New Data Protection Strategies

by Norlynn Nelson & Russ Fellows

June 20, 2008

Technology Insight Series

Evaluator Group

enabling you to make the best storage decisions

Executive Summary

One of today's greatest challenges facing organizations is the protection of corporate data. The issues that complicate data protection are compounded by increasing demand for data capacity, and higher service levels. Often these demands are coupled with regulatory requirements and a shifting business environment impacting applications and infrastructure. IT organizations must meet these demands while maintaining flat budgets.

Protecting data requires a disaster recovery plan that meets business needs, but is able to operate efficiently in order to deliver a competitive advantage for the organization. One technology that can help improve data protection service levels, with minimal impact to operations cost-effectively is Continuous Data Protection (CDP).

IBM with TSM FastBack leverage CDP technology to deliver a set of data protection and recovery tools, which enable IT users to deliver cost effective disaster recovery solutions. The technologies are flexible, allowing businesses to customize solutions capable of delivering higher service and protection levels than were previously possible for each cost point; balancing cost, performance and data protection.

The Delicate Balance

Protecting a business' most valuable resource is the primary challenge facing today's information technology (IT) personnel. The most valuable corporate resource is information, which must be protected from a number of natural and man made disasters. The challenge to not only protect data, but to ensure that disasters may be weathered has many facets, including:

- An ever increasing amount of data
- The importance of application availability, reliability and scalability based on changing business requirements
- Finite backup windows
- Progressively more stringent Restore Point Objectives (RPO) and Restore Time Objectives (RTO)
- Fiduciary and other regulatory compliance issues
- Remote and global geographically dispersed systems

Government statistics show that over 90% of companies who lost access to their data center information for more than 10 days filed for bankruptcy within one year. The issues surrounding backup and data recovery affect all businesses, regardless of size. Whether it is a Fortune 100 company or a small to medium size business (SMB), if essential mission critical business data can not be restored in a timely manner, the impact to the company may be catastrophic.

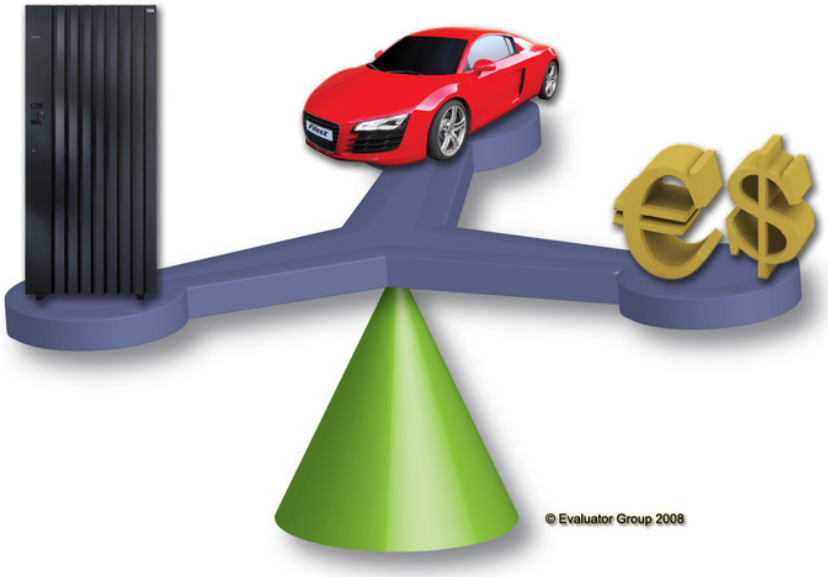

Figure 1: Balancing Capacity, Performance and Cost

Meeting these challenges requires a delicate balance between goals, including financial considerations. Companies must weigh the necessary business requirements for application resiliency with an increasing demand for data storage, all while managing costs.

Any company seeking to minimize the risk of data loss or corruption must have a data protection framework that delivers on the most important aspect –data and application availability. An effective data protection plan will address natural disasters, which have a small probability for

occurrence. A robust and comprehensive disaster recovery (DR) plan will take into consideration other highly probable situations, such as those caused by human error, or other scenarios that may delete or corrupt data. The focus of any DR and business continuity strategy should be the ability to restore data and restart critical applications, quickly and cost effectively, regardless of the cause.

With traditional solutions, there is a direct relationship between data protection service levels and cost – the greater the level of protection, the greater the cost. Costs include hardware, software and implementation, and the costs of managing the DR processes. Costs are compounded when data resides in distributed or remote offices locations; typically site replication and failover is cost prohibitive in SMB environments.

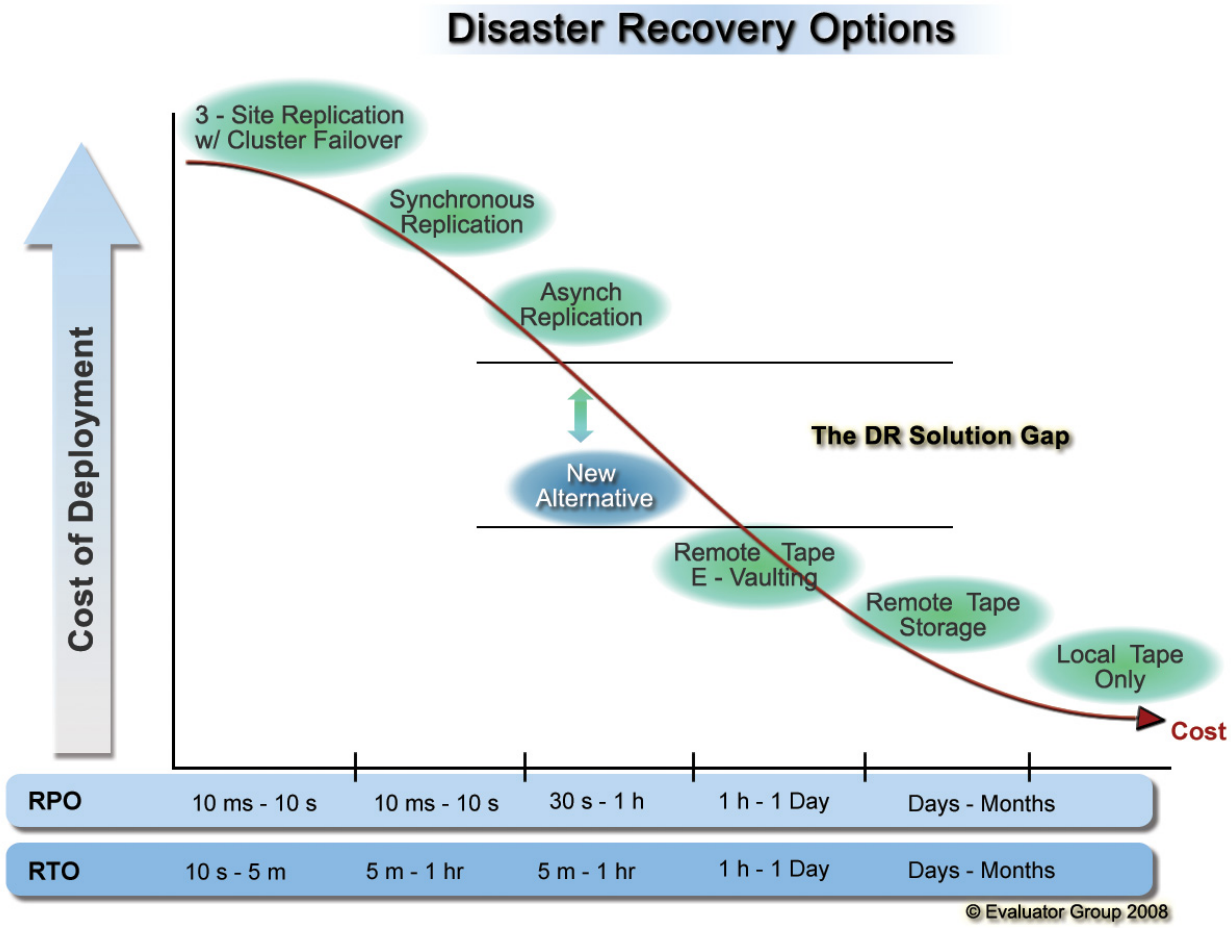

Figure 2: The DR Solutions Gap

As indicated in Figure 2, there is a trade-off between cost and the two primary service levels. For disaster recovery, the RPO (Restore Point Objective – the amount of data at risk since the last backup); and RTO (Restore Time Objective – the amount of time to restore operations) are the primary means of measuring service levels. With current technologies, there is a gap in the solution space, leaving open a wide range of potential solutions. Continuous Data Protection (CPD) disk-based solutions are part of a new breed of solutions that can address this gap, and deliver cost-effective alternatives for companies that require high levels of application availability, and minimal loss of data at cost points dramatically lower than alternatives.

New Generation of Data Protection

With the advent of low-cost, high-capacity disk drives, a new generation of data protection technologies has emerged; including CDP, Snapshots, Disk-to-Disk (D2D), Disk-to-Disk-to-Tape (D2D2T), Remote Office Backup Consolidation (ROBC), Wide Area File Services (WAFS), and Virtual Tape Library (VTL) to name a few. These technologies may be used to supplement or replace current processes. However, their advantage lies in their ability to dramatically increase business RTO and RPO service levels while controlling costs through reduced management overhead coupled with operational efficiency.

These new solutions must still meet the business requirements of a comprehensive data protection strategy, able to meet the restoration needs in the event of: Catastrophic or accidental data loss, Data corruption, and Server or application failure. Solutions must also take into consideration varying protection policies for centralized and distributed systems for different data types, storage utilization, network bandwidth limitations and management issues to be able to reduce costs.

How CDP Addresses the Solution Gap

Recently, as the costs of disk storage have decreased, the industry focus is shifting from traditional tape solutions to those that incorporate disk to improve the costs and efficiency of disaster recovery. CDP has been in the spotlight as an integral part of such solutions filling the DR Solutions Gap as a means to satisfy the increasingly stringent Restore Time (RTO) and Restore Point Objectives (RPO) necessary for critical business applications.

One of the most important differences between CDP and traditional backup and restore technologies is the ability to provide a nearly infinite number of recovery points. Thus, the RPO may be any point in time that provides the view of data desired. With rolling disasters, a problem in one area can create corruption or inconsistencies in other data sets.

A CDP block-level, snapshot, incremental-forever data capture solution provides the ability to meet these requirements by being able to quickly backup and restore data objects regardless of type, amount, location, or any point in time. It allows a comprehensive and flexible data protection strategy to meet the needs of individual applications, central and remote offices, and the business as a whole.

The Evaluator Group and Storage Network Industry Association (SNIA) agree on three main attributes of a CDP solution:

1. data changes are continuously captured or tracked
2. data changes must be stored in a different location than the primary data
3. subjective Restore Point Objectives

Types of CDP

CDP solutions come in block-based, file-based or application-based flavors. Regardless of the specific approach, CDP-based solutions are capable of providing high levels of RTO and RPO while offering better data resiliency through varying granularities of restore points and objects. Block-based CDP is based on a time ordered sequence of block writes of logical devices occurring across the storage network. These solutions provide a view of a past point in time offering high transparency without any application performance effects. The down side with

block only CDP, is that only full logical volumes may be restored. Database objects, application data and individual files cannot be restored individually.

File-based CDP operates by capturing file-system changes and associated meta-data on the file level depending on user defined policies. Policies can be set on a file or logical group which allows differing rules based on the recovery requirements for specific files. Some files may not require CDP with others having more stringent resiliency needs. Restoration of data in file-based solutions is typically faster because unlike block-based which require the volume level view of the recovery point, the user can select and restore from the file's backup instances allowing for a more traditional recovery scenario.

Similar to file-based, application-based CDP offers protection of the application file(s) leaving the application responsible for the journaling necessary to recover from a specific backup point. This allows for greater granularity in recovery, i.e. a specific message in a mailbox or database tables, and additionally allowing a richer set of points in time from which to recover.

A past issue with CDP has been the inability to recover files or application data that was in-use at the time of the data protection. However, there is a new generation of CDP that is able to overcome these past issues.

Regardless of what disaster may cause data unavailability, Xpress Restore, part of IBM's TSM FastBack, provides robust advantages over other CDP solutions by being able to offer data protection for:

- any type of data
- any amount of data
- any location
- any point-in-time

Xpress Restore is a new generation data protection and recovery platform that protects application and other data through block-level, incremental-forever data capture in a non-disruptive manner. TSM FastBack Xpress Restore uses a policy engine that allows data protection strategies to meet the needs of individual applications, central and remote offices, and the business as a whole. TSM FastBack's policy based CDP and/or snapshot scheduled backups offer the ability to adapt, depending upon the business needs, allowing 24 hour operation of systems in mission critical Windows IT environments.

Remote and Branch Office Challenge

This emphasis on quickly restoring data is the key to maintaining mission and business critical applications. By delivering high RPO and RTO levels with lower costs, businesses gain a competitive advantage through increased productivity and performance, while meeting retention mandates. Increasing amounts of corporate data residing in distributed environments such as Remote Offices or Branch Offices (ROBO) compound the data protection issue. Often, remote offices do not have a resident technical staff, or the bandwidth to support the requirements of user or application needs for effective backup and restore. This lack of resources often leads to an inability to correct problems before they lead to loss of data, or reduced service levels.

The common reliance on tape backup is an even greater issue in ROBO environments. While systems may be reliable, data protection processes may not be. Thus, traditional solutions coupled with a lack of dedicated remote staff may result in lower service levels or irrecoverable

data loss. These problems grow as the amount of data and number of systems and sites increases.

ROBO environments also may have different applications that require different data protection strategies. TSM FastBack Xpress Restore enables single or multi-site environments to fully automate and centralize the management of all backup and recovery processes.

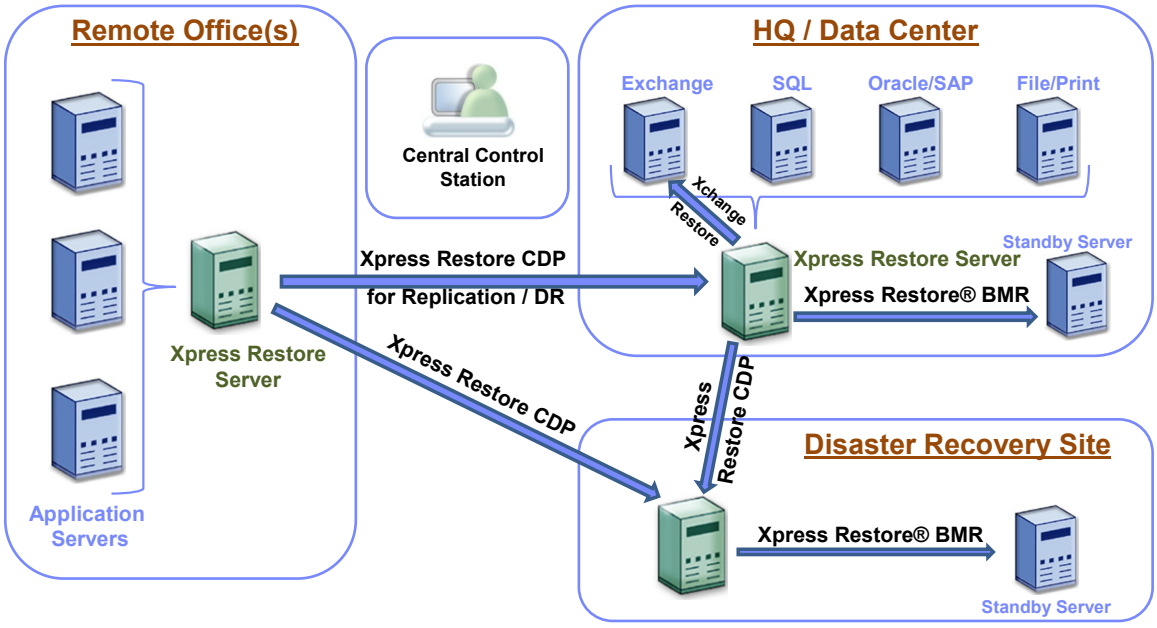

Figure 3: TSM FastBack Xpress Restore Deployment (Source: IBM)

The use of TSM FastBack allows administrators to eliminate tape at remote sites, while maintaining DR capability. Additionally, the unique application integration provides high RPO and RTO levels with lower costs than slower remote tape and vaulting solutions. It also integrates with existing tape backup by utilizing intermediary disk (D2D2T) providing significantly faster backups and recovery, without having to shutdown applications.

Fast Restore is the Key

The speed of backup and restore is the key to the TSM FastBack CDP solution. By implementing incremental forever and block-based backup technology coupled with appropriate protection policies, backup windows are essentially eliminated. After completing the initial full backup, Xpress Restore captures only changed data, as specified by appropriate policies, at the block-level as they are written to disk and builds a logical full copy of data to be protected. This methodology minimizes storage for backup and along with multi-threading, file bundling and compression also significantly reduces the amount of data that must be transported over a WAN to another location.

With the Xpress Mount feature in TSM FastBack's Xpress Restore, administrators can select a point-in-time view of the protected volume and drag-and-drop individual files – effectively seeing a file system view of the data whenever a restoration is desired.

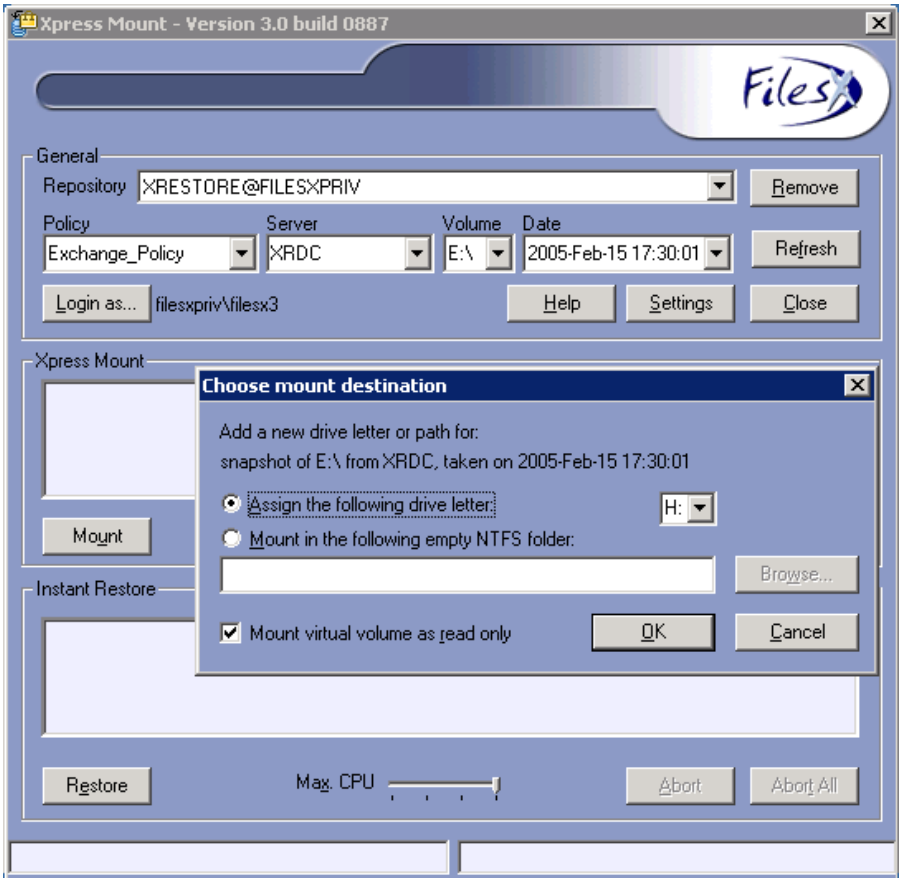

Figure 4: TSM FastBack Xpress Restore Deployment (Source: IBM)

IBM's Xpress Restore provides an instant restore capability that allows users and applications to be running in minutes while full recovery is performed in the background. In the event of an application restore, the application can be brought back online while the restore process is taking place in the background. If an application requests data not yet restored, the restore process prioritizes that data and restores the requested data. Any writes to un-restored data are buffered in cache to be written after the restore has completed.

As a result, Xpress Restore is able to deliver a small recovery time, providing extremely high RTO service levels at a relatively low cost. This type of solution is able to address the gap in DR solutions that exists with most competing solutions today.

Utilizing the policy based backup together with the Xpress Mount feature, which provides an Explorer type view of data in the backup repository, selection and restoration of a file is intuitive, eliminating the need for manual and time-consuming full volume restores. Restore time can be further reduced by allowing individual users the capability to restore their own data objects. Xchange Restore can be used to recover individual mail messages, contacts and other Exchange objects, even from corrupted Exchange Databases. The Bare Metal Recovery (BMR) solution allows server recovery in less than an hour since the OS and data files can be moved between different physical and/or virtual servers.

Case Studies

Case Study1

A company with multiple remote offices, over 100 Windows servers, a critical 24x7 application running on an Oracle database, Exchange 2003 with almost 1,000 mailboxes and a 3.5TB Fibre Channel SAN was experiencing problems with availability, meeting backup windows and satisfying stringent RPO and RTO objectives for their Oracle application, and other business critical IT applications.

They searched for a solution that would offer resilient data protection across the enterprise while meeting their RPO and RTO objectives. TSM FastBack Xpress Restore was selected and installed on their most critical systems, and installed at a remote facility for added DR protection. Prior to using TSM FastBack, traditional tape backup procedures were used which required critical applications to be shut down for up to 2 hours to perform backups. When the Oracle server failed, three days were spent attempting to recover from tape backups. Exchange mailbox recovery was also a major problem and in many cases was not performed due to the complexity and time required.

After TSM FastBack was installed, the backup is now accomplished in less than 5 minutes without having to quiesce applications. When a restore of their Oracle application was required, using TSM FastBack with Xpress Restore allowed the application to resume operations in three minutes. Recovery of individual mailboxes with TSM FastBack is a minor task which can be delegated to selected users eliminating labor intensive recovery procedures from tape.

TSM FastBack has significantly strengthened the company's DR capabilities and recovery is "no longer a worry." The next challenge is a massive consolidation. The ability of TSM FastBack to restore to a completely different system will be utilized in the server and storage consolidation project.

Case Study2

In another instance, a multi-national company with multiple locations faced challenges protecting their ERP applications on SQL Server. Their continued growth and the need to protect data in multiple countries with the requirement to "maintain fast and reliable access to the data at all times" with a centralized DR created unique challenges. After considering a number of options to fulfill these requirements, the company chose TSM FastBack for their data replication and recovery needs.

TSM FastBack was able to provide the mechanisms necessary for continuous protection as well as frequent and scheduled backup requirements offering a "lights-out" solution for the remote offices. The software is also utilized to replicate the remote backup to central sites for "warm" recovery and failover. Because of the company's limited internet bandwidth, TSM FastBack's selective replication technology is paramount in providing necessary data protection without impacting other applications.

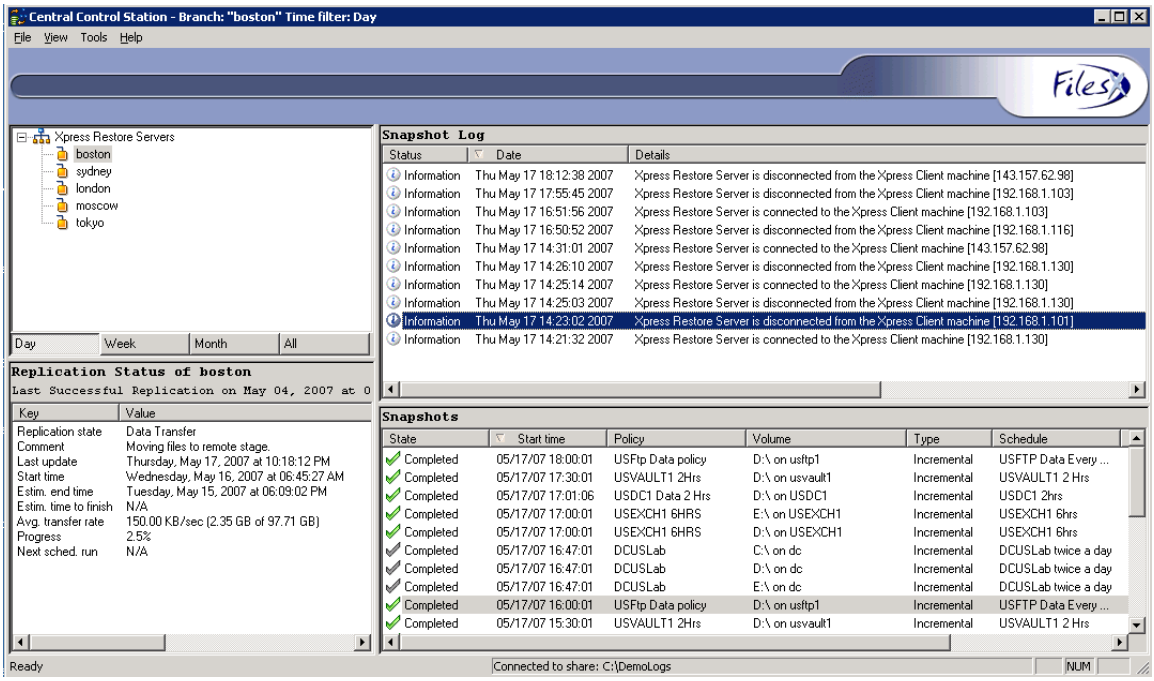

Figure 4: TSM FastBack Monitoring GUI (Source: IBM)

Xpress Restore® with TSM

IBM is incorporating TSM FastBack technology within the Tivoli Storage Manager (TSM) product line. Currently, TSM provides backup, archival and recovery management for file data only. A TSM feature is the flexibility to define various backup, archive and DR 'pools' for data based on business needs and policies. TSM utilizes intermediary disk for quick backup and restore utilizing the database and disk data for creating policy based archive pools.

TSM's policy and incremental forever based data protection compliments TSM FastBack application level CDP technology. The policy and pool concepts of TSM deliver flexibility for disaster recovery, offering backup and archive support for tiered storage media. Business based policies for backup and retention can dictate whether the data is backed up online, nearline or to offline pools. TSM maintains a robust relational database and recovery log of metadata for files that are backed up allowing the ability to easily restore the most recent, or last known good copy of the file. This metadata is also utilized for efficient tape management and reclamation. The incorporation of TSM FastBack technology into TSM's incremental forever pool based backup methodology will enhance the functionality and capability of TSM on the application level and allow greater granularity of the restore functionality of TSM by being able to restore data objects such as individual mailboxes, SQL tables, etc. in the Windows data center environment.

TSM FastBack also adds the ability to manage the protection and consolidation of data in remote and branch offices, providing both local recovery options, as well as incorporating remote data into the TSM storage hierarchy for long-term retention, governance and compliance.

Benchmark Results

TSM FastBack customers provided benchmark information, comparing backup and recovery times between Xpress Restore and their previous full backup-to-tape solution, demonstrating the restore granularity and the vast reduction of restore times.

B / R Operation	Tape	TSM FastBack	Comments
Backup File/Print Server	16 hours	3 min.	Increased frequency from 1/wk to 6/day
Backup E-Mail Server	46 hours	2 min.	Increased frequency from 1/wk to 8/day
Recover Individual File	25 min.	3.5 min.	
Recover Individual E-Mail	Impossible	15 min.	Same for calendars, contacts, tasks, etc.
Restore File Server Volume	10-16 hrs.	6-10 min.	
Restore E-Mail Volume	35-45 hrs.	6-10 min.	

Table 1: The DR Solutions Gap

Xpress Restore is able to effectively meet the broad range of data protection and recovery requirements, by offering:

- local disk-based protection and recovery
- non-disruptive backup – no need to shut down critical applications
- quick granular recovery which can be done in the background not effecting applications
- backup consolidation to a central location and/or secondary site
- centralized monitoring of all backup processes
- seamless integration for traditional backup tape and archiving of data
- regulatory compliance for data protection and archival

TSM FastBack - the Fulcrum for Balance

Regardless of size, today's businesses "Run on Information." The challenge for IT is to provide a balance of storage capacity, data protection and flexible cost options for delivering the RPO and RTO levels necessary to meet business demands.

Many organizations are able to leverage the TSM FastBack application CDP capabilities to provide comprehensive data protection for the entire organization. Some of the unique features delivered by TSM FastBack include:

- Mission-Critical Application Protection and Recovery
- Granular E-Mail Object Recovery
- Off-Site Disaster Recovery
- Business Continuity and Resiliency
- Remote Office Data Protection
- Remote Office Backup Consolidation

IBM's TSM FastBack offers a disk-based adaptive policy backup solution capable of integration with existing storage and tape backup systems, expanding the scope and granularity for recovery. Organizations deploying solutions utilizing TSM FastBack were able to improve the operational efficiency and uptime of production servers by not having to quiesce applications for backups, effectively eliminating backup windows.

More importantly, using application capable CDP provides nearly instant recovery of data objects, application content, files, disk volumes and entire systems with unsurpassed capability and performance to be back up and running within minutes instead of hours or days – without multiple products or complex procedures.

Today's businesses require IT organizations help improve their bottom line while meeting application availability and disaster recovery requirements. In real-world deployments, IBM TSM FastBack customers experience a reduction in Total Cost of Ownership for their DR solutions. This is accomplished by leveraging existing investments in storage and tape systems, reducing disk capacity and network bandwidth requirements, and most importantly improving the efficiency and speed of the data protection and recovery processes. IBM TSM FastBack coupled with Tivoli Storage Manager, provide the ability to achieve the delicate balance between competing goals; providing the necessary data protection and **restore** capability for business data, with minimal downtime and lower costs.