

Protecting Open Databases and Files (Windows Agents Only)

Overview

Active databases and open files are typically your most important data. Therefore, it is critical that you employ a data protection solution that natively supports backing up these mission-critical files. Backup solutions that require complex plug-ins or costly third party tolls to protect open databases and files, such as Open File Manager (OFM) or Open Transaction Manager (OTM), are vulnerable to serious compatibility and operational problems between the backup solution, the database and the database plug-ins. Unfortunately, most of today's backup solutions require these third party tools to back up open databases and files.

Open Database Protection with the SaaS BackupGrid

The SaaS BackupGrid, an automated server backup and recovery solution, removes risk and complexity through built-in support for backing up open files and databases. The SaaS BackupGrid continuously protects open and changing databases such as Microsoft Exchange Server, Microsoft SQL Server, Oracle and Lotus Notes without the need for complex plug-ins or costly third party tools. This is a unique and powerful advantage of the solution.

This document describes two technologies employed by the SaaS BackupGrid that work together to safely and effectively protect open files and databases:

1. Snapshots to ensure access to open files and to ensure transactional consistency of databases.
2. A software agent that utilizes Microsoft's NT Filter Driver APIs to interact directly with file system-level functions to efficiently record changes to files as they occur.

Job Phases

Individual backup processes are referred to as jobs that transition between two phases: initial backup and continuous or scheduled backup.

Initial Backup

The first time a server is configured for backup, an initial backup must first occur. A copy of each file and directory selected to be backed up is efficiently and securely sent to the vault, creating a baseline image of the files and file structure. The initial backup runs until completion. If the initial backup is interrupted (due to server or network outages) the backup will resume automatically from where it left off (under normal circumstances). When the first copy of all data selected for backup has been replicated to the vault, restore jobs can be successfully run to recover data.

Continuous or Scheduled Backups

After the initial backup of a file, the SaaS BackupGrid only needs to transfer the changed deltas within the file. Similarly, after the initial synchronization, the SaaS BackupGrid always performs delta backups and never needs to perform another synchronization (assuming we do not run into any problems, sometimes we request that they re-baseline, and store the old data elsewhere). This enables the SaaS BackupGrid to efficiently back up data 24x7, typically sending changes to data on 15 minute intervals while minimizing network bandwidth usage.

Replication Technologies

Two powerful technologies are used in parallel during the phases described above:

Point-in-Time Versions via Snapshots

Snapshots are a file system function. They provide an isolated, virtual file system for specific applications (such as backup) to access files as they exist at a

moment in time, without having to compete with other applications using those files.

During backup jobs—typically at 15 minute intervals—the SaaS BackupGrid instantiates snapshots. The BackupGrid uses snapshot technology for two primary reasons:

- Snapshots give the software agent full access to a server’s file system without the need for open file managers and without risk of impacting files currently in use.
- Snapshots lock the entire file system at a point in time. This allows for transactional integrity of restored databases.

In practice, when the software agent requests a snapshot, an empty snapshot file is created. After this, if an application modifies a file, a “pre-image” of the blocks of the file being modified are moved to the snapshot file. When SaaS BackupGrid then asks for this file, the snapshotter sends the BackupGrid’s request for this data to the snapshot file. Once the backup job is complete, the BackupGrid deletes the snapshot file.

Efficient Delta Backups via File System Filter

The Microsoft Filter Driver framework is a well-defined set of programming interfaces developed by Microsoft to allow third party services to interact with actions occurring within the file system.

The SaaS BackupGrid uses a file system filter to monitor the changes that occur between snapshots. This way, the BackupGrid is able to efficiently capture and send just the changed deltas after a snapshot is instantiated. The filter removes the need for the BackupGrid to scan the file system for each backup job to determine the changes to the file system. This significantly improves the performance of backup jobs (which typically finish in seconds to minutes).

Not all Microsoft OS versions include the filter driver framework. When the framework is not present, the scanning process is longer, depending on the number of files and the amount of data in the file system. The software agent software scans the file system to identify files and blocks that have changed.

Putting It All Together

By using snapshots and filter drivers, the SaaS BackupGrid is able to natively protect open files while ensuring the transactional integrity of databases. This is accomplished with a minimum of resource consumption. Through the use of these technologies, the BackupGrid addresses two very important obstacles in backing up open and changing files:

- Files open for exclusive read
- Actively changing files

Backing Up Open and Changing Files

Snapshots, by definition, provide complete read access to all files on a system, with no ability to modify them. They are then available to applications using the snapshots.

With the filter driver monitoring file system activities, the software agent is able to:

1. View files as they are opened.
2. Log the location of changes occurring in the files as commands are executed on open files. Changes and their locations are tracked—no data is cached during this process, only records of the changes and their locations.

Database Transactional Integrity

In a 24x7 production environment, many database or application files are always open and actively changing. If the server suffers a power failure and is restarted, one of the functions of the database software is to bring the database into a transactionally consistent state as part of its startup activity by “rolling out” any incomplete transactions. To do this, many databases keep transaction log files as well as the main database files. All modern databases such as SQL, Exchange, Oracle, Sybase and Lotus Notes maintain transactional consistency in their databases.

Backup products typically require either that the database be brought into a quiet state or shut down before backup to ensure that the collection of database files (log files, database files, etc.) is in a consistent state.

In contrast, the SaaS BackupGrid does not require that the database be shut down nor does it necessitate

any pre- or post-processing or the purchase of database plug-ins. Through its snapshot technology, the BackupGrid can guarantee that all the database files are “captured” at a specific instant in time which reflects the on-disk state of all those files at that instant. If the collection of database files is later restored, when the database starts up it will detect files in a state that could have resulted from a power failure or other interruption and, if necessary, will make the database transactionally consistent as part of its startup.

Handling Log Files

With traditional backup, the normal scheme is to retain a database’s transaction log files until after a known good backup has occurred. Often the backup software plug-in must reset the log files or command the database software to do so. Regardless of the method, the size of the log files must be contained so that these files don’t eventually fill up the disk.

As it is unnecessary for the SaaS BackupGrid to interact with database software, the following methods are typically used in conjunction with the BackupGrid to control log file growth, depending on the database:

- SQL – the database consistency checker (DBCC) should be used on a regular basis.
- Exchange – circular logging should be used.
- Oracle (which always uses circular logging) – a simple script or procedure should be used periodically if ARCHIVELOG mode is on to clean the archived redo log files.

Using Native Database Backup Features with the SaaS BackupGrid

Many database packages provide features for handling or assisting with backup. These features can be used in parallel with the SaaS BackupGrid if so desired. A few typical examples are mentioned below:

- SQL provides an SQL Backup capability to make dumps of the database and/or transaction logs into a separate backup directory. The SaaS BackupGrid’s continuous off-site protection often makes the use of SQL backups unnecessary, as many of our customers choose to recover to a 15 minute

point-in-time in the recent past. Others choose to continue to make SQL backups a few times a day as extra protection.

- Oracle offers a logical backup facility as well as the RMAN utility. Oracle also provides ARCHIVELOG mode to enable recovery from online backups. As discussed above, the SaaS BackupGrid performs online backups in a unique way. With the BackupGrid, the main database files are backed up as they change as well as the collection of online redo log files that change. It is unnecessary for customers to backup the archive log files, although they may do so if desired.

Conclusion

Open databases and files are the most difficult to protect – yet they are typically your company’s most valuable data assets. The SaaS BackupGrid helps you overcome this challenge by providing reliable, automated protection of your company’s valuable data, including open databases and files. The BackupGrid’s unique features provide secure, continuous protection for critical open database applications such as Microsoft SQL and Exchange, Lotus Notes, Oracle and other core business applications—without the need for third party add-ons and plug-ins or the extra effort required by more traditional backup solutions.

Please see our other white papers and informational materials for more details on the benefits of online backup and recovery.

About Utility Backup Solutions

Data is the lifeblood of any company, and data backup in the modern era should be as simple as ‘plugging in’ to a utility that is always on, always reliable and can actually deliver when you need it.

Utility Backup Solutions offers best-of-breed, one-stop data protection and recovery solutions designed specifically to address the special needs of today’s small- and-medium-sized businesses. With our automated solutions for servers, our customers can reduce operational costs, focus staff on new initiatives and improve workforce productivity, adhere to regulatory and compliance concerns and achieve total peace-of-mind.

The SaaS BackupGrid service, based on Iron Mountain’s award-winning LiveVault platform, is a comprehensive solution for online protection and recovery of server data for small-and-medium businesses in North America. The SaaS BackupGrid is offered in a utility billing model—that is, the services provided to our customers are consumption-based, pay-as-you-go software-as-a-service subscriptions.

With Utility Backup, you can finally abandon the administrative and financial burden of a legacy magnetic tape backup program, and join the thousands of satisfied customers enjoying the “set-it-and-leave-it” reliability of a modern data protection solution.

We pride ourselves on offering a valuable and effective solution to a real problem for our customers. We believe in transparency, customer satisfaction, and providing true value. If we are successful, we believe we will be rewarded with long customer relationships.

Contact Us

Utility Backup Solutions, LLC
570 El Camino Real #150-501
Redwood City, CA 94063-1262
888-374-3282
www.utilitybackupsolutions.com