VMware Virtual Infrastructure (VI)

Open-E Data Storage Server (DSS)

Overview and Introduction

White Paper Part I

by Erich Popp for Open-E
erich.popp@storage-experts.de

www.open-e.de

TABLE OF CONTENTS

1	**Synopsis**		**4**
1.1	Feasibility		4
1.2	Performance		4
1.3	Manageability		4
1.4	Support		4
2	**Introduction**		**5**
2.1	Customer motivation		5
2.2	Virtualization with VMware products		7
2.3	VMware ESX3 and ESX3i		10
2.4	Data storage with Open-E storage servers		12
2.5	Open-E DSS		12
3	**Architecture**		**14**
3.1	VMware ESX3		14
3.2	VMware ESX3i		14
3.3	Open-E DSS		16
3.4	Infrastructure		17
4	**Preparation of the hardware and software components**		**19**
4.1	Base implementation of VMware ESX3		19
4.2	Base implementation of VMware ESX3i		19
4.3	Base implementation of Open-E DSS		19
4.4	Base implementation of infrastructure		20
5	**FCP connection**		**22**
5.1	Architecture		22
5.2	Requirements (hardware/software)		23
5.3	SAN (zoning, cabling)		23
5.4	Sizing		23
5.5	Open-E DSS		25
5.6	ESX3		28
5.7	VM storage options		31
5.8	Conclusion		31
6	**iSCSI connection**		**32**
6.1	Architecture		32
6.2	Requirements (hardware/software)		33
6.3	LAN (zoning, cabling)		33

6.4 Sizing 33
6.5 Open-E DSS 35
6.6 ESX3 38
6.7 VM storage options 43
6.8 Conclusion 43

7 NFS connection 44
7.1 Architecture 44
7.2 Requirements (hardware/software) 44
7.3 LAN (zoning, cabling) 45
7.4 Sizing 45
7.5 Open-E DSS 46
7.6 ESX3 48
7.7 VM storage options 50
7.8 Conclusion 50

8 Other sources of information 51
8.1 VMware 51
8.2 Open-E 51
8.3 Community 51

9 Indices 52
9.1 Table index 52
9.2 Figure index 53
9.3 Abbreviation index 55

10 Appendices 57
10.1 Tables 57
10.2 Documents 57

1 Synopsis

1.1 Feasibility

All three methods of connection (FCP, iSCSI and NFS) between VMware ESX3 and Open-E DSS can be implemented. VMs can use the relevant files (configuration and vDISKs) on all provided datastores. All storage-specific VMware functions (VMware VMotion, VMware DRS, VMware HA and VMware Storage VMotion) can be used.

1.2 Performance

Achievable performance essentially depends on the storage server used (RAID controller features and disk features). The connection method used is of secondary importance. If scalability (large number of VMs and parallel high-load operation) is paramount, different prioritization may result.

1.3 Manageability

The ESX3 host is either attached directly (VI client, ssh client, VI3 API tools) or attached via a VC server (VI client). The storage server can be controlled via an available server management interface (BMC redirection) and via a web interface. Open-E DSS also offers console access via ssh. The management of storage functions, e.g. snapshots, is not currently possible via Telnet, ssh or any other interface.

1.4 Support

The connection of ESX3 hosts to the Open-E DSS is currently listed in the community certified HCL.

2 Introduction

2.1 Customer motivation

Server consolidation via virtualization is on everyone's lips. Various analysts have called this IT trend one of the most important issues (http://www.virtualization.info/predictions) and predict wide distribution of virtual systems in the future. The market has since experienced very heavy growth, which means an increase in providers and that the advantages of server consolidation through virtualization have achieved a high degree of publicity.

Features and advantages

Server consolidation

Server consolidation through virtualization creates a variety of beneficial features:

- *Hardware-independence (a VM only has certain hardware components and is thus independent and can run on any physical hardware)*
- *Encapsulation (a VM is defined by a set of files on the virtualization host)*
- *Insulation (VMs are insulated from one another, even though the physical hardware is shared)*
- *Hardware load (several VMs on one host lead to a greater load to consolidation ratio)*

Data security

Essentially, VMs can be stored in very different storage systems, including local storage. The use and thus the storage of VMs on "shared storage", especially in VMware ESX3 architectures, is almost obligatory and sometimes necessary for certain functions (VMotion, HA and DRS). This "forces" the user to switch to more robust and secure storage architectures. This gives rise to the following attributes:

- *The use of storage appliances (special hardware/software with precisely custom-tailored functions)*
- *The use of any available "layered applications" (snapshot and replication functions)*
- *Very easy to use DR offer through shared storage (starting a VM anew on another host of the same datastore). The storage of many VMs in a single "shared storage" unit very quickly leads to a single point of failure, however.*

Provision of virtual systems

The provided management tools (especially of VMware) enable administrators to very quickly create new VMs (e.g. rolling out template VMs). Both traditional and new (via VM management) methods can be used to create VMs:

- *Installation of guest OS via data medium*
- *Via OS deployment mechanism (MS RIS, Altiris, NetInstall OSD etc.)*
- *By copying deactivated VMs (unique OS features generally must be reset)*
- *By cloning activated VMs with third-party tools (cloning tools)*
- *Using management functions (e.g. VMware templates)*

Agility

Virtualization enables especially quickly-growing companies to respond to IT needs, i.e. to roll out new VMs in a short period of time and to respond to other load requirements.

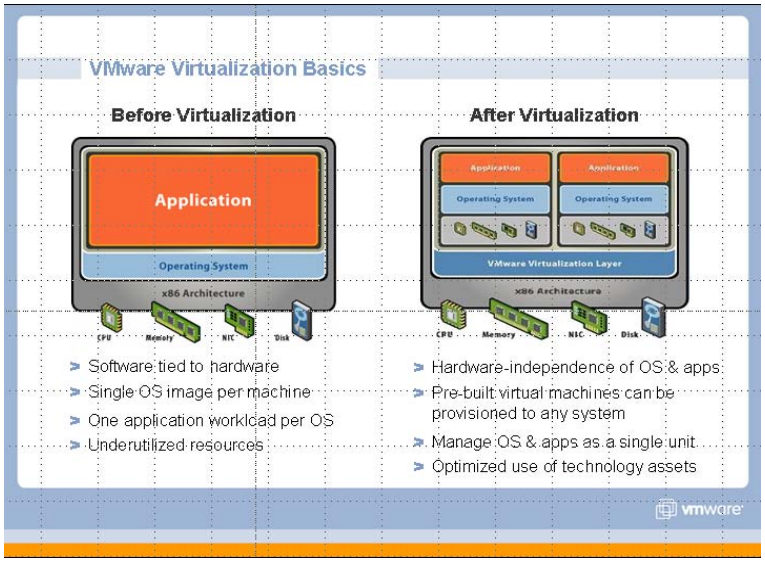
2.2 Virtualization with VMware products

*VMware makes a whole array of products available. This includes products for **virtualization**:*

- *VMware Workstation*
- *VMware Fusion*
- *VMware ACE*
- *VMware ESX3*
- *VMware ESX3i*

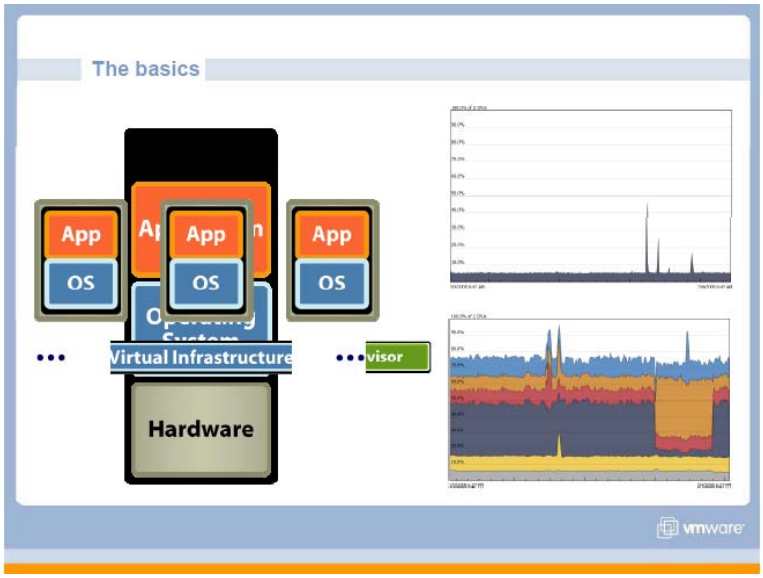
The following products are free to use:

- *VMware Player*
- *VMware Server*



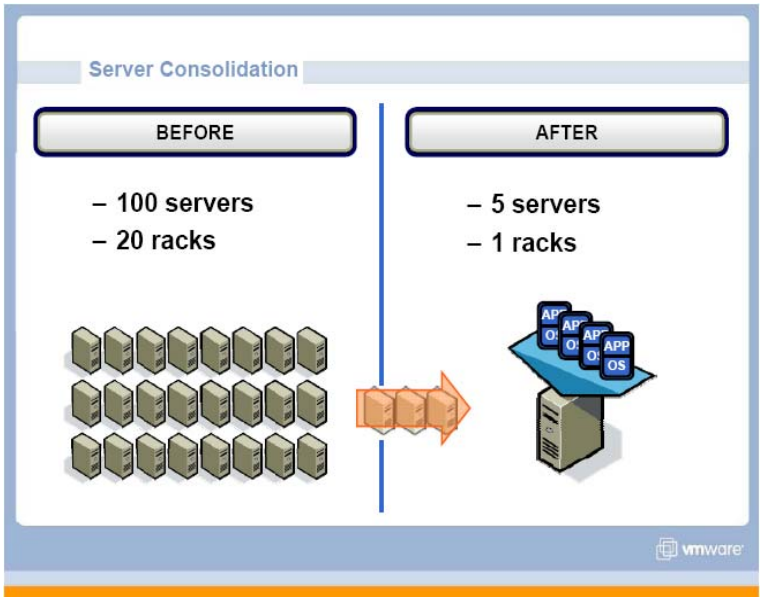
Representation of the advantages of virtualization

Figure 1 – VMware basics



Increase in the load

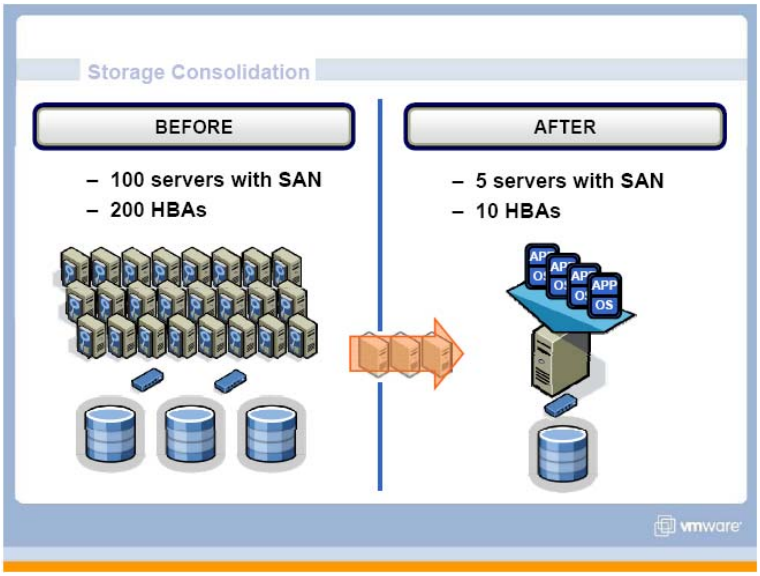
Figure 2 – Load increase



Increase in the load

Figure 3 – Load increase

Server consolidation reduces the number of physical servers. The ratio of the sum of all vCPUs to the number of host pCPUs is called the consolidation ratio. VMware runs on Quad-Core CPUs with a ratio between 2 and 4 per pCPU, i.e. with a standard 2-socket server with a Quad-Core CPU, between 16 and 32 vCPUs can be operated. With an assumed vCPU quantity of one in a VM, this corresponds to between 16 and 32 VMs.



Reduction of the host connections to the storage system

Figure 4 – Reduction of storage connections

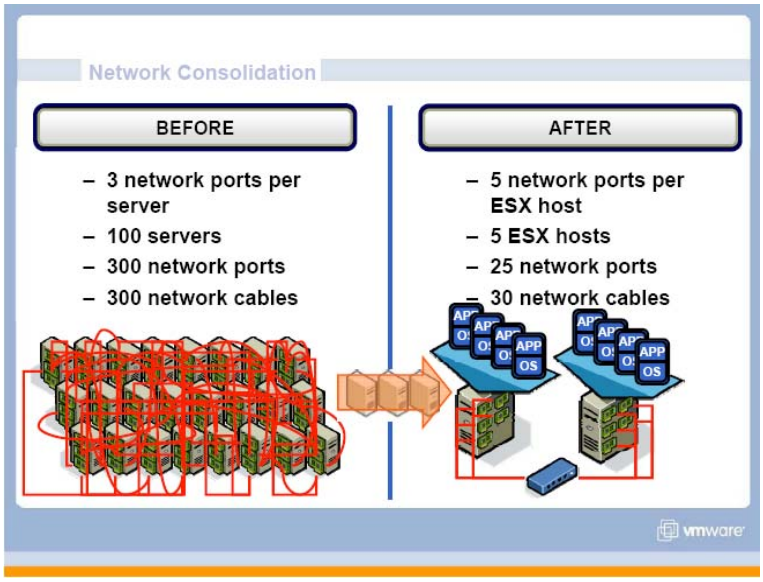
An important, but sometimes counterproductive, side effect is so-called storage consolidation. This means that the system and data disks of, for example, 100 physical servers is consolidated on just a few host disks, i.e. this can also be called storage virtualization. This pleases a great number of administrators, as only existing (available) disk space need be allocated for new VMs.

The following advantages are gained:

+ *Simple storage allocation with VMware tools*
+ *Resizing (enlargement) of VM vDisks in VI client is possible (note: partitions only grow with dynamic data mediums)*
+ *Resizing (span with so-called extends) of VMFS3 file systems*
+ *The number of required ports (FCP and GE) is heavily reduced (10–20 times)*

There are also disadvantages, however:

- *Disk load measurements on the VM level are relatively difficult to repeat on the storage system*
- *vDisk prioritization with VMFS3 is not possible on the storage system (it is possible with the so-called RDM, however)*
- *Existing server disks must be converted to vDisks (not with RDM)*



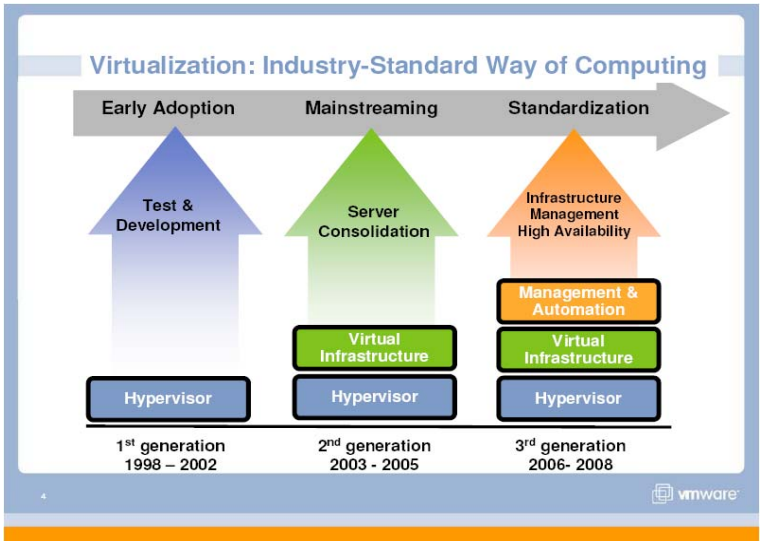
Reduction of the host connections on the network

Figure 5 – Reduction of network connections

Network consolidation has another effect. It defines the reduction in the required host connections, the additional functions (VLANs, security, redundancy and load balancing) and the common use of network resources.

Additional products and functions are primarily provided in the VI3 field and are described on the following pages.

2.3 VMware ESX3 and ESX3i



Currently in the third generation, emphasizing management and automation

Figure 6 – Product evolution

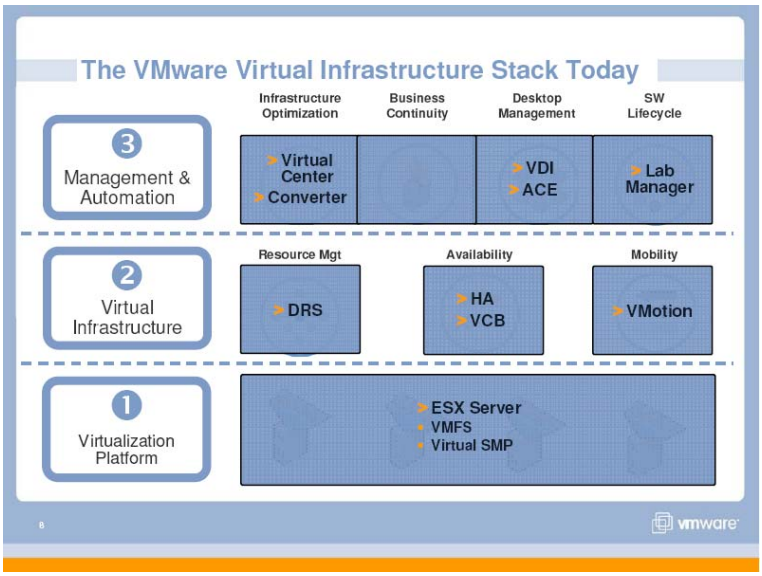
VMware ESX3 has been used in the data-processing center many times via many important functions and features.

Features

- *Extremely reliable and robust*
- *Support for many x86 servers, storage systems and guest operating systems*
- *Very high scalability (32 logical CPUs, 128 GB RAM, 32 Intel GEs, 20 Intel FEs, 20 Broadcom GEs, 64 TB VMFS3 LUNs, 32 hosts in a cluster, 32 hosts on VMFS3)*
- *Open architecture with documented API for third-party tools*
- *Very good management environment with VC server*
- *Additional new features with ESX3.5 (TSO, JumboFrames, 10 GEs and NPIV)*

Functions

- *VMFS (VMware file system)*
- *Virtual SMP (1, 2 or 4 vCPU in a VM)*
- *DRS (VM load balancing)*
- *HA (VM restart after host failure)*
- *VCB (consolidated backup solution)*
- *VMotion (live migration of VMs between ESX3 hosts)*
- *Storage VMotion (VMotion with storage migration)*
- *Converter (P2V and V2V)*
- *Virtual Center (centralized management solution)*
- *Update Manager (patching of ESX3/3i hosts and VMs)*
- *VDI/VDM (desktop hosting with management functions)*
- *LabManager/StageManager (solution for VM workflow management)*



VI software stack (version 3)

Figure 7 – VI software stack (version 3)

A detailed presentation of these functions is not part of this white paper.

2.4 Data storage with Open-E storage servers

All Open-E products are based on an operating system that is completely pre-installed and configured on a USB DOM (Disk On Module). This module is then simply inserted into the server system. Pre-installation of the operating system on the USB DOM enables companies to build up their storage infrastructure very quickly and easily. Downtimes are also reduced considerably, and the storage installations are highly secure thanks to the encapsulated operating system.

Two versions of the OS are present, as with dual bios, where the previous version of the OS is still available and can be called up if necessary. The Open-E storage solutions detect the hardware automatically and install the required drivers of the SAS and RAID controllers, Fiber Channel HBAs, Ethernet cards etc. automatically.

2.5 Open-E DSS

The Open-E DSS (Data Storage Server) is a complete IP and FC (Fiber Channel) storage operating system that offers companies of any size NAS, iSCSI and FC functions (target and initiator) in a separate application that is easy to use and that is highly failure-proof. Data Storage Servers provide a fast, reliable and scalable platform for IP storage that enables common file access, memory consolidation and backup and restoration. Virtualization, or replication, is another highly relevant area these days and is covered by DDS.

Since the operating system has been optimized for environments with dedicated storage and company networks, the solution is especially suitable for network environments with many clients or applications with a high demand for storage, such as multi-video streaming, HDTV and others that also require high data throughput and I/O. Support for Windows, NIS and LDAP domains enables trouble-free expansion of the existing IT infrastructure. The web-supported graphical user interface for management and administration allows reliable control of the storage unit and the backup process for backing up important data.

NAS components

NAS systems (Network Attached Storage) are data servers through which the network can be accessed. NAS enables multiple users to use the same storage space at the same time, whereby the overhead is often minimized through centralized management of the hard disks and energy costs can be reduced considerably through centralized data management and efficient use of resources.

The Open-E DSS software provides simple centralized management, variable scalability, high reliability, availability and business-oriented performance. The software is an optimized operating system with which an NAS solution can be quickly and simply set up on a server. Pre-installation of the operating system on the USB DOM ensures minimum downtime and protection of the operating system from viruses.

Solution highlights include the support of an integrated anti-virus application for the protection of drives and a variety of backup and snapshot functions for easy data backup. The system can also be easily integrated into existing infrastructures and enables the authentication of the user via Active Directory etc.

iSCSI components

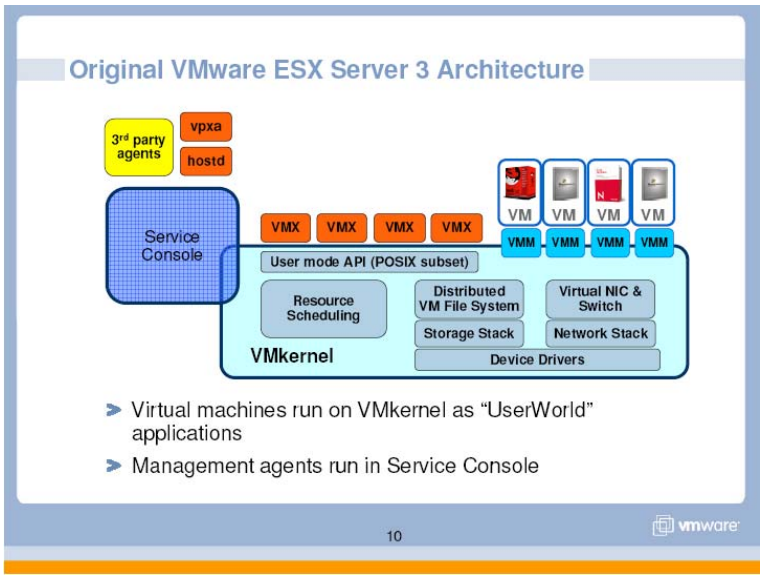
The new iSCSI standard is currently a hot candidate for domination in the storage solution arena. iSCSI is an IP-supported storage network standard for remote storage management. iSCSI is used to link data storage equipment and to simplify data transmission via the intranet and Internet. The iSCSI Target Modules (ITMs) are managed like local hard disks by the administrators. Due to the high availability and performance of IP networks, iSCSI enables site-independent storage and retrieval of data.

Open-E DSS enables cost-effective, centralized storage without the usual expenditures and incompatibilities normally involved with FC SANs (Storage Area Networks). This makes particular sense, since solutions that enable fast and efficient I/O data transmission on the block level can be implemented in this way. This makes the solution optimum for databases.

In contrast to an FC SAN, iSCSI can be executed via existing Ethernet networks and enables easy migration and administration, since existing knowledge on Ethernet and iSCSI can be used. Open-E software belongs to a new generation of applications that expand the possibilities and performance of storage data transmission on the SAN market. It is an optimum combination of data throughput, security, compatibility, management, cost-effectiveness and flexibility.

3 Architecture

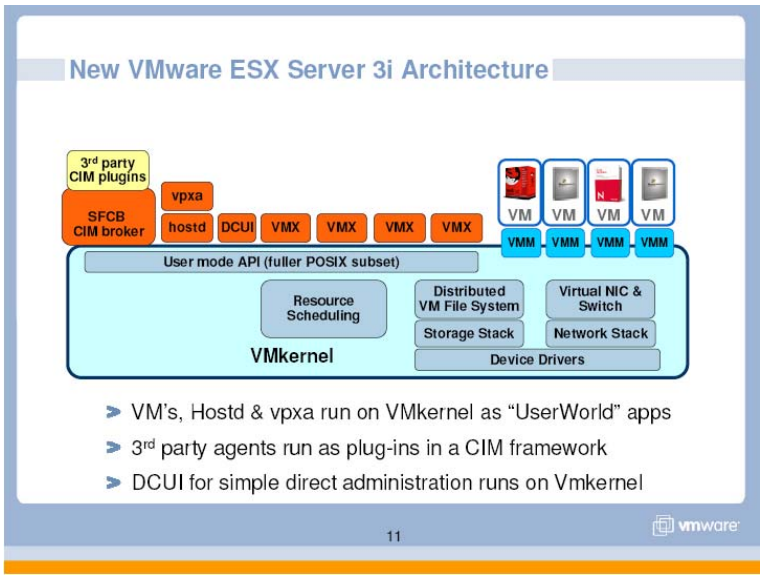
3.1 VMware ESX3



VMware ESX3 architecture

Figure 8 – ESX3 architecture

3.2 VMware ESX3i

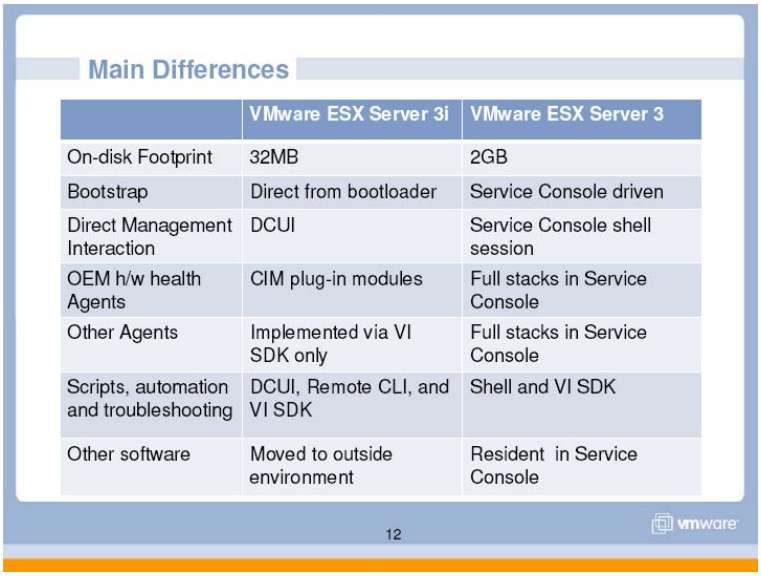


VMware ESX3 architecture

Figure 9 – ESX3i architecture

The architecture of both ESX3 host platforms essentially consists of the following components:

- *VMkernel (hypervisor for specific server models, see HCL)*
- *Resource scheduler (CPU, RAM, DISK and LAN)*
- *Storage stack (multipathing with redundancy and load balancing)*
- *Network stack (VLAN support, security, redundancy and load balancing)*
- *Support for distributed VMFS3 (up to 32 hosts can actively access a VMFS3 LUN)*
- *Support for VM worlds*
- *Service Console with RedHat Linux (ESX3 only)*
- *Support for host and VC agents*
- *Support for third party agents (hardware management agents)*
- *Support for web access (ESX3 only)*
- *Support for API access*



Main differences between ESX3 and ESX3i

Figure 10 – Main differences

3.3 Open-E DSS

The architecture of the Open-E Data Storage Server software can be divided into several main parts:

- *Storage level*
- *Storage virtualization level*
- *NAS and SAN*
- *Network level*

Storage level

The core of every storage server is the storage areas. In general, the Open-E Data Storage System can use either Direct Attached Storage (DAS; hard disks connected directly to the system) or external storage (iSCSI matrices connected via iSCSI initiator or Fiber Channel). Here, a DSS can be the storage of another system based on the FC and iSCSI targets.

Storage virtualization level

This level introduces two abstraction levels between the core storage and the operating system, volumes and logical volumes. This allows users to create, delete and expand partitions in NAS, iSCSI and FC operation without the unnecessary demand for downtime or management expenditure. Another important point is that capacity can also be expanded during running operation without deactivation of the production system if the storage systems no longer have available space and the hardware platform used is hot-plug-capable.

NAS and SAN level

The NAS and iSCSI level lies above the virtual storage level. This layer is transparent to users and functions as a bridge between their data and the storage medium. It enables users to export their data and also provides a variety of security options. Open-E DSS offers users asynchronous data replication on the data level, synchronous block level volume replication and integrated backup utilities that represents a complete storage server for companies of any size.

In addition to other options for data backup on the software level, Open-E supports the DSS hardware RAID controller, the teaming (bonding) of network cards (NIC; Network Interface Controllers), UPSs (Uninterruptible Power Supply) and network UPSs. Support for Microsoft domain manager (PDS/ADS) enables administrators to integrate Open-E DSS into existing heterogeneous architectures.

All features listed above are easily accessible and are configured via a web-based, password-protected (Secure Socket Layer, SSL) GUI on the administrator level.

3.4 Infrastructure

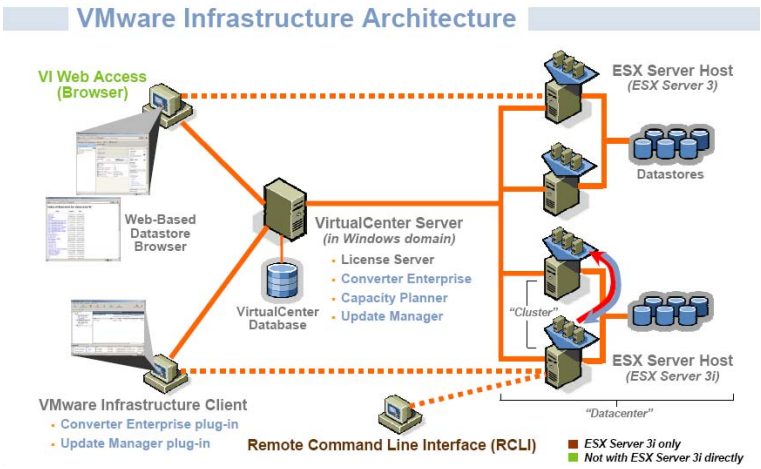


Figure 11 – VI architecture

VI architecture with all important components:

- *VC server*
- *VI client*
- *ESX3 hosts*

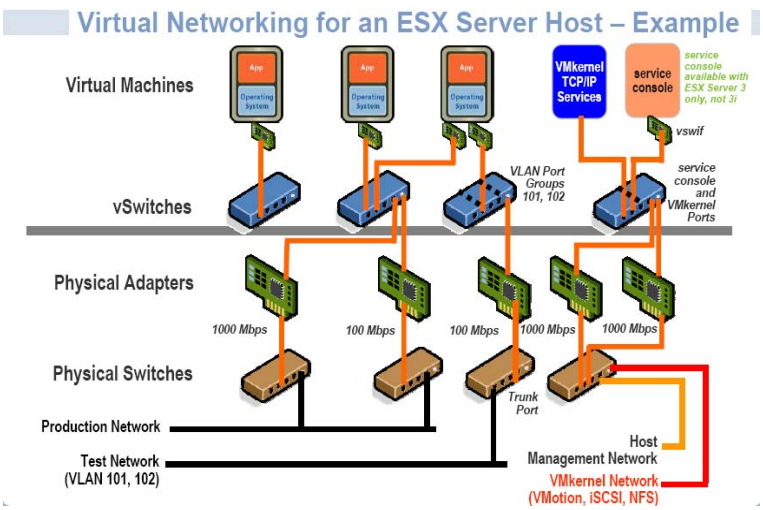


Figure 12 – VI LAN architecture

VI LAN architecture with all important components:

- *pSwitches*
- *pNICs*
- *vSwitches*
- *Portgroups*
- *vNICs*
- *VLANs*
- *Server Console port*
- *VMkernel port (VMotion, IP storage)*

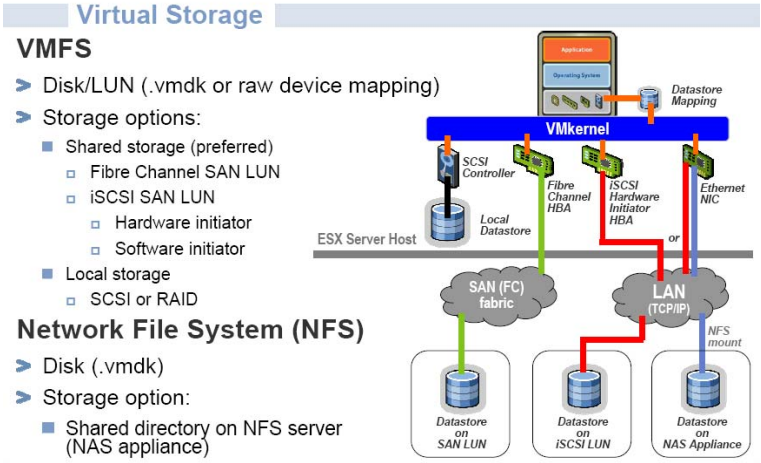


Figure 13 – VI storage architecture

VI storage architecture with all important options:

- *FCP (VMFS or RDM)*
- *iSCSI (VMFS or RDM)*
- *NFS*

Provision of the infrastructure is one of the most important points for the operation of a virtual environment.

Areas of architecture, such as:

- *Management architecture (services, ports etc.)*
- *LAN architecture (VM port groups)*

are outside the focus of this white paper.

In addition, the storage architecture is delved into deeper, as are parts of the LAN architecture for IP storage.

4 Preparation of the hardware and software components

4.1 Base implementation of VMware ESX3

The following is involved in planning a VMware ESX3 implementation:

- *Compatibility guides (HCL, approval for x86 systems and I/O adapters)*
- *Component updates (BIOS, RAID HBA, FC HBA and iSCSI HBA)*
- *PCI configuration (prevention of PCI interrupt sharing)*
- *BIOS configuration (NX/XD, VT/AMD-V, time, APIC mode)*
- *Checklist (keyboard table, boot disk in SAN or local, partitioning boot disk, system information and password)*
- *Post configuration (ntp, security, storage, network and data center)*

The installation of an ESX3 host is comparable to RedHat Linux, Fedora and CentOS installations and has been documented many times. It can be carried out in about 15 minutes and is guided by the user.

The ESX3 host can also be installed on a SAN disk (FCP LUN and iSCSI LUN).

4.2 Base implementation of VMware ESX3i

The following is involved in planning a VMware ESX3i implementation:

- *Compatibility guides (HCL, approval for x86 systems and I/O adapters)*
- *Component updates (BIOS, RAID HBA, FC HBA and iSCSI HBA)*
- *PCI configuration (prevention of PCI interrupt sharing)*
- *BIOS configuration (NX/XD, VT/AMD-V, time, APIC mode)*
- *Post configuration (ntp, security, storage, network and data center)*

An ESX3i host can be commissioned in two different ways:

- *Embedded (included as the boot medium by the server manufacturer)*
- *Installable (can be installed from the medium, approx. 5 minutes)*

4.3 Base implementation of Open-E DSS

The following is involved in planning an Open-E DSS implementation:

- *Compatibility guides (HCL, no limitation for x86 systems, approval for I/O adapters)*
- *Component updates (BIOS, RAID HBA, FC HBA and iSCSI HBA)*
- *PCI configuration (prevention of PCI interrupt sharing)*
- *BIOS configuration (NX/XD, VT/AMD-V, time, APIC mode)*
- *RAID controller configuration (hardware RAID, software RAID)*
- *Checklist (system information)*
- *Post configuration (network, storage and management)*

The commissioning of Open-E DSS is very simple and easy to carry out. All x86 servers are supported; for storage adapters, the compatibility guide is essential. The Open-E DSS module can be connected to either an available USB interface in or on the server. It features a write-protection switch that can be activated once the configuration is complete.

4.4 Base implementation of infrastructure

The following are part of the infrastructure (some are optional):

- *LAN*
- *SAN*
- *VC server*
- *VI client*

Certain preliminary work is required, depending on the connection variant:

FCP

- *LAN*
  VI management
  Server management
  Storage management
  SAN management
  VM networks
  VMotion network
  VMware HA heartbeat network
  These LANs can be implemented physically and/or via VLAN. With VMs, redundant connection to two switches is recommended.
- *SAN*
  FC switches

NAS (iSCSI/NFS)

- *LAN*
  VI management
  Server management
  Storage management
  Storage network
  VM networks
  VMotion network
  VMware HA heartbeat network
  These LANs can be implemented physically and/or via VLAN. Redundant connection to two switches is recommended for the VM network.
  Redundant connection to two switches is recommended for the storage network.

VC server

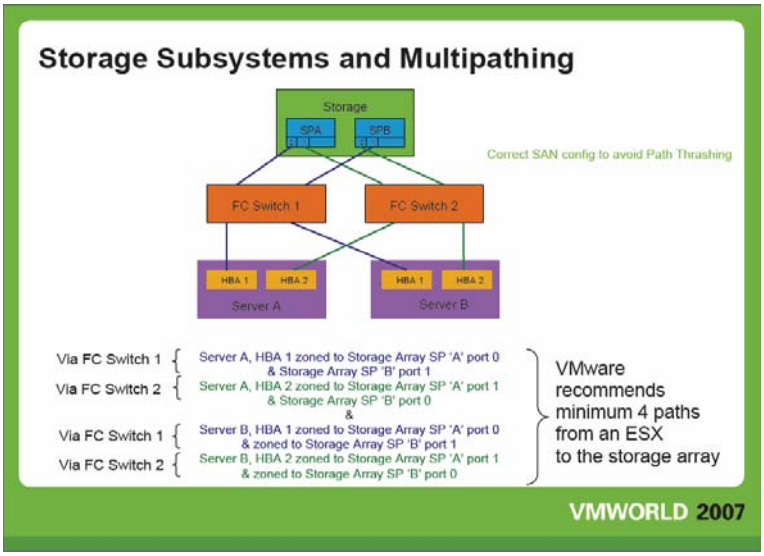
- *Consideration of the minimum requirements (see release notes)*
- *Physical installation or installation as VM*
- *Operation as a windows AD member or standalone*
- *Selection of functions (license server, update manager, converter, consolidation)*
  We recommend installing all functions on one system

VI client

- *Consideration of the minimum requirements (see release notes)*
- *Installation of ESX3 host or VC server homepage (via browser)*

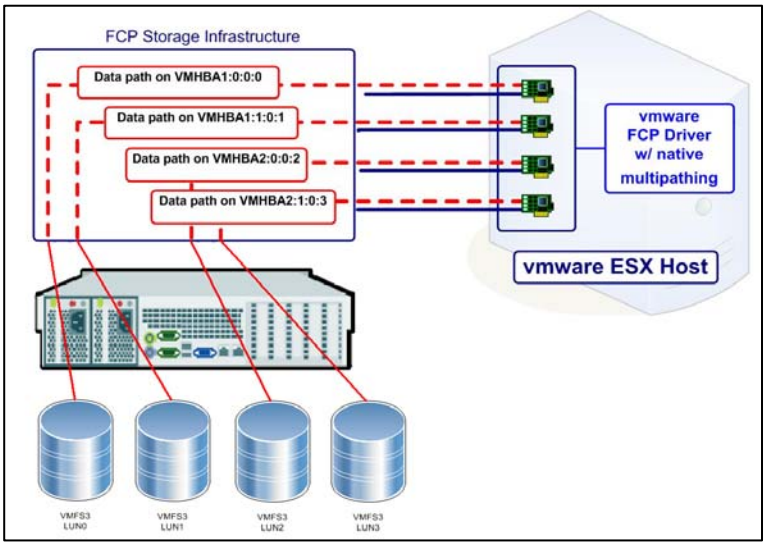
5 FCP connection

5.1 Architecture



Typical VI SAN architecture

Figure 14 – Typical SAN architecture



FCP access paths:

Native multipathing via the VMkernel

One path per LUN

Bandwidth aggregation

Figure 15 – FCP access paths

5.2 Requirements (hardware/software)

Open-E DSS:

- *Two FC HBA (Qlogic FCP target mode)*
- *Current firmware/BIOS version*

SAN:

- *Two FC switches with GBICs*
- *Cabling*
- *Management and function licenses (NPIV?)*
- *Current Fabric OS version*
- *Fabric connection is mandatory for ESX3*

ESX3/ESX3i host:

- *Two FC HBA (Qlogic or Emulex)*
- *Current firmware/BIOS version*

5.3 SAN (zoning, cabling)

Zoning:

- *Port or WWPN zoning (with port zoning, the target paths can be better defined)*
- *One initiator and one target per zone*

Cabling:

- *Always connect cabling from the low target ports to low switch ports (zoning)*

5.4 Sizing

The following attributes must be taken into account for proper planning of the required LUN sizes:

VMFS3:

- *Minimum size: 1,200 MB*
- *Metadata overhead: approx. 600 MB*
- *Maximum size (one extent): 2 TB*
- *Maximum size (32 extents): 64 TB*
- *Maximum: 32 hosts*

Physical server LUNs:

- *Number and size*
- *Shared SCSI cluster*

VMs:

- *Small files (configuration files, log files and NVRAM file) up to a certain size are stored in sub-blocks*
- *vDisk files are created pre-allocated in the standard (fixed reservation of the configured disk space)*
- *Delta files (VMware snapshots, multiple) grow in 16 MB increments up to the vDisk size*
- *Memory snapshots are as large as the VM storage configuration*
- *VMkernel swap file is as large as the VM storage configuration*
- *RDM files are 1 MB in size*

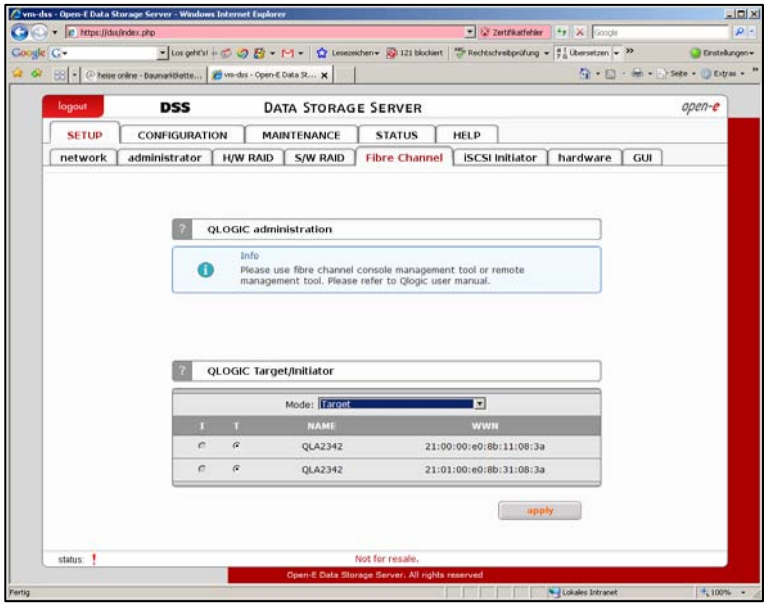
In addition, decisions must be made with regard to:

- *Large LUNs (many VMs) and small LUNs (few VMs)*
  Performance (LUN with many VMs <-> few VMs)
  Backup/restore (large LUN <-> small LUN)
  Storage snapshot function (LUN with many VMs <-> one VM)
- *Use of existing LUNs with RDM*
- *Replication relationships with other storage servers (LUN with many VMs <-> one VM)*
- *Number of possible paths (load balancing: MRU, fixed and RR)*

To sum up, one could say that a compromise must be found between management expenditure, performance and restoration time.

Some best-practice values have already been mentioned. There is a large range involved here, however. Oftentimes, between 20 and 40 VMs are stored on a VMFS3 LUN and thus the LUN generates between 300 and 600 GB.

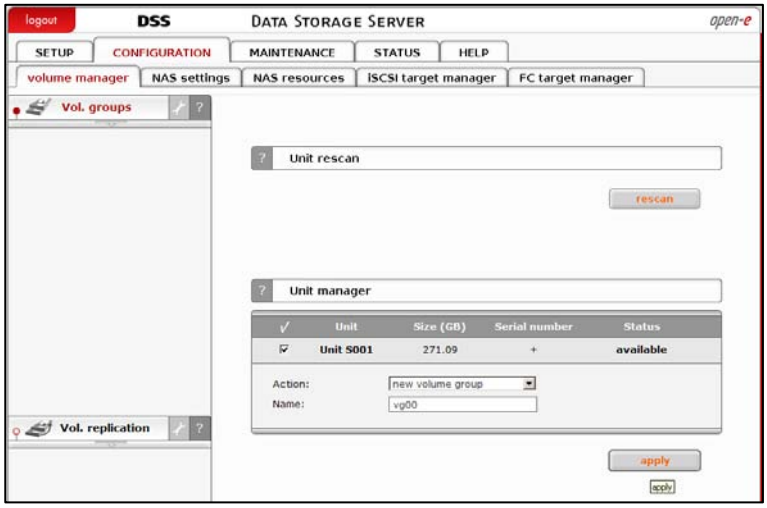
5.5 Open-E DSS



Configuration of Qlogic target mode

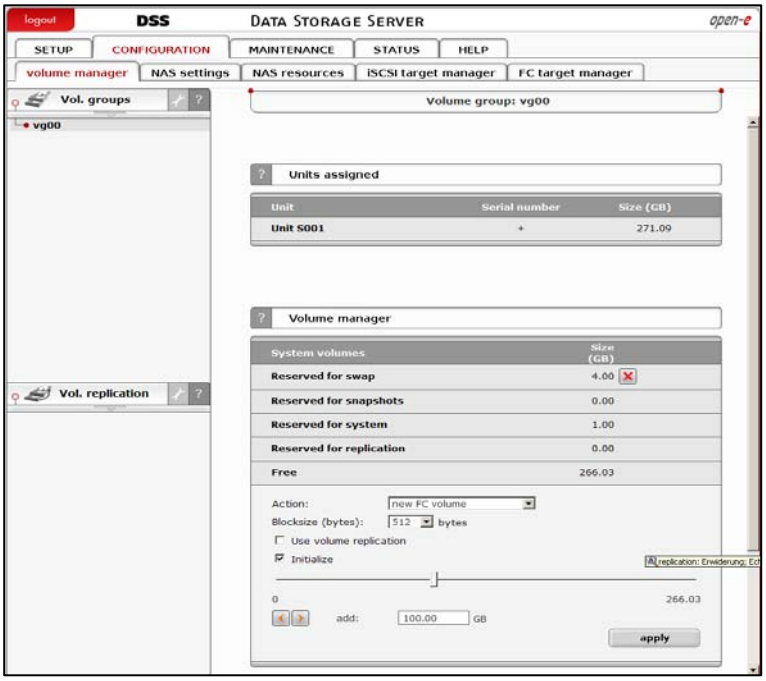
A reboot is then necessary

Figure 16 – Configuration of Qlogic target mode



Configuration of volume group

Figure 17 – Configuration of volume group



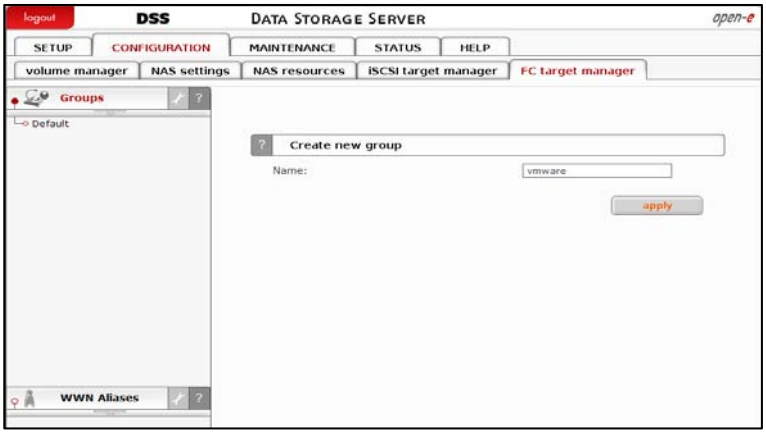
Configuration of FC volume (LUN)

- New FC volume

- Block size: 512 bytes

Initialization may take some time

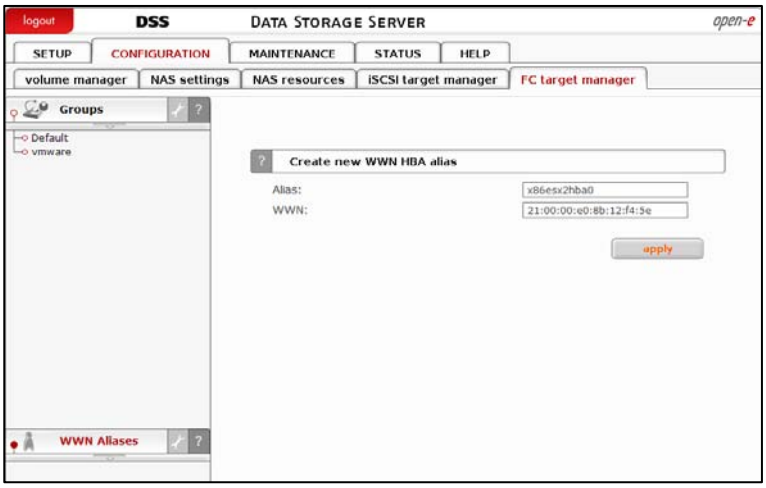
Figure 18 – Configuration of FC volume (LUN)



Configuration of FC target group

Any name can be selected

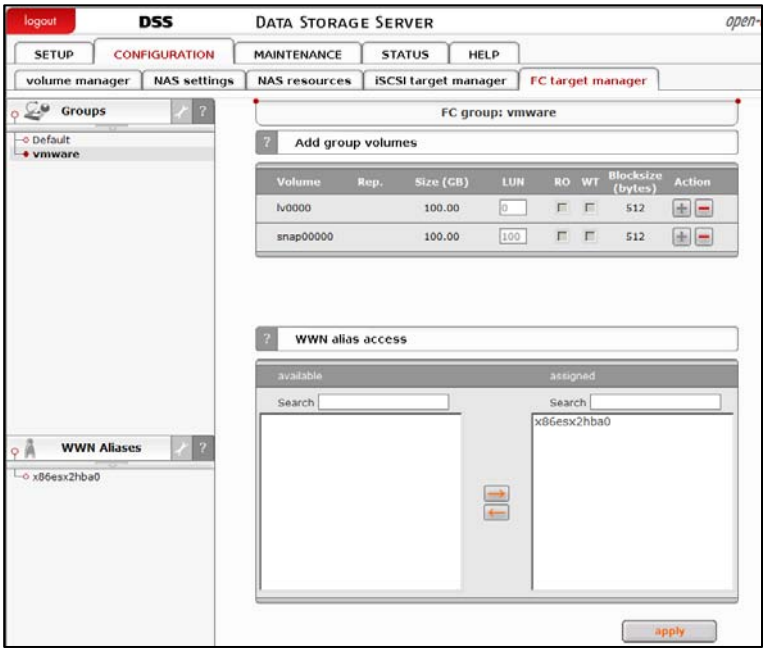
Figure 19 – Configuration of FC target group



Configuration of WWPN HBA alias

Aliases are defined for the host HBA here

Figure 20 – Configuration of WWPN HBA alias

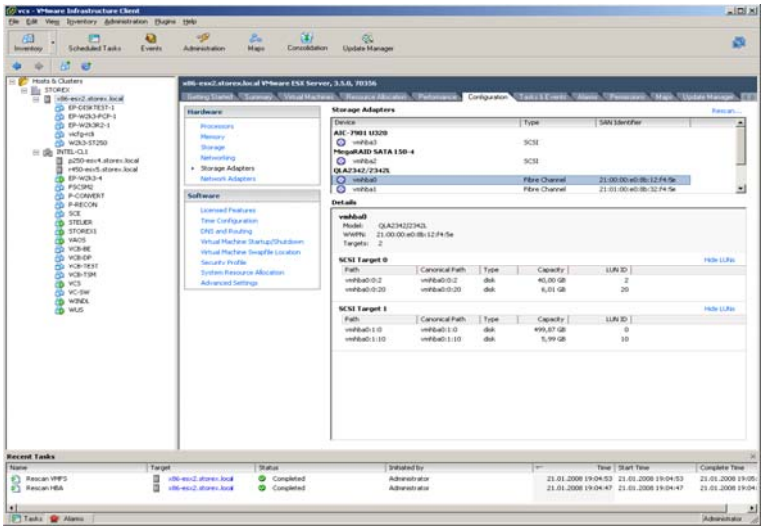


Configuration of volume mapping

Mapping of existing volumes (LUNs) with LUN ID (e.g. 0) and properties (RW and WB) on a target group

Figure 21 – Configuration of volume mapping

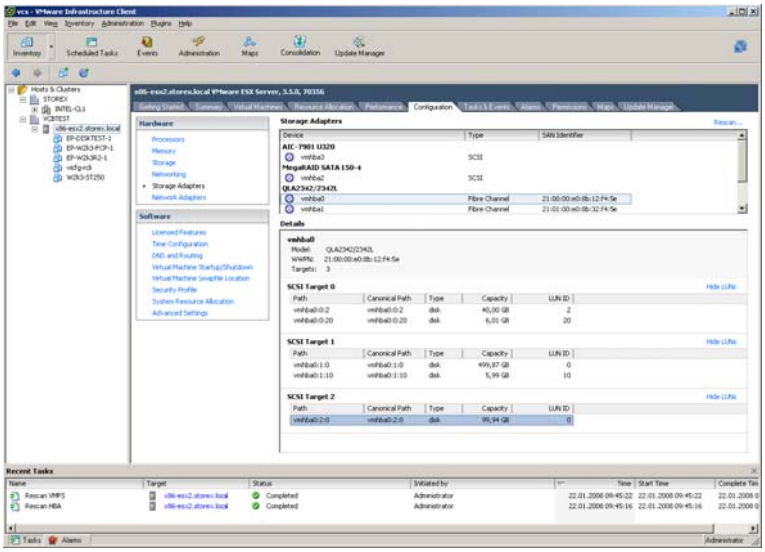
5.6 ESX3



After logon with VI client at VC serve.

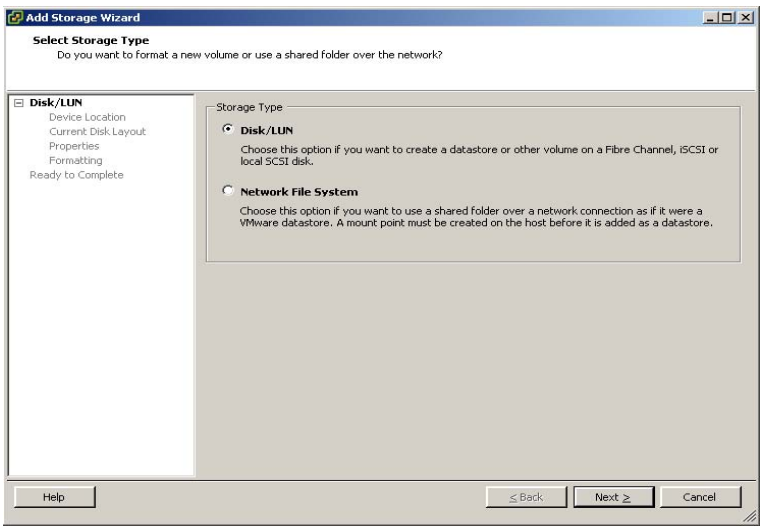
Selection of storage adapter

Figure 22 – Configuration -> Storage Adapters -> Before Rescan



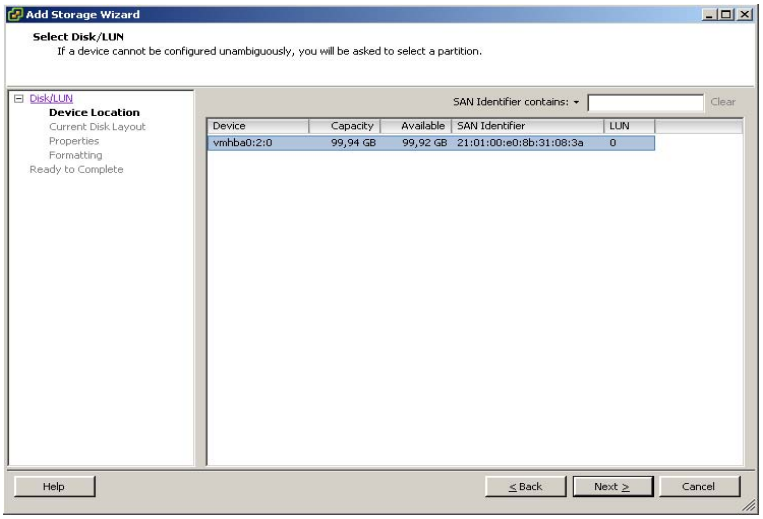
After the rescan, the LUN created on the DSS is displayed and can be formatted with VMFS3

Figure 23 – Configuration -> Storage Adapters -> After Rescan



Formatting of an LUN with VMFS3

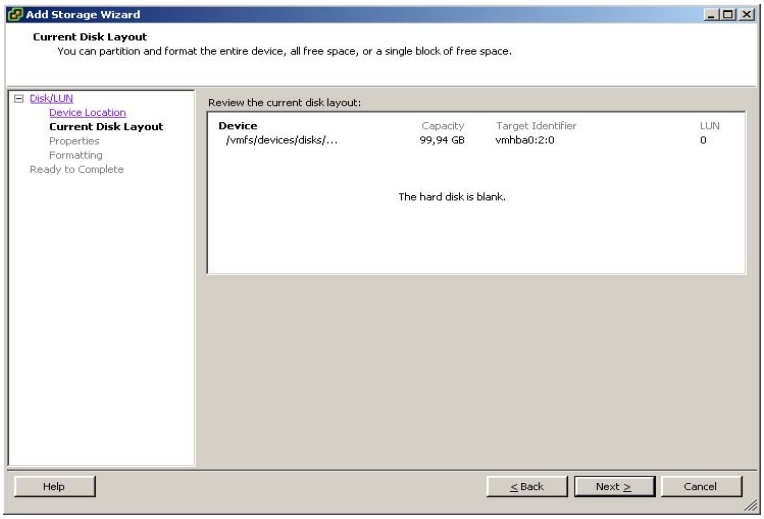
Figure 24 – Configuration -> Storage -> Add Storage -> Disk/LUN



Selection of device

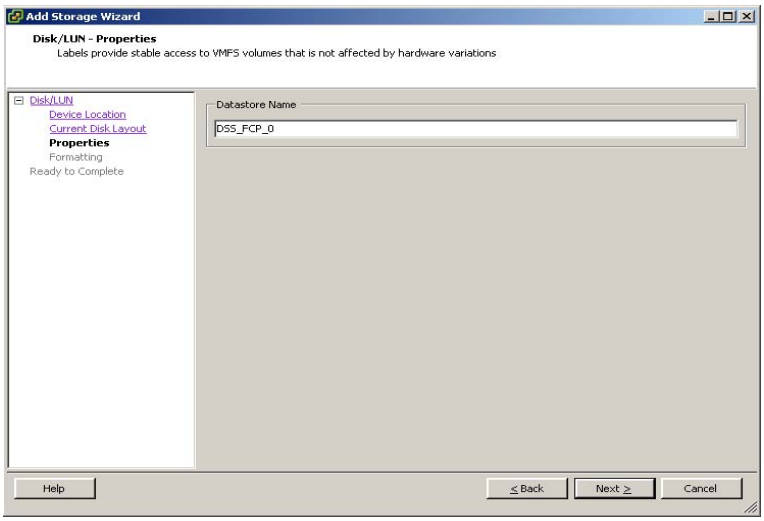
SAN identifier = WWPN target HBA

Figure 25 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Device Location



The disk is not partitioned

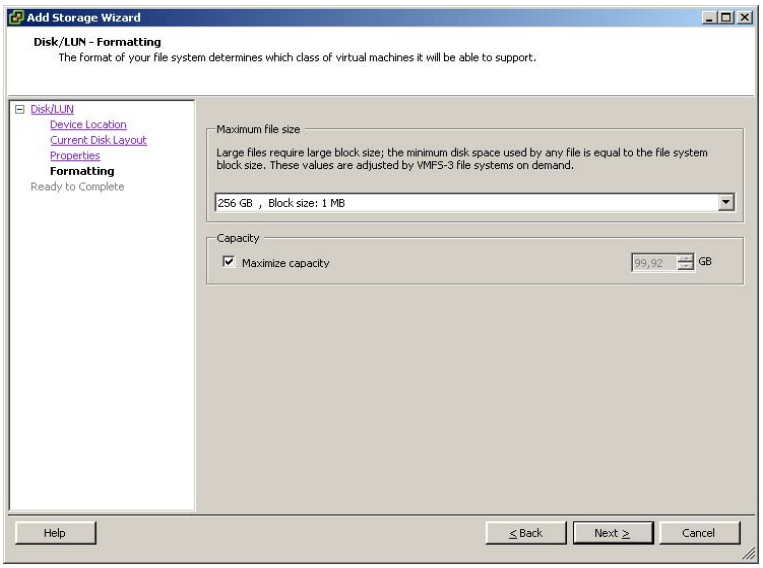
Figure 26 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Current Disk Layout



Datastore name

E.g. Storagename_-Protocol_No

Figure 27 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Datastore Name



Maximum size of a file (vDisk) in a VMFS3:

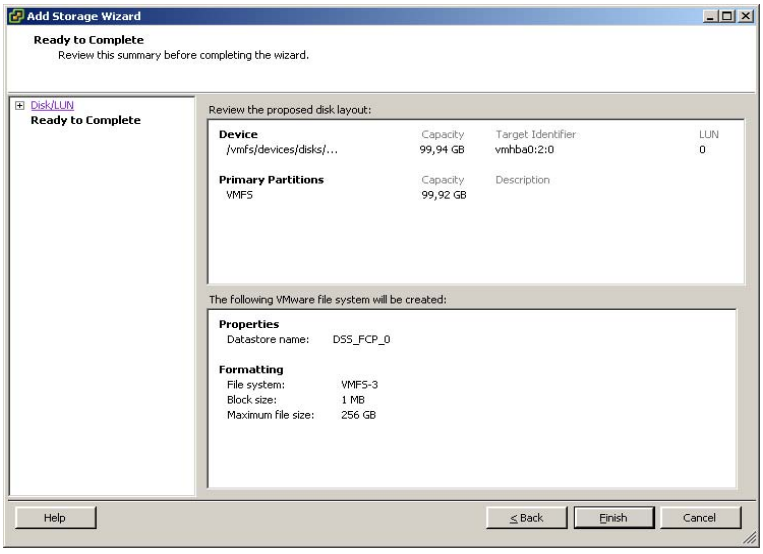
256 GB -> 1 MB

512 GB -> 2 MB

1,024 GB -> 4 MB

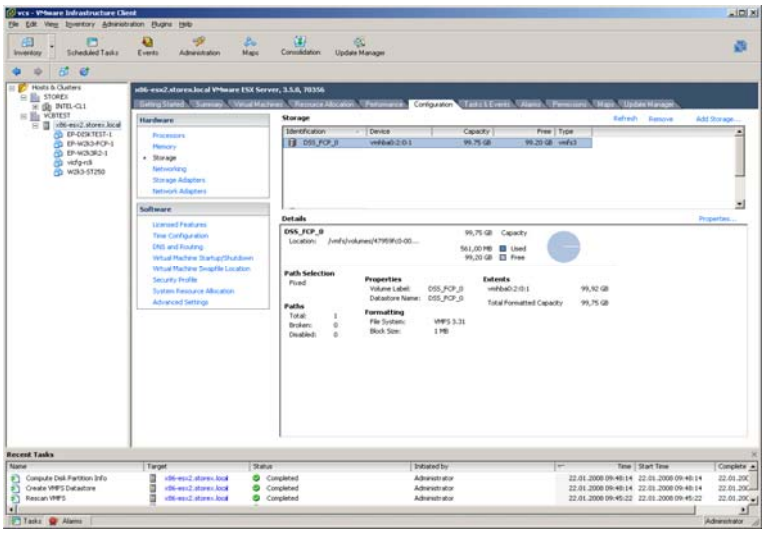
2,048 GB -> 8 MB

Figure 28 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Maximum File Size



Review

Figure 29 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Review



That's it, VMs can now be stored on this datastore

Figure 30 – Configuration -> Storage

5.7 VM storage options

VMs can be stored on an FCP datastore in the following variants:

vDisk as a VMFS3 file

+ *Very easily, set hard disk sizes in the VM configuration are transferred to vDisk file sizes one to one*
+ *VMFS3 file-level locking (by default a vDisk can only be used by one VM)*
+ *vDisk size can be changed in the VI client or via CLI*
+ *Is generated with clone and template deploy processes*
+ *Many VMs (vDisks) are possible*
- *VMFS3 datastores can be relatively easily overbooked (I/O load)*
- *Monitoring of the disk load (I/O load and reservation) is recommended*
- *Operation of storage software (e.g. in-band snapshot control) in the VM is not possible*

vDisk as RDM

+ *Each vDisk has its own LUN*
+ *LUN traits can be controlled via the storage system (RAID level etc.)*
- *These vDisks cannot be created via cloning or template deployment (but CLI or third-party tools -> white paper II is possible)*

5.8 Conclusion

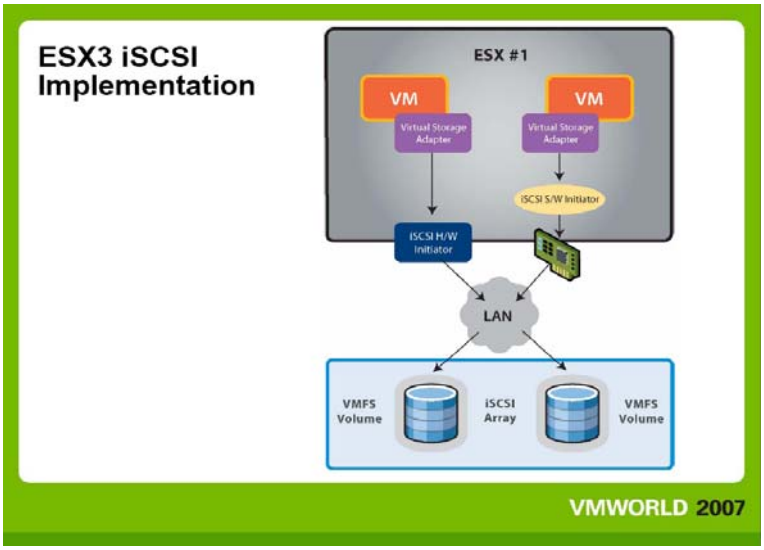
The FCP connection method is very easy to implement. Since this variant is not listed in the VMware compatibility guides, responsibility lies with the owner. The following points should be carefully planned and checked before use:

- *The number of ESX3 hosts to be connected to Open-E DSS*
- *Support of LUN reservation via ESX3 host*

Testing revealed this connection to be absolutely stable.

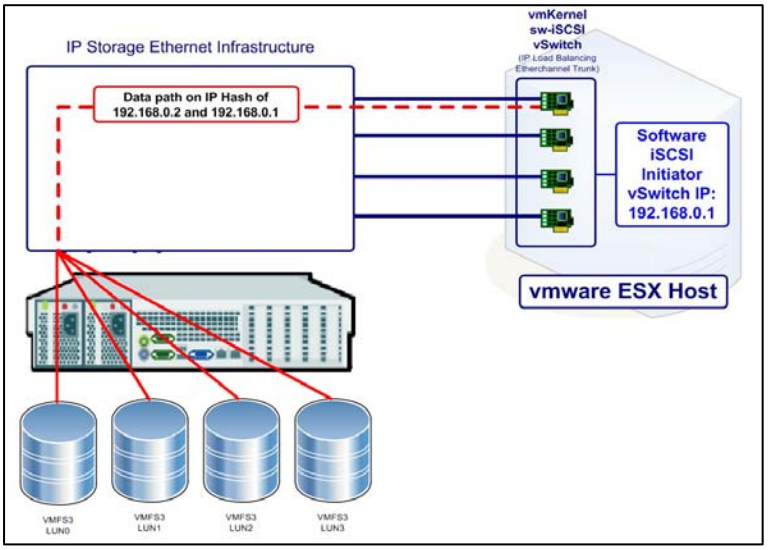
6 iSCSI connection

6.1 Architecture



Representation of both iSCSI implementation variants

Figure 31 – iSCSI architecture



iSCSI access paths:

Redundancy through VMkernel

One path for all LUNs

No bandwidth aggregation

Figure 32 – iSCSI access paths

6.2 Requirements (hardware/software)

Open-E DSS:

- *Two GE ports*
- *Current firmware/BIOS version*

LAN:

- *Two GE switches*
- *Cabling*
- *Management and function licenses (VLAN)*
- *Current OS version*
- *Necessary for ESX3*

ESX3/ESX3i host:

- *Two iSCSI HBAs (Qlogic) (cannot be combined with software iSCSI)*
- *Two GE ports for software iSCSI (cannot be combined with iSCSI HBA)*
- *Current firmware/BIOS version*

6.3 LAN (zoning, cabling)

Zoning:

- *If a dedicated IP storage LAN is not feasible, segmenting can be implemented with a VLAN*

Cabling:

- *Fixed port settings for GE ports*

6.4 Sizing

The following attributes must be taken into account for proper planning of the required LUN sizes:

VMFS3:

- *Minimum size: 1,200 MB*
- *Metadata overhead: approx. 600 MB*
- *Maximum size (1 extent): 2 TB*
- *Maximum size (32 extents): 64 TB*
- *Maximum: 32 hosts*

Physical server LUNs:

- *Number and size*

VMs:

- *Small files (configuration files, log files and NVRAM file) up to a certain size are stored in sub-blocks*
- *vDisk files are created pre-allocated in the standard (fixed reservation of the configured disk space)*
- *Delta files (VMware snapshots, multiple) grow in 16 MB increments up to the vDisk size*
- *Memory snapshots are as large as the VM storage configuration*
- *VMkernel swap file is as large as the VM storage configuration*
- *RDM files are 1 MB in size*

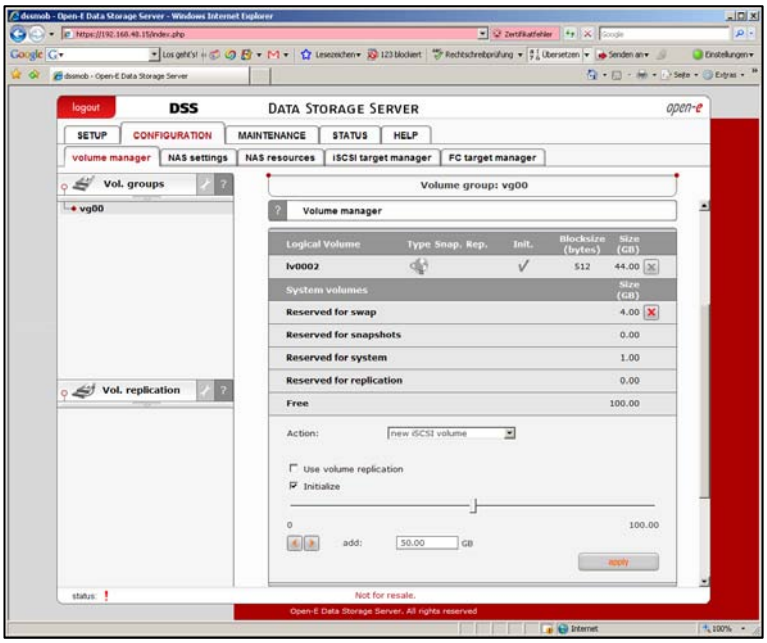
In addition, decisions must be made with regard to:

- *Large LUNs (many VMs) and small LUNs (few VMs)*
  Performance (LUN with many VMs <-> few VMs)
  Backup/restore (large LUN <-> small LUN)
  Storage snapshot function (LUN with many VMs <-> one VM)
- *Use of existing LUNs with RDM*
- *Replication relationships with other storage servers (LUN with many VMs <-> one VM)*
- *The number of possible paths (only with iSCSI HBA, load balancing: MRU, fixed and RR)*

To sum up, one could say that a compromise must be found between management expenditure, performance and restoration time.

Some best-practice values have already been mentioned. There is a large range involved here, however. Oftentimes, between 10 and 20 VMs are stored on a VMFS3 LUN and thus the LUN generates between 150 and 300 GB.

6.5 Open-E DSS

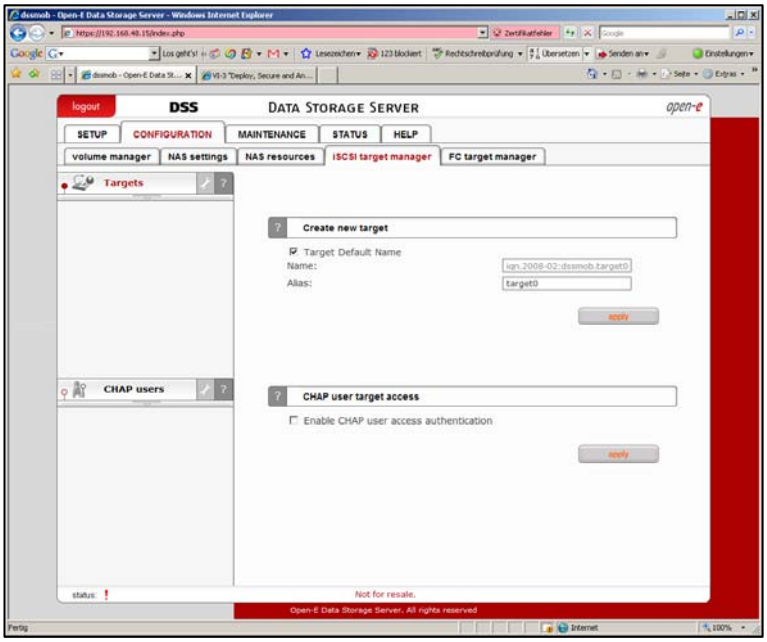


Configuration of iSCSI volume (LUN)

- New iSCSI volume

Initialization may take some time. The volume group was already created in the FCP part

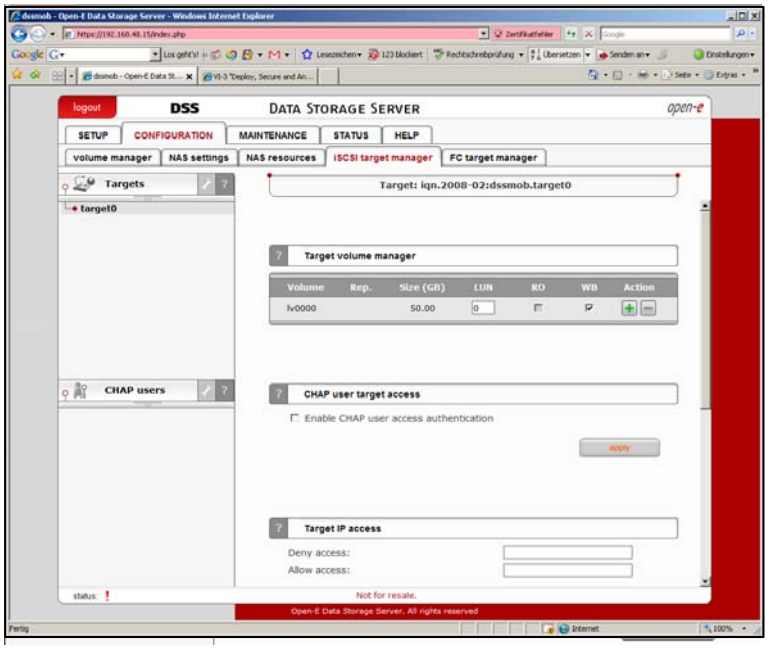
Figure 33 – Configuration of iSCSI volume (LUN)



Configuration of iSCSI target

Any name can be selected

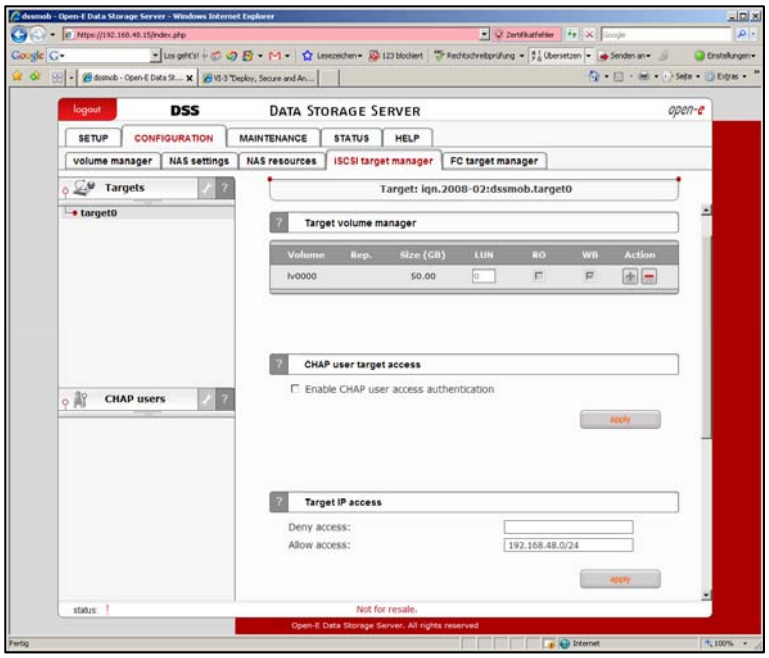
Figure 34 – Configuration of iSCSI target



Configuration of volume mapping

Mapping of existing volumes (LUNs) with LUN ID (e.g. 0) and properties (RW and WB) on a target

Figure 35 – Configuration of iSCSI target mapping

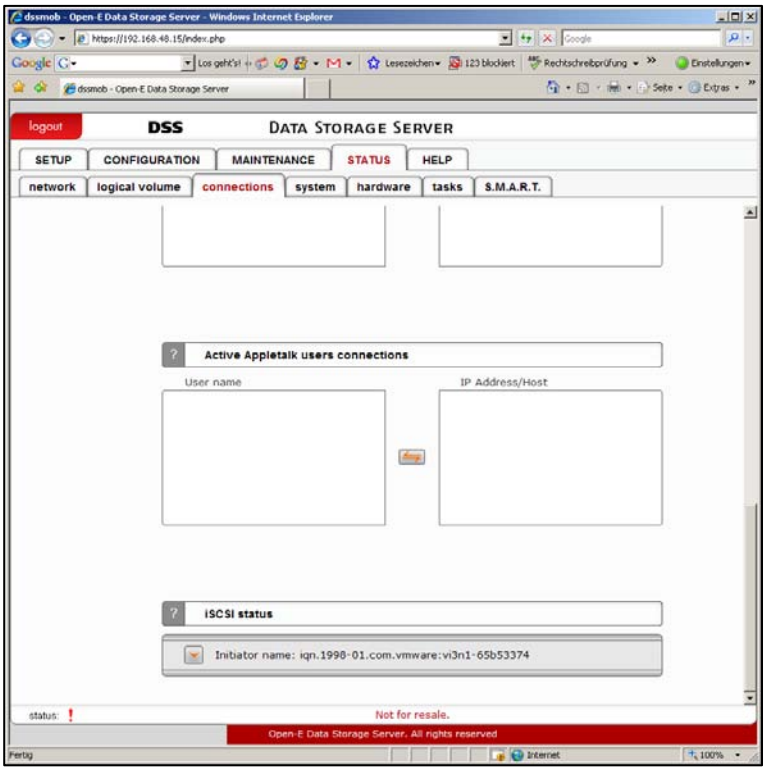


Configuration of iSCSI target access

Definition of impermissible-/permissible initiator addresses

In addition, session handling can be secured with CHAP (CHAP must then also be activated on the iSCSI target)

Figure 36 – Configuration of iSCSI target access



Active iSCSI connections

Figure 37 – Active iSCSI connections

6.6 ESX3

The configuration on the ESX3 side is different for the iSCSI HBA and the software iSCSI. It also differs between the ESX3 and the ESX3i. The following shows the configuration for software iSCSI and an ESX3 host is displayed.

Differences between iSCSI HBA and software iSCSI:

iSCSI HBA:

- *Requires no configuration under networking*
- *Settings in BIOS or in VI client for storage adapters*
- *Adapter is displayed as vmhba standard*
- *Supports dynamic and static discovery*
- *Does not require a Service Console connection to IP storage*
- *Does not require an available 3260 iSCSI port*

Software iSCSI:

- *Requires explicit configuration under networking (VMkernel port and, if applicable, an additional Service Console port)*
- *Settings only in VI client for storage adapters*
- *Adapter is displayed as vmhba32 (for ESX3: as vmhba40) or vmhba33*
- *Only supports static discovery*
- *Explicitly requires a service console connection to IP storage (via router or, better yet, with a second interface)*
- *Requires an available 3260 iSCSI port*

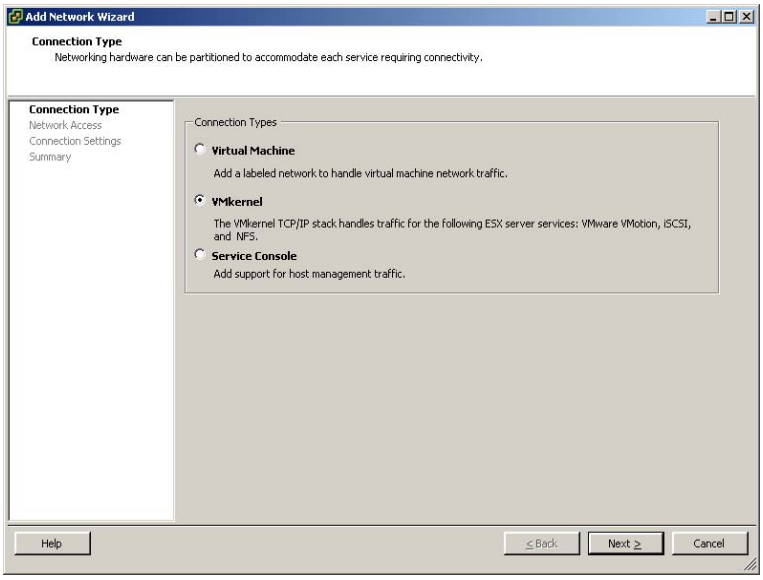
Differences between ESX3 and ESX3i:

ESX3:

- *Requires a Service Console port for software iSCSI*
- *Software iSCSI adapter is displayed as vmhba32 (for ESX3: as vmhba40)*

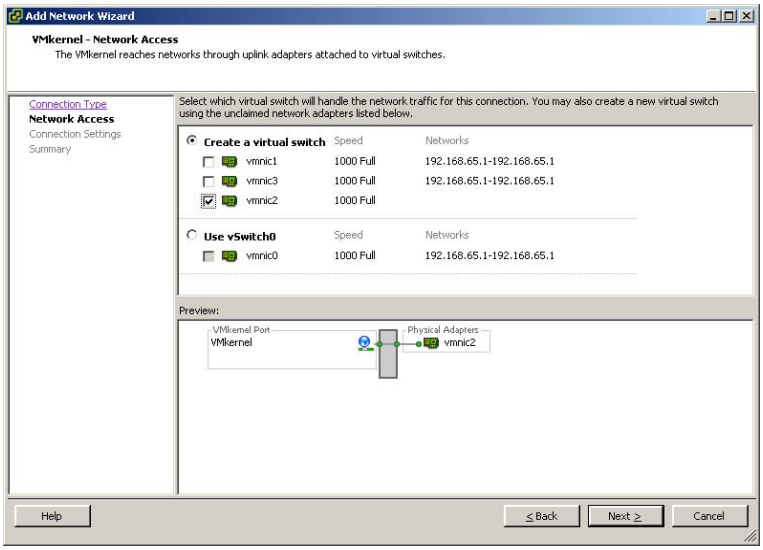
ESX3i:

- *Does not require a Service Console port for software iSCSI*
- *Software iSCSI adapter is displayed as vmhba33*



Configuration of VMkernel port at an existing or additional vSwitch

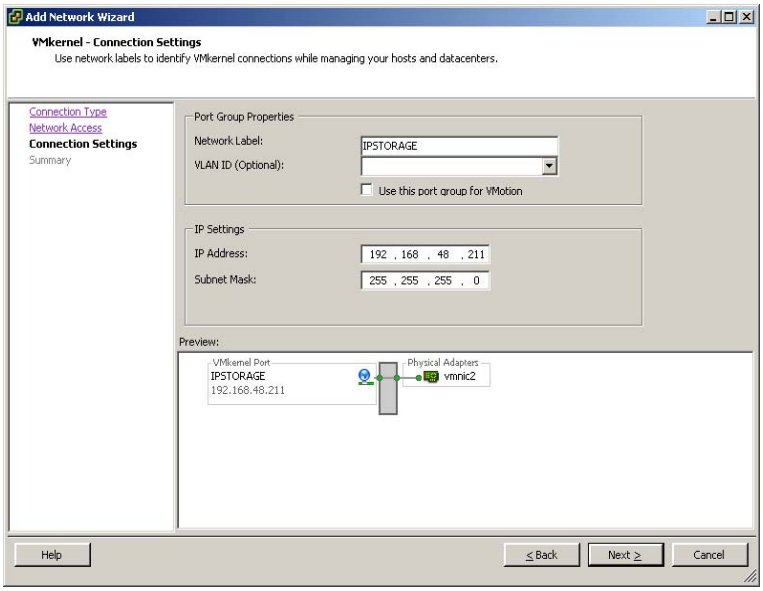
Figure 38 – Configuration of VMkernel port for IP storage



Selection of the physical port

This is displayed as vmnic. The VMkernel also displays found network info, if applicable. We recommend selecting at least two adapters for redundancy here

Figure 39 – Configuration of vSwitch for IP storage



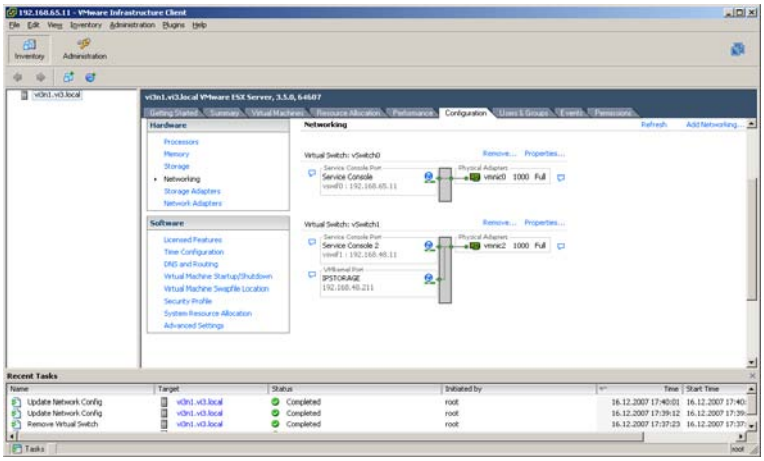
Configuration of IP information for the VMkernel port

Any name can be selected

VMotion can also be activated

Segmenting with VLAN is also possible

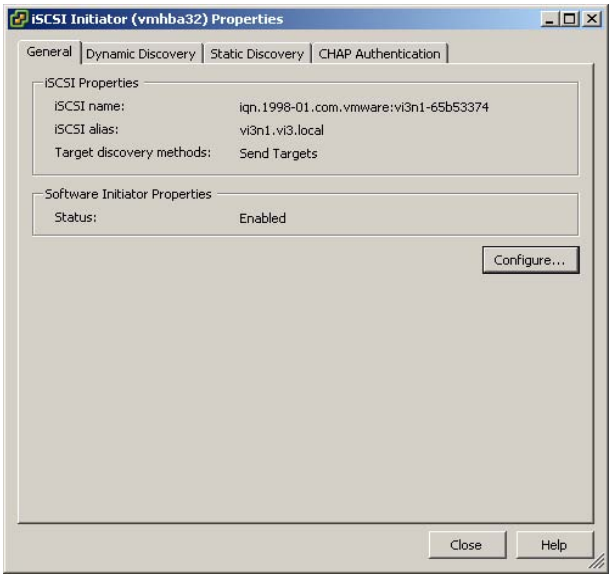
Figure 40 – Configuration of software iSCSI initiator address



Completed software iSCSI network configuration

Additional Service Console port

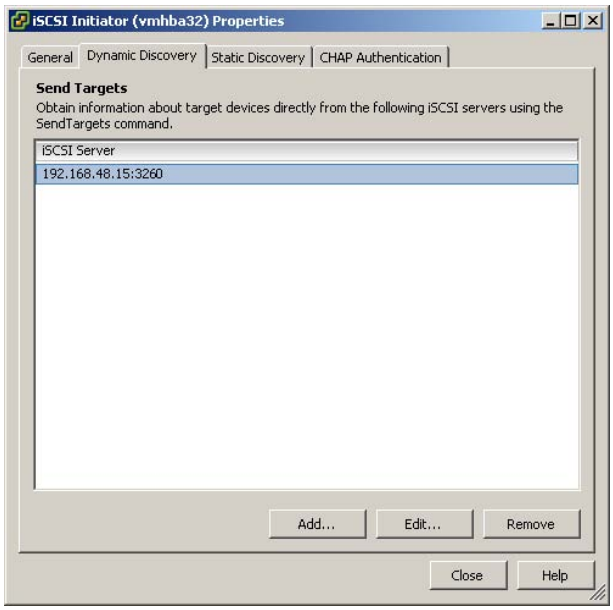
Figure 41 – Configuration of software iSCSI network



Activation of the software iSCSI initiator

The initiator name can be changed. An ESX3 host reboot is then necessary, however

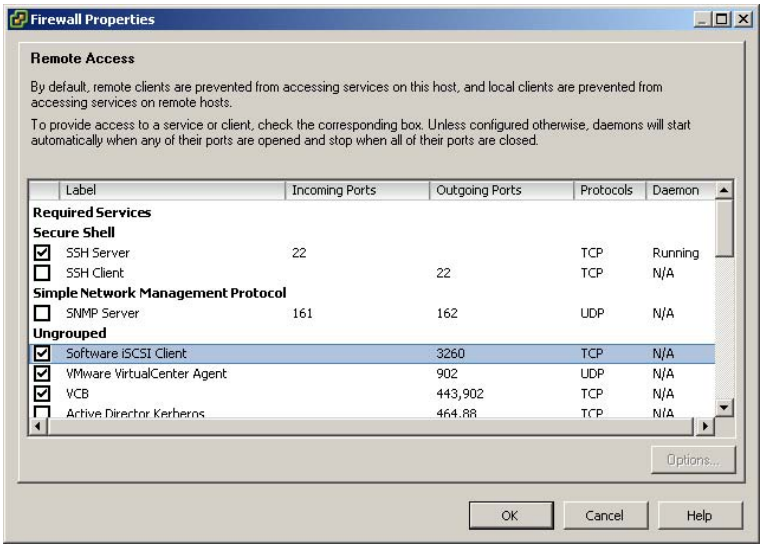
Figure 42 – Configuration of software iSCSI initiator



Configuration of target discovery

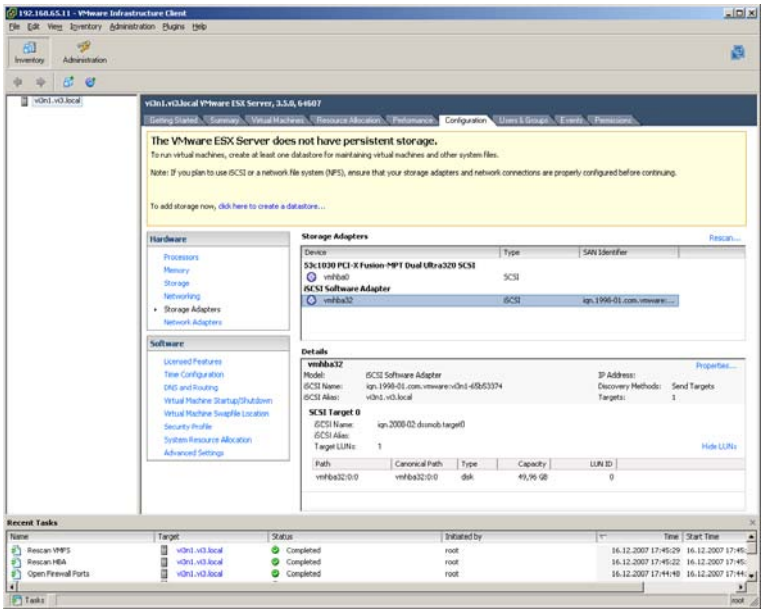
Definition of target IP address

Figure 43 – Configuration of target discovery



With ESX3, it is necessary to open port 3260 in the Service Console

Figure 44 – Configuration of Service Console firewall

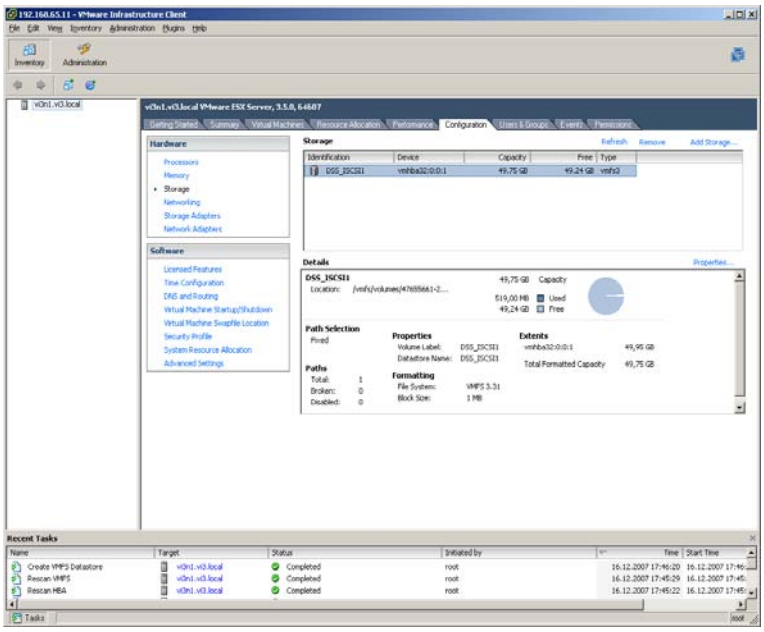


After completion of the configuration

Rescan

The located LUN is displayed

Figure 45 – Result after the adapter rescan



The creation of a VMFS3 file system is identical for FCP LUNs and iSCSI LUNs. If creation of the partition in the VI client is not possible, the fdisk command can be used for ESX3

Figure 46 – Result after the creation of a VMFS3

6.7 VM storage options

VMs can be stored on an iSCSI datastore in the following variants:

vDisk as a VMFS3 file

+ *Very easily, set hard disk sizes in the VM configuration are transferred to vDisk file sizes one to one*
+ *VMFS3 file-level locking (in the standard, a vDisk can only be used by one VM)*
+ *vDisk size can be changed in the VI client or via CLI*
+ *Is generated with clone and template deploy processes*
+ *Many VMs (vDisks) are possible*
- *VMFS3 datastores can be relatively easily overbooked (I/O load)*
- *Monitoring of the disk load (I/O load and reservation) is recommended*
- *Operation of storage software (e.g. in-band snapshot control) in the VM is not possible*

vDisk as RDM

+ *Each vDisk has its own LUN*
+ *LUN traits can be controlled via the storage system (RAID level etc.)*
- *These vDisks cannot be created via cloning or template deployment (but CLI or third-party tools -> white paper II is possible)*

6.8 Conclusion

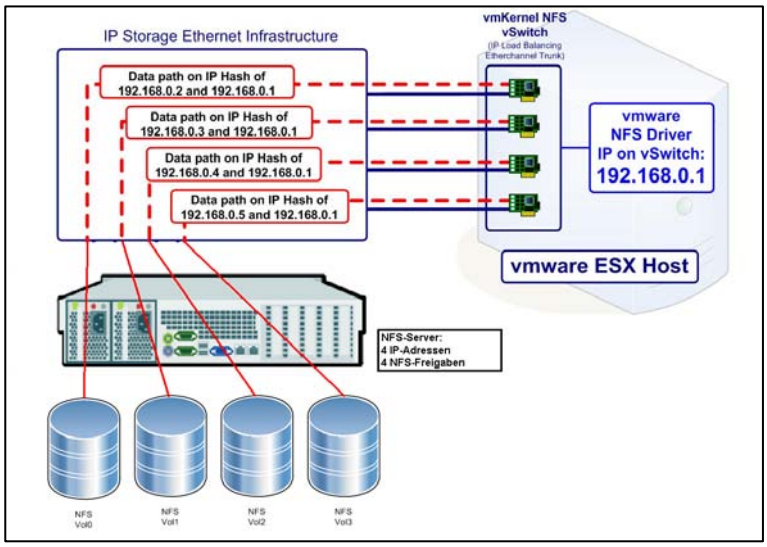
The iSCSI connection method is very easy to implement. Since this variant is not listed in the VMware compatibility guides, responsibility lies with the owner (There is an entry in the community HCL). The following points should be carefully planned and checked before use:

- *The number of ESX3 hosts to be connected to Open-E DSS*
- *Support of LUN reservation via ESX3 host*

Testing revealed this connection to be absolutely stable.

7 NFS connection

7.1 Architecture



NFS access paths:

Redundancy through VMkernel

Each volume via its own release (its own IP address)

Possible bandwidth aggregation

Figure 47 – NFS access paths

7.2 Requirements (hardware/software)

Open-E DSS:

- *NFS V3 over TCP*
- *Two GE ports*
- *Current firmware/BIOS version*

LAN:

- *Two GE switches*
- *Cabling*
- *Management and function licenses (VLAN)*
- *Current OS version*
- *Necessary for ESX3*

ESX3/ESX3i host:

- *Two GE ports for IP storage*
- *Current firmware/BIOS version*

7.3 LAN (zoning, cabling)

Zoning:

- *If a dedicated IP storage LAN is not feasible, segmenting can be implemented with a VLAN*

Cabling:

- *Fixed port settings for GE ports*

7.4 Sizing

The following attributes must be taken into account for proper planning of the required volume sizes:

VMs:

- *Small files (configuration files, log files and NVRAM file)*
- *vDisk files are created thin in the standard (dynamic reservation of the configured disk space)*
- *Delta files (VMware snapshots, multiple) grow in 16 MB increments up to the vDisk size*
- *Memory snapshots are as large as the VM storage configuration*
- *VMkernel swap file is as large as the VM storage configuration*

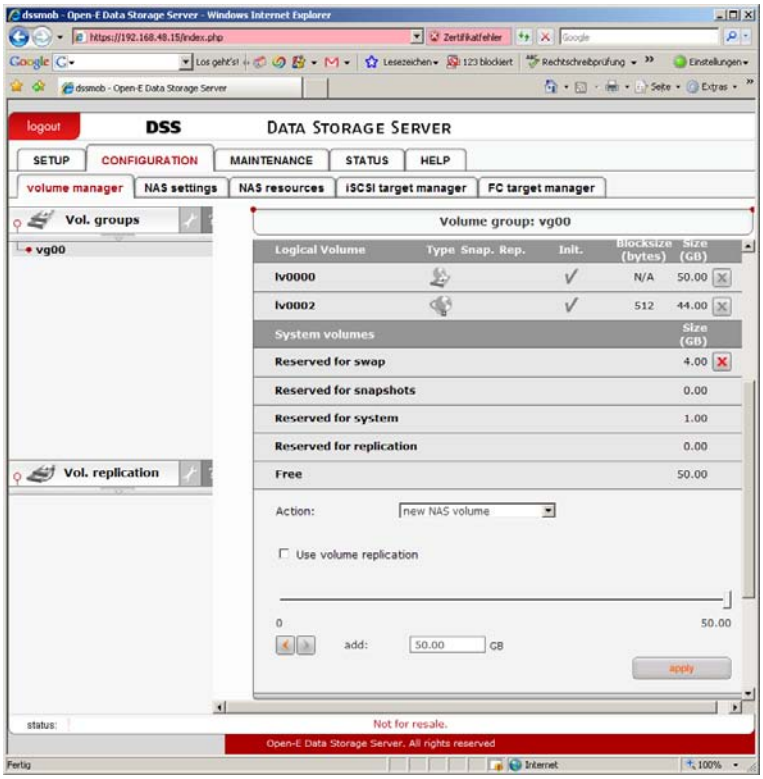
In addition, decisions must be made with regard to:

- *Large volumes (many VMs) and small volumes (few VMs)*
  Performance (volume with many VMs <-> few VMs)
  Backup/restore (large volume <-> small volume)
  Storage snapshot function (volume with many VMs <-> one VM)
- *Replication relationships with other storage servers (volume with many VMs <-> one VM)*

To sum up, one could say that a compromise must be found between management expenditure, performance and restoration time.

Some best-practice values have already been mentioned. There is a large range involved here, however. Oftentimes, between 10 and 20 VMs are stored on an NFS datastore and thus the volume generates between 150 and 300 GB.

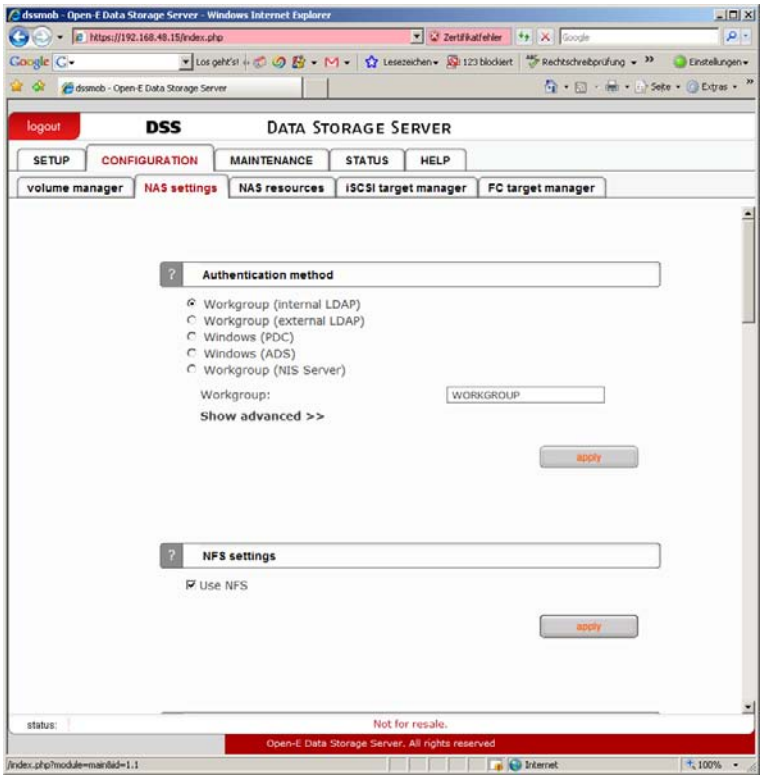
7.5 Open-E DSS



Configuration of NAS volume for NFS access

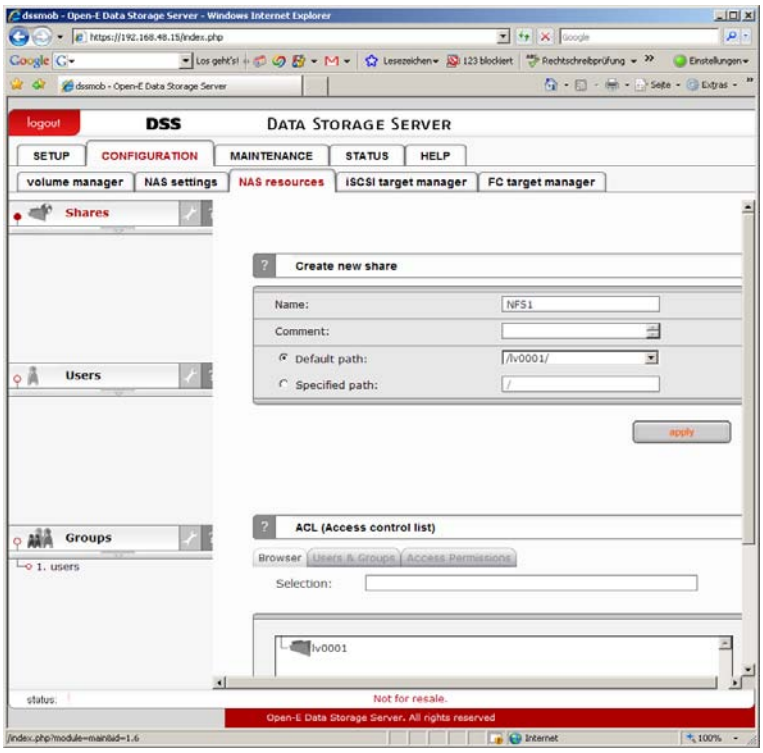
Volume group already exists

Figure 48 – Configuration of NAS volume



Activation of NFS protocol

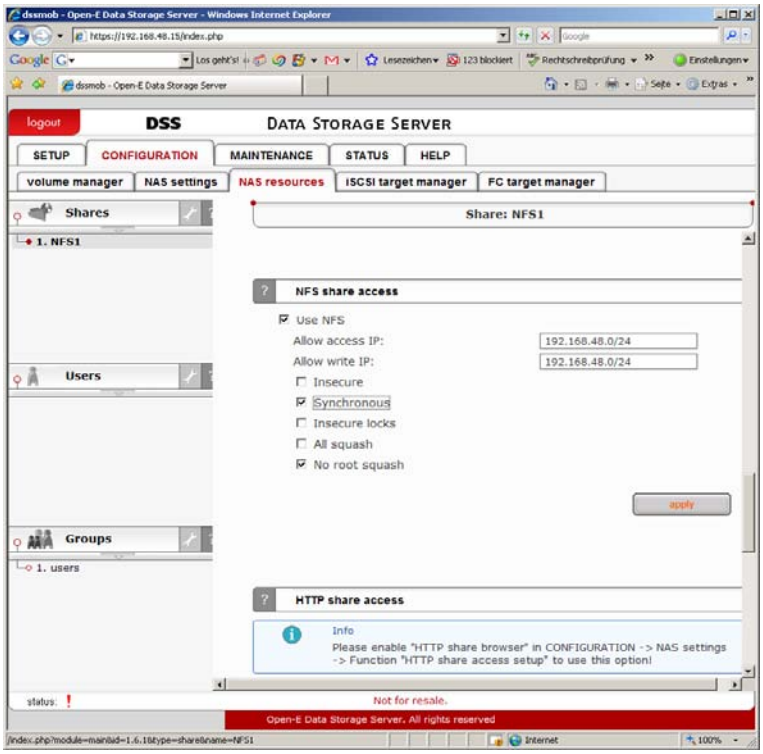
Figure 49 – Activation of NFS



Configuration of NFS release

Definition of the release name

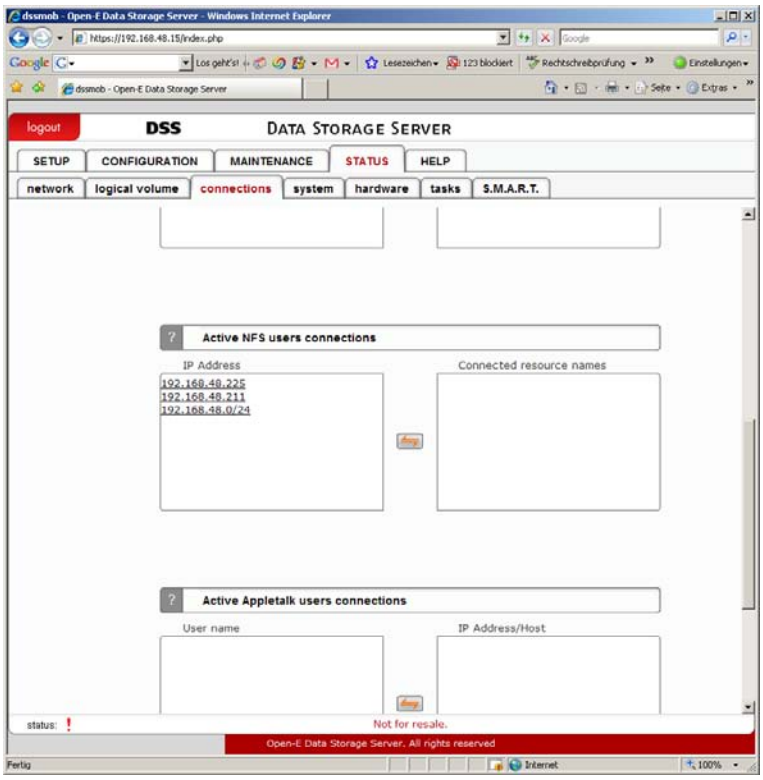
Figure 50 – Configuration of NFS release I



Configuration of NFS release

Definition of the access attributes

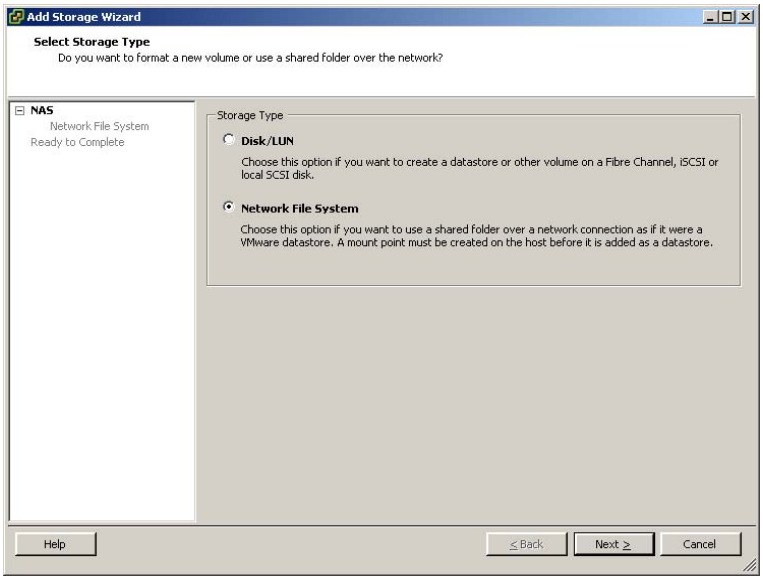
Figure 51 – Configuration of NFS release II



Display of the active NFS connections

Figure 52 – Active NFS connections

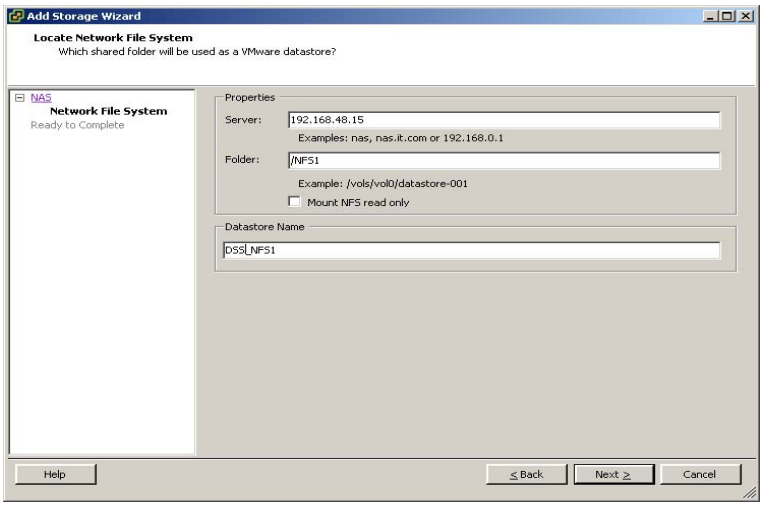
7.6 ESX3



Configuration of NFS on ESX3 or ESX3i

Add storage

Figure 53 – Configuration of NFS



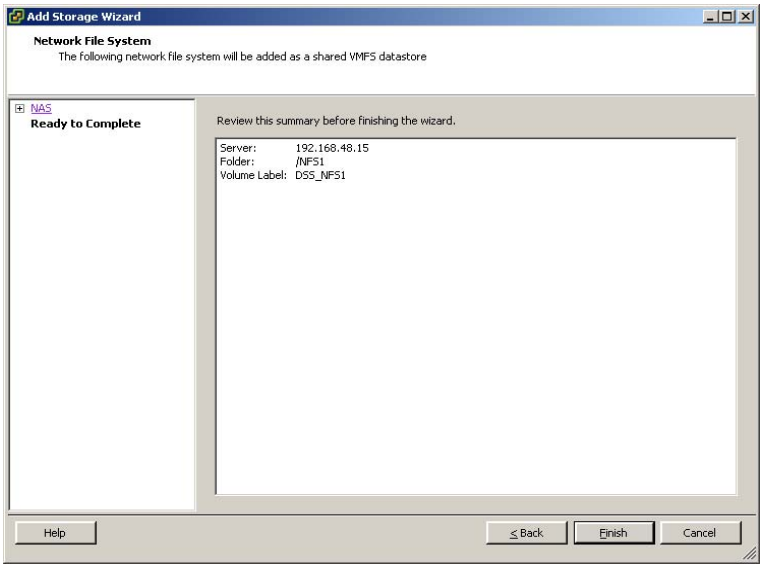
Configuration of NFS properties

Server name or IP address

NFS share name (MUST be identical to release)

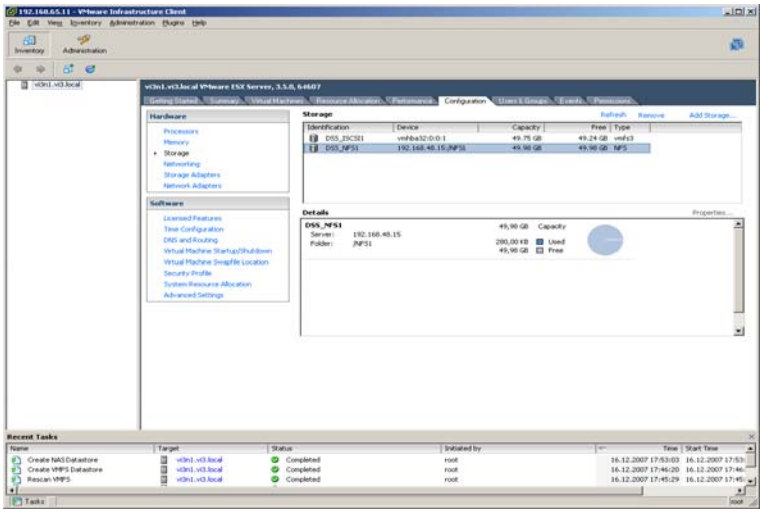
Datastore name (MUST be the same on all ESX3 hosts)

Figure 54 – Configuration of NFS properties



Automatic completion of system

Figure 55 – Configuration of NFS



Complete, set-up datastore for NFS

Figure 56 – NFS datastore

7.7 VM storage options

VMs can be stored on an NFS datastore very easily:

vDisk as a file

+ *Very easily, set hard disk sizes in the VM configuration are transferred to vDisk file sizes one to one.*
+ *VMFS3 file-level locking (by default a vDisk can only be used by one VM).*
+ *vDisk size can be changed in the VI client or via CLI*
+ *Is generated with clone and template deploy processes*
+ *Many VMs (vDisks) are possible*
+ *Simple DR processes possible*

7.8 Conclusion

The NFS connection method is very easy to implement. Since this variant is not listed in the VMware compatibility guides, responsibility lies with the owner (There is an entry in the community HCL). The following points should be carefully planned and checked before use:

- *The number of ESX3 hosts to be connected to Open-E DSS*

Testing revealed this connection to be absolutely stable.

8 Other sources of information

8.1 VMware

http://www.vmware.com

http://communities.vmware.com

8.2 Open-E

http://www.open-e.com

http://forum.open-e.com

8.3 Community

http://www.vmware-land.com

9 Indices

9.1 Table index

Table 1 – Tables and configuration diagrams 57

Table 2 – Documents 57

9.2 Figure index

Figure 1 – VMware basics 7

Figure 2 – Load increase 7

Figure 3 – Load increase 8

Figure 4 – Reduction of storage connections 8

Figure 5 – Reduction of network connections 9

Figure 6 – Product evolution 10

Figure 7 – VI software stack (version 3) 11

Figure 8 – ESX3 architecture 14

Figure 9 – ESX3i architecture 14

Figure 10 – Main differences 15

Figure 11 – VI architecture 17

Figure 12 – VI LAN architecture 17

Figure 13 – VI storage architecture 17

Figure 14 – Typical SAN architecture 22

Figure 15 – FCP access paths 22

Figure 16 – Configuration of Qlogic target mode 25

Figure 17 – Configuration of volume group 25

Figure 18 – Configuration of FC volume (LUN) 26

Figure 19 – Configuration of FC target group 26

Figure 20 – Configuration of WWPN HBA alias 26

Figure 21 – Configuration of volume mapping 27

Figure 22 – Configuration -> Storage Adapters -> Before Rescan 28

Figure 23 – Configuration -> Storage Adapters -> After Rescan 28

Figure 24 – Configuration -> Storage -> Add Storage -> Disk/LUN 28

Figure 25 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Device Location 29

Figure 26 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Current Disk Layout 29

Figure 27 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Datastore Name 29

Figure 28 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Maximum File Size 30

Figure 29 – Configuration -> Storage -> Add Storage -> Disk/LUN -> Review 30

Figure 30 – Configuration -> Storage 30

Figure 31 – iSCSI architecture 32

Figure 32 – iSCSI access paths 32

Figure 33 – Configuration of iSCSI volume (LUN) 35

Figure 34 – Configuration of iSCSI target 35

Figure 35 – Configuration of iSCSI target mapping 36

Figure 36 – Configuration of iSCSI target access 36

Figure 37 – Active iSCSI connections 37

Figure 38 – Configuration of VMkernel port for IP storage 39

Figure 39 – Configuration of vSwitch for IP storage 39

Figure 40 – Configuration of software iSCSI initiator address 40

Figure 41 – Configuration of software iSCSI network 40

Figure 42 – Configuration of software iSCSI initiator 40

Figure 43 – Configuration of target discovery 41

Figure 44 – Configuration of Service Console firewall 41

Figure 45 – Result after the adapter rescan 42

Figure 46 – Result after the creation of a VMFS3 42

Figure 47 – NFS access paths 44

Figure 48 – Configuration of NAS volume 46

Figure 49 – Activation of NFS 46

Figure 50 – Configuration of NFS release I 47

Figure 51 – Configuration of NFS release II 47

Figure 52 – Active NFS connections 48

Figure 53 – Configuration of NFS 48

Figure 54 – Configuration of NFS properties 49

Figure 55 – Configuration of NFS 49

Figure 56 – NFS datastore 49

9.3 Abbreviation index

APIC	*Advanced Programmable Interrupt Controller*
BMC	*Baseboard Management Controller*
CLI	*Command Line Interface*
CPU	*Central Processing Unit*
DRS	*Distributed Resource Scheduler*
DSS	*Data Storage Server*
ESX	*Name of the VMware enterprise hypervisor product*
ESX3i	*Name of the embedded hypervisor*
FC	*Fiber Channel*
FCP	*Fiber Channel Protocol*
HA	*High Availability*
HBA	*Host Bus Adapter*
HCL	*Hardware Compatibility List*
HV	*“Hochverfügbarkeit” (high availability)*
iSCSI	*SCSI over TCP/IP*
MRU	*Most Recently Used*
NFS	*Network File System*
NIC	*Network Interface Card*
NPIV	*N Port ID virtualization*
NX	*No execution*
OS	*Operating System*
pCPU	*physical CPU*
pDISK	*physical DISK*
pHBA	*physical HBA*
pMEMORY	*physical MEMORY*
pNIC	*physical NIC*
pSWITCH	*physical SWITCH*
P2V	*Migration from physical to virtual*

RDM	*Raw Disk Map*
RR	*Round Robin*
SAN	*Storage Area Network*
TSO	*TCP/IP Segment Offloading*
VC	*Virtual Center*
VCS	*Virtual Center Server*
VI3	*Virtual Infrastructure 3*
VM	*Virtual Machine*
vCPU	*virtual CPU*
vDISK	*virtual DISK*
vHBA	*virtual HBA*
vMEMORY	*virtual MEMORY*
vNIC	*virtual NIC*
vSWITCH	*virtual Switch*
V2P	*Migration from virtual to physical*
V2V	*Migration from virtual to virtual*
VT	*Virtualization Technology*
WWN	*World Wide Number*
WWNN	*World Wide Node Number*
WWPN	*World Wide Port Number*
XD	*Execution Deny*

10 Appendices

10.1 Tables

File	Version	Remark

Table 1 – Tables and configuration diagrams

10.2 Documents

File	Version	Remark

Table 2 – Documents