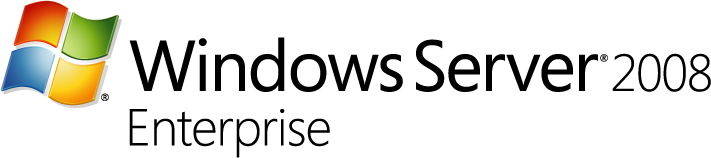
ms-logo_cgray6



Overview of Failover Clustering with Windows Server 2008

White Paper

Published: November 2007

For the latest information, please see   
[www.microsoft.com/windowsserver2008/failover-clusters.mspx](http://www.microsoft.com/windowsserver2008/failover-clusters.mspx)

Contents

Introduction 1

Installation, Management, and Administration 2

Architecture 4

Storage Improvements 6

Closer Integration with VSS 6

Security Improvements 7

Conclusion 8

Related Links 9

# Introduction

Organizations put a lot of value on mission-critical servers and rely on them heavily to run their businesses. As a result, server downtime can be very costly. A heavily used e-mail or database server can easily cost a business thousands or tens of thousands of dollars in lost productivity or lost business for every hour that it is unavailable. For every benefit and advantage brought to an organization by an IT solution, technology and business decision-makers should also think about how to deal with the inevitable downtime of these solutions.

Server workload availability is a trade-off between performance and cost. The ‘round-the-clock pace of global commerce makes uninterrupted IT operations vital to an increasing number of industries, from financial services and logistics to manufacturing and travel and tourism. Even seemingly low-tech business such as trucking suffer when crucial server workloads are unavailable, yet achieving the degree of reliability and availability demanded by mission-critical business requirements is expensive to create and support—both in terms of hardware and software costs as well as in the worker time required to manage the solution. The challenge for organizations is to learn what level of IT service availability is justified by their own price of downtime.

The term “high availability” refers to the characteristics and redundancies of IT infrastructures that make them available to users even in the event of a disruption. Such disruptions can be unexpected and range from anything as localized as the failure of a network card on a single server to something as dramatic (and improbable) as the physical destruction of an entire data center. Service disruptions can also be entirely routine and predictable, such as for server maintenance. Businesses do not prepare for all disruptive possibilities because they are necessarily likely, but because the benefit of such preparation outstrips the cost of operational downtime should such events ever occur.

Clustering can be used as a way to achieve high availability. Simply defined, a cluster is a group of computers working together to run a common set of applications and to present a single logical system to the client and application. The computers in the cluster are physically connected by local-area network (LAN) or wide-area network (WAN) and programmatically connected by cluster software. These connections allow workloads to fail over to another computer in the cluster in the case of network failure of scheduled maintenance, for example. This is an option not available to stand-alone computers.

In Windows Server® 2008, the next major release in the Windows Server® family of operating systems, Microsoft is introducing many new features and technologies that will help to increase the security of computers running Windows Server 2008, increase productivity, and reduce total cost of ownership (TCO).

Failover clustering is an important feature of the Windows Server platform that can improve availability. When one node fails, another node begins to provide service instead. This process is called failover.

Failover clusters in Windows Server 2008 provide high availability and scalability for mission-critical applications such as databases, messaging systems, file and print services, and virtualized workloads. Multiple servers (nodes) in a cluster remain in constant communication. If one of the nodes in a cluster becomes unavailable (such as a result of failure or having been taken down for maintenance), another node immediately begins providing service. Users who are accessing the service continue to access the service and are unaware that it is now being provided from a different server (node).

This document describes these new features and improvements for the failover clustering in Windows Server 2008. Foremost among these improvements is the vastly simplified user interface for creating and managing clusters.

# Installation, Management, and Administration

Clustering in Windows Server 2008 has been radically redesigned to simplify and streamline cluster creation and administration. Rather than worrying about groups and dependencies, administrators can create an entire cluster in one seamless step via a wizard interface; all they have to do is supply a name for the cluster and the servers to be included in the cluster and the wizard takes care of the rest. The end result is that you don’t have to be a cluster specialist or have in-depth knowledge to successfully create and administer Windows Server 2008 failover clusters. This means a far better total cost of cluster ownership for you.

Currently, clusters too often fail because of the complexity of configuring them. Hence Windows Server 2008 clustering comes with a built-in cluster Validate Tool (formerly known as “ClusPrep”). Validate runs a focused set of tests for both functionality and best practices on the servers that are intended to be in a given cluster as a part of the cluster configuration process. Validate performs a software inventory, tests the network, and validates system configuration.

The Validate inventory includes:

* Operating system binary consistency (to ensure that cluster nodes are running the same versions of the operating system as well as have the same hotfix and service pack level)
* Architecture (CPU architecture and memory information)
* Configuration (node domain membership and role; analysis of unsigned drivers)
* Devices (plug-and-play devices, host bus adapters, and network interface cards [NICs])

The Validate verification includes:

* Infrastructure (inter-node communication and SCSI compatibility with Persistent Reservations [PRs])
* Hardware (multiple NICs per node; shared disks accessible from all computers and uniquely identifiable)
* Software (each NIC has different IP address on a dissimilar subnet)
* Functionality (network and disk I/O latencies; failover simulation)

Validate test results are HTML-based for easy collection and remote analysis. The time Validate takes to run can be just a few minutes, though this is a function of how many nodes are in the cluster and how many LUNs are exposed to the servers and may take longer as a result. The minimum number of nodes in a given cluster configuration to run Validate is two.

Running Validate is a required part of cluster creation.

**Note:** When you run Validate, some tests may not pass, but clustering may still install and function. For example, not conforming to a cluster configuration best practice (such as having only one NIC in each node) will raise a warning rather than an error, but the cluster would function. However, passing Validate is the standard for support for clusters in Windows Server 2008: **If a cluster does not pass Validate, it is not supported by Microsoft.**

Once the cluster is created, Validate can also be used as a powerful diagnostic tool in maintaining the cluster. This is because you can run Validate at any time, even after you have created the cluster.

Windows Server 2008 includes a new, easy-to-use, management interface. The previous cluster administrator interface has been replaced with a Microsoft Management Console (MMC) 3.0 snap-in, CluAdmin.msc. This new interface is accessible from within Administrative Tools. (It is also possible to open a blank MMC and then add this snap-in along with any others.) The Cluster Administration Console (CluAdmin.msc) is designed to be task oriented: rather than playing with knobs and dials, administrators select the clustering task that they want to undertake (such as making a file share highly available) and supply the necessary information via the wizard. Administrators can even manage Windows Server 2008 clusters from Windows Vista® client computers by installing the Remote Server Administration Tools (RSAT).

Administrators can access advanced cluster administration options via the cluster.exe command-line interface. Moreover, Windows Server 2008 Failover Clusters are fully scriptable with Windows Management Instrumentation (WMI).

To significantly improve cluster security, Windows Server 2008 cluster nodes cannot co-exist in a legacy cluster. Windows Server 2003 server cluster nodes and Windows Server 2008 failover cluster nodes cannot be on the same cluster. In addition, failover cluster nodes must be joined to an Active Directory®–based domain (not a Windows NT® 4.0–based domain).

The process in moving from Windows Server 2003 clusters to Windows Server 2008 failover clusters will be a migration. The migration functionality can be accessed from a wizard in the Windows Server 2008 in the cluster management snap-in named Migrate Services and Applications. After the tool is run, a report is created that provides information on the migration tasks.

The migration tool will import critical resource settings into the new cluster registry. The migration process migrates clustered resource configuration information. This involves reading the Windows Server 2003 cluster database information for the resources being migrated and then importing that information into the Windows Server 2008 cluster database, realizing that the location of this information may have changed. The primary examples here are the dependency, crypto checkpoint, and registry checkpoint information has all been relocated within the Windows Server 2008 cluster registry structure.

The installation process in Windows Server 2008 has fundamentally changed as well. Roles and features (and the distinction between them) are now more important in Windows Server 2008 than before. For example, failover clustering is a *feature* because it makes other server *roles* highly available. You can install the failover clustering feature through the Initial Configurations Task (ICT) interface or with the Server Manager snap-in in Administrative Tools. In addition, you can uninstall clustering the same way.

The procedure to install cluster functionality in servers has changed dramatically with Windows Server 2008. Windows Server 2008 is far more compartmentalized than Windows Server 2003; cluster is no longer installed by default as with Windows Server 2003 and you must use the Add Feature Wizard to install the Failover Clustering feature. Windows Server 2008 uses the “componentization” model wherein the pieces and parts are not added until you need them. Be aware, there may be some roles and features that will be added by default on product installation. Also note that some roles and features may be needed prior to configuring a cluster resource. For example, the DHCP server role must be installed prior to clustering the DHCP service. The uninstall procedure also uses the same model—you remove features and/or roles. The new install model is also reflected in the new directory structure under windows\cluster.

# Architecture

Windows Server 2008 clusters can support more nodes than in Windows Server 2003. Specifically, x64-based failover clusters support up to 16 nodes in a single cluster, as opposed to the maximum of 8 nodes in Windows Server 2003, providing even greater scalability.

Clustering in Windows Server 2008 also includes a new networking model. To increase the security of nodes in the cluster, the heartbeat process—the process by which cluster nodes signal their integrity to one another—has changed from a broadcast format to a unicast format. In addition, the request-reply format has changed to increase cluster stability. With this change, nodes can be located on different, routed subnets and you can use multiple IP address resources for dependencies, using “or” logic. The cluster network driver (clusnet.sys) has been replaced with a Microsoft Failover Cluster Virtual Adapter (netft.sys). The cluster service has dependency on netft.sys.

Maintenance mode has changed as well. This mode now basically shuts off health monitoring on a node for a period of time so that it does not fail while you work on it.

Another new feature is support for Internet Protocol version 6 (IPv6). The IPv6 addressing protocol is emerging as an important factor in the growth of the Internet. IPv6 specifies addresses that are 128 bits long, compared to IPv4 addresses, which are 32 bits long. This greater address length allows for a much larger number of globally unique addresses to accommodate the explosive growth of the Internet around the world.

Importantly, IP address assignments via DHCP are also allowed in Windows Server 2008. Windows Server 2008 failover cluster nodes can also be on different logical subnets. If a given cluster node is configured to obtain its IP address via DHCP, then IP address resources will automatically be generated using DHCP for dependencies on network name resources. This is a “per-interface” configuration: if a node uses a static address, then the IP address resource will allow for static addressing only. Interfaces that are not configured with a default gateway will be configured for internal cluster communications only and client access will not be allowed unless a configuration change is made inside the snap-in for that network.

These changes afford much more flexibility in implementing geographically dispersed clusters: administrators no longer have to stretch virtual local area networks (VLANs) across the WAN to accommodate geographically distant servers that are on different subnets. Failover cluster nodes can now reside on completely different subnets. Moreover, the network latency requirements in Windows Server 2003 server clustering have been removed from Windows Server 2008 failover clustering; the failover clustering heartbeat requirement has also become fully configurable. Geographically dispersed clusters are easier to deploy and more technically feasible with Windows Server 2008 failover clustering.

Windows Server 2008 also features a new quorum model: Majority Quorum Model. Nodes in Windows Server 2003 R2 Enterprise Edition server clusters use a quorum to track which node owns a clustered application. The quorum is the storage device that must be controlled by the primary node for a clustered application. Only one node at a time may own the quorum. In contrast, the server clusters in Windows Server 2008 incorporate the benefits of single quorum device server clusters and Majority Node Set clusters from Windows Server 2003 R2 Enterprise Edition.

In Windows Server 2008, the term “quorum” refers to a majority of votes for which cluster node controls the cluster and cluster resources, as opposed to being a reference to a particular disk resource in the cluster. By default, nodes and storage all get a vote in Windows Server 2008 failover clusters. As long as a simple majority of votes are present, the cluster can stay up (for example, a single node and shared storage, or both of the nodes without shared storage, in the case of a two-node cluster). However, this is fully configurable: an administrator could set only nodes to get votes, for example, in which case the quorum would behave similarly to a Majority Node Set cluster in Windows Server 2003 R2 Enterprise Edition.

# Storage Improvements

Windows Server 2008 failover clusters are designed for Storage Area Networks (SANs). To this end, only Persistent Reservations (PR) SCSI commands to the disks are supported. In addition, the bus types most commonly used by SANs are supported: Serial Attached SCSI (SAS), Fibre Channel (FC), or iSCSI. Also, the clusdisk.sys function has been rewritten, resulting in much of its functionality being moved to partmgr.sys. In Windows Server 2008, disks are never left in an unprotected state. This significantly reduces the possibility of corruption. In addition, GUID partition table (GPT) disks are now supported and multiple terabyte storage (that is, larger than 2 terabyte LUNs per partition) is now natively possible.

## Closer Integration with VSS

Windows Server 2008 features a closer integration with Volume Shadow Copy Service (VSS), for easier backups. Failover clustering in Windows Server 2008 has its own VSS writer, which enables VSS backup applications to more easily support clusters.

# Security Improvements

The cluster service no longer runs under the context of a domain user account, also known as the Cluster Service Account (CSA). Instead, the failover cluster service runs under the local administrator account with the same privileges as CSA. This account relies heavily on the Cluster Name Object (CNO) in Active Directory.

The text-file-based cluster log is also gone. Event trace logging (.etl) is now enabled via Event Tracing for Windows (ETW). Default log sizes vary and can be modified with cluster installs logs (Operational and ClusterLog). There is new functionality built into the command line. The **cluster.exe** tool allows you to dump the trace log into a text file. This file looks similar to the cluster log used in previous versions of failover clustering. Use the **cluster.exe Log /Generate** command to see this log. Also, you can create diagnostic views inside Event Viewer. Moreover, the Microsoft Support Diagnostic Tool (MSDT) is a new tool that enables you to run local diagnostics specified by a Microsoft Customer Service and Support (CSS) engineer and to upload results online to CSS for faster analysis to help resolve issues

Legacy NT LAN Manager (NTLM) is no longer required, and instead Kerberos will be the primary authentication. All network name resources will have Kerberos enabled and a computer object created, such as CNO or Virtual Computer Object (VCO). In addition, now all communications between nodes are secure; encrypted inter-node communication is configurable, and off by default. The CNO corresponds to the cluster name and is created in Active Directory when the cluster is created. Since a domain user account is no longer used for the cluster service, the CNO is used for all operations that require security inside the cluster. For example, when a network name is created and the resource is brought online, the CNO is used to create the VCO for that Network Name resource in Active Directory.

# Conclusion

A major obstacle to traditional clusters was their complexity build, configure, and manage. This too often resulted in higher-than-necessary costs for organizations, both in lost opportunities to make more applications highly available and from retaining expensive administrative resources exclusively for maintaining the clusters that they did deploy. A major facet of the improvements to failover clusters in Windows Server 2008 is aimed squarely at these challenges, producing a clustering paradigm that radically simplifies and streamlines the cluster creation, configuration, and management process.

Changes to clustering in Windows Server 2008 also improve cluster performance and flexibility. Clusters are x64-based and now support up to 16 nodes. Clusters nodes can have their IP addresses assigned by DHCP and geographically dispersed clusters can span subnets. Windows Server 2008 failover clusters are designed to work well with SANs, natively supporting the most commonly used SAN bus types. And Windows Server 2008 failover clusters are built around a more resilient and customizable quorum model.

# Related Links

For further information, please see the following resources:

Windows Server 2008 home page   
[www.microsoft.com/windowsserver2008/default.mspx](http://www.microsoft.com/windowsserver2008/default.mspx)

Windows Server 2008 failover clustering  
[www.microsoft.com/windowsserver2008/failover-clusters.mspx](http://www.microsoft.com/windowsserver2008/failover-clusters.mspx)

Windows Server 2003 home page   
[www.microsoft.com/windowsserver2003/default.mspx](http://www.microsoft.com/windowsserver2003/default.mspx)

Microsoft Multi-Path I/O Frequently Asked Questions [www.microsoft.com/WindowsServer2003/technologies/storage/mpio/faq.mspx](http://www.microsoft.com/WindowsServer2003/technologies/storage/mpio/faq.mspx)

This is a preliminary document and may be changed substantially prior to final commercial release of the software described herein.

The information contained in this document represents the current view of Microsoft Corporation on the issues discussed as of the date of publication. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information presented after the date of publication.

This white paper is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS DOCUMENT.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in, or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Microsoft Corporation.

Microsoft may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Microsoft, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

© 2007 Microsoft Corporation. All rights reserved.

Microsoft, Active Directory, Windows, Windows NT, Windows Server, Windows Vista, and the Windows logo are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.