KPLABS Course

Splunk 2021 - Beginner to Architect

Domain 8

ISSUED BY

Zeal

REPRESENTATIVE

instructors@kplabs.in

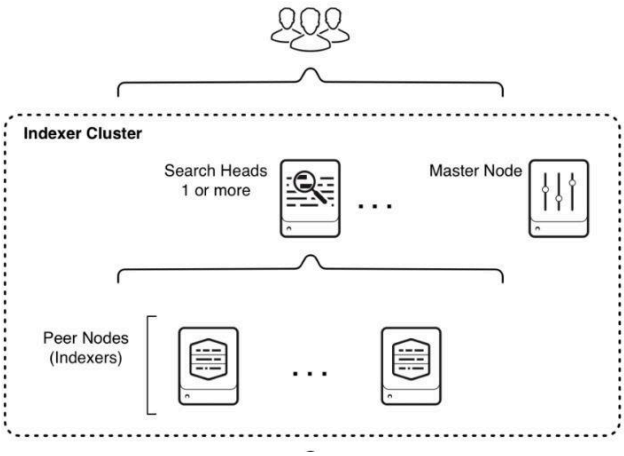
Domain 8 - Indexer Clustering

Module 1: Overview of Indexer Clustering

1.1 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.

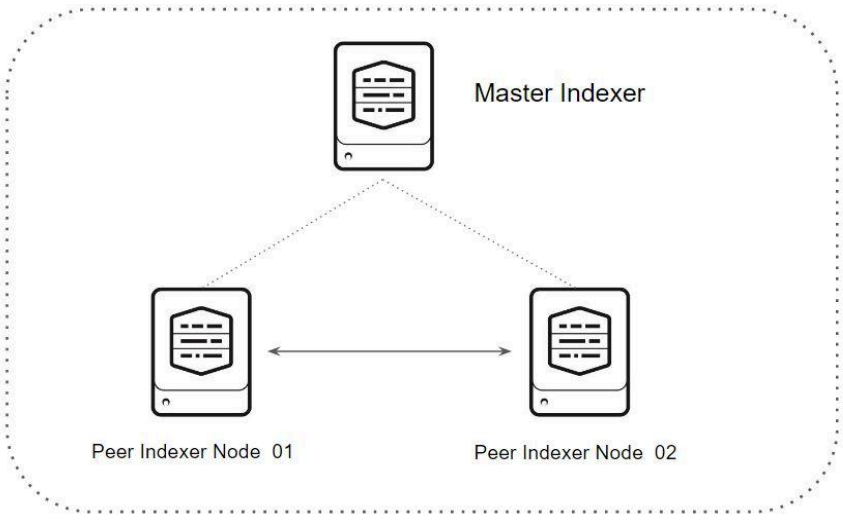
To ensure high availability of data, we can deploy an indexer cluster.



1.2 Master Node and Peer Node

Master Node coordinates the activities of the peer nodes.

Peer Nodes are the nodes which has the actual data and performs the replication-related activities. Any search-related activity requests are sent to the peer nodes.



Module 2: Overview of Indexer Clustering

2.1 Overview of Master Indexer

A cluster has one, and only one master node.

Master node coordinates activities of the peer nodes.

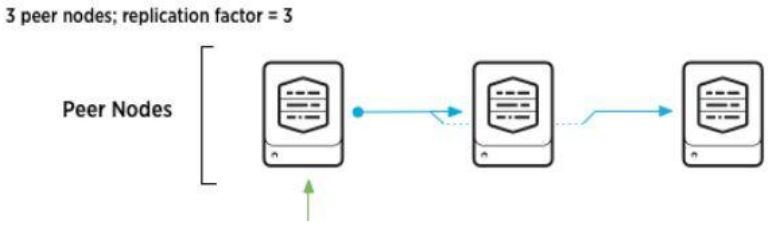
It does not itself store or replicate data.

2.2 Replication Factor

Replication Factor determines how many copies of data the cluster maintains.

This is a key factor since it determines the cluster's fault tolerance.

For example, if we want to ensure that system can handle the failure of two peer nodes, we must configure a replication factor of 3, which means the cluster will store 3 identical copies of data on separate nodes.



2.3 Search Factor

Search Factor determines the number of immediately searchable copies of data the cluster maintains.

Searchable copies of data require more storage than non-searchable copies.

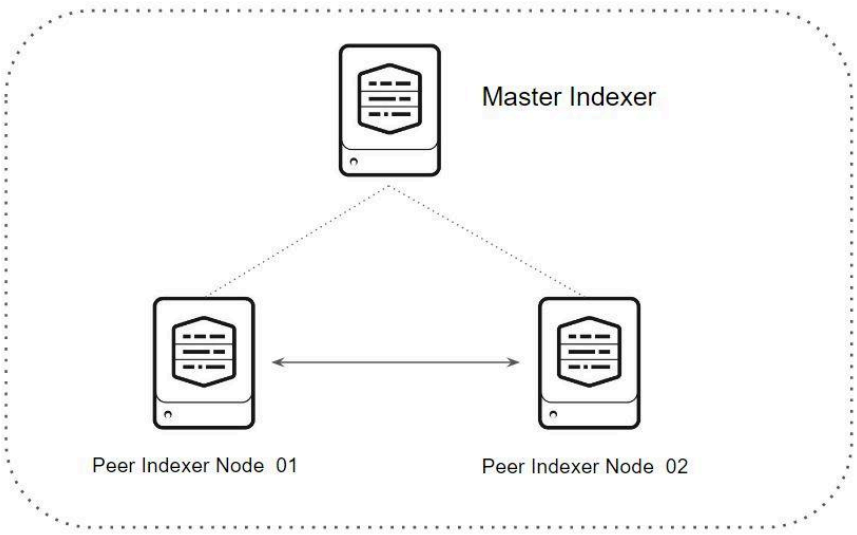
A non-searchable copy is basically plain raw data without any index files.

Module 3: Configuration Bundle

Configuration Bundle is a set of configuration files and apps common to all the peers.

Managed by the master and distributed to peers via bundle push operation.

It is important to note that we should never configure things directly within the peer nodes.



Module 4: Forwarding Logs to Indexer Cluster

Once we have the Indexer cluster built, the universal forwarder would need to start sending logs to the peer indexer nodes.

There are two ways to connect forwarders to indexer nodes:

i) Using the Indexer Discovery Feature.

ii) Connect forwarders directly to the peer nodes.

Module 5: Indexer Discovery

With the Indexer Discovery method, each forwarder queries the master node for a list of peer nodes within the cluster.

It then uses load balancing to forward data to a set of peer nodes.