KPLABS Course

Splunk 2021 - Beginner to Architect

Domain 3

ISSUED BY

Zeal

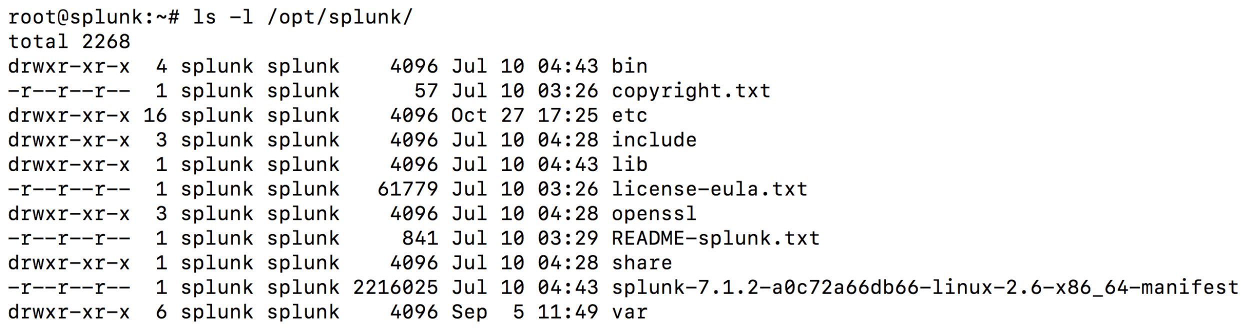
REPRESENTATIVE

instructors@kplabs.in

Domain 3 - Splunk Architecture

Module 1: Directory Structure of Splunk

By default, splunk installation happens in /opt directory.



```
root@splunk:~# ls -l /opt/splunk/
total 2268
drwxr-xr-x  4 splunk splunk    4096 Jul 10 04:43 bin
-r--r--r--  1 splunk splunk      57 Jul 10 03:26 copyright.txt
drwxr-xr-x 16 splunk splunk    4096 Oct 27 17:25 etc
drwxr-xr-x  3 splunk splunk    4096 Jul 10 04:28 include
drwxr-xr-x  1 splunk splunk    4096 Jul 10 04:43 lib
-r--r--r--  1 splunk splunk   61779 Jul 10 03:26 license-eula.txt
drwxr-xr-x  3 splunk splunk    4096 Jul 10 04:28 openssl
-r--r--r--  1 splunk splunk     841 Jul 10 03:29 README-splunk.txt
drwxr-xr-x  1 splunk splunk    4096 Jul 10 04:28 share
-r--r--r--  1 splunk splunk 2216025 Jul 10 04:43 splunk-7.1.2-a0c72a66db66-linux-2.6-x86_64-manifest
drwxr-xr-x  6 splunk splunk    4096 Sep  5 11:49 var
```

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

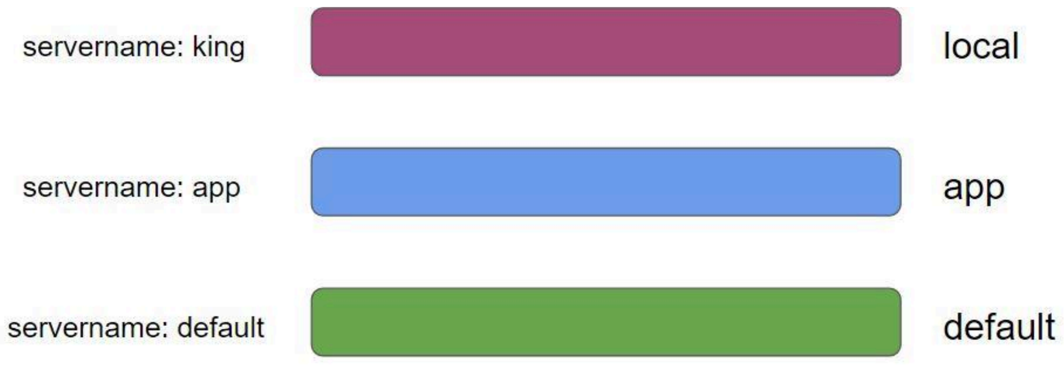
2.1 Overview of Splunk Configuration Directories

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

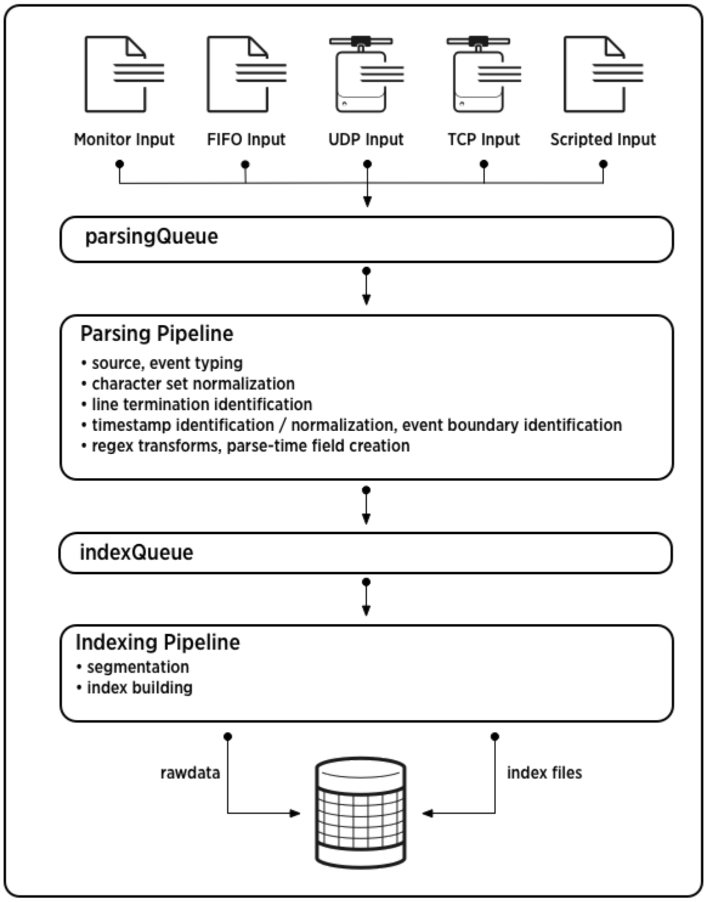
The index is a repository of Splunk data.

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

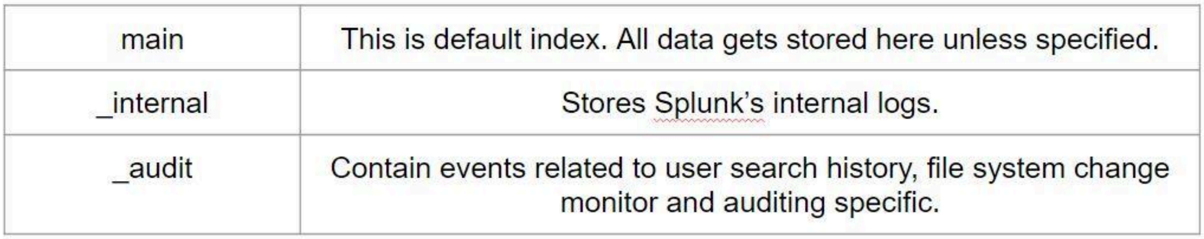
When Splunk indexes your data, it creates a number of files. These files fall into two main categories:

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.


Monitor Input
FIFO Input
UDP Input
TCP Input
Scripted Input
parsingQueue
Parsing Pipeline
• source, event typing
• character set normalization
• line termination identification
• timestamp identification / normalization, event boundary identification
• regex transforms, parse-time field creation
indexQueue
Indexing Pipeline
• segmentation
• index building
rawdata
index files

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



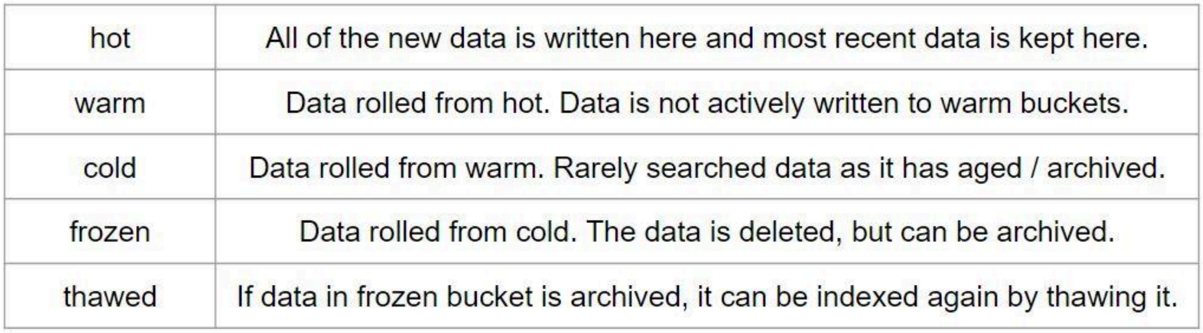
main	This is default index. All data gets stored here unless specified.
_internal	Stores Splunk's internal logs.
_audit	Contain events related to user search history, file system change monitor and auditing specific.

Module 5: Bucket Lifecycle

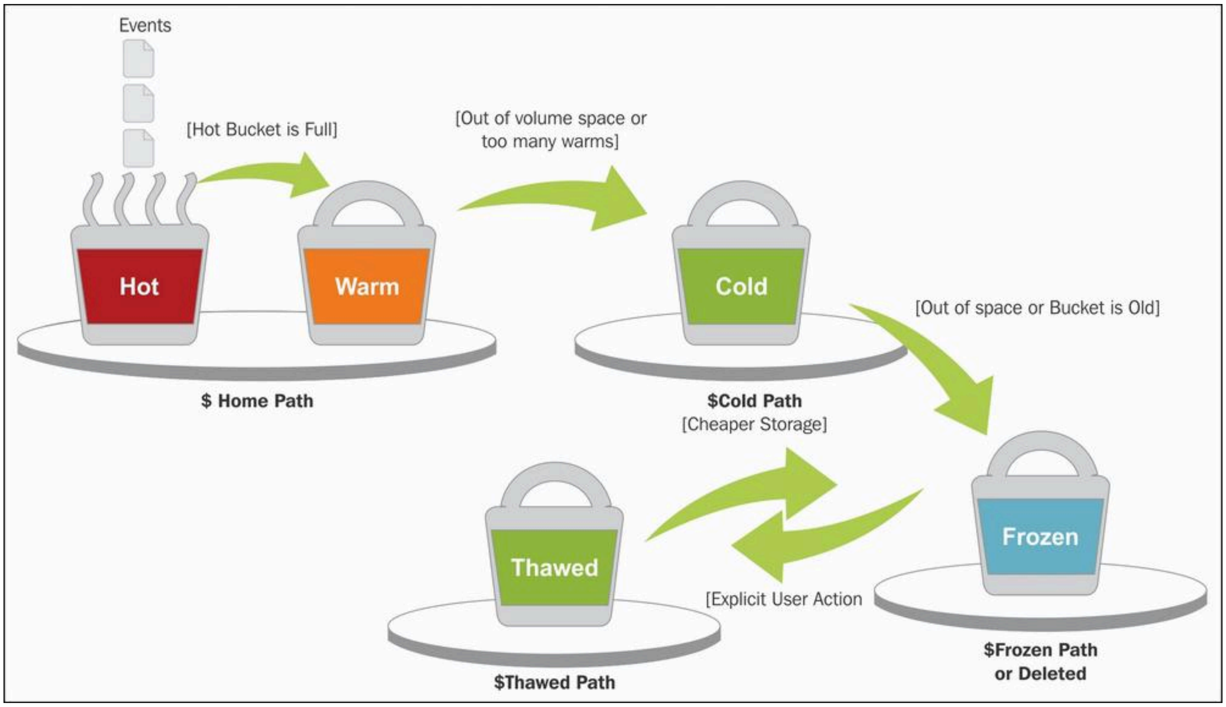
5.1 Overview of Bucket Lifecycle

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



hot	All of the new data is written here and most recent data is kept here.
warm	Data rolled from hot. Data is not actively written to warm buckets.
cold	Data rolled from warm. Rarely searched data as it has aged / archived.
frozen	Data rolled from cold. The data is deleted, but can be archived.
thawed	If data in frozen bucket is archived, it can be indexed again by thawing it.



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:

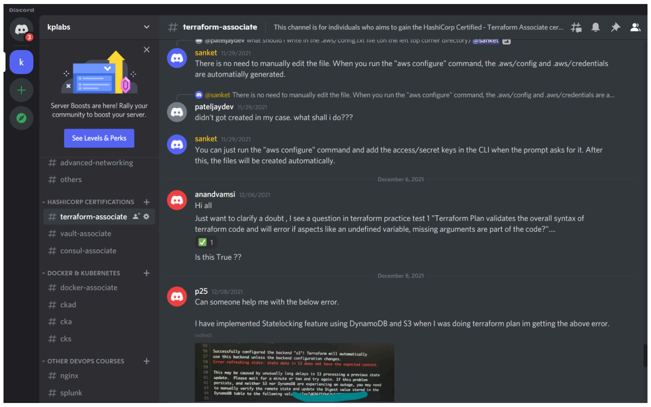
- There is a field called as clientip in access_combined log file.
- You can add option for “Whois Lookup” based on the IP address in clientip field.

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



terraform-associate | This channel is for individuals who aims to gain the HashiCorp Certified - Terraform Associate cer...

@pateljaydev what should i write in the .aws/config.txt file (on the left top corner directory) @sanket

sanket 11/29/2021
There is no need to manually edit the file. When you run the "aws configure" command, the .aws/config and .aws/credentials are automatially generated.

@sanket There is no need to manually edit the file. When you run the "aws configure" command, the .aws/config and .aws/credentials are a...

pateljaydev 11/29/2021
didn't got created in my case. what shall i do???

sanket 11/29/2021
You can just run the "aws configure" command and add the access/secret keys in the CLI when the prompt asks for it. After this, the files will be created automatically.

December 6, 2021

anandvamsi 12/06/2021
Hi all
Just want to clarify a doubt , I see a question in terraform practice test 1 "Terraform Plan validates the overall syntax of terraform code and will error if aspects like an undefined variable, missing arguments are part of the code?"....
✅ 1
Is this True ??

December 8, 2021

p25 12/08/2021
Can someone help me with the below error.

I have implemented Statelocking feature using DynamoDB and S3 when I was doing terraform plan im getting the above error.
(edited)

```
Successfully configured the backend "s3"! Terraform will automatically
use this backend unless the backend configuration changes.
Error refreshing state: state data in S3 does not have the expected content.

This may be caused by unusually long delays in S3 processing a previous state
update.  Please wait for a minute or two and try again. If this problem
persists, and neither S3 nor DynamoDB are experiencing an outage, you may need
to manually verify the remote state and update the Digest value stored in the
DynamoDB table to the following valu[redacted]
```