pavilion data

Solution Brief

# Accelerating Splunk Enterprise

Solution Overview

## Key Benefits

- Increase search times up to 2x for faster analysis and real-time decision making
- Leverage standard Ethernet networking for both servers and storage
- Reduce indexer footprint, and associated power and cooling costs, up to 4x

## Splunk Benefits

- Identify and resolve issues, reduce escalations up to 90%
- Monitor systems, infrastructure, and key performance indicators (KPI) in real time
- Proactively detect and investigate security incidents

## Pavilion Benefits

- Fastest block storage for Splunk indexers hot and warm tiers
- Latency of direct-attached SSDs
- Up to 460 TB in 4U
- Frictionless deployment
- Data resiliency & high availability
- Space-efficient instant snapshots and clones
- Thin provisioning
- Pay As You Grow scalability
- Expand for capacity or performance, independently
- Increase storage utilization up to 10X or more

## Make better decisions faster with Splunk and the Pavilion Array. Move analytics from batch to real-time.

### Splunk

Splunk enables organizations to collect, search, analyze, and visualize the machine data generated by your IT, security and business environment. Machine learning analytics can provide custom predictive analytics to optimize operations and business results.

Splunk architecture consists of indexers, forwarders, and search heads. Indexers store the collected data and index it to be used for searches. Search heads distribute searches to indexers. Forwarders forward search requests to remote indexers.


### The Pavilion Array

The Pavilion array delivers 25X the performance and 10X better latency than typical networked All-Flash-Arrays. As a result, big data analytics applications like Splunk Enterprise can now analyze much larger data sets, and therefore deliver more accurate and timely answers to critical queries. Decisions can be made in real-time at the speed of digital business by analyzing more data quickly.

Ultra-low latency (10s of µs) allows you to make provisioning decisions at application run time instead of procurement time. Replacing direct-attached SSDs in Splunk indexer nodes leads to dramatically increased storage utilization and lower costs in these environments.

### Splunk Storage Challenges

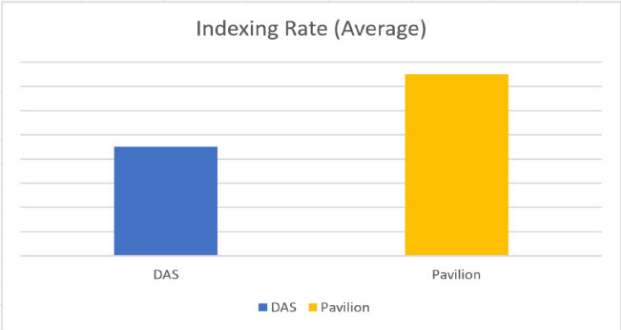
Splunk storage has traditionally used external networked storage arrays with spinning disk in the hot and warm storage tiers. When the high latency of these devices did not support interactive queries for sufficient performance for indexing large datasets, many users turned to direct-attached Flash inside the indexer nodes. However, this means that the storage capacity is stranded in the individual indexer nodes, leading to under-utilization of storage assets, and inflexibility: any growth in storage capacity for the hot and warm tiers requires additional indexer nodes even when additional indexer performance was not required.

A better solution is needed, one with the low latency of direct-attached SSDs and the flexibility of networked storage.

# Real-Time, High-Speed Scalable Analytics Solution With Pavilion and Splunk Enterprise

The Pavilion array is an ideal solution for Splunk storage for organizations requiring the most aggressive analytics solution that can analyze massive amounts of data in real time. By deploying a Pavilion array as centralized storage for Splunk indexers, a Splunk deployment can grow to much larger scale by deploying from a few to hundreds of indexers. These can all use a thinly provisioned shared storage pool with the same or better latency as direct-attached SSDs. When more CPU and memory is needed, administrators can deploy additional indexers at will to handle more search traffic, with the hot and warm tier on the Pavilion array using only the needed capacity.

When the indexers provide sufficient performance, storage capacity can be independently increased to expand the size of the hot and warm tiers. A single Pavilion array can provide up to 460 TB of high-speed storage, at up to 120 GB/s of bandwidth, providing superb performance for Splunk Enterprise.

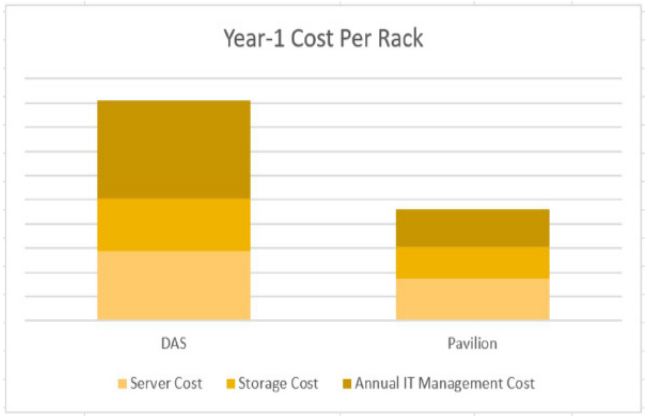



With Splunk Enterprise mid-range indexers, using Pavilion provided an increase of 67% in indexing performance over using direct-attached SSDs at the same processor utilization rate.

This means that 40% fewer indexers will be required to provide the same ingest volumes needed by the organization.

System Specification: Intel 64-bit chip architecture, 16 cores at 2.1GHz, 64GB RAM, 64-bit Linux distribution

Additionally, the type of indexers can be reduced from 2U to 1U, reducing infrastructure, maintenance, management as well as power and cooling costs. Pavilion supports thin provisioning as well, so a large amount of storage can be presented to each indexer, while the thinly provisioned array will only allocate the amount of physical storage needed at a specific time. This will deliver savings on the raw storage footprint required when compared to direct-attached SSDs.




The above comparison demonstrated that using Pavilion allows us to reduce the number of indexing servers, and use 1U servers without direct-attached SSDs, saving a large amount of acquisition costs, rack space, power and cooling. Through the use of thin provisioning and space-efficient snapshots, the total capacity can be reduced for the same usable capacity.

Using the Pavilion array also reduced the operational costs, as a storage admin can manage approximately twice the SAN storage as DAS storage. Together, this reduces the operational costs of the Splunk Enterprise environment dramatically.

The solution is also highly-available so that users can be assured of enterprise level reliable service.

The Pavilion array supports both full RAID and complete hot-swap capabilities for extreme high availability when high availability and resilience is done at the storage level, as well as JBOF configurations when Splunk replication is chosen as the high availability solution. With up to 4 different simultaneous tiers of storage in each array, multiple tiers (Splunk buckets) can all be consolidated on one appliance, alleviating the need for complex storage tiering within Splunk. The high performance and capacity density of the Pavilion array will also allow for reduced infrastructure footprint by requiring fewer indexers to be used, leading to reduced infrastructure footprint and associated power and cooling costs.

The Pavilion array fully supports all types of Splunk Enterprise. By leveraging Pavilion as high-speed, low-latency, high capacity networked storage solution for Splunk Enterprise, real time data analytics can become a reality for today's IT Organizations, allowing you to make better decisions, faster.

pavilion data

Pavilion Data Systems, Inc.
2560 N First St., Suite 220, San Jose, CA 95131
E-mail: sales@paviliondata.com