TRENTON SYSTEMS

JBOF SYSTEM

Rugged High-Performance Computing

WHITEPAPER

24 NVME U.2 DRIVES IN A 2U HIGH
16.2" DEEP, SWAP-OPTIMIZED CHASSIS

BY MICHAEL BOWLING

KEY TAKEAWAYS

PERFORMANCE:

27 GB/s read/write bandwidth

8 drives communicating simultaneously at x4 speeds

All native PCIe Gen3 to 24 NVMe drives

CPU AGNOSTIC:

Two PCIe Gen3 cables from your host computer(s)

Load balance between two Xeon® CPUs

RUGGED:

16.2" short-depth

Tool-less magazines with hot-swap capability

APPLICATIONS:

Local, high-performance

Ideal for Military, AI, Medical, Machine Learning

ABSTRACT

Recent advances with NVMe hard drives are driving a demand for high-performance storage.

Trenton Systems has developed a short depth (16.2″), 2U (3.5″) high hard drive array that tool-lessly holds 24 NMVe U.2 drives (8 drives per cartridge).

There are several things that truly set this product apart, but first and foremost is the performance. Through the use of virtual mode through the PCIe switches, Trenton Systems enables communication to 8 drives simultaneously (all through x4 PCIe Gen3 links). The tested throughput is 27GB/s (8 drives communicating at 3.375GB/s each).

The JBOF system uses two PCIe Gen3 expansion cables to drive the storage array which are cabled from one or more servers (up to 3 meters in distance). This enables all native PCIe connections from the CPU directly (through one of two PCIe switches) to the NVMe drives. There is no protocol translation to/from SATA or SAS which increases your effective throughput.

Also, by using PCIe expansion cables, it allows the user to utilize any host processor board that they would like. Later in this whitepaper we will show some block diagrams which will illustrate best use cases to load balance the bandwidth and get the most out of your JBOF system.

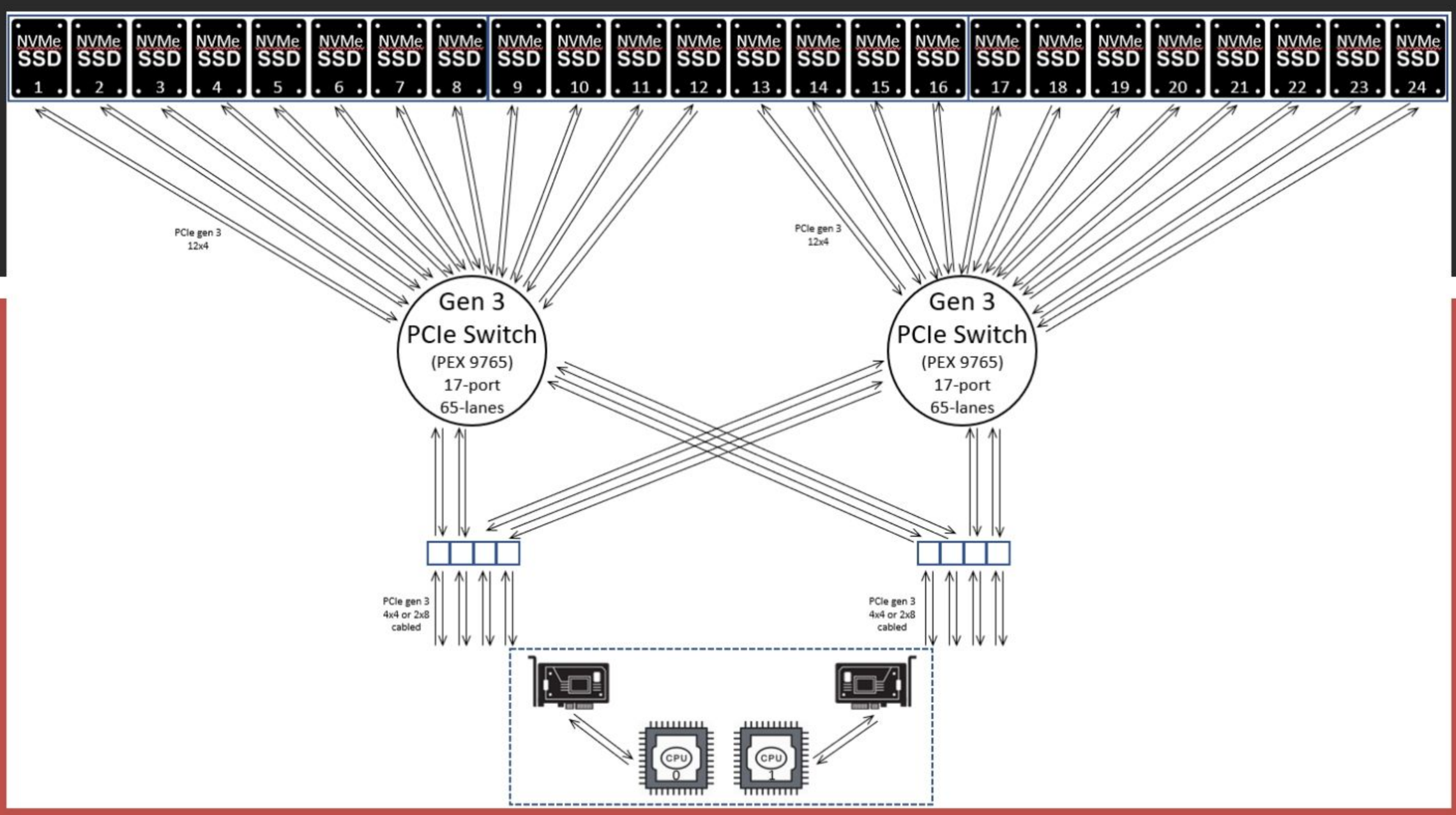
Trenton Systems JBOF solution is designed with military applications in mind. The quick and easy removal of the actual drives allows any operator easy removal and storage of sensitive data on the drives.

Yet, the combination of performance, short-depth, rugged, and configuration flexibility is ideal for any applications that require very high-performance, local storage; such as Military, AI, Machine Learning, Medical, etc.

JBOF SYSTEM

ELECTRICAL

The figure below shows the high-level block diagram of the JBOF system. Note that the PCIe is cabled into the system at the bottom of the diagram (eight x8 PCIe gen 3 links connected to two CPUs). Two links from each cable go to each switch which provides a PCIe mesh that enables multiple paths to each drive. After the PEX 9765 switches, each drives has a dedicated x4 gen 3 link. It is noteworthy that a x4 PCIe gen 3 link has up to 4.0GB/s bandwidth and the higher performance NVMe drives have up to 3.2GB/s performance. This extra bandwidth overhead is important so that each drives does not saturate the PCIe link.

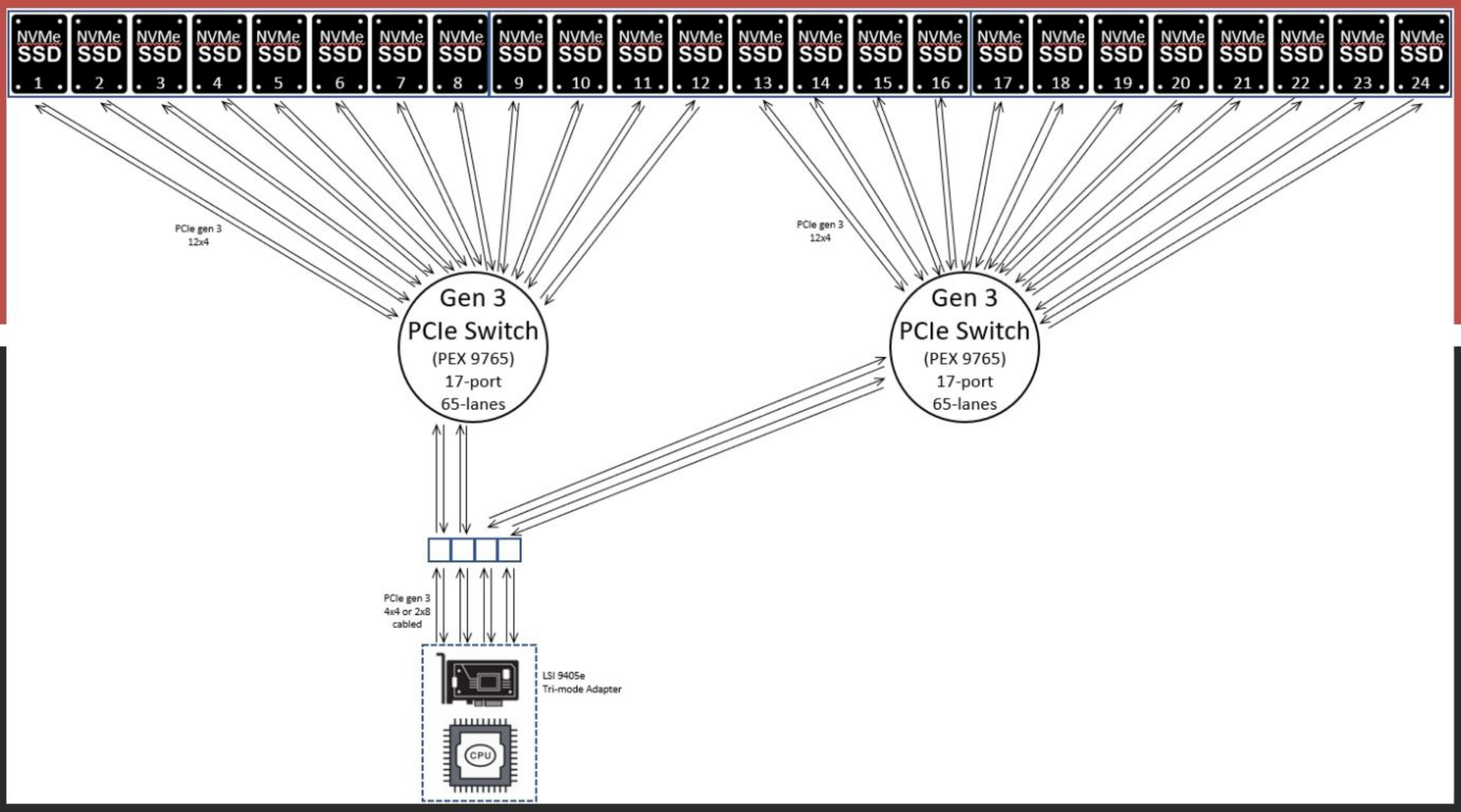


The rectangle at the bottom of the above diagram shows the two Intel Xeon CPUs each connected to a PCIe host expansion board. The host cards could be plugged into almost any rackmount or desktop class computer which makes this JBOF truly agnostic to the host computer. Note that some consumer-class computers may have BIOS limitations (i.e. memory space) so it is important that you consider the proper host processor board.

CONTINUED...

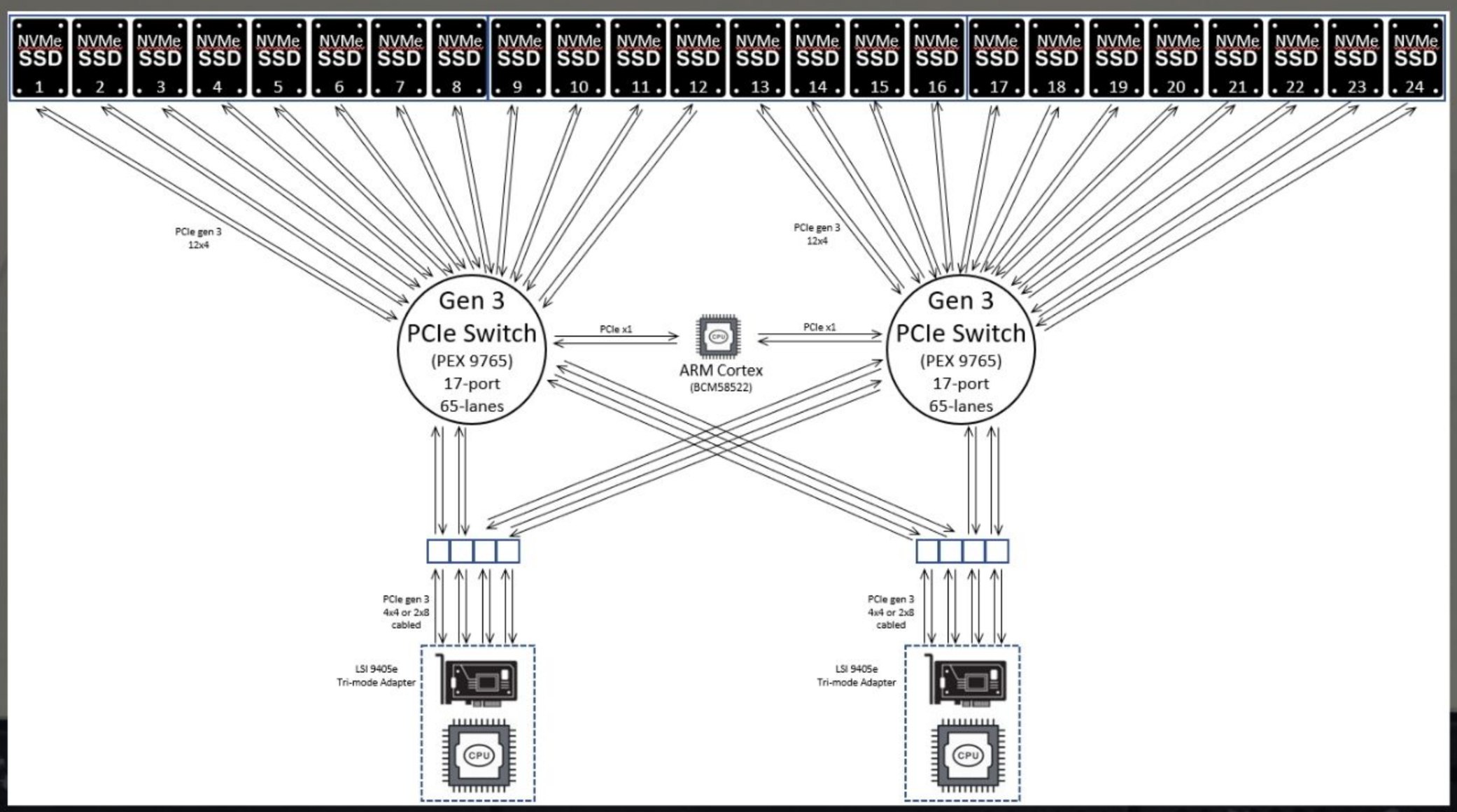
By balancing the PCIe host cards between the two CPUs the application is effectively load balancing the bandwidth. **Note** that there are multiple paths for either CPU to communicate to any of the 24 drives (the mesh connectivity to the switches and/or the QPI bus between the host CPUs). This is important for applications that have sustained communication to particular drives which may saturate some paths.

The low latency PCIe switches (PEX 9765) enables multiple virtual switches. Each switch has up to four virtual switch partitions which is what allows 8 drives to effectively have their own link to the host CPUs. **In addition** to the performance increase, this feature also allows for 1+1 host failover (for example, if one of the host expansion cables is removed).



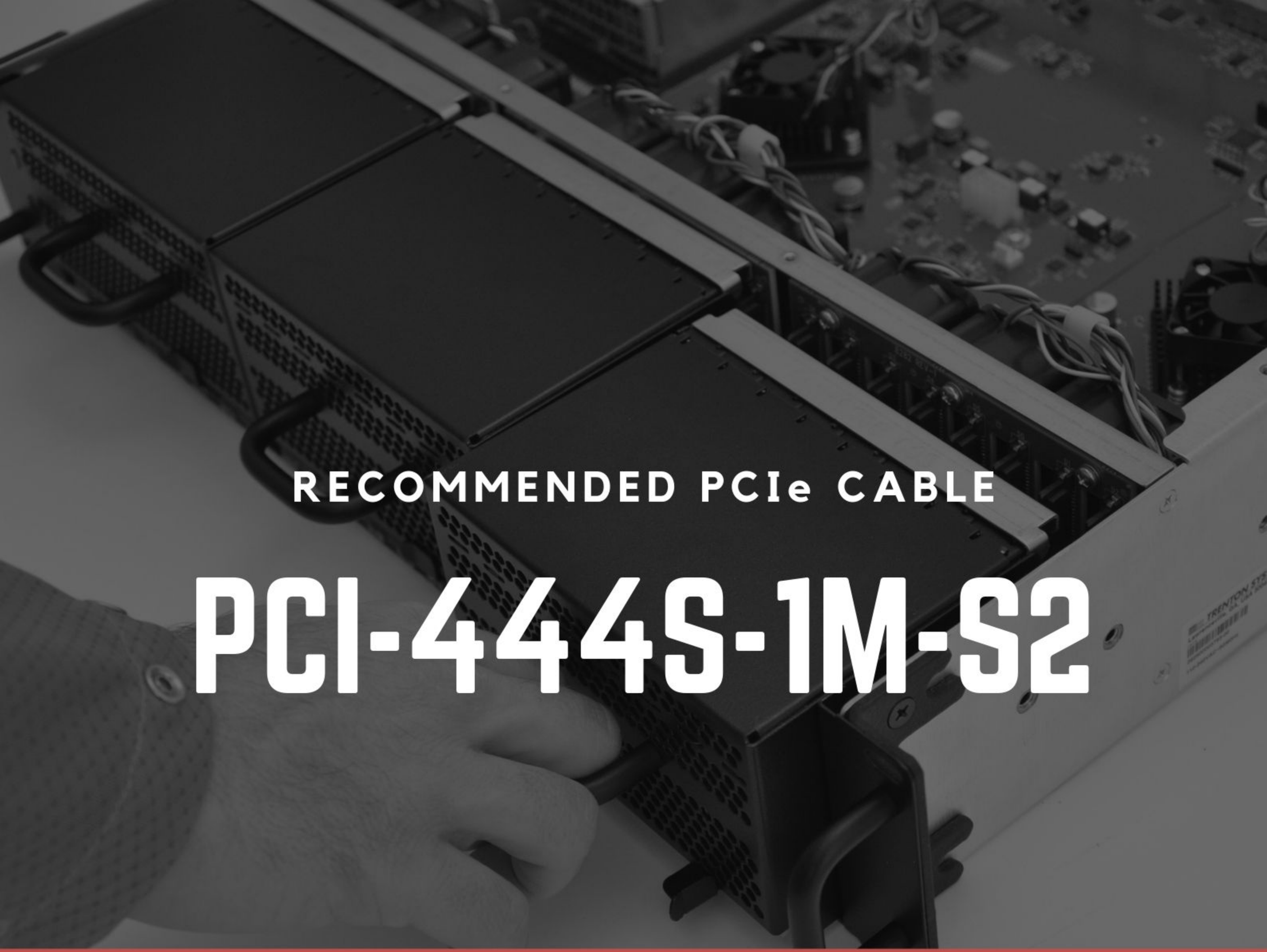
There are several other use cases for the Trenton Systems JBOF System. Here you can see an example where the host computer only has one PCIe expansion card. While not as ideal for bandwidth or redundancy, this lowers the total cost of the system.

ENABLING MULTIPLE, SEPARATE COMPUTERS (DIFFERENT ROOT COMPLEXES) WITH ACCESS TO THE SAME JBOF DOES REQUIRE SOME SOFTWARE, BUT TRENTON SYSTEMS HAS INCORPORATED THIS SOFTWARE **DIRECTLY INTO THE JBOF** THROUGH THE USE OF AN ARM CORTEX FOR THE ARBITRATION AND USING ONE CONVERGENCE CODE FOR THE FABRIC STACK. THE PEX 9765 SWITCHES ARE PUT INTO NON-TRANSPARENT MODE AND THE ARM CPU HANDLES ANY BUS ARBITRATION ISSUES.



While the max theoretical throughput is 27GB/s (8 drives running at 3.375GB/s each), actual results depend on many different real-world factors (i.e. packet size, load balancing between multiple CPUs, drive performance, etc). For many of Trenton Systems tests, we have used the Intel DC P4510 U.2 NVMe drives. These Intel drives are ideal for applications where high-performance and reliability is required, but other U.2 NMVe drives would also perform well in this Trenton Systems JBOF system.

CONTACT TRENTON SYSTEMS AT **(770) 287-3100** FOR ACTUAL BENCHMARK RESULTS



RECOMMENDED PCIe CABLE

PCI-444S-1M-S2

The 1 meter long cable is ideal for PCIe Gen3 speeds (8GHz), is easily removed/serviced (yet also can lock into the mating connector).

Trenton Systems has tested the cable with both LSI 9405e and Dolphin ICS MXH 832 host cards.

Get in touch for actual eye diagram test results

JBOF SYSTEM MECHANICAL

The JBOF is a standard 2U high rackmount system, but the 16.2” depth is unique for these types of storage applications. The chassis is made of light-weight aluminum while also shock and vibe tested for rigidity. The power supply in the system is 900W and removable for easy serviceability. Ideal for SWaP (Size, Weight and Power) applications where the rackmount system must fit into a small environment.

There are 3 magazines each holding 8 drives. There are no screws required to remove the magazines OR to remove each individual drive. [Note that there is another JBOF in development at Trenton Systems that removes the magazines and each drive can individually be removed and that system will be available in 2019.] These tool-less magazines are ideal for applications where the data much be easily and quickly removed and secured.

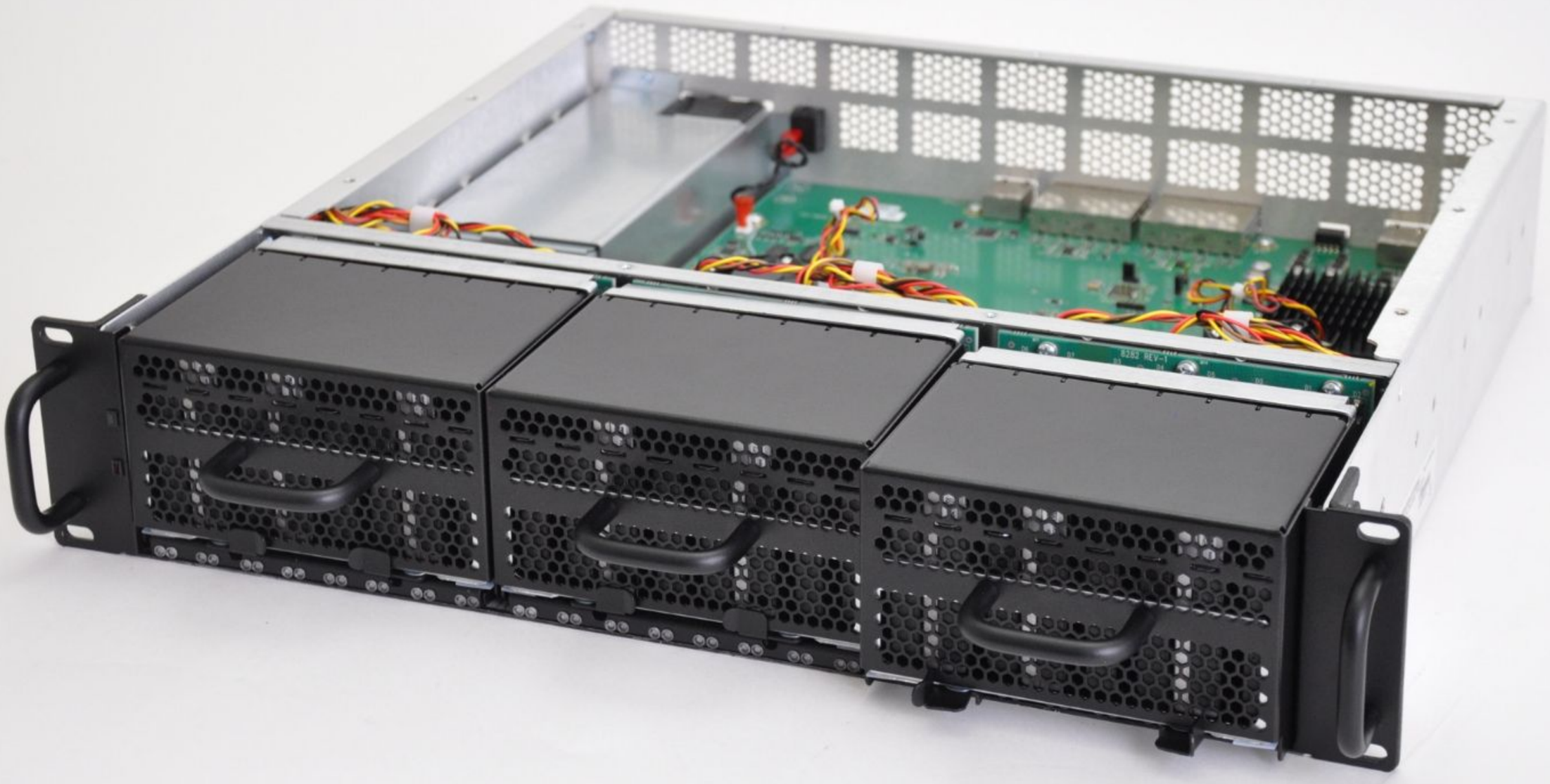
SHOCK & VIBE RESULTS

SPEAK TO ONE OF OUR ENGINEERS FOR ALL TESTS AND POWER CALCULATIONS

Applications

Ideal for military applications such as transit cases deployed in the field, aircraft or shipboard applications that have cabin space constraints, or security applications that require made in US systems that can be quickly and tool-less removal of the drives.

Many applications have a need for high-performance, local, native PCIe, NVMe JBOF storage that need to manipulate high traffic data very quickly and reliably (AI, machine learning, medical, etc).



CONCLUSION

There is a growing need for high-performance, local storage in many markets. The performance combined with the short depth, rugged, yet lightweight system provides a unique solution. The Trenton Systems JBOF system is also unique and flexible to many applications due to the PCIe expansion capability which means the storage solution is agnostic to almost any host computer that the user needs. Native PCIe from CPU to each NVMe drive in combination with 8 drives communicating in parallel provides a high-performance solution that is unmatched in the industry.

TRENTON SYSTEMS®