High Availability in Exchange Server 2007

Product Group - Enterprise

Dell White Paper

By

Ananda Sankaran

Mahmoud Ahmadian

August 2007
## Contents

**Introduction** .................................................................................................................. 3  
**High Availability for Server Roles in Exchange 2007** ........................................ 4  
**Mailbox Server Role Availability Options** ............................................................ 6  
  Single-Copy Clusters ...................................................................................................... 6  
  Local Continuous Replication ........................................................................................ 7  
  Cluster Continuous Replication .................................................................................... 10  
  Comparison ................................................................................................................... 14  
**Availability Options for Other Server Roles** ......................................................... 16  
  Hub Transport Server Role Availability .................................................................. 16  
  Client Access Server Role Availability ....................................................................... 17  
  Edge Transport Server Role Availability .................................................................... 17  
  Unified Messaging Server Role Availability ............................................................. 17  
**High Availability Deployment Considerations** ...................................................... 19  
**Conclusion** .................................................................................................................. 20
**Introduction**

Availability can be crucial for enterprise operations and is a key component in meeting return-on-investment targets. Enterprise messaging systems drive core business communications, and it becomes critical to ensure their availability. New high availability features built into Microsoft® Exchange Server 2007 are significant improvements aimed at providing flexible and cost effective data protection and availability options for Exchange deployments. In Exchange Server 2003, native mailbox availability capabilities were limited to shared storage clustering with Microsoft Cluster Services, which in Exchange Server 2007 is referred to as Single-Copy Clusters (SCC). Exchange Server 2007 introduces the additional mailbox availability options Local Continuous Replication (LCR) and Cluster Continuous Replication (CCR). This article will discuss Exchange Server 2007 high availability features and associated benefits. Performance analysis conducted at Dell labs to characterize the behavior of mailbox availability options will be discussed.
High Availability for Server Roles in Exchange 2007

Exchange Server 2007 provides or distributes its features and functionality through the five different server roles: Mailbox, Hub Transport, Client Access, Edge Transport and Unified Messaging. Figure 1 illustrates the different Exchange Server 2007 roles in an example Exchange Server 2007 infrastructure.

Figure 1: Microsoft Exchange Server 2007 Infrastructure
Each of the five server roles provides different Exchange Server 2007 functionality. Deploying each role in a highly available manner ensures availability of its supported functions. High availability for each server role can be provided using built-in Exchange Server 2007 features and other standard Enterprise methods depending on the server role. Deployments can be made with certain server roles in a highly available configuration and the rest in a standalone configuration. For example, depending on business and IT needs, the decision may be made to deploy the mailbox server role in a highly available configuration and the other roles in a standalone configuration. This choice may be driven by an IT decision that manually recovering from a mailbox server failure is complex and time consuming, with potentially unacceptable down time compared with recovering other server roles. Nevertheless, deployment planning should carefully assess the availability requirements for each server role based on the downtime SLA (Service Level Agreement) for that role’s functionality and associated recoverability parameters. The choice of a particular availability method for deploying a server role will be further influenced by various other factors such as cost, manageability, and scalability.
Mailbox Server Role Availability Options

The Mailbox server role is the most critical role in an Exchange deployment since it hosts the user mailbox data. It is crucial that the mailbox server application and the data it hosts are available and secure. Exchange Server 2007 provides options to achieve high availability for either the mailbox server application or the mailbox data or both. Deployments can employ one of these options based on the availability needs. The three native high availability options available for the Mailbox server role in Exchange Server 2007 are described in the following sections.

Single-Copy Clusters

Single-Copy Clusters (SCC) enable high availability for the mailbox server application. SCC is based on the shared storage Microsoft Cluster Services (MSCS) clustering model that existed with previous Exchange versions. It follows a shared-nothing architecture wherein a single copy of the storage groups and databases reside on shared external storage. Two or more Exchange mailbox servers are connected to this shared storage to form a cluster. At any time only a subset of these servers, called active nodes, host the mailbox databases residing on the shared storage. If one of the active nodes incurs a failure, another server in the cluster, called a passive node, takes over hosting its mailbox databases. This process is termed as a failover and these types of clusters are also popularly called failover clusters. Compared with previous versions, Exchange Server 2007 provides improved deployment setup and management experience with this clustering model. Enhanced management tools for Exchange Server 2007 such as the Exchange Management Shell allow administrators to manage clustered deployments seamlessly. Deployment support for Single Copy Clusters requires the configuration to be certified and listed in the Microsoft Windows Server Catalog under the Cluster Solution category. Deploying Exchange Server 2007 with other versions of Microsoft Exchange Server or SQL Server™ in the same cluster is not supported. However, deploying SQL Server 2005 Express Edition or other Microsoft desktop database applications (Office Access™) is supported on the cluster nodes in a standalone, non-clustered mode. Windows Server® 2003 supports up to eight server nodes in a cluster. It is recommended that each active mailbox server in the cluster have a corresponding passive node in the cluster. This is because at any time a cluster node can host only one active clustered mailbox server and, when multiple active node failures occur, there may not be sufficient passive nodes to handle the failover.
If the Mailbox server role is deployed in a highly available clustered configuration using SCC, no other server role can be consolidated with the Mailbox role on the same server. Thus in SCC deployments, separate servers, in addition to the mailbox cluster servers, need to be provisioned for deploying the Hub Transport and Client Access roles. SCC ensures availability of only the mailbox application. High availability for other Exchange server roles can be provided through mechanisms discussed in later sections.

**Local Continuous Replication**

Local Continuous Replication (LCR) provides availability at the data level to the mailbox databases. LCR does this by creating and maintaining a copy of the Exchange storage groups’ logs and database on a second set of disk volumes connected to the same mailbox server. The copy is maintained asynchronously using transaction log copy and replay on the target databases. The copied version of the storage group is called a passive copy and the corresponding original production version is called an active copy. The passive copy of the storage group is initially created by copying from the active storage group through a “seeding” process. Subsequent updates to the active storage group are synchronized with the passive storage group via a log copy process from the active to the passive. The copied transaction logs are then replayed on the passive storage group’s database. In case of an active copy failure, the mailbox server can be pointed manually to start using the passive copy as the production version. LCR provides an excellent option for quick recovery from user errors and data level disasters that cause database corruption. The passive database copies can be used to offload the required backup activities from the active databases with minimal impact to the end user response time.
Thus the backup window can potentially be extended to handle large data due to the minimal effect on the active copy. The backup schedule can also be modified to remove the requirement of daily full backups due to the availability of passive copy and instead use weekly full backups.

A set of performance tests were conducted on a LCR configuration as illustrated in Figure 3. The objective was to understand the performance impacts of enabling LCR on the mailbox server and storage subsystem. Simulations were conducted on the configuration with LCR disabled and enabled on the mailbox server. The utilization levels of various system resources were recorded, and their average values are displayed in Figure 4 and Figure 5.

![Diagram showing LCR Test configuration](image)

**Figure 3: LCR Test configuration** *(Source – Dell Power Solutions, August 2007)*

**Configuration Details:**

**Mailbox server**
- Dell PowerEdge 2950 with 2 x Intel® Xeon® 5160 dual core 3.00 GHz processors; 8 GB system RAM
- Windows Server 2003 R2 Enterprise x64 Edition with SP2; Exchange Server 2007 Enterprise Edition

**Hub Transport/Client Access server**
- Dell PowerEdge 2950 with 2 x Intel Xeon 5160 dual core 3.00 GHz processors; 16 GB system RAM
- Windows Server 2003 R2 Enterprise x64 Edition with SP2; Exchange Server 2007 Enterprise Edition

**External mailbox storage**
- 1 x Dell PowerVault™ MD3000 with 1 x Dell PowerVault MD1000 attached
  - Database volume: RAID 10 with 10 x 146 GB 15K RPM SAS drives (MD3000)
– Log volume: RAID 1 with 2 x 146 GB 15K RPM SAS drives (MD3000)
– Database copy volume: RAID 10 with 10 x 146 GB 15K RPM SAS drives (MD1000)
– Log copy volume: RAID 1 with 2 x 146 GB 15K RPM SAS drives (MD1000)

Microsoft Loadgen Simulation Tool
- Build version: 08.01.0094.000
- User profile: 1000 heavy users executing 94 tasks per 8 hr user day in Outlook 2007 online mode

As shown in Figure 5, the database disk reads and writes and the log disk reads and writes are at about the same level when LCR is either disabled or enabled on the active copy. The passive copy incurs fewer database reads than the active copy during LCR activity. The log disk activity is also slightly distinct between active and passive copies due to the logs being read from the active copy and replayed on the passive copy. The server shows slightly increased utilization of processor and memory when LCR is enabled (Figure 4). About 4% extra memory and 15% extra CPU is utilized with LCR enabled. Thus with LCR enabled more processor and memory should be allocated to the mailbox server.

Figure 4: LCR processor and memory utilization levels
Generally it is recommended that about 20% extra processor utilization be accommodated and about 1 to 1.5 GB extra memory be allocated on the mailbox server for LCR. The passive copy incurs about 25% less database I/Os than the active copy. The number, type, and RAID configuration of disk drives allocated to the passive copy volume can be adjusted for this factor. If the passive copy will be promoted to production for supporting regular scheduled maintenance of the active copy, then it is recommended that the passive copy be provided the same set of disk resources as the active copy.

**Cluster Continuous Replication**

Cluster Continuous Replication (CCR) enables high availability of both the mailbox server application and the mailbox data. CCR is based on the Microsoft Cluster Services (MSCS) Majority Node Set (MNS) model. This model does not require shared storage and the two mailbox server nodes within the cluster maintain their own copy of mailbox databases. At any time one of the cluster servers acts as the active node and serves client requests. The other server remains as a passive node and automatically takes over if the active node incurs a failure. CCR provides automatic server failover capability similar to SCC. A third server, called witness file share, provides arbitration and allows only one server to function as active node at any time. The passive copy of storage group is kept consistent with the active copy through data replication. Data replication is performed by Exchange Server at the application level and not by the MNS cluster infrastructure. The replication takes place in the form of asynchronous transaction log copy and replay. The passive node copies the transaction logs of the active node from a secure file share on the active node and replays them locally to update its database. Before enabling CCR, the passive copy is created by copying the active database copy through a seeding process to ensure consistency. CCR can provide site resiliency by deploying the active and passive nodes in geographically separate datacenters. CCR eliminates the need for a shared storage, and the storage space allocated to the database and logs on the passive node must be equal or larger than
that of the active node. CCR offers enhancements to data backup and recovery strategy similar to LCR. The passive database copy in CCR can be utilized for offloading required database backup activity and also to reduce backup frequency.

The cluster nodes communicate with each other about their liveliness via private “heartbeats”. When the nodes within the CCR cluster cannot communicate with each other, the witness file share serves as a tie-breaker to avoid split-brain scenarios where each node functions as the active node. The file share witness server must be hosted by a machine that is not part of the CCR cluster but should be part of the Active Directory® domain containing the cluster nodes, and it is recommended that it be the Hub Transport server. The file share witness cluster resource type is a recent additional cluster resource provided in Microsoft Windows Server 2003 Service Pack 1. Windows Server 2003 SP1 also added configurable cluster heartbeats to cluster services. This is especially significant for clusters which could be deployed in geographically dispersed locations. This feature allows configuration of cluster heartbeat parameters so that temporary network problems do not cause unnecessary failovers. The Hub Transport server includes a feature called Transport Dumpster which is utilized to avoid data loss during failover of the active mailbox server node. The Transport Dumpster, maintained in the Hub Transport server, is a queue of recently delivered messages to the active mailbox server from clients and other services. At the time of a failure, the active mailbox server may not have completed processing these recent messages. Thus to ensure their completion after a failover, all Hub Transport servers in the Active Directory site are requested to resubmit the mail in their Transport Dumpster queue to the new active mailbox server. This ensures that all recently delivered messages get recorded and prevents data loss during failures. Certain use cases or scenarios exist where this feature does not provide complete recovery from user data loss.

If the Mailbox server role is deployed in a clustered configuration using CCR, no other server role can be consolidated with the Mailbox role on the same server. Thus in CCR deployments, separate servers, in addition to the mailbox cluster servers, need to be provisioned for deploying the Hub Transport and Client Access roles. CCR ensures availability of only the mailbox application and mailbox data. High availability for other Exchange server roles can be provided through mechanisms discussed in later sections.

During CCR operations, work is performed by the passive node to copy log files from the active node’s secure file share and to replay the logs to its database copy. A set of performance tests was conducted on a CCR configuration as illustrated in Figure 6. The objective was to understand the performance impacts of CCR on the active/passive mailbox server nodes and their associated storage subsystem. Simulations were conducted on the configuration with CCR disabled and enabled on the active node. The utilization levels of various system resources were recorded, and their average values are displayed in Figure 7 and Figure 8.
Figure 6: CCR Test Configuration (Source – Dell Power Solutions, August 2007)

Configuration Details:

Mailbox servers (Active and passive)
- Dell PowerEdge 2950 with 2 x Dual Core Intel Xeon 5160 3.00 GHz processors; 8 GB system RAM
- Windows Server 2003 R2 Enterprise x64 Edition with SP2; Exchange Server 2007 Enterprise Edition

Hub Transport/Client Access server
- Dell PowerEdge 2900 with 2 x Quad Core Intel Xeon X5355 2.66 GHz processors; 16 GB system RAM
- Windows Server 2003 R2 Enterprise x64 Edition with SP2; Exchange Server 2007 Enterprise Edition

External mailbox storage
- 1 x Dell PowerVault MD1000 each on active and passive node
  - Database volume: RAID 10 with 10 x 300 GB 15K RPM SAS drives (MD1000)
  - Log volume: RAID 1 with 2 x 73 GB 15K RPM SAS drives (MD1000)
  - Database copy volume: RAID 10 with 10 x 146 GB 15K RPM SAS drives (MD1000)
  - Log copy volume: RAID 1 with 2 x 73 GB 15K RPM SAS drives (MD1000)

Microsoft Loadgen Simulation Tool
- Build version: 08.01.0094.000
- User profile: 1000 heavy users executing 94 tasks per 8 hour user day in Outlook® 2007 online mode

As shown in Figure 8 the database reads and writes and the log writes remain almost the same when CCR is disabled or enabled on the active node. The log reads show a slight
increase on the active node when CCR is enabled due to the log copy process. The passive node shows increased database write activity and log reads compared to the active node. This is due to the continuous log read and replay process happening on the passive database copy. The processor and memory utilization remain almost at the same levels when CCR is enabled and disabled on the active node as shown in Figure 7. The utilization on the passive node is lower compared to the active node during CCR activity.

Figure 7: CCR processor and memory utilization levels

Figure 8: CCR database and log I/O generated
Generally both active and passive nodes should be provided with the same set of processor and memory resources for CCR. The requirements are similar to the standalone mailbox server requirements for a given user load. The passive node should be provided with the same processor and memory resources as active node, even though during CCR activity it underutilizes those resources. This is because, after failover, the passive node is required to be operated as the active node with the actual user load. The active node incurs about 10% extra database I/Os during CCR activity, and this needs to be considered while designing the storage subsystem. The passive node incurs about 40% more database I/Os than the active node for CCR. This can be considered in designing the storage subsystem for the passive node to meet the ideal I/O performance best practices for the passive node. However, since the passive node does not handle the production load, the additional I/O requirement need not be considered while designing its storage subsystem. The passive node’s storage may exhibit less than optimal performance behavior, which may be acceptable for certain deployments. This tradeoff should be carefully assessed to appropriately design the passive node storage subsystem. In most cases, it should be provisioned with at least the same set of disk resources as the active node.

The network infrastructure between the active and passive nodes should be carefully planned as well. The public network interface on all cluster nodes should be on the same subnet. The private network interface on all cluster nodes should be on the same subnet as well but different from the public network subnet. Virtual LAN (VLAN) implementations can be used between cluster nodes to overcome geographical limitations imposed by the subnet restrictions. However, it must be ensured that the cluster private network point-to-point roundtrip latency is less than 0.5 seconds. The private heartbeat network, used for determining the health of cluster nodes, should be capable of sending and receiving heartbeat information within a certain number of retries which is configurable. If the nodes are distant from each other, the network latency will become a determining factor of how fast the log copy process occurs on the public network and the resultant lag of the passive copy from the active copy. Sufficient network bandwidth should be accounted between the nodes, especially during the initial seeding process, when the database is copied over the network to the passive node. Also, very large databases incurring high update activity will require ample bandwidth to copy the resultant logs from the active node to the passive node within reasonable latency limits. A gigabit Ethernet network infrastructure at a minimum is recommended for such deployments, along with mechanisms such as NIC teaming for the public network on the cluster nodes to ensure bandwidth.

**Comparison**

Figure 9 provides a high level comparison of the three mailbox high availability features described in earlier sections. Specific deployment rules apply for deploying these features. Appropriate product and feature documentation should be consulted before planning these deployments.
<table>
<thead>
<tr>
<th>Feature</th>
<th>SCC</th>
<th>CCR</th>
<th>LCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability level</td>
<td>Application</td>
<td>Application and Data</td>
<td>Data</td>
</tr>
<tr>
<td>Automatic Failover</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Native Data Replication</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Site Resilience for Disaster Recovery</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Microsoft Windows Server Catalog</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Listing for cluster solution hardware</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backup improvements</td>
<td>No</td>
<td>Yes – offload to passive copy</td>
<td>Yes – offload to passive copy</td>
</tr>
</tbody>
</table>

Figure 9: Mailbox availability feature comparison

A new replication feature called Standby Continuous Replication (SCR) is introduced with Exchange Server 2007 Service Pack 1. The SCR feature enables data replication to standby recovery mailbox servers that are not part of an MNS cluster, providing a more flexible data replication topology option. At the time of the publication of this white paper Exchange Server 2007 SP1 was not released.
Availability Options for Other Server Roles

It may be required to deploy other server roles in a highly available configuration in addition to the Mailbox server role. Availability for these server roles can be achieved through a variety of methods. The following sections describe some of these methods.

Figure 10: High Availability for server roles

Hub Transport Server Role Availability

High Availability for the Hub Transport server role can be provided by deploying multiple Hub Transport servers within the site. Internal transport services on the Exchange server roles by default work with multiple Hub Transport servers, and additional configuration or external network load balancing support is not required. The Hub Transport servers deployed in the site
are identified through their configuration information in Active Directory by transport services. If one of the deployed servers fails, other servers in the group will ensure availability of the Hub Transport role. Deployments can provision two or more Hub Transport servers per site based on the availability and scalability needs.

**Client Access Server Role Availability**

High Availability for the Client Access server role can be provided by deploying multiple Client Access servers and using load balancing solutions to distribute the incoming client requests to those servers. The native Windows Network Load Balancing (NLB) services or standard third-party hardware- or software-based load balancing solutions can be utilized. Deployments can provision two or more Client Access servers, configured with a load balancing solution, based on the availability and scalability needs.

Windows NLB Services is included within the appropriate Windows Server 2003 Edition and does not mandate other software or hardware requirements. The service needs to be configured on all the client access servers to form an NLB cluster. The servers in the cluster will be represented by a single virtual IP address and network name. All client requests will be directed to this virtual address, and one of the servers in the cluster will service the client request. If a server fails, the remaining servers in the cluster continue providing services ensuring service availability and scalability. As a best practice, each server can have one network adapter for participating in the cluster and another network adapter for any traffic directly addressed to it. Certain features of the Client Access role require client affinity, where a series of requests from a client need to be satisfied by the same server that maintains state and not by other cluster members. Windows NLB Services includes options to enforce this affinity. DNS Round Robin is another method usually considered for load balancing client requests across multiple servers. A list of server IP addresses is maintained at the DNS and client requests are resolved to one of these IP addresses by going through the list in a Round Robin fashion. DNS Round Robin cannot enforce client affinity, and also requests need to be retried through the list if servers have failed, since failure information is not maintained.

**Edge Transport Server Role Availability**

High Availability for the Edge Transport Server role can be provided by deploying multiple Edge servers in the perimeter network and load balancing requests using DNS Round Robin described in the earlier section or using multiple mail exchangers (MX) records in the DNS. DNS documentation of the implementation deployed can provide further configuration details on these methods. Multiple Edge Transport servers can be deployed using these methods, depending on the load balancing needs.

**Unified Messaging Server Role Availability**

High Availability for the Unified Messaging server role can be provided by deploying two or more Unified Messaging (UM) servers. If VOIP gateway devices are deployed between the UM
servers and the Private Branch Exchange (PBX), then two or more identically configured VOIP gateways may be deployed to ensure their high availability. The UM servers need to be configured as part of the same dial plan. The VOIP gateways can be configured to load balance calls across these servers using Round Robin methods and also reroute calls to another server if one of the servers is down. Deployments where an IP-based PBX is directly configured to route calls to UM servers should be configured for the same load balancing and rerouting functionality. Multiple UM servers and VOIP gateways can be deployed based on availability and scalability needs.
High Availability Deployment Considerations

The LCR, CCR, and SCC features cannot be deployed to coexist on the same mailbox server configuration and are mutually exclusive. A mailbox server can be configured with LCR, CCR, or SCC. If the Mailbox server role is deployed in a highly available clustered configuration using SCC or CCR based on Microsoft Cluster Services (MSCS), no other server role can be consolidated with the Mailbox role on the same server. Thus for SCC or CCR deployments, separate servers need to be provisioned for deploying the Hub Transport, Client Access, and other roles. Consolidating server roles on multiple servers requires careful consideration. An example scenario may be consolidating Hub Transport and Client Access roles on two servers, with both roles consolidated on both servers for HA. Since each of the server roles requires a different method for achieving HA, careful planning is required to ensure their interoperability.

Regardless of specific consolidation and deployment scenarios, hardware for each server role should be planned and sized appropriately for both capacity and performance.
Conclusion

Exchange Server 2007 provides new availability options that enable businesses to effectively meet their availability requirements and protect their messaging systems against hardware failures. Administrators should carefully evaluate the level and type of availability required before deciding which option is most appropriate for their environment. The choice of a particular option will depend on various factors such as cost, downtime, recoverability, scalability, and manageability. Implementing the required availability features and configuring them for optimal performance can help create flexible, highly available systems in enterprise data centers.

Dell PowerEdge servers, Dell PowerVault storage, and Dell | EMC storage provide standard hardware platforms for seamlessly deploying Exchange Server 2007 with the required availability features. More information can be obtained at www.dell.com/exchange. Dell Services include assessment, design, and implementation tailored for those messaging deployments. Dell also offers end-to-end Exchange messaging solutions that include partner offerings for security, archiving, backup, and recovery. More information can be obtained at www.dell.com/secureexchange.