Data Management Forum

ILM Best Practices in Data Recovery for Email

A DMF ILMI Best Practices White Paper

October 1, 2004

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Author:
Edgar St.Pierre, Consulting Engineer, EMC[2]
Co-Chair DMF ILMI Technical Liaison Group
Co-Chair SNIA ILM Technical Work Group

Table of Contents

Table of Figures

1. Introduction

The SNIA definition for ILM: [1]

> The policies, processes, practices, services and tools used to align the business value of information with the most appropriate and cost-effective infrastructure from the time information is created through its final disposition. Information is aligned with business requirements through management policies and service levels associated with applications, metadata and data.

This definition includes how data is to be protected within the IT infrastructure. This white paper is intended to provide very specific examples of data protection from the perspective of "data recoverability" – best practices used to capture a stable set of data that may be restored to the same state at a later point in time.

This paper is one in a series of white papers produced by the SNIA Data Management Forum's ILM Initiative specifically for IT Administrators as the intended audience. It is written with the intent of providing IT Administrators with usable, application-specific, guidelines on deploying ILM solutions using today's technologies.

Each white paper defines best practices for a specific dimension of ILM solutions. In this paper, the topic is data recovery for email. See also the related white papers: Securing Email: Best Practices in Security for ILM [2], Archiving and Tiered Storage for Email [3], and Managing Email for Compliance and Litigation Support. [4]

1.1. Purpose and Scope

1.1.1. Storage Environment

The best practices defined in this paper mostly ignore the storage infrastructure beneath the email server, except to specify what functionality is required from the infrastructure for different solutions. Hence, best practices may be applicable to DAS, SAN, and NAS[1] [5] storage configurations in some cases, or only one in other cases.

1.1.2. Email

This paper provides data recovery best practices that may be applied to several different email server implementations. Specific application notes define the applicability of the following email servers to the data recovery practices described herein:

- Microsoft Exchange 5.5, 2000, and 2003 – includes some practices (such as the use of VSS) that are specific to only Exchange 2003 [6] [7]
- Lotus Domino Server versions 5, 6, and 6.5 [8]
- Unix-based SMTP mail servers

1.2. The Role of ILM: Best Practices at a Glance

Data recoverability is an important aspect of an overall ILM strategy within the data center. This paper should provide two critical ILM perspectives on data recoverability: tiers of service and layers of protection.

Tiers of service: data recovery best practices can be organized in tiers of service with the best performing, most flexible, most complete services at one end of the spectrum – usually with the highest costs. While best practices at the other end of the spectrum will

[1] DAS, SAN, NAS = Direct Attached Storage, Storage Area Network, and Network Attached Storage.

offer lesser capabilities along with lower costs. Aligning the appropriate data recovery services to an email application is dependent on determining the value of that information and the acceptable risk associated with its protection.

One way to approach this is to consider the tiers of service available in the industry today as the tiers of service from which to select, and then deploy needed services on a case-by-case basis. In a larger, more sophisticated IT environment, multiple levels of service for data recovery may be deployed and available within an enterprise, and the IT customers (the lines of business) choose data recovery services commensurate with the value of email to that line of business.

Layers of protection: Operational Recovery and Disaster Recovery, which are defined in Section 1.3, are separate practices to be planned for each application. As a best practice, however, these practices also need to be complementary and coordinated. If the local Operational Recovery practice utilizes disk-based protection, then the site Disaster Recovery practice should leverage those stable copies of data as their source. Likewise, the Disaster Recovery practice should create an extra layer of protection for beyond the window of time protected by local Operational Recovery.

1.3. Defining Data Protection and Data Recoverability

Data protection encompasses aspects of infrastructure availability and recoverability. "Data Availability" addresses the degree to which an application should be able to endure hardware failures yet continue to operate. This touches on configurations for RAID and clustering which are beyond the scope of this paper.

"Data Recoverability" assumes that something physical or logical has gone wrong, and requires that an earlier, stable, copy of the application's data be restored. This may be as a result of hardware or logical failures. Logical failures include software errors, user errors, virus attacks, and more.

Protecting the data in a email application requires that two separate areas of data recoverability be addressed:

- **Local Operational Recovery**: How do I want to recover in the local data center from logical errors and corruptions?
- **Site Disaster Recovery:** How do I want to recover from a site disaster?

These two areas are better defined below. These are also often associated with the concept of "business continuity" planning. This white paper should help in creating such a plan for each business, but it is not intended to replace that planning process. Instead, it is intended to provide very specific recommendations around the data recovery aspects of such a plan.

1.3.1. Local Operational Recovery

Operational Recovery (abbreviated as "OR") is responsible for the vast majority of data recoveries performed in a data center. OR is used to correct "operational" problems such as a corrupt database, user error or hardware failure in the local data center. Essentially, it's everything other than a site disaster.

Businesses are increasingly reliant on email to conduct day-to-day activities. When "email is down", but the rest of the business is "up", then productivity across the business may be severely impacted. Due to this impact, OR may have much more stringent time objectives than those for other applications.

1.3.2. Site Disaster Recovery

The practice of Disaster Recovery (DR) planning is far more complex than can be addressed in this paper. Therefore, in order to provide useful and meaningful best practices, this paper focuses only on aspects of DR data recovery practices and their implications.

1.4. Defining Requirements – How to Determine What is Needed

1.4.1. Data Recovery Performance

These two attributes define data recovery performance requirements:

- **Recovery Point Objective (RPO):** The maximum desired time period prior to a failure or disaster during which changes to data may be lost as a worst case consequence of recovery.[2] Data changes preceding the failure or disaster by at least this time period are preserved by recovery. Zero is a valid value and is equivalent to a "zero data loss" requirement.
- **Recovery Time Objective (RTO):** The maximum[3] desired time period required to bring one or more applications and associated data back to a correct operational state.

Different methods of data protection provide different levels of performance for RPO and RTO. Likewise, each method has associated impacts or costs to consider. Never select a solution based simply on RTO and RPO – examine the whole picture.

1.4.2. Multiple Recoverable Images

It is critical to have multiple recoverable images from which to choose for a complete Operational Recovery solution. As an example, consider the recovery of an email server undergoing a security attack that affects its system files and databases. This recovery requires that the administrator identify the time of the attack, and then select the last OR image taken before that time. The more fine-grained the RPO, then the less time that may have transpired between the last valid recoverable copy and the attack.

Studies[4] have shown that most Operational Recoveries take place within 48 hours of a failure event (in our example, a security attack). The best RPO and RTO capabilities for OR are generally provided by disk-based protection. But current best practices generally limit disk-based protection to fairly short retention and rotation periods measured from hours to days.

If a security attack goes unnoticed, and occurs at the beginning of a long weekend, then it may require additional layers of protection to provide a recoverable image. The first layer of OR may be disk-based in order to provide the best RPO and RTO, but then DR tapes may be used for OR beyond the window of available recoveries from the disk-based replicas. This requires a seamless match of OR and DR planning in order to avoid inadvertent lapses of protection.

---

[2] For email servers with databases, such as Microsoft Exchange and Domino Notes, the worst case recovery assumes that the transaction logs are not available for rolling forward due to corruption or failure. So the entire window of data at risk is the span of time between database and/or logs being copied.

[3] The maximum time for email servers requires time to recover the databases, plus time to playback transaction logs from an entire RPO window.

[4] As an example, a January 2004 Enterprise Storage Group report [9] showed that 56% of all restores occurred within the first 48 hours.

1.4.3. Restore Granularity

What kind of data needs to be restored? If there is a requirement to restore fine-grained objects, such as mail messages, then this can impact the method used to capture data for recovery. Unfortunately, backup methods using fine-grained data capture APIs generally do not provide robust RPO/RTO performance.

At the other end of the spectrum, replica-based strategies for data recovery do provide aggressive solutions for RPO/RTO, but capture data at the volume level. This is not a good match for fine-grained restores since all email (and non-email) files on that volume are restored along with the desired email file(s).

In the case where both fine-grained restores and aggressive RPO/RTO capabilities are desired, then there are still a couple of options:

- Use replicas in combination with email readers: Third party products that read replicas of email databases are able to browse, copy, and paste desired objects from replicas back to production data.
- Use continuous data protection (CDP): Some implementations of this technology, while still evolving, offer the ability to provide the same RPO/RTO times of replicas in conjunction with the ability to mount images from any point in time in order to browse, copy, and paste desired objects.

1.4.4. Cost

The last, but certainly not the least important, requirement to consider for data recovery is cost. To this end, there are two possibilities: a budget is assigned within which the best possible data recovery strategy is deployed; or, an appropriate data recovery strategy commensurate with the value of email to the business is defined and deployed. While the former may be true most often today, the premise of ILM is that it should be based on the latter.

As a first step, consider this: when email is not available, what is the level of productivity in the business? What is the hourly run rate for the business's payroll? If the business has a 1$M/hour payroll, and the loss of email impacts productivity by 30%, then 1 hour of downtime will cost the business $300K.

If a $500K investment in improved data recovery methods pays for itself by reducing the recovery time from 2 hours to 20 minutes after the first email downtime event – would that be considered a justifiable cost?

This can also be viewed in terms of risk tolerance. For a web-based customer service business unit, there is probably a low tolerance for risk. In this case, it may be a company image issue which may raise the risk factor even higher than the financial cost of downtime. Such an email server probably already has a high availability configuration. If the database is corrupted and needs to be recovered, then it should also have a high performance operational recovery method.

While these examples illustrate the ROI for data recoverability solutions based on the cost of downtime to the business, it is also possible to realize a return on investment based on other uses of the same infrastructure. For example, replica-based solutions may be leveraged to create copies of important databases on a regular basis that are then repurposed for use in data mining or business intelligence applications. The

efficiency that can be realized with repurposing for applications such as business intelligence may justify much, or all, of the required investment for data recoverability.

Finally, what if there are multiple lines of business supported by the enterprise? Shouldn't each be able to assess its own impact of email down time? Thus, there may be a need to deploy multiple methods of OR and DR within an enterprise for the same application. Each method would represent a different level, or tier, of service provided by the IT Organization in support of data recoverability. Corporate-wide data classification efforts are often useful in ranking the value of data to the lines of business, and associating those rankings with available and desired tiers of service for storage, protection, and security of data.

2. Methods of Protection for Email

There are many combinations of data sources, OR solutions, and DR solutions that may be combined to provide overall data protection for email. In order to organize these combinations, consider the relationships in Figure 1.

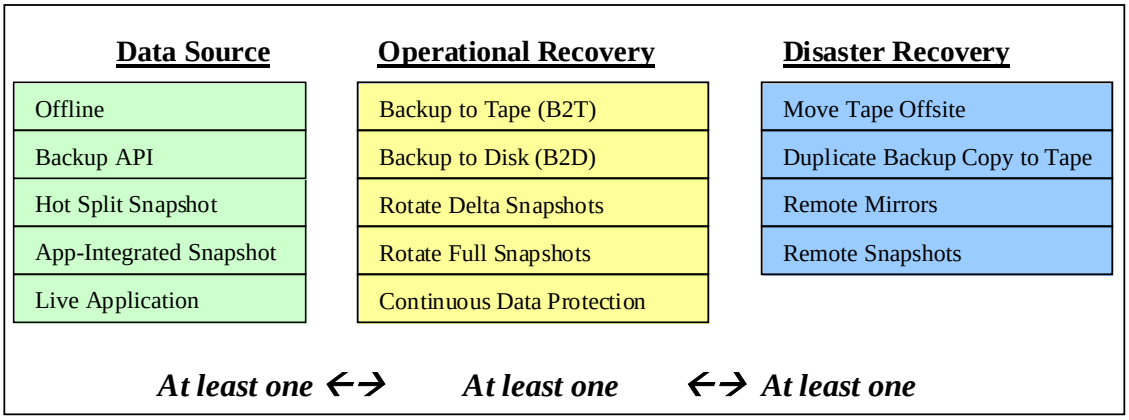

Figure 1: Choosing a Solution

Not all combinations of Source/OR/DR solutions are possible, nor do all combinations make sense. However, as a best practice, *at least* one method of Operational Recovery and one method of Disaster Recovery are recommended. The ideal is to use a single Data Source to do both.

This paper organizes its best practices around Operational Recovery since these sometimes enable the Disaster Recovery solutions. The OR best practices are defined in Section 3 and are intended to bridge the Source Data and Disaster Recovery solutions. The two sections below outline Source Data considerations and Disaster Recovery options. The Section 3 OR best practices each identify source data dependencies, and which Disaster Recovery solutions may be combined with each OR method. Each OR best practice also provides guidance on the type of RPO/RTO performance to expect for that method, the relative cost, as well as the granularity of restore that is possible with each.

2.1. Source Data Considerations

The most important first step of any data recovery best practice is ensuring that there is a stable copy of data available from which to create a recoverable copy. The available options for an email server are outlined below. Some relevant questions to be asked before selecting a source data method include:

- Can the email server be taken offline in order to create backup copies? If so, then the “Offline” method below is certainly the most cost efficient.
- Can the email server endure a performance impact, or is it expected to be operating at I/O levels which can not endure any performance impact whatsoever? The degree to which this is true may influence the replica solutions to use and whether backup copies need to be processed from an alternate server, or can be processed directly from the production email server.

Sources of data for email:

- **Offline:** This is the easiest to implement, but not very amenable to today's 24x7 world. This is valid for any version of email server.
- **Backup API:** A backup API to the email server can be used to perform online full and incremental backups of an email server's data stores. This API is used to produce a stable image of the email database and log files. Other email server APIs, such as object-level APIs, can also be used to capture data at a finer level of granularity during a backup process.
- **Hot Split Snapshot:** This requires mirroring or snapshot capabilities of some kind in a storage array, and must not to be confused with Application-integrated snapshot (see next bullet). This method has multiple steps to it which may require iteration:
    1. Store the email database and logs on synchronous mirrors, or on storage from which snapshots can be created.
    2. "Break" the mirrors, or create a snapshot, in order to create a "frozen" image of the database and logs. These frozen images are essentially "hard crash" copies, or versions of the database and logs that might have been available if someone had suddenly disconnected power from the email server. These are not as easily recoverable for some email servers.
    3. Validate the frozen images using a server-specific utility. This is optional, depending on the specific architecture and capabilities of the email server.
    4. If a frozen image is not usable, then repeat from step 1. If it is usable, then it is a valid stable copy of email.

  This practice is often used to create stable images from which to backup an email server, but can also be used to create multiple online recoverable replicas as well. Hot split snapshot is generally valid for any version of email server, but is not the best option for snapshots if application-integrated snapshots are available.

  Note that Hot Split Snapshots may be either delta snapshots or full clone copies of the email databases and logs. Hot Split Snapshots may be mounted back onto the production server for some configurations in order to create a backup copy; however, when used in this way, then any backup processing may have a performance impact on the host. Delta snapshots may also have a significant performance impact on disk access by the email server if used for backup processing. Hot Split Snapshots may also be mounted onto alternate hosts for backup processing. See Sections 3.3 and 3.4 for more information on clone and delta snapshot cost/performance considerations.
- **Application-Integrated Snapshot**: This solution eliminates the uncertainties associated with Hot Split Snapshots described above. It coordinates the email server application with host- or array-based snapshots to create stable, usable, replicas of the databases and logs.

  Application-Integrated Snapshots may be either delta snapshots or full clone copies of the email databases and logs. They have the same cost/performance considerations as described previously for Hot Split Snapshots.
- **Live Application:** There are two forms of live application protection.

- The first is object level replication, whereby each email object (message, attachment, calendar appointment, etc) is automatically replicated to another secondary server or storage location.
- The second form is using third party tools to copy every disk write from the email server to a secondary server.

Both of these are considered a host-based form of Asynchronous Remote Replication[5] for Disaster Recovery.

In general, these forms of protection may have some impact on email server performance[6], and the choice of server platform and its scalability should be adjusted accordingly. It has the benefit of having an alternate host available for backup processing, if necessary, without impacting the production email server.

2.2. Disaster Recovery Options

The Disaster Recovery options defined below describe data protection methods that can be used to make copies of an email database available offsite. This offsite copy would be used to bring an email server up in the event of a site disaster at the local data center.

2.2.1. Move Backup Tapes Offsite

If the OR practice is to perform a backup to tape (B2T), then one option for DR is to simply move all of the backup tapes to an offsite location after each backup. This may be the least expensive to implement in some cases, but it has a severe impact on the ability to perform Operational Recoveries.

In some cases, this may be acceptable. For example, if the desired RPO for an application is 48 hours, and backups are performed nightly, held for 24 hours, then sent offsite, then the 48 hour RPO can be satisfied. On the down side, any OR that requires an offsite tape will require a significant overhead to retrieve the offsite tape.

The use of remote vaulting to create backup tapes directly over a network is also an alternative, but may require significant network bandwidth to handle all of the data movement.

2.2.2. Duplicate Backup Copies to Offsite

Duplicating backup copies to an offsite location may be duplicating a backup copy from a B2T or B2D operation. These duplicates are then stored offsite. The destination media may also be tape or disk.

Remote vaulting, which is the writing of backup copies over a WAN to a DR vault is most efficient when used to duplicate, rather than originate, backup copies since it provides a utilization of network and backup server bandwidth that does not impact the production servers. This practice also eliminates tape transportation costs, but requires adequate network bandwidth to be able to duplicate the backup copies. Although tape storage in a DR vault is the traditional method associated with remote vaulting, it is also possible to duplicate backup copies to a disk located in a remote DR vault as well.

[5] See Section 2.2 on Page 10 for a description of host-based Asynchronous Remote Mirrors.
[6] This will vary widely with the particular solution; your vendor should provide guidance on level of impact.

The method of duplication is not covered by this white paper. In general, backup applications have built-in duplication capabilities. Using array- or host-based replication or mirroring solutions are also quite feasible.

2.2.3. Remote Mirrors

This may be a host-based or storage array-based operation, as distinguished below. In general, these can also be used as part of a "high availability" solution in addition to use for Disaster Recovery.

- **Host-based Remote Mirrors**: These solutions may be either integrated with an application, or inserted into the data path. In the application-integrated case, each "object" written to the data store is also copied and forwarded to a secondary server. This secondary server also writes the object to its copy of the application data store. This method has the advantage of providing object-level granularity for Disaster Recovery operations.

  Third party solutions can also be used to copy each disk write operation (block level or file-level) to a secondary server, which then writes the same block (or file) to its file system. This method has the advantage of being application-agnostic, but its drawback is that it may only provide crash-consistent images of the source data in some cases.

  Host-based mirrors are limited to DR solutions since they do not provide multiple recoverable copies of data from different points in time, and each solution also propagates any corruption or application errors immediately from the production storage to secondary storage.

  Host-based mirrors are also generally asynchronous in nature because of overhead and latency associated with using TCP for transport. This means there will usually be a small window of data that is vulnerable to being lost.

  The use of host-based remote mirrors must also provide consistency across the application. It is important to maintain order of delivery across multiple mirrored volumes across multiple TCP/IP connections for multiple volumes. Email servers typically store transaction logs on a different volume from the database. A best practice method to do this is to utilize host based mirrors for log mirroring only, then use point-in-time replicas or backup copies for database copies. This ensures that the log files never fall out of synchronization with the database files.

- **Array-based Remote Mirrors**: Like host-based mirroring, these solutions also may be asynchronous in nature (particularly for longer distances), or they may be synchronous. Synchronous remote mirrors ensure that every disk write is completed to both the primary and secondary storage device before acknowledging the write to the application. Due to latency issues, synchronous can only be done over shorter distances.

  Array-based mirrors are limited to DR solutions since they do not provide multiple recoverable copies of data from different points in time, and each solution also propagates any corruption or application errors immediately from the production storage to secondary storage.

  Array-based remote mirrors provide crash-consistent images for DR. Even this

minimum level of consistency, however, requires consistency across the application. Hence, it is important to maintain order of delivery across multiple mirrored disks, and potentially multiple network channels, in order to ensure that an application is able to recover from the remote disk images.

2.2.4. Remote Snapshots

Remote snapshots are storage array-based solutions in which one or more recoverable, stable, copies of data are maintained. These copies may be full copies such as clones, or partial copies such as delta snapshots.

Remote snapshots are tertiary images. That is, if the production data is the primary image, then there needs to be a stable secondary image from which a remote snapshot is produced as a tertiary image. This can be done in at least three different configurations:

- **Copy of a remote mirror**: This can be done by quiescing a remote mirror and creating a clone or delta snapshot from that. As of the writing of this paper, the state-of-the-art for these particular solutions are generally custom-scripted for each environment.
- **Copy of local snapshot**: This can be done by producing a stable clone or delta snapshot locally, then copying that disk image to a remote array.
- **Propagate local snapshots to remote site:** This can be done by producing a stable snapshot locally, then propagating the changes in that snapshot to the remote site.

An additional value of remote snapshots is that they provide multiple copies from which a DR can be performed.

3. Operational Recovery Best Practices and Tiers of Service

3.1. Recovery from a Backup Copy on Tape (OR Method 1)

3.1.1. Data Source

In this OR method, the recovery image is produced using B2T. The source data for the B2T may be from an offline email server, a backup API, a hot split snapshot, or application-integrated snapshot (as described in Section 2.1.).

3.1.2. Operational Recovery Description

In order to capture data for OR, this data protection method uses a traditional backup approach. Depending on the capabilities of the backup program, the data copy may be: A) locally sent to tape that is attached to the email server; B) sent over a LAN to a backup server which writes it to tape; C) written to tape that is shared via a SAN; or, D) a snapshot that is read from and written to tape by a third party device such as a backup server. These are illustrated in Figure 2.

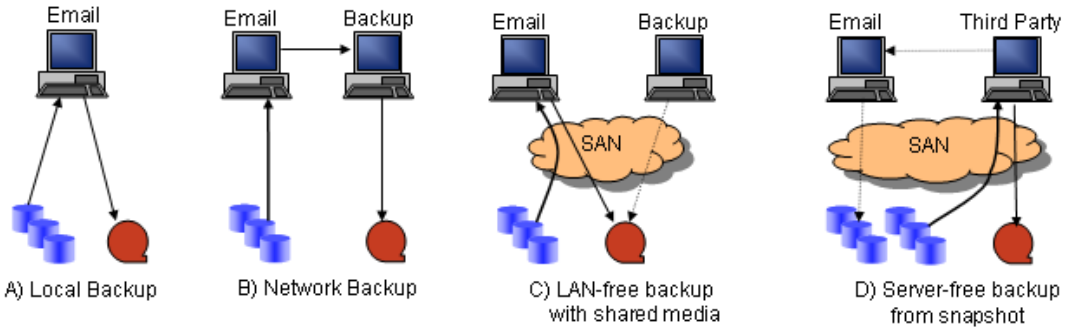

Figure 2: Backup to Tape (B2T)

Note that the variations of backup methods for B2T will impact backup windows, server impact, and other email server criteria during the creation of backup copy. This paper, however, is focused more on recovery criteria. Therefore, no further effort is made to distinguish the benefits and drawbacks of the various B2T methods, and focus is turned instead on recovery criteria.

3.1.3. RPO/RTO

B2T is a fairly time-expensive operation for both data capture and recovery. As such, it offers the poorest performance relative to RPO and RTO, and RTO duration will also vary considerably depending on the size of the email database(s), accessibility of the tapes, and the tape drive technology in use.

B2T is most useful for email applications that allow an RPO of 24 hours or more. One way to reduce RPO using B2T is to only backup the transaction logs (if applicable) on a more frequent basis, or use one of the replica-based OR solutions recommended later. The number of available recovery images will depend on corporate governance for tape retention and email compliance.

Projected RTO: The recovery process is usually several hours in duration – assuming the required tapes are already onsite. The number of hours is dependent on the size of the email database(s), and the tape drive technology in use.

3.1.4. Restore Granularity

Restore granularity is dependent on selected backup granularity and/or use of a third part tool to browse restored email databases.

3.1.5. Cost

B2T can be a very cost-efficient method of protecting data, depending on several factors:

- The number of backup copies retained simultaneously.
- Tape stream parallelism: If faster backups and restores are desired, then multiple parallel data streams need to be used. This increases the overall cost, however, due to additional tape drive acquisition and maintenance costs. From a strictly OR perspective, there should not be an increased per-copy cost for increased parallelism over time. DR costs will increase per-copy, however, since this likely increases the number of tapes sent offsite for each complete copy.
- Tape rotation and retention periods: Tapes can be used several times. The most efficient number of safe uses should be specified by each tape manufacturer.
- Tape technology: different tape technologies offer a variety of density/performance/cost tradeoffs.

3.1.6. Disaster Recovery

There are several DR best practice options available in conjunction with the B2T method of protection for Operational Recovery:

- Move the resulting backup tape offsite, as in Section 2.2.1.
- Duplicate the backup tape and locate one offsite, as in Section 2.2.2.
- Remote mirrors may be used in parallel, as in Section 2.2.3.

3.2. Recovery from a Backup Copy on Disk (OR Method 2)

3.2.1. Data Source

In this OR method, the recovery image is produced using backup to disk (B2D). The source data for B2D may be from an offline email server, a backup API, a hot split snapshot, or an application-integrated snapshot (as described in Section 2.1.).

3.2.2. Operational Recovery Description

In order to capture data for OR, this data protection method uses a traditional backup approach. B2D creates backup copies on disk using the same methods as B2T. Because the backup copy resides on disk, however, the RTO capabilities are typically better than for backup copies sent to tape.

B2D is particularly useful as a short term staging area for backup copies since these images can be used for OR, then Duplications of the backup copies can be made to tape for DR. This process is sometimes referred to as “B2D2T” or “disk-disk-tape” backup.

B2D also encompasses technology referred to as “virtual tape”. The major distinction between virtual tape and other methods of B2D is that a virtual tape solution will emulate tape drives in behavior, and can automatically move backup copies to tape instead of using the backup server to duplicate the backup copy.

3.2.3. RPO/RTO

B2D is recommended for email applications that allow an RPO of 24 hours or more, just like B2T. This is due to the time required to create the backup copy, and the potential impact on the email server. One way to reduce RPO using B2D is to only backup the transaction logs (if applicable) on a more frequent basis, or use one of the replica-based OR solutions recommended later. The number of available recovery images will depend on the disk capacity of the backup server.

Projected RTO: The recovery process for B2D is typically shorter than B2T and can cut the data transfer phase of an email restore process by ¾ or more. Note, however, that the longer RPO will require the same time to replay the transaction log, so the overall effort may still require a significant amount of time.

3.2.4. Restore Granularity

Restore granularity is strictly dependent on selected backup granularity and/or use of a third part tool to browse restored email databases.

3.2.5. Cost

The cost factor for B2D is approximately measured in increased disk cost, which will be slightly more than 1 times the size of the production data for each “full backup” B2D copy retained simultaneously, plus enough disk for simultaneously retained incremental backups.

3.2.6. Disaster Recovery

The DR best practices recommended in conjunction with the B2D method of protection for Operational Recovery:

- Duplicate the B2D backup copy to offsite, as in Section 2.2.2.
- Remote mirrors may be used in parallel, as in Section 2.2.3.

3.3. Recovery from Rotation of Clones (OR Method 3)

3.3.1. Data Source

The source data for a Rotation of Clones is typically performed using a hot split or application-integrated snapshot (as described in Section 2.1.). It can also be done with offline email servers, but this is less common.

3.3.2. Operational Recovery Description

In order to capture data for OR, this data protection method uses software-based or array-based snapshots that create exact, *complete*, replicas of the source volumes to other volumes. An exact copy which occupies as much storage as the source of the copy is referred to as a “clone”. These clones are alternated in and out of a synchronous mirror role. This is illustrated in Figure 3.

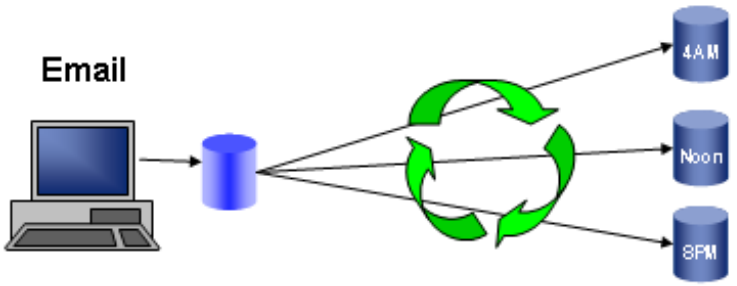

Figure 3: Rotation of Clones

The benefits of this approach vary depending on the specific host- or array-based snapshot software. Here are the potential benefits:

- **Minimal replication time**: the time required to "snap" a clone is negligible if the clone is already active. It is considerably longer if a new clone needs to be started. Impact to the email server is negligible.
- **Multiple clones**: there may be a rotation of multiple clones taken over the course of 8, 24, or 72 hours – or whatever window of protection is desired. A good starting point is the rotation of a couple of clones over a 24 hour period. This reduces the RPO, or risk exposure, and provides 2 recovery points that may be used to recover from logical errors.
- **Potential "instant restore"**: some arrays and software snapshots allow for the snapshot to be restored "on the fly" from the snapshot to the production disk. So the production disk can be mounted immediately after the restore is started. In this case, the RTO is effectively reduced to the setup time required to select the snapshot and start email.

Note that when instant restore is available from a clone, be certain that it is a "protected" instant restore that does not allow writes to be forwarded from the source disk to the clone. This "protection" prevents the corruption of the clone copy.

Cost vs. Performance: This solution is best practice for heavily loaded email servers since the synchronous writes to multiple disks impose minimal overhead. Delta snapshots, as described in Section 3.4, may have an impact on the performance of the server access to the disk in some configurations. This will vary with email server and array vendor. Clones also have the advantage of being usable (such as for running a backup) without having any impact whatsoever on the performance of the email server. The selection of clones vs. delta snapshots usually comes down to a cost vs. performance tradeoff.

3.3.3. RPO/RTO

A rotation of full snapshots offers excellent RPO and RTO performance, with minimal performance impact. The improved RPO is a result of the more frequent snapshots, and since each snapshot restarts the server's transaction log, it also reduces the size of the log that needs to be replayed in a full restore – which therefore reduces the overall RTO.

Recommended RPO: 12 hour clones up to 24 hours old (2 clones total)

Overall projected RTO: 20-40 minutes, depending on the size of the log

3.3.4. Restore Granularity

Clones are volume-based copies, and an entire database, or group of databases, must be on the same volume. (Logs should be on a separate volume.) Hence, the restore granularity of a clone is generally limited to the (group of) databases on a volume. One way around this limitation is to mount the replica to an alternate location and manually copy a single database, mailbox, or message, from the mounted replica. This may sometimes require the use of third party tools.

3.3.5. Cost

The cost of clones is fairly straightforward to calculate: each clone is another copy of the source data. So simply multiply the number of simultaneous clones by the disk storage acquisition cost (accounting for RAID format).

One additional note: there is no recurring cost as there is with tape. However, storage allocations should account for potential growth of the application.

3.3.6. Disaster Recovery

Clones themselves are not a method for site disaster recovery. However, there are scenarios where they may assist with DR:

- Clones may be used as the source of data for a DR copy. This may be a **B2T**, as in Section 2.2.1.
- Clones may be used as the source of data for a DR **copy of a local snapshot** as described in Section 2.2.4.
- If using synchronous remote mirrors, then clones of the remote mirrors may be used to create multiple stable images available for site DR. Remote synchronous mirrors only provide assurance of crash-consistent replicas. A minimum of three snapshots is required to provide assurance of at least one known stable image at all times. (One known stable, one live as a third mirror off the remote replica, and the last one mounted and testing to verify stability.) Today, this is implemented as a custom-scripted solution.

3.4. Recovery from Rotation of Delta Snapshots (OR Method 4)

3.4.1. Data Source

The source data for a Rotation of Delta Snapshots is typically either a hot split or application-integrated snapshot algorithm (as described in Section 2.1). It can also be done with offline email servers, but this is less common.

3.4.2. Operational Recovery Description

In order to capture data for OR, this data protection method uses host-based or array-based snapshots that only require enough additional storage to capture changes to the disk. This is illustrated in Figure 4.

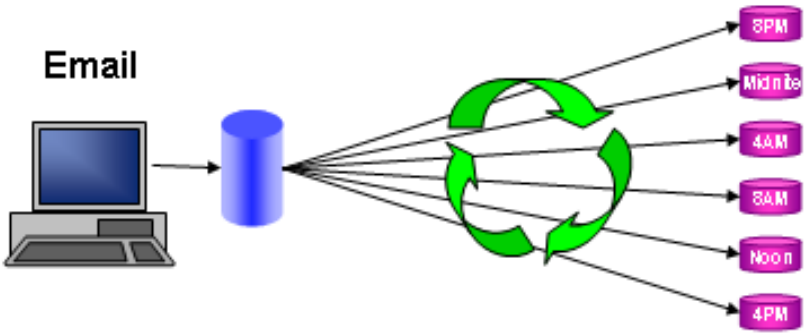

Figure 4: Rotation of Delta Snapshots

The benefits of this approach vary depending on the specific host- or array-based snapshot software. Here are the potential benefits:

- **Minimal replication time**: the time required to setup the delta table is relatively short, and the impact to the email server is negligible.
- **Minimal additional disk space requirements**: since the delta replicas only save changed blocks or tracks, then the amount of additional disk required is minimal.
- **Multiple snapshots**: there may be a rotation of multiple snapshots taken over the course of 8, 24, 72 hours, or whatever is required. A good starting point is the rotation of 6 snapshots over a 24 hour period. This reduces the RPO, or risk exposure to 4 hours. Best practices will also vary with the distribution of load over the course of the day.
- **Potential "instant restore"**: some arrays and software snapshots allow for the delta snapshot to be restored "on the fly" from the delta snapshot to the production disk. So the production disk can be mounted immediately after the restore is started. In this case, the RTO is effectively reduced to the setup time required to select the snapshot and start email. Like rotation of clones, be certain that the instant restore is protected from corrupting the delta snapshot.

Caveats: Some snapshot solutions may cause a performance impact depending on a combination of the particular snapshot implementation, storage configuration, and email volume. Also, because delta snapshots share unchanged disk tracks with the production data, running a backup copy from a snapshot may cause disk thrashing and impact email server performance. This should only be considered an option during low usage periods for the email server. It is also important to note that delta snapshots only provide recovery from logical corruption of the data. If the production volumes are physically damaged, then a recovery from a delta snapshot is not possible.

Cost vs. Performance: This solution is easiest to configure when used with lightly loaded email servers since delta snapshots may have an impact on storage performance in some configurations – but this is not universally true since many solutions will work for heavy email loads if properly configured. This will vary with email server and array vendors. Consult your vendor for recommendations on configuration and load criteria. The selection of clones vs. delta snapshots usually comes down to a cost vs. performance tradeoff.

3.4.3. RPO/RTO

A rotation of delta snapshots offers excellent RPO and RTO performance, with minimal disk utilization. The benefit of running more frequent snapshots is that there is a shorter email recovery log to be run during recovery. Log file recovery generally constitutes the lengthiest recovery step in the snapshot-based recovery process, so a short recovery log can provide significant benefit to recovery time.

Recommended RPO: 4 hour snapshots up to 24 hours old (6 snapshots total)

Overall projected RTO: 15-20 minutes, depending on the size of log

3.4.4. Restore Granularity

Delta snapshots are volume-based copies, and an entire database, or group of databases, must be on the same volume. (Logs should be on a separate volume.) Hence, the restore granularity of a delta snapshot is generally limited to the (group of) databases on a volume. One way around this limitation is to mount the replica to an alternate location and manually copy a single database, mailbox, or message, from the mounted replica. This may sometimes require the use of third party tools.

3.4.5. Cost

The cost of delta snapshots is based on increased disk allocation costs. The calculation for how much additional disk is required will vary with the size of production data and the change frequency to that production data.

If the array or storage software provider defines such a calculation, then use it. If not, then a conservative rule of thumb is that an email database and logs won't change more than 5 to 10% per day. So allocating 10% of the production storage size times the number of days that snapshots are retained should suffice. So any number of snapshots per day for two days of a 60 GB email database could cost up to 12 GB of disk space.

3.4.6. Disaster Recovery

Snapshots alone are not a method for site disaster recovery. However, there are scenarios where snapshots may assist with DR:

- Delta snapshots may be used as the source of data for a DR copy. This may be a **B2T**, as in Section 2.2.1. **Caveat**: delta snapshots essentially share unchanged disk tracks or blocks with the production data. Hence, a DR copy from the delta snapshot will perform significant reads from production disk and affect performance of the email server. This is best done during low usage periods.
- Delta Snapshots may be used as the source of data for a DR **copy of a local snapshot** as described in Section 2.2.4. **Caveat**: delta snapshots share unchanged disk tracks or blocks with the production data. Hence, a DR copy from the delta snapshot will perform significant reads from production disk and affect performance of the email server. This is best done during low use periods.
- If using synchronous remote mirrors, then delta snapshots of the remote mirrors may be used to create multiple stable images available for site DR. Remote synchronous mirrors only provide assurance of crash-consistent replicas. A minimum of three snapshots is required to provide assurance of at least one known stable image at all times. (One known stable, one live as a third mirror off the remote replica, and the last one mounted and testing to verify stability.) Today, this is generally a custom-scripted solution.

3.5. Recovery from Continuous Data Protection (OR Method 5)

Implementations for continuous data protection (CDP) vary widely. This section attempts to generalize expectations around the currently available products and technologies. This is an evolving area of new technology for data recoverability.

3.5.1. Data Source

The source data for CDP is a live application, as described in Section 2.1. This is a "third party" style replication of either every disk write, or of every object, to another storage device.

3.5.2. Operational Recovery Description

In order to capture data for OR, CDP captures every disk write, or every object write, made by the email server. CDP, in one sense, is the degenerate case of rotation of delta snapshots since it essentially creates a "snapshot" that can be mounted for each and every transaction sent to the CDP Service. This is illustrated in Figure 5. Note that the "Dual-write Function" may be placed in any of several locations: on the host, in an array, in a SAN fabric switch, or in an appliance within the SAN.

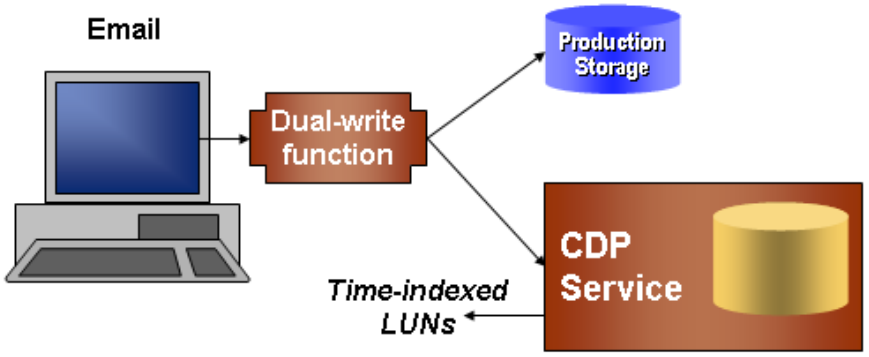

Figure 5: Continuous Data Protection

The strategy with this technology is to capture every write in a non-overwriting form. So even if the same record of a database is updated *n*-times, each of the writes still exists within the CDP Service as being valid for whatever point in time the *n-1* write was made, or the *n-2* write, etc. Each point in time at which a write was made by the CDP-protected email server is then selectable and mountable as a "time-indexed" LUN which is exposed by the CDP Service.

The benefits of this approach vary depending on the specific solution. Some of the benefits, including those associated with different styles of implementation, are:

- **Precision snapshots**: Because every write, or every object, is written simultaneously to a second location in a write-ordered form, it is possible to recover an email server from any point in time. Some consideration needs to be made, however, for whether the exposed LUN is from a "crash-consistent" or "application-consistent" image. Some email servers are more specific as to what is considered a recoverable image. However, every email server and file server needs to be able to recover from crash-consistent images in some manner.

- In another aspect of consistency, it is critical that CDP solutions provide consistency across multiple volumes associated with the email database(s) and log files. A lack of consistency can lead to unrecoverable data for one or more points in time.
- **RPO = 0**: RPO is potentially zero data loss.
- **Short RTO:** In order to recover an email server with a CDP appliance, the desired point-in-time LUN must first be identified, then mounted and verified (either manually or using an automated tool). In a worse case, it may take longer to verify and correct a crash-consistent image. Once verified, then the production data needs to be updated to the state captured in the CDP Service. Performance of this update can vary based on implementation. Some implementations allow the use of the email database in a degraded performance mode while the updates are being applied; some do not. The mounting and browsing of a time-indexed LUN is always possible, which allows copying of smaller objects back to the production data if that is all that is required.
- **Variable retention:** CDP retention algorithms vary and could be as high as several days, weeks, or months, but at significantly coarser granularity. While the precision of the snapshots could be as fine as each write, the retention of each snapshot is not infinite. As time goes on, the older writes that are not required to build a specific point in time view are removed from the CDP Service in order to make room for new writes.

3.5.3. RPO/RTO

Specifics for RPO will vary with crash-consistent vs. application-consistent recovery images, with each offering different versions of what to expect for an RPO = 0. Realistically, both offer solutions that approach RPO = 0 for Operational Recovery.

Specifics for RTO will vary widely based on the particular solutions. Those implementations that can significantly improve RTO time include those that copy the entire point in time image from the CDP Service to production storage while still allowing access to the production storage in the interim.

3.5.4. Restore Granularity

Restore granularity for CDP will range from an entire email server to mail messages.

3.5.5. Cost

The cost for a CDP Service is specific to the particular solution, and will likely have a storage component that will scale at approximately 1.5 to 2 times the size of the protected storage. But again, implementations vary, and environment variables such as change rate and retention period will impact the amount of required storage.

3.5.6. Disaster Recovery

CDP is inherently an OR solution. However:

- CDP can be extended to DR with remote mirroring capabilities added to the CDP Service backing store. This may be inherent to the CDP Service or added as a third party solution. Remote mirroring is described in Section 2.2.3.
- The CDP Service can also be used to create point-in-time images to be used to create backup copies that can be used for DR, as defined in Section 2.2.1.

4. Comparison of Best Practices

4.1. Summary: Tiers of Service for Operational Recovery

The following table summarizes best practices for different Operational Recovery tiers of service available in the industry today. The tiers vary based on cost and performance, but the numbers suggested below are not hard and fast. They can swing above and below estimates for different configurations. For example, it is possible to improve B2T RPO and RTO by increasing the frequency of backups and parallelism of tape drives. But this would significantly increase the cost and application-impact profile and was not considered a best practice.

Explanation of Table headers below:

- **Operational Recovery Type**: As described in Section 3.
- **Typical RPO**: How many hours of email will be lost if the transaction logs for that time *are not* recoverable (worst case condition). Any of these solutions can achieve an RPO of 0 if the transaction logs are available.
- **RTO Estimate**: How many minutes it will take to fully restore the email server data. The worst case for RTO is when transaction logs *are* available[7].
- **Estimated HW Cost**: gross estimate of purchase requirements for each solution. The only recurring cost is for tapes; others are one-time. Each OR Type also has software acquisition and licensing costs associated with it that will vary for each vendor solution.

Note: These are estimates intended to show relative performance between solutions. Assumptions in device performance are listed in the footnotes below. ***Your mileage will vary.*** The best practice is to create your own benchmarks for your environment.

Operational Recovery Type[8]		Typical RPO	RTO Estimate	Estimated HW Cost
Backup copy from Tape	Daily DB & Logs	24 hours	66 -153 minutes[9]	Tapes[10]
	Daily DB, plus Logs every 6 hrs	6 hours	70 -162 minutes[11]	Additional tapes for logs
Backup copy from Disk	Daily DB & Logs	24 hours	62 minutes[12]	Disk: approx 1x data size
	Daily DB, plus Logs every 6 hrs	6 hours	64 minutes	Additional disk for logs
Clone		12 hours	25 minutes	Disk: 1x data size per clone
Delta Snapshot		4 hours	10 minutes	Disk: ~ 0.2x data size
Continuous Data Protection		0 hours	5-30[13] minutes	Disk: varies by change rate, and retention period

Figure 6: Summary of Operational Recovery Tiers of Service

[7] Transaction log "playback" is assumed to be at a linear rate of 45 minutes for 2GB.
[8] Data assumptions: 60GB exchange database plus 2GB/day for log files.
[9] RTO Performance for tape varies significantly, depending on tape location and drive technology. This estimate assumes local tape access and transfer rates of 55 & 10 MB/sec, plus a couple minutes to load and position.
[10] The number of tapes will vary based on size of data, tape rotation schedule, and the tapes re-use policy.
[11] Assumes tape restore of log files are at transfer rates of 20 & 5 MB/sec due to smaller file processing.
[12] RTO Performance varies with specific hardware. This estimate assumes a 65 MB/sec transfer rate.
[13] Recovery time for CDP will vary depending on implementation.

These tiers of service reflect products that are generally available in the industry today. IT organizations may choose to deploy multiple levels of service based on the requirements of the business processes, as adjusted to cost and risk acceptance.

4.2. Summary: Tiers of Service for Disaster Recovery

The following table summarizes best practices for different DR tiers of service available in the industry today. The tiers vary based on cost and performance, and the numbers suggested below are not hard and fast. They can swing above and below estimates for different configurations. Most importantly, these estimates only reflect *data recovery times at the DR site only*. E.g., in the case of removable media, it assumes the media is immediately available and stored on an unlabeled shelf. Overall DR processes and times are much more extensive and address infrastructure, server, data, and personnel recovery activities.

Explanation of Table headers below:

Data DR Recovery Type: As described in Section 2.2.

RPO: How many hours of email will be lost if the transaction logs for that time *are not* recoverable (worst case condition). Note that a DR RPO does not have to be the same as the OR RPO. It should reflect the value of the information commensurate with the risk associated with site disasters. An application's DR RPO ranking relative to other applications will likely be consistent with its ranking relative for OR RPO.

RTO Estimate: How many minutes it will take to fully restore the email server data. The worst case for RTO is when transaction logs *are* available and their "playback" is at a linear rate of 45 minutes for 2GB. The same consistency for OR vs. DR applies to RTO as well. However, RTO must also reflect the overall Disaster Recovery plan with respect to sequencing of applications to be recovered during a site disaster.

Estimated DR Storage Cost: estimate of purchase requirements for each solution. The only recurring cost is for tapes; others are one-time. Each Data DR Method also has software acquisition and licensing costs associated with it that will vary for each vendor solution.

Data DR Recovery Type[14]		RPO	RTO Estimate	Est. DR Storage Cost
From Tape[15]	Tape only – no log playback	24 hours	21-108 minutes	Tapes[16]
	Tape plus mirrors for logs only[17]	0 hours	66-153 minutes	Tapes plus disk for logs
From remote replicas of database	Remote copy of local snapshot	12 hours	5 Minutes	Disk: 1x data size/copy
	Remote copy of local snapshot plus mirrors for logs only	0 hours	25 Minutes	Disk: max log size plus 1x database size/copy
From async replicas of database and logs[18]	App-based async mirror	Minutes	Minutes	Disk: 1x data size
	Host-based async mirror	Minutes	Minutes	Disk: 1x data size
	Array-based async mirror	Minutes	Minutes	Disk: 1x data size
Protected CDP[19]	Mirrored CDP storage	0 hours	5-30 minutes	Disk: 1 x CDP storage size
From sync mirrors	Sync mirrors	0 hours	5 minutes	Disk: 1x data size
	Sync mirrors with cluster failover	0 hours	Automatic	Disk: 1x data size

Figure 7: Summary of Disaster Recovery Tiers of Service

These DR tiers of service reflect products that are generally available in the industry today.

Note: These are estimates. Assumptions in performance are listed in the footnotes below. ***Your mileage will vary.*** You should create your own benchmarks for your environment.

[14] Data assumptions: 60GB exchange database plus 2GB/day for log files.
[15] This assumes immediate tape access and transfer rates of 55 MB/sec and 10 MB/sec.
[16] The number of tapes will vary based on size of data, tape rotation schedule, and the tapes re-use policy.
[17] Host or array-based mirrors.
[18] This assumes that consistency is maintained between database and logs – not all technologies support this. See Figure 6 for sample replica recovery times.
[19] This offers the added benefit of having multiple points in time from which to perform the DR.

5. References

[1] SNIA Dictionary: http://www.snia.org/education/dictionary

[2] ILM Best Practices in Security for Email, SNIA Data Management Forum, October 2004, http://www.snia.org/tech_activities/dmf/docs/Email_Security.pdf

[3] ILM Best Practices in Archiving and Tiered Storage for Email, SNIA Data Management Forum, October 2004, http://www.snia.org/tech_activities/dmf/docs/Email_Archiving_and_Tiered_Storage.pdf

[4] Managing Email for Compliance and Litigation Support – An Overview, SNIA Data Management Forum, October 2004, http://www.snia.org/tech_activities/dmf/docs/Email_Compliance_and_Litigation_Support.pdf

[5] Microsoft has very specific limitations on the supported use of NAS for Exchange – see http://support.microsoft.com/default.aspx?scid=kb;en-us;Q317173.

[6] Microsoft technical notes on Exchange Disaster Recovery: http://www.microsoft.com/exchange/techinfo/administration/2003/DisasterRecovery.asp

[7] XADM: White Paper - Disaster Recovery for Microsoft Exchange 2000: http://support.microsoft.com/default.aspx?scid=kb;en-us;326052

[8] IBM Tivoli Data Protection for Domino: http://publib.boulder.ibm.com/tividd/td/StorageManagerforMail5.2.html

[9] The Evolution of Enterprise Data Protection, Enterprise Storage Group, January 2004.

[10] NetApp SnapManager® for Microsoft Exchange Best Practices http://www.netapp.com/tech_library/3233.html

[11] EMC Legato Networker Module for Lotus: http://www.legato.com/products/networker/modules/lotus.cfm

[12] Gartner Management Update: Best Practices in Business Continuity and Disaster Recovery, IGG-03172004-01, March 17, 2004.

6. Available Products for These Solutions

The following DMF member companies have solutions that support one or more of the best practices defined in this white paper. The following URLs are provided for additional information.

Avamar Technologies: http://www.avamar.com

Avamar is leading the new wave of innovative solutions that deliver increased value for data backup, restore and disaster recovery providing enterprises with disk-based data protection solutions to control the information explosion. The company's flagship offering, Axion is a scalable solution that is setting new standards in data integrity, network and operational efficiency, and cost-effective scalability.

COPAN Systems: http://www.copansys.com

COPAN Systems' Revolution 200T uses patent-pending Power Managed RAID™ and Disk Aerobics™ technologies to provide 224 TB of MAID storage in a single footprint with 2.4 TB/hour throughput. The Revolution 200T is ideal for backup/recovery and active archive applications offering the performance and reliability of disk, at the cost and scale of tape.

EMC^2: **http://www.emc.com/products**

Using a tiered-solution approach, combining high-performance capabilities with lower-cost technologies, EMC^2 helps you meet & exceed varying service levels for both rapid restore and disaster recovery. Recover email databases within minutes—instead of hours or days—with advanced EMC^2 hardware & software solutions, including Replication Manager & Legato Networker.

Hewlett-Packard: http://www.hp.com

HP OpenView Storage Data Protector provides e-mail backup and recovery a broad feature set: on-line backup, zero impact backup, instant recovery of the full e-mail server in minutes - not hours, simple bare metal server recovery, single mail & single object restore, full & incremental backup & recovery, backup on tape & disk.

Hitachi Data Systems: **http://www.HDS.com/Solutions**

Dramatically reduce backup-time, while slashing mission-critical e-mail restore-time from hours to minutes with our Hitachi SplitSecond™ rapid-recovery solutions for Microsoft® Exchange. By incorporating proven Hitachi ShadowImage™ with our Global Solution Services' experience and expertise, SplitSecond™ for Exchange solutions integrates Hitachi Freedom Storage™ systems with our customers' rapid-recovery requirements.

IBM: http://www-306.ibm.com/software/tivoli/products/storage-mgr-mail/

IBM Tivoli Storage Manager for Mail automates the data protection of e-mail servers running either Lotus® Domino® or Microsoft® Exchange. This module utilizes the application program interfaces (APIs) provided by

e-mail application vendors to perform online "hot" backups without shutting down the e-mail server and improve data-restore performance.

Kasten Chase: **http://www.kastenchase.com/**

Kasten Chase's Assurency™ SecureData storage security solution provides security and encryption for stored data throughout the data lifecycle. SecureData enables storage consolidation, streamlines data management and enhances regulatory compliance. SecureData's Lifecycle Key Management provides policy-based, audited key creation, protection and deletion, ensuring efficient data compartmentalization and assured data destruction.

Network Appliance: http://www.netapp.com/solutions

NetApp provides a comprehensive set of data recovery solutions for both operational recovery as well as remote disaster recovery. NetApp's simple to use SnapManager and SMBR products provide lightning fast local recovery. NetApp's feature rich remote disaster recovery products provide both homogeneous support via SnapMirror and heterogeneous support via SnapVault.

About the Author:

Edgar has over 26 years of experience in software engineering, including 8 years in storage software development and the previous 18 years in the communications industry. His experience in the communications industry was focused on network management product development at Prime Computer and Motorola/Codex, applying protocols and data models derived from participation in OSI Network Management Forum activities. At EMC, he has been responsible for requirements analysis and software architecture for several of EMC's portfolio of information management products. Edgar received his BS in Computer Science from Roger Williams University. He is currently co-chair of the SNIA DMF's ILM Initiative Technical Liaison Group, and co-chair of the SNIA ILM Technical Workgroup.

About the Data Management Forum:

The SNIA Data Management Forum is a cooperative initiative of Information Technology Professionals, Vendors, Integrators, and Service Providers formed to define, implement, qualify, and teach improved and reliable methods for the protection, retention, and lifecycle management of electronic data and information.

About the SNIA:

The Storage Networking Industry Association is a not-for-profit organisation made up of more than 300 companies and individuals worldwide spanning virtually the entire storage industry. SNIA members share a common goal: to set the pace of the industry by ensuring that storage networks become efficient, complete and trusted solutions across the IT community. To this end, the SNIA is uniquely committed to delivering standards, education and services that will propel open storage networking solutions into the broader market.

