KPLABS Course

Splunk 2021 - Beginner to Architect

Domain 7

ISSUED BY
Zeal

REPRESENTATIVE
instructors@kplabs.in

Domain 7 - Distributed Splunk Architecture

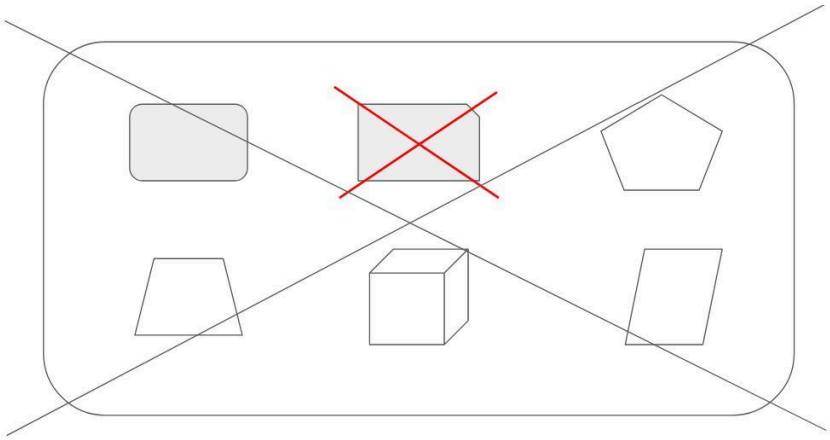
Module 1: Overview of Distributed Splunk Architecture

1.1 Understanding Splunk Components

Splunk Enterprise consist of various other sub-components. These includes:

- Indexer
- Search Head
- Deployment Server
- Forwarders
- License Master
- Monitoring Console

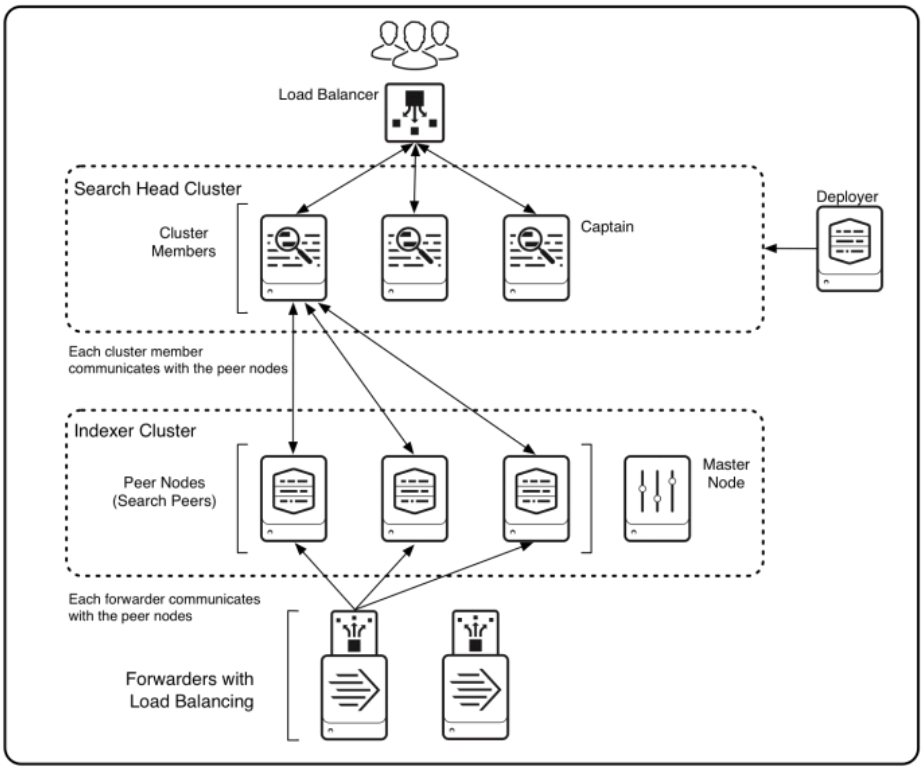
If you have all these components within a single server, the server failure can lead to the entire Splunk going down. This approach is hence not the recommended way.



Splunk Enterprise

The better architecture in production would be to make use of a clustered setup. In this approach, we deploy multiple instances of Splunk Enterprise that join themselves as part of the cluster.

In this architecture, even if one of the servers as part of the cluster goes down, the remaining servers can continue to take the responsibility of serving traffic.



Splunk Enterprise supports clustering features for two major components:

- Indexer
- Search Heads

For components like license master, heavy forward, etc, Splunk supports Active-Passive failover.

Module 2: Understanding License Master

Splunk Enterprise ingests external data, indexes it, and stores it on disk.

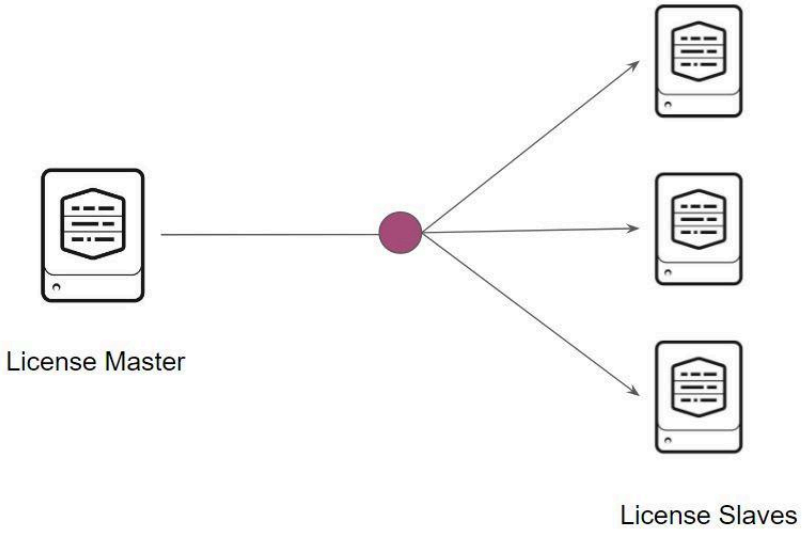
Licenses specify how much external data you can index per day.

All Splunk Enterprise instances require a license.

If you have a standalone indexer, you can install the license locally.

In the case of distributed environment, we need to configure a licens master.

The following diagram depicts the architecture of a License Master.

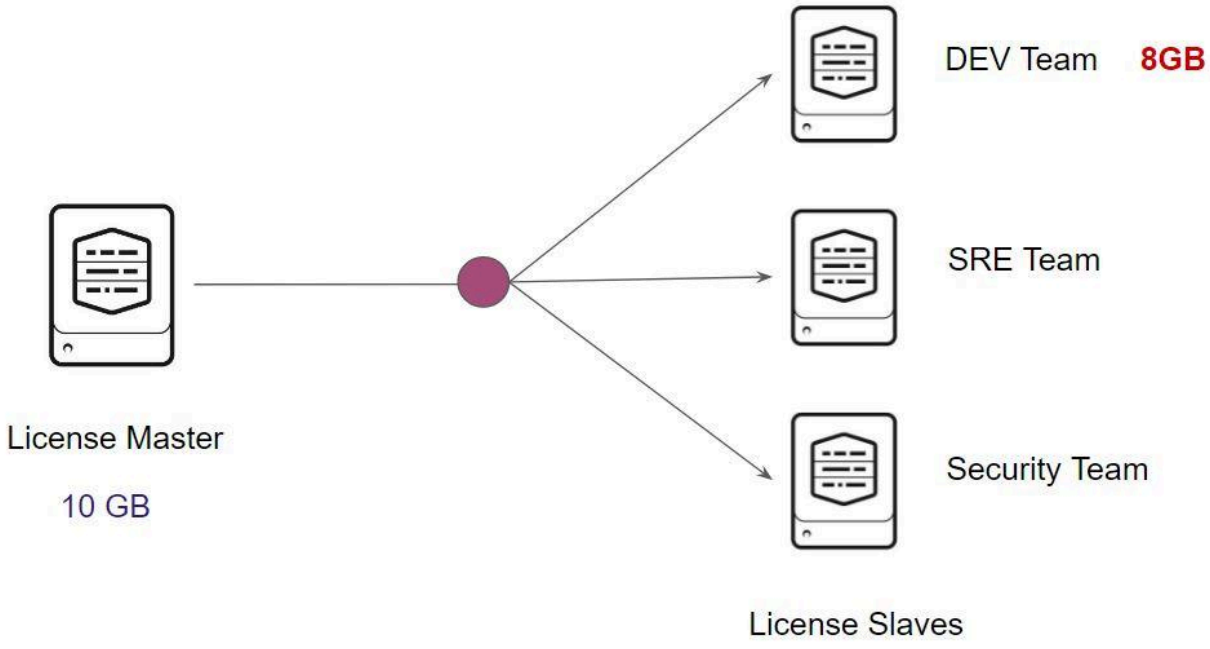


Module 3: License Pool

Splunk License resides in license stack called as Splunk Enterprise Stack.

The stack has a default license pool called auto_generated_pool_enterprise

Any license slave that connects to this license master has access to the default pool.

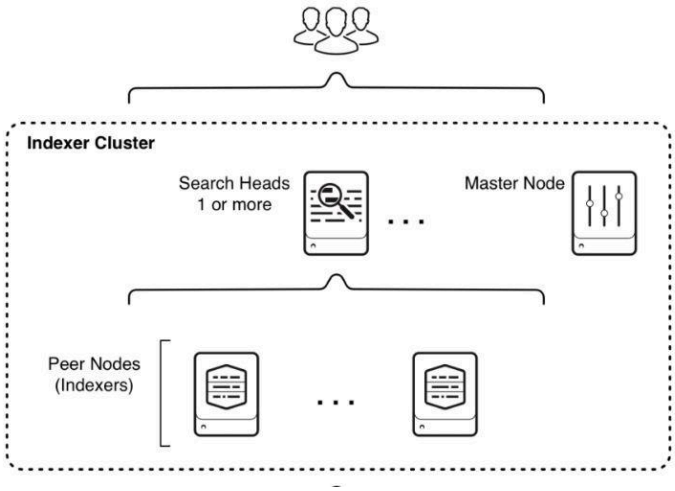


Module 4: Indexer

4.1 Overview of Indexer

Indexer is a component in Splunk Enterprise whose responsibility is to index data, transform data into searchable events, and placing results into an index.

To ensure high availability of data, we can deploy an indexer cluster.



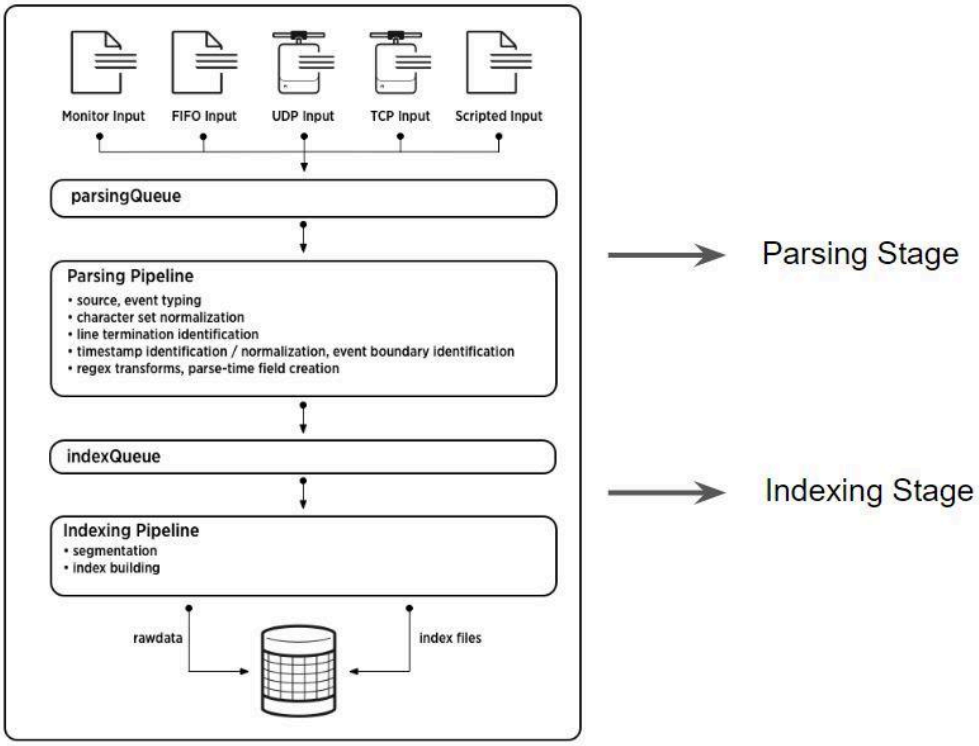
4.2 Parsing and Indexing Stage

While parsing, Splunk Enterprise performs number of actions, including:

- Extracting the set of default fields for each event, including host, source, and sourcetype.
- Configuring character set encoding.
- Identifying line termination using line break rules.
- Mask sensitive details in data like credit card numbers.

During the indexing pipeline, Splunk performs

- Breaking events into segments that can be searched upon.
- Building index data structures.
- Writing raw data and index files to disks.



Module 5: Masking Sensitive Data at Index Time

It might happen that log files would have sensitive information like Credit Cards, SSN, etc.

In such use-cases, you might want to mask such information;

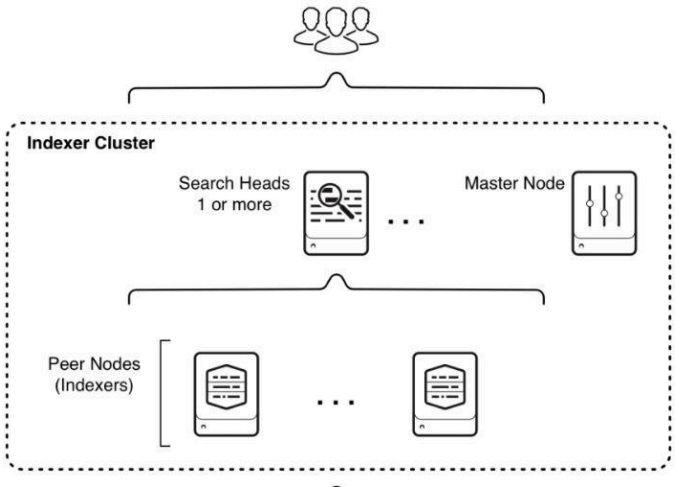
ss=123456789, cc=1234-5678-9012-3456

TO

ss=123456789, cc=xxxx-xxxx-xxxx-3456

Module 6: Search Head

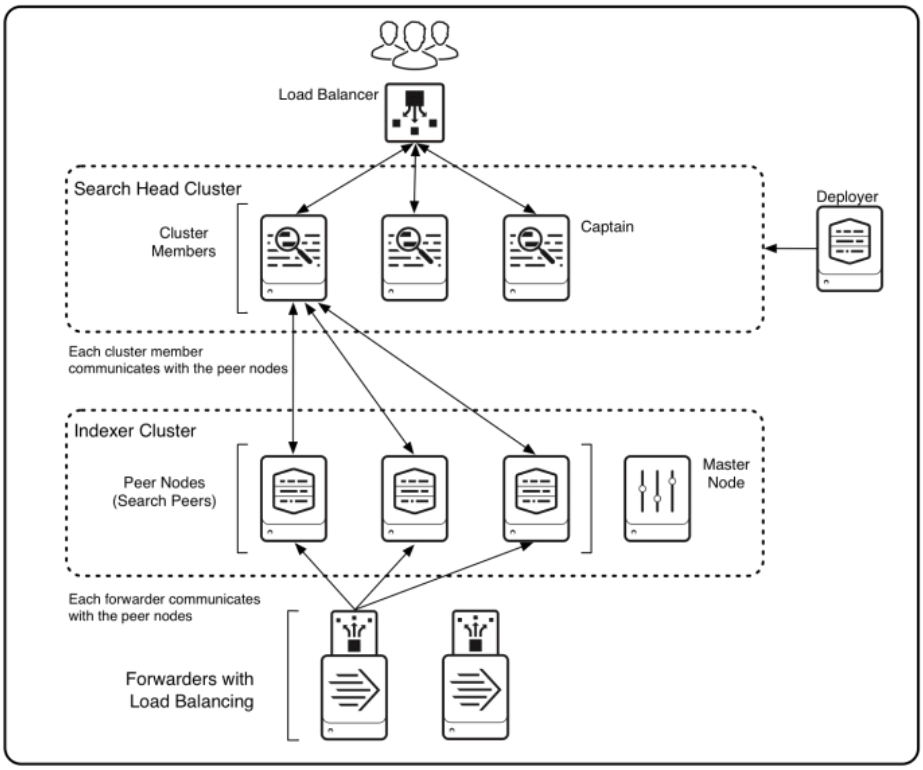
Search Head is a component in Splunk Enterprise whose responsibility is to handle the search management functions, directing search requests to search peers and then merging the results back to the users.



Search Heads are used for a number of functions, some of the primary ones include::

- Search Related Functions.
- Building Dashboards and Reports.
- Data Models.
- Alerting Related Functionality.

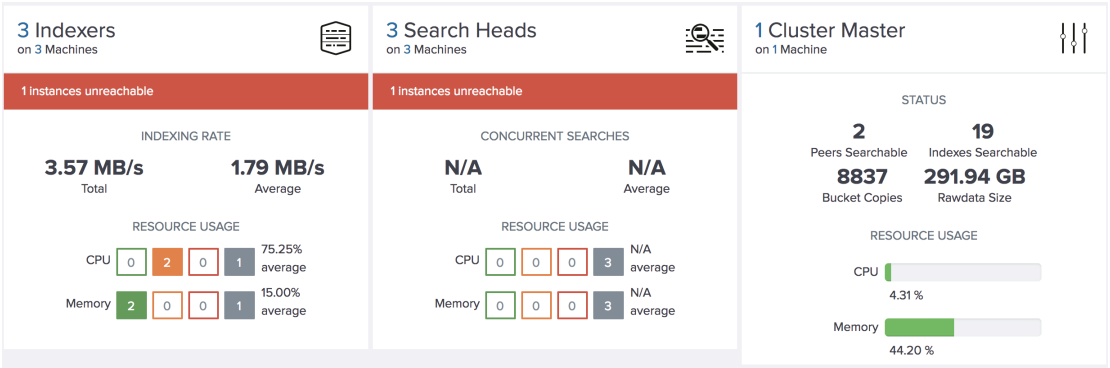
Search Head also supports clustering features. The same can be identified in the following architecture diagram.



Module 7: Splunk Monitoring Console

Monitoring Console allows us to view is detailed information about the topology and performance of your Splunk Enterprise deployment.

The Monitoring Console provides pre-built dashboards that give you visibility into many areas of your deployment, including search and indexing performance, resource usage, license usage, and more.



Pre-Built available dashboards in Monitoring Console provide insights into the following areas:

- search performance and distributed search framework
- indexing performance
- operating system resource usage
- Splunk app key-value store performance
- search head and indexer clustering
- index and volume usage
- forwarder connections and Splunk TCP performance
- HTTP Event Collector performance
- and license usage.

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