A Whitepaper on

Building Data Centers with Dell MXL Blade Switch

Product Management
Dell Networking
October 2012
Abstract

In today’s data-driven and demanding data centers, customers are actively looking to enhance their offerings using cost effective infrastructure and solutions. Data centers are transforming due to data explosion which is driving newer deployment models, better speed and feed, virtualization, multi-tenancy, I/O convergence, workload optimization and simpler management solutions. In this whitepaper, we will discuss the new breed of customer requirements for data centers and how the Dell Force10 MXL blade switch serves the evolving needs.

Data Center is Evolving

Several trends are evident in today’s data centers and will continue to drive the future evolution of data centers as well.

Growing blade servers: Data explosion in the recent years is driving more capable, better and faster compute capabilities in addition to its networking and storage counterparts. In the server market, blade servers are growing faster than the rack and the tower servers (Figure 1). IT managers are optimizing their data centers for efficiency in power, cooling, cabling, management and density contributing to the growth in blade server market.

Figure 1: Server Market Forecast (IDC Q1 2012)

Dell’s extensive server portfolio includes blade, rack and tower options. The new 12th generation servers are best suited for ever increasing workloads in compute intensive environment. Dell M1000e offers an advanced blade chassis that hosts blade servers of full-height, half-height and quarter-height allowing 32 blade servers in 10 rack units.

More I/O on the network: The blade servers are evolving with newer processors, more memory and better capabilities. Many customers are moving from two-socket CPUs to four-socket CPUs. CPUs are evolving from 4-cores to 8-cores to 16-cores. A major shift in the IT industry has already started due to high adoption of server virtualization. Many customers are deploying tens of VMs per physical server.
With server virtualization come higher server utilizations and the need for more I/O. Storage devices are riding the speed and feed trends in addition to addressing more capable features. The evolving servers and storages are driving network connectivity. More processing power drives more traffic and customers are moving from 1G to 10G LOM/daughter card/mezzanine card albeit multiple of them such as dual 1G to quad 1G to dual 10G. The networks adapters are also evolving in their capability to support convergence. Customers now have an option to select Converged Network Adapters thus reducing number of adapters by 50% that otherwise would be needed to support LAN and SAN networks. New technologies like NIC partitioning (NPAR) allow the adapter to be logically partitioned into multiple partitions with each partition dedicated to specific traffic types or VLANs (for example, LAN traffic on partition 1, SAN traffic on partition 2 and so forth). These technologies reduce the need for several adapters, associated cables and upstream switch ports offering a more cost effective alternative.

**Converged Infrastructure and networks:** Many customers are converging their LAN and storage traffic on the same network. Whether the storage is NAS (NFS or CIFS), iSCSI or FCoE, IT managers are looking at converged infrastructure to reduce TCO using Converged Network Adapters (CNA) with fewer cables/optics and switches. The industry analysts expect that over the next 2-3 years there will be greater adoption of converged infrastructure. I/O Convergence drives more traffic onto the adapters and the switch ports they are connected to. This drives more traffic in and out of blade enclosures. The growth in traffic eventually drives the need for highly scalable and high performing connectivity. Hence, customers are migrating from 1G to 10G infrastructure.

**Growing East-West traffic:** Another trend is that the traffic growth in data center is primarily east-west bound. According to Gartner, “The consequence of changing application deployment and increased use of VM migration will mean that traffic patterns in the data center network are changing from being predominantly client/server (north-south) to a significant level of server-to-server (east-west) flows. By 2014, network planners should expect more than 80 percent of traffic in the data center network to be between servers.” (April 2011, “Your Data Center Network is Heading for Traffic Chaos”). Server virtualization, multi-tier applications, distributed applications, cloud based applications are all driving this growth in east-west traffic.

**Investment protection and multi-vendor deployments:** Yet another aspect is future protection and pay-as-you-grow capability. Customers are always looking for future expansion capability without having to fork-lift upgrade their hardware. Deployments using open standards provide more flexibility than proprietary counterparts. Often times, IT managers prefer multiple vendors in their environment and there comes the question of interoperability and seamless integration of multi-vendor gears.

All these aspects discussed so far are contributing to data center transformation and the fundamental components affected by this transformation include server, storage, networking and management. Dell offers a comprehensive portfolio of all these components. Dell solutions offer industry leading data center products, open standard architectures, and best TCO for data centers. Within the data center products, Dell recently introduced the MXL 10/40GbE blade switch. In the subsequent sections, we will discuss how this blade switch meets the evolving needs for data centers.

**Dell Force10 MXL 10/40GbE Blade Switch**

**MXL is designed to handle growing I/O need for your data centers:** Dell Force10 MXL switch is industry's first 40GbE capable, modular and stackable blade switch. It fits into the Dell M1000e enclosure (Figure 2). The MXL blade switch is an addition to the Dell M-Series portfolio and is designed to conform to Dell M-Series IOM specifications. It resides in the back of the 10u M1000e enclosure and does not require additional rack space. It shares M1000e supplied power, cooling, and
management infrastructure and eliminates 100% of traditional external server-to-switch cabling, components, and related infrastructure. This increases data center efficiency and optimizes operations.

Figure 2: High level value proposition of Dell MXL blade switch

The MXL switch supports 32 internal/server facing 10GbE ports and 2 40GbE fixed ports for uplink. It offers the highest number of blade servers in the chassis in the industry offering the highest density solution you can pack in a 10RU chassis. The M1000e chassis can support up to 6 MXLs although typical configuration of 2 MXLs in Fabric A1 and A2 provides fully redundant connectivity for up to 32 blade servers. This is a high performing and high density L2/L3 switch.

**MXL addresses pay-as-you-grow need and offers flexibility:** MXL also supports two optional slots where you can plug in Flex IO modules (Figure 3) addressing pay-as-you-grow need. Three types of Flex IOMs are supported on MXL switch – 4 port SFP+ module, 4-port 10GBASE-T module and 2-port QSFP+ module. The 40GbE can be realized as 4x10GbE using breakout cable. The base switch offers 2 40GbE uplink ports for high bandwidth applications offering a 4:1 oversubscription. With MXL, customers can scale up uplink bandwidth up to 240GbE as their bandwidth needs increase by adding Flex IOMs. In addition, the ability to support 32 internal ports allows customers to “pay as they grow” from a server growth perspective. The flexibility of having different uplinks such as QSFP+, SFP/SFP+ and 10GBASE-T/RJ45 copper allows connectivity into just about any type of upstream networks in data centers.
MXL is architected to better handle East-West traffic: MXL is architected to address the ever increasing east-west traffic within data centers. MXL supports local switching which keeps traffic within the chassis as opposed to forwarding that traffic to the Top of Rack (ToR) switch (Figure 4). MXL also offers stacking features and you can stack up to 6 MXLs where the MXLs can be on the same or different chassis using ring or daisy-chain topology. The stacking feature allows server to server traffic to remain local even across chassis. This can scale up to 192 physical servers across 6 M1000e chassis that can communicate without ever having to pass through a ToR switch. Both these features imply that you would require fewer ToR ports which are typically expensive. For example, if you needed 4x10GE links to connect your blade IOM to ToR, with MXLs, you would need fewer than 4x10GE links as you can retain some traffic within the chassis using local switching and within the rack using stacking given your application allows intra-chassis and intra-rack I/O.

Figure 3: FlexIO Modules for Dell MXL blade switch

Figure 4: With MXL, all traffic does not need to be switched at ToR
**MXL offers converged I/O.** MXL supports Data Center Bridging (DCB) protocols – Priority Flow Control (PFC), Enhanced Transmission Selection (ETS) and Data Center Bridging Exchange (DCBX). The DCB protocols allow lossless Ethernet environment where Ethernet traffic and SAN traffic can be sent over the same physical link and yet maintain the needs for a lossless SAN network. In addition, MXL supports iSCSI and FCoE protocols. EqualLogic iSCSI arrays can be directly connected to the MXL switch or can be connected via Top of Rack switch (Figure 5). MXL supports FIP Snooping Bridge (FSB) and can be configured as a FCoE Transit switch. The FCoE traffic from CNA is simply passed through the FSB to the upstream converged switch where the Ethernet/FC traffic is split. MXL switch can also be directly connected with NAS storage as well.

![Converged LAN and SAN using Dell MXL blade switch](image)

**Figure 5: Converged LAN and SAN using Dell MXL blade switch**

**MXL Enables Data Center in a Chassis:** With the recent launch of EqualLogic PS-M4110 iSCSI blade array, Dell offers a converged data center in chassis M1000e using M420 quarter height blade servers, MXL 10/40GbE blade switch and PS-M4110 arrays (Figure 6). This solution is very compelling for customers who are looking to deploy a simple yet robust solution.
In addition, MXL supports NPAR technology which allows NIC partitioning using logical partitions on the CNAs.

**MXL seamlessly interoperates with other vendors:** MXL brings the Force10 Operating Systems (FTOS) in blade enclosures. FTOS has been battle hardened over the past decade and deployed in major global customer accounts. FTOS offers highly resilient L2/L3 features. The CLI is very similar to Cisco CLI. MXL supports Per Virtual Spanning Tree Plus (PVST+) feature that maintains a Spanning Tree instance per VLAN and uses IEEE 802.1Q trunking. This is readily interoperable with other vendor hardware that supports PVST+ or Rapid PVST+.

MXL is currently managed by CLI and external management tool called Open Manage Network Manager (OMNM). MXL also supports Bare Metal Provisioning which is quite effective to bring up a number of MXLs with zero touch. BMP loads the right FTOS image and also sets up configuration automatically. In future, other Dell management applications will support MXL.

**MXL Deployment:** Now let us take a look at how MXL plugs into the overall data center architecture. Dell networking offers a multitude of platforms and an array of choices using technologies such as Distributed Core and Virtual Link Trunking (VLT). VLT is a L2 multi-pathing technology. MXLs can be connected to ToR or directly to core using its high bandwidth uplinks (Figure 7).
Figure 7: MXL can be connected to ToR or Core network

Figure 8 shows a configuration where 2 S4810 form VLT. Each rack contains 3 M1000e chassis with 2 MXLs each. The MXLs are vertically stacked in each rack. Each MXL is connected to each S4810 using one 40GbE uplinks. This configuration shows 2 racks, 192 10GE blade servers using 4:1 over subscription in a fully redundant way. Note that you can stack 6 chassis addressing up to 192 servers in the same stacking domain.

Figure 8: An example topology with MXL and S4810
Customers can build much larger networks using MXLs connected to distributed core leaf nodes or spine nodes. These architectures address large to extra-large networks offering thousands of 10GE ports for server connectivity.

**Summary**

In this document, we discussed the trends in data centers that are driving the transformation due to data explosion and how MXL blade switch along with other Dell products offers a cost optimized solution. Dell MXL blade switch empowers the switching capabilities using Force10 technology in Dell M1000e blade chassis offering unequalled server density. It offers industry’s highest bandwidth blade switch with 40GbE capability. It supports local switching and stacking to address increasing east-west traffic in data centers reducing the TCO. The MXL switch offers I/O convergence thus reducing the infrastructure cost. It is a high performing blade switch that is ready for today’s demanding data centers and for evolving needs with its pay-as-you-grow offering.