PAC Storage PS Family Best Practice Guide

This guide provides best practices for installing and setting up PAC Storage PS models to deliver high performance and all-time data availability.

Version: 2.0

Updated: April 2020

Contents

Summary

This document focuses on how to propose and configure PAC Storage's PS models to achieve maximum performance and high availability. In this document we cover best practices that may meet generic scenarios for redundant-controller models. We strongly recommend following these principles to deploy PAC Storage products. However, in a real-world IT environment that involves third-party applications, you are advised to take into consideration their respective best practices.

Audience

This guide is intended for the PAC Storage partners, customers and employees who are proposing and configuring the PAC Storage system. We assume the audience is familiar with PAC Storage PS storage usage.

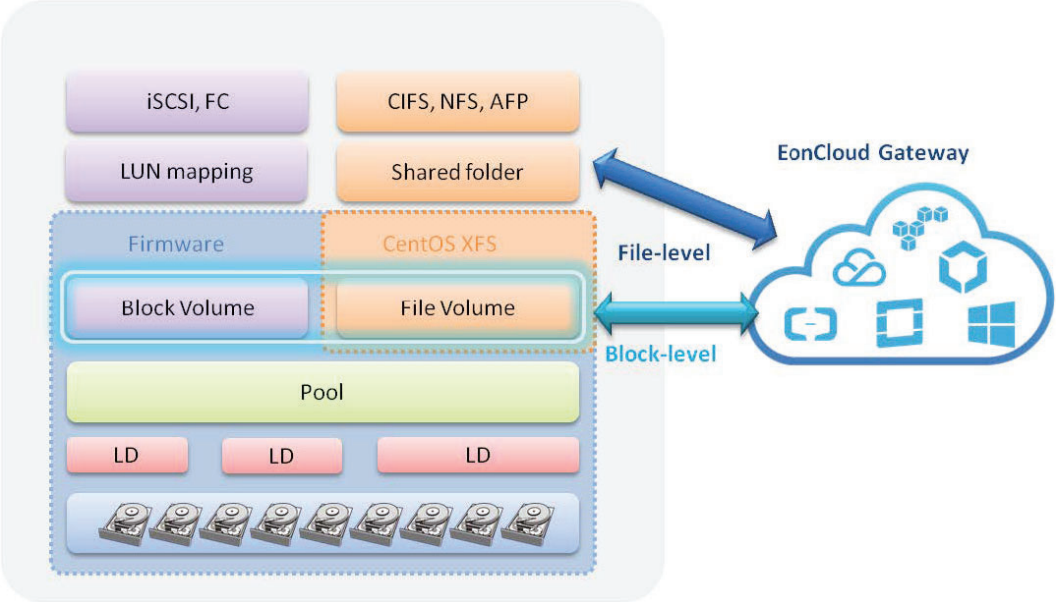
Terminologies

- **Storage controller –** A storage node with computing power to process I/O workload between storage and hosts. A controller comes with a processor, memory, cache backup module and network connectivity
- **Storage system** – An enclosure that includes controllers, fans, and power supplies
- **Expansion enclosure** – An enclosure providing extra storage space to a storage system
- **Asymmetric Logic Unit Access (ALUA) –**The paths from both controllers to any given LUN; one of them can process the active/optimized path
- **RAID (Redundant Array of Independent Disks)** – Storage virtualization technology that integrates multiple drives to provide high performance and data redundancy.
- **Logical Drive** – The basic unit of a RAID group created by a controller.
- **Pool** – An aggregate of logical drives that can host multiple volumes
- **Block-level volume** – A block-level storage space that is accessible via block-level protocols
- **File-level volume** – A storage space with file system that is accessible via file protocols
- **LUN (Logical Unit Number)** – A block-level storage space on the host side
- **File system –** A mechanism to control data storage and retrieval via file protocols such as CIFS, AFP, and FTP
- **Shared folder** – A location for shared storage access via a file protocol.
- **EonCloud Gateway** – A cloud application for backing up volumes and folders to the cloud

PAC Storage PS Architecture

PAC Storage PS series is an enterprise-class storage system which provides an efficient and cost-effective way to allocate storage resources to meet the capacity, availability, and performance requirements of different applications. Basically, users can create a pool on the basis of multiple logical drives (LDs) with different RAID protection levels and a certain number of drives. Moreover, these pools can consist of one or more types of hard drives, allowing for flexible deployment in various scenarios. A volume composed of multiple disk sectors can be created on a pool, and users can choose whether to enable the file system for it.

PAC Storage's proprietary firmware for PAC Storage PS allows creating a block-level or a file-level volume and therefore a storage pool of the same type. Furthermore, as a unified storage system, the PS series allows creating different types of volumes in the same pool at the same time. In addition, tthe PS series's channel Ethernet ports provide file transferring services (CIFS/SMB, NFS, FTP and AFP) and block data transferring services (iSCSI and Fibre Channel), making it a unified SAN and NAS system. Thus, users can enjoy system resiliency by switching the I/O type between block-level and file-level via the web-based management suite – EonOne.



PAC Storage PS architecture

In addition to being a unified SAN-NAS system, the PS series comes with a native cloud storage solution

EonCloud Gateway, an application that transfers data in local volumes and folders to the cloud for backup and restore purposes. This application can fit into IT environments utilizing public cloud services, such as **Amazon S3**, **Aliyun**, **Microsoft Azure**, **Google Cloud** and **OpenStack**.

Generic Best Practices

PAC Storage storage systems come in comprehensive product lines and with flexible modular design that provides accurate and efficient solutions to various needs. In this section we list some customer requirements information and illustrate solution design steps to meet those requirements.

Requirements

Product Families

- PAC Storage PS 3025A Gen2 series is an all-flash array and unified storage system that integrates file-level, block-level, and cloud-integrated features. It features lightning-speed performance and extreme low latency (up to 900K IOPS within 0.5 ms latency) to meet any high density workloads. Moreover, it runs on optimized firmware that is able to extend SSD lifespan with latency-sensitive applications.
- PAC Storage PS family is a hybrid storage system that consolidates file-level, block-level, and cloud-integrated features. This product is suitable for file sharing, especially with cloud services. In this family, there are also turbo models that offer better performance for file-level (NAS) applications.

Capacity

Capacity is one of the most important factors when you determine a RAID protection level. Generally, customers will provide the capacity requirements for raw capacity or usable capacity in advance. The main difference is that the relevant factors of usable capacity are the RAID protection and reserved space overhead, which is a critical issue for capacity-sensitive applications.

Performance

Currently, PAC Storage provides real-world test results of the following two performance tests, available in our official website.

- IOPS workloads by simulating small-block random access
- Bandwidth workloads by simulating large-block sequential access

These industry-standard simulation testing and performance results are used to evaluate the maximum performance with a specific configuration. However, the performance may vary depending on the workload type and the environment. To provide suitable recommendations that match customers' needs, PAC Storage needs to be informed of the type of servers and applications, read/write behavior ratio, data block size, etc.

Drive Type

Each type of hard drives has its unique performance, capacity and cost—a key element for achieving the expected workload.

- SSD drives: Suitable for IOPS-demand and latency-sensitive applications. Especially, the PAC Storage PS 3025A Gen2 all-flash array enjoys boosted performance with SSD drives, leading to outstanding performance results. Both SAS and SATA interface SSDs are supported. However, SSDs typically come at the highest cost among all disk types.
- 10K SAS drives: Suitable for applications with medium bandwidth workloads.
- NL-SAS drives: Suitable for large-capacity applications.
- SATA drives: Cost-effective and commonly used for SMB capacity-oriented. PAC Storage also provides a MUX board solution so that users can configure SATA disks on a redundant-controller model.

Front-end Connectivity

PAC Storage provides a modular host board that simplifies system maintenance and upgrades. Note that the front-end connectivity bandwidth may affect your system performance output. For example, 4 16Gb/s FC channels can provide up to 8GB/s performance throughput. Make sure that the front-end channels' bandwidth meet the demands of your application workload.

The PAC Storage PS series supports multiple protocols (NFS, FTP, and / or SMB) for file-level NAS applications. To achieve better performance, it is recommended to enable the jumbo frame feature.

RAID Protection Level

PAC Storage provides multiple RAID protection levels to choose from. Each RAID protection has its quinit performance, capacity and data critical levels.

- **RAID1** provides the best availability and IOPS performance for random write access with small block workloads. However, its capacity utilization is the lowest since half of the drives will be used for data protection. RAID1 is useful when read IOPS or reliability is considered more important than write performance or storage capacity.
- **RAID5** provides the highest usable capacity, but the lowest availability. Its performance is between RAID1 and RAID6 for general-purpose workload.
- **RAID6** provides more parity drives than RAID5. It provides medium-high protection and storage capacity, with the lowest overall performance.

PAC Storage recommends that you use RAID6 with hot spares for general-purpose workloads. RAID6 has higher availability than RAID5 and better capacity usage than RAID 1. In general, since PAC Storage storage systems can provide outstanding performance, setting the RAID protection level to RAID1 or RAID5 can still achieve same performance as RAID 6.

LUN & File System Number

For block-level volumes, Infortrend's controller assignment mechanism supports Asymmetric Logical Unit Access (ALUA) and Symmetric Active-Active for redundant-controller models. However, these mechanisms come with different levels of manageability and performance.

- If more than two storage pools are deployed, we recommend that you use ALUA to reach optimal performance and each pool be assigned to a different controller.
- The Symmetric Active-Active mode allows a single pool to be assigned to two controllers simultaneously to achieve high availability, and allows one LUN to use both two controllers instead of one. With Symmetric Active-Active, a single LUN's performance can reach 180% of performance in the ALUA mode.

- Generally, the host operating system supports disk striping to combine multiple volumes into a single LUN, for example, as the striped volume in Windows server 2012 R2. ALUA brings less performance impact than Symmetric Active-Active.

On the other hand, file-level volumes do not support Symmetric Active-Active pools. To optimize the overall system performance, it is recommended to create at least 2 pools with each pool assigned to a different controller. Moreover, you should assign each controller to the host server so that the front-end workload of both controllers can be balanced.

Rack Unit

IT administrators are concerned about their limited cabinet space for installing servers, storage devices, and other gadgets. Installing a rack unit may pose a problem to cabinet space use. You should consider the ratio of performance and capacity density of hard drives for efficient space use. For example, a SAS HDD has a higher performance/density ratio than a NL-SAS drive, and therefore deploying SAS HDDs requires less cabinet space.

Configuration Guidelines

In this section, we will provide detailed configuration guidelines based on customers' requirements in the previous sections. Propose appropriate solutions by following the principles.

Product Series

PAC Storage has complete product lines that satisfy different requirements on performance, maximum drive number, and memory cache capabilities. Choose an appropriate product family according to the required RAID level and I/O workload type with reference to our performance report. For example, suppose that the performance requirements is 4K random read 200K IOPS for block-level data. Based on the performance report you are advised to choose models belonging to the PS 2000 series.

Logical Drive (LD)

A storage pool can contain multiple logical drives (LDs) that can achieve stable performance and risk spreading. To optimize system deployment, we recommend that an LD on the 2U12 and 3U16 models should be configured with the same number of drives as the enclosure's bay number (up to 16 disks with a hot spare drive). For models with larger bay numbers, we recommend configuring an LD as a 12-drive unit with a hot spare drive included.

Form factor	Recommendation
2U12bay	11 disks per LD + 1 hot spare
3U16bay	15 disks per LD + 1 hot spare
2U24bay 4U24bay	11 disks per LD + 1 hot spare
2U25bay	12 disks per LD + 1 enclosure spare
2U12bay + 4U60bay expansion enclosure	11 disks per LD + 1 hot spare
4U60bay + 4U60bay expansion enclosure	14 disks per LD + 1 hot spare

- For redundant-controller models, it is recommended to set even number of LDs. Therefore, the LDs can be evenly distributed in even number of pools, which are then assigned to the two controllers to achieve load balancing.
- We recommend that applying the same configurations (e.g. the RAID level, disk number per LD and stripe size) to the LDs as the pool configurations.

Usable Capacity

The usable capacity is determined by RAID protection and reserved system overhead:

Usable capacity = Capacity per disk x (Disk number – RAID parity – Hot spare) x LD number x 0.9 (overhead)

Suppose that we use 32 6TB NL-SAS drives with 2 RAID-6 LDs and 2 hot spare drives. The usable capacity in this environment is 140TB.

The reserved space for data services is the space of a pool where it conserves data such as snapshot images, file system configurations, and metadata. It is recommended to reserve at least 30% of the pool's space for these data services for future needs. By default, the system sends notifications when the reserved space is less than 30% of the pool's capacity.

Performance

The minimum number of drives required for your storage system can be estimated via the PAC Storage performance report. The PAC Storage performance report includes different I/O workload simulation tests with multiple RAID levels. To meet your performance and capacity requirements, contact PAC Storage or its partners for more information.

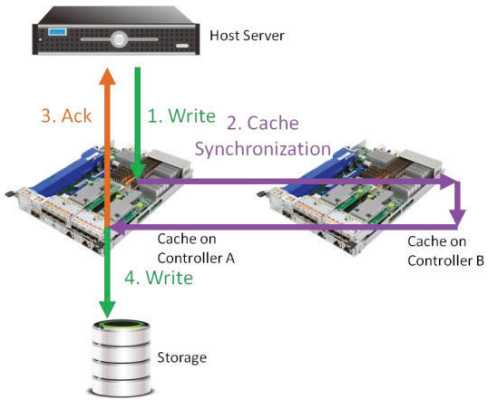
Form Factors

PAC Storage provides multiple form factors of storage devices and expansion enclosures. For space-saving purposes, PAC Storage offers 2U25 and 4U60 high-density expansion enclosures to achieve immense capacity in limited cabinet space. Suppose that the cabinet space is able to accommodate up to 64 NL-SAS drives with 6U rack units. You can easily meet the requirements with a 2U12 bay RAID device and a 4U60 high-density expansion enclosure.

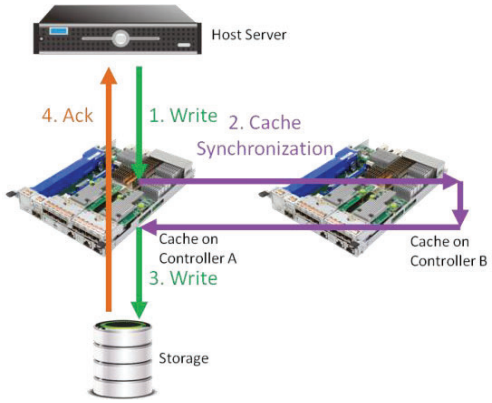
Controller Cache Memory

To achieve maximum performance, we suggest installing at least 16GB memory cache per controller (32GB in total). As more data services are enabled, the system may require more cache memory. We recommend that the cache memory be mounted symmetrically to the DIMM slots since a controller comes with dual channels. For example, an PAC Storage PS 3000 controller has 4 DIMM slots. The performance with dual 16GB DIMM memory is better than that of a single 32GB DIMM. However, the performance of dual 16GB DIMMs equals that of four 8GB DIMMs. Note that you should install the blue DIMM slots first, and that if you install memory of different sizes on a dual DIMM slot, the controller may not boot up properly.

PAC Storage's storage systems feature a mirrored write-back cache policy as the default write-cache policy. For each write request, the data is written to the write cache on the primary controller and then synchronized to that on the secondary controller. After the synchronization is complete, the host server is notified immediately after the data is written to the storage device. As the write cache data is always mirrored and protected between the redundant controllers, the policy has enabled the storage system to provide low latency and high throughput for write-intensive applications. However, users can still switch the write cache policy to write-through cache manually. Moreover, the system can also automatically switch from write-back to write-through cache based on the trigger settings.



The write-back cache mechanism



The write-through cache mechanism

First Installation

In this section we list the basic settings required for the first installation. For more information on advanced features and parameters, refer to the software user manual.

Basic Settings

(1) Access to EonOne

- Use default IP address

  The default IP address of the management interface is **10.10.1.1**. Connect the storage system to the server or PC via EonOne and join them to the same network domain. If there is a firewall in the environment, you need to open the corresponding port to allow access to the management port.

- Assign IP address via the console

  If you cannot access the management interface through the default IP address, go to the console interface (Baud Rate: 38400) via an RS-232 cable. The default login password is **admin**. You can operate via the Enter / Esc and direction keys.

(2) Update Firmware

It is recommended that you update the firmware to the latest version at the first installation. Before updating the firmware, you must turn off the storage system and press the **reset controller** button to ensure that the upgrade process is complete.

(3) System Time

It is recommended to complete the system time settings upon system initialization to obtain time-accurate **event logs**. Accurate time settings also matter for features like utilizing cloud storage, joining an AD domain, and running scheduled tasks.

(4) Licenses

If you have purchased an advanced license, activate it at first installation. A standard license is free for all users and preloaded on your devices.

(5) Channel Modes and Network Services

The network channels default to the block-level mode. You can switch them between the block-level and the file-level modes to suit your applications. Then, restart the system to effect the change. When the file-level mode is in use, remember to enable relevant file protocols on the management interface to allow smooth file sharing and management.

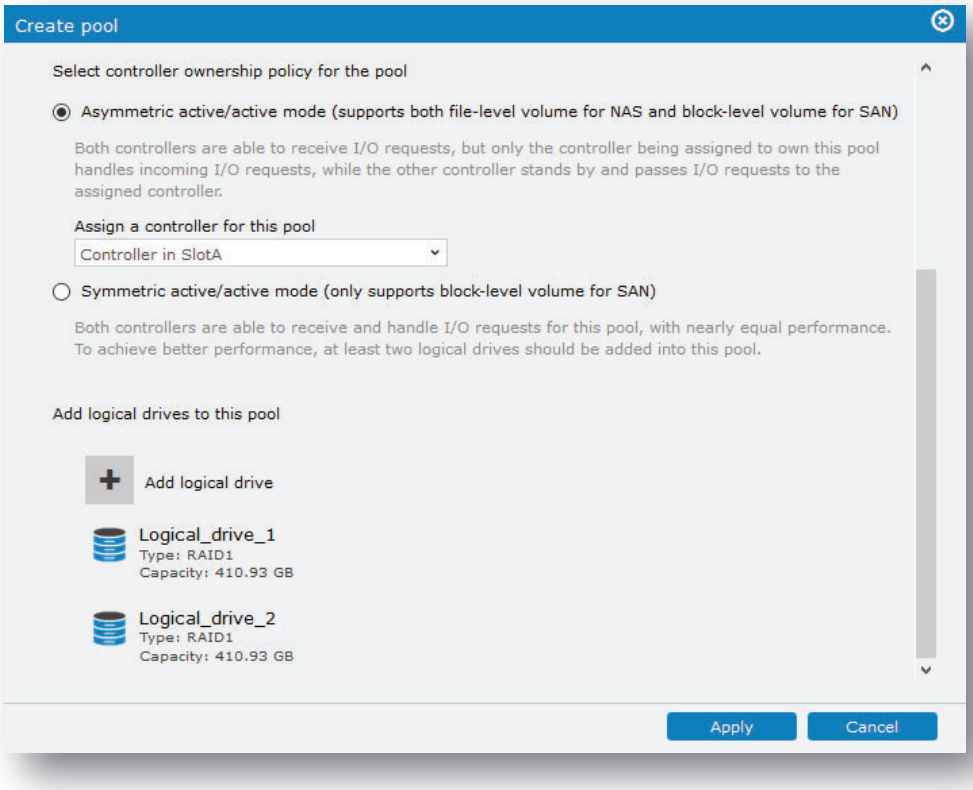
(6) Exporting System Configurations

Export the system configurations for backup after all the system settings are completed. To recover previous system settings after replacing a hardware component or resetting the system to default, import the configuration file back to the system.

Pool Configuration

(1) Create a Pool

Create your first pool based on the deployment plan. Add desired logical drives to the pool and set the stripe size to meet your I/O requirements. The stripe size is 128K by default for general purposes and you may change it to meet your purposes. For example, for an SPC-2 benchmark tool, it is ideal to set the stripe size to 512K; for Microsoft Exchange and Autodesk solutions, it is recommended that you set the stripe size to 256K.



(2) Add Logical Drives to a Pool

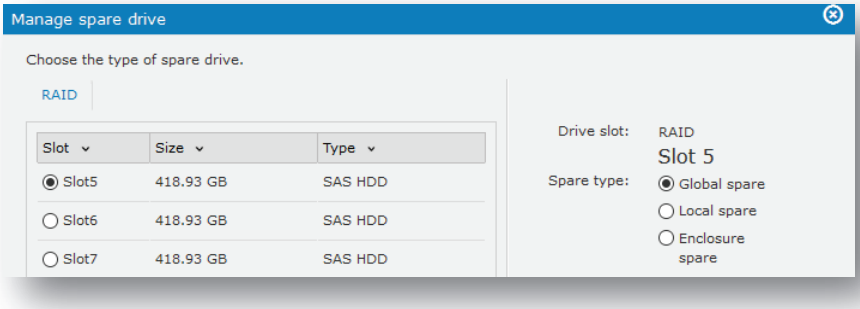
- Add more than one logical drive to the pool. It is recommended to apply the same configurations (e.g. drive number, RAID level and the stripe size) to all logical drives in the pool, which allows for easy storage expansion in the future.

- Run automated storage tiering when you create a hybrid pool that includes logical drives built from SSDs and HDDs (SAS or NL-SAS drives). Automated storage tiering allows a hybrid pool to deliver high performance by utilizing the SSDs' speedy performance and the HDD's large, cost-effective storage capacity. For more information, refer to the application note.

(3) Add a Spare Drive

Spare drives come in three types. A local spare drive is part of a logical drive and it can only replace a member drive in the same logical drive, in the same enclosure or not. A global spare drive does not function for a specific logical drive; it can replace any hard drive in the storage system. Note that rebuilding a logical drive with a local spare drive is more efficient than using a global spare drive.

In general, we recommend that you set a local spare drive for each logical drive, and install a few global spare drives in the system to achieve system-wide protection. In addition, the spare drive and the associated logical drive should match in drive type; that is, an SSD spare drive should be assigned to a logical drive built from SSDs, while an HDD spare drive should be assigned to an HDD-built logical drive.

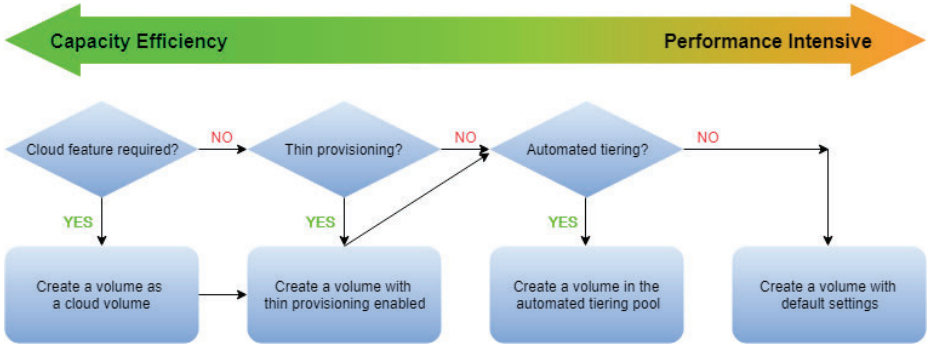


Volume Consideration

Volume Types

The storage system supports block-level and file-level volumes to meet I/O workloads demands. You can enable extra functionality for these two types of volumes:

- Thin provisioning: This functionality allows allocating a large amount of virtual capacity for an application server, regardless of the actual available physical capacity, which renders space usage flexible. Note that, when compared with a full-provisioning volume, a thin-provisioning volume may have lower performance.
- EonCloud Gateway: This EonOne-native application connects a volume to the cloud and manages data traffic between the two sites. Note that thin provisioning is also enabled when this application is in use.

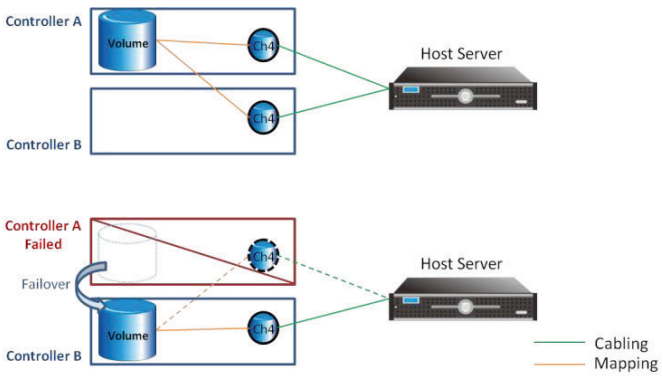


Decision flow for volume creation

Mapping & Cabling for Block-Level Volumes

We strongly recommend that you connect the front-end connection symmetrically to both controllers to ensure high availability and prevent single point of failure in the event of a controller or connection failure. For some data services, symmetric connection of front-end on both controllers is required. For example, a volume assigned to the primary controller (controller A) is LUN-mapped to the channel 4 on both controllers, and these two ports are connected to a host with MPIO enabled.

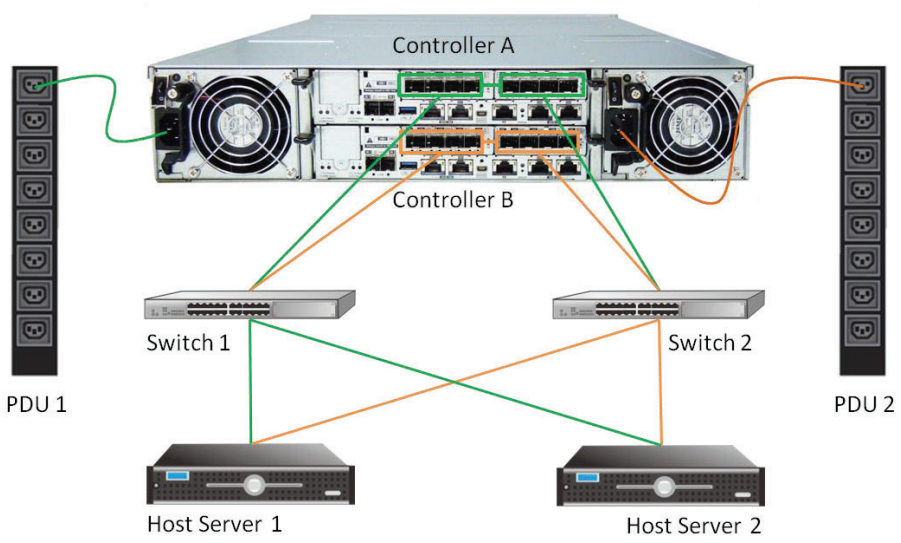
- **MPIO mode:** If the MPIO is in the round robin mode, the IO workload will be balanced via these two paths. Otherwise, the IO workload will only go through the channel of the controller in charge of the volume.
- **Controller failover:** Pools assigned to controller A are re-assigned to controller B when controller A fails. Thus, the IO workload can still be available via the channels on controller B. The failover time for ALUA or Symmetric A-A volumes is approximately 30 to 40 seconds.
- **Channel failover:** The controller monitors whether its channel ports are disconnected. If the channel port of the primary path fails, the ***IO workload will fail over to the passive path via the other controller.***



High Availability Configurations for Block-Level Volumes

To achieve no single point of failure, HA configurations are required. In the following example we leverage the advantage of the PAC Storage PS series redundant-controller model. We recommend that you map the volume to at least two channels of a controller with MPIO enabled. For redundant-controller models, there are at least 4 available paths in this case. You should cross the cabling by setting dual switches between the storage and host servers to avoid a single path or network failure. In the event of a switch or cable failure, the host servers can still allow access via the remaining paths.

Availability	**Solution**
Controller failover / Cable failover for a single channel on both controllers	Symmetrically map the volume to the channel on both controllers (A & B) with MPIO enabled on the host server. Symmetrically connect the cable on both controller ports on the different switches (switch 1 & switch 2).
Cable availability within a controller	Map the volume to more than one channel of a controller with MPIO enabled and connect the cable on the controller ports to different switches.
Power availability	Plug redundant PSU to separate PDUs



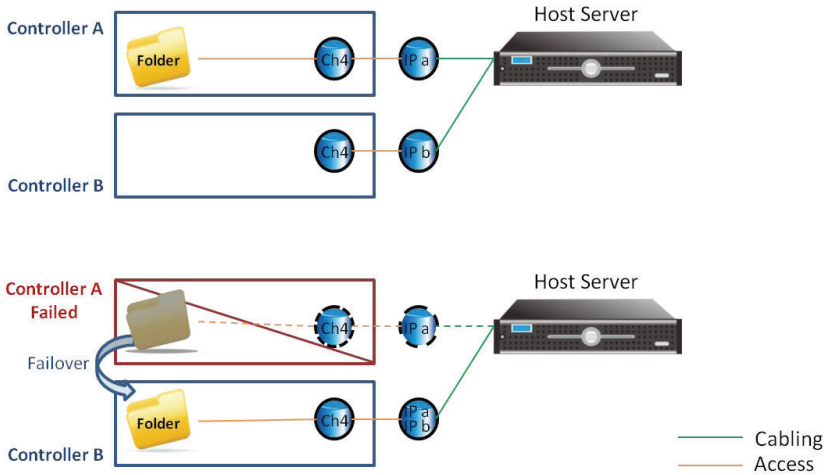
Block-Level HA Configuration

Cabling & Trunking for File-Level Volumes

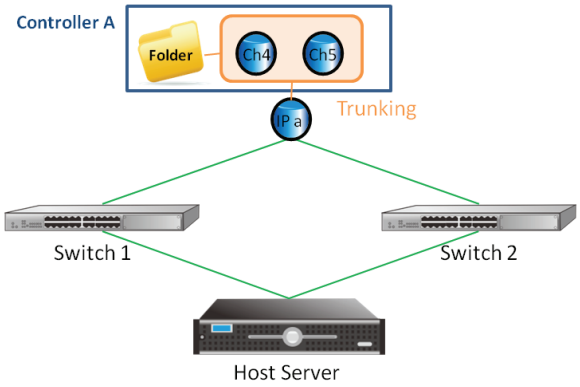
For file-level volumes, we strongly recommend that the front-end connection should go symmetrically to both controllers to implement high availability as well as load-balancing in case a controller or channel fails. For some data services, symmetric front-end connection to both controllers is required.

Generally, a shared folder only allows access via the channels of the controller that holds the belonging pool. Suppose that a shared folder exists in a storage pool assigned to controller A. In this way, the shared folder can only be accessed via the channels of controller A.

- **Controller failover:** If controller A fails, the pool holding the shared folder is then reassigned to controller B. Moreover, controller B is also in charge of controller A's network channels: these channels' original IP addresses remain available to allow access to the shared folder.



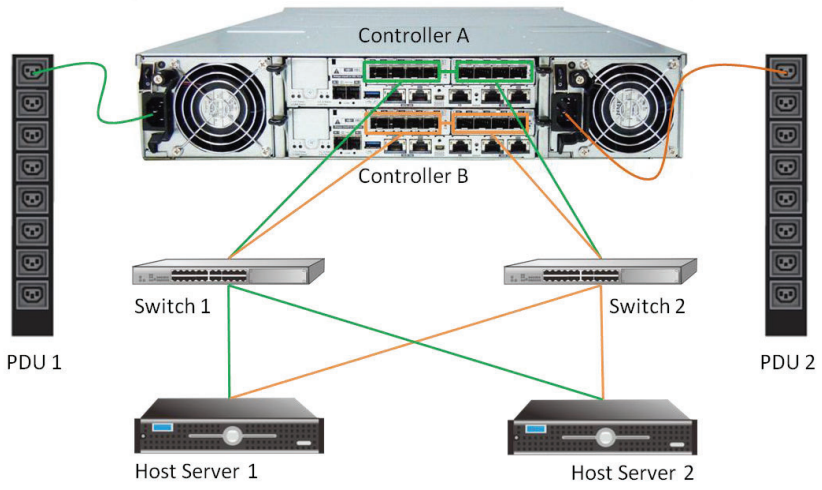
- **Cabling failover:** To avoid single point of failure, we recommend that you deploy more than one cable connection through a single controller and configure trunk groups across multiple ports. For the switch connection, connect the cable to a redundant switch to achieve high availability.



High Availability Configuration for File Sharing

We recommend that you connect the cables to the channels with same port number on both controllers symmetrically to the switch. When a controller fails, the other controller takes over the workload since it is connected to the same switch. Furthermore, we connect channels of a different controller to a different switch with IP trunking (link aggregation) so that the storage remain accessible to the surviving cables if one of the channels fails. To implement high availability, we use dual switches between the storage and host servers to avoid single path or network failure. In case of switch or cable failure, the host servers remain accessible via the remaining path.

Availability	Solution
Controller failover	Symmetrically connect the cables to the channel on both controller (A & B). In case a controller fails, the other controller takes over the failed controller's network channels and their IP addresses.
Cable availability on a controller	Connect at least two channels of a controller to a different switch with trunk group enabled. Thus, the storage system can still be accessed via the same IP address when one of the cables fails.
Power availability	Plug redundant PSU to separate PDUs



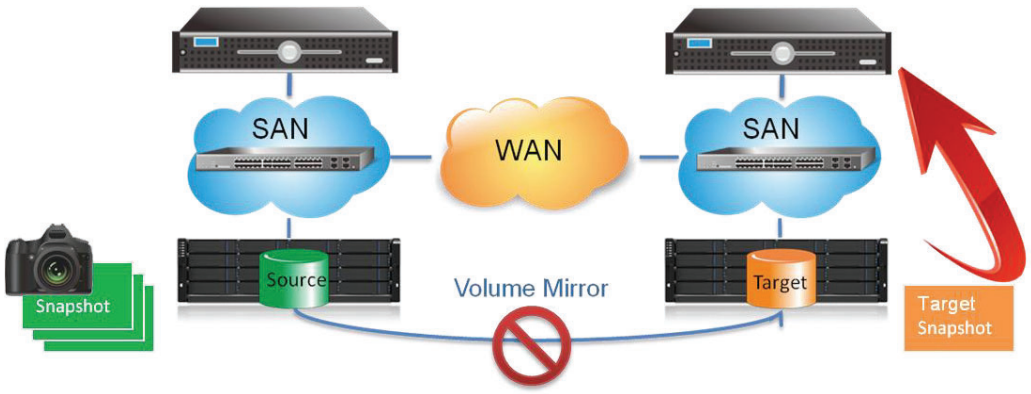
File-Level HA Configuration

Replication

Disaster Recovery Solution

For critical applications like database, replicating data to a remote site is a practical disaster recovery solution. We recommend that you set up a scheduled task to mirror volumes that hold important data, synchronously or asynchronously, to a remote storage to achieve data availability with up-to-date data backup.

Synchronous and asynchronous volume mirroring have their own characteristics. Synchronous volume mirroring requires higher network bandwidth and is suitable for critical data. The host writes data simultaneously to both the source and the target. In this light, synchronous volume mirroring has the best recovery point object but has longer response time. On the other hand, asynchronous volume mirroring has shorter I/O latency by a different snapshot from the source to the target at task creation and at later synchronizations.

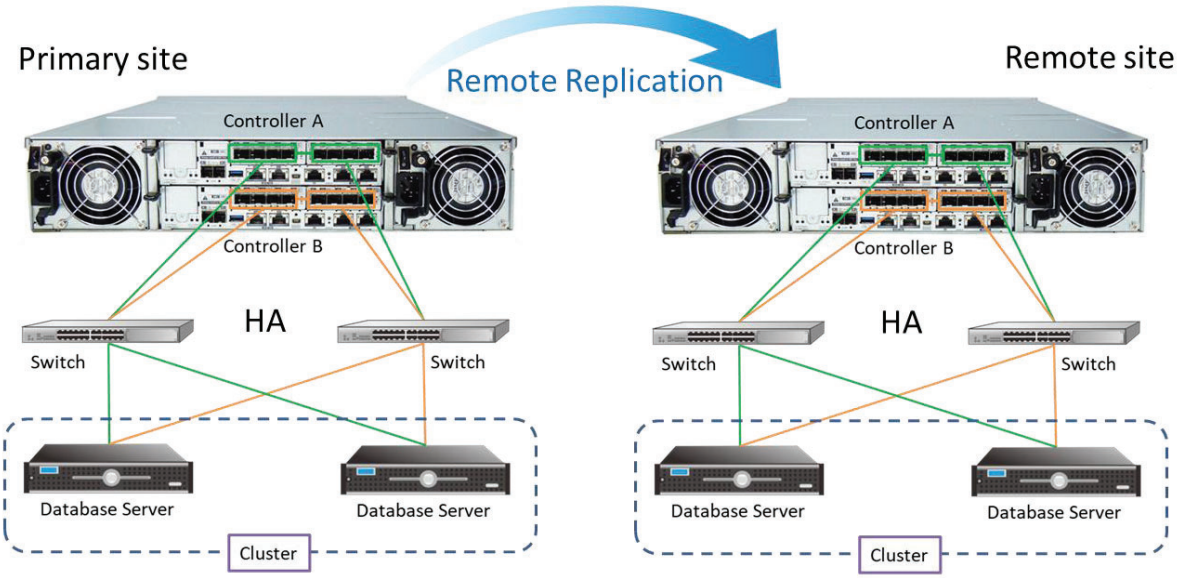


Disaster Recovery with PAC Storage Remote Replication

Example

Disaster Recovery for HA Database Architecture

Database is a critical and latency-intensive application. Asynchronous volume mirroring is suitable as a disaster recovery solution for databases. Both the local and remote sites can join a high availability architecture built with clustered host servers and a redundant storage system. To minimize system downtime if it happens, the storage should be capable of providing failover function for the database cluster and require a remote storage to store its data backup.



Disaster Recovery Solution for Database HA Architecture

Suppose a possible IT scenario that requires 40 TB storage space with automated storage tiering for SQL servers and 100 TB storage space for file servers. Moreover, it requires 1 TB SSD-built cache pool to accelerate the read performance. The requirements can be addressed using a 4U 60-bay storage model with hardware configuration proposed below:

Purpose	Requirement	Pool	Assignment	Logical Drive (LD)	Drive Number	Estimated Capacity
SQL Server	40TB	Pool 1	Controller A	Tier0: SSD RAID1	8 x 400GB SSD	48TB
				Tier1: NL-SAS RAID5	14 x 4TB NL-SAS	
File Sever	100TB	Pool 2	Controller B	LD1: NL-SAS RAID5	15 x 4TB NL-SAS	100TB
				LD2: NL-SAS RAID5	15 x 4TB NL-SAS	
SSD Cache	1TB	SSD Cache Pool	Controller A&B	N/A	4 x 400GB SSD	1.6TB SSD Cache
Global Spare	N/A	N/A	N/A	N/A	2 x 400GB SSD 2 x 4TB NL-SAS	N/A

Possible hardware configuration

Conclusion

By deploying the high availability configuration, users can enjoy data services on the PAC Storage PS series without the hassle of a single point of failure. To ensure data availability, the PAC Storage PS series provides various disaster recovery solutions such as volume-based remote replication and EonCloud Gateway. For NAS file sharing, file replication and EonCloud Gateway provide simple and yet reliable solutions to address the requirements.

Legal Disclaimer and Information

All PAC Storage products, including the product that customers have purchased from PAC Storage or its distributors, are subject to the latest Standard Warranty Policy available on the PAC Storage website.

PAC Storage may from time to time modify, update or upgrade the software, firmware or any accompanying user documentation without any prior notice. PAC Storage will provide access to these new software, firmware, or documentation releases from certain download sections of our website or through our service partners. Customer will be responsible for maintaining updated version of the software, firmware, or other documentation by downloading or obtaining from PAC Storage, and installing designated updated code, including but not limited to firmware, microcode, basic input/out system code, utility programs, device drivers, and diagnostics delivered with PAC Storage product.
Before installing any software, applications or components provided by a third party, customer should ensure that they are compatible and interoperable with PAC Storage product by checking in advance with PAC Storage. Customer is solely responsible for ensuring the compatibility and interoperability of the third party's products with PAC Storage product.

Customer is further solely responsible for ensuring its systems, software, and data are adequately backed up as a precaution against possible failures, alternation, or loss. For any questions of hardware/ software compatibility, and the update/ upgrade code, customer should contact PAC Storage sales representative or technical support for assistance.

To the extent permitted by applicable laws, PAC Storage will NOT be responsible for any interoperability or compatibility issues that may arise when (1) products, software, or options not certified and supported by PAC Storage are used; (2) configurations not certified and supported by PAC Storage are used; (3) parts intended for one system are installed in another system of different make or model.