White Paper

WAN Optimization Technologies in EMC Symmetrix Replication Environments

Applied Technology

Abstract

This white paper outlines and defines the evolution, differences, and best practices for implementing WAN optimization in order to improve the functionality of data center to data center traffic flow critical to today's business continuity processes.

January 2009

Part Number h6066

Table of Contents

Executive summary

The ever-changing technologies that make up a “storage area network” introduce old concepts that sometimes take on new meaning and blend these concepts with new technology. Sometimes these features and functions can be misinterpreted and misused for a particular need in the data center. One example of this is the advances in wide area network (WAN) acceleration, optimization, and everything in between. Mistaking the difference between data optimization, WAN circuit optimization, and application acceleration, though the technologies are complementary, can blur what enhancements a customer may require to improve their environment. The main purposes of WAN optimizers can be any one or a combination of the following: to enhance data retrieval and writing activity across a WAN, giving the user and application the appearance of being on a LAN; to enhance the quality of the end-to-end network, including all LAN and WAN circuits between the appliances (i.e., packet loss, and latency); and to reduce the amount of bandwidth required for data transmission over the WAN through the use of data de-duplication and data compression.

Introduction

This white paper outlines and defines the evolution, differences, and best practices for implementing WAN optimization in order to improve data center to data center traffic flow critical to business continuity. It is not intended to compare the various EMC Select or EMC supported products available for WAN optimization.

This paper provides the framework for EMC customer service (CS), solutions architects, and technical consultants (TCs) to assess a customer’s data center for applying the appropriate WAN and replication services needed to optimize data flow.

This paper covers the essentials for understanding WAN optimization and preparing to choose a product based on features in technology appliances available today. It provides a historical review of the WAN optimization technologies, high-level topology overviews, and guidance to clarify some commonly misunderstood variables. It reviews the most applied technologies today for improving WAN quality of service and performing data de-duplication, as well as where they fit in the storage replication environment, among a few.

Audience

This white paper is intended for CS and EMC technical consultants, and storage administrators who implement replication within organizations that have EMC® Symmetrix® DMX-3 and DMX-4.

Technology overview

EMC Select third-party WAN optimization products are qualified, tested, and proven in the Symmetrix DMX™ environment, among other EMC storage appliances for improving replication solutions over long distances. Critical applications in a WAN environment are dependent on the availability, stability, and consistency of the network provided. Replication over distance has its challenges given the varieties of protocols being used, whether it’s synchronous or asynchronous, amount of data traffic, and type and length of WAN circuit. In synchronous applications, replication traffic will always be stable and limited by distance yet if needed throughput can be enhanced by using features such as write acceleration/fast write and compression. But sometimes these solutions are not enough for the demands of today’s enterprise environments. Asynchronous replication applications such as EMC SRDF®/A as well as multi-site disaster recovery applications such as SRDF/Star provide unlimited distance availability in WAN environments. This, however, does not delineate the fact that asynchronous replication over distance has additional challenges. As the need for replication over distance becomes more prevalent, bandwidth utilization becomes more critical for performance stability.

Varieties of asynchronous protocols being used, amount of data traffic, network medium, and connectivity are some of the factors in which latency and performance can be either introduced or impacted, as shown in Figure 1.

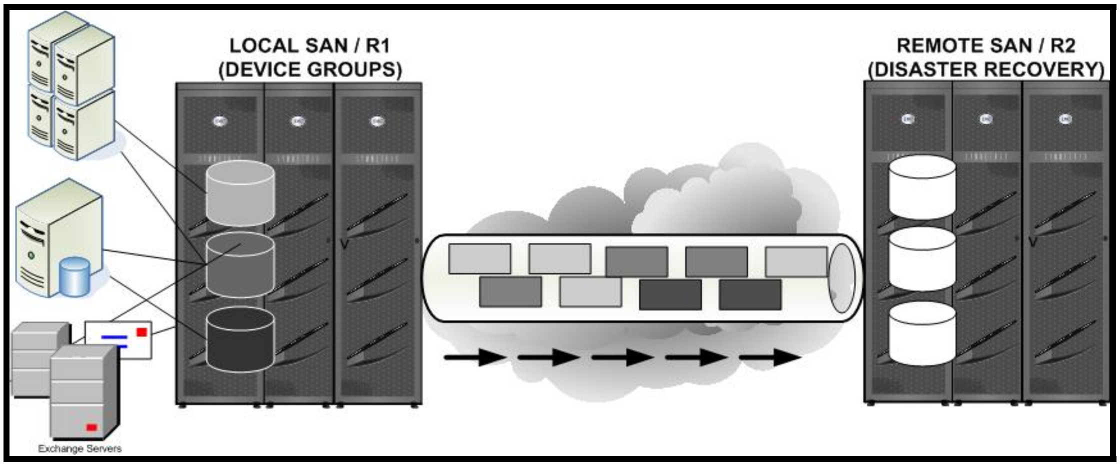

Figure 1. Typical SRDF/A topology

Reducing the recovery point objective (RPO) and recovery time objectives (RTO) are two primary reasons optimizing distance replication over a WAN is essential. RPO is defined as the maximum acceptable level of data loss following an unplanned disruption to business continuance that causes data loss and at what point an administrator can revert to prior to the disruption. RTO is the maximum tolerable length of time it takes to recover from a failure to a given RPO, and represents the amount of revenue lost per unit time measurable in seconds, minutes, etc. As the requirements for RPO and RTO increase, the more strain on bandwidth becomes factored into the possible instability of the environment. As customers lean on these objectives, they become more dependent on the need to increase their WAN capacity; however, increasing bandwidth is not instantaneous and is extremely costly. As there becomes a need to share bandwidth with other applications, stability of critical applications is affected as latency increases and packets are dropped. The existing WAN elements become oversubscribed, thereby adding to the need to share existing bandwidth with other applications and reducing the stability of the most critical ones, as shown in Figure 2.

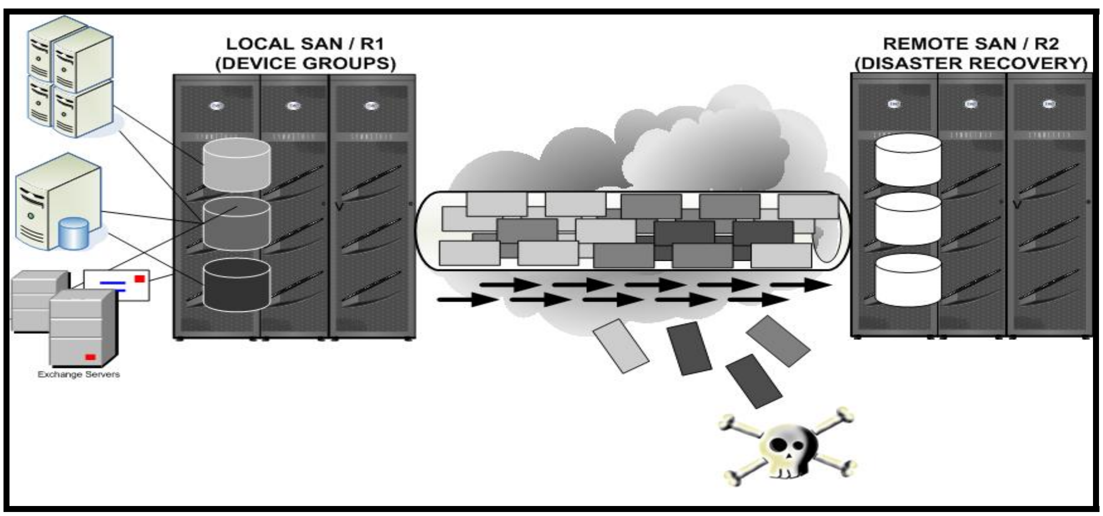

Figure 2. SRDF/A topology with latency, packet drops, and oversubscription

Beyond the need to solve network and growth constraints, moving data faster and more efficiently across the environment is essential to the customer's needs. These layers of complexity are adding up and technology vendors over time have developed ways to adjust to the demands. With WAN optimizers and accelerators, customers can experience many times faster transfers and can depend on disaster recovery applications such as SRDF/A and SRDF/Star to work at the "optimal" level of availability. And this often results in a reduction in the amount of bandwidth required. The total cost of ownership (TCO) of adding these non-intrusive hardware elements to the WAN is essential to the expansion and extensibility of the EMC replication environment. It is also the difference between spending hundreds of thousands of dollars a year on higher band links and a one-time purchase of the equipment. Optimizers are hardware devices that can provide the engine in which critical TCP applications can be handled in these environments. Overall, the TCO for the customer is that they can increase performance, decrease recovery time, and use existing WAN infrastructure for replication while experiencing almost LAN-like speeds.

In the global IP marketplace, the primary focus of optimization has been between the remote office and data center, and most discussion has been revolving around this environment. The demand today for cost savings has been limiting what amount of replication can be done between data center to data center, so focus on how WAN products can increase this potential could outweigh the needs for remote office communication. Historically, WAN optimizers have been developed for the IP market for over 10 years and over the past few years some major changes in both capabilities and performance have occurred.

Historical overview

The first generation of WAN "optimizers" helped to reduce the effects that packet loss, high latency, and jitter had on data communications. Typically this would be done with software-based features that help to either eliminate the limits of TCP, or eliminate the effects of potentially negative events within the network.

The second and third generation of WAN "optimizers" came with the addition of the LZ compression engine and/or some other optimization features. This combination of technologies into optimizer devices provides the ability to both allow for better data flow and increase the logical capacity of the network by compressing the data flowing over the network. The major limit here was generally these devices were only capable of handling throughput levels of a T1 link (1.544 Mb/s). Over time as higher speed WAN links became more common, so to did the ability of the WAN optimizer. In lieu of improvements, however, these devices typically used the LZ compression engine implemented in software and not the hardware. This greatly limited the overall throughput of these devices to bandwidth levels generally around 45-90Mb/s when needing to compress the data, and optimize the WAN. These devices are now becoming less prevalent over time as manufacturers move forward to accommodate higher circuits and shelve them to end-of-life. These now are being replaced with either the fourth- or fifth-generation devices.

The fourth generation of WAN "optimizers" added what is known as application acceleration (a.k.a. application awareness) to the third-generation optimizer. These devices are commonly known today as remote office WAN optimizers. These devices are designed for the type and style of traffic flow that is common to the remote office, so they generally include CIFS application acceleration, data de-duplication, LZ compression, and WAN QOS improvements. What these devices rely upon were very *low* throughput requirements with *very repetitive* file data being transferred between the remote office and the data center.

The fifth generation of WAN optimizers was developed with the same concepts that the previous generations optimizers had, but with specific and significant design difference. These devices are designed for *higher* throughput needs and *less repetitive* data flows. These devices were built for the data center to data center optimization needs. As a result of these changes, however, they can sometimes also be used as fourth-generation optimizers performing to a great degree the same functions when needed as the remote office WAN optimizer. The focus of this white paper is to emphasize the culminations of features that these fifth-generation devices adopted and evolved into in order to provide the most efficient optimizer to date for replication.

WAN optimization replication topologies

The typical configurations used for IP WAN with SRDF can be done using Fibre Channel or multi-protocol channel director (MPCD) ports with Gb/E capability out of the Symmetrix to a corresponding port on a given device. In cases where Fibre Channel is used, typically these configurations are connected to a multiprotocol switch with FCIP capability or distance extension hardware. Consistency on both sides of the SAN is very important to troubleshooting and stability among the variety of optical (small factor pluggable) SFPs[1] providing signal translation throughout the connections across the WAN. In most situations with multiprotocol switches and distance extension, EMC and the vendor corroborate on which are compatible and tested. For the purpose of this paper, it is worth mentioning these choices because they could be possible options for those customers who may already have existing components. Redundancy and high availability are typically done via duplicate hardware on both sides of the WAN depending on customer requirements and budget.

The best practices and explanations throughout this white paper are solely focused on native Gb/E from the Symmetrix, like as shown in Figure 3. Understanding that there will be other components on the LAN in between the Symmetrix and the WAN optimizer, possible nuances may cause interference on throughput and the way the component behaves. If the LAN drops links for example, features such as compression, flow control, and other QoS being used within the network may need to be reviewed by the customer network administrator[2].

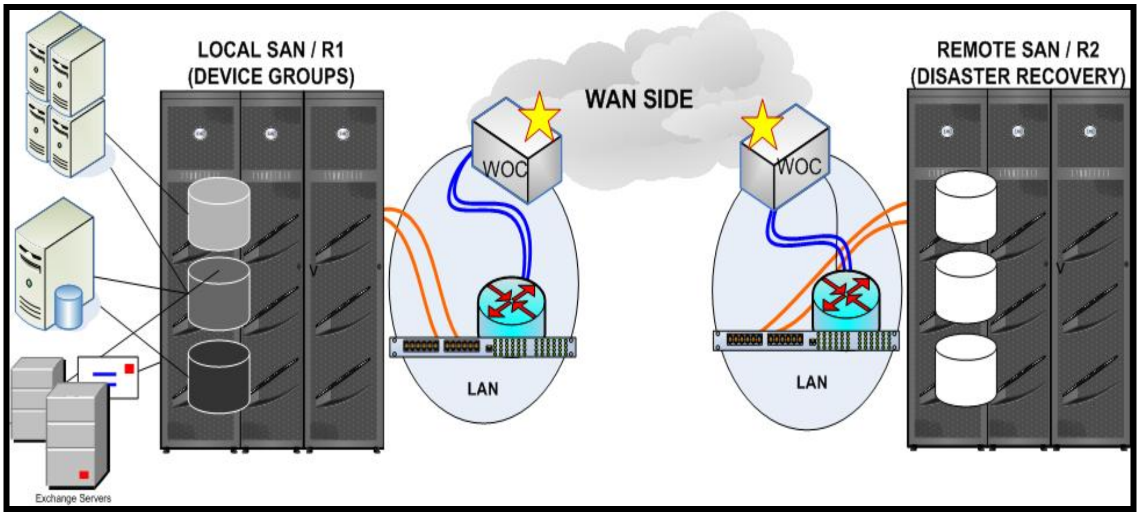

Figure 3. Typical WAN optimization topology

WAN optimization features and best practices

When WAN optimization is called for in a data center to data center environment, the customer is typically looking for a way to increase the amount of available bandwidth or decrease the amount of bandwidth that applications are using by manipulating traffic. This needs to be considered every time an application is added to their WAN connection, and how traffic affects replication needs in the provided IP network also needs to be considered. At first the concept of what is a throughput level seems to be a simple concept; however, when introducing WAN optimization into the IP SAN, there are several numbers that need to be understood as well as in what situation(s) these numbers can be determined. In general, it is incumbent on the EMC CS /TS engineers, and customer network and storage engineers to determine what limitations exist before choosing which vendor to invest in. Not all vendors use the same hardware and software features to provide optimization, but they are designing around the same technologies so their integration will determine what best fits the customer. Also, EMC Select vendors have a variety of offerings that aids in determining how much technology is needed and at what cost. Discussing the features that make up the fifth-generation WAN acceleration devices today may seem as review for many administrators but they are worth mentioning in the concept of how all the features working together provide the most complete

[1] Copper transceivers are used typically in a customer LAN; EMC does not OEM them through the Connectrix® or Symmetrix product lines without an RPQ.
[2] Symmetrix settings may be subjective to each environment and should be reviewed by the Solutions Validation Center – Business Continuance group via Solutions Qualifier (http://gig.corp.emc.com).

optimization solution. These features are *quality of service (QoS) traffic shaping, compression, de-duplication, forward error correction (FEC),* and *packet order correction (POC).* Other features such as security (disk, network), monitoring, and logging are to be considered standard, available to the discretion and design of the vendor, and beyond the scope of this paper.

Quality of service (traffic shaping)

Quality of service (QoS) is the capability to improve traffic flow using the IP-based applications designed to provide traffic shaping and queuing protocols. Two basic functions of QoS are the use of a routing protocol technology that allows for the prioritization of different IP traffic flows, and providing actual quality, or cleanliness, of the network from a perspective of latency, jitter, throughput, and packet loss.

Traffic shaping can be done based on priority via queuing and prioritizing the order of packets in which data can move. By calculating the round trip time (RTT) at its heaviest and then by tuning the traffic appropriately, QoS should be configurable as it is designed by the vendor either via GUI or CLI. Traffic shaping can be based on classifying the data by port, protocol, sensitivity, and also bandwidth allocation, as shown in Figure 4.

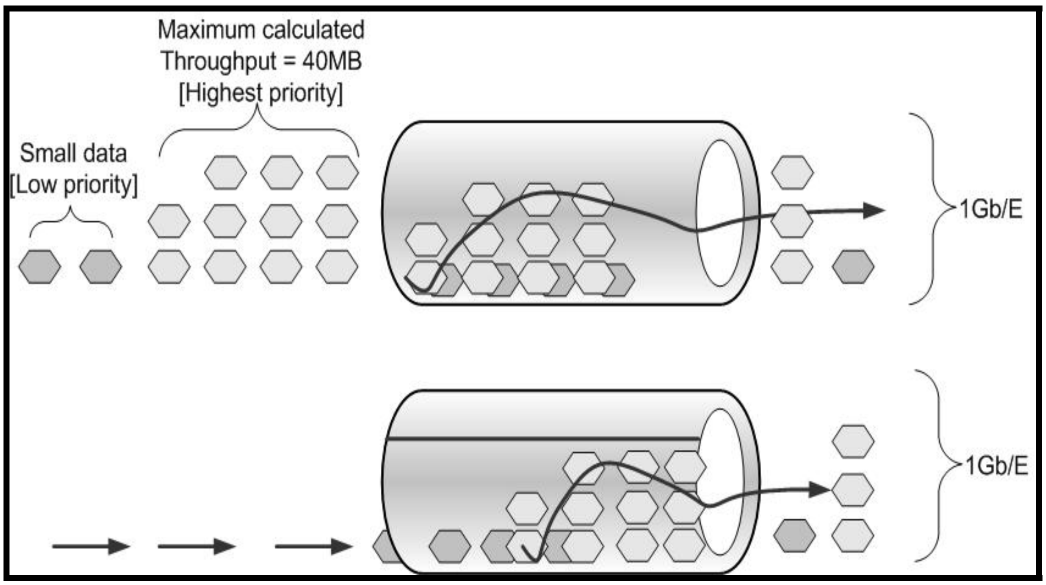

Figure 4. Representation of traffic shaping by priority class

The caveat to using QoS to optimize traffic is that it doesn't delineate the factor of traffic being added to the pipe over time. Building a dependency on only QoS features on the WAN can incidentally cause an administrator to overlook the amount of traffic being added and more likely packets will be dropped due to congestion. The general downside to this with critical applications such as replication is assuring priority is set highest without accidentally setting it at a lower priority level. Since QoS can be overdone, using the appropriate vendor guidelines on setting up QoS and understanding bandwidth needs for the replication technology is a critical first step in building optimization within the device.

Data compression

Compression is common technology that is used to define the reduction of the overall size of an IP packet as it is transported across a network. For a long time compression has been the main resource for reducing data size traffic on a link in order to minimize bandwidth consumed during send and receive between two endpoints. The most common compression algorithm used is Lempel-Ziv (LZ).[3] On average the use of a LZ compression on hard to compress, but compressible, data will yield a reduction ratio of 2.4:1.

[3] There are many different compression algorithms derivative of work done by Lempel and Ziv.

Advanced or session-based compression is widely used in WAN acceleration today. Previous compression algorithms were packet-based transmissions, creating a variety of data packets bundled into single compressed groups. These groups also had to be decompressed and re-assembled on the network. Two-fold problems exist:

- Reduction of ratio throughput possibly as a result of too many data variables, protocols, etc.
- Overhead of re-assembly

Session-based compression takes all of the same data packets and bundles them together in the same streams as to avoid the pitfalls of packet-based compression, as shown in Figure 5.

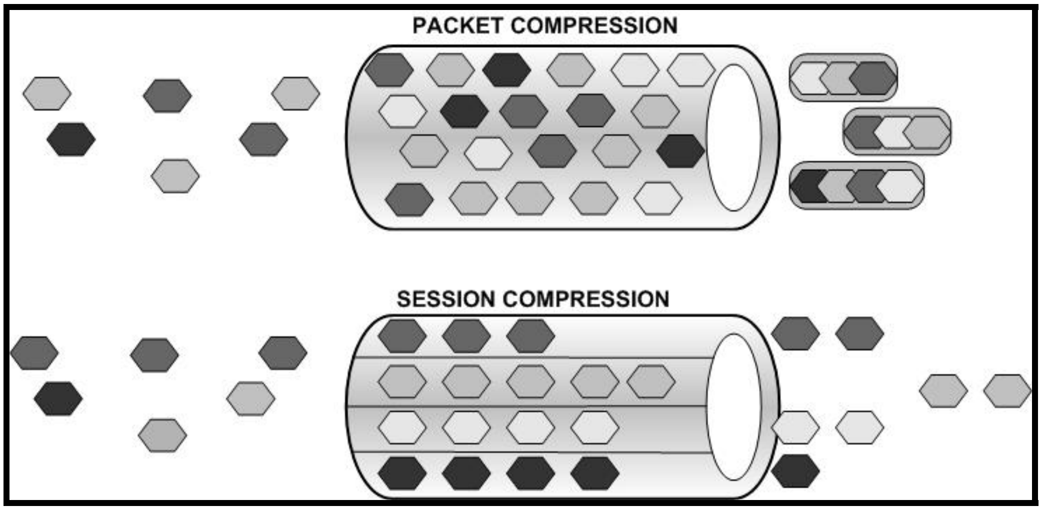

Figure 5. Differences between legacy and advanced compression technology

Compression can be done from the application itself, i.e., configured logically from the Symmetrix host or through the hardware software control components on the link. A ratio of 2:1 is typically seen with native Gb/E out of the Symmetrix without the use of WAN optimization; however, tests have shown that this ratio increases much higher when a hardware device with advanced compression options is introduced to the path of traffic and software compression on the Symmetrix is turned OFF[4]. This is the recommended setting when using WAN optimization devices.

De-duplication

Data de-duplication is a technology that goes beyond the capabilities of the data compression, that is, LZ, engine. It utilizes what are commonly referred to as pattern tables, with corresponding pointers. Duplicate pattern tables are built on both ends of the network, and once a pattern is seen for a second time, instead of sending the data across, the point to the data pattern is sent, allowing for much less actual data to flow over the network. Some vendors may design the pattern tables for data de-duplication via cache-based or disk-based memory depending on the hardware specification. A cache-based system is faster but has smaller tables and will be overwritten continuously, whereas disk-based data processing and storage could add latency because it has larger tables to work with. Yet, over time the patterns remaining become recognizable faster and the overhead of re-identifying the same patterns is reduced. A combination of both cache and disk-based pattern tables can bring even further enhancement. Again, depending on the needs in the environment de-duplication is nonetheless a valuable feature that can accelerate data by eliminating redundancy. Figure 6 provides a graphical depiction of data de-duplication.

[4] Tests vary among vendors and WAN variables; results are not available for public consumption. An EMC customer engineer is needed to turn compression OFF on a Symmetrix.

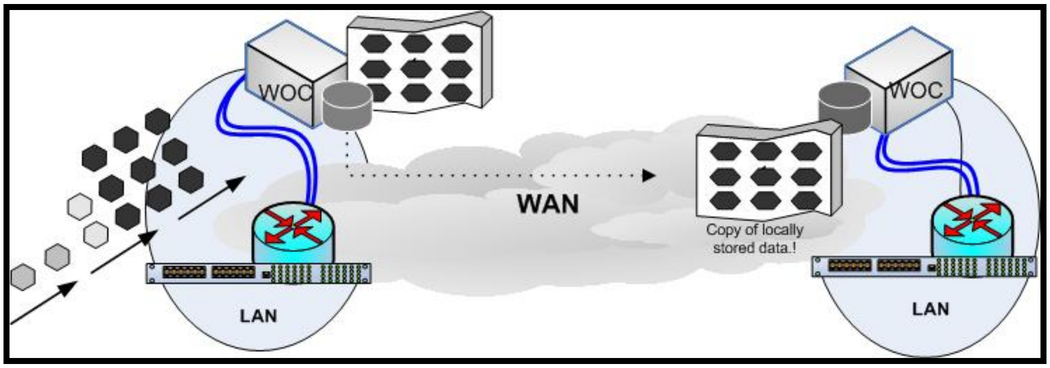

Figure 6. Theoretical representation of de-duplication

De-duplication is being integrated in several ways from host applications to hardware devices. The effectiveness of de-duplication resides in its ability to identify unique "chunks of data" that subsequently can be detected ongoing so that they can be "de-duped" and moved as needed from storage rather than being located redundantly. The variety of applications using de-duplication today by EMC are backup and storage retrieval, archives, and replication. In the latter cases, de-duplication serves the purpose of creating efficiency wherein replication uses the technology to save bandwidth by moving less data.

Forward error and packet order correction (FEC and POC)

Forward error correction[5] (FEC) allows a receiving target to detect and correct errors and transmission without having the initiator resend data. Without FEC, devices have to retransmit data in creating the volume of traffic going across the connection. FEC historically was designed for improving data transmission across a physical link, typically for satellite, SONET circuits, and other analog traffic. Commonly used at layer-1, it provides error control from the sender sending additional data to message, which the receiver can use to detect and correct subsequent errors delimiting the need for retransmission. WAN vendors use FEC at layer-2 or above within an IP network. This technology also allows correction of (rebuilds) packets that are lost or damaged during transmission. Much like the layer-1 (physical) FEC, it adds an error-recovery packet so that the receiving side can reconstruct the packet based on information contained. For example, if a customer knows packet loss is occurring at around 1 percent, they can turn FEC on and assign it to add the recovery packet for every 10 packets transmitted. This ratio of 1:10 would essentially now reduce packet loss to .09 percent. Adding a packet for every five (1:5) transmitted packets would reduce it even more. This adds another level of optimization where retransmission (or the repetition) of information is eliminated. Because of FEC's computational nature, there is some intrinsic overhead; however, it is far from a deterrent, and it is very low comparable to the benefits it provides.

In conjunction with FEC, packet order correction (POC) technology provides the capability to keep track of the order of sent packets and rearrange the out-of-order packets on the target WAN optimizer before sending them forward. When packets are lost during replication, retransmission increases, putting a strain on performance and throughput. By tagging the packets and reordering them performance can be maintained. Even though network routers intrinsically handle out-of-order packets on the IP level, the best practice for situations where replication is affected would be to provide this feature on the ingress/egress points on the WAN by using WAN optimizers. Done in real time, these types of features are really important to Symmetrix replication and provide the capability for more data to be moved onto the LAN at immediate speeds locally before and after the data was optimized during the WAN transmission.

Conclusion

It is obvious that all these WAN optimization features coming together as one holistic mesh can produce the fastest performing, and most accelerating and stable replication environment for Symmetrix applications. Overall analysis shows what environment each feature applies to and where it benefits data replication. Product vendors today are continuously testing the features that create pure optimization for

[5] Not all vendors support FEC or POC. Please refer to their current devices for features and functionality.

applications such as Symmetrix SRDF. Storage administrators, technical consultants, and IT architects alike are all vested in procuring this information from the vendors with assistance from EMC. Though performance numbers are not published from EMC, solutions can be tailored based on the customer's needs. It is, however, indicative of technical consultants and customers to be diligent in educating themselves as well as in producing environmental variables that factor into choosing the right solution!

References

For more information on WAN optimization topics please research EMC's websites (EMC.com and the customer-only Powerlink site) and the following vendor sites:

- Cisco Systems: http://www.cisco.com
- F5: http://www.f5.com
- Netex: http://www.netex.com
- Silver Peak: http://www.silver-peak.com