

**KPLABS Course**

**Splunk 2021 - Beginner to Architect**

**Domain 3**

## **ISSUED BY**

Zeal

## **REPRESENTATIVE**

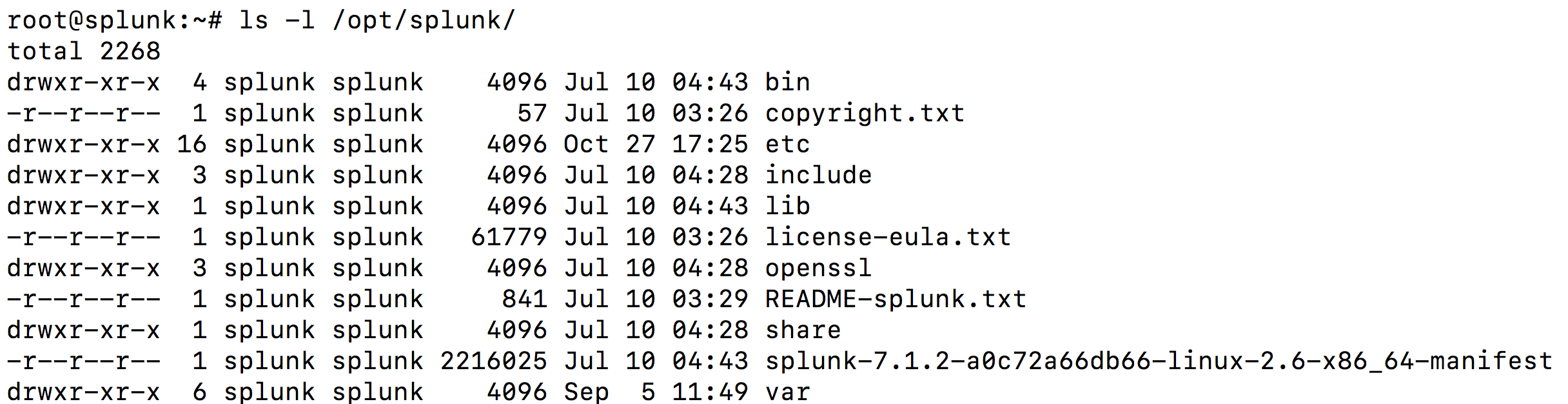
[instructors@kplabs.in](mailto:instructors@kplabs.in)

# 

# **Domain 3 - Splunk Architecture**

## **Module 1: Directory Structure of Splunk**

By default, splunk installation happens in /opt directory.



bin/ directory contains the primary splunk binary as well as various others like btool.

var/ directory primarily contains all the data that gets indexed as well as log files.

etc/ directory contains all the configuration files as well as all the apps and adoons that you install in splunk.

lib/ directory contains necessary libraries needed for splunk to run.

## **Module 2: Splunk Configuration Directories**

2.1 Overview of Splunk Configuration Directories

A single Splunk instance typically has multiple versions of configuration files across several directories within the filesystem.

We can have configuration files with the same names in default, local, and app directories.

This type of structure creates a layering effect that allows Splunk to determine priorities.

The following diagram illustrates the configuration file precedence



2.2 Default Configuration Files

"all these worlds are yours, except /default - attempt no editing there"

The default directory contains preconfigured versions of the configuration files.

Location: $SPLUNK\_HOME/etc/system/default

Important:

Never change or copy anything inside the default directory.

2.3 Local Configuration Files

Changes that you might make under the default/ directory would get overwritten during the upgrade process.

Thus, it’s recommended to create and edit your files in one of the local configuration directories.

Location: $SPLUNK\_HOME/etc/system/local

## 

.

**Module 3: Splunk Configuration Precedence**

3.1 Precedence Order

Following precedence order determines the priority

* System local directory -- highest priority
* App local directories
* App default directories
* System default directory -- lowest priority

Splunk typically searches for attributes within the system/local directory.

Then it looks for any copies of files located in app directories. (ignoring attributes found in s/l)

As last resort, for any attributes not defined above, it looks into system/default.

3.2 Important Pointer

App directory names also affect the precedence.

To determine priority among the collection of app directories, Splunk uses lexicographical order.

Files in an apps directory named "A" have a higher priority than files in an apps directory named "B"

Also, all apps starting with an uppercase letter have precedence over any apps starting with a lowercase letter, due to lexicographical order.

("A" has precedence over "Z", but "Z" has precedence over "a", for example.)

3.3 Summary of Precedence

$SPLUNK\_HOME/etc/system/local/\*

$SPLUNK\_HOME/etc/apps/A/local/\* ... $SPLUNK\_HOME/etc/apps/z/local/\*

$SPLUNK\_HOME/etc/apps/A/default/\* ... $SPLUNK\_HOME/etc/apps/z/default/\*

$SPLUNK\_HOME/etc/system/default/\*

**Module 4: Introduction to Indexes**

The index is a repository of Splunk data.

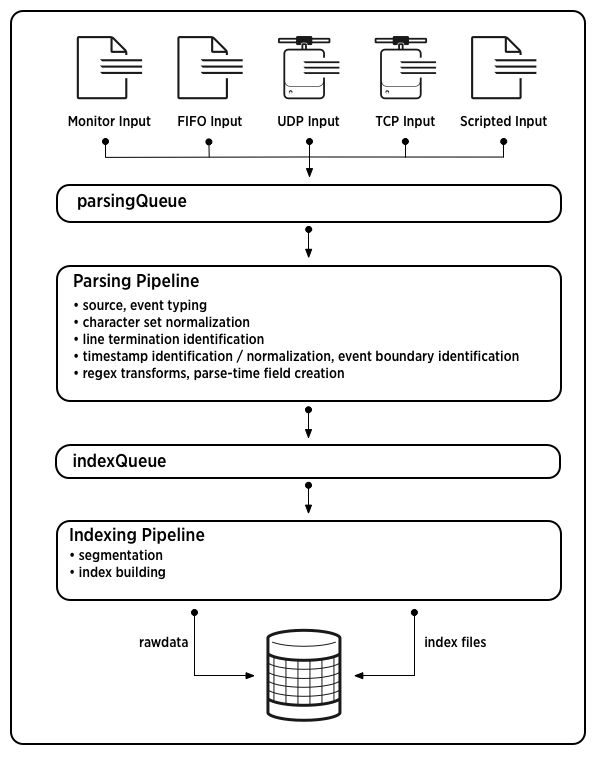
Splunk transforms incoming data into events, which it stores in the indexes.

When Splunk indexes your data, it creates a number of files. These files fall into two main categories:

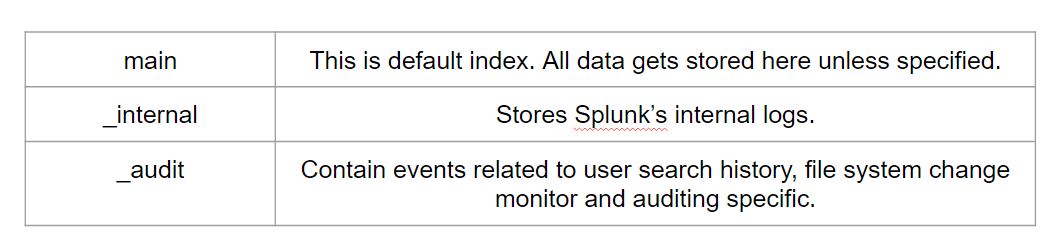
The raw data in compressed form (rawdata)

Indexes that points to raw data (tsidx files), plus some meta-data files.

These files reside in a set of directories organized by age.



Splunk Enterprise comes with a number of pre-configured indexes, including:

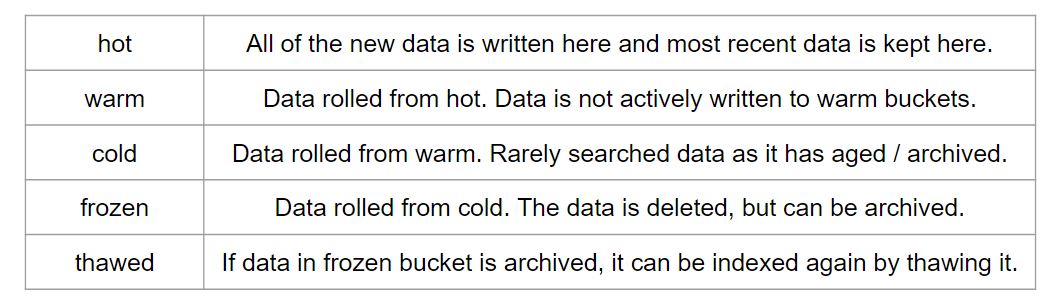


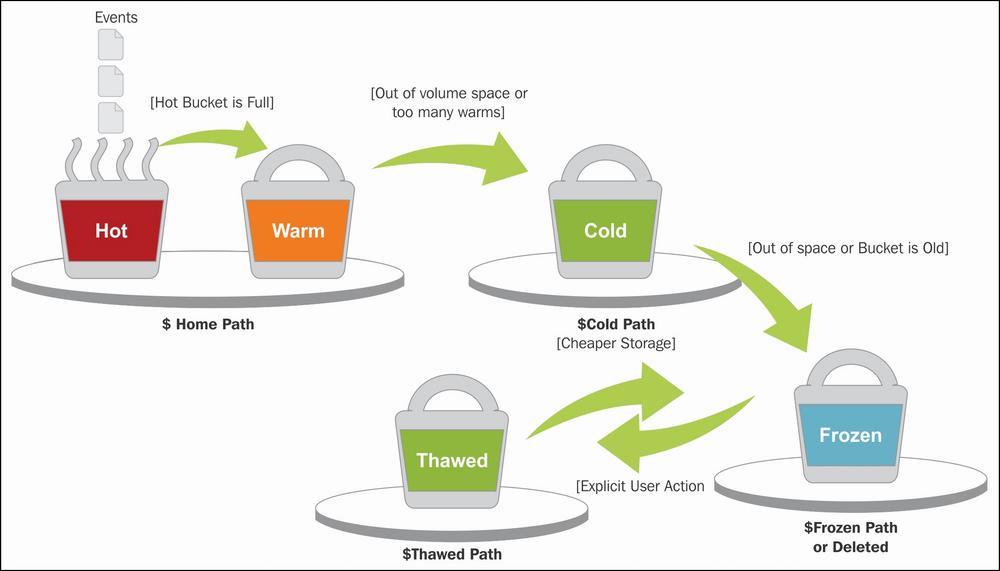
**Module 5: Bucket Lifecycle**

5.1 Overview of Bucket Lifecycle

Splunk stores all its data in directories on the server called buckets

A bucket moves through several stages as it ages – hot, warm, cold, frozen





5.2 Hot Bucket to Warm Bucket

Buckets are rolled from hot to warm in the following condition:

* We get too many hot buckets [maxHotBuckets]
* Hot bucket has not received data since a while
* Timespan of buckets is too large.
* Bucket meta-data files have grown large.
* Index clustering replication error.
* Splunk is restarted

5.3 Warm to Cold Buckets

Ideally, historical data should go here.

Allows us to keep older data on slower (cheaper) storage.

Buckets are rolled from warm to cold when there are too many warm buckets.

[index\_name]

coldPath = $SPLUNK\_DB/$\_index\_name/colddb

maxWarmDBCount = 300

5.4 Cold to Frozen

Data in frozen is no longer searchable.

Data rolls from cold to the frozen bucket when:

* Total size of index (hot+warm+cold) grows too large.
* Oldest event in bucket exceeds specific age.

Config: coldToFrozenDir

In the default process, tsidx file is removed and the bucket is specified to

The destination we specify.

5.5 Thawing Process

This is generally a manual process for restoring archived data.

Overall Steps:

i) mv /tmp/frozendb/db\* $SPLUNK\_HOME/var/lib/splunk/defaultdb/thaweddb/

ii) splunk rebuild $SPLUNK\_HOME/var/lib/splunk/defaultdb/thaweddb/db\*

iii) splunk restart

**Module 6: Splunk Workflow Actions**

Splunk WorkFlow Actions allows us to add interactivity between the indexed fields and other web resources.

Example:

* There is a field called as clientip in access\_combined log file.
* You can add option for “Whois Lookup” based on the IP address in clientip field.

## **Join Our Discord Community**

We invite you to join our Discord community, where you can interact with our support team for any course-based technical queries and connect with other students who are doing the same course.

Joining URL:

<http://kplabs.in/chat>

