

White paper

Software RAID Reveals Opportunities of the NVMe for Enterprise-Grade Tasks



Contents

Executive Summary	3
NVMe Protocol Opportunities and Limitations	3
High Speed and Low Latency	3
NVMe-oF for Network Access	4
Software RAID for NVMe	4
Fast RAID for Linux	5
RAIDIX ERA vs Other Solutions	5
RAIDIX ERA For Enterprise Applications	6
Databases for Analytics and R&D	6
Online Transaction Processing (OLTP)	6
8K Video Production	7
Edge Computing	7
Conclusion	7

Executive Summary

NVMe protocol has significantly expanded data centers opportunities. New drives not only keep all features of the traditional SSDs, but also bring a brand new level of performance and latency.

Due to their functional characteristics, NVMe drives are the right choice for the range of top-notch technologies. OLAP, OLTP, ML & AI, advanced ERP applications, and autonomous production lines still demand a lot of long-term investments, but with NVMe that becomes easier and more affordable.

At the same time, there are some tangible reasons that restrain NVMe drive wide distribution. All of them are very substantial for the storage vendors and system integrators. For NVMe in enterprise-grade data centers two things are crucial: the ability to work within fault-tolerant RAID for data protection and the opportunity to use NVMe-oF as a transport to network-attached drives.

Existing software or hardware RAID technologies either use excessive redundancy in array mirroring or lose up to 65% of the performance in RAID 5 or RAID 6. And not all of these solutions have an ability to support NVMe-oF for network access.

RAIDIX ERA is a software RAID developed with close attention to NVMe features. RAIDIX innovative technologies and focusing on the new protocol's strengths make our software RAID engine capable of demonstrating up to 97% of the total drives performance in RAID 6 and lower latency even for mixed workloads.

In this material we will outline the key NVMe features, evaluate its strengths and weaknesses, and present RAIDIX ERA opportunities and scenarios for work with different applications within enterprise-grade storage infrastructures.

NVMe Protocol Opportunities and Limitations

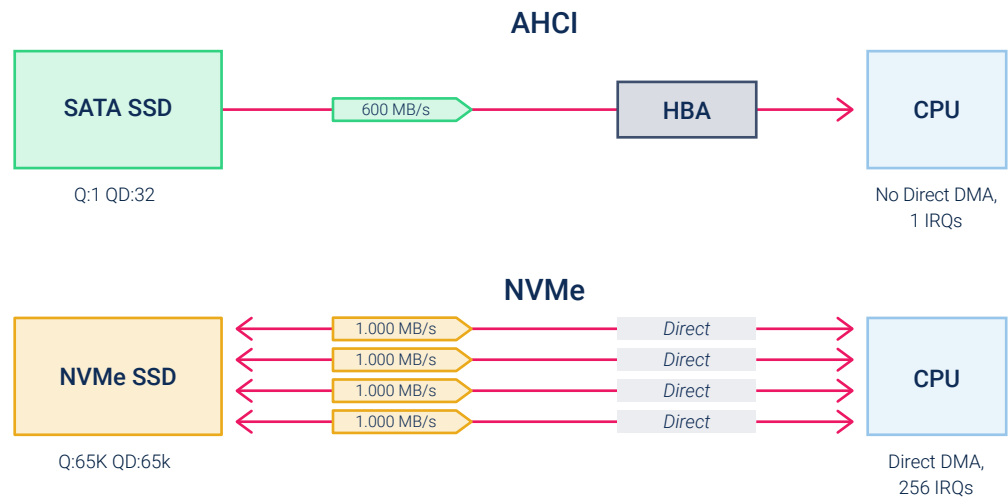
NVMe (Non-Volatile Memory Express) is the first protocol designed specifically for solid state drives. Traditional SAS and SATA were originally developed for HDD complex internal mechanisms. These protocols employ HDD service command set for correct and reliable work of all moving parts within the hard drive.

For the flash memory all these commands are too complicated and redundant. SAS and SATA protocols get access to the data through the hard drive native procedures, ignoring physical traits of the flash memory. NVMe in contrast uses flash drives physics to increase total performance of the media.

High Speed and Low Latency

NVMe is built over the PCIe interface. The protocol natively employs the PCIe command set used by applications to communicate with CPU. In this approach, the HBA stage between drives and CPU is eliminated and the data path is reduced by eliminating SAS/SATA commands adaptation for PCIe bus.

Figure 1.
Comparison
of the AHCI
and NVMe interfaces



Moreover, NVMe drive controller is able to transfer data to CPU using more than 65K queues and 65K commands/queue. SATA/SAS interfaces support only one queue with 32 and 254 commands respectively.

This allows NVMe SSDs to demonstrate two times less latency and several times more read and write speed for mixed and random workloads than traditional SSDs. That's why NVMe drives form the good basis for such speed-demanding technologies like AI & ML, OLTP, Edge computing and others.

NVMe-oF for Network Access

NVMe SSDs have undoubted advantages over SATA/SAS SSDs by means of its direct connection to the server with PCIe bus. But local drives usage is suitable for a very limited range of business tasks. Most of the enterprise-grade data centers need the high-available solutions with flexible scaling and resource allocation.

To review all points of the technology, we should look at this in the context of network environments, where common past-gen protocols eliminate all NVMe performance advantages. The solution for this issue is NVMe-oF extension for NVMe, which allows to use the new protocol with the Ethernet, Fiber Channel and InfiniBand network transports.

NVMe-oF connects application on the server and remote NVMe array at almost local latency rate. This technology is also expected to keep high performance of the NVMe drives for the network infrastructure, but now there is no chance to evaluate it, because modern network adapters still work as a bottleneck for the whole data path.

But anyway, to implement NVMe technology to the enterprise data centers, NVMe-oF support in hardware and software components of existing storage infrastructure is necessary.

Software RAID for NVMe

Talking about NVMe usage beyond the market of consumer devices, we can see that almost all scenarios need a fault-tolerant RAID to protect data in case of drive failure. This can be achieved by using either a RAID engine in the software layer or hardware RAID controllers with direct drives connection. And RAID 5 or RAID 6 will be the most cost-effective choice here.

But RAID parity always consumes system computational resources, so total array performance is often significantly less than sum performance of the used drives.

Also, in RAIDIX ERA we've successfully solved the main challenge of software RAIDs – an excessive consuming of computational resources. Under maximum array loading it shows less than 20% CPU usage, and for full-featured work it needs only 4GB RAM. To make system utilization more effective, RAIDIX ERA provides manual RAM consuming settings.

	Hardware RAID controller	Software RAID	RAIDIX ERA
Performance	Moderate	Moderate	High
Hardware purchase	Yes	No	No
Physical wearing-out and obsolescence	Yes	No	No
Unlimited number of drives	No	Yes	Yes
NVMe-oF support	No	Yes	Yes
RAM consuming	–	High	Low
CPU consuming	–	High	Low
Priority settings for the computational resources	–	No	Yes
Compatibility with any server hardware	No	Yes	Yes
Compatibility with any flash drives from any vendor	No	Yes	Yes
Simple embedding into existing solution	No	Optional	Yes

RAIDIX ERA For Enterprise Applications

RAIDIX ERA makes a highly performant and fault-tolerant array from NVMe drives. It's a perfect choice if you need to run demanding enterprise-grade applications which require low latency and high bandwidth of your back-end infrastructure.

Databases for Analytics and R&D

Analysis of big volumes of data is a key part of the effective decision making in modern enterprises and for conducting massive scientific projects.

Most analytical applications for these purposes have a typical workload pattern to the storage: small and regular write requests and large data read requests. Due to high bandwidth (up to 55GBps) RAIDIX ERA can deliver data payloads to analytical applications faster than other solutions. That reduces data processing time and improves total workflow efficiency.

Online Transaction Processing (OLTP)

OLTP applications are valuable technological solutions in trading, finance, automatic production lines in large manufactures, enterprise-resource planning systems and others. Effective work of these applications is very sensitive to a backend infrastructure performance. It demands extremely low latency and high speed IO processing.

NVMe array powered by RAIDIX ERA as a backend device delivers minimum response time for intensive transactional requests and allows applications to make more IO per second. At the same time, it keeps a high performance rate even in mixed workloads.

8K Video Production

8K resolution, HDR and increased up to 60fps video frame rate significantly raise the requirements to technical equipment of the post-production studios. In this environment, video storage is a bottleneck of the whole infrastructure, as the bandwidth has not only an impact on the production speed and general convenience of the workgroup, but also on security from frame drops during the editing process.

RAIDIX ERA works as a storage subsystem for video production, providing high read/write speed for simultaneous work from many stations. RAIDIX ERA usage provides creative teams with smooth uninterrupted video editing with no risks of frame losses.

Edge Computing

Edge Computing is one of the most popular ways to perform industrial IoT (Internet of Things). In this approach, computational resources of the system are deployed in immediate proximity to the sources of data – the sensors of the active unit. As an example, Edge Computing works in autonomous driving solutions, feeding traffic data to AI applications with extremely low latency.

RAIDIX ERA is used as a part of mobile computational platforms for autonomous vehicles. It manages an NVMe array and provides the lowest response time and robust performance even with a limited number of drives.

Conclusion

NVMe protocol works over PCIe bus, natively using a command set more appropriate for flash memory. It gives an opportunity to transfer data using 65K parallel queues and avoid extra blocks and barriers working with app requests.

To save performance potential of NVMe drives in fault-tolerant array, we've developed RAIDIX ERA – powerful software RAID which demonstrates high speed and low latency even in mixed IO workloads. At the same time, it easily can be embedded in software environments of existing solutions and is compatible with a wide range of drives and server platforms.

For enterprises, RAIDIX ERA unlocks the opportunities of top notch technologies with no extra investments in excessive infrastructure hardware. For storage vendors and system integrators, RAIDIX ERA is a simple and affordable way to get technological advantage for building NVMe solutions.