WHITEPAPER | Persistent Memory

VT-PM8 & VT-PM16 EVALUATION

VT-PM drives, part of Viking's persistent memory technology family of products, are 2.5" U.2 NVMe PCIe Gen3 drives optimized with Radian Memory System's architecture technology. The VT-PM8 and VT-PM16 drives are persistent memory drives that deliver performance and unlimited write endurance similar to that of DRAM, while simultaneously providing the data persistency desired for enterprise applications.

VT-PM drives come with a U.2 NVMe PCIe interface offering enterprise OEMs the ease of deployment for persistent memory unseen in the market today. Alternative developments of persistent memory architecture utilizes NVDIMM technology, a non-volatile memory module that attaches to DDR4 memory sockets, which can limit configurations and utilization of persistent memory due to number of sockets. By enabling the VT-PM8 and VT-PM16 drives behind the industry-standard NVMe interface, OEMs are able to access the drive via normal block or programmed I/O (mmap) byte addressable accesses; without firmware or HW requirements.

FEATURES

- 8GB, 16GB capacity
- Lowest latency NVMe device
- NVMe and byte addressable mmap access
- Dual Port 2x2 or Single Port x4 mode
- NVMe PCIe x4 Gen3 interface
- 2.5" U.2 NVMe Drive form factor
- Hot Swap, Live Insertion, Surprise Remove
- OEM Lifecycle Monitoring Diagnostics

Dual Port Persistent Memory with Unlimited DWPD Endurance

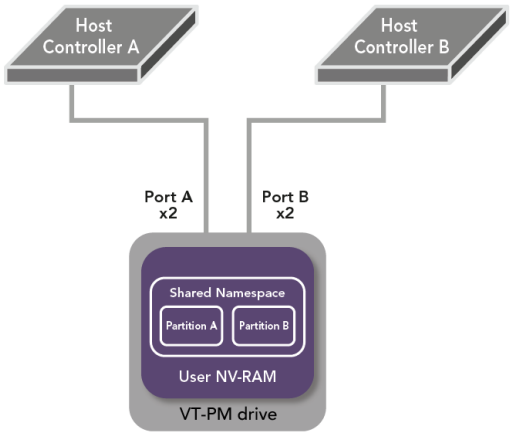
BENEFITS

- Simple, reliable, persistent
- Consuming minimal host CPU resources while providing high performance
- Supports the NVMe command set
- No remote capacitor packs or cabling
- Interface simplicity

UNLIMITED ENDURANCE (DWPD) + PERSISTENT MEMORY

NAND Flash and certain storage class memories have a limited number of Drive-Writes-Per-Day (DWPD) that is often inadequate for write intensive applications like caching, logging, buffering and journaling. The VT-PM drive utilizes DD4 memory that provides unlimited write endurance and is made non-volatile by an internal power backup system that is fully contained in the same U.2 2.5" form factor SSD. No secondary drive slot is required for auxiliary power or tethering cables to a remote power system.

DUAL PORT MODE FOR ACTIVE/ACTIVE CONTROLLERS



- Dual Port 2x2 NVMe device
- Shared NVMe namespace accessible from either port
- Namespace can easily be configured into dedicated partitions for each host
- Partner host can access alternate partition on failover
- Avoids complex inter-host coordination

In dual port 2x2 mode, a single shared NVMe persistent memory namespace is accessible from either port and can be partitioned so that each host can safely write to its own partition without complex inter-host coordination. Both hosts can mount and read from either partition based on sharing permissions. If a host becomes inoperative due to failure or maintenance events, the partner host can access the inoperative host's persistent memory partition by mounting the partition using standard primitives and permissions.

HOT SWAP FUNCTIONALITY

The VT-PM drive includes Surprise Hot-Add and Hot-Remove capabilities to support Hot Swap for High Availability requirements. The SSD's fault tolerance capabilities ensure data integrity and protect against shorn writes through these Hot Swap events and unplanned power interruptions. However, each of these features is platform dependent, requiring complementary support and interoperability from the CPU/motherboard, bios, OS and PCIe fabric.

PERSISTENT MEMORY APPLICATIONS

With exceptional, consistent performance for small random writes and unlimited write endurance, host controlled persistent memory can be the ideal solution for applications such as write-ahead logging, caching, journaling, intent logs or any application requiring low latency, persistent storage regardless of queue depth. The VT-PM drive is visible as a standard block device that supports NVMe transfers or can be memory mapped with a fixed PCIe BAR to support Programmed I/O (PIO) operations based on 4-byte (dword) addressing that are immediately durable.

- Write-ahead logging and caching for low latency, deterministic response times regardless of queue depth
- Persistent non-volatile metadata store to support random writes or high frequency overwrites
- Coalescing and buffering writes in persistent memory, then deterministically scheduling data transfers to large capacity Flash or HDD

ENDURANCE REQUIREMENTS

These data center logging and caching persistent memory applications usually involve near continuous write workloads. The following table extrapolates typical logging and caching write data rates into DWPD:

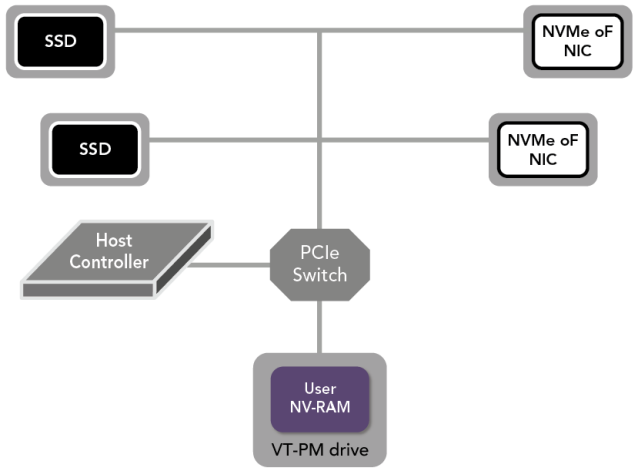
EXAMPLE WRITE DATA RATES	GB/DAY	DWPD WITH 8GB DEVICE
50,000 IOPS (4K)	17,694	2,212
200,000 IOPS (4K)	70,778	8,847
1 GB/s	86,400	10,800
2 GB/s	172,800	21,600

Data center class Flash SSDs generally provide anywhere from 1 to 10 DWPD. Certain new storage class memories are rated at up to 30 DWPD. While significantly better than Flash SSDs, 30 DWPD is dramatically less than the 2,000 to 20,000 DWPD required to fulfill data center write logging and caching applications.

NVME-OF CENTRALIZED BUFFER

As part of an NVMe-oF system, the VT-PM drive can act as a centrally accessible repository on the PCIe fabric. RDMA based transfers or byte addressable memory windows can be partitioned to buffer and interleave data between different I/O, compute and storage endpoints.

Creating fabrics over PCIe and utilizing existing network software drivers can each be simplified and made more performant by having byte addressable memory accessible on the NVMe storage endpoints.



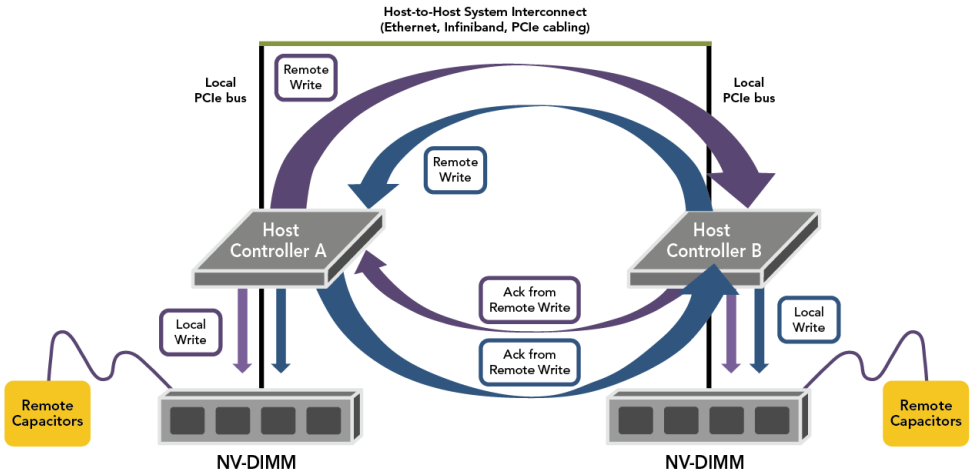
MIRRORING IN ACTIVE/ACTIVE CONFIGURATIONS

The majority of data center storage arrays are deployed in a 'dual head', active/active configuration for high availability. Utilizing local attached persistent memory, such as NV-DIMMs, in these system configurations requires creating a synchronous mirror which often involves considerable complexity, overhead, and inefficiencies.

SYNCHRONOUS MIRRORING

- Overhead from additional operations, acknowledgements and potential copying
- Additional, potentially extensive complexities in system software coherency
- Often requires cabling between controllers and the additional cost of NICs/HBAs
- Additional potential latency of an interconnect protocol and the bridging transition to that protocol
- Difficult to service

The simplicity of the dual-port VT-PM drive overcomes each of these challenges



DIALOG™

The VT-PM drive includes DiaLog (Diagnostic Logging), a host accessible, embedded diagnostic facility that includes various monitoring functions related to predictive/preventive maintenance, reliability, and continuous process and product improvement.

- Measure & Detect
- Diagnose & Predict
- Record & Notify

On-board health monitoring of components and events are tracked utilizing an I2C network and are available to the host. These capabilities target use by OEM customers throughout their Qualification, Production Test, Field Deployment, Repair, and End-of-Life phases.

Upon system power failure, the VT-PM drive switches to an auxiliary power mode provided by on-board supercapacitors and data stored in volatile DRAM is transferred to persistent NAND memory by the Flush-to-Flash firmware. The overall Flush-to-Flash system and underlying NAND array are based on a fault tolerant architecture, utilizing transactional semantics and ACID design principles. In addition, resources such as supercapacitor power and NAND capacity are overprovisioned to address events such as repeated system power blackouts and brownouts, protecting against shorn writes and data integrity errors during the flush process. Extensive monitoring and component checks are performed on an on-going basis during normal operations to discover predictive anomalies in advance of failures.

SPECIFICATIONS

Capacity	16GB (8GB: VT-PM8)
U.2 2.5" NVMe form factor	Length: 100mm, Height: 15.0mm, Width: 69.75mm
Dual Port and Single Port Modes	Dual Port 2x2 lane configuration or Single Port x4 lane configuration
PCIe x4 Gen3	Compliant with PCI-Sig PCIe 3.0 Base specification
NVMe Command Set	NVM Express specification 1.0
DRAM ECC	64-bit data/8-bit ECC code detects double bit errors and corrects single bit errors
NAND ECC	Hardware LDPC engine
NVMe Engines	Supports NVMe command set, submission/completion queues and MSI-X vector interrupts
Programmed I/O (PIO)	Fixed BAR provides support for direct 4 byte addressable (dword) host access by mapping memory (mmap) into host PCI address space with configurable window size
Maximum Payload Size	Configurable to 128B or 256B single packet size
BIST and Health Monitoring	Provides OEMs the ability to monitor environmental status, component health, and log statistics for continuous product life cycle management
Field Upgradeable Firmware Updates	Mechanism for upgrading firmware in the field via host control (no drive removal necessary)
Internal Supercapacitor Module	Field replaceable assembly
Power Requirements (+12V rail)	Typical Maximum: 13.5W at 40°C Recharge Cycle: Up to 17W at 40°C with concurrent r/w operations
Supercapacitor Recharge Time	38 seconds
Temperature	Operating: 0° to 40°C at 100 LFM Storage: 40°C to 85°C
Weight	.295 lbs.
Shock/Vibration	Operating: 5 G Non-Operating: 10 G
ESD	1,500 volts, human body model
MTBF	1M hours
Device Drivers	NVMe Linux 4.10 and above

Radian MEMORY SYSTEMS

Utilizes Radian Memory Systems technology

viking TECHNOLOGY

Global Locations

US Headquarters
2950 Red Hill Avenue
Costa Mesa, CA 92626
Main: +1 714 913 2200
Fax: +1 714 913 2202

Canada Office
500 March Road
Ottawa, ON K2K 0J9
Canada

Texas Office
1201 W. Crosby Road
Carrollton, TX 75006
USA

India Office
A 3, Phase II, MEPZ-Special Economic Zone
NH 45, Tambaram, Chennai-600045
India

Singapore Office
No 2 Chai Chee Drive
Singapore, 109840

For sales information, email us at sales@vikingtechnology.com, or visit our website for all global locations and contact information.

www.vikingtechnology.com

R 1.01