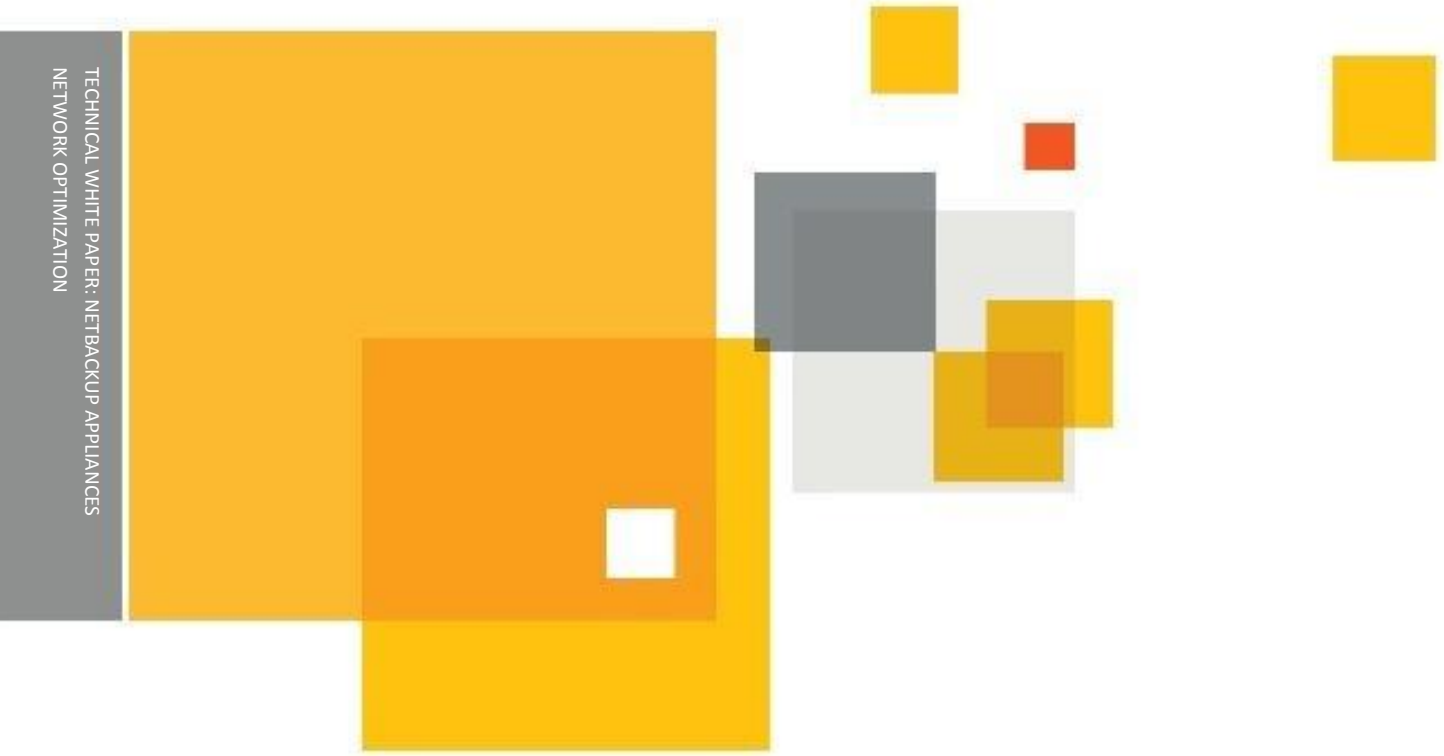


Technical White Paper

NetBackup Appliances
Network Optimization

Who should read this paper

Technical White Papers are designed to introduce Symantec partners and end users to key technologies and technical concepts that are associated with the Symantec Backup and Recovery product family. The information within a Technical White Paper will assist partners and end users as they design and implement data protection solutions based on Symantec Backup and Recovery products.

Technical White Papers are authored and maintained by the Symantec Backup and Recovery Technical Services group. If you have any feedback or questions about this document please email them to DL-SYMC-ASK-SES-NBU@symantec.com stating the document title.

This document outlines the WAN Optimization feature enhancements introduced on the NetBackup 5220 and NetBackup 5020 and applies to:

- NetBackup 5220 & 5230 appliances with version N2.5 and above installed
- NetBackup 5020 & 5030 appliances with version D1.4.2 and above installed

Contents

Introduction

In a world of ever increasing data flow as well as globalization of data centers the effectiveness and utilization of the networks connecting sites is of the highest importance to end users. Even with network enhancement and improvement, the ability of the infrastructure to keep pace with the flow of data has proved not to be in lockstep. To optimize the flow of data verses increasing the pipe that is flows along is seen as critical to keeping operations running and costs minimal. This paper discusses the new WAN Optimization technology that has been introduced in the NetBackup 5220 and 5020 appliances.

Business Value

Local Area Networks and Wide Area Networks

Modern TCP/IP based networks are segregated into one of two forms. Internal and local to a customer data center is the Local Area Network (LAN). External to the data center and leveraged to enable scale and resiliency is a Wide Are Network (WAN). Both networks have three primary characteristics being bandwidth, latency and loss. Corporate networks usually are comprised of LAN's that interact with WAN's to ensure that data is available to a level that allows the business to sustain operations.

Bandwidth

The bandwidth of a network, be it LAN or WAN gives a measurement of the total amount of data that could pass through the network from one point to another. Generally, the network could be a combination of different networking technologies based of Ethernet (100Mbs, 1Gbs, 10Gbs, Infiniband) or other transports (ASL) with various capacities for data transportation - bandwidth. The bandwidth of most corporate networks in the US is general segmented at the 100Mb (megabits) per second range with a WAN performing up to this level. A LAN should be capable of sustaining 100Mb/s up to the theoretical maximums of the leading edge technologies available today.

Latency

The latency of the network gives a measurement of how long a packet would take to be transmitted and returned from one system to another (commonly referred to as Round Trip Latency). Latency is measured using tools like 'ping', which mitigates a lot of processing at the receiving end and therefore skewing the results of the network testing. Another element that plays into latency is that most networks are not a direct connection. The common network environments in the data center is constructed of many devices (switches, routers, gateways, firewalls) which all add some overhead to the packets that are passing through them. The loss is amplified as the network is expanded over a geographical expanse, which is common to a WAN. Typically the latency is measured in milliseconds (msecs) and can range from nearly 0 through to more than 150msecs. The two tables below show the general constraints of a LAN or a WAN as well as some examples of what to expect as the transportation expands across continents and around the world.

Latency	Typical Infrastructure Layout
0 to 20 msecs	Local Area Network
20 to 70 msecs	Intra-Continental Wide Area Network
70 to 150 msecs	Inter-Continental Wide Area Network
Greater than 150 msecs	Extremely Remote Wide Area Network

Table 1: General Latency Guidelines

Locations	Latency
New york to Los Angeles	50msec
New York to London	70msec
Los Angeles to London	110msec
New York to Tokyo	150msec
Los Angeles to Sydney	150msec

Table 2: Example Latency from various global locations

Loss

Network loss is a concept that explains the degradation of network connectivity and the resulting performance forfeiture. Even the best networks can encounter network loss but generally speaking the more dispersed a network is the higher the rate of network loss. Network loss is amplified in a geographically dispersed network due to the fact that most networks are TCP controlled verses UDP and this results in the need to retransmit packets that were received in error or not received at all – the further the distance these retransmitted packets need to travel the more exposure to network degradation. A UDP network doesn't require this but the application that uses this type of network management needs to be capable of handling data loss – media streaming for example. A well performing network would loose or have to have packets retransmitted every 1 in 10000 (0.01%) packet transmissions. A poorly configured or slow WAN this number would be seen as being 50 in 10000 (0.5%).

Other Miscellaneous Characteristics

Network Hops

As mentioned previously the amount of 'hops' in a networks can dramatically affect the performance of the network. Also if the connections in the 'hops' are not designed for performance (i.e. for a packet to get from New York to Los Angeles it should route through Chicago and not London) there will also be performance issues.

Shared infrastructure

In a WAN environment the networking infrastructure is generally not owned or controlled by the customer. A communications company owns the infrastructure, which is then sold on the available bandwidth to various customers. These networks can be throttled either based upon the contract between the customers and the communications provider or by the general utilization of the available networks. An example of shared utilization is if 20 customers may be leveraging the same 100Mb/s connection but one of the customers is occasionally (or continually) using 80Mb/s leaving only 20Mb/s for the other 19 customers.

Jitter

Jitter is the interference of the physical network due to outside influences. This could be electromagnetic interference (EMI) or faulty hardware. The result of a jittery network is many errors received by both transmitting and receiving devices as well as heightened amounts of retransmissions (that could also be effected by the cause of the jitter).

Protection Methods and Technology

NetBackup and Networking

Backup applications leverage both LAN and WAN infrastructure to capture data on systems deployed across the datacenter as well as at remote locations. This data then is then routed to a media server where is will be stored on tape, disk or other storage device for later use. NetBackup has a three-tier architecture that encompasses the control (Master Server), storage management (Media Server) and source (Client). A datacenter may have all three elements or some. In figure 1 below the environment shown demonstrates an architecture where there are two sites. The server connectivity in blue is via LAN and those in red via WAN. Control of the environment is via a Master server located in the data center. Data is routed from the remote clients to a media server in the data center via the WAN. This allows for a copy to be located on the remote site (in the local NetBackup 5220 with Disk Storage) as well as on tape via the media server in the media center.

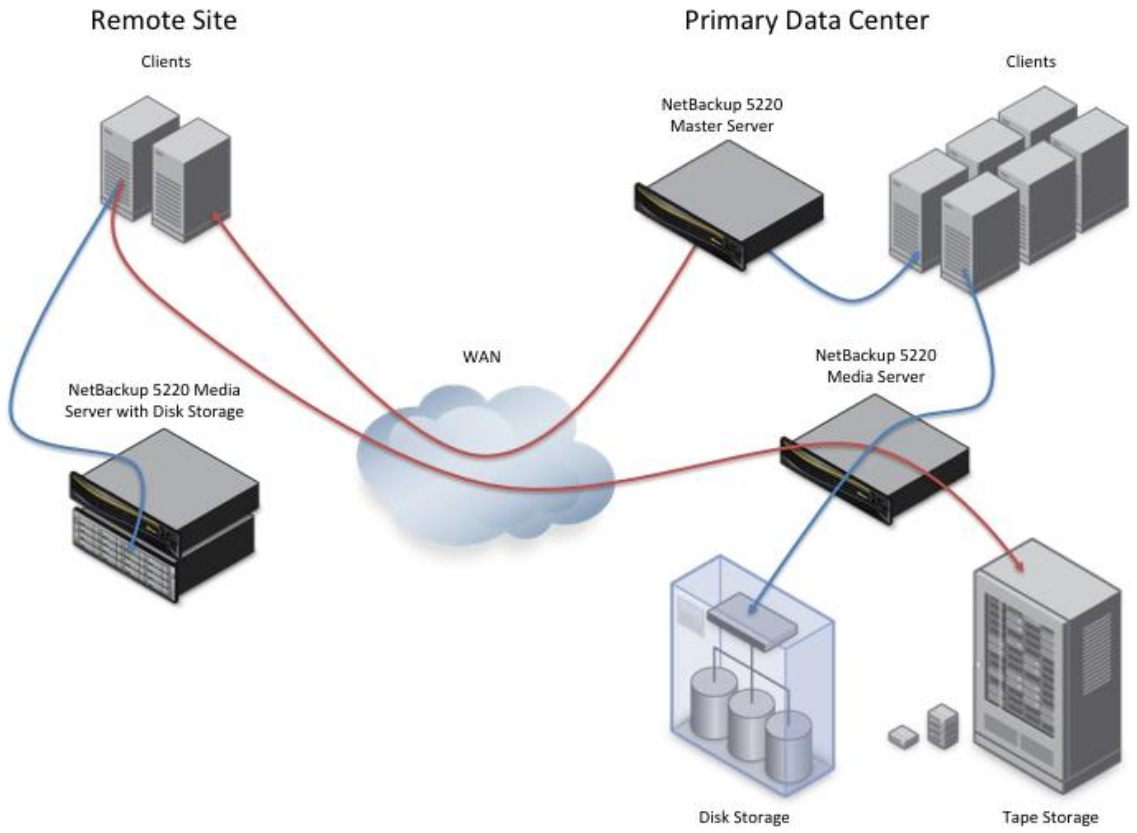


Figure 1: NetBackup Architecture showing Datacenter plus Remote Site

NetBackup leverages TCP/IP communication to provide instructions as to when to commence a backup as well as where the data should be stored; ensuring that the network connectivity between all elements in the environment allows for successful completion of the backups in a timeframe that is applicable to the requirements of the application owner.

Optimizing the Network Interaction

Network management is difficult. The more the networks converge (more LAN's attached to more WAN's) the overhead in keeping the infrastructure growing and running optimally if difficult. Systems interacting with the networks are best severed in being adaptable to present network as opposed to requesting upgrades and changes to increase bandwidth to satisfy the needs of the applications. Symantec has researched and implemented in the NetBackup appliances technologies to enable **WAN Optimization**. This technology works to increase the success rate as well as decrease the backup window but working together is best for the overall performance of the NetBackup environment.

WAN Optimization

Overview

Symantec implemented WAN Optimization technology in the 2.5 release of the NetBackup 5220 as well as the 1.4.2 release of the NetBackup 5020 appliances. The NetBackup 5220 is based upon core NetBackup technologies and can be leveraged in an already implemented NetBackup environment or in a new environment. The 5220 can be configured as a Master Server, a Media Server or as a combination of both. The 5020 appliance is a deduplicated storage appliance that is best suited to environments where storage the minimal amount of backup data is a requirement.

WAN Optimization is a form of PDCC (Packet Delay Congestion Control) technology. PDCC technology monitors and fine tunes network traffic as a method of optimizing network throughput. PDCC technology is analogous to urban freeway traffic metering and congestion control. Though unlike freeway traffic, network traffic has additional unique challenges.

First, unlike freeway traffic that will backup and wait indefinitely, network traffic is subject to timeouts and dropped/lost packets. Resending lost/dropped packets is costly and time consuming, especially for smaller capacity / long distance WANs.

Second, unlike urban freeway traffic, which is under centralized control, WAN traffic over public networks is not under centralized control. PDCC technology adjusts to weaknesses in traditional chaotic WAN traffic patterns, and minimizes timeouts and dropped/lost packets for its managed connections. This in combinations with other more traditional PDCC packet monitoring techniques allows the NetBackup appliances to optimized network traffic.

Implementation

Symantec has implemented the WAN Optimization technology in the NetBackup appliances only. It has been implemented by way of a filter driver in the underlying kernel used in the appliances. The technology, when enabled intercepts and manages TCP traffic with minimal processing overhead. When disabled the traffic completely bypasses the WAN Optimization technology.

There are no tunable parameters as the technology leverages completely heuristically driven logic.

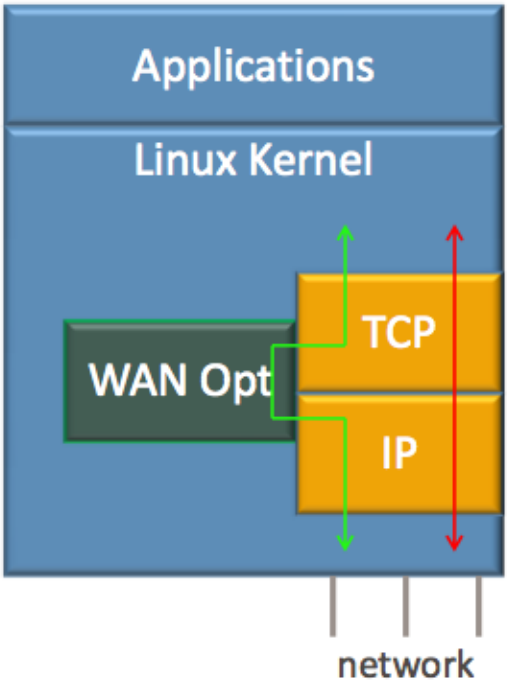


Figure 2 Overview of WAN Optimization Implementation

Enabling and Disabling WAN Optimization

To configure WAN Optimization is easy – it is either on or off. From the Command Line Interface (CLI) in the 5220 or 5020 appliances navigate to the `WANOptimization` menu and select `Enable`. To disable run `Disable`.

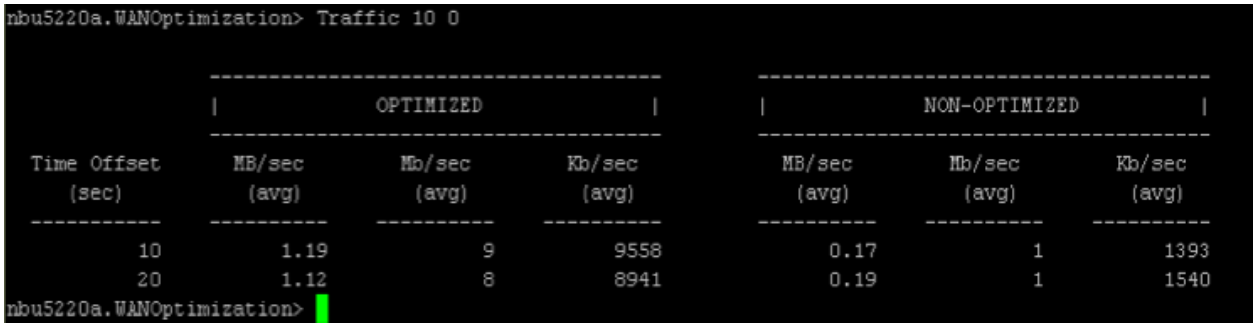
```
nbu5220a.WANOptimization>
nbu5220a.WANOptimization>

DebuggingInfo   Display Debugging Information.
Disable         Disable Network Optimization.
Enable          Enable Network Optimization.
Exit            Logout and exit from the current shell.
Return          Return to the previous menu.
Shell           Shell operations.
Status          Display Network Optimization Report.
Traffic         Display Network Throughput.

nbu5220a.WANOptimization>
```

Figure 3: Enabling WAN Optimization from the Command Line

To monitor the effectiveness of WAN Optimization, from the `WANOptimization` menu run the `Traffic` command. This will show both the optimized and standard traffic flow.



```
nbu5220a.WANOptimization> Traffic 10 0

                 ------------------------------------       ------------------------------------
                 |             OPTIMIZED            |       |            NON-OPTIMIZED         |
                 ------------------------------------       ------------------------------------
  Time Offset      MB/sec        Mb/sec        Kb/sec          MB/sec        Mb/sec        Kb/sec
     (sec)          (avg)         (avg)         (avg)           (avg)         (avg)         (avg)
  -----------    ----------    ----------    ----------       ----------    ----------    ----------
           10          1.19             9          9558             0.17             1          1393
           20          1.12             8          8941             0.19             1          1540
nbu5220a.WANOptimization>
```

Figure 4: Monitoring WAN Optimization from the Command Line

Note: There is no change in WAN Optimization driver implementation with latest NetBackup appliance 5230 running software release 2.6.0.1.

Case Studies

Case Study 1 - Optimized Backup and Recovery

Sites: Roseville, Minnesota to Mountain View, California

Use Case: Optimized Backup and Recovery

Details: Both sites were provisioned with a standard NetBackup 5220 appliance. Backups from the Roseville site were replicated to the Mountain View site using NetBackup Auto Image Replication (AIR). The infrastructure between both sites was regarded as good in that there was only 10 hops, a latency of 54 milliseconds with less then 0.01% loss. On the Roseville site a single 1GB file was backed up 5 times. The backup was replicated to Mountain View. Text (Calibri 11)

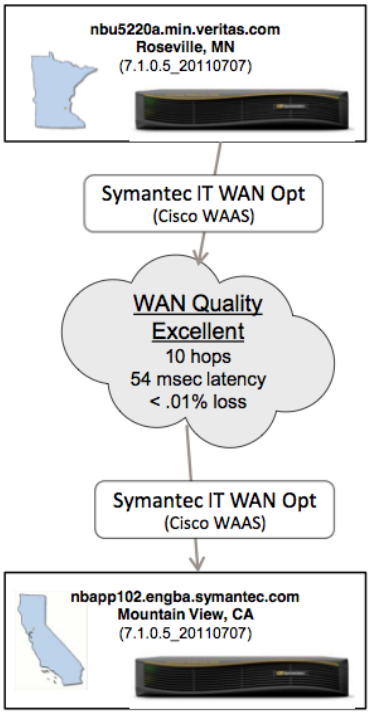


Figure 5: Outline of Case Study 1 Environment

The results of the tests were as follows:

Test Number	With WAN Optimization	Without WAN Optimization
1	43.6 Mb/sec	24.3 Mb/sec
2	50.1 Mb/sec	23.0 Mb/sec
3	55.1 Mb/sec	25.3 Mb/sec
4	63.2 Mb/sec	19.7 Mb/sec
5	55.7 Mb/sec	25.9 Mb/sec

Table 1 Optimized Backup and Replication Case Study Test Results

Results: It can be seen that the performance increase with WAN Optimization is greater than 100%.

Case Study 2 - Optimized Backup to the Cloud

Sites: Roseville, Minnesota to Nirvanix Cloud

Use Case: Optimized Backup to the Cloud

Details: Backing up to the cloud is becoming increasingly popular as the model stabilizes both with regard to costs and technology. Symantec has the ability, by using the Open Storage Technology (OST) to back up directly to the cloud. In this example the Nirvanix cloud was leveraged. NetBackup is presented with some storage from Nirvanix to place backup data on. The System used is the same as in case study 1 – a NetBackup 5220 Appliance. The network infrastructure again was very good in that there were approximately 8 hops, latency of 51 milliseconds and about a 0.05% loss.

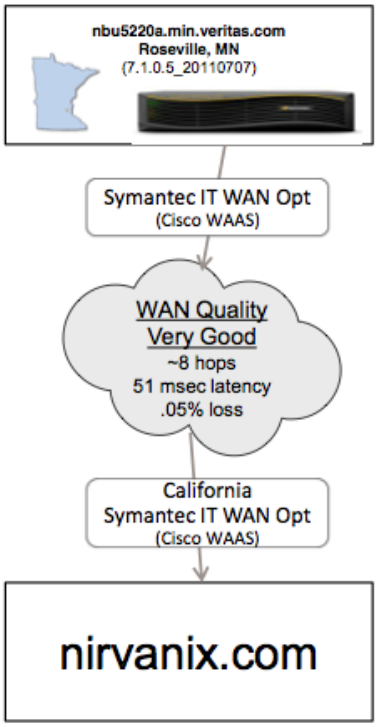


Figure 6: Outline of Case Study 2 Environment

Test Number	With WAN Optimization	Without WAN Optimization
1	5.421 Mb/sec	2.038 Mb/sec
2	3.912 Mb/sec	1.982 Mb/sec
3	4.533 Mb/sec	2.042 Mb/sec
4	1.845 Mb/sec	1.301 Mb/sec

Table 4 Optimized Backup to the Cloud Case Study Test Results

Results: The performance is more variable then in case study 1, but the results are still good – 50% - 150% performance increase.

Case Study 3 - Lab generated results to show variability and utilization of both WAN and LAN

Sites: Internal to the Roseville, Minnesota site (Simulated WAN)

Use Case: None – Lab generated results to show variability and utilization of both WAN and LAN

Details: A network was developed that could be altered to simulate various WAN and LAN environments. This was achieved by throttling various devices to mimic high network usage as well as high latency and loss.

Test number	Bandwidth (Mb)	Latency (msec)	Loss (%)	With WAN Optimization (Mb/sec)	Without WAN Optimization (Mb/sec)	Benefit
1	45Mb	50	0	50.6	11.4	3.42x
2	45Mb	25	0.1	24.7	9.6	1.56x
3	45Mb	50	0.1	12.4	5	1.47x
4	45Mb	100	0.1	6.5	3.9	0.68x
5	45Mb	25	0.5	33.5	6.6	4.07x
6	45Mb	100	0.5	4.6	1.2	2.85x
7	1Gb	50	0.01	99.9	25.6	2.90x
8	1Gb	50	0.1	57.7	8.4	5.85x
9	1Gb	100	0.1	37.2	4.2	7.84x

Table 5 Results of testing on simulated degraded network

Results: From the tests above it can be seen that there is benefit in all cases. Some of the results are minimal (approximately 0.68x increase) but some can be seen as significant increases (approximately 7.84x increase). Most of the results fall in the 2 to 5x benefit range. Generally speaking it is shown that for a WAN the meaningful benefit is when latency is more then 50milliseconds and/or a loss of more then 0.1%. LAN's also see benefit when latency is greater then 20 milliseconds and/or the loss is greater the 0.01%.

Improving the NetBackup Experience

The net result of the implementation of WAN Optimization feature is that of decreasing the amount of failed backups. Providing intelligence in the NetBackup Appliances that allows for the backup product to enable more throughput where needed as well as manage disconnections and unstable networks will enable customers to achieve higher successful backup completion rates. As most customers are measured against a service level agreement that penalizes them when failed backups occur the features discussed in this paper will go some way to mitigating these consequences.

The features are at the core of the drive to make the implementation of NetBackup Appliances simpler and less taxing on time and resources.

Summary

Backup technologies are great at finding issue with networks be them WAN or LAN. The amount of stress that they can place on the networks they leverage can lead to performance and stability issues. Historically the solution was to 'get a bigger pipe' but as budgets diminish this is not the answer in most organizations. Building intelligence into the backup application to enable on the fly management with no oversight allows for customers to use the infrastructure that they currently have in a more efficient manner.

Additional Resources

Link	Description

About Symantec

Symantec is a global leader in providing security, storage, and systems management solutions to help consumers and organizations secure and manage their information-driven world. Our software and services protect against more risks at more points, more completely and efficiently, enabling confidence wherever information is used or stored. Headquartered in Mountain View, 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 website.

Symantec World Headquarters
350 Ellis St.
Mountain View, CA 94043 USA
+1 (650) 527 8000
1 (800) 721 3934
www.symantec.com

Symantec helps organizations secure and manage their information-driven world with data backup and recovery software.