Virtualisation of Critical Network Services

Best Practice Document

Produced by the CESNET-led Working Group on Network Monitoring
(CBPD 120)

Author[s]: Pavel Kislinger kislinger@cis.vutbr.cz,
Vladimír Záhořík zahorik@cis.vutbr.cz

November 2012

Document No:	GN3-NA3-T4-CBPD120
Version / date:	July 2012
Original language :	English
Original title:	“Critical network services virtualisation”
Original version / date:	Version 1.0; July 2012
Contact:	kislinger@cis.vutbr.cz, zahorik@cis.vutbr.cz

CESNET is responsible for the contents of this document. The document was developed by a CESNET-led working group on network monitoring as part of a joint-venture project within the higher-education sector in the Czech Republic.

The research leading to the update and revision of this report has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 238875, relating to the project 'Multi-Gigabit European Research and Education Network and Associated Services (GN3)'.

Table of Contents

Executive Summary

This document describes a way to virtualise the number of network servers that are required for the operation of a large campus network. These servers provide services, including DHCP, DNS, VPN, email, network monitoring, and radius. Most of these services are so important that the network must operate two or more of them at the same time, and this leads to an increase in the number of servers. Usually, these services do not require a great deal of computing power, indicating an excellent opportunity to use virtualisation The document is focused on the different requirements to be considered when choosing the appropriate hardware for the job, with emphasis on the price/performance ratio, while maintaining all the benefits of the Vmware vSphere system, which was selected as the virtualisation platform. The document also describes practical experience and the pitfalls that may be encountered during the installation of the system. It describes the configuration of network devices, the iSCSI storage, and the VMware vSphere hypervisors. The conclusion summarises the results and explains the benefits of virtualisation for the campus network.

The first part of this document describes the advantages and disadvantages of virtualisation for given types of services and explains the purpose of building a virtualisation cluster in two geographically distant locations. The second part is devoted to a specific configuration of network devices and to the preparation that needs to be done before connecting individual parts of the virtualisation cluster. It describes actual experience gained in the operation of these clusters in a situation where it was necessary to revise the network topology. The next section of the document explains how to select the appropriate hardware, especially in the choice of storage and hypervisor hardware. It describes the required properties, with an emphasis on diversity, and in particular, how the virtualisation cluster differs from the usual VMware vSphere cluster. This is followed by a section devoted to the practical configuration of a VMware vSphere and a vCenter. The final section presents the results of the measurement of consumption before and after the virtualisation of the critical servers as well as other operating statistics.

1 Virtualisation of Critical Network Services

The term, virtualisation, is an IT buzzword, referring to technologies that create an abstraction layer between computer hardware and software. This layer creates a transparent, logical structure that masks the physical (real) one. The goals of virtualisation are the simplification of maintenance, easier scalability, higher accessibility, better utilisation of hardware, and improved security. The memory, the processors, the computer network, the storage, the data, or the whole computer system can be virtualised. The virtualisation of a computer enables one physical server to run multiple operating systems. This is called server virtualisation and can be accomplished in different ways. The most frequently used method of server virtualisation is hardware-assisted virtualisation, which requires a special instruction to be set in the CPU (Intel VT-x and AMD-V), but offers the best performance. This method of virtualisation is the focus of this document.

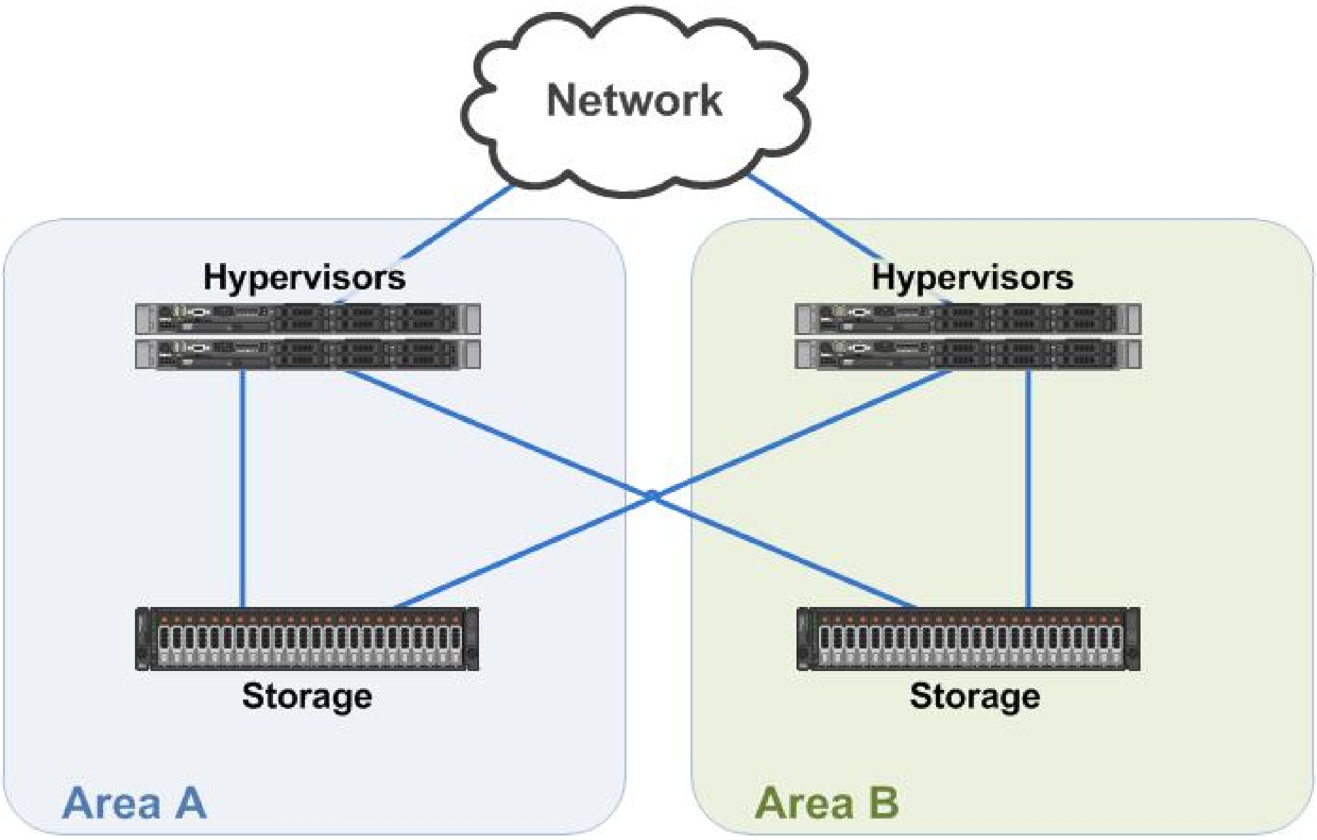
1.1 Resilient virtualisation cluster

To be able to choose the right platform, it is necessary to know and understand the basics of server virtualisation. Everyone who starts with server virtualisation typically installs virtualisation software on a server with large amounts of memory and high capacity drives for all of the virtual servers. This system works well in a test environment. In a production environment, is also necessary to provide protection against various types of accidents that may occur during operation. Current servers have two power supplies, hot swap disk drives in RAID array, multiple CPUs, and many memory modules. However, consideration has to be given to the chance that the server motherboard, the RAID controller, or the power can fail. It is necessary to anticipate infrastructure failures caused by network problems, failure of the cooling system in the server room, or revision of the wiring. All of these cases of failure will result in the unavailability of all of the virtual servers. To counter these potential threats, is necessary to extend the virtualisation system by adding several elements. First, it is necessary to increase the number of hypervisors.

This, in itself, does not contribute substantially, because live migrations of the virtual systems cannot be achieved on two servers alone. Migration of these systems is only possible when all hypervisors have access to shared storage. This storage then becomes a bottleneck, because a failure of the device causes unavailability of the system. Therefore, storage is designed with this in mind: enterprise storage and hard disks. These types of storage have two independent RAID controllers. Each enterprise hard disk has two storage interfaces to connect it to both of the RAID controllers, and this provides resilience in the event of the failure of one of the controllers. The security threats for this type of device are storage power failure, air conditioning failure, or some other disaster. Maintenance of this device is also very complicated, because all actions are performed during run time.

The solution to this problem is to have two storage devices in two, geographically distant locations. All hypervisors have access to both of these storage locations. This makes it possible to move all virtual servers to the first device while maintenance is performed on the second. This ability to move virtual systems to another location is important when there is a planned, structural modification of the wiring, and also in the event of a disaster. A potential bottleneck of this system is an unexpected failure of storage that contains production data. Although this scenario is unlikely, due to the features of enterprise storage, it must be considered. These cases can be solved by restoring the data from a backup server or by using storage facilities with cross-data replication. Unfortunately,these functions are only supported on the top brands of storage hardware, where cost and complexity are much higher than for middle-class storage hardware.

This paper describes a virtualisation cluster, based on two middle-class arrays (without the support of cross-replication) and several hypervisors.



1.2 Selection of the virtualisation platform

There are many virtualisation solutions. Each of them has its advantages/disadvantages and developers always claim that their product is the best. Selection of the best may not be completely obvious. Among other requirements, a virtualisation system must permit the installation of any operating system, live migration of virtual systems between hypervisors, and movement of live virtual systems from one storage subsystem to another (live migration of a Virtual Machine from one storage location to another without downtime). Another important feature is the ability to ensure uninterrupted operation of the virtual systems when the hypervisor fails. The following table compares the characteristics of the best-known virtualisation solutions.

	VMware vSphere	VMware ESXi Free	Microsoft Hyper-V	KVM	XEN	OpenVZ
VM Windows	Yes	Yes	Yes	Yes	Yes	No
VM Linux	Yes	Yes	Partially	Yes	Yes	Partially**
VM Unix	Yes	Yes	No	Yes	Yes	Partially**
VM Migration	Yes	No	Yes	Yes	Yes	Yes
VM Storage M.	Yes*	No	Downtime***	No	No	No

* Storage vMotion is enabled in VMware vSphere Enterprise

** OpenVZ needs a modified kernel for VM

*** Hyper-V online VM storage motion is possible, but with downtime during motion (partially solved in Windows Server 2012)

VMware vSphere is definitely not the most powerful solution, but its flexibility, ease of implementation, and system-support from hardware vendors sets the standard for virtualisation. It provides many features that competitors do not yet offer, such as Distributed Resource Scheduler (DRS), Distributed Power Management (DPM), High Availability (HA), Fault Tolerance (FT), and Network I/O Control. Thus, it is well ahead of its competitors, who are still several steps behind this solution. This dominance entails a disadvantage in terms of price, which is several times higher than the price of other virtualisation platforms.

2 Hardware Selection

Most virtualisation clusters focus on maximum performance, memory size and IOPS. Virtual servers only use the resources related to the services running on them. The performance of most services depends on the speed of CPU (processing video data, simulation, etc.). Other services require maximum IOPS and extra memory (large database systems, web servers). There are also services that never use more than a fraction of CPU and the IOPS-value is not relevant for them because data from storage are rarely loaded. These are also the services necessary for the operation of large computer networks on the campus: DHCP, DNS, VPN, email, radius, and various monitoring tools or web services. The following section explains the selection of hardware suitable for virtualisation of services of this kind.

2.1 Optimal hypervisor hardware

As described in previous sections, the most important hardware parts of the virtualisation cluster are the hypervisors and storage devices. Campus network infrastructure requires their interconnection into one robust VMware vSphere cluster. The selection of hardware must be adapted to the characteristics of the virtualised services, and also to VMware licensing policy. It is necessary to license each physical processor unit and every few gigabytes of memory.

Details about VMware vSphere licensing are available on the website [1]. The best server is equipped with one powerful CPU and up to 64GB memory. This configuration uses all of the resources of one licence of VMware vSphere ESXi 5 Enterprise.

Today's processors are about ten times more powerful than the CPU's of five years ago. This allows a processor to replace several older servers. The best price-performance ratio is offered by the Intel Xeon 5600 or the Xeon E5-2600 family of processors.

A VMware vSphere hypervisor requires about 1GB of disk space for installation. Essentially, this storage is only used at system boot. The optimal solution for hypervisor storage is an enterprise flash memory with a capacity of at least 2GB. Other storage systems are not necessary because the data of the virtual servers are stored on a shared iSCSI storage device. Because there is no need to install additional hard drives, it is possible to fit the server hardware into a small server chassis with a height of 1U (Standard Rack Unit).

The parameters described above correspond to many of the servers from different vendors, and campus networks use servers from HP, IBM, Supermicro, Dell, and others. The best operating characteristics are found in Dell servers, which are better than their competitors in design, operating characteristics, warranty, and low failure rate. Their price for academic institutions is very advantageous and for larger purchases, and discounts of 50-60 percent from the list prices can be obtained. The best offer found was the Dell R610 server in the configuration:

- PowerEdge R610;
- Intel Xeon X5690 Processor;
- 48GB Memory for 1CPU;
- Internal SD Module with 2GB SD Card;
- High Output Redundant Power Supply (2 PSU) 717W;
- Intel X520-T2 10GbE Dual Port Server Adapter, Cu, PCIe;
- iDRAC6 Enterprise.

This is a relatively cheap and powerful solution. This server, with a single license of VMware vSphere 5.0 Enterprise with one-year subscription, cost less than €6000 at the end of 2011. It is possible to equip the server with 48GB of DDR3 memory for one CPU socket (6x8GB), and memory size is more important than CPU performance. Therefore, in the case of limited resources, it is better to use a cheaper CPU with more memory. A server, with less memory than 36 gigabytes will, most likely, be rare in the future. If a failure of one of the hypervisors occurs, it becomes necessary to fit all running virtual systems to the memory of the remaining hypervisors. Otherwise, it is impossible to guarantee their uninterrupted operation. The same restriction applies, not only during failure of a hypervisor, but also for its upgrade. For this reason, it is advantageous for the virtualisation cluster to have at least three or four hypervisors.

VMware vCenter Server is the simplest, most efficient way to manage VMware vSphere. It provides unified management of all of the hosts and VMs in the datacentre from a single console and also provides aggregate performance monitoring. A VMware vCenter Server gives administrators deep insight into the status and configuration of clusters, hosts, VMs, storage, the guest OS, and other critical components of a virtual infrastructure—all from one location [2].

VMware vCenter is a software application designed for Windows. Versions for Linux already exist, but do not yet have all the features of the original Windows version. This application is very important, because without it, it is not possible to migrate both of the virtual servers and their virtual disks. The license for this product is determined by the number of hypervisors in the cluster. If the virtualisation cluster includes three hypervisors and this number is not going to increase in the future, it can be very advantageous to choose a VMware vCenter Academic licence, which provides the full version of central management for up to three hypervisors. In other cases, it is necessary to use the VMware vCenter Standard, which can handle an unlimited number of hypervisors. A single license for a VMware vCenter 5.0 Standard with a one-year subscription cost less than €4000 at the end of 2011.

2.2 Fibre vs. iSCSI storage

The heart of each storage unit is a host bus adapter (HBA) that manages the physical disk drives and presents them as logical units. Connection to the network is also provided by the host bus adapter. Powerful storage devices have additional HBA for load-balancing between both controllers, and these can secure the full functionality of the storage system during a failure of one of them. Selection of the HBA depends on the choice of suitable technology. The dominant technologies in the enterprise storage field are Fibre Channel and iSCSI. Both technologies are supported by Vmware vSphere and offer adequate performance characteristics. The cost of HBAs for both technologies is comparable. The advantage of Fibre Channel technology is its throughput and overall performance. The disadvantage is that it is more expensive and requires a complex network infrastructure. ISCSI is cheaper because it can run on almost any switches.

Most critical services in the network do not require storage systems with extra performance. Therefore, it is not absolutely necessary to deploy Fibre Channel, although it is better in many aspects. ISCSI was chosen primarily because of its lower cost, operational characteristics, and the support from VMware. In addition to the type of HBA chosen, it is also necessary to specify other operating parameters, such as device dimensions, type, and number of power supplies, speed of iSCSI connectivity, form factors, and the capacity of the hard drive and software features.

2.3 Storage parameters

Storage can be connected to the SAN using either 1Gb or 1OGb iSCSI HBA. The difference in price between these HBAs is minimal in comparison with the price of the whole storage system. The advantage of 10Gb is in the acceleration of iSCSI traffic, and thereby, in higher read/write performance of the virtual servers. Sometimes, it may be useful to reduce the interface speed to 1Gb. This allows the connection of the storage system to the existing 1Gb SAN infrastructure, and SAN infrastructure devices can still use the same HBA after an upgrade to 10Gb.

This allows the connection of the storage to the existing 1GbE SAN infrastructure and also makes it possible to continue to use it with higher performance, following an upgrade to 10GbE SAN.

The disk array must be maximally resilient to hardware failures. For this reason, the array must be equipped, not only with two HBA, but also with several power sources. The size of the storage equipment is mostly limited by the height that the storage device would take up in the rack, which should not exceed 2U (Standard Rack Unit).

Twelve classical 3.5" discs or twenty-four 2.5" discs can be accommodated in the space of 2U. 3.5” drives offer the greatest output and capacity. The total capacity of twelve of these disks is 48TB for SATA disks, or roughly 10TB for SAS disks. The total capacity for twenty-four 2.5" SAS disks is 21.6 TB. These disks have a lower tendency to heat, making them more reliable. A greater number of disks provide a higher IOPS and more flexibility with the RAID array [3]. Modern enterprise SSD disks are also 2.5" in size. The following table compares different disk configurations that can be fitted into the 2U disk array.

		Capacity(GB)	Drives in 2U	2U Capacity(GB)	Active Power(W)	MTBF(Mh)
3.5’’	SATA/SAS (NL)	3000	12	36000	13	1.2
	SAS	600		7200	18	1.6
2.5’’	SATA/SAS (NL)	1000	24	24000	7	1.2
	SAS	900		21600	9	1.6
	SATA MLC SSD	512		12288	3	1
	SAS SLC SSD	400		9600	9	2

* values are actual at the end of 2011; [4]

Therefore, it is best to fill the disk array with a higher number of 2.5" SAS disks, mostly because they are more reliable and have a higher capacity. In some cases, it is even better to use several 2.5" SSDs. After considering all the above requirements, the Dell PowerVault MD3620i disk array was chosen as the best, with the following configuration:

- PowerVault MD 3620i;
- 2x HBA 10Gb iSCSI (2x 10GbE port and 1x 1GbE Management port per HBA);
- 24x 600GB 10K RPM SAS 6Gbps 2.5’’ Hotplug Hard Drive;
- 2x Redundant Power Supply (2 PSU) 717W (600W peak output);

3 Network Infrastructure

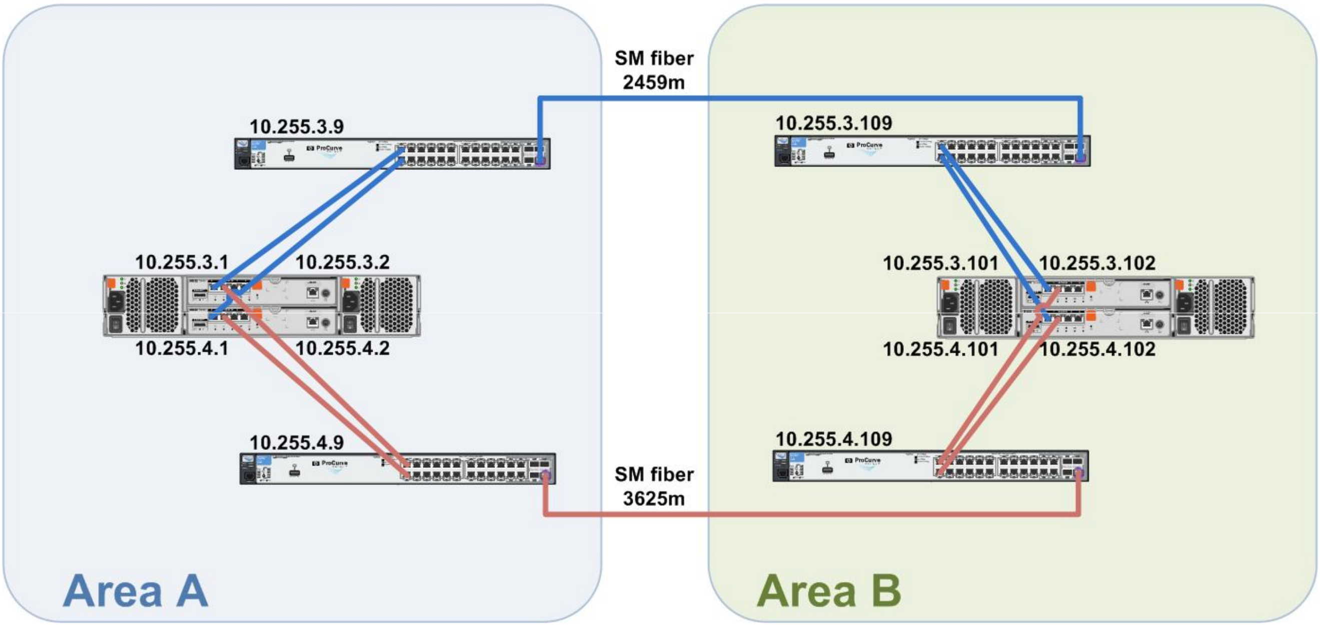
The first step in setting up a virtualisation cluster is to prepare the network infrastructure. This infrastructure provides two primary functions: connection of the hypervisors to the backbone network and to the disk storage.

3.1 Storage interconnection

Switches that connect hypervisors with the disk storage are referred to as SAN infrastructure. A SAN infrastructure could, theoretically, be shared with server access network. This option was tested over a period of several months. This testing determined that it is not the correct choice.

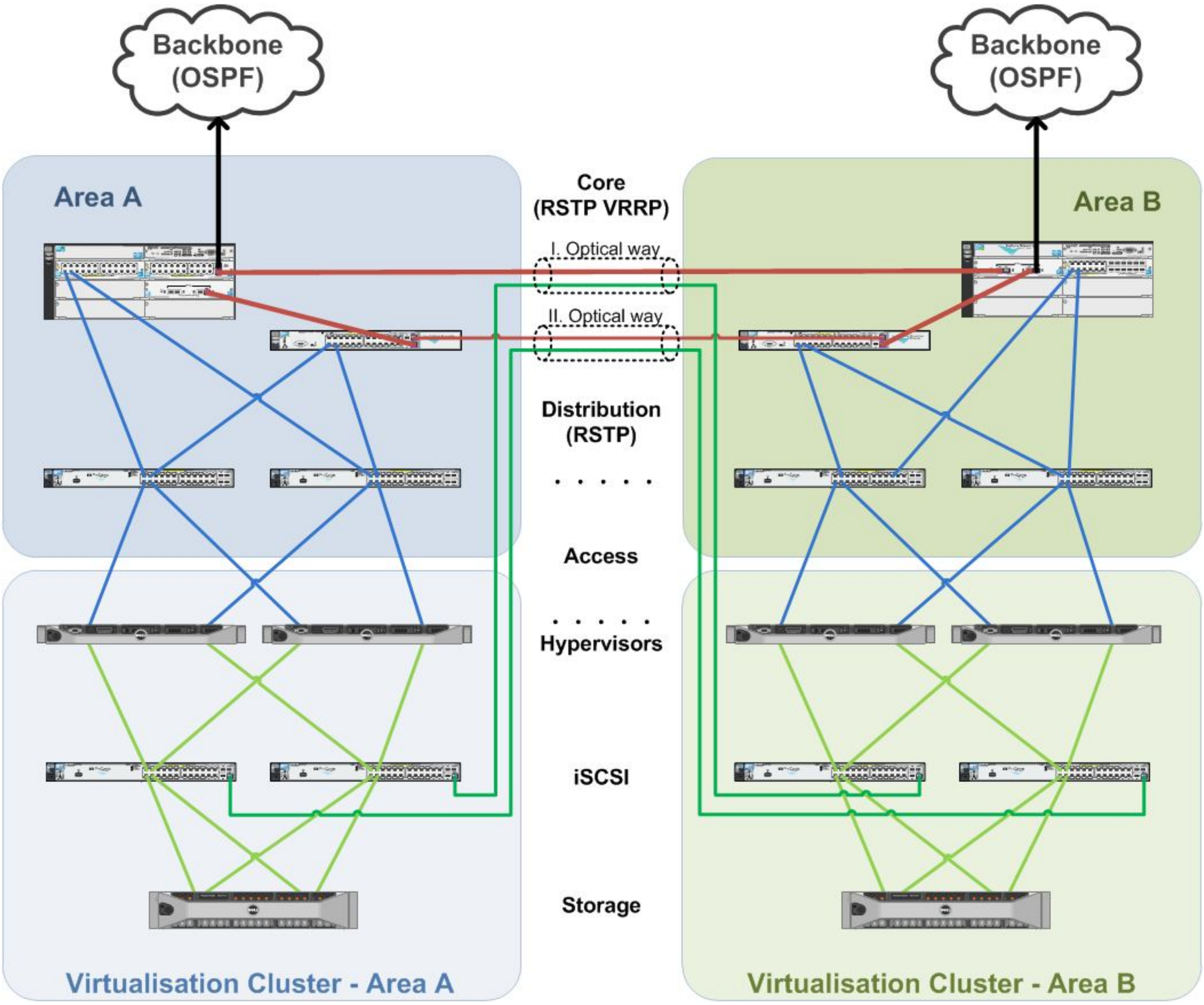
Some network technologies that are used in backbone networks to ensure uninterrupted operation can result in small interruptions in seconds (STP, OSPF, etc.) and cause their convergence time. Another source of potential problems is short-term network peaks, when a link is saturated with a lot of traffic for several seconds (possibly due to DOS attacks or broadcast storms). These problems are characterised by similar behaviour of the virtual servers. A short-term interruption or significant slowing of the SAN network can lead to serious problems with the virtual server. Problems were found on all virtual operating systems, but the most occurred on Linux servers. Their IO system is set up to avoid problems when accessing the hard drive. If the Linux kernel does not get a response from the hard drive in the predefined interval, it connects to the affected part as read-only, making the virtual server entirely dysfunctional. Other systems of connecting to the disk are not read-only, although these too were affected by interruptions. All processes requiring a disk operation at the time had to wait several seconds for the disk operation to be completed. As a result of these problems, it was necessary to set up separate network infrastructures between remote locations to allow direct connection to the SAN switches in both locations.

HP switches are the most used switches on the VUT campus in Brno, which is why the proven HP 2910-24al switch was selected for the SAN infrastructure. This is a standard L2 switch with a management system that allows the connection of four optical transceivers. This switch has sufficient throughput and supports JUMBO packets. Theoretically, the use of JUMBO packets is highly suitable, although the increase in throughput realised is only a few percentage points. A more complex configuration, without diagnostic tools, represents a disadvantage. This is why JUMBO packets are not permitted in the final SAN infrastructure. The following illustrations describe the SAN connection in detail. Two independent L2 segments, ensuring iSCSI connectivity, are highlighted in blue and red. Each of these segments uses a different range of IP addresses: 10.255.3.0/24 and 10.255.4.0/24. All of the hypervisors and the HBA are part of both subnets. This in the only way that resilience against an interruption in one branch can be guaranteed.



3.2 Backbone interconnection

In addition to the SAN infrastructure, the connectivity of hypervisors is also necessary. Each of these must be connected to two backed-up access devices. Through the backbone network, these are then connected to another pair of access devices in the remote location. Hypervisors in all locations must have access to the same VLAN, in such a way so that only one virtual server with a fixed IP address is required to operate any of the hypervisors. This functionality can be achieved using the RSTP protocol for backing up two remote locations, and the VRRP protocol is used to backup the default gateways in each network. The following illustration depicts the resulting topology. The SAN infrastructure is depicted in green. The connections between the active components within the distribution layer are depicted in blue. The backbone links are drawn in red; these provide a high-speed connection to the remote locations. The illustration also includes two separate optical cables, which are crucial to the entire cluster's full redundancy.



3.3 Network configuration

The previous chapter described important aspects for the proper functioning of a virtual cluster. This primarily requires access to the same VLAN for all hypervisors in all locations, as well as back up of the default gates in each subnet in order to keep it accessible even during outages or router upgrades. This functionality is primarily achieved by the backbone network with the STP and VRRP protocols. The configuration of backbone components is examined in detail in [5] and [6]. The VLAN configuration is most important when configuring access points. Every hypervisor must have access to all networks. The names of network interfaces must be same on every hypervizor. This is the only way that problem-free migration of the virtual servers can be guaranteed between the individual hypervisors in the various locations.

The crucial configuration of active components is shown in the following examples. The purpose of the configuration referred to in these examples is to set up three VLAN for the switches. Two of them are set aside to operate the virtual servers (10.0.1.0/24, 10.0.2.0/24) and one is for management (10.255.1.0/24). Besides these three networks, the backbone network must also include two networks reserved for iSCSI operation (10.255.3.0/24, 10.255.4.0/24).

3.3.1 An example of the configuration of a SAN device

Switches in a SAN infrastructure do not require a complicated configuration. Basic commands suffice for their installation. To confirm the functioning of the SFP modules and the stability of the links, it is easiest to use the command *show interface brief* or *show interface 24*, where the number 24 represents the number of the SFP port of the optical module. To confirm the basic connection, the usual *ping* command suffices.

Each component's configuration requires the setting up of IP addresses in the management network and several other commands.

```
hostname "virtual-kou-sw1"
no cdp run
no web-management
vlan 506
   name "mgmt"
   ip address 10.255.1.9 netmask 255.255.255.0
   untagged 1
   exit
snmp-server community "public" operator
snmp-server contact "hostmaster@example.com" location "kou, sal"
```

Properly set values should be applied to the switches. Otherwise, the logs will not make sense. It is a good idea to store the logs remotely, because the reboot of any equipment usually erases the log file. The *ip authorized-managers* command offers, at least, basic security.

```
timesync sntp
time timezone 60
time daylight-time-rule Western-Europe
sntp unicast
sntp server priority 1 10.255.1.1
```

```
logging 10.255.1.1
logging facility syslog

ip authorized-managers 10.255.1.0 255.255.255.0 access manager
crypto key generate ssh rsa
ip ssh
```

The previous configuration is usually the same on all equipment, while the most important configuration for SAN switches is the configuration of the VLAN for iSCSI operation and their IP addresses.

```
vlan 546
   name "vmware-iscsi"
   untagged 2-24
   ip address 10.255.3.109 255.255.255.0
   exit
```

3.3.2 Example of access-device configuration

The basic configuration has already been described in the previous chapter. Besides these basics, a redundant connection to the backbone network and VLAN configuration for connected equipment is important for component access. Configuration of VLAN management and spanning tree protocol (STP) is a basic component of backbone connectivity.

```
vlan 506
   name "mgmt"
   tagged 1-4
   ip address 10.255.1.8 netmask 255.255.255.0
   exit
spanning-tree force-version rstp-operation
spanning-tree
```

Use the *show spanning-tree* command to verify STP functionality. This provides an abundance of useful information. Usually it suffices to verify the state of both uplinks. One should be set to the “Forwarding state”, while the other to “Blocking”. Furthermore, the "Time Since Last Change" value should be the same for all other components in the same STP domain. Once the STP functions properly, it is possible to set up the user VLANs.

```
vlan 3
   name "ant-servers"
   tagged 1-4
   exit
vlan 654
   name "kou-servers"
   tagged 1-4
   exit
```

3.3.3 Example of backbone-device configuration

The configuration of backbone components is sufficiently dealt with in the GN3 documentation [6]. The most important part of the configuration is to set up the STP, GVRP, VRRP, and OSPF protocols.

```
vlan 506
   name "mgmt"
   tagged 1-4
   ip address 10.255.1.3 netmask 255.255.255.0
   exit
spanning-tree force-version rstp-operation
spanning-tree
gvrp
```

To implement the OSPF protocol, it is first necessary to create a point-to-point connection between neighbouring routers and to activate the OSPF to these interfaces. The *show ip ospf neighbour* and *show ip route* commands are the most important for determining the protocol states.

```
ip routing
router ospf
area 0.0.0.2
   redistribute connected
   exit
vlan 240
   name "ext240"
   ip address 147.229.240.2 255.255.255.252
   ip ospf 147.229.240.2 area 0.0.0.2
   tagged B21
   exit
vlan 241
   name "ext241"
   ip address 147.229.241.2 255.255.255.252
   ip ospf 147.229.241.2 area 0.0.0.2
   tagged B22
   exit
```

The last part of the configuration concerns the VRRP protocol. Its configuration must be made on both backbone routers (which perform the backups) at the same time. The configuration should be made for all subnets, whose default gateway should be backed up. The configuration of the primary router can be as follows.

```
vlan 3
   name "ant-servers"
   ip address 147.229.3.1 255.255.255.128
   tagged 1-4
   vrrp vrid 1
      owner
      virtual-ip-address 147.229.3.1 255.255.255.128
      enable
      exit
   exit
vlan 654
   name "kou-servers"
   ip address 147.229.3.254 255.255.255.128
   tagged 1-4
   vrrp vrid 2
      backup
      virtual-ip-address 147.229.3.130 255.255.255.128
      enable
      exit
   exit
```

The configuration of the other router can be as follows.

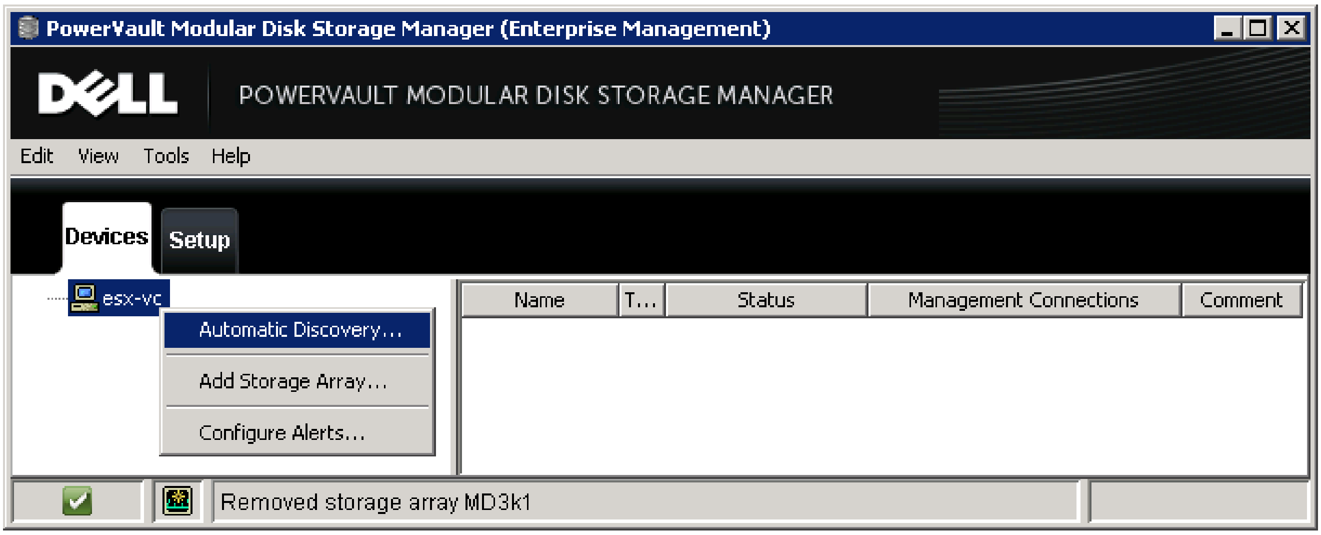
```
vlan 3
   name "ant-servers"
   ip address 147.229.3.126 255.255.255.128
   tagged 1-4
   vrrp vrid 1
      backup
      virtual-ip-address 147.229.3.1 255.255.255.128
      enable
      exit
   exit
vlan 654
   name "kou-servers"
   ip address 147.229.3.130 255.255.255.128
   tagged 1-4
   vrrp vrid 2
      owner
      virtual-ip-address 147.229.3.130 255.255.255.128
      enable
      exit
   exit
```

4 Storage Installation

Preparation of the disk array is another step in the installation process. When mounting into the rack, it is best to install such equipment close to the ground, mostly for stability reasons, because the disk array full of hard drives is usually rather heavy. The temperature surrounding the drives is also important. In some cases, the temperature between the upper and lower sections of the rack can differ by tens of degrees Celsius, depending on the performance of the server room's air conditioning system. After installation into the rack, the disk array management ports can be connected to the same subnet, and both ports should be connected to different switches, for backup purposes. Within the same subnet, a station running on a Windows operating system with the Powervault Modular Disk Storage Manager must be installed in advance. The most up-to-date version of this software is available from the developer [7].

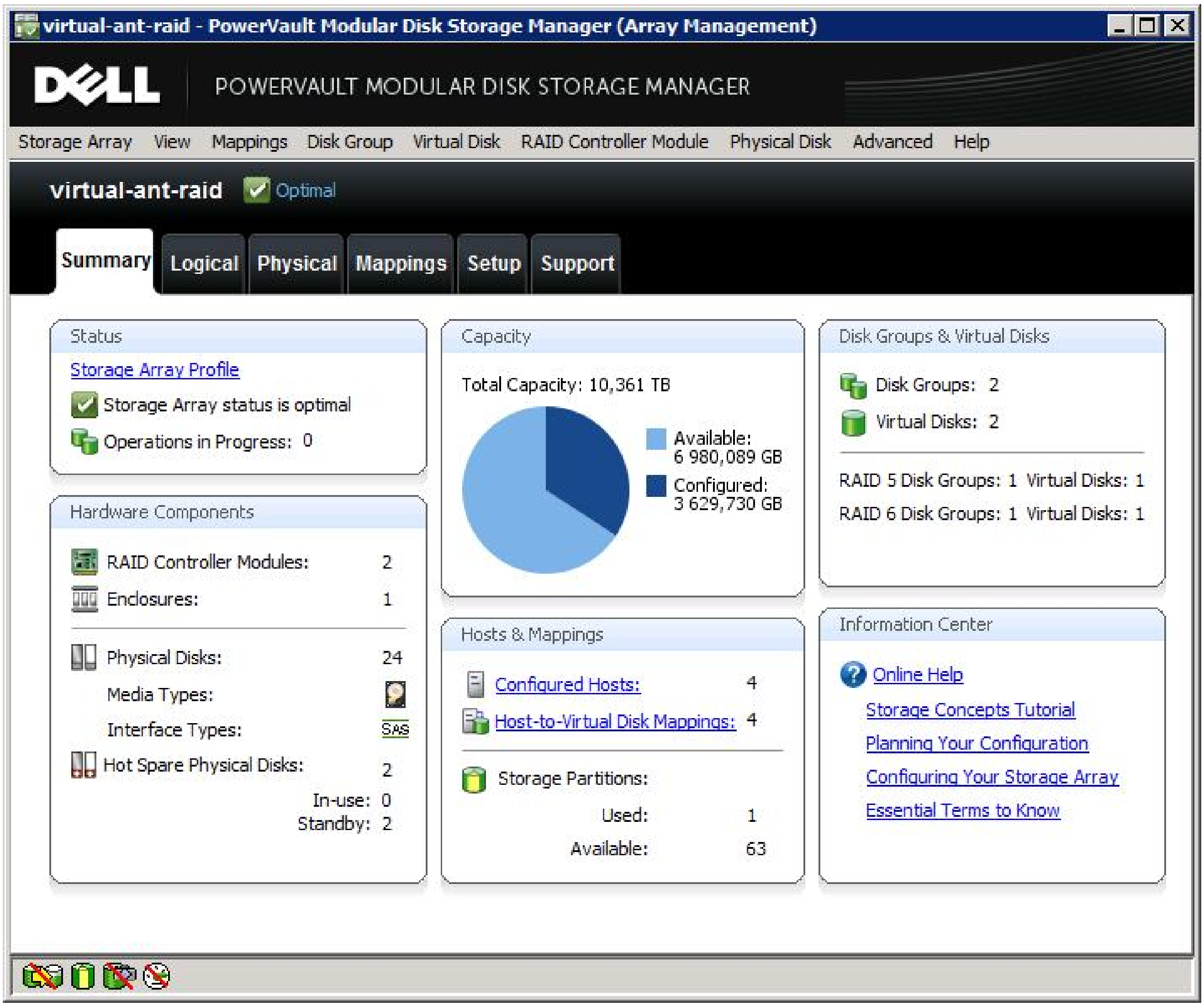
4.1 Connection to the device

The first step in installing the disk array is to connect it to the management software. For arrays with default settings, it is best to use Automatic Discovery. If the management ports are already configured (if they are already assigned to an IP address), it is faster to connect to this array using the assigned IP address.



Removed storage array MD3k1

After connecting to the disk array, the basic status is displayed. Modification of the management port IP addresses is accomplished via *Setup > Configure Ethernet Management Ports*. Configuration of the iSCSI ports is closely related to the SAN infrastructure topology (refer to Chapter 3.1) and is accomplished via *Setup > Configure iSCSI Host Ports*.


virtual-ant-raid - PowerVault Modular Disk Storage Manager (Array Management)
DELL
POWERVAULT MODULAR DISK STORAGE MANAGER
Storage Array View Mappings Disk Group Virtual Disk RAID Controller Module Physical Disk Advanced Help
Summary Logical Physical Mappings Setup Support

virtual-ant-raid ✓ Optimal

Status

Storage Array Profile

✓ Storage Array status is optimal

Operations in Progress: 0

Hardware Components

RAID Controller Modules:	2
Enclosures:	1
Physical Disks:	24
Media Types:	
Interface Types:	SAS
Hot Spare Physical Disks:	2
In-use:	0
Standby:	2

Capacity

Total Capacity: 10,361 TB

Available: 6 980,089 GB

Configured: 3 629,730 GB

Disk Groups & Virtual Disks

Disk Groups: 2

Virtual Disks: 2

RAID 5 Disk Groups: 1 Virtual Disks: 1

RAID 6 Disk Groups: 1 Virtual Disks: 1

Hosts & Mappings

Configured Hosts:	4
Host-to-Virtual Disk Mappings:	4
Storage Partitions:	
Used:	1
Available:	63

Information Center

Online Help

Storage Concepts Tutorial

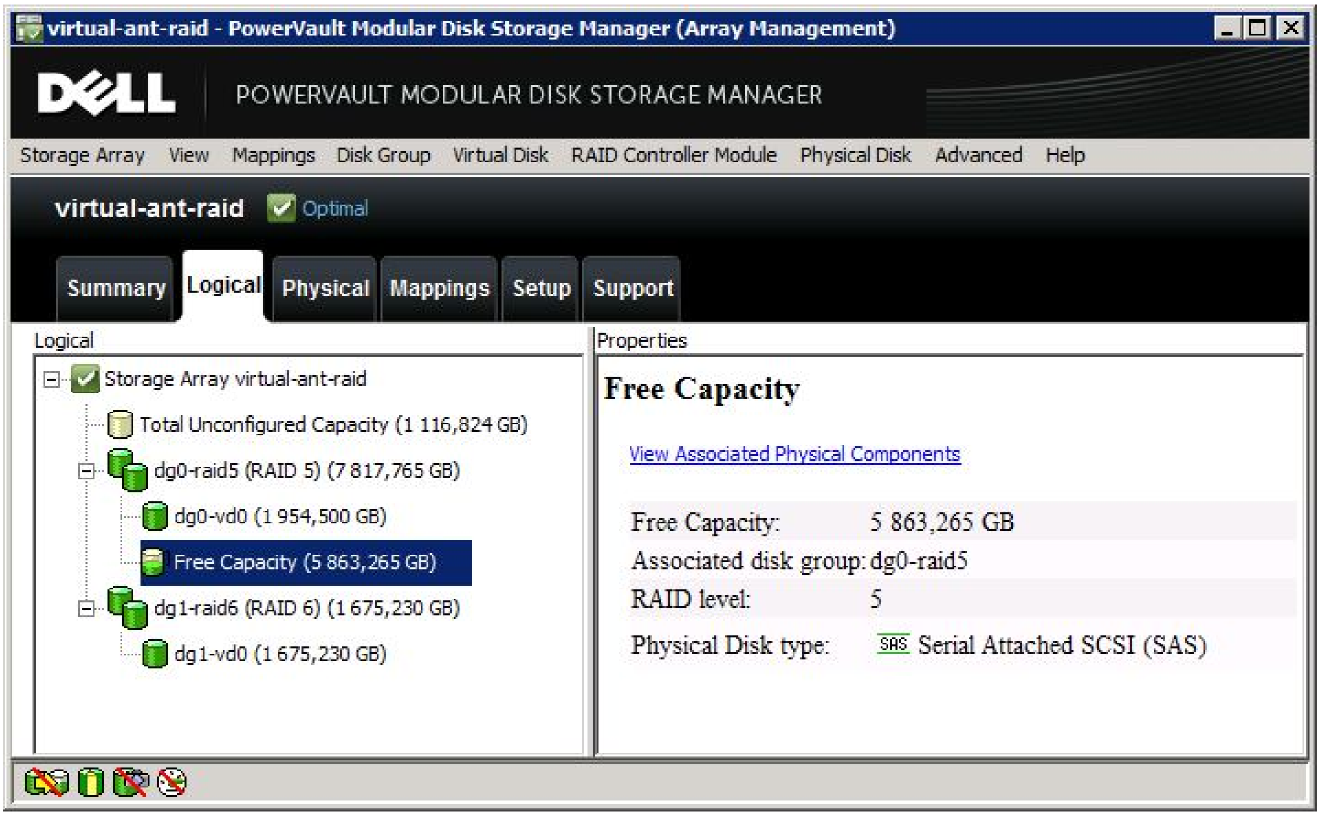
Planning Your Configuration

Configuring Your Storage Array

Essential Terms to Know

4.2 Disk group configuration

The disk array must first be partitioned into disk groups. The size (capacity) of individual disk groups is determined by the number of assigned hard disks, the RAID level, and the set cash of the SSD unit. The total capacity of individual disk groups is then further partitioned into virtual disks.



Storage Array virtual-ant-raid
- Total Unconfigured Capacity (1 116,824 GB)
- dg0-raid5 (RAID 5) (7 817,765 GB)
  - dg0-vd0 (1 954,500 GB)
  - Free Capacity (5 863,265 GB)
- dg1-raid6 (RAID 6) (1 675,230 GB)
  - dg1-vd0 (1 675,230 GB)

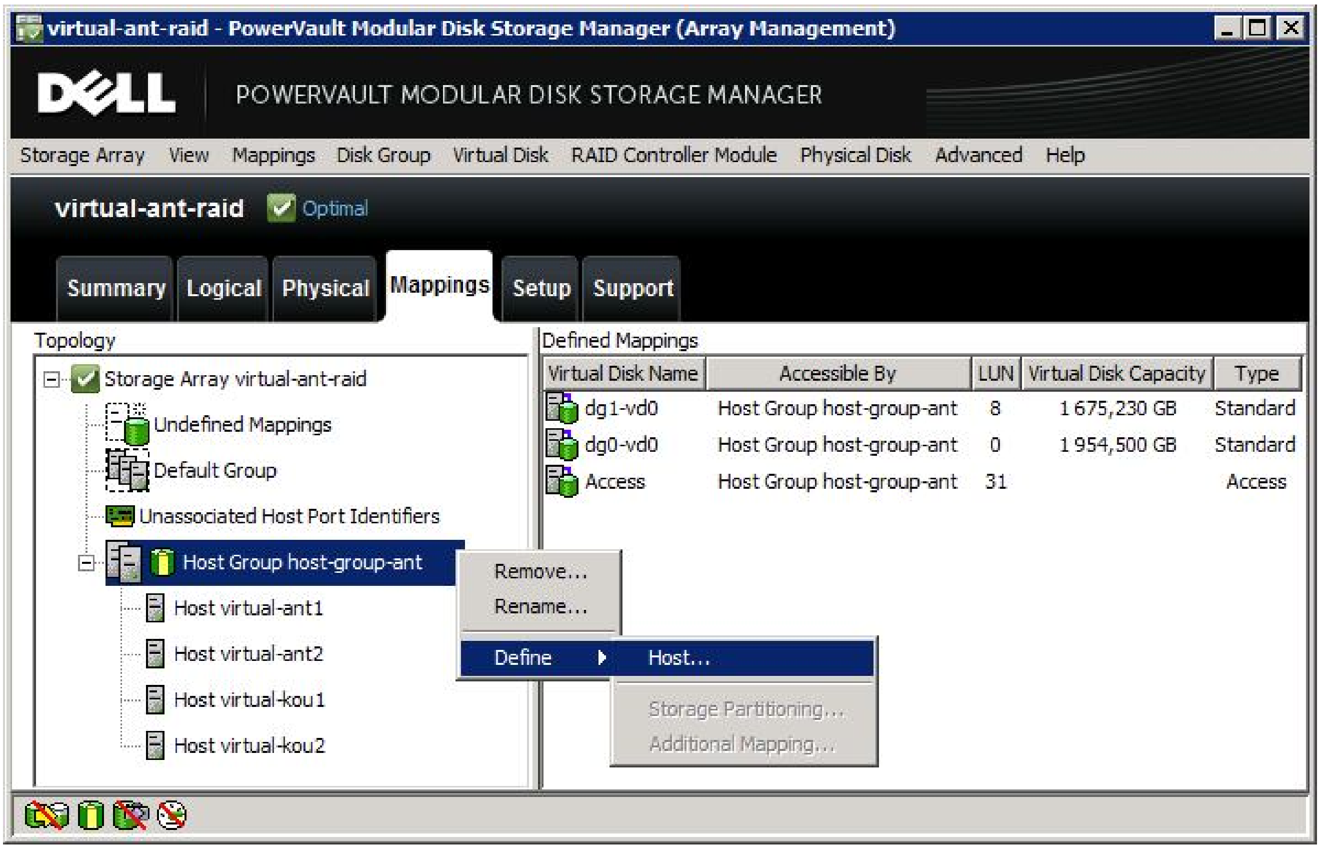
Free Capacity

View Associated Physical Components

Free Capacity: 5 863,265 GB
Associated disk group: dg0-raid5
RAID level: 5
Physical Disk type: SAS Serial Attached SCSI (SAS)

4.3 Virtual disk mappings

These virtual disks are assigned to individual iSCSI clients within the “Mappings” tab. The virtual disk can be shared among clients, provided the sharing is supported by the file operating system. The individual clients are identified, not only by their IP addresses, but are also identified by an "iSCSI Initiator String". For iSCSI clients of the VMware vSphere hypervisor, the "iSCSI Initiator String" is generated once the iSCSI protocol is activated. This is further explained in Chapter 5.5. The mapping of the new hypervisor to the virtual disk or a group of virtual disks is depicted in the following illustration. To add a new host, you will need to know its identifier, as described above. If the host had already attempted to connect to the iSCSI array, its hostname and iSCSI Initiator String have already been stored in the unestablished connection table, where the correct host can be chosen with a single click, without the need to manually input the host's identifier.



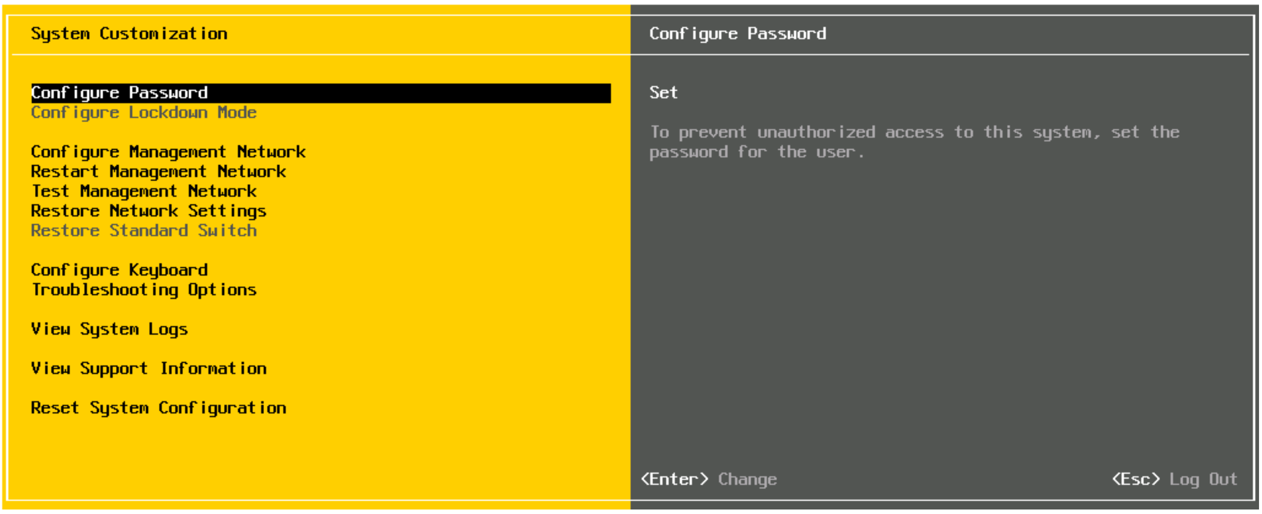
Virtual Disk Name	Accessible By	LUN	Virtual Disk Capacity	Type
dg1-vd0	Host Group host-group-ant	8	1 675,230 GB	Standard
dg0-vd0	Host Group host-group-ant	0	1 954,500 GB	Standard
Access	Host Group host-group-ant	31		Access

5 Hypervisor Installation

A hypervisor's installation is not significantly different from the installation of other operating systems. When booting from the installation disk, it is sufficient to run the Esxi-5.0.0 Installer from the menu. It is necessary to choose the location where the VMware sSphere system should be installed. This location can be on a hard drive, a RAID disk array, an SSD disk, or a flash drive. In Chapter 2.1, the Dell R610 server was chosen, configured for SD memory, and is the server on which the hypervisor is installed. The last step before initiating the installation is to configure the root password. Once it has been rebooted, the system is now set up.

5.1 Connection to the hypervisor

In the default state, the hypervisor is assigned a dynamic IP address from the DHCP server. For problem-free work with the hypervisor, it is better to set a fixed IP address. This is possible from the system's console. Press F2 and enter your login details to show the basic configuration panel.

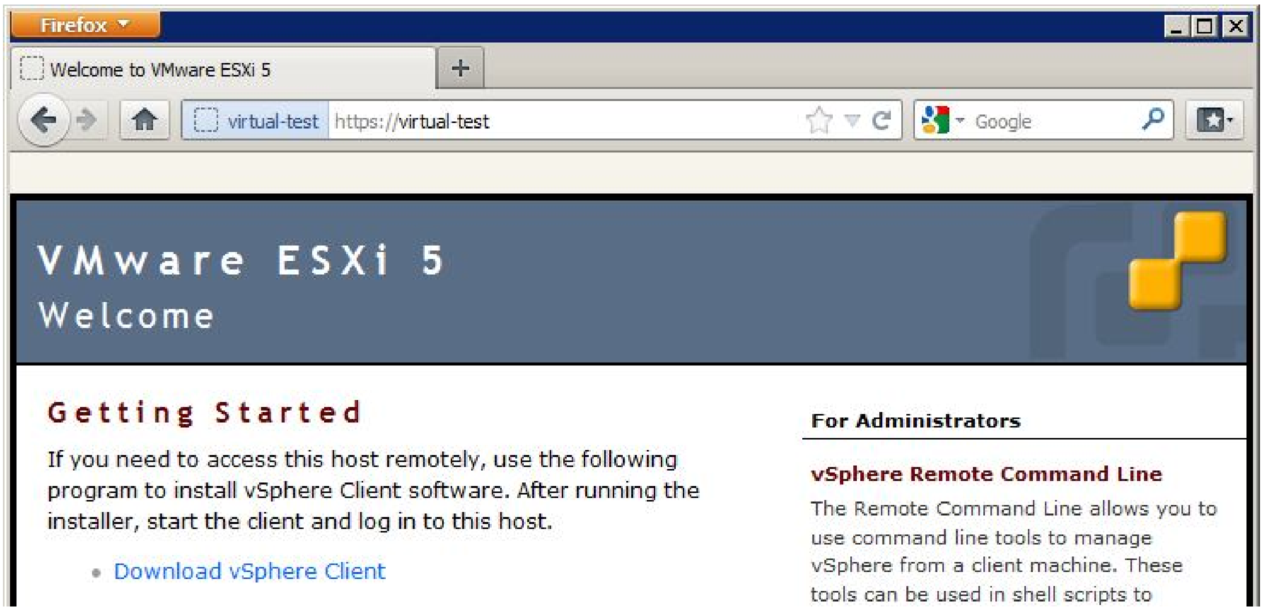


```
System Customization                          Configure Password

Configure Password                            Set
Configure Lockdown Mode
                                              To prevent unauthorized access to this system, set the
Configure Management Network                  password for the user.
Restart Management Network
Test Management Network
Restore Network Settings
Restore Standard Switch

Configure Keyboard
Troubleshooting Options

View System Logs

View Support Information

Reset System Configuration

                                              <Enter> Change                   <Esc> Log Out
```

The purpose of each item should be evident. A fixed IP address can be set under “Configure Management Network”. In addition to the IP address, this item menu can also be used to set up the network interface, VLANid, or the DNS parameters. The Management Network should be restarted and tested according to the following steps. A specialised client should be used for complete configuration.

5.2 VMware vSphere client

This is a separate application used to control and configure the VMware vSphere system. The easiest means of obtaining a client is through the hypervisor's web interface.



VMware ESXi 5

Welcome

Getting Started

If you need to access this host remotely, use the following program to install vSphere Client software. After running the installer, start the client and log in to this host.

- Download vSphere Client

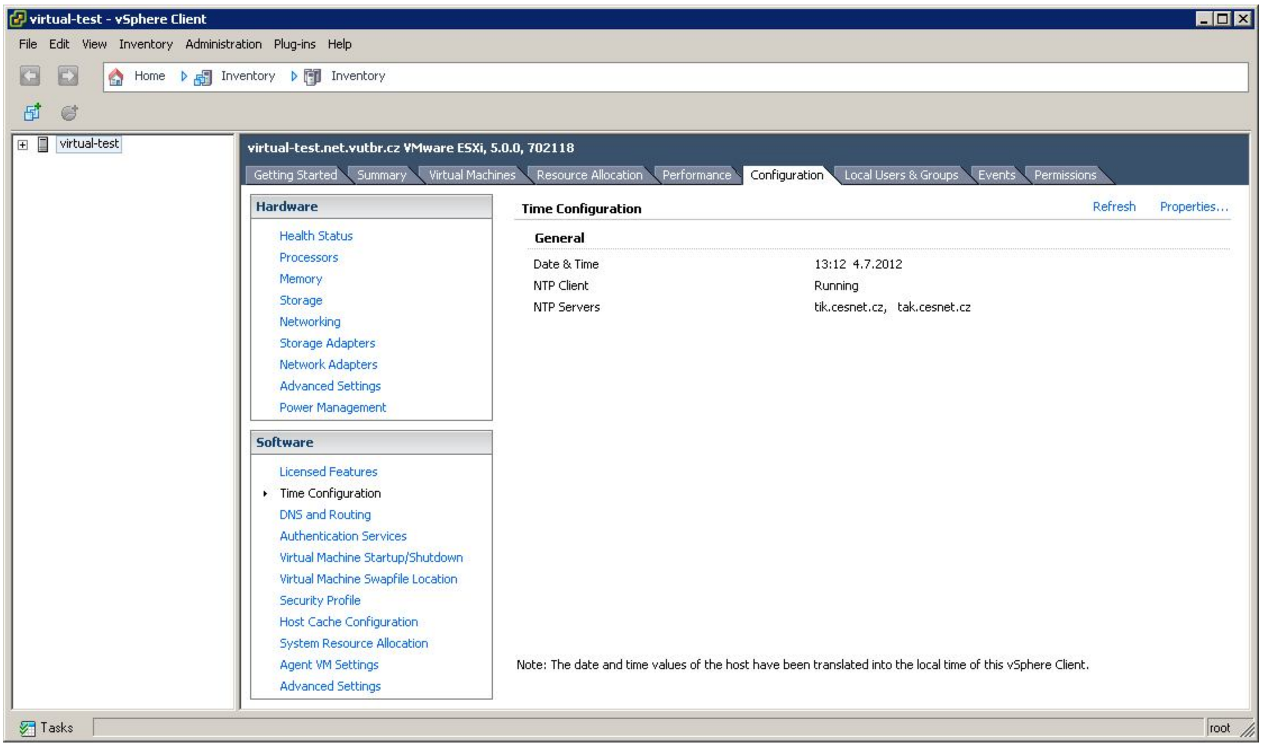
For Administrators

vSphere Remote Command Line

The Remote Command Line allows you to use command line tools to manage vSphere from a client machine. These tools can be used in shell scripts to

5.3 Hypervisor configuration

Configuration of the hypervisor, using the VmWare vSphere Client, requires a preset IP address on the network interface, and login details. All these configuration details are inputted locally on the server (refer to Chapter 5.1).



virtual-test.net.vutbr.cz VMware ESXi, 5.0.0, 702118

Time Configuration

General

Date & Time 13:12 4.7.2012
NTP Client Running
NTP Servers tik.cesnet.cz, tak.cesnet.cz

Note: The date and time values of the host have been translated into the local time of this vSphere Client.

When logging into the hypervisor, it is advisable to configure the NTP server, the DNS parameters, and the SSH server.

- *Time Configuration – Properties – Options – NTP Settings*
- *DNS and Routing – Properties – DNS Configuration – Look for hosts in the following domains*
- *DNS and Routing – Properties – Routing – Default gateway*
- *Security Profile – Services Properties – Options – SSH – Startup Policy – Start and Stop with host*
- *Security Profile – Services Properties – Options – SSH – Services commands – Start*

5.3.1 Network interface configuration

In VmWare, there are two types of network interface: VMKernel and Virtual Machine. VMKernel is a network interface for connecting the following ESXi services: vMotion, iSCSI, NFS, and Host Management. The Virtual Machine interface establishes a connection between the virtual server and the computer network. These interfaces can be configured, either through a VMware vSphere Client or from the command line. The easiest way to configure the network interface is with the VMware vSphere Client. The advantage of this type of configuration is its simplicity; even a beginner can set up the required network interface relatively comfortably. The disadvantage of this method of configuration is the time required. Also, with a larger number of network interfaces, it is more difficult to achieve an identical configuration for all hypervisors (which is required for the proper migration of virtual servers between hypervisors). When configuring multiple hypervisors, it is better to configure the network interfaces using the console. This allows for the configuration of all hypervisor network interfaces using a sequence of commands, which may repeated for all hypervisors that have the same hardware configuration. This method enables an identical configuration of hypervisor network interfaces within a cluster. The following command will show the initial state of network interfaces once the installation of the hypervisor is complete.

```
~ # esxcfg-vswitch -l
Switch Name      Num Ports   Used Ports  Configured Ports  MTU     Uplinks
vSwitch0         128         3           128               1500    vmnic0

  PortGroup Name        VLAN ID  Used Ports  Uplinks
  VM Network            0        0           vmnic0
  Management Network    0        1           vmnic0
```

To understand the virtual network interface in VMware, it is important to understand the following hierarchal concepts: vswitch, vmknic, vmnic, and port group. On the lowest level, physical network interfaces (vmnic) are assigned to individual vswitch objects. These physical network interfaces are used for communication between the hypervisor and the virtual servers with connected systems. More PNIC interfaces in a single vSwitch reveal a higher degree of redundancy (standby or link-aggregation). The lack of a VMNIC interface means that the given vSwitch only handles communication between the virtual servers. Each vSwitch is similar to an L2 switch that supports VLAN. In VMware terminology, "port group" is used instead of VLAN. These port groups are also used to connect virtual servers and for the services of the hypervisor system.

5.3.2 vSwitch configuration

Before configuring vSwitches, you must decide which physical interface should be used to create the vSwitch objects, and which services the vSwitch should stop. In most cases of a VMware cluster with iSCSI storage/array, the correct partition for four physical network cards (vmnic0-4) is as follows:

- vSwitch0 – vmnic0, vmnic1 – back up connection of virtual servers, host management and vMotion;
- vSwitch1 – vmnic2 – iSCSI operation;
- vSwitch2 – vmnic3 – iSCSI operation.

The first step is to configure port groups for host management. At start, a "Management Network" port group is created on the vSwitch0 with vmnic0 uplink. The following commands add a second physical network interface to the virtual switch, set them as standby backup interface, and create port groups for connection to the virtual servers. With the *esxcfg-vswitch* command, the numbers following the -v parameter represent the VLAN ID value, with which the given port group's packets are distributed across the vmnic0 and vmnic1 physical interfaces to the backbone network.

```
esxcfg-vswitch -L vmnic1 vSwitch0
esxcli network vswitch standard policy failover set -s vmnic1 -v vSwitch0

esxcfg-vswitch -A "vpn-server" vSwitch0
esxcfg-vswitch -A "mgmt-ro" vSwitch0
esxcfg-vswitch -A "vlan3" vSwitch0
esxcfg-vswitch -A "vlan3-kou" vSwitch0

esxcfg-vswitch -v 660 -p "vpn-server" vSwitch0
esxcfg-vswitch –v 506 –p "mgmt-ro"    vSwitch0
```

```
esxcfg-vswitch -v 3   -p "vlan3"      vSwitch0
esxcfg-vswitch -v 654 -p "vlan3-kou"  vSwitch0
```

The following steps describe how to create other vSwitch objects and their configuration as iSCSI ports.

```
esxcfg-vswitch -a vSwitch1
esxcfg-vswitch -a vSwitch2
esxcfg-vswitch -L vmnic2 vSwitch1
esxcfg-vswitch -L vmnic3 vSwitch2
esxcfg-vswitch -A "iSCSI1" vSwitch1
esxcfg-vswitch -A "iSCSI2" vSwitch2
```

Each hypervisor has two separate IP address in the subnets, and each address establishes connectivity with the disk array independently.

hypervisor1:

```
esxcfg-vmknic -a -i 10.255.3.10 -n 255.255.255.0 iSCSI1
esxcfg-vmknic -a -i 10.255.4.10 -n 255.255.255.0 iSCSI2
```

hypervisor2:

```
esxcfg-vmknic -a -i 10.255.3.11 -n 255.255.255.0 iSCSI1
esxcfg-vmknic -a -i 10.255.4.12 -n 255.255.255.0 iSCSI2
```

5.3.3 iSCSI configuration

A graphics client is suitable for hypervisor iSCSI configuration.

- *Configuration - Storage adapters - Add - Add software iSCSI adapter*
  A new software iSCSI adapter will be added to the Storage Adapter list. After it has been added, select the software iSCSI adapter in the list and click on Properties to complete the configuration.

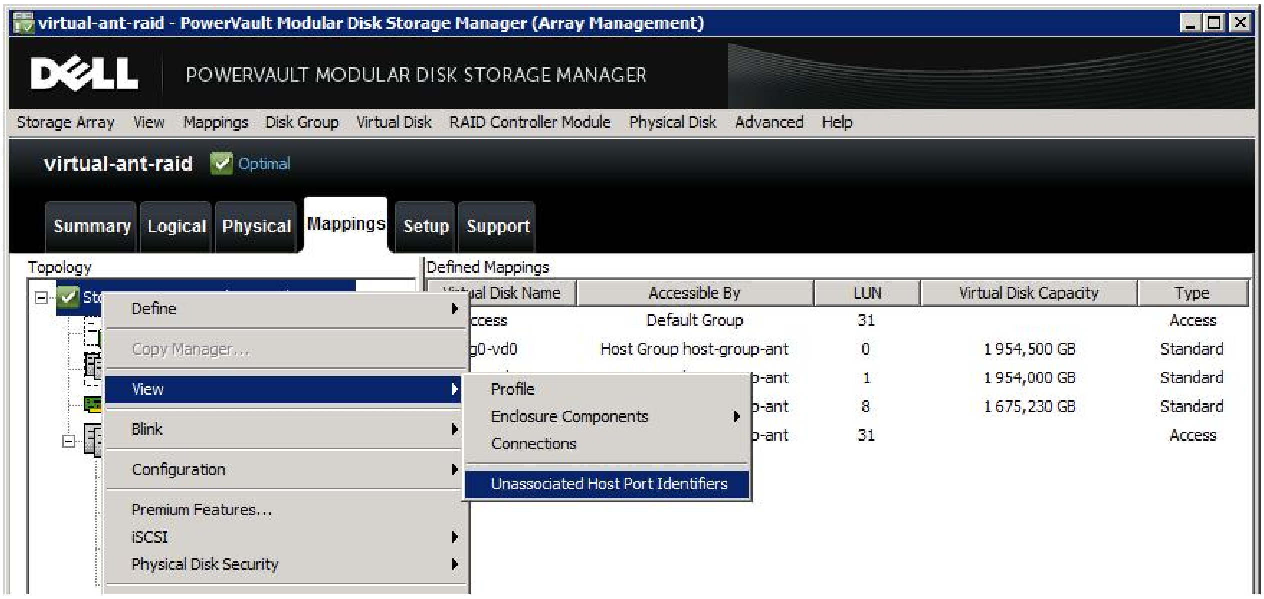
- *OK*
- *iSCSI Software Adapter - vmhba<number> - Properties - Dynamic Discovery / Static Discovery*
  Add IP addresses of iSCSI target. These addresses match topology of SAN infrastructure (Chapter 3.1).

  ```
  10.255.3.1
  10.255.4.1
  10.255.3.2
  10.255.4.2
  ```

- *Next*
  A rescan of the host bus adapter is recommended for this configuration change. Rescan the adapter?

- *Yes*
  This step secures a hypervisor-connection attempt to storage, using its IP addresses and iSCSI name. The iSCSI session must be permited on storage. Here, it is important to check that the storage knows the iSCSI name of hypervisor. The next steps are realised in Powervault Modular Disk Storage Manager and build on information obtained in Chapter 4.3.

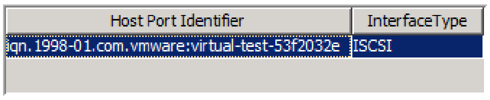
- *Open Powervault Modular Disk Storage Manager*
- *Mappings - Storage(in left window) – View – Unassociated Host Port Identifiers*



Defined Mappings

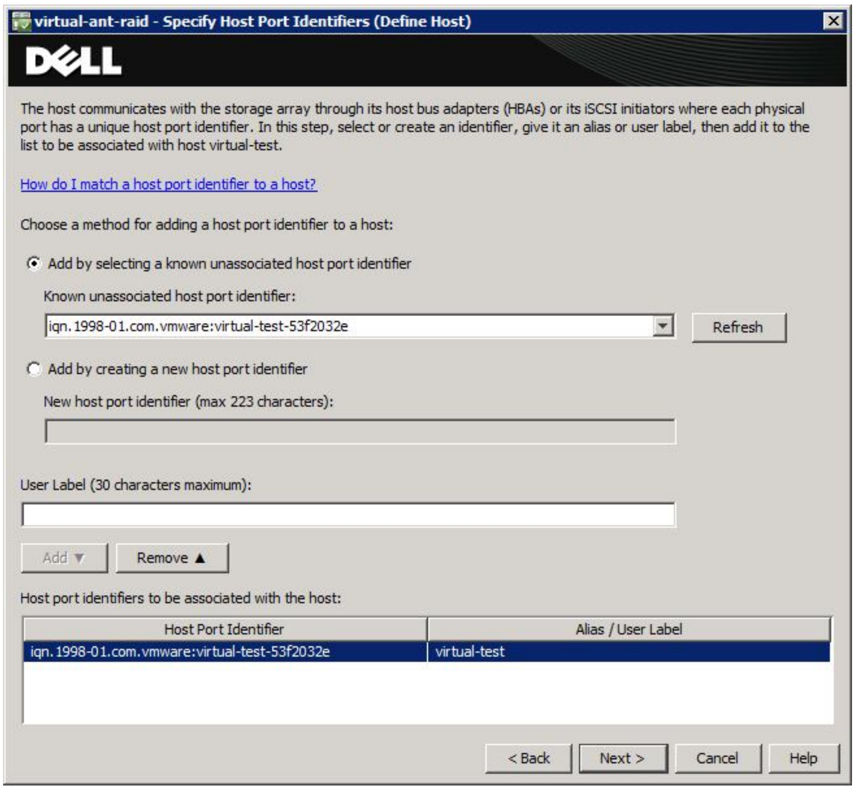
Virtual Disk Name	Accessible By	LUN	Virtual Disk Capacity	Type
ccess	Default Group	31		Access
g0-vd0	Host Group host-group-ant	0	1 954,500 GB	Standard
	p-ant	1	1 954,000 GB	Standard
	p-ant	8	1 675,230 GB	Standard
	p-ant	31		Access

- *List of unassociated host port identifiers.*



Host Port Identifier	InterfaceType
iqn.1998-01.com.vmware:virtual-test-53f2032e	ISCSI

- *Mappings - Storage - Host Group - Define Host - <Host name> - Add by selecting a know unassociated host port identifier <choose right one> - User Laber <write some good string> - Add <check that Host port Identifier and User Label match hypervisors values> - Next*


virtual-ant-raid - Specify Host Port Identifiers (Define Host)
DELL

The host communicates with the storage array through its host bus adapters (HBAs) or its iSCSI initiators where each physical port has a unique host port identifier. In this step, select or create an identifier, give it an alias or user label, then add it to the list to be associated with host virtual-test.

How do I match a host port identifier to a host?

Choose a method for adding a host port identifier to a host:

● Add by selecting a known unassociated host port identifier

Known unassociated host port identifier: iqn.1998-01.com.vmware:virtual-test-53f2032e

○ Add by creating a new host port identifier

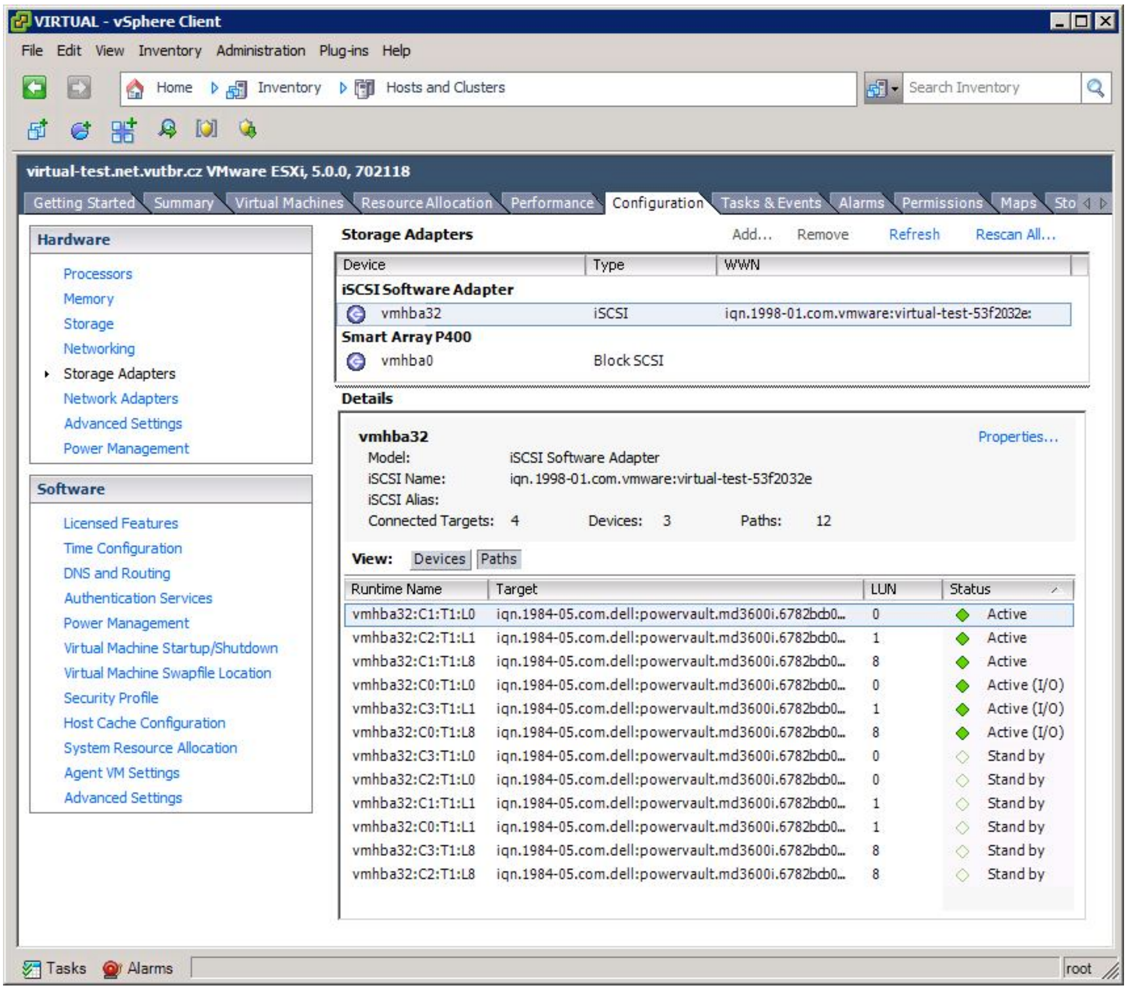
New host port identifier (max 223 characters):

User Label (30 characters maximum):

Host port identifiers to be associated with the host:

Host Port Identifier	Alias / User Label
iqn.1998-01.com.vmware:virtual-test-53f2032e	virtual-test

- *Host Type - VMWARE (or linux) - Next - Finish*
- *Close Powervault Modular Disk Storage Manager*
- *Go back to VMware vSphere Client*
- *Configuration - Storage adapters - Rescan all*
  Now it is possible to see active devices and paths



Device	Type	WWN
iSCSI Software Adapter		
vmhba32	iSCSI	iqn.1998-01.com.vmware:virtual-test-53f2032e:
Smart Array P400		
vmhba0	Block SCSI	

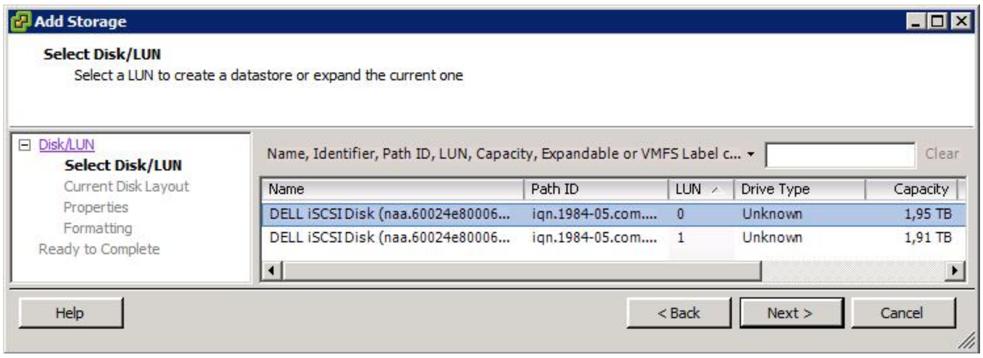
vmhba32
Model: iSCSI Software Adapter
iSCSI Name: iqn.1998-01.com.vmware:virtual-test-53f2032e
iSCSI Alias:
Connected Targets: 4 Devices: 3 Paths: 12

Runtime Name	Target	LUN	Status
vmhba32:C1:T1:L0	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	0	◆ Active
vmhba32:C2:T1:L1	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	1	◆ Active
vmhba32:C1:T1:L8	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	8	◆ Active
vmhba32:C0:T1:L0	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	0	◆ Active (I/O)
vmhba32:C3:T1:L1	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	1	◆ Active (I/O)
vmhba32:C0:T1:L8	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	8	◆ Active (I/O)
vmhba32:C3:T1:L0	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	0	◇ Stand by
vmhba32:C2:T1:L0	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	0	◇ Stand by
vmhba32:C1:T1:L1	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	1	◇ Stand by
vmhba32:C0:T1:L1	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	1	◇ Stand by
vmhba32:C3:T1:L8	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	8	◇ Stand by
vmhba32:C2:T1:L8	iqn.1984-05.com.dell:powervault.md3600i.6782bdb0...	8	◇ Stand by

5.3.3.1 Partitions

The following lines can be skipped if partitions have been created and formatted on the disk array. In other cases, partitions must be created and formatted. Follow these next steps for each newly added partition, according to need.

- *Configuration - Storage - Add Storage - Disk/LUN*



Select Disk/LUN
Select a LUN to create a datastore or expand the current one

Name, Identifier, Path ID, LUN, Capacity, Expandable or VMFS Label c...

Name	Path ID	LUN	Drive Type	Capacity
DELL iSCSI Disk (naa.60024e80006...	iqn.1984-05.com....	0	Unknown	1,95 TB
DELL iSCSI Disk (naa.60024e80006...	iqn.1984-05.com....	1	Unknown	1,91 TB

- *A partition will be created and used - Next*
- *<Enter the partition name> - Next*
- *<Choose optimal Block Size> - Maximum available space - Next*
  1024GB, 4MB Block Size should be good for majority

- *Finish*

5.4 vMotion

The VMware Cluster base is now functional. It has access to the iSCSI array. In order to migrate virtual servers between hypervisors, all hypervisors must be administered centrally by the VMware vCenter Server application. It is also necessary to define the hypervisor's VMKernel network interface (refer to Chapter 5.3.1), across which the vMotion transmissions will be made.

- *Configuration - Networking - vSwitch0 – properties*
- *<choose right VMKernel or define a new one>*
- *Port Properties - vMotion (checkbox on)*

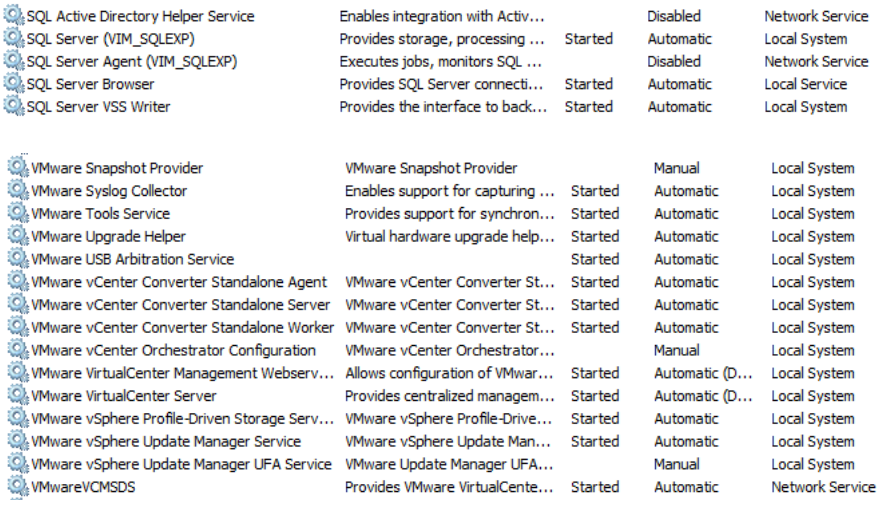
6 vCenter Server

VMware vCenter Server is a software application for the centralised management of a virtual infrastructure. VCenter vSphere 5 comes in either Windows or Linux versions. However, the Linux version is somewhat behind, since it is not possible to run the Update Manager or some of the plugins. For these reasons, it is better to use the full functionality of the Windows version. This version operates on the MSSQL database. MS SQL 2005 is included in vCenter Server's installation. This server is fully functional, but without additional software, it is not possible to backup the database, or to perform more advanced management of the database. However, in most cases, it is sufficient. If you require some of the more advanced functions, simply install MS SQL Server Management Studio. Alternatively, you can migrate to MS SQL 2008, which may be used free-of-charge after fulfilling the licensing conditions.

6.1 vCenter installation

Some conditions must be met before installing the vCenter Server. Above all, you will need a 64bit version of Windows installed on a suitable server. The server may be physical or virtual, but should not be located on any of the hypervisors in the created virtualisation cluster. One of the disadvantages to this would be evident with any hypervisor outage on which vCenter Server is running. If this were the case, then the central management would stop running, as would the arbitrator, which would normally be the server that would determine which virtual server had been affected by the hypervisor outage, and which server should run instead on another hypervisor. From this perspective, it is truly better to run vCenter Server on a separate server, ideally as a virtual server in a standalone installation of VMware vSphere. The advantage of this approach, as opposed to a hardware server, is the VCenter Server's ability to take snapshots, its easy re-installation, and its ability to better manage the resources of the physical servers.

The actual installation of the vCenter Server is trivial and does not require any special effort. The user name used during the installation is the same as the one that will be used to run the server application. However, later users and groups will be added, and their access to virtualisation clusters will be administered. After the server is installed, it is a good idea to also install VMware Update Manager. This is a software application that manages updates of hypervisors and virtual servers. After installing these core applications, you should restart the server and verify, through the services administrator, whether or not the corresponding services have automatically rebooted.



SQL Active Directory Helper Service	Enables integration with Activ...		Disabled	Network Service
SQL Server (VIM_SQLEXP)	Provides storage, processing ...	Started	Automatic	Local System
SQL Server Agent (VIM_SQLEXP)	Executes jobs, monitors SQL ...		Disabled	Network Service
SQL Server Browser	Provides SQL Server connecti...	Started	Automatic	Local Service
SQL Server VSS Writer	Provides the interface to back...	Started	Automatic	Local System
VMware Snapshot Provider	VMware Snapshot Provider		Manual	Local System
VMware Syslog Collector	Enables support for capturing ...	Started	Automatic	Local System
VMware Tools Service	Provides support for synchron...	Started	Automatic	Local System
VMware Upgrade Helper	Virtual hardware upgrade help...	Started	Automatic	Local System
VMware USB Arbitration Service		Started	Automatic	Local System
VMware vCenter Converter Standalone Agent	VMware vCenter Converter St...	Started	Automatic	Local System
VMware vCenter Converter Standalone Server	VMware vCenter Converter St...	Started	Automatic	Local System
VMware vCenter Converter Standalone Worker	VMware vCenter Converter St...	Started	Automatic	Local System
VMware vCenter Orchestrator Configuration	VMware vCenter Orchestrator...		Manual	Local System
VMware VirtualCenter Management Webserv...	Allows configuration of VMwar...	Started	Automatic (D...	Local System
VMware VirtualCenter Server	Provides centralized managem...	Started	Automatic (D...	Local System
VMware vSphere Profile-Driven Storage Serv...	VMware vSphere Profile-Drive...	Started	Automatic	Local System
VMware vSphere Update Manager Service	VMware vSphere Update Man...	Started	Automatic	Local System
VMware vSphere Update Manager UFA Service	VMware Update Manager UFA...		Manual	Local System
VMwareVCMSDS	Provides VMware VirtualCente...	Started	Automatic	Network Service

Sometimes, the order in which the services start may cause conflicts, so if the vCenter does not start up properly, both services should have the following Startup Type value: "Automatic (Delayed Start)" for the following services:

- VMware VirtualCenter Management Webservices;
- VMware VirtualCenter Server.

The VMware vCenter should now be fully functional. For login, the VMware vSphere Client uses the same username and password as the system.

6.2 vCenter configuration

In its default state, vCenter does not yet include any objects that could be administered. Such objects must first be created. The basic object types are as follows.

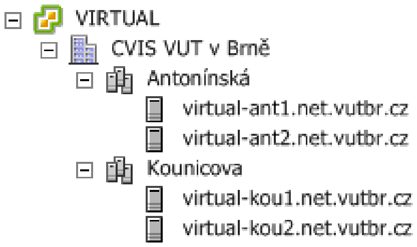
vCenter
Datacenter
Cluster
Host
Virtual Machine

The virtual server (VM) runs on the hypervisor (Host). Hosts are assigned, either to a cluster or directly into the Datacenter. The Datacenter is the basic hierarchical block, which groups clusters with individual hosts. The root of the hierarchy tree is an instance of the VMware vCenter.

The first step in configuring the vCenter is to create a Datacenter object. The importance of this block is that it separates the hardware resources provided by individual hosts and the disk array into functional blocks. Another step is to create a cluster. The last configuration step is to assign the hypervisors to a cluster.

- *<Focus on vCenter instance and show context menu> - New Datacenter - Name - Finish*
- *<Focus on Datacenter instance and show context menu> - New Cluster*
  - *Name - Turn ON vSphere HA - Next*
  - *Host Monitoring Status: Enable Host Monitoring*
  - *Admission Control: Enable*
  - *Admission Control Policy: Host failures the cluster tolerates: 1*
  - *Next*
  - *Cluster Default Settings*
  - *VM restart priority: Medium*
  - *Host Isolation response: Leave powered on*
  - *Next*
  - *VM Monitoring: Disabled*
  - *Monitorign sensitivity: High*
  - *Next*
  - *<Choose right type of CPU>*
  - *Enable EVC for Intel Hosts: Intel Sandy Bridge Generation*
  - *Next*
  - *Store the swapfile in the same directory as the Virtual Machine*
  - *Next – Finish*
- *<Focus on Cluster instance and show context menu> - Add Host*
  - *Connection: <Enter IP or HOSTNAME of hypervisor>*
  - *Authorization: <Enter right credentials>*
  - *Next – Finish*

These steps create a tree structure, such as the one that follows.



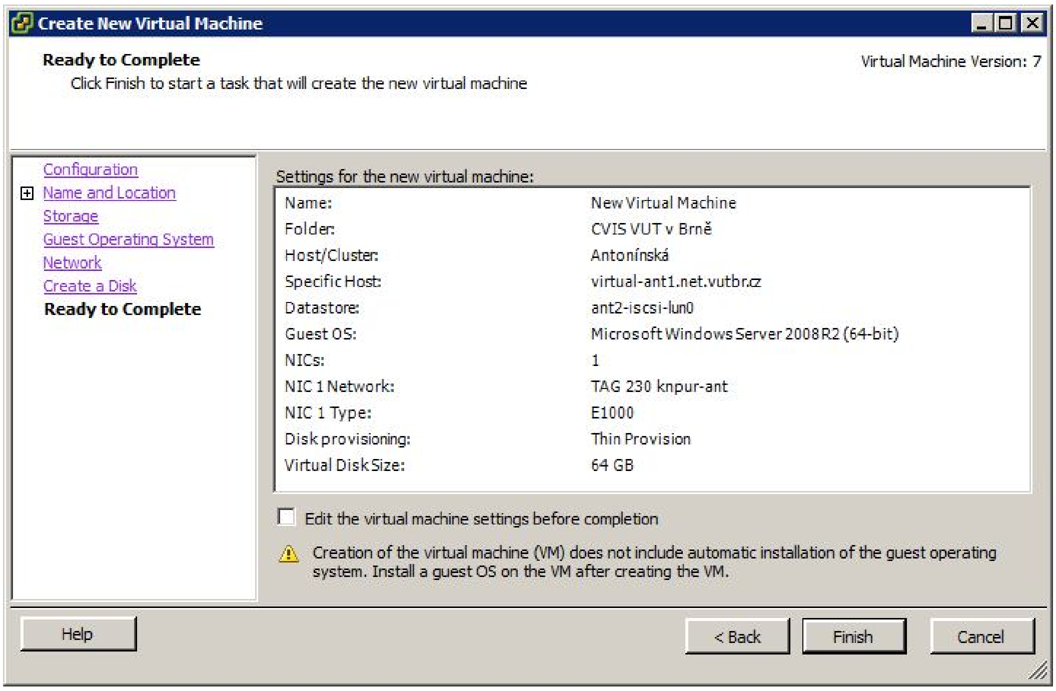
- VIRTUAL
  - CVIS VUT v Brně
    - Antonínská
      - virtual-ant1.net.vutbr.cz
      - virtual-ant2.net.vutbr.cz
    - Kounicova
      - virtual-kou1.net.vutbr.cz
      - virtual-kou2.net.vutbr.cz

6.3 Virtual machines

When creating virtual servers, an item from the “Cluster” or “Host” context menu is usually used.

- *<Focus on Cluster or Host instance and show context menu> - New Virtual Machine*
  - *Configuration: Typical - Next*
  - *Name - Next*
  - *Choose a specific host within a cluster*
  - *Select a destination storage - <iSCSI shared storage must be selected to provide redundancy>*
  - *Next*
  - *Guest Operating System - Next*
  - *Number of NICs*

- *Network: <Choose network name - VLAN ID>*
- *Type of adapter: <Intel E1000 is widely supported network adapter>*
- *Next*
- *Virtual disk size:*
  - *64 GB is good minimal value for WINDOWS 2008 R2/WINDOWS 7 and newer*
  - *8 GB is good minimal value for UNIX/LINUX without X system mounted on /*
    Additional virtual disk for special server functionality is necessary.
    e. g., 64 GB virtual disk mounted on /var/www for webserver
    e. g., 64 GB virtual disk mounted on /var/db for databases
    This schema with every partition on extra virtual disk is very useful for resizing. Instead of resizing a virtual disk and partition and file system is possible to add a new bigger virtual disk to the system a copy all data to a new one. This way save a lot of time and is much saver for your data.
- *Thin Provisioning - <little bit slower, but save a lot of disk space, highly recommended>*
- *Next – Finish*



Ready to Complete
Click Finish to start a task that will create the new virtual machine

Virtual Machine Version: 7

Settings for the new virtual machine:

Name:	New Virtual Machine
Folder:	CVIS VUT v Brně
Host/Cluster:	Antonínská
Specific Host:	virtual-ant1.net.vutbr.cz
Datastore:	ant2-iscsi-lun0
Guest OS:	Microsoft Windows Server 2008 R2 (64-bit)
NICs:	1
NIC 1 Network:	TAG 230 knpur-ant
NIC 1 Type:	E1000
Disk provisioning:	Thin Provision
Virtual Disk Size:	64 GB

☐ Edit the virtual machine settings before completion

⚠ Creation of the virtual machine (VM) does not include automatic installation of the guest operating system. Install a guest OS on the VM after creating the VM.

A virtual server may be installed by running an ISO file from a mapped CV/DVD device, or by PXE. This document does not cover the installation of a virtual operating system.

6.4 Plugins

The following plugins offer advanced functionality of a vSphere Client connected to a vCenter Server. The server part of a plugin is usually installed in the vCenter Server. The vSphere Client displays which plugins are available in the plugin menu.

6.4.1 Update Manager

After installing the VMware Update Manager, the Update Manager plugin becomes available.

- *Plugins - Manage Plugins - VMware vSphere Update Manager Extension*

Download and Install in Status column will begin the installation.

- *Run - Next - Accept - Next - Install – Finish*
- *Install this certificate – Ignore*
  VMware vSphere Update Manager Extension is enabled

The vSphere Client interface now includes an Update Manager menu and an interface for Update Manager Administration. Below is a description of how to upgrade a hypervisor. The first step is to define which patches will be applied to a hypervisor.

- *Home - Update Manager - Download patches and upgrades*
- *Go to Compliance View*
- *Attach*
- *Patch Baselines*
- *Critical Host Patches*
- *Non-Critical Host Patches*
- *Attach*

There are crucial steps to consider whenever patching a hypervisor.

- *Home - Update Manager - Download patches and upgrades*
- *Go to Compliance View*
- *Scan*
  Compliant Host is green. Non-Compilant is red.
- *<Focus on Non-Compilant Host>*
- *<Migrate all VM to another Host>*
- *<Enter the Maintenance Mode>*
- *Remediate – Next – Next – Next – Next – Finish*
  Installation and restart or host will take about 5 minutes.
- *<Exit the Maintenance Mode>*

6.5 vMotion and Storage vMotion

A final step should include verifying the functional migration of virtual servers between hypervisors.

- *<Focus on online VM to migrate> - Migrate*
- *Change Host - Next*
- *<Choose different Host> - Next*
- *Finish*

The migration of a virtual server's data storage may be verified by the following steps.

- *<Focus on online VM to migrate> - Migrate*
- *Change Datastore – Next*
- *<Choose different Datastore> - Next*
- *Finish*

Both of the above tasks should be possible without a VM outage. The time required to make changes to a hypervisor depend on the operating memory of the virtual server. The time required to make changes to a data storage depends on the size of a virtual server's hard disk. A data storage change usually takes some time and may take several dozen hours with large data servers.

7 Conclusion

The purpose of this document is to describe all aspects of operating a virtualisation cluster in order to use it as a manual to design other virtualisation clusters. At the moment, VMware vSphere is the best-tuned platform for virtualisation. This best practice document focuses on this most widely-used tool and on the hardware required for its operation. When choosing hardware and software tools, emphasis is placed on an optimal price/performance ratio so that the chosen software would best serve the conditions of the target environment, and not require the purchase of unnecessary virtualisation software licenses. The final result is proposal for the optimal hardware for this type of virtualisation cluster, as described in the second chapter.

As the following table shows, the migration of the original thirty physical servers to a newly-created virtualisation cluster lowers the power required to operate these systems by 77%.

Previous devices	Count	Power [W]	Total Power [W]
Server PowerEdge 1950 III (2x CPU)	15	295	4425
Server PowerEdge 2950 III (2x CPU)	15	327	4905
Summary			**9330**

After virtualisation cluster installation and migration of all physical systems into a virtual environment, a basic measurement of consumption yielded the following results.

New devices	Count	Power [W]	Total Power [W]
Storage array MD3620i (24x 2.5’’ HDD)	2	452	904
Server PowerEdge R610 (1x CPU, no HDD)	4	228	912
Switch HP ProCurve 2910al-24G	4	82	328
Summary			**2144**

The performance of the redundant virtualisation cluster significantly surpassed all expectation and clearly demonstrates the advantages of virtualisation, especially for older server systems. In addition to the energy savings, there was a savings of about a 60% in the required rack space. However, the greatest advantage is the resilience of virtual systems against hardware failures and the ability to migrate the virtual systems to other locations. Migration to other locations is important during power or temperature-control failures or due to natural disasters affecting the data centre, because it would be able to transfer the virtual servers with their storage to an unaffected location.

The system described is the result of many years of experience with VMware virtualisation at the Brno VUT campus. The maintenance of such a system is easier than dozens of individual servers and the operation of a virtualisation cluster has avoided problems associated with incompatible hardware servers.

8 References

[1] VMware vSphere 5, Licensing, Pricing and Packaging
http://www.vmware.com/files/pdf/vsphere_pricing.pdf

[2] VMware vSphere Documentation
http://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html

[3] Tom's Hardware, 3.5" Vs. 2.5" SAS HDDs: In Storage, Size Matters
Patrick Schmid, Achim Roos, May 2010
http://www.tomshardware.com/reviews/enterprise-storage-sas-hdd,2612.html

[4] Dell Enterprise HDD Specification, August 2011
http://www.dell.com/downloads/global/products/pvaul/en/enterprise-hdd-sdd-specification.pdf

[5] Configuration of HP Procurve Devices in a Campus Environment,
Tomas Podermanski, Vladimir Zahorik, March 2010 (CBPD111, the Czech Republic)
http://www.terena.org/activities/campus-bp/pdf/gn3-na3-t4-cbpd111.pdf

[6] Recommended Resilient Campus Network Design,
Tomas Podermanski, Vladimir Zahorik, March 2010 (CBPD114, the Czech Republic)
http://www.terena.org/activities/campus-bp/pdf/gn3-na3-t4-cbpd114.pdf

[7] Drivers for PowerVault MD3620i, August 2011
http://ftp.euro.dell.com/Pages/Drivers/powervault-md3620i.html

9 List of acronyms

DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
DOS	Denial-of-service Attack (DoS Attack)
GVRP	GARP VLAN Registration Protocol
GARP	Generic Attribute Registration Protocol
IOPS	Input/Output Operations Per Second
IP	Internet Protocol
iSCSI	Internet Small Computer System Interface
L2	Layer 2 - Data link layer of OSI model
L3	Layer 3 - Network layer of OSI model
OSPF	Open Shortest Path First
PSU	Power Supply Unit
RSTP	Rapid Spanning Tree Protocol
SFP	Small Form-factor Pluggable Transceiver
SM fiber	Single-mode Optical Fiber
STP	Spanning Tree Protocol
VLAN	Virtual Local Area Network
VPN	Virtual Private Network
VRRP	Virtual Router Redundancy Protocol

Complete BPDs available at www.terena.org/campus-bp/

campus-bp-announcements@terena.org