Realtime publishers

The Essentials Series: Configuring Microsoft Exchange for Fast Recovery

Addressing Hardware
Needs and Determining
Recovery Objectives

by Ron Barrett

Addressing Hardware Needs and Determining Recovery Objectives	1
Hardware Best Practices	1
Processor	2
Memory	3
Storage Considerations	
Disk Space, Redundancy, and RAID Levels	3
Configuring Disk Space Correctly	4
Configuring Users and Storage Groups	5
Recovery Point Objectives	6
Bandwidth Needs	6
Managing Databases and Log Files	7
Summary	



Copyright Statement

© 2008 Realtime Publishers, Inc. All rights reserved. This site contains materials that have been created, developed, or commissioned by, and published with the permission of, Realtime Publishers, Inc. (the "Materials") and this site and any such Materials are protected by international copyright and trademark laws.

THE MATERIALS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. The Materials are subject to change without notice and do not represent a commitment on the part of Realtime Publishers, Inc or its web site sponsors. In no event shall Realtime Publishers, Inc. or its web site sponsors be held liable for technical or editorial errors or omissions contained in the Materials, including without limitation, for any direct, incidental, special, exemplary or consequential damages whatsoever resulting from the use of any information contained in the Materials.

The Materials (including but not limited to the text, images, audio, and/or video) may not be copied, reproduced, republished, uploaded, posted, transmitted, or distributed in any way, in whole or in part, except that one copy may be downloaded for your personal, non-commercial use on a single computer. In connection with such use, you may not modify or obscure any copyright or other proprietary notice.

The Materials may contain trademarks, services marks and logos that are the property of third parties. You are not permitted to use these trademarks, services marks or logos without prior written consent of such third parties.

Realtime Publishers and the Realtime Publishers logo are registered in the US Patent & Trademark Office. All other product or service names are the property of their respective owners.

If you have any questions about these terms, or if you would like information about licensing materials from Realtime Publishers, please contact us via e-mail at info@realtimepublishers.com.



Addressing Hardware Needs and Determining Recovery Objectives

The term *mission-critical application* does not capture the importance of email in the enterprise. Email is truly a *core application*; it is in many instances the lifeblood of an organization. Such being the case, the need to maintain availability is crucial.

Of course, when you plan and implement your infrastructure, it is with the idea that you will never have the need to recover from a failure. However, if you need to recover Exchange, it is best to ensure that the recovery can be done quickly. This series will focus not only on getting Exchange up and running again but also on ways to make sure it is recovered quickly. Anyone who has ever suffered through an Exchange Server failure knows 10 seconds down is 10 seconds too long. There is no way to promise that you can be back up in 10 seconds, but the information provided here will help shrink that recovery window and get you back online quickly.

Every good recovery plan starts with the hardware you choose. It is equally important, though, to make sure that hardware is properly configured if you are going to achieve fast recovery.

Hardware Best Practices

Minimum requirements for hardware would obviously not be fitting for scenarios where you are looking for fast recovery. These best practices will therefore exceed minimum and match or exceed recommended requirements.

Whether you are planning, designing, or have already implemented an Exchange environment, it is never too late to look at best practices for hardware. This is especially true because hardware costs are so incidental. Moreover, when you compare the cost of additional storage or extra RAM to the cost of an Exchange Server down for the day, you have a very compelling argument for even the stingiest CFO.

When discussing hardware for fast recovery in Exchange, many factors come into play for an organization—such as the number of server roles deployed, and whether these will be on a single server or broken down into multiple servers; how many storage groups will be created; the number of mailboxes per storage group; and mailbox size restrictions (or lack thereof for certain organizations).

A later section will talk specifically about storage considerations. Let's begin by looking at processor and memory considerations.



Processor

As Exchange 2007 works only on a 64-bit processor, this part is simple. Either the Intel EM64T or AMD 64 processor will work.

Warning

The Intel Itanium (IA64) processor will not work with Exchange Server 2007, eliminating this processor as an option.

Recommended configuration for each role is 4x processor cores except the Edge Transport server role, for which 2x processor cores are recommended. The recommended configuration for a server hosting multiple roles is 4x processor cores. Maximum configuration for Client Access, Edge Transport, Unified Messaging, and multiple roles still calls for 4x processor cores; however, for the Hub Transport and Mailbox role, 8x processor cores are recommended.

Of course, people love the idea of throwing as much hardware as possible at a mail server. Consider a few factors to help you decide whether you need the recommended or maximum configurations.

The Hub and Edge Transport server roles processor utilization is based on numerous factors such as number of transport agents, average message size, message rate, configuration of antivirus, and use of third-party applications. The Client Access and Unified Messaging server roles processor utilization is based on the size of the environment. 4x processor cores will be more than sufficient in many environments for these roles.

The rule of thumb for the configuration of the Mailbox role is 1x processor core per 1000 average user mailboxes. Average user mailboxes are classified as 40 messages received and 10 messages sent per user. Even if you classify all your users as very heavy users (120 received and 30 sent), you are looking at 2000 mailboxes for a 4x processor core server.

Note

For servers hosting multiple roles, you should reduce your mailbox count by 20 percent to account for the processor utilization by the Hub Transport and Client Access roles.



Memory

Memory considerations for fast recovery require specific configuration considerations. It is recommended to have 1GB of RAM per core (with a 2GB minimum) for all the server roles except the Mailbox role. The recommended configuration for this role is 2GB plus 5MB per mailbox and 2GB for every four storage groups. As the focus is on fast recovery—and again taking into consideration the incidental cost of RAM—8GB for an average and 16GB for a heavy usage Mailbox role is best practice.

If you are hosting multiple server roles, the recommendation is 8GB plus 5MB per mailbox and 2GB per four storage groups. For fast recovery of Exchange hosting multiple server roles, best practices suggests 16GB for an average and 32GB for a heavy usage Mailbox role.

Note

Even with a new server, going beyond 32GB becomes cost prohibitive for most organizations.

Storage Considerations

Of all the hardware recommendations, storage considerations are one of the most significant. Storage considerations do not simply include how much space you need for Exchange; they also encompass how you will configure and manage that storage space for fast recovery. Therefore, you must look at disk space, disk redundancy, RAID configurations, and other best practice configurations for fast recovery in Exchange.

Disk Space, Redundancy, and RAID Levels

Let's begin with disk space considerations. When speaking about disk space for fast recovery, you are not talking about the size of the disks needed. Rather, you are talking about the percentage of available disk space in correlation to your disk size. Ideally, you want to maintain 40 percent availability for disks in a fast recovery scenario. This number is not an arbitrary value but represents a performance threshold to maintain an environment suitable for fast recovery of Exchange.



Traditionally, for mission-critical applications such as email, capacity and redundancy are considered crucial to successful implementation. Although this is true to an extent, these factors are not as significant as the need to control disk latency and optimize performance—both of which are critical for fast recovery of Exchange. Therefore, in this case, best practice is to opt for a RAID-1+0 solution. Although this setup cuts the available capacity of the disks, it provides for better I/O for read/write operations in Exchange, which is very I/O intensive. In addition, it is recommended to invest in high-performance disks and spindles. Again, this may cut down on the disk capacity (cost being a direct factor for many). Considering that disk subsystems are a greater bottleneck on system performance than processors or memory, high-performance disks are perhaps the best investment you can make in your messaging infrastructure.

Configuring Disk Space Correctly

Having the appropriate amount of available disk space and redundancy is not the only factor to consider. How you use that space is imperative. Exchange Server is I/O intensive, so you need to ensure you use the proper RAID levels and alleviate both the single point of failure as well as bottlenecks in recovery, both of which will defeat all the efforts you put forth.

The first recommendation is one you have heard for years, yet it is widely ignored or incorrectly configured. To start, the operating system (OS) should be installed on a separate volume. Separating the database and transaction logs onto separate physical drives provides both fault tolerance and fast recovery. This reduces risk of data loss if a disk failure occurs by ensuring that databases and transaction logs are not compromised by the loss of a single array—again, no single point of failure.

Another step in configuring hardware for fast recovery is the use of the diskpart.exe tool. Use diskpart.exe to create aligned disk partitions; having the tracks *sectoraligned* will improve performance by nearly 20 percent. Keep in mind that diskpart.exe will not work with dynamic disks and can only be used with basic disks.



Configuring Users and Storage Groups

The configuration of users and storage groups is an area where you will see variance depending on the user base, mailbox quotas, and other configurations. Microsoft recommends one storage group to one database. However, the Standard edition of Exchange allows five storage groups with as many as five databases per storage group. The Enterprise edition allows 50 storage groups with 50 databases per storage group.

In keeping with the objective of fast recovery, it is most beneficial to plan the storage group, databases, and, by extension, user mailboxes in such a way to allow for better performance. For example, a 4x processor core with 12GB of RAM would meet the requirements for hosting 2000 mailboxes in a single storage group on a server hosting only the Mailbox role.

Considering a maximum mailbox size of 250MB per user, you are looking at roughly 504GB of storage. However, that is only part of the equation; add in space for the database dumpster (104MB) and white space (7.3MB), and you now have a mailbox size of 306MB. 2000 mailboxes now require 844GB for databases and 198GB for log files.

In this example, the recommendation is to have eight storage groups and databases with 250 mailboxes per storage group. This would bring you to 130GB per database with log files. Now you can very quickly see the logic in dividing up users into storage groups. Even if you need to recover multiple databases of 105GBs, it will go much quicker than a single 844GB database recovery. Of course, if you only have the standard version of Exchange, eight storage groups is not an option for you, so creating five storage groups of 400 users is the best practice.

You can drive yourself crazy and do the math alone or you can download the Exchange 2007 Mailbox Server Role Storage Requirements Calculator from the Exchange team blog site at

http://msexchangeteam.com/files/12/attachments/entry438481.aspx.



Just to Drive Home the Point

Other storage considerations that have not been included to the individual mailboxes are

- Content Indexing—5 percent database size
- Maintenance—10 percent database size
- Log files—12 to 36MB per mailbox
- Log file overhead—20 percent of total log file space allocated

Although breaking up storage groups does not eliminate the disk space needed to support 200 users, it certainly gives you a much better way of managing them to meet your recovery point objectives (RPOs).

Recovery Point Objectives

Howard Aiken said, "Time has a wonderful way of weeding out the trivial." Not a bad philosophy when you talk about RPOs. Although the goal is to lose nothing at all, the objective of fast recovery is to get up and working quickly with minimal data loss. No data loss would be phenomenal, and achieving that goal comes with a price tag. Albeit the price has dramatically been reduced thanks to many of the technologies introduced in Exchange Server 2007.

Simply put, RPO is your company's policy for the acceptable amount of data loss in a distressed situation. RPO is measured in time, so it is determined in this case by both the amount of acceptable data loss (emails) and recovery time.

Bandwidth Needs

Although RPO is considered more of a hardware issue, it is fitting that to look at the requirements for bandwidth along with RPO. Available bandwidth on both the LAN and WAN side will affect fast recovery. From a LAN perspective, best practice is to have teamed Gigabit network cards for load balancing and fault tolerance. In addition, optimize link speed and duplex settings. In most instances, when you talk about bandwidth needs for fast recovery, you are talking about restoring a mail database(s) that has become corrupt or unreachable locally. Occasionally, you need to think in terms of a complete unrecoverable disaster. The amount of time it takes to back up or replicate data to an alternative off-site location becomes a bandwidth issue. Allocating sufficient bandwidth depends on many factors, including which method or backup is used and the organization's RPO, RTO, and availability of high-bandwidth communications. Large cities have the availability of affordable fiber optic 100Mbit Internet speed. Others are limited by the traditional 1.54Mbit connectivity, so this will be a factor.



Managing Databases and Log Files

One factor to keep in mind for database and log file management in Exchange is the need to back up the environment. Ultimately, the size of the database, backup windows, Service Level Agreements (SLAs), and available resources are going to dictate the size of the database and frequency of maintenance. Another effective method of managing databases and log files is utilizing mailbox quotas. For some organizations, this is easier said than done.

Managing log files is best accomplished by backing up the database and dumping committed logs. Monitoring backup time and performance issues may make moving storage group paths another option. Moving the storage group from a disk that is more than 40 percent full will enhance performance overall and provide a better scenario for fast recovery.

Summary

Addressing hardware needs and considering your RPOs is the first step in successfully implementing a fast recovery solution in Exchange Server. The next article will consider the different backup methods available for Exchange.

