MPS5300 Series Hybrid Storage Datasheet

Overview

The MPS5300 is a mid-to-high-end enterprise-level storage product developed by Maipu. It delivers ideal performance for structured data, unstructured data, and hybrid workloads, and has unique advantages in virtualized environments. Its scalability and performance are leading in the industry.

The MPS5300 is a hyper-converged unified storage product designed and developed for large-scale parallel data storage environments. The storage system supports three network environments, namely IP network, IB network, and FC network, between it and the business side. The storage controller can be configured with storage interfaces and access services such as FC-SAN, IP-SAN, IB-SAN, IP-NAS, and IB-NAS simultaneously, enabling multiple networking modes for a single device. This facilitates flexible networking in the early stage for customers, as well as the reuse of assets and network adjustment in the later stage. The MPS5300 adopts a fully symmetric distributed cluster architecture and can build a storage cluster with up to 256 nodes, providing a unified storage space externally, which is visible to all nodes. All nodes are fully peer-to-peer, achieving load balancing of tasks within the cluster and automatic failover, which greatly enhances the system performance and availability.

MPS5300 Hybrid Storage

.

Features

Product and Architecture Advancement

1. Enterprise-level Distributed Storage Structure:

It adopts the current mainstream distributed storage architecture. By deploying a distributed file system with a tight coupling between software and hardware and a design without a master node, it enables multiple nodes to concurrently provide data services and allows for centralized deployment in the local data center.

All nodes in the entire system provide a globally unique IP mount point externally, offering a single global file system view. Clients can access the entire storage space by mounting the global IP address, which simplifies the deployment process.

Performance can be expanded online by horizontally increasing the number of control nodes, and capacity can be expanded online by vertically adding disks. Compared with a distributed system built solely with servers, it has a greater capacity expansion range, higher cost - effectiveness, and simpler maintenance.

2. Separation of Computing and Storage, Centralized and Shared Data Storage:

The front - end servers of the system can be dedicated to providing services for applications and computing tasks, separating business computing from data storage. Massive data is centrally stored in a single system, enabling shared data access across the entire business process without data silos.

Compared with the integrated computing and storage approach (x86 servers + distributed file system), it has better data sharing capabilities. Moreover, data storage management does not consume the resources of application servers. Even if an application server fails, it does not affect the normal operation of the storage system. Due to the isolation between computing resources and storage resources, different business software can run without interfering with each other, which meets the requirements of data applications in the current complex environment.

3. Client - side Load Balancing, High Bandwidth:

The entire distributed storage system uses a client - side load - balancing mode, which distributes client I/O loads across all distributed storage control nodes, significantly improving the overall I/O performance output.

Compared with the traditional storage - side load - balancing method, it can provide higher aggregated bandwidth and single - stream bandwidth performance. Especially in a Windows environment where a large number of commercial image software is running, it greatly improves the processing efficiency of image software (the native SMB protocol does not support load balancing).

This load - balancing client supports multiple versions of Windows and Linux operating systems. It allows access to the storage system based on the client - side load - balancing mode and supports data access of mainstream personal office software such as Office.

4. Support for Multiple Services, Multiple Users, and Multiple Processes:

This solution effectively supports concurrent operations of multiple services and collaborative data processing among multiple users in complex environments that include hundreds of servers and workstations.

The storage system uses a standard POSIX file system interface, supporting mainstream operating systems, and mainstream image - processing software. It also supports all mainstream office software that can be accessed in the form of a personal network disk.

5. Efficient Storage Management:

The system has high reliability. It supports cache mirroring protection between nodes and power - failure data protection. It also supports the construction of active - active access between two sets of storage (a distributed system built on servers cannot achieve cache data protection).

By configuring a certain number of SSD disks on each node, they can serve as a secondary cache for the system to improve data reading efficiency or as metadata disks for efficient management of local node metadata. Traditional distributed file systems use a centralized metadata management method, which is prone to single - node performance bottlenecks. They also have poor management capabilities for massive small files and fragmented data caused by repeated deletion and writing, resulting in a slow - down over time.

Meanwhile, the system supports a comprehensive permission management function. Different data access permissions can be set according to the different user permission requirements in collaborative office scenarios, ensuring the security of key data from intrusion.

- **System Stability**

1. Write Cache Mirroring Protection

The two controllers in the local control cabinet synchronize write cache mirrors through the high - speed channel on the backplane of the control cabinet. When the host writes data to the write cache of one controller, the written data is automatically transferred to the write cache of the other controller through the PCI - E channel within the control cabinet. Only after the mirroring is successful will the system confirm the successful write operation to the host. This mechanism ensures the consistency of data between the two controllers during the data writing process and prevents data loss and business interruption caused by the failure of a single controller.

2. Cache Grouping

The storage system cache is divided into read cache and write cache, and can also be further divided into control cache and data cache. When the write cache is not in use, all the system caches can be used as read cache. The system reserves a minimum capacity for the read cache to ensure that the read operations still have reasonable cache resources available even under heavy write traffic.

When the occupancy rate of the system read cache reaches the threshold, the eviction algorithm calculates the popularity of data blocks based on the historical and current access frequencies and selects appropriate data for eviction.

Basic IO Processing Flow:

1) The host sends a write I/O request to a certain control node (assumed to be Node 1).
2) Node 1 writes the data to its local cache. Meanwhile, a mirrored cache is generated on another control node (Node 2) in the same control cabinet. Then, it returns a successful write I/O response to the host.
3) Node 1 flushes the data to the disk.

3. Metadata Protection

The storage system uses a metadata mechanism to manage data. Each pair of storage controllers stores a copy of the metadata, which is obtained from the primary metadata node (by default, the first controller in the system is the primary metadata node). The metadata contains relevant information such as data addresses and access permissions. Clients access the metadata information on the controllers to obtain the final data access addresses and permission information, and then access the target data.

All controller nodes in the storage system are in a peer - to - peer relationship. Any change in the local metadata will immediately be notified to other storage controllers for metadata update. At the same time, the global lock mechanism is used to ensure the consistency of metadata across all nodes.

4. Overall UPS Power - failure Data Protection

The storage system adopts an overall UPS power - failure data protection solution to ensure the safety of the entire system. In case of a sudden power outage, the system can be continuously powered for 3 minutes. During this period, the system switches the cache write strategy to the disk mode according to the pre - set policy.

When only the write cache mirroring protection within the first - level chassis is configured, if one of the two power supplies of the system fails, the system switches to the write - through mode according to the policy (if one power supply of the two in a chassis fails, the entire system switches to the disk mode) to ensure business continuity and data safety. If both power supplies fail simultaneously, the data in the cache will be written to the disk, and then the system will shut down in an orderly manner. This not only ensures the safety of the data in the cache but also protects the disks, especially in an environment with poor computer room conditions, maximizing the safety and reliability of the system.

System Reliability

1. System Layer

Support for Active-Active Storage Access: Two sets of storage systems can enable dual copies of data to be online simultaneously, with seamless failover in case of failures. Compared with methods such as backup and disaster recovery, this ensures that the business does not interrupt and data is not lost, maximizing the reliability of the system and the continuity of the business.

Highly Reliable Storage: The storage system is composed of 4 control nodes, which are interconnected through a redundant IB network. The entire system operates in a fully symmetric mode, with all control nodes in the Active state. It implements a two-level cache mirroring protection mechanism. In extreme cases, it can still provide data access services to the front end normally even if 3 control nodes fail, without data loss or business interruption. When any control node fails, other nodes will automatically and uninterruptedly take over the tasks and the managed data of the failed control node completely, and this process is completely transparent to the client.

RAID and N+M Data Protection: All metadata and data of the storage system are stored on the back-end disks. The storage can implement various redundancy modes such as RAID 0, 1, 5, 6, 10, 50, 60 for the disks, and supports hot spare disks and automatic RAID reconstruction. It supports the RAID mode with triple parity checking, allowing any three disk media in the same RAID group to have physical failures without data loss or business interruption.

Data Protection by Multi-replica Technology: The storage system can ensure that multiple copies of data are online simultaneously through the implementation of multi-replica technology, maximizing the continuity of the business and the reliability of the system.

Redundant Configuration of Equipment Components: Components such as power supplies, fans, and host interfaces are redundantly protected and support hot swapping. The redundant equipment does not affect the normal operation of the entire device when a failure occurs. The controller failover time is less than 1 second, and the redundant components can be replaced online within less than 5 minutes.

2. Hardware Layer

All hardware components of the storage system, including controllers, power supplies, fans, and network interfaces, are designed redundantly and support online hot swapping without affecting the system operation.

The storage system uses real-time data checksum, which can complete end-to-end data consistency verification. It automatically checks and repairs potentially damaged sectors on the disk to ensure data integrity.

When a hard disk fails, the data can be quickly reconstructed. 1TB of data can be recovered within 10 minutes, achieving rapid data recovery.

3. Software Functions

The storage system supports a variety of advanced data protection functions such as snapshots, clones, mirror disaster recovery, and WORM, maximizing the security and reliability of the system and data. It also supports advanced functions such as system-level disaster recovery, data backup, and data remote replication....

System High Security

1. Unique Encrypted Secure Storage Technology

The storage system can be configured with an encrypted secure storage function. This function is

currently a qualified and recognized encrypted storage technology product, which supports commercial cryptographic algorithms and the AES algorithm, enabling the storage of data in ciphertext to ensure data security.

This encrypted storage technology realizes the separation of powers among the storage administrator, security administrator, and audit administrator. It is the only storage product in the current industry that can provide an access audit function. It can provide a graphical audit function and user behavior reports with a graphical interface, providing higher security protection for shared storage access methods such as NAS.

2. **Customized Encrypted Security Solution**

The storage system can be combined with the current mainstream encryption technologies to quickly develop encrypted secure storage systems that meet the requirements of different industries, satisfying the data security and confidentiality requirements of different industries.

System High Availability

1. **Unified Storage Technology Architecture**

This storage is one of the very few storage products in the industry that supports providing multiple access methods such as NAS, FC-SAN, IP-SAN, and IB-SAN simultaneously. The unified storage technology can greatly enhance the availability of the storage, making it suitable for various application scenarios where structured and unstructured data are mixed.

This storage supports the above multiple access methods on the same storage controller platform. The NAS function is realized through the combination of non-SAN storage and the NAS file engine. Compared with the SAN + NAS file engine method, the NAS performance, especially the ability to process small files, is more excellent, and the reliability and failover capability of NAS are stronger.

2. **More Advanced Functions**

The storage system can realize functions such as active-active access, snapshots, clones, automatic hierarchical storage, support for third-party software plugins like VMware, deduplication, mirror disaster recovery, and data synchronization. It can manage the entire life cycle of data, improve data usage efficiency, and better serve production systems. In contrast, traditional distributed file systems or similar storage systems lack most of these advanced functions.

Ease of Management

1. **Integrated Centralized Management and Control**

The traditional way of managing distributed systems and storage devices separately increases the management complexity of the entire system, especially when the software and storage are from different brands. This storage method does not require a distributed file system. The unified management and monitoring of the storage side greatly reduce the management cost.

2. **Reduction of Operation and Maintenance Costs**

The traditional distributed system method requires adding I/O servers and management devices for the distributed system. At the same time, due to the poor scalability of traditional disk arrays, more devices are needed to meet the total capacity requirements of the cloud platform. All these will greatly increase the space, power supply, and load-bearing capacity of the computer room, and also bring higher operation and maintenance costs.

The storage system adopts a high-density storage layout. A few node controllers can achieve the performance level that dozens of nodes of traditional storage can only achieve. Correspondingly, it can also significantly reduce the entire storage deployment space and subsequent operation and maintenance costs.

Scalability

1. **Online Performance Expansion**

The storage system can realize the cluster storage function by expanding the controllers online. In the

NAS environment, it can support up to 256 nodes at most, and provide a maximum sustained and stable bandwidth of 1024GB/s.

2. **Online Capacity Expansion**
   The storage system can realize the cluster storage function by expanding the controllers online. At the same time, by expanding the disk cabinets and disks online, the storage space can be expanded to a maximum of 160PB in the NAS environment.

Technical Specifications

Product Model	MPS5300
Controller Expansion	Support horizontal expansion of 2-32 controllers
Cache (per dual-control)	2.0TB
Number of Hard Disks (per dual-controll)	24*3.5” (expansible)
Hard Disks	SSD, SAS, NL-SAS and SATA disks, support mixed insertion, hot swapping
Front-End Port Type	8/16/32Gb/s FC, 1/10/25/40/100Gb/s iSCSI, 45/56/100Gb/s IB,10Gbps FCoE
I/O Slots (per dual-control)	16
RAID Class	RAID0, 1, 5, 6, 10, 50, 60 and Triple parity
OS Supporting	AIX, HP-UX, Solaris, Windows, Linux and so on
Storage Protocol	NFS、CIFS、RDMA、FC、iSCSI、FTP、HTTP、S3、RESTful and so on
Dimension(W*D*H)mm	Host: 437.2mm*775.2mm*176.5mm (4U) 3.5-inch Expansion Enclosure: 483mm*534mm*174.4mm (4U) 2.5-inch Expansion Enclosure: 483mm*511mm*88mm (2U)
Weight (include disk units)	Host: ≤46kg 3.5-inch Expansion Enclosure: ≤29kg 2.5-inch Expansion Enclosure: ≤23kg
Power Consumption	⩽1200W
Power Supply	Host: 100~240V AC ±10% 3.5-inch Expansion Enclosure: 100~240V AC full range 2.5-inch Expansion Enclosure: 100~240V AC full range
Temperature	Operating Temperature: 0℃~35℃ Storage Temperature: -20℃~60℃
Humidity	Operating Humidity: 20%~80% RH (non-condensing) Storage Humidity: 10%~90% RH (non-condensing)

Software Features

Feature	Description
Unified Storage	Integrates SAN and NAS into a unified storage solution, providing both SAN and NAS data services externally.
Distributed File System	Manages multi-node file clusters and provides file storage services externally.
High Availability Multi-SAN Control	Implements multi-node cluster SAN storage functionality, offering cluster SAN storage services externally.
Unified Namespace Access	Provides a global unified mount point for external access, offering a unique IP address for client servers to achieve global access with a unified namespace.

SAN Shared File System	Offers a SAN shared file system, enabling unified namespace management for SAN storage allocated to front-end hosts.
Load Balancing Management Software	Implements load balancing for file access environments, based on either internal storage or host-side load balancing, improving aggregated bandwidth and single-stream bandwidth performance. Supports configurable load balancing strategies like CPU usage, network bandwidth, TCP/IP connections, and polling.
Data Tiering Management Software	Enables data migration across different storage media, automatically storing frequently accessed data on faster disks, and providing tiered storage to improve hot data performance. Supports mixed disk types like SSD, SAS, and NL-SAS within a single storage system.
RAID Protection	Provides RAID 0, 1, 5, 6, 10, 50, 60 protection, with fast data reconstruction and supports global hot spare disks.
Triple Parity Check	Supports triple parity RAID, allowing any three physical disks in a RAID group to fail without data loss.
Erasure Coding Protection	Supports N+M erasure coding protection, providing higher usable space compared to replication, especially in multi-node distributed environments.
Data Multiplexing	Supports data replication, allowing critical data to be configured with two or three copies, with optional user visibility or transparency.
Deduplication	Removes redundant data to reduce storage size and improve available storage space.
Data Compression	Real-time data compression within storage, utilizing hardware compression cards to optimize storage space.
Storage QoS	Allocates storage resources on-demand based on LUN or file system I/O priorities and traffic control, ensuring service quality for data applications.
Data Migration	Implements tiered migration management based on policies, with transparent data access for clients.
OpenStack Compatibility	Supports compatibility with OpenStack cloud platforms, allowing storage space management via OpenStack.
Hadoop Compatibility	Supports compatibility with Hadoop platforms, enabling access to NAS storage through HDFS.
Network and Protocol Support	Supports multiple networks including 40/56Gb IB, 25GE/10GE/GE IP, 8/16Gb FC, with interfaces like NFS, CIFS, iSCSI, FC, RDMA, HTTP, FTP, OpenStack Swift, etc.
Automated Thin Provisioning	Implements flexible storage capacity planning with on-demand allocation, simplifying system management.
Heterogeneous Storage Management	Manages capacity and status monitoring for mainstream open-interface storage systems, enhancing efficiency for heterogeneous platform operations.
Multipath Management	Configures multipath management for SAN environments, enabling load balancing and failover between multiple paths, with bandwidth aggregation and increased redundancy. Supports mainstream Windows, Linux.
Capacity Quotas and Reservations	Provides space quotas for users and volumes, with alerts when data space reaches the quota, ensuring data security. Supports multi-level quota management based on users, groups, and directories.
Cache Partitioning	Divides cache into control and user caches to accelerate business processes, with dynamic adjustments for read and write caches to enhance performance.

	Mirrored write cache between controllers ensures data safety.
Cache Acceleration	Implements a two-tier cache system using SSDs as secondary cache to optimize system caching strategy, improving cache efficiency and storage system performance, especially for small file environments.
Dynamic Cache Optimization	Uses different cache writing methods for aligned full-stripe data and unaligned partial-stripe data to enhance performance.
Multi-Tenant Support	Implements storage virtualization technology to create independent and fully isolated logical partitions quickly and simply.
Remote System Management	Provides secure and comprehensive centralized management for storage devices, with a graphical, unified Chinese management interface. Supports space division, node upgrades, disk and network status monitoring, and logging and fault alarm functions.
Remote Power On/Off	Supports remote power on/off functionality for storage systems.
System Online Upgrade & Expansion	Supports online upgrades for nodes, enabling performance and capacity enhancement without system downtime.
Storage Active-Active Access	Implements active-active access between two storage systems, ensuring continuous service even if one system fails. Supports SAN and NAS integration for active-active access with or without third-party arbitration, and FC link replication.
Snapshots	Provides read-only snapshots of storage data to prevent data loss from accidental operations. Supports directory-level snapshots with manual or scheduled configurations.
Cloning	Creates writable instant copies of LUNs, enabling quick replication of data volumes.
Mirroring	Synchronizes data through FC and IP transmission methods, supporting local and remote volume replication and mirroring in both synchronous and asynchronous modes.
Cache Mirroring	Implements real-time write cache mirroring between two controllers, supporting SAN and NAS write cache data protection.
Failover Software	Enables manual or automatic failover between storage controllers for high availability, ensuring minimal downtime.
Storage Encryption	Integrates real-time encryption and decryption modules to secure data and ensure confidentiality within the storage system.
Legal Compliance	Provides WORM (Write Once, Read Many) disk storage capabilities to enhance data durability and integrity, ensuring compliance with legal requirements.
Security Access Control	Implements comprehensive access control and auditing strategies to ensure secure, concurrent access by multiple clients based on defined permissions. Supports user, group, and directory-based permission management with NIS, Microsoft Active Directory, and LDAP.

Maipu Communication Technology Co., Ltd
No.16, Jiuxing Avenue
Hi-Tech Zone
Chengdu, Sichuan Province
P. R. China
610041
Tel: (86) 28-65544850,
Fax: (86) 28-65544948,
URL: http:// www.maipu.com
Email: overseas@maipu.com