

Non-Volatile Memory Express (NVMe)

The future of high-speed,
low-latency storage

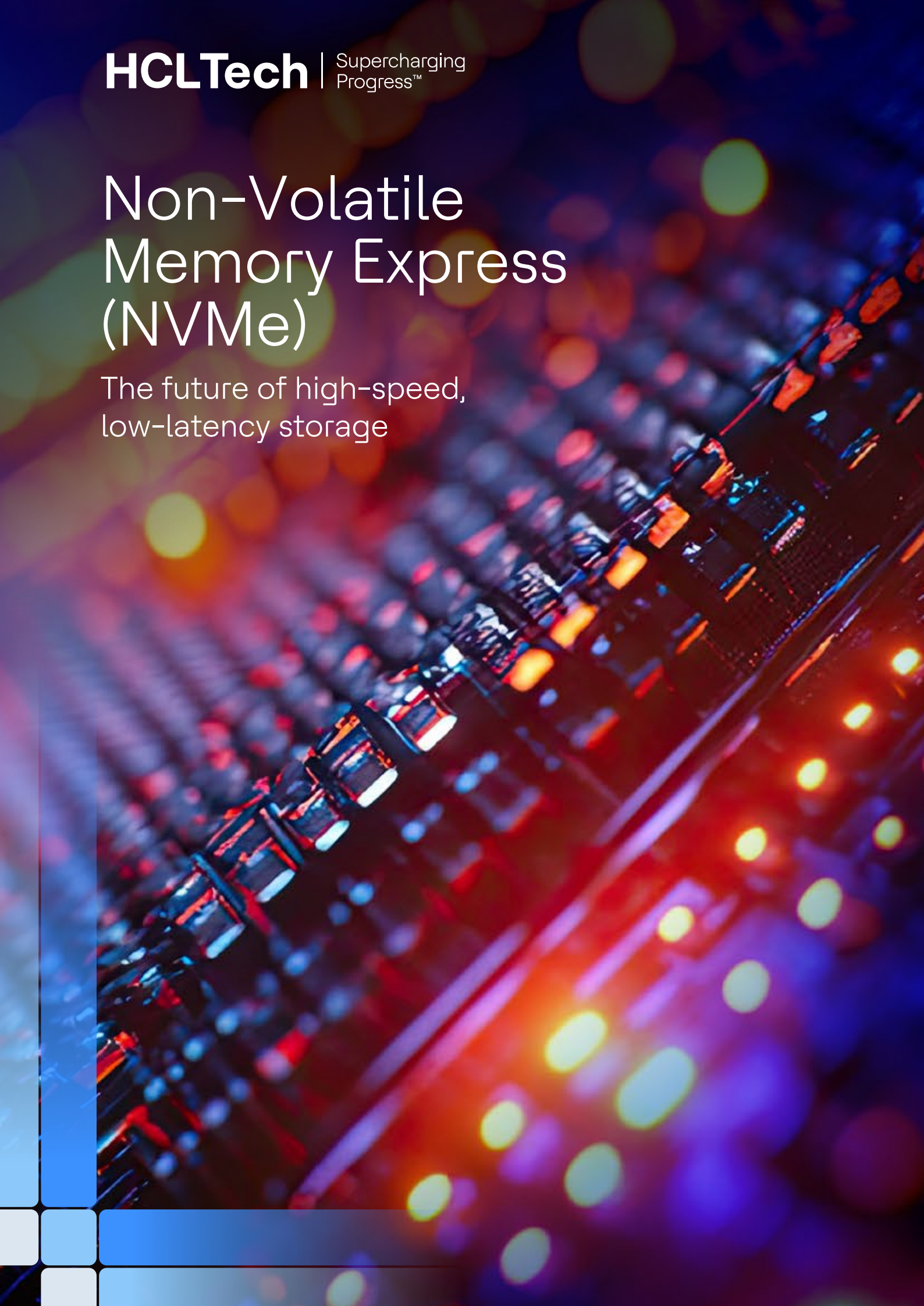


Table of CONTENTS

1. NVMe technological overview	3
• Executive summary	
• Introduction	
2. NVMe front-end	3
• Definition and purpose	
• Key characteristics	
3. NVMe back-end storage	4
• Definition and architecture	
• Key characteristics	
4. NVMe technological benefits	5
5. Use cases and enterprise implementations	6
6. NVMe-oF classifications	6
• NVMe over FC	
• NVMe over TCP	
• NVMe over PCIe	
• NVMe over RDMA(ROCE)	
7. NVMe-oF architecture overview	9
8. NVMe - comparative analysis	10
9. Sustainability benefits	11
10. NVMe operations management consideration	11
• Host multipathing configuration	
• Monitoring and troubleshooting	
• Operational efforts comparison: Traditional FC vs NVMe/FC	
11. Conclusion	12
12. About the Author	13

Non-Volatile Memory Express (NVMe) is a next-generation storage protocol for solid-state drives, delivering industry-leading throughput and ultra-fast response times. Its architecture is optimized for diverse enterprise workloads, enabling efficient, low-latency access to data.

NVMe technological overview

Executive summary

NVMe and NVMe-oF are redefining enterprise storage by replacing legacy protocols such as SATA and SAS with a high-performance, low-latency architecture purpose-built for modern workloads.

NVMe enables direct, parallelized access to flash media over PCIe, delivering exceptional throughput and responsiveness. NVMe-oF extends these capabilities across networked environments, supporting

disaggregated and scalable storage infrastructures optimized for AI/ML, real-time analytics, and cloud native applications.

Key benefits of NVMe protocol

High performance

Delivers sub-millisecond latency and scales efficiently across multiple cores and queues.

Architectural flexibility

Provides a unified protocol for direct-attached (NVMe) and networked (NVMe-oF) storage environments.

Future-ready design

Supports next-generation media technologies (e.g., Storage Class Memory, QLC NAND) and high-speed interconnects (e.g., PCIe 5.0, 400GbE+).

Introduction

NVMe (Non-Volatile Memory Express) is a high-performance storage protocol designed to accelerate data transfer between host systems and storage devices. NVMe front-end and back-end storage architectures are crucial for modern enterprise environments, offering significant latency, throughput, and scalability improvements.

NVMe front-end

Definition and purpose

NVMe front-end storage refers to the interface between host systems and storage arrays using NVMe protocols over high-speed fabrics like PCIe, Ethernet (NVMe/TCP), or Fibre Channel (NVMe/FC). It is designed to reduce latency and increase throughput for enterprise workloads.

Key characteristics



Direct PCIe access

Eliminates SCSI overhead for maximum throughput



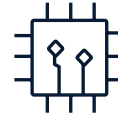
Low latency

Optimized for AI/ML, real-time analytics, and high-speed databases



Scalable architecture

Enables multi-host access and considerable namespace support



High availability

Dual-port SSDs ensure fault tolerance and continuous access



NVMe back-end

Definition and purpose

NVMe back-end storage refers to the internal data path within an array that connects the storage controller to the SSDs using NVMe protocols. It ensures ultra-fast, parallelized access to data stored on NVMe drives, typically over PCIe, Ethernet, or InfiniBand fabrics.

Key characteristics



Massive scalability

NVMe supports multiple parallel I/O queues with deep queue depths, leveraging PCIe for direct, low-latency access. NVMe-oF extends this scalability across networked environments using Ethernet, Fibre Channel, or InfiniBand.



High throughput

Backend performance ranges from 7 GB/s (PCIe Gen3) to 15 GB/s (Gen4/Gen5), depending on the fabric used.



Energy efficiency

NVMe SSDs deliver higher IOPS per watt than traditional SAS/SATA drives, reducing power consumption in performance-intensive deployments.

NVMe technological benefits



Reduced CPU overhead

NVMe's streamlined protocol and efficient driver architecture minimize CPU cycles per I/O operation, enhancing overall system performance



Concurrent workload optimization

Native support for multiple I/O queues enables superior scalability, reducing bottlenecks and ensuring seamless integration with enterprise platforms such as VMware, Oracle, SAP and Microsoft



Low-latency and high IOPS

Achieves sub-millisecond latency (<1ms) and up to 2.1 million IOPS for 8KB read operations, ideal for latency-sensitive applications



High availability

Dual-port NVMe SSDs and bifurcated PCIe links provide robust fault tolerance and continuous data access



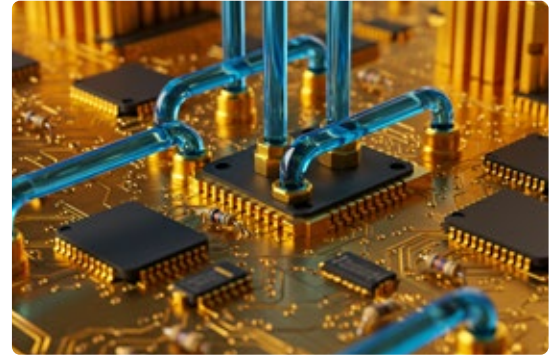
Efficient data protection

NVMe architecture supports high SSD density per stripe group without compromising I/O performance, reducing spare disk requirements and improving rebuild times compared to legacy storage systems



Disaggregated shared everything (DASE)

This technology enables independent scaling of compute and storage resources by leveraging NVMe SSDs, supporting flexible and modular infrastructure design



Operational benefits



Scalable performance

NVMe's parallel I/O architecture and deep queue support deliver consistent, ultra-low latency performance across multi-tenant environments, making it ideal for high-density, mission-critical workloads.



Accelerated provisioning

NVMe back-end SSDs enable rapid volume creation and cloning, streamlining automation pipelines and enhancing DevOps and CI/CD efficiency.



Optimized TCO

With a compact, power-efficient footprint, NVMe reduces energy consumption, cooling requirements and rack space utilization—resulting in lower total cost of ownership and simplified infrastructure management.



Operational intelligence

End-to-end NVMe deployments integrated with telemetry, analytics, and AIOps empower proactive monitoring, predictive maintenance, and faster issue resolution.



Cloud-ready architecture

NVMe-oF supports hybrid and multicloud environments, enabling software-defined storage and seamless scalability for modern workloads, including AI/ML, edge computing, and containerized applications.



Use cases and enterprise implementations

Mission-critical performance

NVMe is engineered for applications requiring consistent low-latency and high throughput, such as real-time analytics, financial systems and transactional databases.

HPE Alletra MP

All-NVMe front and back-end architecture with 32 Gb FC and 25 GbE front-end, and 100 GbE RoCEv2 back-end. Certified for NVIDIA DGX and SAP HANA.

Enterprise workload optimization

Ideal for AI/ML workloads, SAP HANA, Oracle RAC and hypervisor-based environments, NVMe ensures predictable performance and rapid data access.

Pure Storage FlashArray//X

Enterprise-class end-to-end NVMe array delivering sub-500 microsecond latency with unified protocol support.

Dell EMC PowerMax 2000/8000

This product delivers up to 350 GB/s bandwidth through a fully integrated end-to-end NVMe architecture, supporting extreme scalability.

HPE XP8 / Hitachi VSP 5600

High-performance enterprise storage systems offering up to 33 million IOPS and sub-40 microsecond latency.

NetApp AFF series

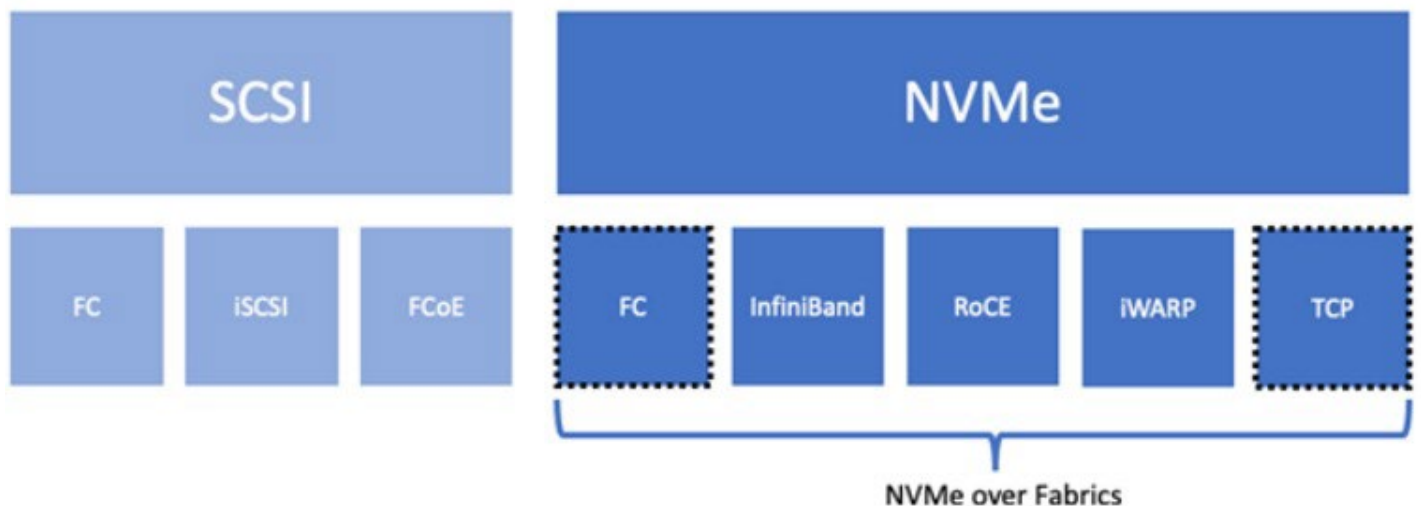
This series offers unified protocol support, including NVMe/TCP and NVMe/FC, with deep cloud integration and AI-driven management for hybrid environments.

VAST Data Universal storage

Combines NVMe over Fabrics with a disaggregated shared-everything architecture, delivering exabyte-scale performance for AI, HPC, and deep analytics workloads.



NVMe-oF classifications



NVMe over FC

NVMe over Fibre Channel (NVMe/FC) is a high-performance storage transport protocol that merges the ultra-fast data access capabilities of Non-Volatile Memory Express (NVMe) with the proven reliability and scalability of Fibre Channel networks. By enabling NVMe command sets to traverse Fibre

Channel fabrics, NVMe/FC facilitates low-latency, high-throughput communication between servers and storage arrays—ideal for mission-critical enterprise environments requiring deterministic performance and robust infrastructure.

Key characteristics



Optimized protocol stack

NVMe/FC encapsulates NVMe command sets within Fibre Channel frames, eliminating the overhead of legacy SCSI protocols. This streamlined approach significantly reduces latency and enhances throughput for high-performance workloads.



High-speed data transport

By leveraging Fibre Channel's low-latency and high-bandwidth capabilities, NVMe/FC enables efficient utilization of modern SSDs and all-flash arrays, ensuring rapid data access across enterprise environments.



Enterprise-grade reliability

Built on Fibre Channel's inherently lossless transport architecture, NVMe/FC guarantees frame integrity and consistent delivery—making it ideal for mission-critical applications that demand fault-tolerant infrastructure.

NVME over TCP

NVMe over TCP is a transport protocol that extends the NVMe standard across conventional Ethernet networks using the TCP/IP stack. As part of the broader NVMe over

Fabrics (NVMe-oF) framework, it enables NVMe command transmission over widely deployed network infrastructures without requiring specialized hardware.

Key advantages



Broad infrastructure compatibility

Unlike NVMe/RDMA, NVMe/TCP operates over standard Ethernet networks, eliminating the need for RDMA-capable adapters and simplifying deployment.



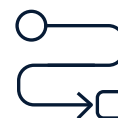
Balanced performance

While not as low-latency as RDMA-based solutions, NVMe/TCP offers substantial performance improvements over legacy protocols such as iSCSI, making it suitable for modern workloads.



Cost-efficient modernization

Ideal for organizations seeking to enhance storage performance without significant investment in new networking hardware.



Flexible deployment scenarios

Commonly adopted in environments where RDMA is impractical, including SMBs, secondary storage tiers, and edge deployments.

NVME over PCIE

NVMe over PCIe represents the native implementation of the NVMe protocol, enabling direct communication between host systems and solid-state drives (SSDs) via PCI Express lanes. This architecture bypasses legacy

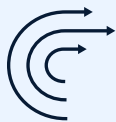
storage protocols such as SATA and SCSI, delivering ultra-low latency and exceptional throughput for performance-critical workloads.

Key advantages



Direct host-to-device communication

NVMe interfaces directly with SSDs over PCIe, eliminating translation layers and reducing I/O overhead.



Massive parallelism

Supports up to 64K I/O queues with 64K commands per queue, enabling highly concurrent data operations.



Sub-millisecond latency

Ideal for real-time applications including AI/ML, advanced analytics, and high-performance databases.



Scalable infrastructure

Easily expands with additional PCIe lanes and NVMe namespaces, supporting flexible and modular system growth.

NVMe over RDMA

NVMe over RDMA (Remote Direct Memory Access) is a high-performance transport protocol within the NVMe over Fabrics (NVMe-oF) framework. It enables direct memory-to-memory data transfers between host

systems and storage devices, bypassing the CPU and operating system kernel to minimize latency and reduce processing overhead.

Key advantages



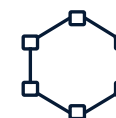
Ultra-low latency

RDMA facilitates direct data movement between memory spaces, making it ideal for latency-sensitive applications such as real-time analytics and high-frequency trading.



High throughput efficiency

Supports parallel I/O operations with minimal CPU involvement, maximizing performance for compute-intensive workloads.



Transport flexibility

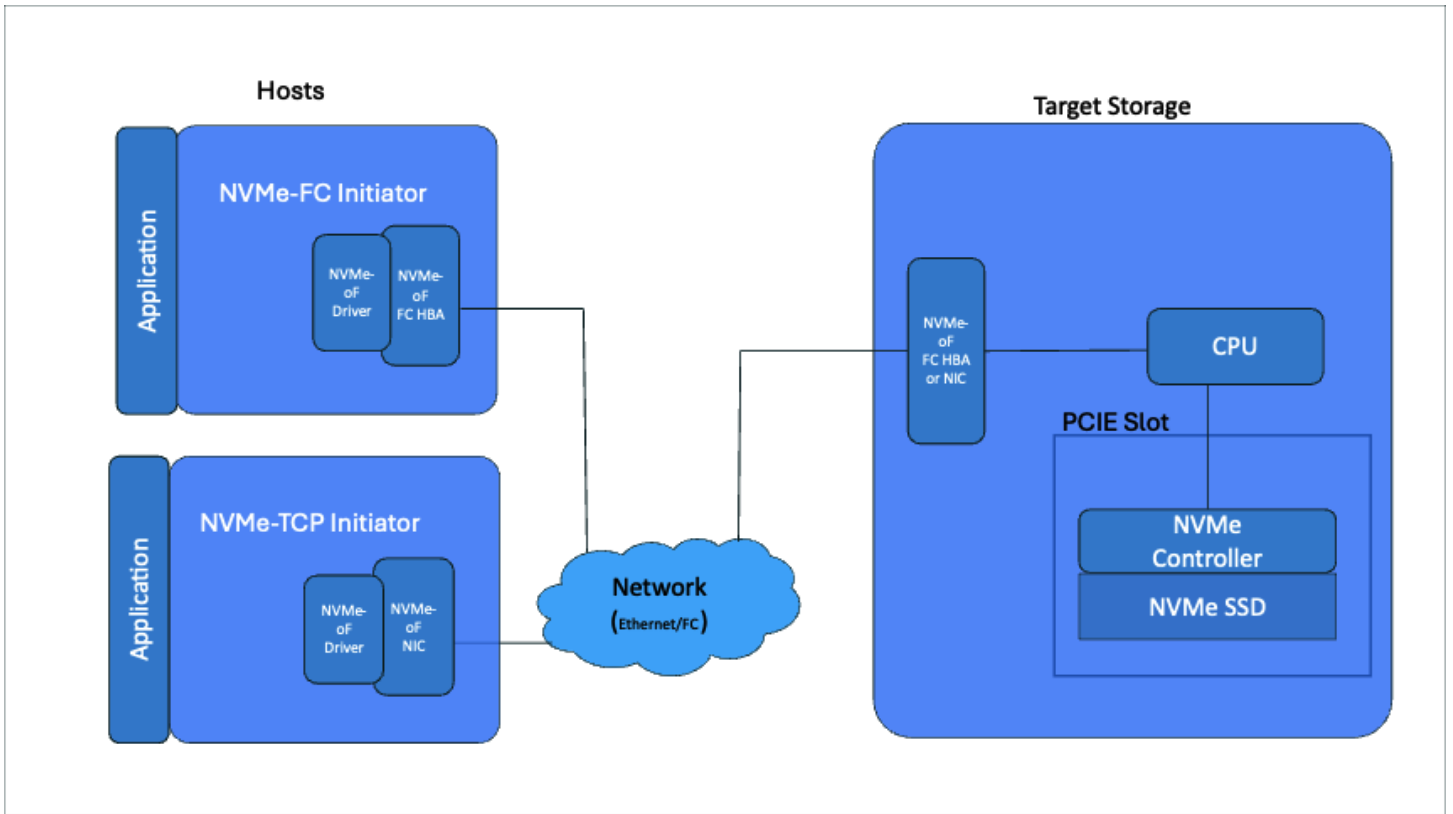
Available in multiple variants including NVMe/RoCE (RDMA over Converged Ethernet), NVMe/IB (InfiniBand) and NVMe/iWARP, allowing deployment across diverse network infrastructures.



Targeted use cases

Suited for AI/ML pipelines, HPC environments, and enterprise applications requiring deterministic performance and low-latency data access.

NVMe-oF architecture overview



NVMe over PCIe represents the native implementation of the NVMe protocol, enabling direct communication between host systems and solid-state drives (SSDs) via PCI Express lanes. This architecture bypasses legacy

storage protocols such as SATA and SCSI, delivering ultra-low latency and exceptional throughput for performance-critical workloads.



NVMe-oF initiator

The host system initiates storage requests, typically from applications or services requiring data access.



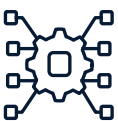
Application layer

Represents the workload or service generating I/O operations, such as databases, analytics engines, or virtual machines.



NVMe-oF driver

Translates application-level I/O requests into NVMe commands optimized for transport across fabric networks.



Network interface (NIC/HBA)

Ethernet NICs, RDMA-capable adapters (RoCE/iWARP), or Fibre Channel HBAs transmit NVMe commands over the selected fabric.



Network fabric

The transport medium—Ethernet, RDMA, or Fibre Channel—used to carry NVMe commands between initiators and targets with minimal latency.



NVMe-oF target

The destination storage endpoint that receives and processes NVMe commands from the host.



NVMe controller

Manages incoming NVMe commands and orchestrates data access to SSDs, ensuring efficient I/O handling.



Target-side NIC/HBA

Similar to the initiator side, this interface receives NVMe commands from the network and passes them to the controller.



NVMe SSDs

High-performance storage media designed for intensive read/write operations, delivering ultra-low latency and high throughput.



NVMe – comparative analysis

Feature	NVMe over PCIe (Direct)	NVMe over Fabrics (RoCE/iWARP)	NVMe over TCP	NVMe over Fibre Channel (FC-NVMe)
Transport medium	PCIe bus (local)	Ethernet (RDMA)	Ethernet (standard TCP/IP)	Fibre Channel network
Use case	Direct-attached SSDs	High-performance SANs	Broad compatibility over IP	FC-based data centers
Latency	Ultra-low (sub-10µs)	Very low (10-50µs)	Moderate (~100µs)	Low (~100µs)
Throughput	Extremely high	Very high	High	High
CPU overhead	Low	Low (offloaded via RDMA)	Higher (TCP stack)	Low (dedicated HBA)
Hardware requirement	NVMe SSD + PCIe lanes	RDMA-capable NIC (RoCE/iWARP)	Standard Ethernet NIC	FC HBA (16G/32G/64G)
Scalability	Limited (local only)	High (over fabric)	High	High
Network fabric cost	N/A	High (RDMA switches, NICs)	Low (standard networking)	Medium to high
Software support	Native in OS	Requires RDMA stack	Supported in Linux and Windows	Supported in modern SANs
Best-fit deployment	Servers, laptops, HCI nodes	HPC, AI/ML, storage arrays	Cloud, edge, general IT	Enterprise SANs, legacy FC upgrades

Sustainability benefits



- **Energy efficiency and thermal optimization**

NVMe SSDs are engineered for high performance with significantly lower power consumption and heat generation than traditional HDDs and legacy SSDs. This translates into reduced energy costs and cooling requirements in data centers, contributing to operational efficiency and environmental sustainability.

- **Lower carbon footprint**

According to industry assessments, NVMe SSDs exhibit a reduced embodied carbon footprint per terabyte, reflecting a more environmentally conscious manufacturing process than older storage technologies.

- **Hardware consolidation and material reduction**

The superior performance of NVMe enables organizations to meet workload demands with fewer devices, minimizing material usage in production and reducing electronic waste. This supports circular economy principles and long-term sustainability goals.

- **Enhanced efficiency and application responsiveness**

NVMe's high-speed, low-latency architecture accelerates data processing and optimizes compute resource utilization. This leads to faster application performance, shorter processing cycles, and lower energy consumption across workloads.

- **Scalability and infrastructure longevity**

Designed for future-ready deployments, NVMe supports scalable architectures that adapt to evolving storage demands. This reduces the frequency of hardware refresh cycles, minimizing resource waste and supporting sustainable IT practices.

NVMe operations management consideration

1101100
10,001
1001101

Host multipathing configuration

Host team needs NVMe-aware drivers and multipathing configurations (e.g., Linux NVMe-cli, Windows NVMe drivers), which differ from legacy FC setups

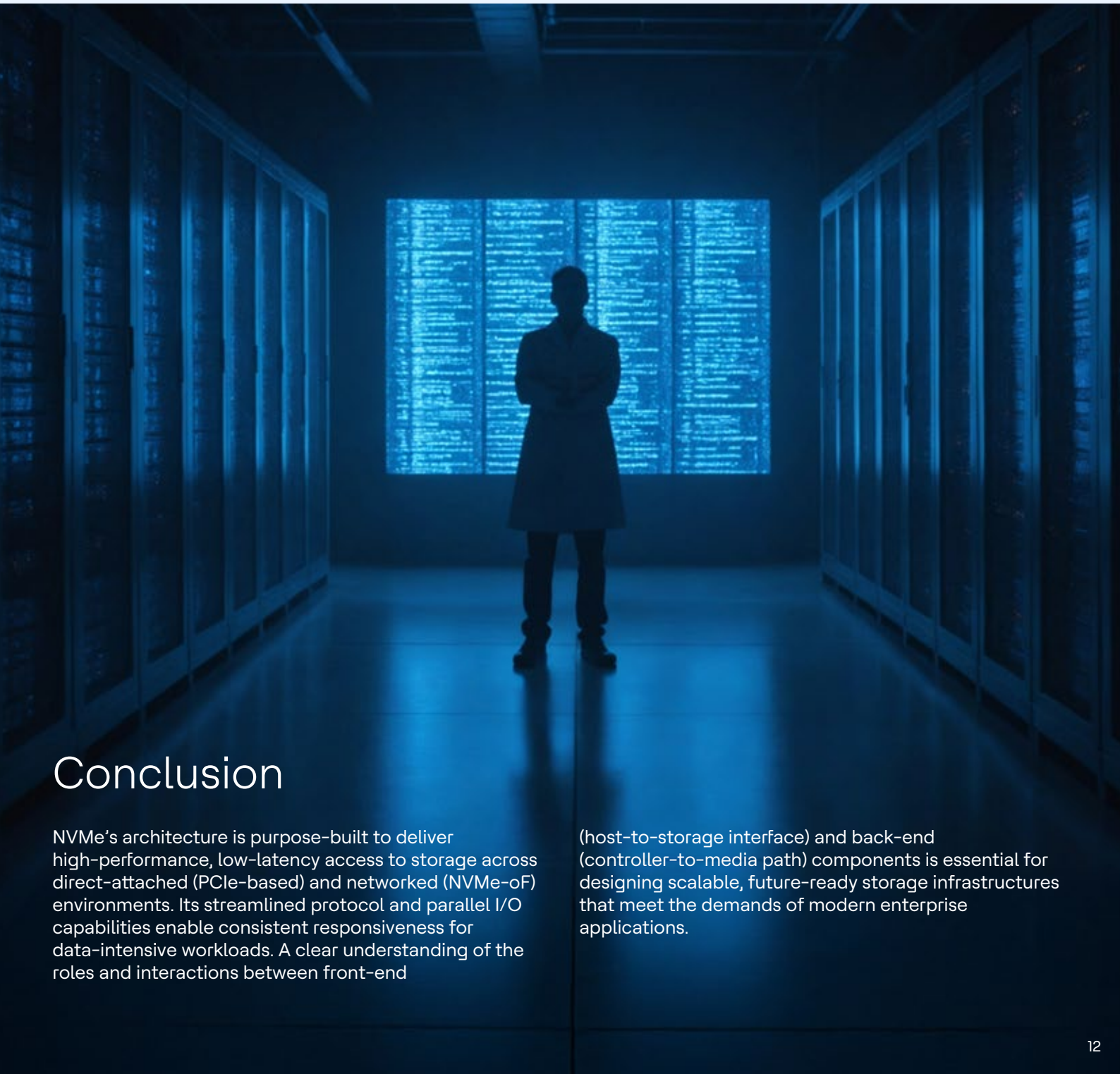


Monitoring and troubleshooting

NVMe/FC introduces new counters and diagnostic tools (e.g., NVMe queue depth, namespace mapping) that require additional skills for NVMe storage management.

Operational efforts comparison with traditional FC vs NVMe/FC

Key Aspects	Traditional FC	NVMe/FC	Complexity Impact
Protocol	SCSI	NVMe	Moderate
Zoning	WWN/TDZ	WWN/TDZ	Same
Multipathing	MPIO	NVMe-MPIO	Slightly Different
Performance	Good	Excellent	Worth the shift
Monitoring	Legacy tools	NVMe-aware tools	Requires update



Conclusion

NVMe's architecture is purpose-built to deliver high-performance, low-latency access to storage across direct-attached (PCIe-based) and networked (NVMe-oF) environments. Its streamlined protocol and parallel I/O capabilities enable consistent responsiveness for data-intensive workloads. A clear understanding of the roles and interactions between front-end

(host-to-storage interface) and back-end (controller-to-media path) components is essential for designing scalable, future-ready storage infrastructures that meet the demands of modern enterprise applications.



Rajan Pillay

Vice President and Head of the
Global Center of Excellence,
Hybrid Cloud Business Unit,

HCLTech

About the Author

With over 28 years of experience, Rajan has held pivotal roles such as Practice Head, Cloud CoE Leader, Solutions Architect, and Offering Leader, contributing across IT architecture, design engineering, operational engineering, resiliency and optimization. He is known for building and leading high-performing technology teams comprising architects, engineers, SREs and SMEs. Rajan aligns cloud strategy roadmaps with customer business goals, emphasizing agility, scalability, and cost efficiency. He drives large-scale hybrid cloud transformation programs and oversees application and infrastructure managed services post-deployment, ensuring SLA-driven operations. Additionally, he leads the deployment of scalable AI and GenAI platforms, enabling intelligent enterprise modernization.



Virendra Singh Shekhawat

AGM and Lead,
Storage and Backup CoE,
Hybrid Cloud Business Unit,

HCLTech

About the Author

Virendra is a seasoned IT infrastructure and hybrid cloud leader with over 18 years of cross-functional experience spanning solution architecture, presales, transformation, consulting, and migration. He has successfully delivered strategic cloud and data center solutions across banking, healthcare, retail and manufacturing sectors. His core expertise lies in Storage, Backup, Disaster Recovery (DR) and Cyber Recovery (CR), with a strong focus on designing resilient, scalable and secure infrastructure strategies.

HCLTech | Supercharging Progress™

HCLTech is a global technology company, home to more than 226,600 people across 60 countries, delivering industry-leading capabilities centered around AI, digital, engineering, cloud and software, powered by a broad portfolio of technology services and products. We work with clients across all major verticals, providing industry solutions for Financial Services, Manufacturing, Life Sciences and Healthcare, High Tech, Semiconductor, Telecom and Media, Retail and CPG, Mobility and Public Services. Consolidated revenues as of 12 months ending September 2025 totaled \$14.2 billion. To learn how we can supercharge progress for you, visit hcltech.com.

hcltech.com

