p +61 1300 113 112
e info@hyperscalers.com

ZSDS
High Performance Software Defined Storage Solution

Based on Ceph using High Performance NVMe Storage by DapuStor

Introduction

HYPERSCALERS

Tuesday, 15 August 2023

INTRODUCTION

Software Defined Storage (SDS) has become the default approach for provision of flexible, available, performant, and secure storage resources in the context of contemporary solution architectures.

Delivery of faster performance is always needed, however, despite other advancements in storage system technology.

Hyperscalers [1] and Dapustor [2] have partnered therefore to jointly co-develop and qualify high performance NVMe storage solution products against the industry standard Ceph software defined storage (SDS) platform.

The result is a high-performance SDS platform engineered to redefine performance, price, versatility, scalability, and availability capabilities in the modern storage solutions marketplace.

This platform delivers all the well-known benefits of Ceph in the context of blazingly fast (NVMe) storage products that integrate industry leading, custom storage controller ASICs with the highest performance flash memory technology available.

We believe that this combination of elements has yielded the fastest software defined storage solution (SDS) available in the marketplace, hence we are referring to it as "ZSDS".

Hyperscalers have developed the ZSDS platform with an all-flash Non-Volatile Memory express (NVMe) Ceph storage cluster using market leading NVMe technology from Dapustor [3] in 1U servers provided by Hyperscalers [4].

This unique combination of well-known, trusted hardware elements and the Ceph platform is the perfect combination for delivery of highly available, mature and flexible Block, File and/or Object-storage services provided by Ceph.

In addition to raw performance, Dapustor NVMe technology provides end-to-end data protection, power level configuration, multiple namespace support, device and capacitor health detection, command priority control, enhanced secure erasure and SR-IOV (virtual compute support). These capabilities are supported by custom engineered, high performance storage controller ASICs developed in-house by Dapustor. Dapustor is a global leader in the field of storage controller ASIC engineering.

The ZSDS platform is highly configurable with our minimum base implementation able to support 15 GBps cluster throughput against 334.2 TB total raw capacity, all in 3 RU. If density is a requirement, this base design can easily be doubled in capacity or throughput (or both) with no additional server hardware required. Conversely, capacity and/or throughput can be expanded across additional cluster nodes for greater availability – the choice is yours.

ZSDS is a very fast, reliable, low cost and turnkey SDS platform that has been rigorously qualified through Hyperscalers' Hardware Appliance Design process to ensure full-stack compatibility across the Ceph, operating system and device driver layers.

Taken in conjunction with Hyperscalers trusted ability to support enterprise class customers across multiple localities, ZSDS should be viewed as a highly engineered, industry leading storage platform solution that can be trusted to support storage requirements across any customer context.

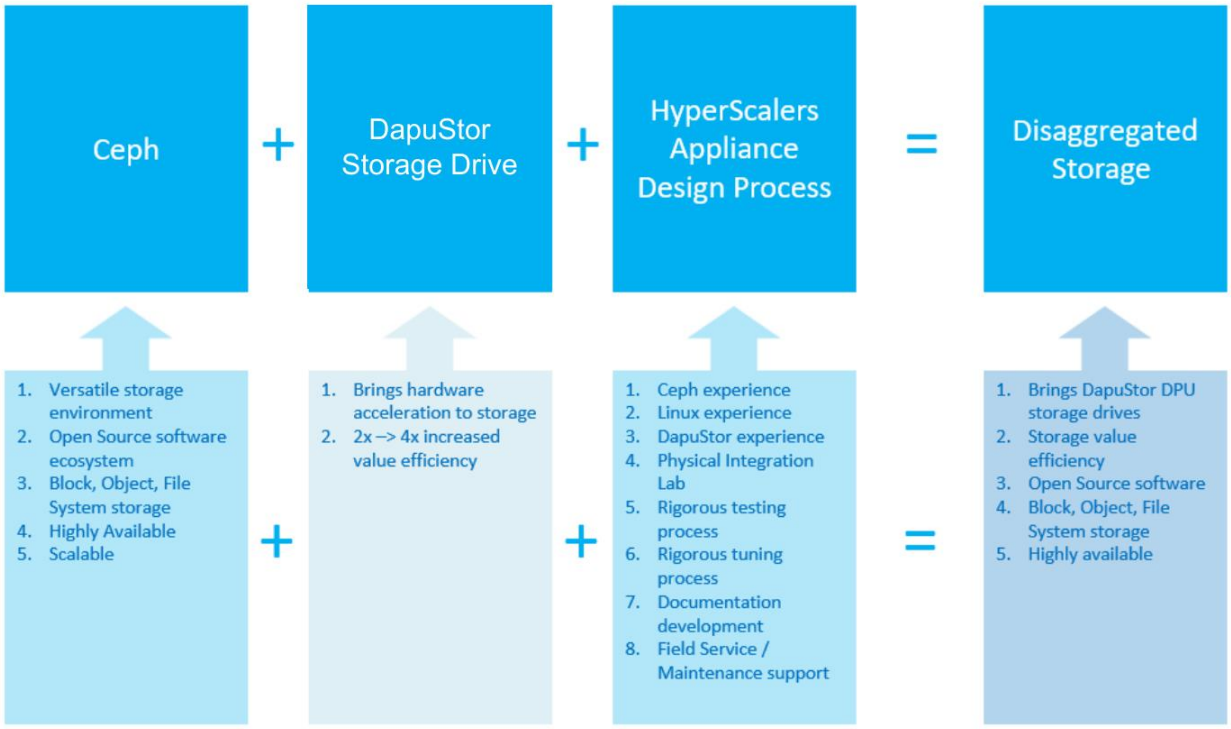


Figure 1 Values of the ZSDS Software Defined Storage Solution

Ceph with hardware acceleration allows you to accomplish more with your data.

Businesses, academic institutions, global enterprises and others all share a common need for fast access to critical data that in many cases is irreplaceable. Hence, we have undertaken the significant engineering effort needed to build the ultimate realisation of software defined storage using a hardware accelerated storage context.

The Ceph platform inherently provides trusted software defined capabilities fulfilling data management requirements across key domains including flexibility (of storage types), configurability, reliability and scalability.

ZSDS takes Ceph capabilities and places them on top of blazingly fast NVMe storage drives by Dapustor. SDS platform speed is derived not only from the use of all-flash NVMe, but in particular through the use of onboard controller devices. These hardware accelerators implement key storage processing functions directly into silicon that were once performed in software running on general purpose processors (for example on RAID management cards).

Silicon based (hardware accelerated) functions allow ZSDS to deliver a combination of performance and compression metrics that are not otherwise achievable in a software defined storage context, even when using all-flash NVMe.

Ceph, via its advanced CRUSH algorithm, supports automated data redundancy, self-management daemons and much more. This means that ZSDS always ensures your data is safely stored, instantly available and optimally distributed for effective disaster recovery.

Engineers within Hyperscalers Canberra based laboratory have worked through an extensive fine-tuning regime to optimise ZSDS for absolute performance. This has included optimisation of numerous critical CPU and network layer attributes, in addition to software stack qualification (done in collaboration with DapuStor). ZSDS utilises only the latest generation Ice Lake Intel CPUs, and very high speed (100gbps) network interfaces.

Data is *invaluable*, so ZSDS keeps it *safe*. Without effective protection, data loss can be highly damaging to business, incurring significant costs and causing irreparable damage to reputation. All data protection features available under Ceph are equally as reliable in a SDS platform context as for a non-accelerated platform context.

ZSDS provides speed and capacity cost advantages transparently, allowing you to focus on the Ceph storage use–cases that are critical to your organisation.

You can continue to streamline data management and compliance processes, reduce downtime and network bottlenecks, and remove traditional barriers to scale.

ZSDS enables any organisation to build out the Ceph storage use-cases that are critical to them using the most highly optimised hardware / cost solution that is currently available.

ZSDS is broadly applicable across all verticals

No matter what type of organisation any of us are working within, we all share a common need for fast access to critical data.

The following examples highlight specific use-case scenarios that are relevant in the context of some well-known environments:

SMEs

ZSDS provides scalable and reliable data storage for object, block and file storage, without the need for the investment in expensive hardware.

Global organizations

ZSDS can be configured to maximise the value of geographically distributed data centres to deliver highly available, resilient data. By placing relevant organisation-wide data physically at the regional locations where it is needed, your people and processes can be empowered in ways that may not have otherwise been easily possible.

Academic institutions

The extensive customisation capabilities of ZSDS enable tailored data curation and management control, in turn supporting the collection of both unstructured and structured data at scale and reliable backups of your research for secure, robust data storage.

Start-ups

ZSDS makes it easy to purchase only what you need initially and to scale on demand later. This means that you can quickly adjust your cluster balance as you add, replace or remove storage media. If your start-up business experiences sudden exponential growth, CSS will grow your cluster right alongside you with no license renegotiation needed.

Developers

The open-source adaptability inherent in ZSDS is perfect for inclusion into your own software environment and architecture, supporting unimpeded experimentation with cloud hosted services and/or operations across large volumes of block, file and object data.

Financial Services

Modern financial services companies are heavily reliant on storage systems to capture, store and process event information in some cases at staggeringly high speeds. This is particularly true within the new breed of Fin-tech style operations where speed and scale flexibility are paramount. ZSDS can help to maximise the cost-performance-reliability equation that is critical within these environments.

Government

Nowhere is the need for cost-optimised, flexible and fast access to large quantities of data more apparent than in the context of government organisations handling critical data relating to an entire population. ZSDS can provide not only the capacity and performance benefits discussed here but also the reliability, maintenance and support services that are always required by government users.

Use ZSDS to achieve…

Intelligent software integration

The technology architecture employed within ZSDS provides a flexible and adaptable foundation for integration to a broad range of services, applications and cloud technologies. From web-scale content repositories to machine learning and artificial intelligence architectures you can gain detailed insights into your data and leverage it to improve business outcomes. Use inbuilt ZSDS protocols to work with block, object, and file storage, or create your own interface using the LIBRADOS API.

Scalable storage

With ZSDS there is no limit to your data growth. As a virtualised storage system ZSDS can scale as you require without the confines of traditional hardware storage. ZSDS is designed to ensure the reliability, performance and availability of your data - even alongside exponential growth.

Undoubted reliability

Advanced algorithms, intelligent object storage daemons and automated self-management capabilities enable ZSDS to provide businesses with high reliability and quick, detailed insights into cluster health. It is well known that manual data management processes can result in operational delays and overlooked cluster issues. In order to mitigate these problems, ZSDS automates round-the-clock monitoring, back-ups and data protection.

Business continuity

There are few businesses that can run smoothly without access to key operational data. Regardless of whether the business context is a customer bookings platform or an automated logistical system (for example), downtime always reduces customer confidence and can have a significant impact on revenue generation. ZSDS will help you to guarantee business continuity by enabling you to implement finely tuned and automated data redundancy across devices, racks and geographic locations.

Controlled investment

In contrast to the hefty payments and costly infrastructure requirements imposed by some storage vendors, ZSDS will run on your existing off-the-shelf hardware. This can make transition to ZSDS far more cost-effective than when dealing with any of the incumbent storage marketplace alternatives. Additionally, by balancing your data replication across a cluster, it becomes easy to monitor cluster capacity and limit any purchasing of additional storage only to the circumstances where it is actually needed.

Choose the hardware that matches your needs

ZSDS provides unparalleled flexibility in your choice of hardware. In fact, it is able to run on just about anything. By freeing organisations from committing to a single hardware vendor, ZSDS supports constant adaptation and innovation, with all hardware components able to be swapped out as needs change. This enables you to provide access to Object, Block or File storage from one unified cluster, while decoupling your data management architecture from any specific hardware elements you may choose to use.

Remove bottlenecks

Bottlenecks that seemed acceptable in smaller deployments may rapidly become costly and unwieldy at scale. As a distributed storage system, ZSDS provides seamless data retrieval by enabling client applications to calculate the location of data within a cluster directly. This capability removes the need for a traditional metadata server. By supporting a direct path for update and retrieval, network traffic is reduced and a critical single point of failure is negated, providing smoother and more reliable service to you and your customers.

A solution you can rely on long-term

Cephs' global community of industry leading developers ensure that Ceph never falls behind the storage technology curve. There are more experts reviewing the Ceph code base on a daily basis than would ever occur for a proprietary solution. This results in quick response to evolving security requirements and new innovations in storage technology. Cephs' core philosophy and open-source model ensures continued development and practicality. Ceph is built to address the needs of its users, and it continues to offer additional features and efficiencies with every new release.

Built to scale

ZSDS offers *high density storage, scalable to petabyte* levels within a smaller footprint (minimum 3 x 1U[1]). With typical cost reduction of 250% per TB of NVMe, ZSDS is a fraction of the traditional costs associated with NVMe class storage.

Fast

It is possible to *deploy and scale* a POC[2]/ production ready storage server in no more than few hours[3]. With Dapustor high-performance NVMe storage, ZSDS achieves throughput up to *15 GB/s*[4].

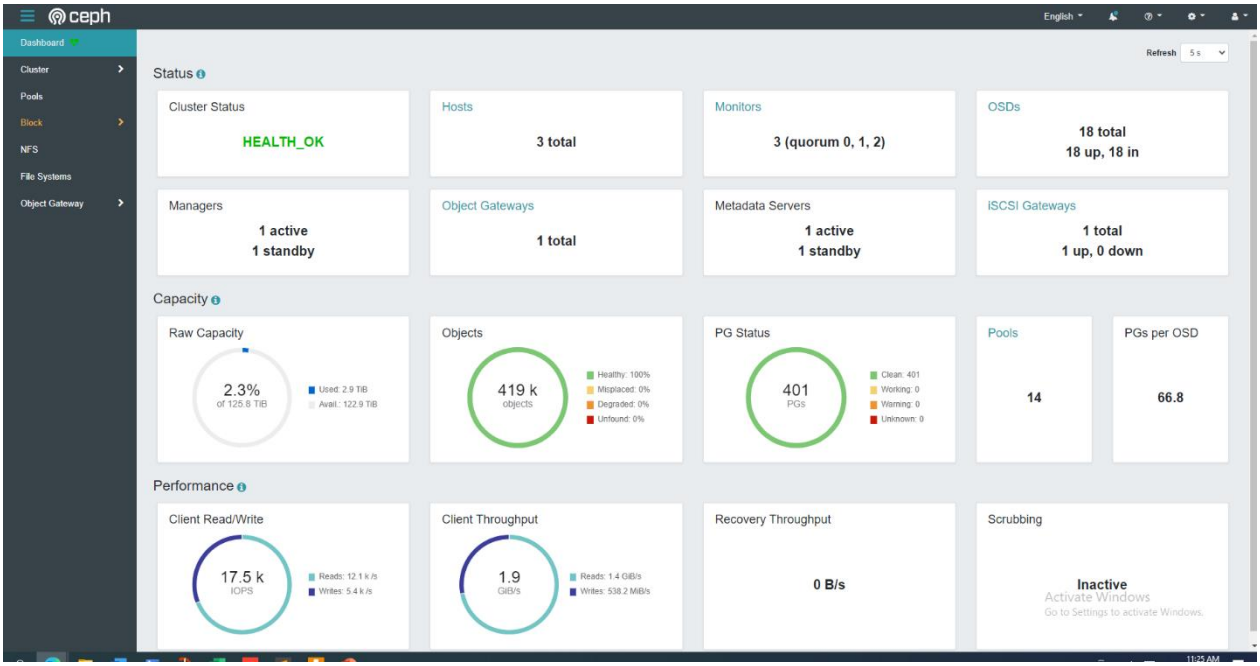
Intuitive Management Interface

ZSDS provides an intuitive and easy to use management user interface / dashboard that enables you to comprehensively manage and monitor all aspects of your environment including:

- Monitoring & audit of cluster logs,
- Addition, configuration and removal of hosts, object storage drives, monitors and managers,
- Creation and configuration of storage pools (with replica/erasure coding) to support Object, Block and File system storage,
- Creation and management of storage volumes and images,
- Creation and management of image mirroring,
- Creation and management of object buckets.

The following screenshot of the Ceph management interface depicts high level Ceph environment status including cluster health, number of Ceph system elements and key capacity / performance metrics:

[1] Refer to ZSDS Reference guide for details on the hardware.
[2] Contact Hyperscalers for a demo/ test drive with LaaS.
[3] For Hyperscalers tested hardware only.
[4] Refer to ZSDS Reference guide for more information.



High Availability and Low risk

Courtesy of the highly evolved Ceph architecture, ZSDS offers superior reliability and availability. This in turn allows storage providers to deliver storage services with minimal downtime. ZSDS *mitigates risks and protects data through replication, mirroring, and erasure coding*[5].

Storage solution for every application

ZSDS offers block, object, and file system storage on demand. This enables it to dynamically serve any combination of hypervisor, container, web/monolithic application and direct clients.

ZSDS Storage Types

Object storage

Object storage is a data storage architecture that manages data as individual binary objects. Each object includes data, metadata (providing reference information on the data) and a unique ID. Object storage allows customers to connect web applications to uniquely identified storage objects acting as storage points, and from there to store any form of unstructured data within them. *In ZSDS, object storage is deployed in the form S3 / Swift buckets which can either be attached to by internal consumers or published as a public or private cloud accessible item.*

File System storage

File System (FS) storage is a data storage structure in which data is stored and maintained as files within folders. This approach will be familiar to any desktop computer user navigating their file system folder structure via Explorer (Windows) or the Finder (MacOS). The File System

[5] Refer to ZSDS Reference guide to know more on replication and erasure coding.

storage structure provides readability and convenience. *File System storage under ZSDS can be deployed as a network file system (NFS) or as CephFS*

Block storage

An abstract form of File System storage is known as Block storage. Block storage supports storage of structured data to be used by VMs, applications and end users. *Block devices provided by ZSDS can be attached to clients as a RADOS block device (RBD) or through iSCSI.*

Support

Hyperscalers offers full hardware and software support to ZSDS customers across multiple geographic localities. Depending on severity[6], analysis of any specific issue can range between a few business hours to three business days.

[6] Refer to your Support agreement with Hyperscalers or contact Hyperscalers to know more about support.

Infrastructure

The following figure shows how the ZSDS architecture relates to some example contemporary storage use-cases, and therefore how it can empower ZSDS users to store, manage and utilise their data:

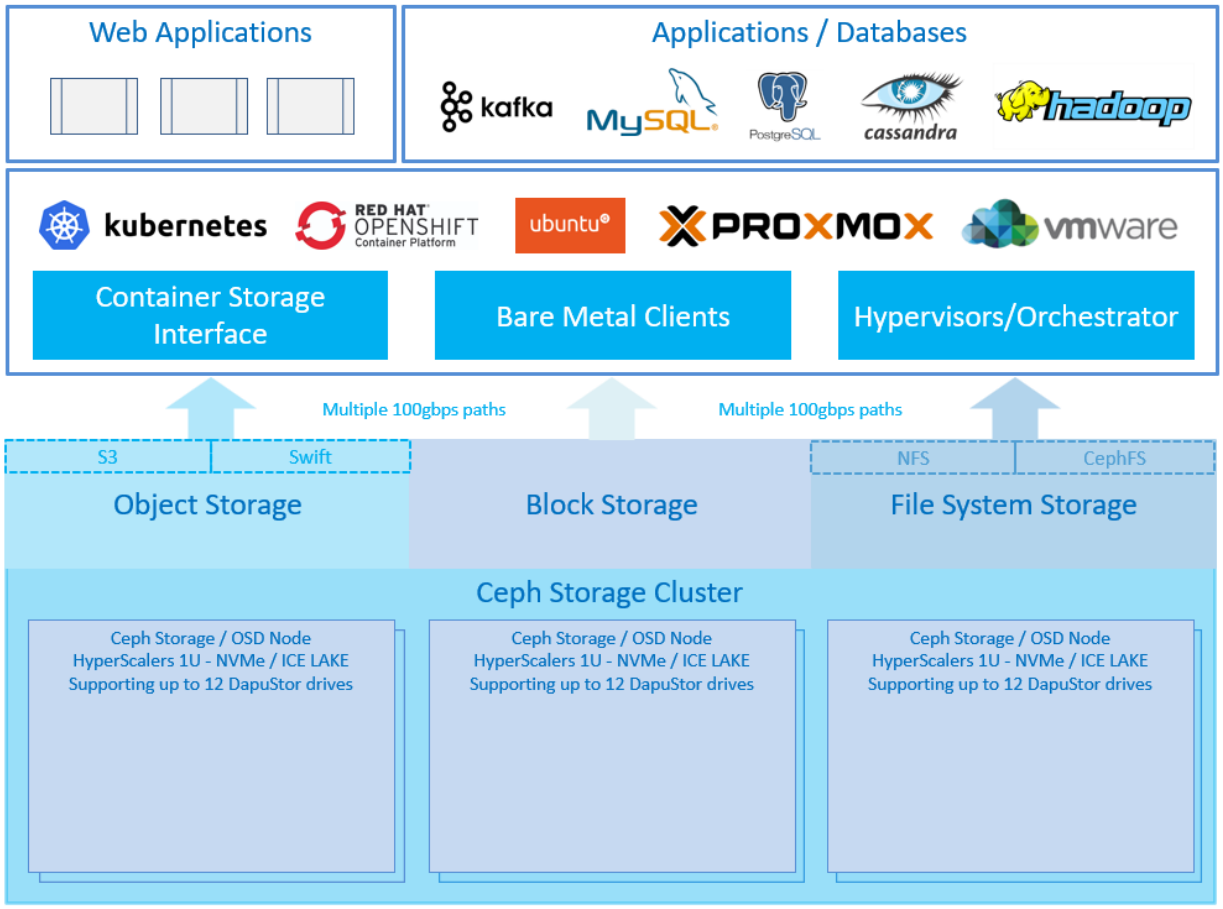


Figure 2 ZSDS Software Defined Storage Solution generic structure

The requirements of this architecture are mentioned at Page 11 of the ZSDS Reference Guide. Interconnect between ZSDS storage elements and any applications/clients depends on various factors - please refer to the ZSDS Reference Guide to understand specific technical requirements for storage interconnect.

Why DapuStor high performance Storage Drive

DapuStor Roealsen5 eSSD (Enterprise Solid State Drive) technology delivers high performance, high reliability and low latency for enterprise and data center applications.

DapuStor is a global leader in the development and production of custom engineered ASICs. DapuStor performs all research & development, design and qualification of specialised ASICs in-house. DapuStor engineers have generated significant amounts of intellectual property, resulting in more than 300 patents being granted. DapuStor owns and operates its own chip fabrication and product manufacturing plants, ensuring that strict quality outcomes are maintained across all stages of production.

DapuStor utilises only the highest performance Flash memory devices in its Roealsen5 product line. These devices combined with the leading-edge capabilities of DapuStor ASIC controllers place DapuStor eSSD (NVMe) products at the forefront of the storage technology marketplace.

Key advantages of DapuStor Roealsen5 eSSD drives include:

1. End-to-end data protection, through Variant Sector Size (VSS) + Protection Information

VSS is a technology that allows SSDs to support multiple sector sizes, which can help to improve performance and reduce the amount of wasted space on the drive. By supporting multiple sector sizes, SSDs can be optimized for different types of workloads and provide better performance for specific applications.

Protection Information (PI) is a feature that adds extra data to each sector to detect and correct errors that may occur during data transmission or storage. This helps to ensure data integrity and prevent data loss or corruption.

DapuStor SSDs use a combination of Variant Sector Size and Protection Information to provide end-to-end data protection. This ensures that data is protected from the moment it is written to the SSD, throughout the data transfer process, and during storage on the SSD. Additionally, DapuStor SSDs feature power loss protection, which uses capacitors to ensure that data is not lost in the event of a power failure.

2. Device heath and capacitor health detection

SMART (Self-Monitoring, Analysis, and Reporting Technology) - All DapuStor SSDs include an integrated SMART monitoring system. It is used to track various metrics related to the drive's health and performance, such as the number of write cycles, the temperature of the drive, and the amount of available spare blocks. When a drive's SMART data indicates that certain thresholds

have been exceeded or errors have been detected, it can be a sign that the drive is experiencing issues or is at risk of failure.

Health Monitoring Software - DapuStor provides health monitoring software that can be used to track the status of SSDs. These tools often provide more detailed information than SMART and can alert users to potential issues before they become critical.

Capacitor Health Detection - The Roealsen5 Series have capacitors that provide power during data writes in the event of a sudden power loss. Capacitor health detection is used to ensure that these components are functioning properly and can provide the necessary power when needed. This is typically done through monitoring the voltage and capacitance of the capacitors and comparing it to expected values.

3. Six levels of power consumption/optimisation

DapuStor Roealsen5 series comes with a power management feature that allows users to adjust power level utilisation based on their specific workload requirements, thereby reducing overall power consumption. The Roealsen5 series supports six different power modes, ranging from the lowest power consumption level (P0) to the highest power consumption level (P5).

The power levels adjustment feature allows users to customize the power consumption of their SSD according to their specific needs, such as read or write-intensive workloads. For example, users can set the SSD to a lower power consumption level during periods of low activity, reducing overall power consumption and extending the life of the device.

This feature is especially beneficial for data center and enterprise customers who need to optimize their power consumption while maintaining high-performance levels. By adjusting the power levels of their SSDs, users can reduce their energy bills and lower their carbon footprint, while still achieving high levels of performance and reliability.

4. Multiple-namespace support (up to 32 namespaces)

Multiple-namespace support is a feature that allows a single physical SSD to be divided into multiple logical namespaces, each with its own independent storage space and access controls. This can be useful in a variety of scenarios, such as virtualization, multi-tenancy, and data isolation.

DapuStor SSDs offer support for up to 32 namespaces, which is a significant number compared to other SSDs on the market. This means that users can create up to 32 independent namespaces on a single DapuStor SSD, each with its own unique identifier and access controls. This provides a high level of flexibility and customization, allowing users to tailor their storage environment to their specific needs.

In addition to multiple-namespace support, DapuStor SSDs also offer other advanced features such as power-loss protection, wear-leveling, and encryption, making them a popular choice for enterprise and data center environments.

5. Weighted round robin command arbitration and high priority command support

Weighted round robin command arbitration is a technique used to manage the ordering of commands that are submitted to an SSD. With this technique, each command is assigned a weight or priority, and the SSD processes commands in a round-robin fashion, taking into account their assigned weight. This ensures that higher priority commands are processed more quickly than lower priority ones.

DapuStor SSDs do support weighted round robin command arbitration, which means that users can assign priorities to their commands and ensure that high-priority commands are processed with a higher priority than lower-priority ones. This can be particularly useful in scenarios where certain commands need to be processed urgently, such as in a real-time application or a database server.

Overall, the support for weighted round robin command arbitration is another example of the advanced features that DapuStor SSDs offer, making them a popular choice for high-performance computing environments.

6. Enhanced secure erase

DapuStor SSDs offer enhanced secure erase features, which are designed to clear existing data safely and protect data security and privacy before drives are reused.

The secure erase feature in DapuStor SSDs is a block erase/block overwrite operation that can be initiated by the user or via a software command. When initiated, the secure erase process overwrites all user data on the drive with a predefined pattern, rendering the data unrecoverable. This is done by overwriting the entire user addressable space on the SSD, including any hidden or reserved areas.

The enhanced secure erase feature in DapuStor SSDs provides a high level of data security, which is particularly important for enterprise and data center environments where sensitive data may be stored. By securely erasing data before a drive is reused, DapuStor SSDs help to prevent the possibility of data leakage or other security breaches.

Overall, the enhanced secure erase feature is another example of the advanced features that DapuStor SSDs offer, making them a popular choice for users who require high-performance storage solutions with robust data security and privacy features.

7. SR-IOV support (virtual environment support).

DapuStor SSDs come with SR-IOV (Single Root Input/Output Virtualization) support, which can provide high performance and low latency in virtual compute environments.

SR-IOV is a technology that allows a single physical device, such as an SSD, to appear as multiple virtual devices to virtual machines. This can provide several benefits in virtualized environments, including improved performance, reduced latency, and better resource utilization.

By supporting SR-IOV, DapuStor SSDs can help to reduce the overhead associated with virtualization, allowing virtual machines to access the SSD directly and achieve high levels of performance and low latency. This can be particularly useful in scenarios where high-performance storage is required, such as in database servers or high-performance computing environments.

Overall, the SR-IOV support in DapuStor SSDs is another example of the advanced features that these SSDs offer, making them a popular choice for users who require high-performance and low-latency storage solutions in virtualized environments.

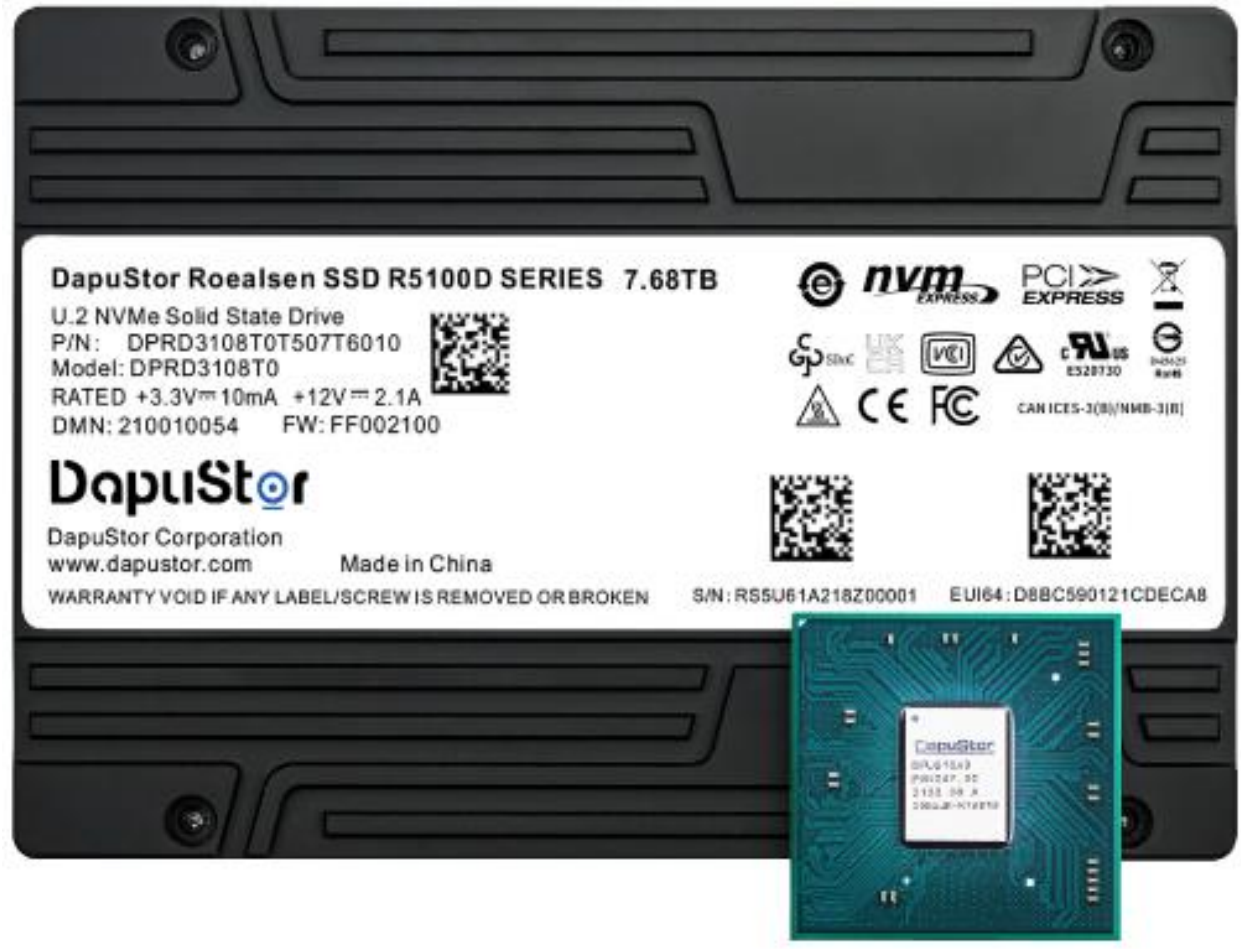


Figure 3 DapuStor eSSD Storage Drive, and high performance ASIC

Why Ceph

The Ceph platform delivers Object, Block and File System storage in one unified system.

Advantages of Ceph technology include:

- Open-source approach and ecosystem
- Block storage (ideal for Host/VM clients) and File-level storage (ideal for web application clients)
- Object storage with Swift and S3 API (Ideal for application development)
- Fault tolerance
- Self-healing and self-managing
- Ceph orchestrator (Cephadm) for easy expansion management of Ceph clusters
- High Availability
- Redundancy via multiple network paths

Why Hyperscalers and DapuStor

Hyperscalers [1] is the world's first open supply chain Original Equipment Manufacturer- (OEM), solving Information Technology challenges through standardization of best practices and hyperscale inspired practices and efficiencies. Hyperscalers offers choice across two open hardware architectures:

- Hyperscale - high efficiency open compute equipment as used by macro service providers
- Tier 1 Original – conventional equipment as per established Tier 1 OEM suppliers.

Each of these architectures is complete with network, compute, storage, and converged GP GPU infrastructure elements, and is open / free from vendor lock-in.

Hyperscalers' appliance solutions are packaged complete with hardware, software and pre-built (customisable) configurations. These were all pre-engineered using an in-house IP Appliance Design Process and validated in partnership with associated major software manufacturers. Many can be "test-driven" using Hyperscalers Lab as a Service (LaaS). Hyperscalers appliance solutions are ideally suited to IaaS, PaaS and SaaS providers looking to implement their services from anywhere.

DapuStor Corporation (DapuStor) is a leading expert in premium enterprise solid-state drives (SSD), SOC, and edge computing related products. DapuStor has comprehensive capabilities in storage system ASIC controller chip design, fabrication and manufacturing.

This deep experience has placed Dapustor at the leading edge of SSD/NVMe product development globally. Dapustor products are used widely across the telco, service provider and data center marketplaces.

DapuStor is an established, large-scale and award-winning manufacturer.

The Dapustor Roealsen5 R5100 series is a family of premium performance, data center grade storage products supporting end-to-end data protection, power efficiency management, multiple-namespace support. command arbitration / prioritisation, support for virtual compute environments and enhanced secure erase security protection. [5].

1 REFERENCES

[1] Hyperscalers, "About HS," [Online]. Available: https://www.hyperscalers.com/about-us-hyperscalers.

[2] Dapustor, "About-Overview," [Online]. Available: https://en.dapustor.com/about.html. [Accessed 2023].

[3] Dapustor, "Roealsen Series" [Online]. Available: https://en.dapustor.com/product.html. [Accessed 2023].

[4] Hyperscalers, "S5X 2.5" | D53X-1U," [Online]. Available: https://www.hyperscalers.com/storage/storage-servers/hyperscalers-S5X-D53X-1U-ice-lake-densest-hyperscale-server-nvme-drives-buy. [Accessed 2022].

[5] Dapustor, "Roealsen R530X [Online]. Available: https://en.dapustor.com/product/2.html. [Accessed 2023].

[6] Ceph, "Ceph Homepage," [Online]. Available: https://ceph.com/en/. [Accessed 2023].