Best Practices and Better Practices for Users
...while you get settled...

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

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

Load Feedback ------------------------------->
Best Practices and Better Practices for Users

Presented by Splunk Blueprints

Burch | Senior Best Practices Engineer

.conf2017 | Version 0.0
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
1. How I Learned

2. Searching
   • Pretty Searches
   • Search Performance
   • Accuracy

3. Evolved Ideas
How I Learned

Now this is a story all about how…
Search Tutorial
Free Search Tutorial -> docs.splunk.com -> Search Tutorial

- Downloads & Installs Splunk
- Add tutorial data
- Local sandbox
Quick Reference Guide

Search “splunk quick reference guide”

**Concepts**

**Events**
An event is a set of values associated with a timestamp. It is a single entry of data and can have one or multiple lines. An event can be a text document, a configuration file, an entire stack trace, and so on. This is an example of an event in a web activity log:

```
```

You can also define transactions to search for and group together events that are conceptually related but span a duration of time. Transactions can represent a multi-step business-related activity, such as all events related to a single customer session on a retail website.

**Host, Source, and Source Type**
A host is the name of the physical or virtual device on which an event originates. The host field provides an easy way to find all data originating from a specific device. A source is the name of the file, directory, data stream, or other input from which a particular event originates. Sources are classified into source types, which can be either well known formats or formats defined by the user. Some common source types are HTTP server logs and Windows event logs.

Events with the same source type can come from different sources. For example, events from the file `/var/log/nginx/messages` and from a syslog input port `/opt/standalone` often share the source type `source_type=linux_xlogy`.

**Index-Time and Search-Time**
During index-time processing, data is read from a source on a host and is classified into a source type. Timestamps are extracted, and the data is parsed into individual events. Line-breaking rules are applied to segment the events to display in the search results. Each event is written to an index on disk, where the event is later retrieved with a search request.

When a search starts, referred to as search-time, indexed events are retrieved from disk. Fields are extracted from the raw text for the event.

**Core Features**

**Search**
Search is the primary way users navigate data in Splunk software. You can write a search to retrieve events from an index, use statistical commands to calculate metrics and generate reports, search for specific conditions within a rolling time window, identify patterns in your data, predict future trends, and so on. You transform the events using the Splunk Search Process Language (SPL). Searches can be saved.

**Additional Features**

**Data Model**
A data model is a hierarchically-organized collection of datasets that Pivot uses to generate reports. Data model objects represent individual datasets, which the data model is composed of.

**Pivot**
Pivot refers to the table, chart, or other visualization you create using the Pivot Editor. You can map attributes defined by data model objects to data visualizations, without manually writing the searches. Pivots can be saved as reports and used to power dashboards.

**Apps**
Apps are a collection of configurations, knowledge objects, and customer designed views and dashboards. Apps extend the Splunk environment to fit the specific needs of organizational teams such as Unix or Windows system administrators, network security specialists, website managers, business analysts, and so on. A single Splunk Enterprise or Splunk Cloud installation can run multiple apps simultaneously.

**Distributed Search**
A distributed search provides a way to scale your deployment by separating the search management and presentation layer from the indexing and search retrieval layer. You use search to facilitate horizontal scaling for enhanced performance, to control access to indexed data, and to manage geographically
Commands by category

The following tables list all the search commands, categorized by their usage. Some commands fit into more than one category based on the options that you specify.

Correlation

These commands can be used to build correlation searches.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>append</td>
<td>Appends subsearch results to current results.</td>
</tr>
<tr>
<td>appendcols</td>
<td>Appends the fields of the subsearch results to current results, first results to first result, second to second, etc.</td>
</tr>
<tr>
<td>appendpipe</td>
<td>Appends the result of the subpipeline applied to the current result set to results.</td>
</tr>
<tr>
<td>arules</td>
<td>Finds association rules between field values.</td>
</tr>
<tr>
<td>associate</td>
<td>Identifies correlations between fields.</td>
</tr>
<tr>
<td>contingency, counttable, ctable</td>
<td>Builds a contingency table for two fields.</td>
</tr>
<tr>
<td>correlate</td>
<td>Calculates the correlation between different fields.</td>
</tr>
<tr>
<td>diff</td>
<td>Returns the difference between two search results.</td>
</tr>
<tr>
<td>join</td>
<td>Combines the results from the main results pipeline with the results from a subsearch.</td>
</tr>
<tr>
<td>lookup</td>
<td>Explicitly invokes field value lookups.</td>
</tr>
</tbody>
</table>
Splunk! The Book
www.splunk.com/goto/book

Exploring Splunk
SEARCH PROCESSING LANGUAGE (SPL)
PRIMER AND COOKBOOK

By David Carasso, Splunk’s Chief Mind
Splunk User Groups
usergroups.splunk.com

Find a Splunk User Group

Connect with like-minded people who are passionate about Splunk technology
Free Education!
splunk.com/education

Splunk Fundamentals 1

Take this course, get Splunk User certified, and be eligible to win up to $4500!

Once you complete and pass this course, you are eligible to take the Splunk Certified User certification exam. The person or persons with the highest score on the certification exam by August 31st, 2017 will win the prize. In the event of a tie the $4,500 will be shared among the winners. Must not have any other Splunk Certification. One passing entry per person. Competition closes on August 31st, 2017.

This course teaches you how to search and navigate in Splunk, use fields, get statistics from your data, create reports, dashboards, lookups, and alerts. Scenario-based examples and hands-on challenges will enable you to create robust searches, reports, and charts. It will also introduce you to Splunk’s datasets features and Pivot interface.
But now I have questions...
answers.splunk.com
Community Q&A

- E-mail notifications
- Fast answers
- Larger distribution
Searching
Pretty Searches

I feel pretty, oh so pretty...
Keep it kosher
Params, cosmetics, combine (rename, evals)

Weak:

... | rename machine as “host for later” | rename net as Subnet | sort “host for later” | timechart count by “host for later” span=1h

Strong:

...| timechart span=1h count by machine | sort machine | rename machine AS “host for later”, net AS Subnet
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_runtime 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

### 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
What the?!
Speaking of UI...

Soooo sssllllloooowwww

Dude! Where’s my fields?!
Search Mode
Speaking of UI…

Soooo sssllllloooowwww
Dude! Where’s my fields?!
Event Types & Tags

Weak:

```
index=oidemo host=dmzlog.splunktel.com sourcetype=access_combined
source=/opt/apache/log/access_combined.log iphone
user_agent="*iphone*"
| stats count by action
tag=iphone_event
```

or

```
eventtype=web_logs
```

Strong:

tag=iphone_event

or

```
eventtype=web_logs
```
Dereference Finesse

Weak:

index=internal
| eval ERROR = case( log_level == "ERROR" , message )
| eval WARN = case( log_level == "WARN" , message )
| eval INFO = case( log_level == "INFO" , message )

Strong:

index=_internal
| eval {log_level} = message

Selected Fields
- ERROR 100+
- INFO 100+
- WARN 100+
- WARNING 92
Pretty Searches: coalesce’s cooler than if

Weak:
...
| eval size = if( isnull(bytes) , if( isnull(b) , "N/A" , b ) , bytes )

Strong:
...
| eval size = coalesce( bytes , b , "N/A" )
Macros
Keyboard Shortcut: Control-Shift-E or Command-Shift-E

Repeatable Code

Definition *
Enter the string the search macro expands to when it is referenced in another

```
"HandleInfoMaxTime` | head _time>($info_max_time - $alert_value_window=false<($info_max_time-$search_$entity_statop=$threshold_field) AS alert_value b $service_statop=alert_value) AS alert_value by alert_value_window="current_window" | eval window_d "none") | `gettime`
```

Use eval-based definition?

Arguments
Enter a comma-delimited string of argument names. Argument names may be:

- entity_statop
- service_statop
- threshold_field
- entity_field
- search_alert

Expand in UI

```
1  index=internal host=leah source=license_usage.log
  | eval _time=time - 43200
  | bin_time span=1d
  | stats latest(b) AS b by slave, pool, _time
  | timechart span=1d sum(b) AS "volume" fixedrange=false
  | join type=outer _time
  | [ search index=internal host=leah source=license_usage.log
  |   | eval _time=time - 30d
  |   | bin_time span=1d
  |   | stats latest(b) AS b by slave, pool, _time
  |   | timechart span=1d sum(b) AS "volume" fixedrange=false
  |   | join type=outer _time
```
Time and Units

Weak:

…| eval new_time = <ridiculous string edits>

Strong:

…| convert ctime(duration)  …| bin span=1h _time
…| eval pause = tostring( pause , “duration” )
…| rename new_time as _time
Search Performance
He's a demon on wheels
Search Performance Improvement

docs.splunk.com “Search Job Inspector”

- events per second = events / seconds
- results per second = results / seconds

This search has completed and has returned **1,000** results by scanning **22,696** events in **1.049** seconds.
Less is more

Weak:

iphone
  | stats count by action
  | search action=AppleWebKit

Strong:

iphone action=AppleWebKit
  | stats count
NOT NOTs

Weak:

\[ \text{index=burch NOT blah=yay blah=cool} \]

Strong:

\[ \text{index=burch blah=duh} \quad \text{index=burch blah!=yay} \]

\[ \text{Implies ( blah!=yay blah=* )} \]
stats vs dedup/transaction

Weak:

... phone=* 
  | dedup phone 
  | table phone 
  | sort phone 

Strong:

... phone=* 
  | stats count by phone, host 
  | fields - count 

Pro Tip:

• Table is cosmetic 
• Fields is reducing
Avoid Subsearches

Weak:

index=burch | eval blah=yay
| append [ search index=simon | eval blah=duh ]

Strong:

( index=burch ... ) OR ( index=simon ... )
| eval blah=case( index=="burch" , "yay" , index=="simon" , "duh" )

(format and return commands for returning results)
foreach FTW!

Weak:

...| timechart span=1h limit=0 sum(eval(b/pow(1024,3))) as size by st

Strong:

...| timechart span=1h limit=0 sum(b) by st
| foreach * [ eval <<-FIELD>> = ' <<-FIELD>> / pow(1024, 3) ]
Weak:

... | transaction host

Strong:

... | transaction maxspan=10m maxxevents=100 ...
metadata

Weak:

index=*  
| stats count by host

Strong:

| metadata index=* type=hosts
eventcount

Weak:

index=*  
  | stats count by index

Strong:

  | eventcount summarize=false index=*

splunk>  .conf2017
Dashboard Performance

If your dashboard...

- ...has many similar searches
- ...is viewed by many
- ...is viewed by few

Then use...

- “Post-process”
- Scheduled report (cache)
- Inline Searches
Pretty SimpleXML
Keyboard Command: CTRL + Shift + F or Command + Shift + F
Metrics Explorer for Splunk

https://splunkbase.splunk.com/app/3726/
Accuracy

A redshirt and a Stormtrooper get into a firefight.
The Stormtrooper misses every shot.
The redshirt dies anyway.
Require Fields

Weak:

iphone
| stats count

Strong:

phone=iphone
| stats count

Wrong Results:
Pulls both phone=iphone and user_agent=*iphone*

Remember:
‘iphone’ is not the same as ‘iphone6s’
Be specific
Time Selector!

Weak:

iphone
  | stats count by action

Strong:

index=oidemo host=dmzlog.splunktel.com
sourcetype=access_combined
source=/opt/apache/log/access_combined.log iphone
user_agent="*iphone*"
  | stats count by action
SPL Testing
Regex, strptime, etc...

Weak:

index=*  
| head  
| eval blah = "try to do stuff"

...  

Strong:

| makeresults  
| eval blah = "try to do stuff"

...
Create New Fields?!
rex different for each person

Interesting Fields
a component 49
# date_hour 1
# date_mday 1
# date_minute 16
a date_month 1
# date_second 60
a date_wday 1
# date_year 1
# date_zone 1
a eventtype 3
a group 37
a idx 100+
a index 1
# ingest_pipe 2
Extract Fields

Search: docs.splunk.com field extractor
Play it Safe

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!
### Hidden Fields: Time

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>
</tbody>
</table>

What does a big difference mean?
### Hidden Fields: Time

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>
Snap-To Times

Weak:

- **Time range**
  - Start time: `-60min`
  - Finish time: 
  - Time specifiers: `y, mon, d, h, m, s`

- **Acceleration**
  - Accelerate this search

- **Schedule and alert**
  - Schedule this search

- **Schedule type**
  - Basic

- **Run every**
  - `hour`

Strong:

- **Time range**
  - Start time: `@hour-1hour`
  - Finish time: `@hour`
  - Time specifiers: `y, mon, d, h, m, s`

- **Acceleration**
  - Accelerate this search

- **Schedule and alert**
  - Schedule this search

- **Schedule type**
  - Basic

- **Run every**
  - `hour`
Time Fields

Weak:

Search
earliest=-24hours latest=now

Strong:

Time range
Start time
@hour-1hour
Finish time
@hour

Acceleration
Accelerate this search

Schedule and alert
Schedule this search

Schedule type *
Basic

Run every *
hour
Alerts

Blueprints for Actionable Alerts

Wednesday, September 27, 2017 | 3:30 PM-4:15 PM

Burch !, Senior Best Practices Engineer, Splunk Inc.

Do you receive too many alerts from your Splunk environment and don't know which to focus on? Do you have so many alerts that you no longer see through the noise? Do you fear that your Splunk investment is losing its purpose and value because users have no choice but to ignore it? I’ve been there. I inherited a system like that. This is an updated version of the popular session from .conf2016 covering the evolution of how I improved those alerts and shifted Splunk from spam to glam. Come to this session to learn from my experiences and approaches, which will provide you with more confidence and actionable alerts.
Evolved Ideas
## Acceleration Options

Knowledge Manager Manual > Use data summaries to accelerate searches > Manage report acceleration

<table>
<thead>
<tr>
<th></th>
<th>Summary Indexing</th>
<th>Report Acceleration</th>
<th>Data Model Acceleration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benefits</strong></td>
<td>• Save disk space</td>
<td>• Backfill</td>
<td>• Backfill</td>
</tr>
<tr>
<td></td>
<td>• Control on impact to system</td>
<td>• Simple</td>
<td>• Simple</td>
</tr>
<tr>
<td></td>
<td>• Backfill</td>
<td></td>
<td>• Extensible</td>
</tr>
<tr>
<td></td>
<td>• Simple</td>
<td></td>
<td>• Search Agnostic</td>
</tr>
<tr>
<td><strong>Limits</strong></td>
<td>• Gaps</td>
<td>• Requires transforming</td>
<td>• Massive if misused</td>
</tr>
<tr>
<td></td>
<td>• Intellectually difficult</td>
<td>• Specific to search</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Backfill</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Don’t Scare Your Admins
Impress Them!

- Accelerations
- Scheduled Searches
- Real Time Searches
- Search Limits

Selection of Impacting Capabilities

<table>
<thead>
<tr>
<th>Group</th>
<th>Capabilities</th>
<th>Why</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accelerations</td>
<td>accelerate_datamodel</td>
<td>Compute &amp; Storage costs</td>
</tr>
<tr>
<td></td>
<td>accelerate_search</td>
<td>Rarely cleaned up</td>
</tr>
<tr>
<td></td>
<td>output_file</td>
<td></td>
</tr>
<tr>
<td>Scheduled Searches</td>
<td>schedule_search</td>
<td>Compute and concurrent load</td>
</tr>
<tr>
<td></td>
<td>schedule_rsearch</td>
<td>Rarely cleaned up</td>
</tr>
<tr>
<td>Real Time Searches</td>
<td>rsearch</td>
<td>Rarely necessary</td>
</tr>
<tr>
<td></td>
<td>schedule_rsearch</td>
<td>Impact on SH + ALL Indexers</td>
</tr>
<tr>
<td>Search Limits</td>
<td>srcrJobsQuota</td>
<td>Boundaries</td>
</tr>
<tr>
<td></td>
<td>srcrMaxTime</td>
<td>Careful, could be annoying</td>
</tr>
<tr>
<td></td>
<td>srcrTimeWin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>srcrDiskQuota</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rtSrchJobsQuota</td>
<td></td>
</tr>
</tbody>
</table>
Medium Matters

There are known knowns; there are things we know that we know.

There are known unknowns; that is to say, there are things that we now know we don’t know.

But there are also unknown unknowns – there are things we do not know we don’t know.

-Donald Rumsfeld
# Search Consumption

## Alerts vs Dashboards vs Searches

<table>
<thead>
<tr>
<th></th>
<th>Root Cause Unknown</th>
<th>Root Cause Known</th>
</tr>
</thead>
</table>
| **Unaware Issue Exists** | **Screens (Dashboards/GT)**  
  - Changes in behavior  
  - Data Driven KPIs  
  - Notice patterns | **Alerts until Fixed**  
  - Monitor for known symptoms  
  - Adaptive Response  
  - Actionable |
| **Aware Issue Exists**  | **Investigations (SPL)**  
  - Go Spelunking! | -                                                          |

---

*Image: Splunk Inc.*
# Search Consumption

## Alerts vs Dashboards vs Searches

<table>
<thead>
<tr>
<th>Unaware Issue Exists</th>
<th>Root Cause Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screens (Dashboards/GT)</td>
<td>- Changes in behavior</td>
</tr>
<tr>
<td></td>
<td>- Data Driven KPIs</td>
</tr>
<tr>
<td></td>
<td>- Notice patterns</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aware Issue Exists</th>
<th>Root Cause Known</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigations (SPL)</td>
<td>- Monitor for known symptoms</td>
</tr>
<tr>
<td></td>
<td>- Adaptive Response</td>
</tr>
<tr>
<td></td>
<td>- Actionable</td>
</tr>
</tbody>
</table>

- Go Spelunking!

- Alerts until Fixed
## Search Consumption

**Alerts vs Dashboards vs Searches**

<table>
<thead>
<tr>
<th>Unaware Issue Exists</th>
<th>Root Cause Unknown</th>
<th>Root Cause Known</th>
</tr>
</thead>
</table>
| Screens (Dashboards/GT) | • Changes in behavior  
  • Data Driven KPIs  
  • Notice patterns | ![Image] |
| Aware Issue Exists | Investigations (SPL) | Alerts until Fixed |
| • Go Spelunking! |
| • Monitor for known symptoms  
  • Adaptive Response  
  • Actionable |

---

**Splunk > .conf2017**
1. How I Learned

2. Searching
   - Pretty Searches
   - Search Performance
   - Accuracy

3. Evolved Ideas
What Now?

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
Search with Burch: Sourcetypes
Search with Burch: Index