Exchange Server Cluster Deployments Best Practices

Bob Combs

Program Manager, Exchange Server
Microsoft Corporation

Agenda

- Overview of Clustering
- Storage Choices
- Exchange 2000 and 2003 Clustering
- Building an Exchange 2003 Cluster
- Building an Exchange 2000 Cluster
- Microsoft Usage of Exchange
- Migrating and Upgrading
- Managing and Monitoring
- Removing an EVS or an Exchange Cluster
- Backup and Disaster Recovery
- Caveats and Resources

What Clustering Is

- Windows 2003 high availability technologies
  - Network Load Balancing (NLB)
  - Microsoft Clustering Service (MSCS)
- NLB
  - Load balances client requests
  - Good for Web farms, etc.
  - Not good for databases
- MSCS
  - Monitors applications
  - Failover when service fails
  - Exchange supports clustering
  - “Shared Nothing” architecture

What Clustering Is Not

- Does not protect against:
  - Shared storage failures
  - Network service failures
  - Application/operational failures or database corruption
  - Site disasters (unless geoclustered)
- Does not provide scalability
- Does not load balance mailboxes
- Cannot move running applications; shared state is lost!

High Availability

- Failover mitigates outages
- Strengthened by fault tolerant design
- Downtime measured in 9s

Availability %	Yearly downtime
90%	876 hours (36.5 days)
95%	438 hours (18.25 days)
99%	87.6 hours (3.65 days)
99.9%	8.76 hours
99.99%	52.56 minutes
99.999% (“5 nines")	5.256 minutes
99.9999%	31.536 seconds

Risk Analysis

- Types of failures
  - Storage failure
  - Network failure
  - Component failures
    - NIC cards
    - RAM
    - CPU
  - System failures
    - Application crash
    - O/S
    - Virus
- Cost of Downtime
  - Productivity
  - Revenue
  - Financial performance
  - Damaged reputation
  - Other expenses
- Level of impact of downtime
  - Minimal; inconvenience
  - Significant; measurable loss
  - Substantial loss
  - Critical business risk

Cluster Terms

- Active/Passive
- Active/Active
- Failover & Failback
- Heartbeat
- Quorum
  - Majority Node Set
- Shared Storage
- Resources and Resource Groups

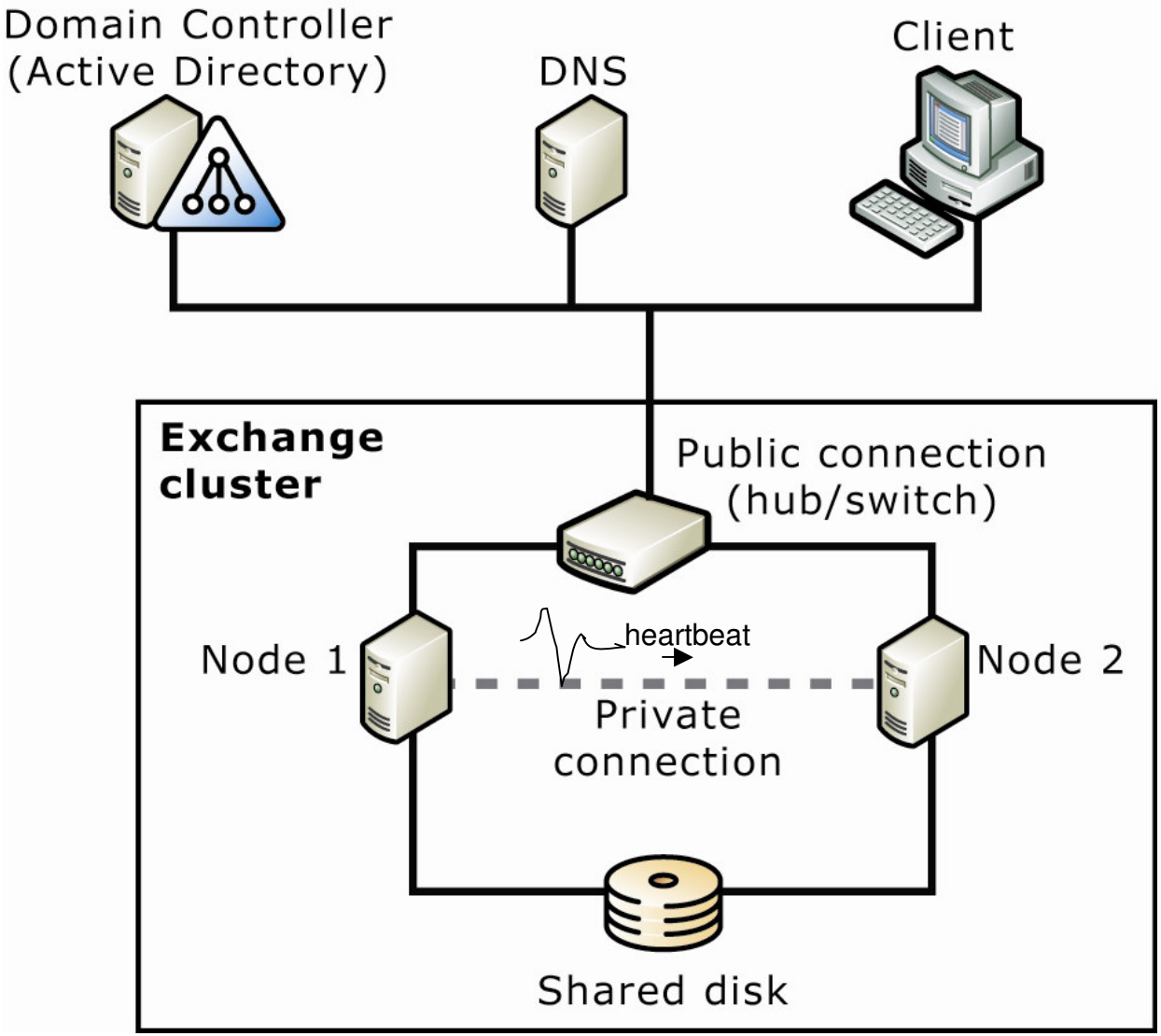
Active/Active and Active/Passive

- **Active/Active**:
  - All of the nodes can have an active EVS (no requirement of spare node);
  - A single node can have multiple EVS at same time;
  - Supported up to 2 nodes only;
  - Single node cluster is A/A.

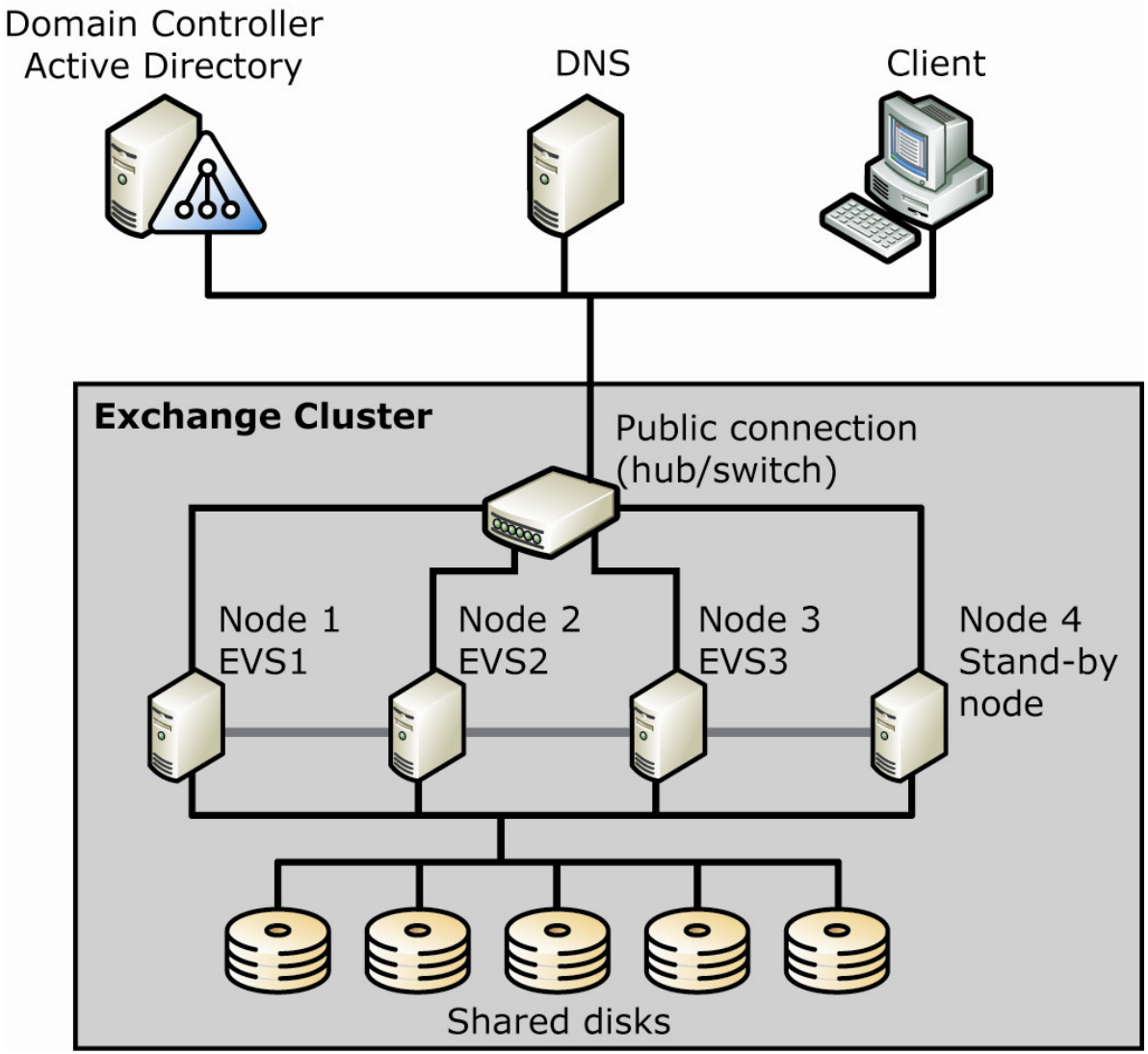
- **Active/Passive**:
  - At least one node is free at any time (spare node);
  - No more than one EVS is active in one node at any time;
  - Enforced when 3 or more nodes (exres fails online for SA resource);
  - Exres tries to stop Exchange services (best effort);
  - 2 nodes and 1 EVS is A/P.

- Windows 2003: Group's common property AntiAffinityClassNames helps failover in A/P. This is not strict, just best effort.

2-Node Cluster

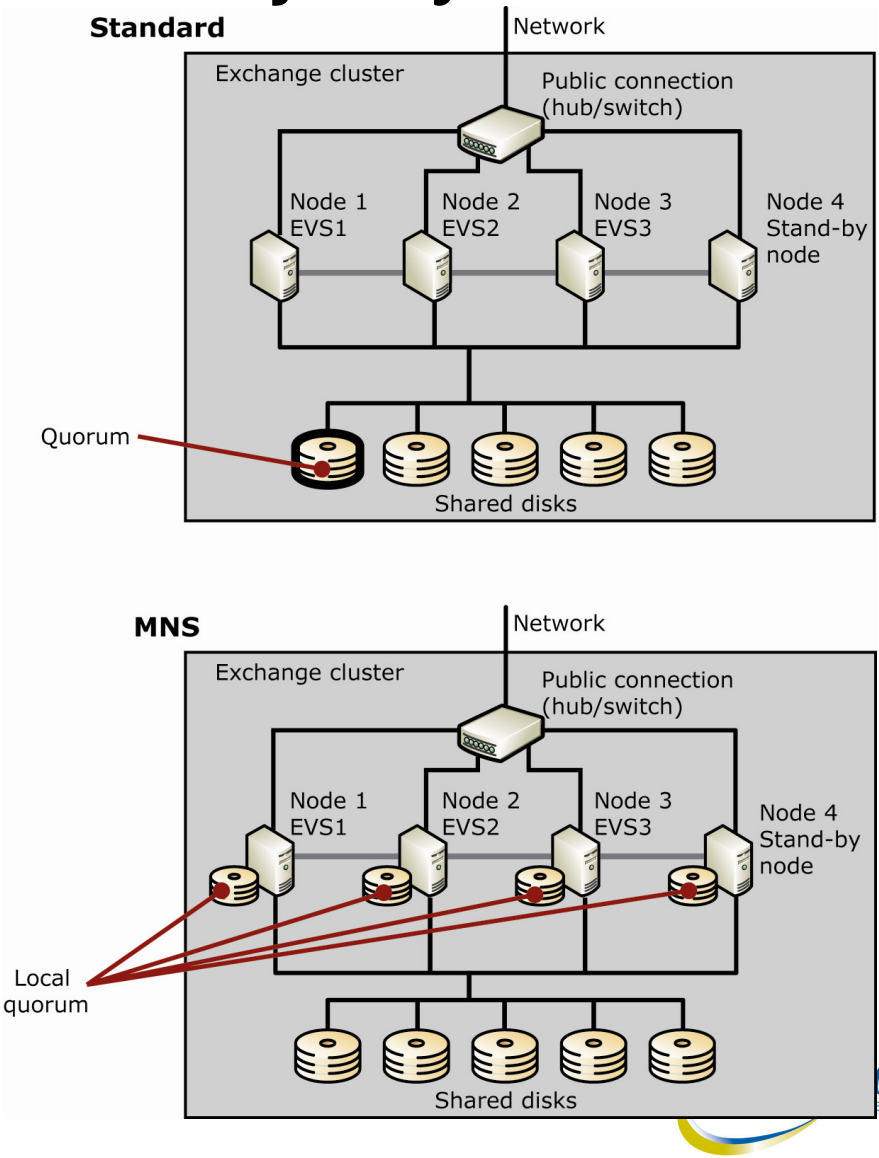


4-Node Cluster



Quorum: Standard vs. Majority-Node Set

- Stores most current configuration data
- Standard Quorum data is in share
- MNS each node has local copy
  - MSCS keeps copies in sync



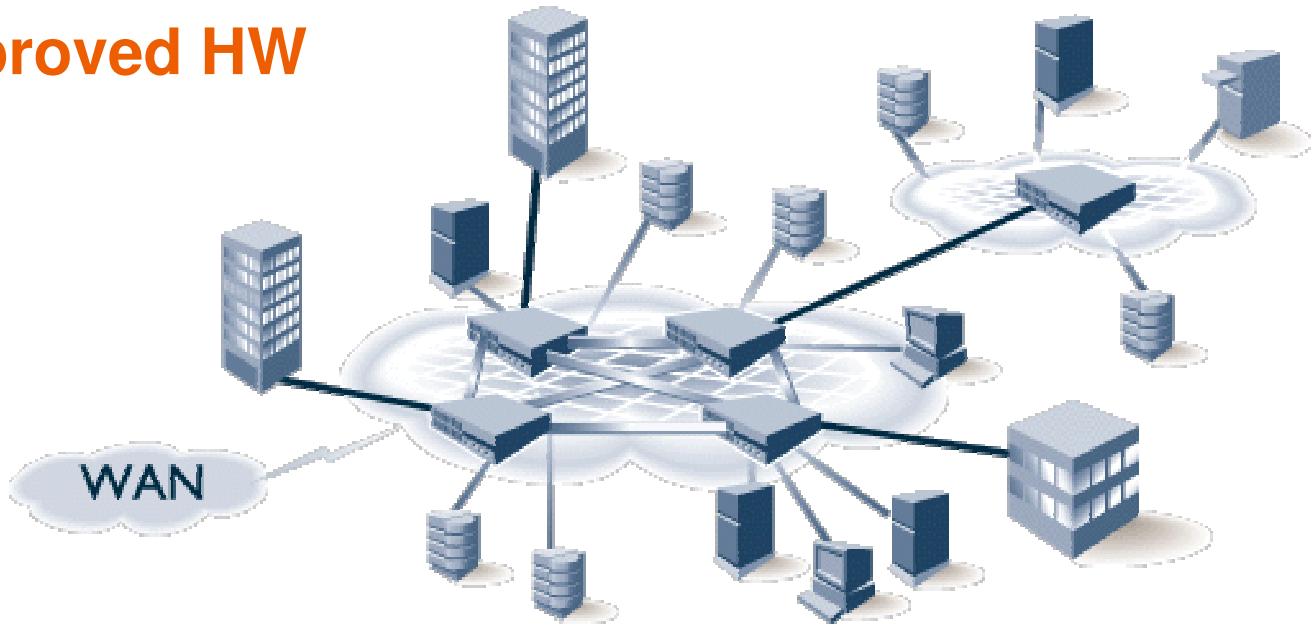
RAID Storage Solutions

- RAID-1: Mirrored disks; optimized performance
- RAID-0+1: Mirrored striped disks; optimized capacity
- RAID-5: Striping w/parity; improved read performance
- If using RAID, use:
  - Separate RAID for .edb/.stm and separate RAID for transaction log files
  - RAID-5 for public folders

RAID solution	Number of drives	Cost	Reliability
RAID-0	10 9-GB disks	High	Low
RAID-1	2 45-GB disks	Low	Low
RAID-0+1	20 9-GB disks	Very high	Very high
RAID-5	11 9-GB disks	High	High

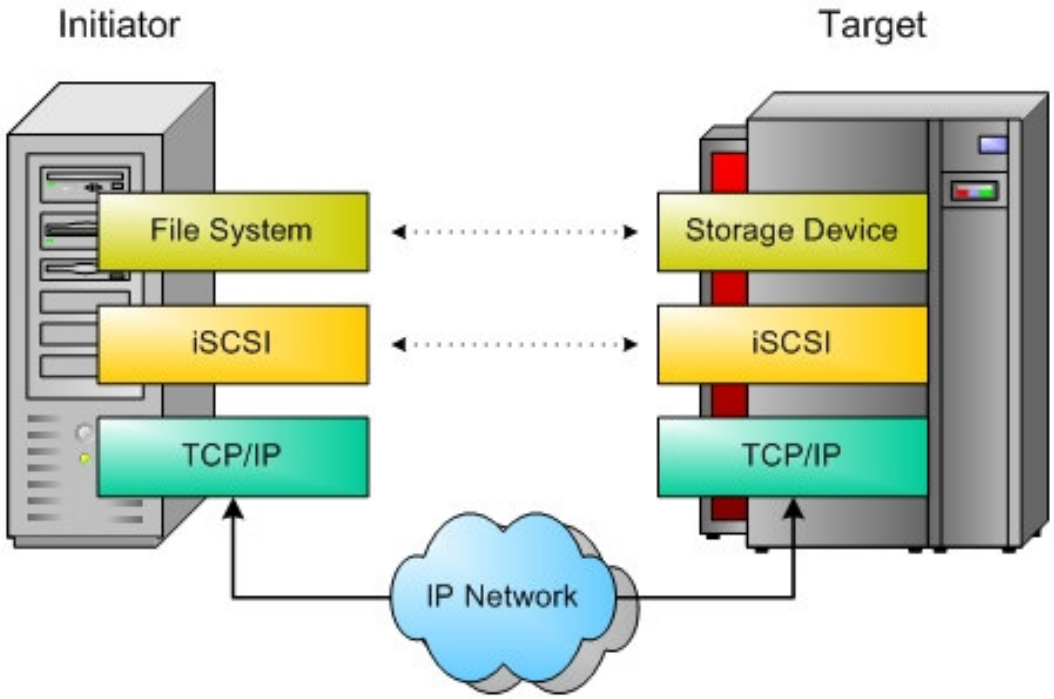
SAN is Recommended

- High I/O Performance and Reliability
- SCSI, Fibre Channel, iSCSI
- SCSI Standards (only SCSI-3 cluster friendly)
- **Use only approved HW**

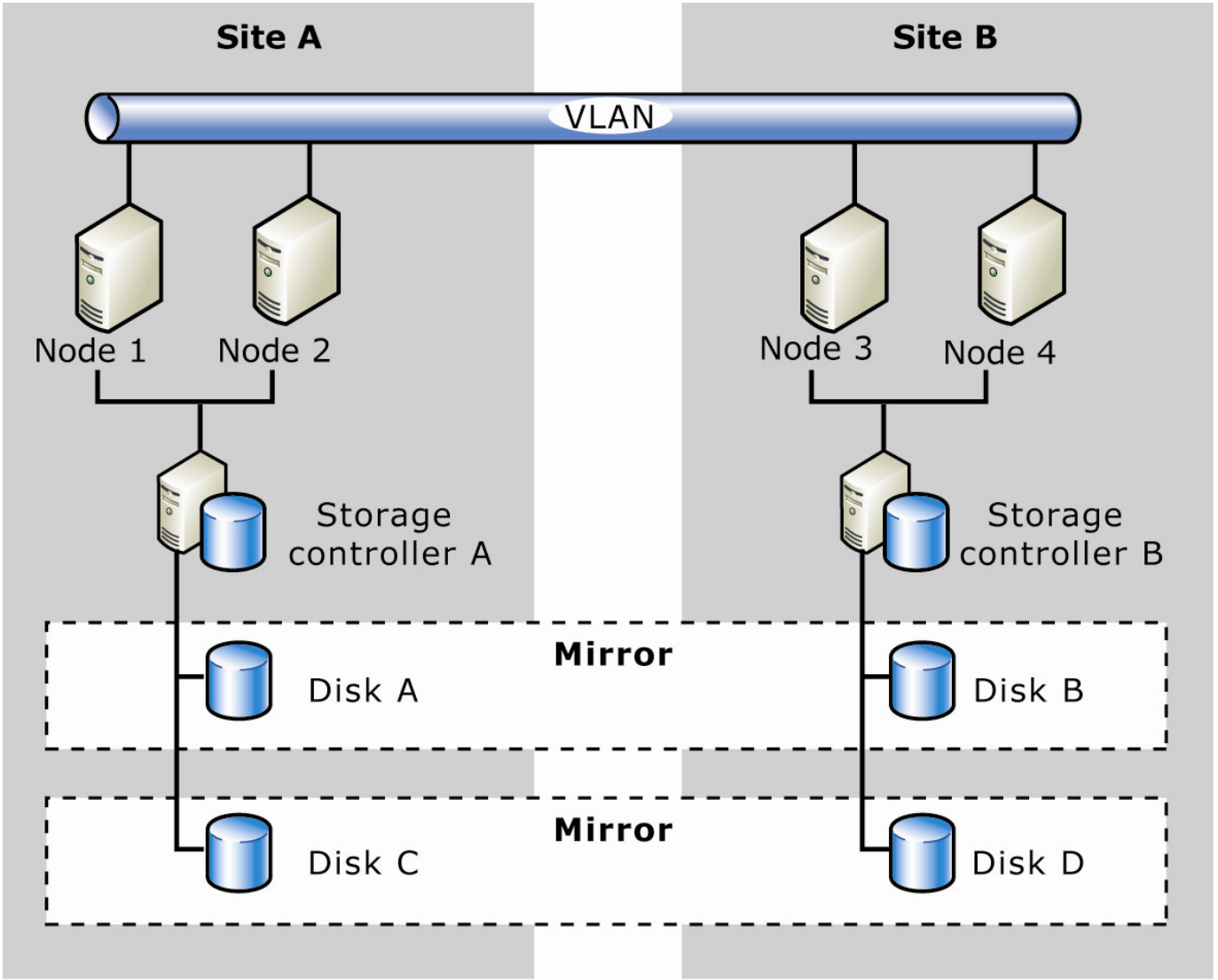


iSCSI

- IP-based SCSI transport
  - Leverages Commodity IP technology
- SAN friendly protocol



Geoclustering



Exchange Clustering

- Exchange Clustering takes advantage of MSCS to provide services that are more available than single-computer server.
- Exchange Virtual Server (EVS) is composed of:
  - one cluster group
  - one or more physical disk resources (SCSI or Fibre Channel)
  - one IP address resource
  - one network name resource
  - several Exchange resources
  - one Exchange server object in the AD
- For clients (Outlook, OWA), a clustered server is indistinguishable from non-cluster server.

Exchange 2000 Server Enterprise Cluster Support

Windows version	Exchange version	Nodes supported
Windows 2000 Advanced Server	Exchange 2000 Enterprise	Up to 2-node
Windows 2000 Advanced Server	Exchange 2000 Enterprise & SP3	Up to 2-node *
Windows 2000 Datacenter	Exchange 2000 Enterprise & SP3	Up to 4-node
Windows 2003 Enterprise	Exchange 2000 Enterprise & SP3	Up to 4-node

* Active/Active only supported on 2-node E2KSP3 configurations

Exchange 2003 Server Enterprise Cluster Support

Windows version	**Number of Nodes**
Windows 2000 Advanced* or Windows 2000 Datacenter* or Windows 2003 Enterprise or Windows 2003 Datacenter	2-node any O/S mix, A/A or A/P
Windows 2003 Enterprise Server or Windows 2003 Datacenter Server	Up to 8-node, A/P

* Windows 2000 must have SP3 & hotfix 329938, or SP4

Exchange Server 2003 Editions

Feature	Exchange 2003 Standard Edition	Exchange 2003 Enterprise Edition
Storage group support	1 storage group	4 storage groups
Number of databases per storage group	2 databases	5 databases
Total database size	16 gigabytes (GB)	Maximum 8 terabytes, limited only by hardware
Exchange Clustering	Not supported	Supported when running Windows Server 2003, Enterprise Edition or Windows Server 2003, Datacenter Edition

What Exchange 2003 Added

- 8-node clustering
- Improved cluster failover time
- Overcome 26-drive letter limit
- Restore storage group
- IPSec between FE & BE
- FE & BE Kerberos authentication
- /USERVA memory switch (W2k3)
- RPC over HTTP for Outlook 2003 (W2k3)
- Outlook 2003 performance (MAPI compression)
- Cross forest Kerberos authentication
- OWA UI improvements
- OWA attachment blocking
- Intelligent Message Filtering
- Real-time Safe & Block lists
- Inbound recipient filtering
- DL caching improvements
- PF replication improvements
- Mbx recovery tool
- Suppression of OOFs to DLs
- Restrict DL to authenticated users

Requirements For Clustering Exchange 2003

- Windows 2000
  - Advanced Server
    - 2-node Active/Passive or 2-node Active/Active
  - Datacenter Server
    - 2-node Active/Active
    - 2-node, 3-node, 4-node Active/Passive
  - Windows 2000 must have SP3 & hotfix from 329938 or SP4
- Windows Server 2003
  - Enterprise Edition and Datacenter Edition
    - 2-node Active/Active
    - Up to 8-node Active/Passive

Requirements For Clustering Exchange 2003

- Exchange Cluster Models
  - Active/Passive is the STRONGLY preferred model
    - Fewer EVS than number of nodes
  - Active/Active
    - Number of EVS' equal to or greater than number of nodes.
    - Limit number concurrent MAPI users per node to 1,900
    - Limit avg CPU utilization on each node to 40%
    - Two instances of store running in one memory space; not enough contiguous VM to bring resource online
- Exchange Virtual Server Limits
  - With two nodes, you can have up to two EVS'
  - With three or more nodes you can have *n-1* where *n* = number of nodes in cluster

Not Supported with Exchange 2003 Clustering

- ADC
- Exchange Connectors
- Exchange Event Service
- Foreign Mail System Connectors
- NNTP
- Site Replication Service

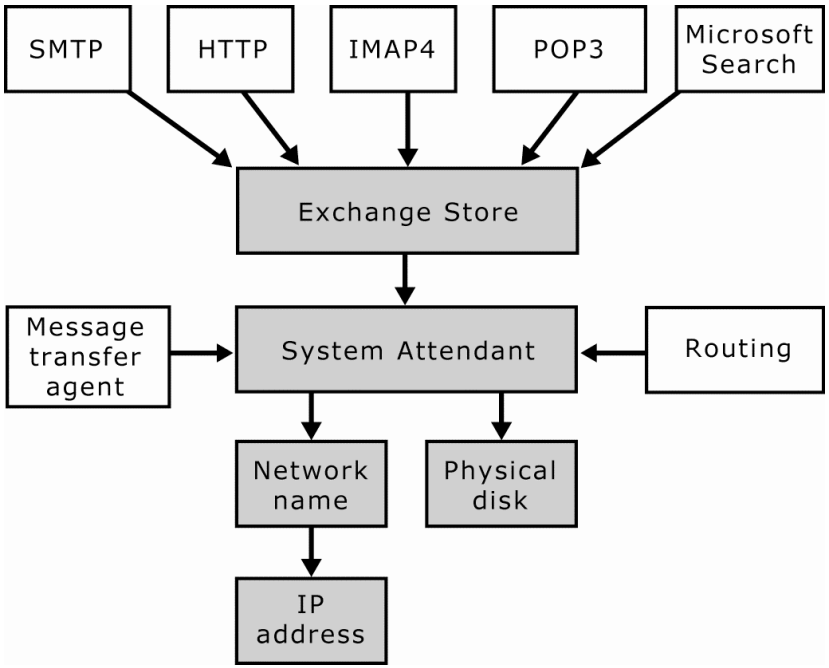
Requirements For Clustering Exchange 2003

- HCL hardware only (HCL/Windows Server Catalog – Cluster or Geographic Cluster)
  - http://www.microsoft.com/windows/catalog/server
  - SCSI or fibre channel external storage
  - **Identical hardware for all nodes**
- OS – 32-bit only
  - Windows 2000 Advanced Server
  - Windows 2000 Datacenter Server
  - Windows Server 2003 Enterprise Edition
  - Windows Server 2003 Datacenter Edition
    - Microsoft Distributed Transaction Coordinator (MSDTC) installed (COMCLUST).
- Exchange Server 2003 Enterprise Edition

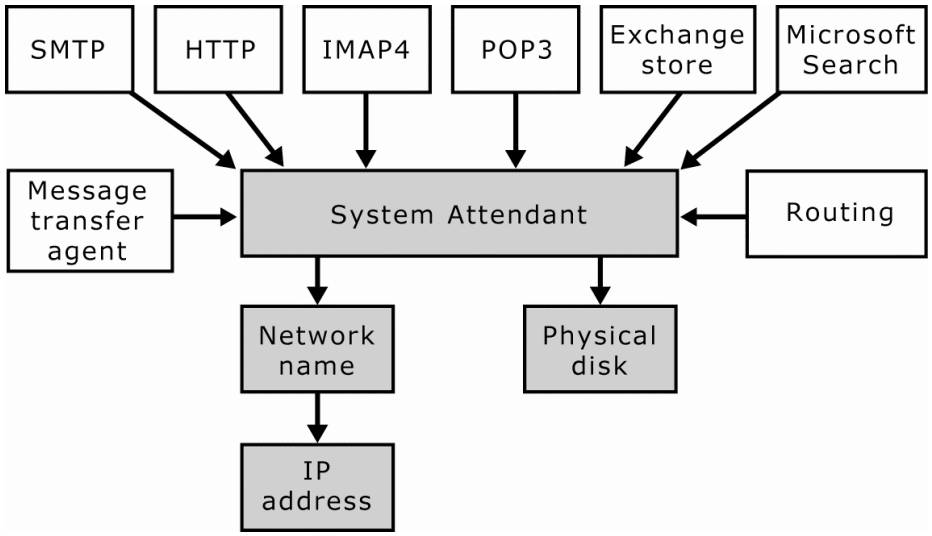
Dependency Hierarchy

- Flattened dependency tree = Faster fail over

Exchange 2000

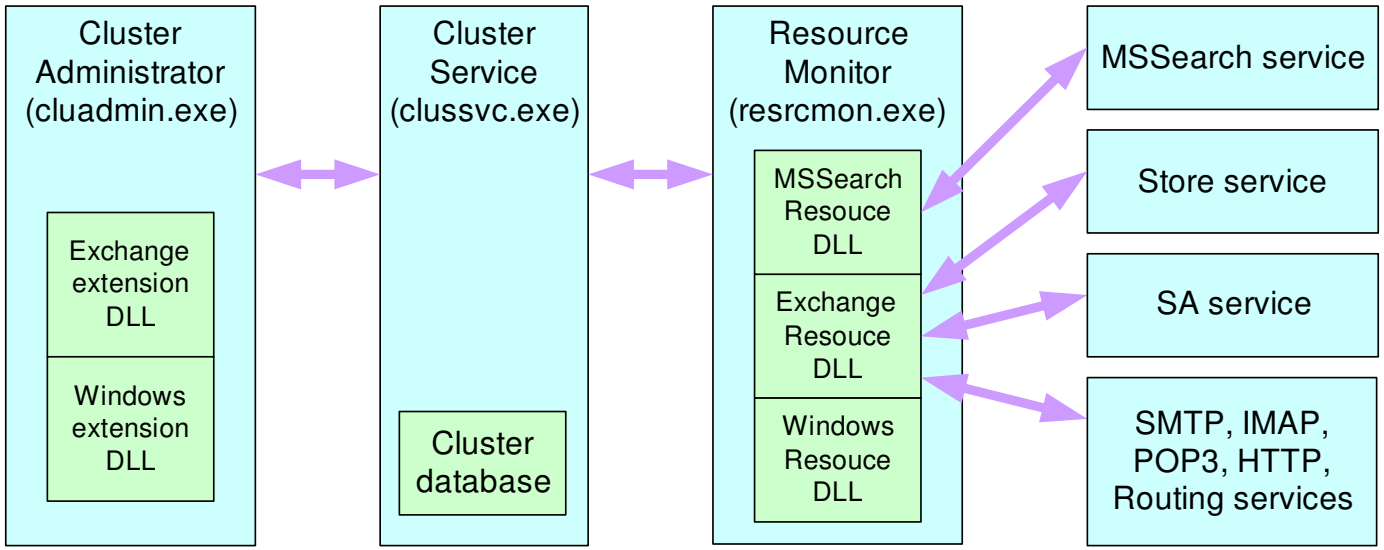


Exchange 2003



Clustering Software Components

- Cluster database
- Cluster service
- Resource monitor
  - Accepts extensions known as resource DLLs
- Cluster Administrator tool
  - Accepts extensions known as Cluster Administrator extension DLL



IsAlive and LooksAlive

- Cluster service needs to know state of resources in order to perform ‘corrective’ actions (stop dependent resources, retry, failover, etc.)
- Cluster service learns state of resource by calling these functions for each resource:
  - **IsAlive**: meant to be more detailed but less frequent check.
  - **LooksAlive**: meant to be less detailed but more frequent check. Cluster service can avoid this call if resource DLL promises to signal event object as soon resource is failed.
- Interval between checks is controlled by a property of each resource. Defaults are:
  - IsAlive: 60 seconds
  - LooksAlive: 5 seconds

Building An Exchange Cluster

- Design Storage
  - 4 storage group maximum on node
    - 5 databases per storage group
  - RAID arrays (RAID 1 or RAID 0+1)
  - Shared disks must be NTFS
  - Separate Physical Disk resources for logs and databases in same EVS
  - Separate Resource Group for Quorum
  - Volume mount points supported on Windows 2003 (318458); useful for 4/8-node clusters

Building An Exchange Cluster

- Design Network
  - Use multiple networks with dedicated private network for internal cluster communications
  - Do not use teaming or DHCP
  - Need an IP address and Network Name resource for
    - Each physical node
    - The cluster resource group
    - Each Exchange Virtual Server
    - Use consistent naming standards

Building An Exchange 2003 Cluster

- Step 1 - Prepare Hardware
  - Apply latest system BIOS
  - Apply latest device firmware
  - Gather latest software drivers
  - Disable unnecessary hardware

Building An Exchange 2003 Cluster

- Step 2 – Install OS and Pre-Reqs
  - Install desired OS (Windows Server 2003 preferred)
    - SMTP, W3SVC and NNTP services
  - Add Nodes to Domain as member servers
    - DCs no longer supported in Exchange 2003 clusters
  - Windows Update / Security Hotfixes
  - Administration Tools – ADMINPAK.MSI

Building An Exchange 2003 Cluster

- Step 3 – Prepare Nodes for Cluster Service
  - Disable unnecessary services
  - Configure Networks
    - Rename connections: Private Network and Public Network
    - Disable NetBIOS and DNS on private (heartbeat) interface
    - Disable Media Sense on NICs – Hard-code (MSKB 258750)
    - Use 10MBs/Half-Duplex if not sure what speed to use
    - Give private network highest binding order
  - Create/Select cluster service account
    - Domain account w/local Administrator rights on each node
      - Does NOT need Exchange Full Admin role
  - Create Quorum partition on shared disk
    - 50MB min; 500MB recommended
  - Create and format additional disks/arrays

Building An Exchange 2003 Cluster

- Step 4 – Install Cluster Service on each node. Set Quorum Log file to reset at 4096KB.
- Step 5 – Install MS DTC on each node by running COMCLUST.EXE
- Step 6 – Install Exchange 2003
  - Unattended setup not supported
  - Binaries installed locally in same location on each node
  - Install one node at a time and reboot each node when finished

Building An Exchange 2003 Cluster

- Step 7 – Install Exchange 2003 Service Packs and Updates
  - Always update one node at a time
- Step 8 – Create Exchange Virtual Server
  - Create Resource Group
    - Disk Resource
    - IP Address Resource
    - Network Name Resource
    - Exchange System Attendant Resource

Building An Exchange 2003 Cluster

- Step 9 – (Optional) Repeat Step 8 if using Active/Active model
- Step 10 – Configure EVS Resources
  - Increase pending timeout on A/A clusters
- Step 11 – Bring Resources Online

Building An Exchange Cluster

- Extra Steps for FE/BE Topology
  - Create HTTP virtual servers in ESM
  - Create virtual directories to match those on FE server
  - Add HTTP virtual servers to EVS
  - Take EVS offline and restart IIS
- You must create one HTTP virtual server for each FE namespace
- SMTP domain must match users of FE server

Building An Exchange Cluster

- Prior to Putting into Production
  - Test failover policies
  - Test hardware (simulate failures)
  - Stress Tests – LOADSIM 2003
    - Test under heavy network loads
    - Test under heavy disk I/O
    - Test under heavy services load
    - Test under large number of simultaneous logon attempts

Building An Exchange Cluster

- Additional Best Practices
  - Do not install applications into the default Cluster Group
  - Do not delete or rename the default Cluster Group or remove any resources from that resource group
  - Do not use APM/ACPI power-saving features
  - Give the Cluster service account full rights to administer computer objects if Kerberos authentication is enabled for virtual servers
  - Do not set the Cluster service account to be a member of the domain administrator group

Building An Exchange 2000 Cluster

1. Prepare Active Directory, if 1st in org
   - Run ForestPrep
   - Run DomainPrep
2. Install Exchange 2000 on each node
   - Verify Cluster Service is running
   - Install MSDTC
   - Run Exchange setup
   - Grant Cluster Administrator Service full Administrator rights

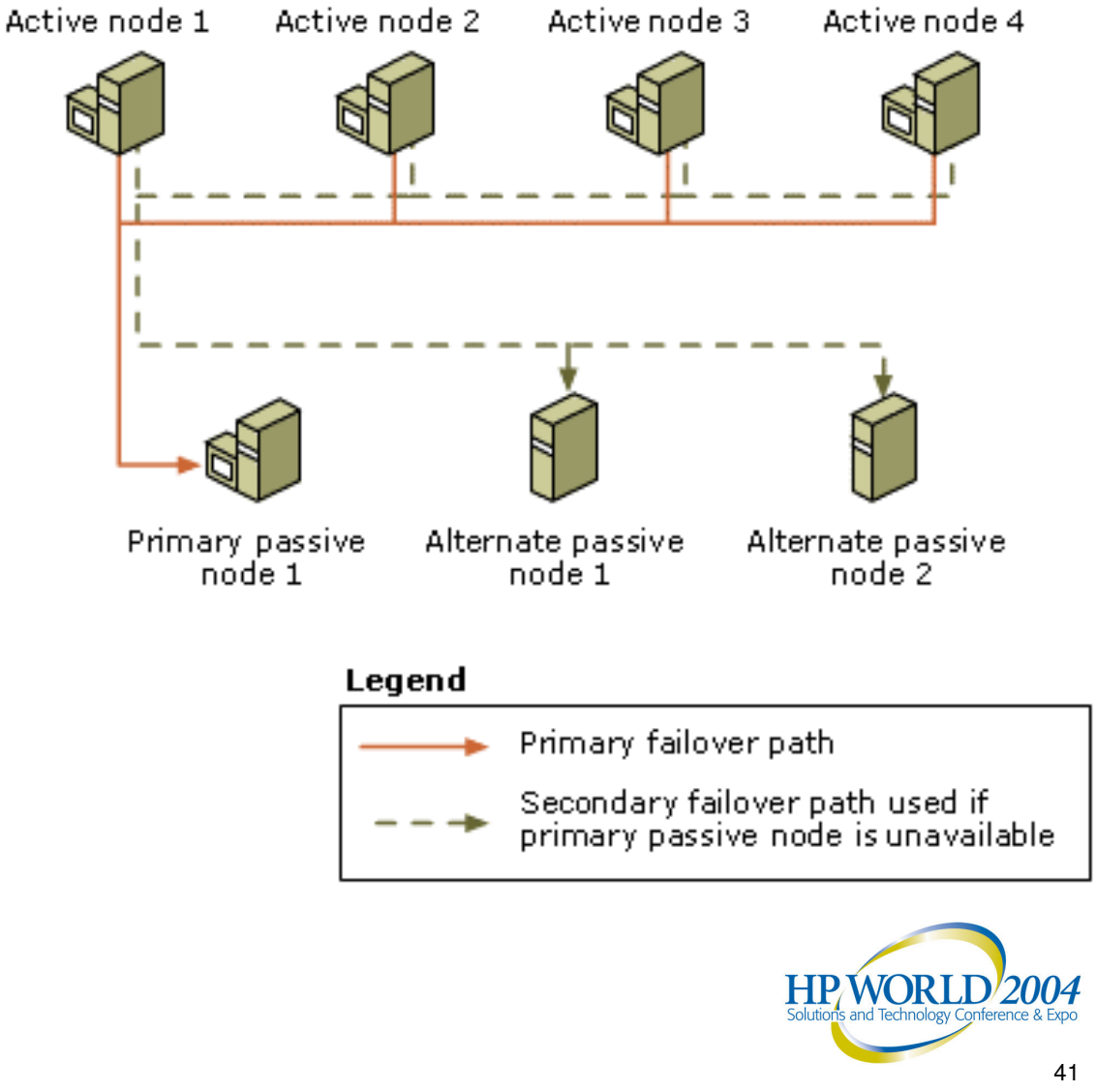
Building An Exchange 2000 Cluster

3. Create Exchange Virtual Servers

- Create EVS
- Create IP Address resource
- Create Network Name resource
- Add a disk resource to EVS
- Create System Attendant resource
  - SA will create
    - IS instance
    - MTA instance (1st EVS only)
    - Routing service instance
    - SMTP virtual server instance
    - HTTP virtual service instance
    - IMAP4 virtual server instance
    - POP3 virtual server instance
    - MS Search instance

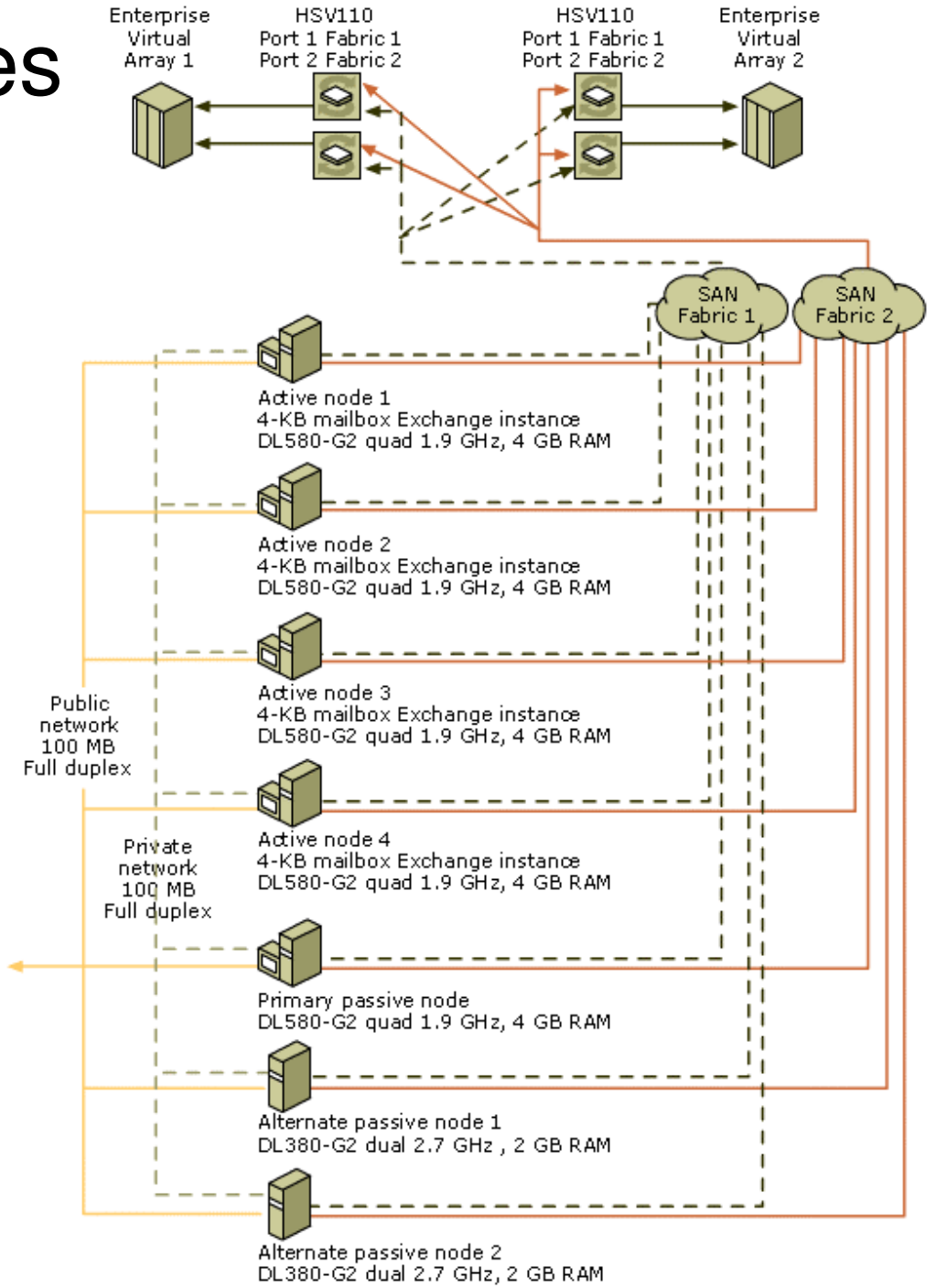
How Microsoft Uses Clustering

- 16,000 mailboxes/ cluster
- 200MB/mbx
- 4 EVS
- 4,000 mbx/EVS
- Redundant SAN
- Active nodes
  - 4 CPU
  - 1.9 GHz HT
  - 4GB RAM
- Alt Passive nodes
  - 2 CPU 2.4GHz



How Microsoft Uses Clustering

- 4 active nodes
- 1 passive node
- 2 alt passive nodes
- 2 SAN
- Duplicate network connections



Migrating From Exchange 5.5

- In-Place Upgrade Not Supported
- Move Mailbox Method - 316886
  - Introduce standalone Exchange 2003 server into Exchange 5.5 Site
  - Build and tune Exchange 2003 cluster and add EVS to Site
  - Move mailboxes to EVS
  - Decommission Exchange 5.5 cluster/stand-alone server

Upgrading From Exchange 2000

- Perform a Rolling Upgrade
  - If existing nodes are DCs, run dcpromo and demote them
  - In A/A model, move *both* EVS' to one node
  - Run Setup.exe on passive node (restart as necessary)
  - Move EVS to upgraded node
    - Exchange resources may have to be offline temporarily for this move to be successful, because the upgrade is not automatic, as it was in Exchange 2000.
  - On the upgraded node, right-click the System Attendant Resource in Cluster Administrator and select "Upgrade Exchange Virtual Server"
  - Run Setup.exe on other node (restart as necessary)
  - Redistribute EVS' (if applicable)

Important Setup Changes

- Block removal of node if EVS is running
- /disasterrecovery switched blocked
- Prevent Exchange 2003 from being first non-legacy server in Exchange 5.5 site
- By default, POP3 and IMAP4 resource not created. See 818480.

Security Changes

- Cluster service account no longer requires Exchange full admin rights. See 821834
- IPSec now supported between front-end servers and back-end clusters. See 821839
- Require Kerberos-enabled Network Name resource. See 235529

Managing An Exchange Cluster

- Cluster Administrator
- Other Objects Managed Within MMC
  - Exchange System Manager
  - Internet Services Manager
  - DNS
  - AD Users & Computers

Monitoring An Exchange Cluster

- Performance (MSExchangeIS)
  - VM Largest Block Size
  - VM Total 16MB Free Blocks
  - VM Total Free Blocks
  - VM Total Large Free Block Bytes
  - Make sure VM Total Large Free Block Bytes ALWAYS exceeds 32 MB
- Event Logs
- IIS Logs
- Cluster Logs

Removing An EVS

- Cannot remove first EVS until all other EVSs have been removed
- Designate a replacement system if bridgehead
- 5 Steps
  - Move all content (mailboxes/PFs) to another EVS
  - Take the SA resource offline
  - Use “Remove Exchange Virtual Server” option
    - Context option on both EVS group, and on System Attendant
  - Verify that EVS server object is gone from AD
    - Because of cluster service account permission changes, your interactive logon account must have the permissions to remove the Active Directory objects
  - Delete remaining cluster resources

Removing Exchange From A Cluster

- Create a temp folder outside Exchsrvr structure
- Copy four DLLs from Exchsrvr\BIN to temp folder: Dsaccess.dll, Exchmem.dll, Expoxy.dll, Pttrace.dll
- Uninstall via Control Panel
- Prompt: Is this the last one in the cluster?
- Reboot
- Copy DLLs from temp folder back to original location
- Run exchmgmt.exe \uninstall

Backup

- Use Exchange-aware backup software
- Location of backup
  - Backup device/software on each node
    - Use cluster-aware backup software
  - LAN-based backup solution
- See MSKB 286422 for details on how to backup and restore a Windows 2003 cluster

Disaster Recovery

- Recovery from Loss of Node
  - Evict the old node and verify eviction
  - Build a new cluster node and join cluster
  - Restore System State only on new node
  - Install Exchange on new node
  - Make new node Possible Owner of resources
- See http://go.microsoft.com/fwlink/?LinkId=1714 for Exchange Disaster Recovery info

Disaster Recovery

- Recovery from loss of shared storage
  - Make sure disk signatures are correct (MSKB 305793 and 280425)
- Recovery from Quorum Failure
  - Use DumpConfig to recreate disk signature
  - Perform a node restore
  - Use ClusRest to restore quorum
  - See MSKB 245762 and 248998 for more information

Disaster Recovery

- Recovery from Total Failure
  - Rebuild all nodes
  - Create original cluster group
  - Restore System State on each node
  - Install Exchange on each node
  - Recreate Exchange resources (use same names)
  - Restore databases from backup

Caveats

If you are installing the Exchange 2003 cluster into an existing Exchange 5.5 organization, the Exchange 2003 cluster may not be the first Exchange 2003 server in a site, and the Exchange 2003 cluster may not be a bridgehead server. Exchange 2003 requires the Site Replication Service (SRS) in a mixed environment. SRS is not supported in a clustered environment

Caveats

- 241626 – XADM: Do Not Use Secure Sockets Layer Only for Cluster Virtual Servers
- 263060 – XADM: Full Exchange Admin Cannot Create a New Store on the Second Virtual Server in a Cluster
- 271449 – XADM: Message Transfer Agent Stacks Service Incorrectly Logs an Error When It Runs in a Cluster Environment
- 241627 – XADM: Cannot Access Additional HTTP Virtual Server on Cluster Servers
- 271407 – XADM: Default SMTP Log File Directory Is Incorrect for Clustered Servers
- 251525 – XADM: Registry Checkpoints for DSProxy Need Manual Configuration
- 266689 – XADM: The "ESEUTIL /CC" Command Does Not Work on Cluster Server

Resources

- Exchange Server 2003 planning guide: http://www.microsoft.com/technet/prodtechnol/exchange/Exchange2003/proddocs/library/MessSyst.asp
- Exchange Server 2003 Deployment Guide: http://www.microsoft.com/technet/prodtechnol/exchange/Exchange2003/proddocs/library/DepGuide.asp
- “Managing Exchange Server Clusters" in the *Exchange Server 2003 Administration Guide* http://go.microsoft.com/fwlink/?LinkId=21769
- Exchange Server 2003 Performance and Scalability Guide http://go.microsoft.com/fwlink/?LinkId=28660
- Exchange Server 2003 Technical Documentation Library: http://www.microsoft.com/exchange/library/

HP WORLD 2004

Solutions and Technology Conference & Expo

Co-produced by:

interex

shared knowledge • shared power

encompass

AN HP USER GROUP