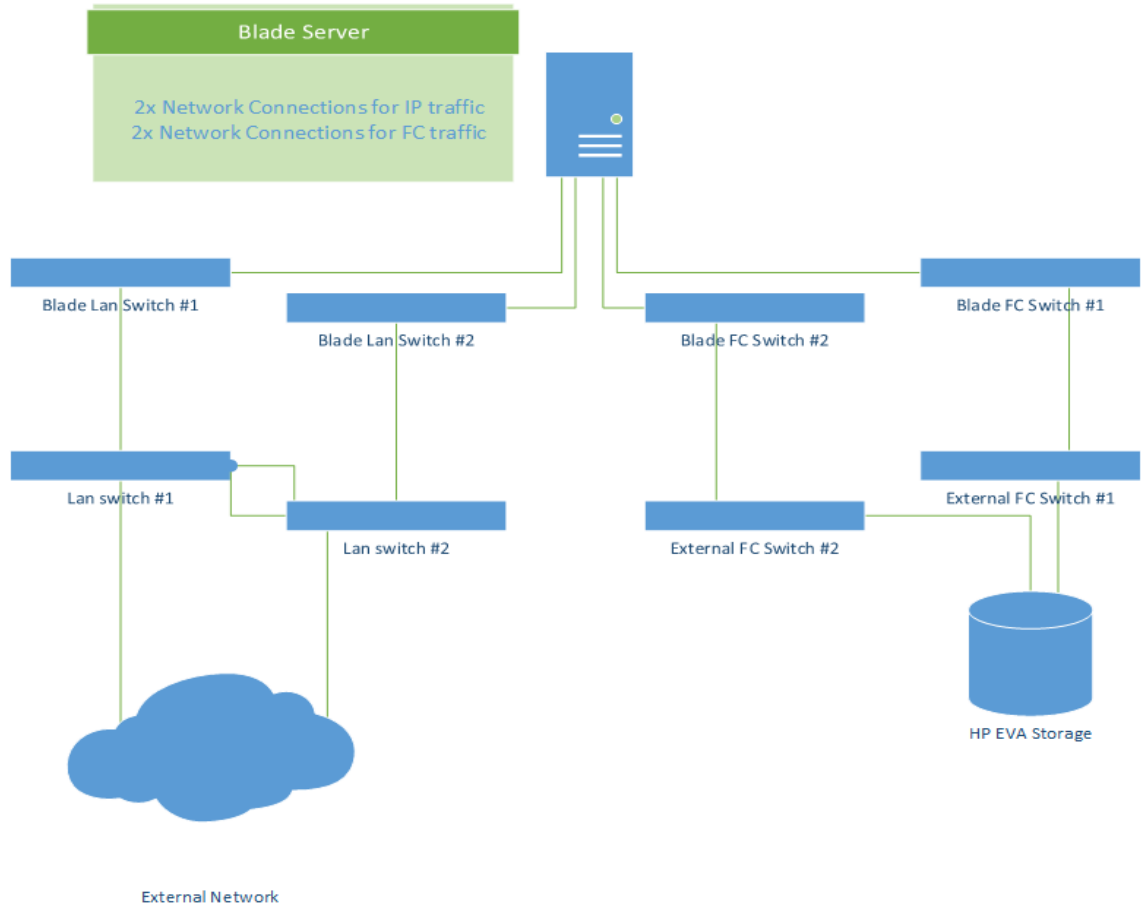


Figure 6. Blade server physical configuration.

3.1.2 Storage

For central storage, a HP EVA HSV200 SAN was selected which support contract had expired. The SAN was still in fully functional order but due high costs for the support contact, it was no longer viable as a production system. Compared to the production infrastructure, iSCSI would not be used as the available hardware only supports Fiber Channel connectivity. This caused a minor difference to how the system was initially configured, but any virtual machines operating on the VMware platform would not notice the difference.

3.2 Design Considerations

As mentioned before, there are a few differences in the hardware capabilities. One goal of virtualization is to abstract the physical hardware from the guest operating system, and thus these hardware changes have no effect on the virtual machines themselves. None the less, the hardware available does mean that some adjustments need to be made to the design of the development environment.

The production environment consists of 3 vCenter server installations; management, Salo and Forssa. As we do not have two physical locations to work with in the development environment, we will call these three locations management, Resource1 and Resource2. Although they are physically in the same rack, we can still logically separate the computing capacity in a similar manner using three vCenter servers as in the production environment.

Table 1 describes the available blade servers, and the notes column contains the vCenter server grouping of the computing capacity into clusters with similar naming. Note that the management cluster consists of 5 blade servers with more computing capacity than the resource clusters. This is to allow for more usable memory for the VMware infrastructure while leaving some capacity for other testing using the two additional clusters.

Although there was more capacity available to the VMware environment, additional changes were needed. The databases could not be separated into three different Microsoft SQL server virtual machines, but rather one VM was to be created to consolidate CPU and ram usage.

Unfortunately similar consolidation was not possible for VMware SSO, vCenter, vShield Manager nor vCloud Director, as any changes to these services would have ended up changing the topology too much compared to the production environment.

The vCenter server does come in two variants, the vCenter Server appliance (VCSA) which services run on a SUSE Linux VM and the vCenter Server binaries that are to be installed on a Microsoft Windows Server. Although it would technically be possible to utilize the VCSA and a Microsoft SQL server, Windows Server was chosen for vCenter Server to keep the installation consistent. This also allows to test any upcoming migration path from a Windows Server installation with Microsoft SQL to VCSA with built-in PostgreSQL server.

An identical IPv4 subnet was chosen for the development network but a different Microsoft Active Directory domain. This makes it possible to use the same or similar IPv4 addresses on the servers and VMs to keep the arrangement simple, but gives an easy visual identification with a different domain. Using a different domain does not prevent administrative errors due to being connected to the wrong environment, but does make it less likely. Using the same IPv4 subnet is enabled with VLAN separation and not routing configuration.

3.2 Implementation

Building the development environment was divided into several stages. Firstly, the physical hardware needed to be mounted to the appropriate rack and cabling was required. Once the physical hardware was powered on, network connectivity was ensured by properly configuring the switches, blade servers, and storage.

Once the underlying hardware was configured appropriately, and tested, the ESXi hypervisors were installed. This was the first step towards a fully functional virtual environment and allows the creation of virtual machines that host additional services.

Authentication and domain name resolution (DNS) are mandatory for a functional cloud infrastructure and were installed once the hypervisors were ready. After this it was possible to install VMware services in the sequence defined by VMware installation and configuration guides.

The following sections contain a detailed description of each step towards the goal of a working public cloud infrastructure.

3.2.1 Hardware installation and configuration

One free full height rack was selected for the hardware as the blade server chassis, storage controllers and storage shelf's required more than 20 units of rack space. In addition, space was required for the two LAN switches, the two Fiber Channel switches and one physical rack server that is used to manage the fiber channel storage using HP EVA management software.

The hardware mentioned was rack-mounted using the appropriate rack mounting rails to secure all equipment in place. Multimode fiber cables were connected between the blade chassis, SAN and fiber channel switches. HP branded multimode 800 nm wavelength lasers were used to connect the equipment to the fiber channel switches according to manufacturer specifications.

The blade server chassis, SAN controllers and SAN shelves have each redundant power supply units (PSU) built in. As such, half of the PSUs were connected to a UPS-backed power distribution unit (PDU) while the other half were connected to a PDU not backed by UPS. This ensures continued operations for a short while even during power outages. The switches do not have redundant power supplies, and thus one FC switch and one LAN switch was connected to UPS-backed power while the other switches were not. As Figure 6 shows, all blade servers are connected in a redundant fashion that tolerates physical switch failure once the ESXi hypervisors have been configured appropriately.

Regular Category 5e Ethernet cables were used for the LAN network connectivity, as there were no special requirements.

Before any configuration, all hardware was reset to their factory defaults. This ensured that no previous configuration would interfere with the desired functionality. The blade chassis, blade server remote control and switches were given a management IP address to allow remote configuration. The individual blade remote control IPs (HP Integrated Lights-Out, ILO) are also required for virtual image mounting to allow ESXi hypervisor installation as no physical cd/dvd-drive exists on the blade servers or the blade chassis itself.

The HP Brocade Fibre Channel switches were not configured to use zoning, which would be a recommended best practice in a production environment (Brocade 2014, p.3). This simplifies management and does not provide any required functionality for the development environment. However this does mean that any server connected to the fibre channel storage network can potentially access any traffic passed in the system and would, therefore, not be a suitable solution in a production environment. Logical unit numbering (LUN) masking is used to target specific storage devices to specific hosts, which, for example, prevents blade servers from booting ESXi of a different host.

LAN Switches were configured with the appropriate VLANs by providing untagged ports to the blade chassis. This configuration allows for a more flexible configuration on the hypervisor side as different virtual machines can share the same physical network connectivity by tagging their IP traffic with the appropriate VLAN tag. Untagged ports were used to allow network connectivity to the physical hardware that did not support VLAN tagging.

The HP EVA configuration was also implemented prior to any hypervisor installation, as the blade servers themselves did not include storage medium to allow hypervisor installation locally. Therefore, each blade server was assigned a small 20 GB LUN for ESXi hypervisor installation. Each blade server was then configured to boot from SAN using their designated LUN. Shared storage was also configured for each cluster for virtual machine installation. Shared storage is a requirement for virtual machine failover in the case of a physical blade server failure. If a virtual machine were to be installed on server local storage, the other servers in the cluster would not be able to run the VM during fail-over.

3.2.2 Hypervisor installation and configuration

VMware vSphere ESXi hypervisor installation provides the core functionality which is used by other VMware components to build upon. Default installation of ESXi is a simple matter and the installation was carried out using the Virtual Disk functionality of the Blade server's remote management. The Hypervisor operating system was installed individually on each and every blade server that was selected for this purpose.

ESXi hosts were named according to their cluster grouping, with a number-sequence starting from 1 identifying the hosts from each other. Each host was given a unique IPv4 management address from the subnet used.

Each host has a total of two physical network interfaces for LAN network traffic which were assigned to one virtual switch (vSwitch). This vSwitch utilizes both physical network interfaces in an active-active manner which provides both redundancy and higher throughput when multiple virtual machines are assigned to the same vSwitch. A single virtual machine is still limited by the physical capacity of a single link as no link aggregation is configured.

The hosts in the management cluster are not configured to utilize distributed virtual switches, while the hosts in resource cluster are. Virtual distributed switches (vDS) create a single place to manage network port groups over multiple hosts and is a requirement for vCloud Director. vDS configuration and purpose is described in greater detail later on. vDS switches are only available in Enterprise Plus licensing of VMware vSphere.

All hosts were also configured to use round-robin as their storage protocol when communicating with the HP EVA storage. This is the recommended best practice in an active-active storage configuration, which was selected in this particular setup. Round-robin utilized all active storage paths allowing Input/output (I/O) operations to be queued on all available controller ports in the storage array. If a path becomes unavailable, the other active paths are used to provide a disruption free operation in case of a failover event.

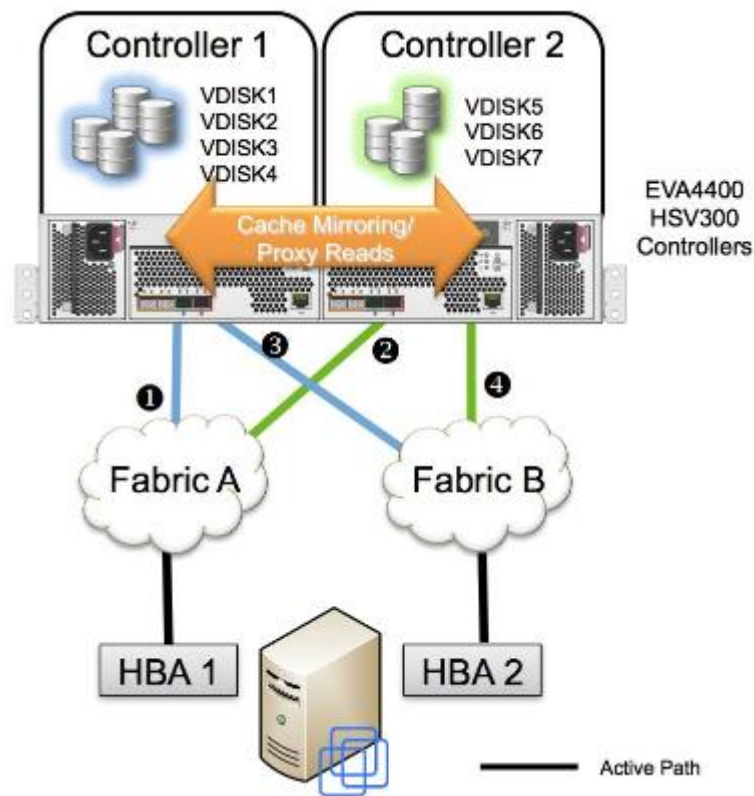


Figure 7. HP Fibre Channel SAN Configuration Diagram (Hewlett-Packard Development Company, L.P. 2014, p.11).

No advanced features are available at this point like HA, FT, DRS and unified management. These features become available once additional components have been installed, namely the vCenter Server. VMware ESXi installation does provide the physical hardware to run multiple guest operating systems using virtualization. The vSphere ESXi hypervisor abstracts the physical hardware by providing virtual machines with a software equivalent and translates any request by the VM to its equivalent in the physical hardware. Without virtualization, only a total of 11 operating systems would be able to run on this particular hardware, compared to 84 running VMs at the time of writing this thesis.

3.2.3 Core services

For any IT environment to function, some core services are required. Some of these services exist to make the work of an IT administrator simpler, while others are required by the software itself to function in a desired way. This section will describe these core services required by a functional VMware public cloud infrastructure.

3.2.3.1 Authentication

Identity management and authentication is required by all the components to be installed. Microsoft Active Directory (AD) provides identification for both user and computer accounts, and work for both VMware services once configured and Microsoft Windows Servers when joined in an AD domain.

3.2.3.2 Domain Name Services

Microsoft Active Directory requires DNS to function properly, and is installed in conjunction with AD. The DNS service role is also usable with other services for fully qualified domain name (FQDN) lookups. VMware recommends to configure any VMware component using its FQDN rather than using IPv4 addresses. If IPv6 were to be used, then DNS would be a mandatory requirement.

3.2.3.3 Databases

VMware vCenter Server requires either an IBM DB2 or Microsoft SQL server for its database, when not using the appliance deployment method. As the production environment uses Microsoft SQL server as its database backing, the same choice was made for the development environment to keep configuration consistent. As mentioned in design considerations, only one database server is to be installed. This saves on vCPU, vRAM resources as all vCenter server databases are located on the same server. This solution does not provide as good performance as separate SQL server installations, nor would it be practical in multiple physical datacenter locations. When considering the development environment, a shared database server meets the requirements.

3.2.4 VMware vCenter

The VMware vCenter Server management solution provides advanced features for a VMware vSphere ESXi environment through different components explained in this section. The vCenter allows the use of high availability, fault tolerance, vMotion to name a few.

A total of three vCenter servers were installed, to create a similar topology as with the production environment. These three vCenter servers were then linked together using linked mode to gain a single plane of view for the systems administrator.

A vCenter server custom installation needs to be installed in the following manner:

1. vCenter Single Sign-On
2. vSphere Web Client
3. vCenter Inventory Service
4. vCenter Server

VMware-vCenter-Server-5.5-Technical-Whitepaper

3.2.4.1 Single Sign-On

A vCenter SSO provides a secure authentication service for many VMware components. SSO will be tied together with Microsoft AD to provide user authentication from an external source, making it possible to use the same username and password combination on several different services. SSO uses tomcat to provide security tokens for any successful authentication that works on all services which are integrated to the SSO domain.

SSO may also be installed as a multi-site configuration, which makes it possible to install multiple SSO servers that can, if necessary, act independently. As the original design requires two physical locations to function independently in case of a network error between the sites, multisite deployment has been chosen. The development environment will also need two SSO servers, likely configured in a multi-site configuration. Although this does not provide a geographical advantage, this will provide a crucial design choice to be tested in case of a system wide upgrade that requires vCenter Server to be upgraded to a newer version. Both SSO servers have their own VM with Windows Server installed.

The development environment uses the following names for the SSO servers and instances:

- SSO-RES1
- SSO-RES2

3.2.4.2 vSphere Web client and vCenter Inventory Service

Each vCenter server instance requires its own vSphere Web Client and vCenter inventory service. Even though it is possible to install separately, the two components was installed on the same server as the vCenter Server. This saves resources and provides the capabilities required. A total of three servers was installed, each for its own vCenter with Web Client and Inventory service.

vSphere Web Client provides a way to manage the vCenter Infrastructure using the web browser, removing the dependency for VI Client which was only available for Windows client computers.

The vCenter Inventory Service provides a file-level (xDB format) cache that speeds up common queries and basic administrative tasks. In addition, since version 5.1, the Inventory service provides object tagging which is used for example with storage policies. Storage policies can be used to create different tier storage in vCloud Director which abstracts the storage layer from the end user. The vCenter Server then makes the appropriate storage vMotion based on these object-tags and the user selection.

Management cluster Inventory service and Resource1 Inventory services are registered to the SSO-RES1 single sign-on instance while Resource2 is registered to the SSO-RES2 single sign-on server.

3.2.4.3 vCenter Server installation and configuration

The vCenter Server installation is carried out separately in a custom installation like this, and as such each vCenter Server instance is configured separately. The shared database server contains a separate database for each vCenter Server, and the ODBC data source is configured on each server prior to vCenter installation. ODBC defines the database connection configuration to an external Microsoft SQL database. Each vCenter server is registered in an identical manner as their Inventory services to their respective SSO servers.

Once the vCenter servers were installed, a linked-mode bond was created to allow unified administrative view of the environment.

A cluster is added to each vCenter server so that the advanced features may be enabled on the ESXi servers installed before. Although all Windows Server installations have been placed so far on the hosts that have been designated for the management cluster, they yet lack the functionality to operate in consensus.

Firstly a Datacenter is created on each vCenter server, and a cluster is created within the datacenter. Appropriate vSphere ESXi hosts are added to the clusters and HA+DRS is enabled. Only when these operations are completed, the hosts may operate to provide fail-over capabilities to the virtual machines running on them. When a HA cluster is created, a master is automatically elected between the hosts joined to the cluster. The hosts as well as the virtual machines running on the cluster are monitored, and in case of a host failure, virtual machines are automatically restarted on an available host on the cluster. The master server monitors the other hypervisor servers in the cluster mainly using network and datastore heartbeat and reports to its vCenter server.

The cluster has to have sufficient resources available for the virtual machines, and thus capacity needs to be calculated to allow failover. In a case when all hosts in the cluster are identical, the capacity can easily be calculated using the following formula: $(\text{ESXi hosts}-1) \times \text{capacity}$ (VMware, Inc. 2013. vSphere Availability ESXi 5.5 vCenter Server 5.5., p19-25).

3.2.5 vShield Network and Security and vCloud Director

The last parts of a fully functional public cloud infrastructure are vCloud Networking and Security (vCNS) and vCloud Director (vCD). Both provide services to the end customer which allows customer administrators to make changes to their own environment without service provider interference.

vCNS provides networking and security aspects to a public cloud infrastructure with the implementation of routing and firewall services. vCNS is controlled by one VM containing required components for a single vCenter server installation. vCNS functions as central management for customer firewall VMs between vCD and the VMs running. Any configuration changes made to the firewall are stored in the vCNS installation, and thus the firewalls can be re-deployed from a firewall template (Virtual Machine template) and vCNS migrates the configuration to the new VM.

vCNS allows cloud customers to manage their own private virtual network with an integrated solution from a single front-end interface. Network Address Translation (NAT), Routing, Firewall Services, Virtual Private Network (VPN) and Load-Balancing are a few of the services that can be configured and managed using vCNS.

Unfortunately vCNS has been announced to be End of Life by VMware, and thus VMware will not release any further updates to the product. VMware NSX is set to replace vCNS in a public cloud infrastructure, but as the production environment still utilizes vCNS, it needed to be installed in the development environment. Two instances were installed matching the resource cluster configuration made earlier. The management cluster does not contain a vCNS instance as vCloud Director will not be used to manage it. The virtual machines, on the other hand, were placed in the management cluster with vCenter Servers and other internal services.

vCloud Director provides the administrator with an easy-to-use interface to provision and manage virtual machines. Each customer is separated to its own organization, with its own users, templates, networks and virtual machines. Although all VMs share a common infrastructure on the underlying virtualization infrastructure, they are logically separated with vCloud Director in a secure manner.

A total of two vCloud Director server virtual machines were deployed to provide cross site availability. This deployment matches the production environment and allows for upgrades to be completed with minimal downtime. Provisioning two CentOS 6 virtual machines takes a minimal amount of resources and thus was not a factor to consider. A single vCD node would have been sufficient as both nodes share a single database, but any upgrades using two vCD cells would not be possible.

A single organization may have one or more virtual datacenters (vDC), which create a logical equivalent in vCenter Server in the form of a Folder and Resource Pool. Virtual datacenters can be provisioned in three different ways.

- 1) Pay-As-You-Go; which means that the resource pool does not have any limitations to the resources it may use.
- 2) Allocation Pool; where the resources pool has a fixed amount of resources available using the resource pool limit values. Virtual Machines can not start if the resource pool does not have enough resources.
- 3) Reservation Pool; where the resources defined for the virtual datacenter are reserved within vCenter Server resource pool and are not shared with other virtual machines even if they are not used.

Any virtual machine created in a particular virtual datacenter, will be placed in its own resource pool and folder within vCenter Server. All VMs created in virtual datacenters provisioned in the same provider virtual datacenter will share the same cluster of ESXi hosts, even though the end user will not see any of the other VMs sharing the capacity. Any networks created in a vDC will also have their equivalent in vCenter Servers distributed virtual switches. A port group within the distributed switch can only be used for the VMs connected to it, and the network is only visible to a particular virtual datacenter. This prevents any other organization administrator from sharing the network with another customer, and thus compromising the security of virtual machines and/or its network traffic.

The Development environment has two provider vDCs matching the two vCenter servers created for resource clusters corresponding the sites Salo and Forssa. As no production customers will ever use the system, a few organizations and virtual datacenters were created to for testing purposes. These organizations have no business critical virtual machines running within them, but they simulate any customer VMs by being powered on. If any changes to the infrastructure cause the VM's to power off, reset or lose network connectivity, then these changes can be taken into account when making similar changes to the production environment. Any actions that disrupt client virtual machine operation can be scheduled after business hours and proper notification can be made in advance.

3.3 Real world usage

The development environment has been used extensively for the past year in miscellaneous cases. During this time, the environment has been used to do a complete major version upgrade, and a second major version upgrade is being planned. Several minor version upgrades have been completed and several VMware PowerCLI scripts have been developed. These scripts include automatic consolidation of virtual machines, virtual machine statistics gathering and automatic provisioning of networks, organizations and vShield Edge firewalls.

The development environment also serves as a company internal testing platform for several internal software, such as monitoring solutions and backup.

4 CONCLUSION AND FUTURE WORK

The development environment was designed to simulate the production environment as closely as possible. Some changes to the scope of the environment were made to save resources as older hardware was at our disposal. In the end, these changes were minimal and a fully functional environment was created successfully.

Changes to name services, IP-Addressing had no impact on the functionality of the services installed. Although the hardware chosen differentiated in age, technology and capacity, this had no effect on how the environment operated. Changes to configuration and drivers were required, but these changes are invisible to the functionality of the cloud end user.

As the development environment is installed and configured as closely as possible compared to the production environment, any changes will with reasonable certainty have the same effect when done again in the production environment as they did when made in the development environment. The main difference comes from the update history, as the production environment was created using ESXi 5.0 and its equivalent software while the development environment was installed using ESXi 5.1 and its equivalent. This difference had no impact during installation and configuration of the software mentioned in this thesis but during the upgrade from versions 5.1 to 5.5, there were problems in the upgrade of the production system when the upgrade on the development environment went mostly flawlessly. These problems were related to Single-Sign-On which was introduced with vCenter 5.1 and the production environment was installed prior to this. The upgrade from 5.0 to 5.1 caused differentiating starting conditions while upgrading to 5.5 and thus these errors were not encountered during the upgrade in the development environment.

As stated earlier, vCNS has been declared end-of-life by VMware and no further product upgrade is available. VMware did release version 6.0 of vCenter server and ESXi, but no upgrade for vCNS was made available. VMware recommends to migrate to VMware NSX, and this may be tested with the development environment to create a plan for migration in the production environment. Any side-by-side migration from an existing product to a new one is typically time-consuming and prone to faults. Therefore, thoroughly testing a migration like this is extremely valuable.

While large-scale testing yields the best results from a development environment, even small upgrades may cause problems. Minor version upgrades may from time to time cause issues due to changes in configuration, product interoperability or other issues and, therefore, testing even the smallest changes is recommended.

The planning and creation of this development environment was completed under a time span of 4 months in 2014 and has been actively used since then. The availability of this resource has been extremely useful in our day-to-day work with the cloud infrastructure and has saved us time troubleshooting while having possible downtime in a business critical infrastructure.

REFERENCES

- Brocade. 2014. Secure SAN Zoning Best Practices. [online] Available at: <http://www.brocade.com/downloads/documents/white_papers/Zoning_Best_Practices_WP-00.pdf> [Accessed 11 May 2014]
- Hewlett-Packard Development Company, L.P. 2014. Hp Enterprise virtual array family with VMware vSphere 4.0 , 4.1 AND 5.0 configuration best practices. [online] Available at: <<https://www.vmware.com/resources/techresources/10234>> [Accessed 11 May 2014] doi: 4AA1-2185ENW
- King, J. 2013. VMware® vCenter Server™ 5.5 Deployment Guide. [online] Available at: <<http://www.vmware.com/vmtn/resources/>> [Accessed 12 May 2014] doi: VMW-TWP-vCTR-SVR-5.5-DEPLY-GUIDE-USLET-101
- VMware, Inc. 2009. VMware® vNetwork Distributed Switch: Migration and Configuration. [online] Available at: <<http://www.vmware.com/resources/techresources/10050>> [Accessed 21 May 2014] doi: VMW_09Q2_vSphere_WP_vDSMigration
- VMware, Inc. 2010. VMware vCloud® Architecture Toolkit Architecting a VMware vCloud. [online] Available at: <<http://www.vmware.com/cloud-computing/cloud-architecture/vcat-toolkit>> [Accessed 12 May 2014] doi: VMW_11Q1_WP_ArchitectingvCloud_p100_R2
- VMware, Inc. 2012. VMware vSphere® Distributed Switch Best Practices. [online] Available at: <<http://www.vmware.com/vmtn/resources/>> [Accessed 21 May 2014] doi: VMW-vSPHR-DIST-SWITCH-PRCTICES-USLET-101
- VMware, Inc. 2013a. vSphere Availability ESXi 5.5 vCenter Server 5.5. [online] Available at: <<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>> [Accessed 21 May 2014] doi: EN-001254-00
- VMware, Inc. 2013b. vSphere Networking vSphere 5.5 ESXi 5.5 vCenter Server 5.5. [online] Available at: <<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>> [Accessed 21 May 2014] doi: EN-001074-02
- VMware, Inc. 2013c. vSphere Security ESXi 5.5 vCenter Server 5.5. [online] Available at: <<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>> [Accessed 21 May 2014] doi: EN-001164-04
- VMware, Inc. 2013d. vSphere Storage ESXi 5.5 vCenter Server 5.5. [online] Available at: <<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>> [Accessed 21 May 2014] doi: EN-001260-06
- VMware, Inc. 2013e. vSphere Virtual Machine Administration. [online] Available at: <<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>> [Accessed 21 May 2014] doi: EN-001518-01