

XtremIO Management Student Guide

Education Services
March 2015



XtremIO Management

The EMC² logo, consisting of the text 'EMC²' in white, sans-serif font on a blue square background.

© Copyright 2015 EMC Corporation. All rights reserved.

Support Contact: [Education Services](#)

1

Welcome to XtremIO Management version 3.0.2.

Copyright ©2015 EMC Corporation. All Rights Reserved. Published in the USA. EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license. The trademarks, logos, and service marks (collectively "Trademarks") appearing in this publication are the property of EMC Corporation and other parties. Nothing contained in this publication should be construed as granting any license or right to use any Trademark without the prior written permission of the party that owns the Trademark.

Acartus, Access Logix, AdvantEdge, AlphaStor, ApplicationXtender, ArchiveXtender, Atmos, Authentica, Authentic Problems, Automated Resource Manager, AutoStart, AutoSwap, AVALONidm, Avamar, Captiva, Catalog Solution, C-Clip, Celerra, Celerra Replicator, Centera, CenterStage, CentraStar, ClaimPack, ClaimsEditor, CLARiiON, ClientPak, Codebook Correlation Technology, Common Information Model, Configuration Intelligence, Configuresoft, Connectrix, CopyCross, CopyPoint, Dantz, DatabaseXtender, Data Domain, Direct Matrix Architecture, DiskXtender, DiskXtender 2000, Document Sciences, Documentum, eInput, E-Lab, EmailXaminer, EmailXtender, EMC2, EMC, EMC Centera, EMC ControlCenter, EMC LifeLine, EMC OnCourse, EMC Proven, EMC Snap, EMC SourceOne, EMC Storage Administrator, Enginuity, eRoom, Event Explorer, FarPoint, FirstPass, FLARE, FormWare, Geosynchrony, Global File Virtualization, Graphic Visualization, Greenplum, HighRoad, HomeBase, InfoMover, Infoscapes, Infra, InputAccel, InputAccel Express, Invista, Ionix, ISIS, Max Retriever, MediaStor, MirrorView, Navisphere, NetWorker, nLayers, OnAlert, OpenScale, PixTools, Powerlink, PowerPath, PowerSnap, QuickScan, Rainfinity, RepliCare, RepliStor, ResourcePak, Retrospect, RSA, the RSA logo, SafeLine, SAN Advisor, SAN Copy, SAN Manager, Smarts, SnapImage, SnapSure, SnapView, SRDF, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix, Symmetrix DMX, Symmetrix VMAX, TimeFinder, UltraFlex, UltraPoint, UltraScale, Unisphere, VMAX, Vblock, Viewlets, Virtual Matrix, Virtual Matrix Architecture, Virtual Provisioning, VisualSAN, VisualSRM, Voyence, VPLEX, VSAM-Assist, WebXtender, xPression, xPresso, YottaYotta

Revision Date: March 2015

Revision Number: MR-1CP-XIOMGMT

Course Overview

Description	This course provides an in-depth understanding of how to manage, monitor, and tune the performance of an XtremIO storage system. Managing the system with the XtremIO Management Server (XMS), provisioning storage to heterogeneous hosts, and configuring XtremIO features are covered in detail. Management with both the GUI and CLI is discussed. This course also covers XtremIO performance analysis, potential bottlenecks, performance monitoring tools and techniques, as well as practices to tune performance.
Audience	This course is intended for IT professionals involved in positioning, implementing, managing and supporting XtremIO storage systems.
Objectives	Upon completion of this course, you should be able to: • Manage the array using the GUI and the CLI • Provision XtremIO storage to hosts • Describe and manage Data at Rest Encryption, Deduplication, Compression, and Snapshot technology • Monitor system metrics, alerts, and performance data • Perform data collection and interpret performance data • Perform basic troubleshooting on XtremIO storage array



© Copyright 2015 EMC Corporation. All rights reserved.

Course Overview

2

This course provides an in-depth understanding of how to install, manage, monitor, and maintain a functional EMC XtremIO storage array. Key topics covered in the course are hardware and software component review, XMS operations, provisioning storage to heterogeneous hosts, Snapshots, monitoring and managing system performance, and basic troubleshooting.

Agenda

Module/Lessons	Labs
Day 1 • XtremIO Hardware and Software • XtremIO Product Overview • XtremIO Hardware Components • External Network Requirements • XtremIO Management Server (XMS) • XtremIO Management • XMS GUI Management • XMS CLI Management • XMS User Management • XtremIO Provisioning • Volume Creation and Sizing • Host Connectivity – VMware • Host Connectivity – MS Windows • Host Connectivity – Linux	• Getting started • XMS Management with the GUI • XMS Management with the CLI • XMS User Account Management • XtremIO Storage Provisioning • XtremIO VMware ESXi Host Configuration • XtremIO Virtual Machine Data store Workload Test

© Copyright 2015 EMC Corporation. All rights reserved.

Course Overview



3

This slide covers the agenda for the first day of this three-day course.

Agenda

Module/Lessons	Labs
Day 2 • XtremIO Provisioning (continued) • Host Connectivity - Solaris • Host Connectivity - AIX • XtremIO Features • Data at Rest Encryption • Compression • Deduplication • XtremIO Snapshots • XtremIO Monitoring with XMS • Monitoring Operation Overview • Monitoring for Performance • Performance Management Overview • Performance Management Overview • Architectural Constructs • Environment Tools	• XtremIO Virtual Machine RDM Workload Test • XtremIO Advanced Monitoring • XtremIO Snapshots



© Copyright 2015 EMC Corporation. All rights reserved.

Course Overview

4

This slide covers the agenda for the second day of this three-day course.

Agenda

Module/Lessons	Labs
Day 3 • Performance Management Overview (continued) • Host Tools • Benchmarking Tools • XtremIO Performance • Optimizing XtremIO Performance • Optimizing Connectivity • Optimizing Hosts • Benchmarking • XtremIO Troubleshooting and High Availability • High Availability • Support Resources and Updates	• XtremIO Performance Exercises

This slide covers the agenda for the third and final day of this course.

This slide intentionally left blank.



Module: XtremIO Overview

Upon completion of this module, you should be able to:

- Identify XtremIO product offerings
- Identify the hardware components of an XtremIO array
- Define network requirements for XtremIO
- Describe the XtremIO Management Server (XMS)



This module introduces the XtremIO hardware and XtremIO management requirements. Starting with an overall product review, the topics discussed include detailed hardware review, networking pre-requisites, and the XtremIO Management Server (XMS).

This slide intentionally left blank.



Lesson: XtremIO Hardware and Software

This lesson covers the following topics:

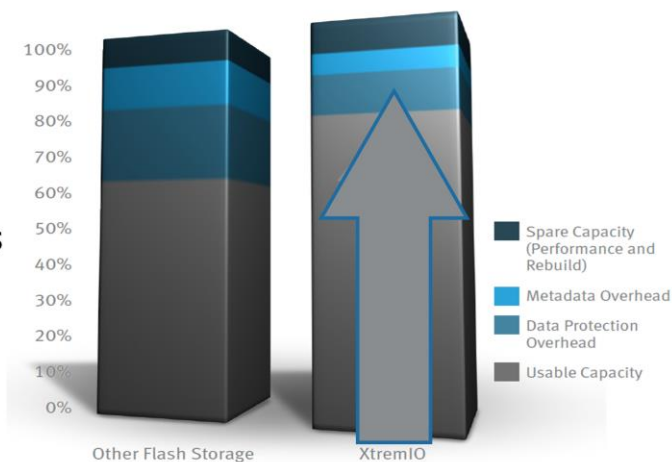
- XtremIO attributes and features overview
- Placement considerations
- Model types



This lesson covers the XtremIO attributes and features, proper product placement, and model types that are available.

Designed Specifically for All-Flash

- Scalable Flash-specific design
- More than just superior performance
- Rich array of data services and storage capabilities
- No system-wide garbage collection



EMC²

XtremIO Overview

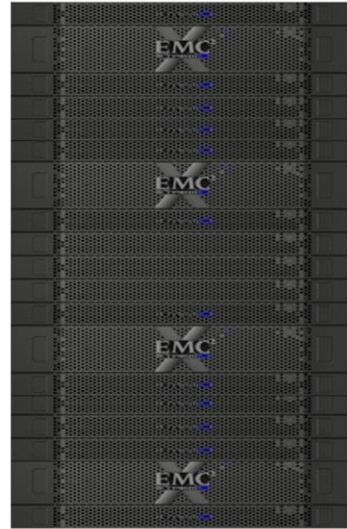
4

The fundamental difference between XtremIO and other all-Flash array architectures is that XtremIO has developed an entirely new array architecture that was built specifically to take advantage of [true] random access media.

Today, this random access media is Flash. XtremIO has a unique data protection schema that leverages the random access capability of Flash drives to offer consistently high performance at a reduced overhead level. It also protects the Flash endurance, does not require system-wide garbage collection, and its scalability allows it to extract the maximum throughput and performance from the Flash media.

XtremIO – More Than Just a Flash Array

- Superior performance
- Inline data reduction
 - Deduplication
 - Compression
- Data protection
- Thin provisioning
- Load balancing
- SSD wear leveling



EMC²

XtremIO Overview

5

Though the use of all Flash drives is a key attribute of an XtremIO system, it is not its sole defining factor.

The XtremIO approach is to use solid state disks, or Flash drives, and proprietary intelligent software to deliver unparalleled levels of performance. Achievable performance ranges from hundreds of thousands to millions of IOPS, and consistently low latency of under one millisecond for small I/Os. The system is also designed to provide minimal planning with a very simple user interface that makes provisioning and managing the array simple.

Some of the key XtremIO product features are:

- High performance
- Line speed deduplication and compression
- Advanced data protection
- Thin provisioning
- Load balancing
- SSD wear leveling

Feature Highlights

Performance

- Consistent and predictable
- Sub-millisecond latency
- Very high IOPs

Efficiency

- Inline deduplication and compression
- Thin provisioning
- XtremIO Data Protection (XDP)

Simplicity

- Easy to use user interface
- Minimal design and implementation efforts
- Simple volume provisioning

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

6

XtremIO is designed for high performance. Supplying consistent, predictable sub-millisecond latency with high IOPs is the foundation of XtremIO, but there are many other benefits as well.

The XtremIO optimized in-memory metadata and true inline data reduction efficiently store only unique 8 kB blocks of data, thus providing a huge savings on storage cost and avoiding writes of duplicate blocks to Flash, which extends eMLC drive endurance.

Since XtremIO is a fully thin provisioned storage system, even more storage savings are realized.

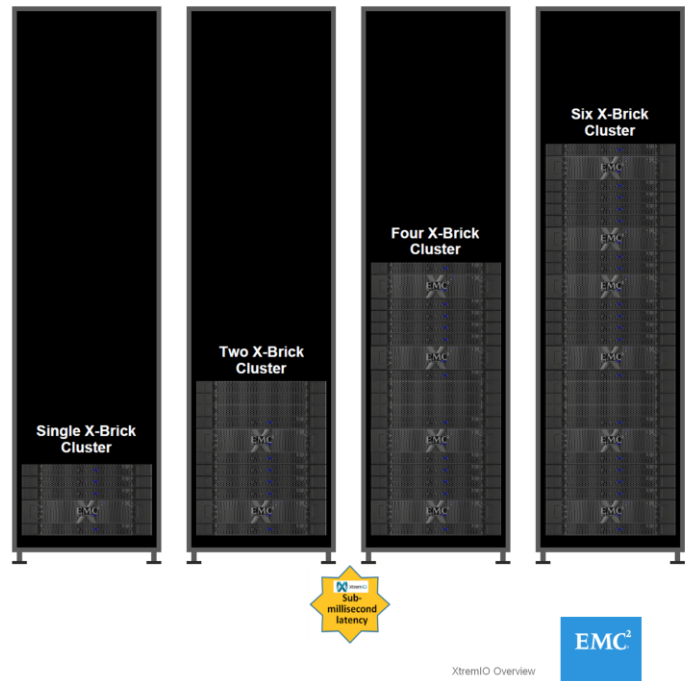
Data protection is of course crucial to every environment. XtremIO not only uses dual parity protection, but XtremIO's Data Protection (XDP) is faster than RAID1 and has the protection of RAID6 with the efficiency of RAID5.

Provisioning storage is a simple three-step process that does not require complicated RAID group configuration, cache sizing, or drive selection.

System Offerings

- 2 X-Brick Models
 - 10TB X-Brick **
 - 20TB X-Brick
- Similar functionality
- Similar hardware components
- Single, dual, quad, and hex clusters are supported [20 TB X-Brick only in hex clusters]

** A half populated starter 10 TB X-Brick is also available



XtremIO is currently offered in one, two, four, and six X-Brick cluster configurations. Each X-Brick cluster is configured as an Active/Active highly-available clustered storage array, without any single point-of-failure. It is important to note that I/Os and space provided by the X-Brick cluster scale simply by increasing the number of X-Bricks in the offering with no impact on latency.

Note that there are 2 X-Brick models, a 10 TB and 20 TB X-Brick, that can be used in multi-X-Brick clusters. The starter 10 TB X-Brick is a half populated [13 SSDs instead of 25] X-Brick which can be expanded to a 10 TB X-Brick, though physical capacity is slightly lower than a native 10 TB X-Brick.

Note that the 6 X-Brick cluster is only available with 20 TB X-Bricks.

XtremIO Environment Use Cases

- Virtual Desktop Infrastructure (VDI)
 - High number of I/O requests
 - Small random I/O
 - Low latency
 - High degree of duplicate data
- Server virtualization
 - Achieving higher virtualization rates
 - Mission-critical applications with demanding I/O load profiles
 - Management simplicity
 - Rapid provisioning of VMs and applications
- Database analytics and testing
 - High performance production databases
 - Support for test and development activities without performance penalties
 - Use of full clones (high duplication rate)
 - High small block random I/O activity
 - Requirement for very low and consistent response times



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

8

There are many environments, applications and solutions that benefit from the addition of an XtremIO storage array. These include Virtual Desktop Infrastructure (VDI), Server virtualization, and database analytics and testing. Some technical requirements for each topic are shown in the slide.

XtremIO Placement Considerations

- Match XtremIO capabilities to ensure implementations meet expectations
- XtremIO is not a storage array designed for all environments
- Best Return On Investment (ROI) might not be achieved:
 - If performance is not a core requirement
 - IOPs are not considered to be important
 - High latency is acceptable
 - If large amounts of storage are required
 - 100s of TB to multiple PB
 - If the data has a low duplication ratio



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

9

XtremIO is not a storage array designed for all environments. When you examine the implementation for XtremIO, there are solutions that do not extract the maximum benefit that XtremIO has to offer.

Solutions that require some or all of the items listed here may represent improper positioning of the XtremIO storage array:

- High latency is acceptable/tolerated
- Low IOPs are acceptable/tolerated
- Large capacity requirements
- Small amounts of duplicate data

Lesson Summary

During this lesson the following topics were covered:

- XtremIO attributes overview
- Placement considerations
- Model types



This lesson covered the XtremIO attributes, placement considerations, and model types that are available.

Lesson: XtremIO Hardware Components

This lesson covers the following topics:

- XtremIO hardware components



This lesson covers the XtremIO hardware review. During this lesson, we will perform a detailed review of the XtremIO components. This includes physical port description and usage as well as naming conventions.

XtremIO Hardware Components

- Storage Controller (SC)
- Disk array enclosure (DAE)
- Battery backup unit (BBU)
- InfiniBand switch (IBSW)
 - Only present in multi-X-Brick cluster configurations
- XtremIO Management Server (XMS)
 - Can be physical or virtual



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

12

The hardware components that comprise an XtremIO storage array are listed in this slide. The purpose is to expose you to all the components and then explain each of these components in detail.

Note: Not all components are required for all cluster configurations.

Storage Controller: CPU and Memory

- Two Intel Xeon CPUs
 - Eight cores each
- 256 GB RAM
- Runs XtremIO Operating System (XIOS)



EMC²

XtremIO Overview

13

The Storage Controller is an Intel ® server. It has two 8-core Intel ® Xeon ® CPUs, 256 GB RAM, and runs the XtremIO Operating System (XIOS). The Storage Controller needs 1U of rack space.

Terminology note:

Storage Controller (SC) and Node (N) refer to the same hardware component. The component name 'Node' might still be displayed in command outputs or older documentation. Storage Controller is the proper component name and should be used.

Storage Controller: Front Panel



After system initialization Reset and Power are disabled

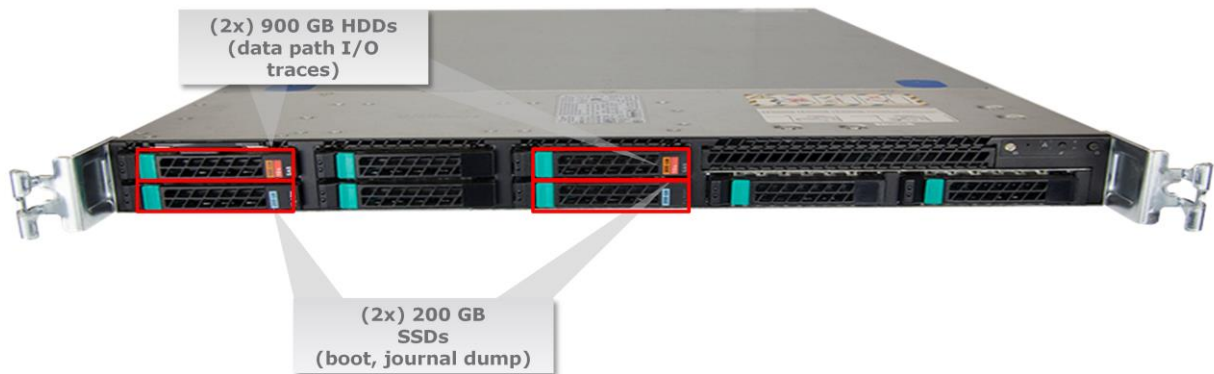
EMC²

XtremIO Overview

14

The front panel has indicator lights and buttons. It includes the Identification LED, NIC Activity LED, System Reset button, and Power button. Once the XtremIO storage array is initialized, the SC front panel Reset and Power buttons are disabled to prevent the user from pressing these buttons and accidentally powering off or resetting the SC. A user will never need to control the SC power through the physical buttons.

Storage Controller: Disk Drives



© Copyright 2015 EMC Corporation. All rights reserved.

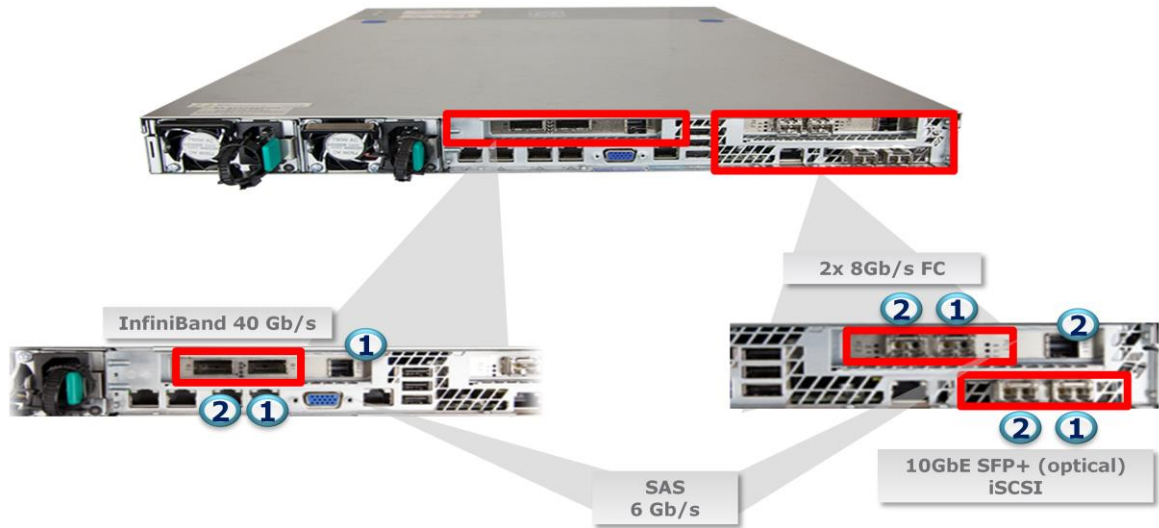
XtremIO Overview

15

In the SC, there are two solid state drives (SSDs), each with a capacity of 200 GB. One is used for the boot partition and the other for the journal dump in case of an emergency power down situation.

The two hard disk drives (HDDs) have a capacity of 900 GB each. They are used to store data path I/O traces and logs. Other disk slots have disk slot fillers only.

Storage Controller: Rear Panel (1 of 3)



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

16

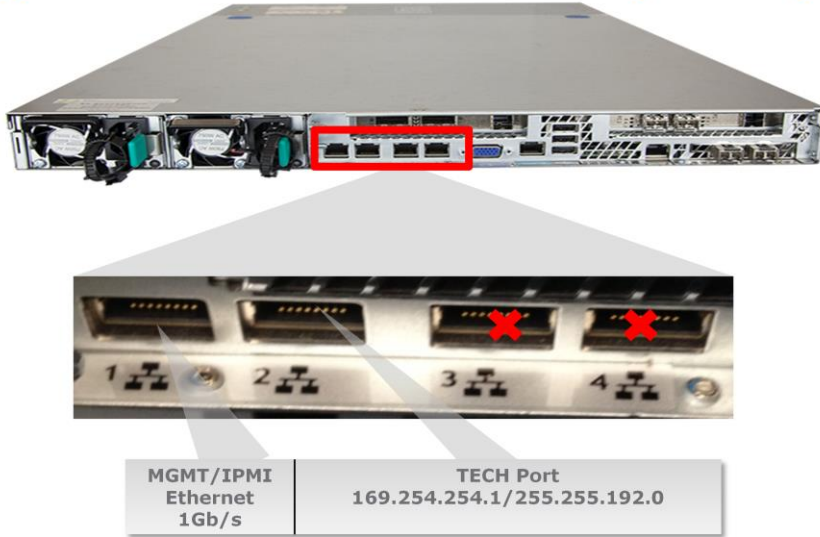
Each Storage Controller has two InfiniBand (IB) ports for the internal XtremIO array communication, named IB1 and IB2.

XtremIO storage arrays provide Fibre Channel (FC) and iSCSI protocol support. The two 8 Gb/s FC ports are named FC1 and FC2.

The two iSCSI 10GbE SFP+ ports are located below the two FC ports. They are named iSCSI1 and iSCSI2.

The two 6 Gb/s SAS ports are used to connect to the X-Brick DAE. They are named SAS1 and SAS2.

Storage Controller: Rear Panel (2 of 3)



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview



17

The Management (MGMT) and Service (TECH) ports are located in the center of the rear panel and consist of 4 RJ45 ports.

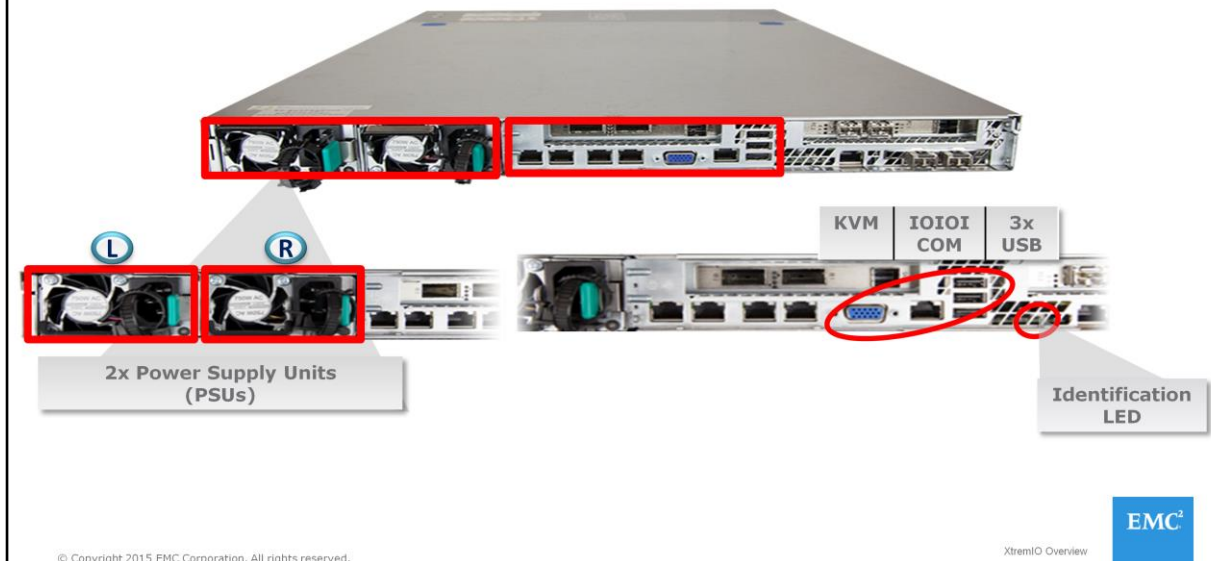
The 1GbE MGMT/IPMI (Intelligent Platform Management Interface) port is labeled network 1, and is connected to the administration network.

MGMT and IPMI interfaces have their own IP addresses on the same physical port, set during cluster creation.

The TECH port is labeled network 2. This is for EMC service personnel only. It provides console access through Secure Shell (SSH).

The remaining two ports (3 & 4) are not used.

Storage Controller: Rear Panel (3 of 3)



The remaining connections are as follows:

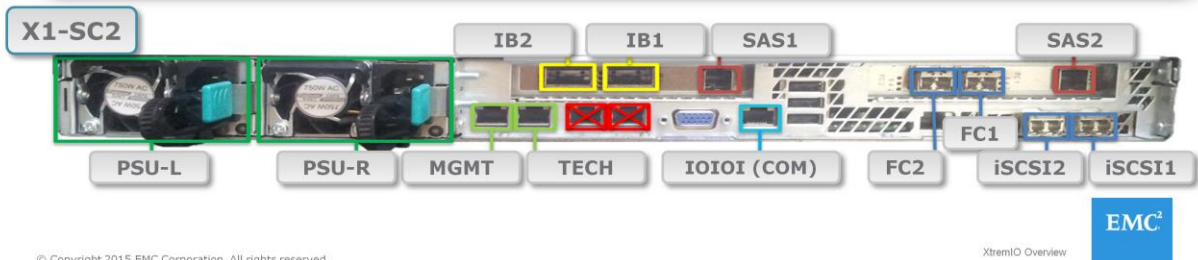
- A KVM/VGA connection for a monitor which might be used in the installation process
- Three active USB connections which might be used for the software installation
- A COM port (RJ45) for the BBU serial connection

There is an identification LED on the rear panel. If you beacon the identification LED, it will beacon both at the front panel and the rear panel.

The Storage Controllers have two redundant Power Supply Units (PSUs), named PSU-L and PSU-R.

Storage Controller: Component/Port Names

Component	Port Name	Example
Power	X<x>-SC<y>-PSU-<L/R>	X1-SC2-PSU-L
MGMT	X<x>-SC<y>-MGMT	X1-SC2-MGMT
InfiniBand	X<x>-SC<y>-IB<1/2>	X1-SC2-IB1
SAS	X<x>-SC<y>-SAS<1/2>	X1-SC2-SAS1
COM	X<x>-SC<y>-IOIOI	X1-SC2-IOIOI
External FC	X<x>-SC<y>-FC<1/2>	X1-SC2-FC2
External iSCSI	X<x>-SC<y>-iSCSI<1/2>	X1-SC2-iSCSI1



Understanding the naming convention that is used in the XtremIO storage array is important. You will need to know this in order to comprehend the documentation, error messages, and the command line output, GUI output, and hardware displays.

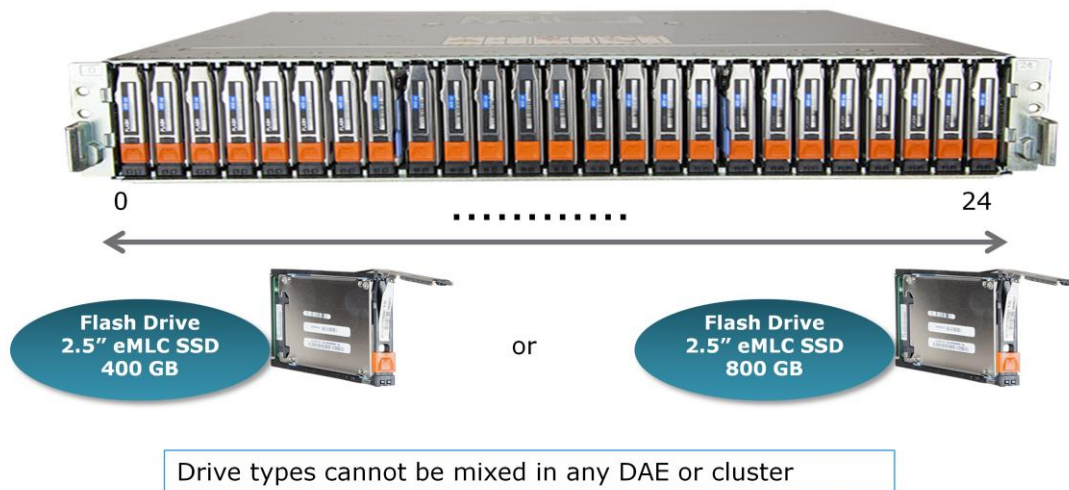
The SCs have PSUs, MGMT/TECH/InfiniBand/SAS/COM ports, and FC/iSCSI ports.

When referencing components and ports, the port and component naming conventions will include the X-Brick number, the X-Brick component (SC, DAE, BBU, etc.), and the port type and number or sub-component.

The examples in the slide use the second SC (SC2) in the first X-Brick (X1).

- If the left PSU of SC2 in X-Brick1 needs to be referenced, then the proper name is X1-SC2-PSU-L.
- If the left MGMT port of SC2 in X-Brick1 needs to be referenced, then the proper name is X1-SC2-MGMT.

Disk Array Enclosure: Front Panel



EMC²

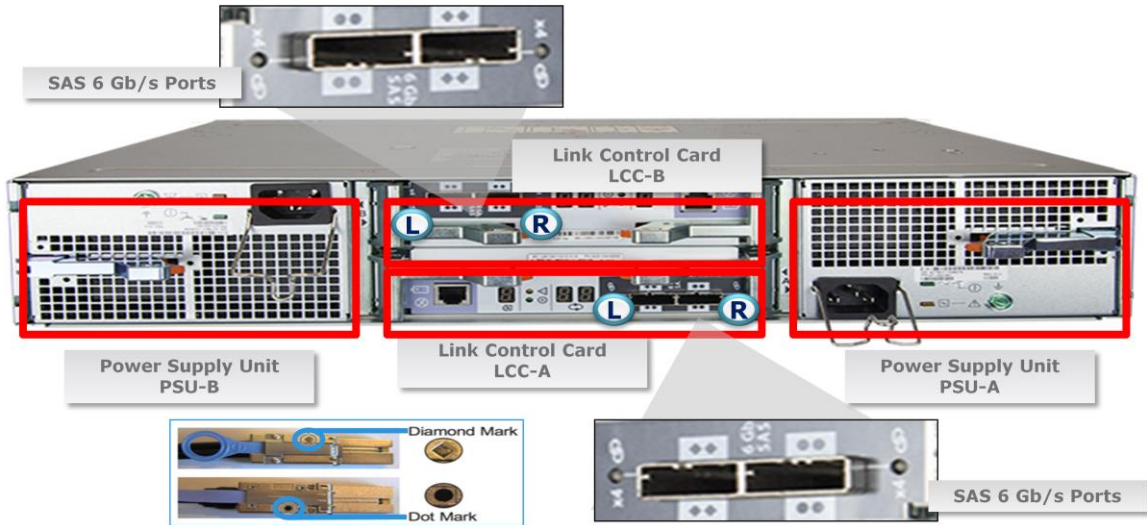
XtremIO Overview

20

The disk array enclosure (DAE) needs 2U of rack space. It is fully-populated with 25 Flash drives in drive slots numbered from 0 to 24. In the half-populated starter X-Brick, there are 13 SSDs in slots 0 through 12; the remaining 12 slots have filler panels to ensure proper cooling.

The drives are Enterprise Multi-level Cell (eMLC) Flash drives (SSDs) with a capacity of 400 GB or 800 GB. Sizes can not be mixed in a DAE or a cluster. eMLC is a type of multi-level cell Flash drive that has been enhanced to accommodate more write cycles than consumer grade MLC Flash offerings.

Disk Array Enclosure: Rear Panel



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

EMC²

21

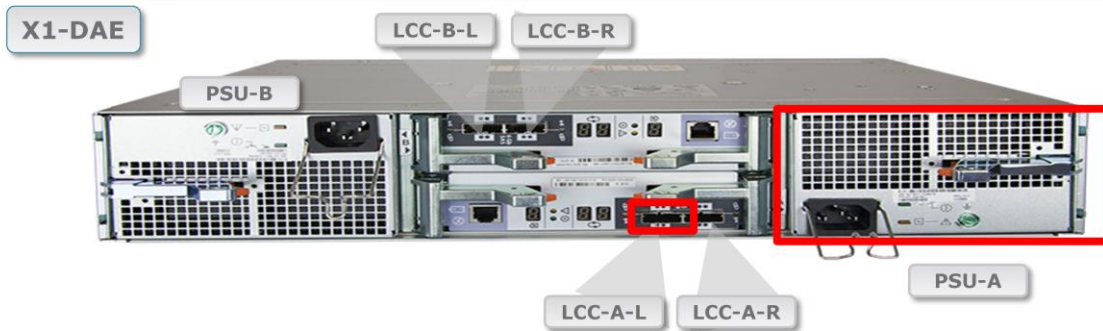
In the rear of the DAE are redundant Power Supply Units (PSUs) named PSU-A and PSU-B. Note that the placement of PSU-A is inverted with respect to PSU-B.

For connecting to the DAE with 6 Gb/s SAS, there are two redundant Link Control Cards (LCCs) named LCC-A and LCC-B. Note that the placement of LCC-A is inverted with respect to LCC-B.

Each LCC has two SAS ports (left and right) that connect to each Storage Controller. In other words, each Storage Controller of an X-Brick is connected to both LCCs. The ports are marked with diamonds and dots that will match the keyed SAS cable labeling. The names of the ports are LCC-A-L (Left) , LCC-A-R (Right) , LCC-B-L (Left), and LCC-B-R (Right).

DAE: Component/Port Names

Component	Port Name	Example
Power	X<x>-DAE-PSU-<A/B>	X1-DAE-PSU-A
SAS	X<x>-DAE-LCC-<A/B>-<L/R>	X1-DAE-LCC-A-L



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

22

The DAE port and component naming is similar to the port and sub-component naming of the SCs.

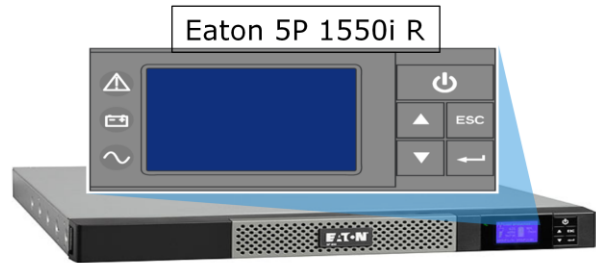
The DAE in the first X-Brick of an XtremIO storage array is used for these examples. Its name is X1-DAE.

The example shows X1-DAE-PSU-A, which references the PSU on the right.

Referencing a SAS port includes the LCC, A or B, and the port, Left or Right. The example references the highlighted SAS port.

Battery Backup Unit: Model Types

- Eaton 5P 1550i R
 - Ships with new orders
 - Technology refresh for the Eaton 1550 Evolution
- Eaton 1550 Evolution
 - End-of-life
 - Still supported
- Battery Backup Unit (BBU) models may not be mixed



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

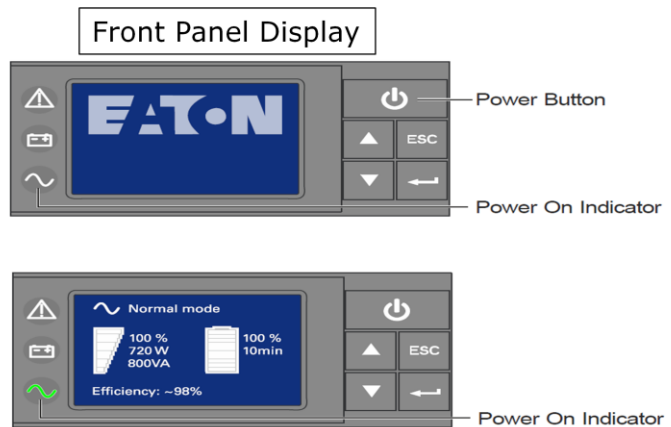
XtremIO Overview

23

XtremIO systems now use the Eaton 5P 1550i R Battery Backup Unit as a technology refresh for the Eaton 1550 Evolution, which has reached EOL.

BBU: Front Panel 5P 1550i R

- LCD display is lit
 - BBU is receiving power
- Required to start cluster
 - Must be charged to not less than 70% of its full capacity
 - Indicated on the LCD display



EMC²

XtremIO Overview

24

The BBU LCD provides information on the status of the BBU. The BBU must be charged to not less than 70% of its full capacity in order for the cluster to start. This information will be displayed on the LCD. If the LCD display is lit, then the BBU is receiving power.

BBU: Rear Panel 5P 1550i R

- At least two in every Cluster configuration
- Provides power to Storage Controllers
- Required to start cluster
 - Both must be charged to at least 70% of full capacity



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

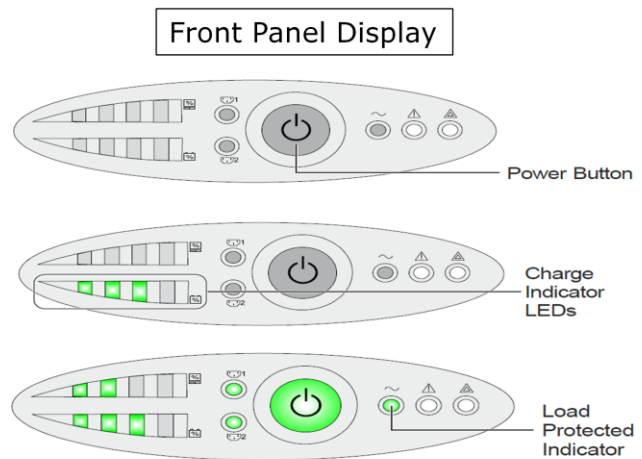


25

This is the rear view of the 5P 1550iR BBU. There needs to be at least two BBUs in every Cluster configuration in order to provide power to Storage Controllers. The COM ports are used for BBU communication and monitoring.

BBU: Front Panel 1550 Evolution

- Two bar graphs
 - Top indicates power consumed from BBU
 - Bottom one shows the battery charge level
- Two LEDs left of power button
 - Status of programmable power outlets
- Power Button
- Three LEDs on right hand side
 - Load-Protected
 - Downgraded-Operation LED
 - Load-Not-Protected



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview



26

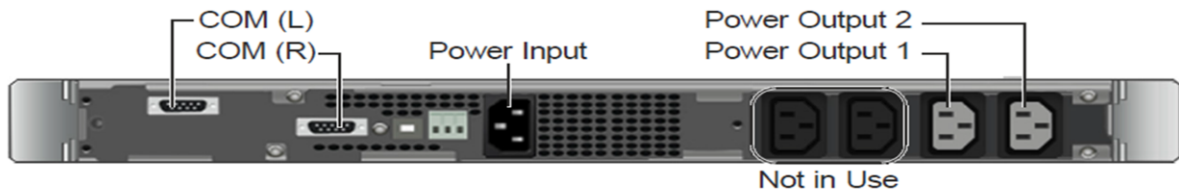
One BBU contains two batteries. It requires 1U of rack space.

The front control panel has a two-bar graph. The top one indicates the percent load (power consumed from BBU), and the bottom one shows the battery charge level.

The next two LEDs represent the status of the two managed (programmable) power outlets. Next is the On/Off button for the BBU and outlets. The next three LEDs are the Load-Protected LED, Downgraded-Operation LED, and the Load-Not-Protected LED.

BBU: Rear Panel 1550 Evolution

- At least two in every configuration
- Provides power to Storage Controllers



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview



27

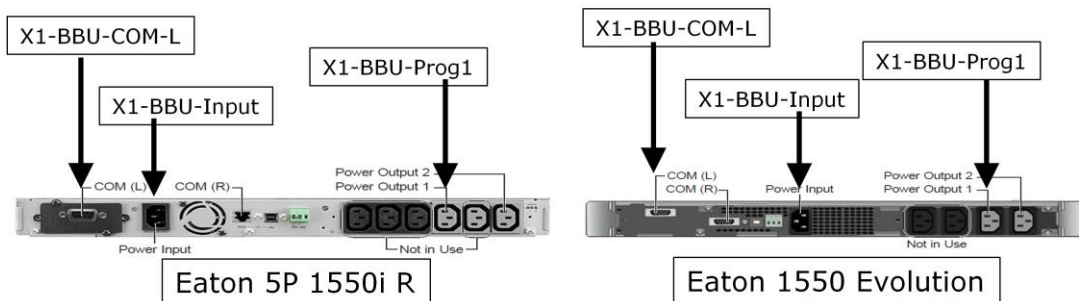
All the connections of the BBU are in the rear panel. Here we find two COM ports, the left and right RS232 serial ports named COM-L and COM-R. These ports are used for the communication connection between the Storage Controllers and the BBU. For these ports, a DB9-RJ45 cable is used.

- The BBU has the DB9 connecting to it
- The RJ45 connector plugs into the COM port of the Storage Controller (marked with IOIOI).

Furthermore, there are power connections: AC from the Power Distribution Unit (PDU) in the rack, as well as power output connections. Of the four output connections, only the two grey managed (programmable) ports are used to connect Storage Controllers, and are named Prog1 and Prog2. The other two black power outputs are not used.

BBU: Power Connection/Port Names

Component	Port Name	Example
Power Input	X<x>-BBU-Input	X1-BBU-Input
Power Output	X<x>-BBU-Prog<1/2>	X1-BBU-Prog1
COM	X<x>-BBU-COM-<L/R>	X1-BBU-COM-L



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

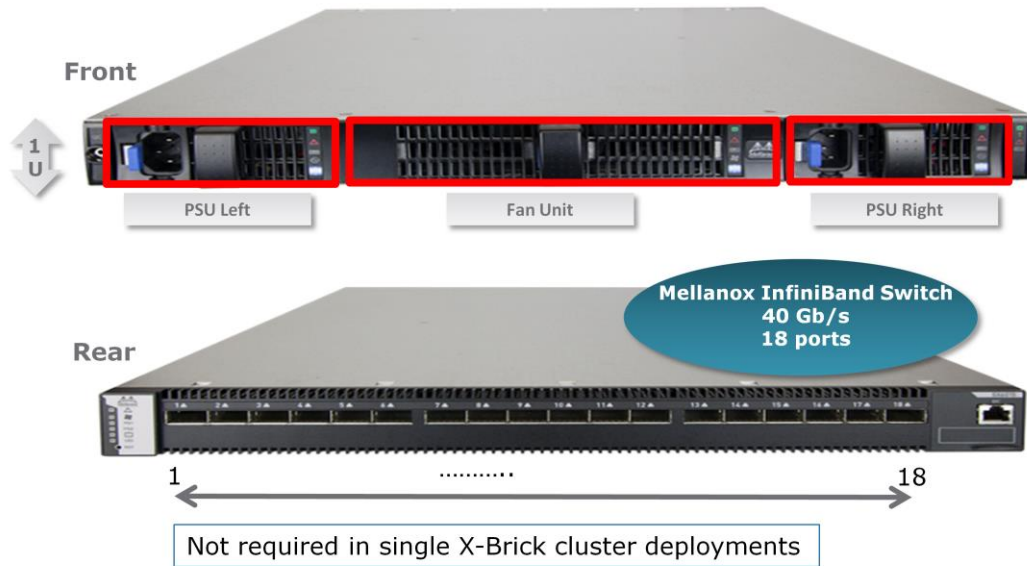
XtremIO Overview

28

The BBUs have serial ports, power input, and power outputs connections. The power connection and port names use the component name BBU (BBU1 or BBU2) and then add the port and number or position. The slide illustrates BBU2 from a single X-Brick.

- The name for the power input would be X1-BBU2-Input
- The name for the two power outputs named Prog1 and Prog2, the example shows X1-BBU2-Prog1
- The name for the COM ports are L for Left and R for Right, the example shows X1-BBU2-COM-L

InfiniBand: Inter X-Brick Communication



The InfiniBand switch needs 1U of rack space. It has 18 ports, which are facing the rear of the cabinet. This allows for short, easy cabling from the SC ports, which are also facing the rear side.

The InfiniBand switch has two redundant Power Supply Units, PSU-L and PSU-R, and a fan unit.

© Copyright 2015 EMC Corporation. All rights reserved.

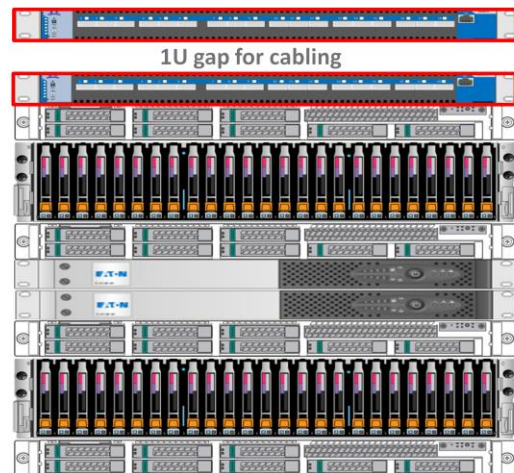
XtremIO Overview

EMC²

29

X-Brick Connectivity: InfiniBand Switches

- Up to six X-Bricks clustered using InfiniBand switches
- Clusters with two or more X-Bricks ship with two 18-port 40 Gb/s switches
- Provide ultra-low latency RDMA communications between X-Bricks
- No management necessary



EMC²

XtremIO Overview

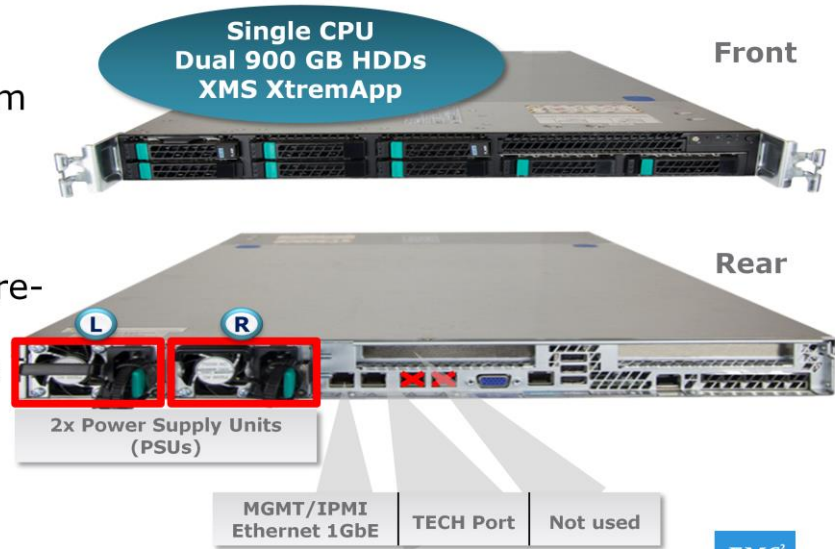
30

Multiple X-Bricks can be combined to form a single system providing linear scalability of both performance and capacity. On a single X-Brick cluster, the SCs are directly connected to one another without the need for InfiniBand switches. With two or more X-Bricks, two 18-port Mellanox InfiniBand 40 Gb/s switches are provided with the system. The switches provide a highly available back-end network that enables high speed, ultra-low latency Remote Direct Memory Access (RDMA) communication between all Storage Controllers in the cluster.

By leveraging RDMA, the XtremIO system is in essence a single, shared memory space spanning all Storage Controllers. Note that the InfiniBand switch is an unmanaged switch, so there is no configuration or management by EMC or the customer other than connecting the cables.

Physical XtremIO Management Server (XMS)

- XMS software requires a platform to run on:
 - Physical
 - Virtual
- Physical XMS is pre-installed with required software
 - May require upgrade



Each XtremIO cluster needs a management server, which can be physical or virtual. If the physical solution is chosen, then the XtremIO Management Server (XMS) requires 1U of rack space. The XMS image and the XMS Xtremapp software are pre-installed on the physical XMS.

On the rear panel, you will find two redundant Power Supply Units (PSUs) – Left and Right – and the MGMT/IPMI and TECH port. Other connections are not used.

If the virtual appliance approach is chosen, an OVA file is used to install the virtual XMS in a VMware environment.

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

31

XtremIO X-Brick Clusters

- Available in 1, 2, 4, and 6 X-Brick configurations
- Number of components varies depending on cluster size
 - Multi X-Brick clusters share some components
 - BBUs and InfiniBand switches

	Single Cluster	Dual Cluster	Quad Cluster	Hex Cluster
# of X-Bricks in cluster	1	2	4	6 [20 TB only]
# of storage controllers	2	4	8	12
# of DAEs	1	2	4	6
# of BBUs	2	2	4	6
# of InfiniBand switches	0	2	2	2
# of XMS servers	1	1	1	1

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

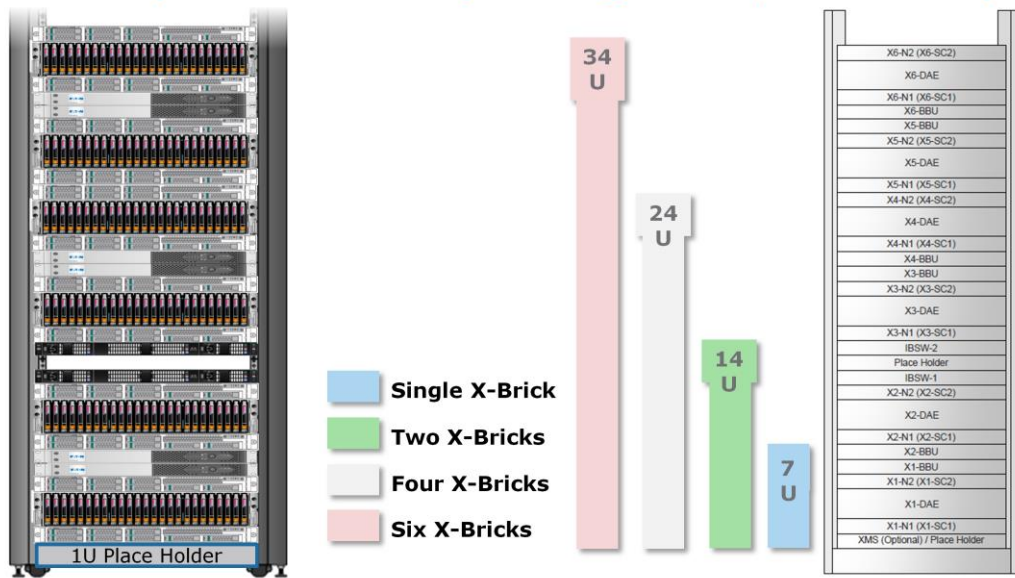
XtremIO Overview

32

Current supported configurations for the XtremIO storage array are one, two, four, and six X-Brick clusters. The chart in this slide represents the required components for each of the different system offerings.

Note that a half-populated starter X-Brick can only be used in a single X-Brick cluster.

Rack Space: Factory Integrated/Field Deployed



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

EMC²

33

The XtremIO storage array ships as factory integrated in an EMC cabinet, or is field deployed in a user-supplied cabinet.

The first 1U is reserved for the physical XMS. If a virtual XMS is deployed, the space is left open. When an array consists of more than one X-Brick, two 1U InfiniBand switches are added, and a 1U space is left between them for cabling.

A one X-Brick array uses 7U of rack space; a two X-Brick array uses 14U of rack space; a four X-Brick array uses 24U of rack space; and a six X-Brick array uses 34U of rack space.

Lesson Summary

During this lesson the following topics were covered:

- Pre-site considerations
- XtremIO hardware components



This lesson covered the XtremIO hardware components. We performed a detailed review of the XtremIO components including physical port description and usage as well as naming conventions.

Lab Introduction: Getting Started

This lab covers the process of connecting to the lab equipment and obtaining product documentation. Operations are:

- Connecting to the VDC
- Examining documentation
- Testing system connectivity

The EMC logo consists of the letters "EMC" in white, with a superscript "2" on the "C", set against a blue square background.

XtremIO Overview

35

In order to complete the lab exercises in this guide there are certain pre-requisite tasks that must be completed. You must be able to connect to the remote equipment being presented and obtain product documentation. The product documents you obtain will be referenced throughout this course.

Lab Debrief

This lab covered:

- Connecting to the VDC
- Examining documentation
- Testing system connectivity

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview



36

This lab covered connecting to the remote equipment being presented and obtaining product documentation. The product documents you obtained will be referenced throughout this course.

Lesson: External Network Requirements

This lesson covers the following topics:

- Fiber Channel zoning
- Ethernet connectivity



This lesson covers external network requirements which include Fibre Channel zoning and iSCSI connectivity.

Site Requirements: Network

Configuration	Fibre Channel Connectivity	iSCSI Connectivity*	Ethernet Ports RJ45 1GbE**
One X-Brick cluster	4 x 8 Gb/s FC ports on the switch	4 x 10GbE LC optical ports on the switch	2 (+1)
Two X-Brick cluster	8 x 8 Gb/s FC ports on the switch	8 x 10GbE LC optical ports on the switch	4 (+1)
Four X-Brick cluster	16 x 8 Gb/s FC ports on the switch	16 x 10GbE LC optical ports on the switch	8 (+1)
Six X-Brick cluster	24 x 8 Gb/s FC ports on the switch	24 x 10GbE LC optical ports on the switch	12 (+1)

* Either Fibre Channel or iSCSI are required for a host

** An additional port is required if a physical XMS is implemented

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

38

XtremIO offers storage connectivity via Fibre Channel and/or iSCSI therefore the proper cables must be supplied and correctly configured in order to successfully present storage. In addition to this XtremIO requires connectivity for management via Ethernet. Note that any single host should use FC or iSCSI, but not both.

Network Firewall Settings

- Allocate two internal network IP addresses for each Storage Controller
- Make sure that the XMS can access these addresses
- It is recommended that all XtremIO IP addresses be protected against external access
 - If the XMS IP address is to be exposed only to specific hosts, use a firewall and open ports to designated hosts

- ◆ Set the FW rules prior to installation.
- ◆ If the XMS is on a different subnet than the X-Bricks, the following firewall ports must remain open in both directions:
 - TCP ports 22 and 11111
 - UDP ports 623 and 123
 - ICMP
- ◆ Open the following TCP ports between the XMS and the managing desktop running the XMS GUI:
 - 22
 - 25
 - 80
 - 443
 - 42502
 - 42503
- ◆ Open the following services from the XMS to the relevant target systems if you want to enable them on the XMS:
 - Syslog UDP 514
 - LDAP TCP 389
 - LDAPS TCP 636
 - LDAP Global Catalog 3268
 - SMTP TCP 25
 - SNMP UDP 162
 - DNS USP and TCP 53, outbound

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

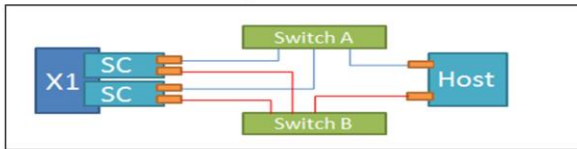


39

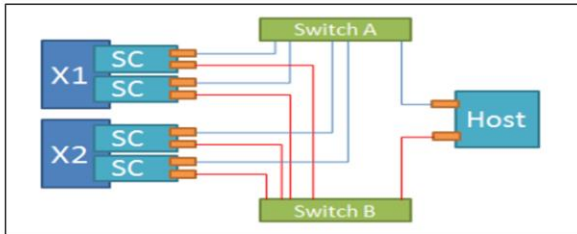
In order to ensure the XtremIO will function properly ensure the firewall settings listed in this slide are configured.

Zoning and IP

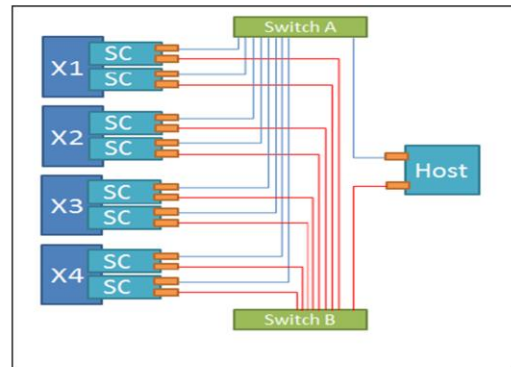
Single Cluster



Dual Cluster



Quad Cluster



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

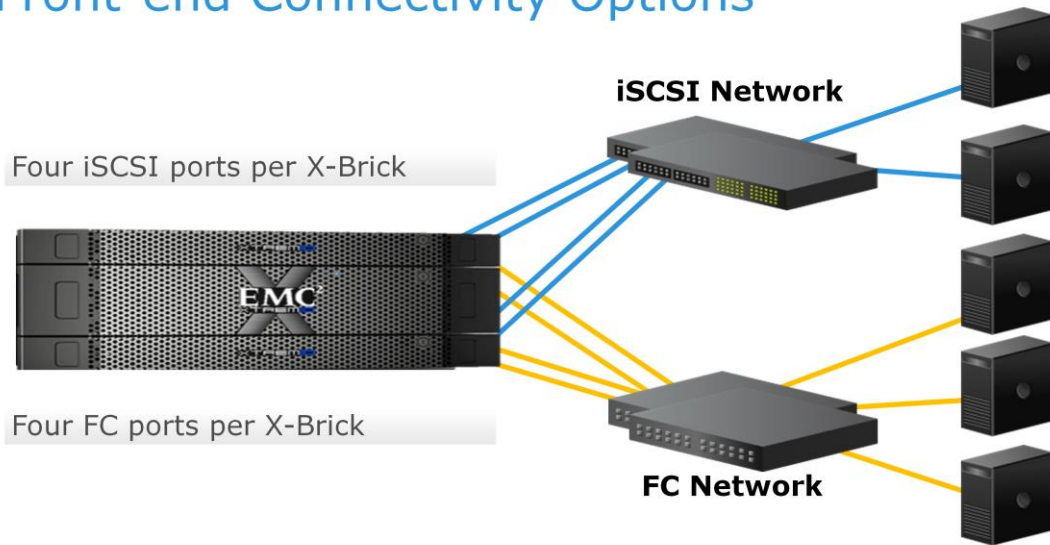
XtremIO Overview

40

When performing the Fibre Channel and Ethernet cabling the ideal configuration is to have redundant paths and redundant switches as shown in the slide.

Note that a single host will not connect to all ports on a 6 X-Brick cluster; doing so would violate the rule that no more than 16 paths should be active between a host and any volume. Since a 6 X-Brick cluster has 24 FC front-end ports, hosts connecting via FC will share connections across the X-Bricks. As an example, one host may connect to X-Bricks 1, 2, and 3, while another connects to X-Bricks 1, 3, and 4, and a third host connects to X-Bricks 4, 5, and 6.

Front-end Connectivity Options



© Copyright 2015 EMC Corporation. All rights reserved.

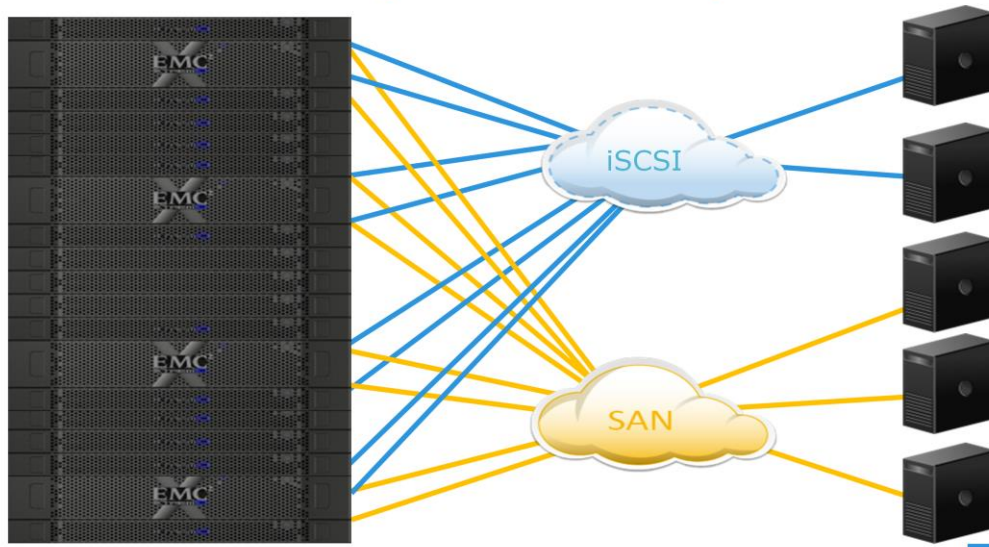
XtremIO Overview



41

XtremIO supports both 8/4/2 Gb/s Fibre Channel and 1/10 Gb/s iSCSI with SFP+ optical connectivity to the hosts. Each X-Brick provides four FC and four iSCSI front-end ports. Hosts can use either of the two protocols. No single host should use both at the same time.

Front-end Ports: Equal and Always Active



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

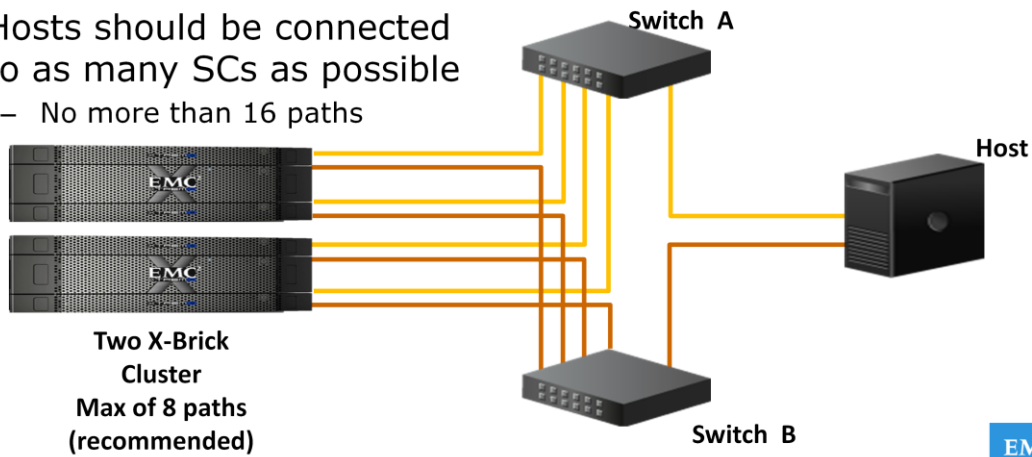


42

A key feature of XtremIO is that all volumes are accessible via all of the front-end ports. Every port provides high performance access to any data anywhere in the system. XtremIO controllers are N-way active, meaning that they are all active at the same time. There is no ALUA, primary controller, or LUN trespasses. All FC and iSCSI front-end ports have access to all LUNs and use the resources of the entire XtremIO cluster for maximum performance.

FC Connectivity Considerations – Zoning

- Only one Host initiator per Zone
- Hosts should be connected to as many SCs as possible
 - No more than 16 paths



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

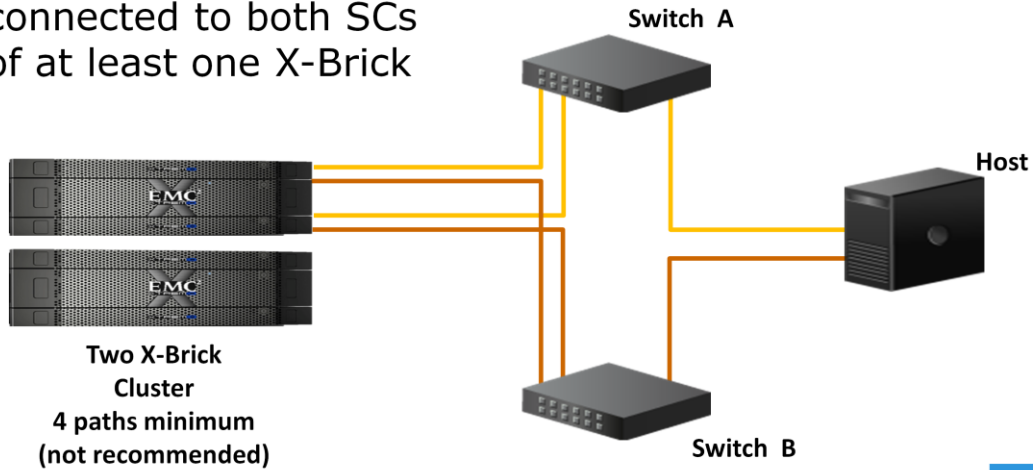
43

As in every SAN storage environment, a highly available environment requires at least two HBA connections per host with each HBA connected to a separate Fibre Channel switch, as shown here. Connecting the XtremIO cluster to the FC switches is also straight forward. Each Storage Controller has 2 FC ports and therefore you should connect each Storage Controller to each Fibre Channel switch.

Each X-Brick of an XtremIO system can lose one Storage Controller and still remain fully functional. In general, every host should be connected to as many Storage Controllers as possible on an XtremIO cluster, as long as the host and multipathing software supports that number of connections. Best practice indicates 4-8 paths per host for dual X-Brick clusters, up to four paths for single X-Brick clusters, up to 16 paths for quad X-Brick clusters, and never to include more than one host initiator in a zone. To avoid multipathing performance degradation, do not use more than 16 paths per device. All volumes are accessible via any and all of the front-end ports.

FC Connectivity Considerations: Zoning (2)

- Hosts should be connected to both SCs of at least one X-Brick



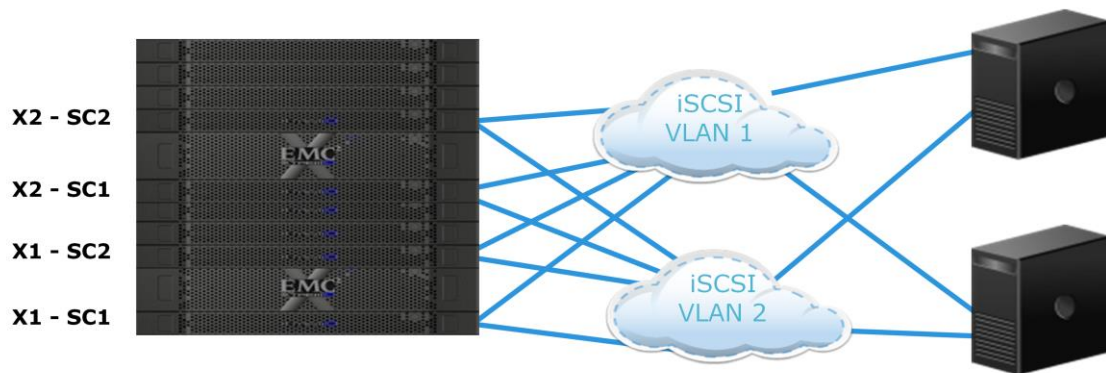
XtremIO Overview



44

At a minimum, every host should be connected to both Storage Controllers of at least one X-Brick. For example, in a two X-Brick cluster, if only 4 paths will be used, then the host needs to be connected to both SCs of one of the X-Bricks. When it comes to a four or six X-Brick cluster, if only 4 or 8 paths will be used, then you must decide which X-Brick will have all 4 connections, or which pair of X-Bricks will have all 8 connections. Please refer to the XtremIO User Guide for more details on zoning.

iSCSI Connectivity Considerations



iSCSI portals cannot be in the same subnet as the management network

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

45

General best practice for highly available iSCSI environments is for every host to have two physical adapters and for these adapters to be connected to separate VLANs, as shown here. As with FC connectivity, connecting the XtremIO iSCSI ports is easy. Since each SC has two iSCSI ports, simply connect each SC to a separate iSCSI subnet or VLAN.

Note that the iSCSI portals cannot be in the same subnet as the management network.

Multi-pathing

Host Platform	Path Management	Cluster
Oracle Linux	PowerPath, Native MPIO, VERITAS VxDMP	Y
Red Hat Linux	PowerPath, Native MPIO, VERITAS VxDMP	Y
Microsoft Windows	PowerPath, MPIO	Y
Solaris	PowerPath, MPxIO, VERITAS VxDMP	Y
VMware	PowerPath/VE, NMP, VERITAS VxDMP	Y

Refer to the Support Matrix for up-to-date information



Multi-pathing software significantly improves performance and redundancy. XtremIO supports a wide array of host-based multi-pathing solutions as well as EMC PowerPath. Refer to the latest support matrix for the most up-to-date information.

Lesson Summary

During this lesson the following topics were covered:

- FC requirements
- Ethernet requirements for iSCSI



This lesson covered external network requirements which included Fibre Channel zoning and iSCSI.

This slide intentionally left blank.



Lesson: XtremIO Management Server (XMS)

This lesson covers the following topics:

- Pre-installation requirements
- XMS attributes
- XMS configuration

This lesson covers XtremIO Management Server (XMS) physical and virtual installations.

Provide DNS, Date, and Time

- Ensure DNS, timezone, date and time, and NTP information are provided for XMS
- The results can be verified with these CLI commands:
 - `show-dns`
 - Primary
 - Secondary
 - `show-datetime`
 - Timezone
 - Date and time
 - NTP server

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

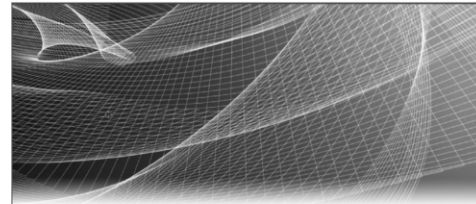


50

To configure the optional DNS name resolution and other network services, open an SSH session to the XMS and log in to the Linux shell as xmsadmin. Then, log in to the XtremIO system as the admin user.

Power On All Components

- The power-up procedure requires running CLI commands with appropriate privileges
- Use the procedures in the XtremIO Storage Array User Guide for:
 - Starting the cluster for the first time
 - Restarting the cluster following a logical shutdown
 - Powering Up the XMS
 - Powering Up the Battery Backup Units
 - Powering Up the Storage Controllers
 - Starting the Cluster



EMC
XtremIO Storage Array

Versions 3.0.1 and 3.0.2

User Guide

P/N 302-001-530
REV. 05

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

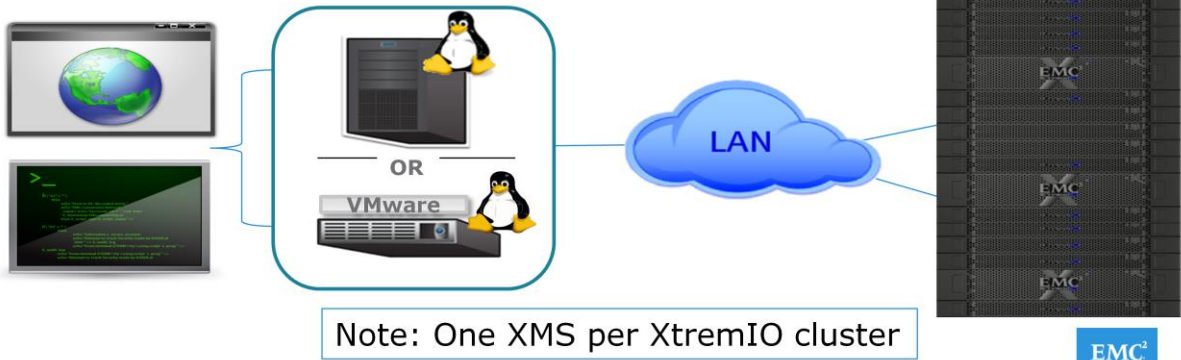
XtremIO Overview

51

The system has power on procedures and shutdown procedures which must be followed. Information on those procedures is available in the User Guide.

XtremIO Management Server (XMS)

- Required Linux-based server
- Management station for an XtremIO system:
 - CLI management via an SSH session
 - GUI management via a Java WebStart application



The XtremIO Management Server (XMS) is a required server used to manage the XtremIO storage system. The XMS is preinstalled with the CLI and GUI via a Java WebStart application. It can be deployed on a dedicated physical XMS server, or as a virtual machine on VMware using an OVA template. It is not supported on Hyper-V. The XMS must access all management ports on the X-Brick Storage Controllers, and must be accessible by any GUI/CLI client host machine. Since all communications use standard TCP/IP connections, the XMS can be located anywhere that satisfies these requirements.

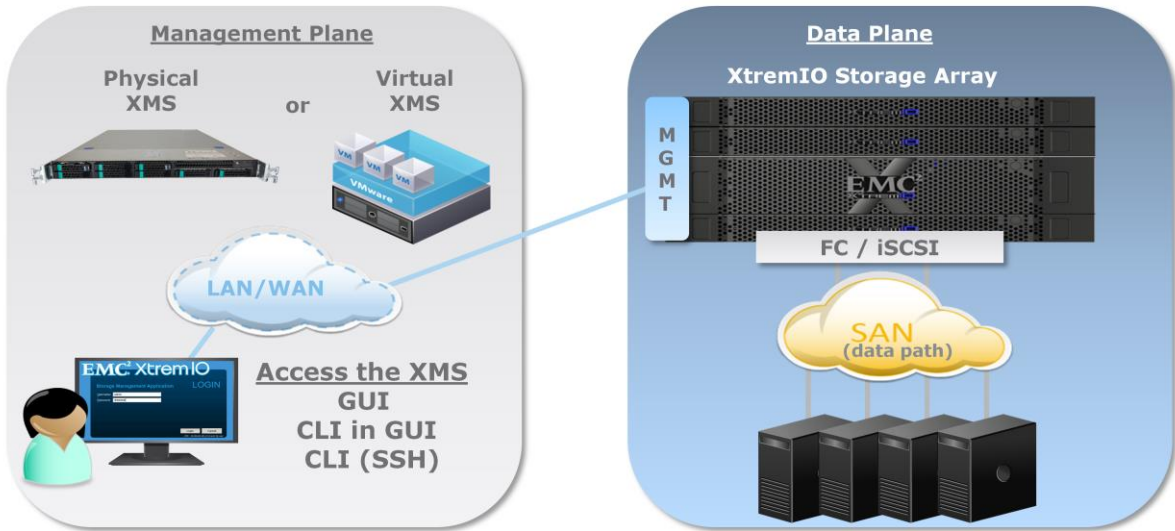
Note: Only one XMS can manage a cluster, and one XMS is required per cluster. If you have two single X-Brick clusters, then you must have two XMS instances, one for each. If you have one quad X-Brick cluster, only one XMS is required.

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

52

XtremIO Management Server (XMS)



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

EMC²

53

The XtremIO Management Server (XMS) enables you to control and manage the cluster. It is part of the Management Plane and TCP/IP network connected to the management ports of each SC in the cluster.

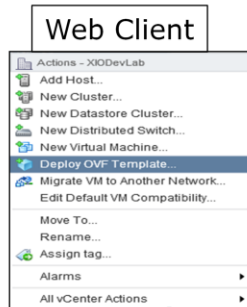
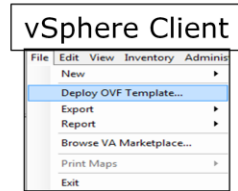
XMS provides:

- Cluster health, events, and cluster performance information
- Performance statistics history database
- Volume management
- Data protection groups operation logic
- Stopping, starting, and restarting of the cluster

The CLI allows cluster administrators and operators to perform supported management operations. It is pre-installed on the XMS, and can also be accessed using standard SSH programs.

Virtual XMS Requirements

- The virtual XMS VM should be deployed on a single host
 - ESXi 4.1 or 5.x
- ESXi host must be able to provide:
 - 8 GB vRAM
 - 2 vCPUs
 - 1 vNIC
 - Single 900 GB disk (thin provisioned)
- The virtual XMS VM can be protected with VMware HA
 - Shared storage in this case should not originate from an XtremIO cluster



XtremIO Overview

54

The virtual XMS VM should be provisioned on RAID protected storage.

The OVA package from which the virtual XMS is deployed contains VMware tools. Therefore, no VMware tools upgrade is required following virtual XMS deployment.

Since the XMS is not in the data path, it can be disconnected from the XtremIO cluster without affecting the I/O. An XMS failure only affects monitoring and configuration activities, such as creating and deleting volumes. However, when using a virtual XMS topology, it is possible to take advantage of VMware vSphere HA features to easily overcome such failures.

Deploy XMS VM OVA

- This step is not required if a physical XMS is implemented
- Performed by EMC or Partner during installation
- Provide required template information
 - Machine name
 - Datastore location
- Power ON

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview



55

After the Storage Controllers are configured, you need to install an XMS. If the virtual appliance option is chosen, navigate to the Downloads page of support.emc.com, and locate the OVA file. The OVA image includes all the required packages for installing the XMS.

To deploy the OVA image, launch the appropriate vSphere interface and deploy the OVA template. You will need to supply a machine name and select a datastore for it. The VM disk on the datastore requires 900 GB of thin-provisioned capacity.

EMC Secure Remote Support (ESRS)

- Support for ESRS GW and ESRS IP
- Provides Connect-In and Connect-Home functionality
- Recommended for all XtremIO environments
- If customer refuses ESRS as a connectivity solution
 - Exception that should be pre-approved
 - Configure XMS to Connect-Home only
 - Email or FTPS
 - These do not provide Connect-In functionality to the XMS
 - EMC receives regular configuration report and product alert information from the customer's XtremIO environment



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Overview

56

EMC Secure Remote Support (ESRS) provides a secure, IP-based, distributed remote support solution that enables command, control, and visibility of remote support access. ESRS fulfills the requirements for authentication, authorization, and auditing with a secure, highly scalable, fault-tolerant system. This IP-based, firewall-friendly remote access architecture initiates all connections from the customer's site.

ESRS provides rapid remote support through two phases:

1. Automated recognition and notification from the customer's site to EMC.
2. Interpretation and response from EMC.

ESRS configuration options with XtremIO are ESRS GW and ESRS IP Client. These two configuration options provide Connect-In and Connect-Home functionality.

Since these configuration options provide the best level of remote support to the cluster, it is highly recommended that you select one of them when completing XMS installation. If the customer refuses to use ESRS as a connectivity solution, you must configure the XMS to Connect-Home only. The Connect-Home only options with XtremIO are Email or FTPS. These two options do not provide Connect-In functionality to the XMS, but merely ensure that EMC receives regular configuration information.

Lesson Summary

During this lesson the following topics were covered:

- Pre-installation requirements
- XMS attributes
- XMS configuration



This lesson covered XtremIO Management Server (XMS) physical and virtual installations.

You should have learned that the XtremIO Management Server (XMS) is required to manage the XtremIO storage system; the XMS is preinstalled with the CLI and GUI, and it can be deployed on a dedicated physical XMS server, or as a virtual machine on VMware. You also saw how only one XMS can manage a cluster, and one XMS is required per cluster.

We also reviewed the virtual XMS requirements, and how the XMS VM should be provisioned on RAID protected storage. You should have learnt that since the XMS is not in the data path, it can be disconnected from the XtremIO cluster without affecting the I/O. An XMS failure only affects monitoring and configuration activities, such as creating and deleting volumes. However, when using a virtual XMS topology, it is possible to take advantage of VMware vSphere HA features to easily overcome such failures.

Module Summary

Key points covered in this module:

- Identify XtremIO product offerings
- Define the hardware components of an XtremIO array
- Define external network requirements
- Define the XMS server



This module introduces the XtremIO hardware. Starting with an overall product review, the topics discussed include detailed hardware review, cabling, and some site pre-requisites.

Module: XtremIO Cluster Management

Upon completion of this module, you should be able to:

- Interact with the XMS GUI interface
- Interact with the XMS CLI interface
- Perform user management



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

1

This module focuses on the management of the XtremIO cluster system through GUI and CLI options.

This slide intentionally left blank.



Lesson: XMS GUI Management

This lesson covers the following topics:

- XMS GUI prerequisites
- XMS navigation



This lesson covers management of an XtremIO storage array using the XMS GUI interface.

XMS GUI Prerequisites

- Check EMC Simple Support Matrix
 - support.emc.com
- IP connectivity to XMS
 - Physical or virtual
- Internet browser
 - Internet Explorer 9; Internet Explorer 10
 - Google Chrome (Latest version)
 - Mozilla Firefox (Latest version)
- Java 7u71 or 8u25 (See Release Notes)



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management



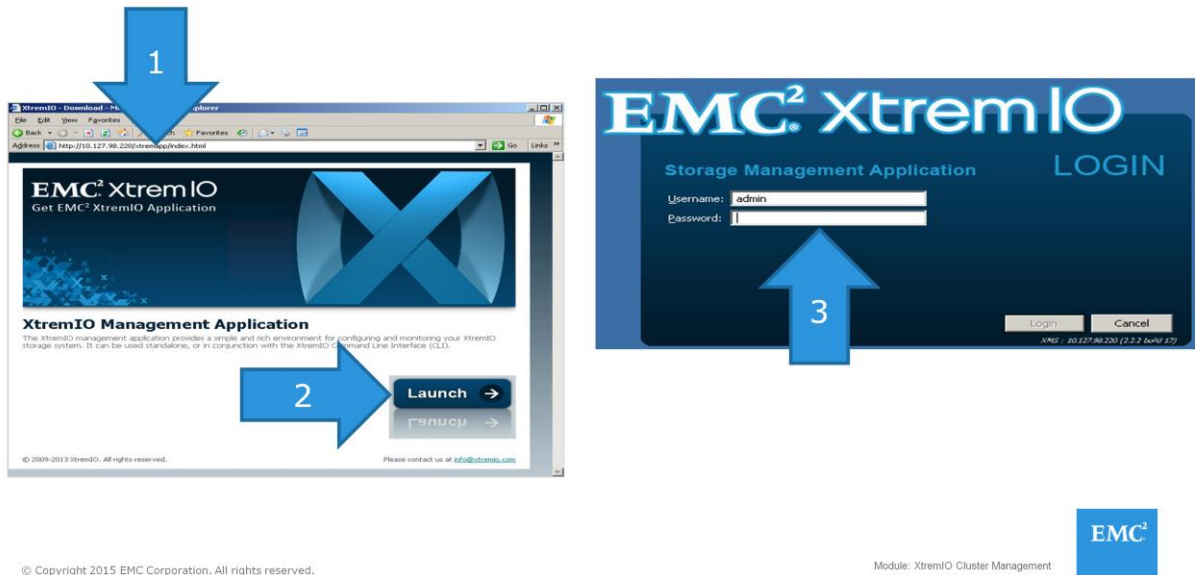
4

Before you access the XtremIO Management Server (XMS) GUI, confirm that you meet the requirements listed here.

The GUI is implemented using a Java client. The GUI client software communicates with the XMS using standard TCP/IP protocols, and can be used in any location that allows the client to access the XMS. The GUI provides easy-to-use tools for performing most cluster operations.

Certain management operations must be performed using the CLI. Operations on multiple components, such as creating multiple volumes, can only be performed using the GUI.

Accessing the GUI



To manage the XtremIO cluster, an administrator must first access the XMS server, whether physical or virtual. In a browser window, enter the IP address or hostname of the XMS server. This will open the XtremIO Management Application download window from the server. The XMS software includes a pre-installed self-signed certificate that enables the initial secure communication between the client machine and the XMS. A link to the XMS server root certificate will be displayed on the Launch page. Simply click on the link and save the certificate to continue.

Click the Launch icon on the page to start downloading the management application. Once the download has completed, the Login window should appear. Enter the username and password of your account and click the Login button to enter the cluster.

Initial Login Screen

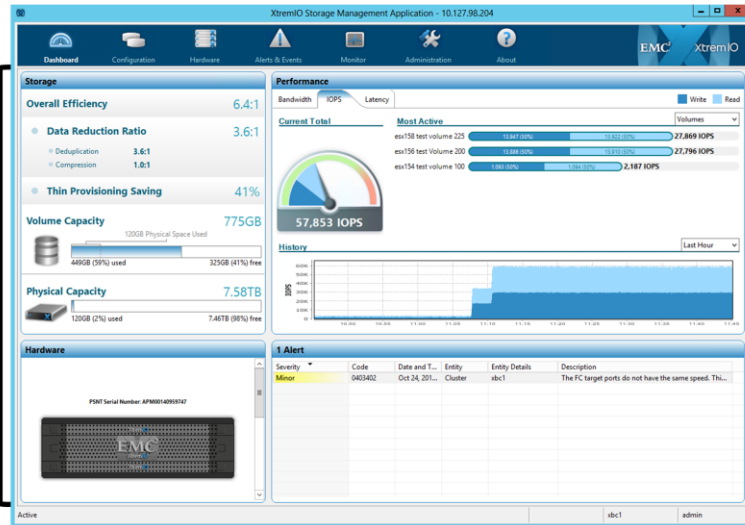
Menu Bar



Workspace



Status Bar



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

6

Using the proper terminology is important for effective communication. The terminology in the slide is used to provide a locational reference throughout this course. This image shows the initial login screen, and identifies the three major areas of the GUI screen.

GUI Menu Bar



- Dashboard – System status and performance data
- Configuration – Creates LUNs and initiator groups and maps them together
- Hardware – Installed components, status, properties, and wiring
- Alerts & Events – Shows system alerts and allows filtering of system events
- Monitor – Monitors specific aspects of system overall performance
- Administration – User administration, notifications, CLI terminal, and configuration



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

7

The menu bar provides:

Dashboard: Home screen that displays an overview of system status, alerts, usage, and performance data.

Configuration: Used to create volumes and volume folders, Initiator Groups and IG folders, show volume and connectivity information, and map volumes to IGs.

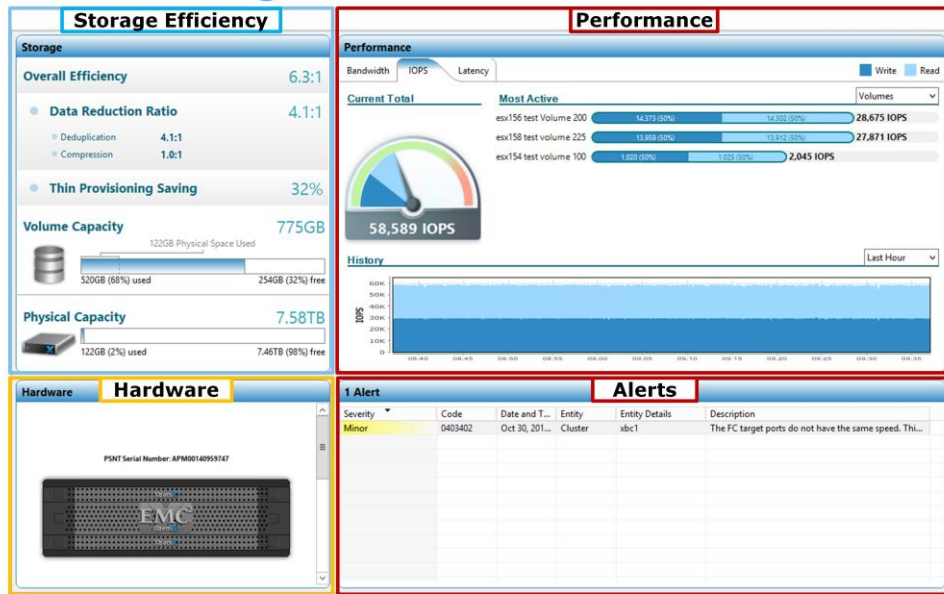
Hardware: Shows installed components, their status, properties, and cable connectivity diagrams.

Alerts & Events: Shows system alerts and allows filtering of system events, based on multiple criteria.

Monitor: Allows users to monitor specific aspects of the system's overall performance and health.

Administration: Manages users, configures iSCSI portals and routes, configures SNMP and email settings, and runs CLI commands.

Dashboard Page



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

EMC²

8

Shown here is the Dashboard Workspace that is first displayed when the Management GUI is opened. Across the top ribbon is the Menu Bar with icons that take users to all the workspaces in the application. Here, you can see the Dashboard icon, highlighted in red. Clicking this icon from any other window will bring the user back to this location.

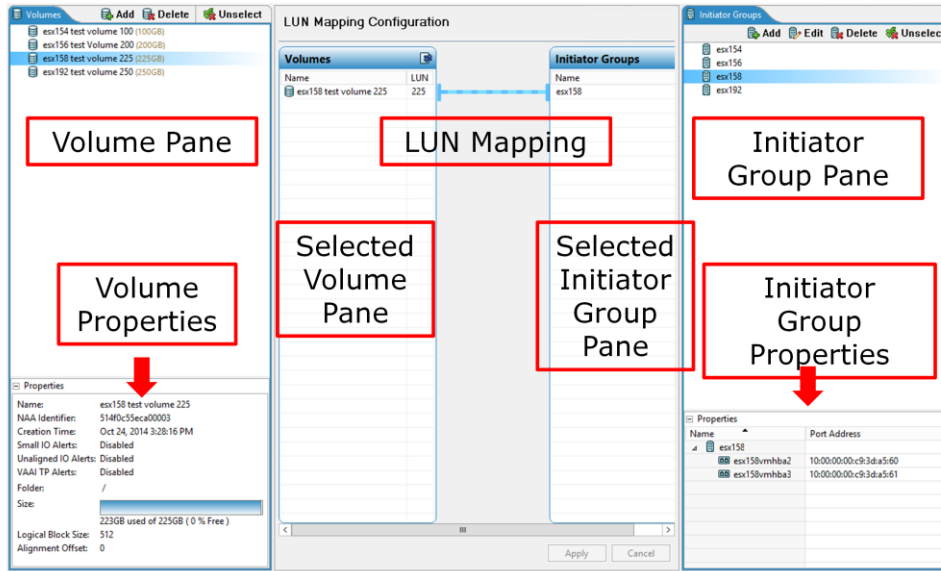
The upper left section of the page displays the storage efficiency for the cluster.

The upper right section of the page displays the general performance of the cluster with a performance history graph.

The lower left pane provides a quick overview of the status of the hardware along with the PSNT (Product Serial Number Tag) if set for the cluster.

The pane in the lower right displays all alerts that have not yet been acknowledged beginning with the most important.

Configuration Page



9

Shown here is the Configuration Workspace.

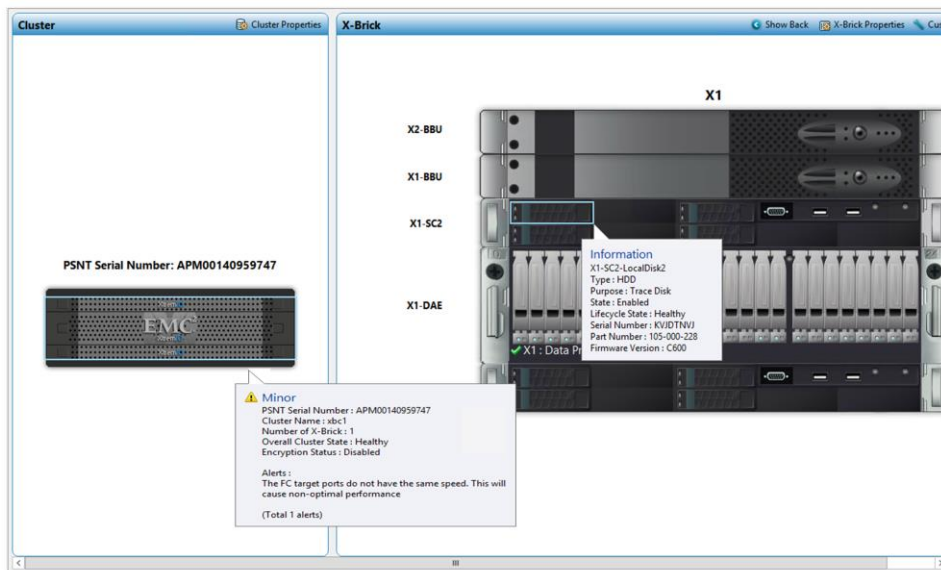
On the left, is the Volume pane. The Volume pane is used to add, delete, and manage volumes. If a volume is selected in this pane, the properties of the volume will be displayed in the Volume Properties pane below.

On the far right is the Initiator Group pane. This pane is used to add, delete, edit, and manage Initiator Groups, which are groupings of WWNs or IQNs from a host that can access the LUNs. If an initiator is selected here, the properties of the group will be displayed below, including HBA WWNs and iSCSI IQNs.

The pane in the center left is the Selected Volumes pane. Any volumes that are selected in the Volume pane will be displayed here so that they can be mapped/unmapped. In the center right, is the Selected Initiator Group pane. As with volumes, any initiators selected in the Initiator Group pane will be displayed here to be mapped/unmapped from the selected volumes.

The last item of note in the Configuration Workspace is the LUN Mapping Connector. This connector shows which LUNs and initiators have been mapped together.

Hardware Page



The Hardware Workspace can be accessed two ways. The first is by clicking the cluster image in the Dashboard Workspace. The second is by selecting the Hardware icon from the Menu. The Rack pane on the left, displays all the components that make up the cluster. If the cursor is passed over parts of the cluster, any alerts for that component will appear.

For a single X-Brick cluster, the X-Brick will be shown in the right pane. In a multiple X-Brick cluster, whichever X-Brick or InfiniBand (IB) switch is selected in the Rack pane, appears in the right pane. In this example, X-brick X1 is selected on the left and appears in the pane on the right. Similar to the left pane, if the cursor is hovered over any main assembly in the right pane, any alerts for it are displayed. When hovering over individual components, their properties are displayed.

Events and Alerts Page - Alerts

The screenshot displays the EMC XtremIO Cluster Management interface, specifically the Alerts page. The main window shows a table of alerts with columns for Severity, Code, Date and Time, Entity, Entity Details, and Description. Two alert icons (a green checkmark and a yellow warning triangle) are highlighted in the top right corner. Two modal windows are open: 'Alert Properties' on the left and 'Edit Alert Definition' on the right. The 'Alert Properties' window shows details for alert 'Inconsistent Fc Tar Speed' (Severity: Minor, Code: 0403402, Time: Oct 30, 2014 8:20:06 AM, Entity: Cluster, Entity Details: xbc1). The 'Edit Alert Definition' window shows the same alert with fields for Description, Activity Mode (checked), Clear Mode (radio buttons for Auto Clear and Acknowledgment Required), and Severity (dropdown menu).

The Alerts and Events Workspace contains two separate tabs, one for Alerts and the other for Events.

The Alerts tab can be used to display all the alerts on the cluster. If an alert is double-clicked, it will open the Alert Properties Page for that alert.

The alert can also be modified by using the Edit Alert Definition button. This will open the Edit Alert Definition window where the Description, Activity Mode, Clear Mode, and Severity can be modified.

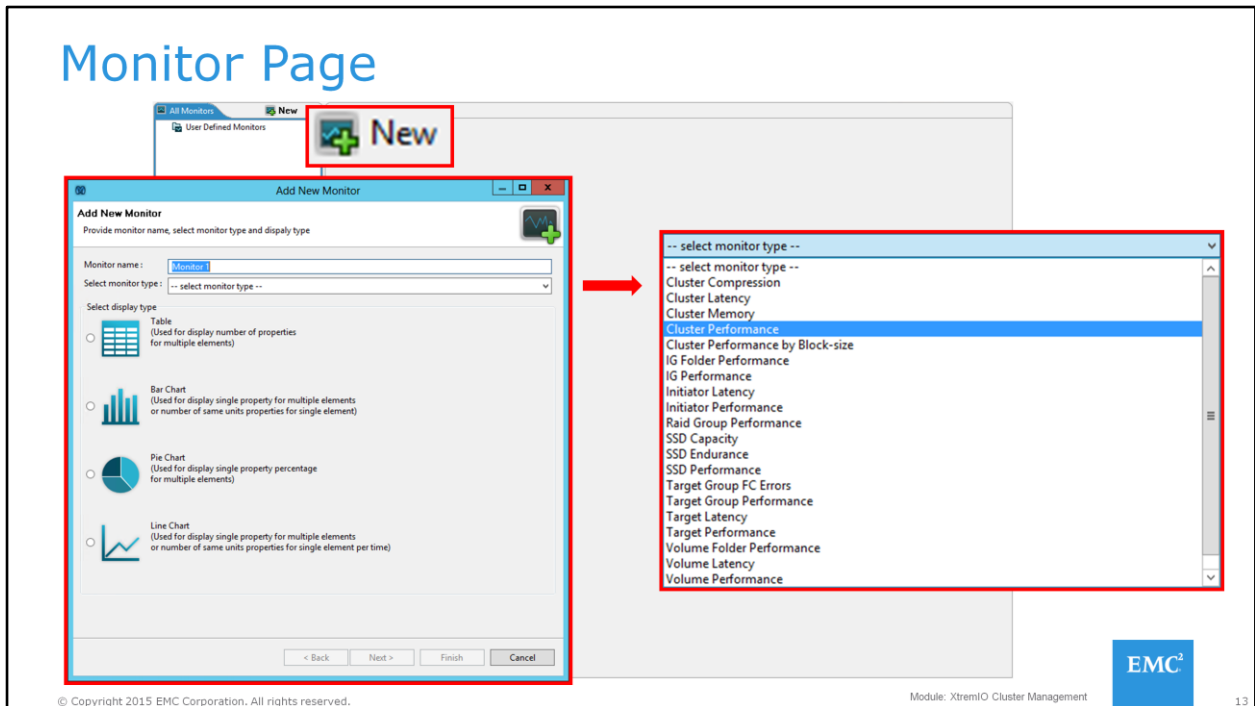
The alerts listing can be sorted by clicking the headers for each column. Specific alerts can be searched for by using the search window. The icons in the top right of the Alerts pane can also be used to acknowledge alerts or to select definitions to be edited.

Events and Alerts Page - Events

The screenshot displays the 'Events' tab in the XtremIO Cluster Management interface. The top section contains filter controls: severity (Critical, Major, Minor, Information), date/time ranges (From Date/Time: 10/30/2014 2:15 PM, To Date/Time: 10/30/2014 2:15 PM), and categories (Software, Hardware, Audit Log, Activity, Security, State Change, Life Cycle). Below the filters is a table of events. The 'Entity' column has a drop-down menu open, showing a list of entities including Alert, Alert Definition, Battery Backup Unit, Cluster, DAE, DAE Controller, DAE PSU, Data Protection Group, Infiniband Switch, Infiniband Switch PSU, Initiator, Initiator Group, Initiator Group Folder, Iscsi Route, Local Disk, Module, SSD, Storage Controller, Storage Controller PSU, and Target. The bottom of the page shows pagination (Page 1 of 8) and the EMC logo.

The Events pane is used for viewing cluster events. The top part of the pane is used to select different criteria for filtering the events to view, such as severity, date, time, category, and entity. The Entity field has a drop-down list that provides all the entities that can be selected.

Monitor Page



To access the Monitor Workspace, select the Monitor icon in the Menu Bar. The left pane contains any User Defined Monitors. The right pane displays the output from the user defined monitors in separate tabs.

To create a new monitor, click the New icon at the top of the User Defined Monitor pane. In the Add New Monitor page, select the display type: Table, Bar Chart, Pie Chart, or Line Chart. Next, from the drop-down window for Monitor Type, select one of the listed objects to monitor, and give the monitor a name.

Administration Page

The screenshot displays the Administration Page of the XtremIO Cluster Management interface. The top navigation bar includes tabs for Cluster, Security, Notification, and CLI Terminal. The Cluster tab is active, showing sub-tabs for iSCSI Network Configuration, iSCSI Ports Configuration, Cluster Encryption, and Cluster Configuration. Three red boxes highlight specific configuration areas:

- iSCSI Network Configuration**: Includes iSCSI Ports Configuration, Cluster Encryption, and Cluster Configuration.
- Users Administration**: Includes LDAPs Configuration and iSCSI Security Configuration.
- Email Configuration**: Includes SNMP Configuration and Syslog Configuration.

The main content area shows the iSCSI Network Configuration table with columns for Target Port, Port Address, IP/Subnet, Port, and VLAN. Below this, there is a table for Destination Subnet and Gateway IP Address. The EMC logo is visible in the bottom right corner.

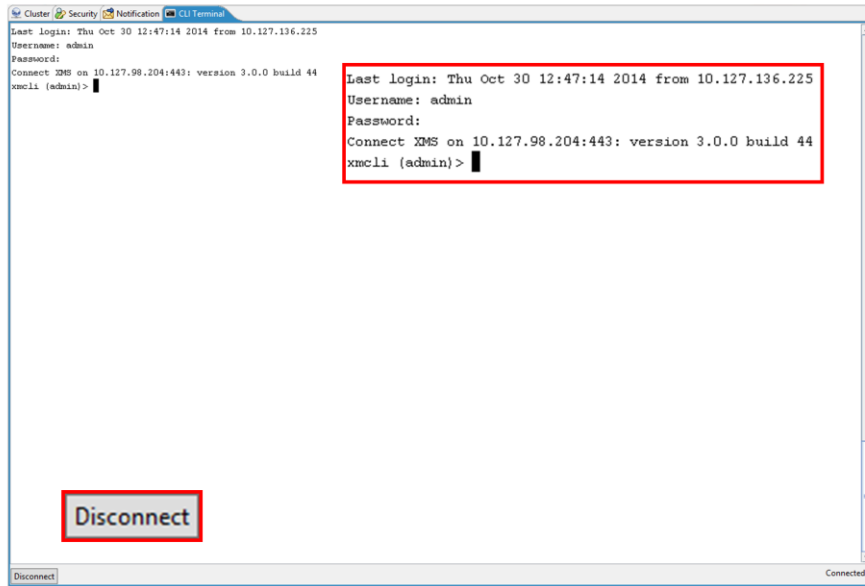
The Administration Workspace can be accessed by clicking the Administration icon in the Menu Bar at the top. This page has four tabs, three of which are discussed on this slide: Cluster, Security, and Notification.

The Cluster tab allows the configuration of iSCSI ports and iSCSI network settings, configuration of cluster encryption, and configuration of the cluster. The latter is currently limited to configuring the VAAI TP limit.

The Security tab can be used to perform several administration functions. The User Administration option allows the addition, removal or modification of local users. The LDAPs Configuration tab allows the addition, removal, or modification of Domain users, and the Notification tab allows the configuration of email, SNMP, and syslog notification.

The fourth tab, CLI Terminal, is discussed on the next slide.

Administration Page – CLI Terminal



The last management option in the Administration Workspace is the CLI Terminal. From here, users can run commands to manage the cluster by either viewing configurations or setting them. When users open the CLI Terminal, they will immediately be logged in to it as the user with which they logged in to the GUI.

Click the Disconnect button to log out. The Connect button will then appear to allow reconnection at a later time.

Lesson Summary

During this lesson the following topics were covered:

- XMS GUI prerequisites
- XMS navigation



This lesson covered management of an XtremIO storage array using the XMS GUI interface. You should have learnt which prerequisites are needed in order to access the XtremIO Management Server. We also reviewed how to access the GUI, along with an overview of the GUI and key menu items within the GUI. You also saw key XMS navigation pages, such as the Dashboard, Configuration, Hardware, Events and Alerts, Monitor, and the Administration pages.

Lab Introduction: XMS Management with the GUI

This lab covers the use of the GUI to manage an XtremIO cluster. A number of commands are introduced, and the command output is used to document the system configuration.



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

17

In order to interact with the XtremIO Graphical User Interface an operator must have situational awareness. This lab will walk you through the tabs provided by the interface and some sub-components. This is an exploratory exercise with little configuration, where you will record important information. Future lab exercises will require you to navigate to the locations experienced in this lab.

Lab Debrief

This lab covered the use of the XMS GUI:

- Explore the tabs and sub-components
- Record important information



This lab covered the use of the XMS GUI for XtremIO cluster management.

Lesson: XMS CLI Management

This lesson covers the following topics:

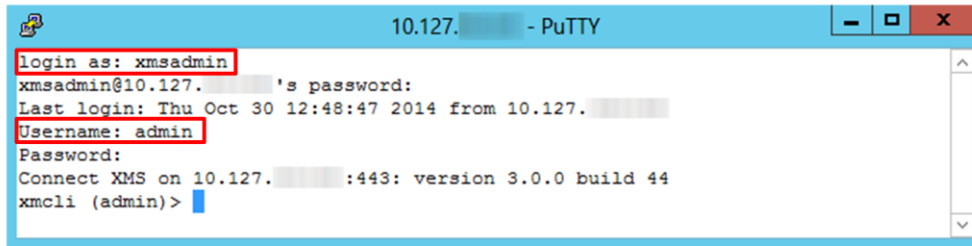
- XMS CLI client prerequisites
- XMS CLI navigation



This lesson covers XtremIO CLI management.

Access the XMS with CLI

- Use SSH programs like PuTTY
- Login process:
 - Log in as: `xmsadmin` and required password
 - This is the user that starts the XMCLI shell
 - Provide a valid user account ID and password
 - This determines the level of access you will receive



```
10.127.1.1 - PuTTY
login as: xmsadmin
xmsadmin@10.127.1.1's password:
Last login: Thu Oct 30 12:48:47 2014 from 10.127.1.1
Username: admin
Password:
Connect XMS on 10.127.1.1:443: version 3.0.0 build 44
xmcli (admin)>
```



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

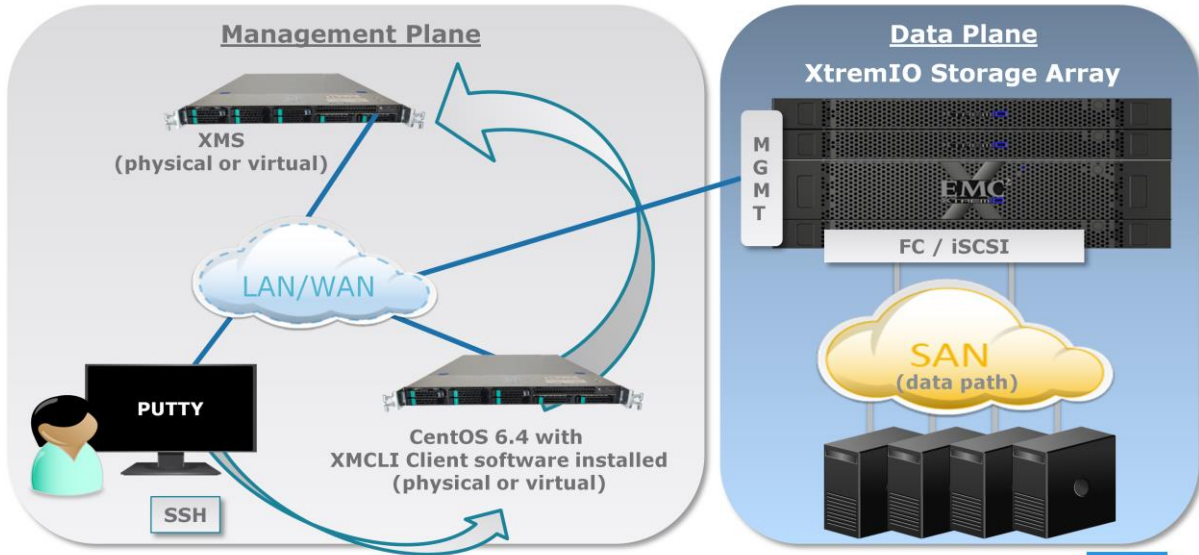
20

Connect to the XMS with an SSH tool such as PuTTY.

- Provide the username `xmsadmin` and the required password
 - This is the user that starts the XMCLI shell
- Provide a valid user account ID and password
 - This determines the level of access you will receive

The CLI allows administrators and other cluster users to perform supported management operations.

XMCLI Client on CentOS 6.4 Server



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

EMC²

21

An XMCLI client package is also available. It needs to be installed on a physical or virtual CentOS 6.4 64-bit server. The XMCLI client communicates with the XMS server via standard TCP/IP connections.

All functionalities that are supported via the XMS GUI console are supported via the XMS client. The XMCLI client can be used to execute CLI commands, using external user scripts.

Detailed information on the CentOS system requirements, and on using the XMCLI client, may be found in the ***XtremIO Storage Array User Guide***, 302-001-288.

Categories of XMCLI Commands

Command – Main Categories	Description
> show	Show cluster objects and configuration
> add	Adds new cluster objects
> modify	Modify an existing object
> remove	Delete an object from the cluster
Other	Cluster maintenance, call-home commands



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

22

The CLI commands are divided into five main categories as follows:

- **Show** commands - Show cluster objects and configuration.
- **Add** Commands - Add new system objects.
- **Modify** Commands - Modify an existing object.
- **Remove** Commands - Delete an object from the cluster.
- **Other** Commands - System maintenance, Call-Home etc.

Displayed data is removed upon exiting the CLI terminal and does not appear when the terminal is reopened.

Use the latest version of the XtremIO Users Guide for the information on using CLI commands with this version of the software.

XMCLI Autocomplete

Use the `Tab` key on the keyboard when entering commands

`Tab`

Completes command

> add-volume

`Tab`

Completes command parameter

> add-volume p

`Tab`

`Tab`

Lists commands or parameters

> add-

```
xmcli (admin)> add-
add-event-handler-definition  add-initiator-group  add-ssd
add-folder                    add-iscsi-portal    add-user-account
add-initiator                 add-iscsi-route   add-volume
xmcli (admin)> add-
```

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

23

When entering commands, the `Tab` key on your keyboard is helpful.

Start entering a command then press the `Tab` key. This automatically fills in the rest of the command, if it is unambiguous. If more than one command has the same spelling, continue manually to enter the command, and then press the `Tab` key once again. Once the command / part becomes unambiguous, press the `Tab` key to complete the command entry.

In the same way, the `Tab` key also helps when entering command parameters. When you start entering a parameter, the `Tab` key will complete the parameter entry, if unique.

Using the `Tab` key twice displays all the XMCLI commands, similar to the `Help` command. When a partial command is entered, press the `Tab` key twice to display all possible commands starting with the entered string. For example, enter `add-` and then press the `Tab` key twice. This displays all the Add commands. This can also be used when entering parameters.

help Command in XMCLI

- > help

```
xmcli (admin)> help
Documented commands (type help <topic>):
=====
acknowledge-alert          show-xenfs
activate-storage-controller show-xms
add-event-handler-definition shutdown
add-folder                shutdown-xms
add-initiator             start-cluster
add-initiator-group       stop-cluster
add-iscsi-portal          stop-cluster-unorderly
add-iscsi-route           unmap-lun
add-ldap-config
```

- > help <command>

- > help add-volume

```
xmcli (admin)> help add-volume
Usage: add-volume property=value list
PROPERTY                                MANDATORY    DESCRIPTION                                VALUE
-----
alignment-offset                       No           Alignment Offset                          range 0-15
cluster-id                             No           Cluster ID                               name or index
lb-size                                No           LB Size                                  $12 or 4096
parent-folder-id                       No           Parent folder ID                         name or index
small-io-alerts                        No           Small IO Alerts                          enabled or disabled
unaligned-io-alerts                   No           Unaligned IO Alerts                      enabled or disabled
vaai-tp-alerts                        No           VAAI TP Alerts                          enabled or disabled
vol-name                               No           Volume Name                              string
vol-size                               Yes          Volume Size                              integer suffixed by [mgtp]
```

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

24

The `help` command displays the list of all available CLI commands, depending on the user role. The `help` command can also be used to display parameters when specifying a single command. This example shows **help add-volume**.

With this command, a list of parameters is displayed. These parameters might be mandatory, as in this example of `vol-size` (volume size). Others might not be mandatory, but can be used, if, for example, you want to set a `vol-name` (volume name) other than the cluster's created default. A short description for each parameter and the value type and range are also provided.

show-clusters-info

- XtremIO SW version
- PSNT

The `show-clusters-info` command displays general cluster information useful to identify the version of the cluster.

```
xmcli (admin)> show-clusters-info
Cluster-Name Index State Conn-State Activation-Time Start-Time SW-Version
xbc1 1 active connected Mon Nov 3 06:01:35 2014 Mon Nov 3 06:01:35 2014 3.0.0-44

PSNT Encryption-Mode Encryption-Supported Mode-Switch-Status Compression-Mode
APM0014xxxxxxx disabled No none enabled

SSH-Firewall-Mode
Unlocked

xmcli (admin)>
```

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

25

The `show-clusters-info` command displays general cluster information similar to the `show-cluster` command. In addition, you can find the software version and build information of the cluster and product serial number tag information, as well as the cluster activation date/time and last start-time in the output.

show-storage-controllers

The show-storage-controllers command displays the cluster's Storage Controller information and status.

```
xmcli (admin)> show-storage-controllers
```

Storage-Controller-Name	Index	Mgr-Addr	IB-Addr-1	IB-Addr-2	IPMI-Addr	Brick-Name
X1-SC1	1	10.127.xx.yy	169.254.0.1	169.254.0.2	10.127.xx.yy	X1
X1-SC2	2	10.127.xx.yy	169.254.0.17	169.254.0.18	10.127.xx.yy	X1

Index	Cluster-Name	Index	State	Health-State	Enabled-State	Unorderly-Stop-Reason	Conn-State
1	xbc1	1	healthy	healthy	enabled	none	connected
1	xbc1	1	healthy	healthy	enabled	none	connected

IPMI-State	Powered-State
connected	on
connected	on

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

26

The `show-storage-controllers` command displays the information of each SC in the cluster, such as IPs, names, connectivity state, and status. It is especially useful to get all the IP addresses for all the SCs plus their state (health state) and enabled-state (configuration status).

show-initiators

- Registered initiators
- Do not need to be connected

The `show-initiators` command displays initiator data.

```
xmcli (admin)> show-initiators
Initiator-Name  Index  Port-Type  Port-Address  IG-Name  Index  Parent-Folder-Name  Index
esx154vmhba2   1      fc         10:00:00:00:c9:3d:a6:b6  esx154   1      /                   1
esx154vmhba3   2      fc         10:00:00:00:c9:3d:a6:b7  esx154   1      /                   1

Certainty-State
ok
ok

Chap-Authentication-Initiator-User-Name  Chap-Discovery-Initiator-User-Name  Chap-Authentication-Cluster-User-Name  Chap-Discovery-Cluster-User-Name
```

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

27

The `show-initiators` command displays registered initiator data, which includes the assigned initiator name when it was registered, the Index, Port Type (FC or iSCSI), Port Address (WWPN or IQN), and Initiator Group Name it belongs to. The Certainty-State column displays whether there is a pending command associated with the initiator. CHAP is not enabled in this environment, and as a result there is no output identifying the CHAP configuration.

Initiators can be registered while they are not connected to the cluster.

show-bbus

The `show-bbus` command displays battery backup unit information.

```
xmcli (admin)> show-bbus
```

Name	Index	Serial-Number	Power-Feed	State	Connectivity-State	Input	Battery-Charge	BBU-Load
X1-BBU 1		DV0P380AW	feed_a	healthy	connected	on	100	22
X2-BBU 2		DV0P500EE	feed_b	healthy	connected	on	100	9

Voltage	FW-Version	Part-Number	Brick-Name	Index	Cluster-Name	Index
214	9901DC	078-000-114-00	X1	1	xbc1	1
213	9901DC	078-000-114-00	X1	1	xbc1	1

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

28

The `show-bbus` command displays BBU information. Among other details, this includes the Name in the cluster, State (healthy, disconnected, failed or transient), and Connectivity-State.

Further details in the command output include the Input, which indicates whether there is input power to the supply, the Battery-Charge level, and BBU-Load (current load). The Part-Number is a string identifier of the part and is assigned by EMC.

show-discovered-initiators-connectivity

The `show-discovered-initiators-connectivity` command displays undefined initiators and the initiators-targets connectivity map.

```
xmcli (admin)> show-discovered-initiators-connectivity
Discovered Initiator List:
Port-Type      Port-Address
xmcli (admin)>
```

Num-Of-Conn-Targets

**No undefined
initiators found!**

```
xmcli (admin)> show-discovered-initiators-connectivity
Discovered Initiator List:
Port-Type      Port-Address
fc             10:00:00:90:fa:1c:e9:fb
xmcli (admin)>
```

Num-Of-Conn-Targets
2

**Undefined
initiator on
two ports**

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

29

The `show-discovered-initiators-connectivity` command displays discovered initiators that are visible by SC ports (FC or iSCSI) but are not members of an initiator group. For example, they have been FC zoned to the XtremIO storage array ports, but are not defined on the array.

This command can be useful when trying to understand who is zoned to the cluster. In the first example, no undefined initiators were found. All initiators that can reach the cluster have been added to initiator groups.

The second example shows one undefined FC initiator, its Worldwide Name, and the number of targets it can reach (2). To determine the detail of target ports it can reach, use the `target-details` parameter. The output then lists the target names in the cluster and their index, as well as the previous information.

This command is useful when setting up, verifying, and/or troubleshooting a cluster and connectivity issues.

show-targets

The `show-targets` command displays the cluster target information.

```
xmcli (admin)> show-targets
```

Name	Index	Port-Type	Port-Address	Port-State	Health-Level	Storage-Controller-Name
X1-SC1-fc1	1	fc	21:00:00:24:ff:57:a8:e6	up	level_1_clear	X1-SC1
X1-SC1-fc2	2	fc	21:00:00:24:ff:57:a8:e7	up	level_1_clear	X1-SC1
X1-SC1-iscsi1	5	iscsi	iqn.2008-05.com.xtremio:001e6793a7d4	down	level_1_clear	X1-SC1
X1-SC1-iscsi2	6	iscsi	iqn.2008-05.com.xtremio:001e6793a7d5	down	level_1_clear	X1-SC1

Index	TG-Name	Index	MTU	Jumbo-Frames	Certainty-State
1	Default	1	1500	False	ok
1	Default	1	1500	False	ok
1	Default	1	1500	False	ok
1	Default	1	1500	False	ok

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

30

The `show-targets` command displays the cluster target information. The port names (FC and iSCSI) are listed with their index, port type, and port address (WWPN or IQN). The port state could show up or down [two of each are shown in the example]. The health level in the example is shown as `level_1_clear`, which indicates that diagnostics performed on the target port have shown it to be healthy.

The port addresses are useful for zoning/networking troubleshooting. The port state and health level information could indicate failed components resulting in performance degradation.

Lesson Summary

During this lesson the following topics were covered:

- XMS CLI client prerequisites
- XMS CLI navigation



This lesson covered management of an XtremIO storage array using the XMS CLI. We reviewed the XMCLI commands that are divided into five main categories: Show, Add, Modify, Remove and Other Commands such as System maintenance, Call-Home etcetera. We also provided details on the XMS CLI navigation such as: **help**, **show-clusters-info**, **show-storage-controllers**, **show-initiators**, **show-bbus**, **show-discovered-initiators-connectivity** and **show-targets**.

This slide intentionally left blank.



Lab Introduction: XMS Management with the CLI

This lab covers the use of the CLI to manage an XtremIO cluster. A number of commands are introduced, and the command output is used to document the system configuration.



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

33

The XtremIO cluster can be managed by the command line interface. The lab will introduce you to some basic commands in which you can obtain information and configure the XMS.

Lab Debrief

This lab covered the use of xmcli:

- Connect to XMS via PuTTY
- Use various commands and record important information



This lab covered the use of xmcli for XtremIO cluster management.

Lesson: XMS User Management

This lesson covers the following topics:

- XMS users and authentication
- XMS roles
- XMS account creation



This lesson covers XMS user management and provides an overview of the XMS users and authentication, XMS roles and XMS account creation.

XMS User Roles

- Technician
 - Perform maintenance tasks, clusters initialization, create other users
 - For support personnel only
- Administrator
 - All general administration functions
 - Create volumes, initiator groups, alerts, etc.
 - Creates other admin and user accounts
- Configuration
 - Used for general storage configuration
 - Cannot create other users
 - No destructive actions (powering off)
- Read Only
 - Enables user to view configuration and performance data



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

36

The Technician role is for support personnel only. It is the most powerful user class. It should only be used when specific instructions require it.

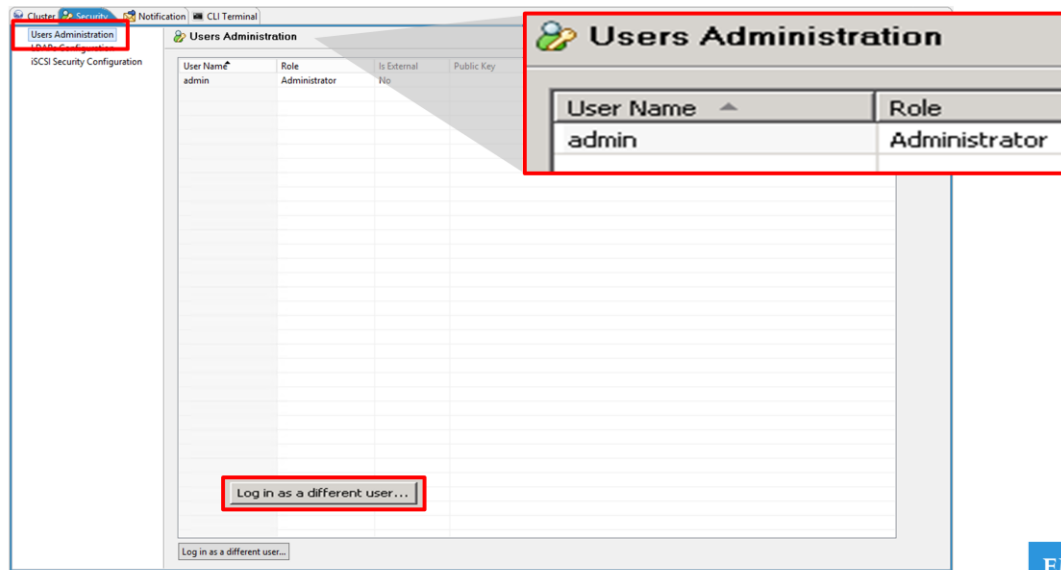
The Administrator role is used to perform all general administrative functions, including creating other users with the Administrator Role. This user can perform all functions needed to successfully manage the cluster such as creating LUNs, mapping initiators, and monitoring performance and alerts.

The Configuration role can perform most cluster activities other than creating other users, acknowledging alerts, and destructive activities like powering off the cluster.

The Read-Only role allows users to monitor the activities of the cluster, but not make any changes.

When the XMS is installed, two user accounts are pre-installed on it. The Tech and Admin users are from the Technician and Administrator role respectively. Configuration and Read Only users are not present by default, and must be created.

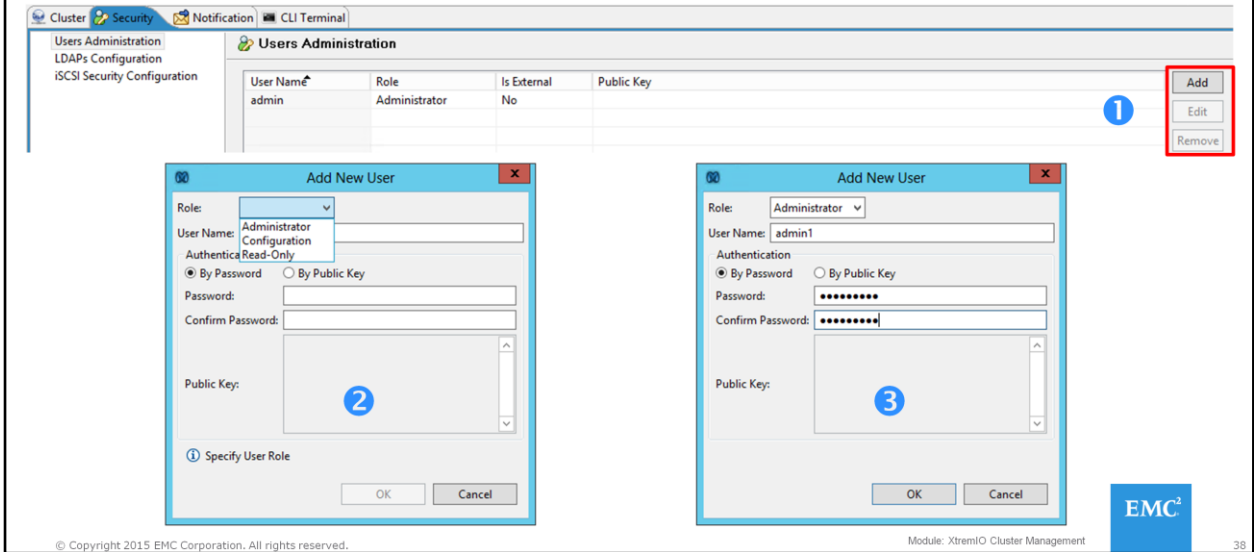
XMS GUI User Management



Users can be created in the administration workspace which can be accessed by clicking the **Administration** icon in the Menu Bar.

Shown here is the **Users Administration** pane, which is the default pane when this page is first opened. Once other users have been created, you can use the **Log in as a different user** button to sign in as another user.

Creating an XMS GUI User



To create a user perform the following:

- Click *Add*
- Select the proper role
- Choose a username
 - Spaces are not permitted
- Choose a password
 - Confirm password

Configuring LDAP Authentication

The screenshot displays the XtremlIO Cluster Management interface for configuring LDAP authentication. The main window shows the 'LDAPs Configuration' tab with a table for managing LDAP entries. The 'Add' button in the top right of the table is highlighted. Below the table, the 'Add New LDAP Configuration' dialog is open, showing fields for Bind DN, Bind password, Search base, Search filter, User to DN rule, Timeout in msec, Cache expiration in hours, and Server URLs. The 'Add' button in this dialog is also highlighted. To the right of the dialog, the 'Add Server URL' dialog is shown, with a field for Server URL and an 'Add' button. Below that, the 'Add Role' dialog is shown, with a dropdown for Role (set to 'Administrator Configuration Read-Only') and an 'Add' button. The EMC logo is in the bottom right corner.

XtremlIO supports Lightweight Access Protocol (LDAP) user authentication. Once configured for LDAP authentication, the XMS redirects user authentication to the configured Active Directory (AD) servers and allows access to authenticated users only. User XMS permissions are defined based on a mapping between the user AD groups and XMS roles.

To configure LDAP authentication perform the following:

- Click Add
- Enter the appropriate Active Domain information
- Enter the AD server URL
- Select the Active Directory group and Role

CLI User Account Manipulation

Command	Description
<code>add-user-account</code>	Adds a new user account.
<code>remove-user-account</code>	Removes a user account.
<code>modify-user-account</code>	Modifies user account parameters.
<code>modify-password</code>	Used to modify one's own password, or for entitled users (Configuration and Admin) to modify others' passwords.
<code>show-user-accounts</code>	Displays user accounts information.
<code>show-ldap-configs</code>	Displays the LDAP users' authentication configuration data.
<code>modify-ldap-config</code>	Adds or modifies the LDAP user authentication.



Here we see an example of the CLI commands used to create user accounts and configure LDAP authentication.

Create Audit Log

The screenshot displays the 'Alerts & Events' workspace. The 'Events' tab is active, showing a table of events. The 'Event Properties' dialog is open for event ID 566, which is an 'Information' severity event with code 50001, occurring on Nov 4, 2014 at 5:24:20 AM. The event is categorized as 'Audit' and involves the 'User' entity. The description states: 'User: admin, Command: modify_password, Arguments: [password=xxxxxxx, usr_id='']'.

Severity	Event Code	Date and Time	Category	Entity	Entity Details	Description
Information	50001	Nov 4, 2014 5:24:20 AM	Audit	User		User: admin, Command: modify_password, Arguments: [password=xxxxxxx, usr_id='']
Information	50001	Nov 4, 2014 3:48:10 AM	Audit	User		User: admin, Command: discover_initiator_connectivity, Arguments: []
Information	50001	Nov 4, 2014 3:48:10 AM	Audit	User		User: admin, Command: maplun, Arguments: [vol_obj_id=['e4b9083fc73a437faaf787624085...], sys_obj_id=['0b10d6cd7e1b42d081e853a...]]
Information	50001	Nov 4, 2014 3:48:10 AM	Audit	User		User: admin, Command: addvolume, Arguments: [sys_obj_id=['0b10d6cd7e1b42d081e853a...]]

A cluster Audit Log can be created in the Alerts & Events Workspace under the Events tab. Check the **Audit Log** option in the **Category** area. Events may be filtered by severity if the checkbox for the appropriate severity level is checked. Every activity and configuration change on the cluster is recorded in an Audit Log. The Audit Log is stored on the XMS and is maintained for 90 days. The Audit Log is cyclic and has a limited file size. When the file size exceeds its allocated disk space, older entries are overwritten. The most recent recorded audit entry, expanded by double-clicking to generate the dialog at bottom left, shows that the user admin ran the **modify_password** command.

Lesson Summary

During this lesson the following topics were covered:

- XMS users and authentication
- XMS roles
- XMS account creation



This lesson covered the XMS users and authentication, XMS roles, such as: Technician, Administrator, Configuration and Read Only; along with detailed steps required to create an XMS account as well as configuring LDAP authentication.

Lab Introduction: XMS User Account Management

This lab covers user account management with the GUI and the CLI



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Cluster Management

43

User account management is a key task for any system administrator. This lab demonstrates user account management via the GUI and CLI by creating different users, assigning them different roles, and verifying their privileges.

Lab Debrief

This lab covered user account management:

- Create different users
- Assign them different roles
- Verify their privileges



This lab covered XtremIO user account management.

Module: Summary

Key points covered in this module:

- Interacting with the XMS GUI interface
- Interacting the XMS CLI interface
- Performing user management



This module covered the XMS GUI and CLI interface as well as XMS user management.

This slide intentionally left blank.



Module: XtremIO Host Configuration

Upon completion of this module, you should be able to:

- Configure volumes, initiators, and volume mapping
- Identify host connectivity requirements for:
 - VMware
 - Windows
 - Linux
 - Solaris
 - AIX



This module covers the XtremIO provisioning process. During this module we'll cover how to configure volumes, initiators, and you'll see how volume mapping works. We'll also identify host connectivity requirements for:

- VMware
- Windows
- Linux
- Solaris
- AIX

Lesson: Volume Creation and Sizing

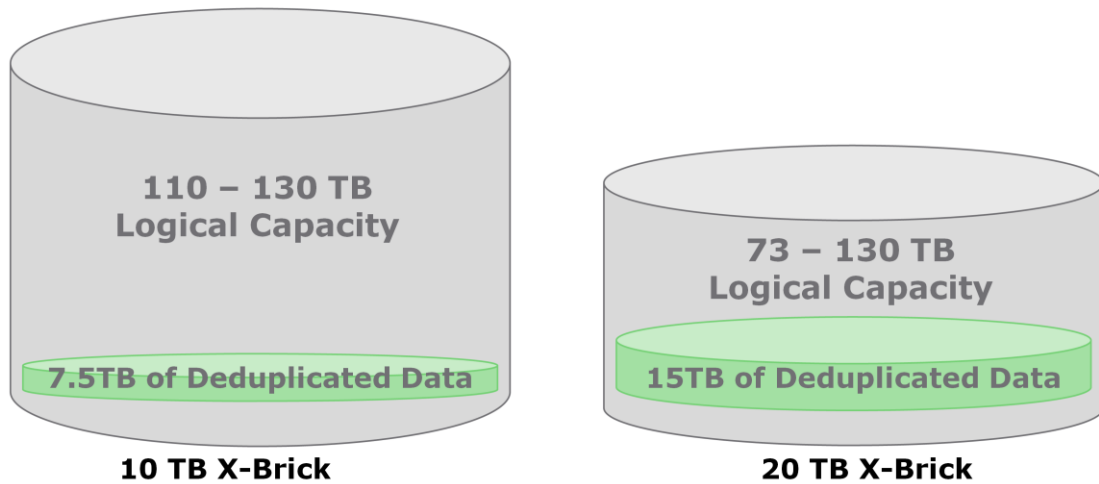
This lesson covers the following topics:

- Volume sizing
- Volume creation
- Volume provisioning



This lesson covers volume sizing, creation and provisioning.

Sizing XtremIO for Capacity



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

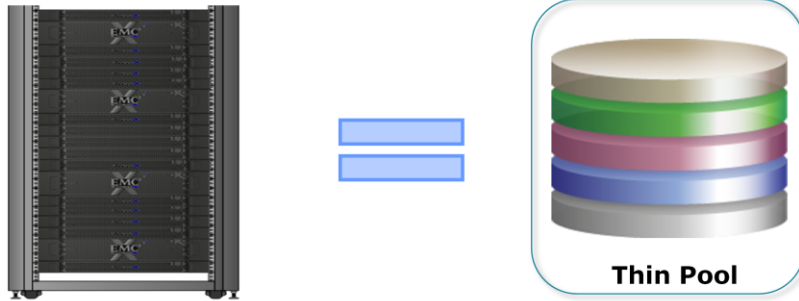
Module: XtremIO Host Configuration

3

Sizing XtremIO arrays for capacity is relatively simple. Each 10 TB X-Brick (containing 400 GB SSDs) can hold up to 7.5TB of deduplicated data. That means if the deduplication ratio is 5:1, each 10 TB X-Brick can hold up to 37.5TB of user data. The logical capacity per 10 TB X-Brick is 110-130 TB. For 20 TB X-Bricks (containing 800 GB SSDs), up to 15 TB of deduplicated data can be stored, and the logical capacity is 73-130 TB. This value is different to what one might expect [double the 10 TB X-Brick numbers] because the Storage Controllers hold more SSD [physical storage] metadata in memory, and therefore have less space for logical storage metadata. The 20 TB X-Bricks enable more physical capacity for applications that contain mostly unique data (low deduplication ratios), such as databases.

Sizing Volumes

- Inherently thin provisioned with no performance loss or storage overhead penalty



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

EMC²

4

XtremIO inherently uses thin provisioning. All volumes are thin provisioned. This is not configurable. Unlike traditional methodologies, XtremIO thin provisioning results in no loss of performance, and requires no capacity overhead.

Scalability

Definition	Max Value (10 TB)	Max Value (20 TB)
Maximum number of X-Bricks per cluster	4	6
Number of Storage Controllers per X-Brick	2	2
Number of management stations per cluster	1	1
Number of Clusters per XMS	1	1
Number of initiators defined per cluster	512	512
Number of initiators defined per initiator group	64	64
Number of initiator groups defined per cluster	512	512
Number of volumes defined per cluster	8192	8192
Volume size (per cluster)	256 TB	256 TB
Logical capacity (per X-Brick)	130 TB	130 TB
Number of initiator groups mapped per volume	64	64
Number of iSCSI portals per cluster	16	24

Refer to the XtremIO Release Notes for the most up-to-date information



These are the scalability numbers for XtremIO. Refer to the latest Release Notes for the most up-to-date information.

Note that the starter 10 TB X-Brick may only be used in a single X-Brick cluster.

Administration: iSCSI Ports

The screenshot displays the 'iSCSI Network Configuration' window. At the top is the 'iSCSI Ports Table' with columns: Target Port, Port Address, IP/Subnet, Port, and VLAN. The first row shows 'X1-SC1-iscsi1' with Port Address 'iqn.2008-05.com.xtremio:001e675b9538' and IP/Subnet '10.0.0.1/24'. To the right of the table are 'Add', 'Edit', and 'Remove' buttons. A callout box labeled 'Expanding entry displays network info' points to the first row. Another callout box labeled 'Add iSCSI Portal' points to the 'Add' button. Below the table is the 'Add iSCSI Portal' dialog box, which has fields for 'Target Port' (set to 'X1-SC1-iscsi1'), 'IP Address/Subnet bits' (set to '10.0.0.1/24'), 'Port' (set to '3260'), and 'VLAN' (set to '0'). Below these fields are 'OK' and 'Cancel' buttons. To the right of the dialog is an 'Error' dialog box with a red 'X' icon and the text: 'Operation Failed. Failed to add iSCSI portal 10.127.98.100/24. Reason: invalid_subnet'. A red box highlights the error message with the text 'Must be different subnet than Management IP'.

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

EMC²

6

The iSCSI section of the Administration workspace is broken into two panes; the top pane, shown here, is for configuring iSCSI portals; the bottom pane shows iSCSI routes. To add an iSCSI port configuration, click the Add icon at the top right of the pane to open the Add iSCSI Portal window. In the Target Port drop-down list, you will find all the iSCSI ports installed on the cluster. Select the port to configure. You will need to enter the IP address for the port in the format IP address/Subnet bits, for example 10.0.0.1/24.

The Port and VLAN sections are grayed out, as they only use the default values of 3260 for the port and VLAN tagging is not supported at this time. If you try to enter an IP address which is on the same Subnet as the Management Port, the cluster will return an error. The iSCSI ports must be on a separate Subnet than the cluster Management ports. Once the correct values have been entered, the OK icon will become active. Shown here is the new iSCSI Portal configuration. Notice that the IQN value can be found here. If the listing for the portal is expanded, users will be able to see the IP address, VLAN, and port entries.

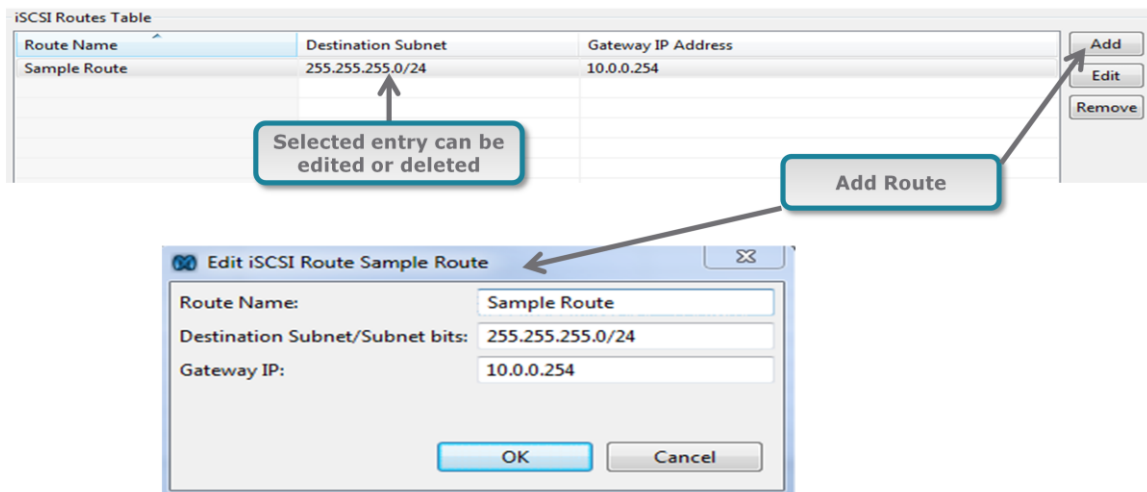
XMCLI commands which can be used for iSCSI when logged into the XMS server are:

`add-iscsi-portal`: Maps a portal to a target.

`remove-iscsi-portal`: Deletes a portal mapping from a target.

`show-iscsi-portals`: Displays a list of iSCSI portals and their properties.

Administration: iSCSI Management Routes



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

7

In order for the iSCSI targets to be reached by initiators on other networks, it may require the use of static routes or an IP router. To add iSCSI routes to the cluster, click the Add icon in the top right of the iSCSI Routes Table pane. Enter a name for the route, such as the network location and enter the netmask of the remote network in the Destination Subnet/Subnet bits format, as in the iSCSI portal configuration. Then, click OK. The new route will now appear in the iSCSI Routes Table pane. If a route is selected, it can be edited or removed using the buttons on the right.

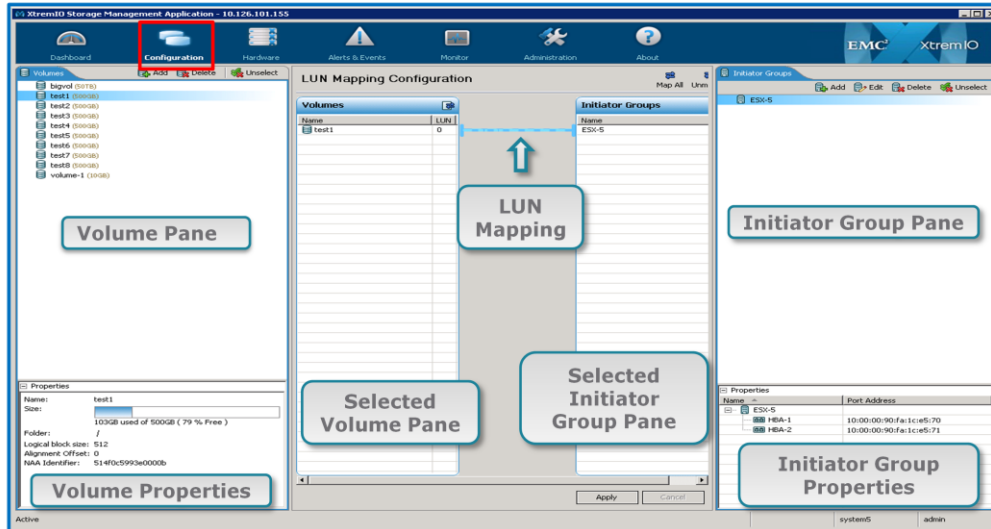
Here are some XMCLI commands which can be used for iSCSI routes when logged in to the XMS server:

`add-iscsi-route` - Adds and configures iSCSI route parameters.

`remove-iscsi-route` - Deletes an iSCSI routing configuration.

`show-iscsi-routes` - Displays a list of iSCSI routes and their properties.

Configuration Page – Overview



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

8

Shown here is the Configuration Workspace of the XtremIO Management GUI. Let's review what each section is used for.

On the left, is the Volume pane. The Volume pane is used to add, delete, and manage volumes. If a volume is selected in this pane, the properties of the volume will be displayed in the Volume Properties pane below. On the far right is the Initiator Group pane. This pane is used to add, delete, edit, and manage Initiator Groups which are groupings of WWNs or IQNs from a host, used to access the LUNs. If an initiator is selected here, the properties of the group will be displayed below, including HBA WWNs and iSCSI IQNs.

The pane in the center left is the Selected Volumes pane. Any volumes that are selected in the Volume pane will be displayed here so that they can be mapped/unmapped. In the center right, is the Selected Initiator Group pane. As with volumes, any initiators selected in the Initiator Group pane will be displayed here to be mapped/unmapped from the selected volumes.

The last item of note in the Configuration Workspace is the LUN Mapping Connector. This connector shows that the LUNs and initiators have been mapped together.

Volume Creation

Add Volume

The screenshot illustrates the XtremlIO Host Configuration interface for volume creation. On the left, the 'Volumes' pane lists 10 sample volumes (Sample1 to Sample10), each 50GB. The 'Add' button is circled in red. The 'Add New Volumes' dialog is open, showing a table of 10 volumes with columns for Name, Size, Type, Small IO Alerts, Unaligned IO Alerts, and VAAI TP Alerts. The 'Add Multiple' button is also circled in red. Below the dialog, the 'Add Multiple Volumes' sub-dialog is shown with fields for Number of Volumes (10), Name (Sample), Size (50G), Type (Normal (512 LBs)), and checkboxes for Small IO Alerts, Unaligned IO Alerts, and VAAI TP Alerts. The 'Properties' pane at the bottom shows details for Sample1, including NAA Identifier, Creation Time, and various alert settings.

This slide illustrates how to create a volume from the Configuration Workspace. The first step is to click the Add icon at the top of the Volume pane. This will open the Add New Volume window. To add a single volume, simply enter a volume name in the Name field; the volume size, such as 4 MB, 5 GB, or 6 TB in the Size field; and the type of volume, such as Normal (512 LBs), 4KB LBs or Legacy Windows (offset: 63) from the drop-down list in the Volume Type field. LBs are the Logical Blocks used to create the volumes. There is an option to enable certain types of alerts on a per-volume base. The alerts are small IOs, which are IOs less than 4 kB, Unaligned IOs, and VAAI TP alerts. VAAI TP alerts are used to notify when thin provisioned storage capacity reaches a set limit.

Note: Default volume size value is GB if none is entered.

To create multiple volumes at once, click the Add Multiples icon in the top right of the Add New Volume page. The Add Multiple Volumes page allows users to set the number of volumes to create the name to use for the series of volumes, along with the starting index number, size of the volumes, and the type of volumes to create. Once the information has been entered and the OK icon selected, the information will populate the Add New Volume page.

The next decision is whether to place the volumes in a folder. If the Next icon is selected, users will be taken to a Create Folder page where they can create a new folder in which to place the volumes. This feature allows users to act on multiple volumes at once, rather than having to select each one individually. To bypass using folders, users can click the Finish icon as the final step. Once the Finish icon has been selected, the new volumes will appear in the Volumes pane. If any of the volumes are selected, their properties will be displayed in the Properties pane at the bottom of the Volumes pane.

Some of the XMCLI commands that can be used when logged in to the XMS for volumes are:

- show-volume: Displays the specified volume information.
- show-volumes: Displays a list of volumes and their information.
- show-volume-folders: Displays the specified folder information.

Initiator Group Creation

Add Initiator

Folders View

Now that volumes have been created, the next step is to create Initiator Groups, which are groupings of host FC WWNs, or iSCSI IQNs, to which to connect the volumes. Initiators from one host or several hosts may be in the same Initiator Group. Just as for volumes, the first step is to click the Add icon at the top of the Initiator Groups pane. When the Add Initiator Group page opens, enter a name for the new group in the Initiator Group Name field. Next, click the Add button to open the Add Initiator screen. Enter a name for the new initiator in the Initiator Name field, and then from the drop-down list in the Initiator Port Address field, select one of the registered initiators. This drop-down list will only display WWNs and IQNs which have already been presented to the cluster and that have successfully logged in. If none are found, they can be manually entered. Initiators can only belong to one Initiator Group at a time. Once the OK icon has been selected on the Add Initiator screen, the new initiators will appear in the Add Initiator Group screen.

As with the volume creation, users now have the choice of placing the Initiator Group into a folder for managing multiple groups at once, or just placing it in the Initiator Group pane. If folders are used, the entry in the Initiator Group pane will be displayed. If folders are not used, the new group will be displayed at the top level. If an Initiator Group is selected in the pane, its properties will be displayed in the Initiator Properties pane at the bottom of the window.

XMCLI commands that may be useful in managing Initiator Groups when logged in to the XMS server are:

show-initiators - Displays initiator data.

show-initiator-group - Displays information for a specific initiator group.

show-initiator-groups - Displays information for all initiator groups.

show-initiators-connectivity - Displays initiators-port connectivity and the number of connected targets.

show-discovered-initiators-connectivity - Displays the initiators-targets connectivity map.

LUN Mapping

The screenshot shows the EMC XtremIO LUN Mapping Configuration interface. The interface is divided into several panes:

- Volumes:** A list of volumes (Sample1 to Sample10) with their LUN IDs (10 to 18). An annotation "Select volumes to map" points to the list.
- Initiator Groups:** A list of initiator groups (Sample IG) with their LUN IDs (18). An annotation "Click to show existing mappings" points to the "Show Mapping" icon.
- Properties:** A pane showing properties for the selected volumes and initiator groups. An annotation "Select to change LUN ID" points to the LUN ID column in the Volumes pane.
- Map All:** A button at the top right of the Initiator Groups pane. An annotation "Select Initiator Groups to map" points to the list of initiator groups.

The "Apply" button is highlighted with a red circle at the bottom of the interface.

Now that volumes and initiator groups have been created, users can map them together to allow access by the hosts. To start the process, users need to select one or more volumes in the Volume pane that they want to assign to a host. This can be done one at a time, or if they are in a folder, all can be chosen by right clicking the folder and choosing Select all volumes in folder. Once selected, the volumes will appear in the Selected Volumes pane. Next, in the Initiator Groups pane, an Initiator Group needs to be selected. One IG or multiples may be selected using the same process as with volumes and folders. The selected initiator will now appear in the Selected Initiator Group window.

At the top of the workspace users will now click on the Map All icon. This will cause a connection to appear between the volumes and the initiator showing the potential configuration. At this time, users can click on the LUN ID number and change it to whatever value is needed to be seen by the host as long as it is not currently in use. If volumes are added at a later time, the system will try filling in missing LUN values first. Once all the settings are configured appropriately, click the Apply icon to finish the mapping process. If at a later time the user needs to see which LUNs are mapped to a specific host, or which host a LUN is mapped to, simply select either one and click the Show Mapping icon, as displayed here.

Configuration

The screenshot displays the 'XtremIO Storage Management Application - 10.126.101.146' interface. The 'Configuration' tab is active, showing the 'LUN Mapping Configuration' workspace. The interface includes a top navigation bar with icons for Dashboard, Configuration, Hardware, Alerts & Events, Monitor, Administration, and About. The main workspace is divided into several sections:

- Volumes:** A list of volumes (system4-01 to system4-14) with a 'Show Volume Mappings' icon circled in red.
- Initiator Groups:** A list of initiator groups (system4-init-group01) with a 'Show Initiator Group Mappings' icon circled in red.
- Properties:** Two property panels at the bottom. The left panel shows details for 'system4-16' (Size: 100GB, 4KB used, 99% Free). The right panel shows details for 'system4-init-group01' (Port Address: 10:00:00:90:fa:1c:e9:fa).
- Connection Map:** A central area showing a connection map between volumes and initiator groups, with a blue dashed line indicating a connection.

The bottom of the interface includes the EMC logo, copyright information (© Copyright 2015 EMC Corporation. All rights reserved.), the module name (Module: XtremIO Host Configuration), and a page number (12).

The Configuration Workspace is used for working with Volumes, Initiator Groups, and LUN Mapping Configurations. It is possible to display the LUN/Initiator Group relationship. Using the Show Volume Mappings icon for a selected LUN/multiple LUNs displays the mapped IG. Using the Show Initiator Group Mappings icon displays all the LUNs that are mapped to the selected IG.

The connection map shows the current connections between initiators and targets, allowing a better visibility of the cluster configuration and assisting with troubleshooting when zoning or subnet errors occur.

Furthermore, the property information for each selected volume (name, logical block size, alignment offset, and NAA identifier) , as well as the property details for each Initiator Group (name, port names, and WWPN) can be displayed at the bottom of the Configuration Workspace.

Connection Map

Right-click on IG to show map

Display ports with active Initiators

Connections	X1-SC1-fc1	X1-SC1-fc2	X1-SC2-fc1	X1-SC2-fc2
2				
2				

Another useful tool is the Connection Map. It can be accessed by right-clicking an Initiator Group, and from the drop-down menu selecting Show Connection Map. This map will display all the ports on the SCs to which this IG's WWNs or IQNs are connected. This is very useful for checking communication and performance issues if zoning is thought to be a problem.

Host Configuration

- Confirm the host is supported
- Host installed and configured with a supported operating system
- Physically connected to the XtremIO cluster through either FC or iSCSI
 - Proper driver/firmware level
 - Proper Ethernet/Zoning



EMC Simple Support Matrix
EMC XtremIO

Platform Support
AIX 6.1 and 7.1 ^{mi}
HP-UX 11i v3 (11.31) ^{i, ab}
Linux: Citrix XenServer 6.x
Linux: OL 5.7 with Oracle Enhanced Red Hat Kernel
Linux: OL 5.8 and later (UEK)
Linux: OL 6.x and later (UEK)
Linux: OL 7.0 (UEK)
Linux: RHEL 5.x
Linux: RHEL 6.x
Linux: RHEL 7.0 ^{no}
Linux: SLES 11 SP2 and SP3
Microsoft Windows 2008 SP2 ^{1, y}
Microsoft Windows 2008 R2 SP1 ^{1, y}
Microsoft Windows 2012 ^{1, y}
Microsoft Windows 2012 R2 ^{1, 4, y}
Solaris 10 x86
Solaris 10 SPARC
Solaris 11 x86
Solaris 11 SPARC
Solaris 11.1 x86
Solaris 11.1 SPARC
Solaris 11.2 x86
Solaris 11.2 SPARC
VMware ESX/ESXi 4.1 (vSphere) ^{4, u, y}
VMware ESXi 5.0 (vSphere) ^{4, y}
VMware ESXi 5.1 (vSphere) ^{4, y}
VMware ESXi 5.5 (vSphere) ^{4, y}



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

14

Before diving into the details of each host type, let's review some prerequisites. The following tasks should be performed before making the configuration changes described in the rest of this module.

The host should be installed and configured with a supported operating system and be physically connected to the XtremIO cluster through either FC or iSCSI. Whether using an HBA to use storage via Fibre Channel, or a NIC or HBA to use storage via iSCSI, the appropriate driver and firmware should be installed on the host. Volumes should be created and presented to the host. Refer to the most recent EMC Simple Support Matrix for XtremIO for supported operating systems, drivers, and firmware.

Lesson Summary

During this lesson the following topics were covered:

- Volume sizing
- Volume creation
- Volume provisioning



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

15

This lesson covered volume sizing, creation, and provisioning.

You saw how sizing XtremIO arrays for capacity is relatively simple. Each X-Brick containing 400 GB SSDs (10 TB X-Brick) can hold up to 7.5TB of deduplicated data. For X-Bricks containing 800 GB SSDs (20 TB X-Bricks), up to 15 TB of deduplicated data can be stored, but the logical capacity is limited to 48 TB. You may also recall the 800 GB X-Bricks enable more capacity for applications that contain mostly unique data (low deduplication ratios), such as databases.

We also covered how to create single and multiple volumes from the Configuration Workspace. We reviewed some XMCLI commands that can be used when logged in to the XMS for volumes:

- `show-volume`: Displays the specified volume information.
- `show-volumes`: Displays a list of volumes and their information.
- `show-volume-folders`: Displays the specified folder information.

You also saw how XtremIO inherently uses thin provisioning. All volumes are thin provisioned and it is not configurable. Unlike traditional methodologies, XtremIO thin provisioning has no performance loss and requires no capacity overhead.

This slide is intentionally left blank.



Lab Introduction: XtremIO Storage Provisioning

This lab covers the required operations needed to create an XtremIO LUN and present it to an existing host. Operations are:

- Validate connected equipment
- Create a volume
- Create an initiator group
- Map a volume



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

17

Successful presentation of storage from an XtremIO X-Brick to an operating system consists of several tasks that need to be performed. This lab will guide you through the required operations needed to create an XtremIO LUN and present it to an existing host.

Lab Debrief

This lab covered:

- Validate connected equipment
- Create a volume
- Create an initiator group
- Map a volume



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

18

This lab covered the required operations needed to create an XtremIO LUN and present it to an existing host.

Lesson: VMware Host Connectivity

This lesson covers the following topics:

- Fibre Channel HBA configuration
- Multipathing software configuration
- Host and vCenter server parameter configuration
- vStorage API for Array Integration (VAAI)
- Using XtremIO storage on ESXi hosts



This lesson covers considerations when presenting XtremIO storage to a VMware ESXi host.

vSphere: HBA Queue Depth

- Adjust for optimal performance
- Increasing queue depth may cause hosts to over stress other arrays connected
 - Resulting in performance degradation
- In mixed environments compare the XtremIO recommendations with those of other platforms before applying them

From the User Guide

HBA Vendor	Command
Qlogic	<code>esxcli system module list grep ql</code>
Emulex	<code>esxcli system module list grep lpfc</code>

Example (for host with Qlogic HBA):

```
# esxcli system module list | grep ql
qla2xxx                true          true
```

To adjust the HBA queue depth, use one of the following commands:

HBA Vendor	Command
Qlogic	<code>esxcli system module parameters set -p ql2xmaxqdepth=256 -m qla2xxx</code>
Emulex	<code>esxcli system module parameters set -p lpfc0_lun_queue_depth=256 -m lpfc820</code>

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

20

The queue depth of Fibre Channel HBAs should be adjusted to the maximum value for optimal performance. The maximum value is typically around 256.

Note that changing the HBA queue depth is designed for advanced users. Increasing the queue depth may cause the host to over stress other non-XtremIO arrays in the environment and cause performance degradation.

HBA Queue Depth: vSphere 5.1 and Earlier

Adjust for optimal performance

- Set `Disk.SchedNumReqOutstanding` to 256
 - Set `Disk.SchedQuantum` to 64
 - Set `Disk.DiskMaxIOSize` to 4096
- Increasing queue depth may cause hosts to stress other arrays connected
 - This can result in performance degradation
 - In mixed environments compare the XtremIO recommendations with those of other platforms before applying them



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

21

vSphere 5.1 and earlier allow queue depth to be specified for the volumes accessible to the host via the `Disk.SchedNumReqOutstanding` parameter. The value should be set to the maximum value, which is 256.

The `Disk.SchedQuantum` parameter, which specifies the maximum number of consecutive "sequential" I/O's allowed from one VM before we force a switch to another VM, should be set to 64.

`Disk.DiskMaxIOSize`, which specifies [in kB] the largest I/O allowed to a volume, should be set to 4096.

HBA Queue Depth: vSphere 5.5

- In vSphere 5.5, the queue depth can be set on a specific volume rather than on all volumes presented to the host
 - Set only after XtremIO volumes are presented to the ESXi host
 - The ESXi command line must be used to make the change
- Run this command to obtain the NAA [ID] for XtremIO LUNs

```
esxcli storage nmp path list | grep XtremIO -B1
```

- Run this command to set SchedNumReqOutstanding for the device to the maximum value (256)

```
esxcli storage core device set -d naa.xxx -O 256
```



When using vSphere 5.5, the number of outstanding I/Os parameter can be set on a specific volume rather than on all volumes presented to the host. Therefore, it should be set only after XtremIO volumes are presented to the ESX host, using the ESXi command line. Note that the volume will be identified by its NAA ID.

vSphere Native Multipathing

- Set the volume to round robin (all VMware interfaces)
- Adjust the vSphere NMP round robin path switching frequency
 - CLI only
 - From the default value (1000 I/O packets) to 1

From the User Guide

Version	Command
vSphere 5.x	<code>esxcli storage nmp psp roundrobin deviceconfig set -d naa.devicename --iops 1 --type iops</code>
vSphere 4.1	<code>esxcli nmp roundrobin setconfig --device=naa.devicename --iops 1</code>

vSphere Client

The screenshot shows the vSphere Client interface. The 'Path Selection' dropdown is set to 'Round Robin (VMware)'. Below it, the 'Storage Array Type' is 'VMW_SATP_DEFAULT_AA'. A table titled 'Paths' shows four active paths with green diamond status icons. A red box highlights the 'Status' column.

Runtime Name	Target	LUN	Status
a:cf 50:01:43:80:21:e1:ba:ce		0	Active (I/O)
d7 50:01:43:80:21:e3:f1:d6		0	Active (I/O)
a:cd 50:01:43:80:21:e1:ba:cc		0	Active (I/O)
d5 50:01:43:80:21:e3:f1:d4		0	Active (I/O)

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

23

If you are using VMware native multi-pathing, it is important to set the proper multi-pathing policy. To achieve best performance, it is recommended to use the round robin path selection policy for XtremIO volumes. Changing the Path Selection to round robin will show that all paths are Active as shown here. This will ensure optimal distribution and availability of load among I/O paths to the XtremIO storage.

Alternatively, this can be done from the ESXi host command line.

Use the `esxcli storage nmp path list | grep XtremIO` command to get the naa_id for the disks.

Then, to set the Path Selection policy to round robin, use the following OS specific command.

For ESXi 5.x, use `esxcli storage nmp device set --device <the naa_id> --psp VMW_PSP_RR`

For ESX 4.1, use `esxcli nmp device setpolicy --device <the naa_id> --psp VMW_PSP_RR`

You can also set the default path selection policy for any arriving storage in ESX 5.x by first identifying the SATP (Storage Array Type Plug-in) with the `esxcli nmp satp list` command and then modifying the default load balancing algorithm with

the `esxcli storage nmp satp set --default-psp=VMW_PSP_RR --satp=<your_SATP_name>` command.

EMC PowerPath/VE Support

- XtremIO supports EMC PowerPath/VE on ESXi hosts
 - Please refer to the latest XtremIO Simple Support Matrix for more information

Platform	PowerPath/VE Version
ESX/ESXi 4.1	PP/VE 5.4 SP2
ESXi 5.0/5.1	PP/VE 5.8, 5.9, 5.9 SP1
ESXi 5.5	PP/VE 5.9, 5.9 SP1



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

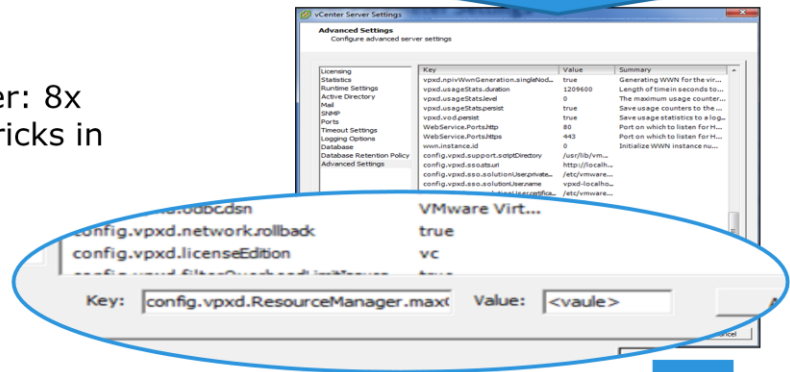
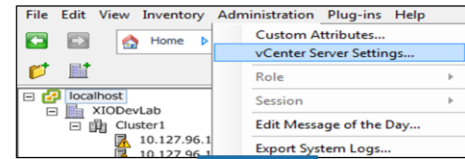
24

The table shows supported PowerPath/VE versions. Please take a moment to review them. Refer to EMC support page for the most updated information on PowerPath/VE support with XtremIO Storage. PowerPath/VE 5.9 SP1 treats XtremIO devices as native devices. Prior versions do not, but treat XtremIO volumes as generic devices. As a result, some functionality such as optimal failover and load balancing may not be present. In addition, the following limitations exist:

- powermt does not display the XtremIO array ID.
- powermt does not display the XtremIO storage interface IDs.
- powermt port-disable is not supported.

vCenter Cloning Operation Considerations

- Adjust the concurrent full cloning operations
 - Single X-Brick cluster: the default value (8) is recommended
 - Multi-X-Brick cluster: 8x the number of X-Bricks in the cluster



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

25

The maximum number of concurrent full cloning operations should be adjusted based on the XtremIO cluster size.

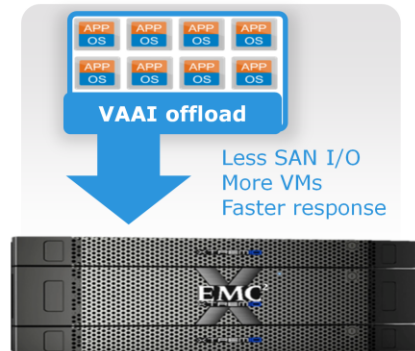
The vCenter Server parameter `config.vpxd.ResourceManager.maxCostPerHost` determines the maximum number of concurrent full clone operations allowed (the default value is 8). Adjusting the parameter should be based on the XtremIO cluster size as follows:

- Single X-Brick: 8 concurrent full clone operations
- Two X-Bricks: 16 concurrent full clone operations
- Four X-Bricks: 32 concurrent full clone operations
- Six X-Bricks: 48 concurrent full clone operations

If the parameter does not already exist, it can be added manually.

vStorage API for Array Integration (VAAI)

- In vSphere version 5.x, VAAI is enabled by default
- In vSphere version 4.1, VAAI is disabled by default
 - Install and enable the "XtremIO ESX 4.1 VIB" on the ESXi server
- XtremIO supports the following VAAI features:
 - Extended Copy (xcopy)
 - Block zeroing
 - Atomic test and set
 - Space reclamation
 - Thin Provision Stun (TPSTUN)



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

26

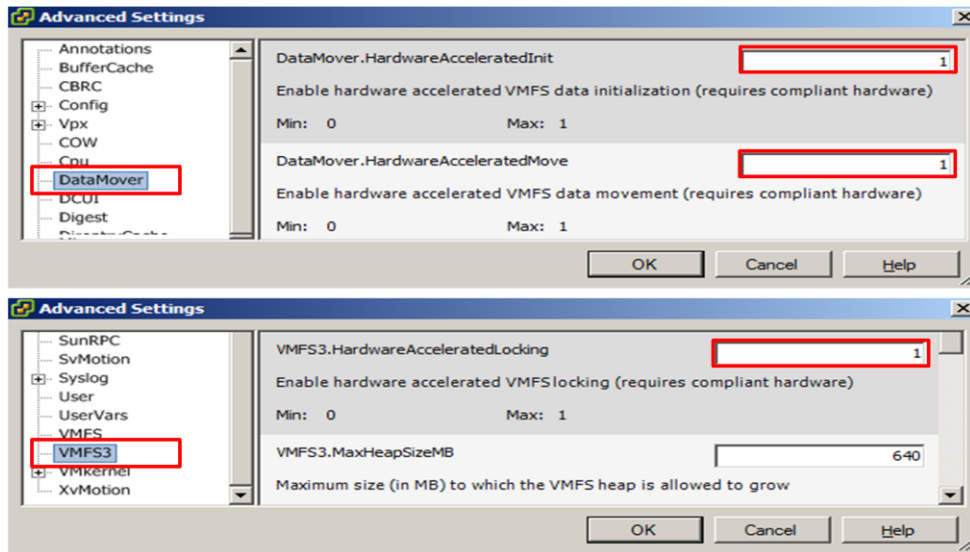
VAAI allows VMware to offload common supported operations from the ESXi hosts to the VAAI-enabled storage system. This reduces the host workload, lessens network traffic, and improves performance.

The VIB [vSphere Installation Bundle] that contains VAAI should be installed for vSphere 4.1; VAAI is enabled by default in vSphere 5.x.

With XtremIO, VAAI becomes even more efficient and provides even better performance. The fact that random IOPS pose no problem for XtremIO and that the Virtual Desktop Infrastructure (VDI) has high deduplication rates, also makes XtremIO the best storage for VDI-specific environments and for VMware environments in general.

In particular, XtremIO supports Extended Copy (xcopy), Block Zeroing, Atomic Test and Set, and Space Reclamation.

Verifying that VAAI is Enabled



EMC²

Now that VAAI for use with XtremIO is installed, it is always a good idea to double check that VAAI is enabled. To do this, use the vSphere Client to navigate to the host, go to the Configuration tab and then choose Advanced Settings under the Software section. Select the Data Mover section and check that the DataMover.HardwareAcceleratedMove and the DataMover.HardwareAcceleratedInit parameters are both set to "1".

Select the VMFS3 section and check that the VMFS3.HardwareAcceleratedLocking parameter is set to "1". If any of these parameters are set to "0", set them to "1".

Note that if the ESX host is also attached to non-XtremIO arrays that do not support VAAI, enabling VAAI may cause issues with those arrays.

VAAI Features

- Extended Copy
 - Offloads copy operations to the array
- Block Zeroing
 - Offloads zeroing activity to array
- Atomic Test and Set
 - Safe and efficient way of guaranteeing only one ESX host can access a LUN sector at any time
- Space reclamation (5.1 and later)
 - ESXi host informs the storage array that space can be reclaimed
- Out of Space VM Suspend and Notification with Thin Provisioning (TPSTUN)
 - Running out of space due to thin provisioning over-commit
 - The command causes suspending VMs, prevents blue screens



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Host Configuration

28

The Extended Copy feature not only avoids the reading and writing to and from the host (as is true for all VAAI-enabled arrays), but it also avoids reads and writes to the SSDs completely and simply creates metadata pointers in memory. The speed is amazing. XtremIO can clone VMs and perform storage vMotion at 6-8GB/s per X-Brick.

When the ESXi server sends the block zeroing command (trim) to XtremIO, XtremIO discards the SCSI command and does not create a mapping. This is because all LUNs are thin LUNs, and are only constructs in the metadata tables where the blocks have not yet been assigned.

If the host tries to read a block that XtremIO has no mapping for, the array simply returns zeros. Zeros consume neither physical SSD space nor logical metadata space in memory, and block zeroing takes place at over 10 GB/s. As a result of XtremIO's efficiency with VAAI block zeroing, it is recommended to create virtual disks with the thick provision eager zero property.

Although not a differentiating factor, it is important to know that XtremIO also supports VAAI Atomic Test and Set, as well as Space Reclamation. Note: This is important, since especially for boot storms there is no need to lock the whole volume to reserve a specific data block.

EMC Virtual Storage Integrator (VSI)

- EMC VSI Unified Storage Management (USM)
 - Provision XtremIO storage from within vSphere Client
 - Virtual Machine File System (VMFS) datastores
 - Raw Device Mapping (RDM) volumes on XtremIO

USM requirements

- One or more storage volumes are available
- One or more initiator groups are created
- One or more storage groups are created
- Network ports are configured
- Network connectivity exists between one or more storage controllers and ESXi hosts



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

29

EMC Virtual Storage Integrator (VSI) Unified Storage Management can be used to provision Virtual Machine File System (VMFS) datastores and Raw Device Mapping volumes on XtremIO arrays from within the vSphere Client. Furthermore, EMC VSI Storage Viewer extends the vSphere Client to facilitate the discovery and identification of XtremIO storage devices allocated to VMware ESX/ESXi hosts and virtual machines.

For further information on using these two vSphere Client plug-ins, refer to the EMC VSI Unified Storage Management [300-012-100] product guide and the EMC VSI Storage Viewer [300-013-072] product guide.

VMware: Using XtremIO Storage

- When creating volume from XtremIO use 512 LB
 - VMware does not support 4KB for datastores
- For VM best performance use RDMs
 - Physical compatibility mode
- Use Eager Zeroed Thick when presenting virtual disks to VMs

The screenshot shows three parts of the VMware configuration interface:

- Volume Type:** A dropdown menu with 'Normal (512 LBs)' selected. Other options are '4KB LBs' and 'Legacy Windows (offset:63)'.
- Select the type of disk to use:** A section with three radio button options: 'Create a new virtual disk', 'Use an existing virtual disk', and 'Raw Device Mappings'. The 'Raw Device Mappings' option is selected and highlighted with a red rectangle. Below it is a descriptive text: 'Give your virtual machine direct access to SAN. This option allows you to use existing SAN commands to manage the storage and continue to access it using a datastore.'
- Disk Provisioning:** A section with three radio button options: 'Thick Provision Lazy Zeroed', 'Thick Provision Eager Zeroed', and 'Thin Provision'. The 'Thick Provision Eager Zeroed' option is selected and highlighted with a red rectangle.



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

30

Shown here are a few considerations when using XtremIO with VMware. Since VMware does not support 4K logical block volumes, be sure to create the XtremIO volume using the default 512 LB size. For the best performance, it is recommended to use physical compatibility RDMs for data volumes. If presenting space via the datastore, it is recommended to create virtual disks with the thick provision eager zeroed provisioning policy. When cloning VMs, full clones are recommended.

Note: VAAI commands and some vCenter features are not supported with RDMs.

Lesson Summary

During this lesson the following topics were covered:

- Fibre Channel HBA configuration
- Multipathing software configuration
- Host and vCenter Server parameter configuration
- vStorage API for Array Integration (VAAI) Settings
- Using XtremIO storage



This lesson covered considerations when presenting XtremIO storage to a VMware host. We covered Fibre Channel HBA configuration and recommended adjusting the queue depth of the Fibre Channel HBA to the maximum value for optimal performance. The maximum value currently is 256.

When using VMware native multipathing, we highlighted the importance of setting the proper multipathing policy. Best performance is achieved by selecting the round robin path selection policy for XtremIO volumes. Changing the Path Selection to round robin will show that all paths are Active. Doing this will ensure optimal distribution and availability of load among I/O paths to the XtremIO storage. This can also be achieved from the ESX host command line.

During the Host and vCenter Server parameter configuration, we saw the maximum number of concurrent full cloning operations should be adjusted based on the XtremIO cluster size. The vCenter Server parameter `config.vpxd.ResourceManager.maxCostPerHost` determines the maximum number of concurrent full clone operations allowed (the default value is 8). Adjusting the parameter should be based on the XtremIO cluster sizes. If the parameter does not already exist, it can be added manually.

You saw how VAAI allows VMware to offload common supported operations from the ESXi hosts to the VAAI-enabled storage system, reducing the host workload, lessening network traffic, and improving performance.

When applying the VAAI settings, you should have noted that there are a few considerations when using XtremIO with VMware.

Noting the restrictions associated with 4K logical block volumes, you saw that the recommended XtremIO volume size is 512 LB.

You should also have noted that using physical compatibility RDMs for data volumes is recommended for best performance.

Note: VAAI commands and some vCenter features are not supported with RDMs.

This slide is intentionally left blank.



Lab Introduction: ESXi Host Configuration

This lab covers the additional optional steps that can be performed or validated to ensure high performance.

- Modify an ESXi host for optimal performance



An XtremIO volume can be presented to an ESXi and while functional, it is not optimal. This lab guides you through the additional optional steps that can be performed or validated to ensure high performance.

Lab Debrief

This lab covered this topic:

- Modify an ESXi host for optimal performance



An XtremIO volume can be presented to an ESXi and while functional, it is not optimal. This lab guides you through the additional optional steps that can be performed or validated to ensure high performance.

Lesson: Microsoft Windows Connectivity

This lesson covers the following topics:

- Fibre Channel HBA configuration
- Multipathing software configuration
- Using XtremIO storage on Windows



This lesson covers considerations when presenting XtremIO storage to a Microsoft Windows host.

Configuring Windows Hosts

Areas that may need attention are:

- HBA queue depth and execution throttle
- Multipathing support parameters
- File system alignment
- Miscellaneous Windows parameters

These may require Registry changes

- May be performed automatically by HBA utility



The configuration of Windows hosts will require adjusting various queue depth settings. Multipathing software should be used, and the file system should be correctly aligned. Subsequent slides will show more detail on these topics.

HBA Driver Parameters: Emulex

Port Information | Statistics | Maintenance | DHCPAP | Target Mapping | **Driver Parameters** | Diagnostics | Transceiver Data | VPD

Installed Driver Type: elxfc

Adapter Parameter	Value
CoalesceRspCnt	8
DriverTraceMask	0x0
EnableAck0	Enabled
EnableAUTH	Disabled
EnableFDMI	0
EnableNPIV	Disabled
ExtTransferSize	3
FrameSizeMSB	0
InitTimeOut	15
LimTransferSize	0
LinkSpeed	Auto Det...
LinkTimeOut	30
LogErrors	3
NodeTimeOut	30
PerPortTrace	0
QueueDepth	254
QueueTarget	0
RmScsiPort	16
ScanDown	1
SlMode	0
TargetOption	0
Topology	2
TraceBufSz	2000000

Modify Adapter Parameter

Parameter: QueueDepth

Value: 254

Range: 1 - 254

Default: 32

Activation Requirements

None. Parameter is dynamically activated.

Description

Outstanding Requests on a per Lun or Target Basis (see QueueTarget)

☐ Make change temporary

☐ Make all changes temporary (if possible)

Restore Defaults Globals Apply Save...

Modify Adapter Parameter

Parameter: ExtTransferSize

Value: 3

Range: 0 - 3

Default: 0

Activation Requirements

Reboot required.

Description

Extended transfer size: 0 = 512Kbytes: 1 = 1Mbyte: 2 = 2Mbytes: 3 = 4Mbytes

☐ Make change temporary

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

37

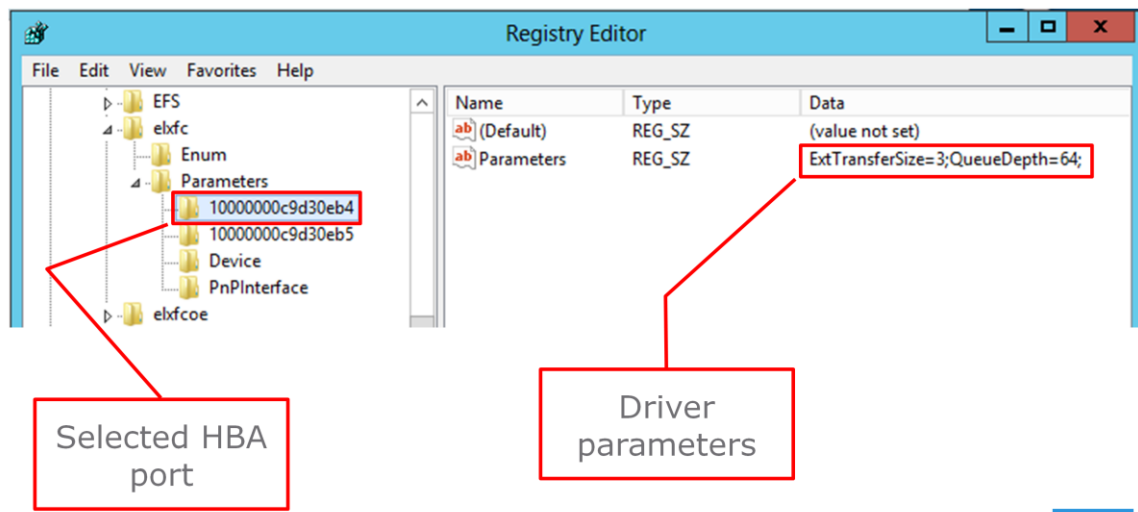
The Emulex HBA management utility, OneCommand Manager [OCManager], allows easy modification of HBA driver parameters. Where the change requires a reboot to take effect, a note to that effect is displayed in the **Activation Requirements** pane, as seen in the screenshot on the right-hand side of the slide.

The **ExtTransferSize** parameter determines the maximum I/O size that can be performed by the HBA. The maximum value, 3, allows I/Os of up to 4 MB, and is the setting that should be selected.

The **QueueTarget** parameter determines whether queue depth settings apply to individual LUNs or to targets. The default value, 0, applies the queue depth setting to individual LUNs.

The **QueueDepth** parameter sets the maximum queue depth per LUN [see **QueueTarget** above]. This should be set to its maximum value, 254.

Queue Depth: Windows 2012 and LPe11002-E



This slide shows the Windows registry entry associated with queue depth for an Emulex LPe11002-E Fibre Channel HBA. The highlighted area identifies the Registry entry at

Computer\HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\elxhc\Parameters\<WWPN>. <WWPN> is the World Wide Port Name, which identifies the specific port on the [in this case, 2-port] HBA.

Note that the parameter is configured on a per-port basis. The queue depth is not modified by changing the registry entry directly, though it could be. Instead, it is modified in the OneCommand GUI, which allows configuration of a number of HBA driver parameters. Changes made there are applied to the Registry. Note also that OneCommand has another parameter related to queue depth, QueueTarget, which determines whether the configured queue depth will be applied on a per-target or per-LUN [default] basis. The parameter takes values in the range 1 to 254. The maximum value may be different depending on the version of the OS and the HBA driver. Changing this parameter may be dynamic; if so, a reboot of the host will not be required.

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

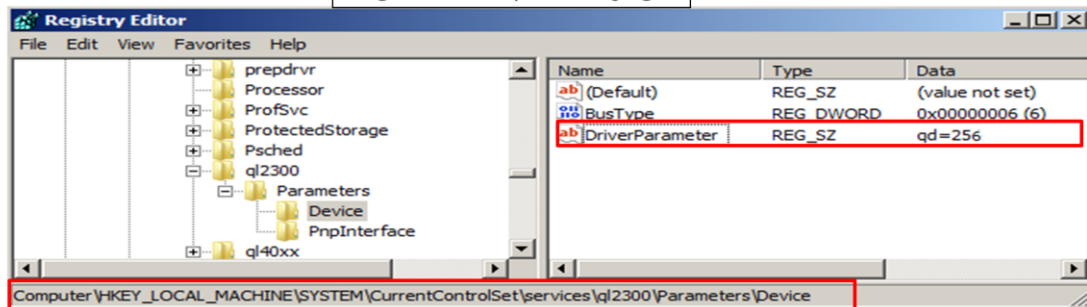
Module: XtremIO Host Configuration

38

Windows: HBA Queue Depth

- On Windows, the HBA queue depth can be adjusted via the Windows Registry
- Set queue depth to 128 for Emulex, 256 for Qlogic

Regedit example for Qlogic



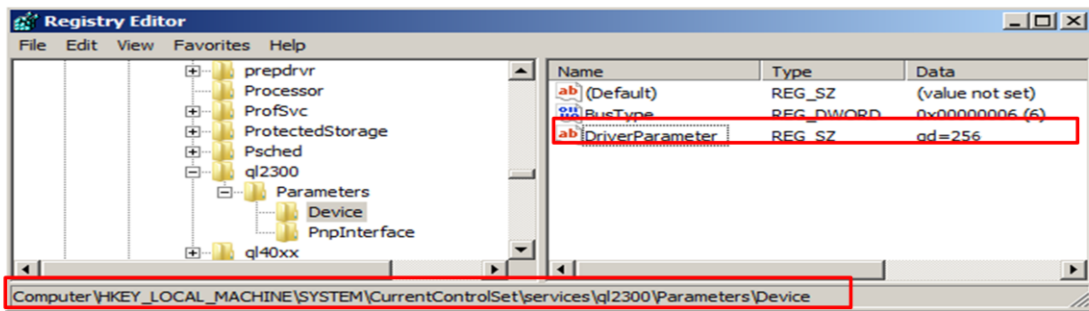
EMC²

When using Windows with XtremIO, it is recommended to adjust the queue depth of the Fibre Channel HBA to the maximum value for optimal performance. The maximum value currently is 256 for QLogic HBAs and 128 for Emulex HBAs.

Note that changing the HBA queue depth is designed for advanced users. Increasing the queue depth may cause the host to over stress other arrays in the environment and cause performance degradation while communicating with them. Therefore, especially in mixed environments with multiple array types connected to the Windows host, you should compare the XtremIO recommendations for HBA queue depth against that of other platforms before deciding if and how to adjust it.

HBA Queue Depth and Execution Throttle

- The HBA queue depth can be adjusted via the Windows Registry
 - The execution throttle is changed with the HBA utility
- Emulex HBAs: set queue depth = 254, execution throttle = 8192
- Qlogic HBAs: set queue depth = 256, execution throttle = 65536



EMC²

When using Windows with XtremIO, it is recommended that the queue depth and execution throttle of the Fibre Channel HBA be adjusted to the maximum value for optimal performance. The maximum queue depth value is currently 256 for QLogic HBAs and 254 for Emulex HBAs. The maximum execution throttle value is currently 65536 [default 32] for QLogic HBAs and 8192 [default] for Emulex HBAs.

Note that changing the HBA queue depth is an advanced activity. Increasing the queue depth may cause the host to over stress other arrays in the environment and cause performance degradation while communicating with them. Therefore, especially in mixed environments with multiple array types connected to the Windows host, you should compare the XtremIO recommendations for HBA queue depth against that of other platforms before deciding if and how to adjust it.

The execution throttle is changed with the HBA CLI or GUI management utility. It should be higher than the queue depth setting.

Microsoft MPIO [Native Multi-Path I/O]

- XtremIO supports Microsoft MPIO
 - Enabled via Server Manager
- The Least Queue Depth policy [policy 4] is recommended
 - Sends I/O down the path with the fewest outstanding I/O requests
 - Set this policy for all XtremIO devices

```
mpclaim -l -t "XtremIO XtremApp" 4  
mpclaim -l -m 4
```



There should be exactly eight (8) spaces between the word XtremApp and the double quotes (").

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

41

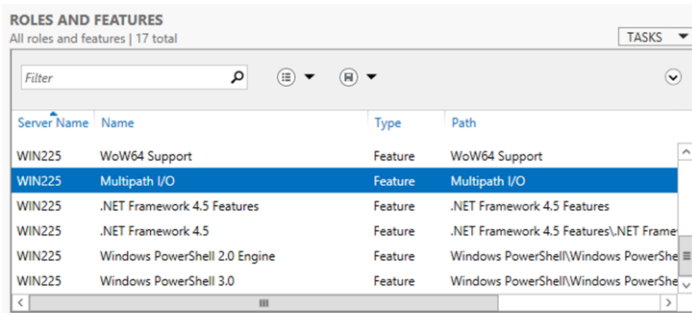
Windows servers should have MPIO enabled when XtremIO arrays are attached. The recommended MPIO policy is 'Least Queue Depth' - policy number 4 – which should be applied to all XtremIO volumes.

This policy may be configured from a high-level CLI, such as PowerShell [but not Command Prompt] with the **mpclaim** command. The command has several options, among which are **-l** [letter 'ell', not number 'one' or letter 'eye'] to set the load-balancing policy.

Two examples of the command are shown on the slide. The first specifies that all devices identified as having vendor ID XtremIO [an 8-character field] and product identifier XtremApp [a 16-character field, hence the spaces] should use policy number 4, the Least Queue Depth policy. The second example assigns policy 4 to all devices controlled by MSDSM, the Microsoft Device Specific Module. Other vendor DSMs are available.

Checking Windows MPIO Installation

- Open Server Manager to verify that MPIO is enabled
- Check current load balancing policy from PowerShell
 - Use `mpclaim -s -d`



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

42

XtremIO supports Microsoft's native Microsoft Multipath I/O (MPIO). This a feature that must be enabled.

- To enable MPIO, go to Server Manager, choose Features and then click Add Feature to launch the Add Feature Wizard.
- Select Multipath I/O and click Next.
- Then click Install to install MPIO on Windows. If requested, reboot the host upon conclusion of the MPIO feature installation.

Configure the Least Queue Depth policy for all the XtremIO devices. This policy sends I/O down the path with the fewest currently outstanding I/O requests. To configure this as the default policy to all LUNs presented by the XtremIO array, run the following two `mpclaim` commands from the command prompt:

```
mpclaim -l -t "XtremIO XtremApp" 4
```

Note that there must be exactly eight blank spaces in the first command before the end quote and that this command will automatically reboot the host without prompting.

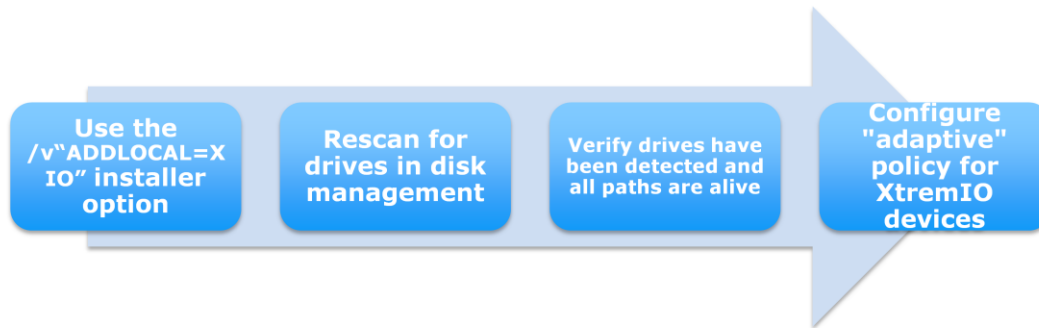
```
mpclaim -l -m 4
```

Use the command `mpclaim -e` to query the hardware IDs of connected devices.

Use the command `mpclaim -s -t` to query the current default MPIO load-balancing algorithm.

EMC PowerPath Support

- PowerPath version 5.7 SP2 is supported
 - Please refer to the latest XtremIO Simple Support Matrix for more information
- Install PowerPath after enabling MPIO



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration



43

To install and configure PowerPath for Windows with XtremIO, enable Window's native MPIO and install PowerPath using the command line with the /v "ADDLOCAL=XIO" installer option. Then rescan for drives and open the PowerPath console to verify that the drives have been detected and that all paths are alive. Lastly, configure the adaptive policy for all the XtremIO drives. To disable this functionality, use the command line and run the powerpath installer with the /v "REMOVE=XIO" installer option.

Please refer to the XtremIO Storage Array User Guide for more details on setting up PowerPath for Microsoft Windows and XtremIO connectivity.

Windows Space Reclamation

- Windows Server 2012 has a space reclaim feature
 - Instructs the XtremIO storage array to reclaim deleted storage
 - Beneficial for reclaiming space from Thin volumes
 - Similar to a garbage collection operation
- Feature is enabled by default
 - Sends space reclaim commands (TRIM or UNMAP) to the storage array
 - Feature can be disabled temporarily if required
 - Quick format operations take a long time to complete
 - Asynchronous TRIM command will lock the entire disk (MS PhysicalDisk) until all its space is reclaimed



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

44

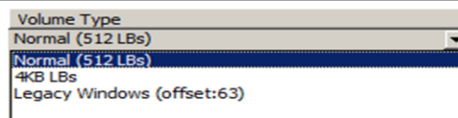
A new feature in Windows Server 2012 allows a Windows host to instruct a storage array to reclaim deleted storage. This feature will have the greatest effect when Thin provisioned volumes are being used. The Windows space reclaim feature sends TRIM or UNMAP commands to XtremIO to notify the array that certain files have been deleted on the host, and they should also be deleted on the array. This feature is enabled by default and is fully supported on XtremIO systems. Note that a standard Windows file delete operation simply marks the file as deleted and modifies the volume directory, but does not delete the data associated with the file.

The space reclamation feature is very similar to a garbage collection operation. In some cases, there may be temporary performance issues because space reclamation uses additional resources. The effect may be minimized by temporarily disabling the feature, and re-enabling it when resource-intensive operations (e.g. formatting) are complete.

Windows: Using XtremIO Storage

- Default is to use 512 LB size
 - Windows 2008 (and earlier) does not support 4KB volumes
 - Use 4KB LBs volumes if application uses a block size which is a multiple of 4 kB
 - Allows easy alignment of file systems
 - No disk alignment is required for Windows 2008 and newer versions
 - Manually set alignment for legacy Windows systems

XtremIO Volume Creation Options



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

45

One important consideration when using XtremIO with Windows is the logical block size of the XtremIO volume. In general, when the applications using the volume have a 4 kB or a multiple of 4 kB block size, it is recommended to configure the XtremIO volume with a 4 kB block size. In all other cases, the 512 byte block size should be used. Note, though, that Windows 2008 (and earlier) does not support 4KB volumes.

Another important consideration is file system alignment. Unaligned disk partitions may substantially impact I/O to the disk. With Windows 2008 and newer, disk partitions are aligned by default to a 1 MB disk alignment. Therefore, no further action is required to align a newly created Windows disk partition when using newer versions of Windows.

However, with past Windows versions, such as Windows 2003, disk partitions in Windows have a 63 sector offset by default. In these cases, disk partitions should be created and aligned.

Lesson Summary

During this lesson the following topics were covered:

- Windows Fibre Channel HBA configuration
- Windows multipathing software configuration
- Using XtremIO storage on Windows hosts



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

46

This lesson covered considerations when presenting XtremIO storage to a Microsoft Windows host.

Covered in the Windows Fibre Channel HBA configuration section was the following:

- Configuring Windows hosts
- The Emulex HBA management utility, OneCommand Manager [OCManager]
- Windows registry entry associated with queue depth for an Emulex LPe11002-E Fibre Channel HBA
- HBA Queue Depth and Execution Throttle

We also reviewed the recommended MPIO policy that should be applied to all XtremIO volumes, while noting that the policy can be configured from a high-level CLI with the `mpclaim` command.

You saw the importance of the logical block size of the XtremIO volume when using XtremIO with Windows; the significance of considering the file system alignment, since unaligned disk partitions can substantially impact I/O to the disk. You should have also noted that with past Windows versions, such as Windows 2003, disk partitions in Windows have a 63 sector offset by default, and when these cases exist disk partitions should be created and manually aligned.

Lesson: Linux Host Connectivity

This lesson covers the following topics:

- Fibre Channel HBA configuration
- Configuring I/O elevators
- Multipathing software configuration
- Using XtremIO storage



This lesson covers considerations when presenting XtremIO storage to a Linux host.

HBA Configuration: Qlogic

SuSE/Red Hat 5.x and earlier	SuSE/Red Hat 6.x and later
- Edit the file <code>/etc/modprobe.conf</code> - Set the <code>ql2xmaxqdepth</code> parameter value [default 32] to 256	- Create or edit the <code>/etc/modprobe.d/qla2xxx.conf</code> file - Set the <code>ql2xmaxqdepth</code> parameter value [default 32] to 256 - Rebuild the RAMdisk

- After the parameter is set, reboot the host
- Change the execution throttle value [default 32] to 65536
 - Use the Qlogic HBA BIOS or CLI utility



To modify the HBA queue depth value, select one of the following options according to the Linux version:

- For SuSE and Red Hat version 5.x and earlier:
Edit the file `/etc/modprobe.conf` and add or modify the following parameter:
`options qla2xxx ql2xmaxqdepth=<new_queue_depth>`
- For later SuSE and Red Hat releases:
 - a. Create or edit the `/etc/modprobe.d/qla2xxx.conf` file and add or modify the following parameter:
`options qla2xxx ql2xmaxqdepth=<new_queue_depth>`
 - b. Rebuild the RAMdisk using the following command:
`mkinitrd -v --with=qla2xxx -f
/boot/initramfs-w-qla2xxx-`uname -r`.img `uname -r``

After the parameter is set, reboot the host. To verify that the parameter is set, run the following command:

```
cat /sys/module/qla2xxx/parameters/ql2xmaxqdepth
```

To change the execution throttle value from the default of 32 to 65536, use the Qlogic HBA BIOS or CLI utility.

HBA Configuration: Emulex

SuSE/Red Hat 5.x and earlier	SuSE/Red Hat 6.x and later
Edit the file /etc/modprobe.conf	- Create or edit the /etc/modprobe.d/lpfc.conf file - Rebuild the RAMdisk

- Increase queue depth to maximum for the two parameters
 - lpfc_lun_queue_depth (default value 32)
 - lpfc_hba_queue_depth (default value 8192)
- After the parameters are set, reboot the host



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

49

For Emulex cards there are two parameters that need to be increased to the maximum value:

- lpfc_lun_queue_depth (default value 32)
- lpfc_hba_queue_depth (default value 8192)

XtremIO recommends setting the LUN queue parameter to the maximum value (128). To modify the HBA Queue Depth value, select one of the following options according to the version:

- For SuSE and Red Hat 5.x and earlier:
Edit the file /etc/modprobe.conf and add or modify the following parameter:
options lpfc lpfc_lun_queue_depth=<new_queue_depth>
- For later SuSE and Red Hat releases:
 - Create or edit the /etc/modprobe.d/lpfc.conf file and add or modify the following parameter:
options lpfc lpfc_lun_queue_depth=<new_queue_depth>
 - Rebuild the RAMdisk using the following command:
mkinitrd -v --with=lpfc820 -f
/boot/initramfs-w-lpfc820-`uname -r`.img `uname -r`

After the parameter is set, reboot the host. To verify that the parameter is set, run the following command:

```
cat /sys/class/scsi_host/host*/lpfc_lun_queue_depth
```

Configuring I/O Elevators

- Linux implements I/O algorithms called elevators to accommodate different workloads
- Set the I/O elevator parameter to "noop" for all-Flash arrays
- GRUB: modify /boot/grub/menu.lst
- GRUB2: modify /etc/default/grub
 - append "elevator=noop" to the list of GRUB_CMDLINE_LINUX_DEFAULT values



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

50

Linux implements various I/O algorithms, called elevators, to suit different workloads.

Elevators can help to reduce seek operations, can prioritize I/O requests, or make sure an I/O request is carried out before a given deadline. Choosing the best suited I/O elevator not only depends on the workload, but on the hardware, too. NL-SAS, SSDs, RAID arrays, or network storage systems, for example, each require different tuning strategies.

The "noop" I/O elevator is an I/O scheduler that just passes down the I/O that comes to it. This elevator is useful for intelligent storage, or devices that do not depend on mechanical movement, like SSDs.

To configure any I/O elevator, follow one of these procedures according to GRUB [**GR**and **Un**ified **B**ootloader] version:

- GRUB (versions prior to GRUB2):
Modify the "/boot/grub/menu.lst" file and append "elevator=noop" as an option to the end of the kernel line.

Example:

```
kernel /vmlinuz-2.6.32-358.el6.x86_64 ro ... elevator=noop
```

- GRUB2:
1. Modify the /etc/default/grub file and append/write "elevator=noop" to the list of GRUB_CMDLINE_LINUX_DEFAULT values.

Example:

```
GRUB_CMDLINE_LINUX_DEFAULT=" ... elevator=noop"
```

- 2. Rebuild GRUB2 configuration files using the following shell command:

```
sudo update-grub
```


Linux Native Multipathing

- The Linux native multipathing requires installing the device-mapper-multipath version 0.4.9-72 (or later) RPM
- Volumes should be configured as multipath devices whenever possible
 - Even when only a single path is initially used
- The configuration file for the multipath daemon is `multipath.conf`
- For iSCSI devices, modify `/etc/iscsi/iscsi.conf`



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

51

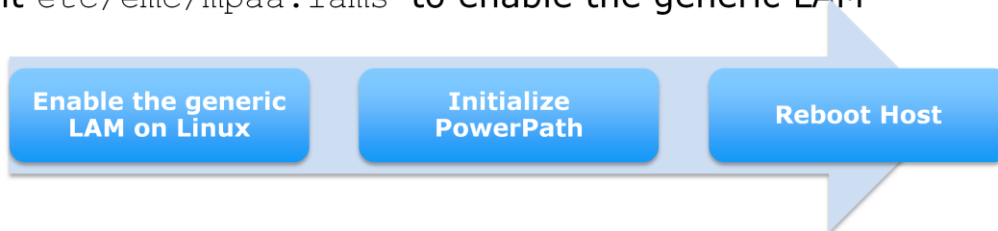
Even when the default multipath configuration appears to be functioning and provides path failure, key values must be set to enable the cluster to survive controller failover. Volumes should be configured as multipath devices whenever possible, even if only a single path is initially used.

The Linux native multipathing requires installing the device-mapper-multipath version 0.4.9-72 (or later) RPM. The configuration file for the multipath daemon is *multipath.conf*. It is used to overwrite the built-in configuration table of the multipath daemon.

If XtremIO volumes are accessed through iSCSI, the *iscsi.conf* file is modified in order to overwrite iSCSI-specific multipathing-related settings.

EMC PowerPath Support

- PowerPath version 5.7 SP1 (or above) is supported
 - XtremIO volumes show up as generic devices
- Edit `etc/emc/mpaa.lams` to enable the generic LAM



- PowerPath considerations:
 - `powermt` does not display the XtremIO array ID or storage interface IDs
 - The `powermt` port-disable feature is not supported



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

62

EMC PowerPath supports XtremIO volumes as generic devices. PowerPath for Linux version 5.7 SP1 (or above) is required.

To enable the generic LAM (**L**oadable **A**rray **M**odule) on Linux:

1. Append the following text to the `etc/emc/mpaa.lams` file:

```
managed:generic:XtremIO:XtremApp
```

2. Run the following shell commands:

```
powermt init
/etc/opt/emcpower/powercf -m
```

3. Reboot the host.

To disable the generic LAM on Linux:

1. Remove the following text from the `etc/emc/mpaa.lams` file:

```
managed:generic:XtremIO:XtremApp
```

2. Run the following shell commands:

```
powermt init
/etc/opt/emcpower/powercf -m
```

3. Reboot the host.

Unlike PowerPath's array-customized LAMs, PowerPath's generic LAM may not feature optimal failover and load balancing behaviors for the array that it manages. When volumes are managed with the generic LAM, note that `powermt` does not report the XtremIO array ID, XtremIO storage Interface IDs, and that the port-disable feature is unsupported.

Linux: Using XtremIO Storage

- When creating volumes in XtremIO for a Linux host, the default logical block (LB) size of a new XtremIO volume is 512 bytes
- This parameter can be adjusted to 4 kB
- Set to 4KB if application using this volume has a 4 kB (or multiple of 4 kB) block size
 - This will better align I/O with the block size of the XtremIO cluster
 - In all other cases, 512B

XtremIO Volume Creation Options

Volume Type
Normal (512 LBs)
Normal (512 LBs)
4KB LBs
Legacy Windows (offset:63)

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

53

When you create volumes in XtremIO for a Linux host, the logical block (LB) size of a new XtremIO volume is 512 bytes. This parameter can be adjusted to 4 kB. When applications using the created volume have a 4 kB (or a multiple of 4 kB) block size, it is recommended to present a volume with a 4 kB block size to a Linux host, to better align I/O with the block size of the XtremIO cluster.

In all other cases, use a 512 byte block size.

Lesson Summary

During this lesson the following topics were covered:

- Considerations when presenting XtremIO storage to a Linux Host



This lesson covered considerations when presenting XtremIO storage to a Linux host. During this lesson we reviewed the HBA configuration for Qlogic and Emulex; how to configure I/O elevators; Linux Native Multipathing configuration; and creating volumes on XtremIO storage.

Lesson: Solaris Host Connectivity

This lesson covers the following topics:

- Fibre Channel HBA configuration
- Multipathing software configuration
- Using XtremIO storage



This lesson covers considerations when presenting XtremIO storage to a Solaris host.

HBA Queue Depth and Execution Throttle

- Execution throttle controls the amount of outstanding I/O requests per HBA port

QLogic	Emulex
• Modify /kernel/drv/qlc.conf • Set execution-throttle parameter to 256	• Modify /kernel/drv/lpfc.conf • Set lun-queue-depth parameter to 256 • Set tgt-queue-depth parameter to 256/<# LUNs>

- Reboot the host after the changes are made
- Changing these parameters requires a reboot



Execution throttle setting controls the amount of outstanding I/O requests per HBA port. For Qlogic HBAs, set the execution throttle to its maximum value (256). To apply this modification to the Solaris host at the HBA level, edit the /kernel/drv/qlc.conf file and set the execution-throttle parameter to 256.

For Emulex HBAs, edit the /kernel/drv/lpfc.conf file and set the lun-queue-depth and tgt-queue-depth parameters as follows:

- lun-queue-depth: The driver uses this value as the default limit for the number of simultaneous commands to issue to a single logical unit on a single target on the loop. It is recommended to set lun-queue-depth to 256.
- tgt-queue-depth: The driver uses this value as the default limit for the number of simultaneous commands to issue to a single target on the loop. A value of 0 results in no target throttling. It is recommended to set tgt-queue-depth to 256/<number of LUNs>.

Solaris: SSD Parameters

- Edit `/etc/system` and configure both `ssd` parameters
- Set the timeout value for a single I/O operation for any `ssd` disk in the cluster
 - Set `ssd_io_time` parameter to 120 seconds (expressed in hex)
 - `ssd:ssd_io_time=0x78`
- Set the depth of the execution throttle per LUN
 - Set the `ssd_max-throttle` parameter to `256/<# LUNs>`
 - `ssd:ssd_max_throttle=256`
- Changing these parameters requires a reboot



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

57

For all devices: edit the `/etc/system` file and configure both `ssd` parameters as follows:

- To set the timeout value for a single I/O operation for any `ssd` disk in the cluster, set the `ssd_io_time` parameter to 120 seconds as follows:

```
set ssd:ssd_io_time=0x78
```

- To set the depth of the execution throttle per LUN, set the `ssd_max-throttle` parameter to `256/<number of ssd luns>` as follows:

```
set ssd:ssd_max_throttle=256
```

StorEdge Traffic Manager Software (STMS)

- Edit the /kernel/drv/scsi_vhci.conf file and set the values
 - load-balance="round-robin"
 - auto-failback="enable"

Solaris 11 x86	All other Solaris hosts
• Add the following to the scsi_vhci.conf file <pre>device-type-scsi-options-list ="XtremIO XtremApp", "symmetric- option"; symmetric-option = 0x1000000;</pre>	• Add the following to the scsi_vhci.conf file <pre>scsi-vhci-failover-override = "XtremIO XtremApp", "f_sym_emc" ;</pre>

- To enable STMS, boot the system using the proper command
 - stmsboot -e



STMS (formerly known as MPxIO), is the built-in Path Management software used by Solaris and supported by Oracle. This section describes the required steps for configuring STMS on the Solaris host.

To configure STMS on the Solaris host:

- Edit the /kernel/drv/scsi_vhci.conf file and set the values of the following parameters:
load-balance="round-robin"
auto-failback="enable"
- Add the following option to the file:
device-type-scsi-options-list ="XtremIO XtremApp",
"symmetric-option";
symmetric-option = 0x1000000;

Note: For Solaris 11 on x86 platform, set load-balance and auto-failback the same as for SPARC system. However, instead of device-type-scsi-options-list and symmetric-option add only:

```
scsi-vhci-failover-override = "XtremIO XtremApp", "f_sym_emc";
```

To enable STMS, boot the system using the following command:

```
stmsboot -e
```


Veritas Dynamic Multipathing (DMP)

- Configuration modifications are required in order to manage XtremIO storage under Veritas Storage Foundation Suite
1. Disable STMS and reboot host
 2. Install Veritas DMP and verify that the configuration is active
 3. Add XtremIO to the jbod list to get values for vid and pid
 4. Create a disk group under VxVM and allocate the disk to it
 5. Create a VxFS file system and mount it



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

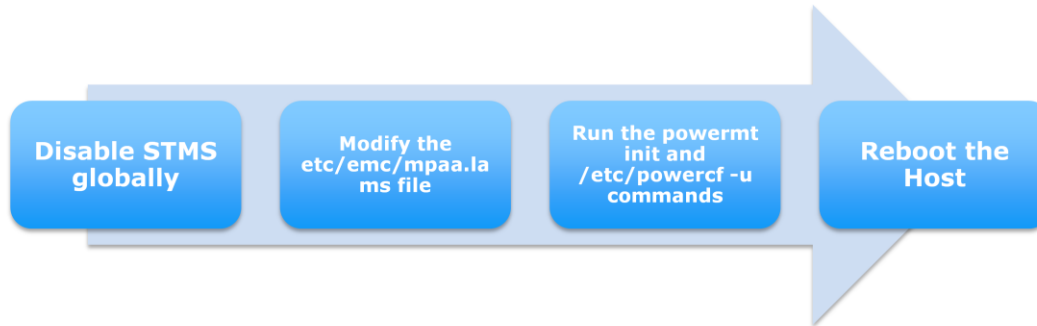
59

To manage XtremIO storage under Veritas Storage Foundation Suite:

1. Disable STMS by modifying the entry `mpxio-disable="yes"` in the `/kernel/drv/fp.conf` file.
2. Run the following command:
`stmsboot -d`
3. Reboot the host
4. Install Veritas DMP. Refer to Veritas Dynamic Multi-Pathing Installation Guide for details.
5. Verify that DMP configuration is active by running the following DMP commands:
`vxdisk path` - to display active paths to the XtremIO volumes
`vxdisk list <disk_name>` - to display disk information and paths
6. To get values for vid and pid by Veritas DMP commands, add XtremIO to the jbod list using the following command:
`vxddladm addjbod vid=XtremIO pid=XtremApp policy=aa`
7. Create a disk group under VxVM and allocate the disk to it.
8. Create a VxFS file system and mount it.

EMC PowerPath Support

- PowerPath versions 5.5 P02 and 5.5 P03 are supported
 - Please refer to the latest XtremIO Simple Support Matrix for more information



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

60

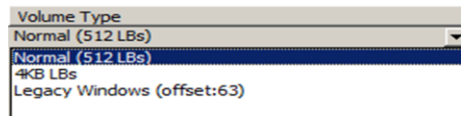
XtremIO supports multipathing using EMC PowerPath. PowerPath handles XtremIO LUNs as generic LAMs. PowerPath for Solaris versions 5.5 P02 and 5.5 P03 are supported with the following Solaris platform versions on XtremIO storage: Solaris 10 SPARC, Solaris 10 x86, Solaris 11 SPARC, Solaris 11 x86.

To enable the generic LAM on Solaris:

1. Disable STMS (MPxIO) globally before configuring PowerPath's generic LAM for XtremIO support. To do this, replace all occurrences of `mpxio-disable=no` with `mpxio-disable=yes` in configuration files `/kernel/drv/fp.conf` and `/kernel/drv/iscsi.conf`
2. Append the following text to the `etc/emc/mpaa.lams` file:
`managed:generic:XtremIO:XtremApp`
3. Run the following shell commands:
`powermt init`
`/etc/powercf -u`
4. Reboot the host.

Solaris: Using XtremIO Storage

- When creating volumes in XtremIO the default logical block (LB) size of a new XtremIO volume is 512 bytes
- Set to 4KB if application using this volume has a 4 kB (or multiple of 4 kB) block size
 - This will better align I/O with the block size of the XtremIO cluster
 - In all other cases, use 512B
 - 4KB LB volumes are not supported for use with UFS



- LUN ID 0 is usually used for the storage array control device
 - Best practice is not to use LUN ID 0



When creating volumes in XtremIO for a Solaris host, the default logical block (LB) size of a new XtremIO volume is 512 bytes. This parameter can be adjusted to 4 kB. In general, when the application using this volume has a 4 kB (or multiple of 4 kB) block size, it is recommended to present a volume with a 4 kB block size to a Solaris host to better align I/O with the block size of the XtremIO cluster. In all other cases, 512B should be used.

LUN ID 0 is typically used for the storage controller; best practice is to avoid the use of LUN ID 0.

Solaris: Disk Alignment

- For proper alignment to the newly allocated volume, format it to EFI partition and create a partition starting at sector 40
- This alignment method is the best when used with UFS – the Solaris native file system
- Using this method with ZFS results in misalignment since ZFS writes metadata to the disk with a different offset
 - For ZFS, align the user partition to start at sector 256

File System	Start Sector
UFS	40
ZFS	256



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

62

Correcting disk alignment after data has been populated requires erasing and restoring data, which can significantly affect performance. It is therefore recommended to verify that the volume is aligned before populating it with customer data. For proper alignment, format the newly allocated volume as an EFI partition starting at sector 40. As a result, the partition is aligned in accordance to the EFI label and all I/Os will start with an offset of 40 sectors.

This alignment method is the best when used with UFS – the Solaris native file system. Using this method with ZFS results in misalignment since ZFS writes metadata to the disk with a different offset. The best practice with ZFS is to align the user partition to start from sector 256.

Note: In Solaris on the x86 platform, it is necessary to create the aligned partition with fdisk before creating a file system. Create an aligned disk according to the table above and place the file system on the formatted slice (i.e. slice 2).

Lesson Summary

During this lesson the following topics were covered:

- Solaris Fibre Channel HBA configuration
- Solaris multipathing software configuration
- Using XtremIO storage on Solaris



This lesson covered storage provisioning to a Solaris host.

During this lesson we reviewed the Execution throttle setting controlling the amount of outstanding I/O requests per HBA port; setting the Solaris SSD parameters; the StorEdge Traffic Manager Software (STMS); Veritas Dynamic Multipathing (DMP); creating volumes in XtremIO for a Solaris host; and finally, correcting disk alignment after data has been populated which can significantly affect performance.

This slide is intentionally left blank.



Lesson: AIX Host Connectivity

This lesson covers the following topics:

- Fibre Channel HBA configuration
- Fast I/O Failure configuration
- Multipathing software configuration
- Using XtremIO storage on AIX



This lesson covers the provisioning of storage to AIX hosts.

Execution Throttle and I/O Transfer Size

- The number of outstanding requests per HBA port [execution throttle] should be increased to 2048
 - Run the command `chdev -l <fcs#> -a num_cmd_elems=2048`
 - Reboot the host
- Maximum I/O transfer size should be increased to 16 MB
 - Run the command `chdev -l <fcs#> -a max_xfer_size=0x1000000`



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

66

The HBA queue depth and execution throttle need to be changed to support XtremIO storage systems with AIX hosts. To change the execution throttle, **for each HBA** in the host, run the `chdev -l fcs# -a num_cmd_elems=2048 -P` command. Replace `fcs#` with `fcs0`, `fcs1`, etc.

In addition, the maximum I/O transfer size allowed should be modified for each HBA. The modified setting allows I/O sizes of up to 16 MB to be used.

Note: For further details on setting queue depth and execution throttle in AIX, refer to the *AIX/VIOS Disk and Adapter IO Queue Tuning* document:

<http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105745>

Path Queue Depth and System Buffer Size

- Increase path queue depth to 256
 - Run the command
`chdev -l <hdiskpower#> -a queue_depth=256` for each volume
- Modify the size of the system block I/O buffer cache
 - Run `chdev -l sys0 -a maxbuf=1000`



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

67

To change the queue depth for each volume [a PowerPath volume is used in the example], run the `chdev -l hdiskpower# -a queue_depth=256` command. Replace `hdiskpower#` with `hdiskpower0`, `hdiskpower1`, etc.

The number of pages in the system buffer should be changed with the `chdev -l sys0 -a maxbuf=1000` command.

Fast I/O Failure Configuration

- Set the Fast I/O Failure value for each fscsi device
 - Run `chdev -l <fscsi#> -a fc_err_recov=fast_fail -P`
- Turn on dynamic tracking for each fscsi device
 - Run `chdev -l <fscsi#> -a dyntrk=yes`
- Verify modified settings with `lasttr -El <fscsi#>`
- Requires a reboot



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

68

When the FC adapter driver detects a link event, such as a lost link between a storage device and a switch, it waits for the fabric to stabilize (approximately 15 seconds). If the device is not on the fabric, the FC adapter driver fails all new I/Os or future retries of the failed I/Os until the device rejoins the fabric.

Fast I/O Failure is controlled by the fscsi device attribute, `fc_err_recov` (default value is `delayed_fail`). To enable Fast I/O Failure for all fscsi devices, set this attribute to `fast_fail`.

If dynamic tracking of FC devices is enabled, the FC adapter driver detects when the Fibre Channel N_Port ID of a device changes. The FC adapter driver then reroutes traffic destined for that device to the new address while the devices are still online.

Dynamic tracking of FC devices is controlled by a new fscsi device attribute, `dyntrk` (default value is `no`). To enable dynamic tracking of FC devices, change all fscsi device attributes to `dyntrk=yes`.

Object Data Manager (ODM)

- The two minimum ODM versions that are supported are:
 - EMC.AIX.6.0.0.4.tar.Z => For AIX 7.1
 - EMC.AIX.5.3.0.9.tar.Z => For AIX 6.1

Installing the EMC XtremIO ODM fileset

- Verify the AIX version number
- Download the correct ODM fileset from EMC
- Uncompress the file
- Create a table of contents file
- Install the ODM fileset

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

69

The Object Data Manager (ODM) is a database of system and device configuration information. The AIX Object Data Manager [ODM] must be updated with configuration attributes for EMC storage devices. To do so, verify the AIX version used by the host, and, based on that version, download the correct ODM file set from EMC. Uncompress the .tar.Z file. Create a table of contents, and install the ODM file set.

The details of the procedure may be found in the XtremIO Storage Array User Guide.

EMC PowerPath Support

- AIX native multipathing is not supported
- PowerPath version 5.7 SP1 is supported
 - Please refer to the latest *XtremIO Simple Support Matrix* for more information
- Install PowerPath
- Verify the installed PowerPath version
 - `powermt version`

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration



70

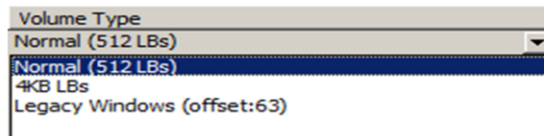
XtremIO supports EMC PowerPath multi-pathing on AIX using PowerPath 5.7 SP1 (or above). PowerPath handles XtremIO devices as native devices and there are no requirements for a generic LAM implementation. To install PowerPath, uncompress, untar, and install the PowerPath fileset by running the following commands:

- `gunzip EMCPower.AIX.5.7.SP1.b013.tar.gz`
- `tar -xvf EMCPower.AIX.5.7.SP1.b013.tar`
- `rm .toc`
- `inutoc .`
- `installp -ad . EMCpower`

Using XtremIO Storage with AIX

- XtremIO volumes for AIX must have a 512 byte block size
- 4 kB block sizes are not supported on AIX
- Configure new volumes with `cfdmgrp`
- Display configured volumes with `lsdev -Cc disk`

XtremIO Volume Creation Options



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Host Configuration

71

When you create volumes in XtremIO for an AIX host, the default logical block (LB) size of a new volume is 512 bytes. This parameter should be left at the default value; 4 kB blocks are not supported on AIX.

Volumes can be configured with the `cfdmgrp` command, and verified with the `lsdev -Cc disk` command after configuration.

Lesson Summary

During this lesson the following topics were covered:

- The provisioning of storage to AIX hosts.



This lesson covered storage provisioning to AIX hosts.

During this lesson you saw how to change the execution throttle for each HBA in the host; how to increase path queue depth and modify the size of the system block I/O buffer cache; how to configure Fast I/O Failure; XtremIO SolutionPack Basics; and finally, you saw how to create volumes in XtremIO for an AIX host.

Module Summary

Key points covered in this module:

- Configuring volumes, initiators, and volume mapping
- Identifying host connectivity requirements:
 - VMware
 - Windows
 - Linux
 - Solaris
 - AIX



This module covered the XtremIO provisioning process.

We reviewed configuring volumes, initiators, and volume mapping; in addition, we also looked at identifying host connectivity requirements for VMware, Windows, Linux, Solaris and AIX.

This slide is intentionally left blank.



Module: XtremIO Features

Upon completion of this module, you should be able to:

- Describe Data at Rest Encryption [D@RE]
- Discuss XtremIO compression
- Discuss XtremIO deduplication
- Configure snapshots on an XtremIO volume
- Identify XtremIO snapshot considerations and limitations



This module focuses on XtremIO Data at Rest Encryption, compression, and deduplication, as well as XtremIO Snapshots.

This slide intentionally left blank.

Lesson: Data at Rest Encryption

This lesson covers the following topics:

- Data at Rest Encryption Overview
- Data at Rest Encryption implementation



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

3

This lesson covers Data at Rest Encryption and how it is implemented on an XtremIO storage array.

Data at Rest Encryption Overview

- XtremIO uses Self Encrypting Drives (SEDs) to protect data
 - Encrypts data on DAE and Storage Controller SSDs
 - Protects data even if SSD is removed from array
- Available on both 20 TB and 10 TB models
- Once encryption is enabled, an Authentication Key PIN is generated on cluster
 - Used to lock/unlock SSD encryption
 - PIN can be changed without data loss if required



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

4

Data at Rest Encryption protects data stored on the XtremIO. It encrypts the data on the DAE and Storage Controller Solid State Drives (SSDs) and continues to protect the data even if the SSD is removed from the cluster. Encryption is disabled by default.

The encryption is based on Self Encrypting Drives (SED). These drives have a specially designed hardware component that performs the encryption without causing any performance degradation. The encryption key, or Data Encryption Key (DEK), is generated on the drive and never leaves the drive. When encryption is enabled, the cluster generates an Authentication Key (AK) PIN that is used to lock and unlock the SSD encryption. The AK is kept securely in the cluster and is used to unlock the SSD during the system start. Removing an SSD from the cluster locks the SSD and renders the media on it unreadable.

If there is a suspicion that the PIN was compromised or when a periodical key rollover is required by regulations, a cluster re-encryption is performed to change the SSD AK PIN without any data loss.

Encryption States and Modes

Encryption Support

- Enabled: All DAE and Storage Controller SSDs are SEDs
- Disabled: At least one SSD is not an SED

Encryption Mode

- Self: All SEDs are locked
- Disabled: All SEDs are unlocked

• Mode switch status

- None
- During Switch
- Failed
- Incomplete

• Mode switch new mode

- Self - cluster currently locking SEDs
- Disabled - cluster currently unlocking SEDs

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

5

For a cluster to be enabled for encryption, all of the DAE SSDs and the Storage Controller SSDs must be SEDs. If at least one SSD is not an SED, then the encryption support field will display Disabled. Once encryption has been enabled, the Self mode will be displayed, meaning that all SEDs are locked.

When changing modes between self and disabled, you will see one of these Mode Switch Status information messages:

- None: No state change is currently in progress
- During Switch: A state change is in progress
- Failed: A state change has failed (resume required)
- Incomplete: A state change has completed. However, not all components were updated.

The Mode Switch New Mode value provides information if the cluster is currently unlocking the SEDs with a **disabled** value, or currently locking the cluster with a self value.

Cluster Encryption GUI Page

Icon	Description
	Encryption is disabled on the cluster.
	Encryption has failed on the cluster.
	The cluster is encrypted.

To access the Cluster Encryption GUI page, select the Cluster tab and then select the **Cluster Encryption** option from the left pane of the Cluster tab.

Note: The Cluster Encryption screen is grayed out (inactive) because encryption is performed at the Technician user level. Customers are informed by the User's Guide that in order to configure the cluster encryption, they should contact EMC support. The encryption state of the cluster is indicated by a 'lock' icon in the Hardware window.

Changing the encryption PIN is also an EMC support activity, and cannot be performed by customers.

Encryption Enabled

The screenshot displays the EMC XtremIO management interface. The 'Cluster Encryption' tab is active, showing 'Encryption is supported' and 'Current encryption status: Encrypted (internal key)'. The 'Hardware' tab shows the PSNT Serial Number: XIOAPLAB00014 and a shield icon with a key, indicating encryption is enabled. Below the interface, the output of the 'show-clusters-info' command is shown, with key fields highlighted in red boxes.

```

xmcli (tech)> show-clusters-info
Cluster-Name Index State Conn-State Activation-Time Start-Time SW-Version
Cluster02 1 active connected Sun Nov 2 11:36:02 2014 Mon Nov 3 19:15:47 2014 3.0.0-44

PSNT
APM0014xxxxxxxxx Encryption-Mode Encryption-Supported Mode-Switch-Status Mode-Switch-New-Mode
self Yes none self
  
```

© Copyright 2015 EMC Corporation. All rights reserved. Module: XtremIO Features 7

In this example, encryption is possible because all SSDs in the cluster are SEDs. Therefore the current encryption status is "Encrypted". Note on the Hardware page a shield with a key is shown in the upper right corner above the hardware.

This information can also be found using show-clusters-info command. Here is the output from the command. Notice the Encryption-Mode is **self**, Encryption-Supported is **Yes** and Mode-Switch-New-Mode is **self**.

Encryption Disabled

Cluster Encryption

Cluster Encryption

Encryption is not possible. Encryption can be applied only if supported by all SSDs in the cluster (including the Local Disk SSDs).

Current encryption status : Disabled

Switching to :

Switching status :

Set Encryption to

Disabled

Apply

Undo

Re-Encrypt

Change Encryption PIN

Hardware

PSNT Serial Number: LAB123456789



```
xmcli (tech)> show-clusters-info
```

Cluster-Name	Index	State	Conn-State	Activation-Time	Start-Time	SW-Version
Cluster02	1	active	connected	Sun Nov 2 11:36:02 2014	Mon Nov 3 19:15:47 2014	3.0.0-44

PSNT	Encryption-Mode	Encryption-Supported	Mode-Switch-Status	Mode-Switch-New-Mode
APM0014xxxxxxxxx	disabled	No	none	disabled

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

8

In this example, encryption is not possible because all SSDs in the cluster are not Self Encrypting Drives (SEDs). Therefore the current encryption status is **disabled**.

This information can also be found using show-clusters-info command. Notice the Encryption-Mode is **disabled**, Encryption-Supported is **No** and Mode-Switch-New-Mode is **disabled**.

SSD Encryption Status

- Verify the encryption status of each SSD

```
xmcli (tech)> show-ssds
```

SSD-Name	Brick-Name	Slot #	...	SSD-Size	DPG-Name	State
wwn-0x5000cca02b03061c	X1	0		745.223G	X1-DPG	in_rg
wwn-0x5000cca02b030030	X1	1		745.223G	X1-DPG	in_rg
wwn-0x5000cca02b02fdbbc	X1	2		745.223G	X1-DPG	in_rg
wwn-0x5000cca02b048654	X1	3		745.223G	X1-DPG	in_rg

Position-State	Endurance-Remaining-%	Certainty-State	SSD-Encryption-Status
good	100	ok	enc_unlocked
good	100	ok	enc_unlocked
good	100	ok	enc_unlocked
good	100	ok	enc_unlocked

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

9

Shown here is an abbreviated output of the show-ssds command. With this command you can verify which SSDs are locked or unlocked.

Lesson Summary

During this lesson the following topics were covered:

- Data at Rest Encryption Overview
- Implement Data at Rest Encryption



This lesson described Data at Rest Encryption and explained how it is implemented on an XtremIO storage array.

Lesson: Compression and Deduplication

This lesson covers the following topics:

- XtremIO compression
- XtremIO deduplication



© Copyright 2014 EMC Corporation. All rights reserved.

Module: XtremIO Features

11

This lesson focuses on XtremIO compression and deduplication, two features that improve data storage efficiency on XtremIO systems. Note that both features are covered in more detail in pre-requisite material, and in white papers, and that both features are always enabled, and are not manageable by the end user.

XtremIO Compression Overview

- Compresses data only when it has been deduplicated
 - No wasting of cycles on duplicate data
- Inline and always on
- Contributes to XtremIO overall data storage efficiency

Storage	
Overall Efficiency	6.4:1
• Data Reduction Ratio	3.6:1
• Deduplication	3.6:1
• Compression	1.0:1
• Thin Provisioning Saving	41%

EMC²

© Copyright 2014 EMC Corporation. All rights reserved.

Module: XtremIO Features

12

XtremIO compression operates on already deduplicated data to ensure that no cycles are wasted on the compression of duplicate data. This feature is always on, is performed inline, and is not manageable by the end user.

Compression, along with deduplication and thin provisioning, contributes to XtremIO data storage efficiency.

XtremIO Deduplication Overview

- Inline, always-on feature
- Makes use of the XtremIO block fingerprinting mechanism
- No overhead or performance impact
- Contributes to the XtremIO overall data storage efficiency

Storage	
Overall Efficiency	6.4:1
• Data Reduction Ratio	3.6:1
• Deduplication	3.6:1
• Compression	1.0:1
• Thin Provisioning Saving	41%

EMC²

© Copyright 2014 EMC Corporation. All rights reserved.

Module: XtremIO Features

13

XtremIO deduplication is an inline process that makes use of the fingerprinting mechanism used to identify data blocks. Duplicate data is not written to the SSDs; instead, the in-memory metadata is updated to indicate that there are multiple references to the data block.

Since the process is performed in memory, and makes use of existing processes, it adds no latency or performance overhead to the system. This feature is always on, cannot be stopped, and cannot be managed by the end user.

This slide intentionally left blank.

Lesson: XtremIO Snapshots

This lesson covers the following topics:

- Snapshot configuration
- Snapshot considerations and limitations



This lesson focuses on XtremIO snapshots. We will cover the snapshot creation process, management, and snapshot considerations and limitations.

XtremIO Snapshots Overview

- Captures the state of data in a volume at a particular point in time
- Snapshots can be accessed like any other volume in the cluster
- Snapshots are read/write, but may be mounted read-only from host
- Use cases for Snapshots
 - Backup and data recovery
 - Development and test
 - Clones
 - Offline processing



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

16

XtremIO snapshots are created by capturing the state of data in volumes at a particular point in time and allowing users to access that data when needed, even when the source volume has changed. Snapshots can be accessed like any other volume in the cluster, and can be read and written. Users can choose to mount the snapshot in read-only mode from the host in order to maintain its consistency.

Snapshots can be used in a number of use cases, including:

- Data recovery
- Backup
- Development and Test
- Clones where persistent writeable snapshots are used
- Offline processing to offload the processing of data from the production server to another less utilized server

Snapshot Creation Overview

- When a snapshot is created, two new volumes are created from the snapshotted volume
 - Ancestor volume (source of snapshot)
 - Host will read and write to this volume from this point on
 - Snapshot volume

Available snapshot types

- Snapshot of a volume
- Snapshot of a snapshot
- Snapshot of a set of members (volumes and/or snapshots)
- Snapshot of a folder containing volumes or snapshots

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

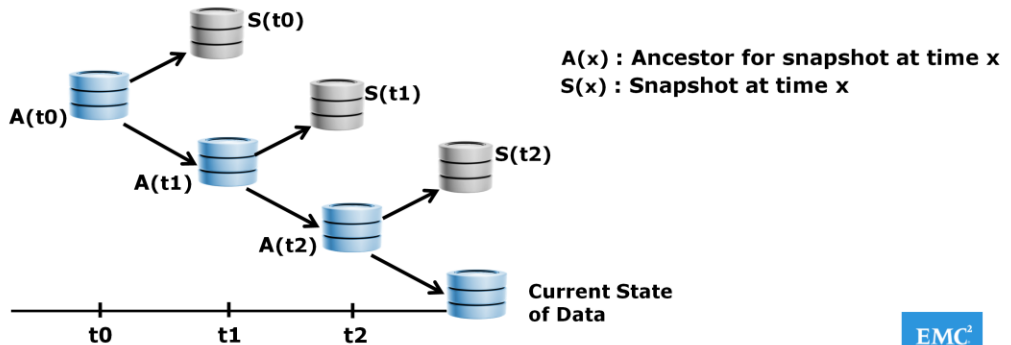
17

When a new snapshot is created, XtremIO always creates two new volumes from the snapshotted volume. One of the volumes is the snapshot itself, and the other volume becomes the source volume for the snapshot, also known as the ancestor volume. This new ancestor volume is now the volume that will be accessed by the host. You can create snapshots from the following volumes:

- Snapshot of a volume
- Snapshot of a snapshot
- Snapshot of a set of members (volumes or snapshots or a combination of both)
- Snapshot of a folder containing volume(s) or snapshot(s)

Capacity Considerations

- Creating a snapshot is immediate and consumes no space
 - A pointer is generated for the snapshot to the ancestor metadata
 - No performance degradation
 - Capacity consumption only occurs when writing a unique block



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

EMC²

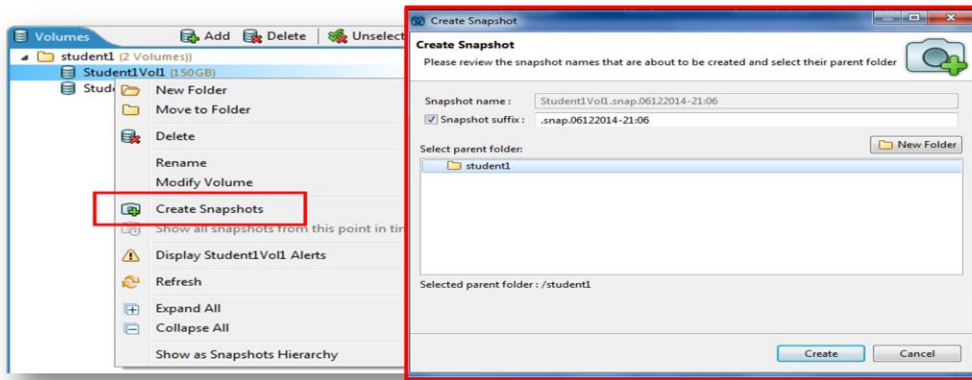
18

When creating a snapshot, XtremIO generates a pointer to the ancestor metadata (of the actual data in the system). Therefore, creating a snapshot is a very quick operation that does not have any impact on the system and does not consume any capacity. Snapshot capacity consumption occurs only if a change requires writing a new unique block. Snapshots are implemented using redirect-on-write methodology, where new writes to the source volume (or the snapshot) are redirected to a new location. Only metadata is updated to point to the new data location. The metadata is shared between all snapshot blocks that have not been changed from the snapshot's ancestor. The snapshot maintains unique metadata only for an LBA (Logical Block Address) whose data block is different from that of its ancestor.

On this slide, we see three snapshots that were taken of a volume during different times. Notice that every snapshot has its own ancestor volume.

Creating a Snapshot – GUI

- Right-click the desired volume and select Create Snapshot
 - Multiple volumes can be selected with the Ctrl key or the Shift key



EMC²

Module: XtremIO Features

19

To create a snapshot using the XMS GUI, from the menu bar, click **Configuration**. In the Volumes pane, select a folder. Expand the list of members (volumes and/or snapshots) and select one or several members from the list. To select multiple members listed sequentially, select a member, hold SHIFT and select the additional members. To select multiple individual list members, select a member, hold Ctrl and select the additional list members.

Right-click the selected volume(s), and from the drop-down menu, select **Create Snapshots**. In the Create Snapshot screen, you will need to define the following a snapshot name, select a parent folder (optional), or create a new folder (optional). Then click **Create**. The new snapshot(s) is/are created and appear(s) in the selected folder within the Volumes pane of the Configuration workspace.

Creating a Snapshot – CLI

- Snapshot creation on a single volume

```
xmcli (admin)> create-snapshot ancestor-vol-id="Vol1" snap-vol-name="Vol1_Snap"
```

- Snapshot creation on a list of volumes

```
xmcli (admin)> create-snapshot snap-list=[ancestor-vol-id="Vol1"  
    snap-vol-name="Vol1_Snap", ancestor-vol-id="Vol2" snap-vol-name="Vol2_Snap"]
```

- Snapshot creation on volumes within a folder

```
xmcli (admin)> create-snapshot source-folder-id="/Volumes" suffix="_Snap"
```



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

20

To create snapshots using the XMCLI, the create-snapshot command will need to be used with one of the following command switches:

- the ancestor-vol-id and snap-vol-name to create a snapshot on a single volume
- snap-list to create snapshots on a list or set of volumes/snapshots
- source-folder-id to create snapshots on volumes within a folder

Viewing Snapshot Information

- Snapshot information is displayed per volume group
 - The ancestor volume is also included

```
xmcli (admin)> show-volume-snapshot-groups
```

Index	Num-of-Vols	Vol-Size	Thin-Provisioning-Ratio	Logical-Space-In-Use	Num-Of-Active-Merges
1	3	450G	0.042	18.808G	0
2	2	200G	0.071	14.263G	0

- Snapshot information per ancestor (output has been shortened)

```
xmcli (admin)> show-snapshots ancestor-vol-id="Vol1"
```

Volume-Name	Vol-Size	LB-Size	VSG-Space-In-Use	Ancestor-Name	Cluster-Name	Parent-Folder-Name
Vol1.snap1	150G	512	18.808G	Vol1	xbc3	/student1
Vol1.snap2	150G	512	18.808G	Vol1	xbc3	/student1



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

21

Snapshot information can be viewed via the XMS GUI or the XMCLI. The GUI will display general volume information. The XMCLI will provide more information regarding the volume snapshot group (VSG) with the show-volume-snapshot-groups command and provide a list of snapshots per ancestor volume via the show-snapshots ancestor-vol-id command. Once a snapshot is created, a snapshot tree size is calculated. This includes the entire logical space of the original volume and all its snapshots.

Snapshot Considerations

- Snapshot deletions have no impact on system performance and SSD media endurance
 - Asynchronous process
 - Deleting a snapshot will result in the merging of metadata if that snapshot is deleted out of order

Description	Maximum Value
# of volumes per cluster (includes snaps)	8192
# of snapshots per production volume	64
# of snapshots from another snapshot	64
# of members in a single snapshot operation	18 for GUI/32 for CLI

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

22

Snapshot deletion is an asynchronous task; it is performed in the background. Deleting a snapshot will result in its metadata being merged if that snapshot is deleted out of order. In other words, if you have three snapshots taken from the same volume at different times, and the middle snapshot is deleted, that metadata will be needed by the first snapshot to keep track of any new blocks that were written at that specific snapshot time. So the deleted snapshot's metadata will need to be merged with the first snapshot's metadata.

The table lists the maximum value for the number of volumes, snapshots, snapshots of another snapshot, and members in a single snapshot operation. The cluster allows taking a snapshot of up to 32 volumes/snapshots simultaneously when using the XMCLI. When using the XMS GUI, only 18 volumes/snapshots are allowed at once. This is to avoid long GUI operations that may time out. Please refer to the latest Release Notes document for up-to-date information.

Snapshot Size Considerations

- It's not possible to view the size of a single snapshot
 - Space is displayed on a snapshot tree basis, or a volume snapshot group (VSG)
- A VSG is made up of the following:
 - One ancestor volume
 - Any number of supported snapshots

```
xmcli (admin)> show-volume-snapshot-groups
```

Index	Num-of-Vols	Vol-Size	Logical-Space-In-Use	Num-Of-Active-Merges
1	3	6T	6T	0
2	3	15T	5T	2
3	3	1.465T	200G	0

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

23

When dealing with snapshots, it's not possible to view the size of a single snapshot. Once a snapshot is created, a snapshot tree size is calculated. This includes the entire logical space per the ancestor volume and all of its snapshots. This snapshot tree is also known as volume snapshot group (VSG).

Lesson Summary

During this lesson the following topics were covered:

- Snapshot configuration
- Snapshot considerations and limitations



This lesson covered snapshot creation process, management, and snapshot considerations and limitations.

Lab Introduction: XtremIO Snapshots

This lab covers

- Snapshot configuration
- Data recovery using a snapshot



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

25

In this lab exercise we will be implementing XtremIO Snapshots on an ESXi volume. We will take a snapshot of a Datastore, then use that snapshot for data recovery.

Lab Debrief

This lab covered the process of creating an XtremIO snapshot on an ESXi volume and using that snapshot for data recovery.



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Features

26

This lab covered the creation of a snapshot of an ESXi datastore, and the recovery of that datastore after data is deleted.

Module Summary

Key points covered in this module:

- XtremIO encryption keeps data secure if drives are removed from the system
- Compression and encryption improve data storage efficiency
- Snapshots instantaneously capture the state of data in a volume at a particular time and consume no space
- New writes to volume or snapshot are redirected to a new location



This module covered the snapshot creation process, management, and snapshot considerations and limitations.

This slide intentionally left blank.



Module: XtremIO Monitoring with XMS

Upon completion of this module, you should be able to:

- Define monitoring operations
- Monitor the performance of an XtremIO array



This module focuses on monitoring of an XtremIO storage array.

This slide intentionally left blank.



Lesson: Monitoring Operations Overview

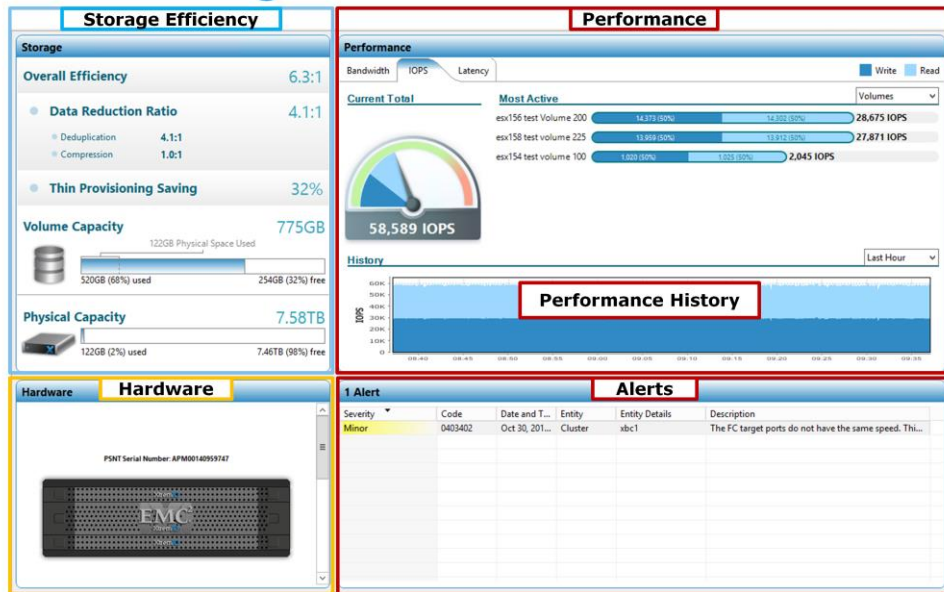
This lesson covers the following topics:

- Common monitoring statistics



This lesson provides an overview of the common monitoring statistics with an XtremIO storage array, and how to navigate the GUI to complete performance monitoring tasks, such as checking overall efficiency, data reduction ratio, thin provisioning savings, and volume capacity.

Dashboard Page: Review



This is the Dashboard Workspace that is first displayed when the Management GUI is opened.

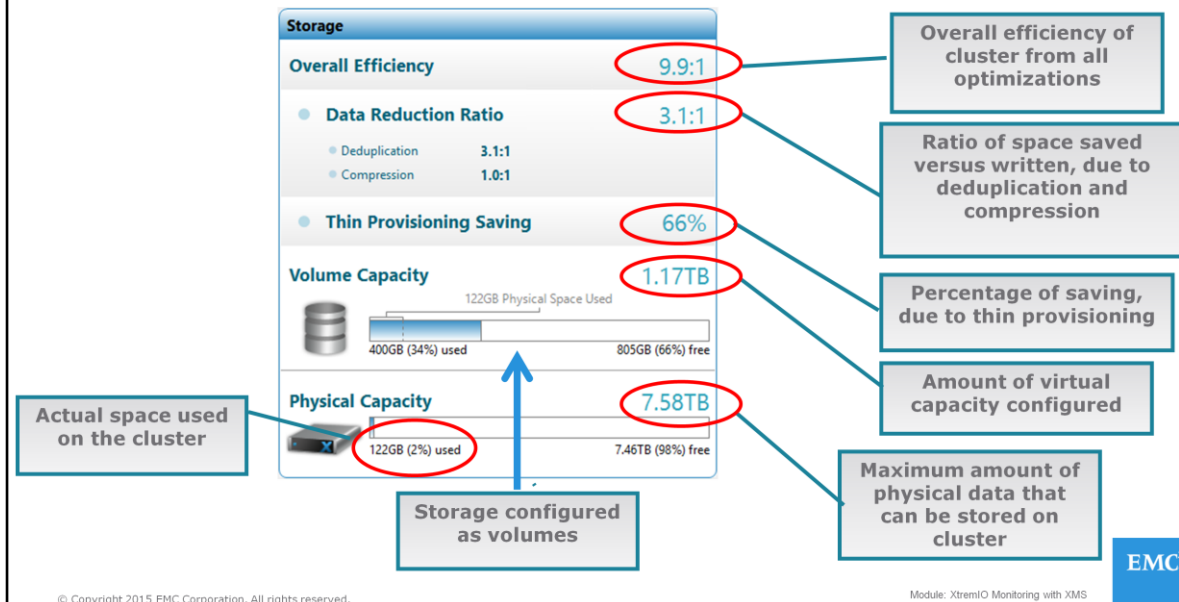
The upper left section of the page displays the storage efficiency for the cluster.

The upper right section of the page displays the general performance of the cluster with a performance history graph.

The lower left pane provides a quick overview of the status of the hardware along with the PSNT (Product Serial Number Tag) if set for the cluster.

The pane in the lower right displays all alerts that have not yet been acknowledged beginning with the most important.

Dashboard Page: Storage Statistics



In the Storage Pane, users can monitor the overall usage of the cluster and its efficiency in storing data using the following parameters:

- **Overall Efficiency** is the ratio between the provisioned space and the physical space.
- **Data Reduction Ratio** is the ratio between the data written to unique addresses and the physical space used to store that data – which will be less due to deduplication and compression. The overall data reduction ratio is broken down further into a deduplication ratio and a compression ratio.
- **Thin Provisioning Savings** is the ratio between the unused address space and the provisioned space.
- **Volume Capacity** is the total amount of disk space that has been configured as LUNs or volumes.

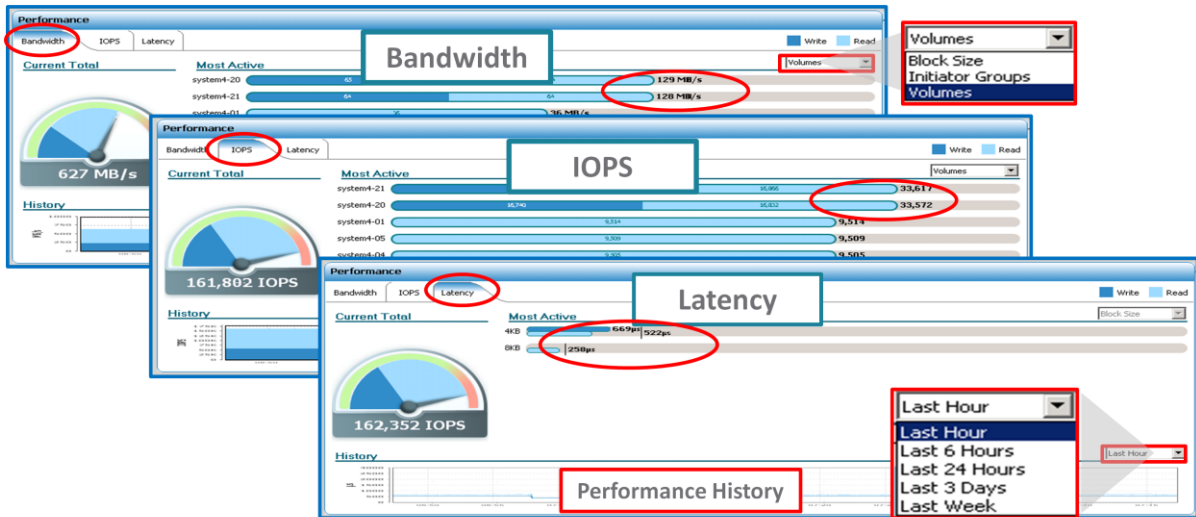
The bar depicts how much space has been assigned to LUNs. The shaded area indicates how much logical volumes space has been used. The dashed line inside the shaded area of the bar graph displays the amount of physical space that has been consumed with unique data. In this example, a volume capacity of 1.17 TB has been configured. 400 GB of logical space has been used, and 122 GB of physical space.

- **Physical Capacity** is the amount of actual physical data that can be stored on the cluster.

Let's look at an example of the calculations for these features: A user has provisioned 5 X 500 GB LUNs to a host, for a total of 2.5 TB. The host has written 1.5 TB of data to the cluster of which 1 TB is unique. The data is compressed at a 2:1 ratio, meaning that 500 GB of data is written to disk.

DRR = 1.5 TB / 500 GB = 3:1, TPS = 2.5 TB – 1.5 TB / 2.5TB = 40%, and Overall Efficiency = 2.5 TB / 500 GB = 5:1.

Dashboard Page: Performance Statistics



The Performance pane on the Dashboard reports real-time performance data in three main categories. The Current Total gauge indicates the total transfer rate for all volumes by MB/s or IOPS, depending on which tab is being accessed.

- The Bandwidth tab displays the system performance with bandwidth as the unit of measure (in MB/s). The Most Active section can display up to seven Volumes, Initiator Groups, or active Block Size, based on the option chosen in the drop-down list
- The IOPS tab displays the system performance with IOPS as the unit of measure. Its measurement also can be based on Volumes, Block Size, or Initiator Groups.
- The Latency tab displays the system performance with latency as the unit of measure in micro-seconds (µs).

For all tabs, the pale blue color indicates reads, and the dark blue color indicates writes.

The History Graph at the bottom of each tab displays performance history per Block Size, Initiator Groups, or Volumes, as selected in the item Selector. The Time Period selector enables users to select the desired time period from the drop-down list with choices from the Last Hour, up to the Last Week.

Some useful CLI commands to use when logged in to the XMS that show similar performance data are:

- `show-most-active`: Displays the most active volumes and initiator groups.
- `show-most-active-initiator-groups`: Displays performance data of the most active initiator groups.
- `show-most-active-volumes`: Displays performance data of the most active volumes.
- `show-volumes-performance`: Displays volume performance data.
- `show-volumes-performance-small`: Displays volume performance data for small (under 4 kB) blocks.
- `show-volumes-performance-unaligned`: Displays volume performance data for unaligned blocks.
- `show-volume-folders-performance`: Displays volume folder performance data.
- `show-volume-folders-performance-small`: Displays volume folder performance data for small (under 4 kB) blocks.

Dashboard Page: Alerts

The screenshot shows the '2 Alerts' dashboard. The table lists two alerts: a Minor alert from Oct 9, 2013 10:16:08 AM and an Information alert from Oct 8, 2013 3:59:47 AM. Both alerts are related to 'Storage Co...' and 'X1-SC2 in X1'. The 'Alert Properties' window shows details for the 'Node Temp Level 2 Unknown' alert, including its severity (Minor), time, entity, and description. The 'Edit Alert Definition' window allows users to modify the alert's description, activity mode (Enabled), clear mode (Auto Clear), and severity (Minor).

Severity	Date and Time	Entity	Entity Details	Description
Minor	Oct 9, 2013 10:16:08 AM	Storage Co...	X1-SC2 in X1	Temperature information is unavailable. The system is unable to read the temperature information.
Information	Oct 8, 2013 3:59:47 AM	Storage Co...	X1-SC1 in X1	Storage Controller's firmware version is incorrect. The system will try to use the Storage Controller.

Alert Properties

Alert : Node Temp Level 2 Unknown
 Severity : Minor
 Time : Oct 9, 2013 10:16:08 AM
 Entity : Storage Controller
 Entity Details : X1-SC2 in X1
 Description : Temperature information is unavailable. The system is unable to read the temperature information.

System Impact :
 Repair Flow :
 Acknowledge Alert Edit Alert Definition OK

Edit Alert Definition

Alert : Node Temp Level 2 Unknown
 Description : Temperature information is unavailable. The system is unable to read the temperature information.

Activity Mode : ☒ Enabled
 Clear Mode : ☒ Auto Clear ☐ Acknowledgment Required
 Severity : Minor

OK Cancel

© Copyright 2015 EMC Corporation. All rights reserved. Module: XtremIO Monitoring with XMS 7

The Alerts pane displays the current system alerts, which are color coded according to severity. Green is a Cleared alert, blue is an Information alert, yellow indicates a Minor alert, orange is a Major alert, and red is a Critical alert.

For each alert, the following details are listed: Severity, Date and Time, Entity, Entity Details, and Description. You can sort the displayed alerts by each of the column headings. Double-clicking an alert will open the Alert Properties page. This will display all the information listed above and will allow the user to either acknowledge the alert or edit the alert definition. To acknowledge the alert, click the Acknowledge Alert icon. Once acknowledged, the alert will no longer appear on the Dashboard page.

The user can change the alert definition by selecting the Edit Alert Definition icon. From here, users can edit the definition, set Activity Mode, set Clear Mode, and change the severity of the alert. The user might decide to do this for a particular alert to make it more meaningful or to set the alert's severity to a higher level depending on their needs.

Dashboard Page: Hardware

The screenshot shows the 'Hardware' pane in the dashboard. It displays a graphic of a two-X-Brick cluster. A tooltip for a 'Major' alert is shown, indicating that the InfiniBand Switch is disconnected and its firmware version is incompatible. Below the hardware graphic, a 'Levels of Severity Color Codes' legend is provided, showing four categories: Critical (Red), Major (Orange), Minor (Yellow), and Information (Blue). A second tooltip for a 'Minor' alert is shown, indicating that the BBU state is uninitialized. The EMC logo is visible in the bottom right corner of the dashboard area.

Hardware

Major
IB-SW1
Health State :Uninitialized
Lifecycle State : Disconnected
Serial Number :
Part Number : ****

Alerts :
InfiniBand Switch is disconnected.
InfiniBand Switch firmware version is incompatible. InfiniBand Switch cannot be used.
(Total 2 alerts)

Levels of Severity Color Codes

Critical **Major** **Minor** **Information**

Minor
X1
BBU state UninitializedBBU X1-BBU1
Alerts :
XMS has detected an X-Brick that is not configured to be part of a system.
(Total 1 alerts)

EMC²

© Copyright 2015 EMC Corporation. All rights reserved. Module: XtremIO Monitoring with XMS 8

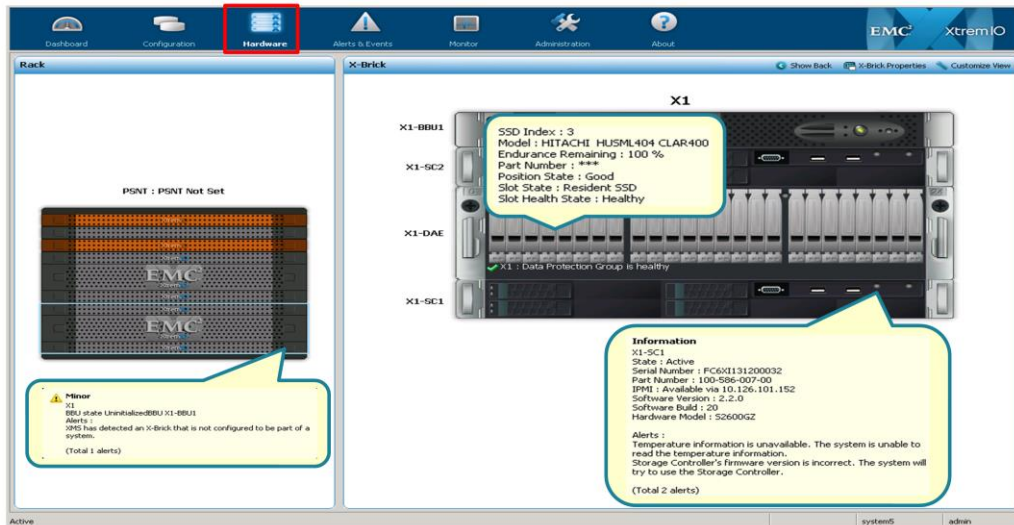
Shown here is the Hardware pane in the Dashboard Workspace that displays a graphic illustration of the system hardware components in a two-X-Brick cluster. If a hardware component suffers from an alert, its color changes according to the alert's severity. Hovering the mouse over the hardware component displays a ToolTip with information on the hardware status, as shown in this Major alert.

Here we see a single X-Brick cluster that is colored yellow and the Minor alert that caused the display to change color. The levels of alert severity include:

- Blue - Information alert
- Yellow - Minor alert
- Orange - Major alert
- Red - Critical alert

Clicking the hardware graphic displays the hardware workspace.

Hardware Page: Overview



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

EMC²

9

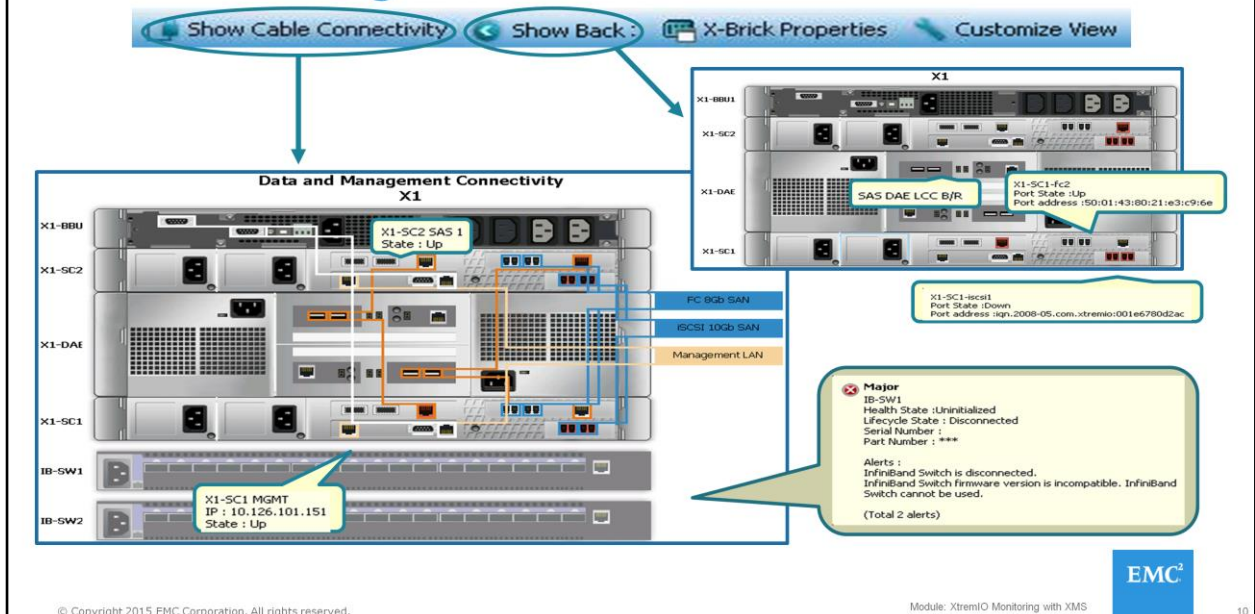
The Hardware Workspace can be accessed two ways. The first is by clicking the cluster image in the Dashboard Workspace. The second is by selecting the Hardware icon from the Menu. The Rack pane on the left, displays all the components that make up the cluster. If the cursor passes over parts of the cluster, any alert for that component will appear. For a single X-Brick cluster, the X-Brick will be shown in the right pane.

In a multiple-X-Brick cluster, whichever X-Brick or InfiniBand (IB) switch is selected in the Rack pane, appears in the right pane. In this example, X-Brick X1 is selected on the left and appears in the pane on the right. Like the left pane, if the cursor is hovered over any main assembly in the right pane, any alerts for it will be displayed. If you mouse-over individual components their properties will be displayed.

Some useful XMCLI commands when logged in to the XMS server are:

- `show-ssds` - Displays a list of SSDs in the cluster and their properties.
- `show-ssds-performance` - Displays the SSDs performance data.
- `show-slots` - Displays a list of SSD slots and their properties.
- `show-storage-controllers` - Displays the cluster SCs information and status.
- `show-storage-controllers-fw-versions` - Displays the SCs firmware version information.

Hardware Page: Cable and Hardware Details



The right pane in the Hardware Workspace can be used to see other views of the selected hardware. The Menu Bar above the right pane has buttons for different viewing options. The Show Back icon will show the rear view of the selected component. Moving your cursor over the ports will show the settings for the ports, such as the 8 Gb/s FC port WWN, or the 10 GbE port IQN information, as well as the port state.

To determine which ports are for FC, iSCSI, or management use, click the Show Cable Connectivity icon from the Menu. This will display the port connectivity scheme. As with other views, if you mouse-over each component it will display its settings or other information.

Some useful XMCLI commands when connected to the XMS server are:

- `show-storage-controllers-psus` - Displays the SCs power supply units information.
- `show-storage-controllers-info` - Displays the cluster SCs information.

Hardware Page: X-Brick Properties

[Show Cable Connectivity](#)
[Show Front](#)
[X-Brick Properties](#)
[Customize View](#)

Protection Group

ID	Name	Size	Health	Position
0	In Data Protection Group	372GB	Healthy	Good
1	In Data Protection Group	372GB	Healthy	Good
2	In Data Protection Group	372GB	Healthy	Good
3	In Data Protection Group	372GB	Healthy	Good
4	In Data Protection Group	372GB	Healthy	Good
5	In Data Protection Group	372GB	Healthy	Good
6	In Data Protection Group	372GB	Healthy	Good
7	In Data Protection Group	372GB	Healthy	Good
8	In Data Protection Group	372GB	Healthy	Good
9	In Data Protection Group	372GB	Healthy	Good
10	In Data Protection Group	372GB	Healthy	Good
11	In Data Protection Group	372GB	Healthy	Good
12	In Data Protection Group	372GB	Healthy	Good
13	In Data Protection Group	372GB	Healthy	Good
14	In Data Protection Group	372GB	Healthy	Good
15	In Data Protection Group	372GB	Healthy	Good
16	In Data Protection Group	372GB	Healthy	Good
17	In Data Protection Group	372GB	Healthy	Good

SSD List

ID	Name	Size	Health	Position
0	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
1	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
2	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
3	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
4	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
5	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
6	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
7	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
8	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
9	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
10	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
11	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
12	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
13	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
14	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
15	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
16	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good
17	Hitachi HUSPL404 CLAR400	372GB	Healthy	Good

SC-1

Lower Node
Name: X1-SC1
Status: Active
Manager IP: 10.126.101.151
Infiniband I2 IP: 169.254.0.1 / 169.254.0.2
BBU Status: OK (last 0, 105-000-293-00)
BBU Battery: Available via 10.126.101.152
BBU Voltage: OK (last 0, 105-000-293-00)
Boot disk: OK (last 0, 105-000-293-00)
Journal disk: OK (last 0, 105-000-293-00)
Trace disk: OK (last 0, 105-000-293-00)

SC-2

Upper Node
Name: X1-SC2
Status: Active
Manager IP: 10.126.101.153
Infiniband I2 IP: 169.254.0.17 / 169.254.0.18
BBU Status: OK (last 0, 105-000-293-00)
BBU Battery: Available via 10.126.101.154
BBU Voltage: OK (last 0, 105-000-293-00)
Boot disk: OK (last 0, 105-000-293-00)
Journal disk: OK (last 0, 105-000-293-00)
Trace disk: OK (last 0, 105-000-293-00)

Customize View

Hardware
☒ Show Fault Leds

Cable Connectivity
Select line type for display connectivity :
☒ Rectangular
☐ Curly
☐ Direct

Select what connectivity type group to show :
☒ Data/Management Connectivity
☐ Power Connectivity

Select what connectivity type to show :
☒ IB Connection
☒ COM Connection
☒ FC Connection
☒ iSCSI Connection
☒ MGMT Connection
☒ SAS Connection

☐ All connectivity type show with a same default color
☒ Each connectivity type show with a system defined color

Keep changes Set default Cancel

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

11

Two other options are available from the right Hardware pane.

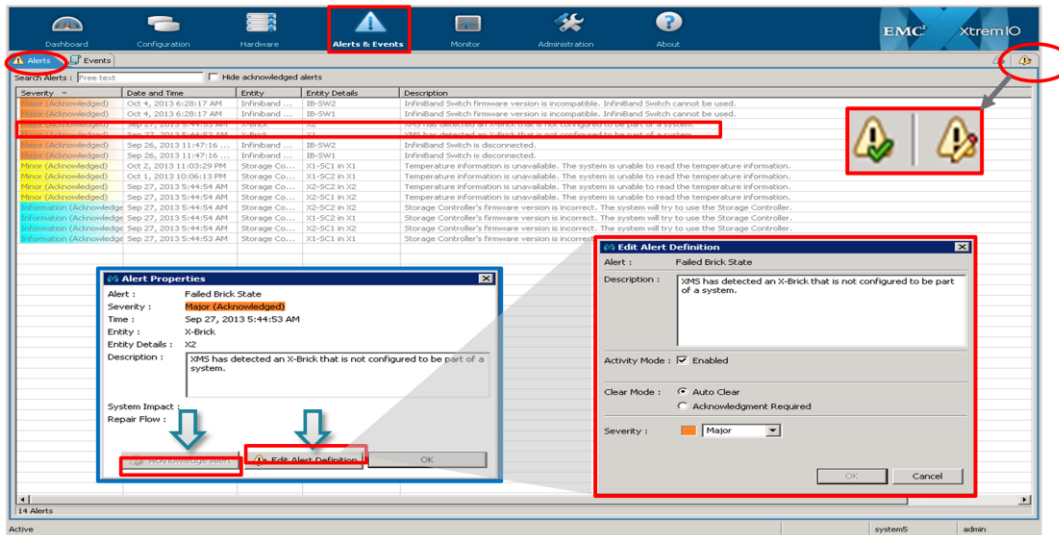
The first is the X-Brick Properties option. This will display the properties and status of the X-Brick selected, including the Protection Group, the SSD list, and SC 1 and SC 2.

The last menu item is the Customize View option. From here users can select whether to show the Fault LEDs, the type of cable connections to display connectivity with, and the type of connectivity to display. Users can choose whether to have different colors for the different types of connectivity or make them all the same.

A useful XMCLI command for similar information when logged in to the XMS server is:

- `show-bricks` - Displays a list of X-Bricks and their associated cluster.
- `show-ssds` - Displays a list of SSDs.
- `show-slots` - Displays drive slots.

Alerts and Events Page: Alerts



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

12

The Alerts and Events Workspace contains two separate tabs, one for Alerts and the other for Events. The Alerts and Events Workspace is accessed by selecting the Alerts and Events icon from the Menu Bar.

The Alerts tab of the Alerts and Events Workspace can be used to display all the alerts on the cluster. If an alert is double-clicked, it will open the Alert Properties Page for that alert which will display the name of the alert, the severity of the entity the alert was on, the name of the entity, and a detailed description. If the alert has been acknowledged already, the Acknowledge Alert button, shown here, will be grayed out. Otherwise, this button can be used to acknowledge the alert and remove it from the Dashboard's Alert pane. The alert can also be modified by using the Edit Alert Definition button. This will open the Edit Alert Definition window where the Description, Activity Mode, Clear Mode, and Severity can be modified.

The alerts listing can be sorted by clicking on the headers for each column. Specific alerts can be searched for by using the search window. The icons in the top right of the Alerts pane can be used to acknowledge alerts or to select definitions to be edited.

Some useful XMCLI commands for similar functionality when logged in to the XMS server are:

- `show-alerts` - Displays a list of active alerts and their details.
- `show-alert-definitions` - Displays a list of pre-defined alerts and their definitions.

Alerts and Events Page: Events

The screenshot shows the EMC XtremIO Alerts and Events page. The 'Events' tab is selected. The top section contains filters for severity (Critical, Major, Minor, Information), date/time, category (Software, Hardware, Audit Log, Activity, Security, Life Cycle), and entity. A search description field is also present. Below the filters is a table of events with columns for severity, date/time, category, entity, and description. A dropdown menu for the entity filter is open, showing a list of entities including Alert, Alert Definition, Infiniband Switch, Initiator, Initiator Group, Initiator Group Folder, DAE, DAE Controller, DAE PSU, Local Disk, Module, Storage Controller, Storage Controller PSU, Data Protection Group, SSD, Snapshot Group, Cluster, Target, Target Group, Battery Backup Unit, and User Account.

The Events pane of the Alerts and Events Workspace is used for viewing cluster events. The top part of the pane is used to select different criteria for filtering the events to view, such as severity, date, time, category and entity. The entity field has a drop-down list that provides all the entities that can be selected.

Data Collection for EMC Technical Support

- To collect XtremIO dossier log files:
 - At **xmcli (admin)>** type **create-debug-info**
 - Use a web browser to navigate to the HTTP address link provided
 - Save the XtremIO dossier log file package
 - Attach the file package to a Service Request
 - Use **show-debug-info** to view listing of log packages available

```
xmcli (tech)> show-debug-info
Name Index System-Name Index Debug-Level Start-Time Create-Time Output-Path
1 XtremIO-SVT 1 medium Thu Mar 14 13:57:20 2013 Thu Mar 14
14:01:40 2013 http://< XMS IP Address >/XtremApp/
DebugInfo/0cae4b1fff88433d9fca7b052b7bb43a.tar.xz
2 XtremIO-SVT 1 medium Thu Mar 14 13:57:20 2013 Thu Mar 14
14:01:40 2013 http://< XMS IP Address
>/XtremApp/DebugInfo/ea5ff1f0f9e54adf9f0784c3d1645311.tar.xz
```



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

14

When the investigation of an XtremIO system is required by EMC, a Global Technical Support representative, may ask you to collect a set of log files. Consult the instruction set provided by the representative or the EMC KnowledgeBase solution *Article Number:000017095 Version:4*. These instructions explain how to collect XtremIO dossier log files and forward them to EMC Global Technical Support.

Lesson Summary

During this lesson the following topics were covered:

- Common monitoring statistics



This lesson covered an overview of the common monitoring statistics with an XtremIO storage array.

This slide intentionally left blank.



Lesson: Monitoring XtremIO Performance

This lesson covers the following topics:

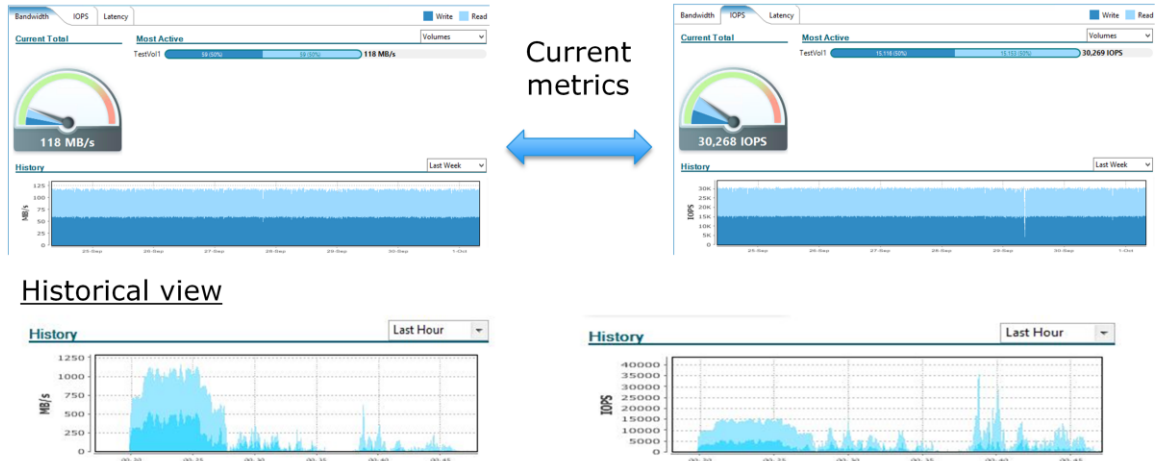
- Monitoring the performance of XtremIO arrays



This lesson covers monitoring an XtremIO storage array for performance.

Basic Performance Monitoring with the GUI

Monitors Bandwidth, IOPS, and Latency



Historical view

These are some example of screenshots from a running system. You can view current system latency, bandwidth, and IOPS. The historical view shows performance from the last hour, 6 hours, 24 hours, 3 days, or 1 week.

Viewing bandwidth, IOPS and latency may not fulfill all monitoring needs. User-defined monitors can be created to allow viewing of other performance and operational parameters.

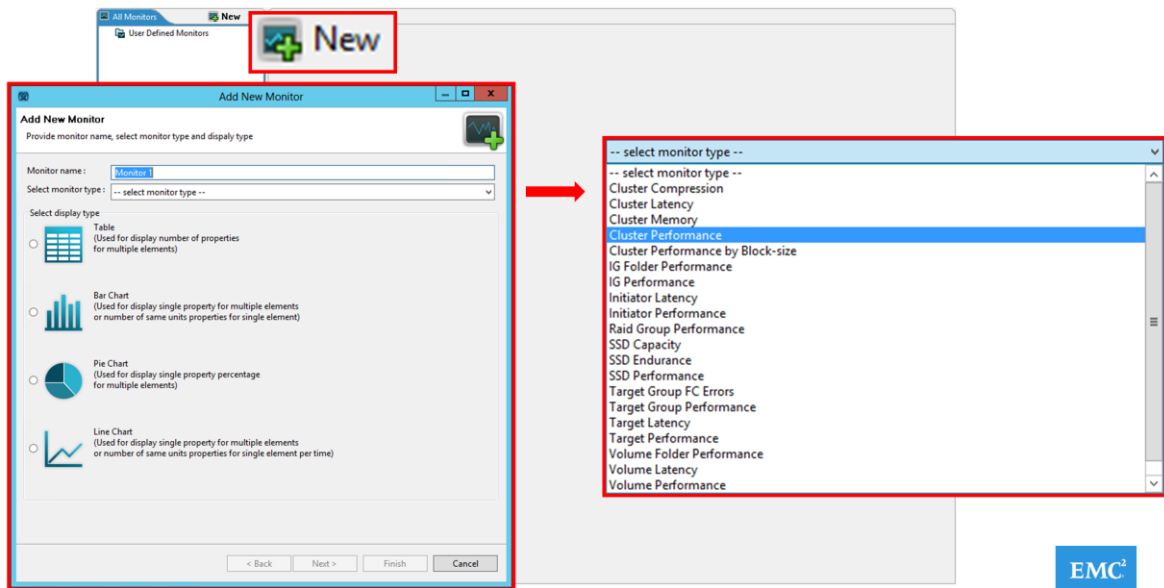
EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

18

Monitor Page: Overview



The Monitor Workspace is accessed by selecting the Monitor icon in the Menu Bar. The left pane contains any User Defined Monitors. The right pane displays the output from the user defined monitors in separate tabs.

To create a new monitor, click the New icon at the top of the User Defined Monitor pane. In the Add New Monitor page, select the display type: Table, Bar Chart, Pie Chart, or Line Chart. Next, from the drop-down window for Monitor Type, select one of the listed objects to monitor, and give the monitor a name.

Creating a Monitor

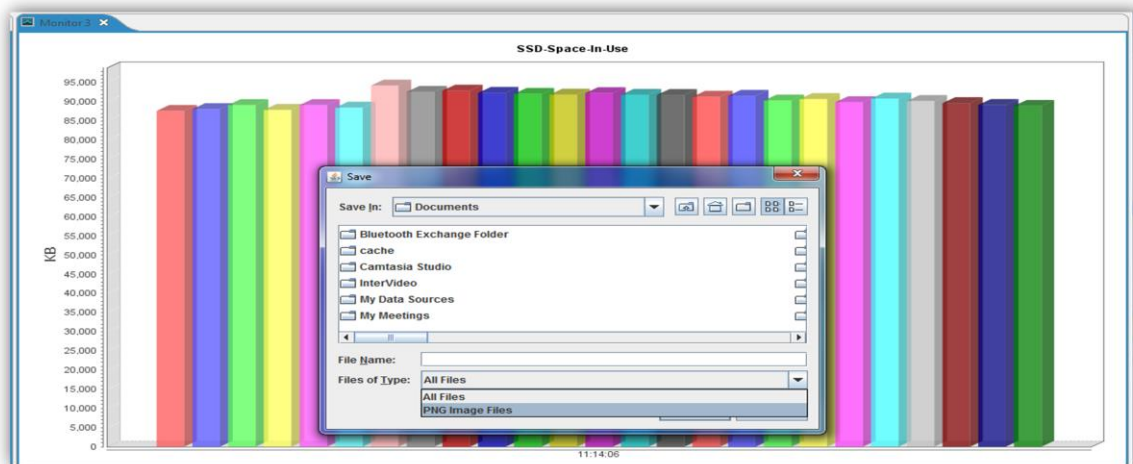
The screenshot illustrates the 'Add New Monitor' process in the EMC XtremIO Monitoring GUI. On the left, the 'Select Objects' pane lists various SSDs, with 'wwn-0x5000cca01306b164' highlighted. The 'Select property to monitor' dropdown is set to 'Endurance-Remaining-%'. The 'Monitor 1' tab is active, displaying a pie chart titled 'Endurance-Remaining % at 83:28:39' showing the distribution of endurance across various SSDs. The chart is divided into segments representing different SSDs, with labels indicating the percentage of endurance remaining for each. A legend at the bottom identifies the SSDs by their WWN values.

The next step to create this monitor type is to select which of the SSDs the user wants to monitor. In the left side of the pane, click the selection boxes to accomplish this. Depending on the type of monitor being created, there may be multiple objects that can be selected. In this example, there is only a single object. When the monitoring template is created ("finish" is pressed); **the monitoring template** will automatically start to run.

Running or stopping the monitoring template can be **achieved by right clicking the task**. If multiple monitors have been created, they will appear on separate tabs. To view several at the same time, the templates can be dragged and dropped for a split view, but when re-opening the GUI the default is the Tabs view.

The monitor display is able to provide up to 30 minutes of information. For volumes, the user can choose to display the read, write, and/or average latency. For initiators and targets, the user can choose to display read, write, and/or average latency per block size.

Saving Performance Data



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremlIO Monitoring with XMS

21

You can save any chart image as a .PNG file. Right click on the desired image to save the file. Tables can also be saved as .CSV files, though that operation can only be performed with the CLI.

Performance Monitoring using XMCLI

- Enter the xmcli for specific commands to monitor performance
 - `show-most-active-initiator-groups`
 - `show-most-active-volumes`
 - `show-ssds-performance`
 - `show-systems-performance`
 - `show-systems-performance-unaligned`
 - `show-targets-performance`
 - `show-volumes-performance-unaligned`
 - `show-xenvs`
- And many other monitor commands



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

22

The XtremIO monitor command with the appropriate extension can be used to monitor various performance parameters.

Show commands can also be used to view performance information. For example:

- `show-most-active`: Displays the most active volumes and initiator groups.
- `show-most-active-initiator-groups`: Displays performance data of the most active initiator groups.
- `show-most-active-volumes`: Displays performance data of the most active volumes.

There many other show commands. For all commands refer to the XtremIO Storage Array User Guide for your version of XtremIO.

Monitoring with show commands

Input Parameters	Description	Value	Manadatory
Duration	Monitor duration	Seconds	No
Frequency	Monitor intervals	Seconds	No

- Input parameters for monitoring using show commands
- Monitoring initiator performance for 3600 s every 5 s

```
xmcli (admin)> show-initiators-performance duration=3600 frequency=5
Initiator-Name  Index Write-BW (MB/s) Write-IOPS Read-BW (MB/s) Read-IOPS
ESXiHBA0       1     113.703       29108     205.773     52678
ESXiHBA1       2     111.965       28663     206.621     52893
Initiator-Name  Index Write-BW (MB/s) Write-IOPS Read-BW (MB/s) Read-IOPS
ESXiHBA0       1     113.940       28964     204.406     52238
ESXiHBA1       2     112.234       28732     205.221     52792
```

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

23

For monitoring with the CLI, the show commands are used with the input parameters duration and frequency. The used values are in seconds. If only one of the parameters is used, the other uses its default.

In this example, both parameters are used. The duration is set to 3600 seconds and the frequency to every 5 seconds. If only the duration parameter is used, then the frequency is using its lowest possible value of five seconds. If only the frequency parameter is used, then the duration is infinite.

The `show-initiators-performance` command displays initiator performance data. The show command is used with the parameters duration and frequency, which is monitoring the initiator performance for the specified duration. Displayed are the registered initiators, their names, index, and write and read bandwidth, as well as their write and read IO/s performance. The total read and write information is the total of I/Os incurred since the initiator was registered.

The first two initiators have almost only performed reads. The third initiator shows no I/O at all.

- Output Parameter - Description
Initiator-Name - Initiator name
Index - Initiator index number

These properties indicate the current IOPS and bandwidth:

- Write-BW (MB/s) - Write bandwidth in MB/s
Write-IOPS - Write IO/s
Read-BW (MB/s) - Read bandwidth in MB/s
Read-IOPS - Read IO/s
BW (MB/s) - Total bandwidth in MB/s
IOPS - Total IO/s
Total-Write-IOs - Total write IOs since initiator registration
Total-Read-IOs - Total read IOs since initiator registration

Exporting Performance History

- Performance history can be exported to a csv file
 - Run the `export-performance-history` command
 - Cannot be run by a user with Read-Only role

```
login as: xmsadmin
xmsadmin@10.127.98.214's password:
Last login: Thu Jul 10 18:36:28 2014 from 10.12.247.85
Username: admin
Password:
Connect XMS on 10.127.98.214:443: version 2.4.0 build 25
xmcli (admin)>
xmcli (admin)> export-performance-history
Exported performance history can be accessed via
https://10.127.98.214/xtremapp/performance_history.csv
xmcli (admin)>
```



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Monitoring with XMS

24

Exporting performance data to a CSV file cannot be performed directly from the GUI. The user must open a CLI session [which can be performed from inside the GUI], and run the **export-performance-history** command. External applications can support the analysis. The data can be compared to baseline history data or any other collected data.

The **export-performance-history** command cannot be run by a user with the Read-Only role. After the file is successfully created, the download information is displayed. The command can be used with an argument for the filename. The default filename is **performance_history.csv**. The file contains one week's worth of data collected at 5 second intervals.

Lesson Summary

During this lesson the following topics were covered:

- Monitoring performance of an XtremIO array



This lesson covered performance monitoring of an XtremIO storage array.

This slide intentionally left blank.



Lab Introduction: Virtual Disk Workload Test

This lab covers workload testing on a VM virtual disk. Operations performed are:

- Deploy a VM
- Generate a workload on the VM Virtual Disk
- Verify workload statistics



This lab guides you through process of placing a workload on the Virtual Disk of a VMware virtual machine and monitoring the results.

Lab Debrief

This lab covered the process of placing a workload on the Virtual Disk of a VMware virtual machine and monitoring the results.



This lab covered the process of placing a workload on the Virtual Disk of a VMware virtual machine and monitoring the results.

Lab Introduction: RDM Workload Test

This lab covers the process of creating and editing Monitors in order to analyze XtremIO volume and initiator performance. Operations performed are:

- Create a new LUN on an XtremIO storage array
- Map the new LUN to a VM as an RDM
- Perform a workload test on the RDM



This Lab covers the process of placing a workload on an RDM assigned to a VMware virtual machine and monitoring the results.

Lab Debrief

This lab covered the process of creating and editing Monitors in order to analyze the volume and initiator performance of an XtremIO LUN used as an RDM.



This lab covered the process of placing a workload on a VMware virtual machine RDM and monitoring the results.

.

Lab Introduction: XtremIO Advanced Monitoring

This lab covers the process of creating and editing Monitors in order to analyze XtremIO volume and initiator performance. Operations are:

- Create XMS Monitors
- Edit XMS Monitors



This Lab covers the process of placing a workload on a VMware virtual machine and performing advanced monitoring of the results.

Lab Debrief

This lab covered the process of creating and editing Monitors in order to analyze XtremIO volume and initiator performance.



This lab covered the process of placing a workload on a VMware virtual machine and performing advanced monitoring of the results.

.

Module Summary

Key points covered in this module:

- Defining monitoring operations
- Monitoring for performance



This module introduced XtremIO monitoring. You were introduced to monitoring functions on the XtremIO array, such as basic performance with the GUI; selecting user-defined monitors in the Monitor workspace; saving performance data; performance monitoring using XMCLI and how to export performance history.

During the Virtual Disk workload test lab, we reviewed the process of placing a workload on the Virtual Disk of a VMware virtual machine and monitoring the results. The RDM Workload Test lab illustrated how to placing a workload on an RDM assigned to a VMware virtual machine and monitoring the results. Finally, the XtremIO Advanced Monitoring lab demonstrated the process of placing a workload on a VMware virtual machine and performing advanced monitoring of the results

This slide intentionally left blank.



Module: XtremIO Performance Management Overview

Upon completion of this module, you should be able to:

- Discuss performance analysis issues
- Relate basic performance measurements to I/O operations
- Monitor performance of XtremIO storage systems
- Describe tools to measure and benchmark performance
- Perform data collection on an XtremIO storage array

This module focuses on the basic principles of performance management, the XtremIO architecture from a performance perspective, and on the monitoring and testing of performance. The module covers the concept of Little's Law and its impact, and EMC best practices for XtremIO array performance. The module discusses where to monitor performance and collect performance data, and introduces several benchmark tools that can be used on the XtremIO array.

This slide intentionally left blank.



Lesson: Performance Management Overview

This lesson covers the following topics:

- Why manage performance?
- The performance life cycle
- Performance indicators
- Little's Law



This lesson gives an overview of performance management, looks at the performance life cycle, and discusses the implications of Little's Law.

Why Analyze Performance?



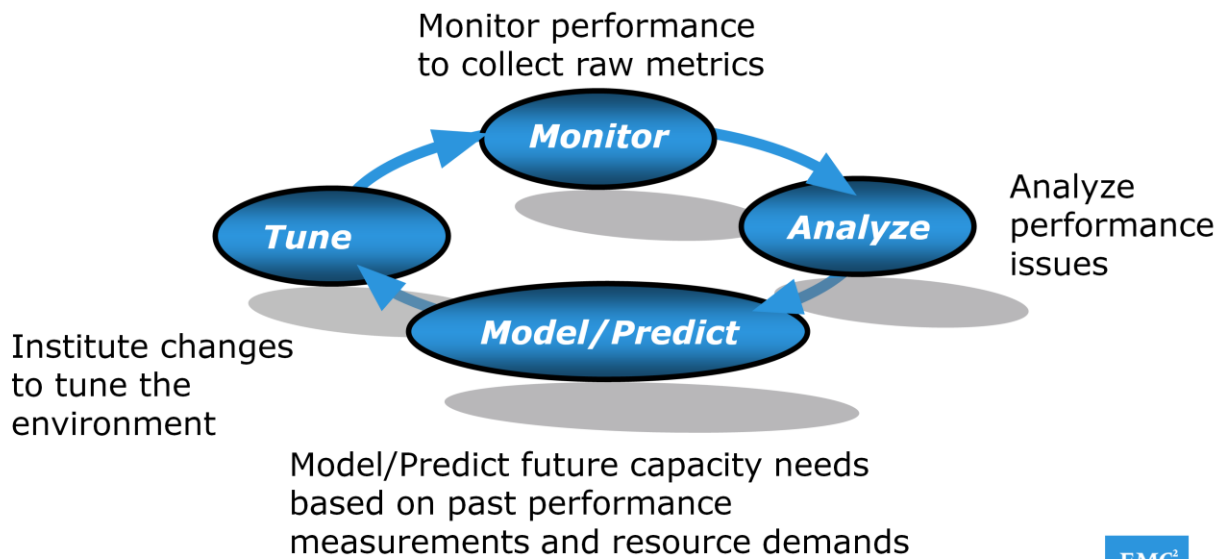
There are several reasons why performance analysis is an important part of managing a complex data environment. Short term problem identification can often be performed simply by looking for abnormal data patterns. If delays are being experienced and one component is showing twice the normal I/O rate, it is probable that, that is causing the issue. Performance analysis can quickly trace this issue to its root cause.

As a medium term measure, performance analysis can be used to better allocate resources. When new drives must be mapped, new file systems must be created, or new applications installed, a quick performance analysis of the components involved will show what drives or data channels are least used. Allocating the new traffic to the unused components helps minimize the performance impact.

Complex provider-consumer environments will often use Service Level Agreements (SLA) to define a guaranteed level of performance. Simple performance analysis reporting must then be put in place to show compliance with the SLAs. Performance Management seeks to maintain this situation. SLAs can also be informal, unwritten expectations. A system which had been functioning at a particular speed, but has suddenly slowed down can be considered to have violated an informal service agreement. Whenever users complain of poor access speed, it can be considered a request for performance (capacity) improvement. The expectation of a certain level of service delivery is just as important as written SLA agreements in the eyes of a customer.

In the long term, detailed performance analysis can help in traffic capacity management. By analyzing the current I/O patterns and predicting the maximum performance thresholds, you can predict how much longer an environment can grow. Good capacity management helps eliminate problems that make short term analysis necessary.

Performance Lifecycle



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

EMC²

5

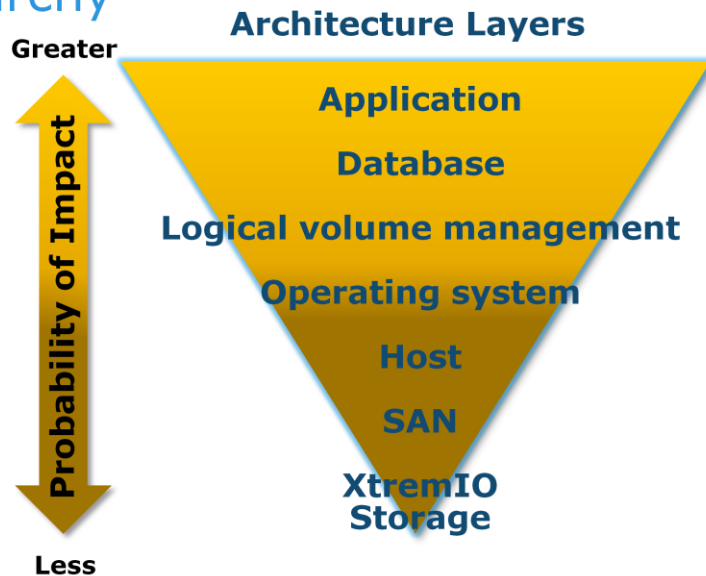
Performance Management, like all service management, should be an ongoing process. Continuous improvements should be made to counter the ever increasing demands for services that every IT department faces. Anticipating performance issues in this way reduces the number of emergencies that send IT staff scrambling.

Basic performance archiving is a type of monitoring—most performance tools can monitor the environment on a continuous basis. Relevant performance metrics can be incorporated into periodic reports that show the status of the storage environment. These reports are analyzed to detect repeated trends or issues. Solutions to these performance issues can be proposed and modeled. Performance histories can be consulted to predict future I/O capacity needs, and those needs also planned for. Finally, the necessary changes are made to tune the environment.

After a change has been made, the results must be observed, and if necessary, improved upon. The process repeats again at this point.

Performance Hierarchy

- More opportunities to optimize performance higher in the architecture
- XtremIO storage array least likely to be source of performance issue



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

6

There are many factors that influence the speed of data access, as the figure above illustrates. Data transfers through user applications, databases, volume managers, operating systems, server hardware, and finally storage systems. Any of these components can create inefficiencies and bottlenecks to reduce performance.

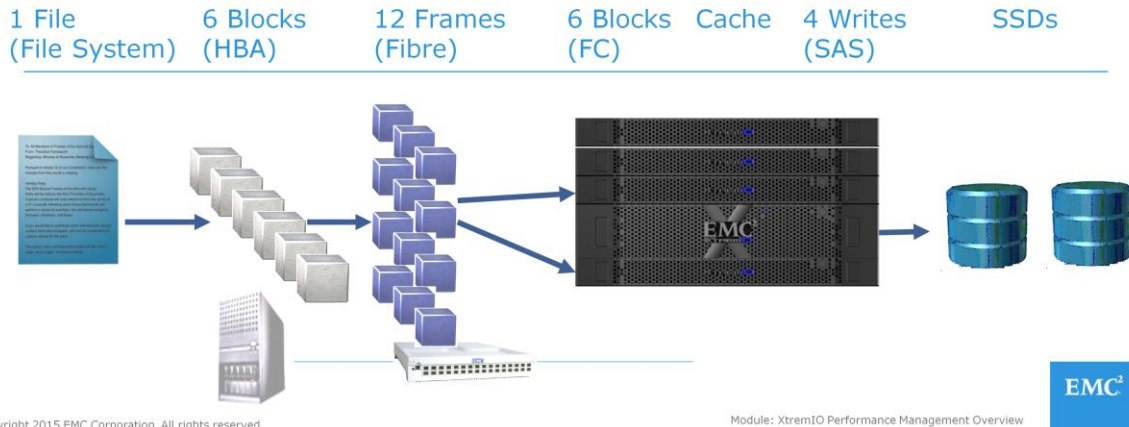
The figure also roughly approximates the relative amount of probability of impacting performance. The larger parts of the pyramid toward the top are more likely to contribute to performance problems. Look for changes and warnings on the top before looking at the components below.

Storage is shown as the smallest of the components. Storage problems are few compared to the other sources of performance issues. Sophisticated storage arrays are designed with performance in mind. Unfortunately, advanced storage systems are often difficult to monitor and improve compared to other components, making it important to proactively manage them.

This class covers the XtremIO performance issues only. While it may be a small part of the performance hierarchy, it is a complex storage array which requires very detailed knowledge to fully analyze.

What is an I/O?

- An I/O is a complete transfer, or unit of work, between End Points
- I/O might be broken into smaller units, or re-combined depending on the protocol of the End Points



An I/O is a unit of work, or a complete transfer, between two end points. There are many end points in a single pathway of the Enterprise storage environment. Each one uses a different protocol, which may break the received I/O into multiple I/Os before passing it on to the next end point. The illustration shows how a single file might be broken up during the process of storing it on an XtremIO array. The file will frequently be broken up into multiple I/Os (or blocks) when the File System Manager transfers it to the Host Bus Adapter (HBA). The HBA perceives the activity as a series of smaller work units.

The HBA will conform to Fibre Channel protocol, and transfer each I/O it receives as a series of frames of up to 2 KB to the next connectivity device. The frames are routed through the fabric to the XtremIO SC, where they are re-assembled into the original I/O form seen by the HBA, and placed into XtremIO cache. The data in cache will then be transferred via a SAS interface to the physical drives.

In every data transfer protocol, several tasks must be performed when sending an I/O. These other tasks add time to the overall process and increase the amount of data on the transmission channel. The components are:

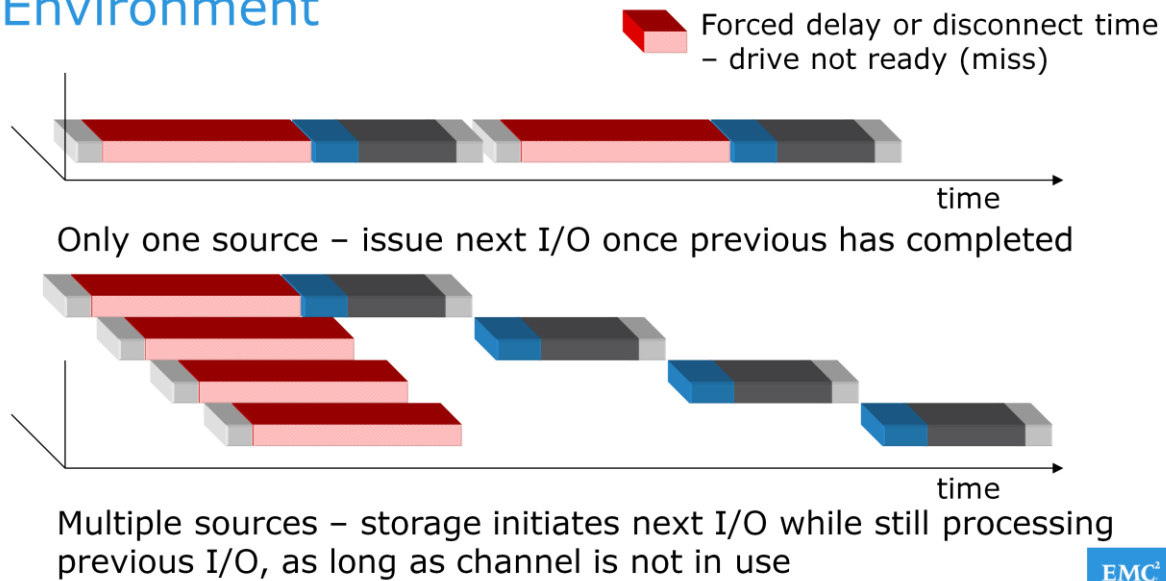
Negotiation, Acknowledgement – Both end points must agree to the transfer and manage the operation. This includes any “handshaking” tasks that processors use to schedule and organize the activity.

Header – Some identifier or address value that the receiving end point uses to properly utilize the data. Any CRC or parity bits used to error check the data can be considered header also, since this overhead adds to the information load on the channel.

Data – The actual data. All of the other components are added to the I/O during transfer and removed before the data can be utilized.

The Negotiation, Header, and Acknowledgement components are largely fixed in size regardless of I/O size. A change in the I/O size typically means a change in the Data component only.

Storage Array Delays in a Multi-Source Environment



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

8

The previous page showed the parts of an I/O as a single continuous sequence of events. This is how an I/O will be processed much of the time, but in other cases, delays might cause a more complex situation. A forced delay will often be introduced by the array when the I/O can not immediately be processed; this will typically happen in a read or write miss. While the array is preparing for the transfer (locating the data on a drive, making space available in cache, etc.), the I/O channel is unused.

If only one host or process source is using that data channel, the delay will turn into wasted time. Each I/O will only be issued when the one before it is completed, so no alternative sequencing can occur. This is illustrated in the top diagram. Note that only two I/Os are serviced.

If multiple hosts or processes are using the same channel, the delay time can be used more efficiently. Under the SCSI protocol used for data transfer, when one source is informed of a delay, it must be disconnected from the channel until notified that the data is ready. During this time, a different source can initiate or service a different I/O. As shown in the bottom diagram, this allows the channel to be used nearly all the time, resulting in more overall I/Os per second. Note that since no more than one actual data packet can be transferred on the channel at a time, the diagram shows only one Negotiation, Header, Data, or Acknowledgement happening at once. Ultimately, this makes for a longer total response time for some of the I/Os even as the overall system shows an improvement in data processing.

This concept is true in the SAN that connects the hosts to the array, as well as within the XtremIO itself. Testing XtremIO performance by sending a single-source I/O stream does not take full advantage of this multi-source optimization.

IOPs, Bandwidth, and I/O Size Relationship

IOPs, Bandwidth, and I/O Size are closely related

Small I/O

IOPs: 5

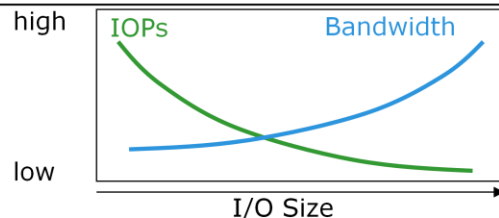
Through:



Large I/O

IOPs: 2

Through:



- As I/O size increases, IOPs decreases
- As I/O size increases, bandwidth increases

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

9

The I/O Per Second, bandwidth, and I/O Size measures are closely related.

When small I/Os are transferred, less time is taken on each I/O, increasing the number that can be moved in a given time period. The IOPs measure is typically large for small I/Os. However, the data payload is small when compared to the header and negotiation factors. So while many I/Os are being exchanged, the total data bandwidth is small.

When large I/Os are transferred, more time is taken on each I/O, decreasing the IOPs measure. However, a larger percentage of the channel resources are spent on data rather than header and negotiation. This increases the overall bandwidth. This is illustrated in the block diagrams above. The combined data parts of the large I/O example is roughly 50% larger than the small I/O example.

The bottom graph in the illustration shows this relationship. As the I/O size being transferred increases the IOPs measure drops. If the number of transfers per second is the important measure of performance in this environment, then smaller I/O size is beneficial. Databases typically set a size for their I/O transfers that maximizes this benefit.

The graph also illustrates how bandwidth will increase as the I/O size increases, even as the total number of I/Os goes down. File copy and backup operations will take advantage of this relationship by increasing the I/O size.

Utilization

Clock tick measures

- Count busy and idle clock tick cycles of processor, disk, (object) etc.
- Accurate measure of object utilization

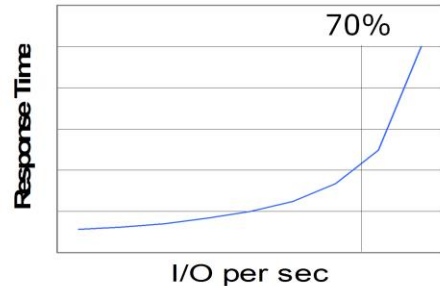


|||| Busy

|||| Idle

Expected value measures

- Divide key performance indicator over the maximum benchmarked
Adjusted for component model and type, but not other performance factors



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

10

Each clock tick that the processor, disk, or other object can be doing something is counted as either busy or idle depending on its activity. If the object has many idle ticks, it is under-utilized. If all of the ticks are busy on the other hand, it cannot possibly be doing more. By comparing the number of busy and idle ticks, you can express the utilization of the object.

Clock tick measures may be available for many components in a data center environment, and are considered the most accurate way to measure utilization.

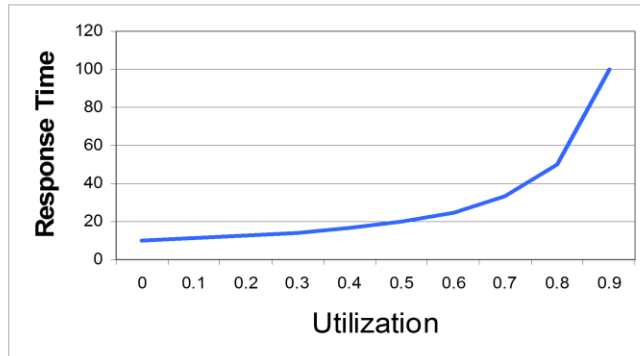
Without a way to measure clock ticks, utilization measures are calculated by dividing the key performance indicator over the maximum expected performance for a component. For example, if benchmarking shows that a component typically reaches a limit at 1,000 I/Os per second, and the current I/O rate is 700 I/Os per second, you can express the utilization as 70%. Note that the benchmarked value is used, not the ideal rating. Benchmarks show that a 4 Gb per second port will reach a limit at somewhat less than 4 Gb per second. Since this calculation does not account for changes in the I/O characteristics or other tasks the component might be performing, it is considered a less accurate way to measure utilization.

Little's Law

- Little's Law predicts non-linear increases in response time with increasing utilization of shared access queuing systems

$$\text{Response} = \frac{\text{Service}}{1 - \text{Utilization}}$$

- **Service:** Time to service a single work unit
- **Utilization:** Proportion of time spent servicing work units (0.0 to 1.0)
- **Response:** Actual time to initiate, service, and return a single work unit



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Performance Management Overview

11

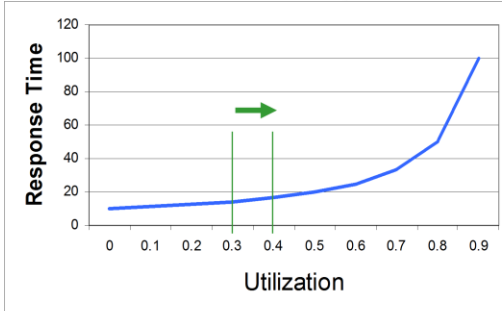
A linear increase in traffic load will typically NOT produce a linear change in performance measures such as response time.

Little's Law provides a theoretical formula for a well known performance relationship seen in queuing systems (take inputs and produce outputs). As shown in the graph above, the law predicts that response time for a queuing system will increase non-linearly as utilization increases. Furthermore, the law states that 100% utilization cannot be achieved without experiencing infinite response time.

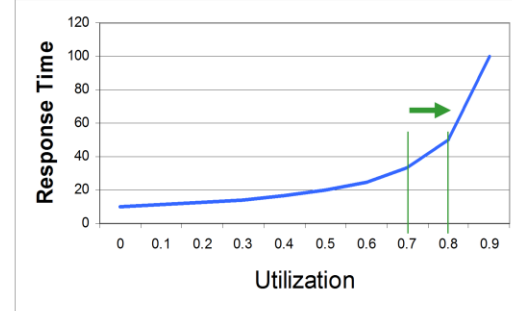
This formula should not be used for exact predictions in complex environments. However, Little's Law is very useful for rough estimates and as a general illustration of the relationship between response time and utilization.

Implications of Little's Law

- A small increase in traffic may cause a small or big increase in response time, depending on current utilization
- Each element in the pathway might be experiencing a different utilization, making predictions even more difficult



Increase in utilization from 0.3 to 0.4 (+10%)
Increase in response time from 14 to 16 (+2)



Increase in utilization from 0.7 to 0.8 (+10%)
Increase in response time from 33 to 50 (+17)

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

12

An important implication of Little's Law is that a fixed increase in traffic will not always cause a fixed increase in response time. Both of the above charts illustrate an increase in utilization of 10% (to keep the math simple, let's assume this was a result of a 10% increase in traffic). The two charts clearly show that the utilization increase has a much greater effect on the response time when utilization is already high. Keeping utilization low while still performing needed data operations will be a repeated theme throughout this course.

Little's Law applies to all elements in a data path: host software, connectivity channels and end points, caching systems, and drives all respond to increases in utilization in this way. Predicting response time changes, as a result of traffic increases, is an extremely difficult task in light of this complexity.

Lesson Summary

During this lesson the following topics were covered:

- Why manage performance?
- Performance management styles
- The performance life cycle
- Performance indicators
- Little's Law



This lesson covered an overview of performance management, the performance life cycle, and the implications of Little's Law.

We reviewed the many reasons for managing performance, such as monitoring abnormal data patterns, better allocating resources, allocating new traffic to unused components to assist in minimizing performance impact, ensuring compliance with the SLAs (Service Level Agreements), help with traffic capacity management, predict the length of time the environment can continue to grow. Finally, we saw how the use of short term analysis provides a solid foundation for good capacity management helping to eliminate future problems.

This slide intentionally left blank.



Lesson: Architectural Constructs

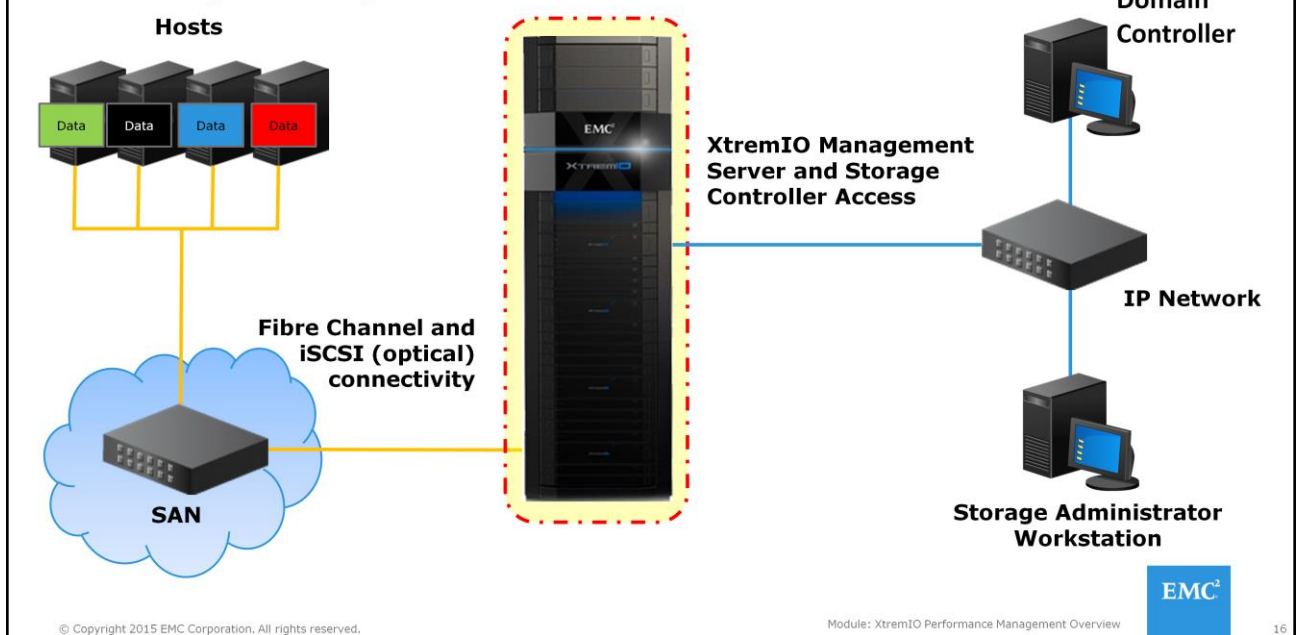
This lesson covers the following topics:

- Theory of operation
- System software components
- Factors that affect performance



This lesson reviews the XtremIO system components and how they relate to performance. It also shows how the components are assembled to build an X-Brick and how X-Bricks can be connected to form an XtremIO cluster. This lesson covers how data is received and written. Data reads are also detailed. We will also cover how performance can be increased, along with how data is positioned on the system as the hardware is scaled. This lesson explains where potential bottlenecks in the environment can occur and what performance indicators are used. Parameters that affect performance are discussed. These fall into the categories of parameters that can be adjusted and parameters that cannot be adjusted.

Theory of Operation



The XtremIO Storage Array automatically deduplicates and compresses data as it enters the system, processing it in data blocks. Deduplication is global (over the entire system), is always on, and is performed in real-time (never as a post-processing operation). After the deduplication the data is compressed inline, before it is written to the SSDs.

XtremIO uses a global memory cache, which is aware of the deduplicated data, and content-based distribution that inherently spreads the data evenly across the entire array. All volumes are accessible across all X-Bricks and across all storage array host ports.

The system uses a back-end InfiniBand network that provides high speeds with ultra-low latency and Remote Direct Memory Access (RDMA) between all storage controllers in the cluster. By leveraging RDMA, the XtremIO system is in essence a single, shared memory space spanning all storage controllers.

- The effective logical capacity of a single X-Brick varies depending upon the data set being stored. For highly duplicated information, which is typical of many virtualized cloned environments, such as Virtual Desktop Integration (VDI), the effective usable capacity is much higher than the available physical flash capacity. Deduplication ratios in the range of 5:1 to 10:1 are routinely achieved in such environments.
- For compressible data, which is typical in many databases and in application data, compression ratios are in 2:1 to 3:1 range.
- Systems benefitting from both data compression and data deduplication, such as Virtual Server Infrastructures (VSI), commonly achieve a 6:1 ratio.

Mapping Table

	LBA Offset		Fingerprint		SSD Offset / Physical Location	
Data →	Address 0	→	20147A8	→	40	→ Data
Data →	Address 1	→	AB45CB7	→	8	→ Data
Data →	Address 2	→	F3AFBA3	→	88	→ Data
Data →	Address 3	→	963FE7B	→	24	→ Data
Data →	Address 4	→	0325F7A	→	64	→ Data
Data →	Address 5	→	134F871	→	128	→ Data
Data →	Address 6	→	CA38C90	→	516	→ Data
Data →	Address 7	→	963FE7B	—	Deduplicated	Do not write Point to 24

EMC²

© Copyright 2015 EMC Corporation. All rights reserved. Module: XtremIO Performance Management Overview 17

In a typical write operation, the incoming data stream reaches any one of the Active-Active Storage Controllers and is broken into data blocks. For every data block, the array fingerprints the data with a unique identifier.

Each Storage Controller maintains a table that manages the location of each data block on SSD and is used to determine if incoming writes already exist within the array. The fingerprint is used to determine the storage location of the data. The LBA to content fingerprint mapping is recorded in the metadata, within the Storage Controller's memory. The table has two parts:

- The first part of the table maps the host Logical Block Addressing (LBA) to a content fingerprint.
- The second part of the table maps the content fingerprint to its location on SSD.

Using the second part of the table provides XtremIO with the unique capability to distribute the data evenly across the array and place each block in the most suitable location on SSD. It also enables the system to skip a non-responding drive or to select where to write new blocks when the array is almost full and there are no empty stripes to write to.

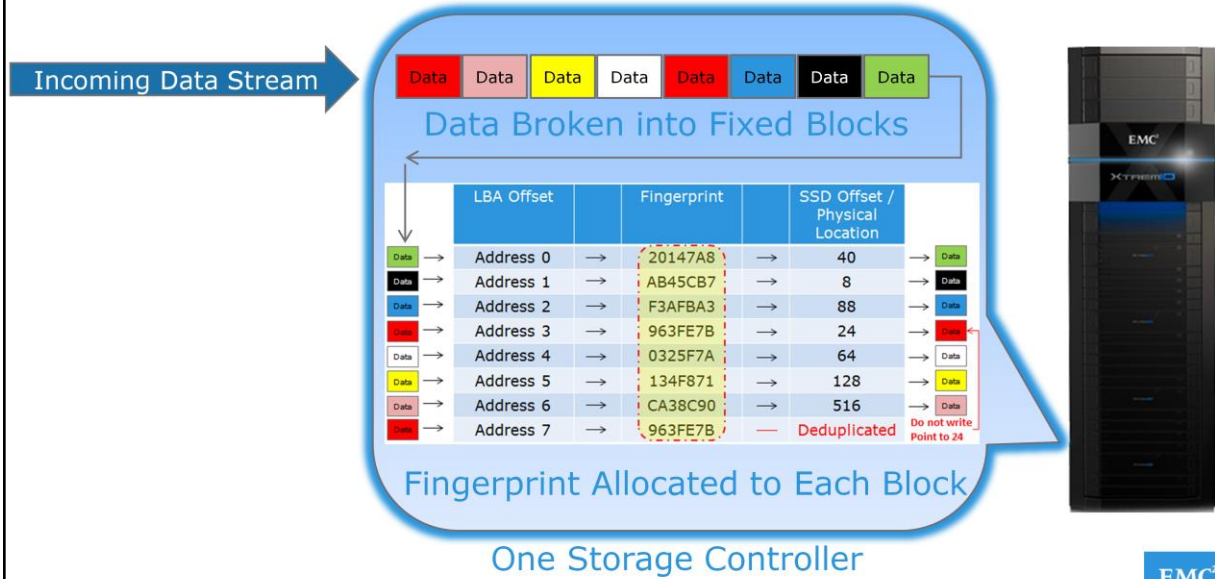
The system checks if the fingerprint and the corresponding data block have already been stored previously. If the fingerprint is new, the system:

- Compresses the data.
- Chooses a location on the array where the block will go (based on the fingerprint, and not the LBA).
- Creates the "fingerprint to physical location" mapping.
- Increments the reference count for the fingerprint by one.
- Performs the write.

In case of a "duplicate" write, the system records the new LBA to fingerprint mapping, and increments the reference count on this specific fingerprint. Since the data already exists in the array, it is neither necessary to change the fingerprint to physical location mapping nor to write anything to SSD. All metadata changes occur within memory. Therefore, the deduplicated write is carried out faster than the first unique block write. This is one of the unique advantages of XtremIO Inline Data Reduction, where deduplication actually improves the write performance.

The actual write of the data block to SSD is carried out asynchronously. At the time of the application write, the system places the data block into the in-memory write buffer (which is protected by replicating to different Storage Controllers via RDMA), and immediately returns an acknowledgement to the host. When enough blocks are collected in the buffer, the system writes them to the XDP (XtremIO Data Protection) stripe(s) on SSDs. This process is carried out in the most efficient manner.

Write I/O Flow – Steps 1 & 2



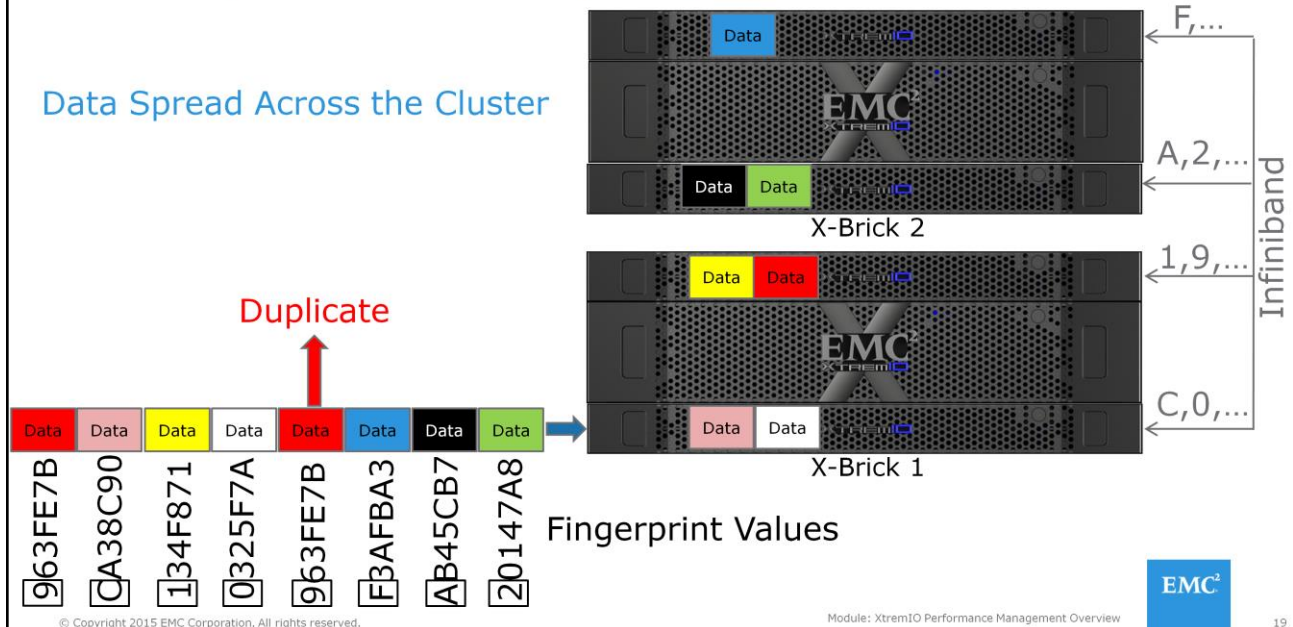
1. The system analyzes the incoming data and segments it into data blocks.
2. For every data block, the array allocates a unique fingerprint to the data.

The array maintains a table with this fingerprint to determine if subsequent writes already exist within the array.

- If a data block does not exist in the system, the processing Storage Controller journals its intention to write the block to other Storage Controllers, using the fingerprint to determine the location of the data.
- If a data block already exists in the system, it is not written as shown by the red X above.

Write I/O Flow – Steps 3, 4, & 5

Data Spread Across the Cluster



3. The array increases the reference count for each data block.
4. Using a consistent distributed mapping, each block is routed to the Storage Controller that is responsible for the relevant fingerprint address space.

The consistent distributed mapping is based on the content fingerprint. The mathematical process that calculates the fingerprints results in a uniform distribution of fingerprint values and the fingerprint mapping is evenly spread among all Storage Controllers in the cluster, as shown above. Data transfer across the cluster is carried out over the InfiniBand network using RDMA. [The example is simplified by using only the first hex digit of the fingerprint; the actual algorithm is more complex]

5. The system sends back an acknowledgement to the host.

Write I/O Flow – Steps 6, 7, & 8

Additional Blocks Populated Full Stripes



X-Brick 2

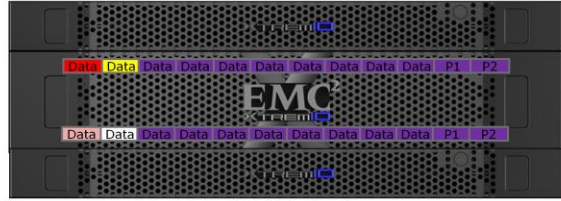


X-Brick 1

Stripes Committed to SSDs



X-Brick 2



X-Brick 1



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

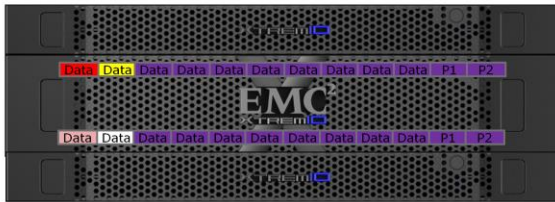
20

6. Due to the even distribution of the fingerprint function, each Storage Controller in the cluster receives an even share of the data blocks. When additional blocks arrive they populate the stripes.
7. The system compresses the data blocks to further reduce the size of each data block.
8. Once a Storage Controller has enough data blocks to fill the emptiest stripe in the array (or a full stripe if available), it transfers them from the cache to SSD.

Read I/O Flow



X-Brick 2



X-Brick 1

	LBA Offset		Fingerprint		SSD Offset / Physical Location	
Data →	Address 0	→	20147A8	→	40	→ Data
Data →	Address 1	→	AB45CB7	→	8	→ Data
Data →	Address 2	→	F3AFBA3	→	88	→ Data
Data →	Address 3	→	963FE7B	→	24	→ Data
Data →	Address 4	→	0325F7A	→	64	→ Data
Data →	Address 5	→	134F871	→	128	→ Data
Data →	Address 6	→	CA38C90	→	516	→ Data
Data →	Address 7	→	963FE7B	→	— Deduplicated	→ Data

System performs look-up of logical address in the LBA to fingerprint mapping

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

21

In a data block read operation, the system performs a look-up of the logical address in the LBA to fingerprint mapping. Once the fingerprint is found, it looks up the fingerprint to physical mapping, and retrieves the data block from the specific physical location. Because the data is evenly written across the cluster and SSDs, the read load is also evenly shared.

XtremIO has a memory-based read cache in each Storage Controller.

- In traditional arrays, the read cache is organized by logical addresses. Blocks at addresses that are more likely to be read are placed in the read cache.
- In the XtremIO array, the read cache is organized by content fingerprint. Blocks whose contents (represented by their fingerprint IDs) are more likely to be read are placed in the cache.

This makes the XtremIO read cache deduplication aware, which means that its relatively small read cache appears much larger than traditional caches of the same size.

If the requested block size is larger than the data block size, XtremIO performs parallel data block reads across the cluster and assembles them into bigger blocks, before returning them to the application.

A compressed data block is decompressed before it is delivered.

Read I/O is issued to Array

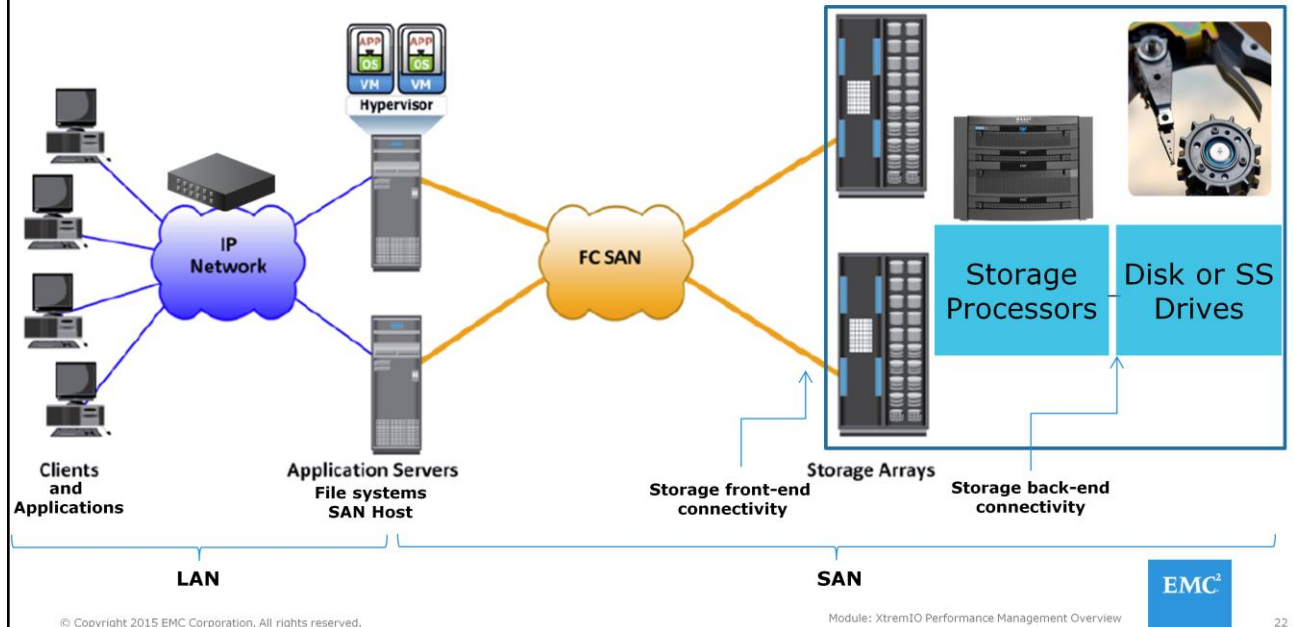
1. The system analyzes the incoming request to determine the LBA for each data block and creates a buffer to hold the data.
2. The following processes occur in parallel:

For each data block, the array finds the stored fingerprint. The fingerprint determines the location of the data block on an X-Brick. For larger I/Os (e.g. 256 kB), multiple X-Bricks are involved in retrieving each data block.

The system transmits the requested Read data to the processing Storage Controller, via RDMA, over InfiniBand.

3. The system sends the fully populated data buffer back to the host.

Traditional Storage Potential Bottlenecks



Here is an example of a SAN environment. The entire system is only as fast as its slowest component.

The SAN environment consist of the following components:

- Application/File Servers (File systems) (SAN Host)
- Interconnecting Devices (FC Node Ports (HBAs), FC Switches or Hubs)
- Storage Controllers (Arrays) (Storage Processors)
- Storage back-end connectivity (SAS)
- Disk Drives (Flash Drives)

Be aware that SAN Hosts also usually attach to LANs, if so, the follow components will also influence overall performance:

- Client and Applications (IP Host)
- IP Node Ports (NICs), Connectors and Cables
- Client IP network (Ethernet Switches)

Any of these components or networks can be a potential bottleneck.

What Affects Performance 1

What <u>cannot</u> be tuned - Internal factors	
Data Protection (XDP)	Data path for IO (R-C-D)
SC are fully Active/Active	SSD response time
No Volume to RAID Group affinity	IB Messages and Journaling
No Storage Pools	RAID processes
No Volume to Target Port Affinity	Low level S/W locks
Fixed Cache Size	



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

23

Factors that can affect performance can be categorized as internal or external to the storage array. There are many places in the entire environment (external) that can become a bottleneck; however, with XtremIO being so fast, it will almost never be the bottleneck. Compared to other storage arrays there is virtually nothing to be tuned on the XtremIO array. Here are some reasons why: There is only one RAID level. The storage controllers work in an active/active configuration. There is no need to re-balance the LUNs by changing their ownership from or to a different storage controller because of suspicion that one is being over utilized and another is not. There are no RAID groups or storage pools. There is no affinity between a RAID group or a volume and a specific I/O port. When a volume is created, it is always available on all target ports. Also, the cache is set to 1 GB per X-brick.

There are other factors that can affect performance. These are considered internal factors and these are the internal array data paths, SSD response times, IB messaging and journaling, data placement processes and low level S/W locks. These S/W locks determine which addresses are available for data writes or rewrites on the SSDs. Most internal factors can not be tuned by changing configuration parameters. These performance factors are only modified by updating the XtremIO firmware or operating system. EMC, XtremIO technical support will have to be involved if these parameters are suspected of causing performance issues.

However, there are parameters that can be tuned; these are shown next.

What Affects Performance 2

What <u>can</u> be tuned - External factors	
HBA Execution Throttle	CPU Utilization
Number of Paths	Memory Utilization
Multipath Configuration	S/W IRQs (CPU Cores to HBA Affinity)
VAAI Support (for ESXi servers)	Topology & Zoning Configuration
Number of Volumes	Data Alignment from application on host
Network Bottlenecks	

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview



24

External factors that can affect performance are included in the list above. Many of the factors have parameters that can be adjusted.

HBA Execution Throttle (Q-Logic calls it Execution Throttle, Emulex calls it Q Depth) can be modified. The number of paths and multi-path configuration such as Adaptive, Round Robin, or Most Recently or Least Used Path can be checked. VAAI support for ESX and ESXi servers can be enabled or disabled. The number of volumes can be controlled. Smaller volume usually perform better than larger ones. Fibre Channel (FC) and IP network speeds should be known. Bottlenecks should be located and some can be mitigated by upgrading to higher speed HBAs or adjusting the number and/or speed of any Inter Switch Links (ISLs). The network director or switch parameters can be adjusted which may have some affect on speed especially with large blocks but also small blocks.

A two X-Brick array, which can handle about 500K read IOPS, will probably never be overloaded by a single host. If it is, you should monitor the CPU and memory utilization and Software Interrupts (IRQs) on the host. Check to see how many CPU cores are being used and which are assigned to handle HBA load on the host.

Check the FC network topology and zoning configuration. Also, the data block size of the application on host should be checked for alignment with the block size of the array.

Lesson Summary

During this lesson the following topics were covered:

- Theory of operation
- System software components
- Factors that affect performance



This lesson covered the XtremIO system software components and how they relate to performance.

This slide intentionally left blank.



Lesson: Environment Tools

This lesson covers the following topics:

- EMC Connectrix Manager Converged Network Edition (CMCNE)
- Data Center Network Manager (DCNM)
- ViPR SRM
- EMC M & R



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

27

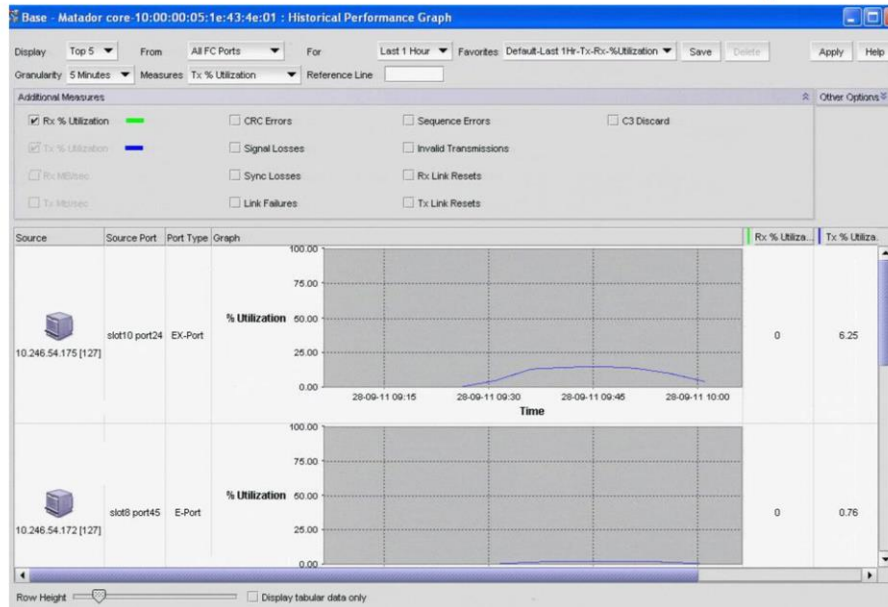
The monitoring of performance in an XtremIO environment involves numerous complex functions. These include monitoring and maintaining the network devices to provisioning the services, data center network infrastructure troubleshooting and capacity planning, and assessing the impact of scheduled network maintenance or migration.

EMC Connectrix Manager Converged Network Edition (CMCNE) is a management application capable of managing both traditional SAN environments and switches as well as the newer converged Ethernet technology, Fibre Channel over Ethernet (FCoE), and IP.

The Connectrix MDS Data Center Network Manager (DCNM) is a tool that can manage storage and data networking over the converged, virtualized data center. DCNM licenses are hosted on a management server.

ViPR SRM and EMC M&R allow monitoring and reporting on not only storage arrays, but the hosts and infrastructure that connects hosts to the storage.

CMCNE: Historical Performance Graph



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview



28

The screenshot shows EMC Connectrix Manager Converged Network Edition (CMCNE), where users can complete most fabric and switch configuration and perform fabric monitoring. The management application provides centralized management of the SAN, as well as quick access to all product configuration applications. Using this application, you can easily configure, manage, and monitor networks.

CMCNE performance monitoring provides details about how much traffic and errors a specific port or switch generates on the fabric over a specific timeframe. You can monitor a switch's real-time performance through a performance graph that displays transmitted and received data. To get an idea of port performance over time, you can also refer to the historical performance chart or report.

Data Center Network Manager (DCNM)



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

29

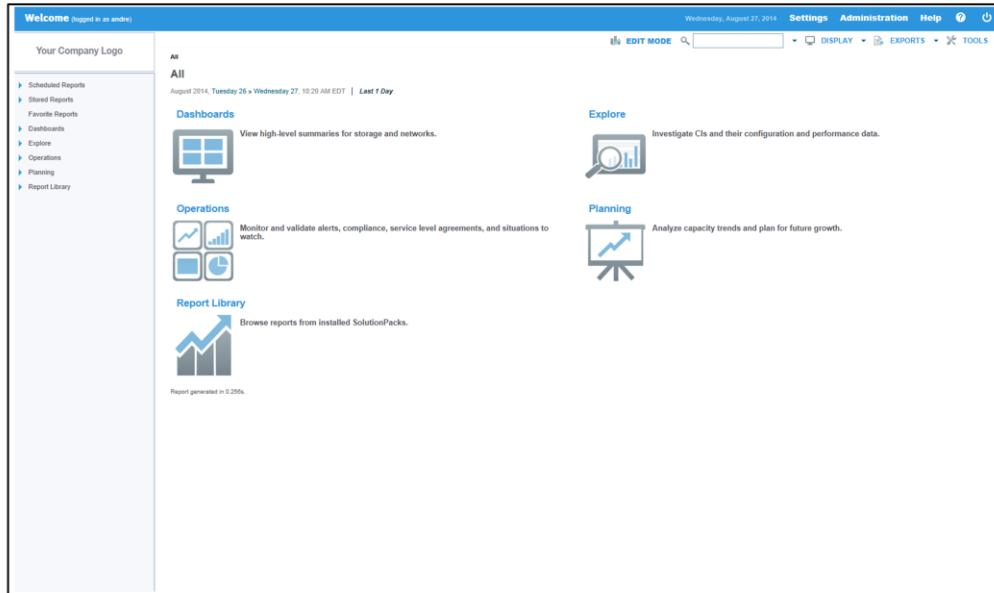
DCNM simplifies management of the data center by offering:

- Virtual Machine-aware path management
 - Enables management of the entire path through the physical to the virtual network across the entire data center environment using VMpath (identifies bottlenecks) and VM-aware (shows dependencies) views.
- Performance and troubleshooting
 - Monitors and provides alerts for fabric availability and performance.
- Interactive dashboard
 - Provides capability to view more details of key performance indicators (KPIs). Proactively measures, analyzes, and predicts performance of SAN infrastructure.
- Scalability
 - Uses federation to scale to large and distributed data center deployments.

This window is referred to as the Dashboard. You can view all the dependencies from the virtual machine out to the physical host, through the fabric, and to the storage array using the virtual machine-aware (VM-aware) topology view. This view allows easy access to a detailed view of the path attributes.

The DCNM SAN window shows a map of the SAN environment. It can be used to manage an entire connectivity environment and perform tasks such as discovery, configuration of SAN zoning, viewing of alerts, and monitoring the environment.

ViPR SRM



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

30

ViPR SRM is storage resource management software that enables IT personnel to visualize, analyze and optimize storage resources. ViPR SRM allows the end-to-end monitoring and analysis of storage services through analysis of real-time metrics for application, hosts, switches, and arrays.

Configuring ViPR SRM Monitoring

- ViPR SRM allows end-to-end monitoring
 - Host, application, switch, and array monitoring
- ViPR SRM can monitor a wide range of storage arrays
 - Each needs a specific 'SolutionPack' module
 - A data collection VM retrieves the device-specific data
- XtremIO is supported by the XtremIO SolutionPack
 - Requires XtremIO version 2.4 or newer
 - Integrates collector configuration, report presentation, and deployment scripts



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

31

ViPR SRM monitors a number of object types in a storage environment. These range from applications and hosts through to the storage systems where application data resides.

Modules called SolutionPacks act as 'device drivers' for the specific objects, and allow them to be monitored in a consistent manner. SolutionPacks exist for not only EMC storage, but a number of third-party storage systems as well. One of the SolutionPacks available for EMC storage is the XtremIO SolutionPack, which supports XtremIO systems at version 2.4 or newer.

XtremIO SolutionPack Basics

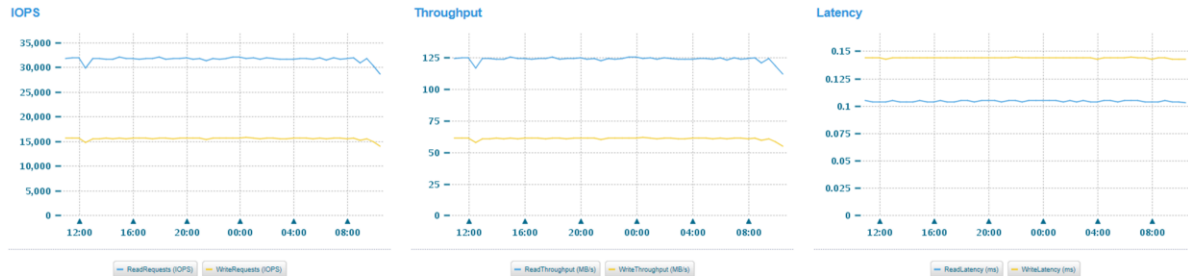
- XtremIO SolutionPack reports:
 - XtremIO information and details
 - XtremIO health
 - XtremIO capacity dashboard, de-duplication and thin provisioning
 - XtremIO performance metrics
 - XtremIO path details and connectivity
 - XtremIO alerts



A ViPR SolutionPack allows the collection of information from a specific supported platform, XtremIO in this case, and generates reports. For XtremIO arrays, the reports generated include general information and system health, connectivity, and performance information. The three important performance metrics are the ones seen in native XtremIO reporting – IOPs, bandwidth, and latency.

Standard Reports: Performance

- IOPs, throughput, and latency vs. time
- Legend not displayed by default
 - Mouse over the top right hand corner to see the controls



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

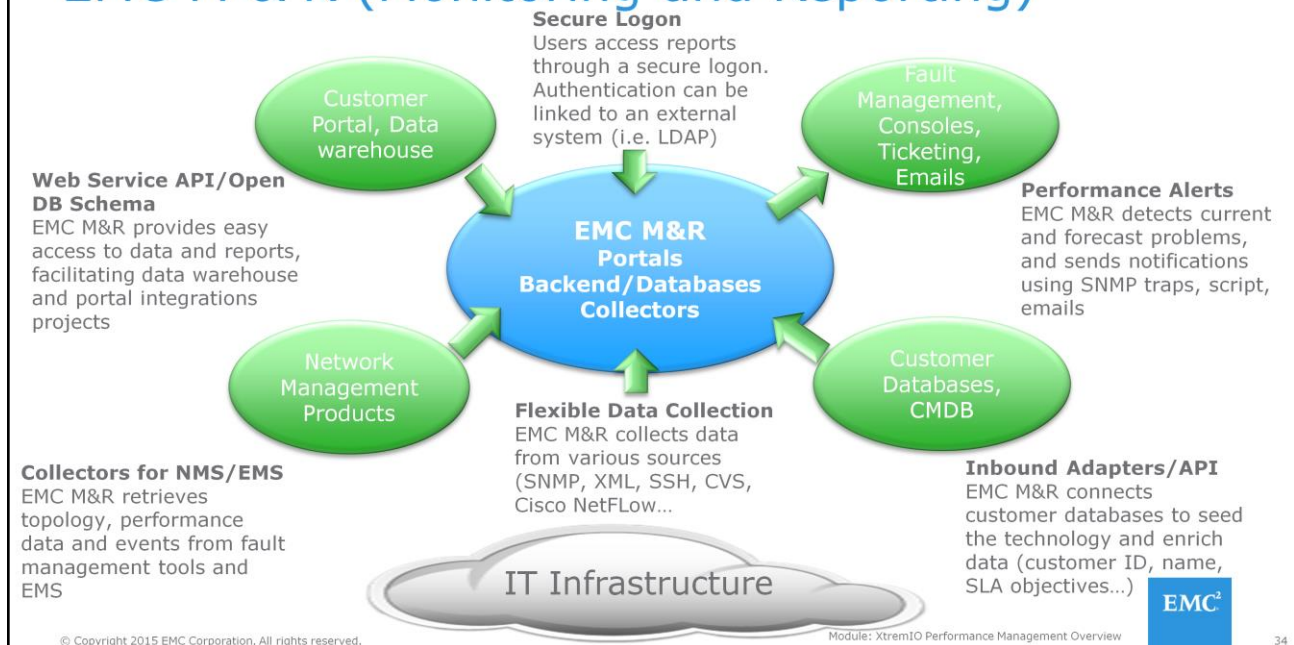


33

ViPR SRM displays XtremIO as charts of the three most important performance parameters: IOPs, throughput, and latency. [Note: in the VNX space, 'IOPs' is known as 'throughput', 'throughput' is called 'bandwidth', and 'latency' is referred to as 'response time'. The parameters are equivalent.]

Latency is likely to be the parameter which will determine the success of an implementation – since XtremIO should deliver IOPs at sub-millisecond latencies, there is a potential problem if that latency is not being maintained. This latency is measured at the array itself, so host processing may add additional time.

EMC M & R (Monitoring and Reporting)



EMC M & R (Monitoring and Reporting - formerly known as EMC Watch4net) is a technology platform that can be used to measure and manage the performance of an entire IT infrastructure. EMC M & R automatically retrieves information for various data sources to provide real-time, historical, and projected visibility into the performance of the infrastructure. EMC M & R:

- Provides visibility of network, systems, applications, storage and services
- Reduces downtime and improves infrastructure performance
- Identifies trends and deviations from normal behaviors
- Publishes dashboards for operations teams, engineers and customers
- Supports complex multi-customer infrastructure
- Reduces OPEX, optimizes CAPEX and increases ROI

Lesson Summary

During this lesson the following topics were covered:

- EMC Connectrix Manager Converged Network Edition (CMCNE)
- Data Center Network Manager (DCNM)
- ViPR SRM
- EMC M & R



This lesson covered non-native tools for monitoring an XtremIO environment.

During this lesson we reviewed EMC Connectrix Manager Converged Network Edition (CMCNE) where users can easily configure, manage, and monitor networks. We looked at Data Center Network Manager (DCNM) which allows for simplified management of data through virtual machine-aware path management, performance and troubleshooting, interactive dashboard, and scalability.

We reviewed the features of ViPR SRM, as a storage resource management software that provides IT personnel the ability to visualize, analyze and optimize storage resources and allows for end-to-end monitoring and analysis of storage services through real-time metrics for application, hosts, switches and arrays.

Finally, we covered EMC M & R (Monitoring and Reporting), a technology platform used to measure and manage the performance of an entire IT infrastructure.

This slide intentionally left blank.



Lesson: Host Tools

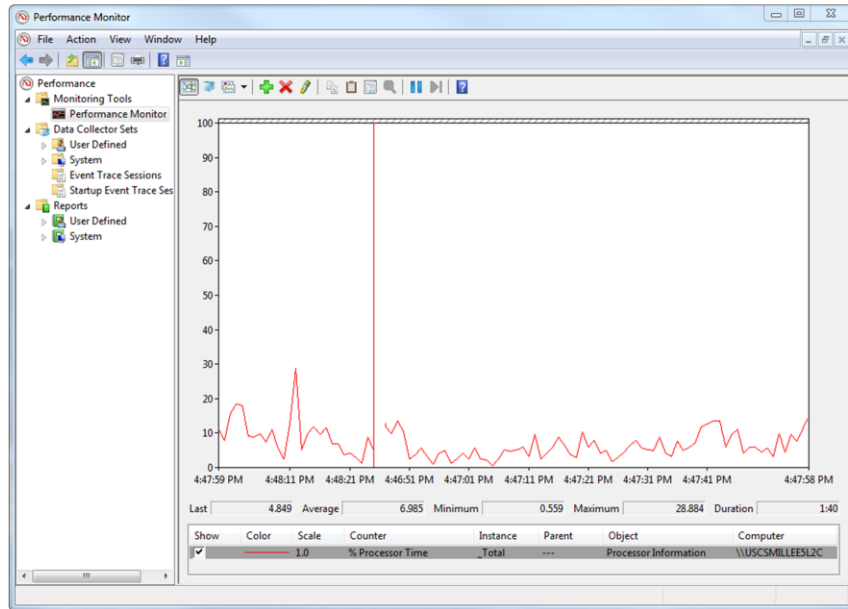
This lesson covers the following topics:

- Perfmon (Windows)
- sar (Unix/Linux)
- iostat (Unix/Linux)
- vCenter (ESXi)



This lesson covers hosts tools that can be used to monitor and collect performance data from the host perspective.

Microsoft Performance Monitor



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview



38

You can use the Microsoft Performance Monitor to view performance data on a Microsoft host either in real time or from a log file. You can create Data Collector Sets to configure and schedule performance counters, perform event traces, and collect configuration data so that you can view reports and analyze the results.

sar

sar is a Unix/Linux command used to monitor I/O performance

```
# sar -d
Linux 2.6.32-431.el6.x86_64 (RH6.andrelab.emc.com)    09/19/2014    _x86_64_    (16 CPU)

01:20:01 PM      DEV      tps  rd_sec/s  wr_sec/s  avgrq-sz  avgqu-sz   await   svctm    %util
01:30:01 PM    dev8-0      0.24    0.00    2.80    11.67    0.01    21.94    9.44    0.23
01:30:01 PM dev253-0      0.36    0.00    2.86    8.00    0.01    18.71    6.36    0.23
01:30:01 PM dev253-1      0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
01:30:01 PM dev253-2      0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
```

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

EMC²

39

The Linux/Unix **sar** command displays the contents of selected cumulative activity counters in the operating system. Based on the values in the count and interval parameters, sar writes information the specified number of times spaced at the specified intervals in seconds. If the interval parameter is set to zero, the sar command displays the average statistics for the time since the system was started. If the interval parameter is specified without the count parameter, then reports are generated continuously. The collected data can also be saved in the file specified by the -o filename flag. If filename is omitted, sar uses the standard system activity daily data file, the /var/log/sa/sadd file, where the dd parameter indicates the current day.

The command syntax is: **sar [options] [<interval> [<count>]]**

Main options and reports of interest for storage:

- -b I/O and transfer rate statistics
- -B Paging statistics
- -d Block device statistics
- -q Queue length and load average statistics
- -S Swap space utilization statistics
- -W Swapping statistics

For more detail, see the **man** page for sar on your system.

iostat

iostat is a Unix/Linux command to monitor I/O performance

```
$ iostat
```

Linux 2.6.32-100.28.5.el6.x86_64 (dev-db)				10/09/2014		
avg-cpu:	%user	%nice	%system	%iowait	%steal	%idle
	5.68	0.00	0.52	2.03	0.00	91.76
Device:	tps	Blk_read/s	Blk_wrtn/s	Blk_read	Blk_wrtn	
sda	194.72	1096.66	1598.70	2719068704	3963827344	
sda1	178.20	773.45	1329.09	1917686794	3295354888	
sda2	16.51	323.19	269.61	801326686	668472456	
sdb	371.31	945.97	1073.33	2345452365	2661206408	
sdb1	371.31	945.95	1073.33	2345396901	2661206408	
sdc	408.03	207.05	972.42	513364213	2411023092	
sdc1	408.03	207.03	972.42	513308749	2411023092	



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

40

The Linux/Unix **iostat** command is used for monitoring the performance of system input/output devices. It generates reports that can be used to change system configuration to better balance the input/output load between physical disks.

The command syntax is: **iostat [options] [device [...] | ALL] [-p [device [,...] | ALL]] [interval [count]]**

The **iostat** command generates three types of reports, the CPU Utilization report, the Device Utilization report and the Network Filesystem report. **iostat** without any argument displays information about the CPU usage, and I/O statistics about all the partitions on the system. The Device Utilization report is relevant to storage environments.

Options

- d Display the device utilization report.
- k Display statistics in kilobytes per second instead of blocks per second. Data displayed are valid only with kernels 2.4 and later.
- m Display statistics in megabytes per second instead of blocks or kilobytes per second. Data displayed are valid only with kernels 2.4 and later.
- N Display the registered device mapper names for any device mapper devices. Useful for viewing LVM2 statistics.
- p [{ device [,...] | ALL }] The -p option displays statistics for block devices and all their partitions that are used by the system. If a device name is entered on the command line, then statistics for it and all its partitions are displayed. Last, the ALL keyword indicates that statistics have to be displayed for all the block devices and partitions defined by the system, including those that have never been used. Note that this option works only with post 2.5 kernels.
- t Print the time for each report displayed. The timestamp format may depend on the value of the S_TIME_FORMAT environment variable (see below).
- V Print version number then exit.
- x Display extended statistics. This option works with post 2.5 kernels since it needs /proc/diskstats file or a mounted sysfs to get the statistics. This option may also work with older kernels (e.g. 2.4) only if extended statistics are available in /proc/partitions (the kernel needs to be patched for that).
- z Tell iostat to omit output for any devices for which there was no activity during the sample period.

VMware vCenter – Hosts and Clusters View



VMware vCenter Server provides a centralized platform for managing VMware vSphere environments. Vsphere is VMware's platform for building virtualized cloud infrastructures. vCenter can be used to deploy and manage virtual infrastructures. This is the Hosts and Clusters view just under **Inventory >> Home**. As you can see in the tree in the left pane, there is a virtual data center. Under the virtual data center there are clusters. A cluster is a group of hosts. When a host is added to a cluster, the host's resources become part of the cluster's resources. The cluster manages the resources of all hosts within it.

VCenter Server also provides resource provisioning and performance evaluation of virtual machines residing in a distributed virtual data center. Performance information of machines in the data center can be seen under the Performance tab.

Lesson Summary

During this lesson the following topics were covered:

- Perfmon (Windows)
- sar (Unix/Linux)
- iostat (Unix/Linux)
- vCenter (ESXi)



This lesson covered hosts tools that can be used to monitor and collect performance data from the host perspective.

Lesson: Benchmarking Tools

This lesson covers the following topics:

- Benchmarking tools
 - btest
 - Iometer
 - dbench
 - vdbench
 - IOzone
 - SLOB [Silly Little Oracle Benchmark]
 - VMware View Planner

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

EMC²

43

Many traditional benchmark tools write the same data repeatedly on the disks. However since systems like XtremIO automatically perform in-line deduplication, the disks are not really being written to. This results in less I/O activity on the disks and the deduplication ratio increases exponentially.

A new breed of benchmark tools have been developed. These included EMC's btest, the open source Iometer, and vdbench. This lesson introduces these benchmark tools, and others, along with VMware View Planner.

Benchmark Tool Issues

- Old versions of vdbench, Iometer do not support deduplication
- SQLIO writes all zeroes
- Confirm that other tools are deduplication/compression aware

Note again that, in order for XtremIO testing to be representative of real-world performance, the tools must be able to generate workloads with unique, rather than endlessly repeated, data patterns.

btest

- Developed by EMC
- Can run against raw devices or file systems
- Open Source: available at <http://btest.sourceforge.net/>
- No installation required
- Copy the btest executable file to a Linux machine and start test
- **NOTE:** Beware of using the tool on your local disk!

btest is an I/O benchmarking utility for Linux hosts which produces synthetic random and sequential IO patterns. This application was developed by EMC, and is available as free shareware. **btest** allows you to control the deduplication and compression ratios for testing modern storage devices like XtremIO. It supports multiple file and block devices and supports sync, direct, aio, and sg. **btest** supports both spinning drives and SSDs. It has data verification and provides latency histograms. It supports subtotal, differential and CSV reports.

btest Example

```
./btest -D -b 4k -B 0 -t 0 -T 32 -w 2 R 50 /dev/sdb  
/dev/sdc /dev/sdd /dev/sde
```

Improper use can overwrite hard drive!

- Use direct I/O with 4K size and ignore timeouts
- Run indefinitely with 32 threads and two workers per thread
- Use 100% random I/O with a 50/50 read/write split
- Use disks /dev/sdb, /dev/sdc, /dev/sdd and /dev/sde
- Shortcut to specify many drives: /dev/sd[bcde]

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview



46

Here, we see an example of using btest that will use direct I/O with 4K block size and ignore timeouts. It will run indefinitely with 32 threads and two workers per thread. The test will run on devices /dev/sdb, c, d, and e and use 50 percent random read and 50 percent random write I/O.

A shortcut for specifying multiple disks is to group them with brackets. The example shown here would use disks sdb through sde.

Note that btest can **overwrite** your hard drive, if not used carefully.

Some other items to consider when using btest are that using the 4K block size will give the best performance numbers but using larger block size can also be useful when testing the fabric and for bandwidth oriented tests. You can tweak the number of workers and threads according to the host throttle point. Lastly, using the multipath device addresses such as /dev/mapper/35 may give better performance and will prevent overwriting local drives as well.

Iometer

- Iometer is an EMC approved tool used to generate workloads for comparison and to measure performance
- Used for the measurement and characterization of:
 - Performance of disks and network controllers
 - Bandwidth and latency capabilities of buses
 - Network throughput to attached drives
 - Shared bus performance
 - System-level hard drive performance
 - System-level network performance
- Free download at iometer.org
- Runs on Windows, Linux, and Unix platforms



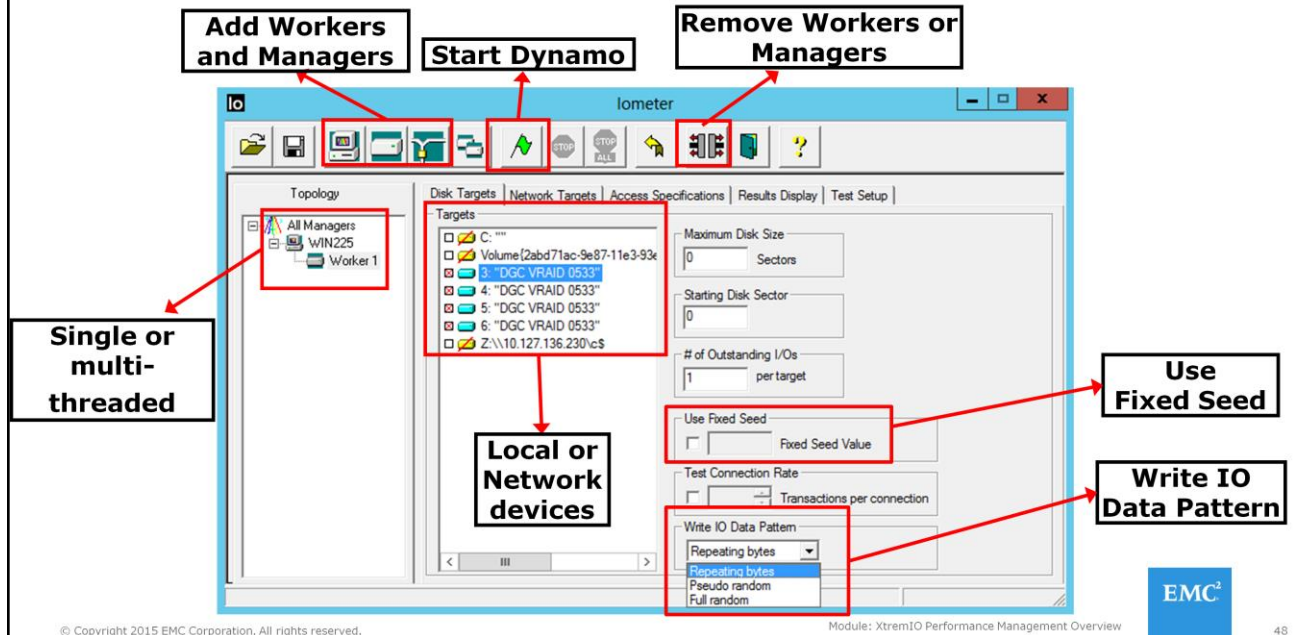
© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

47

Iometer is both a workload generator and a measurement tool. It creates a file and performs I/O operations on that file in order to stress the file system. It then evaluates and records the performance of its I/O operations. **Iometer** can be configured to emulate the disk or network I/O load of any program or benchmark, or can be used to generate entirely synthetic I/O loads. It can generate and measure loads on single or multiple systems. Iometer can be run on most Windows, Linux and Unix platforms. **Iometer** has been enhanced to write random data.

Iometer Interface



This is the **Iometer** graphical user interface. This is what you will see every time you access the tool. In the top left-hand corner you will find the open and save icons. Next to these you will find more icons to add managers and workers for multi-threaded workloads. Here is also where you will find the icon to remove any workers or managers, and a green flag to start Dynamo and generate the workload.

In the middle pane, you find the disk targets, which can be raw devices or file systems. When you map a drive to a NAS CIFS server, for example, it will show up here.

Formatted drives are accessed by writing to a file called **\iobw.tst**. If this file does not exist, the drive's icon will have a red diagonal line through it. At the start of the test the file will be created and grown until the file system or disk is full.

This initialization can take several minutes to hours, depending on the size of the file system/disk. The **iobw.tst** file will fill the entire disk target, unless a maximum disk size is used.

Iometer allows user configuration of access specifications. The user has control over the read/write ratio, I/O sizes, randomness, and burstiness of data.

dbench

- Filesystem and CIFS and NFS benchmark
- Generates load patterns
- Has become the de-facto standard for generating load on a Linux VFS

```
dbench [OPTIONS] <num-procs>
```

```
dbench [-B --backend=<dbench backend>] [-c --loadfile=<filename>] [-D --  
directory=<string>] [-F --fsync] [-R --target-rate=<double>] [-s --sync] [-S --  
sync-dir] [-t --timelimit=<integer>] [-T --tcp-options=<string>] [--run-once]  
[--fake-io] [--scsi=<scsi-device>] [--server=<hostname>] [--export=<string>] [-  
protocol=<string>] [--clients-per-process=<integer>] [--trunc-io=<integer>] [-  
stat-check] [--skip-cleanup] [--per-client-results] [--iscsi-portal=<ip-  
address>] [--iscsi-port=<port>] [--iscsi-target=<iqn name>] [--iscsi-lun=<LUN>]  
[-? --help] [--usage]
```



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

49

DBENCH is a tool to generate I/O workloads to either a filesystem or to a networked CIFS or NFS server. It can also communicate with an iSCSI target. **DBENCH** can be used to stress a filesystem or a server to see which workload becomes saturated. It can also be used for prediction analysis to determine how many concurrent clients and/or applications performing a workload can a server handle before response starts to lag.

DBENCH is free software. It is considered a de-facto standard for generating load on the Linux VFS.

vdbench

- Java-based, multi-host, compression and deduplication aware, CLI/Config file based
- Can be run on raw disk and file systems
- Originally open source but now owned by Oracle
- Located at: <http://www.oracle.com/technetwork/server-storage/vdbench-downloads-1901681.html>
- Part of the PoC Toolkit for testing AFAs



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

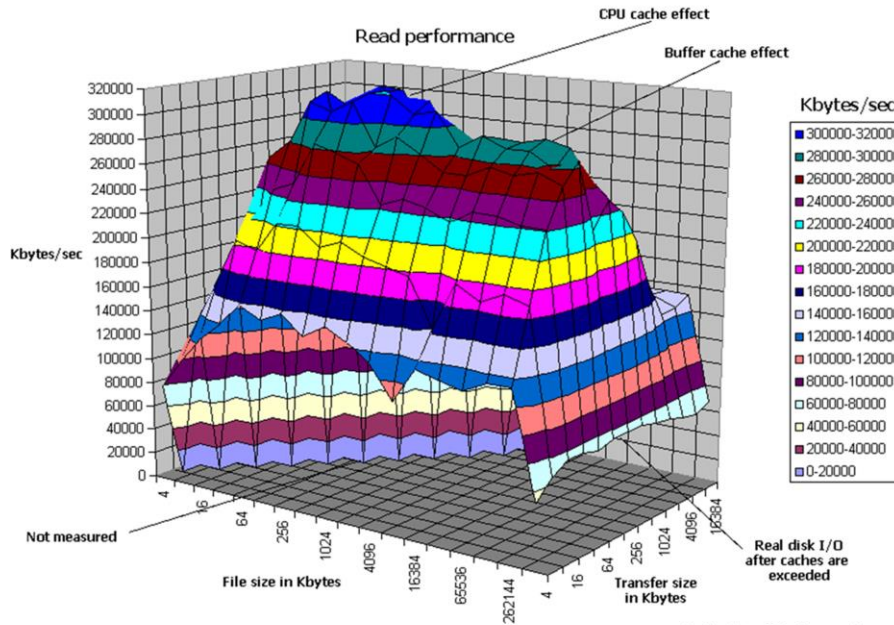
50

Vdbench is a command line utility that can be used generate disk I/O workloads. **Vdbench** can be used to validate storage performance and data integrity. **Vdbench** execution parameters may also specified via an input text file.

Vdbench was a open source software that was available at:
<http://sourceforge.net/projects/vdbench/>.

However, at the release time of this course, Oracle is the new owner and the rules for use will probably change. **Vdbench** has been moved to Oracle Technology Network (OTN) at **<http://www.oracle.com/technetwork/server-storage/vdbench-downloads-1901681.html>**. There is now an OTN License Agreement that must be accepted to download the software.

IOzone



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

EMC²

51

IOzone is a filesystem benchmark tool. The benchmark generates and measures a variety of file operations such as reading and writing to a file. **IOzone** can perform sequential and random operations to the file. **IOzone** is useful for performing a broad filesystem analysis of a vendor's computer platform. The benchmark tests file I/O performance for the following operations:

- Read, wwrite, re-read, re-write, read backwards, read strided, fread, fwrite, random read, pread, mmap, aio_read, and aio_write

IOzone is available for most Windows, Linux, FreeBSD and Unix platforms. It can be downloaded from <http://www.iozone.org>.

In the diagram, various file sizes are tested with a number of different I/O sizes. When the file is small, the entire data set fits into CPU cache or the host buffer cache; little to no disk I/O is required, and bandwidth is very high. As the size of the file being tested increases, more physical disk access is performed, and performance drops off rapidly, as shown by the steep cliff on the right-hand side of the chart.

SLOB

- Silly Little Oracle Benchmark
- Load generator for Oracle DB
 - All IO actually occurs via Oracle DB
- Configurable database size, update rate, number of virtual users
- Results are collected by the Oracle Automatic Workload Repository (AWR)

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Management Overview

EMC²

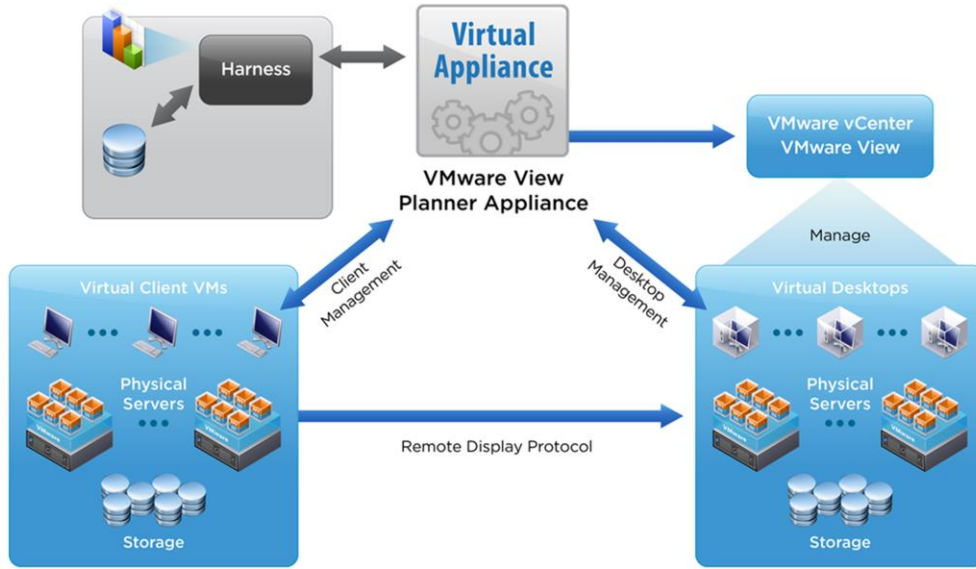
52

SLOB is a popular test when Oracle databases will be used in the production environment. It generates loads by simulating database users, and allows a number of options to be configured.

The Oracle Automatic Workload Repository [AWR] is used to collect performance statistics including:

- Wait events used to identify performance problems.
- Time model statistics indicating the amount of DB time associated with a process.
- Active Session History (ASH) statistics.
- Some system and session statistics.
- Object usage statistics.
- Resource intensive SQL statements.

VMware View Planner



VMware View Planner is a comprehensive standard methodology for comparing virtual desktop deployment platforms. View Planner generates a realistic measure of client-side desktop performance for all desktops being measured on the virtual desktop platform. View Planner uses a rich set of commonly used applications as the desktop workload.

It can be download at <http://www.vmware.com/products/view-planner.html>.

Lesson Summary

During this lesson the following topics were covered:

- btest
- Iometer
- dbench
- Vdbench
- VMware View Planner



This lesson covered benchmarking tools for use in XtremIO environments.

Module Summary

Key points covered in this module:

- Performance analysis issues
- Relating basic performance measurements to I/O operations
- Monitoring performance of XtremIO storage systems
- Tools to measure and benchmark performance
- Performing data collection on an XtremIO storage array



This module focused on the basic principles of performance management, the XtremIO architecture from a performance perspective, and the monitoring and testing of performance on an XtremIO Storage Array.

This slide intentionally left blank.

Module: XtremIO Performance Analysis

Upon completion of this module, you should be able to:

- Discuss expected XtremIO performance
- Analyze XtremIO performance data
- Locate performance bottlenecks
- Resolve XtremIO performance issues



This module focuses on the analysis and resolution of XtremIO performance issues.

This slide intentionally left blank.



Lesson: Optimizing XtremIO Performance

This lesson covers the following topics:

- Metrics used to measure performance
- Expected performance from infrastructure components
- How I/O size and type affects performance

This lesson covers the optimization of XtremIO performance. We will begin by looking at the expected maximum performance in the environment, then look at ways to optimize the environment.

Key Performance Metrics

- The 3 main metrics are IOPs, bandwidth, and latency
- IOPs [throughput]
 - The number of I/O operations completed in one second
 - Measured in operations per second
- Bandwidth
 - The amount of data transferred in one second
 - Measured in megabytes per second [MB/s]
- Latency [response time]
 - Time taken to complete an I/O operation
 - Measured in milliseconds [ms] or microseconds [μs]

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

4

XtremIO storage systems can report on a number of performance metrics. The three most important metrics are IOPs [throughput], bandwidth, and latency.

IOPs is a measure of the number of I/O operations per second performed by the storage system. There are a number of places where there may be limitations on the number of IOPs that can be performed. We will investigate these bottlenecks in subsequent slides.

Bandwidth is a measure of the total amount of data that is transferred in a given time. Bandwidth is the product of IOPs and the average I/O size.

Latency is a measure of the time taken to complete an operation. It is typically measured in milliseconds or microseconds. In the XtremIO arrays the latency as measured at the array, will be in the sub-millisecond range.

Expected Performance

- Performance depends on the environment and the array
- Connectivity plays a large part in the calculation
 - Latency seen by the host includes the path to the storage
 - Path bandwidth limitations could act as a bottleneck
- Hosts need to be correctly configured
 - HBA driver configuration
 - File system queue depths, coalescing

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

5

In order to determine whether or not the observed performance is acceptable, we need to know what the expected performance is. Performance depends not only on the array, but also on the environment in which it is installed. The connectivity between hosts and the array will play a large role. The hosts themselves will also need to be configured correctly in order to achieve optimum performance.

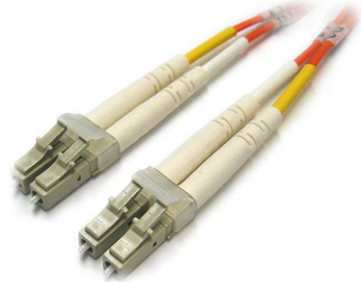
Fibre Channel Performance

4 Gb/s link = 4250 Mb/s

- 8b/10b Encoding: $4250 * (8/10) = 3400$ Mb/s
- Maximum Throughput ≈ 425 MB/s
- Maximum IOPs (4 kB) $\approx 108,800$ IOPs

8 Gb/s link = 8500 Mb/s

- 8b/10b Encoding: $8500 * (8/10) = 6800$ Mb/s
- Maximum Throughput ≈ 850 MB/s
- Maximum IOPs (4 kB) $\approx 217,600$ IOPs



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

6

Here we see the maximum performance available from 4 Gb/s and 8 Gb/s Fibre Channel connections. The relationship is linear as the technology improves, with 8 Gb/s speeds double that of 4 Gb/s speed; 16 Gb/s Fibre Channel will have double the bandwidth of 8 Gb/s FC. It is important to note that the speeds shown include overhead related to frame headers, etc., so the transfer speed achieved for user data on a link will be slightly lower than the IOPs value shown here.

iSCSI Performance

10 Gb/s link

- 64b/66b Encoding: $10000 \times (64/66) = 9697 \text{ Mb/s}$
- Maximum Throughput $\approx 1.18 \text{ GB/s}$
- Maximum IOPs (4 kB) $\approx 310,300 \text{ IOPs}$



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

7

For iSCSI connections, note that a 10 GbE link has a transfer rate of 10×10^9 bits/s [this is a serial link, using decimal values for 'Giga']. After compensating for the 64/66 bit encoding used, that translates to roughly 9697 Mb/s, which is 1.18 GB/s [binary 'Giga']. That gives a theoretical maximum number of 310,304 IOPs at 4 kB I/O sizes, which does not take into account the overhead due to frame and packet headers, etc. The maximum rate under normal operating conditions is likely to be appreciably lower than this.

Comparison of Fibre Channel and iSCSI

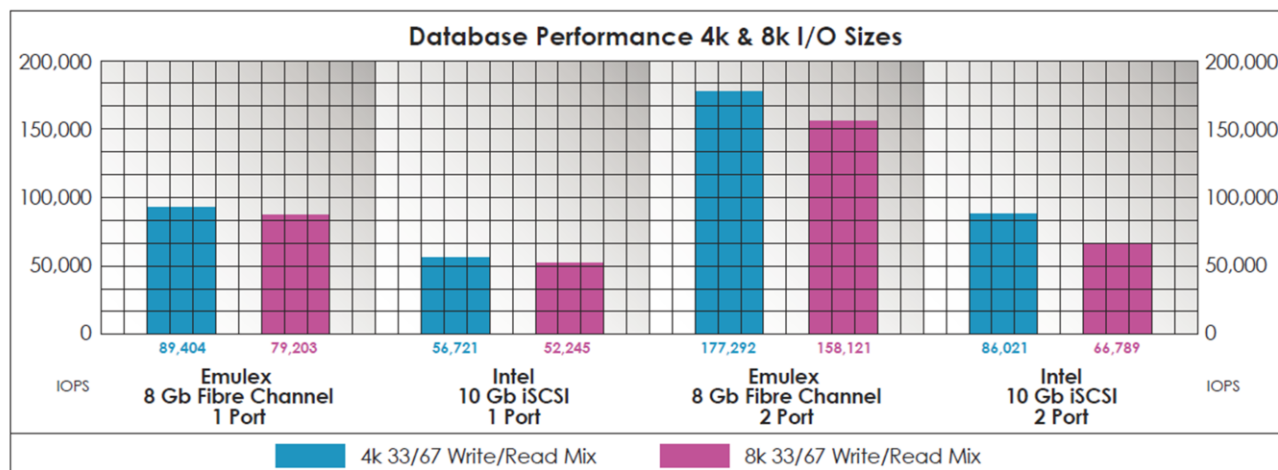


Chart from “**An Analysis of 8 Gigabit Fibre Channel & 10 Gigabit iSCSI** in Terms of Performance, CPU Utilization & Power Consumption” by Third IO, Inc.



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

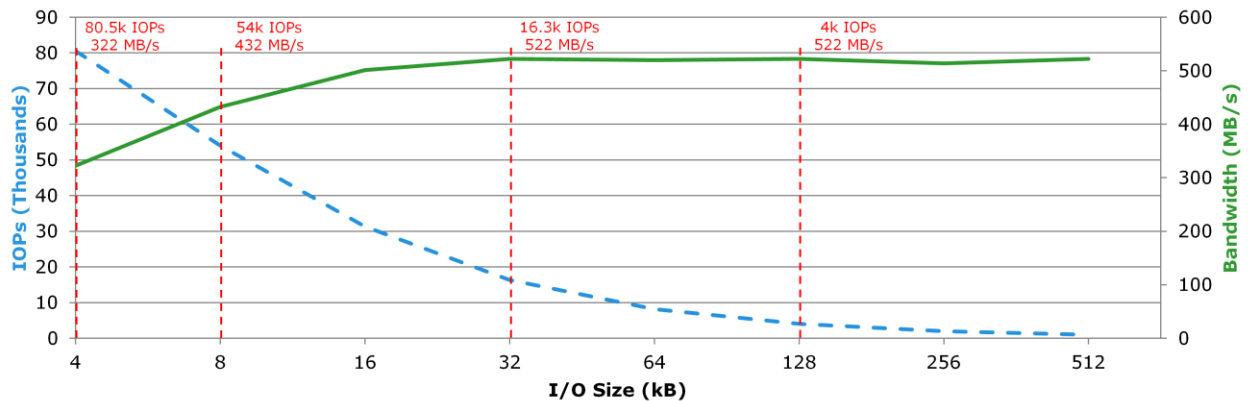
8

The chart was taken from the white paper “**An Analysis of 8 Gigabit Fibre Channel & 10 Gigabit iSCSI** in Terms of Performance, CPU Utilization & Power Consumption” published by Third IO, Inc., and available on their website, www.thirdio.com. Though the paper is old [around 2009], the technologies have not changed, and results are still relevant.

It is clear that, theoretical maxima notwithstanding, 8 Gb/s FC outperforms 10 Gb/s iSCSI by a considerable margin. FC performance scales well from 1 to 2 ports; iSCSI performance does not. Note: The numbers achieved by even a fairly powerful host are well below theoretical maximum numbers. In order to achieve the maximum performance of a fast array, such as the XtremIO, multiple hosts are required.

In XtremIO environments with powerful hosts, iSCSI shows 8% - 35% lower performance than FC, depending on I/O size and R/W ratio.

I/O Size, Bandwidth and IOPs



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

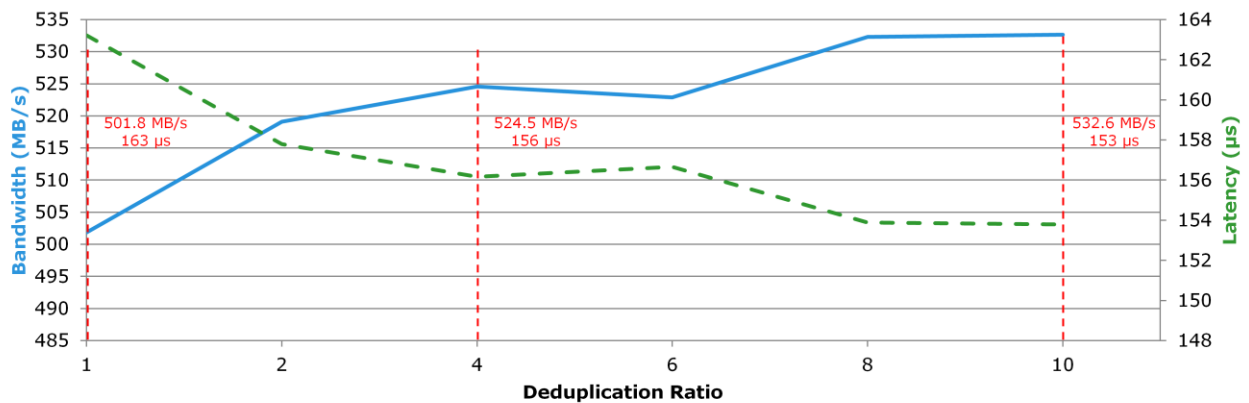
EMC²

9

The chart shows a single host running random writes of different I/O sizes. IOPs decrease while bandwidth increases, and then levels out. At around 522 MB/s we've hit a bottleneck, which could be the host or the connection. The array is certainly capable of servicing a greater workload than the one illustrated here.

These curves are typical of any storage system – as I/O size increases, throughput will decrease and response time [not shown here] will increase as well. The increase in response time is seen because larger I/Os take longer to traverse the link, and take longer to service at the storage system. Bandwidth will eventually plateau – the point at which this happens depends on which component is the bottleneck for the environment.

Deduplication, Bandwidth, and Latency



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

10

This chart shows the example of a single host running random writes of 4 kB. Reads were not performed in this test because they don't affect deduplication. As the deduplication ratio increases, bandwidth increases while latency decreases. This is expected behavior, because less data is actually written to the Flash drives as the level of deduplication increases. Since deduplication is always active on XtremIO systems, there will always be some performance improvement as long as data is even slightly deduplicable.

Data Alignment

- I/O sizes should be multiples of 8 kB where possible
- File system data should be aligned at 8 kB boundaries
- Newer Windows, Linux, and ESXi systems do this automatically
 - Older systems may require the use of fdisk or diskpart

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

11

Data alignment in XtremIO environments refers to two different types of alignment: I/O size, and alignment of user I/O with the structure on the volume.

The first requires that I/O sizes be a multiple of 8 kB where possible. This allows the I/Os to match the internal structure of the system, and will produce improved performance as seen in the previous slide.

The second requires that user I/Os start at a boundary on the volume which is a multiple of 8 kB. Most modern operating systems will perform this alignment by creating a reserved area at the beginning of the volume. A size of 1 MB is used by default in Windows and ESXi environments; other environments may use a different size. Where volumes are created manually, host tools such as fdisk or diskpart may need to be used to perform the alignment. A value of 64 kB will be sufficient for a file system on a host with an Intel x86-based architecture.

Lesson Summary

During this lesson the following topics were covered:

- The three main metrics used to measure performance are
 - Bandwidth, throughput [IOPs], and latency [response time]
- Fibre Channel and iSCSI have vastly different performance
- I/O size, alignment, and deduplication affect performance
- Optimizing the XtremIO environment must always involve the SAN and host systems

This lesson covered the metrics that are used to measure and characterize performance. The connectivity between the array and the host plays a large role. Hosts must be correctly configured to support the massive I/O numbers that XtremIO arrays can process.

Lesson: Optimizing Connectivity

This lesson covers the following topics:

- Zoning best practices
- VSANs
- ISLs
- NPIV

This lesson covers host to array connectivity, and what can be done to optimize performance.

FC Zoning

- An FC zone is configured on an FC switch and contains WWNs of initiators and targets
- A zone allows devices to communicate over the SAN
- Best Practices for FC zoning are:
 - WWPN based zoning (opposed to port number based)
 - Single initiator per FC zone (1:many relationship)
 - A host should at minimum be zoned to two SCs from the same X-Brick
 - A host should be zoned in an HA topology
 - A host should have no more than 16 paths to a LUN

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

14

FC zoning best practices are shown here. Note that these best practices are commonly accepted in the IT industry [though other storage arrays may not call their storage controllers SCs].

FC VSANs

- A VSAN is configured on a FC Switch and contains switch ports. It is used to isolate the SAN environments (similar to VLANs)
- Each VSAN is a separate self-contained fabric using distinctive security policies, zones, events, memberships, and name services
- VSANs may be used to isolate an ISL to specific traffic
- Best Practice for VSANs:
 - If XtremIO has a dedicated VSAN, use a dedicated ISL

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

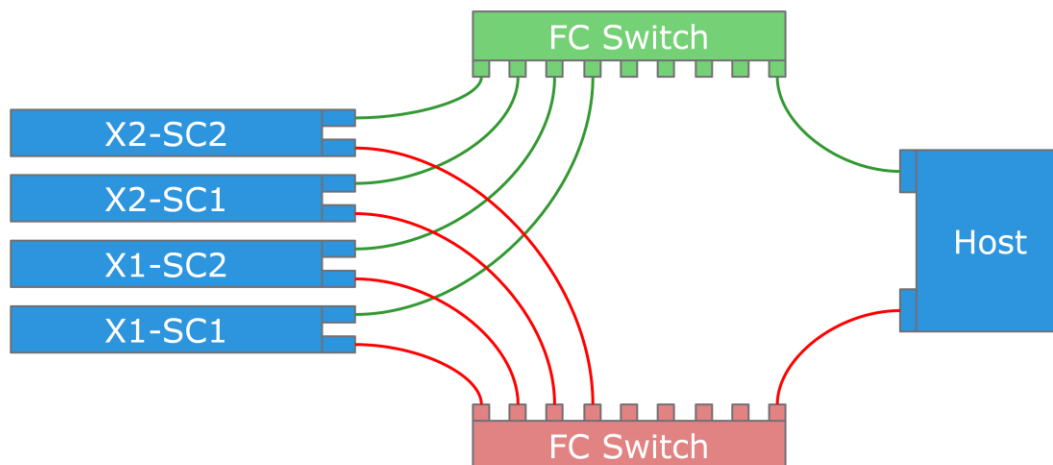
EMC²

15

A VSAN [Virtual Storage Area Network] allows logical separation of SAN traffic where no physical separation exists. Each VSAN behaves as though it is an independent SAN, and events on one VSAN do not affect any others.

In VSAN environments where there are multiple switches, and therefore ISLs, a dedicated ISL should be used if the XtremIO traffic is allocated to a dedicated VSAN.

Zoning Best Practices



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

16

This slide shows the zoning best practices for a dual X-Brick XtremIO environment. An HBA should ideally be connected to one port from each SC. If that is not possible, then at minimum the HBAs should be connected to one port on each SC in one of the X-Bricks in the cluster.

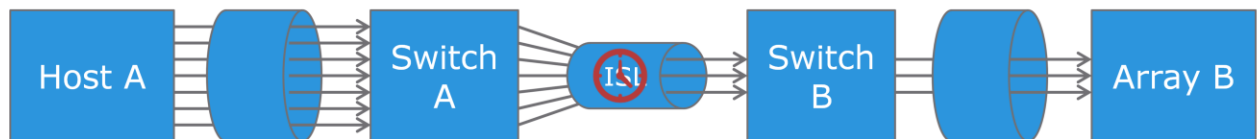
Each X-Brick of an XtremIO system can lose one SC and still remain fully functional. For example, in this two X-Brick system, we can lose two SCs – one in each X-Brick – and all of the data will still be available. A zoning scheme that connects half of the hosts to SC1 on each X-Brick and the rest of the hosts to SC2 would be unsound. In the very unlikely event that SC1 on each X-Brick were to fail, half the hosts would unnecessarily lose connectivity.

In general, every host should be connected to as many SCs as possible on an XtremIO cluster, as long as the host and multi-pathing software supports that number of connections. Best practice is 4-8 paths.

This is an example with Fibre Channel connectivity, but the same philosophy applies to iSCSI as well.

SAN ISLs

- Large environments often have multiple FC or iSCSI switches
- The links which connect these switches are Inter-Switch Links (ISLs)
- ISLs may become over utilized and may cause frames to be discarded



© Copyright 2015 EMC Corporation. All rights reserved.

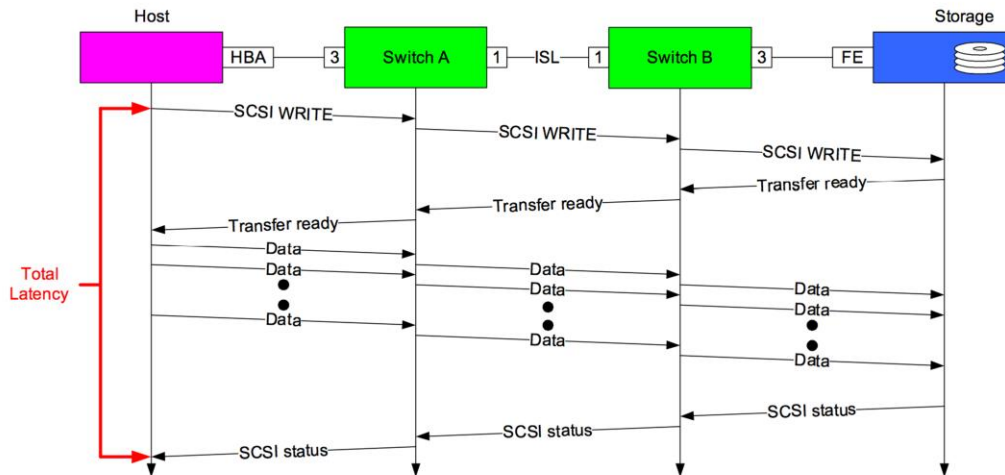
Module: XtremIO Performance Analysis

EMC²

17

Where ISLs are used, they should be carefully sized to ensure that they can carry all the required traffic without delay. A best practice is to keep busy initiator ports close to the target ports that they use. Since the targets ports [typically storage array ports] are shared across many initiator ports [typically HBA ports], the targets should be on the fastest switch ports, and initiators placed on slower switch ports, or ports with shared bandwidth.

SAN ISLs: Writes



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

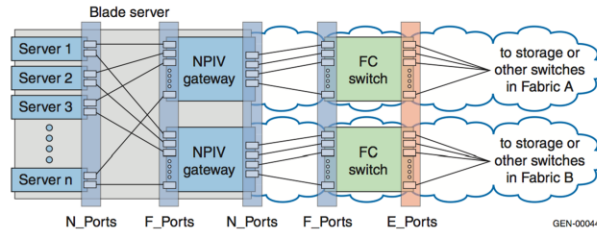
18

When data is written, the SCSI WRITE command must be sent and the storage port must respond with a transfer ready. The data will then be transmitted by the host and a status returned from the storage.

When you include the latency each way with the SCSI WRITE command, the command completion time will be delayed by twice the round trip latency, since the command cannot be considered to be completed by the host until the status is received.

NPIV

- A number of vendors supply switches that allow NPIV:
 - Cisco/Brocade
 - HP
- Not all NPIV features are supported by all vendors



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

19

Note that vendors may have some proprietary features available for NPIV environments, and that those features may not be supported by other vendors. These features include port channeling – the ability to combine several physical links into one aggregated logical link – and trunking – the ability to carry traffic [using the EISL format] for multiple VSANs across a link.

SAN ISL Issues

CRC Errors indicate corruption inside the FC data frame

- Usually a physical problem (cable/SFP) within the FC link

Class 3 Discards usually indicate a very busy fabric

- Can be due to timeouts or invalid/unreachable destinations
- This counter may increment during normal operation as well

BB Credit represents the number of frames a port can store

- A low value can slow down the environment
- Usually only modified by vendors
- Faster or long distance networks use a larger value

Here are some areas where issues can arise in a SAN environment. Some errors may indicate a physical problem in the environment, while others are related to the level of traffic.

Similar issues can arise in LAN environments, and would affect the iSCSI traffic to an XtremIO array.

Lesson Summary

During this lesson the following topics were covered:

- Choice of connection protocol is important
- Inter-switch connections must be robust
- VSANs or VLANs can help to isolate traffic

This lesson covered host to array connections, and how they should be configured for best performance.

This slide intentionally left blank.

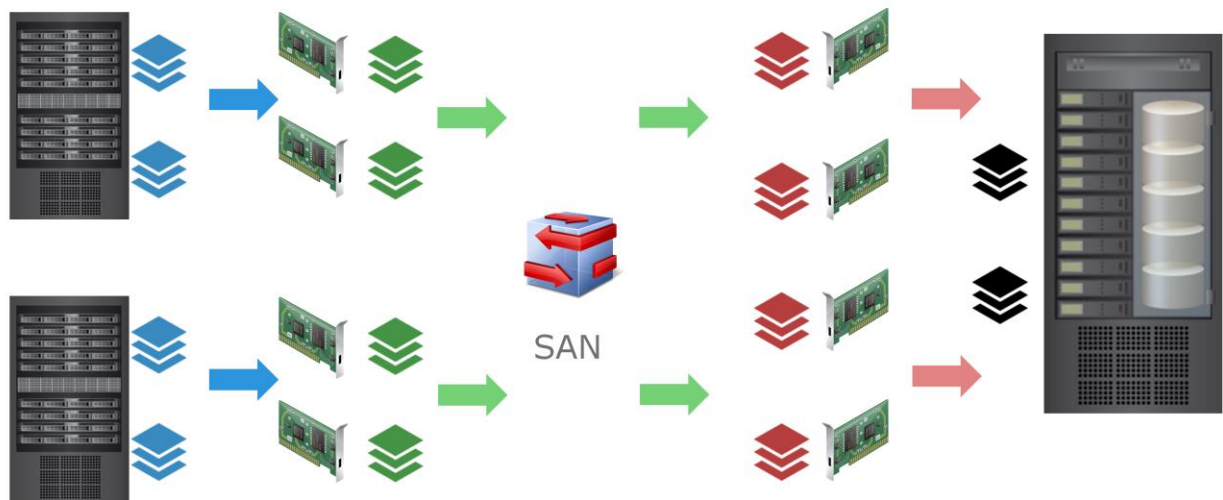
Lesson: Optimizing Hosts

This lesson covers the following topics:

- Queueing in a storage environment
- Configuring HBA queues
- Configuring host settings for storage optimization
- Aligning data to the storage system structure
- Optimizing specific operating systems
 - Windows, Linux, ESXi, Solaris, AIX

This lesson covers host optimization. Hosts will have configuration parameters associated with storage which can be changed to improve performance. An example is the HBA queue depth. Some host behavior which is optimized for mechanical hard drives will need to be disabled or adjusted for all-Flash systems. Different operating systems will have different configuration requirements.

Queuing In a Storage Environment



© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Performance Analysis

EMC²

24

The slide shows queuing of IOs at a different components in a typical storage environment.

Typical host environments have many points at which queuing can take place. The host will have application and operating system queues, which may be user-configurable.

Host connectivity to the SAN involves host HBA ports, FC or iSCSI, which have their own queues associated with them. Those queues, which are configurable through the HBA driver software, often allow the user to select queue depths for each target port and for each LUN/volume.

At the storage system, front-end ports will have their own queues; these are not typically configurable by the user.

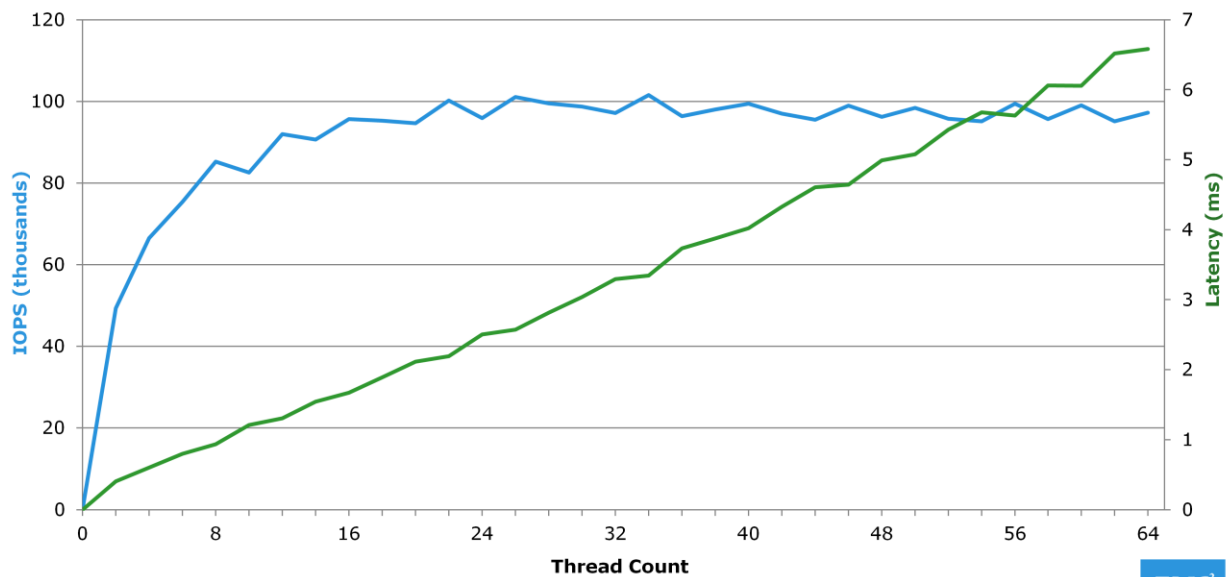
The storage system will also have queues at the controller, LUN/volume, and disk level, and these are also typically not user-configurable.

Configuring Queues – An Introduction

- Queue Depth refers to the number of pending I/O requests per volume path
- Execution Throttle refers to the number of outstanding I/O requests per target port
- XtremIO recommends setting these to a high value
- Default values are often too low for high-performing storage
 - It is not always true that increasing these values causes better performance for any given host

The terms discussed here are used in XtremIO documentation. Note that in other environments, other terms may be used for queue depth and execution throttle. Since the former refers to LUN/volume queuing, it may be referred to as LUN queue depth; whereas, execution throttle may be referred to as target queue depth.

IOPs and Latency vs. Thread Count



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

26

This chart shows a single host running small random I/Os. Note that as the thread count increases, the IOPs value reaches a peak, and then decreases slightly as additional workload is added. Latency increases linearly with thread count.

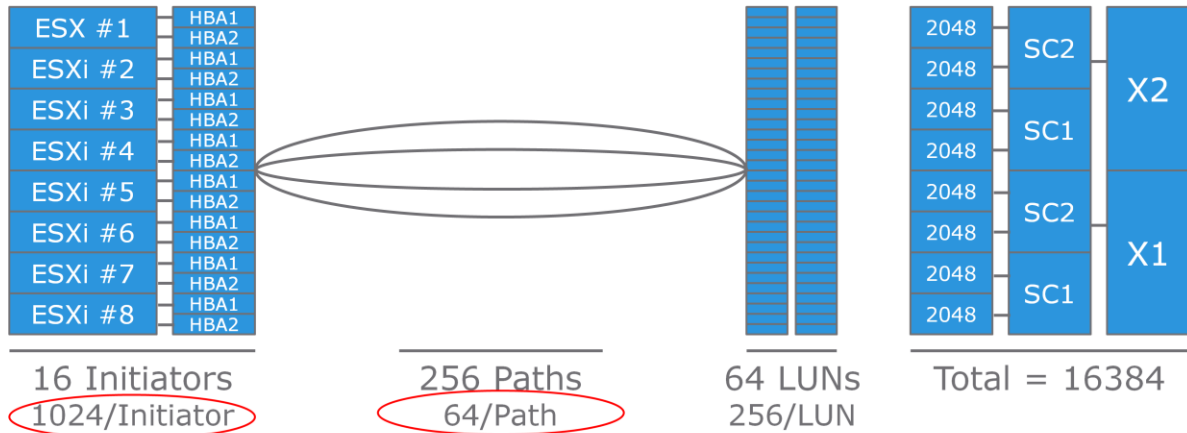
Note that the leveling off of IOPs is [in this case] a bottleneck at the host rather than at the storage system. Also note that at a thread count of 32, which is the default queue depth value for many HBAs, the IOPs value has already reached a peak, and increasing the queue depth would not improve performance.

Example 1: Execution Throttle/Queue Depth

- The Storage Controller front-end ports have a queue of 2048
- Example environment:
 - Dual X-Brick XtremIO Array
 - ESXi farm of 8 ESXi hosts
 - Each ESXi host has 2 single port HBAs
 - Each ESXi host has 8 LUNs allocated to it
 - Each LUN is visible through 4 paths

This is the first of two examples which will illustrate the procedure used to calculate the execution throttle and queue depth for a given host.

Execution Throttle Configuration Example 1



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

28

In this example, there are 2 X-Bricks, each with 4 front-end ports. Since each port has a queue depth of 2048, the aggregate queue depth is 16384.

At the host side, there are 16 initiators, so we can set the queue depth at the initiator to 1024.

Because there are 64 LUNs each with 4 available paths, there are a total of 256 paths. The queue depth per path can therefore be set to 64.

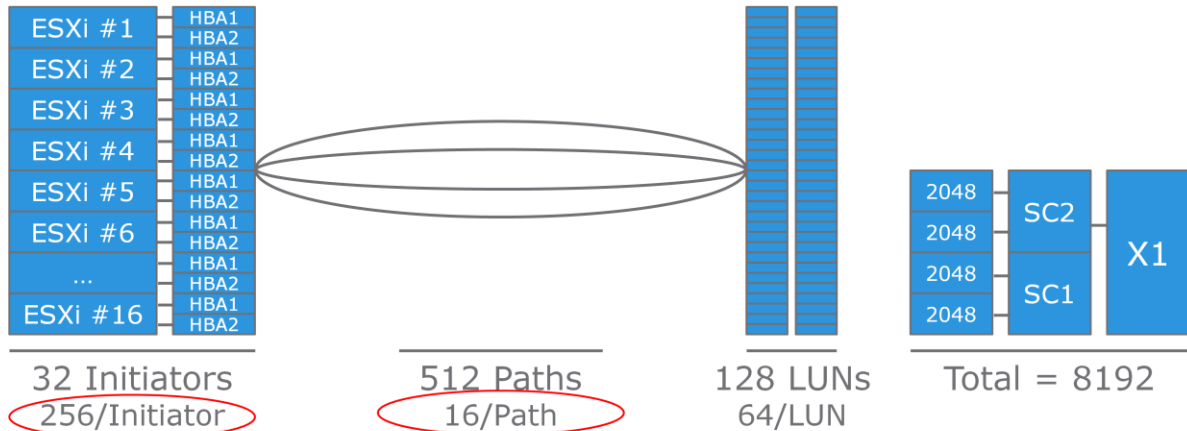
The values for the hosts will be an execution throttle of 1024 and a queue depth of 64.

Example 2: Execution Throttle/Queue Depth

- The Storage Controller front-end ports have a queue of 2048
- Example environment:
 - Single X-Brick XtremIO Array
 - 2 ESXi farms of 8 ESXi hosts each
 - Each ESXi host has 2 single port HBAs
 - Each ESXi host has 8 LUNs allocated to it
 - Each LUN is visible through 4 paths

This is the first of two examples which will illustrate the procedure used to calculate the execution throttle and queue depth for a given host.

Execution Throttle Configuration Example 2



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

30

In this example, there is 1 X-Brick with 4 front-end ports. Since each port has a queue depth of 2048, the aggregate queue depth is 8192.

At the host side, there are 32 initiators, so we can set the queue depth at the initiator to 256.

Because there are 128 LUNs each with 4 available paths, there are a total of 512 paths. The queue depth per path can therefore be set to 16.

The values for the hosts will be an execution throttle of 256 and a queue depth of 16.

Queuing Considerations

- Changing queue values
 - Slow storage systems may be adversely affected
 - Different HBA vendors may not allow the same range of values
- Be aware of the environment as a whole
- Check host queue values if storage systems are underperforming

In environments where hosts are connected only to XtremIO systems, performance of the environment will be improved by making changes to queue depths and the execution throttle. In cases where there are also other storage systems, particularly those that use spinning HDDs, changing those values may cause performance degradation. The SCSI QFULL condition can result when too many I/Os are sent to a storage system port which is unable to service them. This condition is often not handled gracefully by the HBA driver/OS, and can lead to I/O being suspended for several seconds each time it occurs.

I/O Elevators in Host Systems

- I/O elevators are re-ordering schemes used to improve performance
 - They are designed around the use of spinning disks
 - The simplest, no re-ordering, allows FIFO ordering
 - Other schemes base their decision on a number of factors
 - I/O size, data position, queued requests, etc.
- Disable elevators for XtremIO environments
 - Use noop policy

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

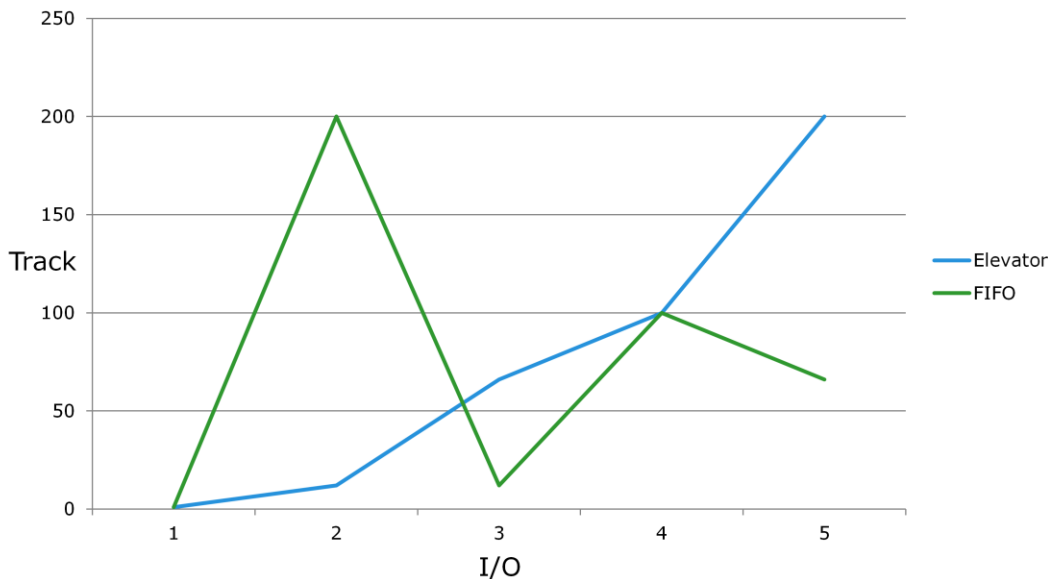
EMC²

32

Disk scheduling algorithms, also called elevators, are designed to improve the performance of storage devices. Since spinning disks are not true random-access devices, their performance is improved when seeks can be reduced. The elevators attempt to reduce the number of seeks.

In environments like XtremIO, where there are no spinning disks, these elevators perform no useful function at all, and in fact may introduce latency. No elevators should be used; instead, the policy should be set to noop – no operation.

FIFO and Elevator I/O Policy



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

33

This chart shows how I/O requests would be handled in 2 different environments – one which services I/Os as they arrive [FIFO], and the other which implements intelligent re-ordering of the requests [Elevator].

Requests are made for data at tracks 200, 12, 100, and 66, with track 1 as the starting track.

The FIFO system moves the R/W head in the sequence 1-200-12-100-66, for a total seek distance of 509 tracks.

The Elevator system moves the R/W head in the sequence 1-12-66-100-200, for a total seek distance of 199 tracks. Since the system has to wait for a certain number of I/Os to arrive and be re-ordered, the start of the sequence is delayed. The time saved by more efficient seeking is usually much more significant than the initial delay, so this system is very widely used in arrays that contain rotating media. This optimization is irrelevant in all-Flash environments.

Multipathing Software

- Multipathing drivers have three main jobs:
 - Unite paths from the same device into one multipathed device
 - Perform load balancing between paths
 - Manage path failover and failback
- XtremIO recommends using multipathing drivers
- Most native Multipathing drivers have been certified
- EMC PowerPath is also supported by XtremIO
 - Recommended because of advanced PowerPath features

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

34

Multipathing software not only resolves conflicts that arise when LUNs are seen across multiple paths, but also manages load balancing across those paths, and path failover and failback when errors occur.

XtremIO recommends the use of multipathing software; without it, any host would only be able to have 1 connection to the XtremIO cluster in order to avoid conflicts. This would lead to single points of failure in the environment.

PowerPath implements optimized load balancing and intelligent path management, and is the recommended multipathing software for environments that support it.

Load Balancing Algorithms

- Any LUN may have more than one path
- Multipathing software chooses which path will be used for the next I/O operation
- A few examples of load balancing algorithms:
 - Fixed path: use the same path until it fails, then move to the next path
 - Round Robin: put all paths in a list and send I/Os to them sequentially
 - Least Queue Depth: send I/O to the path with the lowest queue depth
 - Note: Algorithms may have different names in different operating systems

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

35

Here are examples of load balancing algorithms that may be implemented by multipathing software. In many cases, the round robin algorithm works well with XtremIO, particularly if the option can be configured to change paths with each I/O. Specific recommendations exist for different operating systems; refer to the relevant documentation. For PowerPath revisions that natively support XtremIO volumes, PowerPath will choose the correct algorithm.

Unwanted Application/OS Features

Features to disable:

Read Ahead

- Some operating systems and applications will try to utilize the array and predict what blocks are going to be read next.

Write Caching

- Some operating systems and applications will try to utilize the array and save the writes for a time when the load is reduced.

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

36

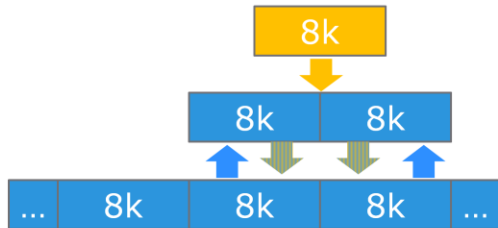
Features such as read-ahead [also called prefetching] and local write caching are often implemented by operating systems and applications in an attempt to optimize the performance of spinning disks. In XtremIO environments, they will lead to increased latency, and should be disabled.

File System Alignment

- Misalignment is the major reason for poor file system performance



Aligned data touches only one file system block and is written directly



Misaligned data touches two file system blocks, and causes read-modify-write operations

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

37

Where possible, ensure that volumes, and file systems, are aligned on an 8 kB boundary. Misalignment affects not only regular reads and writes, but will decrease the benefit of compression and deduplication as well.

Optimizing Specific Operating Systems

- Implement practices listed in the User Guide
- ESX/ESXi: use VAAI commands
- Windows: check alignment
- Linux: check alignment

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

38

For all operating systems, implement the practices listed in the XtremIO User Guide.

VAAI commands, discussed in detail in an earlier module, improve the overall performance of an ESXi environment by offloading I/O work which was traditionally performed by the host to the storage system. That frees up host CPU cycles for other processing, including the processing of I/O.

In Windows and Linux environments, file system alignment is very important, especially with older versions of the OS.

Lesson Summary

During this lesson the following topics were covered:

- Configuring HBA queue depth
- Configuring host settings for storage optimization
- Aligning data to the storage system structure
- Optimizing specific operating systems
 - Windows, Linux, ESXi, Solaris, AIX

This lesson covered the configuration of hosts for optimal performance with XtremIO arrays.

During this lesson we review what a queue depth and execution throttle refers to, along with two examples of execution throttle and queue depth and queuing considerations. We also reviewed Disk scheduling algorithms (or Elevators); the three main functions of multipathing software; load balancing algorithms, and file system alignment and optimizing specific operating systems.

This slide intentionally left blank.

Lesson: Benchmarking

This lesson covers the following topics:

- All-Flash array characteristics
- Preparing for a benchmark
- Running benchmark tests with the PoC Toolkit

This lesson covers benchmarking of XtremIO arrays.

All-Flash Array (AFA) Characteristics

- Flash drive performance depends on its space utilization
 - Empty Flash drives perform much better than full drives
- Flash drives have limited endurance
 - Writes to the drives need to be carefully managed
- Flash drives cannot overwrite existing data
 - New data goes to new locations; old locations must be emptied
- No penalty for random I/O
 - No seeks

© Copyright 2015 EMC Corporation. All rights reserved.

XtremIO Performance Analysis

EMC²

42

All-Flash arrays depend on Flash drives for their performance, and behave very differently to arrays that make use of spinning media. While the performance of spinning media is unaffected by how full the drive is, that is not true of Flash drives. The housekeeping activities performed inside Flash drives move data around (the process is invisible to the end user), so when the drive is empty it is easier to find free space into which to move the data. As the drive fills (as is also true of file systems, for example), the process takes longer, and performance decreases.

In addition, the memory cells used to store the data can support only a limited number of write cycles – tens of thousands to hundreds of thousands of cycles. The drives (or the controllers in some systems) try to spread the writes evenly across all cells – a process known as write leveling or wear leveling. This process involves moving data around in order to free up blocks of cells, and can be time-consuming.

Unlike spinning drives which have variable-length seeks based on the location of the data, Flash drives are true random-access devices, where any read or write takes the same time as any other read or write. As a result, there is no need to try to coalesce data in order to make it more sequential. Testing AFAs therefore concentrates on random data access patterns.

Benchmark Utilities

- Application testing rarely shows bounds of an AFA
- Dedicated benchmark utilities test more thoroughly
 - If they are aware of the array options
- Multiple tools available to generate traffic
 - Iometer, vdbench, btest, FIO, LoadDynamix, etc.
 - Not all tools are deduplication and compression aware

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

43

XtremIO arrays are best tested using a specially designed benchmarking utility. The goal is not to show unrealistic performance by catering to specific array features, but rather to be able to stress the array for extended periods of time to show real-world limits.

Tests that are not compression-aware and deduplication-aware often use all zeroes, or use patterns of data that are identical on each disk sector. Those patterns lead to unrealistic performance, since very little real writing is performed, and lead to inflated deduplication ratios.

IDC Technology Assessment

- IDC published a vendor-neutral guide to All-Flash Array testing
- EMC recommends following this guide for all AFA testing

<http://idcdocserv.com/241856>



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

44

The IDC, a vendor-neutral body, has created guidelines for testing all-Flash arrays. Those guidelines are available at the location shown on the slide. EMC recommends that the guidelines be followed, and in fact they've become the basis for testing of XtremIO arrays in PoC [proof of concept] environments.

XtremIO Proof-of-Concept Toolkit

- Standard methodology for testing an All-Flash Array
- Based on IDC testing recommendations
- vdbench used for workload generation
 - Industry-Standard I/O toolkit from Oracle
- Linux Virtual Appliance based
 - Distributed as OVAs
 - Step-by-step instructions included to build the environment
- Freely available, including documentation, from <https://community.emc.com/docs/DOC-35014>

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

45

The XtremIO Proof of Concept Toolkit is based on IDC testing guidelines, and uses vdbench to generate the workload.

The documentation, and VMs required for the testing [available for download as OVAs] can be found at the location listed on the slide.

Proof-of-Concept Toolkit Page

The screenshot shows the EMC2 website's 'COMMUNITIES' section. The breadcrumb trail is 'All Places > Product Communities > XtremIO > Documents'. The article title is 'How to Test an All-Flash Array', version 11, created by Miroslav Klivansky on May 2, 2014, and last modified by TinaOchoa on Sep 22, 2014. The text describes the challenges of testing All-Flash Arrays and introduces the PoC Toolkit, which includes OVF files, scripts for configuration validation, and a script for automating testing steps. It mentions that performance data is automatically analyzed and graphed. A resource link is provided for a demonstration video showing the whole process from start to finish, with a YouTube URL: <http://youtu.be/mthHweycj9Tw>. Below the link is a text link: 'Download File (600MB HD video in MP4 format)'. The footer of the page includes the copyright notice '© Copyright 2015 EMC Corporation. All rights reserved.', the module name 'Module: XtremIO Performance Analysis', and the page number '46'.

EMC² PRODUCTS & SOLUTIONS SERVICES SUPPORT HOW TO BUY COMMUNITIES PARTNERS EMC+

FIND A COMMUNITY EXPLORE EVENTS CREATE▼ scotthoward▼ Help

All Places > Product Communities > XtremIO > Documents

How to Test an All-Flash Array

Version 11

created by Miroslav Klivansky on May 2, 2014 9:39 PM, last modified by TinaOchoa on Sep 22, 2014 6:54 AM

Over the last few month we've seen an increasing number of requests for help with evaluating All-Flash Arrays. Testing traditional storage can be hard enough, but flash evaluations are even more complicated. For example, just consider requirements to precondition flash prior to testing, coordinate multiple hosts for generating enough load to saturate the array, and selecting the right tools to use with dedupe and compression. Fortunately, the IDC laid out a [comprehensive framework for testing All-Flash Arrays](#). The framework can help guide storage professionals looking to make a clear decision about which flash array is likely to meet their enterprise needs.

But a framework doesn't do the actual tests, and here's where we can help. We've built a Proof-of-Concept (PoC) Toolkit that makes the testing process much easier. The toolkit includes OVF files to quickly deploy virtual machines that coordinate the testing and generate load. There are top level scripts that simplify validating the configuration prior to testing, and a top level script that automates all the steps required to precondition, profile, and evaluate steady-state performance for an All-Flash Array. Best of all, the performance data collected during the testing is automatically analyzed and graphed. That makes it simple to evaluate the array's capabilities and compare it to others.

To get started, check out the following resources:

- Demonstration video showing the whole process from start to finish (YouTube URL: <http://youtu.be/mthHweycj9Tw>)
[Download File \(600MB HD video in MP4 format\)](#)

© Copyright 2015 EMC Corporation. All rights reserved. Module: XtremIO Performance Analysis 46

This is the landing page of the PoC Toolkit. Note that a pre-recorded demo is available.

Proof-of-Concept Toolkit Methodology

- Testing for worst case numbers
 - Write non-deduplicable data (IDC Page 10)
 - Overwrite the array several times to precondition (IDC p11)
 - Age the array by writing non-sequential small block size data across the array (IDC p10)
 - Access the entire array evenly (IDC p14)
 - Use random access (non-sequential) for all testing (IDC p13-14)
 - Test various read/write ratios, and multiple block sizes (IDC p12-13, p14)

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

47

The methodology used in the PoC Toolkit is shown here. Note that the testing consists of a number of phases, which include initial

preparation of the array.

In order to ensure that all regions of the Flash drive are touched, the LUN capacity must approach the capacity limit of the array [90% is recommended], and LUNs must be overwritten at least twice with sequential data. Once the LUNs are filled, random writes are performed against the LUNs to simulate regular system use over a period of time.

The testing involves the use of I/Os of various sizes and various R/W ratios. Once that phase is complete, the array is tested for steady-state performance.

Test results are displayed as graphs, though the raw vdbench data points are also available. Note that this test is vendor-neutral, and may be used on any array. It is useful as a comparison, since the workload will be identical for each run.

PoC Toolkit Environment

VMware normally used due to ease of deployment

- OVAs available for download from community.emc.com
- Can also be configured on physical hardware

Number of VMs dependent on performance of tested array(s)

- 6-8 Workload Generators normally used, plus 1 manager

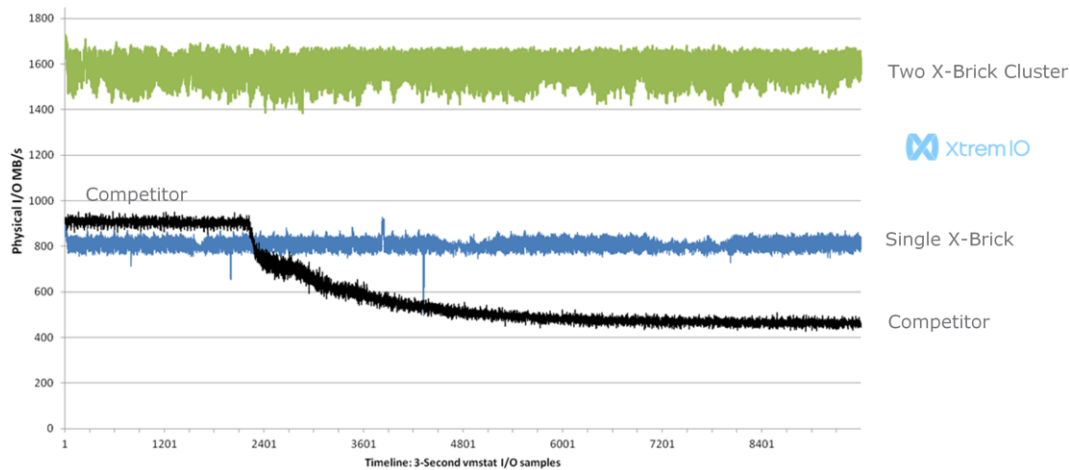
Number of LUNs dependent on array performance

- Minimum of 4 LUNs per workload generator, 8 LUNs recommended
- Assigned to VMs as Raw Device Mappings (RDMs)

The PoC Toolkit environment is simple to implement if VMware is available in the environment. One or more VMs can easily be deployed from the OVA files available with the toolkit. The documentation defines the environment to be used. Note that RDMs are used for testing, since they achieve somewhat better performance than virtual disks.

Why Run Extended Benchmarks?

XtremIO vs. Competitor
Oracle SLOB Workload, 25% SQL Update



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Performance Analysis

EMC²

49

The chart shows a benchmark of a competitor flagship array against XtremIO arrays. The competitor's array initially looked as though it was outperforming the single X-Brick cluster. At a point around 2 hours into the testing [~2400 3-second samples in], the competitor array's performance starts to drop, possibly because of 'write cliff' activity. The XtremIO array shows consistent performance even after 8 hours of continuous testing. A dual X-Brick cluster easily outperforms the competitor array.

Lesson Summary

During this lesson the following topics were covered:

- All-Flash array characteristics
- Preparing for a benchmark
- Selecting benchmark tests

This lesson covered the benchmark process to be used with XtremIO systems.

During this lesson we reviewed the specific benchmarking characteristics for the XtremIO storage array. We covered use of benchmark utilities and tools and their restrictions; the proof-of-concept toolkit, methodology, environment, preparatory and testing steps; and the purposes of running extended benchmarks.

Exercise Introduction: XtremIO Performance

This exercise covers various aspects of XtremIO performance:

- Optimizing host configuration
- Checking queue depths
- Investigating the SAN

This exercise covers various aspects of XtremIO performance.

Exercise Debrief

This exercise covered various aspects of XtremIO performance:

- Optimizing host configuration
- Checking queue depths
- Investigating the SAN

This exercise covered the XtremIO environment as a source of performance issues.

Module Summary

Key points covered in this module:

- Expected XtremIO performance
- Analyzing XtremIO performance data
- Locating performance bottlenecks
- Resolving XtremIO performance issues

This module covered the analysis and troubleshooting of XtremIO performance problems.

During this module we reviewed and set expectations of XtremIO performance; we looked at what it takes to analyze XtremIO performance data, identified performance bottlenecks, and resolved XtremIO performance issues.

This slide intentionally left blank.

Module: XtremIO Troubleshooting and High Availability

Upon completion of this module, you should be able to:

- Explain XtremIO High Availability protection
- Perform basic XtremIO troubleshooting
- Locate XtremIO support resources and updates



This module covers the topics of high availability, troubleshooting, FRUs, and support resources.

Lesson: High Availability

This lesson covers the following topics:

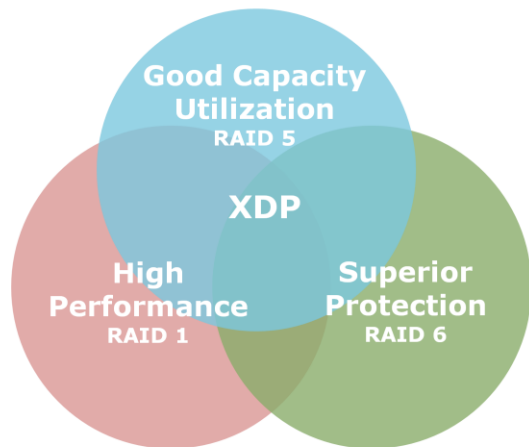
- XtremIO XDP
- Multipathing
- VMware HA XMS



This lesson covers XtremIO Data Protection, multipathing, and protecting a virtual XMS server with VMware High availability.

XtremIO Data Protection (XDP)

- Reinvented data protection
- Combines and extends best attributes of various traditional RAID levels
- Enhances endurance of underlying Flash media
- Designed for partial stripe updates
 - Maintains performance over long term
 - Optimized even when no contiguous free space remains



EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

3

Another key component to the XtremIO data flow is XtremIO's Data Protection (XDP). XtremIO data protection leverages the specific properties of Flash media. The XtremIO Data Protection algorithm simultaneously combines the most desirable traits of traditional RAID algorithms, avoids their pitfalls, and brings entirely new and previously impossible capabilities to the XtremIO storage system.

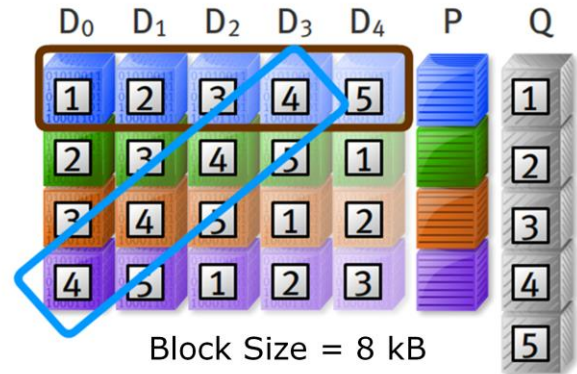
The algorithm places and accesses data in any location on any SSD. Past RAID algorithms had to consider how to keep data contiguous, so as to avoid disk drive head seeks. XDP presumes random access media, such as Flash, is present in the array, and it can lay out data and read it back in highly efficient ways that would heavily penalize a disk-based RAID algorithm, but with no ill effects in XtremIO all-Flash architecture.

XtremIO Data Protection also significantly enhances the endurance of the underlying Flash media compared to any previous RAID algorithm. This is an important consideration for an enterprise Flash array.

XDP has been designed for partial stripe updates. This is critical to maintaining performance over the long-term in a production setting, once the array has been filled and file systems and applications are overwriting old data with new data. Traditional RAID has a heavy performance penalty in this scenario, as it is most efficient when performing full stripe writes into available free space. XtremIO is optimized for the situation every block array eventually reaches, which is no contiguous free space remaining. XDP is vastly superior in performance and Flash endurance in these long-term, steady-state production situations.

XDP (Cont)

- N+2 parity - row and diagonal
- Automatically rebuilds parity upon drive failure
- No hot spares needed
- Tolerate sequence of up to six SSD failures
- Leverages both row and diagonal parity for fast rebuilds



Simplified graphic. Actual XDP stripe is built from 28 rows, 23 data columns, and 2 parity columns.

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

4

Each XtremIO X-Brick contains one DAE with 25 SSDs. The SSDs are configured with N+2 row and diagonal parity, as shown here. If one of the SSDs in an X-Brick fails, XDP automatically rebuilds the data to unused space on the remaining SSDs, using distributed parity and data on the remaining SSDs, while dynamically reconfiguring incoming new writes into a 22+2 stripe size to maintain N+2 double failure protection for all new data written to the array. XDP always maintains a reserved capacity on each drive to perform this drive rebuild.

This reserved capacity is factored out of XtremIO's usable capacity per X-Brick. The first drive failure does not lessen the usable space per X-Brick. Because of this methodology, not only does XtremIO not require hot spares, but every available SSD is always active and adding to the array's performance.

Once the rebuild completes, and the failed drive is replaced, incoming writes will once again be written to the standard 23+2 stripe. This adaptability allows XDP to tolerate a sequence of up to six SSD failures, without having to rush to the data center to replace failed SSDs. XtremIO can continue to rebuild failed drives again and again, while serving user I/Os, as long as there is still available capacity in the array.

This is a very unique feature that allows administrators to defer drive replacements until a convenient time. This is especially important in remote, secure, and 'lights out' data centers.

Another important aspect of XDP is that both row and diagonal parity are used for fast drive rebuilds.

Note that although XtremIO uses N+2 parity and maintains data integrity during two simultaneous drive failures, this situation will cause data unavailability.

Refer to the XtremIO Data Protection and XtremIO Flash Optimized Data Protection white papers for more detailed information about XDP.

Multi-pathing

Host Platform	Path Management	Cluster
Oracle Linux	PowerPath, Native MPIO, VERITAS VxDMP	Y
Red Hat Linux	PowerPath, Native MPIO, VERITAS VxDMP	Y
Microsoft Windows	PowerPath, MPIO	Y
Solaris	PowerPath, MPxIO, VERITAS VxDMP	Y
VMware	PowerPath/VE, NMP, VERITAS VxDMP	Y



Refer to the Support Matrix for up-to-date information



Multi-pathing software significantly improves performance and redundancy. XtremIO supports a wide array of host-based multi-pathing solutions, including EMC PowerPath. Refer to the latest support matrix for the most up-to-date information.

Protecting XMS with VMware HA

- When XMS fails:
 - I/O to the XtremIO storage array will not be impacted
 - No monitoring and/or configuration changes
- The virtual XMS should be configured to withstand the following failures:
 - Power loss or hardware malfunction to the physical host running the virtual XMS
 - Network connectivity loss to the physical host running the virtual XMS (loss of connection between the XMS and the X-Brick)
 - Software failure or inactivity of the VM's operating system



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

6

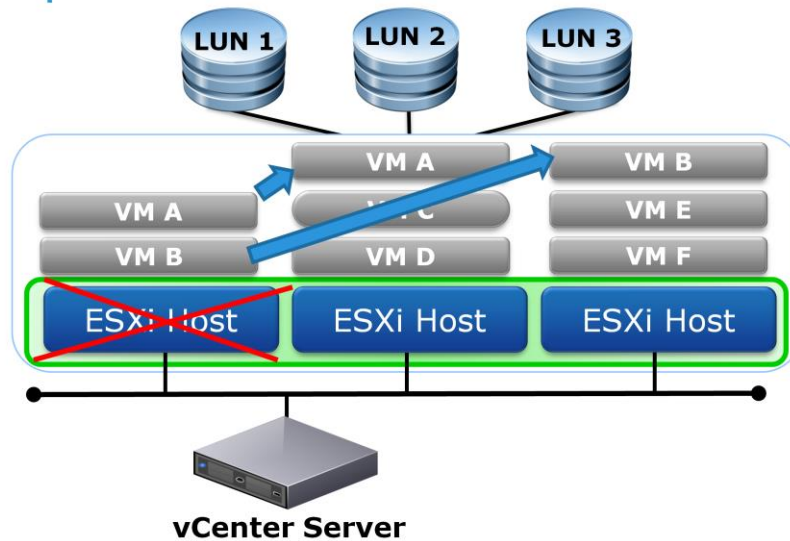
The XMS, as a single component in the XtremIO cluster, is vulnerable to host and network failures. When such failures occur, I/O to the XtremIO storage array will not be impacted. However, no monitoring and/or configuration changes (creating/deleting volumes etc.) will be possible.

When using a virtual XMS topology, it is possible to take advantage of vSphere HA features to easily overcome such failures.

The Virtual XMS should be configured to withstand the following failures:

- Failure due to a power loss or hardware malfunction to the physical host running the Virtual XMS
- Failure due to a network connectivity loss to the physical host running the Virtual XMS (loss of connection between the XMS and the X-Brick)
- Failure due to a software failure or an inactivity of the VM's operating system

vSphere HA



When a host fails, vSphere HA restarts the affected virtual machines on other hosts.

 = vSphere HA Cluster

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

7

This slide is a visual representation of VMware HA in action.

Lesson Summary

During this lesson the following topics were covered:

- XtremIO XDP
- Multipathing
- VMware HA XMS



This lesson covered XtremIO Data Protection, multipathing, and protecting a virtual XMS server with VMware High Availability. We saw how XtremIO Data Protection leverages the specific properties of Flash media; how multi-pathing software significantly improves performance and redundancy; and lastly, we looked at VMWare HA XMS using a virtual XMS topology where it's possible to take advantage of vSphere HA features to easily overcome host and network failures.

Lesson: Troubleshooting

This lesson covers the following topics:

- Power up and power down procedures
- Troubleshooting



This lesson covers troubleshooting, using the GUI, checking alerts and status.

Power On/Off

- Power up procedures
 - CLI commands
 - Multi-step process
 - Must be completed in a specific order
- Shut down procedures
 - Planned
 - Emergency



The XtremIO cluster has a specific power on and power off procedure that can vary depending on the activity, or event, you are presented with.

Cluster Management – Start/Stop

Command	Description	Example
> start-cluster	Starts a stopped cluster and enables it to respond to host IOs and process data	xmcli (admin)> start-cluster The process may take a few minutes. Please do not interrupt.
> shutdown	Shuts down a Storage Controller (identical to the Linux shutdown command), a defined set of Storage Controllers, or all the Storage Controllers of a specified cluster	xmcli (admin)> shutdown cluster-id=1 Executing shutdown... xmcli (admin)> shutdown sc-id-list=[1,2] Executing shutdown...
> stop-cluster	Stops an active cluster (active X-Envs only) and disables data processing in an orderly manner	xmcli (admin)> stop-cluster Are you sure you want to stop Cluster? (Yes/No):

Refer to the XtremIO Storage Array User Guide for specific details.

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

11

Listed here are some basic cluster management commands. Refer to the XtremIO User Guide for more detailed information when powering off or starting a cluster.

The start-cluster command starts a stopped cluster and enables it to respond to host IOs and process data.

The shutdown command shuts down a SC (identical to the Linux shutdown command), a defined set of SCs, or all the SCs of a specified cluster.

The stop-cluster command stops an active cluster (active X-Envs only) and disables data processing in an orderly manner.

Storage Controller – Start/Stop

Command	Description	Example
> power-off	Shuts down a Storage Controller	xmcli (admin)> power-off sc-id=2 Executing Power-off ...
> power-on	Powers up a Storage Controller Powers up multiple Storage Controllers Powers up a cluster	xmcli (admin)> power-on sc-id=2 Executing Power-on ... Power-on Storage Controllers Status: Storage Controller X1-SC2: ok xmcli (admin)> power-on cluster-id=1 Executing Power-on ... Power-on Storage Controllers Status: Storage Controller X1-SC1: ok Storage Controller X1-SC2: ok
> power-cycle	Powers down and up a Storage Controller	xmcli (admin)> power-cycle sc-id=2 Executing Power-cycle ...

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

12

Listed here are some basic SC management commands. Refer to the XtremIO User Guide for more detailed information when powering off, starting or power-cycling an SC.

The power-off command shuts down an SC. Powering off an SC disconnects all connected hosts from the paths to this SC, while powering off an entire cluster means that the entire cluster does not respond to host IOs requests.

The power-on command can power up one SC, multiple SCs, or a cluster.

The power-cycle command powers down and powers up an SC. Power-cycling an SC disconnects all connected hosts from the paths to this SC, while power-cycling the entire cluster means that the entire cluster will not respond to host IOs requests.

Using the XMS GUI

Hardware

Major
 IB-SW1
 Health State :Uninitialized
 Lifecycle State : Disconnected
 Serial Number :
 Part Number : ****
 Alerts :
 InfiniBand Switch is disconnected.
 InfiniBand Switch firmware version is incompatible. InfiniBand Switch cannot be used.
 (Total 2 alerts)

Levels of Severity Color Codes

Critical **Major** **Minor** **Information**

Minor
 X1
 BBU state UninitializedBBU X1-BBU1
 Alerts :
 XMS has detected an X-Brick that is not configured to be part of a system.
 (Total 1 alerts)

EMC²

© Copyright 2015 EMC Corporation. All rights reserved. Module: XtremIO Troubleshooting and High Availability 13

Shown here is the Hardware pane in the Dashboard Workspace that displays a graphic illustration of the system hardware components in a two-X-Brick cluster. If a hardware component suffers from an alert, its color changes according to the alert's severity. Hovering the mouse over the hardware component displays a ToolTip with information on the hardware status, as shown in this Major alert at the top of the slide. The bottom of the slide shows a single X-Brick cluster that is colored yellow and the Minor alert that caused the display to change color. The levels of alert severity include the following:

- Blue - Information alert
- Yellow - Minor alert
- Orange - Major alert
- Red - Critical alert

Clicking the hardware graphic displays the hardware workspace.

Check Alerts

- The `show-alerts` command displays a list of pre-defined alerts with details
- Examine output and run more commands in area of concern

```
xmcli (admin)> show-alerts
```

Index	Description	Severity	Raise-Time
15	Battery Backup Unit is overloaded. Invalid value.	major	Sat Sep 28 21:28:4
14	Battery Backup Unit output power is off. Invalid error.	major	Sat Sep 28 21:28:4

Entity	Name	Index	Alert-Type	State
UPS	X1-BBU1	1	ups_overid_failure_unexpected_value	outstanding
UPS	X1-BBU1	1	ups_output_off_failure_unexpected_value	outstanding

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

14

The `show-alerts` command displays a list of pre-defined alerts and their details on the cluster. When starting to investigate an issue, it is useful to know what alerts exist on the cluster.

Each alert has a unique ID and a description. The severity level, the time it was registered, the entity (component), and its unique name within the cluster are listed. Furthermore, the command output displays the alert type and state, and whether outstanding or acknowledged.

Output Parameter - Description

Index - Alert ID

Description - Alert description

Severity - Alert severity level (critical, major, minor, informational)

Raise-Time - Date and time at which the alert was raised

Entity - Cluster component to which the alert refers

Name - Component name

Index - Component index number

Alert-Type - Alert type

State - Alert status (outstanding or acknowledged)

Check Cluster Status

- The `show-clusters` command displays general cluster information

```
xmcli {admin}> show-clusters
```

Cluster-Name	Index	State	Conn-State	Num-of-Vols	Vol-Size	Address-Space
system4	1	unknown	disconnected	32	3.125T	2.828T



SSD-Space-In-Use	SSD-Space	Total-Writes	Total-Reads	Stop-Reason
1.326T	7.475T	38.561T	245.864T	none

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

15

The `show-clusters` command displays general cluster information and the current state. The output in the example shows that the state of the cluster is unknown and that the XMS currently is disconnected from the cluster. It needs to be investigated why the XMS cannot reach the storage array. It could be network related, bad Ethernet cables, or all the SCs in the XtremIO storage array could be down.

Check Storage Controller

- The `show-storage-controllers` command displays the cluster's Storage Controller's information and status

```
emcli (admin)> show-storage-controllers
```

Storage-Controller-Name	Index	Mgr-Addr	IB-Addr-1	IB-Addr-2	IPMI-Addr
XL-SC1	1	10.126.101.141	169.254.0.1	169.254.0.2	10.126.101.141
XL-SC2	2	10.126.101.142	169.254.0.17	169.254.0.18	10.126.101.142

Brick-Name	Index	Cluster-Name	Index	State	Enabled-State	Unorderly-Stop-Reason	Conn-State	IPMI-State
xl	1	system4	1	healthy	enabled	none	connected	connected
xl	1	system4	1	disconnected	enabled	lost_connectivity_with_node	disconnected	disconnected

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

16

The `show-storage-controllers` command displays the information of each SC in the cluster, such as IPs, names, connectivity state, and status. It is especially useful to get all the IP addresses for all the SCs plus their state (health state) and enabled-state (configuration status).

The negative output shows that one SC is disconnected from the XMS. Its connectivity was lost. Network connectivity (MGMT/IPMI Port) should be checked.

Check BBUs

- The `show-bbus` command displays battery backup unit information

```
xmcli (admin)> show-bbus
```

Name	Brick-Name	Index	Cluster-Name	Index	Serial-Number	Power-Feed	State	Connectivity-State
X1-BBU1		1			AV8N38033	feed_b	healthy	connected
X2-BBU1		2			AV8N38036	feed_a	disconnected	disconnected



In the negative output, we see the State and the Connectivity-State as being disconnected. Most likely it's that the BBU connectivity is disrupted. This could be a problem with the BBU-SC sense cable.

Check Initiator Connectivity

- The `show-initiators-connectivity` command displays initiator-port connectivity and the number of connected targets

```
xmcli (admin)> show-initiators-connectivity
Existing Initiator List:
Name                Index  Port-Type  Port-Address          Num-Of-Conn-Targets
esx-fc1             1      fc         10:00:00:90:fa:1c:e9:fa  2
esx-fc2             2      fc         10:00:00:90:fa:1c:e9:fb  2
esx2-fc1            3      fc         10:00:00:90:fa:1c:e9:f0  0
esx2-fc2            4      fc         10:00:00:90:fa:1c:e9:f1  0
```

```
xmcli (admin)> show-initiators-connectivity target-details
Existing Initiator List:
Name                Index  Port-Type  Port-Address          Num-Of-Conn-Targets  Target-Name
esx-fc1             1      fc         10:00:00:90:fa:1c:e9:fa  2                    X1-SC1-fc1
esx-fc2             2      fc         10:00:00:90:fa:1c:e9:fb  2                    X1-SC1-fc2
esx2-fc1            3      fc         10:00:00:90:fa:1c:e9:f0  0                    X1-SC2-fc1
esx2-fc2            4      fc         10:00:00:90:fa:1c:e9:f1  0                    X1-SC2-fc2
```

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

18

The `show-initiators-connectivity` command displays the names of the registered initiators, their index, port type (FC or iSCSI), and port address (WWPN or IQN). Furthermore, it lists the number of connected targets for the specific initiator. For example, in a Fibre Channel SAN, this number should be identical to the number of targets that this initiator has been zoned to.

If you want to go further in the analysis to see which ports on the cluster of a particular initiator is connected to, use the `target-details` parameter. The target (cluster port) names are listed. The displayed mapping between initiators and targets can help with zoning and performance issues.

It is useful when trying to understand if a defined initiator is already connected or not, that is, if an initiator is successfully zoned to the specified targets. As mentioned before, initiators can be pre-defined (without being connected). The initiators `esx2-fc1` and `esx2-fc2` are registered, but are not yet connected to any targets. This can also become interesting, if an HBA has been swapped, or performance becomes an issue. Is it zoning or adding the new initiator to the initiator group? Is the initiator connected to the expected targets?

Check FC Errors

- The `show-targets-fc-error-counters` command displays Fibre Channel error counters per target.

```
xmcli (admin)> show-targets-fc-error-counters
```

Name	Index	Dumped-Frames	Sync-Loss	Signal-Loss	Invalid-Crc	Link-Failure	Prim-Seq-Err
X1-SC1-fc1	1	0	0	0	0	0	0
X1-SC1-fc2	2	0	0	0	0	0	0
X1-SC2-fc 1	11	0	0	0	0	0	0
X1-SC2-fc2	12	0	0	0	0	0	0

```
xmcli (admin)>
```



The `show-targets-fc-error-counters` command displays Fibre Channel error counters per target. All the cluster port names and their index are listed with counters for dumped frames, synchronization losses, signal losses, invalid CRC errors, link failures, and prime sequential protocol errors. Knowing about FC errors is useful when troubleshooting Fibre Channel issues and/or performance degradation.

Check InfiniBand Counters

- The `show-storage-controllers-infiniband-counters` command displays InfiniBand errors and link down states.

```

xcli (admin)> show-storage-controllers-infiniband-counters
Storage-Controller-Name  Index  Port-Index  Symb-Errs  Symb-Errs-pm  Recovers  Recovers-pm  Lnk-Downed  Lnk-Downed-pm
Xl1-SC1                 1      1           0          0             0          0            0            0
Xl1-SC1                 1      2           0          0             0          0            0            0
Xl1-SC2                 2      1           0          0             0          0            0            0
Xl1-SC2                 2      2           0          0             0          0            0            0

```



```

Symb-Errs  Rcv-Errs-pm  Rmt-Phys-Errs  Rmt-Phys-Errs-pm  Integ-Errs  Integ-Errs-pm  Rate-Gbps
0          0             0             0             0          0             40
0          0             0             0             0          0             40
0          0             0             0             0          0             40
0          0             0             0             0          0             40

```

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

20

The `show-storage-controllers-infiniband-counters` command displays InfiniBand errors and link down states. The output lists each InfiniBand port of each SC. The symbol errors counter and symbol errors counter per minute are minor link errors and minor link errors per minute respectively. The other counters shown here are all standard InfiniBand errors. Each are displayed in total and per minute counters. Refer to more specific IB documentation, if necessary.

The InfiniBand network is a component of the XtremIO cluster which is not user managed. Administrators do not need to have specialized skills in InfiniBand technology.

This command shows errors in the InfiniBand network, which carries the inter-SC communication. Errors on the InfiniBand network can have a huge impact on cluster performance. Hardware failures could be SC IB ports, IB cables or IB switches (in two or four X-Brick clusters).

Lesson Summary

During this lesson the following topics were covered:

- Power up and power down procedures
- Troubleshooting



This lesson covered troubleshooting, using the GUI, checking alerts and status. You saw how the XtremIO cluster has a specific power on and power off procedure that can vary depending on the activity, or event, you are presented with. We also covered the various CLI commands used for troubleshooting.

This slide intentionally left blank.



Lesson: Support Resources and Updates

This lesson covers the following topics:

- Support resources



This lesson covers some of the support resources available to you.

Product Support Page

- support.emc.com
- System documentation
 - Site preparation guide
 - User guide
- Software downloads

The screenshot displays the EMC Online Support Site interface. At the top left, a blue header bar contains the text 'Welcome, [XtremIO Troubleshooting](#)', 'Logout', 'Manage my profile', and a language dropdown set to 'English'. Below this is a 'My Tasks' section with an 'Edit' button and links for 'Create SR', 'Forums', and 'Live Chat'. To the right is a search bar labeled 'Search EMC Support' with a search icon and an 'Advanced »' link. The main content area is titled 'Welcome to the EMC Online Support Site' and includes a welcome message. Below this are several featured sections: 'My Support' (customizable support experience), 'Support by Product' (one-stop location for product information), 'Downloads' (access to product downloads), 'Community' (product-specific discussions), and 'Service Center' (home base for managing support requests). The 'Service Center' section includes links for 'View and manage service requests' and 'Get and Manage Licenses'. At the bottom, there is a Twitter follow prompt for @EMCSupport and the EMC logo. The footer contains copyright information, the module name 'Module: XtremIO Troubleshooting and High Availability', and the page number '24'.

The XtremIO Product Support page contains all the information needed for installation and management of an XtremIO array. Note that this site is intended for support personnel, and may not be accessible by others.

The User Guide includes instructions on how to power on and power off the installed system, and contains complete information on array management using the GUI and CLI.

Available software includes the XtremIO Management Server (XMS) Open Virtual Appliance (OVA) image, for use with the virtual XMS option.

EMC Simple Support Matrix

- Provides basic platform and component support for EMC XtremIO
- Complete product support statements in a simplified, all-in-one format
- All supported configurations for a product or solution that are listed in the ESSM format do not require any additional E-Lab Navigator validation.

EMC²

EMC²
ELAB
PROVEN

EMC Simple Support Matrix
EMC XtremIO

EMC²

© Copyright 2015 EMC Corporation. All rights reserved.

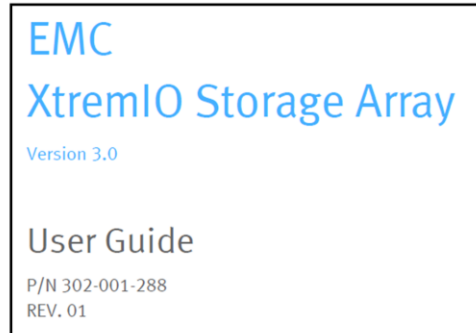
Module: XtremIO Troubleshooting and High Availability

25

This document provides basic platform and component support for EMC XtremIO. The EMC Simple Support Matrix [ESSM] represents complete product support statements in a simplified, all-in-one format. All supported configurations for a product or solution that are listed in the ESSM format do not require any additional E-Lab Navigator validation.

User Guide

- Consult the newest XtremIO Storage Array User Guide
 - Use the graphical user interface (GUI)
 - Monitor the cluster
 - Manage alerts and events
 - Use the command line interface (CLI)
 - Other topics
- Download from EMC support page support.emc.com



© Copyright 2015 EMC Corporation. All rights reserved.

Module: XtremIO Troubleshooting and High Availability

26

Refer to the latest XtremIO Storage Array User Guide when working with the XtremIO storage array. It describes how to use the GUI, monitor the cluster, manage alerts and events, use the CLI, as well as other topics.

Current Release Notes Review

- XtremIO is a new product and currently maintains a high frequency update cycle
- This course might be older than the current GA code version
- Review the most recent Release Notes for product changes



Changes made to the XtremIO array will be listed in the most recent Release Notes. Please refer to them for the most accurate information regarding XtremIO features and limits.

Lesson Summary

During this lesson the following topics were covered:

- Support resources



This lesson covered some of the support resources available to you, such as support.emc.com; the EMC Simple Support Matrix document; the site preparation guide; the user guide, and current release notes.

Module Summary

Key points covered in this module:

- High Availability
- Power Procedures
- Troubleshooting
- Support resources and updates



This module covered the topics of high availability, troubleshooting, FRUs, and support resources.

Course Summary

Key points covered in this course:

- XtremIO hardware and software components
- Using the XtremIO GUI and CLI interface for management
- Provisioning storage from an XtremIO array
- Monitoring an XtremIO storage array
- Maintenance procedures and troubleshooting methods



This course provided an in-depth understanding of how to install, manage, monitor, and maintain a functional XtremIO storage array. It covered operational aspects such as hardware and software review, cabling requirements, cluster definitions, and the XtremIO Management Server (XMS). The emphasis was on provisioning storage from XtremIO arrays to various host types.

This slide intentionally left blank.





This concludes the Training. Thank you for your participation.