splunk>

Splunk Professional Services Technical Assessment Report

Capital One, Richmond, VA, 01/12/2017 – 01/18/2017 – US EAST DIGITAL ONLY

Contacts

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

Executive Summary

This document details the results of the Splunk Professional Services Technical Assessment (SPSTA) of the Splunk environment. 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.

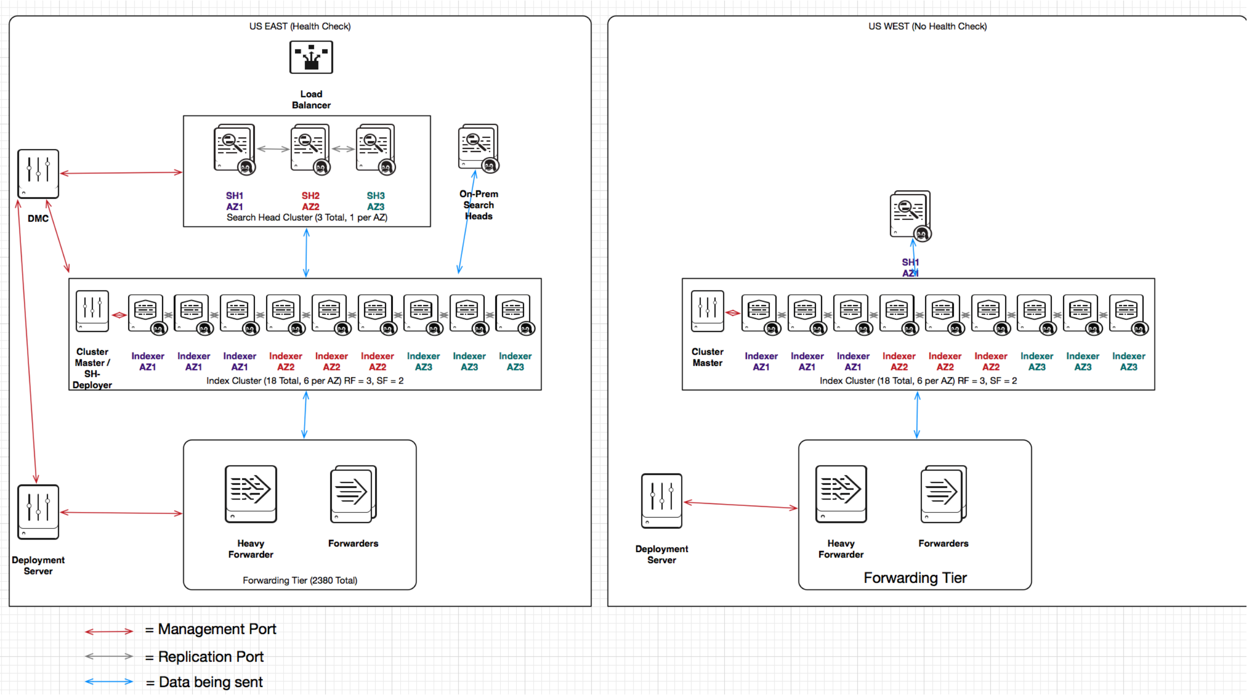
A summary of the findings is as follows:

1. Currently running Splunk 6.3.2. Please upgrade to 6.5.1 or newer.
2. Search heads do not meet Splunk requirements.

   http://docs.splunk.com/Documentation/Splunk/6.5.1/Capacity/Referencehardware
3. Sourcetypes have timestamping, line breaking and truncation issues.
4. Indexes are not following the retention policy. HOT→WARM→COLD→FROZEN / DELETE
5. Forwarders are sending to much data and thruput needs to be increased.
6. Some forwarders are not sending data to all of the indexers within the cluster.
7. Review users assigned to the admin and can_delete roles.
8. Indexers web interface is enabled. This should be disabled.

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 digital architecture:



Instance Inventory

Instance Type	Count	OS	CPU	Memory	Disk
Indexer	2	Linux, x86_64	18	64 GB	10 TB, EC2
Search Head	8	Linux, x86_64	8	32 GB	100 GB, EC2
Deployment Server	1	Linux, x86_64	8	32 GB	30 GB,
License Master	1	Linux, x86_64	4	8 GB	30 GB
Cluster Master / SH Deployer	1	Linux, x86_64	4	16 GB	10 TB, EC2

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 Recommendations

List any architecture related recommendations below.

Finding	Recommendations / Next Steps
1. Multiple individual search head cluster and index clusters.	Implement single SHC and Index Cluster
2. SH-Deployer and Cluster Master on the same instance	Move the SH-Deployer onto a separate instance or different instance than the CM. Can use the DMC instance.
3. Search heads are not able to use hyper threading. Found only 8 cores usable by the search heads.	Splunk recommended at least 16 CPU / 12 GB of RAM with 300 GB 10,000 RPM SAS hard disks. Example, Linux (2 CPU, 2 GB RAM) + Splunk (16 CPU, 12 GB RAM) = 18 CPU, 14 GB of RAM min

Splunk Reference Hardware Requirements: Reference Hardware Requirements.

System Overview

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.

- All the indexers are between 65% - 78% of storage utilization
- SH3 seems to be the most used search head
  - May indicate a load balance issue
  - There was a bug with distributing searches evenly with a SHC please consider upgrade

System Overview Recommendations

List any system overview related recommendations below.

Finding	Recommendations / Next Steps
1. Indexers between 65% and 78% of storage utilization	◆ Consider increase the EBS volume

2. SH3 seems to be heavily utilized. SHC may not be evenly distributing the searches.	◆ Recommend upgrading to 6.5.1 or newer. A lot of SHC bugs have been fixed since 6.3.2.

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. Indexing is not evenly distributed. I see some indexers receiving 5MB per second and others receiving less than 1 MB per second. In addition to that, any indexers indexing more than 1.5 MB per second the indexing queue starts to fill up. This is typically an indication that more IOPS are needed to handle the load.
2. Data ingestion issues:
    a. Timestamp:
        i. 30 sourcetypes are not able to parse their timestamps correctly
    b. Truncation:
        i. 21 sourcetypes are exceeding the truncation limit (May be losing important data)
    c. Line merging issues
        i. 25 sourcetypes are not linebreaking correctly
    d. Sourcetypes are creating dated files instead of a single rolling file.
        i. Currently State: filename.DD:MM:YYYY.log
        ii. Recommended: filename.log → rolls to → filename.log.x
3. Index Issues:
    a. ciaudit is full consider increasing size if not meeting retention period
    b. eapi_cloud1_summary at 95% on digitalidx11ae while other indexers are around 75%
    c. codemobile is between 77% to 67% full
    d. most indexes have to many hot buckets
    e. warm buckets are not rolling to cold
    f. many indexes have data past their expected retention
    g. The aggregator is consuming about the same amount of CPU as indexing

Indexing and Data Overview Recommendations

List any indexing and data review related recommendations below.

Finding	Recommendations / Next Steps
1. Review retention policy for indexers	◆ Verify that buckets are set to auto_high_volume. Also update the retention policy so that older buckets are removed.
2. Fix sourcetype issues	◆ Review bad sourcetypes and create props.conf and transforms.conf entries to fix the issues.
3. Update to a newer version of Splunk	◆ Update to a newer version of splunk and enable forced load balancing on the UF and HF to help evenly distribute the indexing load.
4. 62 forwarders are not sending data to all indexers	◆ 62 forwarders are not sending data to all the indexers.
5. Multiple forwarders are reaching their max thruput.	◆ Consider increasing the [thruput] for each forwarder to help keep data moving to the indexers in real-time.

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. 180+ saved searches that haven't returned results in the last 30 days
2. 1 orphaned saved search / not defined by an owner (NOBODY)
3. 8 Search Heads are fetching remote search logs
4. 851 scheduler errors in the last 24 hours
5. 9/12 incomplete data models
6. 8 eventtypes without indexes

Search Overview Recommendations

List any searching related recommendations below.

Finding	Recommendations / Next Steps
1. 180+ saved searches not returning results	🔶 See if these searches still need to run
2. 1 orphaned saved search	🔶 A user likely left the company so the search should be removed or given to another user
3. 8 Search Heads are returning remote logs	♦ In the limits.conf add these configurations fetch_remote_search_log = false and remote_timeline_fetchall = false
4. 9/12 incomplete data models	♦ The host_mapping lookup table needs to be added for the RTM data models and the Mobile data models have a permissions issue
5. 8 eventtypes without an index	🔶 Indexes are recommended to help with search performance but not required

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. There are currently 42 users that have admin privileges and are able to delete indexed data.
2. Hundreds of dashboards have greater than 20 inline searches.

Security Overview Recommendations

List any security related recommendations below.

Finding	Recommendations / Next Steps
1. Total Admin Users: 42 Total Users that can_delete: 42	♦ Review system access and determine if all 42 users require admin and can_delete rights.
2. 64 dashboards have greater than 20 inline searches	♦ PS recommends having less than 20 inline searches per dashboards.

3. Splunkweb is enabled on all the indexers.	◆ Please disable splunk web on all the indexers.