Cloud Service Provider Delivers Higher and More Reliable Service Levels Using Mellanox InfiniBand Products

Commodity server and storage components are used to deliver enterprise class reliability and high availability, improving ROI

Situation

Logicworks is a leading provider of highly reliable cloud hosting solutions. The company offers its clients managed cloud services with high service levels in the areas of high availability and increased transaction throughput.

Logicworks' cloud data center infrastructure is required to deliver the following important success criteria:

1. **Server agility and utilization** using VMware server virtualization software and commodity server hardware. Ability to deliver more VMs per server.

2. **Increased storage area network and storage array performance** to deliver higher levels of SLA (service level agreement), without compromising high availability and reliability. Ability to service growing customer base without sacrificing SLAs.

3. The infrastructure is required to **consolidate I/O transparently** to reduce capital and operational expenses. Applications in virtual machines (VM) are required to run transparently over virtual network and storage adapter interfaces provided by VMware server virtualization software.

4. To further reduce capital expenditure and pass on their savings to cloud customers, Logicworks further determined that they should **use commodity hardware and open source software platforms** to the farthest extent possible without compromising customer SLAs.

Logicworks explored multiple I/O fabric and storage options. Gigabit Ethernet and Fibre Channel-based LAN and SAN I/O options were considered versus consolidated I/O options using InfiniBand and Fibre Channel over Ethernet (FCoE). Enterprise-class FC storage array options were considered for highly available storage infrastructure and high SAN performance from VMs.

Logicworks performed exhaustive analysis of all options and decided to utilize the following off-the-shelf hardware and software components to meet the four goals described above:

- VMware vSphere on Supermicro servers based on Intel Xeon Processor 5500 series

- 40Gbps Mellanox InfiniBand adapters and switches to connect the server and storage infrastructure

- Open source OFED (OpenFabrics Enterprise Distribution) based high performance storage initiator and target software

- Commodity storage hardware

- LINBIT's DRBD software for storage replication and high availability

The rest of the case study discusses how the above four goals were addressed using the above hardware and software components.

Solution

The solution is presented from the perspective of the four goals Logicworks wanted to achieve with stress on I/O requirements and considerations.

Server agility and utilization:

Since the requirement was to use VMware server virtualization software, all I/O fabric options were considered and it was determined that disparate fabric technologies such as Gigabit Ethernet and Fibre Channel I/O technologies are supported, as

well as consolidated and higher performance I/O such as InfiniBand, 10GigE/iSCSI and FCoE.

Use of 40Gbps InfiniBand with VMware ESX Server enabled replacing multiple GigE and FC adapters, switches, ports and cables with dual InfiniBand ports/adapters, switches, fewer cables. Price/port considerations compared to 10GigE/iSCSI and FCoE (see table below), made InfiniBand the obvious choice for consolidating I/O. Higher LAN and SAN throughput over InfiniBand enabled hosting of more VMs per VMware ESX server. Besides, cost and performance, another key benefit InfiniBand delivered is its hardware based multi-pathing, fast fail-over and reliability characteristics that are not available over Ethernet fabrics used with 10GigE/iSCSI and FCoE. This is discussed more in detail in the sections below.

	10GigE/iSCSI	10GigE/FCoE	40Gbps InfiniBand
Adapter per port cost	$790	$1,100	$755
Switch per port cost	$377	$410	$200

Table1: January 2010 pricing for a single port adapter and a switch port
(Source: Colfax Direct, CDW and SAN Direct)

Increased storage area network and storage array performance:

Logicworks was able to improve VM SAN performance and at the same time reduce costs significantly by taking the following steps:

1. Utilized SCSI RDMA Protocol (SRP) implemented in VMware ESX server over a single converged I/O InfiniBand adapter to deliver 4 times the performance compared to a single 4Gbps FC adapter. This higher performance is a result of combination of an efficient initiator stack, a 40Gbps high bandwidth pipe connecting server and storage, use of SRP target on the storage systems with adequate number of backend disks to deliver the needed storage capacity.

2. Utilized the InfiniBand hardware based multi-pathing and incredibly fast failover capabilities to create a highly available and reliable storage target array configuration using commodity hardware platform. This is shown in figure 1 below. Two such storage arrays were created using InfiniBand front end connectivity and open source

SRP target implementations, one of them acting as primary and the second acting as a standby. LINBIT's DRBD software was used on the target storage systems to actively mirror data written on the primary to the secondary using the IPoIB protocol over the InfiniBand adapters. In case of failure in the primary storage system, Mellanox InfiniBand's hardware based multi-pathing capabilities delivered an instant failover – less than five seconds –to the secondary storage system, without causing any disruption to the applications running in the VMs and therefore services delivered to cloud customers.
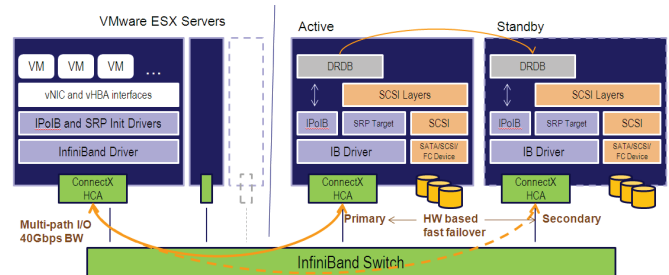

Figure 1: Logicworks' highly available storage solution

Consolidate I/O transparently:

LAN and SAN connectivity for VMs in the VMware ESX server or vSphere environment is delivered using IPoIB and SRP initiator drivers integrated into VMware's VMkernel or hypervisor. Using Mellanox Virtual-IQ technology, they provide multiple vNIC and vHBA interfaces respectively that from the VMs perspective are the same as what is available over multiple GigE NICs and FC HBAs respectively. Also, some of the vNIC interfaces over IPoIB may be allocated to VMotion or other management traffic. Configuration of such vNICs and vHBAs enabled by Virtual-IQ technology can be accomplished using standard VMware management tools such as the Virtual Center. In essence, from the VMs and VM management perspective, InfiniBand as the underlying fabric is completely hidden, helping consolidation of LAN, SAN, VMotion, Console and other traffic in a transparent fashion, helping reduce operational costs.

Use commodity hardware and open source software platforms:

As explained earlier, LogicWorks is able to use commodity hardware for both its server and storage system infrastructure, thanks to InfiniBand. On commodity servers, it is able to deliver more performance for applications running in VMs and it is able to support more users per server through installation of more VMs per server. It is able to eliminate the need to use expensive and power hungry FC storage arrays by using commodity hardware, open source software such as SRP target software and DRBD software to build highly reliable and high performance storage targets.

Results

Deploying Mellanox InfiniBand connectivity solution proved to be the right decision and helped Logicworks to meet their four goals.

Capital expenditure was reduced significantly. Only one InfiniBand dual port adapter is needed to replace four GigE NICs and two 4Gbps FC HBAs per ESX server. This consolidation saved $1,500 per ESX server.

Logicworks saved $40,000 or more per storage system by not purchasing expensive FC storage arrays and building the high performance, reliable and redundant storage infrastructure by itself using a commodity available server that cost $5000.

The InfiniBand hardware-based automatic and fast "failover" to the DRBD replicated storage has been proven to be very reliable and takes less than 5 seconds, enabling non disruptive service to its customers.

Cutting the number of network adaptors from six to one per server, reduced the number of switch ports and cables further lowering overall energy costs and reducing cabling complexity. Overall it simplified the cloud management, reduced power and space and results in a significant saving in operational expenses.

As a result, Logicworks provides its cloud customers a competitive service offering at a significant lower total cost of ownership compared to cloud services that utilize traditional Ethernet and/or Fibre Channel based network solutions.

Mellanox
TECHNOLOGIES

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