IP NETWORK

Brocade MLX Series Architecture

The Brocade MLX Series uses an innovative Clos-based fabric architecture, which provides a high level of scalability, redundancy, and performance.

BROCADE

The Brocade MLX Series of routers is the most powerful suite of IPv4/IPv6/MPLS/Multi-VRF switching routers in the industry. It is a cost-efficient solution that is purpose-built to handle the most demanding applications with non-blocking, wire-speed performance. The Brocade MLX Series has a robust system architecture and a versatile feature set, and it is available in four different sizes, making it capable of scaling from the edge to the core.

INTRODUCTION

Designed with state-of-the-art network processing technology, the Brocade® MLX® Series has a non-blocking switching capacity of 15.36 terabits per second (Tbps), a data forwarding capacity of 12.8 Tbps, and total routing performance of 9.6 billion packets per second (Bpps). Its advanced distributed hardware architecture with fine-grained Quality of Service (QoS) support enables uncompromised full-duplex, wire-speed performance for any mix of IPv4, IPv6, and Multiprotocol Label Switching (MPLS) services. These capabilities are made possible by an innovative system architecture with these distinguishing characteristics:

- Clos-based self-routing, distributed, non-blocking architecture provides the foundation for a robust, scalable platform.

- Distributed network processing and advanced QoS capabilities across the system allow a rich set of features to be implemented at wire-speed rates.

- The Brocade MLX router features a high-availability architecture with a clear separation between control plane and data plane.

- This architecture is fully redundant, with redundant power supplies, management modules, fan trays, and Switch Fabric Modules (SFMs), to avoid any single point of failure.

The Brocade MLX Series is available in the following two product lines:

Brocade MLXe® Series Routers

- Brocade MLXe-32 (32-slot chassis)
- Brocade MLXe-16 (16-slot chassis)
- Brocade MLXe-8 (8-slot chassis)
- Brocade MLXe-4 (4-slot chassis)

Brocade MLX Routers

- Brocade MLX-32 (32-slot chassis)
- Brocade MLX-16 (16-slot chassis)
- Brocade MLX-8 (8-slot chassis)
- Brocade MLX-4 (4-slot chassis)

The management and interface modules can be interchangeably used across any of these systems, which decreases inventory and maintenance costs across the network. All modules are hot-pluggable.

HIGH-DENSITY CORE ROUTING

The Brocade MLX Series is scalable to high-density of 32 × 100 Gigabit Ethernet (GbE) ports, 128 × 40 GbE ports, 768 × 10 GbE ports, or 1536 × 1 GbE ports in a single chassis. In a standard 7-foot (2.1 meter) telco rack, the Brocade MLXe Series can support up to 1152 × 10 GbE ports, 2304 × 1 GbE (MRJ-21) ports, or 960 × 1 GbE (RJ-45 or Small Form-Factor Pluggable [SFP]) ports. The Brocade MLX Series also supports up to 256 × OC-12/48 ports or 64 × OC-192 ports in a single chassis.

SCALABLE CLOS FABRIC ARCHITECTURE

The Brocade MLX Series uses a Clos fabric architecture, which provides a high level of scalability, redundancy, and performance. As shown in Figure 1, there are multiple high-speed Switch Fabric Modules (hSFMs) in the system. A Switch Fabric Module has multiple fabric elements, each of which has multiple connections to every interface slot.

NP = Network Processor
TM = Traffic Manager
FE = Fabric Element
hSFM = Switch Fabric Module
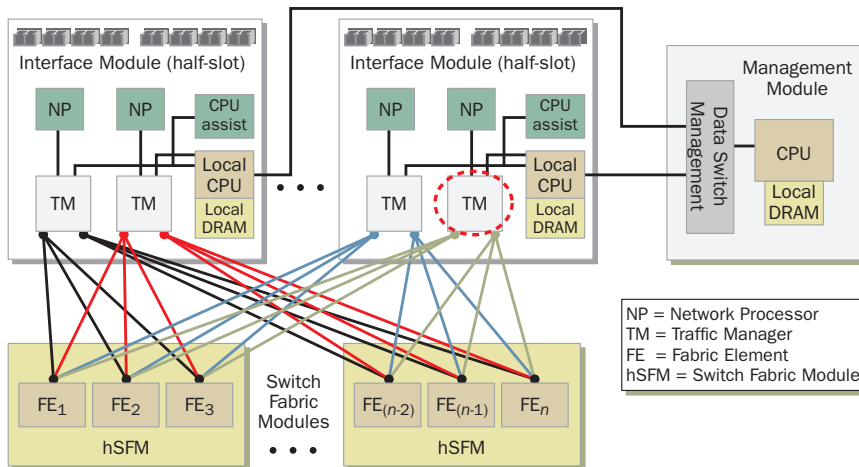
Figure 1.
Clos fabric architecture.

The Clos architecture uses data striping technology to ensure optimal utilization of fabric interconnects. This mechanism always distributes the load equally across all available links between the input and output interface modules. By using fixed-size cells to transport packets across the switch fabric, the switching architecture of the Brocade MLX Series ensures predictable performance with very low and deterministic latency and jitter for any packet size. In addition, the Brocade MLX Series offers a "turbo mode" that increases switch fabric efficiency by using variable-size cells to transport packets. The presence of multiple switching paths between the input and output interface modules provides an additional level of redundancy.

Advantages of a Clos architecture over a traditional architecture include:

- Common architecture across the product family, because the same fabric elements are used on all chassis of the Brocade MLX Series. This demonstrates the superior scalability of the architecture from a small 4-slot system to a large 32-slot system.

- No head-of-line blocking at any point, irrespective of traffic pattern, packet size, or type of traffic.

- Optimal utilization of switch fabric resources at all times. The data striping capability ensures that there is fair utilization of the switch fabric elements at all times without overloading any single switch fabric element.

- "Intra-SFM" redundancy. An hSFM can withstand the failure of some of the fabric elements and still continue to operate with the remaining fabric elements. This unique capability provides a very high level of redundancy, even within an hSFM.

- Exceptional high availability. The Brocade MLX Series supports redundant hSFMs. This allows the Brocade MLX Series to gracefully adapt to the failure of multiple switch fabric elements. Moreover, because there are multiple fabric elements within an hSFM, the failure of a fabric element does not bring down the entire hSFM.

INNOVATIVE PACKET PROCESSING TECHNOLOGY

Built with a state-of-the-art, sixth-generation architecture, terabit-scale switch fabrics, and Brocade MaxScale and VersaScale Packet Processor technology, the Brocade MLX Series has a proven heritage in more than 2,500 networks worldwide. The Brocade MLX Series features the coexistence of Field Programmable Gate Array (FPGA)-based and Application-Specific Integrated Circuit (ASIC)-based packet processors for maximum customer choice (in other words, the "best of both worlds"). Brocade VersaScale Packet Processors provide programmability and feature velocity, without compromise to performance. The Brocade MaxScale-160 Packet Processor provides massive memory, power, and efficiency—delivering industry-leading density and low Total Cost of Operation (TCO). Brocade MLX Packet Processors are designed from Brocade intellectual property in core packet processing and are purpose-built for high-performance provider networks. Brocade VersaScale Packet Processors and the Brocade MaxScale-160 are Software-Defined Networking (SDN)-enabled, supporting the OpenFlow protocol, which enables programmatic control of the network—allowing providers to respond to business needs and dynamic traffic patterns.

Brocade MaxScale-160 Packet Processor

Brocade MaxScale technology is based on 4th-generation high-performance packet processing innovation. Brocade has designed its industry-leading Fibre Channel ASICs for performing switching functions in its Storage Area Network (SAN) switches. Today, Brocade is delivering similar innovation to its high-performance IP routing platform, the Brocade MLX Series. The Brocade MaxScale-160 is designed and tailored specifically for high-capacity service provider and data center core networks with massive memory and high-performance fat pipe. With a 160 gigabits per second (Gbps) capacity, it enables a single router to scale to 12.8 Tbps of data forwarding capacity and a 9.5 Bpps packet routing performance. The Brocade MaxScale-160 has rich multiservice support for Layer 2 and Layer 3 services and provides massive MPLS scaling.

The Brocade MLXe 24-port 10 GbE module is built on the Brocade MaxScale-160 Packet Processor, providing industry-leading 10 GbE density—tripling the existing MLX density to 768 ports in a single router. The Brocade MLXe 24 × 10 GbE module is the most green and efficient module in the Brocade MLX family, due to the extremely low power consumption of the Brocade MaxScale-160 (<45W power consumption per 80 Gbps). This enables network operators to save on operational expenses (cooling and power), while also consolidating the number of devices in the network because of the number of 10 GbE ports supported in a single system.

Figure 2.
Brocade MaxScale-160 Packet Processor.

Brocade VersaScale Packet Processors

The Brocade VersaScale Packet Processor family is designed to enable service innovation through programmability and flexibility, but without compromise to performance. The Brocade interface modules that are built on the Brocade VersaScale Packet Processor technology (such as the 2 × 100 GbE, 4 × 40 GbE, and 8 × 10 GbE modules) provide leading density, zero-packet loss, and line speed for all packet sizes—supporting up to 16K simultaneous multicast groups at line rate. Brocade VersaScale Packet Processors are designed for provider and large-enterprise networks (service-driven IT) and provide a balance of scalability and feature richness. With the advent of virtualization and cloud computing, networks are seeing unpredictable traffic spikes. Brocade VersaScale Packet Processor technology provides extremely deep packet buffering to handle the dynamic nature of this traffic. It also supports large-scale Equal-Cost Multi-Pathing (ECMP), which is ideal for cloud service providers and Web 2.0 companies that need scalable solutions to handle explosive bandwidth growth and to optimize the core for efficient packet transport. The Brocade VersaScale Packet Processors have distributed network processing and advanced QoS capabilities, helping providers tighten their Service-Level Agreements (SLAs) for traditional and value-add cloud services. As customers begin to demand network virtualization through SDN, having the ability to easily add new services becomes vital. Brocade VersaScale Packet Processor technology is SDN-enabled with OpenFlow support today and is ready for future overlay technologies such as Virtual Extensible LAN (VXLAN).

MORE ABOUT THE ARCHITECTURE

Distributed Forwarding

The Brocade MLX Series has a distributed forwarding architecture that combines state-of-the-art network processing technology with a very fast switch fabric to ensure uncompromised, full-duplex, wire-speed performance at any packet size. The use of fast network processors on each interface module allows wire-speed performance to be maintained, independent of the features that have been enabled.

These are several of the capabilities that have been implemented in the multiservice Brocade IronWare® OS software to facilitate distributed packet forwarding and security:

- Distributed Layer 2 MAC address table on each interface module. The management module maintains all the learned MAC addresses and distributes the information that is to be locally maintained on the interface modules. Each interface module locally manages aging of its local MAC addresses and updates the management module in order to keep the MAC table consistent across the entire system.
- Brocade Direct Routing (BDR) technology stores the entire forwarding table in each interface module to allow for hardware forwarding of all traffic.
- Distributed Access Control List (ACL). Each interface module can support up to 224,000 input ACL entries and 128,000 output ACL entries for ACL rules that are applied to local interfaces.

High Availability

Both the hardware and software architecture of the Brocade MLX Series are designed to ensure very high Mean Time Between Failures (MTBF) and low Mean Time To Repair (MTTR). Cable management and module insertion on the same side of the chassis allows ease of serviceability when a failed module needs to be replaced, or a new module needs to be inserted.

The ability to handle the failure of an hSFM, as well as elements within an hSFM ensures a robust, redundant system ideal for non-stop operation. The overall system redundancy is further bolstered by redundancy in other active system components such as power supplies, fans, and management modules. The passive backplane on the Brocade MLX chassis increases the reliability of the system.

Temperature sensors on the system are used to automatically adjust the speed of the fans to maintain an optimal operating temperature. Additionally, there is the ability to automatically power off a module, if the configured temperature threshold is crossed.

The Brocade MLX Series also supports the ability to gracefully shut down a Switch Fabric Module with zero packet loss for a scheduled maintenance event. When this is invoked, the system does not use the links between the interface modules and the decommissioned hSFM.

The modular architecture of the multiservice Brocade IronWare Operating System features these distinguishing characteristics that differentiate it from legacy operating systems that run on routers:

- Industry-leading cold restart time of less than a minute
- Support for a hitless software upgrade
- Hitless Layer 2 and Layer 3 failovers
- Subsecond switchover to the standby management module, if a communication failure occurs between active and standby management modules

Distributed Queuing for Fine-Grained QoS
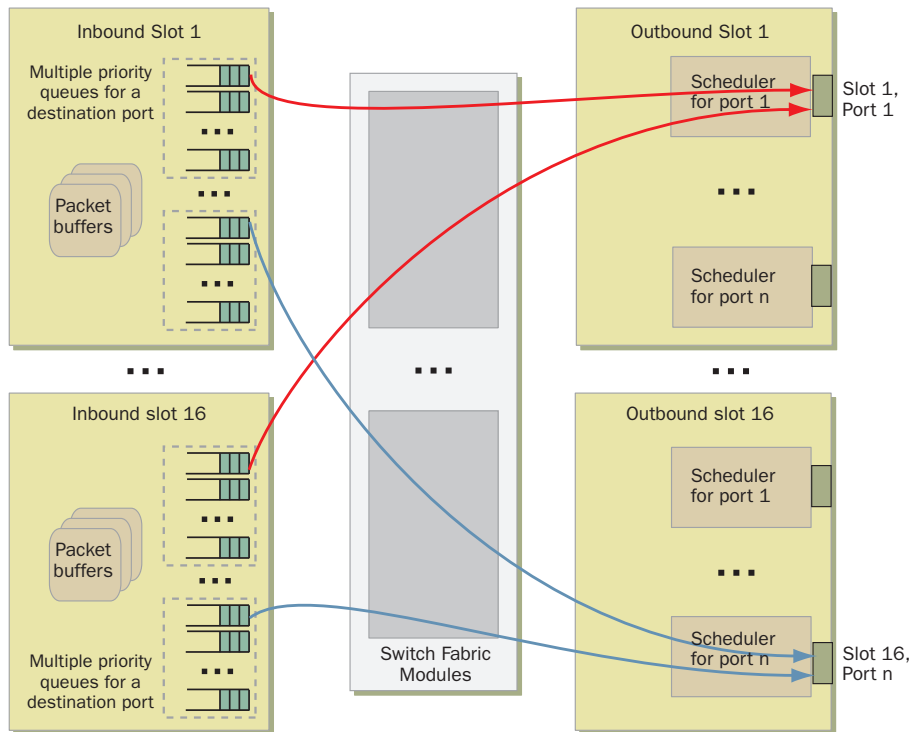
Figure 3.
Distributed queuing scheme on the Brocade MLX Series.

A unique characteristic of the Brocade MLX Series is the use of a distributed queuing scheme that maximizes the utilization of buffers across the entire system during congestion. This scheme marries the benefits of input-side buffering (virtual output queuing) with those of an output-port-driven scheduling mechanism. Input queuing using virtual output queues ensures that bursty traffic from one port does not use too many buffers on an output port. An output-port-driven scheduling scheme ensures that packets are sent to the output port only when the port is ready to transmit a packet.

Each interface module maintains multiple distinct priority queues to every output port on the system. Packets are "pulled" by the outbound interface module when the output port is ready to send a packet. Switch fabric messaging is used to ensure that there is tight coupling between the two stages. This closed-loop feedback between the input and output stages ensures that no information is lost between the two stages. The use of virtual output queues maximizes the efficiency of the system by storing packets on the input module until the output port is ready to transmit the packet. In all, there are up to one million virtual output queues on the Brocade MLX chassis.

Congestion avoidance is handled by applying a Weighted Random Early Detection (WRED) or tail-drop policy. On the output ports, a variety of scheduling mechanisms such as strict priority, weighted fair queuing, or a combination of these approaches can be applied to deliver tiered QoS guarantees for several applications.

The QoS subsystem on the Brocade MLX Series has extensive classification and packet marking capabilities that can be configured as follows:

- Prioritization based on Layer 2 (IEEE 802.1p), Type of Service (ToS), Differentiated Services Code Point (DSCP), or MPLS Experimental bits (MPLS EXP) of an input packet

- Mapping of packet and frame priority from ingress encapsulation to egress encapsulation

- Re-marking of the priority of a packet based on the result of the 2-rate, 3-color policer

HIERARCHICAL QOS

Hierarchical QoS (HQoS) is supported on the egress of 10 GbE ports of the Brocade NI-MLX-10GX8-M and BR-MLX-10GX8-X modules. HQoS allows a carrier to consolidate different services on the same physical device running on the same physical infrastructure. HQoS is a valuable tool, especially for networks that support multiple business customers who are running multiple applications with different prioritization and scheduling requirements, over the same infrastructure. HQoS uses an advanced scheduling mechanism, with multiple levels and instances of queuing, shaping, policing, and marking, to prioritize different services over the same connection and to combine all services into an overall SLA.

HQoS organizes a scheduler policy into a hierarchical tree that consists of a root node, branch nodes, and leaf nodes, where:

- The root node is the convergence point for all traffic and corresponds to a scheduler.

- A branch node is located in the middle of the hierarchy and corresponds to a scheduler.

- A leaf node is at the bottom layer and corresponds to a scheduling queue.

A scheduler can schedule multiple scheduling queues or schedulers. HQoS scheduling levels do not need to support any packet field matching capabilities. Packets are inspected only once before being queued. Once packets go into a queue, everything beyond that point is a sequence of rate shapers and schedulers that were previously configured for the ingress queues. HQoS supports a number of scheduling and shaping levels. Each level performs scheduling and shaping functions. The Brocade implementation of HQoS is one for which any level can be bypassed (partially or completely). By configuring actions oriented to different levels (customer level, service level, and logical port, for example) hierarchical traffic management can be achieved.

Traffic Policers and ACLs

All interface modules support a large number of inbound as well as outbound traffic policers in hardware. Up to 256,000 traffic policers can be concurrently configured in the system. The 2-rate, 3-color policers meter subscriber flows by classifying them as Committed Information Rate (CIR) or Excess Information Rate (EIR). This capability is especially useful when mixing traffic flows with different characteristics on the same port.

For security purposes, input ACLs as well as output ACLs are supported by the system on every interface module. Up to 224,000 input ACL entries and 128,000 output ACL entries for ACL rules can be applied to local interfaces on every interface module.

DoS Guards

Layer 2 services such as Virtual Private LAN Services (VPLS) require support for efficient replication of packets to the entire broadcast domain. For example, traditional architectures handle Ethernet frames with unknown MAC addresses by sending them to a processor to replicate the packet to the broadcast domain. The involvement of the CPU makes the system vulnerable to a potential Denial of Service (DoS) attack. In contrast, the Brocade MLX Series handles this scenario very efficiently by performing the flooding in hardware.

The Brocade MLX Series has a dedicated out-of-band management link between each interface module and the management module, to isolate control traffic from data traffic. Multiple queues to the management module allow different types of control traffic to be prioritized. These capabilities, together with secure management and ACLs, are immensely useful in protecting the system from potential DoS attacks in the network.
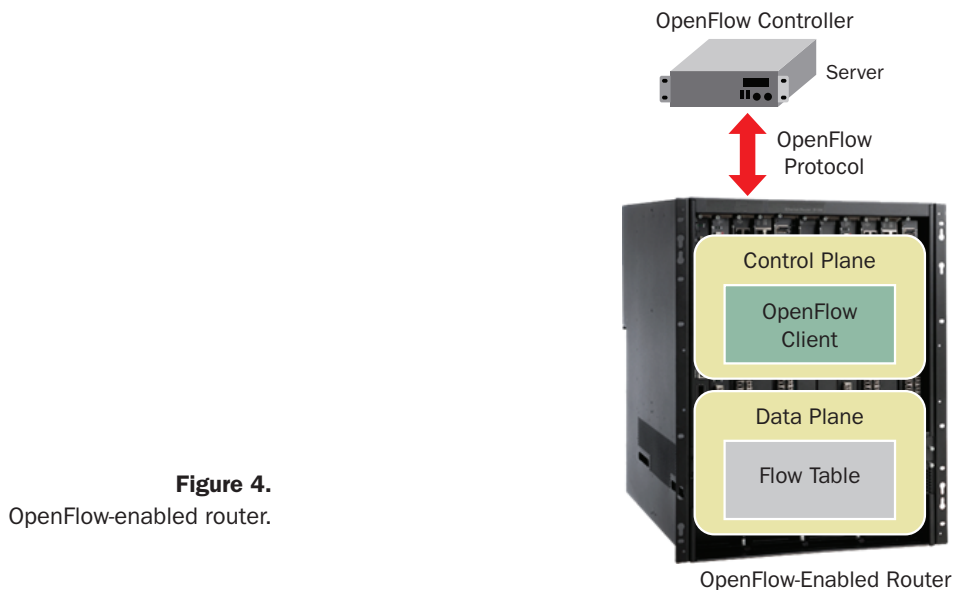
Spatial Multicast Support

The Brocade MLX architecture has native support for spatial multicast, a critical requirement for offering video services in a network. The input interface module sends one copy of an incoming multicast packet to the switch fabric. The switch fabric then replicates the packet within itself to multiple output interface modules in the system, which in turn replicate the multicast packet to the destination ports.

OPENFLOW SUPPORT

Software-Defined Networking (SDN) is a powerful new network paradigm designed for the world's most demanding networking environments. The Brocade MLX Series enables SDN by supporting the OpenFlow protocol, which allows communication between an OpenFlow controller and an OpenFlow-enabled router. Using this approach, organizations can control their networks programmatically, transforming the network into a platform for innovation through new network applications and services.

The OpenFlow-enabled Brocade MLX Series supports an OpenFlow client (control plane software), which communicates with an OpenFlow controller using the OpenFlow protocol. The OpenFlow controller runs on a server or a server cluster. OpenFlow-enabled routers such as the Brocade MLX Series support the abstraction of a flow table, which is manipulated by the OpenFlow controller. The flow table contains flow entries. Each flow entry represents a flow (packets with a given destination IP address). The flow table is sorted by flow priority, which is defined by the controller. The highest priority flows are at the top of the flow table.

Incoming packets on an OpenFlow-enabled port are matched (in order of priority) against the flow entries defined for that port by the OpenFlow controller. If the packet matches a given flow entry, the flow matching process stops, and the set of actions defined for that flow entry are performed. Packets that do not match any flow entry are dropped by default. The Brocade implementation supports an option to send such packets to the controller.

Figure 4.
OpenFlow-enabled router.

OpenFlow Controller

Server

OpenFlow Protocol

Control Plane

OpenFlow Client

Data Plane

Flow Table

OpenFlow-Enabled Router

The Brocade MLX Series supports enabling OpenFlow Hybrid Port Mode. With Brocade Hybrid Port Mode, organizations can simultaneously deploy traditional Layer 2/3 MPLS forwarding with OpenFlow on the same port. This unique capability provides a pragmatic path to SDN by enabling network operators to integrate OpenFlow into existing networks, giving them the programmatic control offered by SDN for specific flows, while the remaining traffic is routed as before.

INDUSTRY-LEADING MULTISERVICE FEATURE SET

Brocade Network Advisor software leverages the cumulative experience that Brocade has gained in powering networks over the last decade. The software complements the Brocade MLX architecture to offer the following capabilities:

- Support for Border Gateway Protocol version 4 (BGP4), Open Shortest Path First (OSPF), Intermediate System-to-Intermediate System (IS-IS), and Routing Information Protocol (RIP) in IPv4 networks

- Support for IPv6, including Multiprotocol Border Gateway Protocol version 4 (MP-BGP4), OSPFv3, IS-IS, and RIP next generation (RIPng) routing protocols

- Support for MPLS, including signaling protocols such as Resource Reservation Protocol-Traffic Engineering (RSVP-TE) and Label Distribution Protocol (LDP)

- Extensive traffic engineering support for MPLS

- MPLS Fast Reroute support

- Layer 2 Virtual Private Network (VPN) using VPLS or Virtual Leased Line (VLL)

- Layer 3 VPN using RFC 2547bis or Multi-Virtual Routing and Forwarding (Multi-VRF)

- Internet Group Management Protocol (IGMP), Multicast Listener Discovery (MLD), Protocol Independent Multicast Sparse Mode (PIM-SM) and Dense Mode (PIM-DM), PIM Source-Specific Multicast (PIM-SSM), and Distance Vector Multicast Routing Protocol (DVMRP) support to power IPv4 and IPv6 multicast applications

- Hitless Layer 2 and Layer 3 failover with support for hitless software upgrades

- Layer 3 redundancy protocols such as Virtual Router Redundancy Protocol (VRRP) and Virtual Router Redundancy Protocol-Extended (VRRP-E)

- Layer 2 redundancy protocols such as Virtual Switch Redundancy Protocol (VSRP)

- Support for MAC layer service protection protocols such as Metro Ring Protocol (MRP) and Rapid Spanning Tree Protocol (RSTP)

- Support for secure management via Secure Shell (SSH) version 1 and version 2, Secure Copy Protocol (SCP) version 1 and version 2, or Simple Network Management Protocol (SNMP) version 3

- sFlow-based Layer 2 to Layer 7 traffic monitoring of activity on the node with underlying hardware support for reliable packet sampling

In contrast to some systems that limit the capabilities that can be concurrently enabled, the Brocade MLX architecture allows both Layer 2 and Layer 3 services to be offered on the same device and the same port concurrently. This ability gives unprecedented flexibility in tailoring the system to meet the needs of the end user.

Scalability

The Brocade MLX Series of routers is a highly scalable family of routers. These are some examples of its industry-leading scalability:

- Up to 16,000 VPLS instances and up to one million VPLS MAC addresses

- 48,000 VLLs per system

- Support for 4,094 VLANs and up to two million MAC addresses

- One million IPv4 routes in hardware Forwarding Information Base (FIB)

- 240,000 IPv6 routes in hardware FIB

- 10 million BGP routes

- 2,000 BGP/MPLS VPNs and up to one million VPN routes

Investment Protection

The Brocade MLX chassis uses a half-slot design for interface modules. The divider between two adjacent half slots can be removed to combine them into a full slot to meet future needs. All chassis have 480 Gbps of full-duplex bandwidth per full slot. In addition, with the ability to offer multiple services, including dual-stack IPv4 and IPv6 as well as MPLS services in hardware, the Brocade MLX Series offers excellent investment protection.

CONCLUSION

The Brocade MLX Series is the most advanced service delivery platform in the industry, offering IPv4, IPv6, MPLS, and Multi-VRF services. Its robust, scalable architecture coupled with a rich feature set in multiservice Brocade IronWare OS software makes the Brocade MLX Series the leading router in its class. An industry-leading density of Ethernet ports in a single rack makes the Brocade MLX Series an excellent investment for businesses that are planning to build a converged multiservice network for the future.

ABOUT BROCADE

Brocade networking solutions help the world's leading organizations transition smoothly to a world where applications and information reside anywhere. This vision is realized through the Brocade One™ strategy, which is designed to deliver key business benefits such as unmatched simplicity, non-stop networking, application optimization, and investment protection.

Innovative Ethernet and storage networking solutions for data center, campus, and service provider networks help reduce complexity and cost while enabling virtualization and cloud computing to increase business agility.

To help ensure a complete solution, Brocade partners with world-class IT companies and provides comprehensive education, support, and professional services offerings.

To learn more, visit www.brocade.com

Corporate Headquarters
San Jose, CA USA
T: +1-408-333-8000
info@brocade.com

European Headquarters
Geneva, Switzerland
T: +41-22-799-56-40
emea-info@brocade.com

Asia Pacific Headquarters
Singapore
T: +65-6538-4700
apac-info@brocade.com

BROCADE