Training – Day #1

15 June 2023 04:28

Two dashboards -

(1) classic (built on top of XML)

(2) studio (built on top of JSON)

#**Splunk enterprise** - contains indexer, search head, deployment server, license master.

indexer - Splunk own database (A Splunk Enterprise instance that indexes data, transforming raw data into events and placing the results into an index)

search head - contains tables and queries are similar to SQL

License master - licensing bought on per day basis, it's a prepaid and works like jio 1.5 GB/day, when the limit (50 GB or depends upon the license we got). So once the daily limit exceeded, indexing will be done for the data but search will be disabled for all the existing data set. ***Total of 5 times we can use this feature of indexing extra GB that the license bought***

https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/Configurealicensemanager

https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/HowSplunklicensingworks

Note : Splunk Enterprise software is mandatory for Indexer, Search-Head, Deployment Server, License Master, Heavy Forwarder, Cluster Master, Search Head Cluster

---

#License Calculation :

License Calculation can be of two types:

Customers can purchase a commercial end-user license to Splunk Enterprise **based on either (i) data volume or (ii) infrastructure**. These Splunk Enterprise licenses are the most common license types. They provide access to the full set of Splunk Enterprise features within a defined limit of ***indexed data per day (volume-based license), or vCPU count (infrastructure license).***

From <https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/TypesofSplunklicenses>

Note:

(i) ***The Splunk Enterprise volume-based license*** **cannot stack with an infrastructure-based license**

(ii) ***The Splunk Enterprise infrastructure-based license*** **cannot stack with an volume-based license**

(iii) ***The Splunk Enterprise volume-based license*** *prevents searching if your license stack is less than* ***100 GB of data per day*** and there are a set number of license warnings.

From <https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/TypesofSplunklicenses>

License conditions	Enterprise: with less than 100 GB of data per day license stack	Enterprise: with 100 GB of data per day or larger license stack	Enterprise: Infrastructure (vCPU)
Currently blocks search while in violation	Yes	No	No
Logs internally and displays message in Splunk Web when in warning or violation	Yes	Yes	No
Stacks with other licenses	Yes	Yes	No
Enables full Splunk Enterprise feature set	Yes	Yes	Yes

Types of Splunk Enterprise licenses:

(1) ***Splunk Enterprise commercial end-user licenses***

1.i. The Splunk Enterprise ***volume-based*** license

1.ii.The Splunk Enterprise ***infrastructure*** license

(2) **Splunk developer licenses**

2.i. ***Dev/Test license*** => only available to customers that have acquired a commercial, paid license to Splunk Enterprise and is subject to Splunk General Terms. It's used by customers with pre-production environments to test upgrades and evaluate customized app + Dev/Test license cannot be stacked with other licenses.

2.ii. ***Developer license*** => gives you access to various Splunk developer tools on dev.splunk.com and a full set of Splunk Enterprise features. + **expires after 6 months.**

2.iii. ***Build Partner license***

(3) **Other types of licenses**

3.i. The Splunk Enterprise ***Trial license*** => gives access to all Splunk Enterprise features.+ for standalone, single-instance installations + cannot be stacked with other licenses. + expires 60 days after you install the Splunk Enterprise + allows you to index 500 MB of data per day

3.ii. ***Free license*** => allows a completely free Splunk Enterprise instance with limited functionality + Does NOT EXPIRE + allows you to index 500 MB of data per day.

3.iii. ***Pre-release license*** => subject to certain beta terms

How Splunk Enterprise licensing works

When data is sent to the Splunk platform, that data is **indexed** **and stored on disk.**

Part of the **indexing process** is

(i) to measure the volume of data being ingested, and

(ii) report that volume to the license manager for license volume tracking.

How data is measured?

(i) When ingesting **event data**, the measured data volume is based on the raw data that is placed into the indexing pipeline. It is not based on the amount of compressed data that is written to disk. Because the data is measured at the indexing pipeline, data that is filtered and dropped prior to indexing does not count against the license volume quota.

(ii) When ingesting **metrics data**, each metric event is measured by volume like event data. However, the per-event size measurement is capped at 150 bytes. Metric events that exceed 150 bytes are recorded as only 150 bytes. Metric events less than 150 bytes are recorded as event size in bytes plus 18 bytes, up to a maximum of 150 bytes. Metrics data draws from the same license quota as event data.

Metric data > 150 bytes => volume of data = 150 bytes

Metric data < 150 bytes => volume of data = event size + 18 bytes

From <https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/HowSplunklicensingworks>

Notes :

(1) The indexing volume is measured daily from midnight to midnight using the system clock on the license manager

(2) internal indexes such as `_internal` and `_introspection` + use of **summary indexing** **and metric rollup summaries** do not count against your license volume quota.

License	Violation conditions
Splunk Enterprise license	1 An Enterprise license stack with a license volume of 100 GB of data per day or more does not currently violate. 2 If you have a license stack with less than 100 GB of data per day of license volume, and you generate 45 license warnings in a rolling 60 day period, you are in violation of your license. If that license stack is split into multiple pools, search is disabled for a pool and its license pool member(s) after 45 warnings over a rolling 60-day window. Other pools and their members will remain searchable if the usage across the remaining license pools does not exceed their allocated license. To reenable search, request a reset license from Splunk Sales.
Splunk Enterprise infrastructure license	An Enterprise license based on vCPU usage does not currently violate.

https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/Aboutlicenseviolations#What_happens_during_a_license_violation.3F

#License Pooling :

when multiple servers which their respective license count, then add all those servers to Master and overall license will be applied to master and overall license count will have the flexibility of 20.

The **license manager is a Splunk Enterprise component used to manage licenses and assign license volume.**

Use the license manager to **group licenses, and assign them to stacks. You can create license pools from the stacks, and assign the license peers to a pool so they can use Splunk Enterprise features and have their license usage levied against a pool.**

https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/Groups,stacks,pools,andotherterminology

URL => https://docs.splunk.com/Documentation/Splunk/9.0.4/Admin/Createalicensepool

---

#Forwarder : Act as forwarding agent, which forward data from source to destination + Forwarder will be installed on application.

2 Types of Forwarder : (https://docs.splunk.com/Documentation/Splunk/9.0.4/Forwarding/Typesofforwarders#:~:text=The%20universal%20forwarder%20contains%20only,to%20reduce%20system%20resource%20usage.)

(1) **Universal Forwarder :**

- **(i) Don't do Parsing**
- **(ii) No GUI**
- **(iii) Light Application**
- **(iv) Size : 25 MB (tar) -> 125 MB (untar)**

(2) **Heavy Forwarder :**

- **(i) Do Parsing**
- **(ii) Have GUI**
- **(iii) Heavy Application as GUI is not their and it needs to install Splunk enterprise first.**
- **(iv) Size : 450 MB (tar) -> 5 GB (untar)**

#Why to use Heavy Forwarder ?

Answer : If hundred/thousands of UFs will send data to Indexer and Indexer will have to do all Parsing, so it will lead to **"Indexing Delay"**. After using HF, "indexing Delay" will be minimal but HF installation size will be huge. Hence best way is to -

Install UF on each Application server, then forward UF data to single HF for parsing and minimize 'indexing delay'

The universal forwarder contains only the components that are necessary to forward data.

*A **heavy forwarder is a full Splunk Enterprise instance that can index, search, and change data as well as forward it.** The heavy forwarder has some features disabled to reduce system resource usage*

#Parsing Stage :

Parsing is a stage to remove unwanted data. Parsing is mandatory before storing data in Indexes.

Parsing can happen either at HF level or Splunk Indexer end.

Note : (1)Default parsing is automatic and it happens at Indexer end.

(2) License is based on amount of Indexing, ***not on amount of Parsing.***

#Index Management:

Splunk stores events in indexes under ***SPLUNK_HOME/var/lib/splunk***

3 Main Indexes :

(1) ***Pre-Configured Index:***

(i) Having internal dataset and it is already created.

*(ii) name starts with "_" like _internal, _audit, _fishbucket (exception of nomenclature is "**summary**" index)*

(iii) We cannot store data in these indexes as license is not getting calculated

(2) ***Default Index***

*(i) **main** => used when an input does not specify any index*

(3) ***Custom Index***

(i) We can create Own indexes.

Note:

Let's assume a Server has 10 events. 3 events has been indexed and server goes down. So a "checkpoint" is getting stored in _fishbucket index to avoid the duplication.

If we want to re-index the data again, then we have to delete the _fishbucket.

#Retention Period

https://community.splunk.com/t5/Deployment-Architecture/How-to-deal-with-bucket-sizes-and-retention-policy/m-p/150760

https://docs.splunk.com/Documentation/Splunk/9.0.4/Indexer/Bucketsandclusters

#Buckets:

Input => Hot Bucket => Warm Bucket => Cold Bucket => (i) Archive OR (ii) Delete

<---------- index ---------->

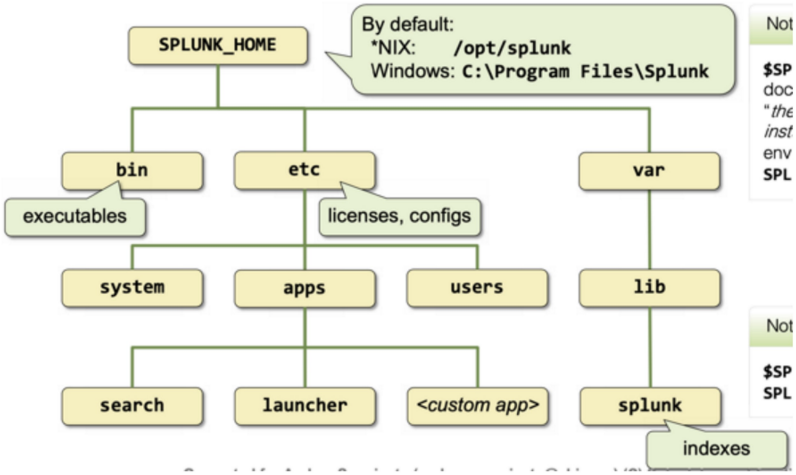
Data moves from one bucket to other based on -

(i) Size of each Bucket type

(ii) Retention Period for each Bucket type

#Splunk Directory Structure:

https://it-learn.io/2020/05/11/splunk-series-iii-system-administrator-class-file-structure-settings-and-cli/



Splunk Configuration files are in 2 folders which are having almost same kind of config files :

(1) ***Default*** => these files are coming with package

(2) ***Local*** => custom/user specific configuration

NOTE :

(i) Local will have ***HIGH PRECEDENCE*** over Default files.

(ii) Application level will have ***HIGH PRECEDENCE*** over System level

Reason for Two kind of folders is that :

(i) it is useful in upgrade and to take Backup.

(ii) Helps in Rollback

 # job management

From <https://docs.splunk.com/Documentation/SplunkCloud/9.0.2303/Search/Aboutjobsandjobmanagement>

Each time you run a search, create a pivot, open a report, or load a dashboard panel, the Splunk software creates a job **in the system**

A job is a process that tracks information about the ad hoc search or saved search.

The information that is tracked includes -

a. the **owner** of the job,
b. the **app** that the job was run on,
c. **how many events were returned**, and
d. **how long the job took to run.**

Each job process creates a search artifact**. The artifact contains the results and associated metadata that are returned at the time that the ad hoc search or saved search was run.**

When you run a new search, a job is retained in the system for a period of time, called the job **lifetime**. **The default lifetime is 10 minutes.** The lifetime starts from the moment the job is run. See Extending job lifetimes in this manual.

Managing long-running jobs:

(1) ***Autopause long-running jobs***

To handle inadvertently long-running search jobs, you can autopause a job.

This feature is **enabled by default only for summary dashboard clicks**, to deal with the situation where a user mistakenly initiates **"all time" searches.**

By default, the limit before autopause is 30 seconds.

(2) **Managing jobs when a computer goes into sleep mode**

When a search is run in Splunk Web from a computer that is not a Splunk server and the computer changes to sleep or hibernate mode, the underlying search process is stopped.

Splunk software interprets the change to sleep or hibernate mode as if the browser tab in which the software is running has been closed and is no longer being used.

To avoid this issue, use one of the following techniques:

(i) Send the job to the background. The job continues to run in the background even when your computer goes into sleep or hibernate mode.

(ii) Save and schedule the search. The search runs independently from the computer that was used to create the search.

(iii) Share the job. The lifetime of the job is automatically extended to **7 days** and read permissions are set to **Everyone.**

Edit search restriction settings

To edit the search restrictions setting for a role:

In Splunk Web, go to Settings > Access Controls > Roles. In Splunk Enterprise you can manually edit search restrictions, which are specified in the **authorize.conf** file, as described in Edit search restrictions manually.

[role_ninja]

rtsearch = enabled

importRoles = user

srchFilter = host=foo

*srchIndexesAllowed = **

srchIndexesDefault = mail;main

srchJobsQuota = 8

rtSrchJobsQuota = 8

srchDiskQuota = 500

srchTimeWin = 86400

srchTimeEarliest = 2592000

 # This creates the role 'ninja', which inherits capabilities from the 'user'

 # role. ninja has almost the same capabilities as power, except cannot

 # schedule searches.

#

*# The search filter limits ninja to **searching on host=foo.***

#

 # ninja is allowed to search all public indexes (those that do not start

 # with underscore), and will search the indexes mail and main if no index is

 # specified in the search.

#

*# ninja is **allowed to run 8 search jobs and 8 real time search jobs***

 # concurrently (these counts are independent).

#

*# ninja is **allowed to take up 500 megabytes total on disk for all their jobs.***

#

*# ninja is allowed to run searches that span a **maximum of one day***

#

*# ninja is **allowed to run searches on data that is newer than 30 days ago***
