Pillar Axiom

System Architecture Overview

ORACLE

PILLAR AXIOM

Part Number: 4420-00029-1200
Pillar Axiom release 4.2
2011 October

Table of Contents

List of Figures

List of Tables

Preface

Audience

This documentation is intended for individuals who plan, purchase, and implement enterprise storage systems.

To use this document successfully, you should have:

- Basic knowledge of the hardware and software used in enterprise storage systems.
- Familiarity with storage planning and implementation concepts.

Before You Read This Document

Being familiar with certain other Pillar Axiom technical documentation helps you succeed in the use of this document.

In particular, we recommend that you refer to the *Pillar Axiom Administrator's Guide* for detailed information on creating and managing storage resources.

Other useful documents include:

- *Pillar Axiom Statistics Tools User Guide*
- *Pillar Axiom SnapDelta FS Reference Guide*
- *Pillar Axiom Pillar Axiom CIFS and NFS Multi-Protocol Planning Guide*
- *Pillar Axiom iSCSI Integration Guide for Windows Platforms*
- *Pillar Axiom NDMP Integration Guide for NAS Systems*
- *Pillar Axiom Service Guide*
- *Pillar Axiom SMIProvider Reference*
- *Pillar Axiom SSF Cabling Reference*
- *Pillar Axiom Support and Interoperability Guide*

- *Pillar Axiom Glossary*

How This Document Is Organized

This document is intended as an overview of the components and features of Oracle's Pillar Axiom 600 storage system.

It can be used as a non-technical introduction to Pillar Axiom storage system concepts, or as a guide for planning and implementing your storage system.

This document is divided into six chapters:

- Chapter 1 provides an overview of the Pillar Axiom system.
- Chapter 2 provides an overview of the Pillar Axiom hardware components.
- Chapter 3 provides an overview of the Pillar Axiom software components.
- Chapter 4 describes the network attached storage (NAS) features of the Pillar Axiom system.
- Chapter 5 describes the storage area network (SAN) features of the Pillar Axiom system.
- Chapter 6 provides an overview of the replication features of the Pillar Axiom system.

Access Documentation

Technical documentation (including installation, service, cabling, integration, and administration guides) for Oracle's Pillar Axiom 600 storage system is available from several sources.

Pillar Axiom GUI	After logging in to the Pillar Axiom Storage Services Manager on the Pilot, navigate to **Support > Technical Documentation** and click on the document of interest.
Web sites	Technical documents (http://www.pillardata.com/techdocs) Customer support portal (https://support.pillardata.com/login.do)

After logging in to the web site, click on **Documents** in the left navigation pane, and then click the appropriate category in the expanded list. Click on the document of interest.

Product CD-ROM

Insert the Technical Documentation CD-ROM that came with your Pillar Axiom storage system into the CD player in a computer. Open the DocMenu PDF and click on the document of interest.

Tip: To search all technical documents on the CD-ROM, click the **Search all PDFs** icon in the top right corner. In the Search dialog, enter the word or phrase for which you would like to search.

Pillar Contacts

Table 1 Contacts at Pillar Data Systems

For help with...	Contact...
Error messages, usage questions, and other support issues	US and Canada: 877-4PILLAR (1-877-474-5527) Europe: +800 PILLAR FS (+800 74 55 27 37) Asia Pacific: +1-408-518-4515 South Africa: +0 800 980 400 Have your system serial number ready. support@pillardata.com Customer support portal (https://support.pillardata.com/login.do)
Training (custom or packaged)	Training and Education (http://www.pillardata.com/support-education/training/)
Sales and general contact information	Company contacts (http://www.pillardata.com/company/contact)
Documentation improvements and resources	docs@pillardata.com Technical documents (http://www.pillardata.com/techdocs) (Log in with your username and password, and select Documents.)

CHAPTER 1

Welcome to the Pillar Axiom System Architecture

Product Overview

The Pillar Axiom system is a full-featured network storage system.

The Pillar Axiom system is a complete and integrated storage system for Network Attached Storage (NAS) and Storage Area Networks (SAN) that supports:

- High performance.
- High availability.
- Reliability.
- Scalability.
- Archival capability.

Feature Overview

Features of the Pillar Axiom storage solution include the following:

Policy-based Quality of Service (QoS)	The Pillar Axiom system employs intelligent storage device management to provide QoS capabilities that yield a very high utilization of the storage resources in the entire system. This QoS functionality enables the system administrator to prioritize data according to the importance of the applications and the required performance for each data store.
Storage Classes	The introduction of Storage Classes makes it possible to configure QoS for each drive type that the Pillar Axiom system supports.
Optimized capacity management	With its built-in capacity planning tool, the Pillar Axiom system helps achieve the most efficient use of available storage capacity.
Advanced backup and recovery methods	A variety of data protection tools are available in the Pillar Axiom system and from Pillar Axiom partners to provide backup and restore capabilities.
Replication	Replication of both NAS filesystems and SAN volumes facilitates automated backup and restore of mission critical data to ensure disaster recovery and business continuance.
Performance management	Numerous application and performance profiles are available to help storage administrators maximize performance.
Scalability	Pillar Axiom systems are designed to scale as business needs expand. Built-in hardware and software flexibility make it easy to add capacity, cache, or CPU power.
NAS and SAN support	NAS and SAN systems can coexist and share the same storage pool.
Flexible storage reallocation	Storage can be reallocated without disrupting normal operations.
Alerts and event management	Easy-to-use software facilitates monitoring and troubleshooting the Pillar Axiom system.

System health information	An intuitive graphical user interface (GUI) provides complete information about system performance at a glance.
Statistical reports	A sophisticated reporting interface facilitates gathering and analyzing all types of system data.
Interoperability	Pillar Axiom systems can be used with a wide variety of popular hardware and software.
Storage volume expansion	Filesystems and LUNs can be expanded to meet growing business needs.

CHAPTER 2

Pillar Axiom Hardware Overview

Pillar Axiom System Components

The Pillar Axiom system provides SAN and NAS connectivity to a common pool of storage.

The system is modular. The three major system components are:

- Pilot management controller.
- Slammer storage controllers.
- Brick storage enclosures.

Slammers and Bricks communicate using a highly redundant Storage System Fabric (SSF).

Slammers and Bricks have:

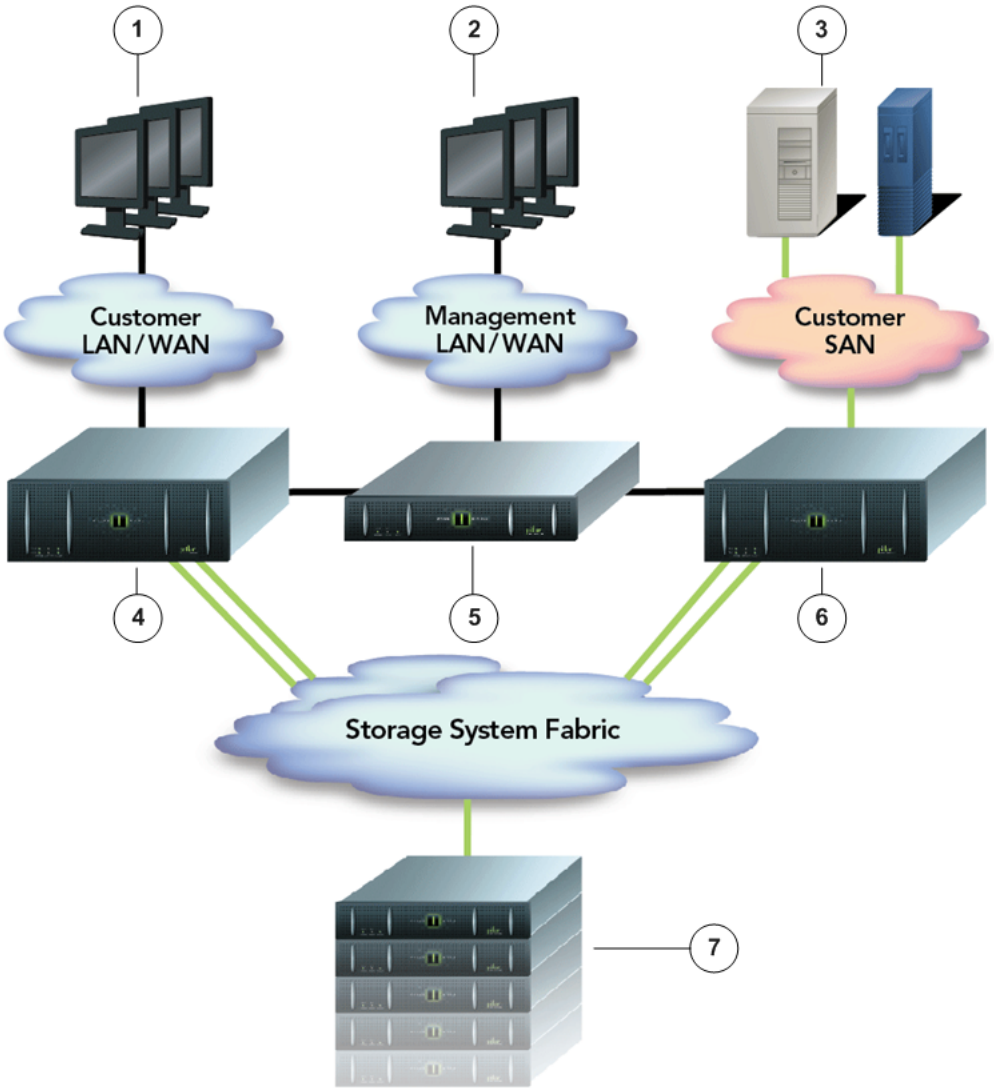
- Redundant power supplies and fans.
- Front and back LEDs that provide system and component identification and status.
- Built-in RAS: reliability, availability, and serviceability.
- Field replaceable units (FRUs).

The system components fit into a Pillar-supplied rack or a standard D-class 4-post 19-inch rack. Larger configurations require multiple adjacent racks.

The Pillar Axiom system continually monitors all hardware components for proper operation and fault status.

The following figure shows the interactions between the Pillar Axiom system components.

Figure 1 Interactions among Pillar Axiom components



Legend

1 NAS clients (NFS or CIFS)	5 Pilot policy controller
2 Management clients	6 SAN Slammer storage controllers
3 SAN clients (FC or iSCSI)	7 Brick storage enclosures
4 NAS Slammer storage controllers	

Storage System Fabric (SSF)

The Storage System Fabric (SSF) consists of a multiply redundant switched Fabric that carries all data traffic among Slammer storage controllers and Brick storage enclosures.

The SSF is defined as:

The protected Fibre Channel fabric internal to Pillar Axiom storage systems that interconnects Bricks and Slammers. The SSF enables communication within the Pillar Axiom system so that all Slammers can connect to any of the Bricks.

On Pillar Axiom 600 systems, each control unit (CU) within a Slammer has a Fibre Channel switch, but Pillar Axiom 300 system CUs each have a Fibre Channel hub. Each Slammer CU has a private interface module (PIM) that contains these Fibre Channel ports, and these ports connect the CU to the SSF. The SSF utilizes either a 2 Gb/s copper or a 4 Gb/s optical Fibre Channel interface to provide sufficient bandwidth.

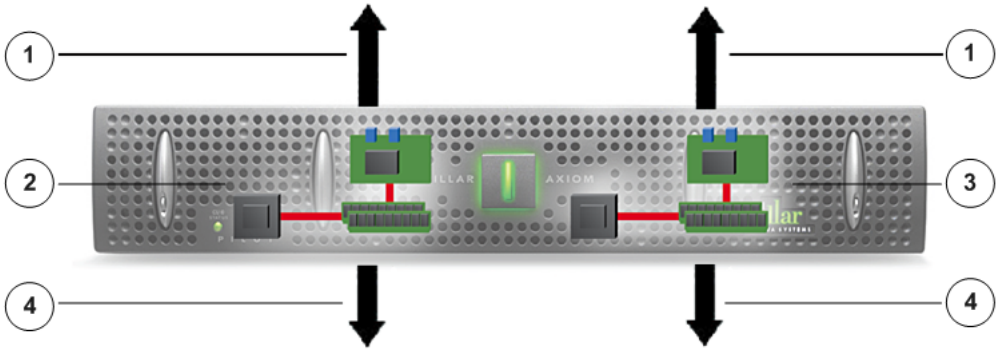
Every port of the switch can simultaneously service data I/O operations. The Fabric Manager controls all data transfers throughout the SSF.

Pilot Management Controllers

The Pilot is an out-of-band management system that directs and manages all system activity. It has no connection to the data path, no access to user data, and no impact on I/O activity. The Pilot communicates with the other components of the Pillar Axiom system over an internal Ethernet pathway that connects the Pilot to the Slammers, and connects the two Pilot control units (CUs). Each Pillar Axiom system has one Pilot, which includes:

- A graphical user interface (GUI) for user configuration and management of the Pillar Axiom system.
- A command line interface (CLI) used in scripts for user configuration and management of the Pillar Axiom system.
- Two independent CUs that operate in active/passive mode.
- A connection to the external network for Pillar Axiom system management, transfer of logs, and generation of alerts.

Figure 2 Pilot components



Legend	
	1 To external management network
	2 Control unit CU0
	3 Control unit CU1
	4 To Slammers

Pilot Functional Description

The Pilot configures and manages storage through Quality of Service (QoS) policies that allocate hardware assets based on need and service level.

The Pilot also:

- Manages backups.
- Manages restoration of backed up data sets.
- Supplies system monitoring, notification, and reporting services.

The Pilot directs the operation of the Slammer based on user settings from the GUI or CLI. Each independent Pilot control unit (CU) has a unique management IP address. A private management interface (PMI) is shared between the Pilot CUs and the Slammer CUs so they can access information from the Slammers.

A heartbeat running between all of the Pilot and Slammer CUs monitors the status of these CUs. In case one CU fails, the other CU becomes active and takes over management. The heartbeat utilizes both the private Ethernet and a backup serial cable.

The Pilot also monitors any replacement Pilot CUs that are added to the system and compares them to the active CU. It revises the software on the new Slammer, Pilot, or Brick CU so that it matches that of the current, active CU.

The Pilot includes the following interfaces:

- Private Management Interface connections for managing the Slammers as well as the active and passive roles of the Pilots.
- An SMI-S Provider interface for monitoring and managing storage resources.
- A VSS Provider interface for managing storage in a Microsoft Windows environment.
- An Oracle Enterprise Manager interface for managing storage in an Oracle environment.
- A Pillar Axiom Path Manager (APM) interface for managing and monitoring storage in combination with APM in connected hosts.

Pilot Software Components

The Pilot management controller includes the following software architecture layers:

Persistent Store — Persistent Store is a subcomponent of Platform Services. Each Pilot has a locally attached drive that mirrors persistent information to both Pilot control units (CUs).

Data Protection Services Data Mover Remote API — The Data Mover provides the interface to the data-mover engine located on a NAS Slammer. The Pilot manages data movement, but the actual movement occurs on a separate NAS Slammer.

NDMP Agent — This agent supports the Network Data Management Protocol (NDMP) that integrates third-party data management applications with system backup functions. It provides the control path to manage backup operations with local and network-attached tape libraries.

MCC Core — Management Configuration and Control (MCC) Core performs the following functions:

- Configures system storage and data paths.
- Manages system startup.
- Performs Slammer CU failover and failback.
- Manages drive replacement, including copyback and assignment of spare.
- Manages RAID CU failures.
- Manages scheduled events.
- Manages Guided Maintenance.
- Monitors the system.
- Manages automatic and manual log collection and transfer.
- Manages and applies Quality of Service (QoS) policies.
- Manages internal log rotation.

- Provides the external interfaces (the MCC UI) to manage the system.

MCC UI

The MCC user interface (UI) component supports the set of user interfaces into the system for management and control. The supported UIs are a graphical, web-based UI, an installable command line application, and an externally accessible API.

SMIProvider

The SMIProvider (Storage Management Initiative provider) provides a Storage Networking Industry Association (SNIA) compliant interface. Clients such as Oracle's Automatic Storage Management (ASM) feature, Microsoft's Volume Copy Shadow Service (VSS), and SMIProvider can be used to manage storage on the Pillar Axiom system.

Note: Unlike SNMP, the SMIProvider makes it possible to modify the Pillar Axiom system configuration.

SNMP Agent

The Simple Network Management Protocol (SNMP) agent provides a standard interface through the external management connection, which supports the SNMP protocol. SNMP GETS (queries) and Traps (event notifications) are supported.

Note: The SNMP agent does not support SET of any system information or configuration.

Pilot Platform Services

The code in Platform Services interacts with hardware and basic OS system functions. It allows other code to function, including:

- Interprocess communication (IPC) and remote procedure call (RPC) mechanisms
- Memory allocation
- Hardware initialization and control
- Network protocol drivers TCP/IP stack
- Topology discovery for storage control units
- Persistent store
- Active or passive state management

Slammer Storage Controllers

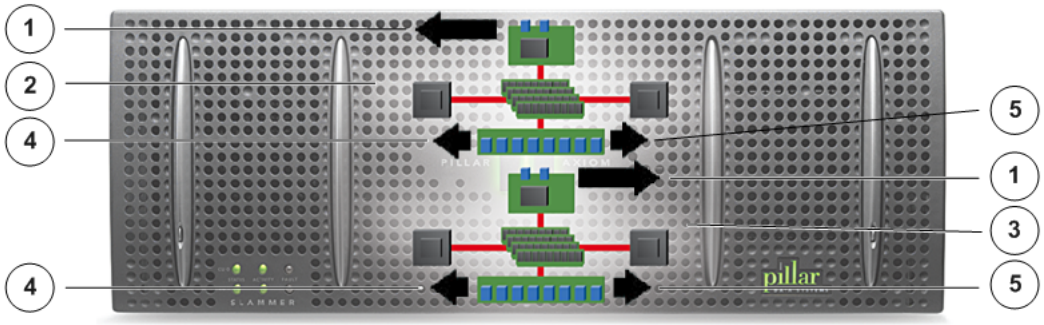
The Slammer provides an external interface to the host storage network. The Slammer processes every I/O request. A Pillar Axiom system can include both NAS and SAN Slammers.

The current architecture allows up to four Slammers for each system:

- Pillar Axiom 300 systems support one Slammer
- Pillar Axiom 600 systems support one, two, three, or four Slammers

The following figure shows the different parts of a Slammer.

Figure 3 Slammer components



Legend	
	1 To host storage network (GbE, FC or iSCSI)
	2 Control unit CU0
	3 Control unit CU1
	4 To Bricks and other Slammer control units (storage system fabric)
	5 To Pilot (private management interface)

Slammer Functional Description

A Slammer contains two control units (CUs) functioning as an active-active asymmetric access pair. The primary components of each Slammer CU include:

- One private interface module (PIM).

- One network interface module (NIM) of one of the types described in the table below.
- One, two, or four processors.
- One or two power supplies.

The characteristics of these primary Slammer CU components, including a breakdown by Pillar Axiom model, are described in the following tables:

Table 2 PIM components in each Slammer CU

Components	Pillar Axiom 300	Pillar Axiom 600
Fibre Channel (FC) controller ports	Two (one internally connected) plus five port bypass hub ports	Four, with one of the following switches: Version 1: 13-port FC switch with a 2 Gb/s copper back end Version 2: 16-port FC switch with one of: • 2 Gb/s copper back end • 4 Gb/s optical back end
Ports for the private management interface (PMI)	Three Ethernet ports	Three Ethernet ports
Status indicator on the Storage System Fabric (SSF)	One bi-color light-emitting diode (LED) for each port	One bi-color LED for each port
Fault, status, and activity indicators	Two LEDs	Two LEDs

Note: Because version 1 (13-port) and version 2 (16-port) Slammer PIMs use different components, version 1 and version 2 PIMs cannot co-exist in the same Slammer. However, a multi-Slammer system can contain a mix of version 1 and version 2 Slammers, as long as the PIMs within each Slammer use the same version. Also, because the presence of a 2 Gb/s component in an otherwise 4 Gb/s loop will force the loop to function at 2 Gb/s, 4 Gb/s optical and 2 Gb/s copper back ends cannot coexist effectively in the same Pillar Axiom system.

Table 3 NIM configurations used in each Slammer CU

Slammer type	Pillar Axiom 300	Pillar Axiom 600
NAS Slammer	Two copper or optical ports with 1 Gb/s GbE SFPs	Four copper or optical ports with 1 Gb/s GbE SFPs
SAN Slammer, FC only	Two optical FC ports, each of which supports 1, 2, or 4 Gb/s FC SFPs	Two optical FC ports, each of which supports 1, 2, or 4 Gb/s FC SFPs, or
		Two optical FC ports, each of which supports 2, 4, or 8 Gb/s FC SPFs
SAN Slammer, Combo FC and iSCSI	Two optical FC ports, each of which supports 1, 2, or 4 Gb/s FC SFPs, plus two copper iSCSI ports, each of which supports 1 Gb/s GbE SFPs	Two optical FC ports, each of which supports 1, 2, or 4 Gb/s FC SFPs, plus two copper iSCSI ports, each of which supports 1 Gb/s GbE SFPs.
iSCSI Slammer	Two copper iSCSI ports, each of which supports 1 Gb/s GbE SFPs	Two copper iSCSI ports, each of which supports 1 Gb/s GbE SFPs

Table 4 Processors used in each Slammer CU

Components	Pillar Axiom 300	Pillar Axiom 600
Processors	Two 2.4 GHz Pentium-4 Xeon processors	One 2.6 GHz Dual-Core or 2.2 GHz Quad-Core AMD Opteron family processor
Memory in each control unit (CU)	3 GB for each CU	12 or 24 GB for each CU

Table 5 Power supplies used in each Slammer CU

Component	Pillar Axiom 300	Pillar Axiom 600
Power supply	One power supply and fan	Two redundant power supplies and fans

Both NAS and SAN Slammers are connected to the system storage pool through up to eight FC ports (four for each Slammer CU). Version 1 (legacy) ports are

capable of 2 Gb/s connections only, and Version 2 ports are capable of 2 Gb/s or 4 Gb/s connections. The internal FC fabric is connected through a set of FC loop switches. The Slammer virtualizes the storage pool, so you can easily increase the number and size of filesystems and LUNs.

Note: The Slammers do not perform RAID processing. RAID processing is handled at the Brick level.

Slammer Software Components

Most of the Pillar Axiom system software functionality resides in the Slammer. The software stack is logically separated by NAS and SAN functionality and management functionality. The Slammer software stack includes the following software architecture layers:

Slammer Platform Services

Platform Services provides the code necessary for the Slammer to interact with hardware and basic system operational functionality. Specific services include:

- Hardware initialization and control
- Interprocess communication (IPC) and remote procedure calls (RPC)
- Memory allocation
- Network protocol drivers
- Small Conputer System Interface (SCSI) protocol drivers
- Fibre Channel (FC) drivers for private interconnect
- RS-232 support
- Private interconnect switch management
- Diagnostic management
- Topology discovery

Management Configuration Manager

The Management Configuration Manager manages system restarts, software updates and activation, and system event services. It manages configuration changes to the system components and to the storage pool. It also communicates with the MCC Core on the Pilot to download and enforce policies.

Configuration Manager	The configuration manager interacts with the Array Manager to manage the shared storage pool.
Distributed Services Component	The distributed services component (DSC) provides a common set of services among all Slammer control units (CUs) in the system. The DSC facilitates NAS and SAN integration. For example, this component makes it possible for a NAS Slammer to take a SAN LUN offline, or for a SAN Slammer to take a NAS filesystem offline. The DSC also allows a SAN Slammer CU to expand storage for a NAS CU.
Data Mover	The Data Mover provides the functions necessary for backup services. An API enables control from the Management Configuration Manager or through the Network Data Management Protocol (NDMP). The Data Mover communicates with the filesystem directly for file-level transfers and with the VLUN layer for block-level transfers. The Data Mover supports the following types of tape library attachments: ○ Direct and SCSI attachment to the Slammer controlled by an internal NDMP daemon ○ Network attachment controlled by an internal NDMP daemon ○ Network attachment controlled by an NDMP daemon running on a third-party software package media server
SCSI Command and Control	This layer includes host FC and iSCSI drivers as well as the command processing function for SAN attachment. The SCSI Command and Control component processes all I/Os through the VLUN layer.
Common Internet File System (CIFS) Protocol	The CIFS protocol provides Windows and other CIFS clients access to the Pillar Axiom filesystem through the GbE ports. Storage is presented as CIFS shares.
Network File System (NFS) Protocol	The NFS protocol provides UNIX, Linux, and other NFS clients access to the Pillar Axiom filesystem through the GbE ports. Storage is presented as NFS mount points.
Meta Filesystem	The Meta Filesystem provides a protocol neutral file system supporting files, directories, and other filesystem objects. It uses a transaction journal and advanced read caching

algorithms to maximize I/O throughput. All filesystem I/Os are processed through the Virtual LUN (VLUN) layer.

Virtual LUN (VLUN)

A VLUN is defined as:

A logical unit of storage where customer data is striped and optionally mirrored across two or more Bricks. In a small system, such as a minimally configured Pillar Axiom 300 system, the mirroring may occur between two primary data segments within a single Brick.

VLUNs support filesystems, LUNs, clones, and snapshots and are internally managed, block-level structures. System administrators manage VLUNs only indirectly when they create or modify logical volumes.

The VLUN layer enables system-wide access to filesystems, LUNs, clones, and snapshots. All storage volumes are accessed through a VLUN. The VLUN interface is block-oriented and supports read and write access through a high-performance block cache. VLUNs form the basis of the Pillar Axiom system's any-to-any access capabilities.

Block Cache

The Block Cache component supplies all read and write cache management. All write operations are mirrored to the redundant control unit's cache, providing full data integrity through a range of restart and failure scenarios. In addition, all snapshot actions are processed by the Block Cache service.

Array Manager

The Array Manager (AM) provides storage pool virtualization functions. It allows all data on Bricks to be treated as a single storage pool. The AM handles LUN mapping to physical drive storage locations.

Generic Data Mover

The generic Data Mover is a platform-supplied service that provides data transfer and messaging requests across the Storage System Fabric (SSF).

Fabric Manager

The Fabric Manager manages all pathways within the SSF. Through the Fabric Manager, system components can send requests to each other without actual knowledge of the channel used to service the request. The Fabric Manager handles path failures by routing requests through alternate channels within the fabric.

InterConnect	The InterConnect communicates directly with the FC circuitry.

Brick Storage Enclosures

There are three types of Bricks:

- SSD (solid-state drive) Bricks.
- SATA (serial advanced technology attachment) Bricks.
- FC (Fibre Channel) Bricks.

Each of these Brick types is a unique Storage Class. All three Brick types can be mixed in a single Pillar Axiom system. All three Brick types can also be mixed in a single Brick chain, but we recommend separating Brick types into distinct chains, if possible.

All types of Bricks support both RAID 5 and Distributed RAID within the same array of the Brick simultaneously. However, there is no advantage to using Distributed RAID on SSD Bricks because of the enhanced speed of the SSD drives.

The drives used in SSD Bricks feature:

- SATA interface.
- SSD media.
- 50 GB capacity.

The drives used in SATA Bricks feature:

- Multiplexed SATA interface.
- Hard disk drive (HDD) media.
- 400 GB (legacy), 500 GB, 750 GB (legacy), 1 TB, or 2 TB capacity.

The drives used in FC Bricks feature:

- Fibre Channel Interface.
- HDD media.
- 146 GB (legacy), 300 GB, or 450 GB capacity.
- Concurrent dual-port transfers.

There are two types of FC Bricks:

- FC RAID Bricks.
- FC Expansion Bricks.

The FC RAID Brick is a head of string controller to which a FC Expansion Brick is attached in order to add additional drives.

Brick Functional Description

The architecture of a Brick provides built-in redundancy of its components. Each RAID controller on a serial advanced technology attachment (SATA) Brick acts as an active controller for one of the two RAID arrays in the Brick and has paths to each of the drives inside the Brick. On Fibre Channel (FC) Bricks, both controllers can access all the drives in the Brick, as well as any attached expansion Bricks. On solid-state drive (SSD), SATA, or FC Bricks, if one controller fails, the other controller continues to process I/Os for all arrays within the Brick. The RAID controllers are hot-swappable.

Data stored in a Brick is accessed by the Slammers through the Fibre Channel-based Storage System Fabric (SSF). A Slammer can read or write data to or from any RAID array on any Brick, because storage resources are shared across all controllers. Standard SCSI commands enable communication between Slammers and Bricks.

Advantages of the Brick RAID architecture include the following:

- Rebuilding from a failed drive is managed at the Brick level, minimizing the performance impact during the rebuild and drastically shortening the rebuild time.
- Because RAID processing is performed at the Brick level, not the Slammer level, Brick RAID controllers free up the Slammer to focus on I/O instead of RAID processing. This improves scalability because the number of I/O operations per second (IOPs) increases as capacity is added.
- If one RAID controller fails, the remaining RAID controller has enough processing power to handle the entire array, so there is no degradation of performance.
- RAID controllers have redundant Fibre Channel cross connections to increase the number of paths from any Slammer control unit (CU) to any RAID array.
- Drives are mounted horizontally in the tray to facilitate drive replacement.
- The Brick firmware can perform enhanced drive error recovery by temporarily removing a suspect drive from the storage array and performing extended recovery actions on that drive. This promotes drive reliability.

The Brick firmware monitors SMART (self-monitoring, analysis, and reporting technology) data for each individual drive. The firmware will proactively remove a drive from service if the drive exceeds SMART error thresholds and there is a spare drive available.

- Redundant power and cooling ensures that a Brick will continue to function normally with only one power supply active.

The controllers monitor the following Brick status information:

- Internal temperatures
- Fan speed control and fan speed feedback
- Drive carrier detection
- World wide name (WWN) and system serial number information
- Power supply status and detection information

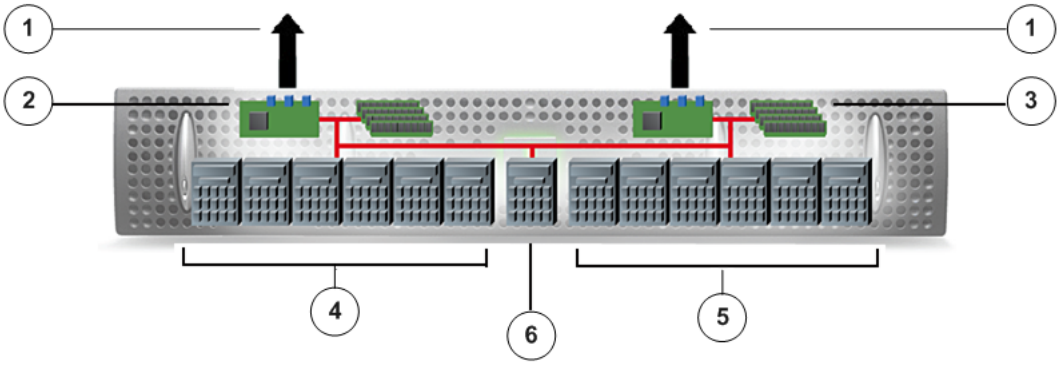
The RAID controller tracks system resets. The RAID controller can also track the removal or insertion of a Fibre Channel cable.

SSD Bricks

Solid-state drive (SSD) Brick enclosures contain 13 SSDs, of which 12 are arranged in two six-drive arrays. The 13th drive is used as a hot spare for automatic failover.

SSD Brick storage enclosures are managed by a pair of version 2 RAID controllers.

Figure 4 SSD Brick components



Legend	
	1 Connects to a Slammer or another RAID controller
	2 RAID controller
	3 RAID controller
	4 RAID group (6 SSDs)
	5 RAID group (6 SSDs)
	6 Hot spare

SSD Bricks use the same SATA interface as SATA Bricks.

The number of SSD Bricks supported depends on the number of Slammers:

- Single-Slammer systems support up to eight SSD Bricks.
- Two and three-Slammer systems support up to 16 SSD Bricks.
- Four-Slammer systems support up to 32 SSD Bricks.

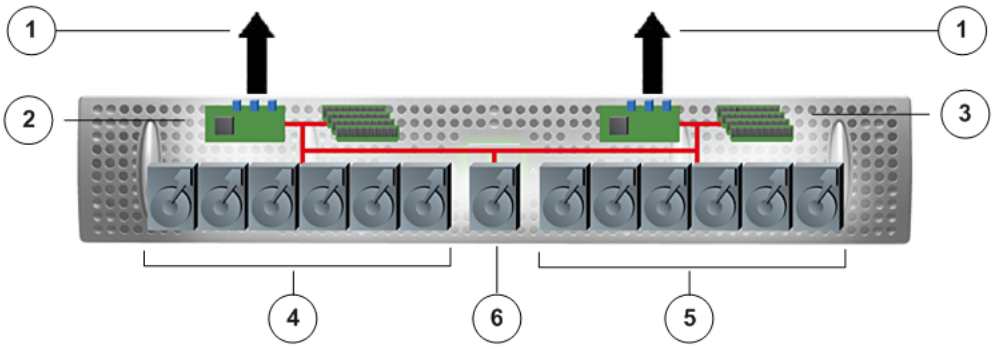
Note: Pillar Axiom 300 systems do not support SSD Bricks.

SATA Bricks

SATA Brick storage enclosures contain 13 hard disk drives (HDDs), of which 12 are arranged in two six-drive arrays. The 13th drive is used as a hot spare for automatic failover.

SATA Brick storage enclosures are managed by a pair of RAID controllers.

Figure 5 SATA Brick components



Legend	
	1 Connects to a Slammer or another RAID controller
	2 RAID controller
	3 RAID controller
	4 RAID group (6 HDDs)
	5 RAID group (6 HDDs)
	6 Hot spare

Under normal conditions, each controller provides access to and control over an array of six HDDs. Under failover conditions, a single controller can control and provide access to both arrays. All Pillar Axiom systems support SATA RAID controllers. SATA Brick controllers come in two types:

- Version 1 (legacy) controllers have one set of four Fibre Channel (FC) ports with high speed serial data connector (HSSDC) connectors, and support only 2 Gb/s copper connections.
- Version 2 controllers have two pairs of FC ports with small form factor pluggable (SFP) connectors, and support either 2 Gb/s copper or 4 Gb/s optical connections. In addition, version 2 SATA controllers employ an updated chipset with greater internal bandwidth, and they support SATA HDDs as well as solid-state drives (SSDs).

Because version 1 and version 2 SATA controllers use different internal communication protocols, these two types of SATA controllers cannot co-exist in the same Brick chassis. In other words, you cannot use a version 2 SATA controller to replace a legacy version 1 controller. A Pillar Axiom system can, however, contain a mix of version 1 and version 2 SATA Bricks.

For a complete list of the rules for configuring SATA Bricks, refer to the *Pillar Axiom SSF Cabling Reference* for the Pillar Axiom system version being configured.

The number of SATA Bricks supported depends on the type of Pillar Axiom system and the number of Slammers:

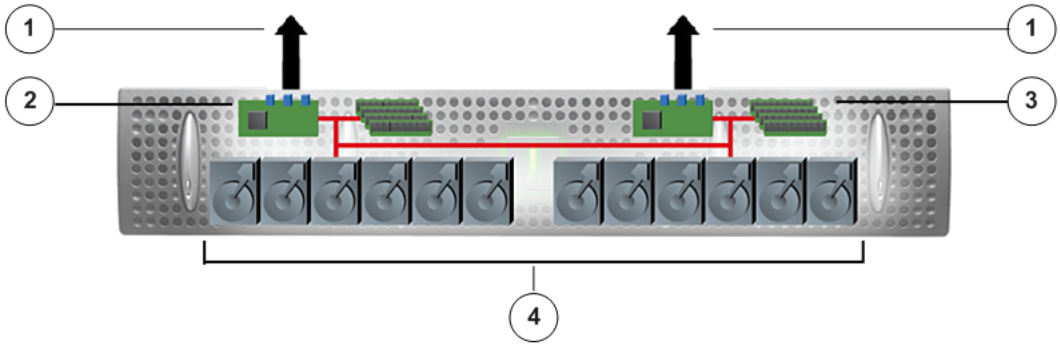
- Single-Slammer Pillar Axiom 300 systems support up to four SATA Bricks.
- Single-Slammer Pillar Axiom 600 systems support up to 32 SATA Bricks.
- Two, three, and four-Slammer Pillar Axiom 600 systems support up to 64 SATA Bricks.

Fibre Channel Bricks

Fibre Channel (FC) Brick storage enclosures contain 12 hard disk drives (HDDs) arranged in a single 11-drive array plus a hot spare. FC Bricks do not have a dedicated hot spare; instead, any drive can be utilized as a spare. If a drive fails, the rebuild occurs on the current hot spare. After the failed drive has been replaced, it becomes the new hot spare.

FC Brick storage enclosures are managed by a pair of RAID controllers.

Figure 6 FC Brick components



Legend	
	1 Connects to a Slammer or another RAID controller
	2 RAID controller
	3 RAID controller
	4 RAID group (12 HDDs including a hot spare)

FC Bricks normally come in pairs of one FC RAID Brick plus one FC Expansion Brick. A FC RAID Brick, however, can be installed without an FC Expansion Brick.

The number of FC Bricks supported depends on the type of Pillar Axiom system and the number of Slammers:

- Single-Slammer Pillar Axiom 300 systems support up to four FC Bricks.
- Single-Slammer Pillar Axiom 600 systems support up to 16 FC Bricks.
- Two, three, and four-Slammer Pillar Axiom 600 systems support up to 32 FC Bricks.

SSD, SATA, and FC Bricks can co-exist on the same Brick string subject to configuration recommendations.

A given Brick string can contain up to a total of four FC Bricks (RAID or Expansion). Only one FC Expansion Brick can be added to an FC RAID Brick, and a Brick string cannot contain more than two FC Expansion Bricks.

For a complete list of the rules for configuring FC Bricks, refer to the *Pillar Axiom SSF Cabling Reference* for the Pillar Axiom system version being configured.

Brick Software Components

The following software resides in the Brick:

Fibre Channel driver

The Fibre Channel (FC) driver serves as an isolation layer between the target FC hardware and the remainder of the Brick firmware. This allows the FC hardware to change without great impact upon the remaining code. The FC driver also translates hardware-specific sets of data structures to simple Small Computer System Interface (SCSI) requests that are queued and eventually processed by the SCSI layer of the Brick firmware.

The FC driver includes a transfer manager that facilitates data transfers from the Brick's data buffer across the FC link to and from the requestor. This includes translating scatter/gather lists into the format expected by the FC hardware.

SCSI layer

The SCSI layer receives queued SCSI requests from the target FC driver. It validates these requests and converts them into the standardized command request format used by the rest of the Brick firmware. This conversion includes correcting command parameters from the big-endian data

	format commonly used in SCSI requests to the native data format of the processor used by the Brick. Once the request is parsed, the SCSI layer sends the newly created command request block on to the command processor.
Command processor	The command processor accepts command request blocks from the SCSI layer and dispatches these requests to the various routines responsible for their timely completion.
Cache management	The cache management code supplies buffer memory for use in processing commands. The cache manager also maintains previously used valid data as long as possible so that this cached data can be used to satisfy future requests without requiring access to slower media types. Should the data needed for a request not be available in cache, the cache code issues the RAID requests needed to gather the data from attached drives.
RAID engine	The RAID engine converts RAID requests into one or more drive requests directed to the storage media. These drive requests are queued and eventually processed by the drive manager. This process requires the RAID engine to calculate the logical-to-physical mapping for various RAID types so that a given RAID unit address always accesses the same physical media address.
Partition manager	The partition manager creates, saves, and restores the information required to define the logical RAID units used by the RAID engine.
Drive manager	The drive manager is responsible for handling all drive requests and any error recovery required to complete those requests. It also maintains the current state for all physical media (drives) in the Brick, and updates that state based upon the results of each I/O operation.
Storage I/O driver	The storage I/O driver provides an interface to the initiator hardware that the FC or SATA driver provides for the target hardware. The storage I/O driver converts the drive requests issued by the RAID engine into the hardware-specific request structures that the initiator hardware requires to communicate with the storage media. This driver also acts as an isolation layer between most of the Brick firmware and the initiator hardware, minimizing code changes when hardware is updated.
Pre-emptive copy	Pillar Axiom Pre-emptive Copy initiates recovery of a drive before the drive fails. By simply copying the data to the

spare drive rather than rebuilding the failing drive, pre-emptive copy shortens recovery time and reduces the chance of data loss.

CHAPTER 3

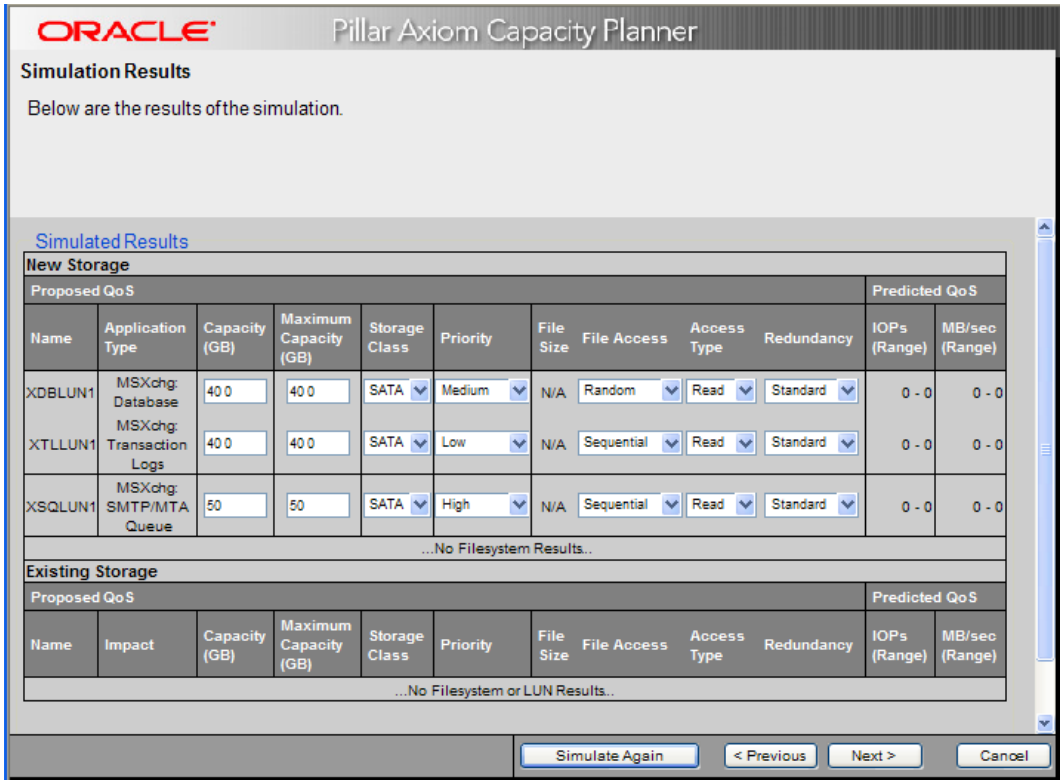
Pillar Axiom Software Overview

Policy-Based Provisioning and Storage Management

The Pillar Axiom system manages storage resources using administrator-defined policies, which are the basis of the storage management system. Performance, utilization, and availability metrics are tailored to individual logical volumes. Policies are established through the Pillar Axiom user interface, using the graphical user interface (GUI) or the command line interface (CLI), and implemented in the core layer of the Pilot. The Management Configuration Manager facilitates the control, monitoring, and reporting of the Slammers to help the Pilot enforce these policies.

Various tools are available in the Pillar Axiom Storage Services Manager (the GUI) to help create filesystems and LUNs. For example, the Pillar Axiom Capacity Planner (shown in the following figure) helps model simulations and create filesystems or LUNs based on those models. When the system administrator enters the information requested by the Pillar Axiom Capacity Planner wizard, the logical volume is built according to those specifications.

Figure 7 Pillar Axiom Capacity Planner



The Pillar Axiom system can create a filesystem or LUN to match data performance, relative priority, and access pattern. Standard or compliance retention policies can also be applied to filesystems. This flexibility ensures that applications ranging from mission-critical to archive receive the appropriate system resources.

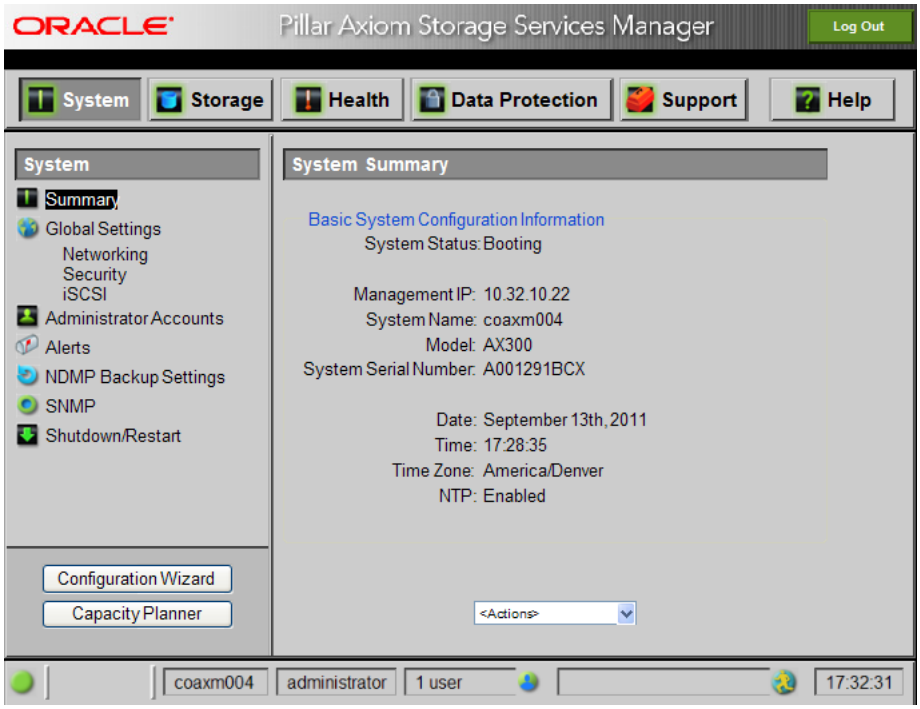
Standard application-aware performance profiles for common application environments are available from the Pillar World Wide Customer Support Center.

The Pillar Axiom Storage Services Manager can also modify the performance, priority, access, or retention properties of existing volumes. If necessary, this process can automatically migrate user data to a different physical location on the storage pool to fulfill the modification request.

Pillar Axiom Storage Services Manager

The Pillar Axiom Storage Services Manager is an easy-to-use GUI. It is organized into sections to help you configure and monitor your Pillar Axiom system. These sections appear in the top row of the following figure.

Figure 8 Pillar Axiom Storage Services Manager



These sections perform the following functions:

- System—Configures global settings, administrator accounts, alerts, Network Data Management Protocol (NDMP) settings, and Simple Network Management Protocol (SNMP) settings. Also, shuts down or restarts the system.
- Storage—Manages volumes (filesystems and LUNs), clones, and snapshots.
- Health—Monitors system hardware, performance, and status events. Provides access to Guided Maintenance, which enables storage administrators to identify and replace failed components, often without interrupting system operation.
- Data Protection—Schedules snapshots and provides immediate replication (Snap FS and clones). Performs remote replication, VSS Provider download, tape device management, and NDMP operations.
- Support—Manages software updates, hardware status, log collection, system tools, utilities downloads, technical documentation, contact information, and system shutdown and restart.

The status bar at the bottom of the Pillar Axiom Storage Services Manager screen provides instant feedback on system performance, running background tasks, or administrator actions that require attention. For detailed information about the items on the status bar, refer to the *Pillar Axiom Administrator's Guide*.

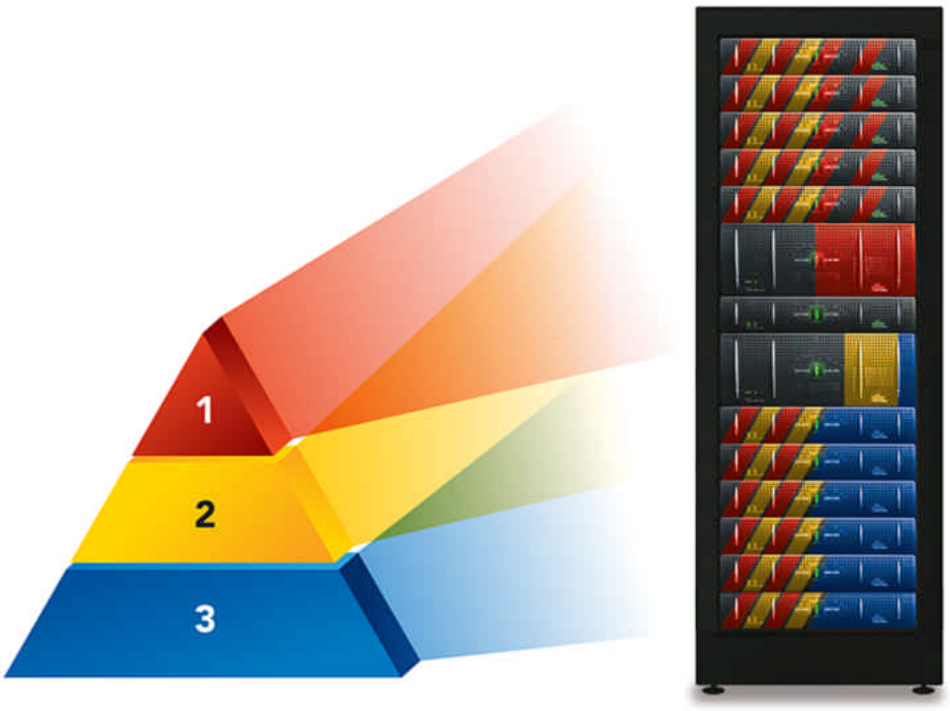
Quality of Service Attributes

The Pillar Axiom system defines a set of policies, or Quality of Service (QoS) attributes, that govern the QoS for the volume (filesystem or LUN). These policies determine how data is stored in the storage pool, the queueing and caching priority of the data, and the type of RAID used for the data, so that the highest performance is achieved.

A Pillar Axiom system allocates storage with different application priorities. System resources are disproportionately applied to deliver the requested Quality of Service (QoS) for each volume. Storage administrators can allocate storage resources and define storage automation parameters through the Pillar Axiom Storage Services Manager, and the Pillar Axiom software takes care of the rest. A single Pillar Axiom system can deliver many tiers of storage, differentiated by performance, availability, data protection, capacity, and scalability. A common implementation has three tiers of storage services:

- Tier 1: Highest performance and availability levels for mission-critical applications, represented as red in the following figure.
- Tier 2: Increased performance and high availability for mid-performance applications, represented as yellow in the following figure.
- Tier 3: Adequate performance and availability to support business-utility applications such as file sharing and archive, represented as blue in the following figure.

Figure 9 Pillar Axiom system data layout



The Pillar Axiom Storage Services Manager makes it possible to assign array resources for each volume (filesystem or LUN), much the same as configuring a virtualized server resource.

Dynamic provisioning and profile-based resource assignments provide the necessary flexibility for quickly and easily adjusting capacities to meet ever changing business storage demands. All of the QoS settings, along with the redundancy attributes, are utilized to determine the RAID type of a volume (filesystem or LUN).

Performance Profiles

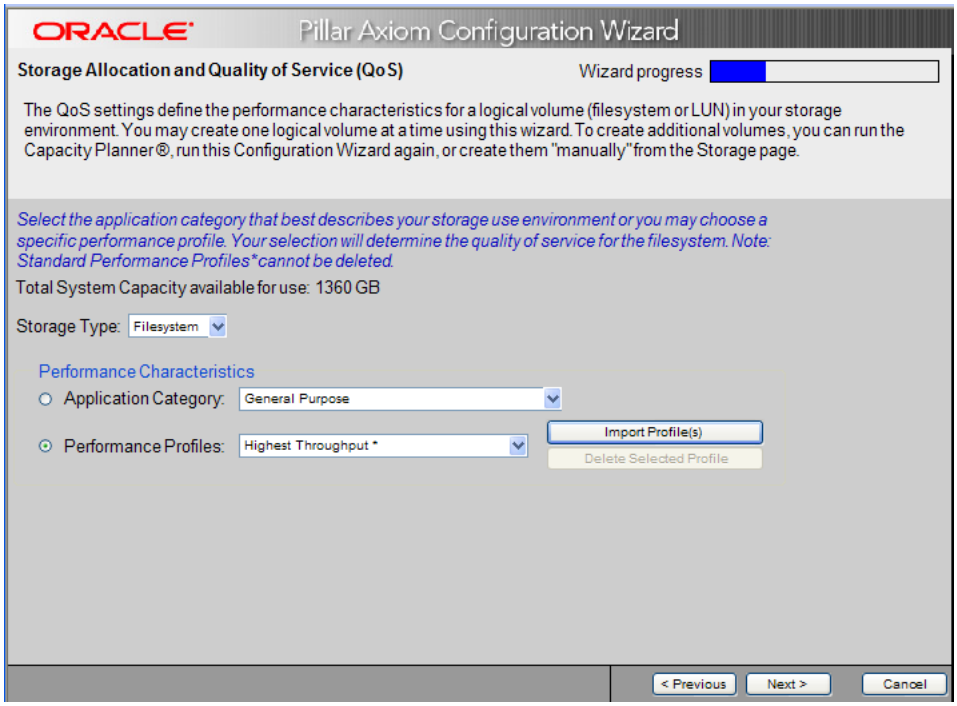
Performance profiles determine the quality of service for logical volumes (filesystems or LUNs).

Pillar Axiom systems provide a variety of pre-set application profiles that leverage best practices already tested by Pillar Axiom users. Alternatively, storage administrators can import Pillar-provided profiles. These policies, combined with Pillar Axiom queue management policies, make it possible to prioritize performance in the storage pool. Administrators can also use oversubscription (thin provisioning) to over-provision resources.

The Pillar Axiom Configuration Wizard allows storage administrators to apply performance profiles to volumes (filesystems or LUNs) during system

configuration. Administrators can select predefined performance characteristics that are based on the category of application that typically accesses the volume or on a more specific performance profile.

Figure 10 Pillar Axiom Configuration Wizard



Administrators can choose between the following application categories:

- Disk-to-disk backup
- Hardware/software development
- Email messaging
- Database and transaction processing
- General purpose

Administrators can request application-specific performance profiles from the Pillar World Wide Customer Support Center, or they can choose one of the following pre-configured performance profiles:

- **Highest Throughput**. This profile stripes data across all available drives in the corresponding Storage Class. The Highest Throughput profile is intended for use in benchmarking or in environments having a small number of logical volumes to be configured.

  Important! Striping data across all available drives in a Storage Class can lead to unexpected contention in larger configurations.

- **Oracle Automatic Storage Management (ASM)**. This profile makes use of a wide stripe (which has a strip depth of 1 MB) for the configured logical volume. A wide stripe minimizes the number of seeks required to service data requests in an Oracle ASM environment by matching strip size to the application request size.

Thin Provisioning

The Pillar Axiom system allows you to provide thinly provisioned volumes (filesystems and LUNs).

Thin provisioning is defined as:

An approach to storage allocation in which a logical volume (filesystem or LUN) appears to be much larger than the storage actually allocated to it. Additional storage is dynamically allocated when necessary. Administrators interact with thinly provisioned volumes when configuring their capacity and growth increments. These types of volumes are sometimes referred to as *sparse filesystems* and *sparse LUNs*.

Note: A Pillar Axiom system uses binary units to calculate and display the capacity of physical storage and the size of logical volumes:

1 MB = 1024^2 (1,048,576) bytes

1 GB = 1024^3 (1,073,741,824) bytes

1 TB = 1024^4 (1,099,511,627,776) bytes

The following sections describe thin provisioning and how it affects storage capacity.

Over Commitment

Traditionally, when storage is allocated to an application, the allocation is dedicated to that application. This assignment prevents other applications from accessing this capacity, even when the amount allocated is never used. Because of this allocation strategy, the capacity is stranded and cannot be leveraged in support of additional needs.

Thin provisioning mitigates these issues by allowing storage administrators to *over commit* a logical volume (LUN or filesystem) by:

- Allocating capacity based on future needs.

- Drawing on a common pool of storage as capacity is consumed.

Thin provisioning allows an administrator to create a logical volume of any size without committing that capacity at that time. Each application has what appears to be all the storage needed for ongoing operations, but without the physical capacity locked to a particular volume.

Administrators can create logical volumes up to the maximum size allowed for the OS with little physical storage assigned to the volume. As data is written to the thinly provisioned volume and capacity is consumed (called *in-fill*), the system automatically allocates additional capacity to the LUN or filesystem in increments.

Note: Solid-state drive (SSD) Bricks do not support thinly provisioned volumes.

Free Capacity in Over-Committed Volumes

A minimum amount of free space is required to create a new logical volume (filesystem or LUN). The actual amount of physical capacity that is consumed from the system free space when you create a new logical volume depends on several factors.

These factors are:

- The RAID geometry of the volume.
- The redundancy Quality of Service (QoS) setting of the volume.

To determine the actual physical capacity needed, the system adds the following:

- To account for parity, the system increases the requested capacity by different amounts, depending on the RAID geometry:
  - 20% for RAID 5 (SATA)
  - 10% for RAID 5 (FC)
  - 100% for Distributed RAID
- If redundancy for the volume is set to Double, the system doubles the physical allocation.

For example, if the requested capacity for a logical volume is 250 GB, and the volume uses RAID 5 geometry in SATA storage, the system allocates an additional 50 GB. If the volume has a redundancy setting of Double, the system allocates an additional 500 GB, for a total physical space allocation of 800 GB.

Provisioning Over-Committed Volumes

The capacity reserved for thin provisioning, which is part of the system overhead, is accounted for in the available capacity that the system reports. In other words, what the system reports as available capacity is fully available for the provisioning of logical volumes.

Unused capacity in the storage array can decrease over time. This decrease is due primarily to two events:

- New volumes are created.
- Over-committed (thinly provisioned) volumes are provisioned (filled in) when they grow. When no unused system capacity remains, the system uses this reserve to fill in the thinly provisioned volumes.

For storage area network (SAN) systems, the degree to which a LUN is over-committed, which defines its thinness, depends on the nature of the host applications that access the LUN. If only specific portions of a LUN are ever accessed by applications, the thinness of that LUN remains the same. As applications attempt to access more and more different areas of the LUN, the system allocates more and more physical space for the LUN, causing the thinness to decrease.

For network attached storage (NAS) systems, the degree to which a filesystem is over-committed, which defines its thinness, depends on the maximum amount of space ever used by this filesystem. As a filesystem consumes more space, it requires more allocation of physical storage to become less thin.

Reducing the space used by a filesystem (by deleting files or snapshots, for example) will not result in physical storage being freed. Thus, reducing the space used by a filesystem will not increase the thinness of the filesystem.

Growth Increments

When the system allocates capacity for a logical volume, the system divides the allocation into slices (called *growth increments*) and uses as many of them as it needs.

Each growth increment is between 1 and 2 GB. For example, if the volume is 2 TB, the system may use between 1024 and 2048 growth increments for the allocation. The exact value depends on the combination of the following choices that characterize the underlying storage for the volume:

- Type of Brick (Fibre Channel or serial ATA)
- RAID geometry (RAID 5 or Distributed RAID)
- Strip size (normal or 1 MB)

Note: When the system needs to grow or in-fill a logical volume, the system returns an error if sufficient capacity does not exist within the Storage Class associated with the volume, even when sufficient capacity exists in other Storage Classes.

Capacity Overhead

Plans for the provisioning of logical volumes must take into account the extra capacity the system allocates to overhead.

To accommodate the level of RAID protection required to allocate a newly created logical volume (filesystem or LUN), the system adds a certain amount of overhead to a request for the capacity of the filesystem or LUN. The capacity consumed and reported for RAID 5 logical volumes includes that overhead. This overhead varies, depending on the RAID geometry and Storage Class assigned to the volume. For RAID 5, the overhead is as follows:

Serial ATA drives and SSDs	20%
Fibre Channel drives	10%

For Distributed RAID, the capacity consumed and reported for logical volumes is twice the requested amount, regardless of Storage Class.

Parity in Reported Capacities

RAID arrays have both physical and virtual capacity.

The physical capacity of a RAID array that is reported includes capacity for parity. Sizes reported in capacity usage summaries and the sizes reported for total, used, and free system capacities are in terms of raw physical capacities.

The virtual capacity of a RAID array that is reported, however, does not include capacity for parity. The ratio between the virtual capacity and the physical capacity depends on whether the storage is RAID 5 or Distributed RAID:

RAID 5: serial ATA (SATA) drives and solid state drives (SSDs)	5:6

RAID 5: Fibre Channel (FC) drives	10:11
Distributed RAID: FC, SATA, and SSD drives	1:2

Reclaiming Capacity

When a user deletes a logical volume (LUN or filesystem), the system reconditions the space (by writing a predefined bit pattern) before freeing it for reuse. As the previously allocated capacity frees up, it becomes available for allocation.

Note: When a large volume is being deleted, the operation can take awhile for all the capacity to be reclaimed. Because of this additional time needed for reconditioning, the amount of used capacity plus the free capacity may not equal the total capacity. During this time, the graphical user interface (GUI) displays the amount of capacity remaining to be reconditioned.

For filesystems, when a user deletes a file that has no snapshots associated with it, the freed blocks appear as free capacity in the snapshot repository that is associated with the parent filesystem. Utilization commands (such as `df` or `du` on Unix systems) show this newly freed capacity.

If the deleted file has snapshots associated with it, the system preserves the blocks in the snapshot repository for that filesystem. In this case, the number of free blocks for the filesystem does not change. Utilization commands show no change in the used and free space for that filesystem. To return these blocks to the free system space, all snapshots in the filesystem must be deleted.

Note: If, however, this deleted file was modified after the most recent snapshot and contained new blocks of data not captured by that snapshot, the system reclaims those new blocks. Utilization commands in this case would show those newly freed blocks as additional free capacity for that filesystem.

High Availability and RAID Geometries

Pillar Axiom systems are designed to provide the highest possible levels of availability.

Many Pillar Axiom components, such as network interfaces, control unit motherboards, power supplies, and fans, are redundant. This redundancy is intended to eliminate single points of failure.

The software layers of the Slammer support high availability. Software processes on each control unit (CU) are in constant communication with each other regarding their status and ability to perform.

The Pillar Axiom system uses a double-safe write system. The system secures the I/O in battery-backed, non-volatile RAM (NVRAM), so that the I/O is safe in case of external power loss. The Pillar Axiom system's redundant CU architecture secures the I/O in both CUs before the write operation is acknowledged as a complete transaction. The write is secured in two places, so only a catastrophic system event can affect the write integrity.

A double-safe write operation stores the data in local memory on the primary CU and in battery-backed memory on the alternate CU. If the primary CU fails completely, the alternate CU already has the data so it can recover and continue.

Storage Classes

The Storage Class feature allows you to specify the preferred storage media to use for a logical volume (filesystem or LUN).

A Storage Class is defined as:

> A categorization of physical storage, each category having distinct characteristics with regard to performance characteristics of data access. Example Storage Classes in a Pillar Axiom system are serial ATA (SATA), Fibre Channel, and solid state drive (SSD). Pillar Axiom 600 systems allow an administrator to explicitly manage volume placement within the overall system storage pool, first by Storage Class and then by relative priority within that Storage Class.

Pillar Axiom systems support the following three Storage Classes:

- FC (Fibre Channel drives, 15K RPM)
- SATA (serial ATA drives, 7.2K RPM)
- SLC SSD (single-level cell, solid state drive)

Note: Which Storage Classes are available on a particular Pillar Axiom system depends on the types of Brick storage enclosures you have installed on the system.

A Storage Class has these attributes:

- A newly created logical volume is associated with a single Storage Class.
- The Pillar Axiom Storage Services Manager graphical user interface (GUI) shows the capacity available within each Storage Class.
- The system will not create a logical volume when the available space for the associated Storage Class is insufficient to accommodate the capacity requested for the volume.

For FC and SATA Storage Classes, the striping of a logical volume is across a number of drives in a collection of RAID groups. The number of drives depends on the Quality of Service (QoS) priority setting for the volume. For the SLC SSD Storage Class, striping for a volume is across all available drives, regardless of the priority setting.

RAID Array Stripes

Pillar Axiom systems support RAID 5 and Distributed RAID geometries within the same Brick array.

RAID 5 arrays support the following strip sizes:

- For wide stripes: 1 MB for each strip.
- For standard stripes:
  - Fibre Channel (FC) Bricks: 64 KB for each strip.
  - Serial ATA (SATA) and solid-state drive (SSD) Bricks: 128 KB for each strip.

Distributed RAID arrays are formed from pairs of standard strips (64 KB strips for FC and 128 KB strips for SATA and SSD) only.

For FC Bricks, a stripe is a collection of 10 data strips and one parity strip. Each strip (64 KB) is written to one of the drives in a FC Brick, which means the stripe is written across 11 drives. For FC Bricks, a stripe also contains 640 KB, but its width is 11. Each FC Brick contains one such array, plus a hot spare.

For SATA and SSD Bricks, a stripe is a collection of five data strips and one parity strip. Each strip (128 KB) is written to one of the drives in a RAID array, which means the stripe is written across six drives. For SATA and SSD Bricks, a stripe contains 640 KB, and its width is six. Each Brick contains two such arrays, plus a hot spare.

For an Oracle Automatic Storage Management (ASM) performance profile, strips contain 1024 KB (1 MB). The number of strips for each stripe remains the same, depending on the type of Brick. Also, the stripe width does not change, only the size of the strip does.

System Maintenance and Problem Detection

The Pillar Axiom system provides several different features to maintain the system and detect problems so that they can be resolved quickly. These features allow you to review system health, set up alert notification, review event log entries, and display performance statistics.

The following sections describe some of these features.

Administrator Actions

Administrator actions are alerts that notify the administrator when a condition needs attention.

Many administrator actions are accompanied by one or more menu selections that help the administrator resolve the condition.

Other administrator actions clear automatically when the issue that generated the administrator action has been resolved.

Call-Home

The Call-Home feature notifies the Pillar World Wide Customer Support Center about issues in the Pillar Axiom system. When a component operates in degraded mode or fails, the system automatically performs failover actions. Although a component failure does not cause downtime, manual intervention is sometimes required to repair or replace the failed component. The system sends a Call-Home message to initiate the repair or replacement process.

Call-Home log collection can be initiated by one of the following methods:

- Manual: The administrator has requested a log collection.
- Event-triggered: An event has triggered the Call-Home.
- Periodic: A specified time has elapsed since the Call-Home was triggered.

The system maintains a directory of data files, each of which captures a Call-Home session. Whenever one of these data files is overwritten or thrown away, a log entry is made noting that fact. The collection of data files represent the ten most recent Call-Home sessions and are listed in the GUI. The system administrator can select a session file and download it to a GUI client machine or send it directly to the currently targeted server.

Call-Home sessions can also be sent to a local Call-Home server. Contact the Pillar World Wide Customer Support Center for details.

Pillar Axiom Pre-Emptive Copy

The Pre-Emptive Copy feature further shortens the time it takes to rebuild a failed drive by doing the bulk of the work before a failure occurs, using a simple copy instead of a rebuild of the entire contents of the drive.

Pillar Axiom Pre-Emptive Copy is a Reliability, Availablity, Serviceability (RAS) feature of the Pillar Axiom RAID firmware that copies the data on a drive (which has been predicted to fail) to the spare drive before the suspect drive fails and is subsequently taken offline for replacement. This feature avoids performance degradation and potential exposure to data loss when the drive does fail.

Guided Maintenance

Guided Maintenance is a feature in the Pillar Axiom system that provides users with a method to identify and replace a field replaceable unit (FRU) in many cases without interrupting system operation and with minimal system degradation. There are four parts to Guided Maintenance:

- First, the system determines the status of the hardware based on events and diagnostics to accurately reflect the state of the system.
- Second, the system helps the administrator correctly identify the faulty component by presenting images of the FRU or beaconing the FRU.
- Third, if required, the system places system components in a condition to prepare for the replacement by redirecting activity from the faulty component to a redundant component.
- Fourth, the system guides the administrator through a set of procedures to replace the FRU.

Providing accurate system status and replacing FRUs are complex operations that involve many lower level components within the system. The process and details on how to maintain system operation are hidden from the user. The Pillar Axiom system is designed to be maintained by the user without requiring support from the Pillar World Wide Customer Support Center.

Each FRU has its own diagnostics which are called by the Pilot to verify that a FRU is accessible and functioning properly. The diagnostics are primarily used to

verify FRUs that have been added or replaced. The system also tracks parts that have failed and been removed to prevent re-insertion of failed components.

Guided Maintenance supports the identification and replacement of FRUs for Slammers and Bricks. For Pilots, Guided Maintenance allows the technician to identify a Pilot CU before replacing it. The *Pillar Axiom Service Guide* provides instructions on all replacement procedures.

Non-Disruptive Software Updates

With rare exceptions, all software updates that do not include a technology refresh are non-disruptive. Non-disruptive software updates help reduce downtime and potential data loss when updating Pillar Axiom software and firmware modules. The system manages application dependencies using a compatibility matrix to ensure all required dependencies are met.

Administrators can schedule software updates up to 72 hours in advance, and the updates can be scheduled to occur during off-peak hours.

Pillar Axiom SMIProvider

SMI-S (Storage Management Initiative Specification) is a set of management interfaces designed to make it easier for storage hardware and management applications from different vendors to work together. Integral to this initiative, an SMI provider serves as a translator between the storage hardware and management applications. With the Pillar Axiom SMIProvider, any management application written natively to the SMI-S standard can manage Pillar Axiom systems.

SMIProvider can respond to requests from SMI-based management applications in supporting network and storage administrators who need to:

- Detect newly added devices on a network.
- Use encrypted SMI connections.
- Provision logical volumes and storage resources.
- Map HBAs to specific LUNs, mask LUNs from specific HBAs, or both.

The Pillar Axiom SMIProvider has the following features:

- Caches much of the information requested by SMI clients to provide enhanced response times.

- Supports Service Location Protocol device discovery using a built-in service agent.
- Supports NAS and SAN environments.
- Supports Fibre Channel protocols.
- Supports SMI version 1.1 (CTP 1.1 certified by SNIA), version 1.2, and version 1.3.

  Note: To access the services of the Pillar Axiom SMIProvider, the client needs to be compliant with SMI-S version 1.1 or later.

Pillar Axiom System Statistics

The Pillar Axiom system collects statistics data on many facets of system operation. Each component of the Pillar Axiom system periodically collects key statistics covering such areas as filesystem read and write performance, block-level read and write performance, and error counts.

Statistical data captured in a binary format on the Pillar Axiom system can be downloaded from the Collect System Information page in the Pillar Axiom Storage Services Manager (GUI) for processing on a client machine.

Application tools provided with the Pillar Axiom system make it possible to process this data into a format suitable for use by statistical applications or spreadsheets. In that format, the statistical data can be used to:

- Analyze the Pillar Axiom system to determine bottlenecks and determine what needs to be tuned in order to optimize performance.
- Track Pillar Axiom system load and capacity.
- Produce reports and graphs for presentation.
- Integrate with existing performance monitoring and reporting applications.

For example, aggregated statistics can help determine whether the system is being fully utilized. An analysis of these statistics could help determine sizing for future expansion of the system.

CHAPTER 4

NAS Overview

NAS Functionality

Pillar Axiom systems provide support for Network Attached Storage (NAS) Network File System (NFS) and Common Internet File System (CIFS) clients. NAS software runs on a NAS Slammer, which can service both NFS and CIFS requests.

Pillar Axiom systems feature a multi-node scalable NAS file server with integrated high availability. Unlike many other NAS systems, the Pillar Axiom NAS file server is able to achieve the highest possible throughput, fault tolerance, and uptime under many different failure scenarios. At the same time, it remains extremely easy to setup, configure, and maintain. By taking into consideration all the configuration options available for your network, you can achieve the highest possible throughput and redundancy.

Up to 1024 filesystems can be configured on the Pillar Axiom system at the same time.

In addition, Pillar Axiom systems allow NAS and SAN systems to share the same storage pool.

Network File System (NFS)

NFS is used primarily in UNIX environments to share filesystems across IP networks. Key features of the Pillar Axiom system's NFS implementation include:

- NFS version 2 and 3.
- Transmission Control Protocol (TCP) and User Datagram Protocol (UDP).
- Network Lock Manager (NLM) for advisory file-level and byte-range locking.
- Quotas for tree, account, and groups of accounts.

Common Internet File Systems (CIFS)

CIFS is used primarily in Microsoft Windows environments across IP networks. The Pillar Axiom system currently supports Microsoft Active Directory Services (ADS) using the Kerberos or NT LAN manager (NTLM) login sequence.

Key features of the CIFS implementation include:

- Share-level and byte-range locking.
- Opportunistic Locks—exclusive, batch, and level II.
- User authentication through consultation with a Windows Domain Controller or Active Directory Server, using Kerberos or NTLM protocol.
- Quotas for trees, accounts, and groups of accounts.

Concurrent NFS and CIFS Access

The Pillar Axiom system supports concurrent access to files through NFS and CIFS by coordinating locking and file permissions and by properly managing the differences between the NFS and CIFS protocols.

During concurrent access, Pillar Axiom system supports collaboration between NFS clients and CIFS clients by using the following rules:

- NFS write requests do respect CIFS mandatory locks.
- NFS read requests do not respect CIFS mandatory locks.
- NFS read and write requests adhere to CIFS ACL settings.
- CIFS read and write requests adhere to NFS file permissions.

Refer to the *Pillar Axiom CIFS and NFS Multi-Protocol Planning Guide* for details.

NAS Networking

The network layer of NAS implementation supports the following:

- Link aggregation (trunking). Pillar Axiom systems support the IEEE 802.3ad standard.

- Virtual local area networks (VLANs).
- Jumbo frames (maximum transmission unit (MTU) 9000 bytes).

Pillar Axiom SecureWORMfs

The Pillar Axiom SecureWORMfs is a type of filesystem used to enforce data retention. Data is stored on a SecureWORMfs in a non-erasable, non-rewritable (protected) manner. SecureWORMfs utilizes Write-Once-Read-Many (WORM) technology that permits use of a read-writable device to store data that, once set to protected, can never be modified and can only be erased when the retention period expires.

There are two types of SecureWORMfs filesystems:

- Standard: Data is retained on the filesystem for a fixed period of time that is specified by the retention period settings; however, data can be deleted at anytime by deleting the entire filesystem.
- Compliance: Stores critical business data as stipulated by various government regulations. Data is retained on the filesystem for a fixed period of time that is specified by file-level retention settings. A SecureWORMfs compliance filesystem cannot be deleted if there are protected files on the filesystem. To prevent malicious manipulation of the Pillar Axiom system clock, this feature requires that the Pillar Axiom system clock rely only on a customer-provided Network Time Protocol (NTP) server.

SecureWORMfs retention is implemented when a protected file is closed.

Note: Pillar Axiom release 4.0 and later permits customers to downgrade a SecureWORMfs compliance filesystem to a SecureWORMfs standard filesystem so that it can be deleted. To delete the SecureWORMfs, the customer must contact the Pillar World Wide Customer Support Center. For your protection, authentication will be required.

CHAPTER 5

SAN Overview

SAN Functionality

LUNs are accessed by host systems through a storage area network (SAN) attachment. The Pillar Axiom system supports Fibre Channel (FC) host bus adapter (HBA) and iSCSI HBA SAN attachments. Supported host operating systems include the following:

- AIX
- HP-UX
- Linux
- Solaris
- Windows

See the *Pillar Axiom Interoperability Guide* for a current list of supported host operating systems.

Up to 1024 host systems of any of the supported operating systems can be connected to the Pillar Axiom system at the same time.

Supported FC topologies include the following:

- Fabric
- Private FC-AL
- Point-to-point
- Public loop

Supported iSCSI features include:

- Optional Challenge-Handshake Authentication Protocol (CHAP) authentication of iSCSI initiators.
- Configurable Pillar Axiom CHAP secret for bi-directional CHAP support.
- Optional access control of iSCSI initiators.
- Explicit or DHCP-based IP address configuration.

- iSNS Internet Storage Name Service.

Refer to the *Pillar Axiom iSCSI Integration Guide for Windows Platforms* for details.

The Pillar Axiom system supports a maximum of:

- 4096 LUNs for each system.
- 255 LUNs for each host.
- 4096 LUNs for each SAN Slammer.
- 256 TCP connections for each iSCSI port.
- 256 iSCSI Initiators for each iSCSI port.
- 32 persistent reservation registration keys for each LUN.
- 512 simultaneous commands for each iSCSI port.
- 2048 simultaneous commands for each FC port.

Pillar Axiom Path Manager

A host system communicates with LUNs through a normal SCSI-over-Fibre Channel (FC) or Internet Small Computer System Interface (iSCSI) initiator driver. Typically, these drivers are provided with the operating system or the HBA manufacturer.

For many operating systems, we supply an optional component called the Pillar Axiom Path Manager to run on the host. The Pillar Axiom Path Manager performs many key functions:

- Enables the use of multiple paths between HBA ports on the host and Slammer ports.
- Balances the load across HBA channels and Slammer ports.
- Monitors path failure and transfers traffic to different paths if a failure occurs.
- Maps LUNs to host volumes.
- Automates the recognition and configuration of the host in the Pillar Axiom Storage Services Manager.
- Reports the status of host drivers and paths in the Pillar Axiom Storage Services Manager.
- Collects logs to include in Call-Home files when SAN hosts is selected.

LUN Configuration

SAN Slammers support both LUN mapping and LUN masking. LUN mapping presents a LUN to a restricted set of hosts as any LUN number. LUN masking can restrict host access to only a defined set of Slammer ports. Initiator access to LUNs is controlled through the initiator's World Wide Port Name (WWPN) or Internet Small Computer System Interface (iSCSI) initiator name. A maximum of 255 LUNs may be exposed to a single host.

Individual LUNs can have different Quality of Service (QoS) settings (such as performance, redundancy, I/O bias, and striping). Multiple LUNs can be mapped to multiple different hosts and multiple hosts can be mapped to an individual LUN, regardless of their respective QoS settings.

About iSNS

The Internet Storage Name Service (iSNS) provides an alternative to iSCSI Software Initiator discovery by establishing an iSNS server on your iSCSI network that automatically discovers iSCSI targets and initiators.

iSNS facilitates automated discovery, management, and configuration of iSCSI devices on a TCP/IP network. iSNS provides intelligent storage discovery and management services comparable to those found in Fibre Channel networks, allowing a commodity IP network to function in a capacity similar to that of a storage area network.

The iSNS feature expects all Pillar Axiom iSCSI ports to have access to the same primary iSNS server. This rule is necessary so that all iSCSI ports can expect the same result when querying the iSNS database for the set of initiators that are members of the Pillar Axiom Discovery Domain Set.

If a Pillar Axiom iSCSI port never has access or loses access to the iSNS server, the Pillar Axiom storage system reports iSNS error events in the Pillar Axiom Storage Services Manager but continues to operate normally. At least one Pillar Axiom iSCSI port must have access to the iSNS server during system startup for iSNS-based initiator access control to function correctly; otherwise, all iSCSI logins are rejected.

Refer to Request for Comments (RFC) 4171 (http://www.rfc-archive.org/getrfc.php?rfc=4171) for complete iSNS information.

CHAPTER 6

Replication Overview

About Replication

The Pillar Axiom system provides Pillar Axiom replication for both network attached storage (NAS) and storage area network (SAN) environments. This replication feature provides efficient operational or disaster recovery (DR) capabilities that improve business data recovery while allowing you to adhere to regulatory or compliance requirements without complex configuration.

The Pillar Axiom replication feature supports short recovery time objectives (RTOs) by making the replicated volumes immediately available, allowing client applications to quickly switch over to and begin accessing the replicated data. When the disaster has passed and the original Pillar Axiom system has been repaired, you can efficiently and quickly restore the original volumes on the repaired system.

The Pillar Axiom replication feature uses block-level operations to transfer data during replication operations over local and wide area networks. Pillar Axiom systems help guard against security exposures by requiring passwords for all replication sessions. If you need security on both ends of the communication link between two Pillar Axiom systems engaged in data transfer operations, you will need to install the appropriate networking hardware to encrypt the data and otherwise secure that link.

Pillar Axiom MaxRep Replication for NAS provides data protection for filesystems. A source filesystem can be replicated to multiple different targets, if desired. Depending on your recovery point objectives (RPOs), synchronizing the source filesystem with any or all of the target filesystems can be performed quickly.

AxiomONE Replication for SAN provides data protection for LUNs. A source LUN can be replicated to a destination on another Pillar Axiom system (local or remote). The destination LUN is updated asynchronously, and checkpoints are taken to ensure that the source and destination LUNs are consistent at any point in time. The replication process can use either Fibre Channel (FC) or Internet Small Computer System Interface (iSCSI) interfaces.

Pillar Axiom Replication for NAS

Pillar Axiom MaxRep Replication for NAS is a native command-line utility that provides data replication among Pillar Axiom systems in a network attached storage (NAS) environment.

Typical usage of this software might transpire in this way:

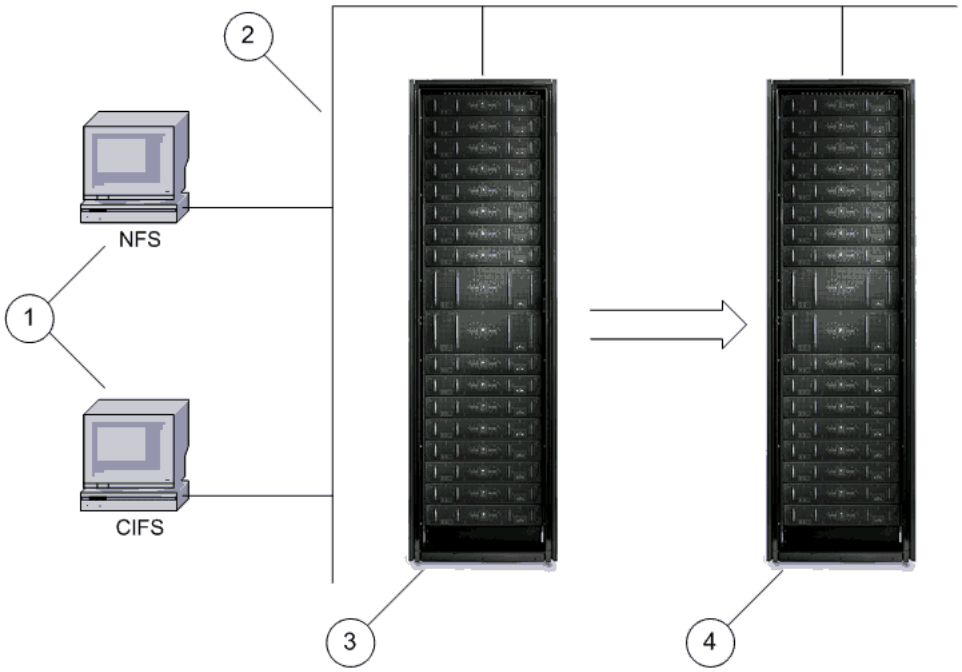
- Establish a relationship between two compatible filesystems. In this relationship, one filesystem is the source of replication and the other filesystem is the target.
- Fully synchronize the content of the two filesystems.
- Perform an incremental synchronization some number of times.
- Break the relationship.
- Mark the target filesystem as being live (readable and writable).
- Use the target filesystem in the same way as any other filesystem.

Note: All replication configuration and management is performed on the command line.

Using block-based read and write operations to perform the replication, the Pillar Axiom MaxRep Replication for NAS utility simulates a backup operation on the source filesystem and a restore operation on the target filesystem. (See the following figure.) The source and target filesystems can reside on the same Pillar Axiom system or on two separate systems that are connected over a local or wide area network.

Replication periodically saves specific information that allows replication to restart as needed. This information is called a *restart point*. For example, if, while the Pillar Axiom software is being updated, the system warmstarts or restarts during the software update, replication uses the restart point to continue the replication operations after the system returns to a Normal state.

Figure 11 System hardware in a NAS replication environment



Legend		
	1 Client hosts	3 Pillar Axiom system A (source)
	2 TCP/IP network	4 Pillar Axiom system B (target)

The Pillar Axiom MaxRep Replication for NAS utility includes client software for host machines that use either Common Internet File System (CIFS) or Network File System (NFS) protocols to communicate with Pillar Axiom servers. You can use the client software for initiating replication operations and for otherwise managing replication objects. This utility allows you to:

- Discover information that is essential to replication operations (such as determining the filesystems that are available for replication).
- Establish a relationship between two filesystems.
- Synchronize two filesystems that have been defined as a replication pair. The synchronization is a controlled, highly organized transfer of data from the source filesystem to the target.
- Place the target filesystem into a *live* mode to make it available for read and write operations by users.

  If the source filesystem fails, the administrator can change the state of the target filesystem from its special passive mode to an active mode that supports user I/O.

To support replication, the system administrator must enable Network Data Management Protocol (NDMP) and configure an NDMP File Server on both the source and the target Pillar Axiom systems. Ideally, this File Server should be dedicated to replication and other NDMP operations and not used for regular Common Internet File System (CIFS) and Network File System (NFS) user traffic.

For performance reasons, an NDMP File Server should have multiple virtual interfaces (VIFs) associated with it, one for each filesystem participating in replication. (A VIF configuration defines a physical network port, including an IP address, on a Slammer control unit (CU). During replication, data transfer occurs through these ports.)

Because a filesystem is homed on a particular Slammer CU, the administrator should consider configuring one NDMP VIF for each NAS Slammer CU that is home to a filesystem that participates in replication. If an NDMP VIF is not defined, the replication software uses an existing VIF that is associated with the filesystem, possibly competing with regular NFS and CIFS user traffic.

NAS replication uses only those VIFs that have been defined for the NDMP File Server. Furthermore, the utility automatically selects the VIFs that are most optimal, which could result in a different VIF for each replication connection.

Note: NDMP traffic and replication traffic share the same File Server and VIFs.

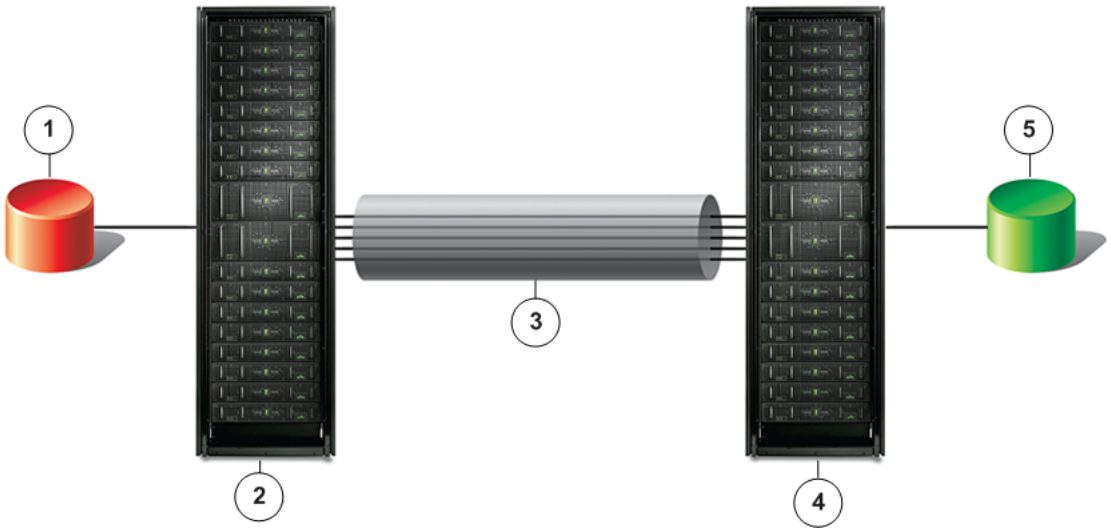
For information on configuring NDMP, File Servers, and VIFs in a Pillar Axiom system, refer to the *Pillar Axiom Administrator's Guide*.

AxiomONE Replication for SAN

AxiomONE Replication for SAN is a utility that enables you to automatically replicate and restore Pillar Axiom system data in a storage area network (SAN) environment.

In SAN Replication, one or more parallel volumes called *replication pairs* are established at primary and secondary locations, typically on separate remotely distributed Pillar Axiom systems. Communication links are established between these primary and secondary locations to maintain consistency of the data on the volumes at both locations. The transfer of data takes place automatically as checkpoint Clone LUNs of the data on the source volume at the primary site are taken at a specified interval, and these clones are duplicated on the destination volume at the secondary site. The replication pair updates continuously as long as the integrity of both volumes and the communication link between the locations are maintained.

Figure 12 Replication pair elements



Legend		
	1 Source volume LUN	4 Pillar Axiom system at the secondary site
	2 Pillar Axiom system at the primary site	5 Destination volume LUN
	3 Communication channel between the primary site and the secondary site	

If a source or destination volume fails, or if the communication link between the primary and secondary sites is interrupted, replication stops until the problem is corrected. When normal operations are resumed, replication resumes automatically.

AxiomONE Replication for SAN includes client software that you can install on Linux or Windows host machines. The client software provides a command-line interface for managing replication. You can use this client software to:

- Initiate replication by starting replication pairs.
- Check the status of replication pairs.
- Isolate replication pairs. This stops the replication process.
- Resume replication after isolation.
- Perform a site reversal to change the direction of replication, if needed.

Data is transferred over communication channels between Fibre Channel (FC) or Internet Small Computer System Interface (iSCSI) ports on the primary and secondary Pillar Axiom systems. Each replication pair is defined in a

configuration file that specifies the primary and secondary sites for the pair. The configuration file also defines the communication channels to be used by the pair.

Data Protection Services

A Pillar Axiom system provides the following forms of replication, each with its own distinct purpose:

- Snap FS—a read-only filesystem-based snapshot.
- Clone FS—a read-write filesystem-based snapshot.
- Clone LUN—a read-write LUN-based snapshot.
- Volume Copy—a read-write full volume replica.

Partial-image snapshots (Clone FS and Clone LUN) require storage space allocated for clones. This space is allocated at the time the volume is created. Only changes made to the source volume or the clone are stored on the system in the storage space allocated for clones.

Partial-image snapshots are recommended for short-term backups when you expect only a moderate amount of data to change.

Full-image snapshots (Volume Copies) are the same size as the source volume. The system stores a complete copy of the source volume from the time the snapshot was taken. The data is copied from the source to the snapshot in the background and the full-image snapshot automatically becomes a regular standalone volume once all of the data has been copied.

Full-image snapshots are recommended when a large amount of data will be changing or if you know that the snapshot will be needed as a standalone volume.

Data Replicas and System Capacity

You can create online data replicas in different ways. Each method consumes the capacity in the storage array differently.

The Pillar Axiom system ensures that all logical volumes (filesystems and LUNs) that are associated with a particular replica tree reside on[1] the same Slammer control unit (CU).[2] If you change the home of any of these logical volumes, the system changes all of them. This feature applies to all of the following objects:

- Snap FSs
- Clone FSs
- Clone LUNs
- Volume Copies
- Active data migrations because of Quality of Service (QoS) changes

Volume Copies and logical volumes being migrated due to QoS changes are present in the original replica tree until they are detached. Once the Volume Copy or the migrated volume is detached, the volume is removed from the original replica tree and becomes the root of a new replica tree.

However, after you start a Volume Copy operation or the system starts a data migration operation, if you re-home anything in the replica tree, the mechanics are a little different. If the system has not yet detached the copy from its source volume, the copy will be re-homed. If, however, the system has already detached the copy, the copy is no longer in the original replica tree and, so, is not re-homed.

Table 6 Capacity usage by online data replicas

Method	Description	Capacity usage
Clone FS	Creates a readable and writable point-in-time copy of a filesystem.	Consumes system space allocated for clones. Only changes to the source or clone are stored.
Snap FS	Creates a read-only copy of a filesystem that users can access	Consumes part of the capacity of the parent filesystem.

[1] Sometimes the term *homed on* or *owned by* is used instead of *reside on*.

[2] This discussion of replica trees on a Slammer CU does not apply to replicated objects created by the Pillar Axiom MaxRep Replication for NAS and Pillar Axiom MaxRep Replication for SAN utilities.

Table 6 Capacity usage by online data replicas (continued)

Method	Description	Capacity usage
	through special directories in the filesystem hierarchy.	
Clone LUN	Creates a readable and writable point-in-time snapshot of a LUN.	Consumes system space allocated for clones. Only changes to the source or clone are stored.
Volume Copy	Creates a block-level, full-image, read-write copy of a logical volume (filesystem or LUN). QoS attributes for a Volume Copy can differ from the QoS attributes of the original.	Consumes free space from system capacity that is equal to the current size of the volume.

The online data replicas identified in the preceding table have the following characteristics:

- They require no prior configuration (other than the initial allocation).
- They are created by explicit one-time operations.
- They are created on the same Pillar Axiom system as the source volume.
- Updates to the source volume are not reflected in the replica. When data changes in the source volume, that change *is not* reflected in the replica.

In comparison, data replicas created by the Pillar Axiom Replication utilities for network attached storage (NAS) and storage area network (SAN) environments are summarized in the following table.

Table 7 Capacity usage by remote data replicas

Method	Description	Capacity usage
NAS-based remote replica	Creates a read-only copy (snapshot) of a source filesystem on a second (or same) Pillar Axiom system. The content of this replica is manually synchronized with the paired source volume. For disaster recovery purposes, a system administrator can make this replica live so applications can continue reading and writing.	Consumes free space from the target filesystem on the target system.

Table 7 Capacity usage by remote data replicas (continued)

Method	Description	Capacity usage
SAN-based remote replica	Creates a clone of a LUN on a second Pillar Axiom system. The content of this replica is automatically synchronized with the paired source volume. A system administrator can use this replica for disaster recovery purposes. The remote replica LUN is normally not accessible. However, when the replica is isolated (replication stopped), a system administrator can map it, which makes it visible to client systems. An isolated replica is readable and writeable.	Consumes system space allocated for Clone LUNs on both systems. Only the changes to the source are stored on the target system.

Remote replicas have the following characteristics:

- SAN-based replicas are configured before they are created.
- NAS and SAN-based replicas are paired with a source volume.
- NAS and SAN-based replicas are created on a remote Pillar Axiom system.
- Updates to the source volume are reflected in the replica. When data changes in the source volume, that change *is* reflected in the replica the next time a synchronization operation takes place.

Note: A Pillar Axiom system uses binary units to calculate and display the capacity of physical storage and the size of logical volumes:

1 MB = 1024^2 (1,048,576) bytes

1 GB = 1024^3 (1,073,741,824) bytes

1 TB = 1024^4 (1,099,511,627,776) bytes

Snap FS

Snap FS is a filesystem snapshot that serves primarily as a file recovery mechanism. Snapshots are point-in-time copies that use the active filesystem as a base. A Snap FS preserves a view of the data at the exact time of the snapshot, allowing users to access older versions of files from a hidden subdirectory within the primary filesystem. Snap FSs are read-only.

A common reason for making periodic Snap FSs is to allow ordinary users to recover files that have been accidentally deleted without any assistance from the Pillar Axiom system administrator.

Another snapshot scenario is when a system administrator changes a software program and then finds the changes do not work. A snapshot allows the administrator to return to a previous version of the program and protects against inadvertent deletion of files.

The Pillar Axiom Storage Services Manager (the GUI), allows you to build and manage snapshot (Snap FS) schedules. A typical approach is to perform snapshots hourly and retain those snapshots for a day or a week.

Clone FS

Clone FS is a point-in-time, read-write copy of a filesystem that you intend to snap (split) from the source filesystem for immediate use. You can create an immediate snapshot of a filesystem. If a filesystem requires a filesystem check (FSCK), you can clone the filesystem, FSCK the clone, and then synchronize the two.

Clone FSs utilize partial-block snapshot technology. Because Clone FSs point to the original data, and have the same QoS parameters as the source filesystem, you cannot change the Quality of Service (QoS) attributes or the Clone FS location on the Slammer control unit (CU). Only data that has been modified on either the source filesystem or the Clone FS is copied. This means that Clone FSs can use significantly smaller storage space than full volume copies.

Clone LUN

Clone LUN is a point-in-time copy of a LUN that you intend to snap (split) from the source LUN for immediate read-write use. You can create an immediate snapshot of a LUN. Clone LUNs utilize partial-block snapshot technology.

Clone LUNs point to the original data and are created using the same QoS parameters as the source LUN. You can, however, change the priority level Quality of Service (QoS) attribute after you create the Clone LUN. Only data that has been modified on either the source LUN or the Clone LUN is stored in the clone storage space. This means that Clone LUNs can use significantly smaller storage space than full volume copies.

Volume Copy

A full-volume copy is a point-in-time, block-for-block replication of the source volume (filesystem or LUN). You can use the copy for backup, testing, reporting, or data warehousing. These workloads are directed to the copied volume, not to the source. Thus, the primary volume can service application I/Os with minimal performance degradation from ancillary activities. Volume Copies have read and write capabilities.

The Volume Copy is made to a new volume which can have its own QoS metrics. This allows system resources to be maximized for the task at hand. For example, a replicated volume that is used for reporting is assigned a lower performance priority and a higher read-centric access pattern than would the source volume.

Clone Storage Space

Clones consume space on the system. This is referred to as the allocated clone storage space or clone repository. The system allocates space for clones during volume creation. When you create clones, sufficient space for the clones must be available on the system. The system stores only the changes made to either the source volume or the clone in the allocated clone storage space.

Reduced Clone LUN Overhead

Most snapshot implementations handle the necessary data manipulation as part of the write process. Pillar Axiom SAN systems handle the cloned data processing during the write de-stage process as data is flushed from protected storage to the physical drives. This technique enables a write request to complete more quickly. This implementation allows for the use of Clone LUNs with minimal degradation of system performance.

Pillar Axiom SnapDelta FS

The Pillar Axiom SnapDelta FS feature lists files that have been created, renamed, or deleted, or files with changed content, within a Pillar Axiom filesystem during the interval between two Snap FS snapshots.

External applications sometimes need to scan the contents of a Pillar Axiom filesystem to extract information. After an initial scan, these applications need to perform periodic rescans to process new, renamed, and deleted files, and files with changed content. Examples of these external applications include:

- File-based replication applications.
- Search and indexing applications.
- Information classification applications.
- Virus scanning applications.

The Pillar Axiom system provides a filesystem change reporting command, `axiom_snapdelta_fs`, to enable external applications to rescan Pillar Axiom filesystems. SnapDelta FS provides efficient access to the set of file changes created during the interval between two Snap FS snapshots.

For example, here is how an external application might use SnapDelta FS to process filesystem changes:

- First, an application that tracks changes to a filesystem creates a snapshot and performs a full scan of the contents of the filesystem in that snapshot.
- Later, the application creates a second snapshot and uses the `axiom_snapdelta_fs` command to find the changes between the first snapshot and the second snapshot.
- Next, the application creates a third snapshot and performs further tracking of filesystem changes by finding the changes between the second and third snapshot, and so on.

Filesystem change reporting works on two snapshots at a time. These snapshots may be immediate snapshots created specifically for the use of `axiom_snapdelta_fs`, or they may be scheduled snapshots created automatically for the filesystem.

Backup Tools

The Pillar Axiom system provides tools to facilitate backup. The following sections describe these tools.

NDMP-Based Backup System Configuration

Network Data Management Protocol (NDMP) is an industry-standard protocol that allows for the use of third-party backup applications to manage the backup and recovery of customer data. An NDMP user account, password, and access port are configured through the Pilot. Pillar Axiom systems support NDMP version 4. Refer to http://www.ndmp.org/info/faq.shtml for details.

NDMP-based backup and restore operations can be integrated into your existing backup and recovery system. When you do this, you can completely automate the backup and restore operations.

The Pillar Axiom storage system supports:

- CommVault Galaxy 6.1.
- Symantec Veritas NetBackup 6.0.
- EMC NetWorker 7.3.
- BakBone NetVault: BackUp 7.
- Oracle Secure Backup 10.1.0.3.

To automate backup tasks, create schedules for backup operations through your data management application (DMA). Refer to the documentation of your DMA for details. Pillar Axiom storage systems allow up to five distinct NDMP operations at the same time. That is, a total of five concurrent backups or restores is allowed. For example, two backups and three restores can be run concurrently, but not three backups and three restores.

Pillar Axiom storage systems allow up to 250 user-defined Snap FS snapshots for each filesystem. Because each block-level or file-level backup requires one free available filesystem snapshot to succeed, no more than 249 snapshots of a given filesystem can be used for a backup. If 250 or more snapshots exist in the filesystem prior to the backup, the backup will fail.

Important! Due to the sensitivity of tape libraries and tape backup, do not run administrative tasks on the Fibre Channel switch to which the tape device is attached while a backup or restore is in progress. This could cause timeouts or failures.

Microsoft Volume Shadow Copy Service (VSS)

Pillar Axiom systems can use the Microsoft Volume Shadow Copy Service (VSS) in their backup and restore solutions.

For a detailed description of Microsoft VSS refer to this article on the Microsoft Developers Network (MSDN): The VSS Model (Windows) (http://msdn.microsoft.com/en-us/library/aa384625.aspx).

Index