

NVMe-oF Storage for Splunk

Accelerating Splunk with no-compromises

Splunk's Storage Problem

Today, Splunk storage uses legacy networked storage arrays in the hot and warm storage tiers. As data grows, applications begin to crawl and customers turn to direct-attached SSDs inside the indexer nodes to satisfy performance.

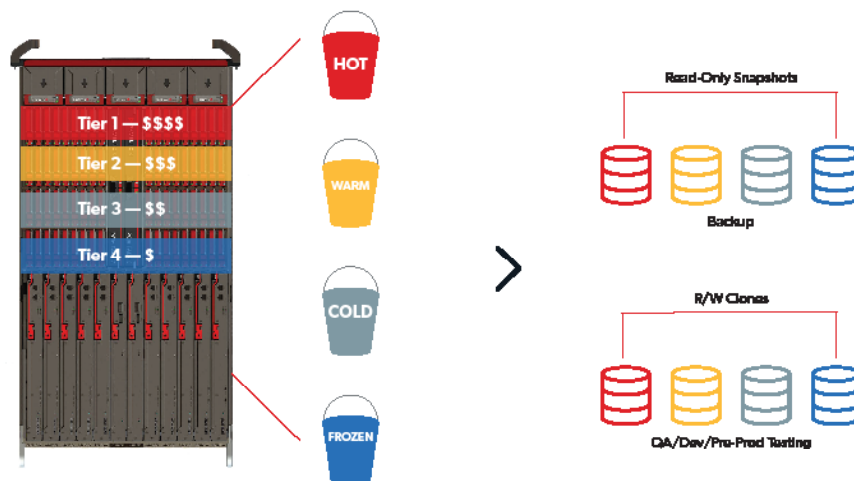
This approach has resulted in stranded storage capacity in each indexer node, leading to inflexibility and severe under-utilization of storage resources. When there is a need for growth in storage capacity, it requires additional indexer nodes even when additional indexer performance is not required.

A more efficient and cost-effective solution is needed, one with the low latency offered by direct-attached SSDs but with the economics and flexibility of networked storage.

The Pavilion HyperParallel Data Platform™

The Pavilion HyperParallel Data Platform delivers never before seen NVMe performance (120 GB/s throughput, 25µs of latency, and 20M IOPS) and density (2.2 Petabyte) in a compact 4U form factor. It provides Spark applications with the performance of locally attached NVMe SSDs, enabling Splunk operations to move to a Composable, Disaggregated Infrastructure (CDI) infrastructure, where application resources are readily available. The Pavilion HFA uses NVMe SSDs and supports multiple block and file protocols with NVMe-oF/RDMA, NVMe-oF/TCP, iSCSI, S3 and NFS.

The Pavilion HyperParallel Data Platform requires no proprietary software to be installed on a server farm and uses standard NVMe, NVMe-oF, Ethernet, and InfiniBand drivers, freeing up host resources for processing and reducing deployment risk.



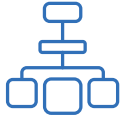
Benefits

- Increases density 2X
- Petabyte scalability, high-performance, low-latency, and linear scaling maximizes Splunk efficiency
- Protect for your data and your business. Meet evolving requirements for data security and compliance
- Double indexing rate and reduce indexer footprint by up to 40%
- Deployment flexibility using concurrent protocols
- Enterprise design and data integrity validation ensure reliable access to data
- Rack scale/CDI management via Web GUI, vCenter, Kubernetes, RESTful API, OpenStack, DTMF/Redfish, and Swordfish

Benefits of Disaggregation

The Pavilion HyperParallel Data Platform delivers high performance and low latency to Splunk applications.

Scalable & Flexible



The Pavilion HyperParallel Data Platform provides up to 2.2 Petabytes to Splunk deployments, which can simultaneously use its NVMe/RDMA (Ethernet, IB), NVMe/TCP, iSCSI, NFS (v3, v4, pNFS, RDMA) and S3 interfaces. Splunk says that disk I/O and throughput are the most common limitation for Splunk infrastructures and the ultra-high performance and ultra-low latency of the Pavilion HyperParallel Data Platform can power even the largest deployment. The array scalability allows Splunk users to focus on the needs for compute instead of storage.

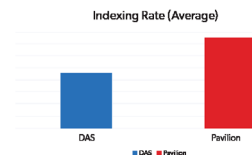
The Pavilion HyperParallel Data Platform thin provisioning feature provides Splunk with significantly better utilization than DAS, since the Pavilion array will only allocate physical space when the Splunk application needs it. This results in physical space savings vs. DAS of up to 75% per server.

Economical



The Pavilion HyperParallel Data Platform is ideal for Splunk and for organizations deploying a few indexers to hundreds. By deploying the platform, customers can take advantage of thinly provisioned shared pools with the same or better latency as direct-attached SSDs.

If more CPU and memory is needed, administrators can seamlessly deploy additional indexers to handle the search traffic, with the hot and warm tier using only the needed capacity. The storage capacity and performance can be independently scaled to expand the size of the hot and warm tiers.



The Pavilion HyperParallel Data Platform's ultra-low latency removes storage as the indexing bottleneck. Lab testing shows that Splunk Enterprise can index 67% more transactions using the Pavilion HyperParallel Data Platform than with direct-attached SSDs. Testing also showed that 40% fewer indexers were required to support the same ingest volume reducing TCO.

With Pavilion, you are no longer constrained by the size of the SSDs. Thin provisioning allows the application to use the required amount of storage at any given time, regardless of how much capacity has been advertised to that specific Splunk server. With no more extra copies of data, the Pavilion HyperParallel Data Platform reduces the amount of raw flash storage deployed in a Splunk environment, by up to 3X, reducing TCO. The Pavilion HyperParallel Data Platform is designed not just for exotic applications, but is the one platform for all your Splunk needs.

Fast & Dense



The ultra-high performance, extreme low latency, and multiple storage controllers of the Pavilion HyperParallel Data Platform accelerates Splunk workflows and boosts time to results. The Pavilion HyperParallel Data Platform lets Splunk applications search terabytes of structured and unstructured data over 20 times faster than a DAS NVMe SSD.

Safe & Secure



Protect the SSDs with RAID-6 and Pavilion SwarmController, which rebuilds a failed SSD at the rate of less than 5 minutes per TB. The Pavilion HyperParallel Data Platform ensures that the failure of an SSD does not impact Splunk operations.

Keep data and snapshots safe with a FIPS-compliant data at rest encryption, Pavilion's always-on encryption keeps data secure without impacting Splunk performance.

Pavilion data assurance features provide self-healing bit-rot support for data, assuring every workflow gets uncorrupted data. Take consistent, instant, zero-footprint, and uncorrupted snapshots, encrypt them and provide them to backup and disaster recovery processes to speed operations and ensure consistency.

Enterprise Strength & Enterprise Support



Get 24/7 proactive support, end-to-end data integrity, a modular chassis, and redundancy throughout the storage array to protect your Splunk applications as infrastructures grow. With the Pavilion HyperParallel Data Platform, all the features come in-the-box, including thin provisioning, snapshots, clones, data at rest encryption and more.

The ultra-high performance, extreme low latency, and multiple storage controllers of the Pavilion HyperParallel Data Platform accelerates Splunk workflows and boosts time to results. Pavilion lets Splunk applications search terabytes of structured and unstructured data over 20 times faster than a DAS NVMe SSD.

Disaggregating flash storage from each server in a rack means no longer dedicating resources to managing local SSDs. Using servers that don't have to support DAS storage can increase the compute density of a rack by 2X and reduce power and cooling requirements. No custom software needs to be installed on Splunk servers, enabling them to take full advantage of host processing resources and reduce risk. The result? The power, simplicity, and density offered by the Pavilion HyperParallel Data Platform provides Splunk deployments increased agility, flexibility, and lower TCO.

Find Out More

Pavilion shatters customer expectations and resulting organizational outcomes by revolutionizing data processing for modern AI/ML, HPC, Analytics, Enterprise Edge and other data-driven applications. The Pavilion HyperParallel Data Platform, powered by Pavilion HyperOS, delivers unmatched performance and density, ultra-low latency, unlimited scalability and flexibility, providing customers unprecedented choice and control. Learn why Fortune 500 companies and federal government agencies choose Pavilion. Visit www.pavilion.io or follow the company twitter at <https://twitter.com/PavilionData>