Hewlett Packard Enterprise

Technical white paper

Check if the document is available in the language of your choice.

SPLUNK WITH HPE PRIMERA, HPE NIMBLE STORAGE, HPE PROLIANT SERVERS, AND SCALITY RING

Solution overview and best practices

CONTENTS

EXECUTIVE SUMMARY

Every IT application, system, and piece of infrastructure in a company, right down to the light switches, generate data at millisecond intervals. This machine-generated data is as complex as it is invaluable. It is also one of the fastest growing areas of Big Data. Machine-generated data contains a detailed record of all user transactions and individual customer and component behavior. It further captures sensor activity, machine performance, security threats, and fraudulent activity, to name a few. The data holds valuable insights, critical to both the security and the profitability of the enterprise.

Splunk's core capabilities include the universal collection and indexing of machine data from nearly any source, while providing real-time monitoring for patterns, thresholds, and alerts, in addition to search and analyze capabilities for historical purposes. Splunk is commonly used for Enterprise Security, but also offers use cases extending to areas like online retail order tracking, insurance claims processing, and marketing analytics. Users can create many different types of visualizations and dashboards to provide real-time operational intelligence, security monitoring, and event management.

While the amount of data is increasing like never before, the data that is most recent is also most valuable. Data must be acted on right away, or it begins losing its value. According to a technical insight paper named Enterprise Storage for Splunk: Architectural Options and Values published by The Evaluator Group, research performed by Splunk indicates that more than 95% of Splunk searches are over data that is less than 24 hours old.

The classic implementation of Splunk uses five tiers or "buckets" for data lifecycle management. A bucket is a directory that contains the indexed events from a specific period of time. When Splunk indexes the incoming data, it extracts events and writes them to a compressed raw data file, and then writes event locality metadata to the index file, both of which are created in a hot bucket. As events age and capacity fills up, data moves to lower tiers in an automatic process based on user-defined policies and available storage capacity. Hot and warm buckets always reside in the same storage tier, which is normally the top tier of high performance storage. Cold buckets can be on a different (lower) storage tier that is typically cheaper and has lower performance. If archive storage has been defined for the oldest data, those events are moved to frozen buckets on your archive storage. If data that has been moved to frozen buckets is needed again, it must be moved to a thawed bucket where it can be reindexed. These thawed buckets can be on the hot/warm tier or the cold tier, depending on user configuration.

A Splunk SmartStore (SmartStore) implementation uses S3 object storage to minimize the amount of data stored on Splunk's indexers. With SmartStore, hot buckets will always be stored on high performance local storage. As soon as data rolls out of the hot buckets, it gets copied to warm buckets on S3 object storage such as HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING (Scality RING). Those warm buckets will also be retained in a high speed local cache, managed by the indexer's cache manager. Splunk's cold buckets are eliminated because the warm buckets have already been moved to less expensive object storage but still remain fully searchable. Because the data comes from many different sources, the storage volumes in either a classic implementation or Splunk SmartStore implementation can quickly consume terabytes of space with billions of files.

To support the heavy write loads of new data being indexed and the significant read loads of queries and visualizations, Splunk requires a high-performance, tiered storage infrastructure with high capacity. HPE all-flash, enterprise-class storage arrays, along with Scality RING, enable organizations to decouple storage and compute so that each can be scaled independently to meet the changing needs of the business.

Storage solutions from Hewlett Packard Enterprise can address deployments ranging from small scale implementations to scale-out deployments requiring high performance with affordable economics. Splunk deployments tend to grow rapidly once implemented and necessitate a storage infrastructure that grows in an economically sensible way. As your deployment grows to accommodate more data and more users, HPE storage scales seamlessly to provide optimal performance at a cost that is affordable.

Through HPE GreenLake Cloud Services (HPE GreenLake), Splunk infrastructure can be acquired in a cloud-like, consumption-based fashion. HPE GreenLake allows cloud-like procurement through a pay-as-you-go, consumption-based model. This allows Splunk users to scale storage and compute capacity up and down while paying only for the capacity consumed.

Target audience: This document is intended for data administrators, architects for Splunk IT environments, storage administrators, presales consultants, and solution architects who are designing, implementing, and maintaining data analytics solutions based on Splunk running on HPE servers and storage.

Document purpose: This whitepaper highlights recognizable benefits to technical audiences, and provides sizing recommendations for deploying a data analytics solution based on three levels of data ingest: 300 GB per day, 1 TB per day, and 2 TB per day, with a significant concurrent search load running at the same time.

INTRODUCTION

Splunk is a powerful tool for analyzing and understanding your data as it streams in from many different sources. Splunk captures and extracts events from real-time data, and stores it in searchable files. Splunk can then generate alerts, reports, and dashboard visualizations to help you identify patterns and find anomalies in your data. The Splunk software makes your machine data accessible across your organization, and provides intelligence for business operations as well as web analytics, security, and regulatory compliance.

Storing data from many different sources presents a number of challenges. Splunk environments can vary widely in size and scope, but the factors that are common to all deployments are the need for performance and the ability to scale. Customers often implement Splunk on a limited scale at first, expecting to grow their deployment over time. While this is a good starting point for Splunk environments, the deployments typically grow rapidly as the solution shows value and customers expand their investment. Splunk has determined from user feedback that storage requirements are growing faster than compute requirements. Organizations sometimes find that they cannot maintain fast response times on their searches, and their need for capacity for cold and frozen data keeps increasing. Thus, they want a solution that scales easily and inexpensively while maintaining peak performance.

For storage of your indexed data, HPE Primera all-flash storage arrays and HPE Nimble Storage arrays provide the flexibility to deploy solutions with high performance while still being economical. HPE Primera and HPE Nimble Storage arrays afford high performance and scalability for Splunk's hot/warm and cold tiers. Splunk's cold and frozen tiers can be moved to a less expensive tier, such as HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING.

Splunk supplies the flexibility to implement physical, virtual, and containerized deployments. The storage requirements vary so widely from one deployment to another that there is no one-size-fits-all architecture. This paper focuses on a flexible and highly scalable solution by combining an all-flash array with a long-retention archival tier.

To gain the most value from Splunk, your organization needs a high-performance storage architecture that minimizes search times, enables storage tiering, and makes data volumes easy to manage. High-performance HPE Primera and HPE Nimble Storage arrays maximize end user productivity by facilitating fast searches without causing a bottleneck for writing the new data being indexed. HPE arrays also provide advanced data protection and manageability with built-in redundancies, storage snapshot backups, HPE Recovery Manager Central (RMC), and HPE Infosight predictive analytics.

A Splunk deployment implemented on HPE Primera or HPE Nimble Storage arrays delivers high performance and low latency without complex tuning or set up. As the volume of data grows and the need to retain it increases, keeping all your data on enterprise-class arrays (especially older, long-term retention data) becomes more and more expensive. Additionally, keeping older data that is no longer accessed frequently on enterprise-class storage will slow down your searches. Ideally, Splunk implementations should keep the most recent data in primary SSD-based storage where it can be searched quickly, with the older and infrequently accessed data stored on a lower tier but still accessible.

HPE enterprise-class storage can be augmented with archival storage on HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING, alongside tiering managed by Splunk, to provide longer retention for data you no longer access frequently but cannot get rid of. This implementation yields all the storage you need for your archived data at a significantly reduced cost, compared to maintaining all data in block-based storage. Government regulations or company policies regarding document retention make hosting Splunk's frozen tier on archive storage a practical alternative. When retention requirements make it necessary to keep large amounts of infrequently accessed data for extended periods of time, it becomes economically attractive to implement Splunk's frozen tier on archival storage with HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING.

Scality RING lets you budget and control costs with no unexpected fees; for example, retrieval fees for accessing your data. For data protection, Scality RING supplies customizable availability and failure domains. Customers can configure the data protection policy at the object level, with replication of up to five copies. Erasure coding can be configured to yield as much as 14-nines of durability with low overhead for larger objects. Data protection options include geo-redundancy, affording additional tolerance of multiple disk, server, rack, and even site failures. For environments that scale to hundreds of terabytes or more, HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING gives the security and protection of an on-premises solution with the flexibility and economics of the cloud. This solution has been fully tested.

The Scality RING software can be deployed on industry-standard x86 servers, such as the HPE Apollo 4000 systems. This type of archive store with virtually unlimited scale can also be used to consolidate other Big Data use cases and data storage to allow for even larger economies of scale.

To run Splunk efficiently and effectively in a mission-critical environment, your servers need plenty of system resources for all instances of Splunk. HPE ProLiant servers have all the resources you need to meet and exceed the minimum hardware requirements for running Splunk. Hewlett Packard Enterprise offers a rack-optimized portfolio of the world's most secure industry-standard servers[1]. This is an agile infrastructure that enables software-defined intelligence. HPE ProLiant servers have world-class performance and supreme versatility for multi-workload computing. If your plan is to virtualize your Splunk deployment, HPE ProLiant servers have plenty of system resources to meet and exceed Splunk's minimum system requirements for running multiple instances of Splunk as virtual machines in a VMware® environment.

As a total solution stack provider, Hewlett Packard Enterprise provides a full range of storage and servers with the ability to provide different solution architectures for different size deployments. This enables you to decouple your storage and compute, allowing you to scale storage and compute resources independently. Hewlett Packard Enterprise has everything you need to implement Splunk environments ranging from a small single instance to multisite, geo-dispersed, scale-out deployments with petabytes of data. With the addition of an archive tier, Hewlett Packard Enterprise gives your Splunk environment the ability to stretch to the cloud, whether it be an on-premises private cloud or the public cloud. In addition, HPE GreenLake provides the ability to grow organically as your Splunk storage or compute needs grow, so that Splunk users only pay for the capacity consumed, thus keeping their costs to a minimum.

SOLUTION OVERVIEW

FIGURE 1 illustrates the environment tested for this solution. Splunk Enterprise was implemented in a distributed deployment using HPE ProLiant DL380 Gen10 servers to run all instances of Splunk. The environment contained one search head and a cluster of indexers. In addition, there were separate instances of Splunk for the index cluster master and license master. A collection of universal forwarders were used to send data to the indexer cluster at varying rates. This testing addressed daily indexing rates of 300 GB per day, 1 TB per day, and 2 TB per day. Later phases of testing included automated searches to simulate a large number of users and scheduled processes running concurrent searches of the indexed events.

During the first portion of testing, primary storage for the indexed data was provided by an HPE 3PAR enterprise-class array configured with tiered storage using both solid-state drives (SSDs) and hard disk drives (HDDs). Subsequent testing used an HPE Primera all-flash array. An HPE Nimble Storage all-flash array could have been used in place of either array with similar results.

[1] Based on an external firm conducting cybersecurity penetration testing on a range of server products from a range of manufacturers, May 2017.

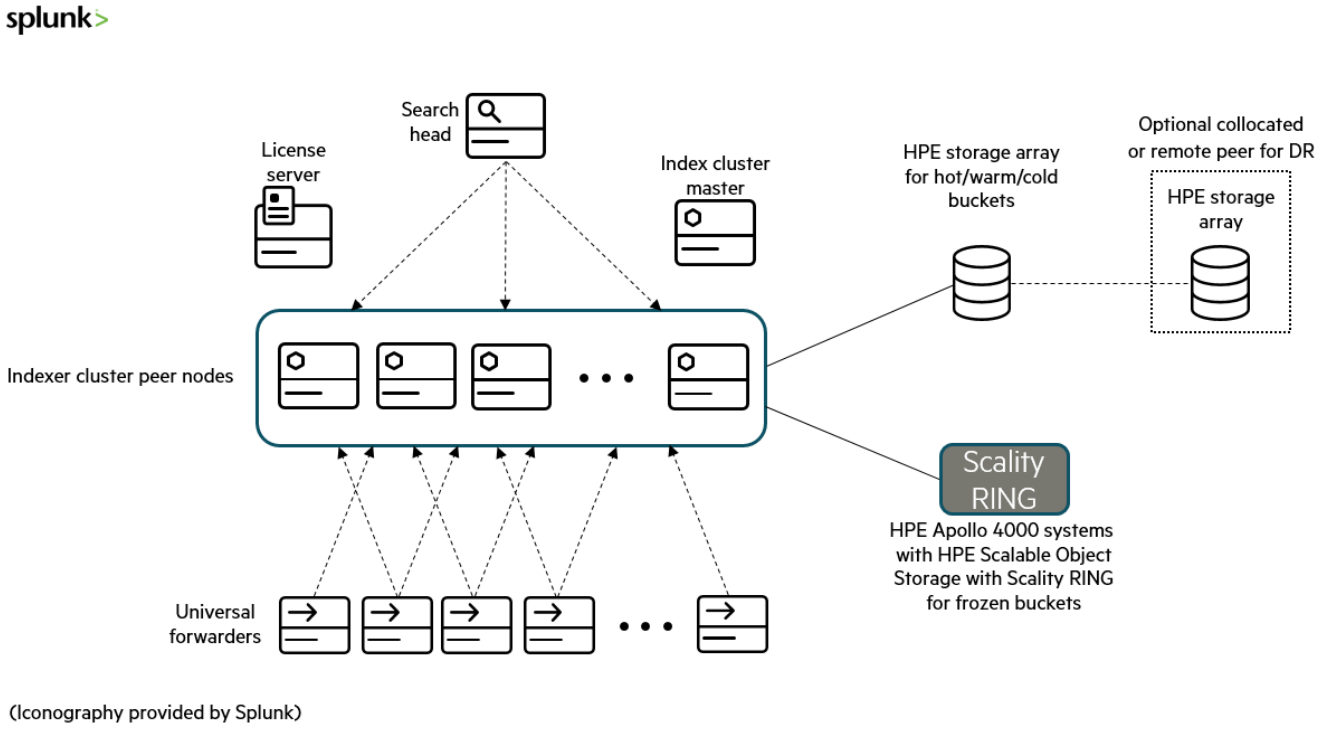

FIGURE 1. Simplified architectural diagram of the Splunk environment tested

In our testing, events came from a variety of sources including Windows event logs, HPE 3PAR array syslog data, and Fibre Channel switch syslog data. Most events, however, were generated by Splunk Event Generator (Eventgen), a utility that can be downloaded from Splunkbase. Eventgen allows the user to generate real-time events which simulate real data in a controlled manner with the precision necessary for obtaining the targeted indexing rates with consistent repeatability.

The data sources were configured to generate a continuous stream of data which equaled the desired ingestion rate over a 24 hour period. There was some variability in the data sources to reflect a nonconstant level of indexing. The data ran continuously over the entire 24 hours, rather than in burst fashion over a much shorter time.

Most of the Splunk indexers were configured as virtual machines, each having 12 vCPU cores and 64 GB of memory. In addition to these virtualized indexers, one physical indexer was used. This physical server had a total of 20 CPU cores and 256 GB of memory. With the workloads used during this testing, no significant difference in performance was noted between the virtual indexers and the physical indexer.

The following tests were performed:

- Initially, all buckets were on SSD-based storage in the HPE 3PAR array.
- The tests were repeated with hot and warm buckets still on SSD, and only the cold buckets moved to HDD-based storage on the same array.
- The tests were then repeated with hot and warm buckets still on SSD, cold buckets still on HDD, and a frozen path which used archive storage on HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING. If desired, the cold buckets could be moved off of the storage array altogether and stored on Scality RING using NFS file shares. For our testing, we chose to keep the cold buckets on a lower tier of storage on the HPE 3PAR array.
- Finally, the HPE Primera all-flash array was tested with Splunk's hot, warm, and cold buckets on SSD-based storage on the same array. During this phase of testing, scripted searches were added to produce a significant number of reads on the index volumes, in addition to the continuous writes from the events still being indexed at a rate of 2 TB per day.

Data protection is built-in with HPE enterprise-class arrays. Because the storage is decoupled from the computing resources, your indexed data is always protected, regardless of what happens in your Splunk environment. Robust disaster recovery (DR) is provided by:

- The redundancy built into the arrays.
- The automatic replication managed by Splunk when the indexers are clustered.
- The ability of HPE Primera and HPE 3PAR arrays, and HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING, to support replication and DR across geographically separated sites.

As an optional extension of the solution, a second HPE array could be added to the configuration as a peer, as shown in FIGURE 1. In addition to the DR advantages this second array would provide, HPE Peer Persistence eliminates silos by using the storage resources at DR sites along with storage at primary sites during normal business operations. This supplies an extremely available, always-on storage solution that can withstand failure of an entire array without any impact to mission-critical applications.

Also, storage snapshots which are scheduled and managed by HPE RMC can be used to create backups of all Splunk's buckets to furnish fast and nondisruptive backups. These storage snapshots can quickly and easily be moved to the peer array using the Peer Copy feature in RMC to further increase data protection and reduce risk.

HPE Primera and HPE 3PAR storage snapshots initially consume no space, and only increases in size as they record changes to the original volume. This allows an earlier state of the original volume to be recreated by starting with the current state and "rolling back" all the changes that have been made since the snapshot was created. Snapshots can be scheduled through RMC, and if configured, these storage snapshots can be moved to the archival storage as retention backups.

If your Splunk environment is virtualized and uses storage space on the HPE arrays to create datastores for the virtual machine's operating system disks, HPE RMC supplies a significant additional benefit of providing the ability to create snapshots of the volumes containing the virtual machine disk files (VMDKs) of your Splunk instances. While Splunk automatically replicates all the indexed data across the index cluster, it has no mechanism for backing up or protecting the Splunk instances contained on the virtual machines (VMs). Using HPE RMC is an easy way to do scheduled backups so that if a VM or VMware ESXi™ server is lost or corrupted, the VMs containing your Splunk instances can be easily and quickly restored.

The use cases tested for this paper were centered mainly on indexing and searching events. With its analytics capabilities, the power of Splunk is most beneficial in use cases that are read intensive. Exercising the HPE Peer Persistence and storage snapshot capabilities of HPE storage arrays is not in the scope of this paper, but these features are included with HPE Primera and HPE Nimble Storage arrays at no incremental cost.

SOLUTION COMPONENTS

Hardware

HPE ProLiant DL380 Gen10 servers

This solution made extensive use of HPE ProLiant DL380 Gen10 servers with two physical Intel® Xeon® Gold CPUs having 18 cores each and 256 GB of RAM. These servers were used as the platform for most instances of the Splunk environment, including the search head, indexers, index cluster master, license master, and so on. The Splunk universal forwarder was installed on a variety of other HPE ProLiant servers to collect events from diverse sources and direct those events to the cluster of indexers.

HPE 3PAR 20800 enterprise-class storage array

The HPE 3PAR 20800 storage array used in this testing provided primary storage to the Splunk indexers. The array was an 8-node system with 24 CPUs per node. It contained both SSD and HDD, and had volumes created on both types of storage for each indexer in the cluster. Initial testing was performed with all Splunk hot, warm, and cold buckets residing on SSD-based storage. Subsequent testing was performed with data tiering in the array where hot/warm buckets resided on SSD-based volumes, while cold buckets resided on HDD-based volumes.

HPE Primera A650 enterprise-class storage array

In our most recent testing, the HPE 3PAR 20800 storage array was replaced with an HPE Primera A650 all-flash storage array. This array was a 2-node system containing only SSDs. Virtual volumes were created on this array to contain the VM system disks for the virtualized instances of Splunk, as well as index storage for each indexer in the cluster. In testing with this array, Splunk's hot, warm, and cold buckets were on SSD-based volumes.

HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING

HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING (Scality RING) is a software-defined storage (SDS), data storage solution that runs on HPE Apollo 4000 servers, and is designed for multi-application environments needing to store unstructured data at petabyte scale. It is perfectly suited to the software-defined data center (SDDC). The Scality RING software is designed to create unbounded scale-out storage systems to consolidate and protect data used by multiple applications and workloads, including file and object applications. The Scality RING software provides a set of intelligent services for data access, data protection, and systems management. The top layer of data access services offers native file and S3-compatible cloud object storage interfaces for applications.

To optimize your business performance with object storage, Hewlett Packard Enterprise and Scality have certified Scality RING on the 24 large form factor (LFF) HPE Apollo 4200 Gen10 and the 60 LFF HPE Apollo 4510 Gen10 servers. HPE Apollo 4000 systems are purpose built for big data analytics, software-defined storage, backup and archive, and other data storage-intensive workloads to help customers meet their data center operation challenges. By deploying your Scality RING solution on Hewlett Packard Enterprise qualified platforms, customers can rapidly adopt new hardware and media innovations with no need for data migration or hardware refresh. Support is simpler and deployments are highly adaptable. Overall, this significantly reduces acquisition and operating costs and increases project flexibility.

The Scality RING's advanced routing capabilities, configurable data management, and software-defined architecture afford full system availability and uptime during planned and unplanned events, including hardware refreshes, capacity upgrades, software upgrades, and hardware failures. HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING is designed to be self-managing and autonomous, thus freeing your administrators to work on other value-added tasks. The software is deployed as a distributed system on a minimum cluster of three storage servers. This system can be seamlessly expanded online to thousands of physical storage servers as the need for storage capacity grows.

Scality RING supports network speeds of 10 GbE, 40 GbE, and 100 GbE, with a load balancer feeding multiple NFS or S3 connectors at the same time. Comparable network bandwidth to a public cloud service would be prohibitively expensive.

With multiple network ports bonded together, a load balancer, and a fast network as previously described, Scality RING and the network connection to it should not become a performance bottleneck. In contrast with public cloud, the bandwidth of the connection to the cloud can easily become a performance limiter, depending on your network provider bandwidth and service-level agreement (SLA).

No special configuration of the Scality RING is needed to use its storage with your Big Data applications, such as Splunk. HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING is delivered from Hewlett Packard Enterprise as a fully functional target, ready to be activated and used immediately. In addition to NFS for storing cold and frozen buckets, Scality RING also supports S3 object access for Splunk SmartStore, or for non-Splunk use cases.

Application software

Testing was begun using Splunk Enterprise version 7.3.1. As testing progressed, Splunk released a later version, so the software was upgraded to version 8.0.3, which was the latest version available at time of testing with the HPE Primera storage array.

Connectivity

For this testing, the indexers were connected to the HPE 3PAR array, and subsequently the HPE Primera array, via 16 Gb Fibre Channel SAN. All network traffic between Splunk instances was transferred across a 1 GbE LAN. With a daily indexing rate of 2 TB per day, our average incoming data rate was less than 25 megabytes per second (MB/s). The maximum throughput of a 1 GbE network is approximately 123 MB/s, so the 1 GbE network had plenty of bandwidth to accommodate all the traffic generated by Splunk, which included:

- New events being indexed
- Replication of indexed data across indexers in the cluster
- Network traffic associated with data movement as buckets aged in Splunk and were moved to a lower tier of storage
- Search traffic generated by users searching the indexed events

In our testing, the network did not become a bottleneck or impede the timely transfer of Splunk data. If your infrastructure will allow it, a more scalable solution would be to have a dedicated high speed network, such as 10 GbE or faster, for your Splunk traffic. A faster network will more easily support scaling to a higher number of indexers and a higher daily indexing rate than used in this testing, and would support faster rebalancing and bucket fixup when indexers have scheduled or unscheduled downtime. A faster network will also scale better to support a larger number of users searching the indexed data, without causing latency issues that could impact search response times.

For this testing, the connection to the HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING was over the same 1 GbE LAN as all other Splunk-related traffic. No network bottleneck was observed during our testing. As described earlier, a faster network connection to the archival storage would be desirable to allow scaling to higher daily indexing rates and higher numbers of users conducting searches.

OVERVIEW AND SUMMARY OF TESTING

Hewlett Packard Enterprise offers a full line of storage arrays and servers allowing customers to build different solution architectures for different sizes of Splunk deployments. The subsequent sections of this paper provide guidance for sizing HPE enterprise-class storage arrays for your Splunk deployment. For determining the number and type of Splunk instances required for a given configuration, Hewlett Packard Enterprise recommends that customers follow the Splunk sizing best practices when determining the number of search heads, indexers, and other instances of Splunk required in their environment. Documentation on planning for indexers, search heads, and so forth in a Splunk deployment can be found at:

- https://docs.splunk.com/Documentation/Splunk/latest/Capacity/IntroductiontocapacityplanningforSplunkEnterprise
- https://docs.splunk.com/Documentation/Splunk/latest/Capacity/Summaryofperformancerecommendations
- https://docs.splunk.com/Documentation/Splunk/latest/Deploy/Distributedoverview

During our testing, the HPE 3PAR and HPE Primera storage arrays, HPE ProLiant servers, and HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING provided plenty of flexibility and performance to support the demands of a distributed Splunk deployment. Configuring Splunk to place all buckets on SSD-based volumes, or hot and warm buckets on SSD-based volumes with cold buckets on HDD-based volumes, proved to be a simple task. The path variables are contained in a file on the index cluster master, which gets pushed to all indexers in the cluster. There is one path variable for the location of hot and warm buckets, and a separate variable for the location of cold buckets. These variables can be set to the same path or to different paths as your configuration and needs dictate. As long as all the indexers in the clusters are of the same operating system, specifying the paths for hot and warm buckets, and cold buckets, across all of your indexers is a simple matter.

The file on the index cluster master that specifies the location of hot and warm buckets, and cold buckets, also has a provision to specify the location for frozen buckets. In our case, frozen data was moved to HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING as the archive storage layer.

Testing was first performed with hot, warm, and cold buckets all contained on the HPE 3PAR storage array, and no frozen path. In this configuration, older events are deleted when they age out of Splunk's cold buckets. Subsequent testing involved moving older event data from Splunk's cold buckets on the HPE 3PAR storage array to frozen buckets on HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING.

Subsequent testing was performed with an HPE Primera all-flash array having Splunk's hot, warm, and cold buckets on the same tier of high performance SSD-based storage.

Performance metrics for the storage arrays were captured at various points during testing, and are presented in tables and charts later in this paper. This performance data was used in conjunction with storage capacity sizing guidelines from Splunk to determine minimum sizing for an HPE enterprise-class storage array appropriately sized for the workloads tested in this solution. Examples of array configurations sized for each daily indexing rate tested are included later in this paper.

During testing it was determined that the HPE storage arrays had plenty of resources to handle both the write and read demands placed upon them. Thus, the addition of a significant number of automated searches running concurrently with the indexing rates tested did not require any increase in the sizing or specifications of the arrays.

As a result of testing this solution, several clear benefits were observed which are described in the following sections. These should be considered best practices for implementing a Splunk deployment on HPE enterprise-class storage arrays and servers.

BEST PRACTICES AND CONFIGURATION GUIDANCE FOR THE SOLUTION

Benefits of indexer clustering

Splunk provides the option of setting up a cluster of indexers in distributed deployments. In testing of this solution, several indexers were grouped together to form an indexer cluster. This spreads out the indexed events over multiple instances of Splunk, allowing you to load balance and index more data in less time. When a Splunk deployment is configured to use a group of indexers clustered together, one of the benefits is replication of indexed data across the different indexers. This replication protects the indexed data, and supplies indexer cluster node failover. In the context of searches, the indexer nodes are referred to as peer nodes. Supporting peer node failover also ensures data availability, even in the event of a peer node failure. Splunk typically recommends that the Replication Factor (RF) be set one digit higher than the Search Factor (SF). This means that you have multiple searchable copies of all indexed data, plus one or more additional non-searchable copies of the raw indexed data. Clustering your indexers also delivers the benefit of allowing Splunk to manage replication of indexed data automatically, resulting in faster searches because the search workload is spread over several indexers.

Benefits of HPE Primera and HPE Nimble Storage arrays

Reduce RF to equal SF

Setting your RF higher than your SF is not needed when the indexed data is stored on an enterprise-class storage array, such as HPE Primera or HPE Nimble Storage arrays. Having the RF higher than the SF is especially useful when all indexed data is stored on local disks in the indexers. However, in the solution described in this paper, the HPE storage arrays were already configured with redundant storage (RAID-6), so the data is already protected by the array itself from server loss as well as drive failure. With this data protection already in place on the array, an extra copy of the raw event data is not needed.

If an indexer node goes down in a configuration similar to that described in this paper, nothing happens to the indexed data stored on the enterprise-class array. When the indexer is restored to functionality, or a new one is installed, you simply mount the storage volume file systems on the server, configure the `indexes.conf` file to point to the existing indexes on the array, and all the indexed data is still there and searchable. If the indexer was part of an index cluster, mount the storage volume(s) on the new indexer, and apply the cluster bundle from the index cluster master to the new indexer. Depending on the circumstances of the outage, the cluster master may detect that there are over-replicated buckets resulting from the bucket fixup activity while the indexer was down. These extra replicas are not deleted automatically, but can be easily deleted through the Splunk GUI or CLI.

Implementing HPE Primera or HPE Nimble Storage arrays to hold the indexed data eliminates the need for additional storage capacity to store an extra raw, unsearchable copy of all indexed events. With this configuration, Splunk's RF can be reduced to the same value as the SF, without any risk of data loss. As added benefits, network congestion is reduced, and data indexing appears to run slightly faster without the overhead of writing an extra raw copy of all indexed events.

To illustrate the capacity savings of eliminating the extra non-searchable copy of raw data by lowering RF to be the same value as SF, consider the example in TABLE 1. These figures are based on the rule of thumb for the size of indexed data after it has been processed and stored by Splunk. For a full explanation of this rule of thumb, see the Sizing a traditional Splunk implementation section later in this paper.

TABLE 1. Example capacity reduction by eliminating the extra non-searchable copy of indexed data

	No Replication	RF = 3, SF = 2	RF = 2, SF = 2
Daily indexing rate	1 TB	1 TB	1 TB
Retention time	365 days	365 days	365 days
Total capacity required	182.5 TB	419.75 TB	365 TB

As seen in TABLE 1, when the daily indexing rate is 1 TB per day and the data is retained for one year, maintaining an extra non-searchable copy of indexed data increases the required capacity by 15% (approximately 55 TB). This percentage is based on the size of the compressed raw data file, and will remain relatively constant even with different rates of indexing and varying lengths of retention.

In our testing for this solution, both RF and SF (see column three in TABLE 1) were set to **2**. This resulted in having two copies of all indexed events spread across the index cluster, and both copies were fully searchable. There were no extra non-searchable copies of indexed events in our testing because they were not needed.

Data protection from media failure

As described in the previous paragraphs, another benefit of using an enterprise-class array is that the impact of disk or SSD failure is significantly reduced compared to configurations which store all data on local disks in the indexers. When all data is stored on internal direct attached storage (DAS) in the indexer, a media failure would result in bucket fixup if the internal disks are not set up in at least a RAID 10 mirror configuration. This can be a lengthy and resource-intensive process. However, with an enterprise-class array, such as HPE Primera or HPE Nimble Storage arrays, the only impact of a media failure is a minimal reduction in performance as the RAID rebuild runs on the array.

Read-ahead capability for faster searches

HPE enterprise-class storage arrays have the advantage of a built-in read-ahead capability. The arrays detect reads of sequential streams, and prefetch data. In the use cases tested here with Splunk, this prefetch function brings indexed data being searched into the array's cache. With this capability, subsequent searches are faster since the indexed events do not need to be read directly from disk for every search.

HPE Priority Optimization software

HPE Priority Optimization software for the HPE Primera storage array enables Quality of Service (QoS) levels for applications and workloads based on business requirements. QoS is an essential element in multitenant storage architectures. By implementing HPE Priority Optimization, storage administrators can deliver a predictable service level for their Splunk environment, and manage bursts of data irrespective of other users on a shared storage array. With the ability to provision storage performance as you provision storage capacity, you can assure the necessary level of performance for Splunk without the need to partition physical resources. In this testing, Splunk was the only workload in use, so HPE Priority Optimization was not used.

HPE Peer Persistence software

HPE Peer Persistence software for HPE Primera arrays leverages the robust, high-availability solutions already accessible on HPE Primera systems, and extends them even further by enabling a peer relationship between two HPE storage systems located in the same data center or at geographically separated sites. HPE Peer Persistence enables a highly available, completely redundant storage infrastructure, and provides continuous access to data even when one storage system has planned or unplanned downtime. If implemented across sites, HPE Peer Persistence allows users to exercise both their primary and secondary sites actively, which results in more efficient utilization of IT infrastructure and reduces overall IT expense. This federated, high-availability feature is implemented on top of existing HPE Remote Copy infrastructure, with no need for additional hardware or appliances.

Storage snapshots

One of the valuable features of HPE Primera and HPE Nimble Storage systems is the ability to create snapshot copies of data. In any of the HPE enterprise-class storage arrays, a volume collection can be configured to create a snapshot of volumes to ensure the consistent backup of all Splunk Enterprise data buckets and virtual machine system disks.

Hot buckets in Splunk are not normally backed up, although they are replicated across the index cluster if such a cluster has been implemented. In order to have backups of data in the hot buckets, it is possible to use the storage snapshot capability of the HPE Primera or HPE Nimble Storage arrays. If an index cluster has not been implemented, there is no automatic replication of hot buckets. Therefore, in an environment where indexed events are only stored on local disks on a non-clustered indexer, the content of the hot buckets is vulnerable to loss when an indexer fails. HPE storage snapshots greatly reduce this risk by capturing the content of all Splunk's buckets, including the hot buckets.

A storage snapshot captures all data in the virtual volume on the array at the time the snapshot is taken, thus including the content of the hot buckets along with the other buckets on the volume. These snapshots are very fast and require very little physical space when created. Because of these advantages, storage snapshots can be generated more frequently than traditional backups, providing more recovery points.

In addition to backing up the content of hot buckets, storage snapshots can also be used to provide robust backups of the VMs in a virtualized Splunk environment. If the system disks of the VMs hosting your Splunk instances are contained in datastores created from storage on an HPE storage array, RMC can be used to schedule and replicate backups of those VMs. This provides more complete protection not only for the indexed data in your environment but also the VMs running Splunk instances as well.

To illustrate the capacity savings enabled by storage snapshots, consider the following example:

A Splunk environment has an indexing rate of 2 TB per day, with a retention period of 20 days before the data is rolled to frozen buckets on archival storage. The indexers are clustered, with RF and SF both set to **2**, as described earlier. This results in a total data size of 40 TB on the HPE array. Further, assume that the SLA requires a daily full restore point, which is retained for seven days. With a traditional backup implementation, the space required for seven full backups would be 40 TB x 7 days = 280 TB. However, using HPE storage snapshots to provide the same level of backup and restore capability would require only 40 TB + (6 days x 2 TB per day) = 52 TB. Therefore, HPE storage snapshots produce an 81% capacity savings compared to traditional backup methods.

These snapshots can be used at a later time, if needed, to restore the content of not only the warm and cold buckets but also the hot buckets, which are not normally backed up in Splunk. The VMs running your Splunk instances can also be restored from storage snapshots to provide rapid and robust recovery. HPE Recovery Manager Central (RMC) ships free with HPE Primera and HPE Nimble Storage arrays, and can be used to schedule and manage storage snapshots. The Peer Copy feature within HPE RMC allows these storage snapshots to be replicated directly between HPE Primera and HPE Nimble Storage arrays. This implementation eliminates the need to use host specific or hypervisor specific tools or expensive network-based appliances to manage the replication between arrays. HPE RMC also supports the ability to move replicas to the cloud for longer retention, or in combination with HPE Cloud Volumes, to facilitate use cases where the Splunk data needs to be accessed from cloud instances.

Data reduction

Data reduction, a combination of deduplication and compression, can be enabled on a virtual volume when it is created on the HPE Primera or HPE Nimble Storage array. Deduplication is supported for SSD-based virtual volumes only. To achieve the best advantage, deduplicated virtual volumes should be formatted on the host using a 16 KiB allocation unit.

We tested with and without data reduction enabled on the volumes containing the indexed data. A small improvement in storage capacity utilization was observed when deduplication and compression were enabled on the array. Splunk compresses the raw events as they are written to the indexes on disk, but does not compress the `tsidx` (index) files containing location metadata about the indexed events that have been written to the raw data file. By default, Splunk uses gzip to compress events as they are indexed. This is set in the `indexes.conf` file in the `$SPLUNK_HOME\etc\system\default` directory. In our testing, Splunk's compression with gzip produced a data reduction of about 12x on the raw data file. Because gzip is such an efficient compression algorithm, the already-compressed data cannot be compressed much further when it is written to a storage array. Therefore, the primary gain that can be obtained with the data reduction provided by a storage array is a small amount from compressing the `tsidx` files which contain metadata and are not, by default, compressed by Splunk.

When deduplication and compression were enabled in our testing, we observed a data reduction factor of 1.3. This gain in efficiency represents the combination of gains in deduplication and compression together. The space savings and performance impact of data reduction are heavily dependent on the content and characteristics of the events being indexed. In other configurations with other types of data, there is evidence that higher values of reduction can be obtained.

In certain cases where the indexers are CPU bound, it may be helpful to switch Splunk's compression algorithm from gzip to lz4 in the `indexes.conf` file. The lz4 compression algorithm uses less server compute resources than gzip, but is less effective at compression than gzip. If compression is changed to lz4, it is important to make sure that data reduction is enabled on the storage volume(s) where the indexed data is written. This has the effect of shifting some of the CPU overhead of compression activity from the indexers to the storage array, and may allow you to index data at a slightly higher rate without having to upgrade your indexers. This configuration was not tested as a part of this activity.

Benefits of supplementing HPE enterprise-class storage with archival storage

Archive storage solutions, such as HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING, enable you to off-load the oldest event data from your enterprise-class storage arrays to an archive, and lower your overall storage costs.

To minimize the cost of storage, Splunk's hot and warm buckets can be kept on SSD-based storage on an HPE enterprise-class storage array, while cold and frozen buckets can be moved to less costly storage on Scality RING, using NFS file shares. This implementation will perform quite well, and demonstrates the value of tiered storage. Using this approach keeps Splunk's most frequently searched data on the highest performing storage tier, with less frequently searched data moving to a lower tier as it ages, while maintaining good search response.

In cases where storage tiering is not required, Splunk's hot, warm, and cold buckets can all be contained in SSD-based storage. This implementation performs exceptionally well, as demonstrated in testing with the HPE Primera all-flash storage array. The fast response time for searches across hot, warm, and cold buckets produced search results with minimal latency.

If the cold buckets and frozen buckets reside on the same storage volume, the frozen path (**coldToFrozenDir**) can be specified in Splunk's `indexes.conf` file, along with **homePath** and **coldPath**. As an alternative, if the frozen path is on a different volume or a different storage system, such as Scality RING, the `indexes.conf` file provides a mechanism for the frozen data movement to be initiated (**coldToFrozenScript**). Splunk provides an example of such a script with the software in **$SPLUNK_HOME/bin/coldToFrozenExample.py**, but it is up to the user to write a script in the language of their choosing to handle moving the data. This script is then called automatically by Splunk every time a bucket needs to roll from cold to frozen. After the bucket is copied to the archive location, Splunk deletes it from the cold path.

In our testing, a script written in Python was used to facilitate moving older indexed events from cold buckets to frozen buckets on the HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING. MD5 checksums were calculated on the raw data files in the cold buckets prior to copying to the frozen location on the Scality RING. As buckets are written to the frozen destination on the Scality RING, an MD5 checksum is included as part of the metadata. These MD5 checksums were read and compared with the checksums calculated on the raw data files in the cold buckets before they were rolled to frozen buckets. In all cases, the MD5 checksums matched exactly, confirming that the buckets were copied accurately.

Scality RING provides the flexibility to support single-site and geo-dispersed configurations. The customizable policies provide data protection at the object level, and include geo-redundancy, so there is protection against failures at the disk, server, rack, and even site levels. Whether single-site or geo-dispersed, the HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING scales virtually without limits, so you can grow your capacity as needed with linear performance in less space and with extreme data availability.

A single Scality RING can simultaneously be used for many purposes, in addition to Splunk's cold buckets and long-term retention of event data in Splunk's frozen buckets. Other non-Splunk uses could include hosting scale-out object storage as a service, nearline archives for media and entertainment companies, medical imaging and records, long-term archives of surveillance information or research data, financial compliance archives, and many other applications.

SPLUNK SMARTSTORE WITH HPE SERVERS AND STORAGE AND HPE SCALABLE OBJECT STORAGE WITH SCALITY RING

The purpose of Splunk's SmartStore feature is to minimize the amount of data stored on local disks, while maintaining fast indexing and search capabilities. As soon as data rolls out of the hot buckets, it gets copied to warm buckets on S3 object storage, such as HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING (Scality RING). The main difference between SmartStore and the traditional implementation of Splunk is that with SmartStore, the master copies of warm buckets are stored on a remote S3 object store, while the indexer's local disks are used to cache copies of warm buckets. The indexer contains a cache manager that maintains a cache of warm buckets on the indexer's local disks, which are usually SSD. This reduces the need for the indexer to fetch warm buckets from remote storage when they are needed for repeated searches. When the indexer needs to search a warm bucket for which there is no cached copy, the cache manager fetches a copy of the master bucket from the object store, and places it in local cache.

With Splunk SmartStore, cold buckets are no longer needed because the warm buckets are already offloaded to less expensive object storage. When buckets age out of the warm tier, they roll directly to frozen if a frozen path is configured. Otherwise, they are deleted.

In versions of Splunk prior to 8.0, when SmartStore is used in a deployment with a cluster of indexers, only the hot buckets are replicated across peer nodes of the cluster. Replication of the warm buckets is managed internally by the object storage system. If several indexers go down simultaneously, no warm data is lost because the object store already contains multiple copies of the warm buckets. Availability is not affected, but search performance can be degraded until local caches of surviving indexer peer nodes are repopulated. Beginning with Splunk version 8.0, Splunk's Replication Factor is enforced for cached warm buckets. These versions of Splunk will manage hot and warm bucket replication in addition to having multiple copies of warm buckets managed internally by the object store.

Using Splunk SmartStore with HPE servers and storage and Scality RING, S3 object storage allows users to decouple their storage and compute layers. This gives you the flexibility to scale each resource independently, and avoid the typical imbalance with demand for storage outpacing demand for compute. The result is higher utilization of all your resources, and lower total cost of ownership (TCO).

SmartStore enables you to optimize your data storage costs based on native, transparent tiering. Only Splunk's hot buckets are stored locally on the indexers, using the top tier of storage capacity, while frequently used warm buckets are kept in the indexer's cache. Older, less frequently searched events which age and roll out of Splunk's warm buckets can be rolled to frozen buckets on the S3 object storage which is significantly more cost effective for infrequently accessed data. By implementing this tiered approach to bucket management, large datasets remain fully searchable at the lowest possible storage cost.

Deploying and managing your storage is simplified with SmartStore, and has higher elasticity. SmartStore's disaggregated architecture is a perfect match for the limitless scalability, high throughput performance, and extreme resiliency of Scality RING.

When an HPE Primera or HPE Nimble Storage all-flash array is used in a Splunk SmartStore environment, SSD-based block storage from the array can be used as primary index storage and as cache for your indexers. This reduces the need for additional SSDs in each indexer to contain the hot buckets and cached warm buckets. In addition, it provides increased data protection for hot buckets that have not yet been written to the object store (RAID 6, compared to using RAID 10 mirroring with internal SSDs on the indexers). Furthermore, configuring and managing storage across multiple indexers with HPE storage arrays and VMWare plugins can be simpler and easier than managing individual SSDs on each indexer.

Scality RING and HPE servers and storage are SmartStore Ready for single-site implementations. Testing SmartStore with HPE servers and storage and Scality RING is beyond the scope of this paper, so this configuration was not implemented during our testing.

CAPACITY AND SIZING

Sizing a traditional Splunk implementation

When sizing the storage for Splunk, there are three main factors to consider:

1. Maximum daily indexing rate
2. Retention time of your data
3. Replication Factor (RF) and Search Factor (SF) for data replication across clustered indexers

See the Splunk documentation Estimate your storage requirements for more information.

The Splunk indexers extract events from incoming data and write them to a file referred to as the "raw data" file, which is automatically compressed. An index (`tsidx`) file is also created, which contains location metadata about the events in the raw data file. This file makes the events searchable, and is not compressed by default. In addition to the raw events written to the raw data file, sufficient metadata is also written to the raw data file to enable recreation of the `tsidx` file when necessary. For example, the `tsidx` file is deleted when data rolls from cold buckets to frozen buckets. If frozen data is needed again and is moved to a thawed directory, the `tsidx` file must be recreated in order to make the events searchable again.

The rule of thumb from Splunk is that once data has been indexed and compressed, one replica occupies on average approximately 50% of the volume of the original data. The actual data footprint can vary based on the type of data being indexed. Using the rule of thumb in an example, if data is being indexed at a daily rate of 100 GB with no replication of indexes, the amount of data for one replica written to disk by the indexer(s) will be approximately 50 GB per day. This includes approximately 15 GB for the compressed raw data file and 35 GB for the `tsidx` file, which is not compressed. If indexes are being replicated through use of a cluster of indexers, the amount of data written to disk per day will be multiplied by the number of copies of the index (specified by the values of RF and SF). Thus, for an indexing rate of 100 GB per day with RF = **2** and SF = **2**, the amount of storage would be about 100 GB because you are keeping two copies of everything. If RF is set higher than SF, the data footprint will be increased by the size of the additional copy (or copies) of the non-searchable raw data file. Hence, if RF = **3** and SF = **2**, the estimated required storage would be about 115 GB.

It is essential to understand how retention time is affected by the capacity of the storage where your indexes reside and by the settings for your indexes. For example, when using the default parameters for maximum size of indexes, retention times, and so forth, data retention time at an indexing rate of 2 TB per day is only a few hours instead of days or weeks. Using these default parameters, it is very possible for new data to be indexed into hot buckets, rolled to warm, rolled to cold, and then moved to frozen in just a matter of hours. If no frozen path is defined, the data is simply deleted when it rolls out of the cold buckets. All this rolling of data is based solely on the extremely small capacity being consumed so quickly. In effect, the index files act as a First In, First Out (FIFO) storage device with very little depth.

Very short retention times, as just described, hold true even if the capacity of the storage is much larger than required. The excess capacity will be unused due to the default settings in Splunk which limit the number and size of hot and warm buckets. To retain data for a longer period of time, it is necessary to understand the capacity needed as well as the parameters that affect retention, and configure them appropriately. Replication of indexed data is handled automatically in Splunk, but you must ensure you have enough capacity for the indexed data along with the data replicated across indexers.

The following settings determine the size and number of your index buckets, the retention times, and how the storage is used. These parameters are set in the `indexes.conf` file, and can be different for each index. Care should be taken in determining appropriate values for these parameters. The default settings are often not acceptable, especially at daily indexing rates higher than 100 GB per day.

- **maxTotalDataSizeMB** defines the maximum size an index can grow to (in MB).
  - The default is **500,000 MB**.
- **maxDataSize** defines the maximum size a hot bucket can reach before rolling to warm (in MB).
  - The default is **auto**, which sets the max hot bucket size to 750 MB.
  - For high volume indexes (an index that gets >10 GB per day), Splunk recommends setting this to **auto_high_volume**, which sets the max hot bucket size to 1 GB on 32-bit systems or 10 GB on 64-bit systems.
  - It can be set to any value up to 1 TB, but typically set it to something between 100 MB and 50 GB.
- **maxHotBuckets** sets the maximum number of hot buckets allowed in an index.
  - The default is **3**.
- **maxWarmDBCount** sets the maximum number of warm buckets allowed in an index.
  - The default is **300**.
- **coldPath.maxDataSizeMB** defines the maximum size the cold buckets of an index can grow to (in MB), if the cold buckets are in a different path from the hot/warm buckets defined by **homePath** and the **maxTotalDataSizeMB** parameter described earlier. There is no default value for this setting. If it is not specified, the maximum size of the directory defined by **coldPath** is not constrained by Splunk.
- **frozenTimePeriodInSecs** defines the maximum length of time events will be retained. This sets the number of seconds until indexed data rolls to frozen. Thus, it includes all time that indexed data is held in hot, warm, and cold buckets combined. If no frozen path is defined, the default behavior will apply which is to delete data when it rolls out of the cold buckets.
  - The default is **188,697,600 seconds** (6 years).

TABLE 2 shows an example of the capacity required for indexed data, assuming there is no replication of the indexed data. If index replication is used due to the implementation of clustered indexers, the need for storage capacity will increase significantly, as shown in TABLE 3.

See How Splunk Enterprise calculates disk storage for details on capacity requirements. Based on this Splunk document, a general rule of thumb for calculating storage capacity where there is no replication of indexed data is as follows:

- Base storage capacity = (daily indexing rate) x (number of days of retention) x 0.5

If indexes are replicated across a cluster of indexers, the base storage capacity determined from the prior calculation should be multiplied by the Search Factor as follows:

- Storage capacity with index replication = (Base storage capacity) x (Search Factor)

NOTE
The above equation assumes that Replication Factor and Search Factor are set to the same value, as recommended elsewhere in this paper.

TABLE 2 and TABLE 3 summarize the total amount of storage required for the indexing rates shown, with varying levels of retention. Specifically, TABLE 2 lists the amount of storage required when there is no replication of indexed data, as would be the case when indexers are not clustered.

TABLE 2. Example storage capacity based on index rate and retention period with no replication of indexed data

Daily indexing rate	Retained for 2 weeks	Retained for 30 days	Retained for 90 days	Retained for 365 days
300 GB/day	2.1 TB	4.5 TB	13.5 TB	54.75 TB
1 TB/day	7 TB	15 TB	45 TB	182.5 TB
2 TB/day	14 TB	30 TB	90 TB	365 TB

TABLE 3 lists the storage required in a configuration with a cluster of indexers using replication, and showing two different settings for replication. The first column under each retention period is for a configuration where RF = **2** and SF = **2**. The second column under each retention period is for a configuration where there is one additional non-searchable copy of indexed data, thus RF = **3** and SF = **2**. The capacity calculations in TABLE 3 and TABLE 4 are based on the rule of thumb for storage sizing discussed previously. Your actual capacity requirements may vary depending on the type of data indexed and how compressible it is.

The sizes in TABLE 3 represent the total capacity that should be distributed across all the indexers in the deployment, not the amount of storage required for each indexer. TABLE 3 and TABLE 2 show sample configurations and recommended storage capacities for different sizes of Splunk Enterprise deployments. Customers should work with their Splunk architects and their Hewlett Packard Enterprise account team to determine the correct solution for their data indexing and performance requirements.

TABLE 3. Storage capacity based on index rate and retention period with replication of indexed data

Daily indexing rate	RF = 2, SF = 2 Retained for 2 weeks	RF = 3, SF = 2 Retained for 2 weeks	RF = 2, SF = 2 Retained for 30 days	RF = 3, SF = 2 Retained for 30 days	RF = 2, SF = 2 Retained for 90 days	RF = 3, SF = 2 Retained for 90 days	RF = 2, SF = 2 Retained for 1 year	RF = 3, SF = 2 Retained for 1 year
300 GB/day	4.2 TB	4.83 TB	9 TB	10.35 TB	27 TB	31.05 TB	109.5 TB	125.925 TB
1 TB/day	14 TB	16.1 TB	30 TB	34.5 TB	90 TB	103.5 TB	365 TB	419.75 TB
2 TB/day	28 TB	32.2 TB	60 TB	69 TB	180 TB	207 TB	730 TB	839.5 TB

Sample minimum configurations for an HPE Primera all-flash array that would accommodate each indexing rate tested, and retain the data for one year, are shown in TABLE 4. Similar capacity and performance can be achieved with an HPE Nimble Storage AF60 or AF80 array, as exhibited in TABLE 5. These configurations are based on having Replication Factor and Search Factor both set to **2**, as described earlier. Many other configurations are possible; these are just provided as examples that would have the minimum capacity required at each indexing rate with tiered storage. These configurations assume that Splunk's hot, warm, and cold buckets are retained on the HPE Primera or HPE Nimble Storage array. If an additional storage tier such as Scality RING with the NFS connector is used for Splunk's cold buckets, the capacity required on the HPE Primera or HPE Nimble Storage array would be reduced accordingly.

The array specifications listed in TABLE 4 and TABLE 5 provide plenty of system resources to handle all the indexing rates tested, while simultaneously servicing a very significant load of interactive, automated, and scheduled searches.

NOTE
Any of the configurations in TABLE 4 or TABLE 5 can be scaled up as necessary by either adding more or larger drives, drive enclosures, controller nodes, or all, as needs increase.

TABLE 4. Sample minimum HPE Primera configurations based on index rate and retention period

	300 GB/day	1 TB/day	2 TB/day
Retention period	365 days	365 days	365 days
Minimum required capacity	110 TB	365 TB	730 TB
Total capacity	114.24 TB usable 153.55 TB raw	367.76 TB usable 475.99 TB raw	736.48 TB usable 951.98 TB raw
Array model	HPE Primera A630	HPE Primera A650	HPE Primera A670
Number of controllers	2	4	4
Host ports	8 x 16 Gb Fibre Channel	16 x 16 Gb Fibre Channel	16 x 32 Gb Fibre Channel
Number of disks	20 x 7.68 TB SSD	62 x 7.68 TB SSD	62 x 15.36 TB SSD
Input/output operations per second (IOPS), 8K random read/write	98K IOPS	250K IOPS	295K IOPS
IOPS, 8K random write	41K IOPS	110K IOPS	115K IOPS
Throughput, 256K sequential write	817.2 MB/s	1.8 GB/s	2.2 GB/s
Throughput, 256K sequential read	7.7 GB/s	17.9 GB/s	23.3 GB/s

TABLE 5. Sample minimum HPE Nimble Storage configurations based on index rate and retention period

	300 GB/day	1 TB/day	2 TB/day
Retention period	365 days	365 days	365 days
Minimum required capacity	110 TB	365 TB	730 TB
Total capacity	136.44 TB usable 184 TB raw	375.2 TB usable 506 TB raw	751.5 TB usable 1 PB raw
Array model	HPE Nimble Storage AF60	HPE Nimble Storage AF60	HPE Nimble Storage AF80
Host ports	8 x 16 Gb Fibre Channel	8 x 16 Gb Fibre Channel	8 x 16 Gb Fibre Channel
Number of disks	24 x 7.68 TB SSD	48 x 7.68 TB SSD 24 x 3.84 TB SSD 24 x 1.92 TB SSD	120 x 7.68 TB SSD 24 x 3.84 TB SSD
Throughput, 256K sequential write	2.6 GB/s	2.6 GB/s	3.8 GB/s
Throughput, 256K sequential read	9.0 GB/s	9.0 GB/s	9.0 GB/s

Sizing with Splunk SmartStore

When sizing HPE storage for a SmartStore implementation, the following factors must be considered:

1. For best performance of both indexing and searches, local storage on the indexers should be SSD. The amount of local SSD storage on each indexer should be sized based on the daily indexing rate and hot bucket retention requirements to accommodate hot buckets with replication. In lieu of local SSD storage in each indexer, a similar amount of SSD-based block storage from an HPE array could be used.
2. In addition to the capacity needed for hot buckets and replication, SSD storage on the indexers (or HPE storage array, if implemented) will also serve as local cache for warm buckets. With versions of Splunk prior to 8.0, the SSD capacity must be sized to include a single searchable copy of each warm bucket to be held in cache. For example, if you want to do fast searches that go back seven days, you would size the cache to hold seven days of searchable warm buckets with no replication, in addition to the capacity needed for hot buckets with replication. Beginning with Splunk version 8.0, the capacity needed for caching of warm buckets must be increased to allow a number of warm bucket replicas equal to the Replication Factor (RF) set in Splunk.
3. Size the HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING for a single, searchable copy of each warm bucket, per warm bucket retention settings. Bear in mind that the object store by design contains multiple copies of the warm buckets.
4. In addition to the capacity needed for warm buckets, the object storage will also be used for frozen buckets (compressed raw data files). Include enough capacity to hold the frozen buckets, per frozen data retention requirements.

For more information about Splunk SmartStore, see the Splunk documentation SmartStore architecture overview.

PERFORMANCE ANALYSIS AND RECOMMENDATIONS

The performance of Splunk is heavily dependent on the performance of the storage. Specifically, indexing performance is contingent on the write speed of the primary storage where the hot and warm buckets are located (the **homePath** location). Search performance is reliant on the read speed and latency of the storage where the hot/warm buckets and the cold buckets are located (the **homePath** and **coldPath** locations, respectively).

Writes to the volumes on the array are made up of several different components, including:

- New events being indexed
- Activity related to index replication across clustered indexers
- Data movement, as indexed events age and roll to lower tiers within Splunk

Reads from the volumes on the array are primarily made up of two components:

1. Searches of indexed events, either initiated by users, or by scheduled or automated reports, or dashboards
2. Data movement, as indexed events age and roll to lower tiers within Splunk

Performing searches over long time periods will search across old data stored on volumes that may be on a lower tier of slower storage. These searches will be slower than doing a specific search on recent events in the highest tier of high speed volumes which reside on your enterprise-class storage. Whenever possible, it is best to limit search scope via the time range picker in Splunk so that only the hot and warm buckets will be searched. The time range picker is highlighted in FIGURE 2.

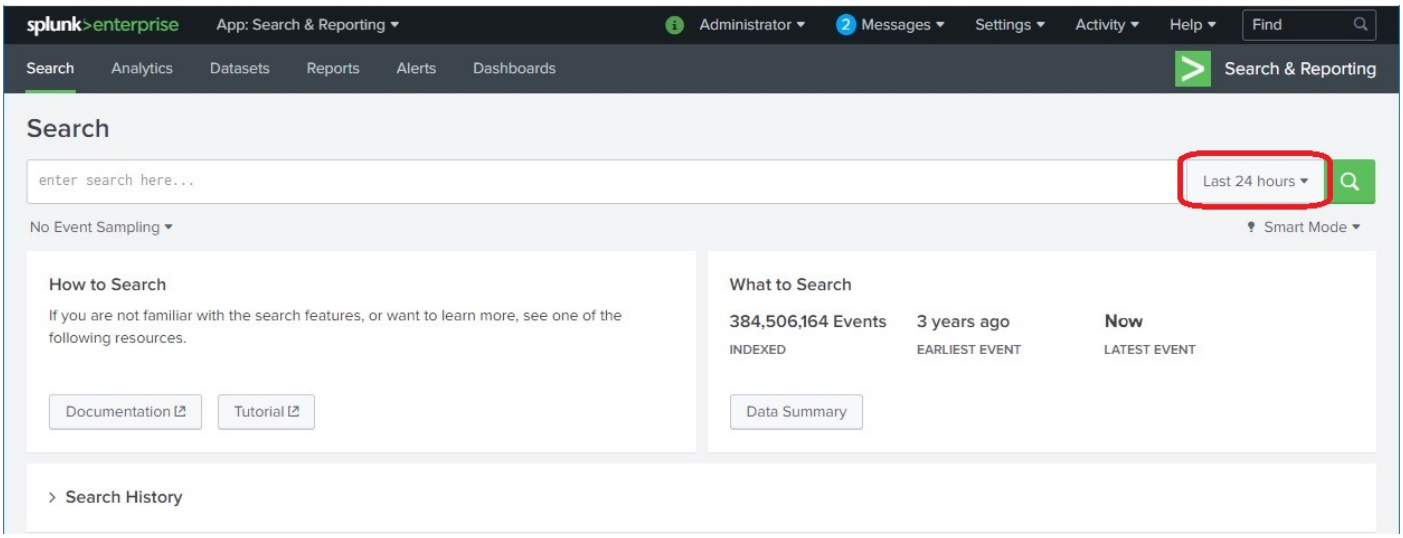

FIGURE 2. Splunk's "Search" page with the time range picker highlighted

Splunk and Hewlett Packard Enterprise recommend defining **homePath** (for hot/warm buckets) on your high speed disks, such as SSD, and defining **coldPath** (for cold buckets) on the slower disks which still have a good read speed, such as SAS HDD. If you are using an all-flash array, **homePath** and **coldPath** should both be on the same tier of high speed storage. If you are using frozen storage for archiving and longer retention, configure the frozen buckets to reside on archive storage, such as HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING, and create an appropriate script to manage copying the data from the cold location to the frozen location.

An HPE Primera all-flash storage array and an HPE 3PAR storage array were tested with varying workloads. One of these workloads consisted of mostly writes, while another higher workload was comprised of both writes and reads. Early testing involved processing events at indexing rates of 300 GB per day, 1 TB per day, and 2 TB per day. During this portion of the testing, the reads on the array were mostly from activity related to replication of buckets across the indexers and data moving to a lower tier of storage as events aged. A small amount of read activity was generated by automated and interactive searches.

Later testing was done using a mixed workload consisting of writes at an indexing rate of 2 TB per day with a simultaneous heavy read load generated by automated concurrent searches. Reads during this portion of the testing still included activity related to replication of buckets and data movement as events aged, along with the automated searches. In addition to the higher workload, this testing added several more indexers to the indexer cluster compared to the number of indexers present in the cluster during the previous phase of testing.

FIGURE 3 shows the average response time of the HPE Primera all-flash array with the mixed workload. The response times observed with the HPE 3PAR array were similar to those shown here. The time displayed in the chart represents write and read response times at each data point averaged together. In this test, the daily indexing rate was kept steady at 2 TB per day, while a heavy read load was added generated by automated searches.

In cases where array-based data reduction is enabled, reads are expected to take slightly longer compared to cases where data reduction is not used. This slight increase is due to the additional metadata lookup necessary to reconstruct the data on each read operation.

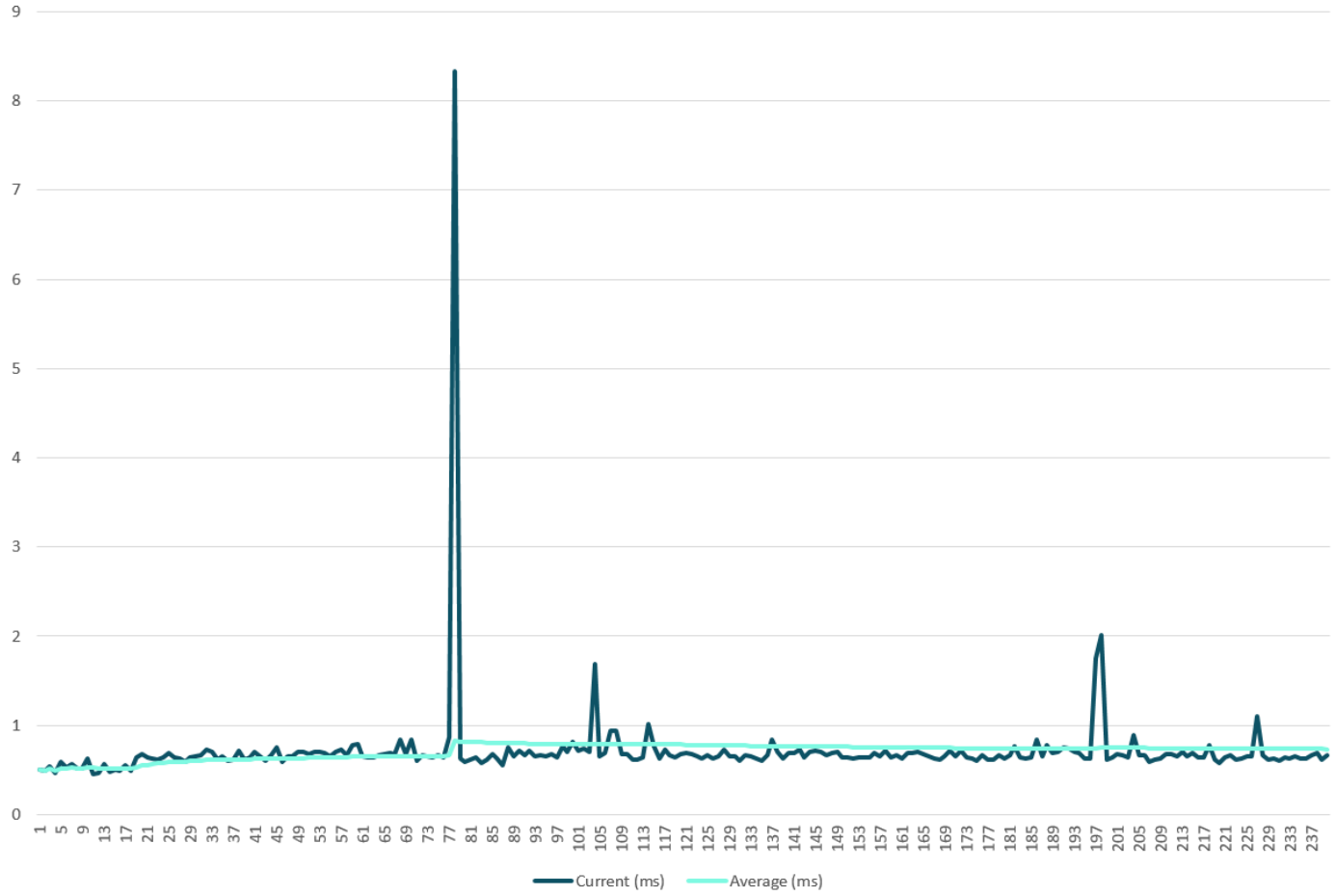

FIGURE 3. Current and average time to service I/O requests with indexing rate of 2 TB per day and simultaneous heavy search load

The write and read response times were being measured in real time on the array, and noise in the sampling caused a few outliers to be reported over the 240 sample periods, which covered a test span of 40 minutes. By far, the largest of these outliers was a response time of 8.33 milliseconds reported in the 79th sample, as seen in FIGURE 3**.** The typical response time of the HPE Primera array was between 600 and 800 microseconds. These very low response times for both reads and writes enable rapid indexing and searching of data without worry that the storage will become a bottleneck.

FIGURE 4 exhibits the aggregate throughput for all indexers in KB/s. The chart indicates that the volume of reads started out relatively low, then increased sharply. This is because when the data collection was begun, the automated searches had not yet been initiated. During those initial moments of the collection period, the only read activity present on the array was that associated with bucket replication and normal movement of events due to aging. As soon as the scripted searches were launched, the read load on the array increased significantly and leveled out at a rate of about 1.2 GB/s, while the write load remained mostly constant with several peaks.

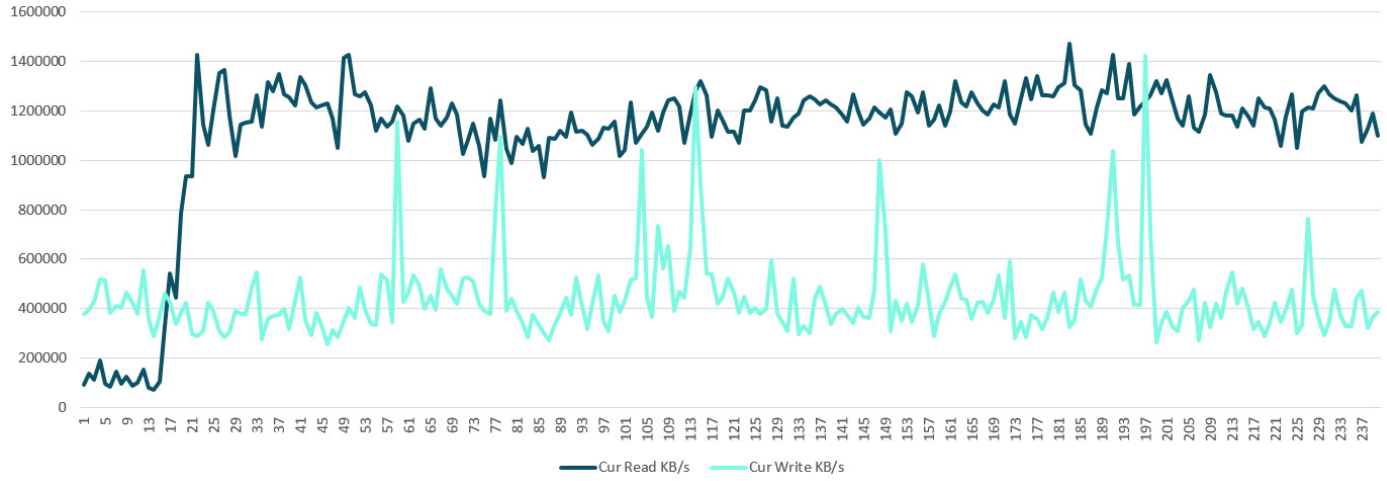

FIGURE 4. Array throughput with HPE Primera all-flash array at an indexing rate of 2 TB per day and simultaneous heavy search load

As FIGURE 4 demonstrates, the HPE storage array is not heavily loaded, even at a daily indexing rate of 2 TB per day with a significant workload of concurrent searches. In FIGURE 5, a sample of the aggregate throughput of all indexers at a moment in time is shown in the **Indexing Performance: Deployment** page of Splunk's **Monitoring Console**.

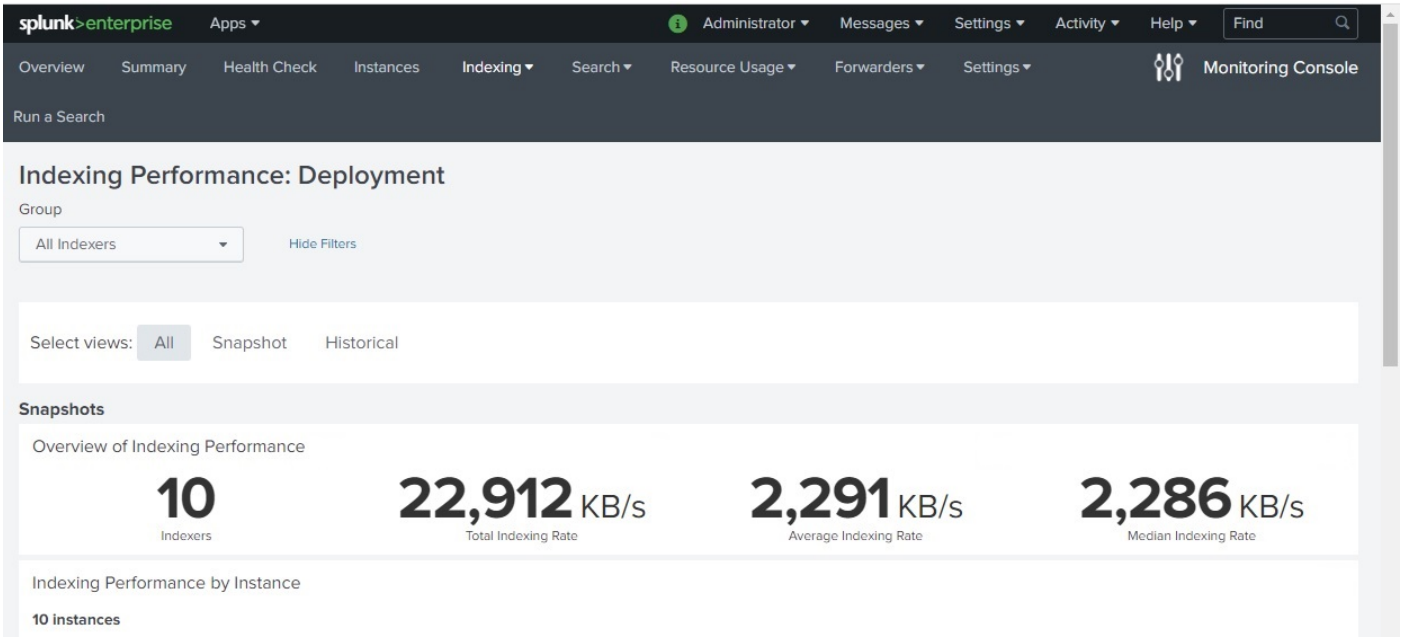

FIGURE 5. Indexing performance at an indexing rate of approximately 2 TB per day

A minimally configured HPE Primera with two nodes, as described in TABLE 4, is capable of servicing the entire workload of writes and reads with resources to spare. Hewlett Packard Enterprise offers several different enterprise-class arrays that can easily carry the workloads described in this paper, with plenty of bandwidth for growth and other uses. HPE storage arrays can readily accommodate multitenant I/O, especially with the implementation of HPE Priority Optimization.

FIGURE 6 shows the CPU utilization for the CPUs in the HPE Primera all-flash array during testing. The line at the bottom of the graph, marked **sys** shows the percentage of CPU utilization for system processes within the array. The middle line, marked **user**, shows the amount of CPU utilization for user processes. Array-based data reduction activity operates in **user** mode on the array, so it is included in this value. The top line of the chart simply displays CPU **idle** time.

This data was captured during the same test run as the response times reported in FIGURE 3, and the throughput data reported in FIGURE 4. As with the data shown in FIGURE 4, the first few moments represented in the chart show that CPU utilization was between 20 and 30 percent until the automated searches were started. Once the searches began, the CPU utilization increased as expected, showing the increased workload on the array CPUs to service the reads and to perform the extra metadata lookup required due to data reduction. The CPU utilization observed in the HPE 3PAR array was almost identical to what is shown in FIGURE 6 for the HPE Primera all-flash array.

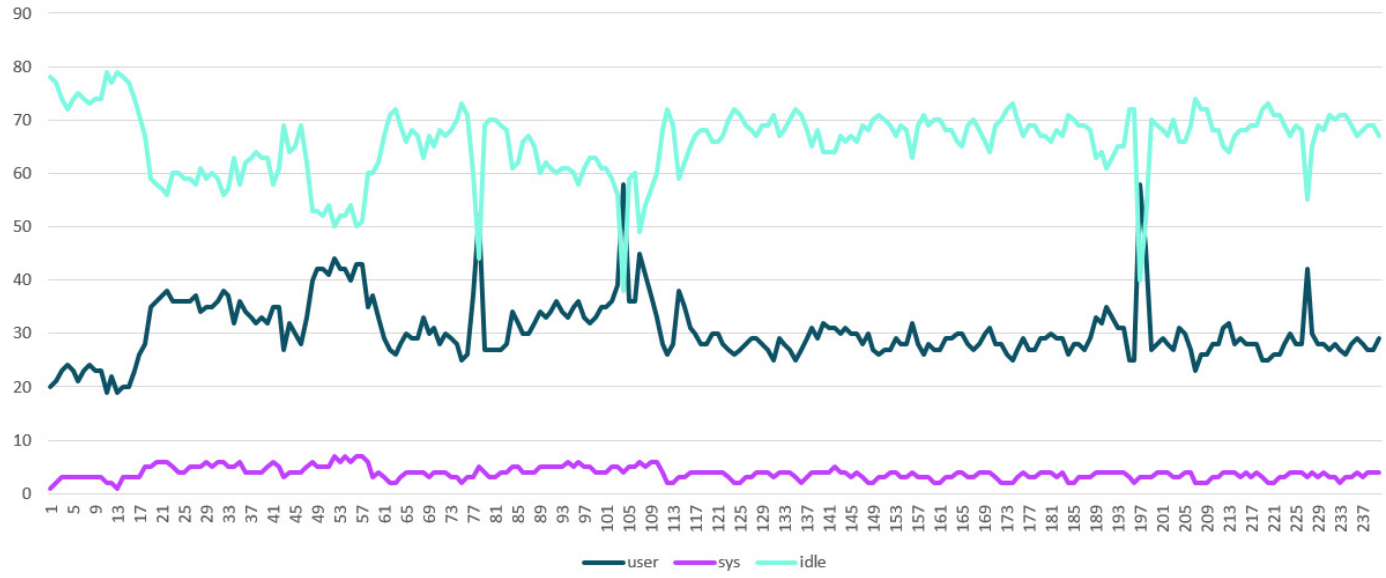

FIGURE 6. CPU utilization of HPE Primera all-flash array at an indexing rate of 2 TB per day and simultaneous heavy search load

As described previously with the throughput results in FIGURE 4, an appropriately sized HPE Primera array still has plenty of system resources available to accommodate higher daily rates of indexing, fast searches initiated by users or by scheduled searches and reports, and any other workloads your needs may require.

The section Sizing a traditional Splunk implementation details the minimum configuration for an HPE Primera array that would accommodate the indexing and search rates tested for this solution. Hewlett Packard Enterprise recommends sizing the capacity and throughput of your storage to match the capacity and I/O workload required. The choice of available arrays, drive types, drive sizes, and number of controller nodes affords you the flexibility to select the right capacity and level of performance at an affordable cost, while still giving you the option to scale up as the daily indexing rate and number of users running concurrent or scheduled searches increases.

SUMMARY

Deploying Splunk is often accompanied by a number of challenges. Every implementation needs to handle a different number of events coming from a different number and type of sources, making every implementation unique to the demands of the business. But, every organization requires an inexpensive way to scale while maintaining high performance.

Hewlett Packard Enterprise offers you the ability to decouple storage and compute, making it easy to deploy just the amount of storage or compute capacity that is called for, no matter what the size or configuration is of your Splunk environment. This also allows you to seamlessly scale storage and compute capacity independently as your data and number of users increase.

HPE Primera and HPE Nimble Storage all-flash, enterprise-class storage arrays and HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING have all the flexibility, capacity, performance, and upgradeability you need to implement a Splunk solution that is not only the right size at the right price now, but can easily and inexpensively grow as your requirements change. With the added flexibility of HPE GreenLake's pay-as-you-go model based on consumption, Splunk users can get the performance and capacity they want without overpaying for resources they are not using.

The solution highlighted in this white paper is an efficient, dependable, scalable, and economical way to unlock the value of your data and provide operational intelligence for your business. Using Hewlett Packard Enterprise award-winning storage, servers, network infrastructure, and support services in conjunction with Splunk's software and support services will help you get the most value out of your data at the best possible price.

IMPLEMENTING A PROOF OF CONCEPT

As a matter of best practice for all deployments, Hewlett Packard Enterprise recommends implementing a proof of concept using a test environment that matches as closely as possible the planned production environment. In this way, appropriate performance and scalability characterizations can be obtained. For help with a proof of concept, contact a Hewlett Packard Enterprise Services representative (hpe.com/us/en/services/consulting.html) or your Hewlett Packard Enterprise partner. Splunk also provides assistance to customers for implementing Splunk software (Splunk Implementation Services) and for providing guidance to help customers extract value from their data (Splunk Adoption Services). Contact your Splunk representative, or see Splunk Services at https://www.splunk.com/en_us/support-and-services/splunk-services.html for more information.

RESOURCES AND ADDITIONAL LINKS

HPE Primera
hpe.com/us/en/storage/hpe-primera.html

HPE Nimble Storage
hpe.com/us/en/storage/nimble.html

HPE Apollo 4000 systems with HPE Scalable Object Storage with Scality RING
hpe.com/us/en/storage/file-object.html

HPE ProLiant DL servers
hpe.com/us/en/servers/proliant-dl-servers.html

HPE Scalable Object Storage with Scality RING on HPE Apollo 4200 Gen10 Technical white paper
hpe.com/v2/getpdf.aspx/4AA5-9749ENW.pdf

Scality RING shines bright in Gartner Critical Capabilities for Object Storage report
https://community.hpe.com/t5/Around-the-Storage-Block/Scality-RING-shines-bright-in-Gartner-Critical-Capabilities-for/ba-p/7000060

Hewlett Packard Enterprise Technology Consulting Services
hpe.com/us/en/services/consulting.html

Splunk Enterprise Capacity Planning Manual: Reference hardware (minimum hardware requirements for Splunk deployments)
https://docs.splunk.com/Documentation/Splunk/latest/Capacity/Referencehardware

Splunk Enterprise Capacity Planning Manual: Introduction to capacity planning for Splunk Enterprise
https://docs.splunk.com/Documentation/Splunk/latest/Capacity/IntroductiontocapacityplanningforSplunkEnterprise

Splunk Demo: Scality RING Storage for Splunk
https://vimeo.com/425789238

LEARN MORE AT

hpe.com/storage

a00090875ENW, July 2020, Rev 2

Hewlett Packard Enterprise