

# Windows<sup>®</sup> IT Pro

## **Replay™ for Microsoft Exchange: Enterprise Quality Protection and Recovery at an Affordable Price**

By Mark Arnold

Published: May 2007



# Replay™ for Microsoft Exchange: Enterprise Quality Protection and Recovery at an Affordable Price

High Availability, Backups, Replication. Today's Microsoft Exchange environments are replete with such solutions. Each one of them has their own niche, their own place in the portfolio of solutions that the Exchange architect may select as part of their Exchange protection design.

Recording data changes at the block level, that is at the level of the physical disk block, is the most efficient way of protecting information but this has hitherto really only been available as a hardware-based solution and even then only one storage vendor, Network Appliance, has the patented method to improve efficiency and performance.

Continuous Data Protection, CDP, is seen by many as the way forward but if you don't have an expensive SAN environment your options were previously limited to such technologies as Double Take and Neverfail. Now there are other options to take efficient changed data blocks and protect not only the Exchange databases but also the entire system from the bare metal upwards.

Appassure Software has developed just such a solution, branded Replay™, to deliver continuous data protection not only to a customer base previously unable to utilise block level protection, but also to SAN customers who wish to have an alternative to expensive licensed options for Exchange protection.

This white paper will present the key challenges facing an Exchange administrator and how Appassure Replay can be deployed to offer comprehensive system protection.

## → Contents

Exchange Server and Continuous Data Protection .....	2
The Importance of Backups .....	2
Backup Drawbacks.....	3
Bit Level vs. Block Level .....	3
The AppAssure Solution.....	3
Recapturing Lost Time .....	4
Rolling Back to Prevent Infection.....	5
Recovery Console Builder .....	6



Copyright 2007 Appassure Software. All rights reserved.

## Exchange Server and Continuous Data Protection

Exchange Server is a mission-critical application to the vast majority of corporations, more than many realise or care to admit. Service Level Agreements, (SLAs) are often out of step with the needs of the users, dating back to when email was an application that people could get along perfectly well without. Very often the “standard” SLA still applies, which might say that next day service is sufficient. Should IT Administrators take that SLA at face value and simply conduct a plain old restore from backup tapes and take the server off line for up to a day or do they work through the problem as if their lives depended on it? Obviously the answer is the latter but the next question is what tools do they use to ensure that the server never becomes unavailable?

The simple answer to that is they simply don't, or more accurately, can't. Servers go down, power fails, databases corrupt, service packs and patches are required. Are all these a part of your SLA matrix? Probably not. It's a good bet that database corruption is missing from your SLA recovery matrix because the support managers got a “well, that all depends” from the Exchange Administrators and they decided to go with something very generic. Information store problems come in all shapes and sizes, from total loss due to hardware failure right through to bizarre happenings that not even a Jet database specialist at Microsoft is able to talk you through. Sometimes a restore is the only option.

But what kind of restore? Isn't it far better to try to make sure that you can recover quickly from anything that happens to your Exchange server using a single application that will go from an entire server failure down to a block on the disk that represents the last change made to the Exchange store. And what about a single-item restore. SLAs rarely, if ever, include a mandate that a single mailbox or a single item must be restored within a specified time.

Highly granular restores are certainly available to you if you have a good Storage Area Network with an innovative snapshot backup technology such as that available within a good Storage Area Network with an innovative snapshot backup technology such as that available within Network Appliance Data ONTAP. But what if you don't have a SAN or you have a SAN from another vendor who doesn't offer you quite such a flexible solution? And of course what are your options when the SAN vendor charges a sum that you simply cannot afford or justify to your business? Other vendors have taken elements of their existing backup technologies and customised them

for various Microsoft products, making some parts of the backup and restore process arguably too complex and time-consuming. However, Replay™ has been purposely designed to enable administrators to control their whole Exchange Server 2003 and 2007 backup, restore, and recovery process from one integrated application.

In addition to “simple” restore requirements when you know what the problem is, even if not how to solve it without a restore, there is the case of audits and other compliance-related situations. Should you be required to maintain additional infrastructure just on the off chance that a need to trawl through numbers of backups in order to restore one or more messages? Similarly, should you be required to hold large amounts of full backups amounting to many dozens of gigabytes of storage, again on the off chance that it might be required? Wouldn't it be far more preferable to maintain a much smaller volume of physical data but nevertheless the same amount of information in order to satisfy the needs of the auditors or investigators?

### The Importance of Backups

Restores are all well and good but what about backups, what about the space they take and what about the time they take? A restore is no good unless the backup is valid and a backup is no good unless it can enable the restore within a very short timeframe. Continuous Data Protection (CDP) is an umbrella term for a wide range of solutions. CDP essentially means being able to conduct backups in a manner that is non-intrusive to the user base, where the data is recoverable quickly, known as the Recovery Time Objective or RTO, and where the Recovery Point Objective (RPO) is as close as possible to the point where the information store stopped – for whatever reason it stopped.

Administrators can deploy a wide range of solutions to assist in both server availability and recovery when that server or component, software or hardware, fails. The various types of Microsoft clustering services can be employed to increase server availability – with varying degrees of success.

Microsoft has a number of clustering and other high availability options available. Single Copy Clusters (SCC) are conventional clusters with two or more physical servers and a single copy of an information store. Local Continuous Replication (LCR) is a method of making a copy of the information store and hosting it on a different set of disks on the same server. Clustered Continuous Replication (CCR) is a new high availability method in Exchange 2007 that enables you to maintain a second copy of an information store on another server. The

new Standby Continuous Replication (SCR) capability in Exchange 2007 Service Pack 1 will let Exchange administrators achieve the same level of off-server information store replication, without the need to have experience in cluster technologies.

## Backup Drawbacks

There is a major drawback to each and every one of these resiliency measures: information store corruption. Each of the three methods, SCR, LCR and CCR, rely on log shipping to take log files generated by the primary information store and replay those log files into the secondary store. A corruption in an inbound message to the primary will be replicated, through the log files, into the secondary. Each of the solutions are categorized as high availability solutions but none offers continuous data protection; each of them rely on backups of the stores and have their own unique issues surrounding the complexity of restores and failovers should it be necessary. When Replay™ takes its block-level backups, it does a checksum on the data as a corruption detection measure, thus ensuring that the information written to the Replay™ server is valid.

Microsoft Data Protection Manager can be deployed throughout the organization to offer a level of CDP. But, as with many products from Microsoft, there are a range of features and add-ons that are not included within the core product because Microsoft sees the benefit in an independent vendor producing either a competing product for a different customer profile or an enhancement package that sits on top of a Microsoft product.

The next problem with continuous data protection is the amount of data being protected and the volume of changes being made. Conventional backup applications look at changes in files so that they can register the change and make a copy of the file. That can be very wasteful in terms of the amount of data and therefore disk and tape space that must be maintained. If only a small portion of a file is changed what is the point of marking the entire file for backup?

## Bit Level vs. Block Level

Microsoft DPM looks at bit level changes, which is generally little better than taking backups using file level changes. All SAN vendors instead use block level backups, meaning that data on individual disk blocks are monitored for changes and the location of those blocks that do change are recorded so that a restore is as simple as changing the pointers in the metadata back to where they

were. Thus an Exchange information store that is dozens of gigabytes in size may be restored to a previous state in a matter of tens of seconds rather than one or two hours.

Ordinarily, if you don't have access to a Storage Area Network your options are very limited. There are applications that will effectively clone the Exchange Server to another physical server by looking at changed blocks and sending them over the network. Some of these applications give you an option to roll back to a previous state should a corruption or other problem occur. However, these applications are not CDP candidates because they are simply used to provide high availability for the application they are associated with. If you chose to pause the replication from the production to the secondary server in order to carry out some work on the secondary, you will lose the capability to quickly fail over to the secondary. It may well be the case that changes are queued up on the primary server while the secondary copy of the information store is being worked on, but having the changed information on the primary is certainly no way to protect the data.

Replay™ from AppAssure is a product that allows you to continuously protect the entire Exchange server, not just the information stores. It does this by regularly sending changed blocks to the AppAssure Replay server while simultaneously allowing Exchange administrators to work on one or more previous instances of an Exchange information store.

## The AppAssure Solution

Let's take a deeper look at AppAssure and how it will help you provide end-to-end comprehensive CDP to your Exchange server.

Continuous Data Protection relies on something fairly obvious; data is continually being monitored and backed up to ensure that it can be recovered quickly should the need arise. To do this, Replay™ monitors data at the disk block level rather than the file level or even part of the file (byte) level. SAN technologies rely on this approach to ensure that the lowest possible level of data can be protected. This means that the agent on the server being protected does not have to be highly complex and deeply application-aware.

Conventionally understood backups are very expensive in either disk space, tape or more usually both. An Exchange full backup will use very large amounts of disk, dozens of gigabytes in most cases. Conducting full backups followed by incremental or differential backups

instantly lengthens the restore process and does nothing to rationalize the amount of data stored on disk at any given time. Since Replay™ works in concert with the Microsoft Volume Shadow Copy subsystem (VSS), it is therefore fully “Exchange Aware” and performs backups of the Exchange information stores on a schedule of your definition. Since the VSS process is involved the number of transaction logs that are consistently resident on the system is extremely small so consideration can be given to handing over more space to the information stores rather than over provisioning log file space.

## Recapturing Lost Time

Securing blocks of data also means that backups are very much more efficient in terms of data secured and recoveries can therefore be very precise in terms of the granularity of the data to be restored.

As more organizations agree to SLAs that commit to providing email services around the clock, the pressure constantly increases on non-client-facing activities such as backup and maintenance cycles as well as the inevitable service pack and patching regimes. Replay™ gives you several hours per day back because the continuous data protection is constantly securing the information on your server. This means you no longer need to assign a block of time after what used to be “normal” working hours. While

there is a penalty for this—typically one to two percent of processor utilization throughout the day—there should be no Exchange server that cannot take this load in stride.

A major shortfall with other backup applications and certainly other CDP solutions is the ability to conduct a full and reliable validation of the integrity of the backup, either on the server that was backed up or on a different server. Not having the confidence that any and all backups taken from the Exchange server are valid means that you have to take other measures to validate the backup. Replay™ gives you the ability to do that verification on the server hosting the backup files and then mount a virtual database store in order to carry out single mailbox and single item restores. Furthermore, Replay™ will achieve the verification very much faster than other applications.

Replay™ takes the last backup and, knowing that the integrity of the database was valid at that point verifies only the changes, or deltas, allowing for quick verification completion. While products such as Network Appliance SnapManager for Exchange will take many snapshot backups, they must complete a full verification of the entire database, resulting in a delay before a database is marked as verified. Replay™ gives you many more backups and makes those backups available to you sooner.

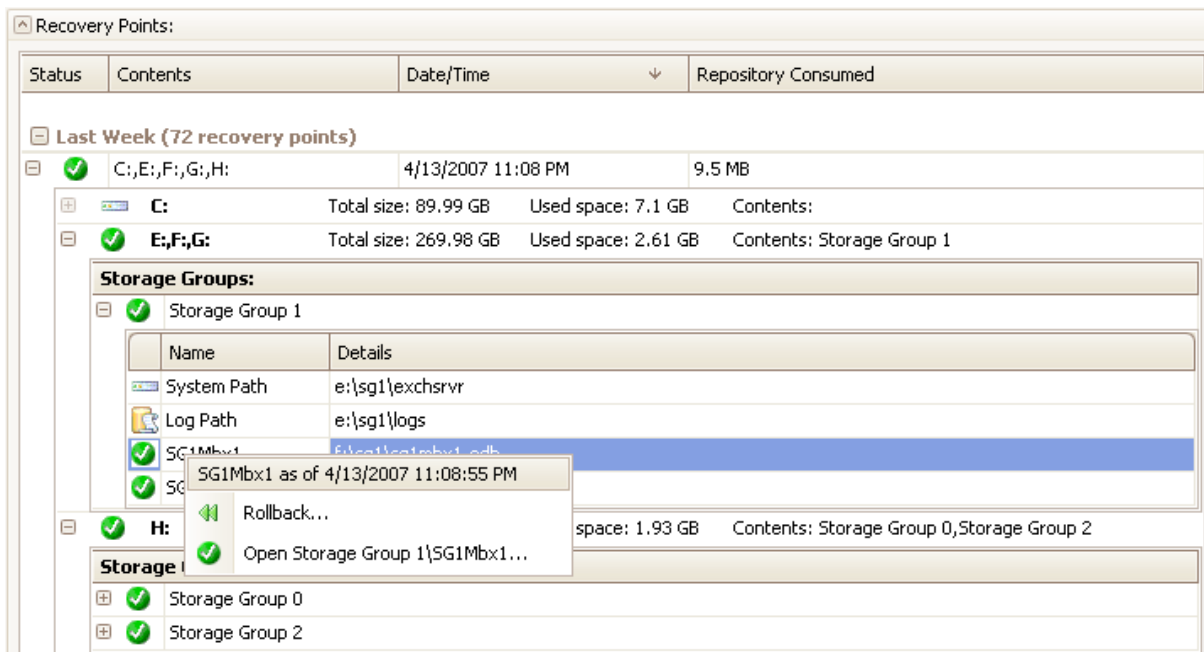


Figure 1. Validated Recovery Points

Recovery of the Exchange server without significant extra investment in hardware is also a challenge for many other applications. Replay™ allows you to mount a backup quickly and easily on the AppAssure Replay Server in order to extract messages and mailboxes directly from the backup, all in a couple of mouse clicks. Other applications exist to restore information directly from a backup, either file or media, but those require the use of a third-party backup application to take the backup in the first place. Other CDP products such as Microsoft DPM require the Recovery Storage Group (RSG) to put data back onto an Exchange server before it can be accessed properly.

## Rolling Back to Prevent Infection

A virus infection is a rare event but serves well as an example for the next piece of functionality that the continuous data protection within Replay™ offers. Should

a non-corrupt but nonetheless undesirable message enter your system and propagate, the Exchange user base, on its own, can do very little about it. Replay™ gives you the capability to take an immediate copy of your current state and then roll back your Exchange stores to a previous point in time. The very final backup snapshot you took allows you to extract legitimate and desirable messages and then play them into the information store again. You have therefore recovered from a virus infection within minutes and with, in all likelihood, no loss of legitimate message flow. Without the assistance of technologies such as AppAssure Replay™ the story would have been very different, involving lots of downtime and many impatient users. Recovering without email loss would have been impossible because the log files that belong to the affected storage group would also have contained the very message that you were trying to expunge from your information stores.

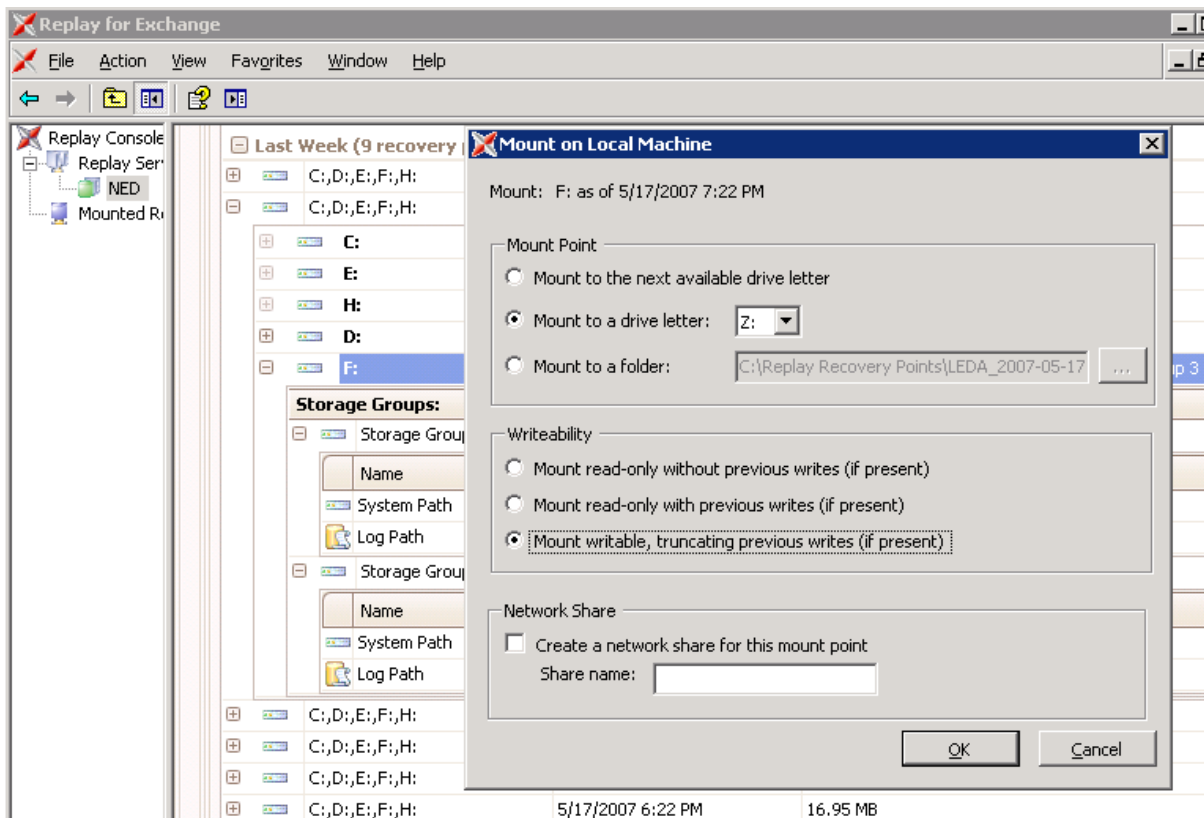


Figure 2. Mounting a Virtual Database

Exchange 2007 is no different from Exchange 2003 in that it is not possible to restore a single public folder item without restoring the entire public folder store onto a separate piece of hardware, usually one that is off the main production network so a restoration of a domain controller is also required. The private information stores can make use of the Recovery Storage Group on any server but the public stores cannot. Replay™ gives you the flexibility to dispense with your non-Microsoft backup application such as Veritas Backup Exec or CA ARCserve and allow the package that secures your entire hardware platform to restore single items, folder trees or indeed an entire public folder store right to the production server or anywhere else that you require.

## Recovery Console Builder

Should your Exchange Server need a complete rebuild due, perhaps, to complete server failure, then Replay™ has a solution to facilitate that. The Recovery Console Builder lets you build a base image to get a new server onto the network but no farther. As with other applications, a complete image of the server can be placed onto the new hardware but this is where Replay™ departs from other applications such as Ghost, Windows Deployment Services, (formerly Remote Installation Services or RIS)

etc. Replay™ goes the extra step by taking the block level changes and applying the Exchange Server application to the hardware and then Exchange information stores. Replay™ lets you select the restore point to which you wish to take the new server, you are not forced to roll forward to the last backup taken. This is particularly useful in non-disaster recovery situations. Taking, for example, the need to move from an old Exchange server to a new one but maintaining the same details such as IP address and computer name as before.

Ordinarily it is not recommended to take a system offline and restore it, en masse, to a new piece of hardware because the simpler and safer method is to introduce the new hardware and migrate Exchange mailboxes and public folders to it progressively and in a controlled manner. However, there are always those situations where you need to maintain the same name for whatever reason. Replay™ allows you to take a final block level backup of a server and then shut the old server down. The Replay™ Recovery Console is then used to recreate the Exchange server on the new hardware and apply all of the applications and information in a controlled manner, right up to the point the old server was shut down.

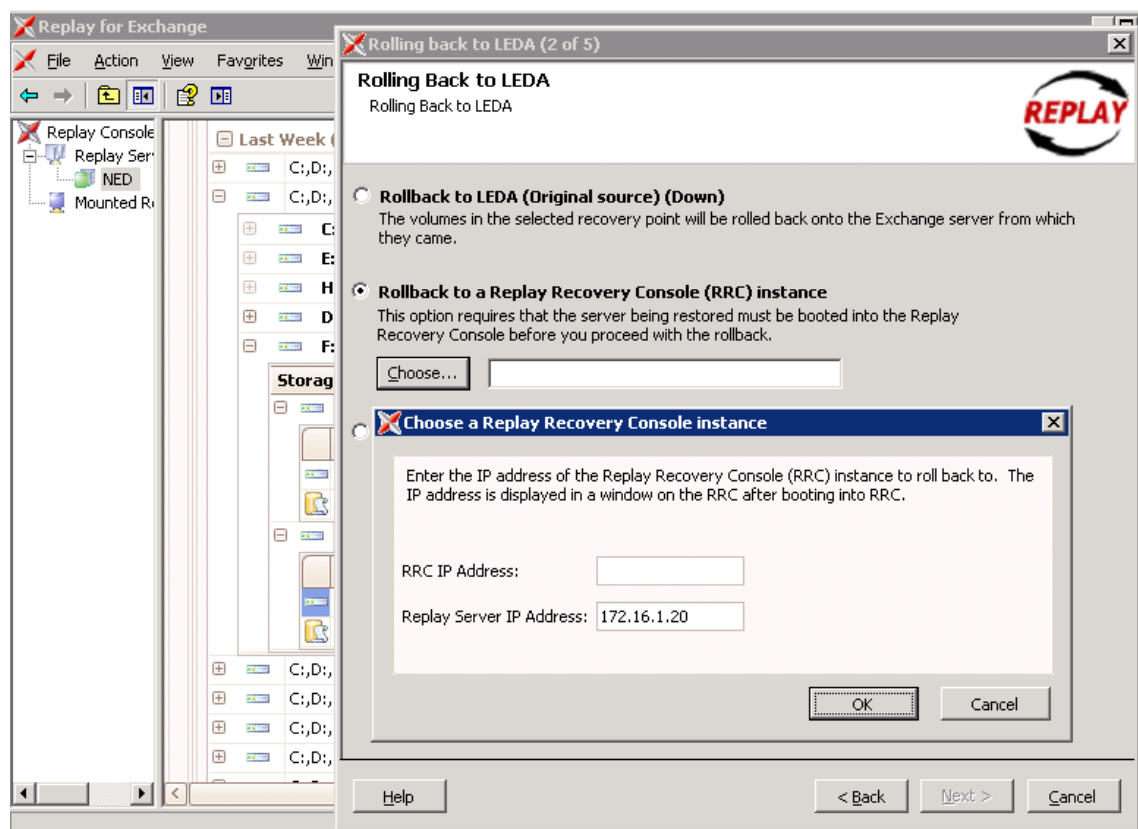


Figure 3. Restoring to a Recovery Server

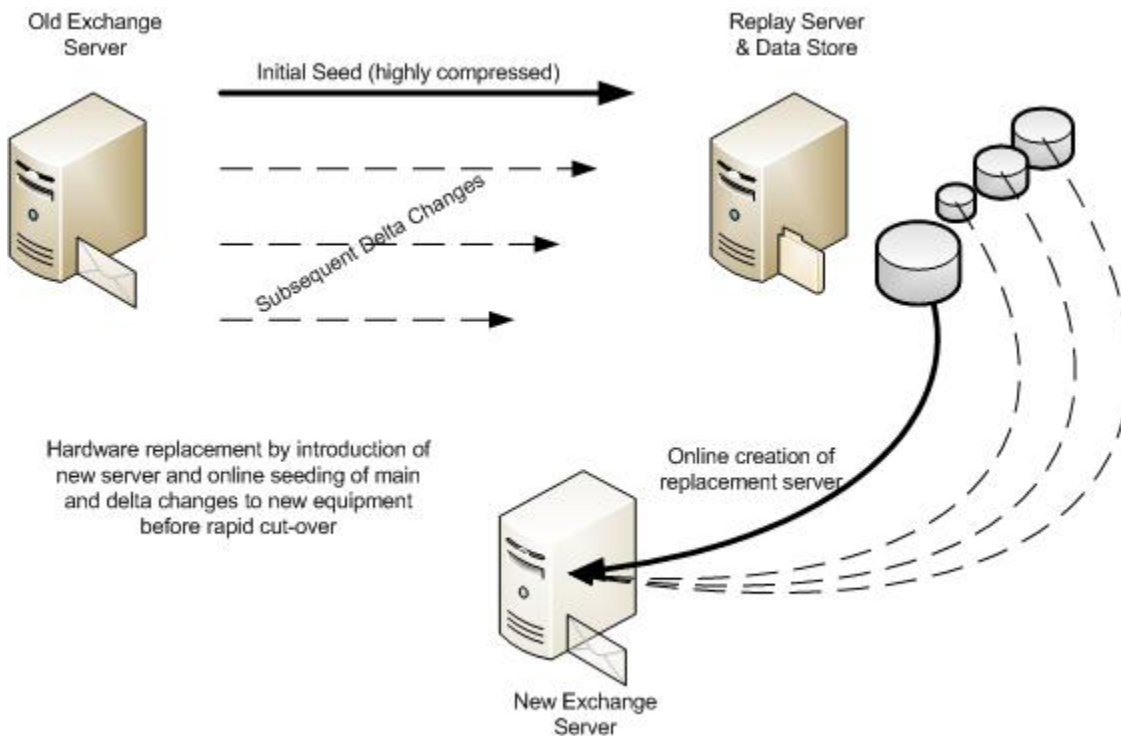
Because the restoration is done at the block level rather than on a file-by-file basis, the recovery is very much quicker and more reliable. While there is no substitute for comprehensive and diligent change control, Replay™ gives you the confidence that the exact configuration of the old has been faithfully reproduced onto the new. This kind of recovery scenario is also useful in situations where testing of upgrades and hardware replacement is necessary. Most of the time an upgrade project will involve the clean installation of a server, application and sample set of data onto a server and then the practicing of the upgrade process onto this clean system. We all know that servers are likely to accumulate little changes and undocumented patches or configuration changes in their lifetime, so the testing of a clean configuration is not necessarily going to yield accurate results when the upgrade is tested. Using Replay™ to apply a complete copy of the old server onto a new server, physical or virtual, gives the confidence that the upgrade procedure and its results can be checked against what is effectively live information.

One of the final items that typical Exchange backup solutions have trouble with is the need to ensure a domain controller is running and available. But what happens if

you are in a lab or some other disconnected environment such as an audit or discovery situation? Replay™ does not require the presence of a domain controller to recover information; it can extract information in a standalone format.

In summary, we can see that by using AppAssure Replay™ you may dispense with your backup application, bare metal recovery solution, test recovery hardware, and a significant amount of worry. Replay™ will save significant amounts of disk and tape space by only recording changed blocks rather than changed files. If you wish to use a clustered environment to provide comprehensive High Availability, then Replay™ will work with that solution to provide the continuous data protection. Replay™ will verify and mount any information store that it is protecting onto the replay server to allow you to extract any Microsoft Exchange item or folder without interrupting the performance of the live Exchange server.

A single Replay™ deployment gives you a significant competitive advantage over the pressures of your business by keeping the backup, recovery, and testing solution down to a single, easy-to-configure integrated application.



**Figure 4. Rebuilding a New Server with Old Server Still Online**