# 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

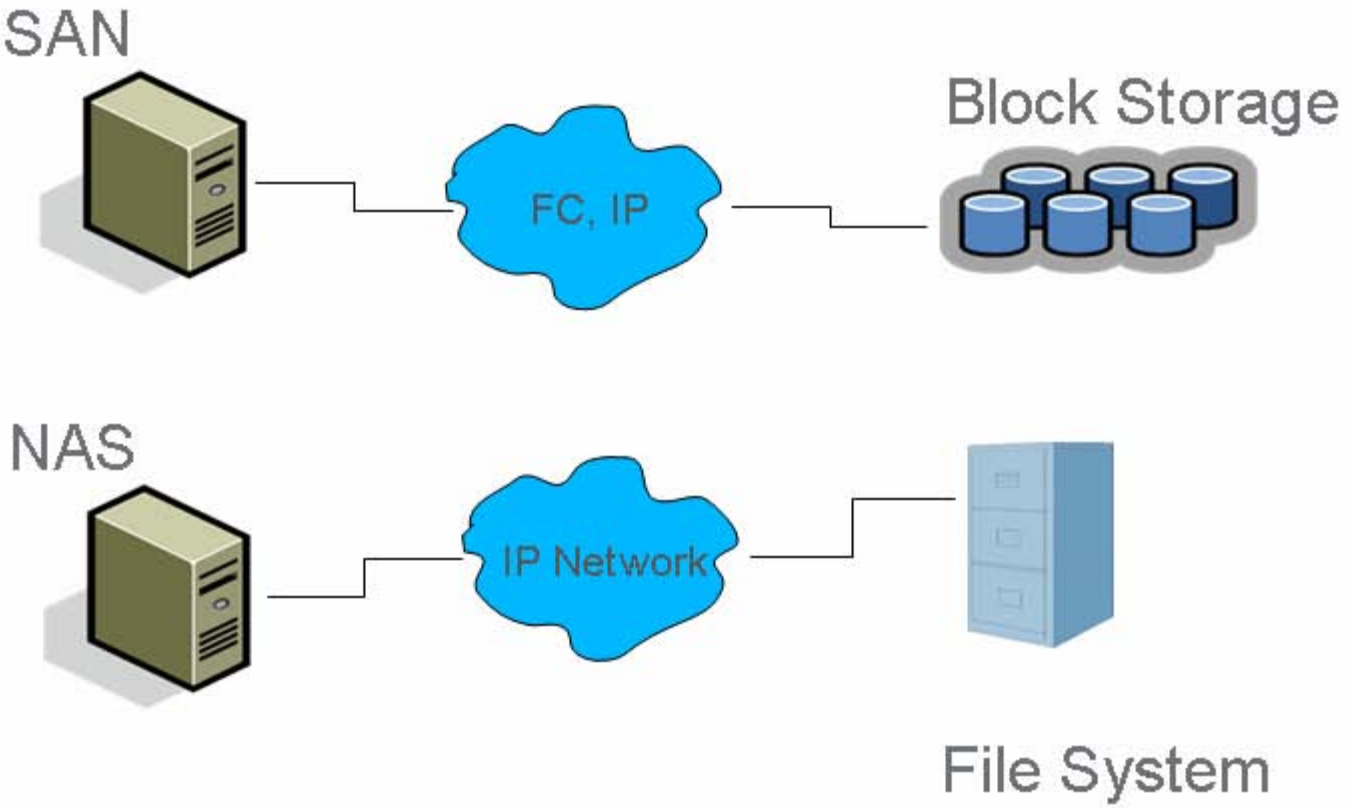
# What is iSCSI?

- Transport of SCSI over TCP/IP
  - Internet “standard”
  - IETF 3720, IEFT 3347
- SAN-oriented
  - Comparable to Fibre Channel in many ways
- Non-priveleged Service
  - Port 3260, anyone can be an iSCSI target
- Target Names
  - iqn.date.com.namingauthority:uniquename
  - 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



# 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 industy backing: Cisco, EMC, NetApp, Micrsoft
  - 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

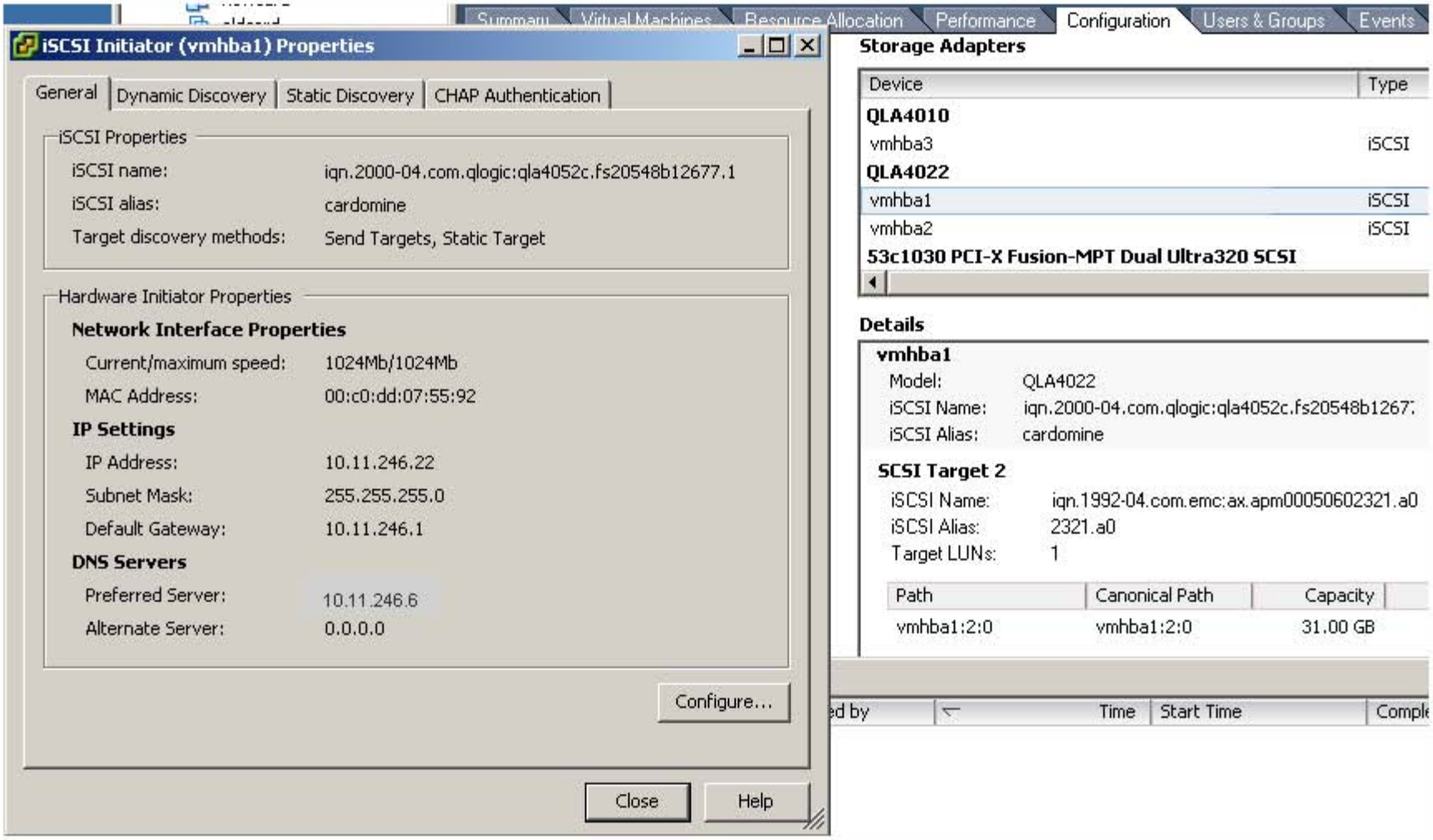
# OK, Where's Our Storage?

- 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

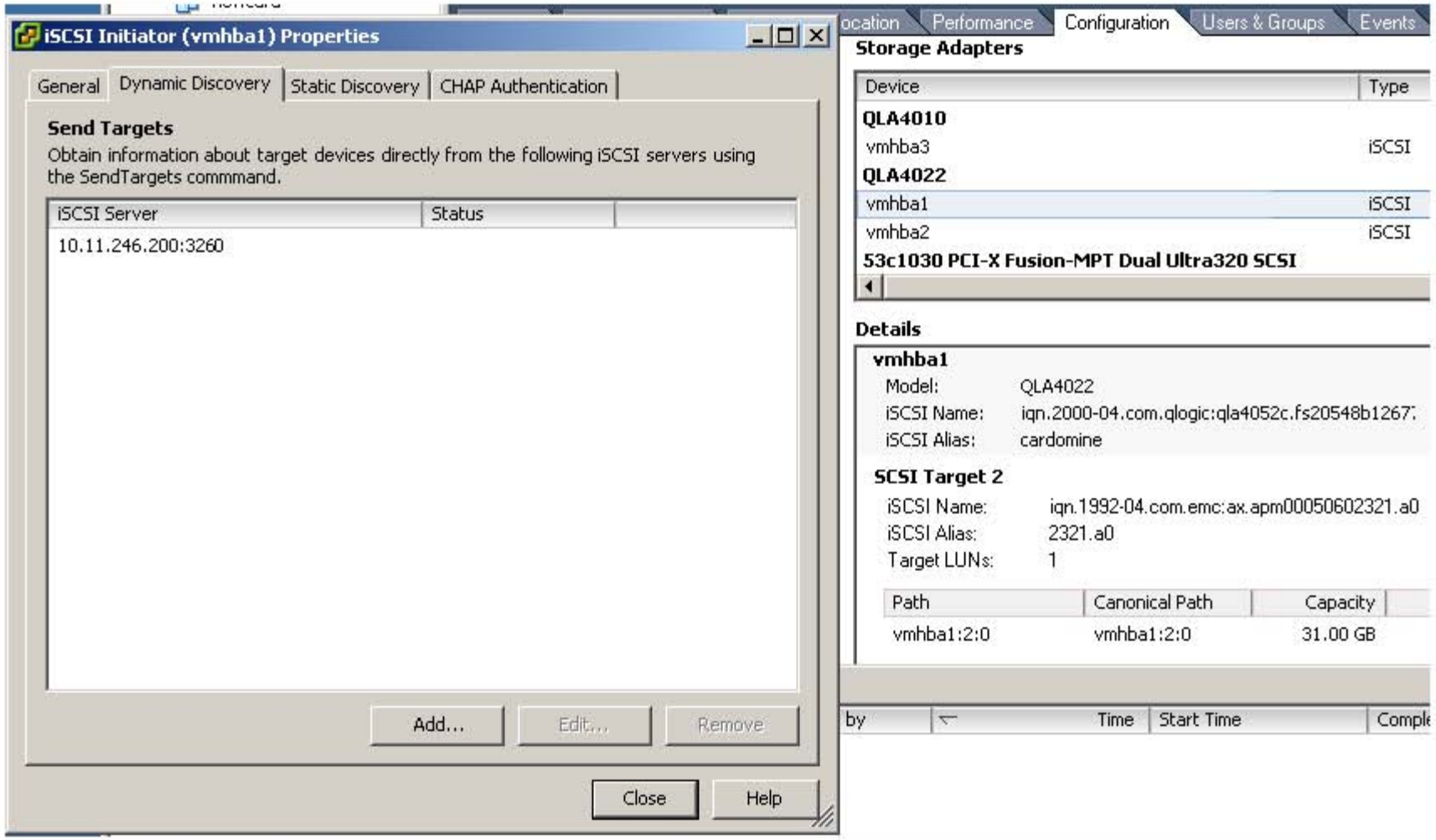
# iSCSI Configuration

- 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

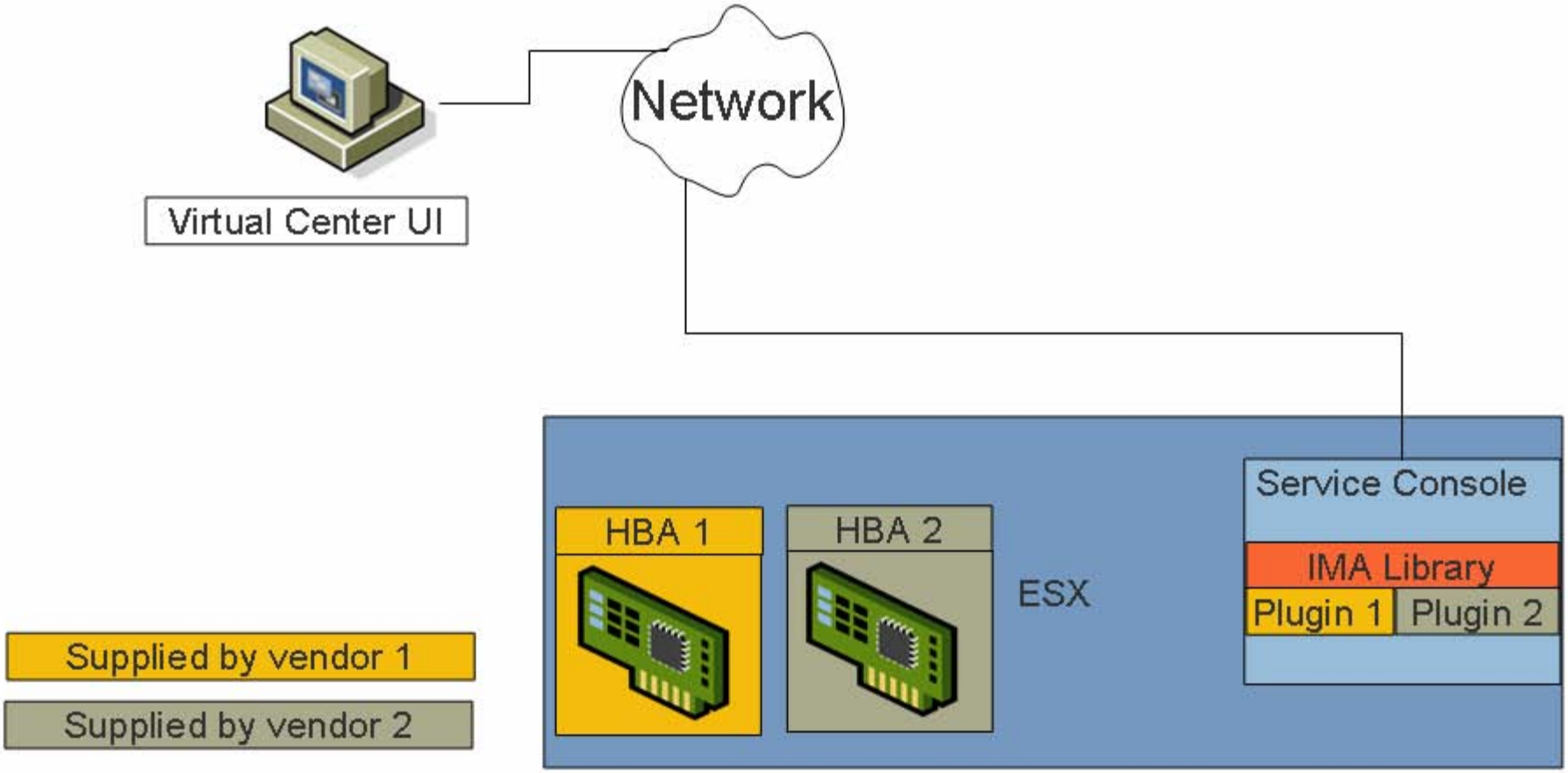
# iSCSI Configuration



# iSCSI Configuration



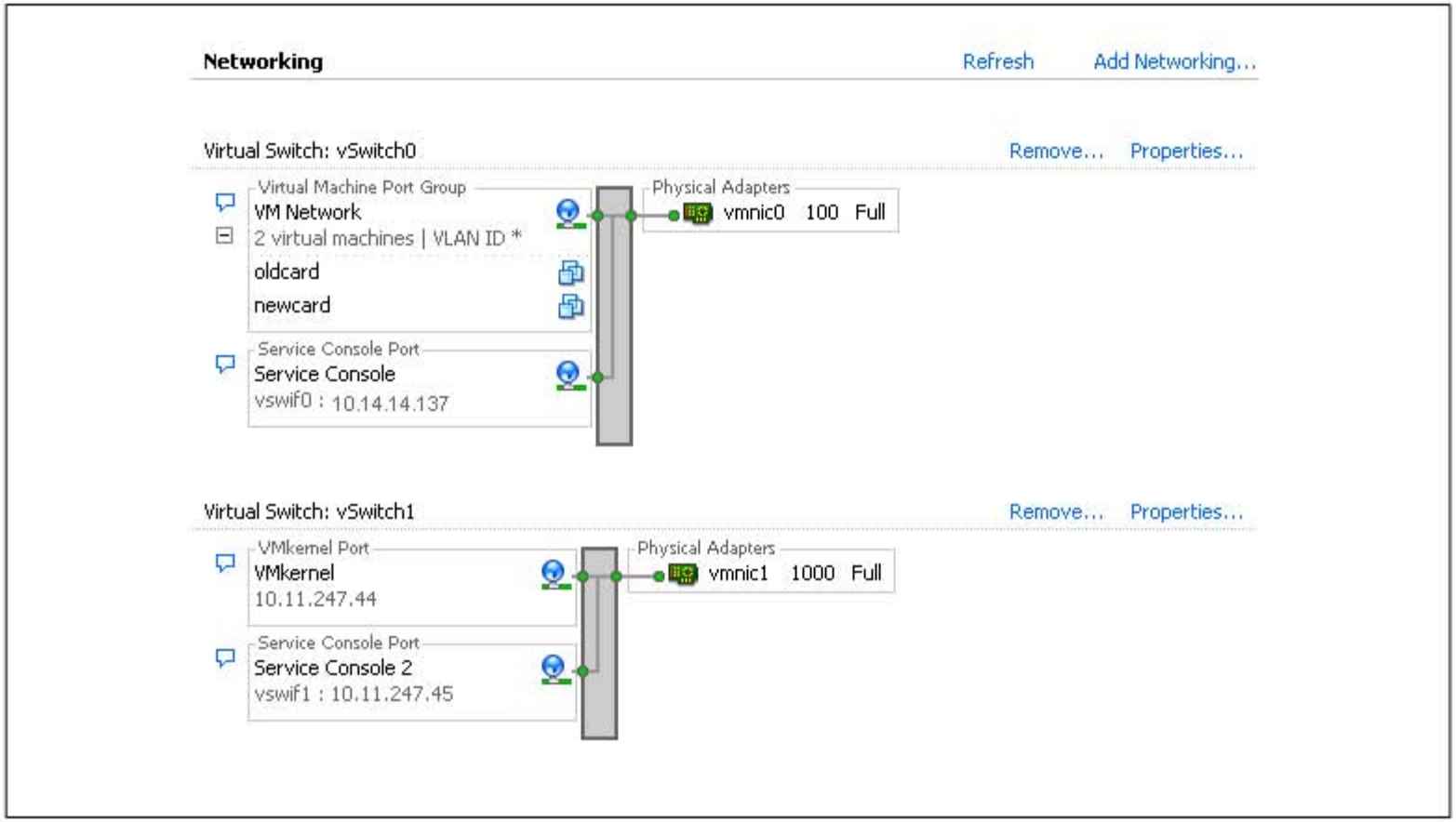
# IMA Plugin Overview



# 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

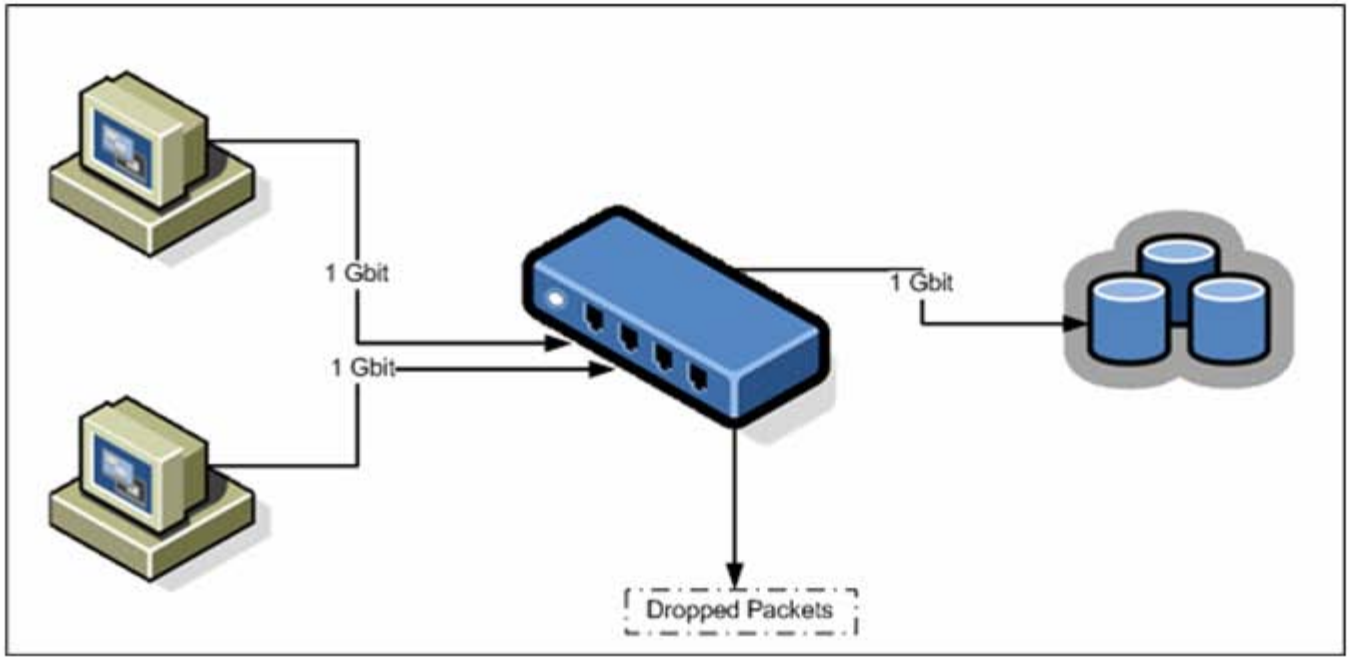
# NFS Configuration



# 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

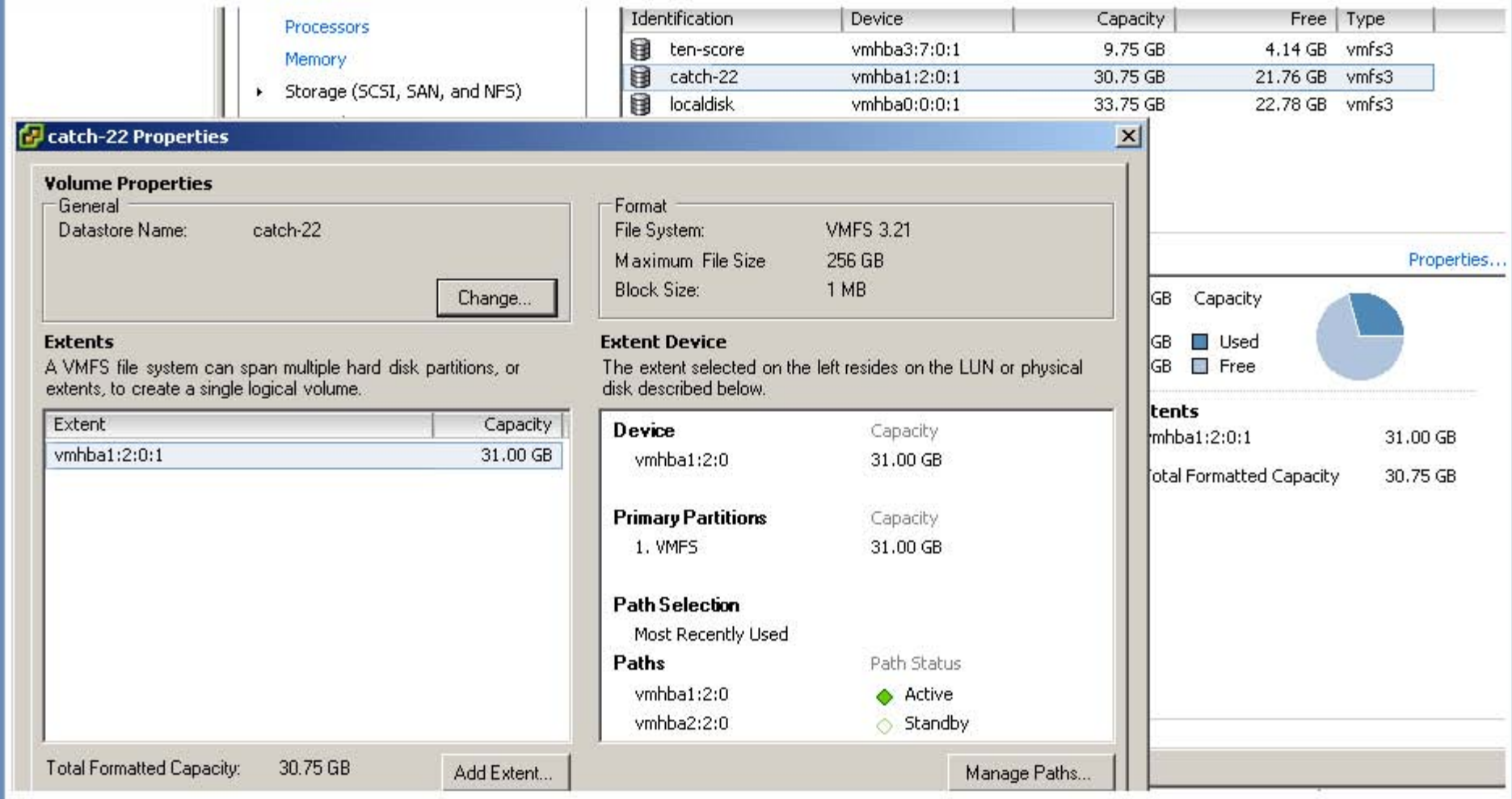
```
*: interface is up
IHQ: pkts in input hold queue       IQD: pkts dropped from input queue
OHQ: pkts in output hold queue      OQD: pkts dropped from output queue
RXBS: rx rate (bits/sec)            RXPS: rx rate (pkts/sec)
TXBS: tx rate (bits/sec)            TXPS: tx rate (pkts/sec)
TRTL: throttle count

  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



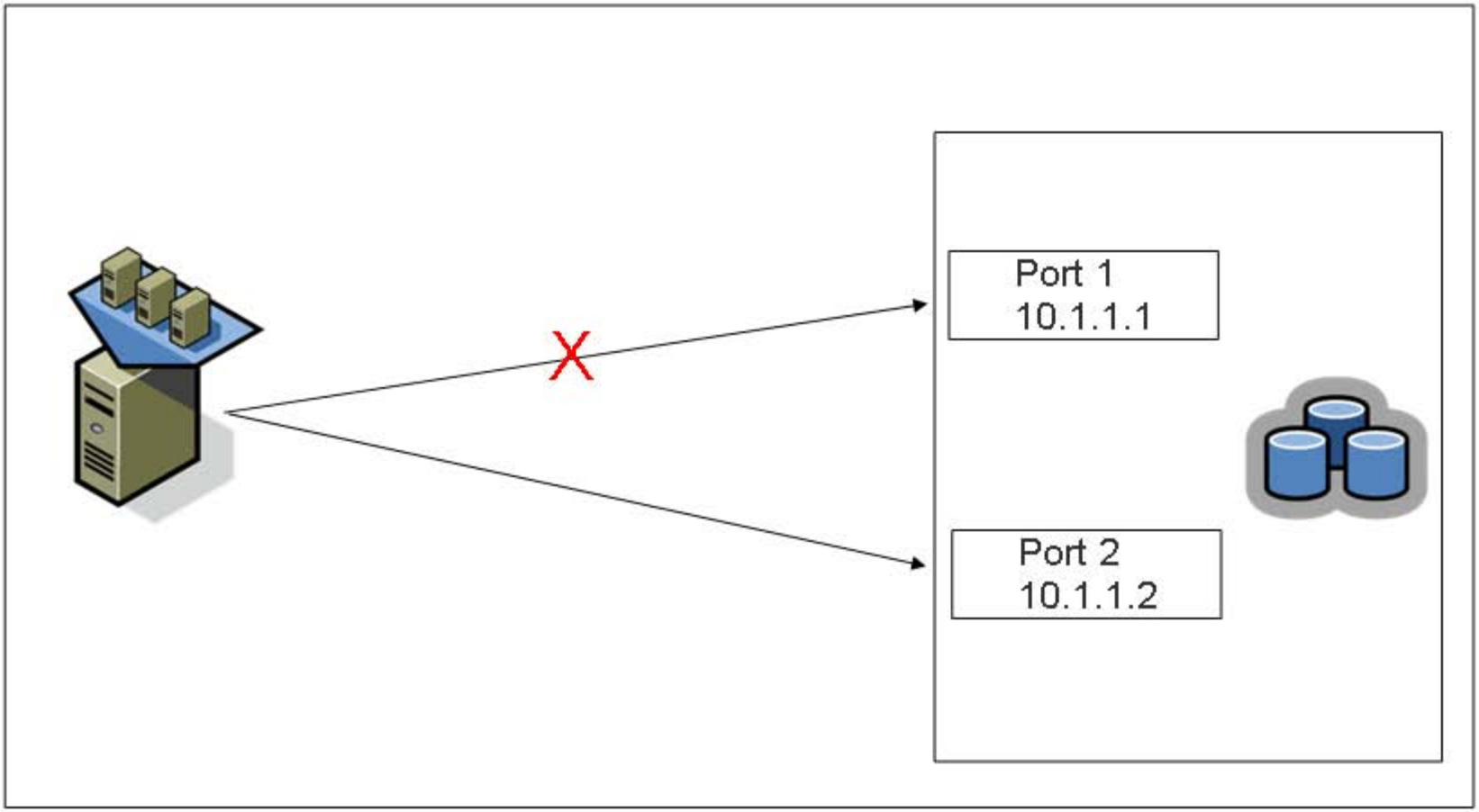
# Network Multipath with NIC Teaming

- 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