Best Practices and Better Practices for Admins
…while you get settled…

► Latest Slides:
  • https://splunk.box.com/v/blueprints-practices-admin

► Collaborate: #bestpractices
  • Sign Up @ http://splk.it/slack

► Load Feedback --------------------------------------->
Best Practices and Better Practices for Admins

Presented by Splunk Blueprints

Burch | Senior Best Practices Engineer
Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Listen to Your Data, The Engine for Machine Data, Splunk Cloud, Splunk Light and SPL are trademarks and registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners. © 2017 Splunk Inc. All rights reserved.
“Scale customer success through the automation of adoption services and best practices”

Blueprint’s Mission
What’s a “Burch”?  

Senior Best Practices Engineer  

- Was a Senior Sales Engineer  
- Before that, Splunk Customer  
- Before that, Middleware Eng  
- Before that, Computer Science  
- Before that, an idea of my parents
Agenda

Are you in the right place?

1. User Management
2. Data Onboarding
3. Splunk Health
4. Config Management
5. App & TA Creation
6. Architecture
7. Search Tier
8. Indexing Tier
9. Securing Splunk
User Management
User Education & Enablement

- Creating Content:
  - Teaching + Videos + Wikis

- Is that your core competency?

- Outsource it to us!
  - Capture unique things
Search Tutorial
Free Search Tutorial -> docs.splunk.com -> Search Tutorial

- Download & Installs Splunk
- Local sandbox
- Add tutorial data
Exploring Splunk
SEARCH PROCESSING LANGUAGE (SPL) PRIMER AND COOKBOOK

By David Carasso, Splunk's Chief Mind
Community Q&A
answers.splunk.com

- E-mail notifications
- Fast answers
- Larger distribution
App as Workspace

Default App with Default Dashboard
- Welcome page

Dashboard for new users
- not search box

Drive their eyes/focus
- Hide other apps – even Search!
- show_in_nav = false
Welcome Page Creator
https://splunkbase.splunk.com/app/2991

Hands-on Labs

Creating Welcome Pages
Tuesday, September 26, 2017 | 2:00 PM-2:15 PM
Wednesday, September 27, 2017 | 11:00 AM-11:15 AM

Burch I, Senior Best Practices Engineer, Splunk Inc.

Users often land in Splunk with no clue where to begin. In this lab, you’ll get hands-on training on how to use the Welcome App Page Creator to create an effective starting page for your users. Check out the associated blog post (http://blogs.splunk.com/2016/09/01/introducing-the-welcome-page-creator) for more details!
Incentive Driven User Onboarding

“I can’t believe those users did those things I let them do!”

- Don’t be a data butler
- Identify & coach & promote to power
- Work with you to implement and learn best practices
Blueprints for Onboarding Teams

Thursday @ 11:35am

STOP HITTING YOURSELF

@JAZMINREDUX
Banner Notifications
docs.splunk.com “Splunk Web messages”

▶ Examples:
  • Scheduled restart
  • Ongoing issues
  • Cool KO to check out

▶ Specific audiences
  • Role
  • Capability
BAU Account

Dog Food!

- Use non-admin account
  - Prevents accidents
  - Live with limitations
  - Appreciate user experience

- Admin on MC
Data Onboarding
“If you log it, then you should Splunk it”

- App/System performance to write logs
- Disk to store logs

cronjobs/scheduled tasks to Splunk

- Scripted Inputs
- standard output/error captured
- Example: Log Rotation crontab
Onboarding != Ingestion
A David Paper Joint!

Onboarding Phases

1. Initial Request
2. Definition
3. Implementation
4. Value
5. Validation
6. Announcement

Ingestion

- Event Breaks
- Time Stamps
- Source
- Sourcetype
- Index
- Host

Why does this matter?
Logging

Search: dev.splunk.com “logging best practices”

I DID ABSOLUTELY NOTHING TODAY

AND IT WAS EVERYTHING I
THOUGHT IT COULD BE
## Hidden Fields: Time

**Search:** docs.splunk.com “search time modifiers”

<table>
<thead>
<tr>
<th>Event Time</th>
<th>Index Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>_time</td>
<td>_indextime</td>
</tr>
<tr>
<td>earliest</td>
<td>_index_earliest</td>
</tr>
<tr>
<td>latest</td>
<td>_index_latest</td>
</tr>
</tbody>
</table>

- What does a big difference mean?
- Search over last 5min every 5min but there’s a 10min delay in _indextime
- When is this ok vs needs attention?
Splunk Sandboxing

Hands-on Labs

Sandboxing with Splunk (with Docker)
Accept it. You're afraid to take risks in Splunk. So was I. That is, until Docker changed my life.
Join the cult and learn how to rapidly create disposable Splunk sandboxes in mere minutes!
Support Tickets

docs.splunk.com “How to file a great Support case”

▶ Open Cases
  • break/fix only
  • Details, details, details
  • Diags everywhere!
    • Remote
    • Upload to case

▶ Schedule webex
  • Delay and much lost in email
Point & Purpose

Renamed from “Distributed Management Console (DMC)”

- Buddy with License Server
- Standalone instance
- Conceptually “Admin Console”
  - No user stuff
  - Only MC apps/jobs
## Health Check

**Add your own!**

<table>
<thead>
<tr>
<th>Check</th>
<th>Category</th>
<th>Tags</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event-processing issues</td>
<td>Data Collection</td>
<td>event_breaking, indexing, timestamp_extraction</td>
<td></td>
</tr>
<tr>
<td>Expiring or expired licenses</td>
<td>Data Indexing</td>
<td>licensing</td>
<td></td>
</tr>
<tr>
<td>Indexing status</td>
<td>Data Indexing</td>
<td>indexing</td>
<td></td>
</tr>
<tr>
<td>Local indexing on non-indexer instances</td>
<td>Data Indexing</td>
<td>best_practices, forwarding, indexing</td>
<td></td>
</tr>
<tr>
<td>Missing forwarders</td>
<td>Data Indexing</td>
<td>forwarding</td>
<td></td>
</tr>
<tr>
<td>Saturation of event-processing queues</td>
<td>Data Indexing</td>
<td>indexing, queues</td>
<td></td>
</tr>
<tr>
<td>License warnings and violations</td>
<td>Data Indexing</td>
<td>indexing, licensing</td>
<td></td>
</tr>
<tr>
<td>Distributed search health assessment</td>
<td>Data Search</td>
<td>distributed_search</td>
<td></td>
</tr>
<tr>
<td>Search scheduler skip ratio</td>
<td>Data Search</td>
<td>scheduler</td>
<td></td>
</tr>
<tr>
<td>Integrity check of installed files</td>
<td>Splunk Miscellaneous</td>
<td>configuration, installation</td>
<td></td>
</tr>
<tr>
<td>KV Store status</td>
<td>Splunk Miscellaneous</td>
<td>kv_store</td>
<td></td>
</tr>
<tr>
<td>Orphaned scheduled searches</td>
<td>Splunk Miscellaneous</td>
<td>configuration, search</td>
<td></td>
</tr>
<tr>
<td>Upgrade opportunity from search head pooling to search head clustering</td>
<td>Splunk Miscellaneous</td>
<td>best_practices, configuration</td>
<td></td>
</tr>
<tr>
<td>Excessive physical memory usage</td>
<td>Splunk Miscellaneous</td>
<td>resource_usage</td>
<td></td>
</tr>
<tr>
<td>Linux kernel transparent huge pages</td>
<td>System and Environment</td>
<td>best_practices, operating_system</td>
<td></td>
</tr>
<tr>
<td>Assessment of server ulimits</td>
<td>System and Environment</td>
<td>best_practices, operating_system</td>
<td></td>
</tr>
<tr>
<td>Near-critical disk usage</td>
<td>System and Environment</td>
<td>capacity, storage</td>
<td></td>
</tr>
<tr>
<td>System hardware provisioning assessment</td>
<td>System and Environment</td>
<td>best_practices, capacity, scalability</td>
<td></td>
</tr>
</tbody>
</table>

**Linux kernel transparent huge pages**

**Description**

This attempt to determine whether Splunk is running on a Linux server where kernel transparent huge pages are enabled. **NOTE**: This check is relevant only for Linux. **NOTE**: This check yields results only for instances that are running Splunk Enterprise 6.5 or higher. Instances running an older version produce search errors that can be ignored.

**Message**

This health check item has not been run or is still running. Check back when it is complete.
Find Impacting Searches

- **Search Activity:**
  - Top 20 Memory-Consuming Searches

- **Search Usage Statistics**
  - Long-Running Searches

- **Great for**
  - Clean up
  - Identifying users to mature
Config Management
To btool, or not to btool

- `btool <configuration> list <stanza|> --debug|>

- **Add to your env path! (source a profile file from an app)**
  - **Linux:** `export LD_LIBRARY_PATH=$SPLUNK_HOME/lib`
  - **Mac:** `export DYLD_LIBRARY_PATH=$SPLUNK_HOME/lib`

- **No “.conf”**

- **Use --debug with | grep -v “system/default”**

- **Not current runtime**
Indent Config

Example:

```
[general]
pass4SymmKey = $1$ShiC+P0X
serverName = elBurcho
sessionTimeout = 30m
```

Benefit

- Easily see system vs hand edits
- Detect hand config updated by system
Simple Version Control

▶ Good: Scripted Input
  • Specific Diag (or just etc dir)
  • Clean old copies

▶ Better: Scripted Input
  • Check in to git

▶ Best: Custom Built Solution
  • Source Control

▶ Targets
  • Utilities
  • SHC Working Folder

▶ Source Control != High Availability
  • VMotion type stuffs
Keep It Clean: Naming Conventions
Handout at Customer Success Studio

▶ Template: <summary>|>_<company>_<function>_<environment>

▶ <company>
• Yours or from a 3rd party/splunk app

▶ <function>
• Nothing that changes (i.e. organization/teams)

▶ <environment>
• PROD, DR, QA, TEST, DEV, etc…

▶ <summary>|>
• Exists as a modifying of corresponding index
App Management
What practices do you notice?

Burch_configbackup_ta
Burch_CustomerOverview
Burch_datacollection_ta
Burch_deployer_ta
Burch_deploymentserver_ta
Burch_dmc_ta
Burch_dreamhost_ta
Burch_es_ta
Burch_forwarder_ta
Burch_heavyforwarder_ta
Burch_indexer_ta
Burch_license_client_ta
Burch_license_server_ta
Burch_master_ta
Burch_multisite_site1_ta
Burch_multisite_site2_ta
Burch_sandbox_ta
Burch_searchheadcluster_ta
Burch_searchhead_distributed_ta
Burch_searchhead_ta
Burch_searchtimeko_ta
Burch_splunkAdmin_nix_ta
Burch_splunkUpgrade_ta
Burch_splunk_admin
Burch_splunk_default
Burch_splunk_developer
Burch_splunk_power
Burch_splunk_user
Burch_stopdeploymentclient_ta
Burch_utility_ta
Burch_zglobal_ta
Bootstrap

Minimal system/local

1. Install Splunk Enterprise

2. Bootstrap
   • Point to DS/Master/Deployer
   • system/local overwritten by apps
   • Centralized control
   • Global App < Function Apps

3. Download app with scripted input
   • Non config changes
   • Risky!
App Development

- No index please!
  - Provide recommendation
  - Volumes vs Retentions vs RBAC etc..

- Inputs disabled
  - Don’t touch my license & storage!

- Remove files
  - .DS_Store
  - .pyc .pyo
  - local.meta

- Macros & Tags
  - easy modification
  - imagine rewriting every search/dashboard
  - Candidates: index, sourcetype, source

- Prebuilt Panels vs Dashboards
  - ./splunk package app <app>
  - Tar non-compatibilities
Certification for Practices
Add-on Builder includes App Inspect

Overview of Splunk AppInspect

Welcome to Splunk AppInspect!

Splunk AppInspect validates your Splunk app against a set of Splunk-defined criteria so that you can be assured of its quality and robustness. AppInspect runs various checks on your app package, and then produces a report that clearly details any missed criteria. AppInspect ensures that your Splunk app is ready for production use on your own Splunk Enterprise instance, or for submission to Splunkbase as either a standard or certified app.

Note: For more information about developing Splunk apps, see the resources page in this documentation.

AppInspect evaluates all of the following for a given Splunk app:

- Structure
- Feature set
- Security
- Readiness for Splunk certification

- Information in the Splunk AppInspect help.
Architecture
If you expect to grow big…

- **Separate Installs:**
  - Easier scalability
  - Avoid reload deploy-server on restart
  - Cheap VMs

- **Keep Utility apps in sync**
  - DS -> Master -> IDXC
  - DS -> Deployer -> SHC
  - Not for faint of heart…
Data Management

“Compare QA & PROD...D’oh!”

- Non PROD data -> PROD SPLUNK!
  - “If a single team depends on it, then it’s production” – Terry Martin
  - Or Search Head traverses

- Logical Separation:
  - Role Based Access Control
  - Separate indexes per env
  - Use eventtypes/tags

- `forwardedindex.filter.disable`
Data Distribution Quirks

Worst Practices...and How to Fix Them

Tuesday, September 26, 2017 | 3:30 PM-4:15 PM  INTERMEDIATE

Jeff Champagne, Staff Architect, Splunk Inc.

We’ve all slowed down to get a glimpse of a car crash on the freeway or tuned in to hear about a celebrity scandal. This session will analyze the Splunk equivalent of a 16-car pileup from an architecture and search workload management perspective. Come hear about real-life Splunk deployments that went bad and how you can avoid those same pitfalls.

- Consolidated data == serial search
- Forwarders:Indexers Ratio
- autoLBVolume + autoLBFrequency
Data Collection Tier
Practices whether push or pull data

▶ Easier to scale
  • Vertical (VM specs)
  • Horizontal (cheaper than indexers)
  • Load balancer (not hardcoded)

▶ Minimize IDX/SH Restarts
Search Tier
Help me?!  

> n00b  
> Ninja
Ninja: Debug This

Where’s Waldo eval max_runtime?!

```
`dmc_audit_get_searches(*)` | stats min(_time) as _time, values(user) as user, max(total_run_time) as total_run_time, first(search) as search, first(search_type) as search_type, first(apiStart_time) as apiStart_time, first(apiEnd_time) as apiEnd_time by search_id | where isnotnull(search) AND search_type ="ad hoc" | search user="**" | stats count median(total_run_time) as median_run_time max(total_run_time) as max_runtime values(user) as user by search | eval median runtime=if(isnotnull(median runtime), median runtime, "+")) | eval max runtime=if(isnotnull(max runtime), max runtime, "+") | sort - count | rename search as "Search", count as "Count", median runtime as "Median Runtime", max runtime as "Max Runtime", user as User | fieldformat "Median Runtime" = `dmc_convert_runtime('Median Runtime')` | fieldformat "Max Runtime" = `dmc_convert_runtime('Max Runtime')`
```
n00b: Debug This
Keyboard Command: Ctrl + \ or Command + \
Search Interface Improvements
user-prefs.conf with export = system

Default

**Search**
Use these properties for assistance with command syntax including example in different colors.

- **Search assistant**
  - Compact
  - Full
  - None

- **Syntax highlighting**
  - Light theme

**Search auto-format**
- On
- Off

**Show line numbers**
- On
- Off

Suggestion

**Search**
Use these properties for assistance with command syntax including example in different colors.

- **Search assistant**
  - Compact
  - Full
  - None

- **Syntax highlighting**
  - Dark theme

**Search auto-format**
- On
- Off

**Show line numbers**
- On
- Off
# SHC Need 2 Knows

“So…I can’t just treat it like a Deployment Server?!”

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Caveats</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ Deployer not critical path</td>
<td>▶ Min 3+ SHs</td>
</tr>
<tr>
<td>▶ Config -&gt; default</td>
<td>• Odd number for consensus</td>
</tr>
<tr>
<td>▶ More effective hardware utiliz.</td>
<td>▶ Same specs</td>
</tr>
<tr>
<td>▶ Eliminates dedicated alerting SHs</td>
<td>▶ No manual conf edits on SHs</td>
</tr>
<tr>
<td>• A.K.A. Job Servers</td>
<td>• Split Brain</td>
</tr>
</tbody>
</table>

---

**Benefits**

- Deployer not critical path
- Config -> default
- More effective hardware utiliz.
- Eliminates dedicated alerting SHs
  - A.K.A. Job Servers

**Caveats**

- Min 3+ SHs
  - Odd number for consensus
- Same specs
- No manual conf edits on SHs
  - Split Brain
Search Head limits.conf

Example:

[scheduler]
max_searches_perc
auto_summary_perc

shc_role_quota_enforcement
shc_syswide_quota_enforcement

Benefit

▶ Defaults to 50%
▶ Ad Hoc takes precedent regardless
▶ Additional controls for scheduling

▶ Quota cluster wide
  • Default is instance specific
Indexing Tier

Trivia: What does an indexer do?
Cluster of One
“We lost that data even though we had replication”

Benefits

▶ “Retroactive” data replication

▶ No additional disk
  • If factors are still 1

▶ summary_replication

Challenges

▶ ONLY IF YOU PLAN TO NEED REPLICATION
  • Multisite
  • Long Retention Times

▶ Administratively difficult
  • Higher chance of errors
  • Conceptually abstract
Indexer Discovery
Search docs.splunk.com for “indexerdiscovery”

Pros
- Dynamic indexer listings
- indexerWeightByDiskCapacity
  - Indexers with different volume sizes

Cons
- Requires network traffic to master node
  - Forwarder silence if master down @ start
- Total Disk != Free Space
- Lead to uneven data distribution
Data Rebalance
Search docs.splunk.com for “Rebalance the indexer cluster”

Data Rebalance

Threshold? 0.9

Max Runtime? optional

Index? All Indexes

Data was last rebalanced 0 days and 0 hours and 0 minutes ago

[Cancel] [Start]
Index Definitions

[volume:home]
path = $SPLUNK_DB
maxVolumeDataSizeMB =

[volume:cold]
path = $SPLUNK_DB
maxVolumeDataSizeMB =

[default]
homePath = volume:home/$_index_name/db
coldPath = volume:cold/$_index_name/colddb
thawedPath = $SPLUNK_DB/$_index_name/thaweddb

[newindex]

Let’s talk about...

- volume:
  - maxVolumeDataSizeMB
    - Indexes compete for storage

- [default]

- [newindex]

- $_index_name
Splunk Data Life Cycle: Determining When and Where to Roll Data

Wednesday, September 27, 2017 | 11:00 AM-11:45 AM

Jeff Champagne, Staff Architect, Splunk Inc.

Splunk has many options for managing data via hot/warm and cold paths, freezing, roll to HDFS, and TSIDX reduce. These features can impact your search performance, retention and resiliency. This session will provide you with an in-depth understanding of the Splunk data life cycle options and how to determine which will work best in your environment.
Security Through Obscurity

- Change default ports
- Change default system account ($SPLUNK_HOME/etc/default/user-seed.conf)

Auditable Logins

- Empty $SPLUNK_HOME/etc/passwd and $SPLUNK_HOME/etc/.ui_login
- Distribute authentication.conf

“Best practices for Splunk Enterprise security” in docs.splunk.com
Wrap Up

1. User Management
2. Data Onboarding
3. Splunk Health
4. Config Management
5. App & TA Creation
6. Architecture
7. Search Tier
8. Indexing Tier
9. Securing Splunk
1. Rate this! (be honest)

2. Collaborate: #bestpractices
   • Sign Up @ http://splk.it/slack

3. Customer Success Studio

4. More talks, search for
   • Blueprints
   • Burch
   • Champagne
   • Delaney
   • Optimization
   • Best Practices
   • Veuve
Questions & Discussion?

Don't forget to rate this session in the .conf2017 mobile app