splunk>

Splunk Professional Services Technical Assessment Report

Capital One, Richmond, VA – 01/17/17 – 01/23/17 – ON-PREM ONLY

Contacts

Name	Email	Phone	Location	Name
Tom Nichter	Tom.Nichter@capitalone.com		Richmond, VA	Tom Nichter

Executive Summary

This document details the results of the Splunk Professional Services Technical Assessment (SPSTA) of the Splunk environment at Capital One. Information herein was collected from interviews with the staff regarding the infrastructure, architecture, applications and use cases and supporting performance metrics and key indicators of the environment were gathered from the Splunk infrastructure directly.

The following legend indicates the relative importance of the various findings:

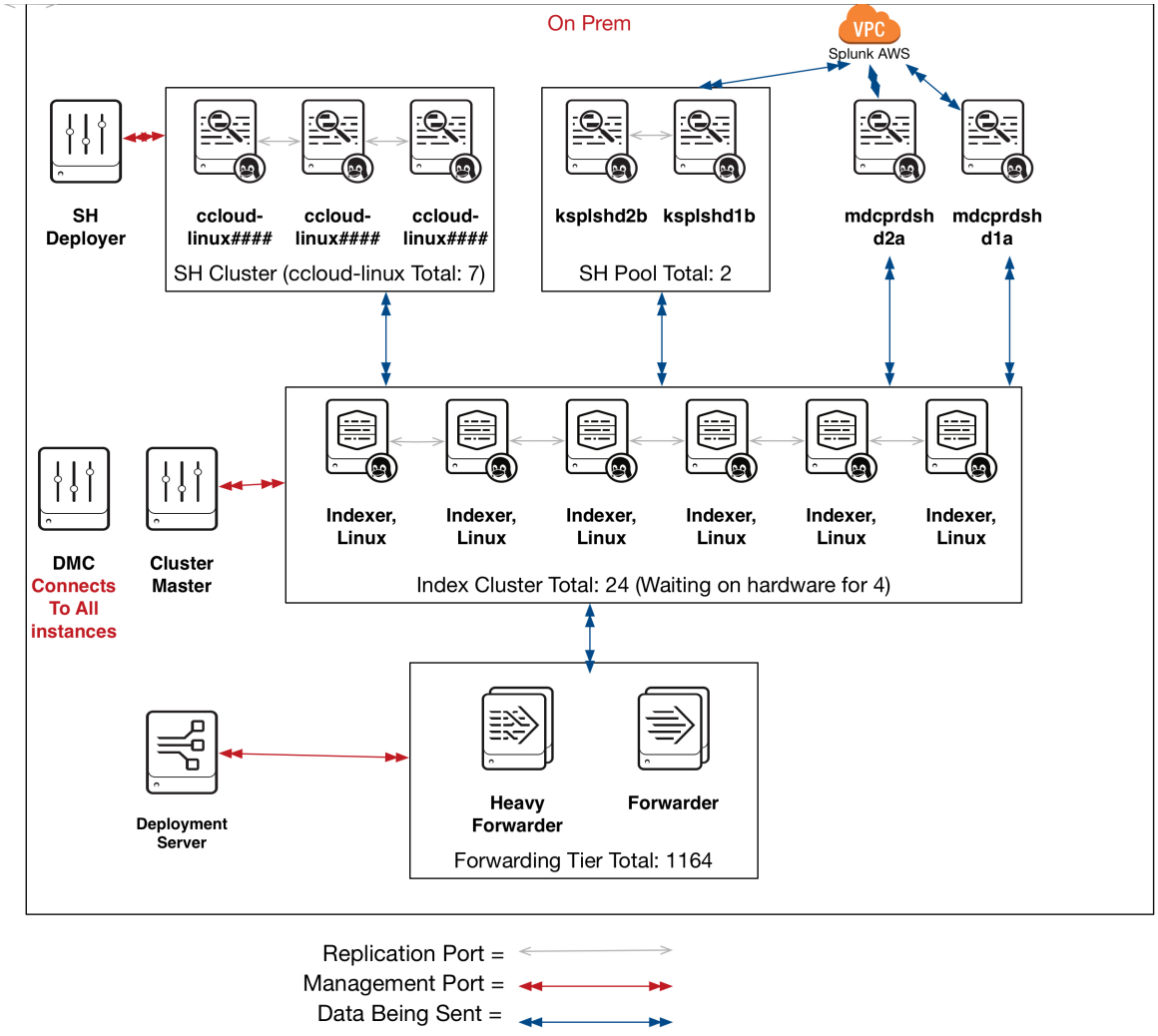
- Items requiring immediate attention and are fixes or changes that can be done immediately with relatively low risk to the current infrastructure.
- Short-term actions that could provide benefits on the existing architecture in the next 3-4 months. These items are changes that will help provide stability and governance in the current architecture.
- Long-term actions that involve major changes to the existing architecture to address scalability or overall ease of maintenance.

A summary of the findings is as follows:

1. SHC members do not meet hardware requirements.
2. All servers should be connected to an NTP server.
3. Indexers consuming a lot of CPU
4. Verify that all files are owned by Splunk.
5. Issues with data ingestion
6. Bundles are hitting timeouts
7. THP is enabled
8. 92 instances aren't managed by the DS
9. retention policy issues
10. Searches are being skipped because the SHs are over-utilized
11. Data Models are not completing
12. Review users with admin and delete privileges.
13. Review reports and dashboards.

Architecture

The Architecture section reviews the customer's current architecture and results from the discovery stage of the STA. Below is a diagram of the current Capital One – On Prem architecture:



Instance Inventory

Instance Type	Count	OS	CPU	Memory	Disk
Indexer	8	Linux, x86_64	20 cores	32 GB	
Indexer	8	Linux, x86_64	24 cores	128 GB	
Indexer	4	Linux, x86_64	36 cores	128 GB	
Search Head	2	Linux, x86_64	24 cores	128 GB	100 GB / 14% used
Search Pool	2	Linux, x86_64	20 cores	32 GB	130 GB / 10% used

Search Head Cluster	8	Linux, x86_64	6 cores	16 GB	40 GB / 50% used
Deployment Server	1	Linux, x86_64	4 cores	8 GB	72 GB / 9.5% used
License Master	1	Linux, x86_64	4 cores	8 GB	70 GB / 11% used
Cluster Master	1	Linux, x86_64	4 cores	8 GB	70 GB / 11% used
Search Head Cluster Deployer					

License Usage

List the current license usage statistics.

Purchased License Capacity (GB)	Average License Consumed (GB) Last 30 Days	Peak License Consumed (GB) Last 30 Days
10 TB	5.2 TB	6.1 TB

Architecture Notes or Findings

List any miscellaneous notes, assumptions, or details regarding the Splunk infrastructure servers.

1. Consider migrating to a single SHC
2. SH are split apart and under spec

Architecture Recommendations

List any architecture related recommendations below.

Finding	Recommendations / Next Steps
1. Consider migrating to a single SHC	◆ This will help reduce overhead for the on-prem environment.
2. Consider all indexers and search heads having the same hardware.	◆ Its recommended that all members within a cluster have the same hardware specs.

Splunk Reference Hardware Requirements: Reference Hardware Requirements.

System Overview

splunk>

The system overview section of the STA reviews system resource utilization, server configuration, application inventory, and error messages. Detailed findings and supporting artifacts can be found below.

System Overview Findings

List any miscellaneous notes, assumptions, or details regarding the system overview. If applicable, provide supporting screenshots.

1. Search head cluster members do not meet Splunk minimum requirements.
2. All servers are not connected to an NTP server.
    a. ksplshd2b
    b. ksplshd1b
    c. mdcprdsplk1b
    d. mdcprdsplk4b
    e. ksplidx1c
    f. ksplidx3d
    g. ksplidx3c
    h. mdcprdshd1a
    i. ksplidx4c
    j. ksplidx2d
    k. mdcprdshd2a
    l. ksplidx2c
    m. ksplidx4d
    n. ksplidx1d
3. Indexers are constantly consuming more than 50% of the available CPU.
    a. ksplidx2d
    b. ksplidx3c
    c. ksplidx3d
    d. ksplidx4c
    e. ksplidx1c
    f. ksplidx4d
    g. ksplidx2c
    h. ksplidx1d
4. Noticed a lot of permission issues in the internal logs. This may indicate that the Splunk's directory will need to be chowned.
5. A lot of truncating, timestamping and linebreaking errors.
6. Bundle replication timeouts should be increased.
7. THP is enabled on 30 servers

   a. If THP is disabled at boot time that's fine
   b. If THP is disabled at runtime then defrag also has to be disabled
8. kspldpy2a doesn't meet ulimit requirements
9. 92 instances are not managed by the deployment server
10. 2 clients haven't phoned home
   a. lobatchprdv2
   b. lobatchprdk2

System Overview Recommendations

List any system overview related recommendations below.

Finding	Recommendations / Next Steps
1. SHC members do not meet hardware requirements.	◆ Provide different hardware to meet Splunk minimum requirements if possible.
2. All servers should be connected to an NTP server.	◆ Connect servers to NTP.
3. Indexers consuming a lot of CPU	◆ Enable force time load balancing on the forwarders to help spread the load evenly. Also consider increase number of indexers. Upgrade to a newer version of Splunk to gain control of bucket rebalancing to help balance search better.
4. Verify that all files are owned by Splunk.	◆ /opt/splunk should be owned by the splunk user
5. Issues with data ingestion	◆ Review datasources and update props.conf and transforms.conf files
6. Bundles are hitting timeouts	◆ Increase timeouts
7. THP is enabled	◆ Disable THP during boot for all splunk instances
8. kspldpy2a	◆ increase ulimits
9. 92 instances aren't managed by the DS	◆ All instances should be managed by a SH-Deployer, Cluster Master or Deployment server.
10. Clients aren't phoning home	◆ See why clients aren't phoning home

splunk>

Indexing and Data Overview

A detailed examination of data on-boarding, forwarder management, and indexer utilization and capacity was conducted including an overview of Splunk data onboarding best practices. Detailed findings and supporting artifacts can be found below.

Indexing and Data Overview Findings

List any miscellaneous notes, assumptions, or details regarding the indexing and data onboarding review. If applicable, provide supporting screenshots.

1. Sourcetype Issues
    a. 72 sourcetypes with data parsing issues
    b. 57 sourcetypes with truncation issues
    c. 35 sourcetypes with line merging issues
2. 186 sourcetypes have a delay of 30+ minutes before being indexed
3. ksplidx2d, ksplidx3c, ksplidx4d, ksplidx1c, ksplidx1d, ksplidx3d have less than 25% free space
4. mdcprdsplk2b is ingesting about half of the data that the other indexers are
5. kdcprdsplk1c – kdcprdsplk4c are not currently ingesting data
6. Some indexes are not following the retention policy

Indexing and Data Overview Recommendations

List any indexing and data review related recommendations below.

Finding	Recommendations / Next Steps
1. Sourcetypes need props.conf and transforms.conf entries.	◆ Review the data and modify the settings to ingest the data correctly.
2. Forwarders are not sending data to all indexers.	◆ Forwarders are not sending data to all indexers. Its possible they are forwarding to intermediate forwarders.
3. Index size	◆ Provide additional hardware as needed
4. Index data load	◆ Verify that all forwarders are sending data to all the indexers and that its being load balanced
5. Indexes not following retention policy	◆ Review the indexes.conf and make corrections as needed.

splunk>

Search Overview

The search overview section of the STA covers searching best practices, optimization techniques and a review of overall search performance and usage patterns. Detailed findings and supporting artifacts can be found below.

Search Overview Findings

List any miscellaneous notes, assumptions, or details regarding the search overview. If applicable, provide supporting screenshots.

1. 17% to 25% of searches are being skipped
2. 2 of 18 data models are not completing
3. 14 search heads have remote search logs enabled
4. 14 eventtypes have no indexes defined
5. Search Activity Dashboard
    a. ksplshd2b is being heavily utilized compared to other search heads
    b. mdcprdshd1a is being over-utilized and running multiple real-time searches
    c. mdcprdshd2a is being over-utilized and running greater than 5 real-time searches
6. ksplshd2b may have storage issues because dispatching a search takes longer than usual

Search Overview Recommendations

List any searching related recommendations below.

Finding	Recommendations / Next Steps
1. Searches are being skipped.	◆ This is caused by a performance issue with multiple search heads. Some search heads have a lot of searches while others are under utilized. May help if a single cluster was implemented.
2. Data models are not completing	◆ Likely due to performance issues with both the search heads and indexers.
3. 14 search heads have remote logs enabled.	◆ Modify the limits.conf to disable this functionality.
4. 3 search heads are being heavily utilized.	◆ Help separate the user load and disable real-time searches

Security Overview

The security overview section of the SPSTA reviews security settings and user and role based access controls. Detailed findings and supporting artifacts can be found below.

Security Overview Findings

List any miscellaneous notes, assumptions, or details regarding the security review. If applicable, provide supporting screenshots.

1. 42 users with admin privileges
2. 42 users with delete privileges
3. 571 scheduled reports with 456 alerting reports.
    a. 138 scheduled reports have returned no results in the last 30 days.
4. There are 497 reports in the search app.
5. There are 1,300 private reports. 153 of these are currently scheduled.
6. 94 dashboards have greater than 20 inline searches. Moreover, 7 dashboards have greater than 100 inline searches.

Security Overview Recommendations

List any security related recommendations below.

Finding	Recommendations / Next Steps
1. Review users with admin and delete privileges.	◆ Update authentication settings.
2. Review users with private scheduled reports	◆ Update authentication settings.
3. 94 dashboards have greater than 20 inline searches.	◆ Dashboards should be split apart to improve Splunk's performance.