

VMware ESX Server as a foundation for HA and DR for the Microsoft Server Platform

Rick Magoon

Senior Solutions Architect, EMC

Microsoft Practice



VMWORLD 2006

Agenda

- Definitions and Acronyms
- Introduction
- HA Layers
 - Server Hardware
 - SAN Hardware
 - Network Backbone
 - VMware Virtual Layer (HA, DRS, VMotion)
 - Applications within Windows Guest OS
- DR Layers
 - VMware Virtual Layer (VCB, snapshots, clones, templates)
 - Applications within Windows Guest OS
 - Electronic Archiving
 - Deep Archiving
- COOP
 - Replication
 - Data Integrity
 - Failback
 - Downtime
- Business Continuity
- Architectural Considerations

Definitions and Acronyms

- HA = High Availability (\$\$)
 - First line of defense
 - The use of redundant hardware and software technologies to provide automatic non-disruptive continuation of services in case of failure
- DR = Disaster Recovery (\$)
 - First or second line of defense
 - Allow for restoration of specific services and data in case of failure, deletion or corruption
- COOP = Continuation of Operations (\$\$\$)
 - Third line of defense
 - Remote failover of critical services to offsite datacenter in case of disaster

Introduction

- Too many data centers have unstable Microsoft environments because High Availability and redundancy was never architected prior to their server farm build-out
- Many companies are unaware of potential faults and lack of redundancy within their own server environment until something happens
- High Availability and redundancy is never a one-size-fits-all solution
 - A major undertaking would be required to experience complete high availability in a physical datacenter. A total automation of server hardware, data availability and application availability would be required
 - Tiers must be established based on criticality of network service to see any ROI

Introduction

- A sample architecture for a highly reliable physical datacenter may include the following components:
 - Microsoft clustering for core applications like Exchange and SQL
 - Load balanced front-end servers like Outlook Web Access (OWA) and Sharepoint web servers
 - SAN software to support snapshots and/or cloning
 - SAN agents that integrate with Microsoft back-end services
 - Cluster aware agents that support database applications like SQL and Exchange
 - Cold standby servers for all other applications that cannot be load balanced or clustered
 - Strict maintenance and patch management for all redundant servers to ensure successful automated failover

Introduction

- More resources would be required for a redundant datacenter in a DR or COOP scenario. In addition to the previous items, the following would be needed for a redundant remote datacenter for critical services
 - Duplicate server hardware
 - Maintain patch management on servers located at DR site
 - Expensive remote replication or mirroring software either host-based or SAN-based
 - For SAN-based data replication, a Fibre Channel to Ethernet bridge (FCIP) is required to transfer storage traffic over the WAN link
 - Expensive and complicated geographic clustering technology to fail-over critical applications to DR site
- In addition to the major requirements stated above, the equipment can consume significant resources like rack space, power, server hardware, and cabling
- Virtualization encapsulates your server infrastructure to reduce costs and provide new methods for performing HA, COOP and DR

HA Layers: Hardware

- Hardware is the most common form of HA that we take for granted in our enterprise equipment. It is the first line of defense
- Server Hardware
 - RAID
 - Dual power
 - Dual NICs
 - Redundant fans
 - Dual HBAs
- SAN Hardware
 - Global hot spares
 - Multiple data paths
 - Redundant storage processors
 - RAID
 - Redundant SAN fabrics
 - Multiple HBAs
- Network
 - NIC teaming
 - Load balancing
 - Redundant switches

HA Layers: Virtual Infrastructure 3

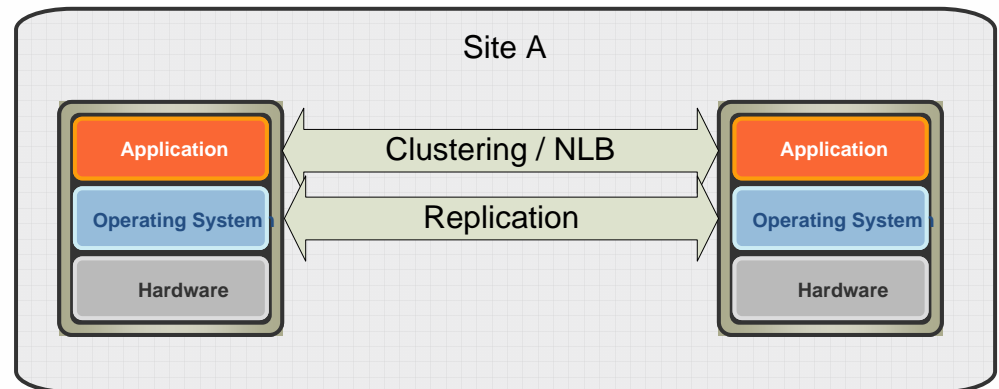
- Encapsulates all guest operating systems with additional HA not otherwise obtainable in a non-virtualized datacenter.
High SLAs, minimal downtime regardless of the network service
- VMotion provides guest OS level redundancy by moving running VMs between ESX hosts. It requires manual operation if not combined with VMware DRS
- VMware DRS (Distributed Resource Scheduler) will automatically migrate a running VM with VMotion if CPU and/or memory thresholds are reached
- VMware HA allows for automated failover in case of VMware ESX host failure. It provides N-to-N failover capabilities for a cluster
- Snapshots – Allow snapshot management for virtual machines that can either be manually triggered or scripted. This allows for easy retrieval of previously known good VM configurations

HA Layers: Application Level HA

- Since VMware HA technologies does not have insight into the guest OS, the above problems will go unnoticed without the addition of Guest OS HA software
- VirtualCenter does not accommodate the following problems in a Windows guest operating system:
 - Operating system failures
 - Bad patches
 - Bad driver upgrades
 - Viruses and spyware
 - Application faults
 - Data corruption

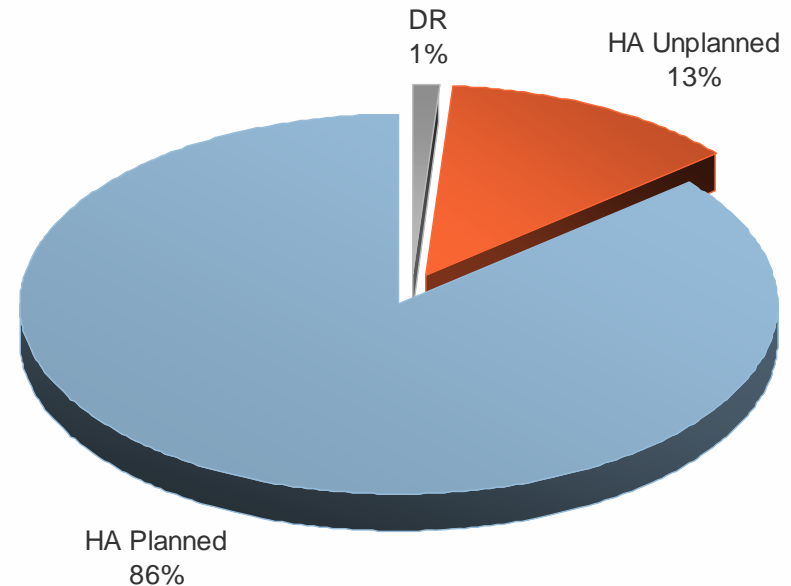
HA Layers: Application Level HA

- If additional application protection is required, there are built-in solutions in the Enterprise Edition of Windows and 3rd party vendors.
- All application HA solutions must meet the compatibility requirements of Windows. Some popular solutions are:
- Clustering Technologies (Multiple nodes on shared disks):
 - Microsoft Cluster Services
 - Veritas Cluster Server
 - NSI GeoCluster
 - EMC Autostart
- Load Balancing (Front end network services):
 - Windows NLB
 - BigIP
- Host Based Replication (Hot standby):
 - NeverFail
 - EMC Replistor



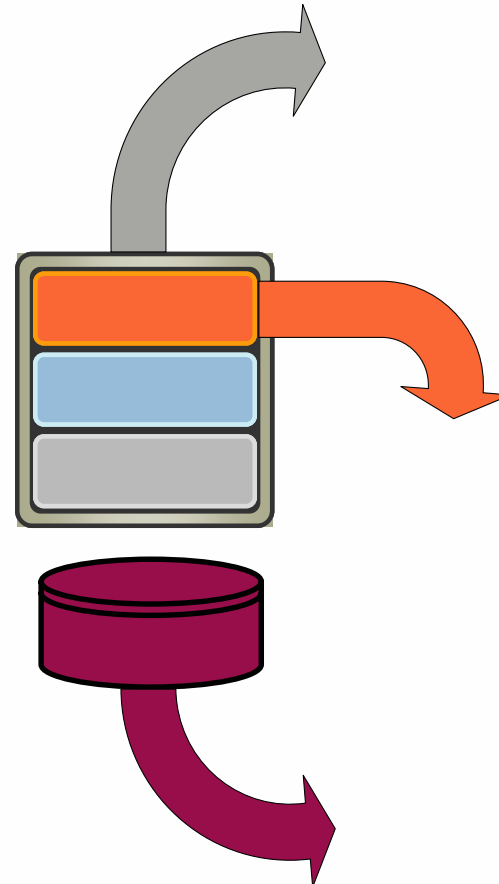
DR Layers

- DR is your insurance for situations like flood, fire, earthquake and contaminated buildings
- Occurs less than 1% of the time
- Tiers of services should be set to establish lowest cost possible for a suite of DR solutions
- Consistency between all VMs is key to usable data
- DR solutions measured in RPO and RTO



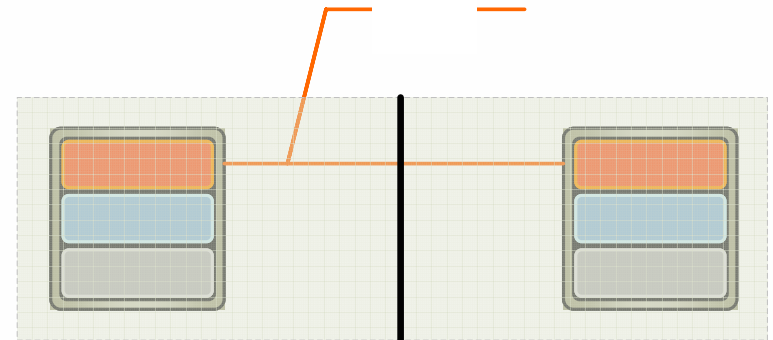
DR Layers

- DR solutions available on SAN, VM and Windows OS/Application layers
 - Multiple layers are required for a complete DR solution
- To restore one or more VMs you must use a technology that utilizes the virtual layer
- To restore application objects or files you must use backup software that recognizes the Windows guest OS
 - VMware Consolidated Backup on a proxy server will mount virtual disks to separate backup server



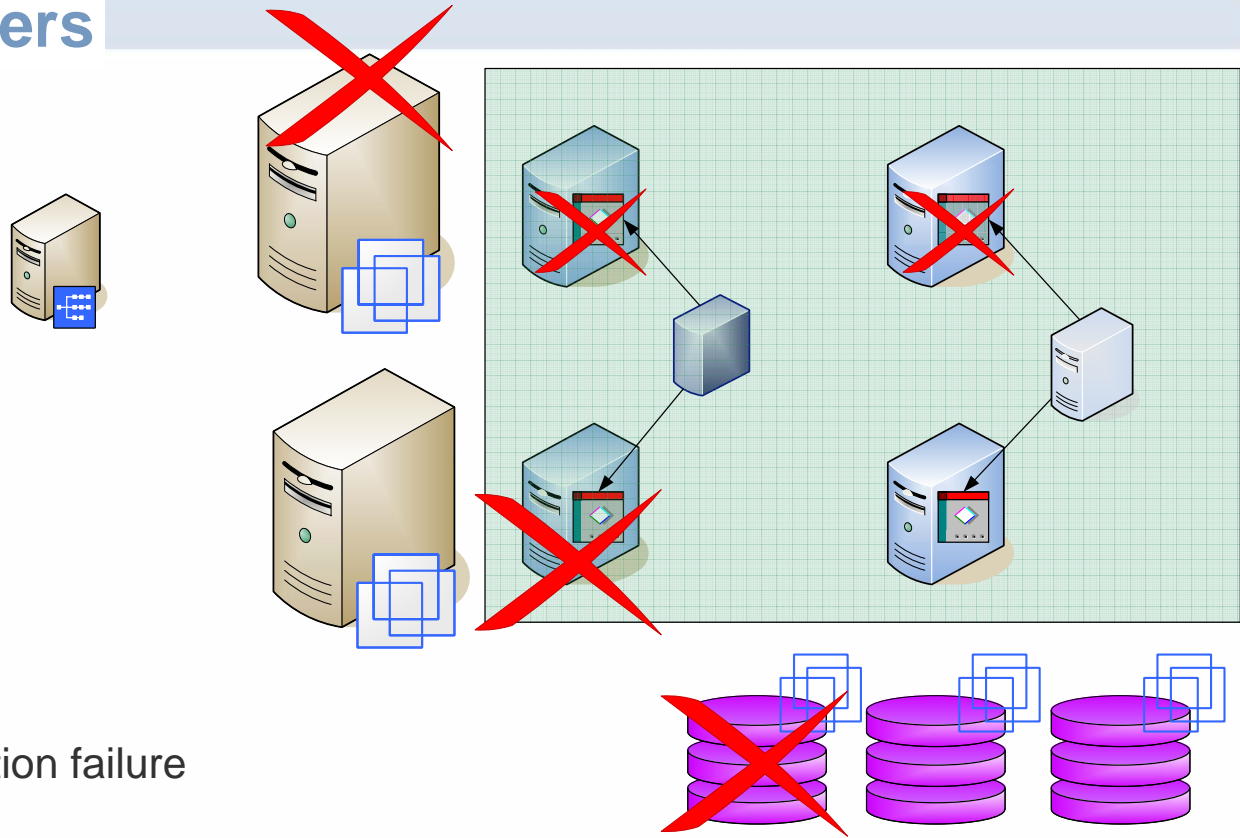
COOP

- Continuity of Operations allows for continuation of network services during the time of disaster
 - It usually represents another datacenter that is triggered with a series of steps for manual failover
 - Only critical services are typically replicated to lower overall costs associated with bandwidth and software
- Because each guest OS is represented by a series of files on the SAN, entire sets of VMs can easily be replicated by technologies within the virtual layer or SAN replication
- Traditional geographic clustering technologies can also be used within the Windows guest OS for a more automated failover

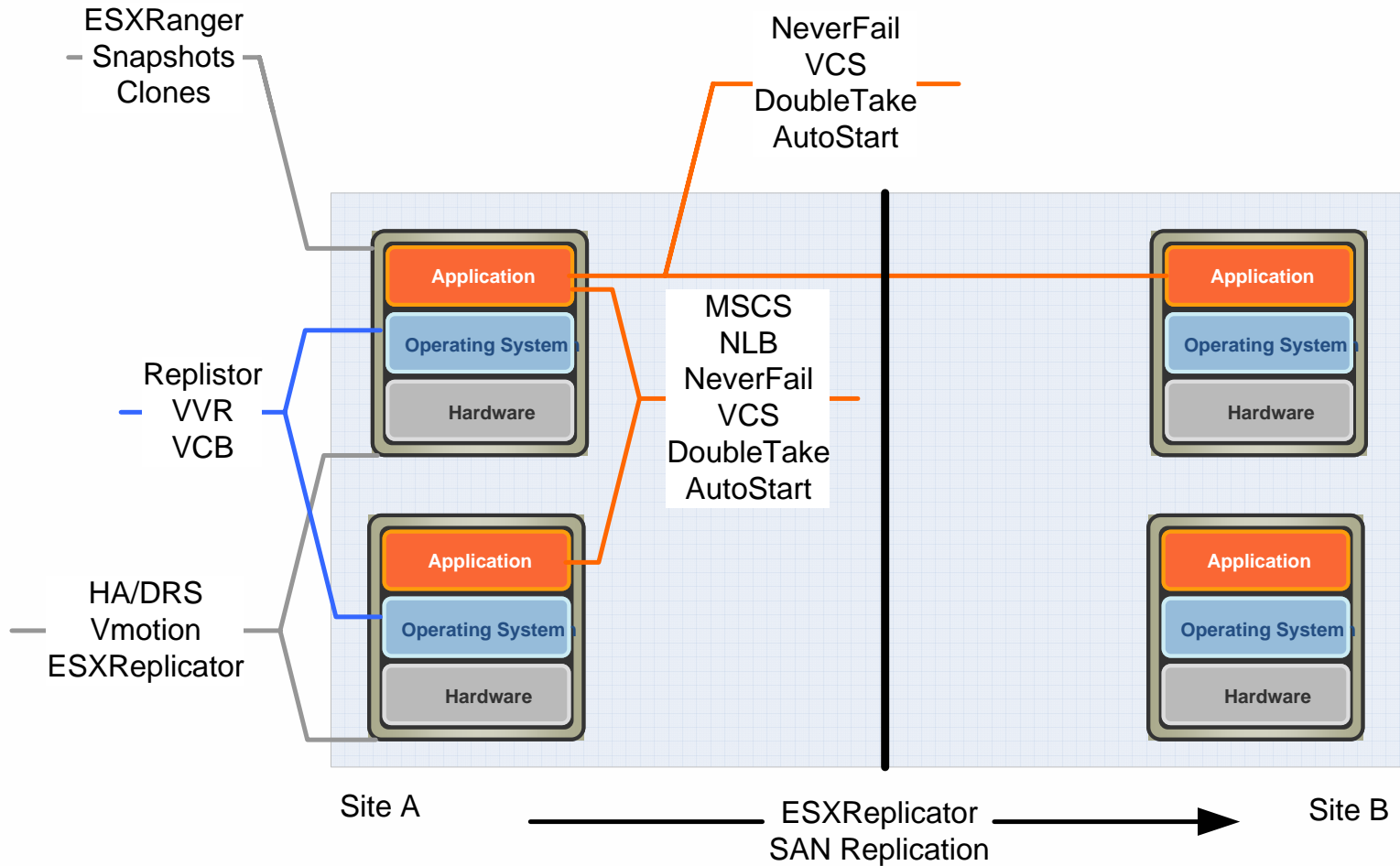


VMware HA Layers

- A – ESX Host Failure
- B – VM Failure
- C – Guest OS application failure
- D – VMFS3 or LUN Corruption
- E – Clustered Application Failure



Summary of HA/DR/COOP Options

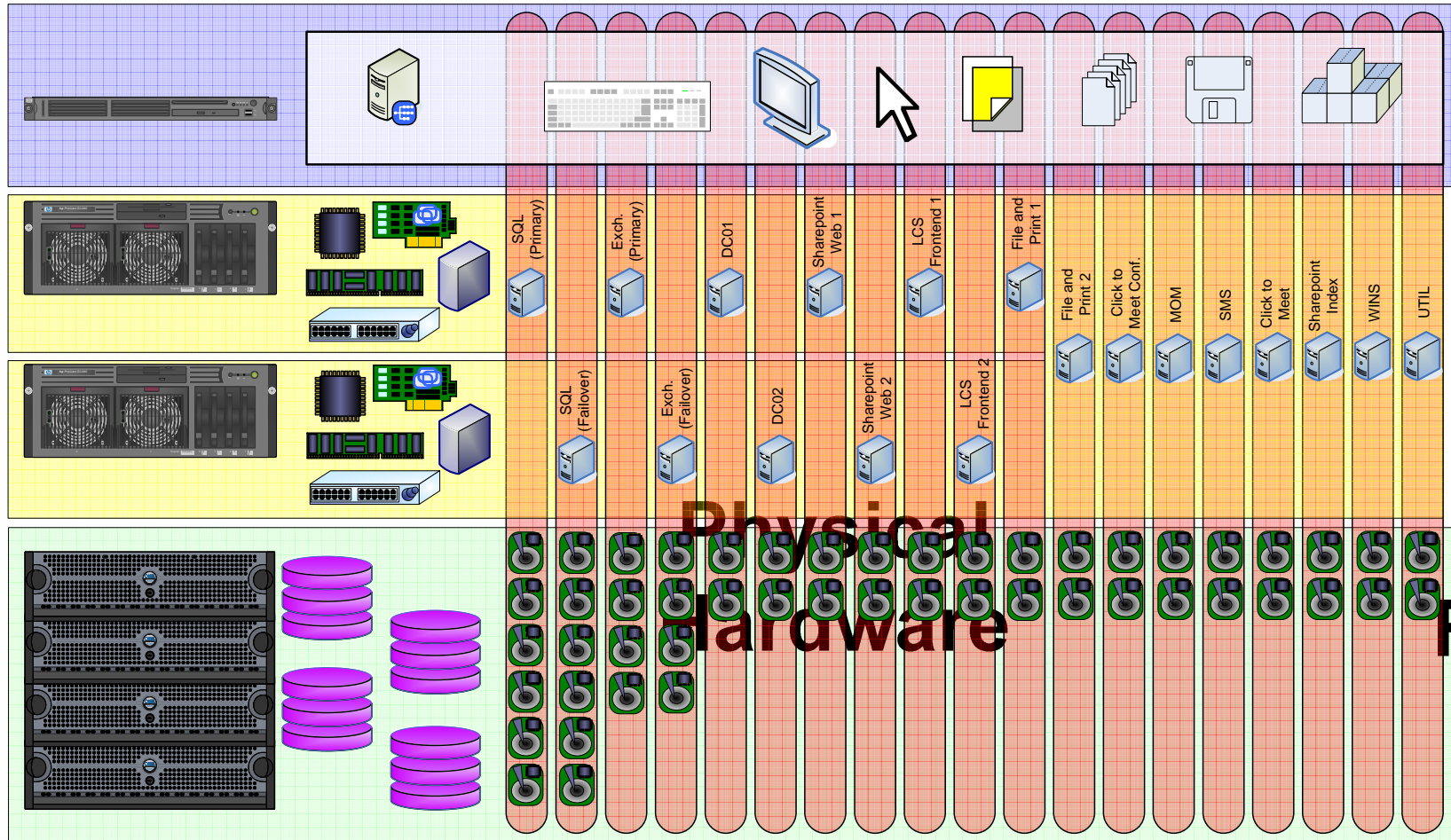


Architectural Considerations: Virtual Disk Layout

■ VMFS3 Volumes

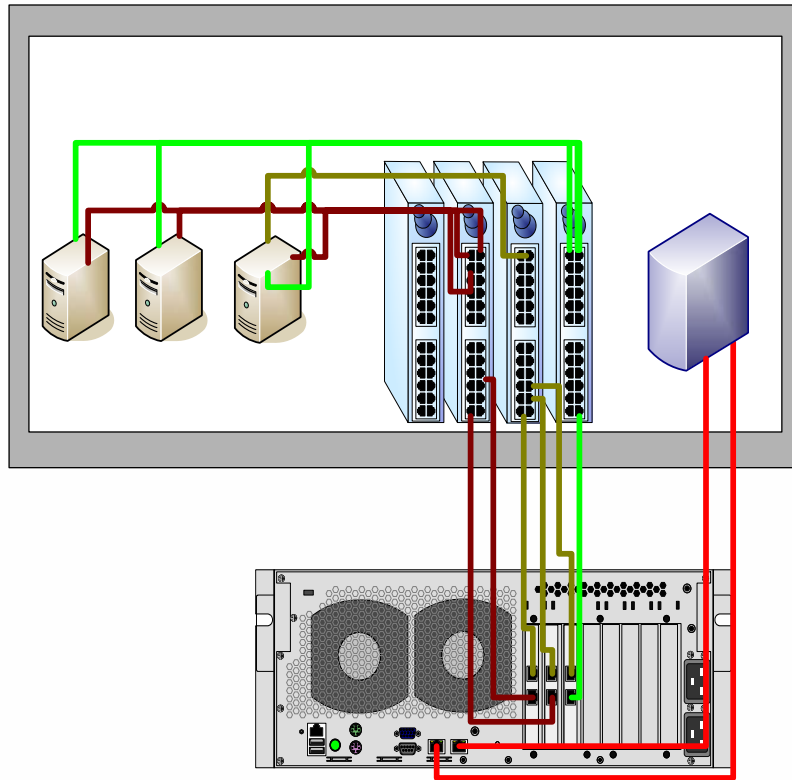
- Keep the size of your virtual disks to a minimum.
 - More fragmentation
 - Longer to backup / restore
 - Limited portability
 - You can always expand later.
 - Downtime required with native ESX commands
 - No downtime with products like Veritas VVM or MS Dynamic Disk)
- Keep the number of disks per LUN/RAID Group to a minimum. This will affect performance and flexibility to perform SAN replication
- Position VMs so clustered and NLB nodes are on separate ESX hosts
 - DRS and HA should not affect positioning of clustered and NLB nodes
 - DRS has anti-affinity rules to prevent cluster and NLB nodes from running on the same host
- If Microsoft Clustering will be used, architect your SAN storage with raw LUNs for shared storage and use raw disk mappings (RDM)

Architectural Considerations: Virtual Disk Layout



Physical
Resources

Architectural Considerations: Virtual Networking



- Always purchase the most Ethernet ports that you can afford
 - VI3 has a limit of 8 physical ports
 - Use port groups if you have a limited amount of ports or your server has limited expandability
 - If limited expandability, consider 4-port Intel Pro/1000 adapters
- Two or more ports should be teamed or load balanced to provide redundancy for critical virtual switches
 - Assign ports across separate physical adapters if possible
 - Only if they're the same type of adapter
- Service console only requires one port unless you are performing backups from the service console
 - You may wish to dedicate two ports if service console is considered critical
- Having a separate port for VMotion will guarantee the best performance

Architectural Considerations: Microsoft Windows

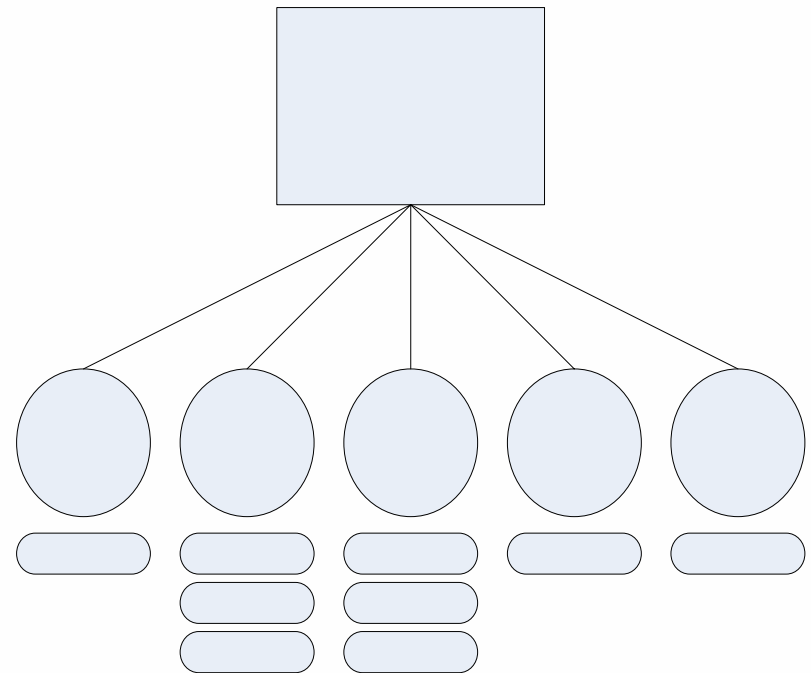
- If you are performing hot backups or snapshots, make sure they are consistent
 - VMware snapshots are not application aware
 - Inconsistent VMs will either not function or potentially destroy your Active Directory environment
 - Domain controllers
 - Live Communications Server
 - Exchange Server
 - SQL Server and any connected database application server
- Microsoft clustering is a good choice for application-level HA but storage must be configured to support clustering (RDMS)
- If snapshots or mirroring is used on the SAN, an application agent must reside within Windows guest OS to maintain data consistency

Summary of HA and COOP Options

Technology	Type	Scope	Applications	Application Aware	Data consistency (When running)	Async/Sync
Hardware	HA	Broad	All	No	No	NA
VMware HA	HA	Broad	All	No	No	NA
VMware DRS	HA	Broad	All	No	No	NA
VMware Vmotion	HA	Broad	All	No	No	NA
Vizioncore ESXRanger	DR	Broad	All	No	No	Async
Vizioncore ESXReplicator	DR / COOP	Broad	All	No	No	Async
VMware Consolidated Backup	DR	Broad	Depends	Yes	Yes	Async
Microsoft Clustering	HA	Focused	Cluster aware apps	Yes	Yes	NA
Microsoft WinNLB	HA	Focused	NLB compatible apps	Yes	Yes	NA
Veritas VCS	HA / COOP	Focused	Compatible Apps	Yes	Yes	NA
Veritas VVR	HA	Focused	Compatible Apps	No	Yes	Async
NSI Doubletake	HA / COOP	Focused	Compatible Apps	Yes	Yes	Async
EMC Autostart	HA / COOP	Focused	Compatible Apps	Yes	Yes	Async
EMC Replistor	HA / COOP	Focused	Compatible Apps	Yes/No	Yes	Async
SAN based replication	DR / COOP	Broad	All	Yes/No	Yes	Async/Sync
VMware snapshots	DR	Broad	All	No	No	NA
SAN snapshots	DR	Broad	All	Yes/No	Yes	NA

Architectural Considerations: Resource Pools

- Group VMs into resource pools that represent AD services
- VMware DRS will allocate VMs to appropriate ESX host to meet resource group requirements
- This feature will optimize your Windows AD environment by giving resources where it matters the most
- All critical VMs should belong to resource pools. Non-critical VMs will not be guaranteed resources



Lessons Learned

- HA, DR and COOP will protect you from unplanned events but preventing operator errors can also help with unplanned downtime and potential corruption
- Lessons Learned:
 - Never extend production VMFS volumes
 - Always zone your fabric
 - This is not necessary for the operation of ESX servers, but will be an issue if Windows servers are SAN-attached to the same fabric
 - This is especially important when installing a Windows server connected to the same fabric(s), because it polls all storage that it sees and will corrupt anything it has access to
 - Never attach a virtual disk file to another VM that's not the original owner of that disk
 - VMware 3.0 uses series of flat files, delta files and snapshot disks to maintain a full logical virtual disk. Any modification to the wrong disk will render your virtual disk file set unusable
 - Revert to snapshot function in VirtualCenter does not have a confirmation dialog box

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

VMWORLD 2006

