

[Close](#)

The role of InfiniBand in the enterprise

By Jeff Boles

May 29, 2009 -- For many IT managers, choosing a computing fabric often falls short of receiving the center stage attention it deserves. Familiarity and time have set infrastructure architects on a seemingly intransigent course toward settling for fabrics that have become the *de facto* norm. On one hand is Ethernet, which is regarded as the solution for an almost ridiculously broad range of needs – from Web surfing to high-performance, every-microsecond-counts storage traffic. On the other hand is Fibre Channel, which provides more deterministic performance and bandwidth, but at the cost of bifurcating the enterprise network while doubling the cabling, administrative overhead, and total cost of ownership.

But even with separate fabrics, these networks sometimes fall short of today's requirements. So slowly but surely some enterprises are waking up to the potential of new fabrics.

There is no other component of the data center that is as important as the fabric, and the I/O challenges facing enterprises today have been setting the stage for next-generation fabrics for several years. One of the more mature solutions – InfiniBand – is sometimes at the forefront when users are driven to select a new fabric.

New customers are often surprised at what InfiniBand can deliver, and how easily it integrates into today's data center architectures. InfiniBand has been leading the charge toward consolidated, flexible, low latency, high bandwidth, lossless I/O connectivity. While other fabrics have just recently turned to addressing next generation requirements – with technologies such as Fibre Channel over Ethernet (FCoE) still seeking general acceptance and even standardization – InfiniBand's maturity and e bandwidth advantages continue to attract new customers.

Although InfiniBand remains a small industry compared to the Ethernet juggernaut, it continues to grow aggressively, and this year it is growing beyond projections. This article examines what is driving some enterprises to consider new fabrics today, and delves into a few use cases for InfiniBand in the data center.

I/O dilemmas

Enterprises face a barrage of new I/O demands from multi-core CPU servers and bandwidth-intensive applications such as sophisticated ERP systems; business intelligence systems that analyze years of data and might even encompass video analytics; virtual servers; content creation applications; web application workloads that handle ever richer content and larger numbers of customer transactions, and many more.

A single server can demand more bandwidth than most fabrics can handle. Case in point: the AMD Shanghai and Intel Nehalem processors, which were designed for intense I/O performance and have effectively doubled the performance of prior generation processors. In conjunction with hypervisor architectures built to harness every ounce of performance from underlying hardware, the virtual infrastructure can almost become a virtual denial of service for existing infrastructure.

The enterprise's traditional answer to more I/O – namely, more adapters and cabling – cannot be the solution in

enterprises now faced with power, cooling, space, and management constraints. Companies that addressed previous I/O challenges by adding more segregated connections to more separate networks, separate fabrics, or even local buses can no longer pursue these practices.

Unified fabrics

The answer is a single unified fabric that can reduce cabling, connect increasingly greater numbers of high-density servers within the power and airflow limits of the data center, and deliver the performance in high bandwidth and low latency to meet the full spectrum of traffic and storage demands in next generation data centers. While InfiniBand may not be the end-all fabric for every enterprise, it is increasingly finding adoption in pockets of enterprises for a variety of reasons.

InfiniBand as a next generation data center fabric doesn't require forklift upgrades in order to meet I/O requirements. InfiniBand can easily be applied in a rack, across virtual servers, within clusters, and in other domains. The I/O challenge is no longer behind a single server, but behind 40 servers (with 160 cores) in a rack, or 16 servers (with 64 cores) in a blade, each of which may in turn support 8 to 30 logical servers running on top of a hypervisor that might demand as many as 400,000 I/Os per second (IOPS).

Few technologies are as well equipped as InfiniBand to handle these I/O challenges. InfiniBand already has loss-less transmission and deterministic quality of service. In addition, InfiniBand is tightly coupled to processor and memory buses via RDMA-based protocols that can deliver high-performance networking.

Following are a few examples of where InfiniBand is finding adoption in data centers.

Scalable computing

InfiniBand often plays a role in enterprises with huge datasets. The majority of Fortune 1000 companies are involved in high-throughput processing behind a wide variety of systems, including business analytics, content creation, content transcoding, real-time financial applications, messaging systems, consolidated server infrastructures, and more. In these cases, InfiniBand has worked its way into the enterprise as a localized fabric that, via transparent interconnection to existing networks, is sometimes even hidden from the eyes of administrators.

Case in point: The demands on database environments continue to grow, and in response Hewlett-Packard, in conjunction with Oracle, created the InfiniBand-based HP Oracle Database Machine, dubbed Exadata.

Exadata is a clustered grid of eight Oracle RAC servers coupled to 14 Exadata storage servers running over a consolidated low-latency, high-bandwidth InfiniBand fabric with four 24-port Voltaire 9024 InfiniBand switches.

Voltaire's switching platform in the Exadata cluster harnesses InfiniBand's high availability to perform path optimization and/or loss-less restructuring of the entire fabric if any significant event or outage occurs.

Core InfiniBand technologies are integrated with the Exadata solution. Specifically, while the Exadata platform distributes some data intelligence to the storage nodes, each of those nodes depends on Remote Direct Memory Access (RDMA) and Reliable Datagram Sockets (RDS) to give the cluster the bandwidth and latency necessary to perform distributed data operations at rates upwards of 7GBps. Those protocols allow the Exadata platform to access each distributed processing node (Oracle RAC or Exadata storage) with minimal latency and host processing overhead.

Server virtualization

Server virtualization continues to be one of the top five IT priorities, and more than one-third of enterprise workloads go into virtual environments. Often, server virtualization proves to be a torture test for existing fabrics, and some enterprises are turning to localized InfiniBand fabrics in these scenarios, where there is the need for both high bandwidth and low latency.

While previous generation server architectures often constrained overall I/O performance, the latest generation of servers -- and the widespread adoption of PCIe 2.0 -- have addressed some of the I/O challenges. Today, virtual guests can readily demand bandwidth in excess of 500MBps, leading virtual hosts to exceed even 10Gbps Ethernet capabilities.

Virtual infrastructures also require increasingly fluid workload movement, as well as sophisticated failover and fault-tolerant features. Every innovation around these technologies moves the virtual infrastructure toward the vision of a data center architecture where workloads can be continuously protected and transparently moved across physical systems on demand. Yet every innovation has underlying impacts on the fabric, whether it is latency-dependent traffic generated by I/O synchronization across multiple systems, or bandwidth consumption from the high-speed movement of server images across a data center.

Finally, the consolidated virtual infrastructure is dense, and compounds problems with power, cabling, and cooling. Virtual server configurations often have as many as 8 Fibre Channel and/or Ethernet adapters in a single physical host. When compared to a single 40Gbps QDR InfiniBand fabric, traditional fabrics can more than double what it costs to operate and manage I/O in virtual server infrastructures.

For these reasons, some enterprises are deploying InfiniBand as a localized, high-bandwidth network behind virtual infrastructures, with vendors such as Voltaire, Xsigo and 3Leaf Systems driving adoption by delivering sophisticated levels of management on top of a unified fabric.

Enterprise clouds

Cloud infrastructures are virtualization at scale, and raise the bar for flexibility while demanding a connectivity layer that can help orchestrate and manage huge numbers of resources. Just as InfiniBand helps make the virtual infrastructure more fluid by setting it free from physicality, InfiniBand is one of the key technologies behind leading cloud services providers.

Cloud workloads demand freedom from physical resources, so that they can be dynamically re-provisioned or moved as business demands change. To deliver fluidity around workload management within a cloud infrastructure, cloud provider A-Server boots hosted virtual server images from a shared storage infrastructure. If failures occur, or if the infrastructure changes, images can be dynamically booted from other servers.

Large-scale cloud infrastructures push the envelope for data center density, and require efficient, high-bandwidth connectivity with minimal cabling, as well as energy efficiency. Simultaneously, multi-tenancy infrastructures demand granular, deterministic management of traffic that can tie QoS to SLAs, and guarantee security. InfiniBand meets these requirements, and as the concept of private cloud infrastructures takes hold in the enterprise, InfiniBand will make more inroads.

Jeff Boles is a senior analyst and director of validation services with **the Taneja Group** research and consulting firm.

Sidebar: The Exadata cluster

In 2008, HP and Oracle jointly released the HP Oracle Database Machine. LGR Telecommunications is one customer demonstrating the type of data demands that require such high performance database solutions. LGR's CDRLive solution is behind real-time data analysis at the biggest telecom companies in the world, and on a daily basis is responsible for helping providers understand traffic patterns, user interactions, and more.

One of LGR Telecom's CDRLive installations endures the constant loading of 40,000 rows of data per second, while supporting the on-going interaction of more than 2,500 users querying the data as it is loaded. By switching to the InfiniBand backbone in an Exadata cluster, LGR found a 20x performance increase over the prior 128-core, highly-tuned HP Superdome / Oracle 10g environment.

Exadata has allowed LGR to package a solution that can be easily dropped into telecom provider data centers as a single rack solution. According to Paul Hartley, general manager at LGR Telecommunications, "One cannot underestimate the huge benefits provided by InfiniBand. It has fundamentally changed our perspective on networking inside the data center."

A number of other vendors leverage InfiniBand for low latency, high performance, availability, consolidation, or other capabilities. DataDirect Networks, for example, delivers native InfiniBand storage systems to ingest large amounts of video content in its SeaChange and S2A9900-based xStreamScaler solutions. Other examples include clustered InfiniBand storage nodes from Isilon serving up geospatial data to compute clusters; the military taking advantage of InfiniBand in LSI's Engenio 7900 storage systems; and other vendors such as Atrato and Fusion-IO employing InfiniBand to exploit the full performance of their solid-state storage systems.

Sidebar: Datacenter-as-a-service

A-Server, based in Lochristi, Belgium, announced its Datacenter-as-a-Service (DAAS) solution early this year. DAAS delivers complete sets of computing infrastructure – storage, CPU, and network – either hosted within an A-Server physical data center or as a fully managed set of equipment at a customer's premises. Together, the components inside an A-Server DAAS look similar to an Amazon Web Services environment, where customers can access the infrastructure over an IP network and fire up virtual server images from shared block storage to build any assembly of application and storage services.

Hosted infrastructures of the scale that A-Server is building require shared storage. Without shared storage, reconfiguration – for availability, load balancing, or routine management - can't be accomplished. A-Server hosts boot images off of high performance centralized storage, but low latency and flexible host attachment is paramount. InfiniBand delivers on both counts. iSCSI over InfiniBand allows A-Server to simultaneously boot huge numbers of servers with low latency and high bandwidth. Simultaneously, the flexible, highly available InfiniBand fabric allows A-Server to move virtual guests for capacity or failure management without complex zone or host- port reconfigurations.

Related articles:

[Mellanox ships InfiniBand switch for I/O consolidation](#)

[Dell to resell Xsigo's virtual I/O products](#)

To access this Article, go to:

http://www.infostor.com/infostor-2/en-us/index/articles/display.articles.infostor.san.the-role_of_infiniband.html