What are the primary steps involved in deploying a Splunk search head cluster?

The deployment process includes:Splunk Documentation

- **Identifying requirements**: Determine cluster size and replication factor.
- **Setting up the deployer**: Choose and configure a separate Splunk Enterprise instance to distribute apps and configurations.
- **Installing Splunk Enterprise instances**: Install on machines designated as cluster members.
- **Initializing cluster members**: Configure each member with necessary settings.
- **Bringing up the cluster captain**: One member becomes the captain to coordinate activities.
- **Performing post-deployment setup**: Connect to search peers, set up load balancers, and verify the cluster's health. Splunk Documentation+6Splunk Documentation+6Splunk Documentation+6Splunk Documentation+11Splunk Documentation+11Splunk Documentation+11Splunk Documentation+4Splunk Documentation+4Splunk Documentation+4Splunk Documentation+8Splunk Documentation+8Splunk Documentation+8Splunk Documentation+7Splunk Documentation+7Splunk Documentation+7

---

2. What is the role of the deployer in a search head cluster?

The deployer is a Splunk Enterprise instance used to distribute apps and certain configuration updates to search head cluster members. It ensures consistency across the cluster by pushing configuration bundles to all members. The deployer must reside outside the cluster and cannot be a cluster member. Splunk Documentation+11Splunk Documentation+11Splunk Documentation+11

---

3. Why is a replication factor important in a search head cluster, and how is it determined?

The replication factor determines the number of copies of search artifacts the cluster maintains. A higher replication factor enhances fault tolerance by ensuring that if one member fails, replicated data is available on other members. The optimal replication factor balances between desired fault tolerance and storage capacity. Splunk Documentation+2Splunk Documentation+2Splunk Documentation+2

---

4. Can a single-member search head cluster be deployed, and what are its implications?

Yes, a single-member search head cluster can be deployed, primarily for testing or initial setup with plans to scale later. However, it doesn't provide high availability. When scaling, it's recommended to expand directly to at least three members to ensure stability and proper captain election. Splunk Documentation+4Splunk Documentation+4Splunk Documentation+4

---

5. What are the system requirements for search head cluster members?

Each member must:Splunk Documentation+2Splunk Documentation+2Splunk Documentation+2

- Run on its own machine or virtual machine.
- Use the same operating system and version as other members.
- Run the same version of Splunk Enterprise.
- Be connected over a high-speed network.
- Have sufficient storage to accommodate replicated search artifacts, considering the replication factor. Splunk Documentation+11Splunk Documentation+11Splunk Documentation+11

---

6. How does the captain function within a search head cluster?

The captain is a designated member that coordinates cluster-wide activities, including job scheduling and replication. If the captain fails, another member is automatically elected to take over, ensuring continuous operation without a single point of failure. Splunk Documentation+2Splunk Documentation+2Splunk Documentation+2

---

7. What configurations are managed by the deployer versus those replicated automatically?

The deployer manages:Splunk Documentation+3Splunk Documentation+3Splunk Documentation+3

- New or upgraded apps.
- Configuration files edited directly, such as indexes.conf or inputs.conf.
- Non-search-related updates.Splunk Documentation+12Splunk Documentation+12Splunk Documentation+12

In contrast, the cluster automatically replicates runtime changes to knowledge objects, like saved searches or dashboards, across all members. Splunk Documentation+10Splunk Documentation+10Splunk Documentation+10

---

8. Why is it crucial to use the deployer for app distribution in a search head cluster?

Using the deployer ensures that all cluster members receive consistent app configurations. Directly installing apps on individual members can lead to inconsistencies and conflicts, especially during cluster expansions or when members rejoin after downtime. Splunk Documentation+4Splunk Documentation+4Splunk Documentation+4

---

9. What is the purpose of the pass4SymmKey in a search head cluster?

The pass4SymmKey is a shared secret key used for authentication between the deployer and cluster members. It ensures secure communication and must be identical across all members and the deployer. A mismatch can prevent successful deployment of configuration bundles. Splunk Documentation+1Splunk Documentation+1

---

10. How does a load balancer enhance a search head cluster deployment?

A load balancer distributes user search requests across multiple search head cluster members, optimizing resource utilization and ensuring high availability. It provides a single point of access for users, abstracting the underlying cluster complexity. Splunk Documentation+3Splunk Documentation+3Splunk Documentation+3

---

11. What is the function of search artifact replication in a search head cluster?

Search artifact replication ensures that search-related data (like search jobs, reports, and user artifacts) are available across multiple members of the cluster. This replication supports failover and high availability—if one member goes down, others still have access

to the artifacts needed to continue operations. The replication process is managed automatically by the captain.

---

12. How does a member join an existing search head cluster?

To join an existing cluster, a member must:

- Be installed and configured with the correct server.conf and shcluster.conf settings.
- Use the splunk init shcluster-config command to initialize the member.
- Run the splunk bootstrap shcluster-captain or use splunk add shcluster-member from another member (usually the captain) to add the new member.
- Share the same pass4SymmKey, replication port, and cluster label as other members.

---

13. What happens if two members are mistakenly initialized with the bootstrap command?

Using splunk bootstrap shcluster-captain on more than one member creates a cluster split-brain condition. This results in two captains and inconsistent state replication. Splunk recommends bootstrapping only one member to avoid this. If a split occurs, you must manually reset and rejoin the cluster properly.

---

14. Can you upgrade Splunk Enterprise in a search head cluster without downtime? How?

Yes, by performing a rolling upgrade:

- Upgrade non-captain members one at a time.
- Let them rejoin the cluster.
- Upgrade the captain last (or force a captain handover to upgrade it).
- This method ensures continuous availability, although some transient inconsistencies in replicated data might occur during the process.

---

15. What are common challenges or misconfigurations during SHC deployment?

Common issues include:

- Mismatched pass4SymmKey between deployer and members.
- Incorrect replication ports or firewalls blocking cluster communication.
- Failing to use the deployer for app deployment, leading to inconsistent configurations.
- Misuse of the bootstrap command on multiple members.
- Overlooking time synchronization (NTP is recommended across all members to avoid replication errors).

---

1. What is the role of the manager node in a Splunk indexer cluster?

The manager node controls index replication, distributes configuration bundles to peers, and informs the search head about which peers to search.

---

2. What is the default replication and search factor in a Splunk cluster?

- **Replication Factor (RF):** 3
- **Search Factor (SF):** 2

---

3. What does the pass4SymmKey do in a Splunk cluster?

It authenticates communication between cluster nodes and must be the same across all cluster instances.

---

4. Can you use a free license for clustering in Splunk?

No. A free license cannot be used in clustered environments.

---

5. What are the three Splunk server roles in indexer clustering?

- **Manager Node**

- **Peer Node**
- **Search Head**

---

6. What configuration file is used to enable clustering in Splunk?

server.conf

---

7. How is data replicated in a single-site cluster with RF=3 and SF=2?

There are 3 copies of raw data, and 2 of them must be searchable (contain tsidx files).

---

8. How is disk usage estimated for clusters?

Estimate using:

- ~15% of daily index data for raw data
- ~35% of daily index data for index files

---

9. What command is used to configure a manager node?

./splunk edit cluster-config -mode manager -replication_factor 2 -search_factor 2 -secret mycluster

---

10. Can search heads belong to multiple indexer clusters?

Yes, they can. You use splunk add cluster-manager to configure them.

---

11. What is a searchable bucket?

A bucket that contains both raw data and index files (tsidx), allowing it to be searched.

---

12. What are the components of a clustered bucket name?

- Newest/oldest time
- Local ID
- GUID (originating peer's unique ID)

---

13. How are generations used in clustering?

Generations track the state of searchable bucket primaries as peers leave and rejoin the cluster.

---

14. What happens during a rolling restart?

A percentage of peer nodes restart sequentially to avoid service interruption.

---

15. What command performs a rolling restart?

./splunk rolling-restart cluster-peers

---

16. What is maintenance mode in Splunk clusters?

A mode that suspends bucket fix-ups and replication during maintenance tasks like upgrades.

---

17. How are configuration bundles pushed to peers?

Through the manager node, using:

./splunk apply cluster-bundle

---

18. What does the repFactor = auto setting do in indexes.conf?

It tells the cluster to replicate the index automatically according to the cluster's RF.

---

19. Can peer nodes be downgraded to non-clustered indexers?

No, once a node is part of a cluster, it cannot be reverted to standalone.

---

20. What are the benefits of multisite clustering?

- Disaster recovery
- Search affinity
- Geographic redundancy

---

21. What is site_replication_factor?

Specifies how raw data is distributed across sites. Example: origin:1,total:2.

---

22. What is site_search_factor?

Specifies how many copies of searchable data are distributed across sites. Example: origin:1,total:2.

---

23. What does the multisite=true setting do?

Enables multisite clustering in the configuration.

---

24. What are the common Splunk cluster ports?

- 8089 (management)
- 9100 (replication)
- 9997 (data receiving)

---

25. How do you convert a single-site cluster to a multisite one?

- Switch manager and peer nodes to multisite mode.
- Restart nodes in maintenance mode.
- Use splunk edit cluster-manager.

---

26. What is search affinity in multisite clusters?

It ensures that searches are run preferentially on local searchable copies, reducing WAN traffic.

---

27. How is search affinity disabled?

Edit the search head configuration and set a default site with site = site0.

---

28. What happens to legacy single-site buckets after migration?

They follow original RF/SF unless constrain_singlesite_buckets=false is set.

---

29. What is the command to show cluster status?

./splunk show cluster-status --verbose

---

30. What are the log files used for cluster diagnostics?

- splunkd_access.log
- splunkd.log
- metrics.log

Would you like this in a downloadable or printable format (e.g., PDF or Word)?

4o

You said:

add another 50 question

---

31. What is the function of "generations" in a Splunk indexer cluster?

Generations track the primary copies of buckets during cluster transitions. When a peer fails, the cluster increments the generation and reassigns primaries accordingly.

---

32. What is a primary bucket in clustering?

A primary bucket is the one actively used for searches. Only one primary exists per bucket group at a time.

---

33. What happens when there are excess buckets?

Excess buckets are redundant copies exceeding RF or SF requirements. The manager node detects and removes them to save space.

---

34. How does the cluster manager determine searchable rolling restart order?

It restarts 10% of peers at a time (configurable), maintaining cluster health throughout.

---

35. How do you initiate a searchable rolling restart?

splunk rolling-restart cluster-peers

---

36. What is the impact of maintenance mode on replication?

Replication and bucket fix-up are paused. This avoids unnecessary data duplication during maintenance tasks.

---

37. How do you enable maintenance mode?

splunk enable maintenance-mode

---

38. How are peers configured in a multisite environment?

Each peer includes its site designation and manager_uri in server.conf under [clustering].

---

39. What is site replication factor format?

Example: site_replication_factor = origin:2,total:3

---

40. Can search affinity be disabled?

Yes. Disabling it spreads search load across all sites, which may increase WAN traffic but improve balance.

---

41. What does splunk show cluster-status do?

It shows peer health, bucket status, and replication/search factor compliance.

---

42. What is a manual detention state for peers?

A state where a peer continues to search but stops indexing and replicating—useful for troubleshooting.

---

43. How do you manually place a peer in detention?

./splunk edit cluster-config -manual_detention on -peers <guid>

---

44. What is indexer discovery?

A feature where forwarders query the manager node for a list of active peers, allowing dynamic load balancing.

---

45. How is indexer discovery configured on forwarders?

In outputs.conf:

[indexer_discovery:cluster1]

manager_uri = https://<manager>:8089

pass4SymmKey = AnotherSecret

---

46. What is polling_rate in indexer discovery?

It's used by the manager node to determine how often forwarders should request peer lists.

---

47. What happens when converting a non-clustered indexer to a clustered peer?

Only new incoming data is replicated. Legacy buckets are not unless manually migrated.

---

48. Can an indexer peer be removed and reverted to standalone?

No. Once an indexer is part of a cluster, it cannot be used as a standalone indexer.

---

49. What are the risks of incorrectly managing bucket fix-up?

Excessive replication can saturate the network, especially after a node loss. Use bandwidth throttling if needed.

---

50. How do you validate and apply a configuration bundle?

./splunk validate cluster-bundle

./splunk apply cluster-bundle

---

51. What log files show clustering activities?

- splunkd.log
- metrics.log
- splunkd_access.log

---

52. What is constrain_singlesite_buckets=false used for?

It forces legacy single-site buckets to follow new

Here are **50 more detailed questions and answers** from the *Splunk 9.0 Cluster Administration* training document, continuing from question 30:

---

31. What command enables maintenance mode in a Splunk cluster?

./splunk enable maintenance-mode

---

32. What is the purpose of maintenance mode?

It temporarily halts bucket fix-ups and reassignment while you perform tasks like upgrades or peer restarts, ensuring the cluster doesn't rebalance unnecessarily.

---

33. What command disables maintenance mode?

./splunk disable maintenance-mode

---

34. What command shows maintenance mode status?

./splunk show maintenance-mode

---

35. How do you perform a rolling restart on multisite clusters?

./splunk rolling-restart cluster-peers -site-by-site true -site-order site2,site1

---

36. What happens when a primary bucket is lost?

The manager node promotes a searchable backup to primary and replicates missing data to maintain RF/SF.

---

37. What is a rawdata-only bucket?

A replicated bucket containing only raw data without index files (.tsidx), not searchable until promoted.

---

38. What is a searchable backup bucket?

A bucket copy with index files, ready to serve search queries when needed.

---

39. What happens when two peers fail in a 4-peer, RF=3, SF=2 cluster?

The cluster may remain valid but not complete—SF may be restored, but RF cannot due to lack of nodes.

---

40. What are the three types of buckets in a cluster?

- Primary (origin)
- Searchable backup
- Rawdata-only replica

---

41. What is the indexer cluster bucket naming convention?

- Clustered origin: db_<newest>_<oldest>_<localid>_<guid>
- Clustered replicated: rb_<newest>_<oldest>_<localid>_<guid>

---

42. What is the default configuration for a new manager node?

Replication factor = 3, Search factor = 2

---

43. What file stores the pass4SymmKey value?

server.conf under the [clustering] stanza

---

44. What command shows the decrypted shared secret?

./splunk show-decrypted --value '<hashed_secret>'

---

45. What is repFactor in indexes.conf?

It determines if the index is replicated across peer nodes:

- repFactor = auto: enable replication
- repFactor = 0: no replication

---

46. What are peer-apps in Splunk?

A read-only directory on peers where manager-distributed config bundles are staged. Do not edit manually.

---

47. What is the configuration path on manager node for pushing config bundles?

$SPLUNK_HOME/etc/manager-apps/

---

48. How do you validate a cluster bundle before applying it?

./splunk validate cluster-bundle

---

49. How do you apply a validated cluster bundle?

splunk apply cluster-bundle

---

50. How do you check cluster bundle status?

./ splunk show cluster-bundle-status --verbose

./splunk show cluster-bundle-status --verbose

---

51. How is config precedence handled in clustered indexers?

1. peer-apps local
2. system local
3. app local
4. peer-apps default
5. app default
6. system default

---

52. What command enables a peer to listen for data?

./splunk enable listen 9997

---

53. What command sets a peer to clustering mode?

./splunk edit cluster-config -mode peer ...

---

54. Can a peer node be downgraded to standalone?

No, a peer node cannot be reverted to non-clustered once joined.

---

55. What is the function of the Monitoring Console in clustering?

It provides performance monitoring and cluster health visualization across the environment.

---

56. What command lists license peers?

./splunk list licenser-peers

---

57. What command adds a license?

./splunk add licenses /path/to/license.lic

---

58. How do you switch a node to license peer mode?

chasplunk edit licenser-localpeer -manager_uri https://<host>:<port>

---

59. What tool is used to convert legacy buckets for multisite?

Splunk recommends contacting **Professional Services**—legacy buckets are not automatically replicated.

---

60. What happens to legacy buckets post-migration?

They retain their original RF/SF behavior unless constrain_singlesite_buckets=false is set.

---

61. Where is the GUID for a node located?

$SPLUNK_HOME/etc/instance.cfg

---

62. What happens when a peer goes offline in multisite?

Buckets may become incomplete; replication may fail depending on the site_replication_factor.

---

63. What is site_replication_factor = origin:2, total:3?

Two copies must remain in the origin site; third copy can be in any other site.

---

64. What command upgrades the manager node?

Follow this sequence:

1. Upgrade manager
2. Upgrade peers
3. Finalize with:

./splunk upgrade-finalize cluster-peers

---

65. What is the upgrade-init command for peers?

./splunk upgrade-init cluster-peers

---

66. How does Splunk track excess buckets?

Using generation IDs. Excess buckets are flagged and can be cleaned up post-rebalancing.

---

67. What log shows peer-to-manager communication?

splunkd_access.log with endpoints like /services/cluster/manager

---

68. What is the minimum number of peers for RF=3?

At least 3 peer nodes are required.

---

69. Can search heads search non-clustered peers and clustered ones together?

Yes, they can search across both types simultaneously.

---

70. How does Splunk ensure high availability in SmartStore-enabled clusters?

By offloading buckets to remote storage while maintaining searchable warm buckets locally.

---

71. Where are searchable bucket status flags available via REST?

http

/services/cluster/manager/buckets

---

72. What command shows peer cluster configuration?

./splunk show cluster-config

---

73. What happens to .tsidx files during primary reassignment?

They are generated on promoted peers to make buckets searchable again.

---

74. What are acceptable system requirements for indexers?

- 800 IOPS disk speed
- 12 GB RAM minimum
- 64-bit OS and CPUs at 2+ GHz

---

75. What command removes a peer from the cluster?

Use:

./splunk offline --enforce-counts

---

76. What is the use of metrics.log in clustering?

Tracks cluster connections and peer health under group=clusterout_connections.

---

77. How does a search head become aware of indexers in a cluster?

The manager node provides a list of active and searchable peers.

---

78. How are apps deployed in SHC vs. indexer clusters?

- **SHC**: via the Deployer
- **Indexer Clusters**: via Manager Node and manager-apps directory

---

80. What is the impact of disabled search affinity?

Searches may span all sites instead of staying local, increasing WAN traffic.

---

1. What is the primary purpose of indexer discovery in Splunk?

It allows forwarders to dynamically discover and load-balance among indexers without having to manually update outputs.conf when indexers change.

---

2. What are the key optional settings for indexer discovery?

- **Polling rate**: Controls how often forwarders query the manager.
- **Weighted load balancing**: Distributes traffic based on indexer weights.

---

3. What configuration file is used to set up indexer discovery on forwarders?

outputs.conf under the [tcpout] stanza.

---

4. What setting enables acknowledgment in forwarders?

useACK = true in the outputs.conf file ensures data is confirmed as indexed before being discarded.

---

5. Why is useACK important in clustered environments?

It ensures data durability by confirming that data has been indexed and replicated according to the cluster's policies.

---

6. What happens when a forwarder sends data to a node that fails mid-transfer?

A bucket roll occurs, and the cluster initiates a fix-up process to preserve data durability.

---

7. What is volume-based load balancing?

It distributes data to indexers based on the volume sent to each, not simply round-robin, improving even data distribution.

---

8. What is the challenge of managing forwarders without indexer discovery?

Manual updates to outputs.conf are required on every forwarder when indexer lists change, often requiring restarts.

---

9. How does indexer discovery simplify forwarder management?

It centralizes indexer list updates at the cluster manager level, eliminating the need to manually configure each forwarder.

---

10. How is indexer discovery enabled on the manager node?

Through the deployment of a discovery endpoint, which forwarders query for available indexers.

---

11. How do you test forwarder site failover?

By simulating a site outage and verifying if the forwarder redirects traffic to indexers in another site.

---

12. What CLI command enables the deployment client on a forwarder?

You update the deploymentclient.conf and restart the forwarder to register it with the deployment server.

---

13. In a lab environment, what port is typically used for Splunkd?

The default is usually 8089, but it can vary (e.g., 8289 for a forwarder in the lab exercise).

---

14. How can you verify a forwarder's app deployment?

Use the Monitoring Console (MC) to check app status or confirm via the forwarder's internal logs.

---

15. What information must be included in the Monitoring Console for forwarders?

You must update the instance's server role and register it as a forwarder in MC.

---

16. How does site failover benefit indexer discovery?

It provides high availability by allowing forwarders to automatically reroute traffic to other sites if the origin site fails.

---

17. What is the polling behavior of indexer discovery?

Forwarders periodically check in with the manager to refresh the list of available indexers.

---

18. What is the server parameter in outputs.conf used for?

It specifies the list of indexer targets or the discovery endpoint if using indexer discovery.

---

19. What does the deployment server do in forwarder management?

It distributes configuration bundles (like outputs.conf) to forwarders for consistent setup and centralized control.

---

20. What is the splunk list deploy-clients command used for?

To verify which forwarders are currently connected to the deployment server.