Global Information Assurance Certification Paper

SANS Security Essentials
GSEC Practical Assignment
Version 1.2e

Centralized Backups
Michael J. Gallagher
July 2001

Introduction

During SANS training one learns the term "Defense in Depth". This phase basically means security at several different levels. This would include defense mechanisms such as firewalls, Virus scanners, Intrusion detection systems, and even backups. One might ask, "What does backups have to do with security". The answer is; if all other systems fail to prevent someone from destroying a company's data that company is out of business unless they can turn to their backups.

Today's reliance on corporate computer systems and the information they contain continues to grow on a daily basis. Corporate systems today also have grown from single Mainframes with terminals to tens, even hundreds of heterogeneous systems spread across the globe. Companies can have Web servers running NT, Unix servers hosting their databases, and Novell File and print servers. Environments like these have brought an end to the days of tape drives attached to every server and standalone backups. Management has a comprehensive understanding of cost related to systems downtime, overwhelming administration of single system backups and the media involved. With these facts in mind, corporations have migrated to the centralized backup scheme, investing sometimes hundreds of thousands of dollars into backup software, tape libraries and training for the staff to implement and manage just such an environment.

Backup definitions

One has to have an understanding of the different type of backups that these software packages use when backing up data. Some packages use the full, differential, or incremental backup schemes such as Veritas Netbackup. Others like Legato's Networker use the different number levels usually associated with Unix ufsdump utilities along with the full and incremental dataset save types. Below are definitions of the different types of backups.

Full Backup - A procedure for backing up all the files on a hard disk by copying them to a tape or another storage medium. It is a good security measure for frequent users to do full backups once a week. http://www.xrefer.com/entry/622843

Differential Backup - A procedure for backing up only files that have been changed or added since the last full backup. Earlier versions of these files will be replaced in a differential backup. http://www.xrefer.com/entry.jsp?xrefid=622612

Incremental Backup - A procedure for backing up only the files that have changed or been added

to a system since doing the last Backup. It is good practice to do an incremental backup from the hard disk onto a tape whenever finish working at a computer system. http://www.xrefer.com/entry.jsp?xrefid=622947

Ufsdump Backup Levels - All files that have been modified since the last ufsdump at a lower dump level are copied to the destination (normally a magnetic tape device). For instance, if a level 2 dump was done on Monday, followed by a level 4 dump on Tuesday, a subsequent level 3 dump on Wednesday would contain all files modified or added since the level 2 (Monday) backup. A level 0 dump copies the entire file system. Valid dump levels are numbers are 0 through 9 http://uwsg.iu.edu/usail/man/solaris/ufsdump.1.html

Who they are and how they work.

There are several software vendors that produce centralized backup software. A few of the popular products today are Veritas Netbackup, Veritas BackupExec, Legato Networker, and Computer Associates ARCserve 2000 Advanced Edition. All of these products work on the same basic principle. That is to backup systems over the network to a server that has some sort of storage device attached. Some of the products support a varying number of protocols but the predominate protocol used by these products is TCP. A centralized backup scheme can contain several different software/hardware modules. We will focus on the 3 most widely used system definitions, the first being the central server. This server controls the entire backup environment. It usually controls indexes, backup schedules, client group definitions and hardware configurations. This server is also responsible for logging problems with any of the backups and reporting them to the system administrator in charge of backups. The second type of server in a centralized backup system would be what is called a media server or storage node. This server is attached to some sort of storage medium for the backup, usually a tape device attached via fiber or SCSI connection. The media server/storage node is responsible for actually putting the data to tape. This server takes it direction from the central server as to what files to backup and the particular media set to put it on. The third and most important is the client. Client software is installed to every system that requires a backup. Even the central server and media servers usually have client software on them so that they may back themselves up.

The features of centralized backups that administrators have the hardest time grasping are tape pools, classes (groups) and retention periods. Administrators tend to be familiar with the Grandfather, Father, and Son scenario on single servers. This is when one server writes one tape per night. That tape is then retained for a period of time depending on which rotation it is in. Once that tape has reached its expiration date the tape can be put back into the rotation and used again. The results of such a scenario are that the full capacity of the media remains unused. A single forty-gigabyte DLT tape could be used to backup a single server with eight gigabytes of data on it. Once the tape expires the tape is rewritten from the beginning and the remaining thirty-two gigabytes of tape remains empty. Data in a centralized backup system is saved by the retention period assigned to it. Retention periods are the time the central backup systems remembers where a file is stored on tape in a save set. Once all of the save sets on a tape have reached the end of their retention period the tape is marked for reused. The data on these tapes is retrievable until that tape gets used again. If a tape has not been reused most backup systems will

allow the tape to be scanned and all of the files on that tape are re-logged into the indexes so that the data may be extracted.

Classes or groups back up data in a central backup system. This is where an administrator defines what is to be backed up and from which servers. A class or group may contain several different schedules. Schedules are also where the administrator defines what pools are to be used. The administrator may have one schedule for daily backup writing to a daily pools, another for weekly backups to a weekly pools and yet another for monthly backups writing to a monthly pool. Centralized backups group tapes together by pools. Using the pool concept a number of backups can be put onto a tape. When a tape is filled the backup continues using another free tape within that pool. When configuring the backup systems tape pools the administrator need to insure that all of the backups using that pool have the same retention period. Once full a tapes within a pool will not recycle until all of the save sets have passed their retention date. An example would be if the system were running daily backups from four servers. The first night all of the backups could conceivable fit onto the first tape in the pool. The following night only two of the servers were able to put there save sets onto that tape before another tape was used. The first tape would not become available until the save sets from the second night had expired. At that point the first tape would be marked as free and available for use again. Mixing retention periods within tape pools will result in tapes never expiring and the system searching for free tapes. For example, let's say the system has a daily pool with retention of two weeks and a monthly pool with a retention level of seven months. If a weekly backup gets written to a tape in the daily pool that tape will not become available for seven months. All of the daily save sets could be expired but the one monthly backup with a long retention period would hold that tape.

What to backup and how long to keep it

Rules on what to backup are a fairly subjective issue. Some tend to argue that backing up a systems operating system is unnecessary. Their argument is based on the premise that if there is a total system failure the system administrator will have to reload the operating system and backup client before any data can restore any data. This tends to be true for most of the Unix clients, however most of these systems tend to offer some sort of disaster recover option for Windows Based and Novell NetWare servers. Also when software packages are installed on Unix servers they may modify files and add files to the Operating Systems mount points. This would mean that the system administrator would have to reload every Software package on the system before the failure. The author is a firm believer in backing up everything. As a Backup administrator decisions will have to be made on what to backup based on backup windows, media cost storage device capacity, and most of all, what can afford to be lost.

How long a company keeps its data stored is another issue that administrators must deal with. First and foremost is following company policies and laws governing the business. Data such as financial records in some cases must be kept seven years. Policies may require corporate proprietary information such as design documents or copyrighted material kept for an infinite amount of time. Other information such as operating system files and e-mail may have a retention period just long enough to insure that one can recover from a system failure. Management may request that e-mail files not only expire but the information is destroyed.

Lawyers have become relentless in their search for data during any litigation process. They can request all e-mail records that may pertain to a pending lawsuit. If the data is stored on tape a corporation can be legally bound to provide any information they still have. One has to remember that Tape is a finite storage option. Tapes can be counted on to retain their contents for about five years.

Features and Options

Most of the central backup packages available start with a base configuration. This would include a central server acting also as a storage node, support for one or two tape drives, and four to eight like operating system client licenses. The author has listed below some of the options available with most centralized backup packages.

Client licenses – These are required to backup servers not covered by the base package. This would include servers of the same operating system as the master server above the number included in the base package number or clients using an operating system other then ones on the central server.

Robotics, Slot and/or Drive licenses – Depending on the package these are needed based on the tape library intended to be use with the system.

SAN Backup Options – This feature is done one of two ways. Either the option is installed on all of the servers on the SAN, thus making the servers into storage nodes. These nodes are only allowed to back the data they contain on the san or internally to a tape storage device also located on the SAN. The other way this is done is that the media server interacts with the storage device such as an EMC disk array. The software will instruct the EMC to create a snapshot of the designated volumes to alternate storage. Then the backup server would backup this snapshot as if it were attached to the original host.

Shared Storage Option – This option allows Storage nodes to access a tape library at the same time. The server and library are connected via a common connection. The connections are usually via a SAN fabric or Multi-hosted SCSI connection. It is configured so that one server acts as the director. This server controls the robotics and tape drive assignment for the other host. Once a drive is assigned to a server, that server has exclusive rights to that drive until the backup it is running has completed.

Database modules – This is software that is added to the central backup system so those databases can be backed up while online. Most vendors offer database modules for Oracle, Sybase, Informix, Microsoft SQL, and Microsoft exchange. These modules tend to work with utilities included with the databases. For example, a central backup software module for Oracle will normally interface with the Oracle RMAN utility to capture an online backup.

Encryption Modules – Provides addition security for you backups. Both the client and server have keys on encryption keys. When a backup is preformed the data is encrypted at the client

before being sent to the server for storage on tape. This added an extra level of security to your backup over the network or if a third party manages to get hold of the backup tape. The keys used to create the encryption should be record and stored in a secure place. In the event of a system failure the data would be irretrievable without the encryption keys.

Archive/Vaulting Options – These options are software or scripts added to the backup software. They can be used to duplicate tapes written during the normal backup schedules. The option then can change the retention period associated with these tapes so that they can be held for longer-term storage. An example would be rather then run a daily, weekly and monthly on the same day the option will duplicate the daily and give the duplicate a weekly or monthly retention time. Thus, daily tapes can be kept onsite so that files can be recovered without having to pull a tape from offsite storage.

Disaster Recovery Options – These are included with larger base packages. Usually these options are only available for windows based systems. To use these options a set of boot disk are created. If a system failed an administrator or operator could boot a new system with the disk that were created. The disk would then walk an administrator or operator through reloading the operating system, installing client software, and recovering from the last full backup.

True Image restores – Most of the higher end products have this feature built into their software. It is used primarily when doing a combination of full and incremental backups. Let say that a full backup runs on a Sunday night. Monday, a file is created. That file is then backed up on the Monday incremental tape. Thursday, this file is deleted and on Friday the system crashes. With true image restores, that files would not be restored as the system recognized the fact that the system was deleted while doing the Thursday night backups. Essentially the system would only be restored with only the files it had on it during the Thursday night backup, even though the file was not on the Thursday backup.

Open File Options – With this feature in place the backup software will backup files on clients that are show as open. For this feature to work the file must remain "quiet" for a few seconds so the backup software can get a complete copy of it. These features do not normally work well with files in a constant state of flux, such as a mail stores.

Concurrent Sessions/Multiplexing – Most of the backup software available today have this feature built in. This feature allows multiple clients to backup to the same tape at the same time. The data from these clients is mixed by the storage node or media server and then feed to the tape drive. This feature is very useful in an environment that has slower network links. However there are some drawbacks to it use. Restore times increases, as more data has to be read from the tapes to locate the files that are attempting to be retrieved.

Firewalls and Clients outside of them

When a company makes a large investment into a total backup strategy management tend to

want the IT staff to us it to backup every server on the network. This would include the ones in the screened network and outside the firewall. It is possible to backup these servers but to do so requires a large amount of port openings on the firewall. The author has listed just a few of the requirements for Veritas Netbackup, Backup Exec and Legato's Networker below.

Veritas Backup Exec for Windows and Unix

- Port 6101 : Backup Exec UNIX and 95/98/ME Agent
- Port 6103 : Backup Exec Agent Accelerator and Remote Agent
- Port 135 (TCP and UDP) for Remote Procedure Call Service
- Port 137 (UDP) for NetBIOS Name Service
- Port 138 (UDP) for NetBIOS datagram
- Port 139 (TCP) for NetBIOS session
- Port 1024 & Above : RPC Communication

http://seer.support.veritas.com/srchengine/sth.dll?Tag&Path=seer%2Esupport%2Everitas%2Ecom%2Fdocs%2F233828%2Ehtm&CiRestriction=ports

Veritas Netbackup for Windows and Unix

Outbound:

- Allow ports 512 - 1024 on the master/media server outbound to port 13782 on the client.
- Allow connections from port 13721 outbound from the master server to ports from 512 - 1024 on all media servers.

Inbound:

- Allow connections to ports 512 - 1024 on the master server inbound from the client.
- Allow connections inbound from the client to port 13720 on the master server.
- If multiplexing is being used (not just on), and streaming multiple jobs to a single tape in unison is occurring, then allow inbound connections from the client to ports 1025 - 5000 on the master server.

http://seer.support.veritas.com/srchengine/sth.dll?Tag&Path=seer%2Esupport%2Everitas%2Ecom%2Fdocs%2F187321%2Ehtm&CiRestriction=ports

Legato Networker for Unix

- Ports 7397 through 9936 (TCP and UDP) for Service Ports
- Ports 10001 through 30000 (TCP and UDP) for Connection Ports

http://www.legato.com/support/documentation/bulletins/354.html

The convenience of a centralized backup system can open up company systems to some risk. Backup software use remote procedure calls to initiate backups and restores. Opening ports on a firewall will only increase the likelihood of intrusion. Some firewall can be configured with time-based rules so that ports required for backup and restore are only open during designated hours. While this may reduce the risk, the author would recommend using the host-based utilities include with your systems operating system for backing up systems outside of the firewall.

Security features and roles

Most centralized backups have security features built into them. For instance, file restores are normally done just from the originating clients. This feature stops just any client from getting access to every file in the corporation. Roles can be defined within the systems also. The main administrator would normally be the system administrator of the central server. Roles such as Tape Operator can be defined. Tape operators can do task such as add and remove tapes from the various devices and tape libraries in a data center. Tape Operators would not be allowed access to change client definitions or backup schedules. Another role might be a systems operations roll. This role would be responsible for task such as enabling or disabling drives, and checking backups status. These features can be very useful in securing the corporate backups. Configured incorrectly the entire system can be at risk.

Off-site Storage

The data is now protected on tape. Where should it be stored? Leaving it in the tape library or in an administrators desk draw is a precursor for disaster. If a catastrophic event destroys the corporate faculties the data will go with it. Storing tape in someone's home is not a good idea either. If a tape is needed the administrator have to try to locate that person. While on-site tapes should be stored in a secure and fireproof vault. There are several inexpensive types available on the market today. For off-site storage the author recommends that a reputable records-storage company be contacted. These companies offer scheduled pick –ups and drop-off. They will also guarantee that tapes will be delivered to the company faculties when they are called for within a pre-determined amount of time.

Summary

There are several very good Centralized Backup systems available on the market today, a few of them the author has listed above. All of these systems are extremely flexible so that they may be tailored to almost any backup need. These systems can take some time to setup to gain the desired results. However, once setup administration should be very minimal, limited only to tape rotation into and out of your storage devices and the addition of any new servers to the backup scheme. These systems are especially helpful with backup windows shrinking and the increasing 24 hour availability requirements being put on IT departments today. When considering a complex backup system a company should review all of the backup requirements and choose the package that best suits their needs. Most of the vendors today are willing to work with potential customers and allow them to test their software on an evaluation basis.

References:

1. xrefer - full backup - http://www.xrefer.com/entry/622843
2. xrefer - differential backup - http://www.xrefer.com/entry.jsp?xrefid=622612
3. xrefer - incremental backup - http://www.xrefer.com/entry.jsp?xrefid=622947
4. How to back up a computer that is protected by a firewall, or isolated in a separate workgroup for security reasons, with Backup Exec for Windows NT and Windows 2000. -

http://seer.support.veritas.com/srchengine/sth.dll?Tag&Path=seer%2Esupport%2Everitas%2Ecom%2Fdocs%2F233828%2Ehtm&CiRestriction=ports

5. What ports need to be open to back up a client that is behind a firewall? - http://seer.support.veritas.com/srchengine/sth.dll?Tag&Path=seer%2Esupport%2Everitas%2Ecom%2Fdocs%2F187321%2Ehtm&CiRestriction=ports
6. The electronic document trail - http://www.nwfusion.com/newsletters/ecomm/2000/0918ecomm1.html
7. The Use of Firewalls with NetWorker® Server Release 5.5 and Later - http://www.legato.com/support/documentation/bulletins/354.html
8. ufsdump - incremental file system dump - http://uwsg.iu.edu/usail/man/solaris/ufsdump.1.html
9. The Hows and Whens of Tape Backups - http://www.networkcomputing.com/1205/1205ws1.html
10. Veritas Netbackup Datacenter 3.4 System Adminstators Guide - http://ftp.support.veritas.com/pub/support/Products/NetBackup_DataCenter/netbackup_dc_adminguide_unixserver_232348.pdf
11. Legato Networker Adminstators Guide Unix Version - http://web1.legato.com/infodev/publications/NetWorker/UNIX/5.5/cd_docs/uxag.pdf