TANEJA GROUP
TECHNOLOGY ANALYSTS

T E C H N O L O G Y B R I E F

InfiniBand and the Enterprise – evolving demands for high performance fabric

May 2009

THE NET NET

For the enterprise IT practitioner, the matter of choosing a computing fabric often falls far short of receiving the center stage attention it seems the matter *should* deserve. Familiarity and time have set the infrastructure architect on a seemingly intransigent course toward settling for fabrics that have become the defacto norm. On one hand is the ubiquitous Ethernet network that is often looked to 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 the Fibre Channel Storage Area Network (FC SAN) that brings more deterministic performance and bandwidth to the table but at the cost of forever bifurcating the enterprise network while doubling the cabling, administrative overhead, and cost of ownership.

But even with separate fabrics, these networks can no longer stand up under today's demands. So slowly but surely, the enterprise is waking up to the potential of new fabrics. In our opinion, it is about time. There is no other component of the data center that is as important as the fabric. But the IO challenges facing every enterprise today have in fact been setting the stage for next generation fabrics for several years, and for that reason, when it comes to maturity and range of capabilities, you can often find the industry pioneers at the head of the pack. In turn, we find one of the most mature solutions – InfiniBand – at the forefront when users are driven to select a new fabric. Frequently, new customers have been surprised at just what InfiniBand can deliver, and just how easily it integrates into today's data center architectures. But it isn't hard to imagine why - among the crowd of next generation fabrics - Fibre Channel, Ethernet, PCIe and others – InfiniBand has longest been leading the charge toward consolidated, flexible, low latency, high bandwidth, lossless IO connectivity. While other fabrics have just turned to addressing next generation needs – with technologies such as FCoE still seeking general acceptance and even standardization – InfiniBand's maturity and massive bandwidth advantages continue to capture new customers.

Although InfiniBand remains a small industry compared to the Ethernet juggernaut, InfiniBand continues to grow aggressively, and this year it has grown well beyond even aggressive projections. Market buzz may be missing the velocity of InfiniBand adoption, and why and where users are turning to the technology. In our observation, the IO drivers are clear. We'll turn to look in more depth at what is driving the enterprise to consider new fabrics today, and examine a few use cases where the enterprise and the InfiniBand fabric are finding they are an ideal match for each other.

87 Elm Street, Suite 900 Hopkinton, MA 01748 Tel: 508-435-5040 Fax: 508-435-1530 www.tanejagroup.com

Today's IO Dilemma – driving the enterprise to new fabric architectures

Today the enterprise is on the frontline of a barrage of IO demands. There are more applications and multi core multi CPU systems generating more IO in every corner of the enterprise. A short list naming only a few might include bigger and more sophisticated ERP systems; business intelligence systems that now analyze years of data and might even encompass video analytics; virtual servers and desktops hammering shared golden boot images; content creation and transcoding applications; availability and protection tools that transfer enormous amounts of data; and web application workloads that handle ever richer content and larger numbers of customer transactions that seem unbounded in growth.

Moreover, compared to merely a few years ago when making the most of IO was an exercise in tuning limited processor and bus architectures, more IO can now be demanded by a single server than most fabrics can handle. Case in point: the AMD Shanghai and Intel Nehalem processor microarchitectures that are currently storming the market in the latest processor bandwidth wars. Each has been designed for the holy grail of IO performance improvement, and has effectively doubled the capabilities 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 the existing infrastructure.

The enterprise's traditional answer to more IO – namely more adapters and cabling – cannot be the solution in an enterprise that is now faced with power, cooling, space, and physical management constraints. The enterprise that addressed previous IO challenges by adding more segregated connections to more separate networks, separate fabrics, or even local buses, can simply no longer pursue these practices. The enterprise is out of power, out of space, and out of muscle to run and manage these tangled webs of traditional connectivity.

The Unified Fabric

In our observance of in-the-trenches IT practitioners groaning under the daily burdens of such infrastructures, it has become clear to us that the only answer is a single unified fabric that can: (1) reduce cabling, (2) connect increasingly greater numbers of ever-denser servers within the power and airflow limits of the data center, and (3) deliver the performance in bandwidth, latency, and losslessness to meet the full spectrum of traffic and storage demands in the next generation of consolidated enterprise computing. As is clear from the range of solutions entering the market, these characteristics equate to nirvana for the fabrics behind next generation computing. But while many solutions are still formulating their plan for delivering these characteristics across every protocol that a single wire fabric may serve, InfiniBand has long had the technological maturity to deliver. While InfiniBand may

not be on a drive to become the end all fabric for everyone, it is increasingly finding adoption in pockets of the enterprise, and for good reason.

InfiniBand as the Next Generation IO Solution

InfiniBand as a next generation data center fabric finds adoption because it doesn't necessitate forklift upgrades in order to tackle localized IO problems, and because it can simultaneously tackle sticky IO problems with greater maturity than competing fabrics. InfiniBand can easily be applied within the rack, across virtual servers, within clusters, or to similar problem domains. And it is just these types of collections of servers and infrastructure where enterprise IT is facing sticky IO problems today. The IO 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 now might demand as many as 400,000 IOs per second.

And when it comes to performance for these intense environments, few technologies are as well equipped as InfiniBand. While organizations like the IEEE wrestle with standardizing Ethernet mechanisms like 802.1Qbb per-priority pause, InfiniBand has losslessness and deterministic quality of service in its genes. Meanwhile, intrinsic to InfiniBand is a deep coupling into the very heart of processor and memory buses via RDMA-based protocols that can deliver high performance networking like no other fabric around. Compared to tunneled or encapsulated approaches like iWARP over Ethernet that vendors are still trying to come to terms with, InfiniBand can deliver enormous efficiency.

At the end of the day, when InfiniBand steps into the picture for these reasons, there is still a place for traditional networks, just not where they are ill equipped to deliver. With seamless, transparent integration into existing data centers, InfiniBand is making the case that there is room for both technologies, old and new, and that with this approach, InfiniBand can solve some sticky challenges.

This year we have looked at the state of InfiniBand adoption in the enterprise to draw out a few illustrative examples of where InfiniBand is finding adoption.

Scalable Compute

The enterprise has fearlessly jumped into the realm of aggressive computing around huge datasets, and InfiniBand has often ended up playing a role. Today, we find a majority of the Fortune 1000 businesses we talk to are involved in high throughput processing behind a wide variety of businesss systems: business analytics, content creation, content transcoding, real-time financial applications, messaging systems, consolidated server infrastructures, and more. In this case, InfiniBand has often 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, HP created the InfiniBand-based HP Oracle database machine (Exadata) as a mainstream platform for the enterprise. That solution is a clustered grid of 8 Oracle RAC servers coupled to 14 Exadata storage servers, running over a simple, consolidated, low latency and high bandwidth InfiniBand fabric of 4 Voltaire 9024 24 port InfiniBand switches.

While the underlying technologies are significant, the practical use of Exadata is dependent upon InfiniBand. Voltaire's switching platform within the Exadata cluster harnesses InfiniBand's high availability to perform path optimization and/or losslessly restructure the entire fabric if any significant event or outage occurs. Such stateless, transparent high-availability is a key requirement for out of the box enterprise platforms.

Moreover, core InfiniBand technologies are deeply integrated with the Exadata solution. Specifically, while core Exadata technology distributes some data intelligence to the storage nodes, each of those nodes is dependent upon 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 7GB/s. Those protocols let Exadata access each distributed processing node (Oracle RAC or Exadata storage) with minimal latency and minimal host processing overhead.

The Exadata Cluster

In 2008, HP and Oracle jointly released the HP Oracle Database Machine. LGR Telecommunications is one customer demonstrating just what type of data demands are calling for 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 – key data that is at the heart of day to day network operations.

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 over 2500 users querying this real-time data as it is loaded. By switching to the InfiniBand backbone in an Exadata cluster, LGR found a 20x out-of-the-box performance increase over the prior 128 core, highly tuned HP Superdome / Oracle 10g environment. Meanwhile, Exadata has allowed LGR to package up a solution that can be easily dropped into telecom provider data centers as a single rack solution, require minimal management, yet integrate seamlessly with surrounding infrastructure. According to Paul Hartley, General Manager at LGR Telecommunications, "One cannot underestimate the huge benefits provided by InfiniBand, and it has fundamentally changed our perspective on networking inside the datacenter."

Other examples of solutions that have turned to InfiniBand for latency, performance, high availability, consolidation, or other capabilities abound, with vendors such as DataDirectNetworks delivering native InfiniBand storage as an to ingest large amounts of video content in their SeaChange and S2A9900-based xStreamScaler solution, clustered InfiniBand storage nodes from Isilon serving up geospatial data to compute clusters, the military harnessing the native InfiniBand in LSI Engenio 7900 storage for various top secret projects, and other vendors like Fusion-IO and Atrato employing InfiniBand to deliver the full performance of their solid state storage solutions.

Server Virtualization

Our research shows that server virtualization continues to be one of the top 5 IT priorities, and that over 1/3 of the enterprise workloads being deployed today go into virtual environments. More often than not, server virtualization proves to be a torture test for existing fabrics, and we have in turn seen increasing numbers of enterprises turn to localized InfiniBand fabrics.

First and foremost behind this adoption, is the need for both bandwidth and low latency in these consolidated infrastructures. While previous generation server architectures often constrained overall IO performance, the latest generation microarchitectures and the widespread adoption of PCIe 2.0 have set IO free. Today, virtual guests can readily demand bandwidth in excess of 500Mbps, leading virtual hosts to exceed even 10Gbps Ethernet capabilities.

Simultaneously, the virtual infrastructure is rife with innovations, and features for increasingly fluid workload movement, as well as sophisticated failover and fault tolerance seem to spring forward like ants marching on. Every innovation around these technologies moves the virtual infrastructure toward the vision of a datacenter architecture where workloads can be seamlessly, continuously protected, and transparently moved across physical systems upon demand. Yet every innovation has underlying impacts upon the fabric, whether it is latency dependent traffic generated by IO synchronization across multiple systems, or the bandwidth consumption from the high-speed pushing and pulling of server images across a data center.

Finally, the consolidated virtual infrastructure is dense, and compounds problems with power, cabling, and cooling. We commonly see virtual server configurations with as many as 8 Fibre Channel and/or Ethernet adapters in a single physical host. When compared to a single 40Gbps QDR InfiniBand fabric that outperforms every other fabric in watt per gigabit, traditional fabrics can more than double what it costs to operate and manage the IO behind the virtual server infrastructure.

For these reasons we find more customers deploying InfiniBand as a localized, high bandwidth network behind their virtual infrastructures, with innovators like Voltaire, Xsigo and 3Leaf Systems driving adoption by delivering even more sophisticated levels of management on top of this one unified fabric.

Enterprise Clouds

Turning to look at InfiniBand use within current hotbeds of innovation calls attention to the cloud. 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 recognized as one of the cutting edge technologies behind leading cloud providers.

Cloud workloads demand freedom from physical resources, so that they can be dynamically reprovisioned or moved as business demands change. Take for example A-Server, a cloud provider we've highlighted in the sidebar. To deliver fluidity around workload management within a cloud infrastructure, 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, minimum

Datacenter-as-a-Service

The company A-Server, based in Lochristi, Belgium announced their Datacenter-as-a-Service (DAAS) solution early 2009. DAAS delivers complete sets of computing infrastructure – storage, CPU, and network – either hosted within an A-Server physical datacenter or as a fully managed set of equipment at a customer's premises. Together, the componentry inside an A-Server DAAS looks 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 imaginable assembly of application and data 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 - simply can't be done. 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 best-in-class low latency and high bandwidth. Simultaneously, the flexible, highly available IB fabric allows A-Server to move virtual guests for capacity or failure management without complex zone or host port reconfigurations.

cabling connectivity, and extreme power efficiency. Simultaneously, multi-tenancy infrastructures at cloud scale demand granular, deterministic management of over the wire traffic that can tie Quality of Service to SLAs, and guarantee no less than concrete security. On these counts InfiniBand delivers better than others, and as the concept of private cloud infrastructures takes hold in the enterprise, we have little doubt we'll see even more InfiniBand enter into the enterprise.

Taneja Group Opinion

While we have always been bullish on next generation fabrics, we have also always held that such solutions will find the majority of their success by solving specific problems in the traditional data center. Whether it is for connectivity within computer clusters that are now common in mainstream computing, the IO throughput behind virtual servers, the dynamic interconnects within cloud infrastructures, or dozens of other use cases, traditional fabrics cannot address the emerging IO problems facing the enterprise. At every turn, additional demands continue to energize the market for the many evolving fabric solutions available today.

But the truth is that long ago this market began evolving on the fringes of mainstream computing, and InfiniBand has the incredible maturity that is born from cutting its teeth through delivering real world solutions behind demanding compute environments. Consequently, while other fabrics coming to market draw attention to IO problems, InfiniBand gets customers. Moreover, this shows up in real world numbers and is the reason that the adoption of 40Gbps QDR adapters is far outpacing even the most aggressive analyst predictions. Sophisticated InfiniBand switches from vendors like Voltaire, Qlogic, and Mellanox have long delivered the same benefits emerging solutions are often still just hoping to deliver with other fabrics.

Administrators should be attentive to how such fabrics are evolving, and investigate where there are justified opportunities to integrate these technologies with their infrastructures. When the case for a new fabric is clear, consider just what is needed from localized high bandwidth networks, without being fooled by excess complexity or the brand name bandied about by traditional vendors of choice. When it comes to maturity, transparency, and sophistication, InfiniBand has long carried the flag. In the evolving marketplace for next generation fabrics, we see this leadership steadily reflected in InfiniBand's adoption in the enterprise.