WHITE PAPER: NETBACKUP ARCHITECTURE OVERVIEW

symantec.

Confidence in a connected world.

NetBackup™ Architecture Overview

Dave High

December 2008

NetBackup™ Architecture Overview

Contents

1.0 Introduction

The purpose of this document is to provide an overview of the requirements for a basic NetBackup™ architecture to help customers understand which components are required in a data protection solution. Where possible, an example of a component will be noted; however, that component could be swapped for a similar component with the same bandwidth characteristics based on the vendor of choice. A much more comprehensive document, the Veritas NetBackup Backup Planning and Performance Tuning Guide, is available for free from Symantec at:

http://seer.entsupport.symantec.com/docs/307083.htm

This overview is not meant to replace the above referenced document or other documents that provide a deeper understanding of the best way to create a data protection environment, nor is it meant to replace an onsite Professional Services architecting engagement. It is intended to be a quick overview of how a NetBackup architecture is created to assist customers who are new to NetBackup or who are considering upgrading/scaling an environment. This document is designed to help you understand what will be required from a hardware, software and personnel standpoint when NetBackup is deployed; the above referenced document can be used as more of a deep dive.

With the deployment of newer, multi-core servers and PCI Express™ (PCIe) bus, server choices have become less of an issue than they were even a couple of years ago. Most modern servers are capable of performing the master server and/or media server tasks very well; therefore the recommendation for NetBackup customers is to choose the hardware they are comfortable with, or a vendor they have a relationship with.

Note that not all features described in this document are available with all currently supported versions of NetBackup on all platforms and with all hardware configurations. You can learn more about what is and is not supported in the various hardware compatibility lists (HCLs) on the Symantec Support website. This link provides a gateway to all current HCL documents:

http://seer.entsupport.symantec.com/docs/303344.htm

1.1 Glossary of Terms

The following are some of the terms with brief explanations that are used in this document:

- Bandwidth—Refers to a component's physical limitation when sending data from one point to another. It also refers to how much work an individual backup administrator (FTE) is capable of. Bandwidth has many variables.

- Master Server—The controller of the NetBackup environment. The master server schedules the jobs, maintains the policies, allocates the resources, stores the catalog metadata, etc.

- Media Server—The workhorse of the NetBackup environment. Data passes through the media server to the final repository in a typical scenario.

- NetBackup Client—Any system that is sending data to a media server. In the case of a SAN media server, the client and media server are the same system.

- Policy—The "who, what, where, when and how" of a backup. The policy can include multiple clients, multiple schedules, multiple file listings, etc.

- Storage Unit—A virtual representation for physical storage where NetBackup will send data.

- Recovery Point Objective or RPO—Determines how much data loss is acceptable. For example, if backups occur only once every 24 hours, then the maximum data loss could be as many as 24 hours. In order to meet a specific RPO, careful planning and testing are required. Short RPOs require backup and recovery scenarios that typically will not involve traditional backup from disk to tape.

- Recovery Time Objective or RTO—Determines how long the environment has until it needs to be back online. In many instances this could be less than an hour, in which case traditional backup to disk/tape is not the right choice for protection. Similar to the RPO, proper planning and extensive testing are needed to document steps for recovery to shorten the amount of time recovery takes.

- Service Level Agreement or SLA—The SLA of an environment is usually based on input from those whose data is being protected and determines how data should be backed up, the RPO/RTO associated with the data, how long it is stored (retention), etc.

1.2 Bandwidth Overview

The term “bandwidth” is used extensively when developing a data protection solution. In general, it refers to the ability to move data from point A to point B, and how fast that data is capable of moving—based on the components that are in use. In the broader sense, bandwidth is also used for the “human” component. For example, only very small environments should have a single data protection administrator. While it is difficult to pinpoint exact FTE requirements, recommendations will be made in this document based on other customers’ experiences and overall size of the solution.

As an environment is developed, various components within the solution will have specific bandwidth numbers they are capable of; for example, speeds. Based on testing and customer experiences, we recommend that any bandwidth numbers follow the “70% rule,” which dictates that any speed claimed by the hardware vendor be reduced to 70% of the top-rated speed. For instance, Gigabit Ethernet (GbE) speed is theoretically 125 MB/sec; however, customer experience and testing have shown that a practical speed for performing bandwidth calculations is closer to 87 MB/sec (125 * 70% = 87.5 MB/sec). If you use 87 MB/sec as the maximum throughput speed of a GbE connection as your baseline when developing the architecture it makes the overall numbers more realistic. Any extra speed that can be achieved just makes the backups faster, but should not be expected.

The only exception to the 70% rule is tape drives. They can be driven to 100% of “native” rated speed with the right hardware and exceed native speeds based on the compressibility of the data being backed up.

1.3 NetBackup Overview

NetBackup is considered a “three-tiered architecture” in that it has a master server that provides configuration services, policy creation, scheduling, reporting, logging, etc. It is the “traffic cop” of the backup solution and allocates resources to the media servers, the second tier.

The master server processes are typically CPU-intensive based on such things as job scheduling, running binaries, metadata gathering (catalog), etc. The disk used for the master server should be “fast” disk based on the need to write the catalog data; however, it does not need to be the fastest and most expensive disk in the data center. It should be configured for fast writes and should have the ability to be expanded (size increased) on the fly without bringing the server down. The master server also needs a disk area to write any log files. This will be discussed in more detail later.

The media server is the workhorse of the NetBackup environment. In high performance environments, it must be a high I/O data mover with multiple NICs, multiple HBAs, connected to disk, tape drives, SAN, and to the LAN depending on the backup requirements. It must be able to move vast quantities of data from point A to point B under the direction of the master server. It is connected to the hardware used in the environment and thus has different requirements than the master server. Media server requirements will be discussed in more detail later.

Finally, there are the clients. These are the systems that have the data that must be protected. Regardless of how the data is being protected, the client is where the data originally resides. Some options blur the line of client and media server; however, from the NetBackup standpoint, when a media server is sending its own data to a device for backup, it is considered a client. This distinction is made to make it easier when opening a support case or discussing your environment with an SE or consultant.

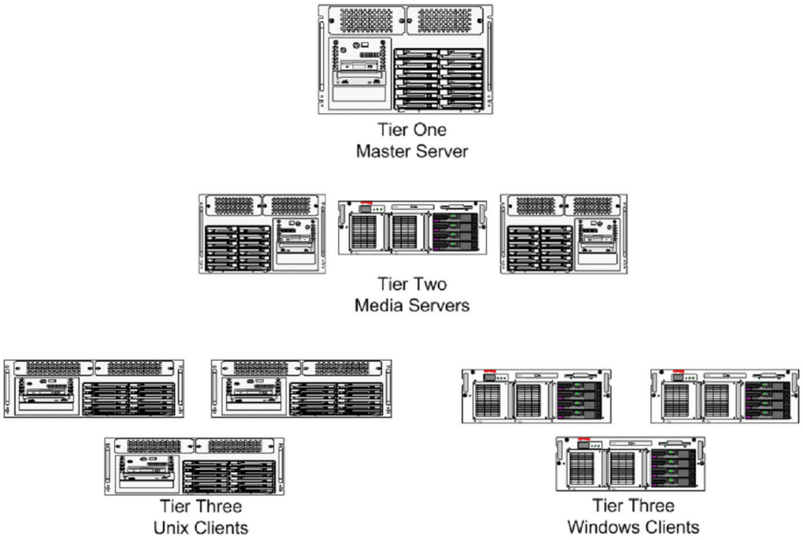

Figure 1: Architecture Example – NetBackup's Three Tiers

1.4 Data Protection in a Nutshell

As a data protection architecture is developed, a number of items that are often overlooked need to be considered.

1. What needs to be backed up? (Production data, etc.)
2. What doesn't need to be backed up? (For example, development data that can easily be re-created from other sources)
3. How often does the data need to be captured? (What is the RPO requirement? See below.)
4. How long should the data take to restore? (What is the RTO requirement? See below.)
5. How long does the backup data need to be saved? (What is the user expectation? Are there compliance requirements that need to be considered—for example, HIPAA or SOX audits—that will determine how long data needs to be retained?)
6. What is your expected turnover of operations staff? (How much do you need to provision for staff training, etc.?)
7. What will happen in the event of a site loss? (Is there a site DR plan, are the backups being stored off-site, etc.?)

Most of this translates into one thing: The establishment of solid, workable service level agreements (SLAs). The key word here is "agreement." Any agreed service level must be realistically achievable even when things are not running 100% smoothly.

Accept this as fact: Backups are going to fail sometimes. Hardware, software and human error can cause failures, so developing a solution that attempts to meet the goals of the company while minimizing risk should be a part of the process.

Ideally, a data protection system should only need to retain information for a short period of time. From a disaster recovery perspective, a backup has no value once a more recent backup of the same data exists. The goal is always to recover the data to a point as close to the point of loss as possible; however, compliance requirements mean that it is becoming increasingly common to retain at least periodic backups for long periods of time. This needs to be factored in when designing the solution because extended retention increases elements such as the amount of storage capacity and the size of the backup catalog.

Once goals and requirements are determined, the size of the overall data protection solution can be determined. Hardware, software and personnel requirements can be determined mathematically; for example, bandwidth requirements. Hardware and software can then be put in place to meet these requirements, and personnel can be allocated to keep the solution running. Standardizing on a suite of products from a specific vendor and minimizing the amount of scripting and/or third-party solutions can help during personnel turnover.

1.5 RPO/RTO/SLA Development and Recovery

Prior to determining which hardware to buy or software to run for any data protection solution, the requirements of the company must be understood. It is impossible to know which server to use, which tape drive or disk array to buy, and how many people will be needed to manage the solution unless RPO/RTOs and/or SLAs have been developed. This is the first step in designing a solution.

While backup SLAs are good, the purpose of a data protection solution is "recovery"; therefore as SLAs, RPOs and RTOs are developed, the recovery requirements must be understood and always considered. Very short RPO and RTOs are possible; however, this methodology is very expensive. A traditional data protection environment—backing up clients over the LAN to a tape drive—is very inexpensive compared to other solutions; however, it will have very high RPO and RTO capability. That is, if you need an RPO of 2 hours and an RTO of 4 hours, you will not be able to meet these requirements if you are simply sending data to tape. Most tape-based solutions provide an RPO/RTO of no fewer than 48 hours. If backups are going to tape and are then being sent offsite, this will also increase the amount of time of a recovery. Typically, offsite tapes require 12–24 hours to be returned, mounted, loaded and the recovery to be started. If you require less than that, you must look at a disk-based option and/or a non-traditional method such as snapshots and/or replication. Short RTOs may require initial disk-based backups that are then sent to tape for longer-term storage. Short RPOs may require multiple backups through the day and may require snapshots or some other means to quickly recover. You must understand your recovery needs rather than your backup needs to properly plan for your environment.

Always be realistic when setting recovery SLAs. It is a popular belief that running a scheduled backup every day guarantees a 24-hour RPO in all cases; however, this is not the case. Even if the backup succeeds, there is a period of time between the start and finish of the backup that potentially falls outside the 24 hour period. Also consider that the possibility that the backup may fail and need to be rerun and that a backup failure may go unnoticed for many hours (particularly where there is no 24x7 operations coverage). It is easy to see that the RPO may be much longer than 24 hours. A more reasonable RPO for sites doing daily backups is 48 hours.

Depending on the size or function of the company, there may already be requirements from the legal department on how data must be backed up and how long it must be retained. Many times this pertains to email retention; however, it could also include documents and other files on network file shares. If the legal department requires that all files be backed up and saved for seven years, and that any data that has been backed up in the last 30 days must be recovered within two hours, then you will need a large capacity library to be able to store all the data for at least 30 days before the tapes are ejected and sent offsite, or you will need enough disk storage to save the data for 30 days on disk before being copied to tape.

If there are no legal requirements, HR may have suggestions for what data should be backed up and how long it should be saved. This will impact the hardware that is purchased and the cost of the solution.

Finally, SLAs need to be developed based on your budget. It is easy to state “yes, we can save everything forever,” but without unlimited budget, this won’t be possible. Therefore when discussing the requirements with the users who are impacted by the SLAs, negotiations may be needed to meet their needs while remaining within a budget.

In summary, you must understand the site requirements for backups and recoveries before you can properly develop an architecture to meet those requirements. You cannot develop SLAs, RPOs and RTOs in a silo. You must talk to the users and other groups within your company to fully understand the requirements. The final piece is budget. You can determine RPO/RTO/SLA requirements by talking to the appropriate users/groups and pricing out what will be required, then you can go back to the users with more realistic expectations as to what you can do for them vs. what they need. A “chargeback model” can be very beneficial. This allows you to charge the groups for the backup service. For more information about chargeback, see section 6.

2.0 NetBackup Hardware

NetBackup requires hardware to do its job. A master server is needed to schedule the jobs, allocate resources, and track the metadata. Also required are media servers and transmission lines to move the data, and ultimately a place to store the data at the end of the path. This section discusses some of these requirements to help you better understand how it all works when building a solution, and to help you make the best decisions when buying hardware. As mentioned in section 1, the Veritas NetBackup Backup Planning and Performance Tuning Guide should also be consulted to help you size the systems once you have decided which hardware to buy.

2.1 Master Server

The master server is the controller of the NetBackup data protection solution. It allocates resources to the media servers, keeps track of the media being used, catalogs the information from the clients, maintains data for basic reporting, and schedules backups among a host of other duties. We recommend that the master server be a dedicated system rather than acting as a Master/Media server, although in very small environments it can act as both. If an environment is small to begin with, but any type of growth is planned—either by the total amount of data to back up, or the overall number of jobs per day—then a dedicated master server should be considered. Master servers should not be used to co-host other applications.

2.1.1 CPU Requirement

The master server functionality is very CPU-intensive, yet it requires less I/O bandwidth than a media server. Modern multiple core hardware options have made choosing a master server for most environments much easier than it previously was. The aforementioned Veritas NetBackup Backup Planning and Performance Tuning Guide has some guidelines for purchasing hardware, including information about sizing a system based on the number of clients, the number of backup jobs and the number of media servers that are part of the environment to be controlled by the master server. Distilling this information shows that any system capable of multiple physical CPUs and multiple cores on those CPUs would be a good master server for most environments.

Testing has shown that multiple physical CPUs perform better than a single multiple core CPU. For example, two dual core processors perform better than a single quad core processor, therefore this should be considered when purchasing. Currently, quad core processors may be less expensive than dual core, so a master server should have a minimum of two physical quad core processors where possible.

2.1.2 RAM Requirement

The RAM on the master server is used for basic computing tasks rather than performance buffer tuning like on the media servers. In most cases, “matching” the RAM to CPUs is adequate; for example 2 GB per CPU or CPU core. Additional RAM may not provide additional performance; however, with the price of systems and system components decreasing, loading up the system with CPUs and RAM can be beneficial, especially in situations where scaling is needed at a later date.

2.1.3 Disk Space: Catalog and Logging

Another frequently asked question regarding the master server is how much disk space is required for the catalog. Now that disk is so inexpensive, this question is less meaningful than it was even a couple of years ago. The size of the NetBackup catalog is a function of the number of files being backed up and the number of copies of the backup being retained and is thus directly influenced by the frequency of backups and the retention periods used. If you want to calculate exactly how much catalog space is needed, refer to the Veritas NetBackup Planning and Performance Tuning Guide for more details on how to estimate the size.

For the purposes of the overview and for most master servers, we recommend that you simply start with 200 GB of catalog space and use some type of disk management software such as Veritas Storage Foundation™ to grow it on the fly if needed. As noted previously, the catalog disk should be fast disk optimized for write performance.

Another high disk usage requirement to consider is the log files. Typically, we recommend that the log files on a UNIX® system be pointed to a disk area—preferably, a separate physical disk area—that is separate from the catalog disk and they should never be on the root of the system disk whether on UNIX or Windows®. The log files reside under /usr/openv/netbackup/logs on UNIX. This could be a separate mount point from the rest of NetBackup.

On Windows, NetBackup should typically be installed on the E:\ drive or other hard drive that is not the C:\ drive. The log subdirectories can grow very large and very quickly especially when troubleshooting using high verbose rates. If the logs and catalogs are placed on the root drives, and they fill up, the machine can lock up and/or crash and may be difficult to restart due to a full root file system.

2.1.4 Master Server Hardware Example

An example of a very good master server is the Sun™ T5220 or the HP ProLiant™ DL580 G5. Both are capable of multiple, multi-core CPUs and lots of RAM. They have very good processing speeds and as a master server are capable of scheduling, collecting metadata, and reporting in busy environments. They should be as fully stocked as budget permits and should have at least two physical CPUs that have multiple cores and at least 8 GB of RAM. A single GbE NIC is sufficient for catalog metadata traffic. The master server should have access to a tape drive for catalog backup, which would typically require an HBA for SAN connectivity to a tape device.

2.2 The Media Server

The media server is the data mover in a NetBackup environment and as such, it requires specific bandwidth considerations. Modern servers—those purchased in the last 18 months—will typically make I/O a non-issue based on the inclusion of PCIe bus in most servers. It is difficult to find a modern enterprise-class server that does not have multiple PCIe slots. In a typical system that includes PCIe, and depending on the type of PCIe, the average speed of the backplane has a bandwidth of 4 GB/sec compared to 256 MB/sec for PCI. Some of the more modern systems are twice as fast (16 lanes compared to 8 lanes); however, for the purposes of this document, consider any server class system with multiple PCIe slots to be adequate for media server usage.

2.2.1 CPU Requirement

When a media server is chosen, the CPU configuration is not as important as that of the master server; however, there are some considerations based on the number of backups that will be sent through the media server either via the LAN or the SAN. However, this being said, once again modern hardware with multi-core processors and inexpensive RAM make most servers acceptable. In most testing, having a system with multiple physical CPUs has shown better performance than a system with a single physical CPU. For example, two dual-core processors provide better performance than a single quad core CPU.

When backups are running through a media server, there is a CPU impact. Extensive testing with the newer multi-core processors has not been completed; however, in general, assume that 5 MHz of CPU capacity per 1 MB/sec of data movement in and out of the media server is typical for backups being written to tape and regular disk storage. Note that some storage options, such as the Media Server Encryption Option (MSEO) and PureDisk™ Deduplication Option (PDDO) place additional load on the media server CPU.

For example, a LAN-based media server (one that services clients across the LAN and has multiple GbE or a 10 GbE interface) that is backing up 20 clients each sending 5 MB/sec to a tape library would need 1,000 MHz of CPU power:

- 20 Clients * 5 MB/sec = 100 MB/sec total = 1,000 MHz
- 500 MHz to receive the data across the LAN
- 500 MHz to send the data to the tape drives

Because modern CPUs are measured in GHz, it is apparent that whether trunked GbE or 10 GbE is used, the CPUs in a modern server will not be a bottleneck. You should also consider whether anything else is running on the media server. When troubleshooting performance issues with Windows backups involve "stuck" processes on a media server (such as a stuck Explorer session on a Windows machine), large quantities of CPU cycles can be consumed, thus decreasing the overall performance of the backup. For this reason, monitoring the CPU usage may be required in the event of performance issues to ensure that something other than NetBackup isn't overloading the CPUs.

2.2.2 RAM Requirement

The media server uses RAM for buffering when sending data from the source to the backup device. Think of buffers as buckets; that is, when you want to empty a pool of water, the number of buckets as well as the size of the buckets are important. Tuning a NetBackup media server requires that the "buckets" be manipulated and performance checked to determine if the change has had an effect on the backup speeds. Over-tuning can occur when the numbers are set too high, thus negatively impacting the media server performance.

For more information about tuning procedures and instructions for determining the correct amount of RAM, see the Veritas NetBackup Backup Planning and Performance Tuning Guide. In general, a media server should have at least 8 GB of RAM for tuning, even though only a small part of that will be used for actual buffers. Remember that most media servers have a limited number of drives that can reach capacity based on bus and HBA card configuration. Typically, a single media server can utilize 4–6 LTO3 or LTO4 drives (as an example) depending on bus type and speed, the number of HBAs, etc. Therefore if you have one media server and 16 drives to tune, you may have incorrectly configured your architecture.

One final note on the media server and PCIe. The proper placement of the HBA and NIC cards should be discussed with the hardware vendor for maximum throughput. While there are any number of slots in a system (PCI, PCIx, PCIe), placing the high speed cards in the correct slots is critical to achieve full throughput.

2.2.3 Media Server Example

An example of a very good and moderately priced media server from a Sun UNIX perspective is the Sun T5220. When used as a media server to move data from disk to tape (for this example), with a 4 GB HBA it is capable of moving 400 MB/sec, which takes into account the 70% rule. This equates to 17 TB data transfer in a 12-hour period. This assumes that the disk and tape drives are capable of this type of sustained read/write speeds, but the bus on the T5220 should never be the bottleneck. The T5220 can also include dual 10 GbE XAUI ports for fast LAN transfers if your switch is 10 GbE capable. Similar systems with PCIe bus will provide similar speeds. The easiest thing to do if you are not a Sun UNIX customer is to ask your hardware vendor for a similar Dell®, Compaq®, HP-UX®, AIX™, etc. machine. If the server has PCI Express, you should be fine.

2.3 Full Media Server

A traditional media server is one that is used to move data from LAN and SAN based clients to a final destination typically across the SAN; however, this type of media server can also send data from SAN attached disk to a final destination as well. It can also be used as a proxy server during off-host backups. This is called a “full” media server. It has also been referred to in the past as a LAN media server to differentiate it from a SAN media server, which is discussed below.

As a general rule, a media server (not a SAN media server, which is described in the next section) should be regarded as a dedicated server and should not also be used to host other applications that could contend for I/O bandwidth.
When being used to back up client systems across the LAN, the bandwidth into the media server needs to be adequate to “match” the SAN component for the devices at the final destination. While SAN speeds have been increasing, LAN speeds have not kept pace and the new 10 GbE connections have not been widely adopted due to cost. In the past, a GbE NIC in a media server was adequate because most clients were at 100 Mbit; however, this is no longer the case. Most clients in enterprise environments are now GbE, therefore the media server bandwidth must be increased in order for these clients to function properly.

All major OS and hardware vendors now have technology for “trunking” the GbE NICs in a server (also called “port aggregation”), which makes it possible to take standard GbE connections in a media server and turn them into a “fat pipe” that is capable of moving large quantities of data across the LAN. A single IP address is presented to the clients, and the switch handles the traffic as if it were a single large pipe. Any media server that is moving data across the LAN requires either 10 GbE or port aggregation technology of some type in order to move enough data to meet backup windows.

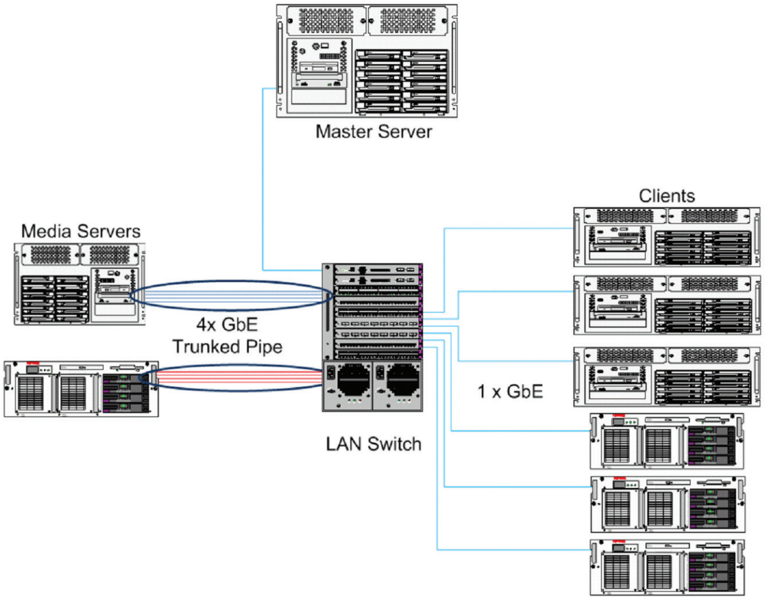

Figure 2: Example of LAN Trunking/Port Aggregation for Media Server Bandwidth

When being used to back up client systems over the SAN (SAN clients), the media server must be configured as an "FT" media server with specific HBA ports configured with the Symantec target mode driver to receive data from the client systems. Operating system and HBA hardware restrictions apply to FT media servers. We recommend that you check the latest NetBackup hardware and operating system compatibility lists to ensure that the hardware and operating system chosen support this feature.

2.3.1 Media Server Load Balancing

If the environment is NetBackup 6.5 or higher, Media Server Load Balancing (MSLB) can be incorporated into the architecture. This increases reliability and ROI of the solution. MSLB allows you to create a pool of media servers so that the clients can be backed up to any of the media servers. The master server runs an algorithm that rates each media server on a number of factors to determine which should be used. Typically, it is the "least used"; however; there are other factors. This negates the need to have clustered media servers. If you have an MSLB group set up, and one media server goes down, the traffic will simply use another media server until the bad one can be replaced. MSLB in conjunction with shared disk pools and Shared Storage Option can make for very high ROI on the hardware investment, minimized management of the backup solution, and more reliable backups.

While extensive testing as to the number of media servers that can be added to a balanced group has not been completed, we recommend that you start with four media servers. As more media servers are added, track how long it takes jobs to start in your environment. When it begins to take too long to start jobs (more than a minute or two), then no additional media servers should be added to that group. The time required to start jobs is due to the algorithm that NetBackup uses to choose the "best" media server in a group to send a backup to.

2.4 San Media Server

In many cases, backing up a client across the LAN is simply not feasible. Typically, this is due to bandwidth restrictions on LAN speed and the need for a backup that is faster than what the LAN can achieve. In this case, it is possible to install the SAN media server code on the client and with an HBA zoned to see tape drives, move data across the SAN directly to shared drives.

The speeds of these backups are typically only limited by the SAN connection and the ability of the hardware and OS to send the data across the SAN. Most systems can send single streams of data at about 15 MB/sec each, which is about 54 GB/hour per stream. In many cases, depending on the server hardware, SAN media server systems can send multiple streams at the same time. While this places more overhead on the system, it can speed up the backups.

While an observed average is in the 15 MB/sec range, some single stream backups can get into the 70–80 MB/sec range, which can back up 252 GB per hour or more without impacting the production LAN. These amounts and speeds depend on the server, the SAN, and what else is running on the server during backups.

There are a couple of drawbacks to the SAN media server and while it is an excellent way to send data to tape/disk quickly, for production systems that cannot have impact on the CPU, a snapshot or “mirror split” technology may be a better, albeit more expensive solution. The drawbacks include the overhead of the NetBackup software on the client, and the fact that the system is attached across the SAN to tape drives and may require a reboot if a drive is replaced, firmware is upgraded on the drive or HBA, etc. To solve the problem of having production systems see hardware across the SAN, the SAN client was introduced in NetBackup 6.5 to allow a backup to run across the SAN to a media server. This is discussed in the next section.

A final note on SAN media servers is that they require buffer tuning similar to regular media servers to achieve optimal throughput. And because they are probably running another application of some type, a SAM media server may have more robust CPU and RAM requirements than a regular media server.

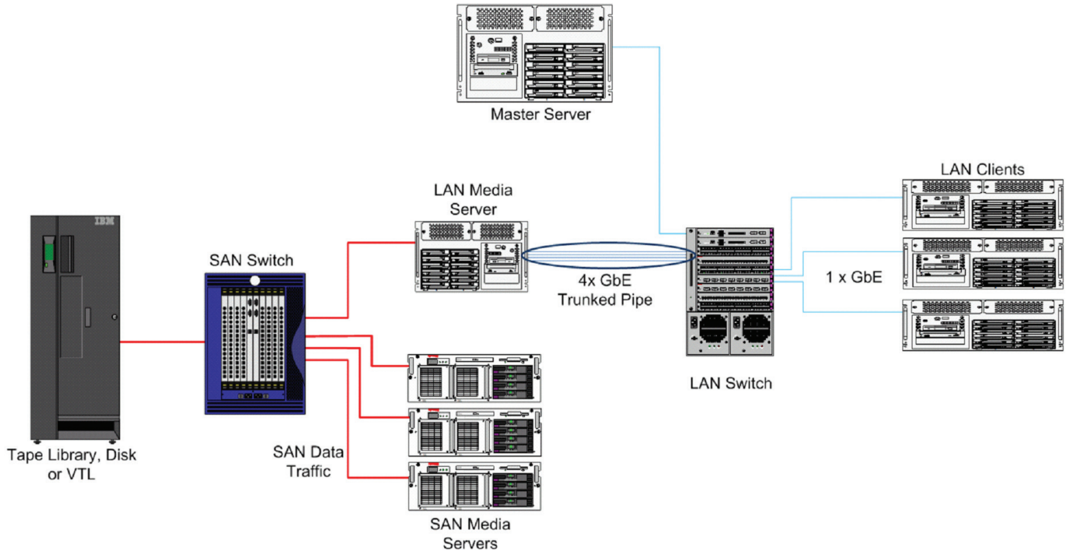

Figure 3: LAN and SAN Media Server Traffic

2.5 LAN Client

The NetBackup client is any system that is sending data that needs to be backed up. Most of the time a “client” is a system sending data across the LAN; however, there is now a SAN client that can send the data faster using the SAN, which is discussed in the next section.

The standard LAN-based client is typically capable of sending 4–8 MB/sec per stream on a 100 Mbit LAN connection and between 15 MB–70 MB/sec per stream on a GbE connection. Testing has shown that due to limitations on hardware—disk speed, CPU usage, and other activities occurring during the backup—even though GbE is typically capable of 80+ MB/sec, most systems cannot send a single stream of data that fast due to overhead on the TCP/IP stack and the operating system limitations. For this reason, systems with NICs faster than 100 Mbit will probably need to be configured to run multiple streams; however, testing is needed in each environment to determine how fast data can be sent, because there are many factors that affect data transfer across the LAN.

Running a standard client-based backup places overhead on the CPU of the system being backed up. Depending on the type of backup, number of files and overall size, this overhead can be considerable and usually a system being backed up will show a noticeable slowdown in any application that is being run at the same time as the backup. In addition, on very fast systems that do not have much running on them, a 100 Mbit LAN connection can quickly be overloaded. NetBackup, with its multi-streaming and multiplexing ability, can overwhelm a poorly configured or overloaded LAN switch as well. It has been said that there is “no better LAN diagnostic tool than a well-configured NetBackup implementation.”

Architecting for LAN clients is all about matching the media servers and the media server LAN bandwidth (NIC cards, trunking, etc.) with the number of clients that need to back up and the amount of time that is available for the backup to occur.

For example, if you have 100 LAN clients and each has 100 GB of data that needs to back up (that is, a total of 10 TB of data) with a full backup over a weekend window of 60 hours, then you will need an aggregate throughput of 167 GB per hour to make the backup window. This is 47 MB/sec. Because most LAN clients are only capable of sending 5 MB/sec across a 100 Mbit link, 10 clients will need to back up at one time and the media server will need to be capable of handling 47 MB/sec, which is easily done with a single GbE NIC.

Simple enough. But let’s increase the size by a factor of 10.

Assume you have 1,000 LAN clients with 100 GB of data for a total of 100 TB of data. This size is typical of a wide range of NetBackup customers. This would require a speed of 1.7 TB per hour, or 470 MB/sec to complete. A single media server with 4xGbE in a trunked configuration provides approximately 320 MB/sec throughput (based on the 70% rule and experience seen in the field), therefore more than one media server would be required if 10 GbE is not available. In this case, two media servers are required, and if placed in a MSLB group sharing as many as 8 LTO3/LTO4 drives, the SLA of 60 hours could be reached.

As stated previously, designing a backup solution is about determining your requirements and developing the bandwidth to meet the requirements. The Veritas NetBackup Backup Planning and Performance Tuning Guide has additional information about properly sizing an environment.

2.6 SAN Client

The SAN client was introduced in NetBackup 6.5 as an alternative to the SAN media server. The traffic path is via the SAN rather than the LAN, and a media server is still used as the path to the final storage device, be it disk or tape. For large clients that can have the overhead of the backup but that do not require the larger media server footprint or do not need to be directly connected to hardware, this is a very good option.

Installation is very easy; however, as mentioned in the previous section about media servers, there are specific hardware and operating system requirements on the media server for running backups from SAN clients. These include a limited range of 2 and 4 Gbit HBAs that support the FT media server’s target mode driver. For the type and model number of HBAs that will work, refer to the latest hardware compatibility list.

SAN clients have been shown single stream backup speeds as fast as 150 MB/sec over a properly configured SAN. This equates to 540 GB per hour. Using multiple streams with 4 Gbit HBAs, SAN clients can achieve aggregated backup rates as fast as 500 MB/sec (1.8 TB per hour), comparable to 10 Gbit LAN connections and SAN media servers.

Once installed and configured, a SAN client is much easier to manage than a SAN media server and offers considerable performance advantages over current GbE LAN clients,

For more information about the SAN client and configuration, visit:
http://seer.entsupport.symantec.com/docs/293110.htm

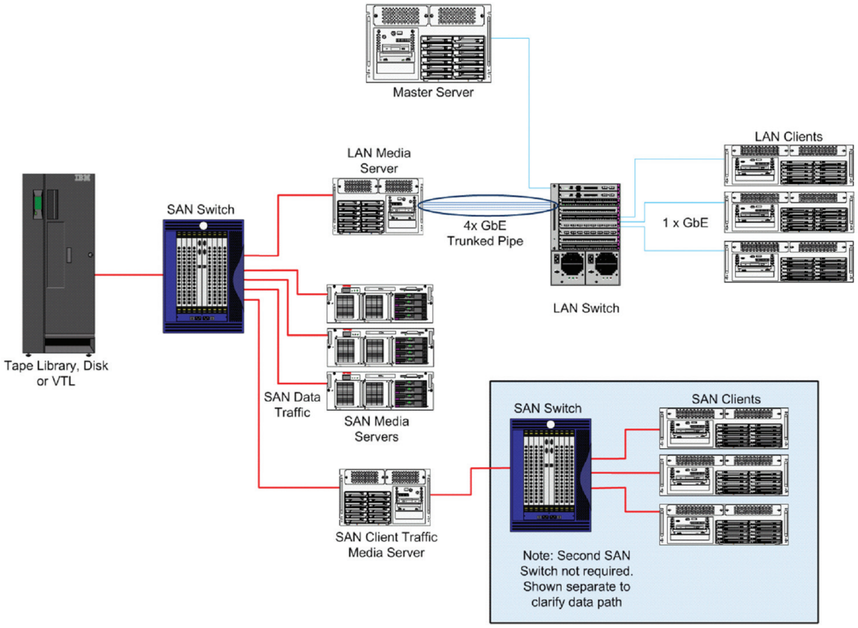

Figure 4: SAN Client Traffic Example

Based on the very fast backups capable with SAN clients, fewer clients typically need to be backed up at a time. This means that a single SAN client media server can handle a large number of SAN clients if the backup window is large enough to accommodate the amount of data that needs to be backed up.

3.0 Moving Data: LAN and SAN

Data being backed up needs a pathway to get from the client to the media server and subsequently to the final storage point. This section discusses the two most common pathways for moving the data—the LAN and SAN.

3.1 LAN

As mentioned previously, most backups in a traditional backup strategy move across the LAN. NetBackup will, in most cases, pull data from clients as fast as possible without regard for issues created by overloading the LAN or CPU of the client. While it is possible to set up “throttling” of clients based on IP addresses, most customers should not use this capability because the purpose of backup is to get the data off the client as fast as possible. For this reason it is important to know when backups can run on which clients. We do not recommend doing a backup job for a very important production server in the middle of the work day because the backup will create performance overhead.

When architecting the backup solution, you must know the current capacity of the LAN and its ability to absorb backup traffic. Older switches are not capable of port aggregation on the media server, therefore anything more than 1 GbE may not be possible. Older clients may still be on 100 Mbit connections, therefore a single stream of data may come close to saturating the available bandwidth. Also, the connection between “edge” and “core” switches needs to be considered. Many edge switches—those that are capable of 12–24 connections that then feed into the larger director-style switches—may only have a 2 Gb connection between the switches. If the ports are GbE and all 24 clients that are plugged into the edge switch start backups at the same time, the link between the switches can be quickly oversubscribed.

In addition, if the LAN is already saturated with production traffic, adding the overhead of backup traffic can cause serious capacity issues. Some customers opt to install a second, dedicated backup LAN to handle backup traffic.

If the clients are all GbE, then the media servers will need to be sized on the LAN accordingly. For example, if the media server has 4xGbE NICs that are using port aggregation, then theoretically each media server can service the traffic of four 1xGbE clients. In actuality, they can do more depending on the speed each client can send data, so in the case of LAN clients the architecture should reflect the average speed of the clients (typically 15 MB/sec across GbE) rather than the theoretical speed of the GbE link. Using this figure, each media server with 4 GbE NICs can handle as many as 21 clients at a time (320 MB/sec average port aggregated speed / 15 MB/sec per client = 21.3 clients). By now you can better understand how bandwidth needs to be matched and how the media server needs to be capable of handling traffic from multiple systems.

If each client can run multi-streaming and send more data, then the number needs to be adjusted; however, using 15 MB/sec is a good estimate. This does not mean that for every 21 GbE clients you need to install a media server. It does mean that based on your backup windows and SLAs, each media server can service 21 clients at one time, therefore if each client requires an hour to back up, and you have a 12-hour window, each media server can handle 251 clients. In many cases, especially when weekend backups are being done as fulls—thus the backup window is longer—more clients can utilize the media server.

Finally, NetBackup can quickly overwhelm a LAN if there are capacity issues. While it is possible to “throttle” the traffic based on IP addresses, in most cases needing to do this indicates an issue with the LAN, which needs to be diagnosed and corrected.

3.2 SAN

Traditionally, SANs have been considered as high speed pathways to devices. HBA speeds have increased rapidly, and modern hardware is typically capable of supporting multiple 4 Gb HBAs. Faster SAN speeds—much faster than the 10 GbE LAN bandwidth increase—have been adopted due to a smaller price tag and the need for speed to faster devices, both disk and tape.

From a NetBackup perspective, the SAN is simply a way to move data, and properly configured media servers can take advantage of the large amounts of bandwidth between the various systems to move data to the final storage point very quickly. With the advent of PCIe and 4 Gb HBAs, the SAN is seldom the bottleneck in a backup strategy.

When architecting the SAN for backup traffic, it is important to note where the data is coming from as well as where it is being sent and then architect accordingly. Data typically resides on disk and needs to be moved to tape/VTL or to a different set of disks. This is usually done via the SAN if the LAN is not used. As noted previously, a SAN media server moves data from a SAN-attached disk to SAN-attached storage using the speeds of the SAN. SAN backups can be very fast if properly architected. If a system is a production system that cannot withstand the performance issues of a backup, then the SAN can also be used to perform snapshot or off-host backups where a proxy server is used to move the data. If this type of backup is wanted, you will need to consult other documents.

NetBackup offers the ability to share tape and disk drives across the SAN to be used for backups. This ability allows the master server to allocate SAN-attached tape drives and SAN-attached disk drives to the various media servers. This can provide a very high ROI on the backup infrastructure in addition to creating a highly available solution without having to run a clustering solution.

3.3 Scaling

Scaling is the ability to take an existing environment and grow it as the need to protect more data grows. From a NetBackup perspective, this occurs by adding hardware to the solution and making it available for use with NetBackup. Hardware must be added in such a way as to not create bandwidth bottleneck issues. This means that the entire solution needs to be understood before hardware is added so that the right hardware is added.

For example, assume there is a customer who is not meeting backup windows. A tape drive vendor tells the customer that by adding 10 more LTO3 drives they can achieve their windows (10 LTO3 drives can write a native combined total of 800 MB/sec). The customer adds the drives, but their windows do not change. This occurs frequently. Adding tape drives without adding media servers will typically not assist with reducing the backup window. Adding media servers without adding tape drives has the same effect.

Another example is adding LAN clients without increasing LAN capacity or the capacity into the media servers. If the LAN is already maxed out and more clients are added, the current backups will have a LAN bottleneck, which results in no performance gains. Many customers add servers to their data centers without considering the expense of adding media servers, SAN ports, high speed LAN ports, additional tape drives, etc. For example, a large database is needed so machines are procured, racked up and the system is powered up. If protecting that system has not been included in the budget, adding it to the backup infrastructure can cause other backups to fail even though they have been running successfully.

A backup strategy has a finite capacity for moving data. Once that capacity is reached, all backups begin experiencing performance issues. A backup success rate of 99.9% can quickly drop as new servers that require protection are added.

A proper scaling strategy is one that determines—based on current capacity—how much more can be added before issues arise. This allows backup administrators to know how much more data or how many more clients can be added before bottlenecks occur. Budgets can be developed in advance for additional or higher capacity tape drives and media servers, for an upgrade to 10 GbE for the media servers, or for higher capacity SAN connections.

Scaling should be addressed proactively and not reactively. This is the best way to maintain a high level of backup reliance. NetBackup is very easy to scale. A new media server can be installed and added to the master server's environment. If a SAN is being used, the media server can be zoned to see the tape drives, the hardware wizard on the master server can be run to see the devices attached to the new media server, a storage unit is configured (also using a wizard), and policies can be modified to begin using the new media server. Adding a media server to an existing NetBackup environment takes less than an hour once the hardware is configured and the SAN is zoned.

The general rule for scaling is simply to look at the entire environment and what is happening with capacity and speeds before deciding which hardware needs to be added. This can prevent the costly mistake of buying expensive new hardware that doesn't solve the problem.

4.0 Backup Storage Devices

Every backup environment needs some place to back up to. This section discusses some of the options as they pertain to development of the overall architecture. Not every option can be explored in this document as there are any number of places to store data that go beyond “basic” backup and recovery. Because this document was developed as an overview, only the basics are discussed.

4.1 Tape Drives

Most customers still use some tape drives even though disk has made inroads as a primary backup destination as noted in the next section. Tape drives and libraries have a number of advantages over disk, including the ability to be easily replaced if one goes bad without data loss, the ability to send tapes offsite, and the ability to expand capacity by removing full tapes and replacing them with blank tapes. The number one complaint about tape backup environments is the cost of the tapes themselves; however, reducing the number of tapes used can be a very simple process of changing the backup paradigm to not back up so much data all the time. While it is beyond the scope of this document to fully discuss backup paradigms, the days of “fulls on the weekends and incrementals daily” may be obsolete. Much of the data being backed up has RTOs that could withstand a single monthly full backup and weekly cumulative incrementals. Another option is deduplication, which is discussed later in this document.

The modern tape drive, such as the LTO3 and LTO4, are capable of writing data at 80–120 MB/sec, which equates to 288–432 GB/hr per drive. Very few disk arrays can match this speed, and if you have 10x drives, the numbers are 10x greater. While it is not a myth that disk-based backups are faster, performance and success depend on the existing configuration and the goals of the strategy as well as the type of disk being used, the SAN configuration, etc. In most cases, tape drives are not currently being driven at 100% capacity due to bottleneck issues. In this case, disk as a primary backup medium can be much faster.

When architecting the number of tape drives that are needed, the most important piece to keep in mind is the number of media servers to drive them. Attaching 16 LTO4 drives to a single media server will not provide faster backups and the ROI for the drives will be unsatisfactory. In general, when doing an initial architecture or performing a scaling exercise, matching the tape drive bandwidth with the media server and HBA bandwidth is required to get proper performance.

For example, an LTO3 drive can write at a native speed of 80 MB/sec and must write at least 30 MB/sec to prevent deteriorating performance from repositioning (called “shoe-shining”). A media server with a PCIe bus and 4 GB HBA can write about 400 MB/sec, therefore each media server could drive a total of 5 LTO3 drives to full capacity. If it’s more than that, then the drive bandwidth is wasted. When new tape drives are added, media servers must be added at the same time to match the bandwidth. Similarly, adding more or faster tape drives without increasing the number of media servers is not cost-effective.

Tape drives fail from time to time and those that run 24/7 fail more often than those that don’t. While trying to run with fewer drives and run them all the time increases ROI, it also increases failure rates. This means the drive needs to be replaced and while most modern robots provide hot-swap capability, sometimes systems need to be rebooted in order for the OS to recognize them. This can be an issue in a very busy production environment and/or where many SAN media servers are in use. In large environments the reboots can be scheduled during a normal maintenance outage; however, this is less than ideal. For this reason, developing an architecture with enough drives (and perhaps spare drives for restores) that does not require 24/7 use of the drives is more reliable, though more expensive.

As a final note on tape drive, modern tape drives can compress the data before it is written, therefore very fast speeds can be achieved when backing up to tape drives using modern media servers. Tape drives are not subject to the 70% rule and modern tape drives are seldom the bottleneck. While the decision of which tape drives to use in an environment is up to you, determining what speed is required and then adding drives and media servers to meet the needs will help you make the best decision.

4.2 Tape Libraries

Tape libraries are available in many sizes. While it is beyond the scope of this document to provide specifics, a general best-practice recommendation is to purchase a library can be scaled by adding more frames, and perhaps one that can be partitioned so that multiple master servers can use a single library if needed. Most modern enterprise libraries can be scaled and partitioned. The library should also be positioned in the data center in a physical location that allows for scaling. Having to move racks of servers by bolting them to a new frame is a very difficult, expensive, and time-consuming task. Placing the library in a location that is not “blocked” and provides easy access can prevent lots of extra work as your environment grows.

The expense of a library is in the robotic arm, therefore scalable libraries are initially expensive; however, adding a new frame with tape storage and/or drive bays is much less expensive than the initial cost. For this reason, buying a library that is adequate but that can be easily expanded is considered best practice.

4.3 Disk as a Backup Destination

Many customers are choosing disk as a backup destination. This is a very good way to perform fast backups and recoveries. There are many different options available for architecting disk into your data protection environment. Most are beyond the scope of this document. NetBackup 6.5 Disk Based Data Protection is a very good white paper that discusses the options for disk with NetBackup 6.5. For more information, visit:

http://www.symantec.com/business/products/whitepapers.jsp?pcid=pcat_storage&pvid=2_1

Performance of the disk depends on the type of disk and more importantly the layout of the target disks. In most cases, disk should be configured for write performance and the NetBackup administrator/architect must work closely with the hardware administrator to properly configure the disk.

Disk can be very beneficial as a primary backup medium when there are performance issues caused by slow data streams. Tape drives have performance issues when they are not supplied with an adequate stream of data. Disk drives do not have this issue. In environments that have slow clients, staging them to disk first can increase the ROI of the tape drives. It is also very beneficial to send numerous small backups—such as Oracle® archive logs—to disk rather than to tape. This saves the time needed to load and position a tape drive. For example, in one environment that was visited by Symantec Professional Services, the customer was running more than 20,000 backup jobs per day and had a 35% success rate. Approximately 15,000 of the backups were archive logs (32 MB files) running to tape drives. These backups required 2–5 minutes to load and position the tapes for less than a second of actual writing of data. When disk was located and these 15,000 backups were sent to disk, the success rate went from 35% to 98% overnight. The backups that needed the drives were available and the archive logs going to short-term disk storage took seconds rather than minutes.

4.3.1 Deduplication

Probably the largest challenge with using disk as a primary backup medium is capacity. A tape library can have its capacity completely renewed by pulling out full tapes and adding new blank tapes. This is not the case with disk. While there are options such as high-water marks and storage lifecycles, typically a considerable amount of disk space is needed when it is used as a primary backup medium. One resolution is to use deduplication when using disk as a primary backup medium.

Deduplication in the data center is available as a NetBackup bolt-on using PureDisk Data Center Option. When configured, backups are sent to a PureDisk storage unit and are “deduped.” For example, multiple copies of C:\Windows across the impacted clients are only backed up once. After the initial backup, only changes within the files are then backed up, therefore as the environment changes over time, the backups become smaller and smaller. While deduplication cannot help in every situation, backing up standard file systems can greatly reduce the amount of data being backed up while still allowing for fast restores. You can create archive sets, which take the data that has been deduplicated and copies it to tape for a specific client. These archive sets can be sent offsite for disaster recovery or longer-term storage.

For more information about NetBackup PureDisk, visit:

http://www.symantec.com/business/support/overview.jsp?pid=52672

4.3.2 Virtual Tape Libraries

Many customers have included Virtual Tape Libraries (VTL) in their overall architectures. The VTL is disk back end that emulates a tape front end. These backups can be faster than a standard tape-based backup due to the decreased time to load and position a tape; however, there can be capacity issues with VTLs as well. It is beyond the scope of this document to go into a lengthy discussion on VTLs; however, once you have determined your RPO/RTO/SLAs, you should know if a VTL is best for your environment.

5.0 Configuration and Management

Once NetBackup is installed, it must be configured and managed. This section explores the basic requirements for configuration and discusses management options to increase reliability and decrease the need for management.

5.1 NetBackup Policy Configuration

While it is beyond the scope of this document to dive deep into the NetBackup overall configuration, no architecture document would be complete without a basic understanding of how NetBackup is configured with regard to moving data. For this reason, this white paper briefly discusses the policy configuration as well as some configuration parameters that can impact performance.

The policy is the “Who, What, When, Where and How” of the backup strategy. Within the policy are the options that are used when a backup starts. There are various checkboxes within the policy GUI that impact what type of backup will be run. A Schedule tab indicates when the backup should run and whether it should be a full or an incremental backup and how long it should be retained. A Files tab indicates what is to be backed up. A Clients tab that indicates which physical clients will be backed up. Choosing the best configuration for your particular environment is critical to the overall success of the backup solution, therefore a thorough understanding of how the policy works is also critical.

In general, best practice on policies, from a high level, dictate the following:

1. Group “like” clients together to reduce the number of policies that are created:
    a. Too many policies can be difficult to manage, especially as growth occurs.
    b. For example, if you have 100 Windows systems that require everything on them to be backed up, create a single policy that backs up all of them and use ALL_LOCAL_DRIVES as the directives.

2. Depending on the SLA/RPO/RTO requirements, many systems can get by with a full backup bi-weekly or once per month with incremental backups daily (either cumulative or differential, depending on how much time is available for recovery). This can minimize resource use, including disk and tape, while still offering acceptable recovery times, especially if disk is used for the incremental backup storage. This will reduce the mount time needed for multiple tapes.
3. A number of options in the Policy Attribute screen can be used for complex backups such as snapshot and off-hosting. This requires a deeper understanding of the hardware requirements in most cases.

4. A thorough understanding of the check boxes on the Attributes tab is needed before configuring a backup for first-time administrators. For instance, you should never check the Compression box unless you are running backups across the WAN. It is much better to use native hardware compression on the hardware rather than software compression.

5. Policies are the most complicated facet of the NetBackup solution, therefore when creating policies, you should refer to the user guides to gain a deeper understanding of how the policy works and how choices can impact the overall backups.

5.2 Storage Units

The storage unit is a "virtual" pointer to physical hardware. For example, if a media server has four tape drives attached to it either directly or across the SAN, a storage unit is created that will use the four drives when a backup is pointed to that storage unit. NetBackup decides which of the four drives to use based on any number of factors including MPX settings, number of clients in the policy, etc.

There are many different types of storage units available within the NetBackup configuration, depending on which options have been purchased and which type of hardware is in place. The type that is used depends on the requirements of the overall environment.

Information about the various types of storage units and when to use each can be found in the NetBackup 6.5 manuals.

5.3 Recovery

The purpose of any data protection solution is the ability to recover the data that has been backed up. Recovery with NetBackup is very easy. A separate GUI is started typically off the main GUI, the Client that backed up the data is selected, and each backup that has been run for that Client can be drilled into to find the files to recover. Date ranges can be specified if desired to narrow down the results. If a full systems recovery is needed, the built-in Bare Metal Recovery (BMR) can be used if this option was selected during the backup cycle.

As a solution is developed, an eye to recovery must be considered. If only enough tape drives to perform backups are configured and they run 24/7, then when a restore is needed there may not be drive resources available to the restore process. For this reason, when developing the number of tape drives that are needed, one or two additional drives should be added and reserved for restores. If the environment will not do a lot of restores, or if restores occur during the day and backups occur at night, then this may not be an issue.

5.4 Personnel

No discussion about architecture and bandwidth is complete without discussing the administrators who are in charge of the environment; however, the number of administrators required for any given environment is a difficult number to establish because there are many factors that influence the overall number. Small, but very complex environments doing snapshots, agent backups, offhost backups, etc. may require more people to manage the solution than a much larger environment doing simple backups.

Most environments grow over time, but IT personnel headcount typically does not grow in the same proportions. Many times failing backups can be attributed to NetBackup administrators who have limited time to manage the solution and have even less time or no time to troubleshoot issues to achieve higher success rates.
In general, data protection FTEs are based many factors, including but not limited to:

- Number of servers under management
- Amount of data under management
- Is the environment 24/7/365
- RPO/RTO/SLA requirements
- Is there an operations staff for monitoring

Typically, for each 8-hour shift there should be an 850:1 ratio of clients to backup staff. For every 850 clients configured to use NetBackup, one FTE is required; however, every environment should have at least two FTEs unless it is a very small environment so that one person is not on call all the time.

Most large enterprise environments will have failures based on situations beyond the control of the administrator. If there is not a support staff (such as night time operations staff) to retry jobs, the administrator could be awakened at all hours or need to troubleshoot on weekends. This could quickly lead to staff burnout and personnel turnover. There should always be more than one person who understands the backup software and can set up new clients, configure new policies, troubleshoot issues, etc. Otherwise, during personnel turnover, the data protection environment could be at risk.

5.5 Agents

NetBackup has the ability to back up databases using agents that work directly with the database software to make sure the backup is constant and reliable. For example, the Oracle agent interfaces with Oracle RMAN to perform operations such as placing the database in Hot Backup mode, or quiescing the database for a cold backup. The data is then sent to NetBackup via RMAN with NetBackup simply acting as the conduit for a tape or disk device.

Using agents is a very effective way to perform backups while maintaining a standardized backup solution. Many times databases are backed up using custom scripts created by the local DBA; however, problems arise when the databases grow, or the DBA leaves the company and the new DBA does not understand the custom scripts.

NetBackup offers agents for most databases and has options that allow the DBAs to continue to control the backup or gives the control to the backup administrator. One issue with non-agent database backups is that scheduling can become a challenge especially in busy environments. If a DBA needs a backup run but all the tape drives are in use, this can cause the database to fail. In some environments with critical database applications, media servers and tape drives can be dedicated to the database if needed.

5.6 Reporting and Monitoring

Proper monitoring of a data protection solution and the ability to produce reports are critical to the success of the strategy. For companies that have audit requirements, the ability to pull reports that prove that data is being protected can make a huge difference. Monitoring the environment so that failures are managed proactively is also critical to meeting SLAs.

From a monitoring point of view, NetBackup includes the NetBackup Operations Manager (NOM). NOM provides a Web-based view into the backup environment and has drill downs to backup and restore jobs as well as an insight into hardware, which allows an operations staff to manipulate drives, stop and restart backups, etc.

Reporting is done using a separate purchased utility called Veritas™ Backup Reporter. Backup Reporter allows for reporting on the environment as well as trending and can offer the ability to produce reports that show the “heavy hitters” in an environment. This can be beneficial for budget cycles to show the capacity of the solution vs. the needs of the solution so that budget can be properly allocated.

Backup Reporter can also help when revising SLAs, RPO and RTOs. For example, if you have historical data on how much the data protection solution is costing (which Backup Reporter can provide) and a group wants to increase the amount of time data will remain on tape, it is easy to extrapolate how many tapes will be needed. This way you can go back to the group with the information about how much extra it will cost to be able to meet their new SLA. Because most data protection strategies are on a shoestring budget, being able to show how much a specific SLA costs is a very good way to either reduce the SLA or get additional funding for it. This is a “chargeback model” that many large customers have found to be very beneficial.

For more information about NetBackup Operations Manager and Veritas Backup Reporter, visit the Symantec Support website.

6.0 Summary

Developing an architecture for data protection is “simple” math.

1. Determine the SLAs, RPO and RTOs.
2. Calculate how much data and the number of clients to be protected.
3. Determine the backup windows.
4. Apply hardware with the appropriate bandwidth to meet the requirements.
5. Try to fit the above requirements into the budget and renegotiate with the users as needed.

Modern servers and other backup hardware (tapes, disk, etc.) make the choices of which server/hardware to buy much less critical than they were 18 months ago. Tape drives are much faster and can store more data than ever before, and disk options are also growing; however, even with all the advances in technology the data protection paradigm is shifting away from standard day-to-day backups to deduplication, continuous data protection, and snapshots to name a few. It is simply too difficult to send petabytes of data from disk to tape and save it for an infinite amount of time, and many Symantec customers have reached and passed the petabyte mark. As data grows, the paradigm of how to protect it may need to change.

A strategy, if properly architected, can achieve a very high success rate with minimal FTE requirements. Conversely, a poorly designed backup solution will continue to operate poorly regardless of how many FTEs are there are.

About Symantec

Symantec is a global leader in infrastructure software, enabling businesses and consumers to have confidence in a connected world. The company helps customers protect their infrastructure, information, and interactions by delivering software and services that address risks to security, availability, compliance, and performance. Headquartered in Cupertino, Calif., Symantec has operations in 40 countries. More information is available at www.symantec.com.

For specific country offices and contact numbers, please visit our Web site. For product information in the U.S., call toll-free 1 (800) 745 6054.

Symantec Corporation
World Headquarters
20330 Stevens Creek Boulevard
Cupertino, CA 95014 USA
+1 (408) 517 8000
1 (800) 721 3934
www.symantec.com