Lab Spotlight

Optimizing Splunk Deployments with NVMe Direct Scale-out Flash from Apeiron

Date: June 2017 **Authors:** Mike Leone, Senior Lab Analyst

Abstract: This ESG Lab Spotlight highlights the benefits of leveraging Apeiron's ADS1000 Non-volatile Memory Express (NVMe) Direct Scale-out Flash (DSF) solution to deliver an optimized storage environment in massive, distributed Splunk environments.

The Challenges

As organizations become more data-driven, the ability to organize, store, and analyze constantly growing datasets is a challenge. This is especially true when leveraging traditional infrastructures, which consist of expensive, disparate silos of storage. These silos are difficult to manage, upgrade, and integrate. Organizations want access to all of their data as fast as possible without sacrificing performance, cost, or potential insight. In today's IT environments, data analytic tools collect a massive amount of data that require processing in real-time; this requires a shift in focus from the tools to the underlying infrastructure supporting these new demands. ESG research shows that event and log data are the most common data sources organizations use for business intelligence and analytics purposes, such as Splunk's operational intelligence platform. Consequently, it is essential to find a future-proof technology solution that can not only support the entire infrastructure, but improve existing resource utilization to deliver higher levels of ROI. [1]

Apeiron ADS1000

Apeiron has created a Direct Scale-out Flash (DSF) solution that leverages "NVMe over Ethernet" (NoE) to create a storage network with the high performance and simplicity of internal or captive storage. NVMe is a new storage protocol designed to use the internal PCIe bus to take advantage of the bandwidth, low latency, and parallel capabilities of flash technology. The system uses any commercially available NVMe drive and provides organizations with freedom of choice to select a drive that properly supports its existing applications, whether based on drive profile or supplier. This integrated storage and networking solution leverages the ease of use and cost-efficiencies of the Ethernet ecosystem to provide extremely high performance with the ability to independently scale compute and storage.

The solution is a combination of driver-level software and hardware. The ADS1000 enclosure is a 2U drive and network chassis capable of presenting 24 NVMe SSDs (38 TB to 192 TBs today); capacities will increase as larger NVMe SSDs come to market. The device is fully HA with 16-port I/O modules, redundant power and cooling. The server component includes dual 40GbE PCIe HBAs, which can be purchased separately or included in Apeiron's x86 1U compute node. A key feature of this design is a fully integrated 40Gb Ethernet switch which eliminates the need to procure, deploy and manage any external storage switching infrastructure.

[1] Source: ESG Research Report, *Enterprise Big Data, Business Intelligence, and Analytics Trends: Redux*, to be published.

Splunk

A leading event and log data analytics vendor, Splunk enables organizations to gain insight from virtually any data source. As machine data is collected, organizations can query, analyze, and visualize their data to gain valuable insights about resources, applications, and security. Further, the ability to monitor and alert based on real-time data enables organizations to proactively address potential problems, while reporting capabilities allow data and insights to be easily shared or integrated with third-party applications.

The Splunk architecture consists of three key software components (indexers, forwarders, and search heads) that can be deployed within a single server or distributed across multiple server instances. Indexers provide indexing and search capabilities for local and remote data and host the primary Splunk data store. Forwarders forward data to remote indexers for indexing and storage, and can either forward unparsed data, or can parse and index data prior to forwarding it. Search heads distribute searches to indexers. In large deployments with high numbers of concurrent users, search heads are important in distributing search requests across the indexers.

Why Apeiron for Splunk?

With traditional storage implementations, architectural limitations exist that can have a direct impact on Splunk performance, scalability, capacity consumption, and data accessibility. This is particularly important in distributed Splunk environments that rely heavily on analyzed machine and log data to make important business decisions. Storage solutions with controller-based architectures and external storage switching introduce a single point of data access that can directly impact processing and networking performance. Disparate resource silos add management complexity, and have a dramatic impact on Splunk performance. Splunk was designed to leverage scale-out architectures where storage is captive in the server. As the business value of Splunk was realized however, the customers began to demand larger data sets and real-time reporting. IT groups began to deploy Splunk on traditional SAN and NAS environments in an attempt to meet the surge in data. Although the capacity needs could be met, controller-based systems with external switching create significant performance bottlenecks, leading to lower productivity.

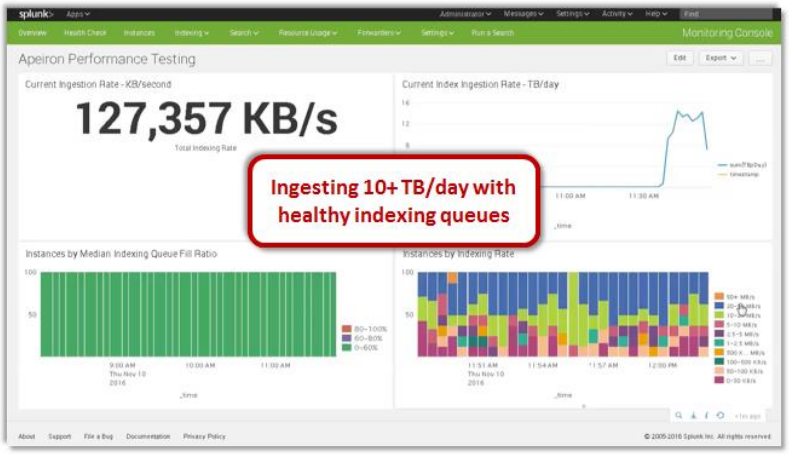
Apeiron delivers extreme performance in a shared infrastructure. To the client, the ADS1000 looks and acts like direct attached storage (DAS), but with all the operational and economic advantages of a storage network. It has the ability to access thousands of NVMe drives at speeds actually faster than internal PCIe connected NVMe drives—without any of the typical network latency. In fact, ESG Labs validated the performance delivered by Apeiron is actually faster than DAS (through the ability to spread workloads across more drives). IOPs, latency, and throughput comparisons yielded indistinguishable differences between the ADS1000 with an external NVMe SSD, and an NVMe SSD internal to the server. NoE switching moves the bottleneck to the NAND drive architecture itself (the ADS1000 storage network introduces less than 1.5 µS of network latency-imperceptable to the application) . In addition to the performance benefits, the elimination of all external storage switching delivers significant consolidation savings over traditional storage arrays.

For Splunk, Apeiron's Direct Scale-out Flash eliminates the management complexities that traditional fabrics and storage require. This not only improves operational efficiency through consolidation, but reduces operational expenditures by providing a single performance tier for even the strictest Splunk performance SLAs. With the improved I/O performance of the ADS1000, Splunk indexing is completed many times faster. The ESG Lab validated Apeiron is more than capable of handling the I/O requirements of Splunk software, which consists of sequential writes for data ingest/collection and random reads for search and indexing. Searches provide instant access to all data placed into Splunk's hot, warm and cold data stages. Because Apeiron's servers come with 36 cores and 128GB of cache in a 1U form factor, more searches can be done on fewer servers. ESG Lab validated that Apeiron delivered a 70%-150% improvement in CPU utilization in its tests. These features mean that with Apeiron, Splunk indexing can be done for a much larger number of forwarders. Queries are accelerated to the point that customers can query years of data in real-time rather than the 90-Days typical with traditional arrays. Apeiron storage not only eliminates I/O-bound queries with the 18M+ IOPS in each ADS1000 enclosure, but enables organizations to independently scale compute or storage - depending on the performance or retention requirements.

ESG Lab Testing

ESG Lab audited results of 3-week performance, scalability, and capacity benchmark completed by World Wide Technology Inc. (WWT), located in St. Louis, MO. WWT configured an ADS appliance in their Advanced Technology Center (ATC). The appliance included seven 1-U x86 servers and a 2-U ADS1000 enclosure with 24 NVMe SSDs. The SSDs were from two suppliers to demonstrate Apeiron's ability to support complete NVMe SSD interoperability. Four servers were configured as indexers, one as the Enterprise Security (ES) search head, one as an Ad Hoc search head and as the Splunk cluster manager. A data set consisting of 70 billion events and spanning two weeks of time was used. Splunk version 6.5 was used for testing and standard Splunk performance and monitoring tools were used to report the results. It should be noted that all 60 out of the box ES correlation searches were run in parallel with Ad Hoc queries, dense queries and data ingestion. Both ingest rates and search times of varying query complexities (density) were monitored throughout testing.

Figure 1. Data Ingestion with Apeiron



The first phase of testing focused on the ingest rate. As shown in Figure 1, while running the concurrent queries, Apeiron was able to yield average ingest rates of 10 TB/day, with some periods producing rates as high as 12 TB/day, while indexing queues remained in a fully healthy state. Without the high number of concurrent queries the system routinely ingested well over 20 TB/day.

In comparison, a similarly configured Splunk reference architecture with one search head, four indexers, and using a traditional storage backend would yield results in the 1 TB/day range. Apeiron was documented with an ingest rate nearly 10X that of traditional storage architecture over several weeks.

For search results, two phases of testing were completed. First, Apeiron was compared to a recommended Splunk reference architecture with expected performance results based on three different search types: rare, super sparse, and sparse. The results are highlighted in Table 1. Apeiron outperformed the reference architecture across all tests. The super sparse results show that Apeiron completed two different searches 58x and 88x faster. The sparse results highlighted an order of magnitude improvement that could be gained with Apeiron: one dense search with the reference architecture took more than 2.5 days, while Apeiron completed the task in just over 3.5 hours.

Table 1. Comparing Apeiron Search Results with Splunk Reference Architecture Guidelines

Search Type	Events Found	Search Time (in seconds)		Times Faster with Apeiron
		Splunk Reference Architecture	Apeiron ADS1000	
Rare	23	11.2	2.1	5
	115	11.2	6.1	1.8
Super Sparse	26,802	1,112.2	12.6	88
	180,850	1,112.2	19.2	58
Sparse	155,459,317	31,091.9	2,842.2	11
	1,126,745,647	225,349.1	14,985.3	15

The second phase of testing consisted of running the same dense search on a data set consisting of 55 million records in three different environments: a bare metal environment with traditional storage, a virtualized environment with traditional storage, and a bare metal environment with Aperion storage. The need to procure and manage virtualization software is eliminated with Apeiron, due to its ability to support the independent scaling of compute and storage. Both the number of records/second searched and the time to complete the search were monitored. The results are shown in Table 2. Apeiron yielded a performance gain that was an order of magnitude greater than that of both Splunk environments that leveraged traditional storage. The dense search on the ADS appliance completed on average 8x faster, while searching an average of 7x more records/second. This performance demonstrates that Apeiron delivers a new level of Splunk scalability and performance.

Table 2. 55 Million-record Dense Search Comparison

Type of Splunk Environment	Records per Second	Search Time (in seconds)
Bare Metal (Traditional Storage)	12,388	4,834
Virtualized (Traditional Storage)	10,528	5,596
Bare Metal (**Apeiron**)	**86,341**	**620**

The Bigger Truth

As data analytics tools such as Splunk continue to mature and add value to the business, a shift in focus will be needed. Data sets continue to grow and traditional infrastructures struggle to meet performance, scalability and cost requirements. In short, the larger the dataset, the heavier the burden on server and storage resources to deliver on performance SLAs. In larger deployments, Splunk components are distributed across multiple server instances, with forwarders consuming data, indexers indexing and searching it, and search heads coordinating searches across indexers. A heavy burden is placed on the underlying infrastructure, and on IT administrators to ensure that a scalable, high performance platform remains online and operating at peak efficiency. Traditional storage approaches introduce the potential for this mission-critical workflow to be interrupted. As an example, a traditional storage infrastructure for Splunk deploys three separate networking protocols (FC, IB and Ethernet). This places a tremendous burden on the IT staff and their budget.

The Apeiron architecture was designed to provide all the performance NVMe offers , and to leverage a single networking protocol; an Ethernet fabric which drives down both costs and risk. NVMe SSDs offer the next major storage performance leap, with much lower latency and greater density. As the world's only native NVMe storage network, Apeiron brings NVMe out of the rigid silos typical of scale-out architectures and into a networked storage environment. The ADS1000 provides a single pool of high performance NVMe that IT can allocate to an unlimited number of servers. With petabyte-scale NVMe storage, no external switching, and significant server CPU benefits, the consolidation of hardware and IT functions provides a compelling ROI/TCO justification.

The benefits of having Apeiron as the underlying storage infrastructure to support a dynamically growing Splunk deployment are obvious. Organizations can make decisions based on analysis of all their data, not a subset; search and index faster; support more users at higher levels of concurrency; keep costs in check with simplified management and resource consolidation; and deliver higher levels of predictable performance that easily scales with the environment.