SAN VALLEY | THE INTERNET SAN COMPANY™

White Paper

Draft

Technical Considerations For Storage Networking Across MAN/WAN

Performance and Reliability Requirements for IP-SAN Extensions

Version 0.2

Technical Considerations For Storage Networking Across MAN/WAN

Executive Summary

As Fibre Channel Storage Area Networks (SANs) become ubiquitous, requirements increase to extend the power of SANs beyond the data center. This creates demand for SAN interconnection to address requirements for business continuance, disaster recovery and remote mirroring applications. In order to deliver the performance and reliability of a SAN across Metropolitan and Wide Area Networks, the end-to-end solution must be able to transport storage traffic across the network without introducing impediments. Understanding these requirements, SAN Valley Systems has delivered breakthrough IP-SAN technology that is responsible for connecting storage across Optical and IP infrastructures. This IP-SAN technology enables the ability to maintain wire speed throughput of storage traffic without impacting the integrity of this data across the entire network. This paper will examine the specific aspects of addressing these customer requirements while leveraging existing infrastructure, transport and data recovery mechanisms to achieve maximum performance.

Introduction

Fibre Channel SANs are currently deployed at a rate that is expected to have a Compounded Annual Growth Rate (CAGR) of over 69% between 2000 and 2004[1]. These Fibre Channel SANs are being deployed as storage "islands," most of which lack connectivity between the data centers where these SANs are currently being deployed. This growth is creating an opportunity to interconnect these SAN islands within a metro or wide area, enabling compelling service opportunities for disaster recovery, backup and resource consolidation. There are however, limited transport options for extending SANs beyond the data center, namely, dark fiber, wave division multiplexing, ATM and IP.

IP continues to be the standard method of data communication within Enterprise and Service Provider market segments; the ability to extend storage services across these networks offers considerable benefits. IT organizations have invested billions to support and grow their IP infrastructures, utilizing this investment to carry storage traffic would make sense from a capital preservation perspective. It must be considered that when transporting storage traffic over distance; enterprise customers would typically not utilize public network infrastructures (big "I" Internet) for transporting mission critical data, but would alternatively utilize private, leased or managed network facilities, which offer superior performance, reliability or scalability.

To satisfy the requirements of transporting storage traffic across networks that could potentially span hundreds of kilometers, it is important to understand the characteristics of the protocols that are transported across each network boundary. Within the data center, Fibre Channel presumes there is a virtually loss-free network to maximize performance, relying on the upper layer protocols (SCSI) to guarantee that data is reliably delivered to the other side. Across the MAN or WAN, the network conditions may not be as predictable. For example, congestion, latency and packet loss become factors that drive decisions about transport technology choices. In order to select the proper transport protocol that IP will run across, there are tradeoffs network designers are faced with. A connection-oriented protocol, such as Transmission Control Protocol (TCP), provides congestion control and data recovery, but at significantly reduced performance, which can seriously impede storage applications. With connection-less transport protocols, such as User Datagram Protocol (UDP), the storage transport can deliver superior performance that matches the requirements of the Fibre Channel network, which leverages the reliability of the upper layer protocols to guarantee data delivery. The task of congestion control and management is now left to the IP-SAN gateway, which is the appropriate place for these functions to reside.

[1] IDC, 2001

Interconnection of FC SANs across IP Networks

There is a very strong demand for transporting storage traffic across IP networks. This is namely proven in major metropolitan areas where issues such as business continuance and disaster recovery are competitive imperatives. In order to address this demand, SAN Valley Systems has delivered the SL1000 IP-SAN Gateway. The SL1000 offers seamless extension of SANs across MAN and WAN infrastructures. As enterprise IT managers and service providers design networks for transporting storage traffic, only networks that offer a high degree of security, reliability and performance will meet the storage networking requirements. In these types of networks, strict controls are placed upon the network in such a manner that congestion, packet loss and latency are limited, predictable and supported by Service Level Agreements (SLAs). The specific network architectures used for SAN connectivity are explored in the following sections.

Dark Fiber and DWDM Networks

In order to provide the performance and reliability necessary for business continuance applications such as synchronous replication, Fibre Channel traffic can be multiplexed onto a fiber or to a specific optical wavelength for transport. Using Dark Fiber or DWDM to transport storage traffic offers the benefits of high performance, reliability and security, but is very costly, unmanaged and has distance limitations as shown in Figure 1.

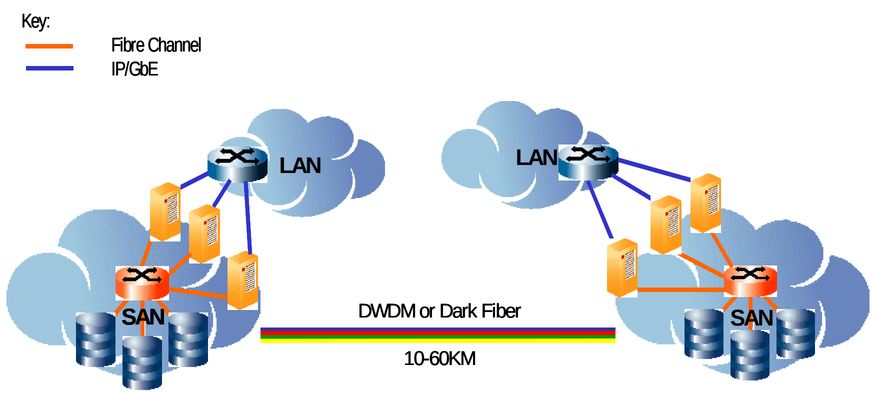

Figure 1: Fibre Channel Storage Traffic Extended Over Fiber

By using these dedicated (or leased) fibre resources, the network operator can effectively transport storage traffic from 10KM to 60KM (when using extended Fibre Channel buffer credits) without concern about packet loss and congestion. However, performance will significantly degrade when attempting to extend SANs across longer distances. If errors occur across the network links, the higher-level protocols (in this case SCSI) perform the error detection and recovery.

By adding the SL1000 to this network topology, the architecture delivers superior performance, reliability and security, but the connections are now fully managed. Another key benefit is the ability to support distances beyond 400KM without any performance degradation as shown in figure 2 .

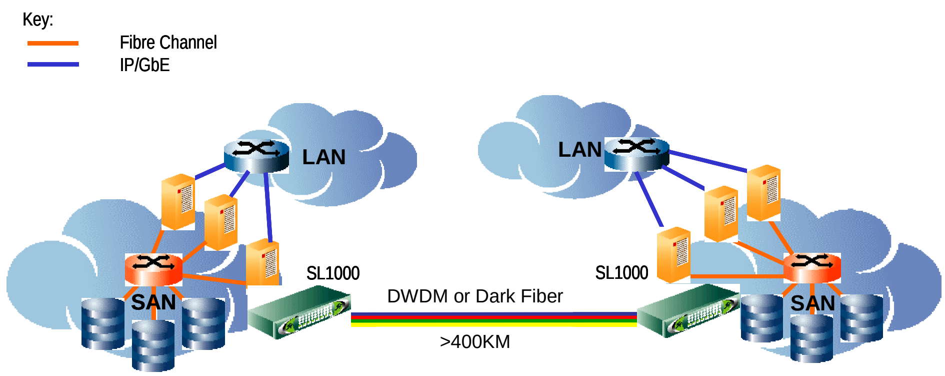

Figure 2: Fibre Networks with IP-SAN Gateway for Transport of SANs

Internet Protocol Networks

In order to support business continuance and disaster recovery applications across IP networks, high reliability and wire speed performance remain as fundamental requirements when analyzing network transport characteristics. While it is typically safe to assume that packet loss, congestion and latency are limited in dedicated fiber networks; the same assertions cannot necessarily be made for many IP network infrastructures. Recall that IP networks are available in various speeds, feeds and qualities. The public Internet, for example, is an ever-changing, dynamic, best-effort infrastructure that offers no guarantees or quality of service. Managed IP network connections, on the other hand, utilize a fully managed network architecture that combines the scalability of a packet switched infrastructure with the reliability of a circuit switched network.

When transporting storage applications across MAN and WANs, these services will be utilizing IP networks that offer high degrees of scalability, reliability and serviceability, typically backed by SLAs. A typical IP-SAN extension is shown in figure 3.

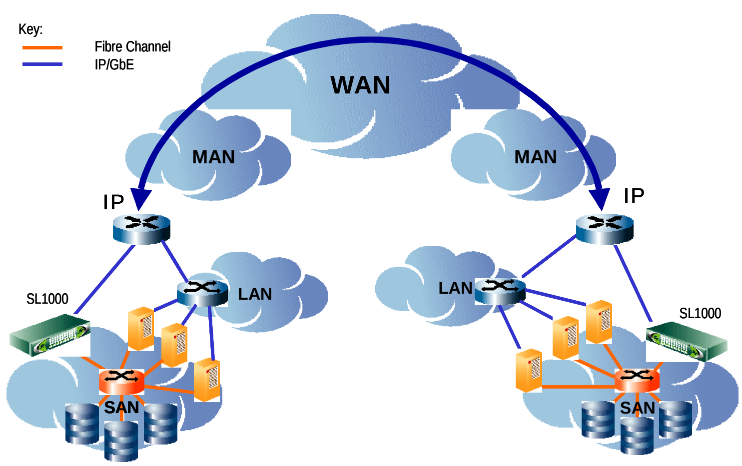

Figure 3: Interconnection FC SANs Across IP Networks

Congestion Control

Even with SLAs, a good network designer never assumes that the network is free from congestion. It is imperative that the network is designed to tolerate periods of over-subscription, which leads to congestion, added latency and even packet loss. Under these conditions, the network must be able to maintain optimal performance, but must also provide a flow control mechanism that reduces the sensitivities to congestion and latency. With the SL1000, the flow control mechanisms that exist are leveraged from the IP network, specifically mapping the pause frames used in Gigabit Ethernet to the FC Layer 2 (FC2) buffer credit mechanisms as shown in Figure 4. By segmenting the flow control mechanism into three distinct domains, the SL1000 achieves the following objectives:

1. Maps the GbE Pause Frame flow control onto the Fibre Channel network, by limiting the number of buffer credits that the Fiber Channel switch is allocating.
2. Reduces the performance sensitivity to latency caused by increasing the distance between the two SANs, enabling Fibre Channel traffic to be transported across longer distances without performance degradation.

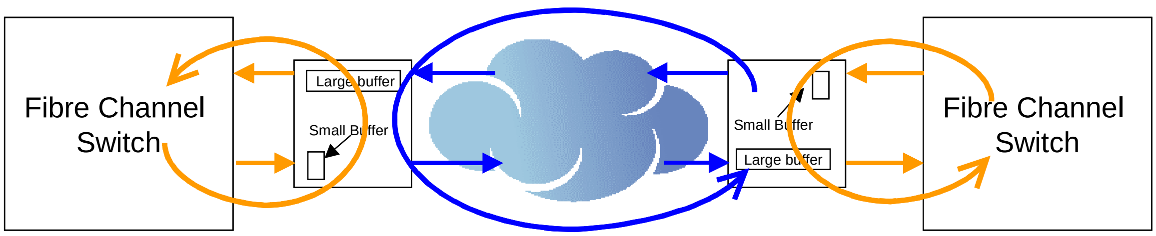

Figure 4: Flow Control - FC Buffer Credits Mapped to Gbe Pause Frames

Rate Shaping

By providing the ability to rate shape the IP-SAN traffic across a MAN/WAN, network administrators can shape the encapsulated storage traffic to the performance of the network. The ability to rate shape the storage traffic is accomplished by sending just enough buffer credits to the SAN to match the exact performance requirements of the IP network. For example, if the SLA on an IP network has been designed such that the network will maintain no fewer than 155Mbps, then the rate shaping algorithm will apply back pressure to the (1Gbps) Fibre Channel network (via buffer credit allocation) to only send 155Mbps of storage traffic on the network as shown in Figure 5.

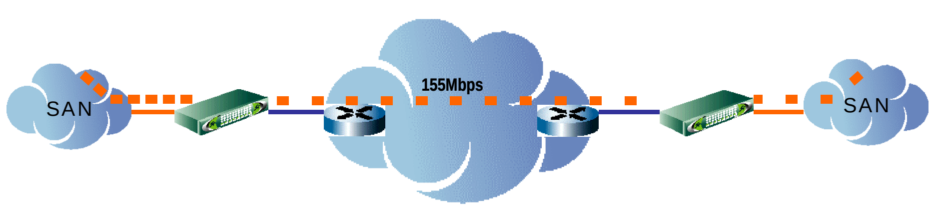

Figure 5: Rate Shaping Applied by Limiting Buffer Credits

Data Recovery

As mentioned above, even in well-designed, well-managed network, allowances must be made for congestion and packet loss. In the case of Fibre Channel, recovery mechanisms to detect and recover lost frames are implemented to prevent losing data, which can cause storage applications to fail. This recovery process ensures that the business continuance applications continue to proceed even when the transport layer is subject to congestion or data loss.

When a stream of serial FC frames that are being sent across an IP network, if a frame is dropped, the Fibre Channel layer 4 (FC4) mechanisms detect that a frame has been delivered out of order. The FC4 flags the next frame as "out of order" which is rejected and a retransmission is requested as the EDTOV timeout expires.

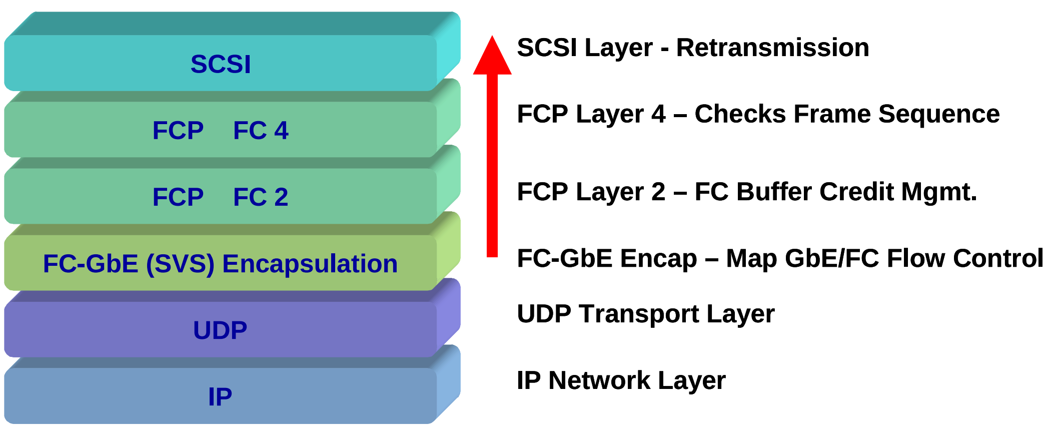

Figure 6: Flow Control and Data Recovery Mechanisms

Technical Merits

Considering the transport mechanisms discussed above, performance of the SL1000 is optimized for lightly congested environments. Once congestion becomes an issue, the encapsulation layer handles the rate shaping, flow control and traffic management, Fibre Channel performs the acknowledgement and retransmission request, while SCSI handles the retransmissions. This is all performed without affecting the performance of the storage application and without windowing. In TCP environments, windowing can severely impact performance, causing business continuance applications such as synchronous mirroring to fail. This TCP flow control mechanism was designed for heavily congested networks, which we have already asserted would not be used for business continuance and disaster recovery applications. When compared to UDP in this type of environment, UDP will continue to stream data at wire-speed, when a frame gets dropped, SCSI will resend, and the throughput will not decrease.

Conclusion

When considering the requirements of extending storage across IP networks, it is critical to consider the network architectures as well as the performance requirements of the applications that are being used on the network. When operating on optical networks that have predictable performance characteristics and reliability, little consideration is required for congestion, packet loss and data recovery. But when operating in packet switched environments, where over-subscription is possible, even with SLAs, an IP-SAN solution must deliver additional value. SAN Valley's SL1000 offers robust traffic management, flow control, traffic shaping and leverages the data recovery mechanisms prevalent in storage networks. SAN Valley has developed a product that offers the ability to deliver wire-speed storage services over longer distances with zero impact on the reliability of the delivered data.