



Unlock In-Server Flash with InfiniBand and Symantec Cluster File System

Introduction 1
 Performance Metrics 2
 Summary 3

Introduction

Solid State Drives (SSD) and flash technologies are putting Fibre Channel (FC) against the ropes. Fibre Channel’s lengthy evolution cannot cope with faster pipe demands, as 8Gb/s bandwidth performance is not good enough to satisfy all application requirements. While 10Gb/s and 40Gb/s Ethernet may look like an alternative, InfiniBand (IB) currently supports up to 56Gb/s, with a roadmap to 100 and 200Gb/s . Both Ethernet and IB have a considerable advantage over FC.

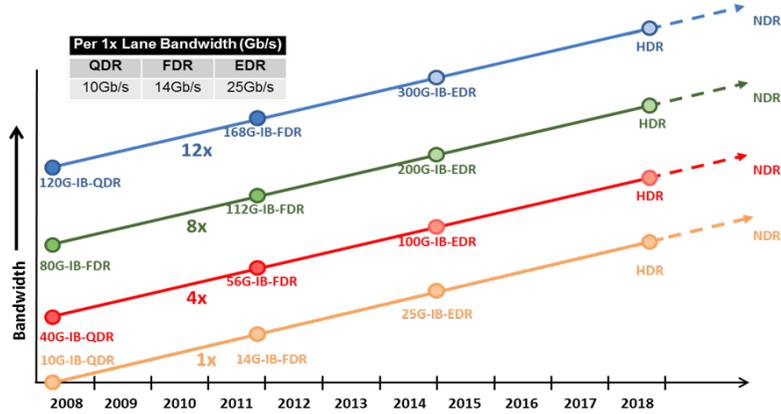


Figure 1: InfiniBand™ Roadmap

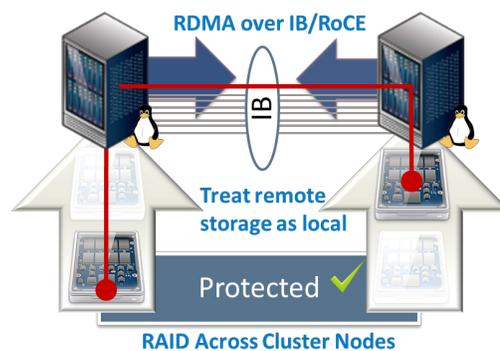
SAN-based storage has typically been used to increase performance and provide data availability for multiple applications and clustered systems. However, SAN systems are now trying to cope with new high-performance application demands by adding Solid State Drives (SSDs) into the storage arrays. Given SAN complexities, this is a costly solution with only partial benefits for applications. While SSDs can offer an incredibly high number of IOPS and bandwidth, FC will be the new bottleneck as the round trips add milliseconds of latency on the top. A millisecond is an eternity, with response times now measured in micro- and nano-seconds.

Solid State capacities and performance are increasing every year, while simultaneously decreasing in \$/GB. These higher capacities and lower costs are making SSDs the default storage medium for enterprise servers, bringing the data closer to the application. By combining terabytes of fast storage with increasing processing power, application performance rises dramatically through reduced latencies.

While in-server flash can provide better performance, it still lacks data availability and data management tools provided by enterprise storage arrays. Local capacity will not be accessible if the server is not available, impacting critical applications that rely on high performance and throughput to generate revenue. To override that limitation and still receive all the benefits of in-server flash, a clustered system that manages data protection across servers while providing high availability for applications is needed. To maximize the performance of in-server flash in a redundant configuration, a faster interconnect is needed to move data between systems.

Mellanox InfiniBand provides a high throughput and low latency interconnect to ship data across servers and storage systems. Although traditionally used in high-performance computing (HPC) environments, InfiniBand, with transfer rates up to 56Gb/s, provides the capability to unlock the potential of in-server flash.

Symantec Cluster File System High Availability 6.1 (SFCFSHA) for Linux introduces support for “shared nothing” configurations with its new Flexible Storage Sharing (FSS) feature, providing a single namespace across nodes using internal storage capabilities. Resiliency and data redundancy for internal storage is no longer a problem with FSS, as it provides mirroring capabilities across servers for resiliency, avoiding any single point of failure in a “shared nothing” architecture. In SFCFSHA 6.1, FSS supports up to eight nodes to share local, SAN, or hybrid storage while providing a single namespace for mission critical applications.



FSS supports Mellanox InfiniBand, using RDMA (Remote Direct Memory Access) to ship data across server nodes. With a possibility of eight InfiniBand interconnects, configurations can achieve a theoretical transfer rate of 54GB/s (432Gb/s) between nodes. That transfer rate would be equivalent to moving 12 DVDs per second across servers.

Using in-server flash in a “shared nothing” environment will benefit typical use-cases for Symantec Cluster File System, such as Oracle RAC, SybaseIQ, Clustered NFS, Business Intelligence, and custom financial applications that require a high number of concurrent reads and writes.

Performance Metrics

Low Latency Transport (LLT), a protocol used to ship I/O between servers, has been enhanced to utilize InfiniBand. Bigger packet sizes will provide better bandwidth, as seen in Figure 2, in which non-RDMA 40Gb/s Ethernet, 40Gb/s RoCE (RDMA over Converged Ethernet) and FDR 56Gb/s InfiniBand is used.

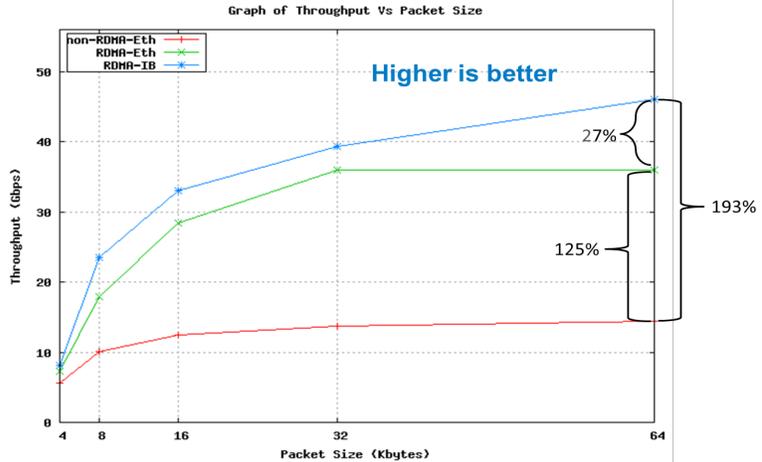


Figure 2: Different Interconnect Bandwidths

A single Solid State Drive providing up to 1.5GB/s can saturate a 10Gb/s Ethernet link. Figure 3 shows the performance difference when accessing data locally or from another node using InfiniBand or 10Gb/s Ethernet. Clearly the last one reaches its bottleneck point at 1.25GB/s while a single InfiniBand link can provide much better capabilities.

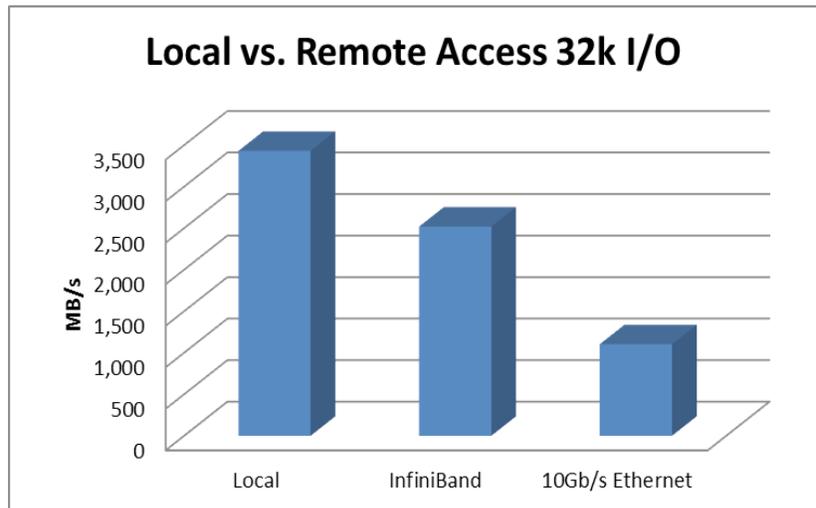


Figure 3: Performance Differences

Summary

The combination of Solid State Drives, InfiniBand interconnects and Symantec Cluster File System HA 6.1 provides a high performance and low cost alternative to expensive and complex SANs. Internal storage will be protected across servers, avoiding any single point of failure in a “shared nothing” architecture. The High Availability component will make sure the application is recovered in seconds. Un-lock in-server flash, improves application performance and reduces costs.



350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
 Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com

