

Storage Management On Cooperative Computing Networks

By Thomas J. Meehan



Over the last few years, the phenomenal growth of LANs, which was, perhaps, over-emphasized by the press, also speculated on the death of the mainframe. We are now reading that they are not quite ready to be buried and there might even be a resurgence for the mainframe.

As with their premature pronouncements of death, these reports on rebirth are the media looking for controversy to sell newspapers. They overlook details of what is really happening. Do not expect the venerable and insulated mainframe systems of the '70s and '80s to reappear.

Today's mainframes are equal partners in the world of cooperative computing — a new world in which mainframes, LAN systems, PCs and workstations (distributed and centralized) all share and exchange data.

Conventional wisdom demands "Backup anything on which you must depend." A corollary to that advice is "A backup is good only if the data can be easily restored." This introduces bothersome details such as keeping track of backups, assuring that they are in a safe place, recycling them and deciding their life expectancy. When disk space becomes scarce, migrating data is the consummate answer. During the past 25 years, sophisticated storage management solutions have evolved for mainframe disk systems.

Basic Principles

Basic principles do not change — the same storage management solutions that worked for mainframe systems will work for cooperative computing networks. Attaching small-frame tape drives is a traditional approach to backing up PCs. Applied in a cooperative computing network, this approach can evolve into a ridiculous scheme that creates hundreds of little data centers complete with tape drives and increased staff.

Using reliable, small-frame tape drives with automatic loaders for LAN file servers can cost

tremendous sums of money. Trying to cut costs by connecting fewer tape drives to centralized servers introduces performance problems associated with transmitting backup data over a network and still leaves redundant clerical and administrative costs.

Capitalize On Existing Resources

Organizations with MVS systems already have a substantial investment in tape and DASD hardware. They also have in place attendant support staff responsible for backup and restore, usually with well developed plans to handle disaster recovery. Ironically, there is less than universal agreement on involving the mainframe in a storage management solution for PCs, workstations and LAN file servers.

Much of this controversy is nothing more than a turf war, easily resolved by good managers. However, network client/server systems have unique requirements stemming from the very technology that makes them possible, and most backup solutions oriented toward mainframe backup servers are not up to the task of network storage management.

Backup

Traditionally, there are two procedures for backup:

- Full-Volume Backup — copies all the volume's files and its file system or directory information.
- Incremental Backup — copies only new and changed files. This takes less time, but makes it difficult to manage a restoration.

To circumvent these problems, a balance is usually struck. After an initial full backup, incremental backups keep copies of files current. To reduce the problem of managing numerous incremental backups, a new starting point is periodically established at regular intervals by taking another full-volume backup.

A traditional approach, however, may not prove feasible on a network. Full backup is easily postponed because of the time it takes. Users are reluctant to run a backup that may take 10 to 25 hours to complete. The deployment of high-tech fiber links may offer future relief; however, without significant improvement, line speed is the major deterrent to network-based backup.

The incremental approach may present problems during restore, but it is attractive in a cooperative computing situation, particularly when only a few files change on any one day. This is a perfect example of technology demanding an innovative approach — one that will produce the same result on cooperative computing networks in a shorter time.

New Approaches

Data in memory and cache techniques are based on the premise that the fastest I/Os are the ones that are not done. Incremental backup applies this concept when it only copies a file if it changes.

Now extend this idea to full-volume backup. Do not transmit anything across the network that is not necessary. Use data previously read (backed up) and already on the mainframe as the source for unchanged files. To create a full backup, merge these copies forward, together with the regularly transmitted copies of new or changed files and a current copy of a volume's directory. With less data moving across a network, an MVS backup server can construct a full backup complete with volume directories.

Forward Merge Backup

Forward merge backup describes both the process and the result. In a normal incremental backup, a client PC or LAN file server ships a copy of new and updated files to an MVS backup server. For merge backup, a copy of a volume's file directory is also transmitted. Files that have not changed are already on the MVS mainframe.

The key to all this is a central inventory control system — an inventory able to identify a file that changes several times, for example, on Monday and then again Tuesday, but not on Wednesday, Thursday or Friday. The Tuesday copy is selected for a merge to the full backup. This means the program files that *never* change (e.g., files in Windows, OS/2 systems directories and the unchanging files in word processor, spreadsheet and other application directories) will be transmitted across the network only once.

Forward merge can produce a more thorough backup and overcome problems encountered with the network backup of PCs and LAN file servers. One shortcoming of traditional PC and LAN full-volume backup is that it may skip open files. This does not happen with a

forward merge backup. AN MVS backup server does not have to depend exclusively on a file's update attribute flag — it can compare a current copy of a volume directory with its own inventory information. When it finds a file entry that does not match its backup, the MVS server requests a current copy on the file from the client PC or LAN file server.

All-Tape Scenario

Variations within merge backup offer different advantages. An all-tape scenario is a backup directly to tape without intermediate DASD and can be a single tape file per week. For example, take an incremental backup on Monday; use the same tape on Tuesday; and have the system automatically piggyback all succeeding incrementals to the same tape. At the end of the week, do a forward merge backup. New and changed files are again piggybacked to the same tape.

All the incremental copies of files backed up that week already on the same tape presents a definite advantage. A file backed up Tuesday but only read Wednesday, Thursday and Friday does not have to be retransmitted across the network. Since the file is already on the tape as part of the Tuesday incremental, it does not have to move. Also a simple update to an inventory tracking database will show that the file is now part of the full-volume backup.

Merge backup is a win-win proposition. If many files are normally backed up by incremental, there is less work creating a full backup. The files are already on the current tape. When there are fewer incremental files, merge may do more copying but all the work is done on the MVS backup server. Unchanged files on the previous full backup are simply copied forward to the new full backup — there is no network transmission. Additionally, only files resident on the volume are merged forward — not deleted files. Recording a copy of the directory on tape make it self-contained for disaster recovery purposes. Old (full merge) tapes may be scratched or kept for years on-site or off-site.

The advantages include:

- Data goes directly to tape with no DASD requirement.
- A full-merge backup produces one backup file per week.
- There is no cluttering of tape management systems.

- You deal with one entity.
- You have one backup to bring to a disaster site.
- Copies of incrementals can be sent off-site daily.
- At the end of a week, all off-site copies may be replaced with a single, full-merge backup.

Tape And Disk Scenario

Users typically schedule backup for network clients when network traffic is lightest. This can put backup into contention for tape units with an overnight production cycle. Putting the incremental to disk in a combined tape and disk scenario eliminates any contention for tape units. It requires some disk space but this is limited, since it is only for incremental copies.

On weekends or whenever full backup is created, incremental disk backups become part of the repository of unchanged files. Merging disk copies to tape under this scenario may require a little I/O but it allows a single full-merge backup to replace individual disk copies and the forward merge backup eliminates network transmission of unaltered files. Additionally, recovery is faster if restores are from current backups because there are no tape mounts.

Hierarchical Storage Management (HSM)

Hierarchical storage is an idea proposed to compensate for the limited capacity and high cost of disk storage. It advocates moving less-important data to more economical and higher capacity but typically slower media. Establishing a hierarchy of storage media and moving data along that hierarchy is known as file migration.

To be successful on a network, a file migration scheme must be easy to use and must overcome transfer speed limitations.

One school of thought is to improve performance by keeping a level of the storage hierarchy on the network, close to the user. This is typically so close to the top of the hierarchy, though, that it defeats the purpose. Identifying files that are infrequently referenced often can take considerable time, so most storage

management systems apply a bit of probability theory. They simply *bet* files that have not been referenced for some time are dormant and will be good migration candidates.

Consider the advantages for migration under a merge backup process — files that already reside in the backup repository on the MVS system are the best migration candidates. Migration specifications such as “How long should it stay on a PC or LAN file server before it migrates and then how long to keep it in the migration system?” can easily be included on the same menus used to define file backup specifications. Marking files that meet these specifications as migrated in the inventory control file can be a processing option.

Files meeting migration specifications are probably on the previous full backup and do not need to be transmitted across the network. Even if migration age is kept tight, a file is on a current incremental backup. Migration will never require separate processing to expire files or consolidate tapes if files are automatically carried forward until they expire or are restored; it is taken care of during a forward-merge backup. Migration has almost no extra overhead; the PC simply scratches the file from its disk.

Inquiry

An MVS resident, central inventory control facility provides a tremendous inquiry capability for both storage administrators and users. With all backups recorded in one place, anything from the latest incremental to a previous full-volume backup and every backup in between is easily found. With all the copies of files migrated from a volume on the same backup, it is as if a PC or LAN file server has an expandable disk volume. An integrated system could display migrated files in line with normal backups. Restoring a file, directory or entire file system is accomplished with only a click.

Restore

“Transfer as little as possible” also applies to restoration. Restore should always use compression and transmission-reduction schemes. Traditionally,


network storage management systems use incremental backup to save time. However, because incremental backups do not account for deleted files, after a while, a volume cannot hold all the files that might be restored to it. This will cause problems unless a restore can automatically bypass deleted files and files recorded as migrated. This requires a restoration system to periodically include a full-volume backup or, even better, maintain an image of a volume's file directory as with the forward-merge technique.

Restore should transmit only a single image for each file that is to be restored. To ensure the fastest recovery, a system should chronologically process physical, sequential files coming from tape or disk. To ensure that the single, most current image of a file is restored, it is prudent to simultaneously process its control information inventory in reverse chronological order.

Conclusion

Cooperative computing has not changed basic principles for storage management—the solutions remain the same:

- Store a copy in a secure place.
- Transfer as little as possible, using the fastest means possible.
- Do only what is absolutely necessary.
- Put less-important data on less-costly media.
- Automate as much as possible and simplify the rest.

Forward merge backup is an example of applying these basic principles in an innovative way. 

ABOUT THE AUTHOR

Thomas J. Meehan, a vice president of Innovation Data Processing, Inc., is responsible for emerging technology. Innovation, well-known for its FAST DUMP RESTORE (FDR) family of products, specializes in backup and restore and DASD storage management. Innovation Data Processing Inc, 275 Paterson Ave., Little Falls, NJ 07424, (201)890-7300.

FDR/UPSTREAM

Centralized Enterprise LAN Backup By Pat Fitzsimmons and Bob Perper

PC and File Server backup is often a dirty deed that no one really wants to perform. Backups are time-consuming and manually intensive, and someone must do them every day. Some servers may not get regular backups and users' PCs are often completely forgotten.

No longer can PCs be thought of as personal. They are a corporate asset containing information and data that goes beyond the simple memo. This asset must be adequately protected.

The corporate mainframe offers facilities unmatched; automated schedulers, robotic tape loaders, operations staff, physical security as well as a controlled environment and disaster recovery make it the ideal environment for the storage of corporate data.

Innovation Data Processing's FDR/UPSTREAM is an industrial-strength PC and LAN file server backup solution. Backups and restores can be completed automated and unattended using either existing host scheduling facilities or a supplied PC scheduler. This relieves the individual departments of the numerous problems associated with the PC-based backups.

FDR/UPSTREAM joins FDR, ABR, IAM and Innovation's other products as a high-performance, cost-effective solution for PC and LAN file server backup needs.

Works Well With Others

FDR/UPSTREAM stores PC and LAN file server data on mainframe disk and/or tape. It integrates smoothly into existing storage management systems (ABR, HSM, etc.), which allow FDR/UPSTREAM to fit nicely into your company's existing operations and disaster recovery plans. Finally, all your critical corporate data is back in a safe, secure place.

Security of your PC's data is guar-

anteed by your existing host security system. FDR/UPSTREAM cooperates with RACF, ACF-2 and TOPSECRET while maintaining PC security information.

Send PC data to the host, using communications facilities, such as SNA APPC/LU6.2, which is included in VTAM or TCP/IP.

No Experience Necessary

FDR/UPSTREAM's newest release offers TSO ISPF panels to help make the installation and configuration a breeze. Status panels allow operations staff to view the state of your backups and restores, and provide statistics and reporting for a complete view and audit trail.

Panels help you set up automation of recurring backup jobs for integration into your scheduling system.

PC files stored on the mainframe can easily be viewed using the ISPF panels, and selected restores can be specified and sent to a PC or server.

No Speed Limit

What makes FDR/UPSTREAM so unique is its speed. Not only is it a fast micro-to-mainframe bulk data product, but it is the smartest.

FDR/UPSTREAM's unique forward merge facility intelligently assures that only the data needed to build a complete backup is actually transmitted. Basically, day-to-day operations are incremental backups.

On The PC Side...

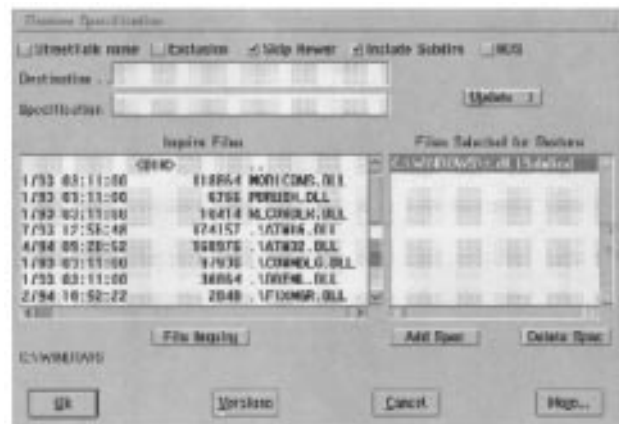
FDR/UPSTREAM PC is as complete as its mainframe partner. Supporting Novell, Banyan, IBM LAN Server, NT Advanced Server, FDR/UPSTREAM backs up and

restores most major types of LAN file servers.

Just as on the mainframe, FDR/UPSTREAM PC can be used by the inexperienced; GUI, point-and-shoot screens make it something that almost anyone can use.

FDR/UPSTREAM is not only for file servers. Its unique ULtra facility makes the automation of backups of any number of workstations simple and inexpensive. A small component resides on the workstation and a central PC backs up the data using sophisticated LAN protocols.

Mainframe-like storage management is here for PCs. FDR/UPSTREAM can detect, report, migrate and store files that



have not been accessed based on rules you define.

Mainframes have been offering automation for many years. Finally extending these services to your PCs is a smart move. For more information, contact Innovation Data Processing, Innovation Plaza, 275 Paterson Ave., Little Falls, NJ 07424, (201)890-7300. □

ABOUT THE AUTHORS

Pat Fitzsimmons is product manager and Bob Perper is a developer of the FDR/UPSTREAM product line from Innovation Data Processing.