



# Using Splunk Enterprise To Optimize Tailored Long-term Data Retention

Tomasz Bania | Incident Response Lead, Dolby

Eric Krieser | Splunk Professional Services

September 2017 | Washington, DC

# 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.

# Overview

## ► Problems To Address

- Limited Retention Capabilities
- Search Completion Time for Extended Searches

## ► Goals

- Maintain key event fields for a specific retention period
- Maintain the smallest footprint possible for archiving this retained content

## ► Challenges

- Storage Budget Limitations
- Significant Bloat in Current Data Set

# Considerations

- ▶ Standard summaries using stats, and sistats adds key=value to raw event causing the summary events to use more space than the original raw event in many cases
- ▶ Standard summaries using stats, and sistats create very heavy tsidx indexes causing additional overhead
- ▶ Standard summaries add additional field content (metrics etc.) to the summary event causing additional bloat
- ▶ Multiple sourcetypes are many times stored across different indexes defined reduced and disparate retention settings

```
130.60.4 - - [07/Jan 18:10:57:153] "GET /category.screen?category_id=GIFTS&JSESSIONID=SD1SL4FF10ADFF10 HTTP 1.1" 404 720 "http://buttercup-shopping.com/cart.do?action=view&itemId=EST-6&product_id=FL-SW-01" "Opera/9.80.20130.0 Windows NT 5.1; SV1: ; [07/Jan 18:10:57:123] "GET /product.screen?product_id=FL-DSH-01&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 404 3322 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product_id=MEX11A74-0" "Chrome/28.0.1467.0" [07/Jan 18:10:56:156] "GET /oldlink?item_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1318 "http://buttercup-shopping.com/cart.do?action=changequantity&itemId=EST-18&product_id=AV-CB-01&JSESSIONID=SD10SL0F2ADFF9 HTTP 1.1" 200 2423 "http://buttercup-shopping.com/cart.do?action=remove&itemId=EST-6&JSESSIONID=SD10SL0F2ADFF9" [07/Jan 18:10:56:189] "GET /category.screen?category_id=FLOWERS&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 200 3865 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product_id=MEX11A74-0" [07/Jan 18:10:56:187] "GET /category.screen?category_id=FLOWERS&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 200 3865 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product_id=MEX11A74-0" [07/Jan 18:10:56:188] "GET /category.screen?category_id=FLOWERS&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 200 3865 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product_id=MEX11A74-0"
```

# Bloat In Standard Summary Event

## ► Original Source Event:

```
Jun 22 12:07:04 1,2017/06/22 12:07:03,001606001116,THREAT,url,1,2017/06/22
12:07:03,192.168.0.2,184.106.31.170,0.0.0.0,0.0.0.0,rule1,crusher,,web-
browsing,vsys1,trust,untrust,ethernet1/2,ethernet1/1,blocked,2017/06/22
12:07:03,37148,1,52586,80,0,0,0x208000,tcp,alert,"modern-design.cn/rex/config.bin",
(9999),not-resolved,informational,client-to-server,0,0x0,192.168.0.0-192.168.255.255,United States,0,text/html
```

```
index=main sourcetype=pan:threat | fillnull value "nan" action dest_ip src_ip | eval orig_sourcetype=sourcetype | sistats
count by _time, host, orig_sourcetype, action, dest_ip, src_ip | fields - count | collect index=summary addtime=0
```

```
index=main sourcetype=pan:threat | fillnull value "nan" action dest_ip src_ip | eval orig_sourcetype=sourcetype | fields _time, host, orig_sourcetype, action, dest_ip, src_ip | collect index=summary addtime=0
```

# Desired Results

## ► Output using sistats/table statement results:

```
search_name=sum_panthreat_sistats, search_now=1498155300.000, info_min_time=1498154880.000,
info_max_time=1498155180.000, info_search_time=1498155301.125, action=blocked, dest_ip="194.106.31.170",
orig_host="127.0.0.1", orig_sourcetype="pan:threat", psrsvd_gc=1, psrsvd_v=1, src_ip="192.168.0.2"
```

## ► Output using fields statement results:

```
Jun 22 11:57:59 1,2017/06/22 11:57:59,001606001116,THREAT,url,1,2017/06/22
11:57:59,192.168.0.2,184.106.31.170,0.0.0.0,0.0.0.0,rule1,crusher,,web-
browsing,vsys1,trust,untrust,ethernet1/2,ethernet1/1,forwardAll,2017/06/22
11:57:59,53995,1,56068,80,0,0,0x208000,tcp,alert,"www.st-resources.net/config.bin",(9999),not-
resolved,informational,client-to-server,0,0x0,192.168.0.0-192.168.255.255,United States,0,text/html
```

## ► Desired output:

```
1498154279,127.0.0.1,pan:threat,blocked,192.168.0.2,184.106.31.170
```

130.60.4 - - [07/Jun 18:10:57:153] "GET /category.screen?category\_id=GIFTS&JSESSIONID=SD1SLAFF10ADFF10 HTTP 1.1" 404 720 "http://buttercup-shopping.com/cart.do?action=view&itemId=EST-6&product\_id=FL-SW-01" "Opera/9.80.2013.10.1.0 (Windows NT 6.0; rv:15.0) Gecko/20100101 Firefox/15.0.1"

128.241.220.82 - - [07/Jun 18:10:57:123] "GET /product.screen?product\_id=FL-DSH-01&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 404 3322 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product\_id=MX-11-01" "Mozilla/5.0 (Windows NT 6.0; rv:15.0) Gecko/20100101 Firefox/15.0.1"

317 27.160.0.0 - - [07/Jun 18:10:56:156] "GET /oldlink?item\_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1316 "http://buttercup-shopping.com/cart.do?action=changequantity&itemId=EST-18&product\_id=AV-CB-01&JSESSIONID=SD10SL9FF2ADFF9 HTTP 1.1" 200 2423 "http://buttercup-shopping.com/cart.do?action=remove&itemId=EST-1&JSESSIONID=SD5SL8FF1ADFF6 HTTP 1.1" 200 3865 "http://buttercup-shopping.com/oldlink?item\_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1316 "http://buttercup-shopping.com/cart.do?action=changequantity&itemId=EST-18&product\_id=AV-CB-01&JSESSIONID=SD10SL9FF2ADFF9 HTTP 1.1" 200 2423 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product\_id=MX-11-01" "Mozilla/5.0 (Windows NT 6.0; rv:15.0) Gecko/20100101 Firefox/15.0.1"

130.60.4 - - [07/Jun 18:10:57:153] "GET /category.screen?category\_id=GIFTS&JSESSIONID=SD1SLAFF10ADFF10 HTTP 1.1" 404 720 "http://buttercup-shopping.com/cart.do?action=view&itemId=EST-6&product\_id=FL-SW-01" "Opera/9.80.2013.10.1.0 (Windows NT 6.0; rv:15.0) Gecko/20100101 Firefox/15.0.1"

128.241.220.82 - - [07/Jun 18:10:57:123] "GET /product.screen?product\_id=FL-DSH-01&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 404 3322 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product\_id=MX-11-01" "Mozilla/5.0 (Windows NT 6.0; rv:15.0) Gecko/20100101 Firefox/15.0.1"

317 27.160.0.0 - - [07/Jun 18:10:56:156] "GET /oldlink?item\_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1316 "http://buttercup-shopping.com/cart.do?action=changequantity&itemId=EST-18&product\_id=AV-CB-01&JSESSIONID=SD10SL9FF2ADFF9 HTTP 1.1" 200 2423 "http://buttercup-shopping.com/cart.do?action=remove&itemId=EST-1&JSESSIONID=SD5SL8FF1ADFF6 HTTP 1.1" 200 3865 "http://buttercup-shopping.com/oldlink?item\_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1316 "http://buttercup-shopping.com/cart.do?action=changequantity&itemId=EST-18&product\_id=AV-CB-01&JSESSIONID=SD10SL9FF2ADFF9 HTTP 1.1" 200 2423 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-26&product\_id=MX-11-01" "Mozilla/5.0 (Windows NT 6.0; rv:15.0) Gecko/20100101 Firefox/15.0.1"



# Implementing The Desired Result

## ► Constructing a summary search using a CSV output:

```
index=main sourcetype=pan:threat | fillnull value "nan" action dest_ip src_ip | eval orig_sourcetype=sourcetype | eval
_raw= _time.", ".host.", ".orig_sourcetype.", ".action.", ".dest_ip.", ".src_ip | fields _time _raw | collect index=summary
addtime=0
```

## ► Build the summary Field Definition to Define Your Fields

props.conf

```
[source::sum_panthreat_csv]
KV_MODE=none
REPORT-parse_sum_panthreat_csv = parse_sum_panthreat_csv
```

transforms.conf

```
[parse_sum_panthreat_csv]
DELIMS = ","
FIELDS = orig_time, orig_host, orig_sourcetype, action, dest_ip, src_ip
```



# Final Result Of Bloat Reduction

Extracted Fields View

6/22/17 12:32:59.000 PM 1498156379,127.0.0.1,pan:threat,blocked,192.168.0.2,74.125.224.195

Event Actions ▾

Type	<input checked="" type="checkbox"/>	Field	Value	Actions
Selected	<input checked="" type="checkbox"/>	host ▾	ekrieser-mbp2	▾
	<input checked="" type="checkbox"/>	source ▾	sum_panthreat_csv	▾
	<input checked="" type="checkbox"/>	sourcetype ▾	stash	▾
Event	<input type="checkbox"/>	action ▾	blocked	▾
	<input type="checkbox"/>	dest_ip ▾	192.168.0.2	▾
	<input type="checkbox"/>	index ▾	summary	▾
	<input type="checkbox"/>	linecount ▾	1	▾
	<input type="checkbox"/>	orig_host ▾	127.0.0.1	▾
	<input type="checkbox"/>	orig_sourcetype ▾	pan:threat	▾
	<input type="checkbox"/>	orig_time ▾	1498156379	▾
	<input type="checkbox"/>	splunk_server ▾	ekrieser-mbp2	▾
	<input type="checkbox"/>	src_ip ▾	74.125.224.195	▾
Time +		_time ▾	2017-06-22T12:32:59.000-06:00	

# Review Of Key Points To The Solution

- Use summary searching to consolidate critical data stored across disparate sourcetypes and indexes
- Construct a single field called `_raw` that contains the desired summary content in some type of a CSV or character delimited form
- Pass `_time`, and `_raw` to the summary index
- Define the field format of the summary event using `DELIM`, and `FIELDS` transforms definitions

# Things To Take Into Consideration

- ▶ Parsing fields with equal (=) signs can be a pain
- ▶ Take particular care in defining delimiter to use (CSV may not be the best in your use case)

# Average Savings Stats



72% Reduction In  
Storage Volume



26% Faster Searches



Over 3x Longer Retention  
at Zero Additional Cost

130.60.4 - - [07/Jan 18:10:57:153] "GET /category.screen?category\_id=GIFTS&JSESSIONID=SD1SL4FF10ADFF10 HTTP 1.1" 404 720 "http://buttercup-shopping.com/cart.do?action=view&itemId=EST-6&product\_id=FL-SW-01" "Opera/9.80.2013.10 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 128.241.220.82 - - [07/Jan 18:10:57:123] "GET /product.screen?product\_id=FL-DSH-01&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 404 3322 "http://buttercup-shopping.com/category.screen?category\_id=GIFTS" "Mozilla/5.0 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 317.27.160.0 - - [07/Jan 18:10:56:156] "GET /oldlink?item\_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1316 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-16&product\_id=RP-LI-02" "Opera/9.80.2013.10 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 130.60.4 - - [07/Jan 18:10:57:153] "GET /category.screen?category\_id=GIFTS&JSESSIONID=SD1SL4FF10ADFF10 HTTP 1.1" 404 720 "http://buttercup-shopping.com/cart.do?action=view&itemId=EST-6&product\_id=FL-SW-01" "Opera/9.80.2013.10 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 128.241.220.82 - - [07/Jan 18:10:57:123] "GET /product.screen?product\_id=FL-DSH-01&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 404 3322 "http://buttercup-shopping.com/category.screen?category\_id=GIFTS" "Mozilla/5.0 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 317.27.160.0 - - [07/Jan 18:10:56:156] "GET /oldlink?item\_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1316 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-16&product\_id=RP-LI-02" "Opera/9.80.2013.10 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 130.60.4 - - [07/Jan 18:10:57:153] "GET /category.screen?category\_id=GIFTS&JSESSIONID=SD1SL4FF10ADFF10 HTTP 1.1" 404 720 "http://buttercup-shopping.com/cart.do?action=view&itemId=EST-6&product\_id=FL-SW-01" "Opera/9.80.2013.10 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 128.241.220.82 - - [07/Jan 18:10:57:123] "GET /product.screen?product\_id=FL-DSH-01&JSESSIONID=SD5SL7FF6ADFF9 HTTP 1.1" 404 3322 "http://buttercup-shopping.com/category.screen?category\_id=GIFTS" "Mozilla/5.0 (Windows NT 6.0; WOW64; rv:31.0) like Gecko" 317.27.160.0 - - [07/Jan 18:10:56:156] "GET /oldlink?item\_id=EST-26&JSESSIONID=SD5SL9FF1ADFF3 HTTP 1.1" 200 1316 "http://buttercup-shopping.com/cart.do?action=purchase&itemId=EST-16&product\_id=RP-LI-02" "Opera/9.80.2013.10 (Windows NT 6.0; WOW64; rv:31.0) like Gecko"



# Q&A

---

Tomasz Bania | Incident Response Lead, Dolby  
Eric Krieser | Splunk Professional Services

# Thank You

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

splunk> .conf2017