Storage Fabric Unification with Ethernet

Realizing the Benefits of Network Convergence with Terminator 5

Executive Summary

This paper illustrates how the various fabrics used in a typical storage network can be unified into a single infrastructure, using the ubiquitous and familiar Ethernet technology. The paper follows a case study where 4 different networks using 3 widely different technologies could be replaced by a single infrastructure, using a single network adapter, with clear CAPEX and OPEX savings. The benefits of infrastructure consolidation, particularly over Ethernet, are manifold: lower acquisition costs of both switches and adapters, improved volume pricing, higher network efficiency and lower power consumption, reduced management and maintenance costs, and improved simplicity.

Introduction

Data storage demand has been on an explosive growth curve for many years, only accelerating as increasingly large datasets are generated (for Physics, DNA mapping, geo-computation/seismic applications, etc) and faster connected mobile users generate rich content at massive scale, worldwide. A significant proportion of this flood of data is still stored in conventional Storage Area Network (SAN) installations. Enterprise data generation has similarly been on an exponential growth path over the past decade, and is expected to continue for the foreseeable future. Most enterprises have moved to SAN based storage installations for the reliability and manageability benefits they provide.

A typical SAN consists of: a frontend network, where clients connect to the storage through file (e.g., NFS, CIFS) or block (e.g. iSCSI, FCoE) based protocols; a backend fabric, where storage servers connect to the storage media (HDD, SSD or tape); and a replication link or fabric for remote mirroring and disaster recovery, where applicable. An additional management network is needed to maintain and configure the devices and fabrics needed.

Copyright 2014. Chelsio Communications Inc. All Rights Reserved.

A Fabric Too Many

The following figure shows the components of a traditional storage system, with typical interconnect technologies in use.

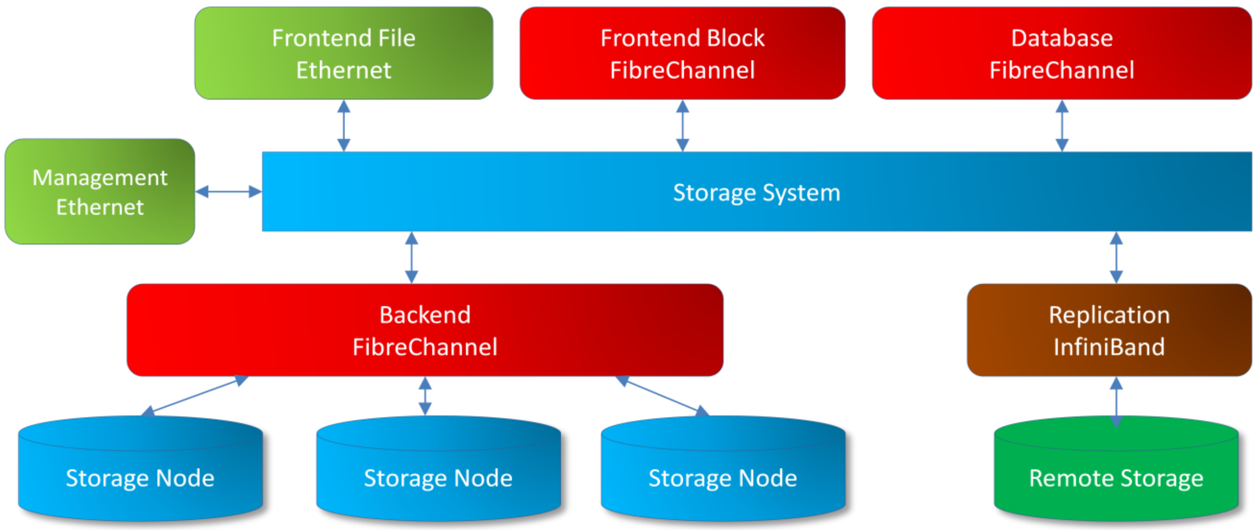

Figure 1 – Traditional Storage System

The backend has typically used Fibre Channel (FC) as a fabric, a legacy dating to the initial head start with high speed fiber that FC enjoyed over Ethernet, along with some storage-oriented features, such as protocol acceleration and data reliability. Recently, InfiniBand has seen use in this area, which can benefit from its Remote DMA (RDMA) capability.

Block storage has also typically used FC in the frontend, although iSCSI has become a popular alternative. Database servers are similarly connected to the storage system through Fibre Channel as well.

Efficient mirroring/replication requires high speed communication and can also benefit from RDMA, which allows transparent data transfer from local memory to the mirror system, with no software involvement and without consuming system CPU cycles. InfiniBand, thanks to high speed RDMA capabilities, has seen some deployment in this area.

The storage frontend is normally connected through an IP/Ethernet infrastructure to the users for file storage access, such as NFS, CIFS/SMB, or HTTP.

Finally, a separate Ethernet based management network is typically used for isolation.

Looking at Figure 1, it is evident that the number of widely different fabrics is unwieldy, and a direct source of complexity, installation and maintenance cost. There is an immediate appeal to replacing

Copyright 2014. Chelsio Communications Inc. All Rights Reserved.

these disparate fabrics with a single, unifying technology. Ideally, this would be Ethernet for its ubiquity, simplicity and cost effectiveness, not to forget that it probably is irreplaceable as a frontend network.

The main concern in using a single technology is preserving the benefits and features of each of the specialized and esoteric fabrics being replaced. The following section shows how this is achieved by leveraging the capabilities of Chelsio's T5 Unified Wire Engine to not only unify the fabric technology, but also utilize a single network adapter in the process.

A Fabric to Unify Them All

Chelsio's T5 engine is the latest (fifth generation) of a line of network processors designed with the express goal of unifying all applications and traffic types over an Ethernet interface. T5 is a highly integrated, hyper-virtualized 4x10GbE/2x40GbE controller with full support of NIC, TOE, RDMA, iSCSI and FCoE applications, and enhanced data integrity protection. T5 provides no-compromise performance with both low latency and high bandwidth, and scales to true 40 Gigabit line-rate operation. T5 thus enables fabric consolidation by simultaneously supporting TCP/IP and UDP/IP sockets, iSCSI, FCoE and RDMA for storage applications at wire speed.

Furthermore, by implementing separate hardware channels, and a hardware based traffic management infrastructure, it provides the appropriate quality of service (QoS) demanded by the various applications even as they simultaneously share the hardware. In effect, T5 enables InfiniBand and Fibre Channel applications to run unmodified and concurrently over standard Ethernet.

The following figure shows how the same system can be made to utilize Ethernet exclusively, leveraging the multi-faceted feature set of T5 to replace FC and IB, while preserving or exceeding their performance and capabilities.

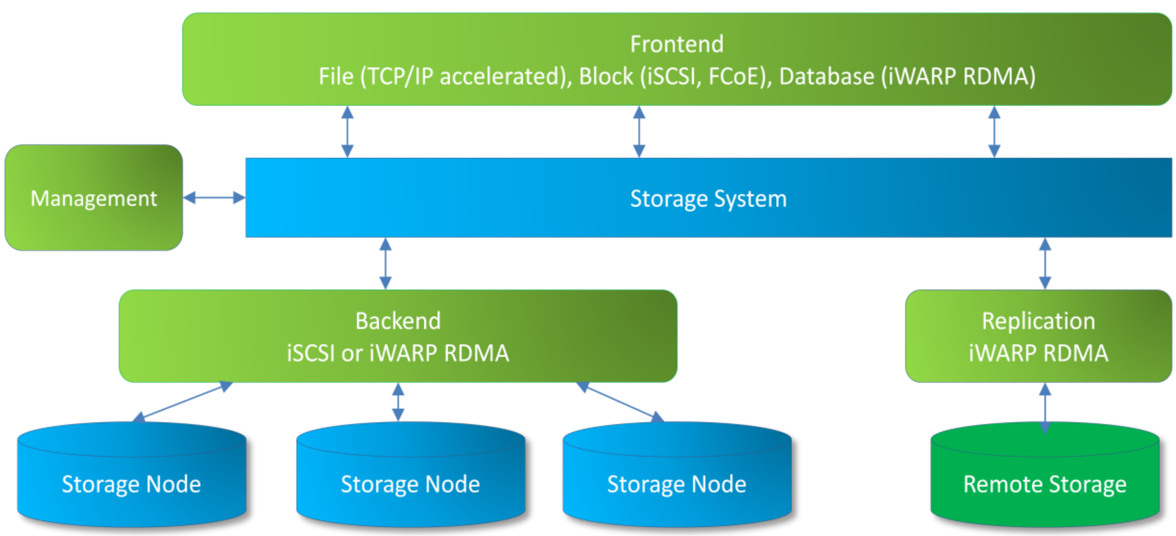

Figure 2 – Unified Fabric with Ethernet

Copyright 2014. Chelsio Communications Inc. All Rights Reserved.

Thanks to the native RDMA implementation of the IETF standard Internet Wide Area RDMA Protocol (iWARP), T5 allows 40Gbps wire speed data transfer, efficient replication even across up to tens of thousands of miles distances (halfway along the globe). Because it shares the same (OpenFabrics) middleware with InfiniBand, it is a drop in replacement wherever IB is used, while scaling to significantly longer distances when utilized for replication.

With a high performance, fully offloaded implementation of iSCSI and FCoE, T5 can similarly drop in as a replacement for the FC fabric, preserving the low CPU utilization of native FC adapters.

The same protocols can be used on the frontend as well, in addition to offloaded NFS/SMB. Furthermore, T5 is the highest performing adapter for the latest SMB version (v3.0), using the RDMA based SMB Direct protocol.

Finally, thanks to the use of TCP/IP as a foundation for iSCSI and iWARP, standard cost effective Ethernet switches can be used across the board.

Case Study

This section takes a look at a Tier-1 OEM's clustered storage system as a model for providing all Ethernet connectivity. This system's impressive capacity, at 5M IOPS and 170GB/sec, places it far ahead of others in its class, with SPECsfs2008_nfs 3x those of the closest competitor. These performance figures allow it to meet the requirements for high performance computing, DNA mapping, video processing, satellite survey, Data Center storage, and other Big Data applications.

Based on a symmetric highly scalable architecture, it allows expansion through simple addition of nodes, with zero downtime. At the core of its connectivity is a high speed fabric, with an Ethernet-only option, where each node is connected through two Chelsio adapters to the frontend and internal networks. The system offers a wide variety of file-based interfaces notably NFS, SMB/CIFS, HDFS, FTP, HTTP and S3, as well as a block-based iSCSI interface for clients, and utilizes RDMA in the backend. Ethernet based convergence is made possible through using Chelsio's iWARP RDMA support to replace the costlier InfiniBand alternative, and leverages Chelsio's TCP/IP offload and iSCSI offload for improved performance and efficiency on the frontend.

Conclusions

This paper shows how a complex multi-fabric storage installation can be replaced by a single, standard Ethernet based infrastructure solution, using a single adapter, realizing tremendous CAPEX and OPEX benefits in the process. The resulting system enjoys all the benefits of Ethernet in simplicity, cost effectiveness and economies of scale while maintaining or exceeding the performance, features and quality-of-service of specialized fabrics.

Copyright 2014. Chelsio Communications Inc. All Rights Reserved.