

CentricStor in a TSM Environment

Issue September 2004

Pages 4

CentricStor

Improving Performance and Reducing Costs in a TSM Environment

There are times when the whole is greater than the sum of the parts. Together, Tivoli Storage Manager (TSM) and CentricStor present just such a proposition. TSM brings capabilities for organizing data. CentricStor increases performance and reduces cost. Together they make an unbeatable storage solution. Many of the unique functions of TSM are enhanced with CentricStor. CentricStor relieves TSM of some data caching functions and performs them faster while generating savings in overall storage resources.

“There is a simple statement that is an undisputed fact that no one individual or company can dispute: Backups do not matter. What does matter is what you can restore. No one wants to make backups. They make backups because they have to. Either government regulation requires you to save data such as payroll or employee information for a specific length of time, or you need to save data to support your business.” (IBM Tivoli Storage Management Concepts; IBM.COM/Redbooks; July 2003, p56)

IBM promotes Tivoli Storage Manager to satisfy user business requirements in the areas of storage consolidation, data sharing, data protection, and disaster recovery. CentricStor enhances each of these requirements while improving recovery, minimizing media reclamation workloads, and supporting aggressive data management through archiving -- all while delivering significant savings.



There are at least seven functional areas where CentricStor will benefit TSM implementations.

Replacement for TSM Disk Pools -- CentricStor can be a direct replacement for TSM "disk pools" (commonly called TSM disk cache in non-TSM circles). Disk storage pools are used to maximize the number of concurrent backup sessions temporarily storing data until removable media can be used effectively. CentricStor automatically eliminates tape drive contention and drive performance issues because it presents a nearly unlimited pool of virtual tape drives. In the electronic vaulting configuration, replacing disk pools with CentricStor expedites the process of getting backup data into protected, offsite storage. CentricStor creates multiple copies of tape volumes by policy without consuming server resources. Physical placement of the volumes created may be managed by CentricStor into a remote location.

CentricStor data flow is similar to TSM: Backup data is stored on disk before being moved to physical tape. However, CentricStor compresses data as it is being stored in the virtual tape cache. If the data can be compressed at 2:1, then only half the disk space is required and subsequent operations occur at twice the speed.

Reduce TSM Reclamation process -- TSM moves backup data to new media continuously in an effort to fully utilize physical tape cartridges. This activity can be overwhelming in large environments. Depending on disk pool sizes, tape volume size, and the number of tape drives, some customers' tape reclamation activity can't finish before the next backup window starts.

When configuring CentricStor for TSM, it eliminates the need for TSM's reclamation process for ONSITE volumes. CentricStor's smaller 2 GB logical volume size is sufficiently small that "mostly" normal attrition would cause volumes to become "empty." Analysts recommend the reclamation threshold should be set to 100, then periodically query TSM's data base for volumes with less than 25% utilization. Most tapes become empty without the need for TSM's reclamation.



CentricStor has a perpetual background policy-controlled process that reorganizes physical media eliminating expired and scratched virtual volumes. This process provides maximum media utilization. CentricStor also performs automatic media refreshing by reading every tape after six months of inactivity and rewriting the data to a different tape to ensure that data is continually available.

More LAN-free TSM Connectivity -- In traditional TSM deployments, TSM Clients send backup data over an IP network to the disk pool attached to a TSM Server. TSM Servers then send client data from their disk pools to physical tape when the resources are available. This is an inefficient way to move large amounts of data (even with faster networks) since IP packets are relatively small and collisions of packets are plentiful on busy networks.

CentricStor allows users to convert large TSM Client servers into TSM Agents with their own connections to (dedicated, virtual) tape drives. As TSM Clients are converted into TSM Agents and more data is transferred across the SAN, the need for dedicated IP network capacity and TSM disk is reduced. Customers that have extensive networks to support TSM can now redeploy this expensive resource.

Dynamic multithreading -- TSM Release 5.2 added performance improvements described as dynamic multithreading that allows simultaneous data transfers to occur from TSM Clients to TSM Servers. Multithread performance claims to improve performance by a factor of 3:1.

Multiple data streams require access to multiple volumes to store the data. CentricStor presents a large number of volumes and devices to TSM servers. Optimizing the deployment of multithreading can best occur with the large number of storage devices available with CentricStor.

Migration -- In TSM, migration is the process of moving data between a hierarchy of storage pools including storage on disk and tape. Migration is a server-intensive application that can move data across various storage locations reducing storage costs for infrequently accessed data. Depending upon TSM policies in place, migration logically delivers data to magnetic tape as a final repository of data.

CentricStor supports migration simply by having TSM make a single data migration from original disk subsystems to virtual tape. This single step is supported by fast retrievals as the data will remain in virtual tape cache. CentricStor always copies data from virtual tape cache onto physical media ensuring the lowest possible cost for storage while retaining an instance in cache for performance.

Copy Pools and Export of OFFSITE tapes – A TSM feature makes simultaneous copies of data from a primary storage pool on disk to secondary ‘copy pools’ on either disk or tape. One of the purposes of copy pools is time-optimized simultaneous creation of redundant copies of tapes for data availability and for export to vaults for disaster recovery.

CentricStor also performs policy-managed creation of multiple copies of physical tapes from single writes of server data into virtual tape cache. In TSM environments where disk pools (primary storage pools) have been eliminated it is critical that multiple tapes exist for recovery. In a disaster scenario, recovery is expected to be performed from OFFSITE tapes. TSM can recover from either ONSITE or OFFSITE tapes. CentricStor tape ‘cloning’ will perform the same functions as TSM copy pools, but without delays waiting for tape resources.

Drive sharing -- TSM currently does not support non-IBM drive sharing or pooling. Third-party applications (such as Enterprise DistribuTape from Gresham Software) have added function to TSM to enable resource pooling enhancing TSM deployment into non-IBM tape environments. TSM Release 5.3 is expected to include Enterprise DistribuTape functionality to simplify licensing and acquisition of software.

With CentricStor, each TSM Server and Agent can have as many virtual drives as it needs. Virtual drives eliminate a sharing requirement because they are so abundant. More drives operating in parallel maximize throughput performance and simplify operation. Keep in mind, however, that Gresham Software is currently needed with CentricStor for other reasons including multiple TSM Agents’ access to a common scratch pool of tapes. Utilities exist to synchronize scratch tapes between the TSM tape catalog and CentricStor ensuring data accuracy and maximum performance.

Fast Backups; Faster Recoveries -- With CentricStor, TSM Clients and Agents send their data directly to (virtual) tape, bypassing the disk pool, improving overall TSM performance. During backups, mount times for tape are reduced to less than one second. During restores, CentricStor eliminates physical tape mounts enabling very fast recalls from cache.

Planning is important – All TSM documentation stresses planning, planning, planning. Deployment of CentricStor will require a review of several policies and parameters in the TSM application. Some examples: (1) In the export of physical volumes, set read-only (R/O) for LV's. When sending cloned tapes offsite, the counter-part onsite logical volumes should be marked read-only to TSM ensuring TSM will not add data to the onsite volumes. Therefore the data contents will match the offsite volumes, so a disaster recovery event (ie, running IMPORT for any reason) will recover all data associated with those affected volumes. (2) Using TSM's re-use delay properly. Some TSM analysts have exploited a discrepancy between their retention period for TSM db versus the data tapes to recover data files that had been expired but not yet overwritten. Effective use of CentricStor requires using TSM functions properly.

Why did customers decide to move to a CentricStor solution?

A major German health insurance company: Before CentricStor they only had 8 physical tape drives for the TSM Servers and the TSM storage agents (LAN-free backup clients). They used TSM together with Gresham software to switch these 8 drives between TSM servers. This solution never worked error-free as there were always problems during switching. They used STK Powderhorn libraries and HSC with Library Station to control the libraries from TSM. Any problems with the mainframe or the interconnect control software left them unable to backup or restore the open systems data. Solving problems during the switching of the 8 * 9940 units required the help from the mainframe experts to bring the devices back online. They have both main and backup data centers. All servers are installed on both sites. All data (disk and tape) are mirrored between the data centers. If they have a disaster they are still able to run their complete workload from the surviving data center. For open system they were using the TSM tape cloning facility at night to copy tape data between the data centers.

The open systems environment is growing very fast at this firm. They knew that they were unable to handle this situation with a physical backup solution. Because they also knew the positive aspects (financial and technical) of using a virtualization solution from their Mainframes, they decided to get the same positive aspects for their open systems servers.

With the CentricStor solution they got 128 "virtual" drives and assigned fixed drives to all of the servers so they didn't need to switch the drives. Operationally, the TSM environment is completely independent from the Mainframe world and able to control its own solution. The resulting solution includes only 6 LTO-2 tape drives with a reduction in tape cartridges from over 900 to 300. CentricStor reduced the time for scratch mounts to about one second and each of the virtual drives operates at server channel speeds significantly improving performance by a factor of three.



With open systems data growing very quickly, they were unable to complete the cloning of this data during the night. Cloning over night meant they will have a data loss for one day if they had to switch to the second data center. Therefore they looked for a solution capable of doing a synchronized tape backup copy between the data centers without using a server. The insurer concluded that CentricStor is the only solution on the market that provides this functionality.

A large equipment manufacturing company: Before CentricStor there were 64 IBM 3590 Models E/H tape drives mounted in 2 IBM 3494 tape libraries. TSM utilized a 10 TB disk pool on high-performance, primary disk subsystems. After CentricStor was deployed, there are 12 'Jaguar' (IBM 3592) tape drives installed into the 3494 libraries and the TSM disk pools were completely eliminated.

A major insurance firm: Among other modifications this insurance company eliminated the 10TB disk pool when they added CentricStor to the TSM environment. The change made all of the 10TB of disk storage available for storage growth, deferring disk purchases for a period of time.

CentricStor delivers dramatic improvements to TSM environments in both performance and savings. Customer measurements of a three-fold improvements in backup are significant. In the case of manufacturing company above, hardware savings were in excess of \$4 million.

Many of the media-related tasks, traditionally performed by TSM, are performed by CentricStor in ways that enhance TSM's overall capabilities. TSM is excellent at managing data and media. CentricStor optimizes use of physical resources (tape drives and media) and increases performance by delivering disk-to-disk speeds when supporting tape operations. Together these products improve the users operating environment.

CentricStor is positioned to facilitate customers' growth by decoupling physical storage changes from the virtual images presented to TSM. Capacity changes are made transparently to TSM. Expected enhancements to TSM both in Release 5.3 and those requested in later releases will make CentricStor easier to use and more effective. CentricStor is not limited to TSM Backup but also supports TSM Archiving through the same system. Performance benefits from CentricStor to TSM Backup are also available to Archiving delivering even more advantages from a single solution.

© Peak Data, Inc., August 20, 2004