White paper
Cisco public

Splunk on Cisco Compute Hyperconverged with Nutanix Infrastructure

February 2025

Contents

Executive summary

To extract meaningful insights from the vast amount of machine data generated by IT systems and applications, organizations need robust analytics platforms that can efficiently handle large-scale data ingestion, storage, and real-time analysis to transform the data into actionable insights, improve operation efficiency, and enhance security and decision making. **Splunk** has emerged as a leading platform for operational intelligence, by centeralizing data collection, indexing, and analysis to provide real-time visibility into IT infrastructure and advanced search, monitoring, and reporting capabilities.

The performance and scalability of Splunk depends heavily on the underlying infrastructure. Traditional IT architectures can often struggle to meet the demands of Splunk's data-heavy workloads, operational bottlenecks, and escalating costs as data volumes continue to grow.

Cisco and Nutanix have formed a strategic partnership to deliver a comprehensive hybrid-cloud solution called Cisco Compute Hyperconverged with Nutanix. Cisco Compute Hyperconverged with Nutanix (CCHN) offers a robust platform for running Splunk, combining the strengths of Cisco's hardware with the strengths of Cisco® and Nutanix software products. By leveraging Cisco Compute Hyperconverged with Nutanix for Splunk, organizations can focus on deriving insights from their data rather than managing the underlying infrastructure. This hyperconverged solution integrates servers, storage, and networking operations, reducing complexity in Splunk deployments.

Deploying Splunk Enterprise on Cisco Compute Hyperconverged with Nutanix (CCHN) offers several benefits, combining the powerful data analytics and monitoring capabilities of Splunk with the scalable, efficient, and resilient infrastructure provided by Cisco and Nutanix. CCHN offers extensive automation capabilities, reducing the operational overhead associated with managing infrastructure. This allows IT teams to focus on leveraging Splunk's data-analytics capabilities rather than managing hardware.

This document outlines the design, configuration, and deployment steps for running Splunk Enterprise in a Single Site Distributed Clustered model (C3/C13) on Cisco Compute Hyperconverged utilizing Nutanix in Cisco Intersight® Standalone Mode (ISM). For more details on "Single Site Distributed Clustered deployment model (C3/C13),"

Please read about Splunk Validated Architectures at https://docs.splunk.com/Documentation/SVA/current/Architectures/C3C13.

Technology overview

Splunk Enterprise is a software product that enables you to search, analyze, and visualize the data gathered from the components of your IT infrastructure or business. Splunk Enterprise takes the data from the servers, applications, databases, network devices, and virtual machines that make up your IT infrastructure. As long as the machine that generates the data is a part of your network, Splunk Enterprise can collect the data from anywhere, whether the data is local, remote, or in the cloud.

After you define the data source, Splunk Enterprise indexes the data stream and parses it into a series of individual events that you can view and search. Once the data is collected, the index segments, stores, compresses the data and maintains the supporting metadata to accelerate searching. Search is the primary way users navigate their data in Splunk Enterprise. Splunk Enterprise allows you to save searches and pivots as reports, and then to add these reports to dashboards as dashboard panels.

Splunk Enterprise provides the end-to-end, real-time solution for machine data delivering the following core capabilities:

- Universal collection and indexing of machine data, from virtually any source
- Powerful Search Processing Language (SPL) to search and analyze real-time and historical data
- Real-time monitoring for patterns and thresholds; real-time alerts when specific conditions arise

Organizations typically start with Splunk to solve a specific problem, and then expand from there to address a broad range of use cases, such as application troubleshooting, IT infrastructure monitoring, security, business analytics, and many others. As operational analytics become increasingly critical to day-to-day decision-making and Splunk deployments expand to terabytes of data, a high-performance, highly scalable infrastructure is critical to ensuring rapid and predictable delivery of insights. The Cisco Compute Hyperconverged with Nutanix solution helps you overcome the challenge of deploying Splunk Enterprise on a global scale with an integrated workflow.

CCHN provides a unified platform that integrates compute, storage, networking, and virtualization. It delivers a highly scalable, resilient, and flexible architecture to support various enterprise workloads.

Key features include:

- **Simplified management:** With Cisco Intersight and Nutanix's Prism management interface offers a unified management platform that simplifies the deployment, monitoring, and management of the infrastructure.
- **Faster time to insights:** With the optimized performance of Cisco and Nutanix, Splunk’s data indexing, searching, and reporting can occur faster, enabling your organization to derive insights in real time.
- **Scalability:** Cisco hardware with Nutanix is designed to scale horizontally, meaning you can easily expand your infrastructure by adding additional nodes, providing the required performance as your Splunk deployment grows.
- **Performance optimization:** Nutanix's software-defined storage, coupled with Cisco’s high-performance hardware, ensures that Splunk can run with minimal latency, improving the speed of indexing and search queries. This results in faster insights from the vast amounts of machine data that Splunk collects.
- **High availability and disaster recovery:** Both Splunk and Nutanix provide built-in redundancy, data-protection, and disaster-recovery options, ensuring that your Splunk deployment remains operational even in the event of hardware failure.

The Cisco Compute Hyperconverged with Nutanix family of appliances delivers preconfigured Cisco UCS® servers that are ready to be deployed as nodes to form Nutanix clusters in a variety of configurations. Each server appliance contains three software layers: Cisco UCS server firmware, a hypervisor (Nutanix Acropolis Hypervisor [Nutanix AHV]), and hyperconverged storage software (Nutanix Acropolis Operating System [Nutanix AOS]). Physically, nodes are deployed into clusters, with a cluster consisting of Cisco Compute Hyperconverged all-flash servers. These servers can be interconnected and managed in two different ways as Cisco Intersight Standalone Mode (ISM) and Cisco UCS Managed Mode (UCSM).

Consideration

To successfully deploy Splunk on Cisco Compute Hyperconverged with Nutanix infrastructure, organizations should consider the following best practices:

- **Sizing the infrastructure:** Carefully assess the required resources (CPU, RAM, storage, and networking) based on the expected data volume and workload. Both Cisco and Nutanix offer tools to help determine the appropriate configuration for your Splunk deployment.
- **Storage optimization:** Nutanix's software-defined storage can be tuned to optimize performance for Splunk's indexing and data-retention requirements. Leveraging Nutanix's hybrid or all-flash configurations can further accelerate data access speeds.
- **Cluster design:** Design your Splunk deployment with scalability in mind. CCHN allow for easy scaling of both compute and storage resources, so you can start small and expand as needed.
- **Network optimization:** Ensure that the network configuration between CCHN is optimized for Splunk's high-throughput and low-latency requirements. Consider segmenting traffic for Splunk's search, indexing, and replication processes to maximize performance.
- **High-availability configuration:** Use Nutanix's integrated disaster recovery and backup capabilities to ensure data resilience. Additionally, ensure that Splunk is configured to take advantage of the underlying infrastructure's high-availability features.

Solution overview

The reference Splunk deployment utilized in this white paper used Cisco UCS C-Series nodes for deploying Cisco Compute Hyperconverged with Nutanix infrastructure running in Intersight Standalone Mode. While this document highlights two evolving technologies into one solution, first as Cisco Compute Hyperconverged with Nutanix infrastructure and, secondly, as data analytics with Splunk Enterprise. It is beyond the scope of this document to cover the detailed information on the first part of a CCHN deployment in Intersight Standalone Mode. Please refer to the following Cisco Validated Design (CVD) for more detailed information on the first part of the deployment of CCHN:
https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/CCHC_Nutanix_ISM.html.

Splunk architecture

Splunk software comes packaged as an all-in-one distribution. Depending on your needs, you can deploy Splunk Enterprise as a single instance, or you can create deployments that span multiple instances, ranging from just a few to hundreds or even thousands of instances. Designing a scalable architecture for the Splunk platform requires knowledge of the Splunk instance roles and how they are intended to scale.

A Splunk Enterprise deployment has many dimensions. Below are some of the important factors that can have a significant impact on Splunk Enterprise performance.

- **Amount of incoming data:** The more data you send to Splunk Enterprise, the more time it needs to process the data into events that you can search, report, and generate alerts on.
- **Amount of indexed data:** As the amount of data stored in a Splunk Enterprise index increases, so does the I/O bandwidth needed to store data and provide results for searches.
- **Number of searches:** If you plan to invoke a lot of searches, Splunk Enterprise needs capacity to perform those searches promptly and efficiently. A higher search count over a given period of time requires more resources.

Note: For more details, please refer to the capacity-planning manual at: https://docs.splunk.com/Documentation/Splunk/9.3.2/Capacity/DimensionsofaSplunkEnterprisedeployment

While these factors have an impact on the basic sizing requirements of your Splunk Enterprise deployment, addressing each of them individually does not guarantee peak performance gains for the deployment. For example, if your Splunk Enterprise deployment calls for a low volume of daily indexing but has a high number of concurrent users, it has significantly different resource needs than a setup with a low number of concurrent users and a high indexing volume. Additionally, as both user count and amount of indexed data rise, you must distribute the environment across multiple servers to maintain a similar performance level.

To support larger environments where data originates on many machines, where you need to process large volumes of data, or where many users need to search the data, you can scale the deployment by distributing Splunk Enterprise instances across multiple machines. This is known as a "distributed deployment." In a typical distributed deployment, each Splunk Enterprise instance performs a specialized task and resides on one of three processing tiers corresponding to the main processing functions:

- Search management tier
- Indexer tier
- Data collection tier

You can, for example, create a deployment with many instances that reside on the data input tier and only ingest data, a deployment with several other instances that reside on the indexer tier and index the data, or one instance that resides on the search management tier and manages searches. These specialized instances are known as "components."

Components of Splunk Enterprise software deployment

(1) Search tier

a. **Search Head (SH):** The search head provides the UI for Splunk users and coordinates scheduled search activity. Search heads are dedicated Splunk software instances in distributed deployments. Search heads can be virtualized for easy failure recovery, provided they are deployed with appropriate CPU and memory resources.

b. **Search Head Cluster (SHC):** A search head cluster is a pool of at least three clustered search heads. It provides horizontal scalability for the search head tier and transparent user failover in case of outages. Search head clusters require dedicated machines with, ideally, identical system specifications. The members of search head cluster can also be virtualized for easy failure recovery, provided they are deployed with appropriate CPU and memory resources.

(2) Indexing tier

a. **Indexer:** Indexers are the heart and soul of a Splunk deployment. They process and index incoming data and also serve as search peers to fulfill search requests initiated on the search tier. Indexers must always be on dedicated servers in distributed or clustered deployments. Indexers perform best on bare-metal servers or in dedicated, high-performance virtual machines, if adequate resources can be guaranteed.

(3) Management tier

   a. **Cluster Manager (CM):** The cluster manager is the required coordinator for all activity in a clustered deployment. In clusters with a large number of index buckets (high data volume/retention), the cluster manager will likely require a dedicated server to run on. To achieve cluster-manager high availability, you can deploy two or more cluster managers in an active/standby configuration. You can configure the managers to support either automatic or manual failover.

   b. **Deployment Server (DS):** The deployment server manages configuration of forwarder configuration. It should be deployed on a dedicated instance. It can be virtualized for easy failure recovery.

   c. **License Manager (LM):** The license manager is required by other Splunk software components to enable licensed features and track daily data-ingest volume. The license manager role has minimal capacity and availability requirements and can be colocated with other management functions. It can be virtualized for easy failure recovery.

   d. **Monitoring Console (MC):** The monitoring console provides dashboards for usage and health monitoring of your environment. It also contains a number of prepackaged platform alerts that can be customized to provide notifications for operational issues. In clustered environments, the MC can be collocated with the cluster manager node, in addition to the license manager and deployment server function in nonclustered deployments. It can be virtualized for easy failure recovery.

   e. **Search Head Cluster Deployer (SHC-D):** The search head cluster deployer is needed to bootstrap a SHC and manage Splunk configurations deployed to the cluster. The SHC-D is not a runtime component and has minimal system requirements. It can be colocated with other management roles.

(4) Data collection tier

   a. **Forwarders:** A forwarder consumes data and then forwards the data, usually to an indexer. Forwarders usually require minimal resources, allowing them to reside lightly on the machine generating the data. There are several types of forwarders, but the universal forwarder is the right choice for most purposes. A small-footprint version of a forwarder, it uses a lightweight version of Splunk Enterprise that simply inputs data, performs minimal processing on the data, and then forwards the data to a Splunk indexer or a third-party system.

You can add components to each tier as necessary to support greater demands on that tier. For example, if you have a large number of users, you can add extra search heads to better service the users. In this solution, we have deployed Splunk Enterprise on Single Site Distributed Clustered with SHC - Single Site (C3 / C13) mode on Nutanix Cluster. The following diagram represents a single site distributed clustered deployment with a Search Head Cluster (SHC) topology:

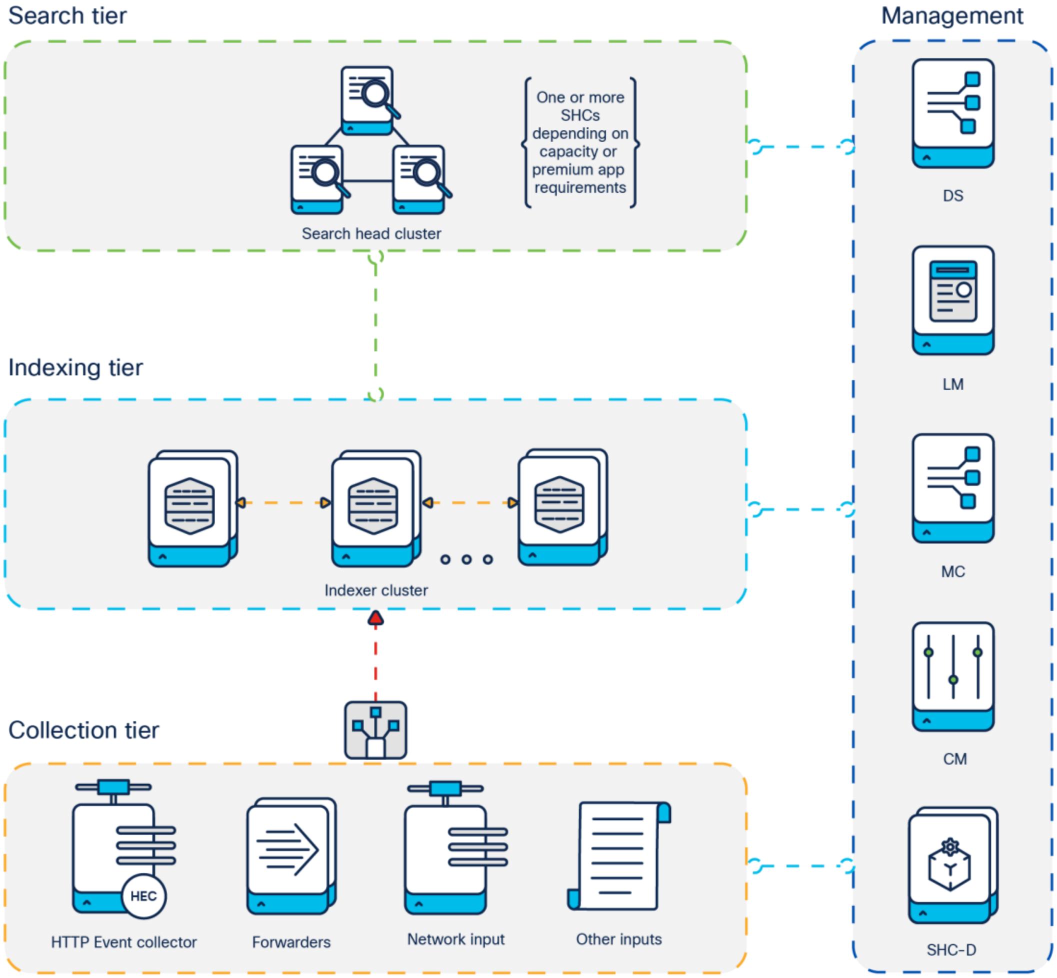


Figure 1.
Distributed Clustered Deployment with SHC - Single Site (C3 / C13)

The reference hardware specification is a baseline for scoping and scaling the Splunk platform for your use. The recommendations are based on Splunk Validated Architectures.

Note: For more details, please refer https://docs.splunk.com/Documentation/SVA/current/Architectures/C3C13

Deployment architecture

The deployment architecture for Splunk running on Cisco Compute Hyperconverged with Nutanix in Intersight Standalone Mode (ISM) is shown below.

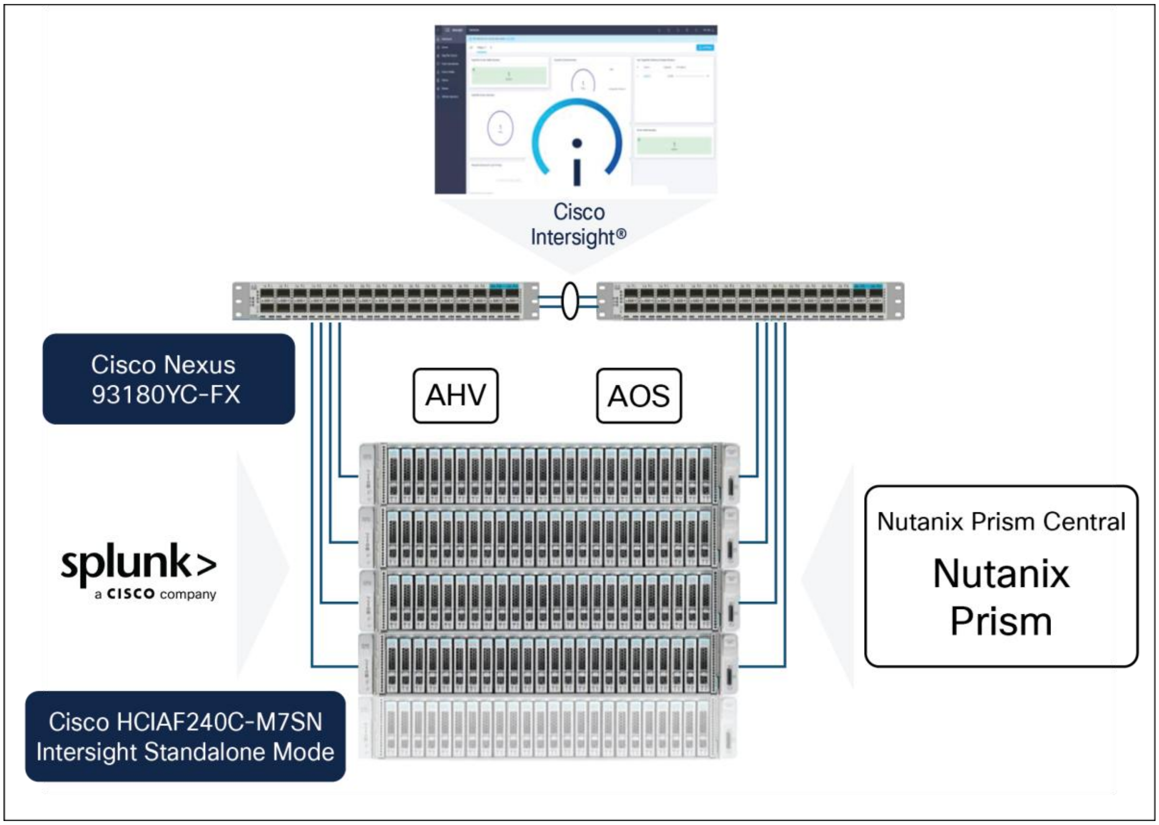


Figure 2.
High level deployment architecture

The entire day-0 deployment is managed through Cisco Intersight and Nutanix Foundation Central, enabled through Prism Central. As shown above, for this solution we used 4x Cisco Compute Hyperconverged HCIAF240C M7 All-NVMe/All-Flash Servers. For more details, go to: [HCIAF240C M7 All-NVMe/All-Flash Server specification sheet](https://www.cisco.com/c/dam/en/us/products/collateral/hyperconverged-infrastructure/hc-240m7-specsheet.pdf).

Each Cisco UCS C240 M7 Rack Server is configured with the following:

- 2x Intel® Xeon® Gold I6448H
- 384 GB DDR5 memory
- 2x 240GB M.2 card managed through M.2 RAID controller
- 24x 1.9 TB NVMe
- 1x Cisco VIC 15425 4x 10/25/50G PCIe C-Series w/Secure Boot

You have several options to configure CPU, memory, network cards, GPUs, and storage as detailed in this spec sheet: **https://www.cisco.com/c/dam/en/us/products/collateral/hyperconverged-infrastructure/hc-240m7-specsheet.pdf**.

Software components

Table 1 lists the software components and the versions validated for the Splunk Enterprise on Cisco Compute Hyperconverged with Nutanix in Intersight Standalone software components.

Table 1. Software components

Component	Version
Splunk Enterprise	9.3.1
Red Hat Enterprise Linux	8.9
Foundation Central	1.6
Prism Central deployed on ESXi cluster	pc.2022.6.0.10
AOS and AHV bundled	nutanix_installer_package-release-fraser-6.5.5.6
Cisco UCS C240 M7 Rack Server	4.3 (3.240043)

Solution deployment

Please refer to the following Cisco Validated Design (CVD) guide to deploy Cisco Compute Hyperconverged with Nutanix in Intersight Standalone Mode, allowing nodes to be connected to a pair of Top-of-Rack (ToR) switches and servers centrally managed using Cisco Intersight: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/CCHC_Nutanix_ISM.html.

Splunk Enterprise can be deployed in physical, virtual, cloud, or hybrid environments. When deploying Splunk Enterprise in a virtualized environment, please consider following the recommendations for the different components that are part of the Splunk deployment.

Reference host specifications for performance on distributed deployments

Distributed deployments are designed to separate the index and search functionality into dedicated tiers that can be sized and scaled independently without disrupting the other tier. Planning system resources and bandwidth to enable search and index performance in a distributed virtual environment depends on the total volume of data being indexed and the number of active concurrent searches (scheduled or other) at any time. The daily data-ingest volume and the concurrent search volume are the two most important factors used when estimating the hardware capabilities and node counts for each tier. The search and indexing roles prioritize different compute resources. The indexing tier uses high-performance storage to store and retrieve data efficiently. The search tier uses CPU cores and RAM to handle ad-hoc and scheduled search workloads.

An increase in search tier capacity corresponds to increased search load on the indexing tier, requiring scaling of the indexer nodes. Scaling either tier can be done vertically by increasing per-instance hardware resources, or horizontally by increasing the total node count.

There are several performance factors to consider when deploying Splunk software inside hypervisor virtual machines. These considerations are CPU, memory, and disk/storage resources:

- **CPU:** Splunk search heads and indexers are CPU-intensive. Sharing CPUs with other virtual machines running on the same host can result in high wait times, which might negatively impact Splunk performance. Splunk indexer and search head virtual machines should have 100 percent of vCPUs reserved to ensure good performance.
- **Memory:** Memory is critical for Splunk search heads and indexers. Splunk virtual machines should have reserved memory available to them. VMware hosts running Splunk Enterprise should not be configured with memory overcommit, because overcommitting memory will likely result in poor performance due to ballooning and/or page swapping to the hard disk.
- **Disk/Storage:** Splunk indexers are usually CPU- and disk I/O-intensive, so a disk exposed to these indexers within virtual machines should be capable of 1200+ random seeks per second. In virtual environments, high-performance, low-latency storage is required for the Splunk distributed indexing and searches. Cisco UCS all-flash server systems provide an excellent choice of hardware and storage for the high-performing virtual infrastructure required for Splunk deployments.

Key considerations for this deployment

This reference solution was designed on four physical Cisco UCS C240 M7 All NVMe Rack Server nodes in a Nutanix cluster, and the following are a few of the key considerations for optimal Splunk performance:

- Up to 1 TB daily indexing
- Indexing less than 150 GB of data per day per indexer
- Data retention was set for 30 days of hot/warm storage and 30 days of cold storage; however, no archived/frozen storage is configured.
- Total hot/warm storage requirement is 16 TB.
- Total cold storage requirement is 16 TB.
- Split the Splunk search and indexing functions – indexers for parallelized indexing and search heads to distribute parallelized searches
- Maintain full reservations on CPU and memory settings of the indexer and search head virtual machines
- No snapshotting for virtual machines running Splunk Enterprise

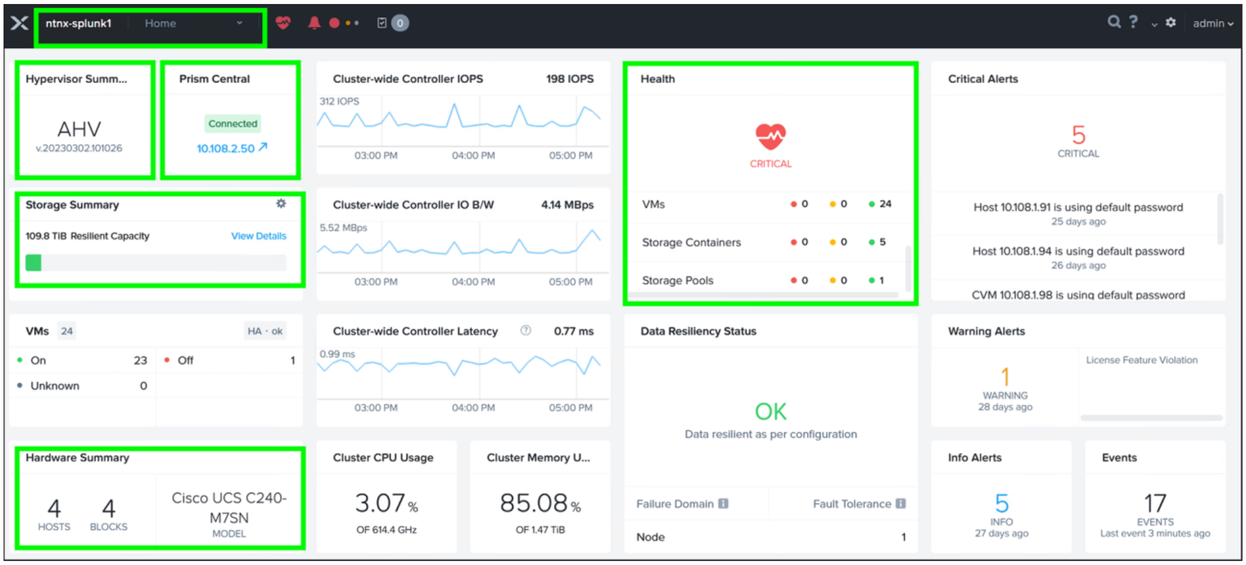
The guidelines provided in this section are targeted for Splunk Enterprise for IT Operations Analytics (ITOA) use cases. When leveraging premium solutions such as Splunk Enterprise Security (ES) and Splunk IT Services Intelligence (ITSI), it may be necessary to increase the allocation of virtual CPU (vCPU) and memory resources. Also, mixing Splunk workloads with other applications in the same Hyperconverged-Infrastructure (HCI) cluster can impact of the performance and needs to be tested by the customers.

Note: The suggested maximum indexing capacities per indexer node are up to 300 GB per day for IT operational analytics, up to 200 GB per day for IT Services Intelligence (ITSI), and up to 100 GB per day for enterprise security. You can scale by adding search heads and indexers to the cluster as needed.

Note: Cluster manager node includes all the administration roles of Splunk Cluster, such as cluster manager, license manager, and monitoring console.

Preparing for Splunk deployment

This section provides the guidelines for deploying the solution. Log in to the Nutanix Prism Element to view the existing cluster. The screenshot in Figure 3 shows a summary of the Nutanix cluster “ntnx-splunk1” that we created and used for this Splunk deployment.



ntnx-splunk1 Home

admin

Hypervisor Summ... AHV v.20230302.101026

Prism Central Connected 10.108.2.50

Cluster-wide Controller IOPS 198 IOPS

Health CRITICAL

VMs	● 0	● 0	● 24
Storage Containers	● 0	● 0	● 5
Storage Pools	● 0	● 0	● 1

Critical Alerts 5 CRITICAL

Host 10.108.1.91 is using default password 25 days ago

Host 10.108.1.94 is using default password 26 days ago

CVM 10.108.1.98 is using default password

Storage Summary

109.8 TiB Resilient Capacity View Details

Cluster-wide Controller IO B/W 4.14 MBps

VMs 24 HA · ok

● On	23	● Off	1
● Unknown	0		

Cluster-wide Controller Latency 0.77 ms

Data Resiliency Status OK Data resilient as per configuration

Failure Domain Node

Fault Tolerance 1

Warning Alerts 1 WARNING 28 days ago

License Feature Violation

Hardware Summary 4 HOSTS 4 BLOCKS Cisco UCS C240-M7SN MODEL

Cluster CPU Usage 3.07 % OF 614.4 GHz

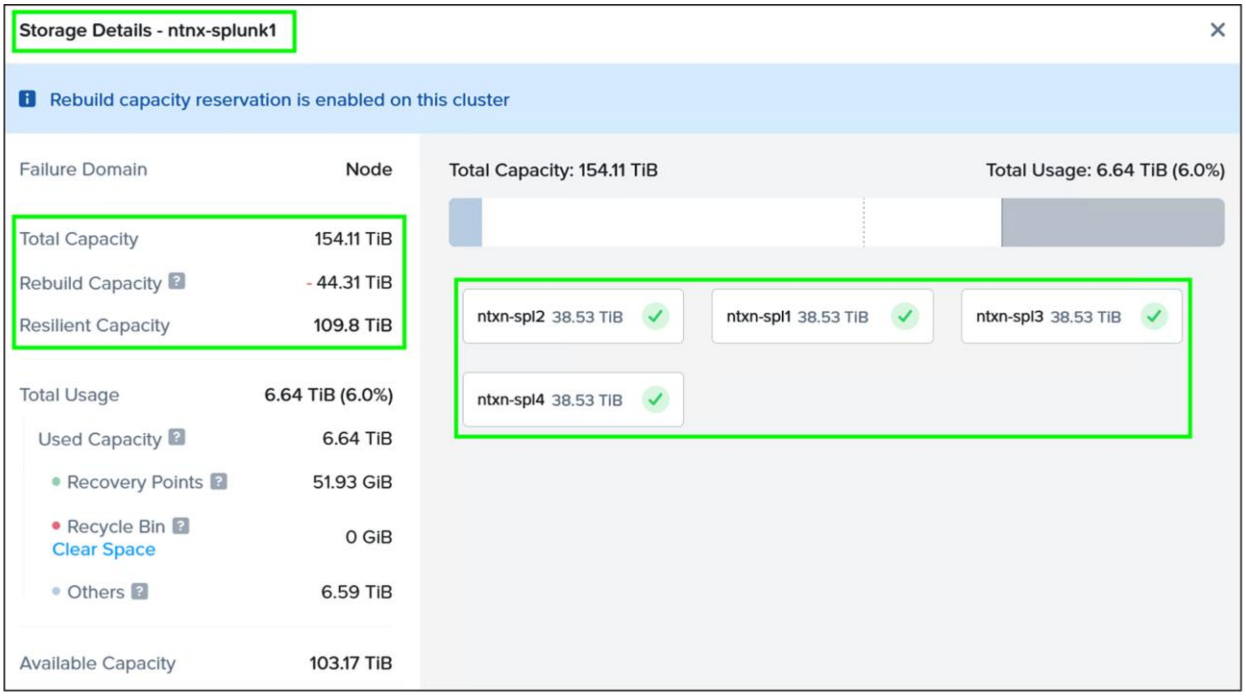
Cluster Memory U... 85.08 % OF 1.47 TiB

Info Alerts 5 INFO 27 days ago

Events 17 EVENTS Last event 3 minutes ago

Figure 3.
Nutanix Prism Element dashboard

The screenshot in Figure 4 shows the storage details of the Nutanix cluster “ntnx-splunk1” that we used for this solution. As mentioned earlier, we used 4 x Cisco Compute Hyperconverged HCIAF240C M7 All-NVMe/All-Flash Servers, with each server having 24 x 1.9 TB NVMe Front drives.



Storage Details - ntnx-splunk1

Rebuild capacity reservation is enabled on this cluster

Failure Domain	Node
Total Capacity	154.11 TiB
Rebuild Capacity	- 44.31 TiB
Resilient Capacity	109.8 TiB
Total Usage	6.64 TiB (6.0%)
Used Capacity	6.64 TiB
● Recovery Points	51.93 GiB
● Recycle Bin Clear Space	0 GiB
● Others	6.59 TiB
Available Capacity	103.17 TiB

Total Capacity: 154.11 TiB

Total Usage: 6.64 TiB (6.0%)

ntxn-spl2 38.53 TiB ✓

ntxn-spl1 38.53 TiB ✓

ntxn-spl3 38.53 TiB ✓

ntxn-spl4 38.53 TiB ✓

Figure 4.
Nutanix storage cluster details

Create and configure Splunk Enterprise virtual machines

For this solution, we configured Splunk Enterprise running on virtual machines with the Red Hat Enterprise Linux 8.9 operating system. Table 2 gives the details of all the virtual machines configured for deploying this solution.

Table 2. Splunk virtual machines

Splunk component	Number of virtual machines	Number of vCPUs	Memory	OS	Disk 1 OS	Disk 2 Hot/warm data	Disk 3 Cold data
Indexer	8	16	64 GB	Red Hat Enterprise Linux 8.9	400 GB	2000 GB	2000 GB
Search heads	3	16	64 GB	Red Hat Enterprise Linux 8.9	400 GB	N/A	N/A
Cluster manager, license manager, and monitoring console	1	4	16 GB	Red Hat Enterprise Linux 8.9	200 GB	N/A	N/A
Deployer	1	4	16 GB	Red Hat Enterprise Linux 8.9	200 GB	N/A	N/A

Table 3 highlights each virtual machine with its host name and role in the Splunk tiers configured in this deployment.

Table 3. Virtual machine and its role

Splunk tier	Virtual machine name	Role
Admin tier	admin1	admin1 – Cluster manager node, license manager node, monitoring console node
	admin2	admin2 – Deployer node
Indexing tier	idx1	idx1 – Indexer node 1
	idx2	idx2 – Indexer node 2
	idx3	idx3 – Indexer node 3
	idx4	idx4 – Indexer node 4
	idx5	idx5 – Indexer node 5
	idx6	idx6 – Indexer node 6
	idx7	idx7 – Indexer node 7
	idx8	idx8 – Indexer node 8

Splunk tier	Virtual machine name	Role
Search tier	sh1	sh1 – Search head node 1
	sh2	sh2 – Search head node 2
	sh3	sh3 – Search head node 3

Post-OS configuration steps before installing Splunk

After installing RHEL OS in each of the virtual machines, you will need to perform some additional steps as part of the post-OS setup before installing Splunk. Most, but not all of the steps required to install the components, are explained in this document. To implement these steps, you need to identify one of the management servers to act as the cluster manager node and then use Ansible or your preferred automation tool to push these settings in the entire setup. The steps are as follows:

1. Edit the "/etc/hosts" file according to nodes part of the cluster deployment
2. Setup Password-less Login to run Ansible script from the admin node
3. Disable SELinux
4. Disable Linux Firewall
5. Install httpd
6. Configure NTP or Chrony services
7. Enable syslog
8. Configure ulimit settings
9. Set the number of TCP retries
10. Configure Virtual Machine Swapping and overcommit
11. Disable IPv6 Defaults
12. Disable Transparent Huge Pages
13. Configure hot/warm and cold storage on each indexer

We configured each indexer with hot/warm and cold data storage for retention periods of 30 days. For this, we configured and added two additional disks to each indexer virtual machine. Then we created a partition using a parted utility, formatted the volume using mkfs, and, after the volumes were formatted, we mounted these two disks as "/data/disk1" for hot/warm storage and "/data/disk2" for cold storage. We also updated the entry in the "/etc/fstab" file.

Installing Splunk Enterprise

In this deployment, three clustered search heads, eight clustered indexers, a license manager, a cluster manager, a monitoring console, and a deployer are configured.

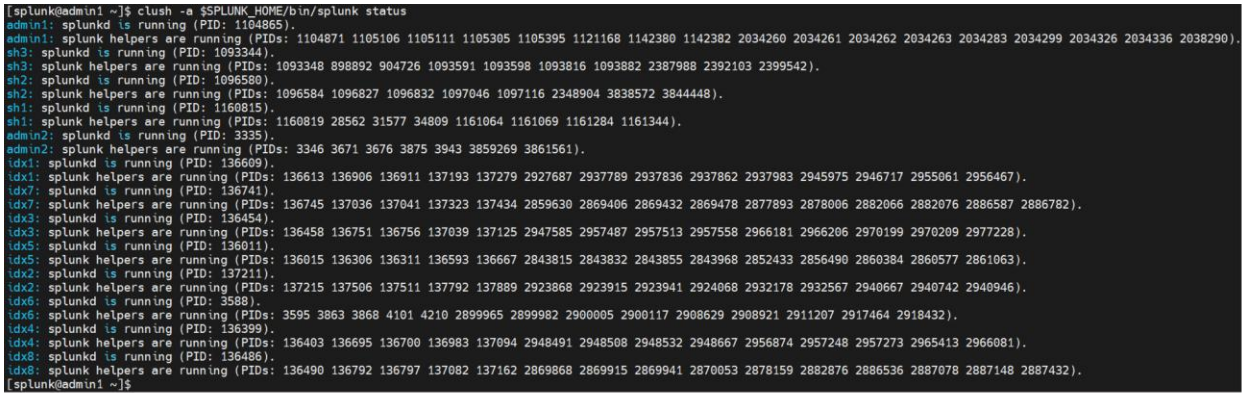
The installation sequence is as follows:

1. Install Splunk Enterprise
2. Configure the license manager node
3. Configure the cluster manager node
4. Configure the indexing cluster
5. Configure the deployer
6. Configure the search head cluster
7. Configure the distributed management console

Install Splunk Enterprise

Splunk Enterprise is a single software package that can be configured to function in a specific role. Installation of Splunk across all nodes is the same, with no specific parameters required; configuration changes then are required for each component. Therefore, a simple installation across every server is used here as the base to build this architecture. See the Splunk documentation for detailed installation steps: https://docs.splunk.com/Documentation/Splunk/9.3.2/Installation/InstallonLinux

Download the Splunk Enterprise software from the Splunk.com website. Copy it to the server admin1 and then copy the Splunk software to all the nodes (two admin servers, three search heads, and eight indexers). We have configured and installed Splunk Enterprise in the "/data/disk1" directory of the indexers, search heads. and admin servers as shown in Figure 5:



```
[splunk@admin1 ~]$ clush -a $SPLUNK_HOME/bin/splunk status
admin1: splunkd is running (PID: 1104865).
admin1: splunk helpers are running (PIDs: 1104871 1105106 1105111 1105305 1105395 1121168 1142380 1142382 2034260 2034261 2034262 2034263 2034283 2034299 2034326 2034336 2038290).
sh3: splunkd is running (PID: 1093344).
sh3: splunk helpers are running (PIDs: 1093348 898892 904726 1093591 1093598 1093816 1093882 2387988 2392103 2399542).
sh2: splunkd is running (PID: 1096580).
sh2: splunk helpers are running (PIDs: 1096584 1096827 1096832 1097046 1097116 2348904 3838572 3844448).
sh1: splunkd is running (PID: 1160815).
sh1: splunk helpers are running (PIDs: 1160819 28562 31577 34809 1161064 1161069 1161284 1161344).
admin2: splunkd is running (PID: 3335).
admin2: splunk helpers are running (PIDs: 3346 3671 3676 3875 3943 3859269 3861561).
idx1: splunkd is running (PID: 136609).
idx1: splunk helpers are running (PIDs: 136613 136906 136911 137193 137279 2927687 2937789 2937836 2937862 2937983 2945975 2946717 2955061 2956467).
idx7: splunkd is running (PID: 136741).
idx7: splunk helpers are running (PIDs: 136745 137036 137041 137323 137434 2859630 2869406 2869432 2869478 2877893 2878006 2882066 2882076 2886587 2886782).
idx3: splunkd is running (PID: 136454).
idx3: splunk helpers are running (PIDs: 136458 136751 136756 137039 137125 2947585 2957487 2957513 2957558 2966181 2966206 2970199 2970209 2977228).
idx5: splunkd is running (PID: 136011).
idx5: splunk helpers are running (PIDs: 136015 136306 136311 136593 136667 2843815 2843832 2843855 2843968 2852433 2856490 2860384 2860577 2861063).
idx2: splunkd is running (PID: 137211).
idx2: splunk helpers are running (PIDs: 137215 137506 137511 137792 137889 2923868 2923915 2923941 2924068 2932178 2932567 2940667 2940742 2940946).
idx6: splunkd is running (PID: 3588).
idx6: splunk helpers are running (PIDs: 3595 3863 3868 4101 4210 2899965 2899982 2900005 2900117 2908629 2908921 2911207 2917464 2918432).
idx4: splunkd is running (PID: 136399).
idx4: splunk helpers are running (PIDs: 136403 136695 136700 136983 137094 2948491 2948508 2948532 2948667 2956874 2957248 2957273 2965413 2966081).
idx8: splunkd is running (PID: 136486).
idx8: splunk helpers are running (PIDs: 136490 136792 136797 137082 137162 2869868 2869915 2869941 2870053 2878159 2882876 2886536 2887078 2887148 2887432).
[splunk@admin1 ~]$
```

Figure 5.
Splunk status summary

Start the Splunk Enterprise cluster

1. Log in to the admin node as user “splunk” and start the Splunk Enterprise services.
   a. `clush -a $SPLUNK_HOME/bin/splunk start --accept-license --no-prompt`
2. Verify the status of the Splunk Enterprise services:
   a. `clush -a $SPLUNK_HOME/bin/splunk status`
3. Log out of the shell user “splunk” and log back in as the root user to configure Splunk Enterprise to start automatically when the system is rebooted:
   a. `clush -a $SPLUNK_HOME/bin/splunk enable boot-start -user splunk`
4. Because Splunk Enterprise was started with no prompts, you need to create the admin user with a seed password to start the system. Be sure to change this password to something more secure when you log in to the web user interface for the first time. From the Command-Line Interface (CLI), using SSH, log in to the admin server admin1 as user “splunk.”
5. Stop Splunk Enterprise on all servers:
   a. `clush -a $SPLUNK_HOME/bin/splunk stop`
6. Start Splunk Enterprise on all servers and verify the status to make sure that the Splunk Enterprise services are running:

```
clush -a "$SPLUNK_HOME/bin/splunk start"
clush -a "$SPLUNK_HOME/bin/splunk status"
```

Log in as the Splunk “admin” the first time

Log in to the Splunk web user interface using your chosen password. The GUI will prompt you to change the admin password. This step needs to be performed on all Splunk instances separately by logging in to the GUI of each Splunk instance.

Launch the admin1 Splunk instance’s web user interface using the IP address and the default port 8000, as in this example: **http://admin1:8000**.

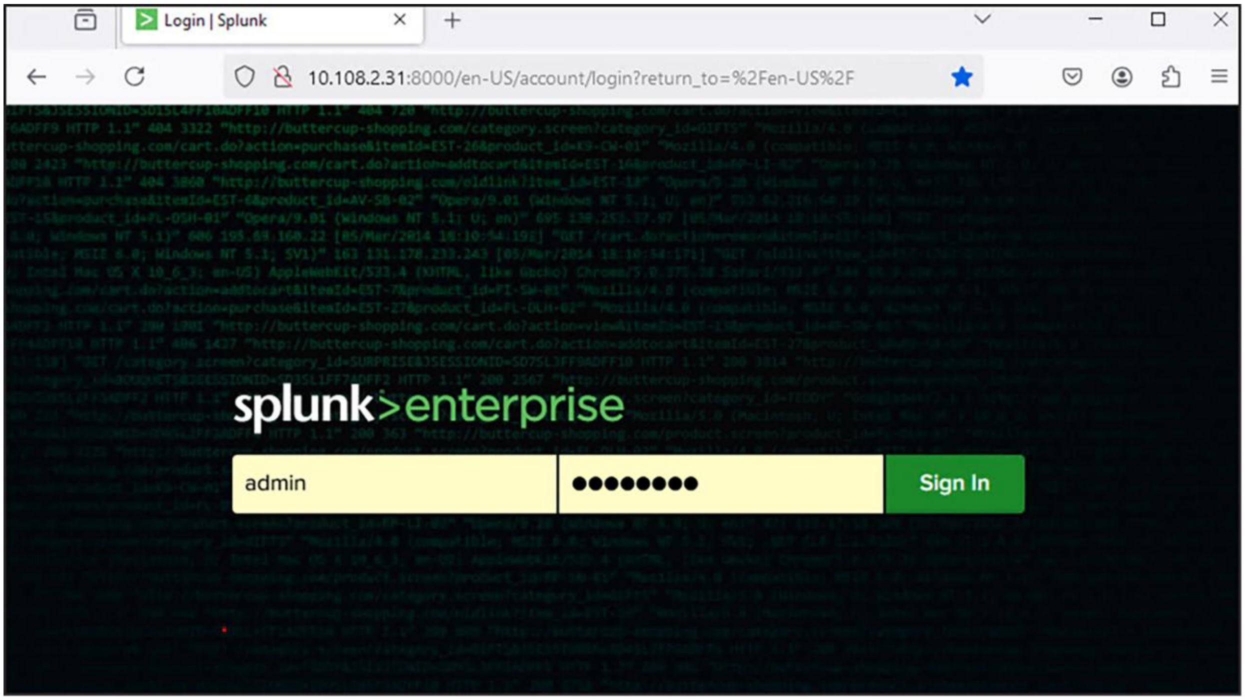


Figure 6.
Splunk login interface

Note: Splunk software uses port number 8000 as the default web interface and port number 8089 as the default management port. Also, reboot the virtual machines and verify that the Splunk software starts upon boot.

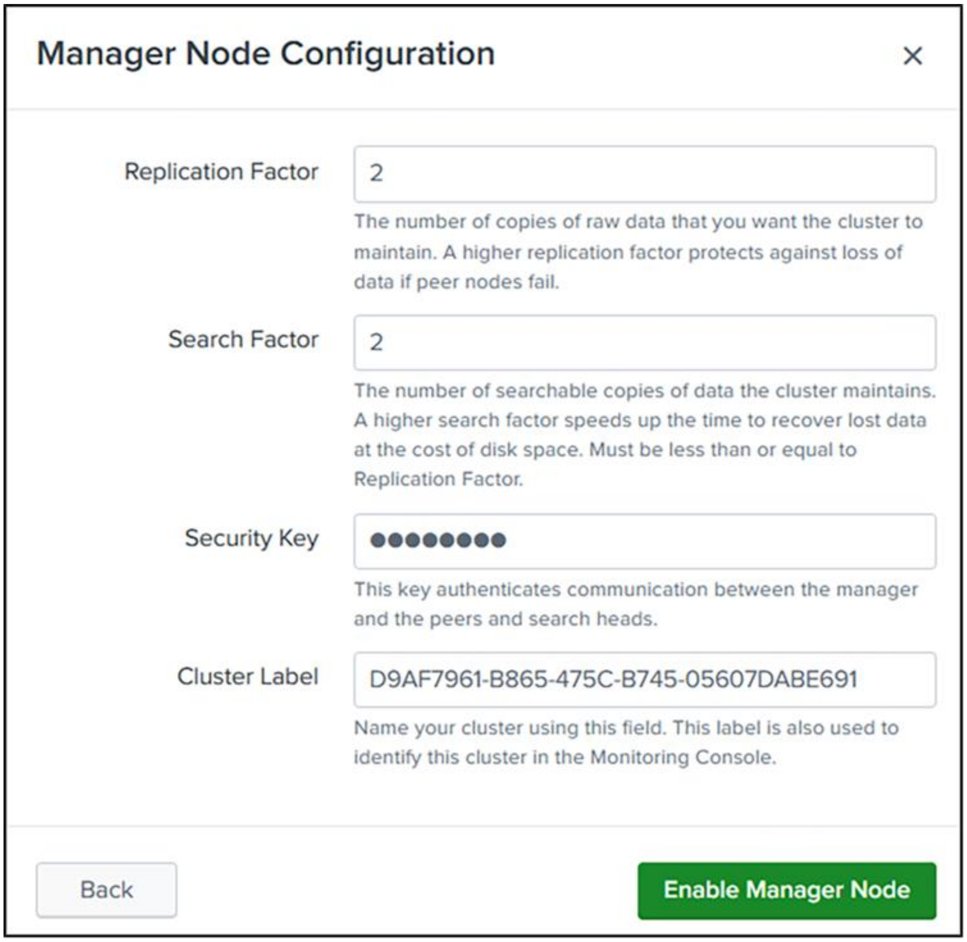
Configure Splunk Enterprise and its components

After you finish installing the Splunk Enterprise software on all of the virtual machines (search heads, indexers, and admin node), please configure the steps below for setting up a Splunk Enterprise cluster.

1. Configure the Splunk Enterprise cluster
2. Install the Splunk Enterprise licenses
    a. Please refer to Splunk website for more information on licensing: https://www.splunk.com/en_us/products/pricing/ingest-pricing.html
        i. https://www.splunk.com/en_us/legal/licensed-capacity.html
    b. The servers in the Splunk Enterprise infrastructure that perform indexing must be licensed. Any Splunk instance can be configured to perform the role of license manager. In this deployment, the admin server (admin1) is configured as the license manager, and all of the other Splunk instances are configured and associated with a remote manager license server.

3. Setup the license manager
    a. Configure the server admin1 and designated this Splunk instance, as the manager license server
    b. Log in to the server admin1 as user admin. Navigate to the licensing screen by choosing Settings > Licensing > Select Change License Group. Select the Enterprise License radio button and then Save. In the Add New License dialog box, click Choose File to select your license file. Install the licenses and then Click Restart Now.
    c. After applying licenses to this node, click on Licensing > Change to peer > and then select "Designate this Splunk instance, admin1, as the manager license server,". This will configure node admin1 as license manager for this Splunk cluster.
4. Configure the indexers, search heads, and admin server as license peers
    a. Configure all of the other Splunk instances from a standalone license server to designate another Splunk instance as the license manager server. Configure each node individually by accessing the respective web user interfaces, and then "Designate a different Splunk instance as the manager license server."
5. Configure Splunk indexer cluster, and provision the indexer cluster manager on admin server node.
    a. The Splunk Enterprise indexers in an indexer cluster are configured to replicate each other's data, so the system keeps multiple copies of all data. This process is known as index replication. The number of copies is controlled by a parameter known as the replication factor. By maintaining multiple, identical copies of Splunk Enterprise data, clusters prevent data loss while promoting data availability for searching.
    b. For a Splunk deployment on Nutanix, the replication factor is tied to the storage and to how Splunk handles both its data indexing and its storage. Splunk cluster replication factor of 2 (RF2) may be acceptable if you're using an indexer cluster with a sufficient number of nodes to handle failures. Also, the combination of Splunk indexer clustering and Nutanix replication factor of 2 (or higher) will ensure that, even if a node fails in Nutanix, the data remains available across multiple nodes. This is critical to avoid data loss and maintain performance during failures.
    c. For a Splunk deployment on Nutanix, a splunk cluster replication factor of 3 (RF3) is typically recommended for production environments to ensure data resiliency and availability, especially in larger or mission-critical environments. For smaller or less critical setups, RF2 may be sufficient, but RF3 provides better fault tolerance. Please refer to both Nutanix and Splunk documentation for the most up-to-date recommendations based on your specific environment and use cases.

d. For this deployment, we configured indexer cluster with a replication factor of 2 on Splunk as shown below:



Manager Node Configuration

Replication Factor: 2

The number of copies of raw data that you want the cluster to maintain. A higher replication factor protects against loss of data if peer nodes fail.

Search Factor: 2

The number of searchable copies of data the cluster maintains. A higher search factor speeds up the time to recover lost data at the cost of disk space. Must be less than or equal to Replication Factor.

Security Key: ●●●●●●●●

This key authenticates communication between the manager and the peers and search heads.

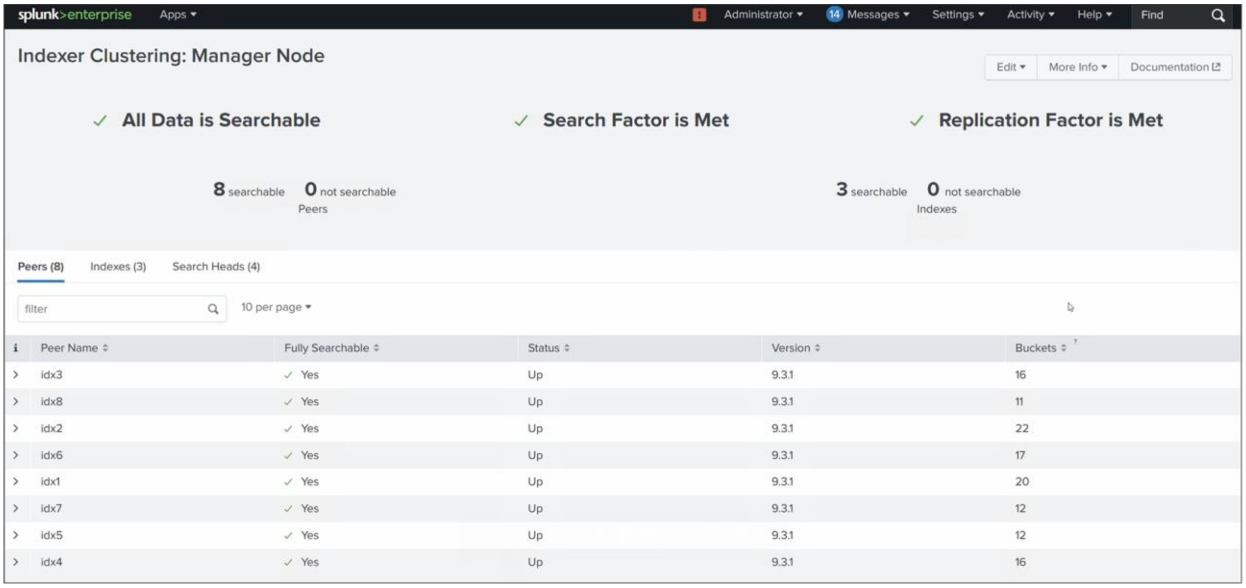
Cluster Label: D9AF7961-B865-475C-B745-05607DABE691

Name your cluster using this field. This label is also used to identify this cluster in the Monitoring Console.

Figure 7.
Replication factor configuration for Splunk cluster

e. **Note:** Replication and search factors vary by deployment. The replication factor indicates the number of copies to be maintained on the indexers. The search factor indicates how many of those copies will return search results. In the configuration described here, one indexer could be down, and searches will still return all results. If the configuration needs to be more resilient, you can increase the replication factor, but this change will also increase disk consumption. Check the Splunk documentation for more information: https://docs.splunk.com/Documentation/Splunk/9.3.2/Indexer/Thereplicationfactor

6. Configure all of the indexing peers and verify the cluster



Indexer Clustering: Manager Node

✓ All Data is Searchable

✓ Search Factor is Met

✓ Replication Factor is Met

8 searchable 0 not searchable
Peers

3 searchable 0 not searchable
Indexes

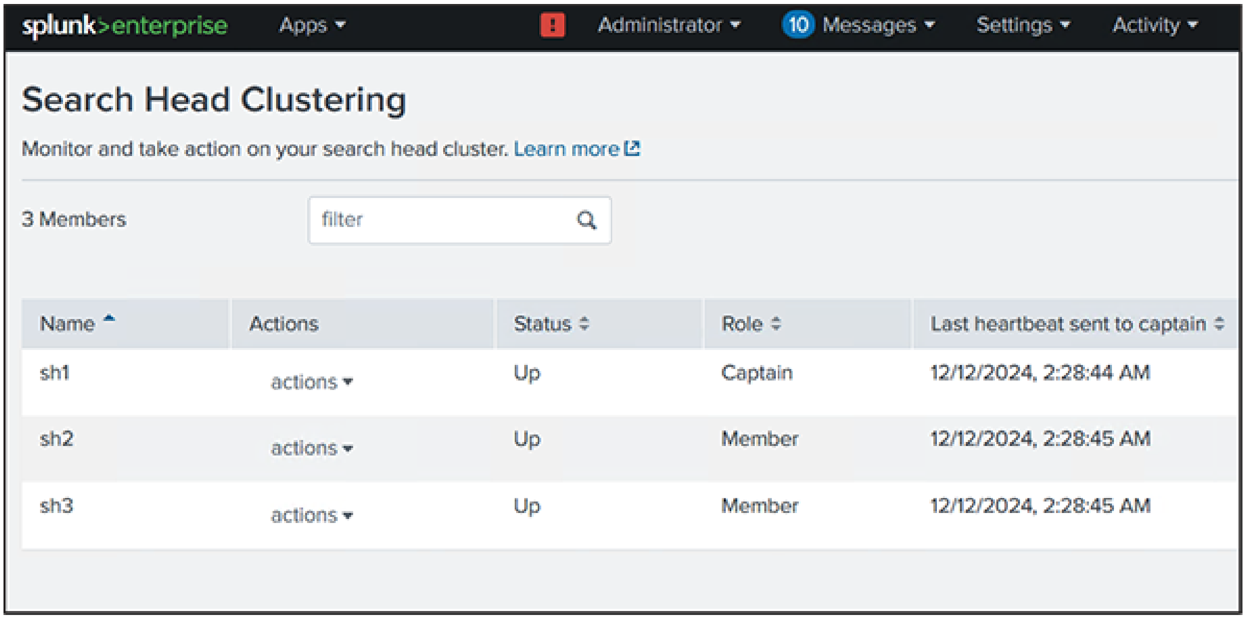
Peers (8) Indexes (3) Search Heads (4)

filter 10 per page

i	Peer Name	Fully Searchable	Status	Version	Buckets
>	idx3	✓ Yes	Up	9.3.1	16
>	idx8	✓ Yes	Up	9.3.1	11
>	idx2	✓ Yes	Up	9.3.1	22
>	idx6	✓ Yes	Up	9.3.1	17
>	idx1	✓ Yes	Up	9.3.1	20
>	idx7	✓ Yes	Up	9.3.1	12
>	idx5	✓ Yes	Up	9.3.1	12
>	idx4	✓ Yes	Up	9.3.1	16

Figure 8.
Splunk indexer cluster summary

7. Configure receiving on the peer nodes
    a. For the indexers (or peer nodes) to receive data from the forwarders, the inputs.conf file of all the indexers needs to be configured with a line to enable TCP port 9997. You accomplish this by editing a special-purpose inputs.conf file in the cluster manager.
8. Configure the cluster manager node to forward all its data to the indexer layer
    a. As a best practice, you should forward all manager node internal data to the indexer (peer node) layer. This step has several advantages. It enables diagnostics for the manager node if it goes down. The data leading up to the failure is accumulated on the indexers, where a search head can later access it.
9. Configure search head clustering, search head cluster members and elect search head captain
    a. A search head cluster consists of a group of search heads that share configurations, job scheduling, and search artifacts. The search heads are known as the cluster members. For this solution, we configured three search head hosts (sh1, sh2, and sh3) and added to the search head cluster.
    b. One cluster member has the role of captain, which means that it coordinates job scheduling and replication activities among all the members.



Search Head Clustering

Monitor and take action on your search head cluster. Learn more

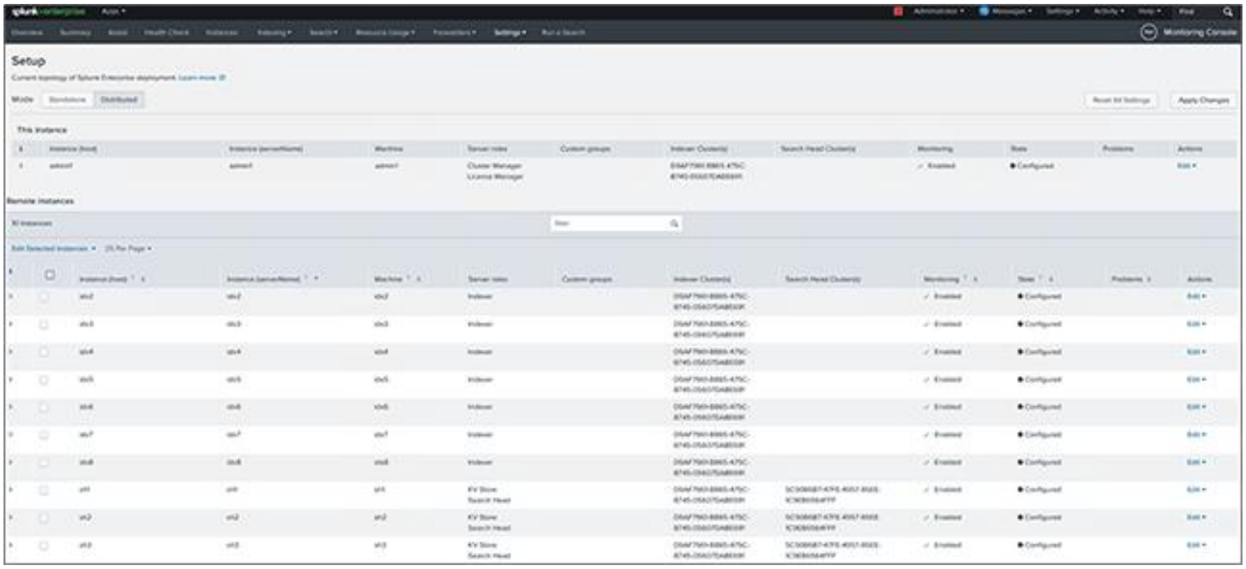
3 Members

Name	Actions	Status	Role	Last heartbeat sent to captain
sh1	actions	Up	Captain	12/12/2024, 2:28:44 AM
sh2	actions	Up	Member	12/12/2024, 2:28:45 AM
sh3	actions	Up	Member	12/12/2024, 2:28:45 AM

Figure 9.
Splunk search head clustering

10. Configure search head load balancing
11. Configure the deployer node
    a. Any Splunk Enterprise instance can be configured to act as the deployer. In this solution, admin2 is selected to serve this function as well.
        i. **Note:** Do not locate deployer functions on a search head cluster member. The deployer instance must be separate from any cluster member, because it is used to manage the configurations of the cluster members
12. Configure search heads to forward their data to the indexer layer
    a. As a best practice, you should forward all search head internal data to the search peer (indexer) layer. This practice has several advantages. It enables diagnostics for the search head if it goes down. The data leading up to the failure is accumulated on the indexers, where another search head can later access it. In addition, by forwarding the results of summary index searches to the indexer level, all search heads have access to them. Otherwise, the results are available only to the search head that generates them.
13. Integrate the search head cluster with the indexer cluster
    a. For the search heads of the search head cluster to be able to search across all the indexer clusters, the search head cluster must be integrated with the indexer cluster.

14. Configure the Splunk monitoring console
    a. We used the admin1 server node to function as the distributed monitoring console for this deployment.



Setup

Mode: Standalone | Distributed

This instance

#	Instance (host)	Instance (serverName)	Machine	Server roles	Custom groups	Indexer Cluster(s)	Search Head Cluster(s)	Monitoring	State	Problems	Actions
1	admin1	admin1	admin1	Cluster Manager License Manager		[illegible]		✓ Enabled	● Configured		Edit ▾

Remote instances

10 instances

Instance (host)	Instance (serverName)	Machine	Server roles	Custom groups	Indexer Cluster(s)	Search Head Cluster(s)	Monitoring	State	Problems	Actions
idx2	idx2	idx2	Indexer		[illegible]		✓ Enabled	● Configured		Edit ▾
idx3	idx3	idx3	Indexer		[illegible]		✓ Enabled	● Configured		Edit ▾
idx4	idx4	idx4	Indexer		[illegible]		✓ Enabled	● Configured		Edit ▾
idx5	idx5	idx5	Indexer		[illegible]		✓ Enabled	● Configured		Edit ▾
idx6	idx6	idx6	Indexer		[illegible]		✓ Enabled	● Configured		Edit ▾
idx7	idx7	idx7	Indexer		[illegible]		✓ Enabled	● Configured		Edit ▾
idx8	idx8	idx8	Indexer		[illegible]		✓ Enabled	● Configured		Edit ▾
sh1	sh1	sh1	KV Store Search Head		[illegible]	[illegible]	✓ Enabled	● Configured		Edit ▾
sh2	sh2	sh2	KV Store Search Head		[illegible]	[illegible]	✓ Enabled	● Configured		Edit ▾
sh3	sh3	sh3	KV Store Search Head		[illegible]	[illegible]	✓ Enabled	● Configured		Edit ▾

Figure 10.
Monitoring console setup dashboard for distributed deployment

15. Configure the deployment server

16. Configure and install a universal forwarder
    a. We did not configure forwarder for this solution.

Summary

Machine data generated from various sources, such as IT systems, applications, and IoT devices, contains critical insights leading to organizational success and efficiency, but mining that data can be complicated without the right data analytics platform. Splunk Enterprise software delivers best-in-class operational visibility and digital intelligence by monitoring all machine-generated data and making it accessible, usable, and valuable across the organization. Splunk deployments typically start small and expand rapidly to address additional use cases.

Cisco Compute Hyperconverged with Nutanix (CCHN) provides optimized hyperconverged infrastructure for any workload at any scale. It simplifies and accelerates the deployment process, hosts data with consistency and resiliency, and can be easily scaled out according to an organization’s requirements. The configurations detailed in this document can be extended to clusters of various sizes depending on requirements. Cisco Compute Hyperconverged with Nutanix (CCHN), integrated with Splunk Enterprise, empowers organizations to effectively build and securely manage a next-generation data center infrastructure. This innovative solution not only scales to meet the growing demands of data but also enhances the potential for achieving smarter business outcomes.

For more information

- Cisco Validated Design for deploying Cisco Compute Hyperconverged with Nutanix in Intersight Standalone Mode
- Splunk Enterprise:
- Cisco Validated Design Zone
- Cisco Unified Computing System (Cisco UCS) servers
- Cisco Compute Hyperconverged with Nutanix
- Cisco Intersight
- Splunk Validated Architecture (SVA)
- HCIAF240C M7 All-NVMe/All-Flash Server
- Nutanix reference documentation

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **https://www.cisco.com/go/offices**.

Printed in USA C11-4908301-00 02/25