Scenario based Splunk Admin Interview Questions (Part – 1 & 2)

Table of Contents

Prerequisites

- Basic understanding of Splunk Components & their functions

Notice: This Document provides a few examples of scenario-based questions one can expect in an interview, but not limited to.

Q1: What happens, If Cluster Master down

If a master goes down, the ***cluster can continue to run as usual***, as long as there are no other failures. Peers can continue to ingest data, stream copies to other peers, replicate buckets, and respond to search requests from the search head, at least for a while. Nevertheless, ***you should treat a downed master as a serious failure***. Cluster master is responsible for bucket replication & maintaining the availability of Cluster, so as soon as possible, the cluster master has to be up.

To deal with the possibility of a downed master, you can configure a stand-by master that can take over if needed.

Q2: What happens, If one of the Indexers down in 3-member cluster

Assumption: ***The cluster has Search factor-2 & Replication Factor-3***

- When a peer goes down, the master can therefore instruct the remaining peers to fix the cluster's set of buckets, with the aim of returning to the state where the cluster has:
  1) Exactly one primary copy of each bucket.
  2) A full set of searchable copies for each bucket, matching the search factor.
  3) A full set of copies (searchable and non-searchable combined) for each bucket, matching the replication factor (the complete state).

- ***Searches can continue*** across the cluster after a node goes down; however, searches will provide ***only partial results*** until the cluster regains its valid state

Q3: What happens, If two of the Indexers down in 3-member cluster

Assumption: ***The cluster has Search factor-2 & Replication Factor-3***

- When two of the indexer peers go down, (if the cluster has Search factor-2 & Replication Factor-3). ***The Cluster master will instruct the last remaining peer to make the copies of all available buckets as searchable ones***, also make them as primary copy (This includes creating the index files from the raw data).
- During this activity some of the searches will give ***incomplete results***. once all the bucket copies are made searchable, then the searches can give complete results

Q4: What happens, If all of the Indexers down in 3-member cluster

- The Cluster master will show all members status as down. ***Forwarders will hold the data*** which can significantly increase the waiting queue on the forwarders.
- This means ***the complete data is inaccessible***.
- ***Search head will not show any data*** for queries, as there are no indexers to take the request.

Q5: What happens, If one of the Search Heads down in 3-member cluster

Assumption: ***The cluster has Search factor-2 & Replication Factor-3***

- If a search head cluster member fails for any reason and leaves the cluster unexpectedly, the ***cluster can usually continue to function*** without interruption.
- All search artifacts resident on the failed member remains available through other search heads
- If the failed member was serving as captain, the remaining nodes elect another member as captain. Since members share configurations, the new captain is immediately fully functional. But here the election requires a minimum of 3 nodes, unfortunately we have only 2. So, we have to go with configuring anyone of the available member as a Static captain & the other one as member

Q6: What happens, If two of the Search Heads down in 3-member cluster

Assumption: ***The cluster has Search factor-2 & Replication Factor-3***

- If a majority of members fail, the ***cluster cannot successfully elect a new captain***, which results in failure of the entire cluster.
- In this case you may have to go ahead with configuring the remaining member as static captain

Q7: What happens, If all of the Search Heads down in 3-member cluster

- This will show nothing to users, probably user will get ***"The site can't be reached"*** or ***"502 Bad Gateway"*** error when they try to access the Instances URL or the Load balancer URL

Q8: What happens, If Deployer is down

- You ***cannot push new configurations*** to the members.
- A member that joins or rejoins the cluster, or restarts, ***cannot pull the latest set of apps tar balls***.

Q9: What happens, If Monitoring Console down

- ***No major issues to the environment's data flow***. searches will work normally; indexing will happen as usual.
- Only issue is the admin or whoever was using the Monitoring Console to track & monitor the health of Splunk Environment, ***will lose the visibility***

Q10: What happens, If Deployment Server is down

- Deployment client (say Forwarders) periodically polls the deployment server, to get the checksum of the apps assigned to them. If the Deployment server is down, ***Deployment Clients will not receive the details, which doesn't interrupt the Forwarders functionality*.**
- The issue here is you can't deploy the latest apps/configurations to deployment clients (say forwarders)

Q11: What happens, If Universal Forwarder is down

- The ***Data from the source will not be forwarded*** to the indexer.
- ***As soon as the forwarder comes up, it will continue the streaming of data***

Q12: What happens, If License Master/Server is down

- A license slave communicates their license volume usage to the license master every minute. If a license slave cannot reach the license master for 72 hours or more, the slave is in violation and ***search is blocked***. A violation still ***allows indexing to continue***. Users can not search the slave in violation until the it reconnects with the license master.

Q13: How to Decide number of Search Heads & Indexers

- The ***table available in Splunk Docs*** shows the number of reference machines that you need to index and search data in Splunk Enterprise, ***depending on the number of concurrent users and the amounts of data that the instance indexes.***
- Link:https://docs.splunk.com/Documentation/Splunk/8.0.10/Capacity/Summaryofperformancerecommendations

Q14: How to Choose a Forwarder? (UF or HF)

- **Universal Forwarder:** When the use case requires ***light weight agent*** which just collects & forward the data, ***no additional changes to the data*** is needed, such as masking, filtering, at the source level
- **Heavy Forwarder:** When the ***use case includes normalizing data*** (masking, filtering, event-based routing, etc.) at the source level (e.g., sensitive data such as IP's need to me masked before sending to Indexer)

Q15: Why do we need an Intermediate Forwarder? (IF)

- Intermediate forwarder, will act as a ***consolidation point from multiple forwarders*** and send data to Indexer as a single stream. However, it can also do parsing, if the data is coming from Universal forwarder, which can ***reduce the workload of Indexer***.

Q16: Do we need License for Forwarders?

- The universal forwarder package includes its own license. The license is enabled or applied automatically. This license allows forwarding but not indexing of unlimited data, and also enables security on the forwarder so that users must supply a username and password to access it.
- The heavy forwarder should have access to an Enterprise license stack ***if you plan to perform indexing on the forwarder or to enable authentication on the forwarder***.

Q17: Why can't we use a single Splunk instance with huge size, instead of Search Head, Indexer, Universal Forwarder separately?

- For POC or Small deployment, Single Splunk Instance will be a good choice
- But for large scale of data and the greatest number of users,
  - ***o*** It will be a ***single point of failure***
  - o When it comes to scaling, the 3 functions of Splunk (Input, Index, Search) ***can't be scaled separately as needed***.
- In the real scenario, the Splunk dashboards can't be created & managed, in the data source itself.

Q18: How Splunk Stores Indexes?

In Simple words, Splunk Stores indexes as "Flat Files"

- Indexer creates a number of files to index data:
  1. The raw data in compressed form (the ***rawdata journal***)
  2. Indexes that point to the raw data (***tsidx files***)
  3. Some other ***metadata files***
- Together, these files constitute the Splunk Enterprise index. Reside in sets of directories, or buckets, organized by age.
- The data in each bucket is bounded by a limited time range.

Q19: Is it possible to open the flat files in notepad++?

- ***No,*** they are stored in some encoded format, and ***only readable by Splunk***

Q20: Is it possible to rename the Index? If yes, how? If not, why?

- ***Yes,*** it's possible to rename the index
- Steps to rename the index:
  - Rename the corresponding stanza in *indexes.conf* file
  - Restart the Splunk

Note: This procedure works only on the standalone instance of Splunk, not in the Indexer Cluster

Q21: How to clean the Index in a Standalone Splunk instance & Indexer Cluster?

- In Standalone instance use below command to clean the index

```
/opt/splunk/bin/splunk clean eventdata -index "indexname" -f
```

- In Indexer cluster,
  - Add lower retention policy (say 10 seconds) to the respective index stanza
  - Deploy the bundle to Indexer Cluster
  - Wait for a few mins for the changes to take effect, after a few mins the data will be cleaned. (You can verify the same from the Splunk Web/UI)

Note: This way of cleaning data is preferred only in dev environments. Please try to avoid this in production environment

Q22: How to migrate index from one Splunk server to another Splunk server?

- Migrate the respective stanza of indexes.conf file from the old server to new server
- Migrate the index directory from old server to new server. (Zip it and transfer)
- e.g: the directory for index named ***"test_index"*** will look like below

```
/opt/splunk/var/lib/splunk/test_index/
```

Note: Make sure the Splunk service is stopped & all the hot buckets are rolled to warm, to avoid any data loss.

Q23: How to take backup of Splunk configuration/data? Why should we do that?

- Take the backup of below directories (In Splunk Enterprise installation)

```
/opt/splunk/etc/

/opt/splunk/var/lib/
```

- Take the backup of below directories (In Splunk Universal Forwarder installation)

```
/opt/splunkforwarder/etc/

/opt/splunkforwarder/var/lib/
```

- The reason behind taking the backup is to have a plan for failure scenarios. This backup will help in restoring the configurations & data stored in Splunk

Note: Taking backup will differ based on the Splunk component, depending upon the purpose it's being used. (Detailed discussion will be needed during the architecture planning to come up with the requirement for backup)

Q24: How to Upgrade the Splunk Enterprise?

- The same way we install Splunk for the first time, when you install the Splunk on top of an existing installation, it identifies it as an upgrade and performs the same. Follow below steps:
  - **Step-1:** Stop the Splunk Enterprise
  - **Step-2:** Take a backup of the files
  - **Step-3:** Download latest Splunk Enterprise package
  - **Step-4:** Extract the package on top of the existing installation
  - **Step-5:** Start the Splunk Enterprise

Note: Make sure you read the prompts in between & understand the impact, before you proceed with the upgrade.

Q25: How to Upgrade the Splunk Enterprise, which is Clustered?

- Stop the manager node.
- Stop all the peers and search heads.
- Upgrade the manager node
- Start the manager node

- Enable maintenance mode in manager node
- Upgrade the search heads, followed by the peer nodes.
- Start the peer nodes and search heads
- Disable maintenance mode in manager node

Note: The above-mentioned procedure is when you choose to upgrade all tier at once, this is not the only way to upgrade Splunk Clustered environment. There are other ways, which you can try by referring the Splunk Docs

Q26: How to Upgrade the Splunk Universal Forwarder?

- The same way we install Splunk Universal Forwarder for the first time, when you install the Splunk Universal Forwarder on top of an existing installation, it identifies it as an upgrade and performs the same. Follow below steps:
  - **Step-1:** Stop the Splunk Universal Forwarder
  - **Step-2:** Take a backup of the files
  - **Step-3:** Download latest Splunk Universal Forwarder package
  - **Step-4:** Extract the package on top of the existing installation
  - **Step-5:** Start the Splunk Universal Forwarder

Note: Make sure you read the prompts in between & understand the impact, before you proceed with the upgrade.

Q27: How to Deploy Apps to Search Head Clusters?

- Place the apps in below directory of Deployer

```
/opt/splunk/etc/shcluster/apps/
```

- Push the configuration bundle to Search Heads using below command

```
/opt/splunk apply shcluster-bundle -target <URI>:<management_port> -auth
<username>:<password>
```

Note: shcluster-bundle command has various parameters, based on your use case you can select the parameters, such as preserving lookups, merging local to default, etc,.

Q28: How to Deploy Apps to Indexer Cluster?

- Place the apps in below directory of Cluster Manager/master

```
/opt/splunk/etc/master-apps/
```

- Validate the bundle and check whether a restart is necessary using below command in Cluster Master

```
/opt/splunk/bin splunk validate cluster-bundle --check-restart
```

- Push the configuration bundle to Indexers using below command in cluster master

```
/opt/splunk/bin splunk apply cluster-bundle --answer-yes
```

Q29: How to connect Forwarders to an Indexer Cluster?

- Create outputs.conf with the indexer IP as the target instances as shown below

```
[tcpout]
defaultGroup=my_indexers


[tcpout:my_indexers]
server=mysplunk_indexer1:9997, mysplunk_indexer2:9996


[tcpout-server://mysplunk_indexer1:9997]
```

(or)

- If it's a Heavy Forwarder you can also try Splunk web/UI, go to settings >> Forwarding & Receiving >> Configure receiving >> Add New >> Enter the details of Indexer and save it.

Q30: Difference between Heavy Forwarders & HTTP Event Collector (HEC)?

HEAVY FORWARDER	HTTP EVENT COLLECTOR
Pull Approach Pulls data when new files or updates to existing files are made.	**Push Approach** Listens on a port and you push data to it
Good for **monitoring files or directories** and transforming data	Good for **receiving data directly** from an application.

A **Full Splunk instance** that handles inputs and sends data to a Splunk Indexer	HEC is **a one way (among many)** in which you can bring data to Splunk
In other words, ***HF is a machine / tool and HEC is a function of this machine / tool***	

Q31: Does the Indexer Cluster work, If Cluster Master is down? If yes, then why do we need a Cluster Master?

- ***Yes, Indexer Cluster works***, even if the Cluster Master is down.
- However, Cluster master is responsible for bucket replication & maintaining the availability of Cluster, so as soon as possible, the cluster master has to be up.
- If the cluster master is down, the newly created hot buckets will not get the target peers from the cluster master, it will use the previous target peers, which ***can unbalance the cluster*** in terms of buckets

Q32: What is meant by Colocation of Splunk Components? Why & when should we do that?

- ***Configuring more than one management component in a single Splunk instance*** is called colocation of Splunk components. (e.g License Master & Monitoring Console can be configured in Single Splunk Instance)
- You might be able to combine two management components on a single Splunk Enterprise instance. In some cases, you can locate a management component on an instance with a processing component.
- In some low-use situations, you might be able to combine more than two management components in a single instance, although it is not generally recommended that you do so. If you intend to do so, monitor the performance impacts closely to ensure that you are not overloading the instance.

Note: Refer the table below (from Splunk Docs)

Management component	Colocate with						
	LM ?	MC?	DS?	CM?	Deployer?	Indexer?	Search head?
License master	-	yes	yes	yes	yes	yes	yes

Monitoring console	yes	-	yes	yes	yes	no	yes
Deployment server	yes	yes	-	no	yes	yes	yes
Indexer cluster manager node	yes	yes	no	-	yes	no	no
Search head cluster deployer	yes	yes	yes	yes	-	no	no

Key: "**LM**" = license master; "**MC**" = monitoring console; "**DS**" = deployment server; "**CM**" = cluster manager.

Q33: Is it possible to use Deployment Server to distribute apps to Search Head Cluster & Indexer Cluster?

- **You cannot use** the deployment server to update indexer cluster peer nodes or search head cluster members **directly**.
- **Indexer Cluster:**
  - Do not use deployment server or forwarder management to manage configuration files across peer nodes (indexers) in an indexer cluster. Instead, use the configuration bundle method.
  - You can, however, use the deployment server to distribute updates to the master node, which then uses the configuration bundle method to distribute them to the peer nodes.
- **Search head clusters:**
  - Do not use deployment servers to update search head cluster members.
  - The deployment server is not supported as a means to distribute configurations or apps to cluster members. To distribute configurations across the set of members, you must use the search head cluster deployer.

Q34: How to reduce license usage in Splunk? How & where the license usage is being calculated?

- To reduce license usage, ***filter the unwanted data*** in the Input phase itself & index only required 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.

Q35: Why do we need a License Master/Server? Why can't we install licenses on each Splunk Server directly?

- If you have multiple Splunk Enterprise instances, you'll want to ***manage their license access from a central location***.
- Installing license on each Splunk Server directly, will require manual effort by login in to each server/uploading it to each server. (Every time when license expires you will have to update all the servers)

Q36: Can the Replication Factor be lower than the Search factor? why?

- ***No.*** Replication Factor can't be lower than Search Factor
- Search Factor talks about how many searchable copies out of the replicated copies (which is decided by Replication factor). So, it can't be like 4 searchable copies out of 3 replicated copies. ***(This is not realistic)***

Q37: How to create Index in indexer Cluster?

- Step-1: Create a stanza in indexes.conf file of your app
- Step-2: Place the app in the below location of Cluster Master

```
/opt/splunk/etc/master-apps/
```

- Step-2: Place the app in the below location of Cluster Master

```
/opt/splunk/etc/master-apps/
```

- Step-3: Validate the bundle and check whether a restart is necessary using below command in Cluster Master

```
/opt/splunk/bin splunk validate cluster-bundle --check-restart
```

- Step-4: Push the configuration bundle to Indexers using below command in cluster master

```
/opt/splunk/bin splunk apply cluster-bundle --answer-yes
```

- Step-5: Login to any one of Indexer & check the indexes list. Settings>>Indexes

Q38: What is meant by Retention Policy? and where it's been used in Splunk?

- Data retention policies are sets of rules that ***determine how long data remains available*** for consumption from a message queue.
- Typically, streamed data is retained in an available state until a specified interval of time has passed or a maximum amount of retained data is exceeded.
- It is defined in indexes.conf as shown below

```
[default]
maxWarmDBCount = 200
frozenTimePeriodInSecs = 432000
rotatePeriodInSecs = 30
```

Q39: What happens, if the network connectivity between the Forwarder & Indexer is lost?

- Universal Forwarder will not be able to connect to the indexer, due to which the ***data flow will be stopped/will be on hold***. (Some of the data blocks will be ***stored in wait queue***, once the wait queue is full forwarder stops reading the data)

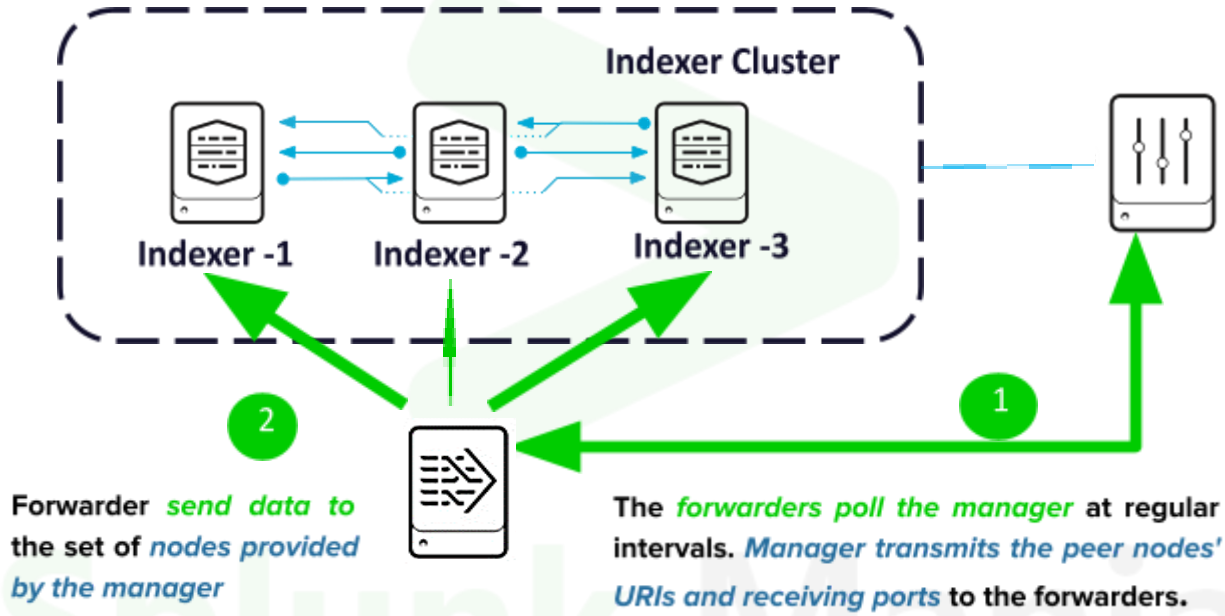
Q40: How to make sure if all data forwarded by Forwarder is received by the Indexer?

- Enable ***Load Balancing*** in outputs.conf of the forwarder
- Enable ***Indexer Acknowledgement (useACK=true)*** in outputs.conf of the forwarder
    - The forwarder sends data continuously to the indexer, in blocks of approximately 64kB. The forwarder maintains a copy of each block in memory, in its wait queue, until it gets an acknowledgment from the indexer. While waiting, it continues to send more data blocks.
    - If all goes well, the indexer:
        1. Receives the block of data.
        2. Parses the data.
        3. Writes the data to the file system as events (raw data and index data).
        4. Send an acknowledgment to the forwarder.
    - The acknowledgment tells the forwarder that the indexer received the data and successfully wrote it to the file system. Upon receiving the acknowledgment, the forwarder releases the block from memory.

Q41: How to avoid overloading of Indexers, while sending the data from Forwarders?

- ***With load balancing configuration***, a forwarder distributes data across several receiving instances. Each receiver gets a portion of the total data, and together the receivers hold all the data.

Q42: What is the Indexer Discovery Method?



Q43: What happens when you remove the stanza from indexes.conf & restart the Splunk?

- ***Index name will disappear*** from the list (in UI, Settings>>Indexes)
- ***The indexed data will still reside in the directory*** where it was indexed. But the ***data will not be searchable.***
- If we create stanza again, for the same index & restart the Splunk, then the data will become searchable.

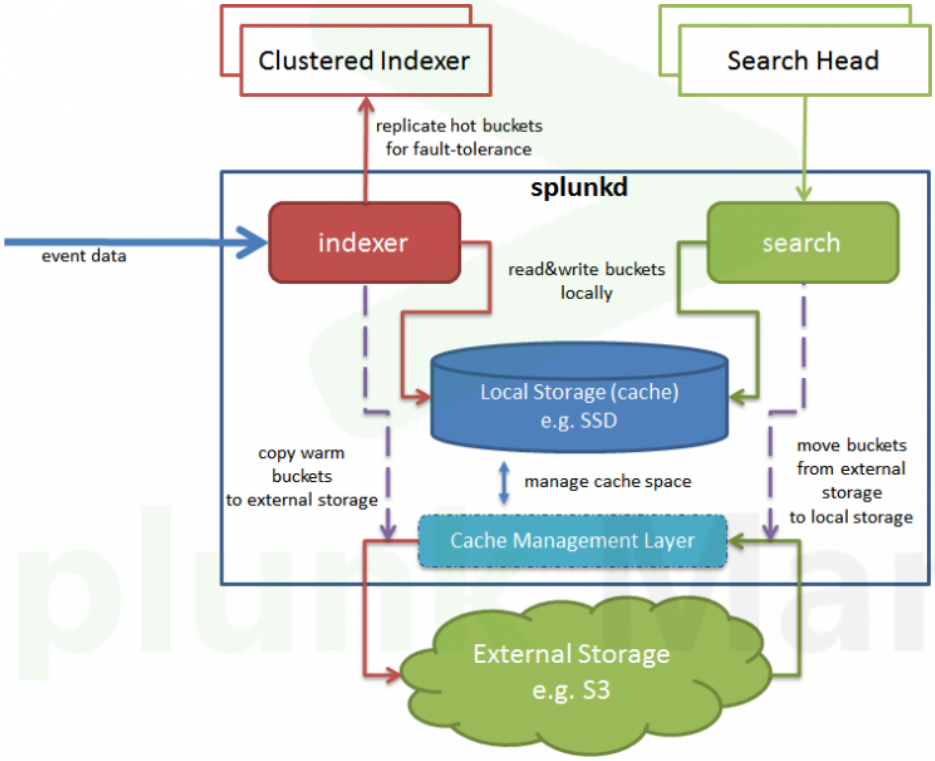
Q44: What happens when the Disk is full? Index is full?

- If the Disk is full, ***indexing pipeline will be stopped/hold***
- If the Index is full, ***the oldest/earliest event will be removed***, as per the retention policy mentioned in indexes.conf.

Q45: What is SmartStore in Splunk?

- SmartStore is an indexer capability that provides a way to use remote object stores, such as Amazon S3 or Google GCS, to store indexed data.

- As a deployment's data volume increases, demand for storage typically outpaces demand for compute resources. SmartStore allows you to manage your indexer storage and compute resources in a cost-effective manner by scaling those resources separately.



Q46: Does Splunk Support SSO/SAML? How to configure them?

- ***Yes. Splunk Does supports SSO/SAML***
- Detailed steps given in Splunk Docs
  - https://docs.splunk.com/Documentation/Splunk/8.2.6/Security/SAMLSHC
  - https://docs.splunk.com/Documentation/Splunk/8.2.6/Security/SAMLSSObestpractices
  - https://docs.splunk.com/Documentation/Splunk/8.2.6/Security/ConfigureSAMLSSO

Q47: How to enable SSL in Splunk? Is it possible to use custom SSL Certificates?

- You can turn on default encryption in Splunk Web using the ***default certificates*** that come with the installation, but ***this does not represent a high level of security,*** since every installation provides those certificates.
    - ***o*** In Splunk Web, select ***Settings > System > Server settings***, and then click ***General Settings.***
    - o Under Splunk Web, for ***Enable SSL (HTTPS) in Splunk Web?*** select the ***Yes*** radio button.
    - o ***Restart Splunk Web***.
    - o After Splunk Web restarts, to access Splunk Web on that instance, ***use "https://<your site name>:<port>" for the URL***.
- To ***use a custom SSL certificate***, follow the step-by-step procedure mentioned in Splunk Document.
- Link: https://docs.splunk.com/Documentation/Splunk/8.2.6/Security/SecureSplunkWebusingasignedcertificate

Q48: How to restrict users to specific index/data?

- ***Create a role with restricted*** access and only assign this role to user
    - o Step-1: Create a role which has access to the specific index. e.g: ***"role_alpha"*** has access to only the ***"alpha_idx"*** index.
    - o Step-2: Assign only this role to the respective user (remove any other roles from the user)

Q49: How to collect logs from AWS? or how to integrate Splunk with AWS?

- Use the ***Splunk Add-on for Amazon Web Services (AWS)*** to collect data on Amazon Web Services. The Splunk Add-on for AWS ***offers pretested add-on inputs for four main use cases (CloudTrail log, performance, billing, and IT and security data)***, but you can create an input manually for a miscellaneous Amazon Web Service.
- App Link: https://splunkbase.splunk.com/app/1876/
- Procedure Document Link: http://docs.splunk.com/Documentation/AddOns/latest/AWS/Description

Q50: How to connect/integrate Databases with Splunk?

- ***Splunk DB Connect*** enables you to combine your structured ***data from databases*** with your unstructured machine data, and then use Splunk Enterprise to provide insights into all of that combined data.

- When you use Splunk DB Connect, you are creating ***additional data inputs*** for Splunk Enterprise. That is, you're giving Splunk Enterprise more sources of data to consume. ***Splunk DB Connect is what connects your relational database*** data to Splunk Enterprise and makes that data consumable by Splunk Enterprise.

- In addition, ***Splunk DB Connect can do the reverse—write Splunk Enterprise data back to your relational database***
- App Link: https://splunkbase.splunk.com/app/2686/
- Procedure Document Link: https://docs.splunk.com/Documentation/DBX/3.8.0/DeployDBX/AboutSplunkDBConnect

Q51: How to integrate ServiceNow with Splunk?

- The ***Splunk Add-on for ServiceNow*** allows a Splunk software administrator to use ServiceNow REST APIs to collect the following types of data:
    - Incident data
    - Event data
    - Change data
    - User data
    - User group data
    - Location data
    - Configuration management database (CMDB) configuration item (CI) data
- After you install and configure this add-on, you can use workflow actions that link directly from events in the Splunk platform search results to relevant ServiceNow incidents, events, and knowledge base articles. You can also use the custom commands, alert actions, and scripts to create new incidents and events in your ServiceNow instance, and update the incidents created from the Splunk platform.
- App Link: https://splunkbase.splunk.com/app/1928/
- Procedure Document Link: https://docs.splunk.com/Documentation/AddOns/released/ServiceNow/About

Q52: How to get logs from Windows (or) Linux?

- The ***Splunk Add-on for Unix and Linux*** collects *nix data from *nix hosts. You can install the Splunk Add-on for Unix and Linux on a forwarder to send data from any number of *nix hosts a Splunk Enterprise indexer or group of indexers. You can also use the add-on to provide data for other apps, such as Splunk IT Service Intelligence or Splunk Enterprise Security.
- The Splunk Add-on for Unix and Linux collects the following data using file inputs:
    - Changes to files in the ***/etc*** directory and subdirectories.
    - Changes to files in the ***/var/log*** directory and subdirectories.
- App Link: https://splunkbase.splunk.com/app/833

- Procedure Document Link: https://docs.splunk.com/Documentation/UnixAddOn/6.0.0/User/AbouttheSplunkAdd-onforUnixandLinux

- The ***Splunk Add-on for Windows*** allows a Splunk software administrator to collect:
    - CPU, disk, I/O, memory, log, configuration, and user data with data inputs.
    - Active Directory and Domain Name Server debug logs from Windows hosts that act as domain controllers for a supported version of a Windows Server. You must configure Active Directory audit policy since Active Directory does not log certain events by default.
    - Domain Name Server debug logs from Windows hosts that run a Windows DNS Server. Windows DNS Server does not log certain events by default, and you must enable debug logging.
- App Link: https://splunkbase.splunk.com/app/742
- Procedure Document Link: https://docs.splunk.com/Documentation/WindowsAddOn/8.1.2/User/AbouttheSplunkAdd-onforWindows

Q53: How to Anonymize/ Normalize/ Mask data in Splunk?

There are two ways to anonymize data with a heavy forwarder:

- ***Use the SEDCMD setting***. This setting exists in the props.conf configuration file, which you configure on the heavy forwarder. It acts like a sed *nix script to do replacements and substitutions. This method is more straightforward, takes less time to configure, and is slightly faster than a regular expression transform. But there are limits to how many times you can invoke the SEDCMD setting and what it can do.

- ***Use a regular expression (regex) transform.*** This method takes longer to configure, but less complex to modify after the initial configuration. You can also assign this method to multiple data inputs more flexibly.

Both of these options are also available in Splunk Enterprise, where you can complete the configuration on either a heavy forwarder or an indexer.

For more details:
https://docs.splunk.com/Documentation/Splunk/8.2.6/Data/Anonymizedata#Anonymize_data

Q54: How to index the same data into 2 different indexers?

On the instance that is to do the routing, open a command or shell prompt.
Edit ***$SPLUNK_HOME/etc/system/local/props.conf*** to set two TRANSFORMS-routing settings: one for syslog data and a default for all other data.

```
[default]
TRANSFORMS-routing=errorRouting

[syslog]
TRANSFORMS-routing=syslogRouting
```

Edit $SPLUNK_HOME/etc/system/local/transforms.conf to set the routing rules for each routing transform.

```
[errorRouting]
REGEX=error
DEST_KEY=_TCP_ROUTING
FORMAT=errorGroup

[syslogRouting]
REGEX=.
DEST_KEY=_TCP_ROUTING
FORMAT=syslogGroup
```

In this example, if a syslog event contains the word "error", it routes to syslogGroup, not errorGroup. This is due to the settings you previously specified in props.conf. Those settings dictated that all syslog events should be filtered through the syslogRouting transform, while all non-syslog (default) events should be filtered through the errorRouting transform. Therefore, only non-syslog events get inspected for errors. Edit ***$SPLUNK_HOME/etc/system/local/outputs.conf*** to define the target groups.

```
[tcpout]
defaultGroup=everythingElseGroup

[tcpout:syslogGroup]
server=10.1.1.197:9996, 10.1.1.198:9997
```

```
[tcpout:errorGroup]
server=10.1.1.200:9999

[tcpout:everythingElseGroup]
server=10.1.1.250:6666
```

syslogGroup and errorGroup receive events according to the rules specified in transforms.conf. All other events get routed to the default group, everythingElseGroup.

Q55: Other than Forwarders, what are all the ways the data can be onboarded to Splunk?

- HTTP Event Collector
- Manual Upload through “Add Data” option from UI
- Custom streaming commands
- Summary indexing
- Splunk DB Connect, etc.

Q56: How to upgrade Splunk Apps?

- Stop the Splunk
- Take the backup of the respective Splunk App
- Extract/Place the latest App on top of the existing installation
- Restart the Splunk

Note: These steps are meant for Standalone Splunk instance, not for the clustered deployments

Q57: How to install Apps from Splunk base?

- ***Method-1***
  - Step-1: In Splunk UI/Web, Go to Manage Apps, click on “Find more apps”
  - Step-2: Find out required app, with help of filters
  - Step-3: click on “Install” button, (This will ask splunk.com credentials)
  - Step-4: Once the app is installed, restart the Splunk (If applicable)

- ***Method-2***
  - Step-1: Download the required app from Splunkbase to your laptop
  - Step-2: In Splunk UI/Web, Go to Manage Apps, click on “Install app from file” & upload the app (if the app already present, select the checkbox which says “upgrade this app, if already exists”)
  - Step-3: click on “Install” button
  - Step-4: Once the app is installed, restart the Splunk (If applicable).

Q58: How to index log files into Splunk?

- **Method-1:** Upload from Splunk UI using "Settings >> Add Data" option.
- **Method-2:** Create stanza in ***inputs.conf*** file as shown below (e.g., to index /var/log/messages log file)

```
[monitor:///var/log/messages]
disabled = 0
index = test_index
```

- **Method-3:** Execute CLI command as shown below (e.g., to index /var/log/messages log file)

```
/opt/splunk/bin/splunk add monitor /var/log/messages
```

Q59: How to index custom CSV files into Splunk?

- **Method-1:** Upload from Splunk UI using "Settings >> Add Data" option.
- **Method-2:** Use inputs.conf & props.conf file
  - *Step-1: Define a new sourcetype **in props.conf,*** by creating a stanza which tells Splunk how to extract the file header and structured file data. e.g.

```
[structuredCSVDataProps]
FIELD_DELIMITER=,
HEADER_FIELD_DELIMITER=\s
FIELD_QUOTE="
```

  - *Step-2: Create stanza **in inputs.conf**,* as shown below (e.g., to index *sample_data.csv* file)

```
[monitor:///opt/test/data/StructuredData/sample_data.csv]
sourcetype=structuredCSVDataProps
```

- **Method-3:** Execute CLI command as shown below (e.g., to index *sample_data.csv* file) (Assuming the sourcetype "**structuredCSVDataProps**" is already created)

```
/opt/splunk/bin/splunk add monitor /
opt/test/data/StructuredData/sample_data.csv -sourcetype
structuredCSVDataProps
```

Q60: How index Json files into Splunk?

- **Method-1:** Upload from Splunk UI using "Settings >> Add Data" option.
- **Method-2:** Create stanza in ***inputs.conf*** file as shown below (e.g., to index /tmp/sample_data.json log file)

```
[monitor:///tmp/sample_data.json]
disabled = 0
index = test_index
sourcetype = _json
```

- **Method-3:** Execute CLI command as shown below (e.g., to index /tmp/sample_data.json log file)

```
/opt/splunk/bin/splunk add monitor /tmp/sample_data.json -sourcetype
_json
```

Q61: How to connect UF with Indexer?

- **Method-1:** Create ***outputs.conf*** file as shown below (e.g., to connect to the indexers with IP – 10.128.72.25 & 10.128.72.26)

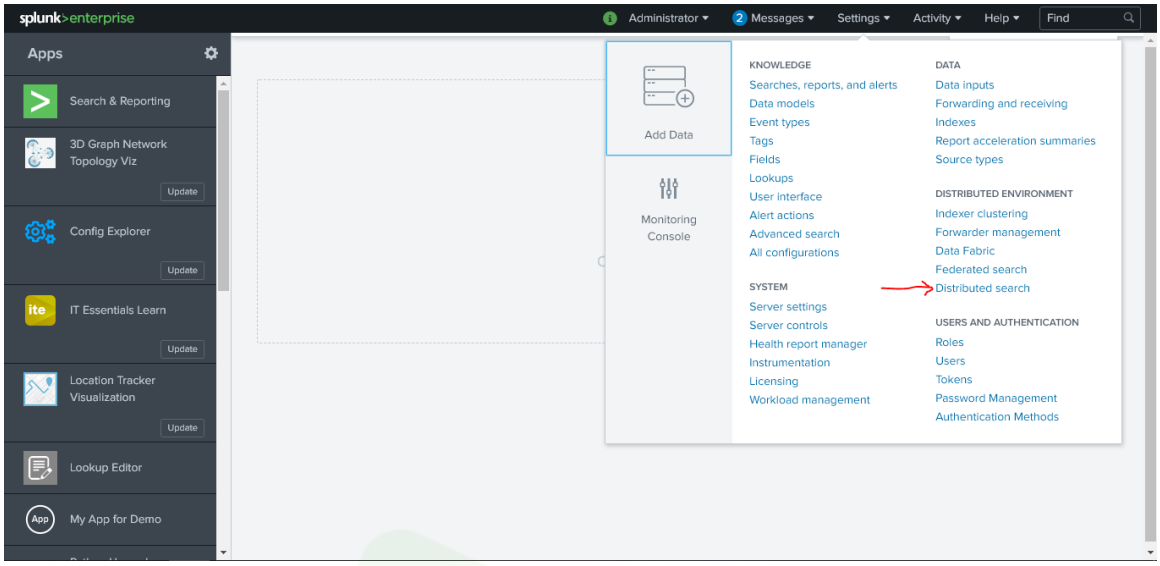
```
[tcpout]
defaultGroup=myIndexerGroup

[tcpout:myIndexerGroup]
server=10.128.72.25:9997, 10.128.72.26:9997

[tcpout-server://10.128.72.25:9997]

[tcpout-server://10.128.72.26:9997]
```

- **Method-3:** Execute CLI command as shown below (e.g., to connect to the indexers with IP – 10.128.72.25 & 10.128.72.26)

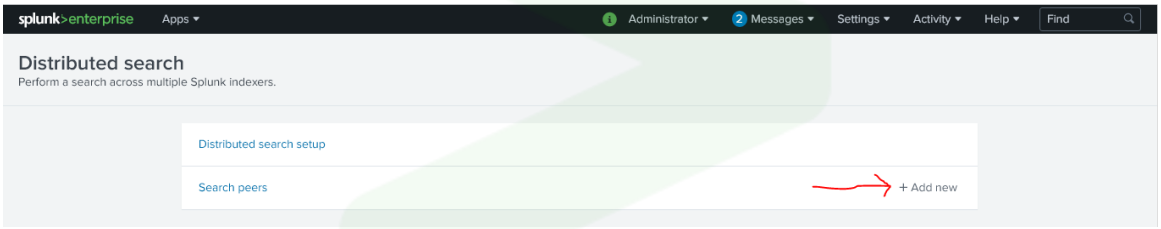
```
/opt/splunkforwarder/bin/splunk add forward-server 10.128.72.25:9997
```

Q62: How to connect Search Head with Indexer?

- Login to Splunk Head node
- Under Settings, select “Distributed search”, as shown below

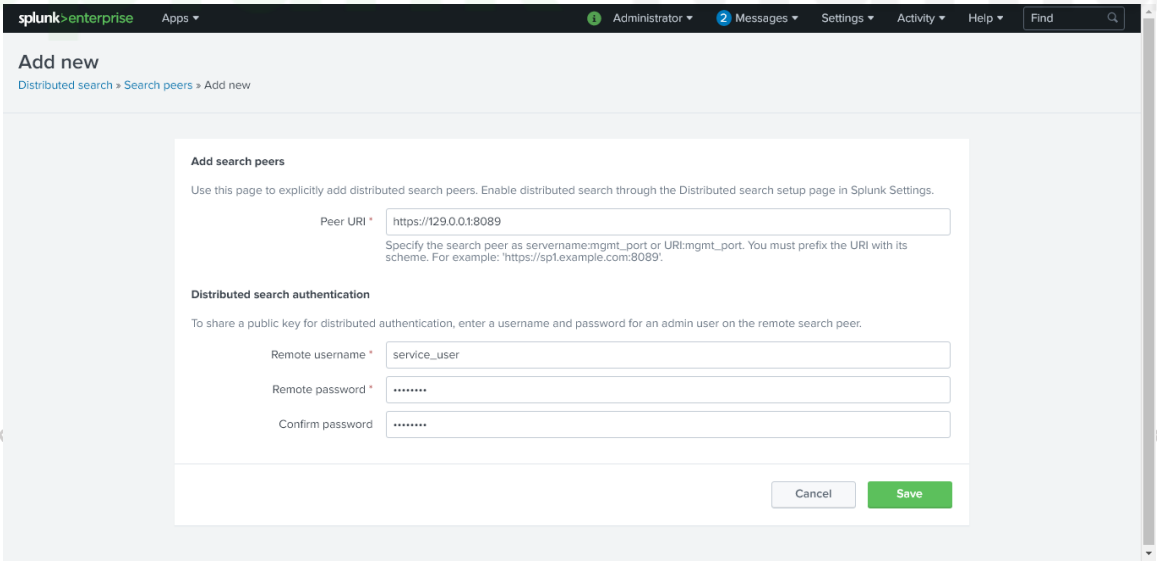


- Then, under “Distributed search”>> “Search peers” >> Click “+ Add new”



- On the next screen, enter the details such as Peer URI (this refers to Indexer URI), and the authentication details (Service account created in Step-2), then click on “Save” button as shown below

- Once the details are added, it will be listed under “Search Peers” page

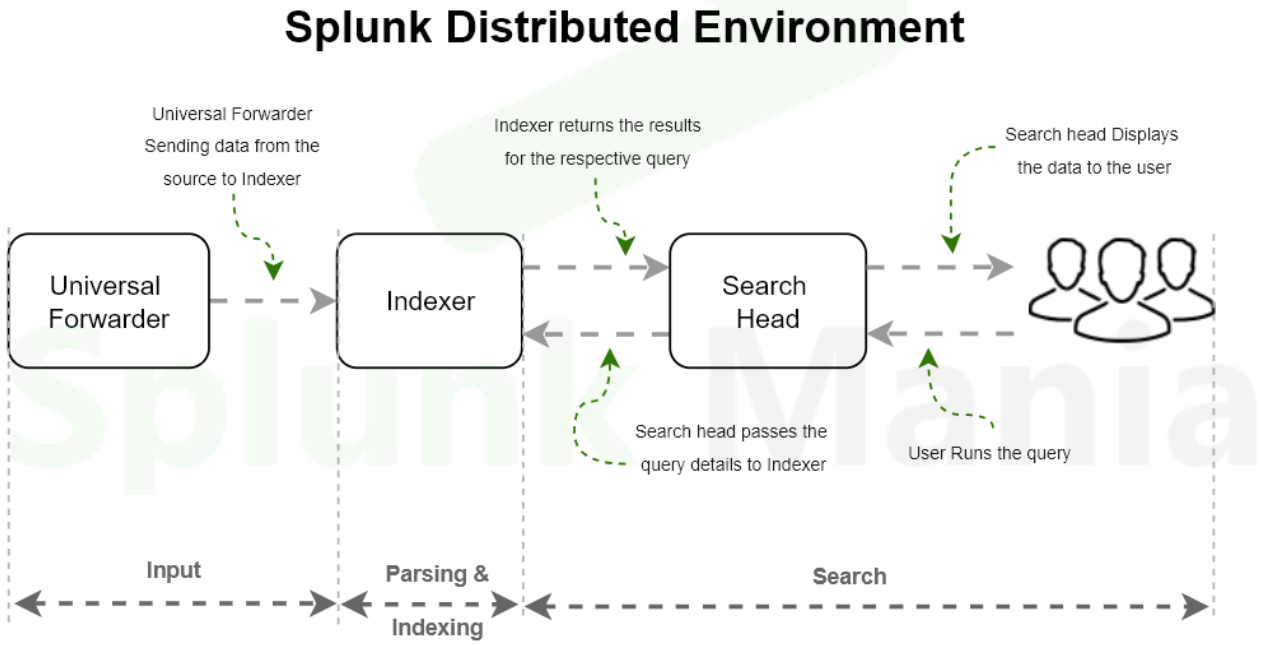


- After completing all above steps, go to Search head, open “Search & Reporting” app, then run below query to see the results from all 3 Splunk nodes (Heavy Forwarder, indexer & Search Head) with Time Range of “Last 15 minutes”

  Index=_internal

- The results of the above query will have data from 3 hosts. This shows the connection between all 3 Splunk nodes (Heavy Forwarder, indexer & Search Head) are live.

Q63: What happens in a distributed environment, under the hood/behind the screen, when a user runs a query in Search Head?



Q64: Can Search Head store/index the data?

- ***By default, yes***. But as a best practice, ***forwarding the data from Search Head to Indexer is preferred***, this will make sure the data is available in index tier alone, thus improving the Search Head performance.

Q65: Can Forwarders store/index the data?

- ***By default, No***. as the forwarders need not to have the data stored locally. But if it’s needed (storing data locally on Forwarders, ***it can be enabled using the property in outputs.conf as*** shown below)

```
[tcpout]
defaultGroup=myIndexerGroup
indexAndForward = true
```

- ***"indexAndForward = true"*** means, the data will be stored in Forwarder also, while the same being sent to indexer

Q66: What is the location of Splunk index?

```
/opt/splunk/var/lib/splunk/

(or)

$SPLUNK_HOME/var/lib/splunk/
```

Q67: How Forwarders avoid duplicates? up to what level?

- ***Using "fishbucket",*** forwarders are avoiding duplicates. The fishbucket ***keeps track of the files being monitored, like a checkpoint***. if the file's older content is not modified & it has additional content at the end of the file, then forwarder will take only the additional content, with reference to fishbucket

Q68: How to re-index data in Splunk?

- **Method-1:** Delete the fishbucket directory of Forwarder

```
/opt/splunkforwarder/var/lib/splunk/fishbucket/
```

  - This method will re-index whole data which forwarder indexed earlier, this method is not preferred in production
- **Method-2:** Reset the fishbucket for the specific file in the Forwarder

```
<Answer here> $SPLUNK_HOME/bin/splunk cmd $SPLUNK_HOME/bin/btprobe -d
$SPLUNK_HOME/var/lib/splunk/fishbucket/splunk_private_db --file
/path/to/file.log --reset
```

- **Method-3:** Adding crcSalt to indexes.conf as shown below

```
[monitor:///var/log/messages]
disabled = 0
index = test_index
crcSalt = sometext
```

Q69: How does fishbucket work?

- When the file monitor processor looks at a file, it searches the fishbucket to see if the ***CRC from the beginning*** of the file is already there.
- If not, the file is indexed as new.
- If yes, then we check the CRC of where we were reading against the saved value in ***seekcrc***.
- If it matches and the file is longer than the saved ***seek pointer***, then there is new stuff at the end to read.
- If the top of the file matches but the seekcrc doesn't, or the seek pointer is beyond the current end of the file, then something in the part we have already read has changed. Since we don't know what might have changed, we just index the whole thing.

Q70: Difference between Modular & Scripted Inputs?

FEATURE	MODULAR INPUTS	SCRIPTED INPUTS
Checkpointing	***Supported***	***Supported,*** but requires manual implementation
Configuration	Parameters defined in inputs.conf. Users can configure inputs using Splunk Web input **Settings** fields. Validation support	Inline arguments Separate configuration outside of the Splunk platform
REST management API	***Supported.*** Access modular inputs using REST	***Not supported***
Event boundaries	***Supported.*** XML streaming simplifies specifying event boundaries.	Available, but ***requires additional script complexity***
FEATURE	MODULAR INPUTS	SCRIPTED INPUTS
Multi-platform support	***Supported.*** Developers can package a modular input script to include versions for separate platforms.	***Not supported***

REST endpoint authorization	***Supported.*** Use capabilities to control access.	***n/a***
Run as user	***Not supported.*** All modular input scripts are run as the user using the Splunk instance.	***Supported.*** You can specify which user can run the script
Single instance mode	***Supported***	***Supported,*** but requires manual implementation

Q71: How to Integrate Jenkins with Splunk?

- The ***Splunk App for Jenkins*** makes it possible to use Splunk solutions to collect, monitor and analyze the wealth of Jenkins data.
- Since the engineering teams started using the Splunk App for Jenkins to gain insight into ***Jenkins health, builds and test data***, they were able to develop and test at greater velocity while ensuring the high quality of Splunk software. They were able to ***reduce build certification times from weeks to just a few minutes***.
- App Link: https://splunkbase.splunk.com/app/3332
- Procedure:

Step 1. Splunk App Installation

- ***o*** Install on a single instance
  1. Download the add-on from Splunkbase.
  2. From the Splunk Web home screen, click the gear icon next to Apps.
  3. Click Install app from file.
  4. Locate the downloaded file and click Upload.
  5. Restart Splunk
- o Install on a search head cluster

  Use Splunk Web to deploy it on search headers or use deployer to install it in a search header cluster.

  http://docs.splunk.com/Documentation/AddOns/latest/Overview/Distributedinstall
- o Install on an indexer cluster

  copy indexes.conf and props.conf from this App's default folder to ***$SPLUNK_HOME/etc/master-apps/_cluster/local/*** on cluster master, then run

```
/opt/splunk/bin splunk apply cluster-bundle
```

Link: http://docs.splunk.com/Documentation/Splunk/6.5.3/Indexer/Updatepeerconfigurations

Step 2. Splunk Jenkins Plugin Installation and Configuration

- Splunk App gets all its data from Splunk's plugin for Jenkins.
- The Jenkins plugin can be downloaded directly in the Jenkins update center.
- Detailed plugin configuration guide can be found here https://wiki.jenkins-ci.org/display/JENKINS/Splunk+Plugin+for+Jenkins

Q72: How to create a Search Head Cluster?

- Step-1: Enable the deployer with pass4symmkey
- Step-2: initiate the search head cluster configuration in all search head members
- Step-3: Bootstrap the captain election (execute the CLI command to trigger captain election) Once the captain election is elected, the search head cluster is formed perfectly
- Step-4: Connect search head cluster with Indexer Cluster

Q73: How to create an Indexer Cluster?

- **Step-1:** Enable Cluster master node
- **Step-2:** Add all the search peers (indexers) to the cluster
- **Step-3:** Once the cluster gets the replication factor number of peers, up and running, the indexer cluster will become valid & complete state

Q74: How to connect Search Head cluster with Indexer Cluster?

- To integrate a search head cluster with an indexer cluster, ***configure each member of the search head cluster as a search head on the indexer cluster***. Once you do that, the search heads get their list of search peers from the manager node of the indexer cluster.
- Configure each search head cluster member as a search head on the indexer cluster. Use the CLI ***splunk edit cluster-config*** command. Also restart the Search head, for example:

```
/opt/splunk/bin/splunk edit cluster-config -mode searchhead -master_uri
https://10.152.31.202:8089 -secret newsecret123 -auth login:password


/opt/splunk/bin/splunk restart
```

Q75: How to Configure Monitoring Console?

- **In Standalone Deployment:**
  - ***Leave the default configuration as it is,*** by default the monitoring is enabled for standalone deployments
- **In Distributed Environment:**

- Log into the instance on which you want to configure the monitoring console. The instance by default is in standalone mode, unconfigured.
- In Splunk Web, select ***Monitoring Console > Settings > General Setup***.
- Click Distributed mode.
- Confirm the following:
  - ✔ The columns labeled instance and machine are populated correctly and show unique values within each column.
  - ✔ The server roles are correct. For example, a search head that is also a license master must have both server roles listed. If not, click ***Edit > Edit Server Roles and select the correct server roles for the instance***.
  - ✔ If you are using indexer clustering, make sure the cluster master instance is set to the cluster master server role. If not, click ***Edit > Edit Server Roles and select the correct server role***.
- Click Apply Changes.

Q76: Difference between Deployment Server & Deployer

DEPLOYMENT SERVER	DEPLOYER
Distributing configurations, apps, and content updates to groups of Splunk Enterprise instances	Distribute apps and certain other configuration updates to search head cluster members
The set of updates that the deployment server distributes is called the configuration files and apps, to deployment clients.	The set of updates that the deployer distributes is called the configuration bundle, to search head cluster members
Mainly used for Forwarder	Mainly used for Search Head Cluster Members
You cannot use the deployment server to update indexer cluster peer nodes or search head cluster members.	

Happy Splunking…!!

*Any help/support required on the Splunk, please contact the Splunk **Mania Team** using any one of the methods mentioned at the end of this document.*

References

https://docs.splunk.com/Documentation/Splunk

Contact:

WhatsApp : +919345372209

Email : splunkmania@gmail.com

LinkedIn : https://www.linkedin.com/company/splunk-mania

Facebook : https://www.facebook.com/SplunkMania

Instagram : Splunk Mania (@splunkmania) • Instagram photos and videos

Slack : https://splunkmania.slack.com/

YouTube : Splunk Mania - YouTube

Website : Splunk Mania