






26. DECUS München e.V. Symposium 2003
1A03



Microsoft Windows 2003 Server Clusters

Josef Beeking
Hewlett-Packard GmbH

Agenda 

- What's new in Clustering for Windows Server 2003

What's New in Clustering for Windows Server 2003 Server Clusters



- **NOTE:** Server clusters is a general term used to describe clusters based on the Microsoft® Cluster Service (MSCS), as opposed to clusters based on Network Load Balancing.

Server Clusters General



- **Larger Cluster Sizes**
- Microsoft Windows® Server 2003 Enterprise Edition supports 8-node clusters (was two), and Windows Server 2003 Datacenter Edition supports 8-node clusters (was four).
- *Benefits*
- *Greater Flexibility*

Server Clusters General



- **Terminal Server Application Mode**
- Terminal Server can run in application mode on nodes in a Server cluster. NOTE: There is no failover of Terminal Server sessions.
- *Benefits*
- *High Availability* - Terminal Server directory service can be made highly available through failover.


Server Clusters General



- **Majority Node Set (MNS) Clusters**
- optional quorum resource that does not require a disk on a shared bus for the quorum device. Geographically dispersed clusters. This mechanism provides a single, Microsoft-supplied quorum resource that is independent of any storage solution for a geographically dispersed or multi-site cluster. NOTE: There is a separate cluster Hardware Compatibility List (HCL) for geographic clusters.

- .

Server Clusters
Majority Node Set Quorums




Number of nodes in the cluster	Number of node failures before the loss of Quorum
1	0
2	0
3	1
4	1
5	2
6	2
7	3
8	3

- $(\lfloor \text{Number of nodes configured in the cluster} \rfloor / 2) + 1$

09.04.2003 page 7

Server Clusters
Majority Node Set Quorums




Cluster

Site 1

Node

Node

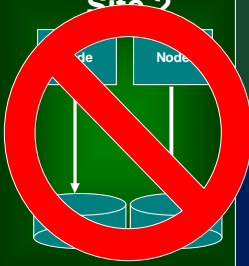
Node



Site 2

Node

Node



09.04.2003 page 8

Server Clusters General



- Low-cost or appliance-like highly available solutions that have no shared disks but use other techniques such as log shipping or software disk or file system replication and mirroring to make data available on multiple nodes in the cluster.
- NOTE: Windows Server 2003 provides no mechanism to mirror or replicate user data across the nodes of an MNS cluster, so while it is possible to build clusters with no shared disks at all, it is an application specific issue to make the application data highly available and redundant across machines.

Server Clusters Installation



- **Installed by Default**
- Clustering is installed by default.
- *Benefits*
- *Easier Administration* – you no longer need to provide a media CD to install Server clusters.
- *No reboot* – you no longer need to reboot

Server Clusters Installation



- **Pre-configuration Analysis**
- *Benefits*
- *Compatibility –*

Server Clusters Installation



- **Default Values**
- *Benefits*
- *Easier Administration –* Server cluster creation asks fewer setup questions, data is collected and the code makes decisions about the configuration. The goal is to get a “default” Server cluster up and running that can then be customized using the Server cluster administration tools if required.

Server Clusters Installation



- **Multi Node Addition**
- Allows multiple nodes to be added to a Server cluster in a single operation.
- *Benefits*
- *Easier Administration* – makes it quicker and easier to create multi-node Server clusters.

Server Clusters Installation



- **Extensible Architecture**
- *Benefits*
- *Third-Party Support* – allows applications to setup Server cluster resources and/or change their configuration as part of Server cluster installations rather than as a separate post-Server cluster installation task.

Server Clusters Installation



- **Remote Administration**
- *Benefits*
- *Easier Administration* – allows for better remote administration via Terminal Services.

Server Clusters Installation



- **Command Line Tools**
- *Benefits*
- *Easier Administration* – much easier to automate the process of creating a cluster.

Server Clusters Installation



- **Simpler Uninstallation**
- *Benefits*
- *Easier Administration* – Uninstalling the Cluster Service is much more efficient as you only need to evict the node through Cluster Administrator or Cluster.exe and the node is unconfigured for Cluster support. There is also a new switch for Cluster.exe which will force the uninstall if there is a problem with getting into Cluster Administrator:

```
cluster node %NODENAME% /force
```

Server Clusters Installation



- **Quorum Log Size**
- *The default size of the quorum log has been increased to 4096 KB (was 64 KB).*
- *Benefits*
- *Large number of shares – a quorum log of 4,096 KB allows for large numbers of file or printer shares (e.g. 200 printer shares). In previous versions, the quorum log would run out of space with this many shares, causing inconsistent failover of resources.*

Server Clusters Installation



- **Local Quorum**
- *Benefits*
- *Test Cluster* – This makes it very easy for users to create a test cluster on their local PC for testing out cluster applications, or for getting familiar with the Cluster Service. Users do not need special cluster hardware that has been certified on the Microsoft Cluster HCL to run a test cluster.
Note: Local quorum is only supported for one node clusters (i.e. lonewolf). In addition, the use of hardware that has not been certified on the HCL is not supported for production environments

Server Clusters Installation



- **Recovery** – in the event you lose all of your shared disks, one option for getting a temporary cluster working (e.g. while you wait for new hardware) is to use the cluster.exe /fixquorum switch to start the cluster, then create a local quorum resource and set this as your quorum. In the case of a print cluster, you can point the spool folder to the local disk. In the case of a file share, you can point the file share resource to the local disk, where backup data has been restored. Obviously, this does not provide any failover, and would only be seen as a temporary measure.

Server Clusters Installation



- **Quorum Selection**
- You no longer need to select which disk is going to be used as the Quorum Resource. It is automatically configured on the smallest disk that is larger than 50 MB and formatted NTFS.
- *Benefits*
- *Easier Administration* – the end user no longer has to worry about which disk to use for the quorum. NOTE: The option to move the Quorum Resource to another disk is available during setup or after the Cluster has been configured.

Server Clusters Installation



- **Integration**
- **Active Directory**
- Cluster Service now has much tighter integration with Active Directory™ (AD), including a "virtual" computer object, Kerberos authentication, and a default location for services to publish service control points (e.g. MSMQ).
- *Benefits*
- *Virtual Server* –

Server Clusters Installation



- To allow Kerberos authentication to services hosted in a virtual server, and
 - For cluster-aware and Active Directory-aware services (such as MSMQ) to publish service provider information specific to the virtual server they are hosted in.
- *Kerberos Authentication* –

Server Clusters Installation



- *Publish Services* –
- NOTE: Cluster integration does not make any changes to the AD schema.

Server Clusters Installation

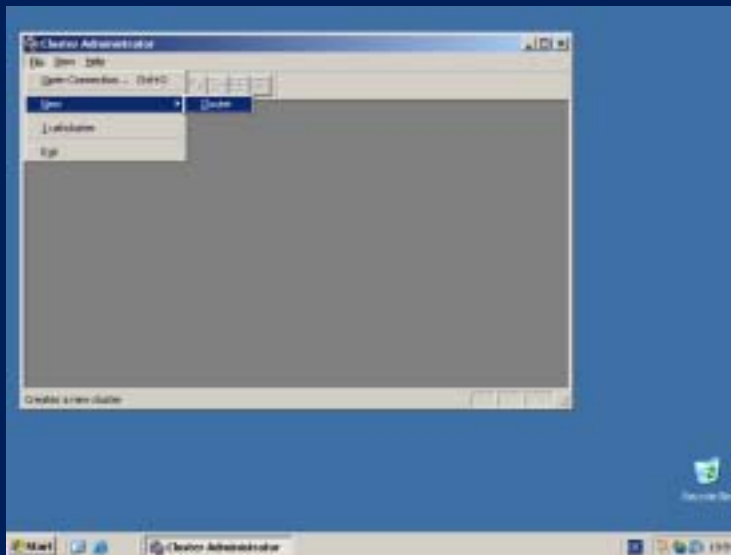


- **Extend Cluster Shared Disk Partitions**
- If the underlying storage hardware supports dynamic expansions of a disk unit, or LUN, then the disk volume can be extended online using the DISKPART.EXE utility.
- *Benefits*
- *Easier Administration* – Existing volumes can be expanded online without taking down applications or services.

Server Clusters Installation



Server Clusters Installation



09.04.2003

page 27

Server Clusters Installation



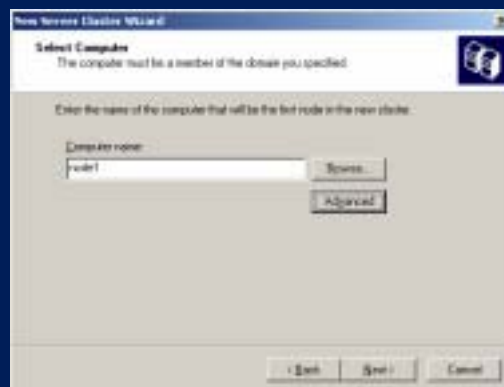
09.04.2003

page 28

Server Clusters Installation



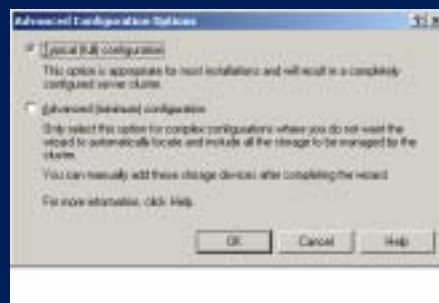
Server Clusters Installation



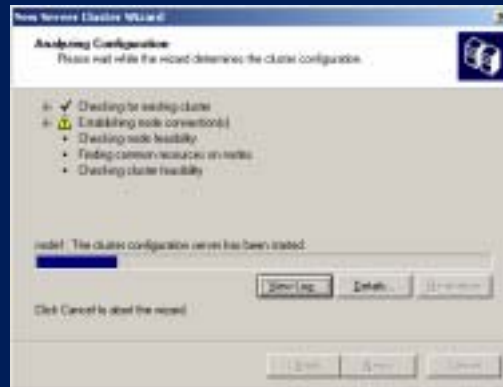
Server Clusters Installation



Server Clusters Installation



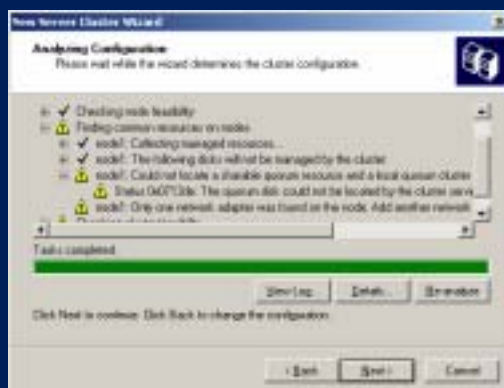
Server Clusters Installation



Server Clusters Installation



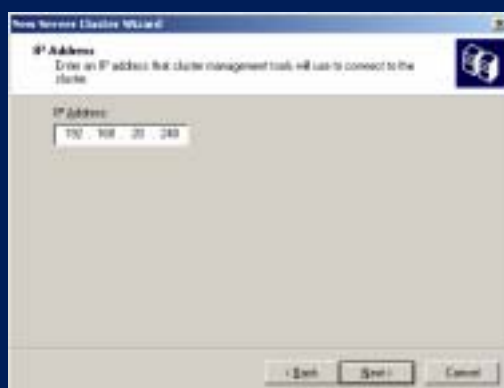
Server Clusters Installation



09.04.2003

page 35

Server Clusters Installation



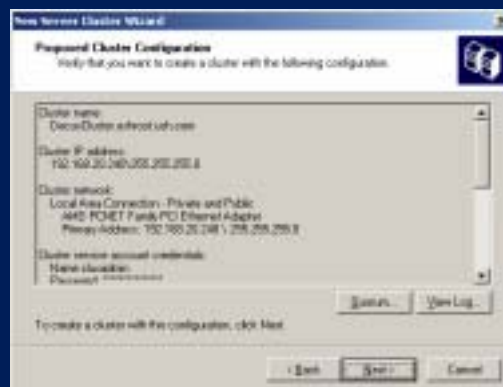
09.04.2003

page 36

Server Clusters Installation



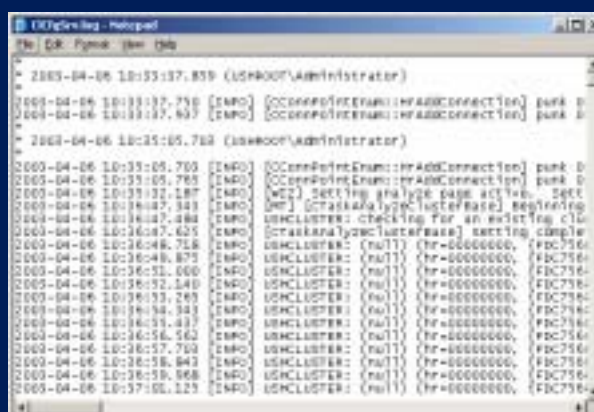
Server Clusters Installation



Server Clusters Installation



Server Clusters Installation



Server Clusters Installation



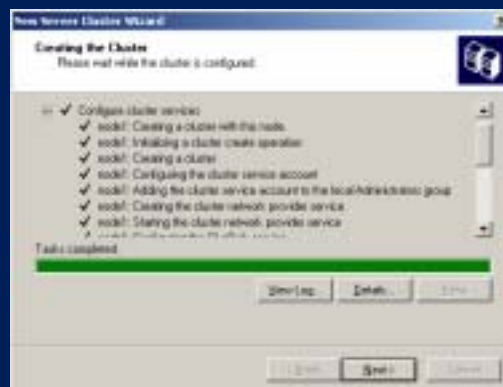
Server Clusters Installation



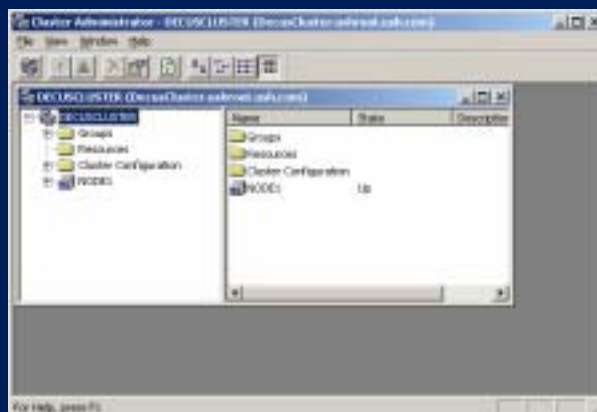
Server Clusters Installation



Server Clusters Installation



Server Clusters Installation



09.04.2003

page 45

Server Clusters Resources



- **Printer Configuration**
- Cluster Service now provides a much simpler configuration process for setting up clustered printers.
- *Benefits*
- *Easier Administration* – To set up a clustered print server, you need to configure only the Spooler resource in Cluster Administrator and then connect to the virtual server to configure the ports and print queues. This is an improvement over previous versions of Cluster Service in which you had to repeat the configuration steps on each node in the cluster, including installing printer drivers.

Server Clusters Resources



- **MSDTC Configuration**
- The Microsoft Distributed Transaction Coordinator (MSDTC) can now be configured once, and then be replicated to all nodes.
- *Benefits*
- *Easier Administration* – in previous versions, the COMCLUST.EXE utility had to be run on each node in order to cluster the MSDTC. It is now possible to configure MSDTC as a resource type, assign it to a resource group, and have it automatically configured on all cluster nodes

Server Clusters Resources



- **Scripting**
- Existing applications can be made Server cluster-aware using scripting (VBScript and Jscript) rather than writing resource dlls in C or C++.
- *Benefits*
- *Easier Development* – makes it much simpler to write specific resource plug-ins for applications so they can be monitored and controlled in a Server cluster. Supports resource specific properties, allowing a resource script to store Server cluster-wide configurations that can be used and managed in the same way as any other resource.

Server Clusters Resources



- **MSMQ Triggers**
- Cluster Service has enhanced the MSMQ resource type to allow multiple instances on the same cluster.
- *Benefits*
- *Enhanced Functionality* – allows you to have multiple clustered message queues running at the same time, providing increased performance (in the case of Active/Active MSMQ clusters) and flexibility.
- NOTE: You can only have one MSMQ resource per Cluster Group

Server Clusters Network Enhancements



- **Enhanced Network Failover**
- Cluster Service now supports enhanced logic for failover when there has been a complete loss of internal (heartbeat) communication. The network state for public communication of all nodes is now taken into account.
- *Benefits*
- *Better Failover* – in Windows 2000, if Node A owned the quorum disk and lost all network interfaces (i.e. public and heartbeat), it would retain control of the cluster, despite the fact that no one could communicate with it, and that another node may have had a working public interface. Windows Server 2003 cluster nodes now take the state of their public interfaces into account prior to arbitrating for control of the cluster.

Server Clusters Network Enhancements



- **Media Sense Detection**
- When using Cluster Service, if network connectivity is lost, the TCP/IP stack does not get unloaded by default, as it did in Windows 2000. There is no longer the need to set the DisableDHCPMediaSense registry key.
- *Benefits*
- *Better Failover* – in Windows 2000, if network connectivity is lost, the TCP/IP stack was unloaded, which meant that all resources that depended on IP addresses were taken offline. Also, when the networks came back online, their network role reverted to the default setting (i.e. client and private). By disabling Media Sense by default, it means the network role is preserved, as well as keeping all IP address dependant resources online.

Server Clusters Network Enhancements



- **Multicast Heartbeat**
- Allows multi-cast heartbeats between nodes in a Server cluster. Multi-cast heartbeat is automatically selected if the cluster is large enough and the network infrastructure can support multi-cast between the cluster nodes. Although the multi-cast parameters can be controlled manually, a typical configuration requires no administration tasks or tuning to enable this feature. If multicast communication fails for any reason, the internal communications will revert to unicast. All internal communications are signed and secure.
- *Benefits*
- *Reduced Network Traffic* – by using multicast, it reduces the amount of traffic in a cluster subnet, which can be particularly beneficial in clusters of more than two nodes, or geographically dispersed clusters.

Server Clusters Storage



- **Storage**
- **Volume Mount Points**
- Volume mount points are now supported on shared disks (excluding the quorum), and will work properly on failover if configured correctly.
- *Benefits*
- *Flexible Filesystem Namespace* - volume mount points (Windows 2000 or later) are directories that point to specified disk volumes in a persistent manner (e.g. you can configure C:\Data to point to a disk volume). They bypass the need to associate each disk volume with a drive letter, thereby surpassing the 26 drive letter limitation (e.g. without volume mount points, you would have to create a G: drive to map the "Data" volume to). Now that Cluster Service supports volume mount points, you have much greater flexibility in how you map your shared disk namespace.
- **NOTE:** The directory that hosts the volume mount point must be NTFS since the underlying mechanism uses NTFS reparse points. However the file system that is being mounted can be FAT, FAT32, NTFS, CDFS, or UDFS.

Server Clusters Storage



- **Client Side Caching (CSC)**
- Client Side Caching (CSC) is now supported for clustered file shares.
- *Benefits*
- *Offline File Access* –Client Side Caching for clustered file shares allows a client to cache data stored on a clustered share. The client works on a local copy of the data that is uploaded back to the Server cluster when the file is closed. This allows the failure of a server in the Server cluster and the subsequent failover of the file share service to be hidden from the client

Server Clusters Storage



– Distributed File System

- Distributed File System (DFS) has had a number of improvements, including multiple stand-alone roots, independent root failover, and support for Active/Active configurations.
- *Benefits*
- Distributed File System (DFS) allows multiple file shares on different machines to be aggregated into a common namespace (e.g. `\dfsroot\share1` and `\dfsroot\share2` are actually aggregated from `\server1\share1` and `\server2\share2`). New clustering benefits include:
 - *Multiple Stand-Alone Roots* – previous versions only supported one clustered stand-alone root. You can now have multiple ones, giving you much greater flexibility in planning your distributed file system namespace (e.g. multiple DFS roots on the same virtual server, or multiple DFS roots on different virtual servers).
 - *Independent Failover* – granular failover control is available for each DFS root, allowing you to configure failover settings on an individual basis and resulting in faster failover times.
 - *Active/Active Configurations* – you can now have multiple stand-alone roots running actively on multiple nodes.

Server Clusters Storage



– Encrypted File System

- With Windows Server 2003, the encrypting file system (EFS) is supported on clustered file shares. This allows data to be stored in encrypted format on clustered disks.

Server Clusters Storage



- **Storage Area Networks (SAN)**
- Clustering has been optimized for SANs, including targeted device resets and the shared storage buses.
- *Benefits*
- *Targeted Bus Resets* - the Server cluster software now issues a special control code when releasing disk drives during arbitration. This can be used in conjunction with HBA drivers that support the extended Windows Server 2003 feature set to selectively reset devices on the SAN rather than full bus reset. This ensures that the Server cluster has much lower impact on the SAN fabric.
- *Shared Storage Bus* – shared disks can be located on the same storage bus as the Boot, Pagefile and dump file disks. This allows a clustered server to have a single storage bus (or a single redundant storage bus). NOTE: This feature is disabled by default due to the configuration restrictions. This feature can/should only be enabled by OEMs and IHVs for specific and qualified solutions. This is NOT a general purpose feature exposed to end users.

Server Clusters Operations



- **Backup and Restore**
- You can actively restore the local cluster nodes cluster configuration or you can restore the cluster information to all nodes in the Cluster. A node restoration is also built into Automatic System Recovery (ASR).
- *Benefits*
- *Backup and Restore* – Backup (NTBackup.exe) in Windows Server 2003 has been enhanced to enable seamless backups and restores of the local Cluster database, and to be able to restore the configuration locally and to all nodes in a Cluster.
- *Automated System Recovery* – ASR can completely restore a cluster in a variety of scenarios, including: a) damaged or missing system files, b) complete OS reinstallation due to hardware failure, c) a damaged Cluster database, and d) changed disk signatures (including shared

Server Clusters Operations



- **Enhanced Node Failover**
- Cluster Service now includes enhanced logic for node failover when you have a cluster with three or more nodes. This includes doing a manual “Move Group” operation in Cluster Administrator.
- *Benefits*
- *Better Failover* – during failover in a cluster with three or more nodes, the Cluster Service will take into account the “Preferred Owner List” for each resource, as well as the installation order for each node, in order to work out which node the group should be moved to.

Server Clusters Operations



- **Group Affinity Support**
- Allows an application to describe itself as an N+1 application. In other words, the application is running actively on N nodes of the Server cluster and there are 1 “spare” nodes available if an active node fails. In the event of failure, the failover manager will try to ensure that the application is failed over to a spare node rather than a node that is currently running the application.
- *Benefits*
- *Better performance* – applications are failed over to spare nodes before active nodes.

Server Clusters Operations



- **Node Eviction**
- Evicting a node from a Server cluster no longer requires a reboot to clean up the Server cluster state. A node can be moved from one Server cluster to another without having to reboot. In the event of a catastrophic failure, the Server cluster configuration can be force cleaned regardless of the Server cluster state.
- *Benefits*
- *Increased Availability* – not having to reboot increases the uptime of the system.
- *Disaster Recovery* – in the event of a node failure, the cluster can be cleaned up easily.

Server Clusters Operations



- **Rolling Upgrades**
- Rolling upgrades are supported from Windows 2000 to Windows Server 2003.
- *Benefits*
- *Minimum downtime* – rolling upgrades allow one node in a cluster to be taken offline for upgrading, while other nodes in the cluster continue to function on an older version. **NOTE:** There is no support for rolling upgrades from a Microsoft Windows NT 4.0 cluster to a Windows Server 2003 cluster. An upgrade from Windows NT 4.0 is supported but the cluster will have to be taken offline during the upgrade.

Server Clusters Operations



- **Queued Changes**
- The cluster service will now queue up changes that need to be completed if a node is offline.
- *Benefits*
- *Easier Administration* – ensures that you do not need to apply a change twice if a node is offline. For example, if a node is offline and is evicted from the Cluster by a remaining node, the cluster service will be uninstalled the next time the first node attempts to join the Cluster. This also holds true for applications

Server Clusters Operations



- **Disk Changes**
- The Cluster Service more efficiently adjusts to shared disk changes in regards to size changes and drive letter assignments.
- *Benefits*
- *Dynamic Disk Size* – If you increase the size of a shared disk, the Cluster Service will now dynamically adjust to it. This is particularly helpful for SANs, where volume sizes can change easily. It does this by working directly with Volume Mount Manager, and no longer directly uses the DISKINFO or DISK keys.
NOTE: These keys are maintained for backwards compatibility with previous versions of the Cluster Service.

Server Clusters Operations



- **Password Change**
- Cluster Service account password changes no longer require any downtime of the cluster nodes. In addition, passwords can be reset on multiple clusters at the same time.
- *Benefits*
- *Reduced Downtime* – In Windows Server 2003, you can change the Cluster Service account password on the domain as well as on each local node, without having to take the cluster offline. If multiple clusters use the same Cluster service account, you can change them simultaneously. In Microsoft Windows NT 4.0 and Microsoft Windows 2000, to change the Cluster service account password, you have to stop the Cluster service on all nodes before you can make the password change.

Server Clusters Operations



- **Resource Deletion**
- Resources can be deleted in Cluster Administrator or with Cluster.exe without taking them offline first.
- *Benefits*
- *Easier Administration* – in previous versions, you first had to take a resource offline before you could delete it. Now, Cluster Service will take them offline automatically, and then delete them.

Server Clusters Operations



- **WMI Support**
- Server clusters provides WMI support for:
- Cluster control and management functions including starting and stopping resources, creating new resource and dependencies etc.
- Application and cluster state information. WMI can be used to query whether applications are online, whether cluster nodes are up and running as well as a host of other status information.
- Cluster state change events are propagated via WMI to allow applications to subscribe to WMI events that show when an application has failed, when an application is restarted, when a node fails etc.
- *Benefits*
- *Better Management* – allows Server clusters to be managed as part of an overall WMI environment.

Server Clusters Supporting and Troubleshooting



- **Offline/Failure Reason Codes**
- These provide additional information to the resource as to why the application was taken offline, or failed.
- *Benefits*
- *Better Troubleshooting* – allows the application to have different semantics if the application has failed or some dependency of the application has failed, versus the administrator specifically moved the group to another node in the Server cluster.

Server Clusters Supporting and Troubleshooting



- **Software Tracing**
- Cluster Service now has a feature called software tracing that will produce more information to help with troubleshooting Cluster issues.
- *Benefits*
- *Better Troubleshooting* – this is a new method for debugging that will allow Microsoft to debug the Cluster Service without loading checked build versions of the dll's (symbols).

Server Clusters Supporting and Troubleshooting



- **Cluster Logs**
- A number of improvements have been made to the Cluster Service log files, including a setup log, error levels (info, warn, err), local server time entry, and GUID to resource name mapping.
- *Benefits*
- *Setup Log* – during configuration of Cluster Service, a separate setup log (%SystemRoot%\system32\Logfiles\Cluster\CICfgSrv.log) is created to assist in troubleshooting.
- *Error Levels* – this makes it easy to be able to highlight just the entries that require action (e.g. err).
- *Local Server Time Stamp* – this will assist in comparing event log entries to Cluster logs.
- *GUID to Resource Name Mapping* – this assists in understanding the cluster log references to GUIDs. A Cluster object file (%windir%\Cluster\Cluster.obj) is automatically created and maintained that contains a mapping of GUID's to Resource Name mappings.

Server Clusters Supporting and Troubleshooting



- **Event Log**
- Additional events are written to the event log indicating not only error cases, but showing when resources are successfully failed over from one node to another. *Benefits*
- *Better Monitoring* – this allows event log parsing and management tools to be used to track successful failovers rather than just catastrophic failures.

Server Clusters Supporting and Troubleshooting



- **Clusdiag**
- A new tool called clusdiag is available in the Windows Server 2003 Resource Kit.
- *Benefits*
- *Better Troubleshooting* – makes reading and correlating cluster logs across multiple cluster nodes and debugging of cluster issues more straight forward.
- *Validation and Testing* – Clusdiag allows users to run stress tests on the server, storage and clustering infrastructure. As such, it can be used as a validation and test tool before a cluster can be put into production

Server Clusters Supporting and Troubleshooting



- **Chkdsk Log**
- The cluster service creates a chkdsk log whenever chkdsk is run on a shared disk.
- *Benefits*
- *Better Monitoring* – this allows a system administrator to find out and react to any issues that were discovered during the chkdsk process.

Server Clusters Supporting and Troubleshooting



- **Disk Corruption**
- When Disk Corruption is suspect, the Cluster Service reports the results of CHKDSK in event logs and creates a log in %systemroot%\cluster.
- *Benefits*
- *Better Troubleshooting* – results are logged in the Application and Cluster.log. In addition, the Cluster.log references a log file (e.g. %windir%\CLUSTER\CHKDSK_DISK2_SIGE9443789.LOG) in which detailed CHKDSK output is recorded.

Microsoft Windows 2003 Cluster Server
Cluster Resources



- **File and print shares**
- **Generic services or applications**
- **Generic scripts**
- **Physical disks**
- **Microsoft Distributed Transaction Coordinator (MSDTC)**
- **Internet Information Services (IIS)**
- **Message Queuing (MSMQ) Triggers**
- **Network addresses and names**

Microsoft Windows 2003 Cluster Server
Cluster Resources



- Windows Server 2003, Enterprise Edition, and Windows Server 2003, Datacenter Edition, include resource DLLs for the following additional services:
 - **Distributed File System (DFS)**
 - **Dynamic Host Configuration Protocol (DHCP) service**
 - **Windows Internet Service (WINS)**

For More Information



- Books
- *Windows NT Microsoft Cluster Server*, by Richard R. Lee, Osborne McGraw-Hill, 1999.
- *Windows 2000 Cluster Server Guidebook*, by David Libertone, Prentice Hall, 2000.
- *Windows NT Backup & Recovery*, by John McMains and Bob Chronister, Osborne McGraw-Hill, 1998.
- *Windows NT Clustering Blueprints*, by Mark A. Sportack, SAMS Publishing, 1997.
- *In Search of Clusters, Second Edition: The Coming Battle in Lowly Parallel Computing*, Gregory F. Pfister, Prentice Hall, 1998, ISBN: 0138997098.
- *The Book of SCSI*, Peter M. Ridge, No Starch Press, 1995, ISBN: 1886411026.
- *Transaction Processing Concepts and Techniques*, Gray, J., Reuter A., Morgan Kaufmann, 1994. ISBN 1558601902, survey of outages, transaction techniques.

For More Information



- Related Links
- For information about the Windows 2000 Server family of products, see:
- <http://www.microsoft.com/windows2000/server/default.asp>
- For information about Windows 2000 Server Reliability & Availability, see: <http://www.microsoft.com/windows2000/server/evaluation/business/overview/reliable/default.asp>
- For information about Windows 2000 Clustering Technologies, see: <http://www.microsoft.com/windows2000/technologies/clustering/default.asp>

For More Information



- Related Links
- For information about what's new in Windows server 2003 clusters, see:
<http://www.microsoft.com/windows/netserver/evaluation/overview/technologies/clustering.asp>
- For documentation about the architecture of Windows base services, including Windows Clustering technologies, see *Windows Base Services* under *Windows Development* in the Platform Software Developer Kit (SDK):
<http://www.msdn.microsoft.com/library>





09.04.2003

page 81