Global Information Assurance Certification Paper

SANS GSEC Practical

Enterprise Storage Management Solution

Peter Kent
April 23, 2003

Version 1.4b (2)

Abstract

With the increased volume of data, and the reduction or nonexistence of backup windows and staff, how do you protect and secure your company's data? This was the problem being faced. There were over 250 Window and 40 UNIX servers spread across Canada and the north east of the US. Of these, there were 60 Windows and 15 UNIX servers in the main computer room for which a backup solution was required first. The data store had increased from Gigabytes to Terabytes and the backup windows had shrunk from 12 hours a day and 36 hours a weekend to less than 6 hours a night and 18 hours on the weekend. This situation had developed over a number of years, and a new solution was required very quickly as a SAN was being expanded and another TB of data was being added. My challenge was to recommend and sell a central backup solution to upper management. Once approved, implement it to backup UNIX and Windows based servers in the main computer room. Future needs would include central backup management for servers in our second computer room and remote servers at our customers' locations.

It was decided by the executive that this was a security issue no different than protecting data on the servers. IT Security would develop the recommendations and oversee the implementation of a new Enterprise Backup Solution. Once in production, IT Security would audit the system to ensure that the company was protected and could recover information when needed, up to and including a catastrophic disaster. An IBM 3584 Tape Library had already been purchased and would be used as part of the solution.

Identifying the Risks

The host operating systems used by the company include Solaris, AIX, N/T and Windows 2000. As there had never been a standard developed for backing up servers, a number of different solutions were used. Veritas NetBackup and Microsoft's N/T backup solution were used for the Windows servers. The backups for the Solaris servers were performed using the O/S utility USF DUMP and AIX employed BACKUP. Only Veritas offered a tape management system, the rest had to be managed manually. As a result, it was very time consuming for Operations to maintain the tape system. As it was manual, it was also prone to mistakes.

As servers were installed, any one of these backup solutions was implemented, and the backup schedules were created by the implementer based on her/his judgment. As a result, there were no standards in place. Without standards or central management, no one could be sure what was and wasn't being backed up. Unfortunately, it was usually when a restore request was made that a deficiency was identified.

In order to support the backups, 45 tape drives were needed. These tape drives were a mixture of internal and external depending on the server. The tape drives used included DLT, IBM 3570, 8MM and 4MM. Due to the number and age of the drives, maintenance costs were high. The tape drives were also unreliable due to the lack of staff to clean and maintain them properly. There were two external tape drives available when a primary tape drive failed, however, with the high failure rate, they were usually in use. Less critical backup would be sacrificed for the more critical ones. As this was a judgment call, at times the wrong backup would be cancelled.

As the backup solutions being used were stand alone on individual servers, there was no central logging. This resulted in the need for the individual server logs to be checked to determine if there were any problems. Due to lack of staff and the number of servers that had to be checked, this was usually not completed until after the backup window had ended. This was a serious issue, as problems were not identified until it was too late to rerun the backup. On occasion, requests to restore information could not be completed due to problems that, if identified in a timely manner, could have been corrected and the backup rerun with no missing data.

The tape media used for was a mixture of DLT, IBM 3570, 8MM and 4MM. As the servers had their own tape drives, a group of tapes was needed to fulfill the backup schedule and any archiving. There was no central scratch tape pool. In addition, the capacities of these tapes are quite small in today's terms, and resulted in the need for large numbers of tapes. As the result, changes to backups (schedule or number of tapes) could create a shortage of tapes. To resolve the immediate problem, the tapes that held the oldest backup were scratched for immediate use. This reduced the possibility of recoverability of older files.

Tapes are not indestructible. Due to the backup window, and the technology we used, only one copy of a backed up file was retained on tape. Should that tape become unusable for any reason, the backup is lost. As time was limited, another copy of the data could not be made and the only copy that was produced was not verified to ensure it was complete and correct. This resulted in media problems occurring on a regular basis, and restore requests not being fulfilled.

The total available disk on the network was 6TB, and 50% utilized, therefore 3TB of data had to be backed up. The storage capacity on the network was growing quickly. The primary SAN was expanding with the addition of another 1 TB (20% more capacity) of disk. It was expected that the disk would be filled within 4 months. The available window for backing up had been reached and backups were being scaled back to meet the window of availability. This resulted in less protection of data and a slower recovery should one be needed. A faster solution was required.

The backup scenarios used were a combination of full, incremental and differential. A full backup is performed on each server during the weekend. The daily backups were either incremental or differential. Incremental backups save any files that have changed since they were last backed up. The differential backup copies any files that had changed since the last full backup. The advantage of incremental backups is that the least amount of data is saved so the backup window is the shortest. The disadvantage is that recovery of a system required more tape loads and would be slower as the full had to be restored first, and then each day's backup. The advantage of differential backups is that only files that have changed since the last full backup was completed are backed up. In addition, restoring only required the full backup and the differential from the day the restore was needed for. The disadvantage over the incremental is that each day after the full backup there could be more to save and the backup window could grow. In our case, the backup window had been reached so incremental backups were used. This slows the recovery process and was deemed unacceptable for disaster recovery.

The retention of data for backup purposes has always been an issue when it came to user data. The schedule used mainly was 12 monthly rotations, 5 weekly rotations and 2 weeks of daily backups. As there was no way of knowing how often files changed, there was no way of knowing when a file was backed up during the week. To improve the possibility of being able to restore a file, 2 weeks of daily and weekly backups were retained. As the result, any file that was changed in the previous 2 weeks could be restored. There were also 3 more weekly and 12 monthly rotations retained. Problems occurred when users wanted an older version of a file restored that fell between the scheduled rotations. A new method was needed to ensure the correct number of copies existed for each file no matter if it changed daily, weekly or yearly.

Due to the issues identified, it was apparent that a centrally controlled, policy driven backup solution and the tape library were the best alternative for us. This would ensure that data would be retained in a complete, safe and secure manner. It would also allow the backup to complete within the window available and recovery from a disaster in the time frame we required.

Search for a Solution

The first step was to identify the requirements that were needed in a centralized Backup Solution for our company.

A database was created that included all of the servers housed in the primary computer room as there was no complete, up to date inventory available. Once completed, an inventory of each server was added to the database that included the hardware, Operating System, databases, disk capacity and current utilization. The current backup requirements and expected growth for the next year were also added to the database. The Operating systems and databases included:

Operating System - N/T 4, Server 2000, 2000 Advanced, Solaris (2.6 and 2.8), and AIX (4.3.3 and 4.1)

Databases - Sybase, SQL (6.5, 7 and 2000),Exchange 5.5, and Oracle

The current backup process was designed mainly based on the backup window. It was decided that the replacement backup solution had to be primarily designed for the quickest recovery possible and yet complete in the Window available Based on these requirements, a search for the appropriate software was started.

Based on this information, we developed the following requirements for a backup solution:

- Work on servers running one of the O/S's listed above.
- The application had to run on Windows 2000 AS.
- Backup the databases listed above.
- Problems encountered during a backup had to be identified to Operation immediately.
- Centralized exception reporting on the logs of any issues.
- Customizable reporting.
- Policy driven to centralize management.
- Centralized scheduling of backups.
- Reduce uniqueness of backups to reduce the exposure of missed files.
- Backup approximately 360 GB from 75 servers in 6 hours.
- Allow manual overrides in the schedule.
- Multiple copies of a backup in case of corruption or destruction.
- To start a recovery in less than 30 minutes of Operator labor.
- To be able to restore the current version of a file 99.9% of the time.
- To reduce Operation's labor cost for backup and recovery by 20%.
- To backup the designated servers on time 99.9% of the time (not including communication or client server problems).
- Automation tools for running scripts.
- Single media for backups.
- Efficient tape media usage.
- Disaster recovery for the backup solution.

Other items to keep in mind for the future were:

- Bar metal restore capabilities
- The ability to backup images of servers
- The ability to backup remote servers

Knowing what was required; Gartner was contacted to assist in identifying the top two backup solutions available that could meet our requirements. Gartner

advise, in a telephone conference call, that Veritas NetBackup and IBM's Tivoli Storage Manager (TSM) were the leaders. Our company already had a good relationship with both Veritas and IBM so both companies were contacted and information on their products requested. In addition, a search was undertaken on the Internet to gather information to backup the idea that these were the best 2 alternatives.

I used the Internet to access the Veritas and IBM Tivoli sites to determine which package would best meet our requirements. Both packages met our requirements with the exception that Veritas didn't manage tape usage. This was not a critical concern; however ongoing cost was always an issue. Further research was needed to determine which system would provide the requirements needed while using the least amount of tapes.

I researched both packages using Google as the search engine, looking for any comparisons between TSM and Veritas. On the Network Computing: Tech Library site, there was a comparison. It appeared both systems were very similar and met most of our requirements. However there were 2 items TSM had that were of interest, TSM had a Web GUI and it was installed at 80 of the Fortune 100 companies. On the Progressive Strategies site, it was shown that based on 1 month of testing that TSM was 36% faster at backing up, 66% faster in restoring and used 25% of the tapes needed by NetBackup.

Upon further research I found that another difference between NetBackup and TSM was the methodology used for backups. NetBackup used the traditional full weekly and daily incremental backups where TSM used progressive incremental backups with versioning. This meant that only files that changed since they were last backed up with TSM were saved. As the result, the smallest amount of data would be saved each night. However, the first time a library is saved; all its contents are backed up. As the servers run 7 days a week, lower bandwidth requirements for TSM was of interest. In addition, fewer tapes and less backup time would be required.

Versioning addressed the problem of how many tape rotations should be kept. In the past, tapes were kept for long periods of time to increase the probability that restore requests could be fulfilled. TSM can be configured to retain a number of versions of an individual file. As the result, no matter if a file changed daily or yearly the number of copies designated would be retained. TSM can also manage the utilization of tapes by moving active file copies to different tapes and freeing up less utilized tapes. This meant fewer tapes would be required.

TSM offered the functionality needed to reduce the risk of media problems. When a server is backed up, the data is sent to a storage pool on the TSM server. As the result, the transfer rate would be 10 times faster than it currently was. Once all the backups were completed, the information was streamed to tapes that would go offsite for disaster recovery. Following this, the data that

existed in the storage pool was also copied to the IBM 3584 tape Library in the computer room. As the result, there would be a complete copy of the current backups onsite should recovery be required, and a full copy offsite in case of disaster. Should a tape become unreadable, TSM would automatically request a backup copy and create a new copy on the tape library.

The tape library protects tapes as they are sealed in a compartment where the robotic arm moves the tapes as needed. In addition, the tape drives are also contained in the same compartment. As the result, there are less contaminants getting to the drives and tapes. This will reduce problems with both. The cleaning of the drives is done automatically.

TSM offered one other function that was not considered in this phase of the project, but was of interest as a future enhancement. This facility allows TSM to manage the contents of hard drives. It can be configured to move files from disk to tape after a designated time of non-use and leave a marker in the directory. When a user tries to access the file, transparently to the user, the file is restored through TSM to the directory for use. The estimate on wasted space due to obsolete files is 30%, managing this could postpone a purchase additional disk in the future.

The SAN can be backed up directly to tape by TSM. In the case of large volumes of data (more than 100GB) TSM can bypass the shared pool of disk and copy straight to the library. This is a significant advantage to us as we will backup the LAN servers to the disk pool and the SAN directly to tape at the same time. Once both backups are complete, a copy will be generated for offsite storage. The LAN backup will then be streamed from the disk pool to the tape library as the onsite copy.

When TSM writes the backup to the tape library, it will stream using as many tape drives as is allocated. The individual tape contains small snippets of data from each system in the storage pool. It does this to increase the throughput but a second advantage was identified. The tapes are unreadable without a copy of the catalog created when it was backed up. As the catalog will not be kept with the tapes, an additional layer of security has been created.

In order to reduce the time spent by Operation monitoring the backups, an alert process was needed. The company was currently using Tivoli Tec Monitor to monitor for network problems. TSM is a member of the Tivoli family and integrates with Tec Monitor easily. As Tivoli experience had already been developed in-house, TSM was attractive.

The next step taken was to bring in each product and evaluate them in our lab. The requirements were the basis for the test plan. Each evaluation ran for 2 weeks so not all features could be confirmed such as tape usage. Operators were used to do the testing as Operations would be the owner of the system. It

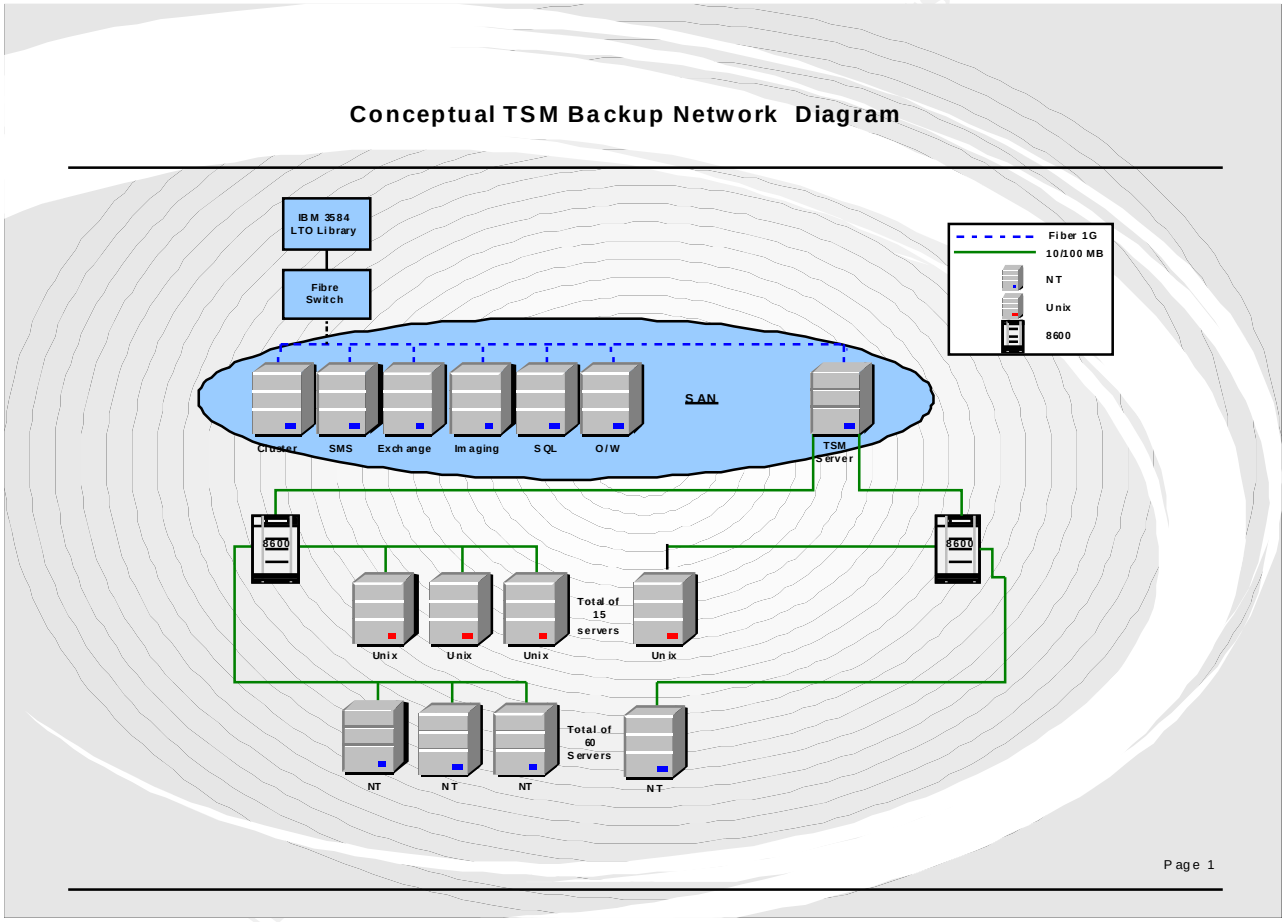
was important to get their feedback on what their impressions of the systems were. They found that both products work very well, however, TSM's GUI was extremely intuitive, if you understood a directory tree structure, you could perform manual backups and restores with no training.

Based on the information gathered, IBM's Tivoli Storage Manager was chosen. Though the research, it became apparent that the initial implementation of TSM was complicated. It was therefore proposed that IBM would be contracted to install and configure the product as there was currently no TSM expertise available in the company. The formal training of the staff would take place 6 weeks after the product was installed so the staff would be familiar with the system, and the training more beneficial.

Result

The proposal was accepted by the Executive. The equipment has been received and the IBM consultant will be arriving in 1 week install TSM and configure the policies.

The following is a conceptual diagram of the Network that will support TSM:



Upon the completion of the implementation of TSM in our primary computer room, the security of the data for which Operations is the custodian, will be protected significantly better than today. The ability for Operations to manage the backing up of all data in the computer room will be handled through a single master server. Through scripting and the scheduler, the only manual tasks will be to remove the tapes that are to go offsite, and respond to any alarms sent to our network monitoring system.

The estimated reduction in labor will be reduced by 25% based on the results of the evaluation. We anticipate that backups will be completed within the new backup window, based on the fact that a minimum of 10 servers will be backed up at one time through the network.

There were 45 tape drives required to backup the servers. These were unreliable. With the new system' only 5 tape dives are needed in the tape library. As TSM can use all the tape drives when generating tapes, the approximate throughput will be 930 GB per hour. As our total backup requirement will be just over 2TB, we will not have a problem with tape throughput.

Tape management will be performed by TSM and when the scratch poll gets below an acceptable number, an alert will be sent to the Operators monitoring system requesting more tapes be ordered. Operations will no longer be able to scratch active tapes early due to tape shortages. As the systems will now use the same media for backups, it will be easier to do capacity planning and identify order points for the year.

The tapes will be unusable by anyone other than those authorized in Operations to read them. This is a much higher level of security than we have hade in the past, and has been a concern to IT Security.

With the use of standard policies that will backup everything on the servers with the exception of the Operating system, temporary files and swap pages, any additions to the server will be captured by the backup. We were not concerned with backing up none volatile data as if it does not change, it will only be backed up the first time. As the result, the assurance of catching any changes outweighs the cost of small amount of extra tape media needed.

As we are also working on a new Disaster Recovery plan, the new tape solution will be an asset. By using TSM, we can restore critical systems more quickly as the Library contains all of the information and is on-site. In the case of a catastrophic disaster where the systems and the backup solution are destroyed, we will use another feature in TSM called collocation. This allows the information for specified servers that would normally be spread over a large number of tapes, to be copied to a small number of tapes that can be managed manually if required. This will allow critical servers to be recovered in the shortest time frame. It will also allow time to recover the TSM system for the restoration of non-critical servers. The other benefit is that a single Operator can manage the whole restoring process on multiple servers.

The final result is that the system is extremely scalable. The server purchased can be expanded by doubling the CPU's, and RAM. The tape library can be expanded to handle up to 992 TB. The number of robotic arms can be increased to 72 drives. Therefore, this system will have a life expectancy of at least 10 years in our operation.

Based on the evidence provided, security has been advanced significantly in the area of data protection.

References

"Automated Disk-Based Data Protection for Heterogeneous Systems". Nov. 2002. URL: http://www.netapp.com/solutions/partners/tivoli.html (14 Jan. 2003)

Spiner, Dan. "IBM Tivoli Storage Manager Demonstrates Critical Advantages Over NetBackup" 11 Nov. 2003. URL: http://www.progstrat.com/press/TSMBSRPR.htm (17 Jan. 2003)

Norris, Craig and Cohen, Barry. "Comparing IBM Tivoli Storage Manager and Veritas NetBackup in Real-World Environments". 25 Oct. 2002. URL: http://www.progstrat.com/techforum/prodreview.html, Product Reviews # 5, (17 Jan. 2003)

"Veritas Backup Exec for Windows Servers – Technical Information http://www.veritas.com/products/listing/ProductTechnicalInfo.jhtml?productId=bews, (10 Jan. 2003)

"IBM Tivoli Storage Manager", URL: http://www-3.ibm.com/software/tivoli/products/storage-mgr/, (10 Jan. 2003)

Network Computing: Tech Library http://networkcomputing.telezoo.com/asp/sc/ic.asp?idCatitems=10347&idCatitems=9874&idCatitems=9947&idCatitems=9886&view=long&history=^709^764^773, (18 Jan. 2003)

Last Updated: July 21st, 2017

Upcoming Training

SANSFIRE 2017	Washington, DC	Jul 22, 2017 - Jul 29, 2017	Live Event
SANS Prague 2017	Prague, Czech Republic	Aug 07, 2017 - Aug 12, 2017	Live Event
SANS Boston 2017	Boston, MA	Aug 07, 2017 - Aug 12, 2017	Live Event
SANS New York City 2017	New York City, NY	Aug 14, 2017 - Aug 19, 2017	Live Event
Community SANS Omaha SEC401*	Omaha, NE	Aug 14, 2017 - Aug 19, 2017	Community SANS
SANS Salt Lake City 2017	Salt Lake City, UT	Aug 14, 2017 - Aug 19, 2017	Live Event
SANS Virginia Beach 2017	Virginia Beach, VA	Aug 21, 2017 - Sep 01, 2017	Live Event
SANS Chicago 2017	Chicago, IL	Aug 21, 2017 - Aug 26, 2017	Live Event
SANS Adelaide 2017	Adelaide, Australia	Aug 21, 2017 - Aug 26, 2017	Live Event
Community SANS Trenton SEC401	Trenton, NJ	Aug 21, 2017 - Aug 26, 2017	Community SANS
Virginia Beach 2017 - SEC401: Security Essentials Bootcamp Style	Virginia Beach, VA	Aug 21, 2017 - Aug 26, 2017	vLive
Community SANS Pasadena SEC401 @ NASA	Pasadena, CA	Aug 23, 2017 - Aug 30, 2017	Community SANS
Mentor Session - SEC401	Minneapolis, MN	Aug 29, 2017 - Oct 10, 2017	Mentor
SANS San Francisco Fall 2017	San Francisco, CA	Sep 05, 2017 - Sep 10, 2017	Live Event
SANS Tampa - Clearwater 2017	Clearwater, FL	Sep 05, 2017 - Sep 10, 2017	Live Event
Mentor Session - SEC401	Edmonton, AB	Sep 06, 2017 - Oct 18, 2017	Mentor
SANS Network Security 2017	Las Vegas, NV	Sep 10, 2017 - Sep 17, 2017	Live Event
Community SANS Albany SEC401	Albany, NY	Sep 11, 2017 - Sep 16, 2017	Community SANS
Mentor Session - SEC401	Ventura, CA	Sep 11, 2017 - Oct 12, 2017	Mentor
Community SANS Columbia SEC401	Columbia, MD	Sep 18, 2017 - Sep 23, 2017	Community SANS
Community SANS Dallas SEC401	Dallas, TX	Sep 18, 2017 - Sep 23, 2017	Community SANS
Community SANS Boise SEC401	Boise, ID	Sep 25, 2017 - Sep 30, 2017	Community SANS
Baltimore Fall 2017 - SEC401: Security Essentials Bootcamp Style	Baltimore, MD	Sep 25, 2017 - Sep 30, 2017	vLive
Community SANS New York SEC401	New York, NY	Sep 25, 2017 - Sep 30, 2017	Community SANS
Rocky Mountain Fall 2017	Denver, CO	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS London September 2017	London, United Kingdom	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS Baltimore Fall 2017	Baltimore, MD	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS Copenhagen 2017	Copenhagen, Denmark	Sep 25, 2017 - Sep 30, 2017	Live Event
Community SANS Sacramento SEC401	Sacramento, CA	Oct 02, 2017 - Oct 07, 2017	Community SANS
SANS DFIR Prague 2017	Prague, Czech Republic	Oct 02, 2017 - Oct 08, 2017	Live Event
Community SANS Charleston SEC401	Charleston, SC	Oct 02, 2017 - Oct 07, 2017	Community SANS