

iSCSI, NAS and IP Storage Configuration for VMware ESX Server

Andy Banta

VMware Storage Engineer



VMWORLD 2006

Overview

- What is iSCSI?
- What is NAS?
- Why are iSCSI and NAS interesting?
- How VMware is doing iSCSI
 - > Implementation
 - > Features
- How VMware is doing NAS
 - > NFS Implementation
- IP Storage Network Configuration
 - > Network Setup
 - > Multi-path Options

What is iSCSI?

- Transport of SCSI over TCP/IP
 - > Internet “standard”
 - > IETF 3720, IEFT 3347
- SAN-oriented
 - > Comparable to Fibre Channel in many ways
- Non-privileged Service
 - > Port 3260, anyone can be an iSCSI target
- Target Names
 - > iqn.date.com.namingauthority:uniqueusername
 - > eui:01234567890ABCDEF
 - > Fully qualified:
 - ibox.vmware.com:860//iqn.1998-01.com.vmware.ibox:target1
 - Hostname port iSCSI name target

What is NAS?

- File systems of TCP/IP
 - > NFS, CIFS most common
- Widely used for sharing files
- NFS Internet Standard
 - > IETF 1094
- NFS naming conventions
 - > `nfs.remote.com:/remote/filesystem`
- More info on iSCSI:
 - > www.ietf.org, www.iscsistorage.com
- More info on NFS:
 - www.ietf.org, en.wikipedia.org/wiki/Network_File_System

SAN vs. NAS

SAN



Block Storage



NAS



File System

Why iSCSI and NAS are Interesting

- Easy access to remote data
- Easy sharing of data
- Inexpensive
 - > No additional hardware
 - Use existing NICs
 - Optionally, iSCSI HBAs, TOE for speed-up
 - > Use existing hardware
 - Any system can be a NAS, iSCSI server
 - > Existing infrastructure
 - Office cabling, routers, switches
 - Administrative expertise
- Use existing network technologies
 - > Authentication, encryption, routing, internet ready

Why NAS is Interesting

- Prolific file-sharing approach
 - > Widespread use
 - > Easy storage conglomeration
- Use existing file system tools
 - > Backups
 - > Resource allocation
- Built-in permissions
 - > File access consistent, network-wide

Why iSCSI is Interesting

- Emerging, Large Market
 - > 50-100% annual growth
 - > About one-third of the SAN market (attach rate)
 - > Lots of industry backing: Cisco, EMC, NetApp, Microsoft
 - > Dominant SAN by the end of the decade
- Different approach than NAS
 - > Raw block storage, not file system
- Different Approach to SAN
 - > Built-in encryption
 - > Built-in discovery
 - > Long haul
 - No bridging

How We do iSCSI

- Software only
 - > Based on the Cisco reference implementation
 - > Uses ESX's vmkernel network stack
 - > Discovery, authentication done through service console
- QLogic iSCSI adapters
 - > QLA4050C, QLA4052C (QLA4010C experimentally)
 - > Looks like any other SCSI adapter
 - > Does its own networking
- Takes advantage of existing storage implementation
 - > Just appears to be more disks to work with
 - > Multipath-capable

iSCSI Features

- ESX Boot Support
 - > Available with QLogic HBAs, only
- Target discovery
 - > SendTargets
 - > Static discovery (QLogic HBAs only)
 - > No SLP or iSNS
- ESX features
 - > VMFS3 or RDMS
 - > VMOTION
 - > Multiple paths
 - SCSI multipath for QLogic HBA
 - NIC teams for SW initiator
- Authentication
 - > CHAP, per initiator
 - > None

NAS Support and Features

- Internal implementation
 - > Uses ESX vmkernel networking stack
 - > NFS version 3 over TCP only (no UDP)
- Internal locking protocol
 - > Mandatory, lease-based for ESX hosts,
- ESX features
 - > No VMFS3 or RDMS
 - > VMOTION
 - > Multiple paths through NIC teaming
 - > No ESX booting
- Security
 - > Standard AUTH_UNIX/AUTH_HOST
 - Hostname, userid, group id for file access

OK, Where's Our Storage?

- iSCSI

- > EMC AX100i, CXx00i Series (and Dell rebadging)
- > EMC Celerra NS, NSX, CNS
- > NetApp FAS 2xx, 9xx, 3020 filers
- > EqualLogic PS Series

- NFS

- > RedHat Linux 3 u5 NFS server
- > Fedora Core 4 NFS server
- > NetApp FAS 2xx, 9xx, 3000, 6000 filers
- > EMC Celerra NS, NSX, CNS

iSCSI Configuration

■ Administration

- > Configured through Virtual Center UI
 - QLogic cards configured completely through Storage configuration
 - SW Initiator takes a few more steps
 - Network setup through Networking config
 - Discovery, Authentication through Storage Configuration
- > Common configuration for all initiators
- > No vendor-specific tools needed
- > SNIA-based IMA library implementation
 - Standard
 - Third-party vendors can use interface for iSCSI management
 - Expandable “plug-in” approach to management

iSCSI Configuration

The screenshot displays the vSphere Client interface with the 'Configuration' tab selected. The left pane shows the 'iSCSI Initiator (vmhba1) Properties' dialog box, and the right pane shows the 'Storage Adapters' configuration.

iSCSI Initiator (vmhba1) Properties

General | Dynamic Discovery | Static Discovery | CHAP Authentication

iSCSI Properties

ISCSI name: iqn.2000-04.com.qlogic:qla4052c.fs20548b12677.1
ISCSI alias: cardomine
Target discovery methods: Send Targets, Static Target

Hardware Initiator Properties

Network Interface Properties

Current/maximum speed: 1024Mb/1024Mb
MAC Address: 00:c0:dd:07:55:92

IP Settings

IP Address: 10.11.246.22
Subnet Mask: 255.255.255.0
Default Gateway: 10.11.246.1

DNS Servers

Preferred Server: 10.11.246.6
Alternate Server: 0.0.0.0

Configure... Close Help

Storage Adapters

Device	Type
QLA4010	
vmhba3	iSCSI
QLA4022	
vmhba1	iSCSI
vmhba2	iSCSI
53c1030 PCI-X Fusion-MPT Dual Ultra320 SCSI	

Details

vmhba1

Model: QLA4022
iSCSI Name: iqn.2000-04.com.qlogic:qla4052c.fs20548b1267.
iSCSI Alias: cardomine

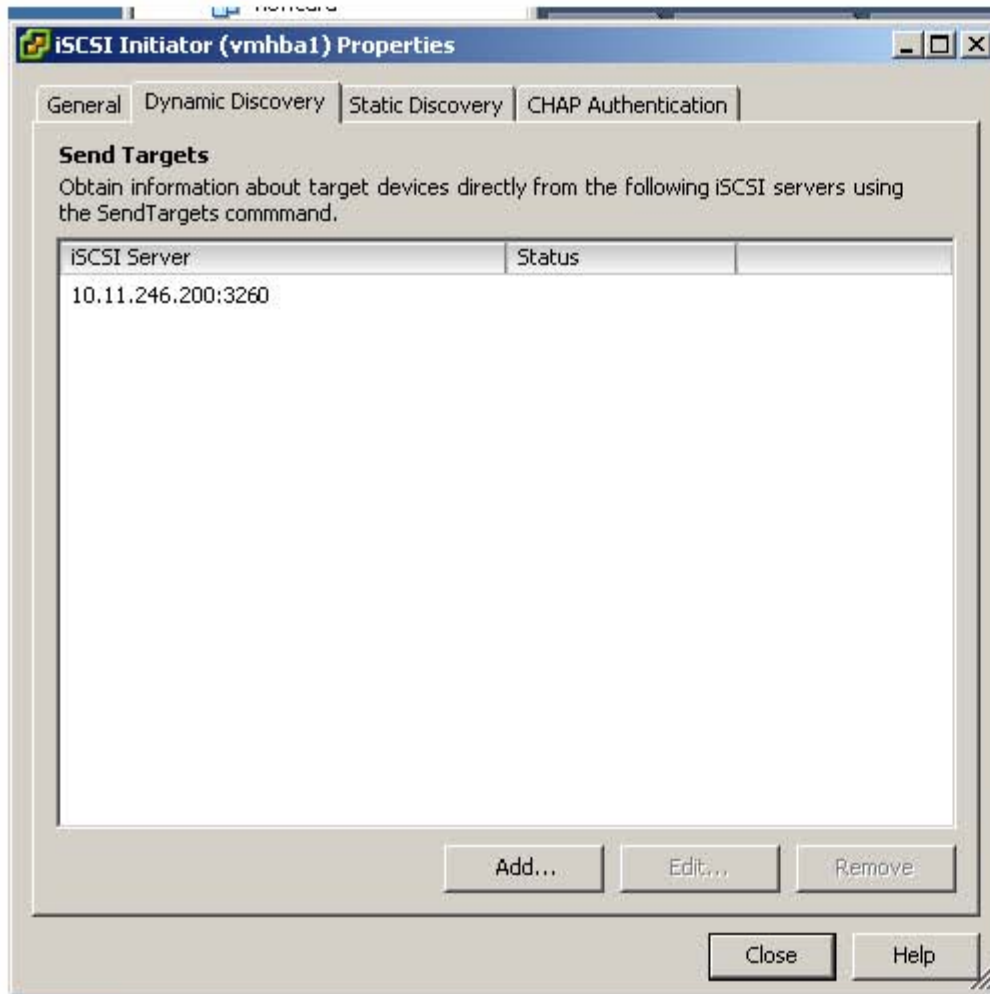
SCSI Target 2

iSCSI Name: iqn.1992-04.com.emc:ax.apm00050602321.a0
iSCSI Alias: 2321.a0
Target LUNs: 1

Path	Canonical Path	Capacity
vmhba1:2:0	vmhba1:2:0	31.00 GB

ed by Time Start Time Comple

iSCSI Configuration



Location Performance Configuration Users & Groups Events

Storage Adapters

Device	Type
QLA4010	
vmhba3	iSCSI
QLA4022	
vmhba1	iSCSI
vmhba2	iSCSI
53c1030 PCI-X Fusion-MPT Dual Ultra320 SCSI	

Details

vmhba1

Model: QLA4022
iSCSI Name: iqn.2000-04.com.qlogic:qla4052c.fs20548b1267;
iSCSI Alias: cardomine

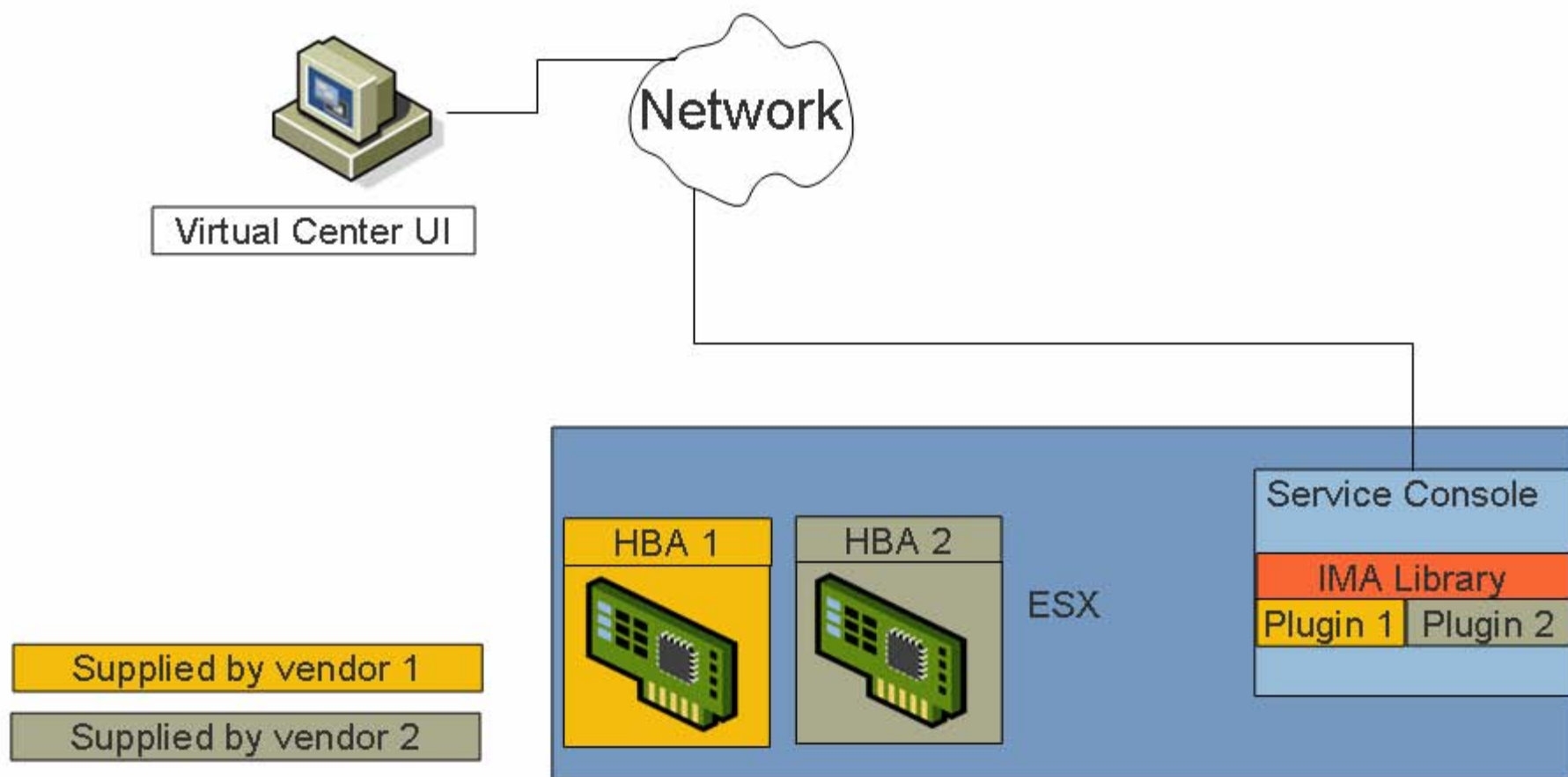
SCSI Target 2

iSCSI Name: iqn.1992-04.com.emc:ax.apm00050602321.a0
iSCSI Alias: 2321.a0
Target LUNs: 1

Path	Canonical Path	Capacity
vmhba1:2:0	vmhba1:2:0	31.00 GB

by	Time	Start Time	Comple
----	------	------------	--------

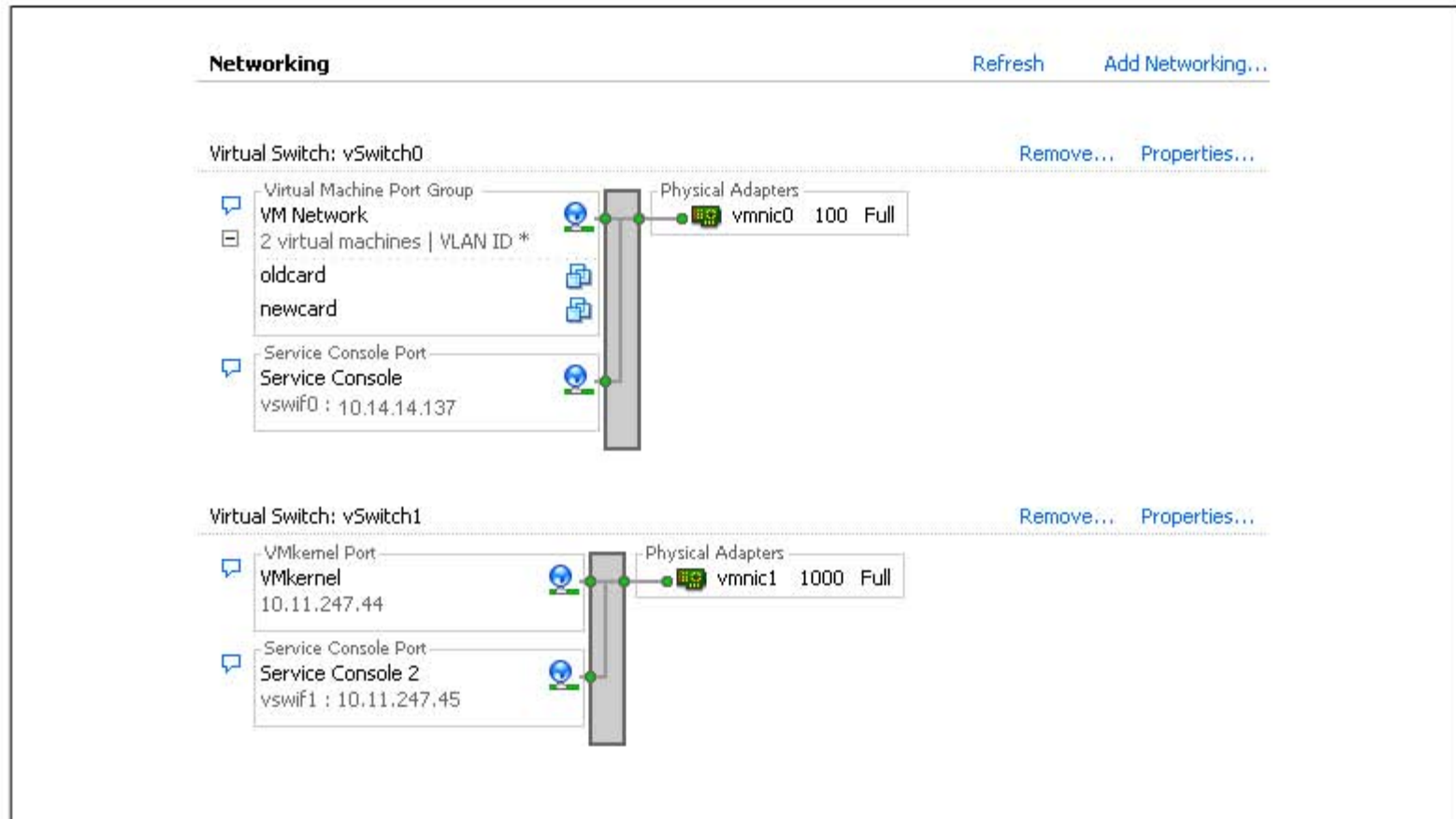
IMA Plugin Overview



NFS Configuration

- Configured with Virtual Center UI
 - > IP Storage Port configured
 - > Storage is selected like any other disk or volume
 - > Datastores used like VMFS2 or VMFS3 volume

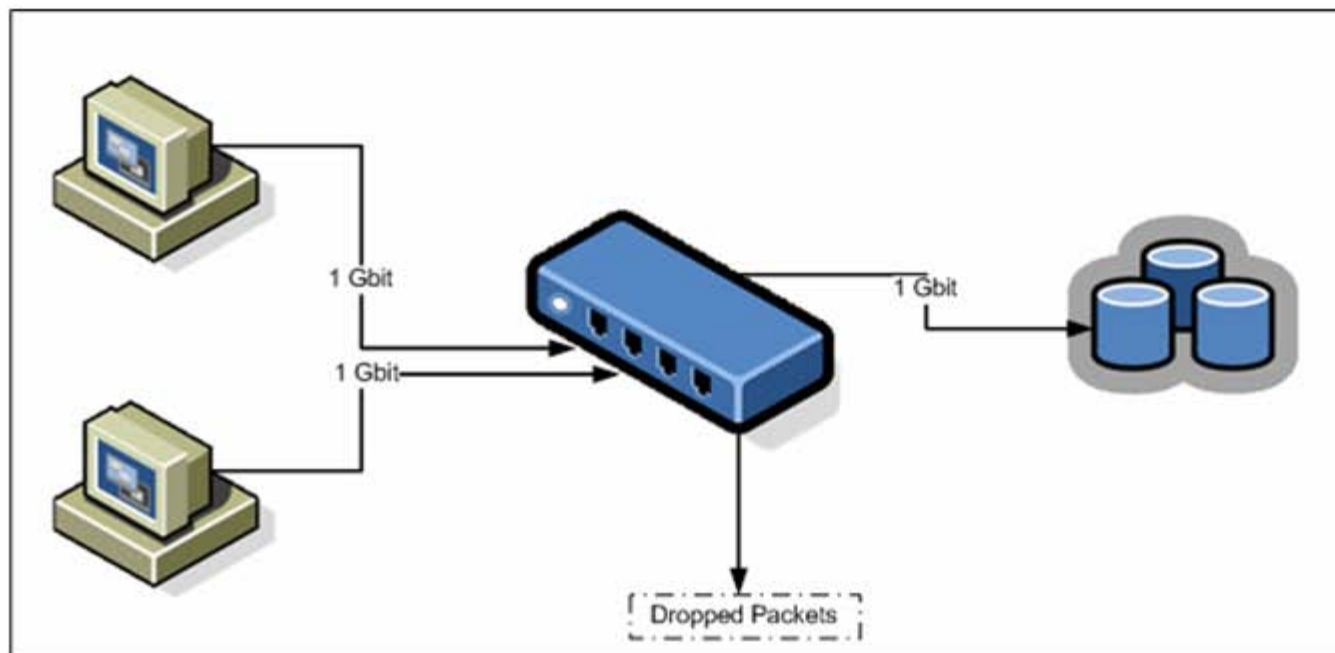
NFS Configuration



IP Storage Network Configuration

- Gig-Ethernet equipment, only
- Separate networks
 - > Security
 - > Separate traffic
 - VLANs still share traffic
 - Separate physical networks for heavy IP-based storage
- Dropped network packets
 - > Dropped because network is shared
 - > Monitoring network errors
 - > Retransmission can be very slow
- Multiple paths for data integrity
 - > iSCSI multi-path configurations
 - > TCP/IP multi-path configuration

Network Oversubscription



2 Gigabits of traffic don't fit on a 1 Gigabit link

Monitoring Dropped Packets

- Available from some storage arrays
- Available from many managed switches
 - > Both directions, more useful

*: interface is up

IHQ: pkts in input hold queue

OHQ: pkts in output hold queue

RXBS: rx rate (bits/sec)

TXBS: tx rate (bits/sec)

TRTL: throttle count

IQD: pkts dropped from input queue

OQD: pkts dropped from output queue

RXPS: rx rate (pkts/sec)

TXPS: tx rate (pkts/sec)

Interface			IHQ	IQD	OHQ	OQD	RXBS	RXPS	TXBS	TXPS
TRTL										

* GigabitEthernet0/1			3	9922	0	0	476303000		62273	
477840000	63677	0								

Storage Multipathing with iSCSI

- ESX maintains path info
- Storage LUNs uniquely identified
- SW iSCSI initiator
 - > Same storage LUN seen multiple times
 - LUN accessed through alternate target for failure
- Multiple iSCSI HBAs
 - > Same as SW initiator
 - > Both HBAs see storage
 - Switch to other HBA for target failure

Storage Multipath with iSCSI

Processors

Memory

► Storage (SCSI, SAN, and NFS)

Identification	Device	Capacity	Free	Type
ten-score	vmhba3:7:0:1	9.75 GB	4.14 GB	vmfs3
catch-22	vmhba1:2:0:1	30.75 GB	21.76 GB	vmfs3
localdisk	vmhba0:0:0:1	33.75 GB	22.78 GB	vmfs3

catch-22 Properties

Volume Properties

General

Datastore Name: catch-22

Change...

Format

File System: VMFS 3.21

Maximum File Size: 256 GB

Block Size: 1 MB

Extents

A VMFS file system can span multiple hard disk partitions, or extents, to create a single logical volume.

Extent	Capacity
vmhba1:2:0:1	31.00 GB

Extent Device

The extent selected on the left resides on the LUN or physical disk described below.

Device

Device	Capacity
vmhba1:2:0	31.00 GB

Primary Partitions

Primary Partitions	Capacity
1. VMFS	31.00 GB

Path Selection

Most Recently Used

Paths

Paths	Path Status
vmhba1:2:0	◆ Active
vmhba2:2:0	◇ Standby

Total Formatted Capacity: 30.75 GB

Add Extent...

Manage Paths...

Properties...

GB Capacity

GB Used
GB Free



Extents

vmhba1:2:0:1	31.00 GB
Total Formatted Capacity	30.75 GB

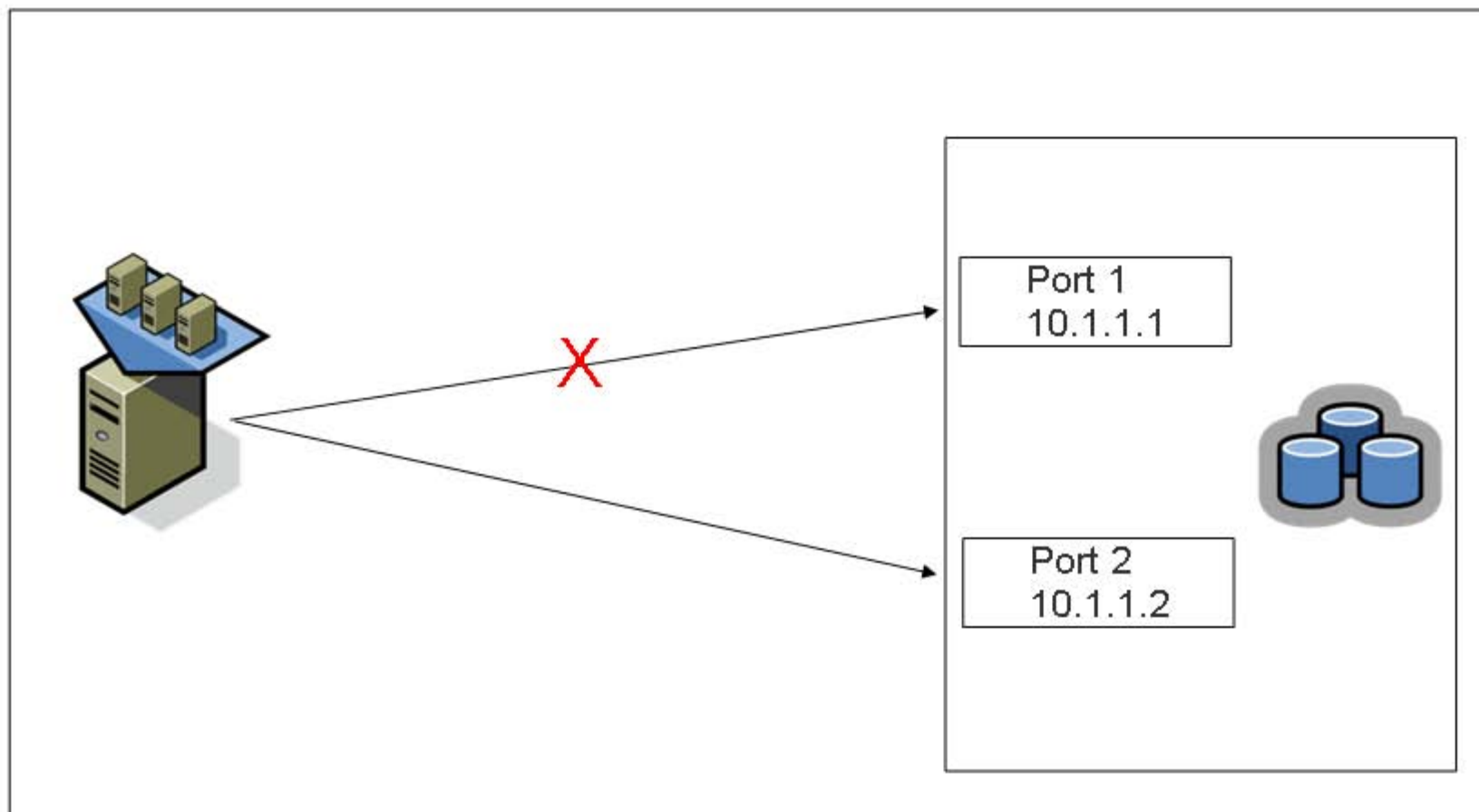
Network Multipath with NIC Teaming

- Useful for SW iSCSI or NFS
 - > Not for HW iSCSI
- Multiple NICs for ESX vswitch
 - > Failover for loss of link
 - > Rudimentary load balancing
 - Configuration-dependent
- Storage multipathing not involved

Network Multipathing with iSCSI Login Redirection

- No configuration needed in ESX
- Both HW iSCSI and SW iSCSI
- Requires support from storage
- Storage array maintains path info
 - > No info in storage multipathing or ESX network config
- iSCSI sessions are redirected to appropriate port
 - > Happens on-the-fly for load balancing

Network Multipathing with iSCSI Login Redirection



iSCSI and NAS Rock!

- Questions?

Presentation Download

Please remember to complete your
session evaluation form
and return it to the room monitors
as you exit the session

The presentation for this session can be downloaded at
<http://www.vmware.com/vmtn/vmworld/sessions/>

Enter the following to download (case-sensitive):

Username: cbv_rep
Password: cbvfor9v9r

Some or all of the features in this document may be representative of feature areas under development. Feature commitments must not be included in contracts, purchase orders, or sales agreements of any kind. Technical feasibility and market demand will affect final delivery.

VMWORLD 2006

