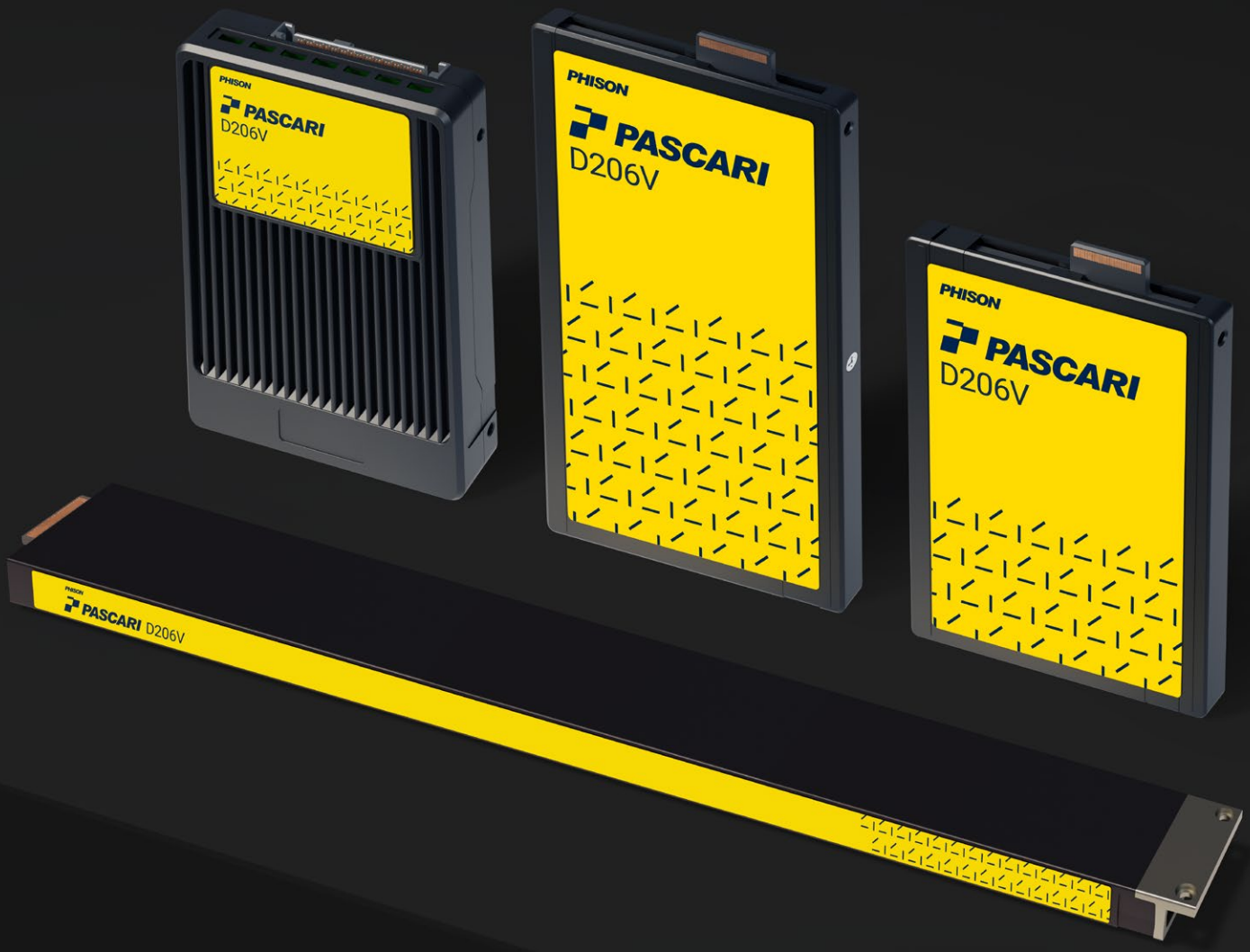
PASCARI

D206V

Exabyte Scaling for AI Storage

Solutions Brief



PHISON

Phison Pascari D206V Ushers in Exabyte Scaling for AI Storage

Phison's Pascari D-series drives are ready to tackle the most demanding storage needs, with the new D206V delivering massive capacities and scalability. The Pascari D206V enterprise SSD represents a major leap forward in storage density and operational efficiency, dramatically increasing both capacity-per-drive and capacity-per-watt for hyperscale AI deployments.

With the accelerating AI revolution, the appetite for high-performance, high-density storage continues to grow. Training frontier-scale large language models, supporting retrieval-augmented generation (RAG), powering vector databases, and sustaining real-time inference pipelines all demand unprecedented levels of data capacity and throughput. Storage infrastructure is no longer just a backend consideration but is now a foundational pillar of AI scalability.

The implications become even more significant when scaling to AI pod architectures. A pod consisting of 12 fully populated storage racks has the potential to deliver 1.4 exabytes of flash storage capacity per pod. This effectively enables exabyte-scale NVMe flash deployments within a relatively compact datacenter footprint.

8X Capacity Increase in the Same Power Envelope

The Pascari D206V delivers an enormous increase in flash density. With individual SSD capacities scaling up to 245.76 TB in a standard U.2 form factor, the D206V delivers eight times the storage capacity of the previous-generation Pascari D200V 30.72 TB.

Even with this exponential jump in capacity, the D206V maintains the same maximum power envelope of 30 watts. That means the D206V achieves an equally dramatic 8X improvement in capacity-per-watt efficiency. (Pascari D206V offers capacities up to 245.76 TB in U.2, 122.88 TB in E1.L and E3.L, and 61.44 TB in E3.S.)

From a datacenter engineering perspective, this is a transformational metric. Traditional scaling methods require adding more drives, more servers, more racks, and more cooling infrastructure to expand storage pools. The D206V changes that equation entirely by radically increasing usable storage density without proportionally increasing power consumption or rack footprint.
In practical terms:

- Previous generation Pascari D200V:
  - 30.72 TB at 30W
  - ~1.02 TB/W
- New Pascari D206V:
  - 245.76 TB at 30W
  - ~8.19 TB/W

This capacity-per-watt efficiency level represents a substantial advancement in flash storage economics for AI infrastructure operators.

AI Infrastructure Demands Massive Storage Density

Modern AI pipelines consume data at astonishing rates. Multi-trillion-token datasets, synthetic training content and results, embedding databases, checkpoint snapshots, inference caches, and multimodal training repositories can easily scale into the petabyte and exabyte range. The challenge isn't simply storing the data; it's storing it efficiently, making it economically and thermally viable.

High-density enterprise SSDs like the Pascari D206V directly address several critical AI infrastructure bottlenecks:

- Reducing rack-level power consumption
- Minimizing datacenter floor space requirements
- Lowering cooling overhead
- Simplifying storage cluster architecture
- Increasing storage density per server node
- Reducing cabling and networking complexity
- Improving overall storage fabric efficiency

By dramatically increasing the usable capacity per drive, organizations can deploy fewer physical devices while achieving substantially larger storage pools. It's not just that the capacity per drive has increased by 8X; the number of racks required for the same capacity potentially drops by the same 8X factor.

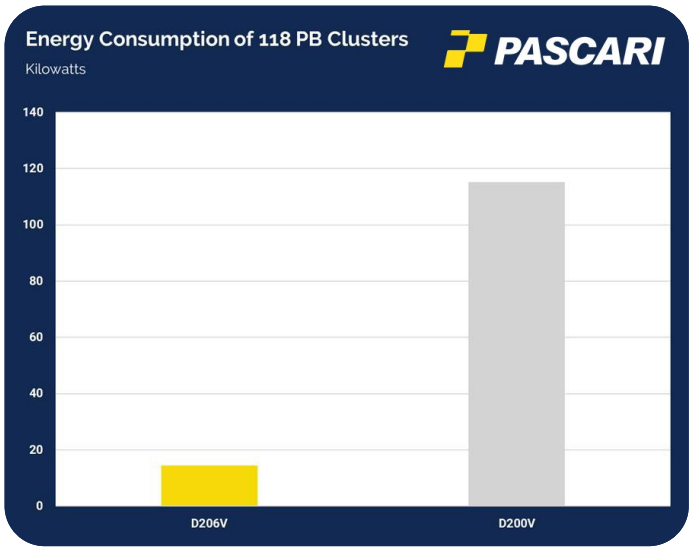
Copyright ©2026 Phison Electronics | All Rights Reserved.

118 PB in a Single Rack, Using Just 14 kW of Power

To understand the scale enabled by the D206V, consider a modern 2U storage server equipped with 24 U.2 SSDs. Using fully populated D206V drives, 24 drives of 245.76 TB yields 5.9 PB per 2U server. Now scale that to a standard 42U rack configuration with 20 servers and the D206V enables a whopping 118 PB per rack.

That's an extraordinary amount of flash storage density packed into a single rack footprint. Even more impressive is the power profile. Each D206V drive consumes a maximum of 30W, with 24 drives in a server requiring just 720W of power. Across 20 servers in a rack, that works out to 14.4 kW total SSD power draw per rack.

More importantly, that's nearly 8 PB of flash capacity per kilowatt from the storage layer. For hyperscale AI clusters, this level of storage efficiency fundamentally changes deployment economics. Rack power budgets can now support dramatically larger datasets without proportional increases in cooling or facility infrastructure.



Exabyte-Scale AI Pods Become Practical

The implications become even more significant when scaling to AI pod architectures. A pod consisting of 12 fully populated storage racks has the potential to deliver 1.4 exabytes of flash storage capacity per pod. This effectively enables exabyte-scale NVMe flash deployments within a relatively compact datacenter footprint.

At this scale, the D206V becomes highly relevant for:

- Frontier AI model training
- Foundation model dataset repositories
- Distributed object storage
- Vector database clusters
- AI inference caching tiers
- GenAI content indexing
- High-speed data pools
- HPC and scientific simulation environments

The convergence of ultra-high-density NAND flash and efficient power utilization is an essential part of data center design as AI datasets continue their exponential expansion.

Pascari D206V Performance Characteristics

Even with the massive increase in capacity, performance remains competitive across the infrastructure. Individual D206V SSDs deliver up to 14,000 MB/s of sequential reads and 3,400 MB/s of sequential writes. In large hyperscale installations, overall throughput is typically determined by the network bandwidth, with the fastest solutions like NVIDIA's Bluefield-4 SuperNIC offering 800 Gbps of bandwidth — that's 100 GB/s of maximum throughput for interserver communications.

With 24 D206V drives installed in a server, the aggregate write speeds reach up to 81.6 GB/s, but combined read speeds reach 336 GB/s, several times faster than current network speeds. As the D206V targets read-heavy workloads, the 800 Gbps network speed of the latest generation hardware will be the limiting factor.

Pascari D206V Performance	
Capacity	Up to 245.76 TB
Sequential Read	Up to 14,000 MB/s
Sequential Write	Up to 3,400 MB/s
Random Read	Up to 2,200K IOPS (4K)
Random Write	Up to 40K IOPS*

*Random write performance on 245.76TB measured with 32KB. All other sizes measured with 16KB.

Copyright ©2026 Phison Electronics | All Rights Reserved.

Redefining Storage Economics for AI

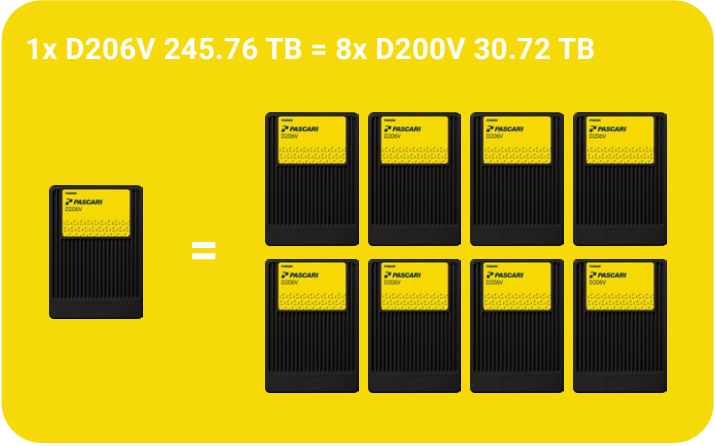
Historically, exabyte-scale storage deployments required enormous amounts of spinning disk infrastructure, consuming massive floor space, power, noise, and cooling resources. Mechanical storage also introduced performance limitations that increasingly conflict with accelerated AI workflows.

The D206V represents a shift toward highly consolidated, all-flash AI storage architectures where:

- Data locality improves
- Latency decreases
- Throughput increases
- Rack density improves dramatically
- Energy efficiency scales more effectively

Storage is one of the largest contributors to infrastructure TCO (total cost of ownership), and increasing capacity-per-watt by 8X can materially impact operational expenditure over the lifespan of a deployment.

This becomes particularly important in GPU-heavy AI clusters, where power budgets are already under intense pressure from accelerators consuming 700W to 1,200W per GPU. Any reduction in storage-layer power requirements directly improves overall infrastructure efficiency.



Built for the Next Era of AI Infrastructure

The Pascari D206V demonstrates how enterprise storage is evolving to meet the demands of AI at hyperscale. By combining massive 245.76 TB capacities with a tightly controlled 30W power envelope, the D206V delivers a rare combination of density, scalability, and energy efficiency.

As AI infrastructure transitions from petabyte-scale environments into exabyte-class deployments, storage systems capable of maximizing both rack density and power efficiency will become increasingly critical. The era of exabyte-scale flash storage is no longer theoretical. It's now here, enabled by the Pascari D206V 245.76 TB enterprise SSD.

Pick a Size and Form Factor

The Pascari D206V is available in multiple form factors and capacities for varying storage requirements. There's a Pascari D206V drive for every storage need whether the server is engineered for U.2, E3.S, E3.L, or E1.L form factors. Capacities of 30.72 TB, 61.44 TB, and 122.88 TB are available in some of the form factors, but the 245.76 TB is U.2 only.

* Up to 245.76 TB capacity available only in the U.2 form factor. Actual user capacity may be less depending on the operating environment. Product specifications subject to change without notice. Performance numbers may vary based on system configuration and testing conditions.

Copyright ©2026 Phison Electronics | All Rights Reserved.

Pascari D206V Enterprise SSD Specifications

Pascari D206V U.2 Specifications

Capacity	30.72 TB	61.44 TB	122.88 TB	245.76 TB
Interface	PCIe 5.0 1x4, 2x2	PCIe 5.0 1x4, 2x2	PCIe 5.0 1x4, 2x2	PCIe 5.0 1x4, 2x2
NVMe Protocol	2.0	2.0	2.0	2.0
NAND Flash	3D QLC	3D QLC	3D QLC	3D QLC
Sequential Read (MB/s)	14,000	14,000	14,000	13,700
Sequential Write (MB/s)	3,300	3,300	3,300	3,100
4K Random Read (IOPS)	2,200K	2,200K	2,200K	2,200K
16K/32K Random Write (IOPS)	34K	34K	34K	21K
Read Latency (µs)	110	110	110	110
Write Latency (µs)	40	40	40	40
Max Power (W)	≤25	≤25	≤25	≤30
Idle Power (W)	≤5	≤5	≤5	≤5
DWPD	0.3	0.3	0.3	0.3
MTBF (million hours)	2.5	2.5	2.5	2.5
Warranty	5	5	5	5

Pascari D206V E3.S Specifications

Capacity	30.72 TB	61.44 TB
Interface	PCIe 5.0 1x4, 2x2	PCIe 5.0 1x4, 2x2
NVMe Protocol	2.0	2.0
NAND Flash	3D QLC	3D QLC
Sequential Read (MB/s)	14,000	14,000
Sequential Write (MB/s)	3,300	3,300
4K Random Read (IOPS)	2,200K	2,200K
16K Random Write (IOPS)	34K	34K
Read Latency (µs)	110	110
Write Latency (µs)	40	40
Max Power (W)	≤25	≤25
Idle Power (W)	≤5	≤5
DWPD	0.3	0.3
MTBF (million hours)	2.5	2.5
Warranty	5	5

(1) The product is still in development stage, all values provided are based on estimation.
(2) 1 TB = 1012 bytes.
(3) Sequential Performance is based on FIO on Linux, 128KB data size, with QD=32, 1 job.
(4) Random Performance is based on FIO on Linux, random read 4KB data size, random write 16KB data size, QD=128, 8 jobs.
(5) Latency is measured with random workloads based on FIO on Linux, random read 4KB data size, random write 16KB data size, QD=1, 1 job.
(6) Power consumption (average RMS) may differ according to flash configuration and platform.
(7) The results of DWPD are obtained in compliance with JESD219A standards and IU size aligned.

Copyright ©2026 Phison Electronics | All Rights Reserved.

Pascari D206V Enterprise SSD Specifications

Pascari D206V E3.L Specifications

Capacity	122.88 TB
Interface	PCIe 5.0 1x4, 2x2
NVMe Protocol	2.0
NAND Flash	3D QLC
Sequential Read (MB/s)	14,000
Sequential Write (MB/s)	3,400
4K Random Read (IOPS)	2,700K
16K Random Write (IOPS)	40K
Read Latency (µs)	110
Write Latency (µs)	40
Max Power (W)	≤30
Idle Power (W)	≤5
DWPD	0.3
MTBF (million hours)	2.5
Warranty	5

Pascari D206V E1.L Specifications

Capacity	122.88 TB
Interface	PCIe 5.0 1x4, 2x2
NVMe Protocol	2.0
NAND Flash	3D QLC
Sequential Read (MB/s)	14,000
Sequential Write (MB/s)	3,300
4K Random Read (IOPS)	2,700K
16K Random Write (IOPS)	34K
Read Latency (µs)	110
Write Latency (µs)	40
Max Power (W)	≤25
Idle Power (W)	≤5
DWPD	0.3
MTBF (million hours)	2.5
Warranty	5

(1) The product is still in development stage, all values provided are based on estimation.
(2) 1 TB = 1012 bytes.
(3) Sequential Performance is based on FIO on Linux, 128KB data size, with QD=32, 1 job.
(4) Random Performance is based on FIO on Linux, random read 4KB data size, random write 16KB data size, QD=128, 8 jobs.
(5) Latency is measured with random workloads based on FIO on Linux, random read 4KB data size, random write 16KB data size, QD=1, 1 job.
(6) Power consumption (average RMS) may differ according to flash configuration and platform.
(7) The results of DWPD are obtained in compliance with JESD219A standards and IU size aligned.

Copyright ©2026 Phison Electronics | All Rights Reserved.