pavilion data

Solution Brief

splunk>

NVME-OF STORAGE FOR SPLUNK

Deploy Disaggregated Storage to Splunk environments.

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.

Splunk's Storage Problem

Today, Splunk storage uses traditional external networked storage arrays in the hot and warm storage tiers. Unfortunately, these devices do not support interactive queries because of their low performance numbers. As latency spikes, applications begin to crawl and customers turn to direct-attached SSDs inside the indexer nodes to satisfy performance.

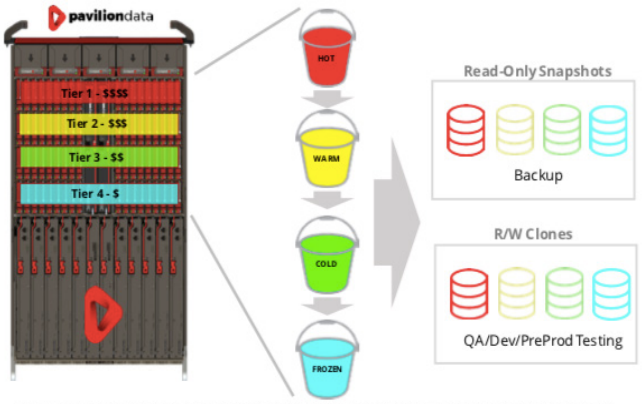
This approach has resulted in stranded storage capacity in each indexer node, leading to inflexibility and severe under-utilization of storage resources. When there is a need for growth in storage capacity, it requires additional indexer nodes even when additional indexer performance is not required.

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.

Pavilion's NVMe-oF Storage Array

Pavilion delivers never before seen NVMe performance and density that allows customers to provision logical flash storage resources over a low latency network. As a result, you can now deploy shared storage in place of direct-attached SSDs in cloud-scale Splunk environments.

The array requires no custom software to be installed on application servers and includes important data management and availability features, including thin provisioning, instant zero-space snapshots and clones, and no single point of failure.



Consolidate All Splunk Data on One Platform and Simplify Backup and Copy Management

Features

Splunk 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

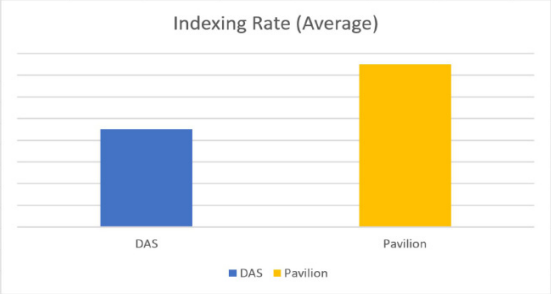
Pavilion Benefits

- Fastest block storage for Splunk indexers hot and warm tiers
- High Performance and High Availability.
- 40 µs Latency
- 14TB - 1 PB in 4U
- Management Ease
- Scale Compute and Storage Seperately.
- Lower TCO by leveraging scale-out architecture
- Frictionless Deployment
- Space-Efficient, Instant Snapshots and Clones
- Thin Provisioning
- Standard Ethernet
- **OPEN**CHOICE Storage

Deliver Disaggregated NVMe-oF Storage for Splunk:

Until now, Splunk indexers were deployed on DAS SSDs because of the performance and fault isolation requirements. The absolute lowest latency requirement, in particular, drove the need for DAS due to its latency fulfillment. However, with new high-speed RDMA-capable networking and efficient storage protocols like NVMe-oF, it is now possible to get the same performance advantages with shared storage.

Pavilion's composable, disaggregated storage is ideal for Splunk and for organizations deploying a few indexers to hundreds. By deploying the platform, we deliver 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. If the indexer provides sufficient performance, the storage capacity can be independently scaled to expand the size of the hot and warm tiers.

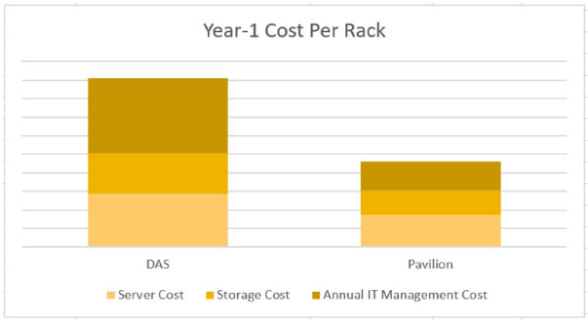


With Splunk mid-range indexers, using Pavilion provided an increase of 67% in indexing performance over using direct-attached SSDs at the same processor utilization rate. This means that 40% fewer indexers are required to provide the same ingest volumes.

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

By disaggregating storage, we allow customers to reduce the number of indexing servers and/or use 1U servers without direct-attached SSDs. This saves them a large amount on acquisition costs, rack space, power and cooling.

Through the use of thin provisioning, we only allocate the amount of storage needed at a specific time and this delivers savings on the raw storage footprint required when compared to direct-attached SSDs.



splunk> Storage Design Considerations

- Insuffucient disk i/o the most commong limitation in Splunk Infrastructure
- Review the disk subsystems requirements before provisionning your hardware
- More disk (specifically) more spindles are better for indexing performance
- Total throughput of the entire system is important
- Ratio of disks to controllers in a particular system should be higher, similar to how you provision a database host

pavilondata

- Delivers 20m IOPS from a 4U Chassis, which can power even the largest Splunk deployment
- Array scalability allows you to focus on needs for compute instead of storage
- 100µs Latency removes storage as the indexing bottleneck.
- Orders of magnitude improvement in performance decision times.
- Performance and Capacity enables easy storage configuration

*Splunk Design Reference: http://docs.splunk.com/Documentation/Splunk/7.0.3/Capacity/Referencehardware

pavilondata

Pavilion Data Systems, Inc.
2550 N First St., Suite 250, San Jose, CA 95131
E-mail: info@paviliondata.io