**PAVILION**

# HyperParallel Splunk

## Performant Scalable SmartStore

**splunk>**

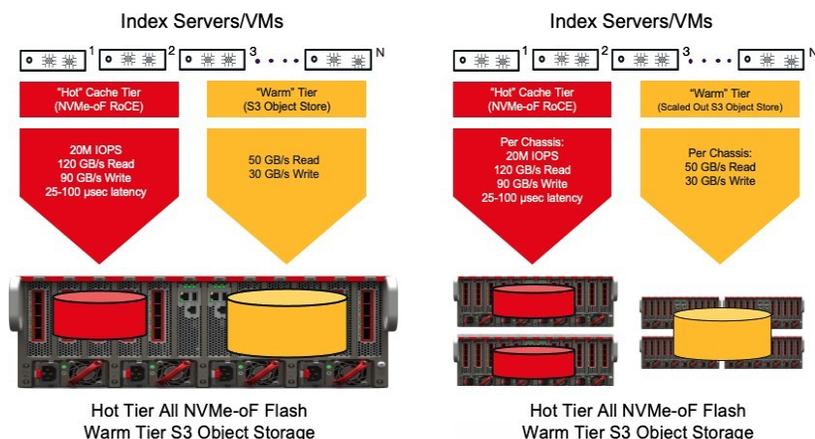## Benefits

### Splunk's Storage Problem

Today, most Splunk SmartStore users build their index caching layer with a few SSDs inside each index server, and preserve their data on an external S3 storage system or cloud. This architecture works for searches using the most recent data or when data ingest rates are low. However, it presents challenges when ingestion rates and data retention requirements exceed the index tier's planned capacity. It also poses a problem with historical long-tail searches and rare data finds, which break the cache system. To overcome these challenges, many users have deployed complex multi-vendor solutions to solve the block storage "hot data" tier and address the S3 object storage "warm data" tier.

While a multi-vendor approach can increase capacity, it reduces performance and increases the latency of "Hot" data. The impact slows ingest and search query while increasing deployment complexity, as you need to manage multiple storage systems (sometimes from different vendors) to handle both tiers.

The storage platform solution that more and more organizations are seeking to deploy consist of using one vendor to deliver all the performance of internal SSD NVMe SSDs for the "Hot" tier with increased capacity, while also providing dense and high-performance S3 Object Storage for the "Warm" Tier.

Pavilion offers the only storage platform that can deliver end-to-end NVMe-oF using RoCE to the indexing servers for the "Hot" Tier, while also delivering high-speed compressed Distributed S3 Object Storage within the same scale-out system. Pavilion shows a performance pipeline that mimics or surpasses internal NVMe SSD performance for the indexing engines while providing a limitless Warm capacity to retain historical data.

Indexers are the lifeblood of the data pipeline and search workflows. All data must be ingested into the index tier, and all search queries are run against the indexers. Pavilion enables customers to scale up to multiple PBs of capacity, allowing the "Hot" caching layer to have longer retention which improves your search cache hit rate.

- Simplify Splunk SmartStore with Index (Hot) and S3 (Warm) Tiers in a single platform

- Double indexing rate and reduce indexer footprint by 63%

- The fastest Ingest & Retrieval Rates available for S3-compliant Object Storage

- Dramatically improved long-tail & rare search speed

- Linear scale-up and scale-out with Block (Hot) & Object (Warm) tiers Independently

- Support more daily ingest with up to 90GB/s write performance (1,000s of indexers)

- Allow longer retention of data without adding indexers

- Scale-up and Scale-out Compressed SObject Storage for fastest "Warm" tier

- Improve performance of searches that exceed cache with high-performance object storage



Hot Tier All NVMe-oF Flash
Warm Tier S3 Object Storage

Hot Tier All NVMe-oF Flash
Warm Tier S3 Object Storage

## Splunkers - Meet Pavilion's HyperParallel Storage Platform

The ultra-high performance, extremely low latency, and multiple storage controllers of the Pavilion HyperParallel Storage Platform accelerates Splunk pipelines and boosts time to results. **The Pavilion HyperParallel Flash Array lets Splunk ingest data at 90GB/s. That's 10x faster than the nearest technology in the same RU.**

A key consideration of deploying SmartStore is the performance impact when a search "breaks" cache and is forced to pull data from the S3 Object Storage. With Pavilion, we provide up to 50GB/s per chassis of S3 Object storage, which scales linearly as you grow. This minimizes the penalty taken to run long-tail searches or rare searches that utilize data, not in the indexers. We make the "Warm" tier almost hot. Pavilion also compresses the data as it's written into the S3 Object Storage, which maximizes the storage efficiency of the data retained by Splunk SmartStore as it moves data from the "Hot" tier to the "Warm" Tier.

**The Pavilion HyperParallel Storage Platform's ultra-low latency removes storage as the indexing bottleneck.** In addition, Pavilion's speed eliminates cache misses with SmartStore, unlocking all of the value in SmartStore tiering.
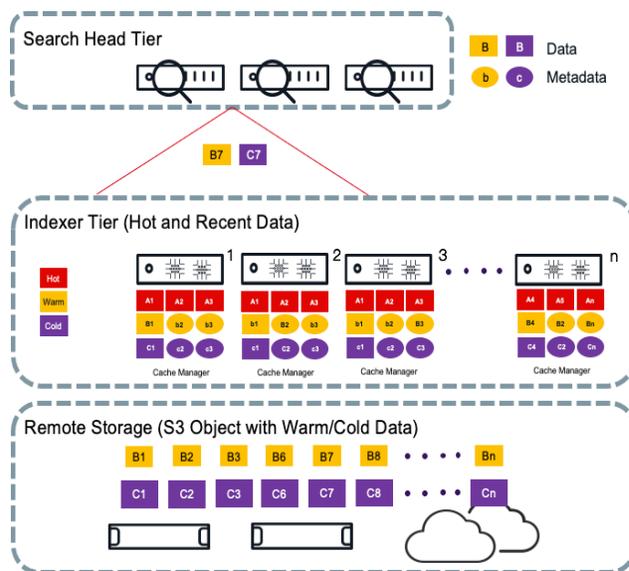


This SmartStore diagram illustrates how to store hot/warm/cold data. All data flows through the indexers and doesn't hit the warm/cold S3 until it is "evicted" by the Cache manager on the Splunk indexer. Splunk keeps metadata on the "Hot" tier after it's evicted

**Pavilion delivers 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. Storage capacity and performance scale independently to expand the hot and warm tiers.

Get 24/7 proactive support, end-to-end data integrity, a modular chassis, and redundancy throughout the storage array to protect 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.**

A Splunk workflow leveraging Pavilion will optimize data ingest streams, store more data and allow for rapid visualization. Pavilion offers universally unmatched storage enabling Splunkers to reimagine what is possible.