HyperIP®: Network Appliance Snapshot Applications Note

SnapMirror, SnapVault, SnapMover, SnapRestore

Introduction

HyperIP® is an appliance that quantifiably and measurably enhances large data movement over high bandwidth (10Mbps to OC12) and lossy (packet loss) IP networks. Any IP network with packet loss is a candidate for HyperIP, regardless of distance.

HyperIP is a working example of RFC3135. RFC3135 describes techniques used to mitigate TCP performance problems over long-distance wide-area networks. These techniques are called "TCP Performance Enhancing Proxies" (PEP). HyperIP is the most recent iteration of a production-hardened transport, deployed in hundreds of Fortune 1000 accounts, that shields TCP applications from performance degradation due to packet loss, latency, and jitter.

Network Appliance Snapshot applications are the gold standard for business continuity and disaster recovery solutions. Snapshot applications include SnapMirror, SnapVault, SnapMover, and SnapRestore.

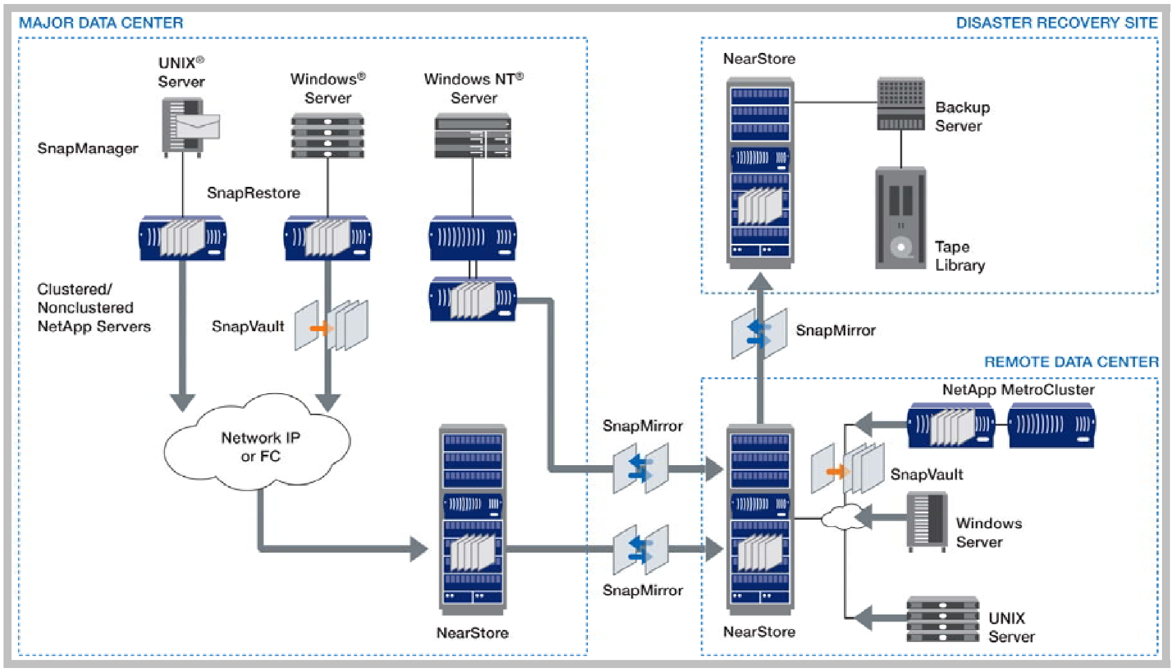

Figure 1: NetApp Logical Diagram for SnapShot Applications

Snapshot applications work over native TCP/IP backbone networks. In many cases multiple Snapshot applications run concurrently, replicating data over the standard TCP/IP WANs to both remote and disaster recovery sites.

Unfortunately, the typical throughput achieved by TCP applications on IP networks over 300 miles and bandwidth at 10Mbps and above, is dismal. The throughput continues to degrade as packet loss increases (a result of bit errors, congestion, and router buffer overflows, etc.). These conditions lead to higher retransmission rates. The net effect is much lower throughput for Snapshot applications. The fault lies squarely with TCP, which was not designed to handle data transfers over high-speed, long distance networks. HyperIP is designed to provide Network Appliance Snapshot applications with the throughput it requires by mitigating the limitations of TCP.

TCP Performance

Several characteristics of TCP cause it to perform poorly over networks characterized by any combination of high bandwidth, packet loss, and long distances:

TCP bandwidth sharing

When individual TCP applications compete for the same bandwidth, congestion is created. As new TCP applications are started, this new workload creates additional demand for the same bandwidth, which initially creates more congestion. The "bandwidth sharing" technique available in TCP is for each application's TCP connection to attempt to send data faster and faster until packet loss occurs, and then back down.

TCP congestion control

Packet loss occurs when packets are damaged or when the capacity of an intermediate network component is exceeded, which causes packets to subsequently be discarded. TCP interprets packet loss as a sign of congestion. The amount of data that TCP can send is flow-controlled by the size of its congestion window. When congestion is detected, TCP dramatically reduces the size of the congestion window and the rate at which it will be allowed to grow. TCP slowly increases the size of the congestion window, but if additional packet loss or packet corruption occurs, TCP will again dramatically reduce the size of the congestion window. If the congestion window is repeatedly smaller than the available network capacity, it will be impossible for TCP to ever fully utilize that capacity.

TCP window scaling

To utilize the full available bandwidth of a data session, enough data must be sent to "fill the pipe". The amount of data that can be "in the air" at any point in time is governed by the window sizing capability of the TCP stacks on both sides of the connection, and the applications. Most TCP implementations (including Network Appliance) support a feature called window scaling, which allows large windows to be used. However, in Snapshot environments consisting of SnapVault clients running on other application servers, this capability may not exist, or it may not be reasonable for the user to tune the TCP stacks on those servers.

Window scaling may still not provide a large enough window, or the server or application may not have enough buffer space available to support the capacity of the network. For example, the total round-trip time for a 3000 mile connection is approximately 60 milliseconds, creating an available data "pipe" at 100 Megabits per second (Mbps) of 750 Kilobytes. A satellite connection (540 milliseconds round-trip time) at 45 Mbps creates a 3 Megabyte pipe. The actual "pipe" will be further affected by latency induced by network hops between application servers. If window scaling is not implemented, or is not tuned to the network requirements, there may still be a large amount of "empty pipe", i.e. unused, but still available, bandwidth.

Implications for Network Appliance Snapshot replication applications on TCP/IP WANs

New regulations, business continuity, and disaster recovery has led to a surge of storage-to-storage replication applications over the WAN. Man-made (9/11, blackouts, human error) and natural (hurricanes, earthquakes, tornados, firestorms) disasters have driven demand. TCP/IP is the protocol of choice for driving data replication over the WAN. There are three reasons for this:

1) The market perceives bandwidth is essentially free. This is because the TCP/IP WAN bandwidth already exists for interactive traffic. Conventional wisdom is that Network Appliance Snapshot replication occurs at night or on weekends. This is when the majority of users are not utilizing the network, thus allowing already existing TCP/IP bandwidth to be leveraged by the Network Appliance replication applications without negatively impacting current applications.
2) The TCP/IP WAN infrastructure already exists, and is viewed as a managed, shared resource. Dedicated, separate Network Appliance Snapshot replication WANs are not required.
3) Additional bandwidth implemented for the Network Appliance Snapshot replication applications can be shared by the interactive TCP/IP applications.

The facts show that TCP/IP bandwidth is neither free nor is there typically enough to accomplish the storage replication in the window of time allotted. TCP/IP bandwidth utilization, packet loss, and latency issues are rarely taken into account in calculating bandwidth requirements. The most likely result is a bandwidth shortfall. This means either the storage replication cannot complete within the window of time allotted, or the user must buy more bandwidth - otherwise known as a conundrum.

The cost effective solution: Network Appliance Snapshot Applications with NetEx HyperIP

HyperIP is designed specifically for applications that transfer large amounts of data over high bandwidth WANS. It is a highly efficient data transport, especially in cases where packet loss, congestion and/or latency exist. HyperIP is a standard TCP/IP network node requiring no modifications to LAN/WAN infrastructures. It provides transparent "acceleration" across high bandwidth WANs.

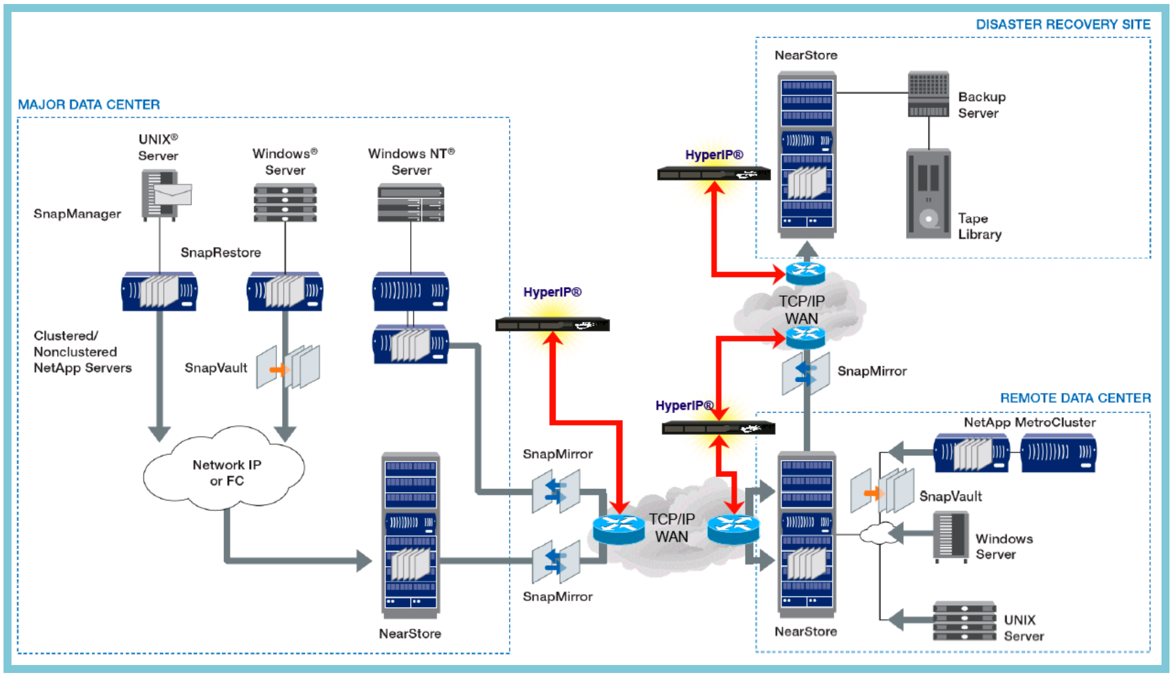

Figure 2: NetApp and HyperIP Logical Diagram (note: HyperIP is implemented in redundant 'paired' configurations

HyperIP Performance

HyperIP is designed to provide the Snapshot applications with the throughput they require by mitigating the effects on performance caused by any combination of packet loss, latency, and jitter. The key HyperIP benefits include:

HyperIP bandwidth sharing

HyperIP serves as the focal point for efficiently managing packet streams over the WAN. When it is the sole driver of the WAN, it has knowledge of the WAN capacity and sustainable link rate.

HyperIP manages multiple LAN Snapshot packet streams, and aggregates them over the HyperIP network. As new Snapshot application connections are started, HyperIP is able to accommodate the additional workload by inserting the new packet stream into the HyperIP connection without creating congestion. As Snapshot applications are stopped, the additional bandwidth capacity is automatically reclaimed by HyperIP for sharing among the remaining Snapshot connections.

This is particularly beneficial in Snapshot environments that consist of multiple NetFilers and/or multiple SnapVault servers, each performing periodic replication activities. From a Snapshot application standpoint, every replication connection that gets started effectively becomes a TCP connection to the local HyperIP appliance. The packet streams from each new Snapshot connection are aggregated with other Snapshot connections already in progress, and the data is carried over the already established HyperIP connection.

HyperIP compression

HyperIP also has the ability to compress data before it is sent over the WAN. Depending on the compressibility of the data, this may reduce the WAN bandwidth usage, and effectively increase the application data throughput over the network. HyperIP can provide compression benefits at rates of up to 155 Mb/s after compression.

HyperIP congestion control

Most networks have some packet loss, usually in the range of .01% to .5% for optical WANs, and from .01 to 1% for copper based TDM networks.

Just as with TCP, HyperIP cannot distinguish between packet loss due to congestion and packet loss due to corruption. However, rather than making window adjustments when packet loss occurs, HyperIP uses rate-based congestion controls to limit throughput. Regardless of whether packet loss is due to congestion or corruption, HyperIP is able to drive throughput up to the capacity of the WAN. HyperIP shields the Snapshot TCP connections from performance variations due to packet loss and latency on the WAN, since the performance over the WAN is managed by HyperIP.

HyperIP window scaling

HyperIP dynamically calculates round-trip times, bandwidth capacity, and transmission rates, and uses that information to calculate the capacity of the network. In contrast to the window scaling issues of native TCP applications, the HyperIP window size is dynamically calculated based on the capacity of the network. This calculation remains totally independent of any application TCP stack capability. This is particularly beneficial in SnapVault environments where application servers may not be running with tuned TCP stacks. With HyperIP, this just doesn't matter, since the server TCP performance capability only affects the LAN segment between the application server and the local HyperIP appliance. The performance gain over the WAN is achieved by HyperIP.

HyperIP's incredible Snapshot test results over IP WANs

Testing results with Network Appliance Snapshot applications have been outstanding. Network Appliance Snapshot applications achieved bandwidth utilization consistently exceeding 90% from distances of hundreds to thousands of miles (even with high packet loss simulation).

What this means to end-users

Running Network Appliance Snapshot applications on native Ethernet TCP/IP fabrics with HyperIP now provides the highest possible throughput option. Time windows for Network Appliance Snapshot replication can now be met. The promise of free bandwidth for Network Appliance Snapshot applications may just turn out to be a reality.

Summary and Conclusion

HyperIP provides the highest possible throughput for Network Appliance Snapshot applications over TCP/IP WANs. It does this with a "Production Hardened Transport" that provides:

- Bandwidth utilization that consistently exceeds 90+% effective data throughput regardless of packet loss from bit errors, congestion, or jitter,
- 2:1 to 15:1 compression,
- Reduction of TCP/IP network latency,
- Network Appliance Snapshot application aggregation.

HyperIP shields the Snapshot applications from TCP performance degradation caused by packet loss, latency, and jitter. HyperIP provides a highly efficient network connection that can provide the WAN connectivity for multiple Snapshot applications. This may allow Snapshot replications to be configured to meet growing business requirements. For example, it may enable replications to be scheduled more often, or it may even enable more servers to participate in Snapshot replications to remote BC / DR sites.

Network Appliance Snapshot applications over TCP/IP WANS with HyperIP are the business continuity and disaster recovery solutions of choice.