



White Paper

Eight Best Practices in Implementing Disaster Recovery Protection for Microsoft Exchange Server 2000/2003

Abstract

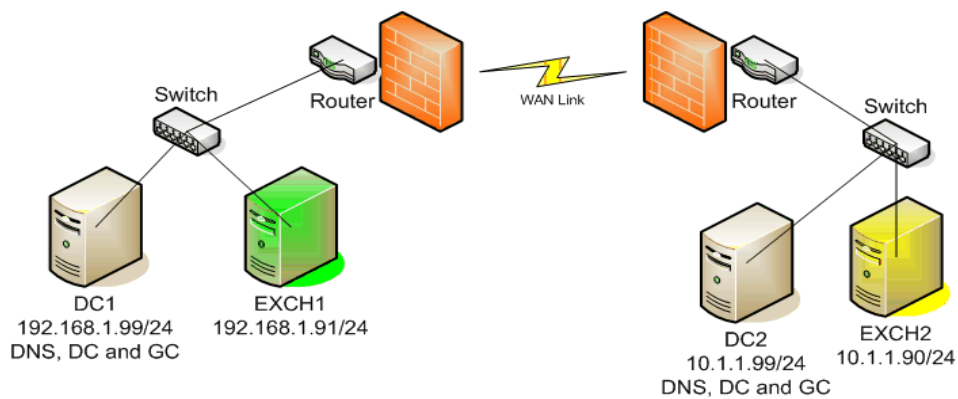
Protection of Microsoft Exchange messaging and collaboration services against unplanned outages due to catastrophic failures at primary IT data centers is becoming a top priority among large and medium corporations. Wide-area clusters which provide replication of mail stores and log files along with monitoring and recovery of critical Exchange services provide this protection. This paper presents eight best practices which, when used as guideposts during cluster implementation, assure the success of any Exchange disaster recovery deployment.

Introduction

Exchange server downtime costs companies around the world millions of dollars a year. Technically-savvy IT organizations work to eliminate or lessen the impact of both planned and unplanned downtime through the implementation of high-availability and disaster recovery solutions. The building of a geographically dispersed Exchange cluster is, however, not a simple task and requires taking several critical factors into consideration. At SteelEye Technology, we have built, deployed and are today supporting wide-area Exchange disaster recovery clusters throughout the world. In this paper, we review eight best practices which SteelEye Technology has developed over five years of deploying Exchange disaster recovery solutions. By understanding the issues raised in this discussion and considering the recommendations during design and deployment, you will deploy a solution that best meets your business objectives.

Disaster Recovery for Exchange

The diagram below shows the primary components of an Exchange disaster recovery implementation. At the primary corporate data center, EXCH1 acts as the primary Exchange server with DC1 providing Active Directory, Domain Controller, DNS and Global Catalog functions. A parallel set of systems exist at the disaster recovery site. Connecting the sites is a WAN connection over which flows all data required to bring Exchange into operation at the disaster recovery site when required. This data is transported across the WAN link typically via data replication software which resides on the Exchange server; storage-based replication can also be used. Certainly most topologies will be more complex with additional services available at each of the two sites. For the purpose of focusing this discussion on Exchange, we will be concerned with only the pieces required to build an Exchange disaster recovery configuration.



Exchange Disaster Recovery Cluster Topology

Two components not shown in the diagram but key to the solution, and which must be considered in the design of any Exchange disaster recovery solution, are the users (clients) which access the Exchange servers and any software which relies on or supports the messaging infrastructure. SPAM filters, virus scanners, PDA connectors, along with other services are likely resident within the Exchange environment and must be accounted for in design. These components should be monitored and recovered as part of the overall solution.

Exchange users may be running Outlook, OWA, or any other third-party email client. They may be directly connected in a local LAN or accessing via WAN or PDA connections. Regardless, the success of an Exchange recovery is measured by the ability of users to continue to send/receive email, access shared folders and any other critical task which may be defined in the project requirements. A client-side perspective must be kept at all times in the design of a successful Exchange disaster recovery solution.

The pieces of the Exchange solution together provide a powerful and mission-critical collaboration and workflow system. The protection of the overall solution against outages resulting from disaster at primary data centers must be a business priority.

The Eight Best Practices

This paper presents eight best practices that SteelEye has developed over our years of building Exchange disaster recovery networks. Following these practices will maximize the likelihood that the deployment will meet your availability objectives and ensure receipt of the highest possible return on your investment.

1 Deploy Sufficient Bandwidth to Disaster Recovery Site

In an Exchange disaster recovery scenario, all mail stores, database and log files must be replicated from the primary Exchange server to the remote backup server so that they are available as needed to operate Exchange during recovery operations. Adequate bandwidth for database and transaction log replication is critical to ensuring that when a failover occurs, the most recent data possible exists at the backup site.

$$\left(\begin{array}{l} \text{Database} \\ \text{.edb file} \\ \text{.stm file} \end{array} \right) + \left(\begin{array}{l} \text{Entries in} \\ \text{Transaction logs} \\ \text{(*.log)} \end{array} \right) = \left(\begin{array}{l} \text{Current} \\ \text{Database} \end{array} \right)$$

If the network connection over which the current database flows is too small, one of two things will occur: either Exchange server performance will degrade waiting for acknowledgements to remote data writes or data will backup in the pipe between locations. Either scenario leads to undesired results, either a poorly performing primary Exchange server or an out-of-date backup Exchange server.

Deciding on the optimal bandwidth between sites is not complicated; it simply takes measuring current usage and understanding the characteristics of a few critical factors. There are two measurements that yield critical factors to network pipe size. The first is Average Rate of Change which tells how much your current database changes on average per hour. We recommend that you measure the database rate of change for two weeks, and then compute the average across this period. The second measurement is Sustained Peak Rate of Change which finds the period of greatest change across the two week sample and notes how much data changed, and would therefore have to be replicated, across the network connection between sites.

As a guide to finding the values needed, follow these steps:

1. Place your Exchange Log Files, Database Files and SMTP Queues onto a dedicated partition(s).
2. Run Microsoft Performance Monitor to collect the "Disk Write Bytes/sec" counter on that logical volume(s). Once you have collected that data over a period of two weeks, use your favorite data analysis tool (Excel, Access, etc) and get the following two values:

Average Rate of Change (in MB per hour) = (Average(Disk Write Bytes/sec))*3600/1,048,576

Sustained Peak Rate of Change = Use the above formula for each 60 minute period and then find the MAX number of the results.

All other factors being equal, the connection that you deploy between sites should be higher than the Average and lower than the Sustained Peak. Of course, anticipated growth in users and other changes in the environment which may lead to increased usage should also be considered. Other factors which also must be analyzed as you ensure optimal network connectivity between the sites include network quality and latency and the efficiency of the replication engine being utilized.

There are aspects of an Exchange Disaster Recovery solution, as we will see later, where cost can be a deciding factor in how to proceed. This is not one. If you cannot afford to invest in a suitable network pipe between your primary IT data center and your backup site, then you cannot afford to build a true Exchange business continuity plan. The measurements that we have discussed above give the minimum that you must deploy; cutting corners on network bandwidth are guaranteed to lead to a failed project.

2 Determine Recovery Objectives; Choose Optimal Protection Solution

The requirements to be met by the Exchange recovery solution should be well known before beginning the design and certainly before beginning the deployment of the solution.

Questions answered during the requirements gathering stage should include:

- ⊗ What is the Recovery Time Objective, or how quickly must Exchange recover following some catastrophic event at the primary data center?
- ⊗ What events classify as serious enough to justify migration of Exchange to the disaster recovery site?
- ⊗ Will automatic recovery of Exchange be allowed on software detection of a problem, or should a human make the final decision on failover? If the latter, how will that human be notified that an error has been detected?

With these requirements understood, an evaluation of software solutions should be undertaken to find the one which will best meet project objectives. Along with meeting the requirements defined in the project plan, the solution chosen should meet the following criteria:

- ü Integrated wide-area data replication and Exchange health monitoring
- ü Automated and configurable failover and failback
- ü Ability to easily perform manual migration of Exchange to allow for administration of primary server without causing downtime
- ü Configurable to meet site-specific requirements such as methods of administrator notification, setting of intervals between heartbeats
- ü Ability to protect adjunct services including SPAM filters, virus scanners and PDA connectors
- ü Supported by an organization with experience in Exchange disaster recovery as demonstrated by time in the market and customer successes

The choice that you make in the software foundation on which the Exchange disaster recovery solution is built is critical to project success. You need a solution which supports you in achieving the recovery objectives and is also flexible in configuration so that it can be optimized for your site and can be changed as objectives change over time. Make your selection carefully and you are well on the way to a successful implementation.

3

Configure for Predictable and Timely Restore

The SteelEye Availability Equation states $T_{RESTORE} = T_{DETECT} + T_{RECOVER}$ and represents that the restoration time of any failed application is determined by the time required to detect the failure plus the time required to complete a recovery procedure. For Exchange, detection of an outage occurs through monitoring the health of:

- the physical server on which Exchange runs
- individual Exchange services
- IP addresses used by clients to access the Exchange server and associated network connections
- volumes on which the mail stores and log files reside
- any supporting components unique to the specific Exchange deployment such as workflow applications, PDA connectors, etc.

These checks should be configured to ensure best operation in your specific Exchange environment. For example, you may want to vary the intervals at which heartbeats are sent between the sites or the number of heartbeats which must be missed before a site is determined to be down based on the latency of your WAN link. You must also decide if you want to allow the monitoring software to take an automatic recovery action on failure detection, or if you instead want a system admin alerted so that a human makes the final decision to failover. Adding an admin notification and confirmation to the recovery process will add time to the recovery, but can prevent false failovers due to intermittent and short-lived network outages.

There may be certain site-specific error conditions for which you want to optimize the detection algorithm; perhaps certain services tend to go down frequently or you have determined that a pattern of virtual memory fragmentation is a precursor to Exchange outages. Monitoring for these conditions should be regular and frequent so that detection and subsequent recovery are as fast as possible.

4

Ensure the Integrity of Replicated Data

To recover Exchange at the remote Disaster Recovery site requires that the Exchange Current Database (see Best Practice 1) be available. In a typical deployment, this is accomplished via data replication between the Primary Exchange server (the source) and the Backup Exchange server (the target). As changes occur in the Exchange database on the source, those same changes are sent to the backup database on the target. In a wide-area deployment, this replication will typically be asynchronous in nature meaning that there will possibly be writes in transit between the two systems – these are writes that have been committed to the database on the source side, but not yet written on the target. The more of these writes which are outstanding at the time of the failover, the longer recovery will take as the log files are used to reconstruct missed messages. To minimize recovery time, make sure that adequate bandwidth exists between the sites so that writes do not queue up in the pipe.

It is also critical that the replication technology guarantee that writes sent from source to target occur in correct time order sequence. This ensures consistency in the log files and mail stores, allows log file replays to rebuild a consistent mail store and minimizes the chance of database corruption if in-transit writes should not be committed at the target side due to some failure and this leads to different database contents on the two sides. A persistent (disk-resident) intent log allows partial resynchronizations following any failure and can greatly reduce recovery time. By tracking the changes in the local Exchange database which have not yet been sent to the remote mirror, only these changes must be sent across the wire during recovery; without a persistent intent log, it is possible that a full synchronization

between the sides would have to occur which can result in extremely long periods of data remaining out of sync.

Make sure that the WAN replication engine which you deploy maintains write order integrity when running in asynchronous mode, replicates only changes after the initial synchronization and implements a disk-resident intent log to minimize resync time. Finally, even though you are implementing a solution which provides off-site real-time replication of the Exchange data, archival to tape remains an absolute requirement. Data replication is not a substitute for a good tape archive solution; it is an adjacent process to facilitate rapid recovery.

5 Remember the Clients; Choose Appropriate Redirection Method

Recovering Exchange at the Disaster Recovery site is only half of the solution; the other critical component is ensuring that all email clients are able to reconnect to Exchange after it is restored at the Disaster Recovery site.

Several methods exist to redirect clients to Exchange running at the Disaster Recovery site. They include:

- **Bridged Network using hardware or VLAN**
This method involves the use of a hardware/software combination to extend the local subnet out to the Disaster Recovery site. It is the most expensive of the automated options but also provides the most seamless recovery method since, in essence, you are migrating Exchange among servers on a LAN and can use virtual IP addresses to float between the active and backup systems. This is the preferred method and delivers the fastest recovery.
- **Route Updates**
An identical subnet is configured in both the primary site and the Disaster Recovery site. The only servers in this DR subnet are the primary Exchange server and the DR Exchange server. When running Exchange in the primary site, the route to this subnet is advertised on the primary site. When a failover occurs, the failover software can programmatically make a SSH or Telnet connection to the router(s) to change where the DR subnet is advertised. Once this change is made and convergence is completed among the routers, clients are automatically redirected to the Exchange server running in the DR site. By utilizing the route update client re-direction method, you can create and protect a virtual IP address to facilitate automatic client redirection without requiring flushing DNS.
- **DNS Updates**
Upon recovering Exchange in the DR site, the DNS A record of the primary Exchange server is updated to reflect the IP address of the Exchange server in the DR site. The DNS update can be accomplished automatically by calling the DNSRECORD.PL script (from W2K Resource Kit, Supplement 1). In the SteelEye solution, this is done from the LifeKeeper recover script. Since all Outlook clients are configured to connect to the primary Exchange server, they will automatically be redirected to the Exchange server running in the DR site. Clients may have to wait for the TTL of the DNS A record to expire or flush their DNS cache before they are redirected to the Exchange server running in the DR site.
- **Alternate Clients**
In a true disaster, it may be the case that regular desktop clients are not available. In these cases, alternate clients such as OWA connections directed to the recovery server can be used for connectivity to the remote site.

The final solution is a means of last resort since it requires human involvement in the recovery process, thereby introducing risk of manual error and also delaying the recovery.

- **Manual reconfiguration**
A documented set of procedures can be provided to all users so that they can reconfigure their email client to point to Exchange running at the Disaster Recovery site following failover. This should only be used if there is absolutely no way that you can automate client redirection.

In deciding which of the methods to use, three primary factors should be taken into account: your recovery time objectives, the types and location of email clients and project budget. The various options each come with trade-offs as reviewed above, making the decision as to which is best for your needs should be done with these characteristics in mind since the end-user experience will be directly related to the solution chosen.

6 Keep Active Directory in Sync with Forced Replication

Exchange 2000/2003 and Active Directory (AD) are very tightly integrated with Exchange having hard dependencies on AD for its operation. Active Directory provides both network security services and all of the services provided by a messaging system — for example, routing, address lookup, and the maintenance and replication of all Exchange attributes. From a client's perspective, Active Directory provides the Global Address List and access to offline address books. On installation, Exchange 2000/2003 extends Active Directory with new Exchange classes and attributes and adds a variety of messaging-related attributes to the user, group, and contact objects in Active Directory — which causes these security accounts to become mail-enabled. In addition, the user object can "own" a mailbox and so receives mailbox-related attributes.

During recovery of Exchange to the Disaster Recovery site, client AD objects which hold the identity of the Exchange server are updated to point to the new active system. Additionally, a number of other objects which point to Exchange services are also modified. These changes are then replicated by Windows to all AD controllers throughout the forest. Depending on the topology and size of the forest, this replication may take from 15 minutes to 24 hours. Obviously, this replication period can become the gating item for recovery times and can also impact the ability to move Exchange back onto the primary server following correction of whatever situation caused the initial outage.

To ensure that AD updates are propagated in a timely manner to facilitate the fastest possible recovery, it is critical that a mechanism exist for forcing replication. An Active Directory replication engine which makes use of AD APIs ensures replication on demand and places updated Exchange server information at locations where remote clients can gain access. Integrating this replication feature directly into the Exchange recovery procedure provides the fastest possible recovery.

It is important that you understand your corporate-wide AD topology and associated replication periods as you compute expected Exchange recovery times. Investigate the use of an integrated AD replication engine to facilitate the propagation of AD updates throughout your entire environment so that all clients will receive updated information and be able to reconnect into Exchange at the disaster recovery site as quickly as possible.

7

Protect and Recover All Pieces of the Solution

Exchange is about more than email. It has become a hub for collaboration, calendaring, information sharing and workflow. Your environment likely incorporates virus scanners, SPAM filters, PDA connectors and workflow applications which connect into the Exchange server.

Whether these pieces reside on the Exchange server itself or on separate gateway systems, they must be protected along with the standard Exchange services. Procedures, ideally automated as part of the restore of Exchange to service, must be in place to migrate the critical services to the backup site during recovery. As additional services are added into the Exchange infrastructure, the Disaster Recovery solution must be easily adaptable to include monitoring and recovery of these new pieces.

The ability to expand your Disaster Recovery solution beyond protection of Exchange to include these adjunct components is critical to recovery of the overall messaging and collaboration system and to the productivity of end-users, and should be a primary consideration in selecting the monitoring and recovery foundation for your deployment.

8

Plan, Plan, Plan, Test, Test, Test, Deploy, Validate, Validate

The most important of the best practices is proper solution planning, testing, deployment and validation. This begins at the initial decision to implement an Exchange disaster recovery solution and continues as long as it is in production.

In the planning phase, you must answer these critical questions which have been previously discussed:

- ⊗ What is the Recovery Time Objective for the solution?
- ⊗ Based on rate of change and expected growth, what bandwidth is required between the primary and disaster recovery site?
- ⊗ What pieces of the solution besides the Exchange server and processes themselves must be monitored and recovered? Is a minimum set of functionality acceptable for some time following recovery and if so, what is that subset?
- ⊗ Are there site-specific error conditions that should be optimized for in the monitoring phase?
- ⊗ Will automatic failover be allowed, or should administrator notification be the first recovery action?
- ⊗ Which email clients must be protected during failover, what access methods do they use and how will these be migrated to the disaster recovery site during recovery?

With these requirements documented, you begin design. If not already known, the choice of a disaster recovery site should be the first decision made. While some organizations decide to use remote corporate offices as backup locations, many look to co-located hosting centers which provide full redundancy for power and network connectivity. If using a remote office location, understand that it is probably not going to be possible to piggyback the replication traffic required for the mail store and log file mirrors onto the existing cross-site network traffic. An analysis of bandwidth requirements and availability must be done. Designing a sufficient connection between the sites – in terms of bandwidth and latency – is critical to deployment success.

At this stage, you should also identify and document the servers, the storage capacity and configuration, the network routing between primary and DR site, and the method of client redirection that will be used.

Personnel with the following skills need to be involved in the entire process from design thru final validation:

- Û Windows server administration
- Û Exchange server administration
- Û Clustering software administration
- Û Network routing and troubleshooting

These skills are critical to building the end-to-end solution and should be involved in developing the initial design, in building the test environment, hands-on in the deployment and subsequent validation.

The testing phase is often difficult because properly emulating the production Exchange environment within a test lab requires simulation of cross-site network connectivity, user load, influence of adjunct programs such as backup jobs and other factors unique to the specific environment. The closer you can get to a mirrored production environment within the test lab, the fewer issues you will see arise during deployment.

The use of Exchange Server Load Simulator (loadsim) allows you to test how a server running Exchange responds to email loads. Having an active loadsim session running while performing failover tests is an effective way to test the consistency of replicated data and the behavior of the failover software when running under loads similar to what you expect to see in production. We have also found hardware-based WAN simulators which can be programmed for various speeds and latencies to be invaluable as both a test and debug tool.

In testing, you are looking to answer several questions:

- Ø Does the failover software perform as expected on both detection and recovery?
- Ø Does the data replication software perform as expected in terms of speed and data consistency?
- Ø Is there any noticeable performance impact on end-users from the presence of the monitoring or data replication software?
- Ø Are the various clients able to seamlessly migrate to Exchange running on the disaster recovery server?

Given your assembled team of experts and a successful testing phase where you have emulated the production environment, deployment should be straight-forward. Of course, you will want to have scheduled sufficient time for the initial synchronization of the cross-site mirror and then subsequent testing of failovers to the disaster recovery site and then back to the production server. Each of these tests should validate that client redirection works as planned and that all services needed for a fully functional Exchange environment are recovered.

Following the initial deployment, it is recommended that failover tests be made after any change (installation of service packs, introduction of new services, etc.) to the Exchange environment. A worst-case scenario is for there to be a disaster at the primary location and the DR solution not bring Exchange into service because of an administrative change that was not accounted for in the recovery process. Not surprisingly, perhaps, most of the error conditions that are reported back to us result from an administrator making a change within the Exchange environment and not realizing the impact on the disaster recovery solution. Even in a static environment, a test should be run at least quarterly to ensure proper monitoring and recovery.

While it is true that deploying an Exchange disaster recovery solution is a complex task, following the eight best practices presented here will lead to a successful deployment which meets the requirements for protecting business critical messaging and collaboration environments.

About SteelEye Technology and LifeKeeper for Exchange

With a singular focus on providing High Availability and Disaster Recovery solutions for Linux and Windows, SteelEye Technology understands the needs of business to minimize mission critical system downtimes and its flagship product, LifeKeeper, has been delivering protection for these mission critical Windows applications since 1996. LifeKeeper was the first high availability clustering solution for Microsoft Exchange with the introduction of an Exchange 5.5 solution in 1998. Since that time, LifeKeeper has been protecting all flavors of Exchange in a variety of configurations throughout the world. Today, the LifeKeeper® Microsoft Exchange Server Recovery Kit provides fault resilience for all flavors Exchange Server 2000 and 2003.

The advantages of SteelEye's implementation of an Exchange Disaster Recovery Solution include:

- LifeKeeper works on all versions of Windows and Exchange server, not just the higher-end (and more costly) enterprise versions.
- When combined with LifeKeeper Data Replication, data replication and high availability clustering are tightly integrated from a single vendor.
- LifeKeeper supports a range of cluster topologies including data replication, shared SCSI, Fibre SAN and iSCSI.
- LifeKeeper Data Replication provides asynchronous volume replication with a single source system going to multiple targets to provide both local and remote copies of Exchange mail stores and logfiles.
- LifeKeeper supports having more than just Exchange active on the protected server and gives the ability to support any application running on that server. If you are a small business who cannot dedicate a server to Exchange, or if you are a business who wants to run "Add-on" programs such as fax servers, etc. on the Exchange server, LifeKeeper supports your configuration.
- Built-in Active Directory replication engine ensures that updates made to AD during recovery processing are immediately propagated to all controllers within the AD forest.
- Through its web-enabled JAVA-based GUI, LifeKeeper is easier to install, configure and administer than any competing clustering technology.

SteelEye LifeKeeper is a flexible and established data and application availability management solution with years of history protecting Microsoft Exchange from planned and unplanned outages. Combined with LifeKeeper Data Replication, it serves as a firm foundation on which to build an Exchange disaster recovery solution.

Businesses must protect their messaging assets from catastrophic failure since the result is a significant loss in productivity, lost revenue, decreased customer satisfaction and increased operational costs. Corporations who care about customer and employee loyalty and retention in the next millennium absolutely cannot afford these failures. The opportunity cost is too great. Following the eight best practices presented in this white paper will ensure a successful deployment of disaster recovery protection for business critical Exchange environments.

SteelEye, LifeKeeper and LifeKeeper Data Replication are registered trademarks of SteelEye Technology, Inc. Other brand and product names used herein are for identification purposes only and may be trademarks of their respective companies.

This document is for informational purposes and is believed to be correct at the time of publication. However, SteelEye does not guarantee the accuracy of the information and reserves the right to change the document at any time. SteelEye makes no warranties, expressed or implied, in this document.

Copyright © 2005
SteelEye Technology, Inc.
Palo Alto, CA U.S.A.
All Rights Reserve