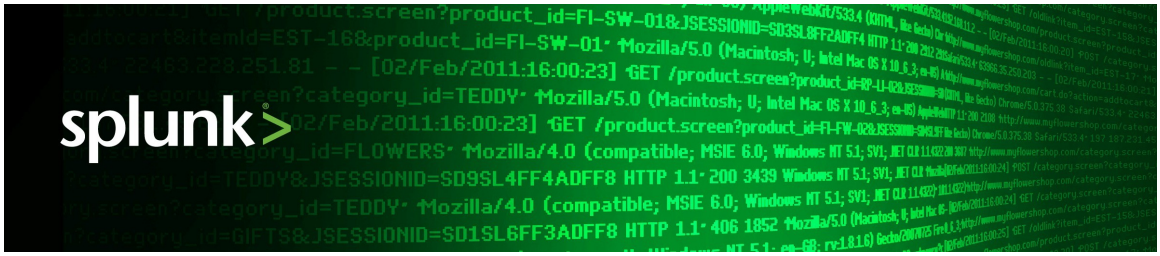

Splunk® Enterprise Capacity Planning Manual 6.5.0

Reference hardware

Generated: 11/04/2016 2:14 pm

Reference hardware

The reference hardware specification is a baseline for scoping and scaling the Splunk platform for your use. It is a performance guide for handling search and indexing loads.

Reference host specification for single-instance deployments

The following requirements represent the basic building block of a Splunk Enterprise deployment.

- Intel x86 64-bit chip architecture
- 12 CPU cores at 2Ghz or greater speed per core
- 12GB RAM
- Standard 1Gb Ethernet NIC, optional second NIC for a management network
- Standard 64-bit Linux or Windows distribution

There are two additional reference specifications that provide higher performance and search concurrency. These specs are described later in this topic.

Disk subsystem

The disk subsystem for a reference machine should be capable of handling a high number of average **I**nput/Output **O**perations **P**er **S**econd (IOPS).

IOPS are a measurement of how much data throughput a hard drive can produce. Because a hard drive reads and writes at different speeds, there are IOPS numbers for disk reads and writes. The average IOPS is the blend between those two figures.

The more average IOPS that a hard drive can produce, the more data it can index and search in a given period of time. While many variable items factor into the amount of IOPS that a hard drive can produce, the following are the most important:

- Its rotational speed in revolutions per minute.
- Its average latency, which is the amount of time it takes to spin its platters half a rotation.
- Its average seek time, which is the amount of time it takes to retrieve a requested block of data.

The drives that are capable of producing the highest IOPS have high rotational speeds and low average latency and seek times. Every drive manufacturer

provides this information, and some provide much more.

For information on IOPS and how to calculate them, see the following articles:

- Getting the hang of IOPS on the Symantec Connect Community.
- Analyzing I/O performance in Linux on CMDLN.ORG (A sysadmin blog).

This specification uses eight 146GB, 15,000 RPM, serial-attached SCSI (SAS) HDs in a Redundant Array of Independent Disks (RAID) 1+0 fault tolerance scheme as the disk subsystem. Each hard drive is capable of about 200 average IOPS. The combined array produces a little over 800 average IOPS.

Insufficient disk I/O is the most common limitation in Splunk infrastructure. For the best results indexing your data, review the disk subsystem requirements before provisioning your hardware.

Maximum performance capability

The maximum performance capabilities measure indexing and search performance independently, and do not represent the combined load of a typical Splunk use case. To review performance recommendations for a reference machine with indexing with search load, see Summary of performance recommendations.

Indexing performance

- Up to 20MB per second (1700GB per day) of raw indexing performance if no searching or other index-related activity occurs.

Search performance

- Up to 50,000 events per second for dense searches.
- Up to 5,000 events per second for sparse searches.
- Up to 2 seconds per index bucket for super-sparse searches.
- From 10 to 50 buckets per second for rare searches with bloom filters.

To find out more about the types of searches and how they affect Splunk Enterprise performance, see How search types affect Splunk Enterprise performance.

Reference host specification for distributed deployments

As the number of active users increases along with the data ingestion rate, the architecture requirements change from a single instance to a distributed Splunk Enterprise environment. The search head and indexer roles have unique

hardware recommendations.

Dedicated search head

A search head uses CPU resources more consistently than an indexer, but does not require the fast disk throughput or a large pool of local storage for indexing.

- Intel 64-bit chip architecture
- 16 CPU cores at 2Ghz or greater speed per core.
- 12GB RAM
- 2 x 300GB, 10,000 RPM SAS hard disks, configured in RAID 1
- A 1Gb Ethernet NIC, optional 2nd NIC for a management network
- A 64-bit Linux or Windows distribution

A search request uses 1 CPU core while the search is active. You must account for scheduled searches when you provision a search head in addition to ad-hoc searches that users run. More active users and higher concurrent search loads require additional CPU cores.

For a review on how searches are prioritized, see the topic Configure the priority of scheduled reports in the *Reporting Manual*. For information on scaling search performance, see How to maximize search performance.

Indexer

When you distribute the indexing process, the Splunk platform can scale to consume terabytes of data in a day. When you add more indexers, you distribute the work of search requests and data indexing across those indexers. This increases performance significantly.

As a reminder, here is the reference indexer specification:

Reference host specification

- Intel 64-bit chip architecture.
- 12 CPU cores at 2GHz or greater per core.
- 12GB RAM.
- Disk subsystem capable of 800 average IOPS. For details, see the topic Disk subsystem.
- A 1Gb Ethernet NIC, with optional second NIC for a management network.
- A 64-bit Linux or Windows distribution.

Splunk has introduced two new specifications that help improve user experience by providing additional CPU cores for better indexing performance and search

concurrency.

A single indexer carries the same disk I/O bandwidth requirements as a group of indexers.

Mid-range specification

The mid-range specification is similar to the base reference specification. This specification improves indexing capacity and search concurrency over a distributed Splunk Enterprise deployment.

- Intel 64-bit chip architecture
- 24 CPU cores at 2GHz or greater speed per core
- 64GB RAM
- Disk subsystem capable of 800 average IOPS
- A 1Gb Ethernet NIC, with optional second NIC for a management network
- A 64-bit Linux or Windows distribution

High-performance specification

The high-performance specification is a further improvement upon the mid-range specification.

- Intel 64-bit chip architecture
- 48 CPU cores at 2GHz or greater speed per core
- 128GB RAM
- A solid state disk (SSD) subsystem as a minimum requirement for hot and warm **index buckets**
- A 1Gb Ethernet NIC with optional second NIC
- A 64-bit Linux or Windows distribution

Disk subsystem information for higher-performance specifications

When indexers retrieve data for searches, they do many disk seeks and bulk reads. At higher daily volumes, local disk might not provide cost-effective storage for the time frames where you want a fast search. Fast attached storage or networked storage, such as storage area networks (SAN) over fiber, can provide the required IOPS for each indexer in these cases.

When you plan your storage infrastructure, understand these key points:

- More disks (specifically, more spindles) are better for indexing performance.
- Total throughput of the entire system is important.

- The ratio of disks to disk controllers in a particular system should be higher, similar to how you provision a database host.

Ratio of indexers to search heads

There is no practical limitation on the number of search heads that an indexer can support, or on the number of indexers that a search head can search against. The use case determines what Splunk instance role (search head or indexer) the infrastructure needs to scale while maintaining performance. For a table with scaling guidelines, see Summary of performance recommendations.

Premium Splunk app requirements

Premium Splunk apps can demand greater hardware resources than the reference specifications in this topic provide. Before architecting a deployment for a premium app, review the app documentation for scaling and hardware recommendations. The following list shows examples of some premium Splunk apps and their recommended hardware specifications.

- Splunk Enterprise Security
- Splunk IT Service Intelligence
- Splunk App for PCI

Virtual hardware

Splunk supports use of its software in virtual hosting environments. An indexer on a hypervisor (such as VMWare) with reserved resources that meet one of the hardware specifications can consume data about 10 to 15 percent slower than a indexer hosted on a bare-metal host. Search performance in a virtual hosting environment is a close match to bare-metal computers.

The performance that a virtual host provides is a best-case scenario that does not account for resource contention with other active virtual hosts that share the same physical host or storage array. It also does not account for certain vendor-specific I/O enhancement techniques, such as Direct I/O or Raw Device Mapping.

For information on how to run Splunk Enterprise in a VMWare virtual machine, see Deploying Splunk Enterprise Inside Virtual Environments on splunk.com.

Splunk Enterprise, self-managed in the cloud

Running Splunk Enterprise in the cloud is an alternative to running it on-premises using bare-metal hardware. Splunk Enterprise delivers similar performance on a

cloud-based infrastructure as it does on bare-metal hardware. Depending upon the vendor and technologies used to provision cloud instances, there can be less resources available than the OS reports.

If you run Splunk Enterprise on an Amazon Web Services (AWS) instance:

- AWS measures CPU power on Elastic Compute Cloud (EC2) instances in virtual CPUs (vCPUs), not real CPUs.
- Each vCPU is a hyper thread of an Intel Xeon core on most AWS instance types. See AWS | Amazon EC2 | Instance Types on the AWS web site.
- As a hyper thread of a core, a vCPU acts as a core, but the physical core must schedule its workload among other workloads of other vCPUs that the physical core handles.

For indexing and data storage, note the following:

- If you choose to use Elastic Block Storage (EBS), the type of EBS volume you choose determines the amount of performance you get.
- Not all EBS volume types have the necessary IOPS to handle Splunk Enterprise operations.
- The "Provisioned IOPS" and "Magnetic" EBS volume types offer the best opportunity to get the IOPS that you need for indexing and searching. See EBS - Product Details on the AWS web site.
- Not every EC2 instance type offers the network throughput to the EBS volume that you need. To ensure that bandwidth you must either launch the instance as "EBS-optimized" or choose an instance type that provides a minimum of 10Gb of bandwidth. See Amazon EC2 Instance Configuration on the AWS web site.

For forwarding, note that the proximity of your cloud infrastructure to your forwarders can have a major impact on performance of the whole environment.

For recommendations on running Splunk Enterprise in AWS, see Deploying Splunk Enterprise On Amazon Webservices on splunk.com.

Splunk Cloud

Splunk offers its machine data platform and licensed software as a subscription service called Splunk Cloud. When you subscribe to the service, you purchase a capacity to index, store, and search your machine data. Splunk Cloud abstracts the infrastructure specification from you and delivers high performance on the capacity you have purchased.

To learn more about Splunk Cloud, visit the Splunk Cloud website.