

# Windows clustering glossary

To configure the Microsoft Cluster Service with Windows 2000 Advanced Server, you need to have a solid grounding in the various terms that are used with the Cluster Service. Because the Cluster Service is a bit complicated, and the setup requires that you grasp a number of important concepts, I have put together this glossary of terms for working with the Cluster Service.

This glossary will cover some of the basic terms you need to know, with a few useful tips thrown in. Please note that this list is not intended to be exhaustive, but to acquaint you with the key terms you need to understand for configuration purposes. The list is also not in alphabetical order, as terms are loosely grouped by subject.

## **Cluster**

One or more servers working together, which appear as a single (virtual) server to the network. Multiple servers in the cluster provide redundant operations in case of a failure on one of the servers.

## **Node**

The name for each server within a cluster to distinguish between a virtual server and the actual physical server.

## **Group**

A container for logically linked resources that are required for an application to run. All resources that list dependencies must reside in the same group and must be online on the same node at the same time to function. Typically, each clustered application or service installed on the cluster is placed in its own group.

## **Resources**

The lowest units managed by Cluster Service. Resources are logical or physical entities managed by the cluster and grouped logically into a cluster group. They can have dependencies—a reliance on another resource to successfully come online before they can go online themselves. (Interrelated dependencies result in a dependency tree.) Some dependencies are required (built-in) when using the Cluster Administrator MMC, but you can manually add logical dependencies where it makes sense. Resources can also have parameters specific to the resource. For example, the actual IP address is the parameter specific to the IP Address resource, whereas the DHCP server resource has parameters for the DHCP database location, backup database location, and audit log.

## **Active/active vs. active/passive**

Terms that describe the server roles with regard to clustering services. With an active/active configuration, each node runs applications but can provide failover capability for the other server. With an active/passive configuration, one node runs the applications and the other server remains unused until failover occurs. Active/active has the benefit of lower cost (both servers are being used) but at the risk of lower performance should one server fail. Active/passive has the benefit of offering better performance after failover, but at a higher cost, since one server remains dormant at all times.

## **Heartbeat**

A single UDP packet sent between nodes in the cluster (usually via the private network) to confirm that they are still online. The first node online in the cluster is responsible for initiating heartbeats to the other nodes.

## **LooksAlive check**

A simple check to verify that a resource is running properly. If it fails or cannot determine the status, the more thorough IsAlive check is used.

**IsAlive check**

An exhaustive check to verify that a resource is running properly. If this check fails, the resource is moved offline and the failover process is triggered.

**Failover**

A transfer of a group (containing resources) from one node to another node when failure is detected.

**Failback**

A transfer of a group back to the original node if this node becomes available within the allowed period. By default, groups do not fail back to their original node but wait to be manually moved back. If configured, failback will happen only with a preferred owner of the group.

**Failover threshold**

The number of times a group will be allowed to fail before the Cluster Service decides that the group can't be brought online anywhere in the cluster and will keep the group offline.

**Quorum (resource)**

A common resource on a shared disk that contains a synchronized version of the current cluster configuration. A quorum must be present for the cluster to function. It requires a minimum of 50 MB (with 500 MB recommended) and typically is assigned to a Q drive.

**Quorum log**

A log that's produced automatically by the quorum resource and is used for transactional events to improve efficiency. The default size is 64 KB, which is often too small for a production cluster and could result in a failure. Since the quorum should have its own disk, and therefore space should be plentiful, consider changing this to a higher value (e.g., 500 MB) to ensure that lengthy transactions, such as dynamic shares, can complete.

**Quorum switches for Cluster Service**

If there are problems loading the quorum, the Cluster Services will fail to load. You can supply several switches as a startup parameter to the Cluster Service (with just one node up) to help remedy this situation:

*ResetQuorumlog*

Use this if the Quorum log is corrupt, which resets the quorum using the local cluster version (some recent changes may be lost).

*FixQuorum*

Use this if the quorum device is not functional (hardware problems). This loads the Cluster Service with all resources offline, including the quorum device, so you can configure a new location for the quorum with the cluster's properties. Use with Cluster Administrator MMC to connect to the local cluster when using this switch.

*NoQuorumLogging*

Use this if there's a problem with the file system because the quorum device is mounted but the quorum log is not used. Either fix the problem (reformat or chkdsk) or delete the corrupted files and restore from backup into the \MSCS folder. Alternatively, change the location of the quorum (and then restore from backup or use the ResetQuorumLog switch).

**Virtual server**

The mechanism by which applications and services in a cluster can be exposed to clients using a logical IP address and a logical server name.

**Shared storage/device**

A disk or disk array that is accessible by more than one node. The Windows 2000 Cluster Service uses the "shared nothing" clustering model, which means that only one node will be able to access a shared device at once. Each node manages its own hardware and requires a Distributed Lock Manager (DLM),

which the Cluster Service does not supply (but could be used with 3<sup>rd</sup> party applications). When first installing the Cluster Service, it's important that both computers do not attempt to access the shared device at the same time. Verify that both servers can access the device and configure the disks with the same drive letter, but make sure that the other server is not booted into the operating system at this point. Install the Cluster Service on one server as a new cluster with the other server powered down, selecting the shared device(s) to be clustered. Then, install Cluster Service on the second server and choose to join the cluster just created. Note that the shared storage must be formatted with NTFS and does not support the following:

- EFS
- Dynamic disks (and therefore dynamic volumes)
- Mounted volumes
- Remote storage
- Reparse points
- Software RAID

### **SCSI vs. Fibre Channel Disk technology**

The two choices of disk technologies supported by Cluster Service. (You can't use IDE drives because they do not support sharing among multiple servers.) Fibre Channel is the better technology. It offers faster transmission over longer distances and supports up to 256 disks that can hold up to 9 terabytes with hot swapping. However, it is more expensive than SCSI and has higher administrative overhead to initially implement.

### **Domainlet**

A limited Windows 2000 domain that can be used so that the servers can be configured as domain controllers (to ensure the Cluster Service authenticates successfully), but without the additional overheads of maintaining standard user accounts or running an Active Directory Global Catalog. To configure a domainlet, remove the Global Catalog setting under the properties of the NTDS Settings in the Active Directory Sites And Services and add a registry key (with no associated value) "IgnoreGCFailures" under HKLM\SYSTEM\CurrentControlSet\Control\Lsa.

### **Cluster-aware applications/services**

Programs that have been written with the Cluster Service in mind (using the Cluster API) to ensure successful installation/configuration and failover with minimal loss of data (e.g., save data before failing over). These applications provide their own resources for configuration.

### **Cluster-unaware applications/services**

Programs that do not use the Cluster API to interact with Cluster Service and, therefore, may not successfully fail over without loss of data. Many cluster-unaware applications/services do work successfully with Cluster Service by using the Generic Application resource DLL or the Generic Service resource DLL.

### **Interconnect**

The private network connection between the cluster nodes, used for the following:

- Heartbeats
- Event log replication
- State changes
- Cluster configuration synchronization
- Application (of generic applications/services) and metadata replication

### **Y cables**

Typically used with a SCSI shared disk array as the shared bus between two servers. Each server is attached to one end of the cable with another end attached to the external disk array; the spare end of the cable is terminated (with terminators). An alternative to Y cables is self-terminating SCSI adapters, but

make sure that they have their own power supply so they continue to offer termination when the computer is powered down.

## SCSI IDs

When using SCSI disks, each SCSI adapter on each server must have a different SCSI ID number. The default shipping value is 7, so standard advice is to change one of the IDs to 6. However, a wise administrator will change both (e.g., one to 5 and another to 6) so that if one fails it can be replaced without worrying about what the original ID was.

## Preferred node

The node that is best suited to running an application. This will be used if failback is configured. It's used for static load balancing and when using asymmetric hardware.

## Cluscfg.exe

The Cluster Configuration wizard, which can run with a number of command-line options so you don't have to interactively supply the configuration information. It can also specify an unattend option to supply an answer file, which suppresses the user interface.

## Cluster Service rights

The Cluster Service account is automatically granted these rights on installation:

- Lock pages in memory
- Log on as a service
- Act as part of the operating system
- Back up files and directories
- Increase quotas
- Increase scheduling priority
- Load and unload device drivers
- Restore files and directories

If you don't manually change this account's user rights, you should not have to worry about these unless you promote or demote your server with `dcpromo`. If you promote a member server to a domain controller, you will need to manually reassign *log on as a service* and *lock pages in memory*. If you demote a domain controller to a member server, you will need to manually reassign the rights *log on as a service*, *lock pages in memory*, and *act as part of the operating system*. Additionally, a demoted server will not include the Cluster Service account in the local Administrators group, so you'll need to manually add this.

## Ownership

Because Cluster Service uses a "shared nothing" model, who owns what is very important. Groups are owned by a node, and resources are owned by groups. You can move resources between groups to satisfy dependencies and application requirements. You can also transfer group ownership by moving a group to another node, but this is usually done only when you need to bring down a node for maintenance (because it involves temporarily taking offline all the resources in the group before they come online again on the other node).

## Possible owner

If a node is not listed as a possible owner of a group, the group cannot fail over to it and will remain failed on the original node.

## Cluster.exe

The command-line utility to manage the Cluster Service as an alternative to the GUI Cluster Administrator MMC (`Cluadmin.exe`). `Cluster.exe` is useful for scripting and is included in the Adminpak.

**Eviction**

Permanently removes a node from a cluster but doesn't uninstall Cluster Service from that server. To rejoin the cluster, you must first uninstall Cluster Server from the evicted node and then reinstall, reconfigure, and rejoin.

**Node states**

Node states can be Up (healthy), Down (cannot communicate with the cluster, so any groups owned by that node will be failed over), Joining (in the processing of joining the cluster; should soon change to Up), Paused (this is often used for maintenance and survives reboots—groups remain active but new resources will not be brought online until status changes to Up), and Unknown (status cannot be determined).

**Cluster database—clusdb**

A hive under HKLM\Cluster that is physically stored under the folder %Systemroot%\Cluster\Clusdb. When a node joins a cluster, it obtains the cluster configuration from the quorum and downloads it to this local cluster database. If this database is corrupt, inaccessible, or unavailable, the Cluster Service may fail to start, so make sure it is backed up. Ensure that the Cluster Service account has full access to this folder.

**Cluster Log**

A complete record of the Cluster Service events that occurred on a specific node (e.g., forming or joining a cluster, creating groups or resources). This is useful for diagnostic purposes and should be used in conjunction with the Windows System event log.

**Clusrest**

A Windows 2000 Resource Kit utility used to restore quorum data to the quorum disk after restoring a node using Windows 2000 Backup.

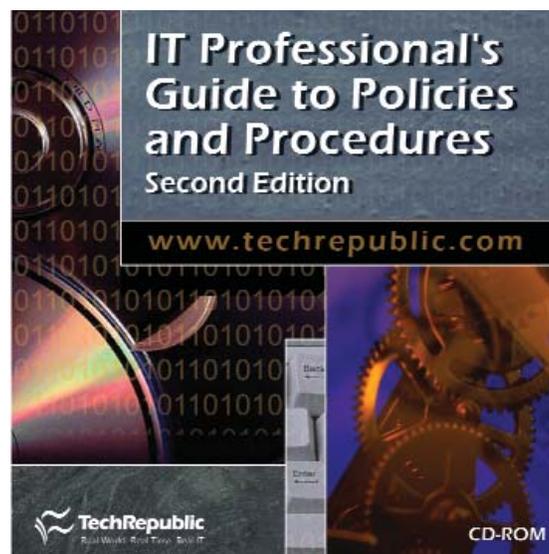
# IT Professional's Guide to Policies and Procedures, Second Edition

Meet the policy and procedure needs of your entire enterprise—and do it with ease with TechRepublic's expanded edition: *IT Professional's Guide to Policies and Procedures, Second Edition*. You'll get over 40 customizable policies and templates to help you create customized systems, procedures, and policies that are relevant, accessible, and usable...all on one CD-ROM!

This comprehensive CD has done almost all the work for you! A complete set of policies and procedures is at your fingertips. You simply select which work for you based on your organization's needs. Plus, you'll get articles and expert advice on how to execute and enforce policies...and learn the repercussions of not having them in place.

You'll soon be deploying comprehensive, customized policies to handle the following:

- E-mail and Internet usage
- Virus prevention
- Information security
- Inventory and equipment use
- Software installation and licensing



**Order Today!**

**Yes!** Please send my copy of *IT Professional's Guide to Policies and Procedures, Second Edition*. I'll receive my CD-ROM at the special Member rate of only \$149 plus \$5.95 s&h (\$8.95 s&h Canada, \$12.95 s&h international). If I'm not completely satisfied, I can return my copy within 30 days for a full refund.