Solution Brief

Building a Multitenant, Scalable, Resilient Storage Solution Using Cisco MDS 9000 Series Switches and EMC XtremIO Flash Arrays

Today's storage administrator has many choices when choosing storage area network (SAN) and storage array solutions. Decisions on current and future solution scale, cost, performance, manageability, security, and multitenancy must be carefully weighed before a solution is chosen. Cisco® MDS 9000 Series 16-Gbps (16G) Fibre Channel switches and directors, along with EMC XtremIO storage arrays, are designed to address these requirements and complement each other to provide industry-leading solutions for today's data centers.

Every day the amount of data processed by hosts, passed onto the wire, and stored and backed up inside the storage array is increasing multifold. This growth requires high-performance storage arrays and storage networking switches with key characteristics such as reliability, stability, and high availability, as well as the flexibility for future scale requirements. The Cisco MDS 9000 Series 16G Fibre Channel switches and EMC XtremeIO storage array provide all of these features to address current and future requirments while meeting the above needs and also providing investment protection.

Cisco MDS 9000 Series 16G Fibre Channel Switches

The Cisco MDS 9000 Series of 16G Fibre Channel switches is made up of the following switches:

- Cisco MDS 9710 Multilayer Director
- Cisco MDS 9706 Multilayer Director
- Cisco MDS 9250i Multiservice Fabric Switch
- Cisco MDS 9396S 16G Multilayer Fabric Switch
- Cisco MDS 9148S 16G Multilayer Fabric Switch

The MDS 9710 and 9706 multilayer director-class switches are designed to meet the critical requirements of large, virtualized data center storage environments. Some of these requirements include high availability, performance, reliability, and resiliency. The 9710 and 9706 provide line-rate non-oversubscribed 16-Gbps Fibre Channel throughput with up to 1.5 Tbps of bidirectional front-panel bandwidth per slot.

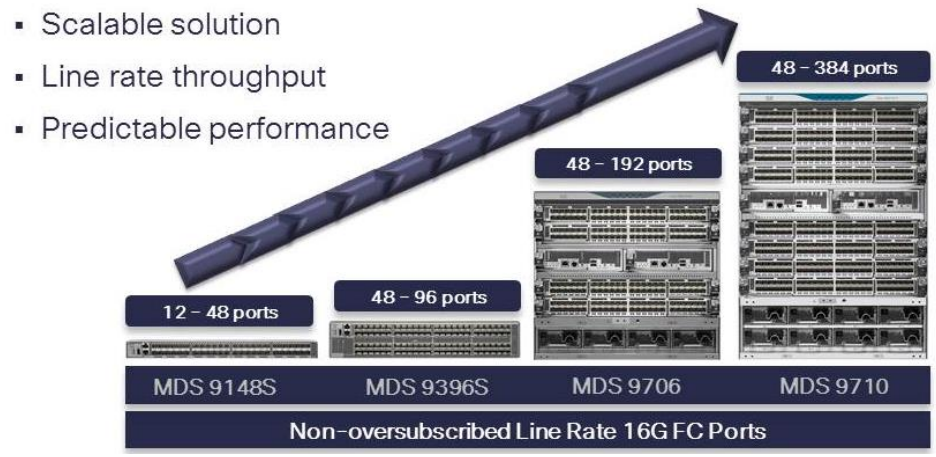
The MDS 9250i multiservice fabric switch provides forty 16-Gbps Fibre Channel ports, eight 10 Gigabit Ethernet Fibre Channel over Ethernet (FCoE) ports, and two 1/10 Gigabit Ethernet IP storage service ports in a 2-rack-unit (2RU) form factor. This switch provides multiprotocol connectivity for the Fibre Channel, FCoE, Fibre Channel over IP (FCIP), and SCSI over IP (iSCSI) protocols. FCIP can be used for data replication across multiple data centers over long-distance IP networks during disaster recovery.

The MDS 9396S is a 16G line-rate high-performance, high-port-density fabric switch in a compact 2RU form factor. This switch has 96 16-Gbps Fibre Channel ports with built-in redundant hardware and large buffer-to-buffer credits for inter-data center connectivity. The base switch model supports 48 enabled ports at 2-, 4-, 8-, 10- and 16-Gbps Fibre Channel speeds, with up to 48 additional ports available with expansion licenses.

The MDS 9148S is a low-cost, high-performance 16G line-rate Fibre Channel fabric switch. It has 48 16-Gbps Fibre Channel line-rate ports, 12 of which are enabled with the base license. The incremental switch license enables licensing of all 48 ports in 12-port increments.

Overall, the MDS 16G switches help redefine the data center with rich key parameters to meet the growth and scale of next-generation data centers across the storage industry (Figure 1).

Figure 1. Cisco MDS Scale-Out Architecture



Scalable Performance

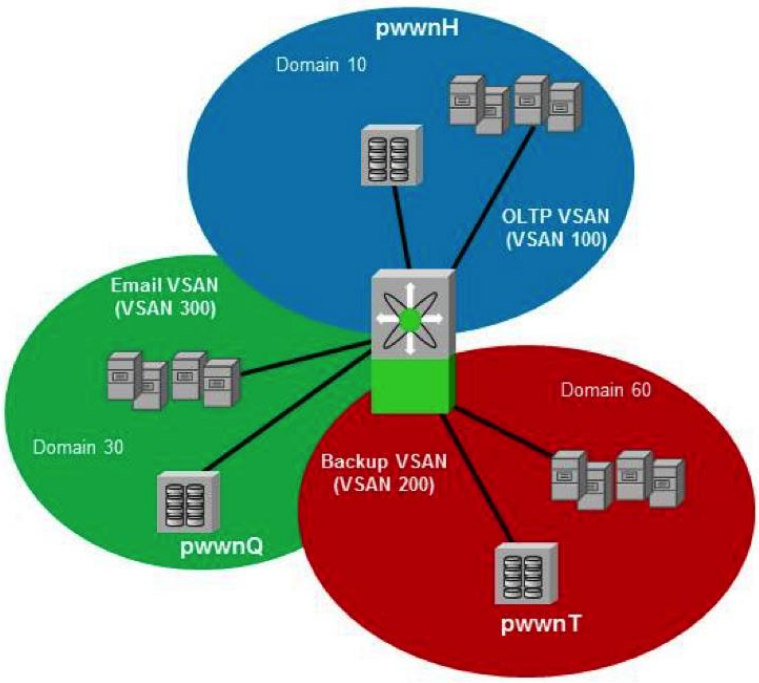
The combination of 16-Gbps Fibre Channel switching modules and fabric switches, along with 10-Gbps FCoE capability, provides the high performance throughput in the SAN industry, with more flexibility and reliability. Based on crossbar fabric architecture with central arbitration, the Cisco MDS architecture provides 16-Gbps line-rate, non-blocking, predictable performance across all traffic conditions for every port in all the chassis. The director-class switches can provide up to 192 16-Gbps Fibre Channel or 384 10-Gbps FCoE ports, whereas fabric switches can support from 12 to 96 ports with 16-Gbps Fibre Channel line-rate performance.

Multitenancy

The Cisco MDS 9000 Series platform supports virtual storage area network (VSAN) technology to provide multitenancy support. Cisco introduced the VSAN concept in 2002 to provide higher scalability and security for deploying multiple virtual fabrics across a single physical infrastructure. Each VSAN has independent and isolated fabric-related configurations such as logins, aliases, zoning, and other Fibre Channel services within its boundaries. Using the Cisco VSAN feature, network architects can build multifabric topologies across a single physical switch or network (Figure 2). Cisco VSANs can also help extend scalability, as they allow the much-needed elasticity and flexibility to allocate or relocate resources as tenant requirements change. Each VSAN can span across multiple switches in a physical network, with independent address space that allows the use of identical Fibre Channel IDs simultaneously in different VSANs. Multiple VSANs can extend across multiple

switches using VSAN trunking. With the release of NX-OS 6.2(13), the maximum limit of VSAN per physical fabric has increased to 80.

Figure 2. Multitenancy using Cisco VSAN technology



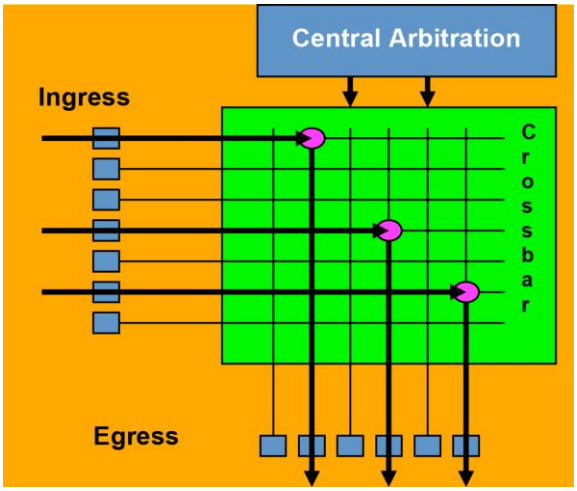
Some of the common classes of criteria that can be applied to separate the traffic among the same or multiple tenants using VSANs are:

- **Storage based:** Tenants can be allocated with a certain amount of storage as needed, with scale to grow using pay-as-you-go flexibility.
- **Performance (bandwidth) based:** Some tenants require high bandwidth throughput and a certain amount of allocated bandwidth with a defined oversubscription ratio. Port channels can be used to allow the higher amount of throughput and bandwidth allocation between switches.
- **Application based:** For latency-sensitive applications, high throughput can be allocated based on VSAN with a low oversubscription ratio. Quality of service (QoS) and class of service (CoS) can be used to differentiate the high-priority traffic.
- **Ad hoc requirements:** Tenants may require temporary storage, bandwidth, and throughput for their test environment during predeployment or testing.
- **Security based:** Tenants can be created based on different classes of departmental security requirements, such as port security, fabric binding, Dynamic Port VSAN Membership (DPVM), etc.
- **Traffic based:** VSANs can help create isolated tenant environments based on different types of traffic – FCIP, FCoE, Fibre Channel, etc.

High Availability and Data Protection

Cisco MDS 9000 Series switches are designed with hardware components that provide redundancy without affecting performance, throughput, or bandwidth. All MDS 9700 Series directors have redundant supervisor modules, multiple fans in a multiple-fan tray, optional N+1 and N:N fabric redundancy, and power supply grid redundancy. MDS fabric switches have grid redundant power supplies and fan trays for resiliency. The Cisco MDS 9000 Series internal architecture includes features such as central arbitration, virtual output queueing (VoQ), built-in cyclic redundancy checks (CRC), and hardware-based slow drain.

Figure 3. Internal architecture of Cisco MDS switches



- The Cisco MDS 16G Fibre Channel switches have built-in data protection provided by CRC technology to detect and drop corrupt packets. Every frame entering the switch receives a CRC before it can be processed. Similarly, every egress frame is examined for integrity before leaving the switch. This mechanism helps make sure that no corrupt frames enter or exit the switch that could affect the performance of the SAN fabric.
- Arbitration: MDS switches have central arbitration technology that helps enable deterministic performance across the switch. For every ingress frame on any port, the central arbiters will allocate the best path for the egress frame on the fabric module of the switch (Figure 3). Central arbiters are capable of scheduling more than a billion frames per second.
- VOQ helps avoid head-of-line blocking to provide line-rate performance on each port. It helps ensure fairness between ports regardless of the port locations within the switch. This helps ensure that if one port is busy, it will not result in blocking of traffic to other ports (head-of-line blocking).
- QoS: Cisco MDS switches can use QoS to provide multitenancy configuration within a SAN. Using QoS, the SAN administrator can assign different priorities and different weights to prioritize the multitenant traffic based on bandwidth, latency, criticality, performance, etc.
- Port channels: Port channels allow users to aggregate up to 16 physical inter-switch links (ISLs) into a single logical bundle to provide redundancy and performance optimization across multiple links between switches. A port channel member can be defined from different line card modules, different application-specific integrated circuits (ASICs), and different port groups on the same switch to provide high availability

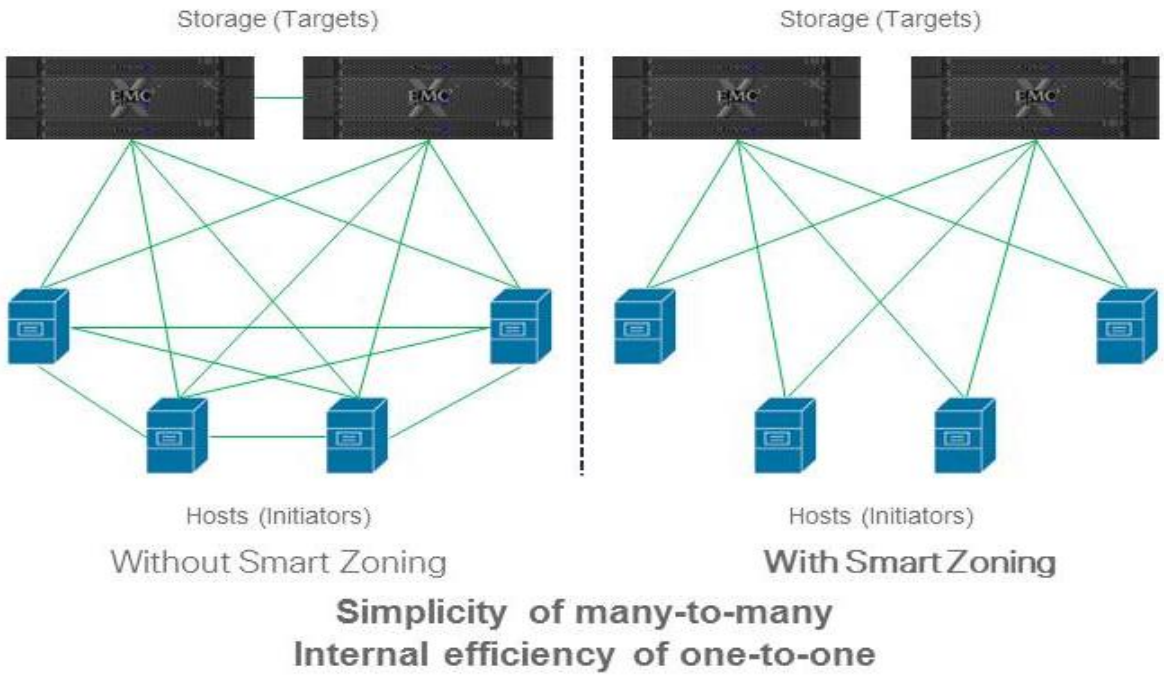
for ISLs. Port channels can load-balance traffic based on source ID, destination ID, and exchange ID. If a link goes down in a port channel, traffic continues on different links without affecting routing tables.

The Cisco MDS 9000 Series also provides non-disruptive software upgrades and downgrades using in-service software upgrade (ISSU) and in-service software downgrade (ISSD) as well as stateful process restart and failover for best-in-class high availability across the portfolio to reduce the impact of upgrading or downgrading the software code on a switch.

Simplified Management

Smart Zoning is a Cisco MDS NX-OS feature that helps with zone management. It allows the simplicity and convenience of many-to-many zones while providing the internal efficiency of one-to-one initiator and target communication. NX-OS provides a built-in command-line interface (CLI) for easy migration to Smart Zoning. Smart Zoning is beneficial in virtualized environments where many initiators access the same targets (Figure 4).

Figure 4. Simplified zone management using Smart Zoning



Device management: Cisco MDS 9000 Series switches provide three different modes of management: the CLI, the web-based Cisco Prime™ Data Center Network Manager, and integration with third-party storage management tools.

The Cisco CLI is consistent across data center Cisco Nexus® and MDS switches. One can manage the SAN using MDS switches and the LAN using Cisco Nexus switches simultaneously. Having a common NX-OS across all the data center switches means using similar commands to deploy, install, configure, troubleshoot, debug, and maintain. This helps optimize the operational expenses involved in managing a dynamic environment consisting of converged infrastructure with multiprotocol connectivity (Fibre Channel, FCoE, iSCSI, IP, IBM Fiber Connection (FICON), and FCIP).

Cisco Prime Data Center Network Manager is the web-based manager for data center switches. It is the industry's first converged SAN and LAN management solution. Data Center Network Manager can manage all Cisco NX-OS devices, including the Cisco MDS 9000 Series and Cisco Nexus family products. The intuitive GUI helps simplify everyday operations of Cisco data center fabrics efficiently.

Cisco Prime Data Center Network Manager offers multiple capabilities across LAN and SAN switches, including:

- Monitoring of events and performance, historically and at scale
- Wizard- and template-based provisioning of NX-OS technologies and services
- Cisco VMpath analytics, with dynamic topology views for extended visibility into virtual infrastructure
- Resource management through trend analysis of inventory and performance
- Rule-based event notification and filtering
- Role-based access control to provide separation between network and storage teams

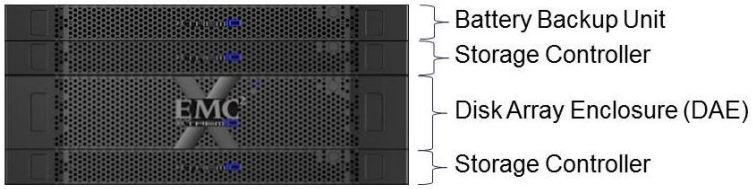
Cisco Prime Data Center Network Manager supports federation of up to eight management servers, to manage up to 120,000 devices using a single management pane. The base management function is available at no charge, and advanced features are unlocked with a license. The application can be installed on Linux and Microsoft Windows operating systems and supports both PostgreSQL and Oracle databases. This application also supports visibility into third party storage disk arrays.

The MDS 9000 Series fabric switches also include Fabric Manager, a built-in management web UI, to support all features available through the CLI or Data Center Network Manager.

EMC XtremIO Product Family

The EMC XtremIO All-Flash Array is a clean-sheet design with a revolutionary architecture. It brings together all the requirements necessary to enable the agile data center: linear scale-out, inline all-the-time data services, and rich data center services for the workloads.

Figure 5. EMC XtremeIO All-Flash Array



The basic hardware building block for these scale-out arrays is the X-Brick (Figure 5). Each X-Brick has two active-active controller nodes and a disk array enclosure packaged together with no single point of failure. The Starter X-Brick with 13 SSDs can be non-disruptively expanded to a full X-Brick with 25 SSDs without any downtime. The scale-out cluster can support up to eight X-Bricks.

The XtremIO platform is designed to maximize the use of flash storage media. Key attributes of this platform are:

- Incredibly high levels of I/O performance, particularly for random I/O workloads that are typical in virtualized environments
- Consistently low (sub-millisecond) latency
- True inline data reduction—the ability to remove redundant information in the data path and write only unique data on the storage array, thus lowering the amount of capacity required
- A full suite of enterprise array capabilities, such as integration with VMware through vStorage APIs for Array Integration (VAAI), N-way active controllers, high availability, strong data protection, and thin provisioning

As the XtremIO array has a scale-out design, additional performance and capacity can be added in a building block approach, with all building blocks forming a single clustered system. XtremIO storage includes the following components:

- **Host adapter ports:** Provide host connectivity through fabric into the array.
- **Storage controllers:** The compute component of the storage array. Storage controllers handle all aspects of data moving into, out of, and between arrays.
- **Disk drives:** Solid-state drives (SSDs) that contain the host/application data and their enclosures.
- **InfiniBand switches:** A computer network communications link used in multi-X-Brick configurations that is switched, high throughput, low latency, scalable, and QoS and failover capable.

XtremIO Operating System (XIOS)

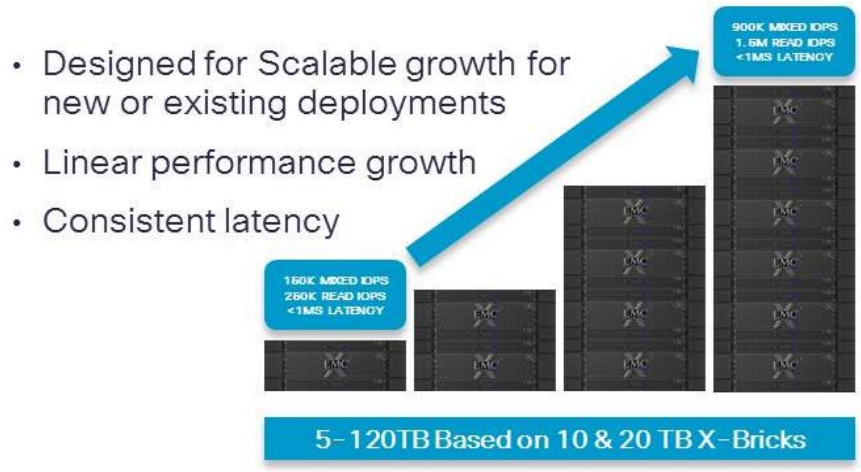
The XtremIO storage cluster is managed by the XtremIO Operating System (XIOS), XtremIO's powerful operating system. XIOS helps ensure that the system remains balanced and always delivers the highest levels of performance without any administrator intervention, as follows:

- Helps ensure that all SSDs in the system are evenly loaded, providing both the highest possible performance as well as endurance that stands up to demanding workloads for the entire life of the array.
- Eliminates the need to perform the complex configuration steps found on traditional arrays. There is no need to set RAID levels, determine drive group sizes, set stripe widths, set caching policies, build aggregates, or do any other such configuration.
- Automatically and optimally configures every volume at all times. I/O performance on existing volumes and data sets automatically increases with large cluster sizes. Every volume is capable of receiving the full performance potential of the entire XtremIO system.

Standards-Based Enterprise Storage System

The XtremIO system interfaces with vSphere hosts using standard Fibre Channel and iSCSI block interfaces. The system supports complete high-availability features, including support for native VMware multipath I/O, protection against failed SSDs, non-disruptive software and firmware upgrades, no single point of failure, and hot-swappable components.

Figure 6. XtremeIO Scale out Architecture



Real-Time, Inline Data Reduction

The XtremIO storage system deduplicates and compresses incoming data in real time, allowing a massive number of virtual machines as well as application data to reside in a small and economical amount of flash capacity. Furthermore, data reduction on the XtremIO array does not adversely affect I/O operations per second (IOPS) or latency performance; rather, it enhances the performance of the virtualized environment.

Scale-Out Design

The X-Brick is the fundamental building block of a scaled-out XtremIO clustered system. Using a Starter X-Brick, virtual server deployments can start small and grow to nearly any scale required by upgrading the Starter X-Brick to an X-Brick and then configuring a larger XtremIO cluster if required. The system expands capacity and performance linearly as building blocks are added, making the virtualized environments extremely simple to size and manage as demands grow over time (Figure 6).

API Integration

The XtremIO array is fully integrated with vSphere through VAAI. All API commands are supported, including ATS, Clone Blocks/Full Copy/XCOPY, Zero Blocks/Write Same, Thin Provisioning, and Block Delete. This integration, in combination with the array's inline data reduction and in-memory metadata management, enables nearly instantaneous virtual machine provisioning and cloning and makes it possible to use large volume sizes for management simplicity.

Massive Performance

The XtremIO array is designed to handle very high, sustained levels of small, random, mixed read and write I/O, as is typical in virtual environments, and to do so consistently with extraordinarily low latency.

Fast Provisioning

XtremIO arrays deliver the industry's first writable snapshot technology that is space-efficient for both data and metadata. XtremIO snapshots are free from limitations of performance, features, topology, or capacity reservations. With their unique in-memory metadata architecture, XtremIO arrays can instantly clone virtual machine environments of any size.

Ease of Use

The XtremIO storage system requires only a few basic setup steps that can be completed in minutes with no tuning or ongoing administration in order to achieve and maintain high performance levels. In fact, the XtremIO system can be deployment ready less than an hour after delivery.

Security with Data at Rest Encryption

XtremIO arrays securely encrypt all data stored on the all-flash array, delivering protection for regulated use cases in sensitive industries such as healthcare, finance, and government.

Conclusion

EMC XtremIO's exceptional performance, capacity savings from unique data reduction capabilities, linear predictive scaling for scale-out architecture, and ease of use leads to breakthrough total cost of ownership in virtualized workload environments. Combining this performance with Cisco MDS 9000 Series switches provides industry-leading technology and feature integration to provide a robust, scalable, reliable, highly available, multitenant storage architecture for current and next-generation data centers.

For more information, please visit http://www.cisco.com/go/mds and http://www.emc.com/storage/xtremio/overview.htm.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C22-736486-00 1/16