

# Backup and Disaster Recovery for Novell GroupWise

---

Written by *Tay Kratzer, Novell Press Author, Developer*  
*tk@gwava.com*

## Introduction

Your GroupWise email system is a critical communication medium and information repository in your organization. It's often taken for granted just how reliably GroupWise performs. The rock solid reliability of GroupWise may create a false sense that disasters that take GroupWise down are just not likely to happen.

If Novell made the server hardware, the server rooms and the buildings that house your GroupWise system, then you might be justified in not worrying about a backup and disaster recovery solution for GroupWise. But in the real world, there are increasing reasons to carefully consider your backup and disaster recovery solution for GroupWise.

## Email is Work Product

Here are a few reasons why you should make sure you have a good backup and disaster recovery solution for GroupWise:

- In many organizations, email has eclipsed the phone as the most used communication medium.
- Information is a vital "product" that all organizations create and share. GroupWise is the vehicle for transporting and storing a significant amount of that information.
- If GroupWise is down, production stops!
- Availability of GroupWise is critical, but so is the information in GroupWise.
- **All** messages in the GroupWise system need to be backed up. It's not acceptable to have a backup that just grabs a moment (snapshot) in time.
- It is no longer acceptable to say that you only restore information for "important people." Most modern organizations operate with significantly reduced human resources. Even users with less visible positions within an organization may need information restored because the information is vital to the organization as a whole.
- Other critical systems tie into the data in a GroupWise system. One example is email and personal address books in the GroupWise system that are replicated to wireless devices such as BlackBerry devices or GroupWise Mobile Server-connected devices.

Because of the reasons just listed, your backup and disaster recovery solution for your GroupWise system should include the following functionality:

- Quick disaster recovery
- Easy GroupWise data recovery
- A complete backup of all message items in GroupWise

## Usable Backups

It may be that you need to shift the way you think about backups from, “we use our backup system only if we have to” to, “our backup system is an integral part of our overall GroupWise system.” Whatever solution you devise needs to address both quick disaster recovery and quick data recovery. If you do not devise such a solution, it is guaranteed that you will have a lot of wasted time and frustration on the part of many people in your organization.

If the backup system that you create for GroupWise gives you quick access to the native GroupWise Message Store, then you may have also created a quick disaster recovery solution. However, if your backup solution has data simply replicated to another disk location (such as a SAN), in the case of a server disaster, you may have GroupWise data with no server around to host the data.

## Backup vs. Archiving

Data kept in another repository, such as an archiving solution, isn't nearly as useful because it's not “actionable.” Actionable data is native GroupWise data that is accessible from the GroupWise client or connected clients for use in the GroupWise environment. Furthermore, there are components of the GroupWise experience that cannot be easily replicated to other repositories. Any GroupWise administrator who has been required to try and restore a GroupWise Address book or GroupWise Calendar knows that there is more to a GroupWise user's account than their e-mail. And it should go without saying that if you have an archiving solution, it certainly cannot function as a disaster recovery solution for your GroupWise system.

## A High-Level Technical Overview of GroupWise Data

GroupWise users' mailboxes are kept at a GroupWise post office. The GroupWise post office data (Message Store) is a database. However the structure of the GroupWise Message Store database is unique when contrasted with other email technologies. Here is what makes the GroupWise Message Store unique:

### Scalability

The GroupWise Message Store database is highly partitioned. The partitioned architecture of the database allows for:

- Significantly less server CPU, network and disk resources needed on the GroupWise server vs. other competing solutions.
- Highly contained damage, meaning; if there is damage to the GroupWise Message Store, it is generally contained in a small subsection of the GroupWise Message Store.

Ten percent of the data in the GroupWise Message Store database is highly dynamic (changing) in nature and 90% is largely static (unchanging).

- The dynamic data (10%) of the GroupWise Message Store is contained in database files.
- The static data (90%) of the GroupWise Message Store is contained in individual record files, in which one file holds only one record apiece.
- The databases (dynamic data) are effectively indexes of all data kept in both the databases and the record files (static data).

### Efficiency

Only one copy of a message exists in the GroupWise Message Store on a GroupWise post office.

- If USERA sends a message to USERB and USERC on the same post office, POST\_OFFICE\_ONE, GroupWise saves only one copy of the message and its associated records are kept and shared by all three users' mailboxes.

NOTE: This design is notable since it doesn't save multiple copies of attachments. It's a guarantee that attachment sizes will continue to grow, especially as bandwidth expands. With GroupWise's single-instance storage, and compression technology, your email system is well suited for the future.

- If USERA on POST\_OFFICE\_ONE sends a message to USERB and USERC on a different post office, POST\_OFFICE\_TWO, there is one copy of the message stored at POST\_OFFICE\_ONE, and one other copy replicated to and stored on POST\_OFFICE\_TWO.

Every GroupWise user has his or her own individual database file in the GroupWise Message Store with the following architecture:

- Most of a user's database is largely an index pointing to the individual message records which are mainly located in, or indexed in, one of 255 message databases.
- User database files contain data that is specific only to the user. This data is as follows:
  - Personal items such as personal appointments, personal tasks and personal notes
  - Personal address books
  - Personal GroupWise rules and preference settings

## GroupWise Message Store Rules

- No individual database will exceed 2 gigabytes in size\*.
  - Because of the highly partitioned nature of the GroupWise Message Store this rule does not actually limit the scalability of users' mailboxes.

\*NOTE: There are some rare cases in which an individual database might bloat beyond 2 gigabytes in size, however; this is considered an error condition that must be corrected.

- Individual records (which contain messages) never change.
- Records exceeding 2 kilobytes in size, are written outside of the \*.db files.
  - These records represent 90% of the static data in a GroupWise Message Store

## GroupWise Message Store on Disk

On disk, the GroupWise message store looks like this:

- A user database directory that contains a database (\*.db) file specific to each user on that particular GroupWise post office.
- A message database directory that contains approximately 250 message databases.
- A static record directory structure under which are approximately 250 subdirectories that contain the records that are over 2 kilobytes in size.

So, if you were to closely examine the message store, you would discover that:

- On larger post offices there are **hundreds of thousands** of static record files.
- The disk space ratio of static record file directories vs. database file directories is approximately 9 to 1. That is, 10% of the data is in the database file directories and 90% of the data is in the static record file directories.
- Individual databases, or message databases, can be in the hundreds of megabytes in size.
- The last accessed times on almost all databases are recent, which means that all databases are dynamic in nature.

*"The biggest hurdle to getting a complete and consistent backup of the GroupWise Message Store is its partitioned nature."*

## Challenges of Backing-up GroupWise Data

Just as the GroupWise Message Store database is unique in its design, backing-up the GroupWise Message Store is unique its implications. Here are the two key points to backup and recovery of the GroupWise Message Store:

- In order for a backup of the GroupWise Message Store to be considered complete and consistent, every database and record file must be backed up.
- Likewise, for the GroupWise Message Store to be wholly useful for disaster recovery or data recovery purposes, every database and record file must be available for complete and consistent recovery.

The biggest hurdle to getting a complete and consistent backup of the GroupWise Message Store is its partitioned nature. For example, if a GroupWise user's personal database is backed up, that doesn't mean that the user's mailbox is actually backed up. It just means that *one* file in the Message Store is backed up, but all other files must also be backed up.

### Solution #1: SmartPurge

Novell has long been aware of the challenges of getting a complete and consistent backup for GroupWise. To help with the problem, Novell has provided a technology called "SmartPurge." If enabled, the SmartPurge technology works in the following manner:

- A message item and all of its supporting records cannot be emptied from a user's GroupWise mailbox trash until they have been backed up.
- If a user archives an item that has not been backed up, a copy of the item goes to their archive mailbox, but a copy of the item is also kept in the user's trash until the item has been backed up.

In my own experience and research with third party backup solutions, many vendors do not take advantage of the SmartPurge API. While some aspects of their technology are very cutting-edge, they miss the boat when it comes to GroupWise Message Store consistency and completeness. Without using SmartPurge, their backups have the following deficiencies:

- Backups are somewhat consistent, but not entirely consistent
  - In the instance when an item comes in during the time of the backup and a static record file is written out, if the backup has already completed backing up the new static record files, before backing up the databases, then content of the backup is considered inconsistent.

- Backups are only a snapshot in time
- Backups are incomplete, because not all items are backed up
  - For example, if an item comes in at 9:00 A.M. and is deleted at 11:00 A.M. and the backup goes on at 8:00 P.M. then that item has escaped the backup.

Vendors who integrate with the SmartPurge API technology are doing a much better job of backing up the GroupWise Message Store.

## Solution #2: Backup Media

To address the need for a “quick” disaster recovery or quick data recovery solution for GroupWise, you will need to determine the backup media you want to use. For long-term storage, tape media cannot be beat for cost effectiveness. However, for near-term backups and disaster recovery, tape media is not viable because it just too slow and cumbersome. Fortunately, hard drive storage has become less expensive and is the best alternative for “hot backups.” By implementing a hot backup solution, your GroupWise Message Store can be available for:

- faster disaster recovery
- faster data recovery

For several years now, customers and third-party vendors have realized that the best way to implement GroupWise backups is by using hot backups with disk-to-disk-to-tape technology. Novell has even provided two utilities called GWBACKUP and DBCOPY to assist in this effort.

The following sections address the backup technologies available, rather than specific vendors. Some of these technologies may or may not be difficult to implement, depending on the skill set of your IT staff. Rather than give my personal opinion on complexity, you can investigate these options on your own.

## Snapshot Technology

This technology grabs the data in the GroupWise message store at one moment in time and stores the snapshot in another location. Here are the advantages and disadvantages of this technology:

### Advantages

- Data is available for relatively quick disaster recovery
- Data is available for relatively quick message restore

- Methods of creating a snapshot are well-documented and as such, you may be able to devise a large portion of your GroupWise backup strategy in-house.
- Given the right configuration and hardware, a snapshot solution can be very fast.

### Disadvantages

- The faster snapshot solutions require a downtime window of some sort.
- Each snapshot is 100% of the size of the post office. So if you identify that most of your restoration needs are from the last two weeks for example, keeping all of those snapshots is going to take a lot of disk space.

NOTE: Virtual Machine (VM) based solutions may help to resolve this issue because VM images can be relatively small and can be combined with master VM sets fairly easily.

- None of the snapshot solutions that I am currently aware of integrate with the SmartPurge technology. As such, these backups should be considered mostly consistent, but incomplete.
- Some snapshot solutions are very slow at creating backups. For example, a university on the east coast of the United States recently reported that their third-party snapshot solution took over 24 hours to perform a backup of one 350 gigabyte post office.
- Some snapshot solutions may not support GroupWise on platforms other than NetWare.

## Full Backup-to-Disk Technology

This technology provides a fully replicated GroupWise message store to another location. Here are the advantages and disadvantages of this technology:

### Advantages

- Data is available for relatively quick disaster recovery
- Data is available for relatively quick restore
- Methods of creating a full backup-to-disk backup with GWBACKUP and DBCOPY are well documented and as such, you may be able to devise a large portion of your GroupWise backup strategy in-house

### Disadvantages

- Each backup is 100% of the size of the post office. So if you determine that most of your restoration needs are from the last two weeks for example, keeping all of those full backups is going to take a lot of disk space.
- Only some of the third-party full-backup-to-disk products integrate with Novell's SmartPurge technology. Those backups that do not use the SmartPurge technology are somewhat consistent, but incomplete.
- Homemade backup solutions that use GWBACKUP and DBCOPY are often slow and very limiting. Furthermore, backups using GWBACKUP and DBCOPY still do not take advantage of SmartPurge integration. As such these backups are incomplete. Neither are they entirely consistent.

### GWAVA Reload

Unlike other GroupWise backup solutions, Reload's technology is unique. GWAVA Reload was written from the ground up in order to address the challenges of backing up the GroupWise Message Store. It is unique because its backups are both consistent and complete. Reload provides a hot backup *and* integration with the SmartPurge technology. Here are the advantages and disadvantages of Reload:

### Advantages

- Reload provides push-button disaster recovery so that a GroupWise post office can be up and functioning on the Reload server in a matter of two minutes. No decisions need to be made; no backups need to be combined. It simply requires the push of a button and the live post office can be hosted at the Reload server. The Reload server loads a GroupWise POA right on the Reload server running against the most current backup, so it is immediately available to service any client connections, including wireless devices and GroupWise WebAccess.
- Reload tightly integrates with Novell's GroupWise SmartPurge technology. Nothing escapes Reload!

- Reload backups are available for data recovery in a matter of two minutes.
- Every Reload backup is effectively a full backup, but on average, it only requires 12% of the size of the post office (1/8<sup>th</sup> of the space required by other hot backup technologies).
- Reload backup times are significantly better than any other backup solution. For example, a 300 gigabyte post office at the State of Utah is being backed up in 45 minutes a night. A Chicago law firm that uses Reload backs up their 850 gigabyte post office in 70 minutes a night.
- "You're only as good as your latest backup." Because Reload backup times are so fast, it is feasible to set Reload to create intra-day backups. Furthermore, those intra-day backups require no additional disk space on the Reload server.
- Reload requires significantly less disk space than other hot backup solutions. Two weeks of data for a 100 gigabyte post office will require 500 gigabytes on the Reload server. For backups using snapshot or full backup-to-disk technology, the space required would be approximately 1.4 terabytes.
- Reload backs up GroupWise post offices and domains on any server platform that GroupWise runs on.

### Disadvantages

- The most common concern with Reload is its perceived complexity. Reload runs on a Novell SLES or OES Linux server. People investigating Reload are most worried about using Linux.

Linux in its native format does have a steep learning curve; however, Novell's SLES and OES products have greatly reduced that learning curve and it's getting better and better with each new release.

Reload itself is 100% menu driven. Reload has a web interface and a C-worthy type interface similar to the NetWare server console menuing system. Installing Reload, configuring and using Reload requires virtually no knowledge of Linux.

## Conclusion

The GroupWise Message Store is partitioned so that it saves a significant amount of CPU, network, and disk resources when compared to other collaboration products.

The biggest hurdle to getting a complete and consistent backup of the GroupWise Message Store is its partitioned nature.

In order for your backup solution for GroupWise to be up-to-date and usable for disaster recovery and data recovery purposes it should provide:

- Hot backups on disk
- Novell GroupWise SmartPurge integration
- Quick access to its hot backups for both disaster recovery and data recovery



[www.GWAVA.com](http://www.GWAVA.com)

+1 (866) GO-GWAVA • fax +1 (646) 304-6250 • questions@GWAVA.com  
100 Alexis Nihon • Suite 500 • St-Laurent • QC • Canada • H4M 2P1