.conf2015

COMCAST

To Xfinity and Beyond:
Mission Critical Metrics and Tips For Managing Any Size Splunk Installation

Kate Lawrence-Gupta & Joe Cramasta
Comcast -- Technology & Product Division

splunk>

Disclaimer

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 the 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.

Lineup

- Intros
- Overview of Comcast/Splunk deployment
- What to Measure & Why
- Planning Ahead
- Measuring Capacity/Utilization

- Using REST commands
- Detecting Latency
- Dealing With High Volume Sources
- Wrap up
- Q & A

Kate & Joe

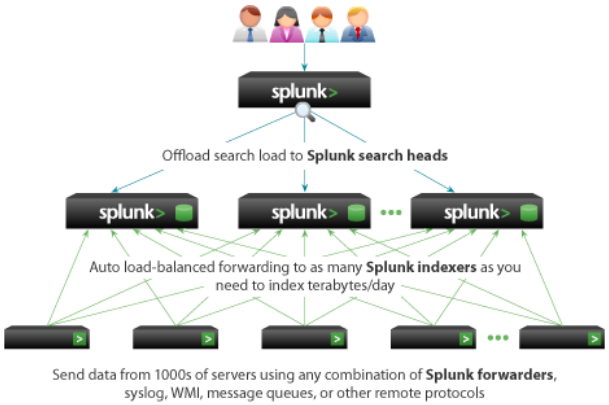
- Managing & Senior Engineer responsible for several Splunk installations at Comcast
- Run a dedicated team providing Splunk as an operational service
- Between the 2 of us over 25 years of experience in operations, monitoring, and systems administration

Comcast Overview

- Global media and technology company consisting of Comcast Cable and NBC Universal
- Comcast Cable: Nation's largest video, hi-speed internet and phone provider under the XFINITY brand
- Creator of the X1 Entertainment Operating System & XFINITY Home Security System
- NBC Universal: One of world's leading media and entertainment companies

Splunk Deployment

- Splunk is “critical path”
- Several thousands of forwarders sending data to Splunk.
- Supporting over a 1000 Splunk Users
- Dedicated Team of 4 Splunk Admins
- Splunk runs on dedicated hardware & storage across multiple datacenters
- 99.95% uptime & less than 10 seconds of indexing latency



Use Splunk to Measure Splunk

How?

By Trending!

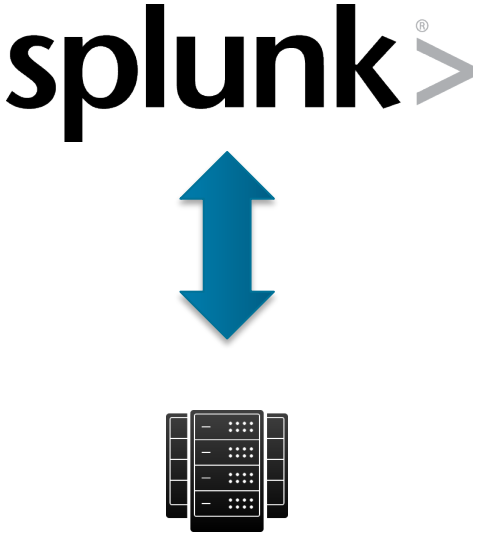
CPU Utilization

License Volume Trends

OVER TIME

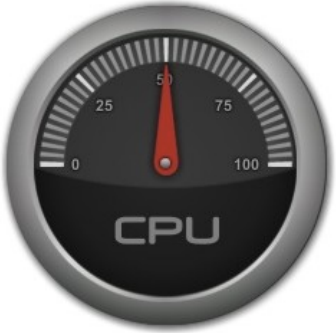
Getting Started

- Setup a Management Search Head
- Peered to:
  - Indexers
  - Search Heads
  - Deployment & Cluster Managers
  - Heavy Forwarders



What to Measure?
Start with the CPU…

- Measure your CPU performance at the individual host level.



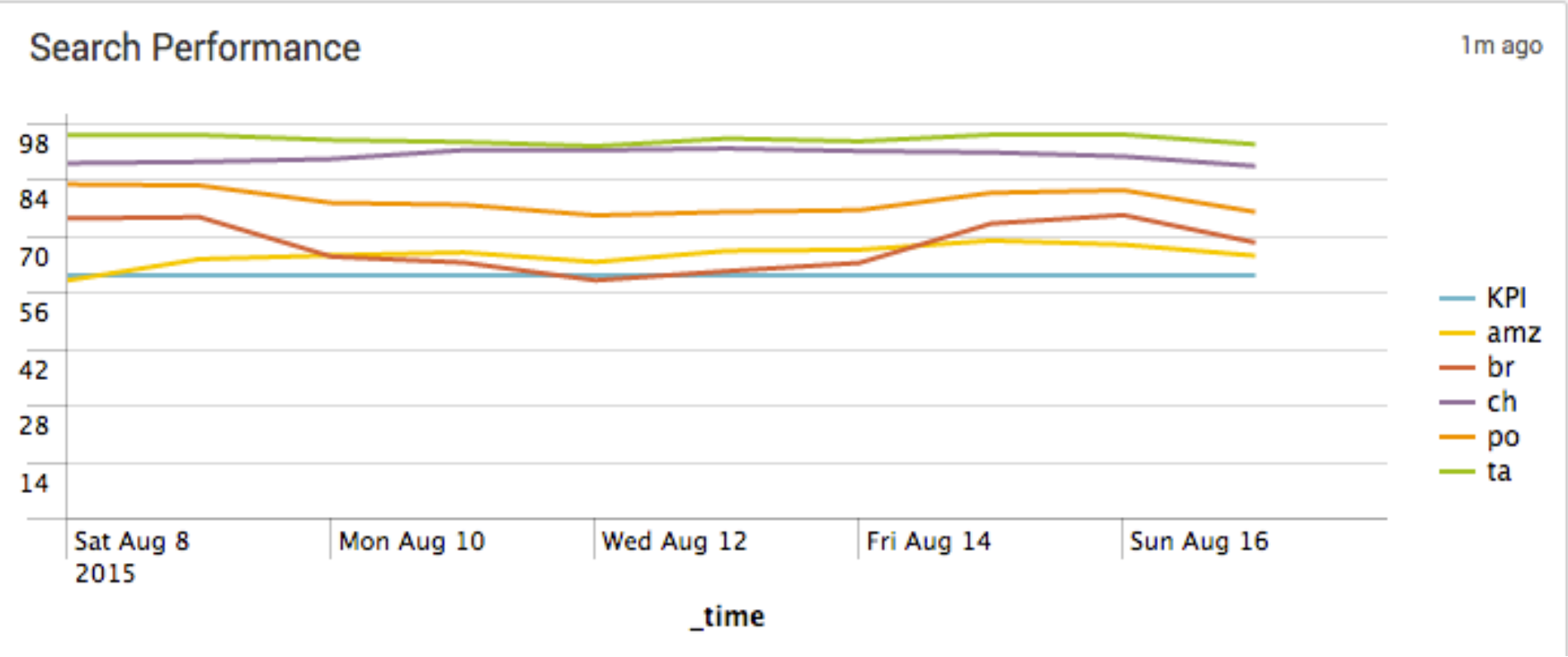
Why CPU?

Basic CPU stats show us how the indexers are performing throughout the day

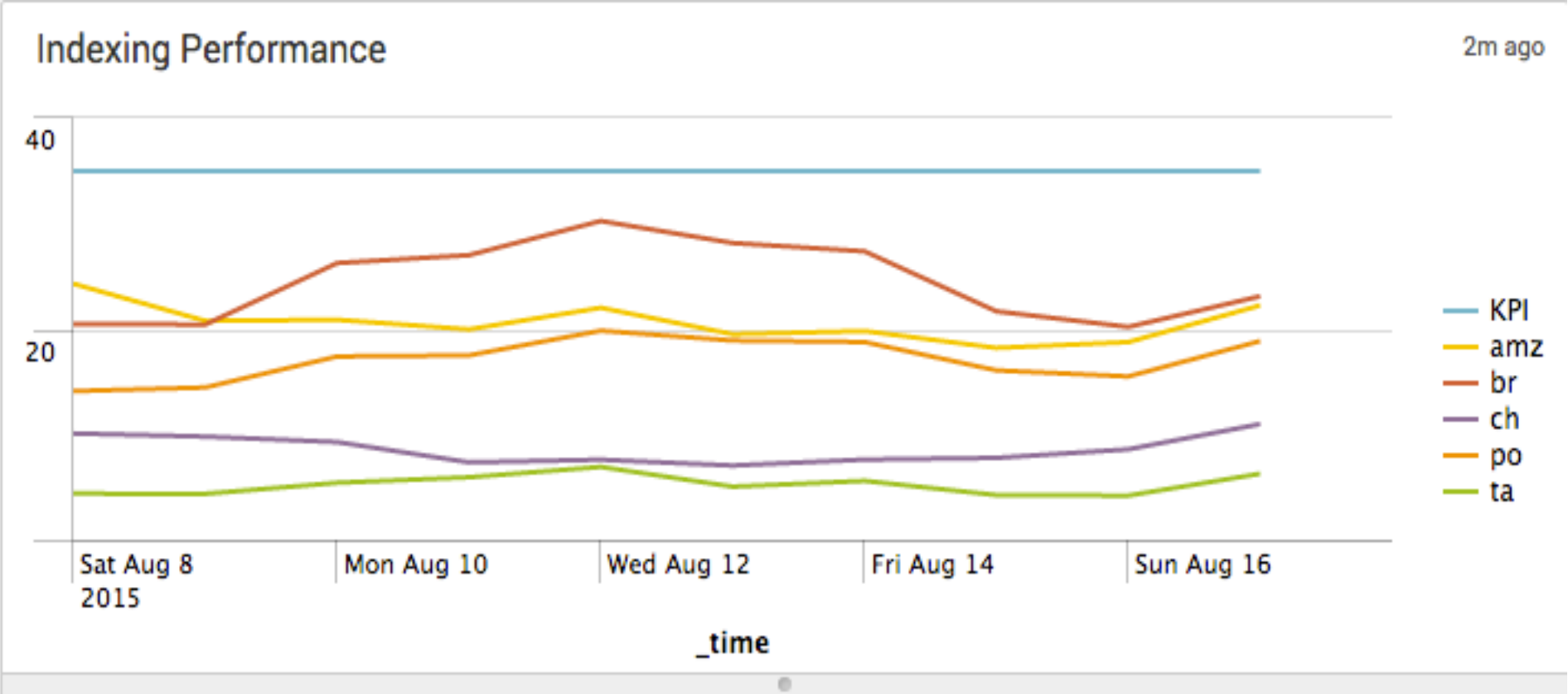
% IDLE – measures your search load.

% USER – measures your index load.

Trend It- %IDLE



Trend It- %USER

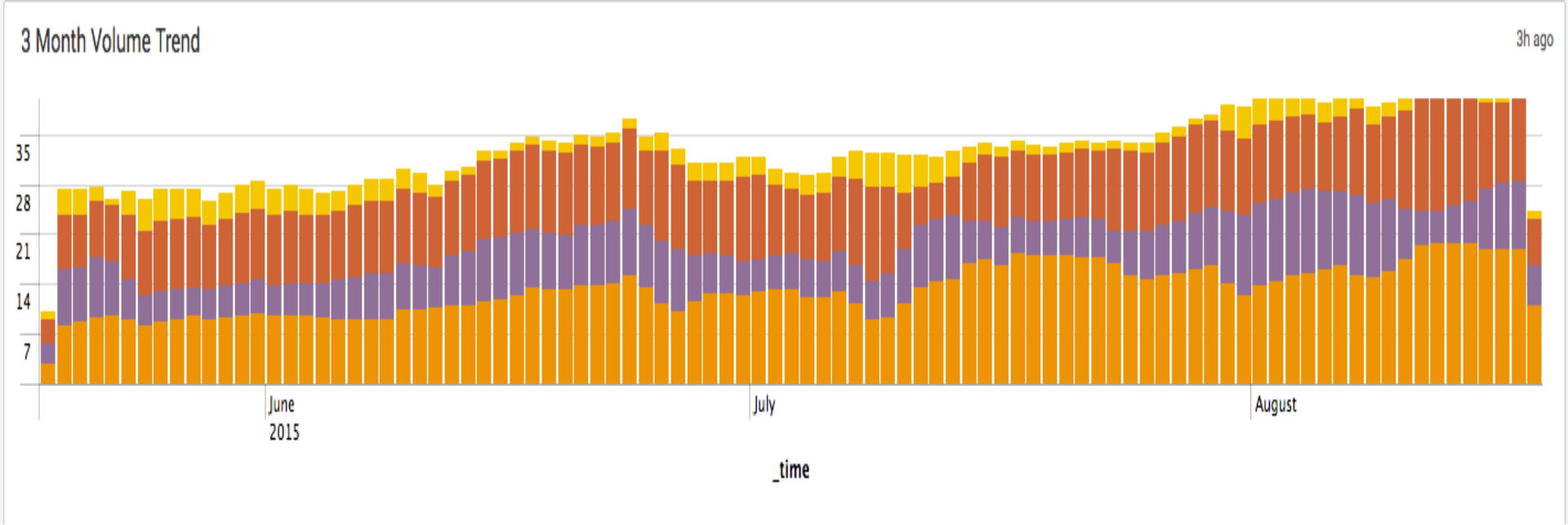


Trend It – License Volume

Take a closer look at that license volume…

- Your licensing data is full of great information that allows you to actually TREND your data volume over time.

Trend It - Volume



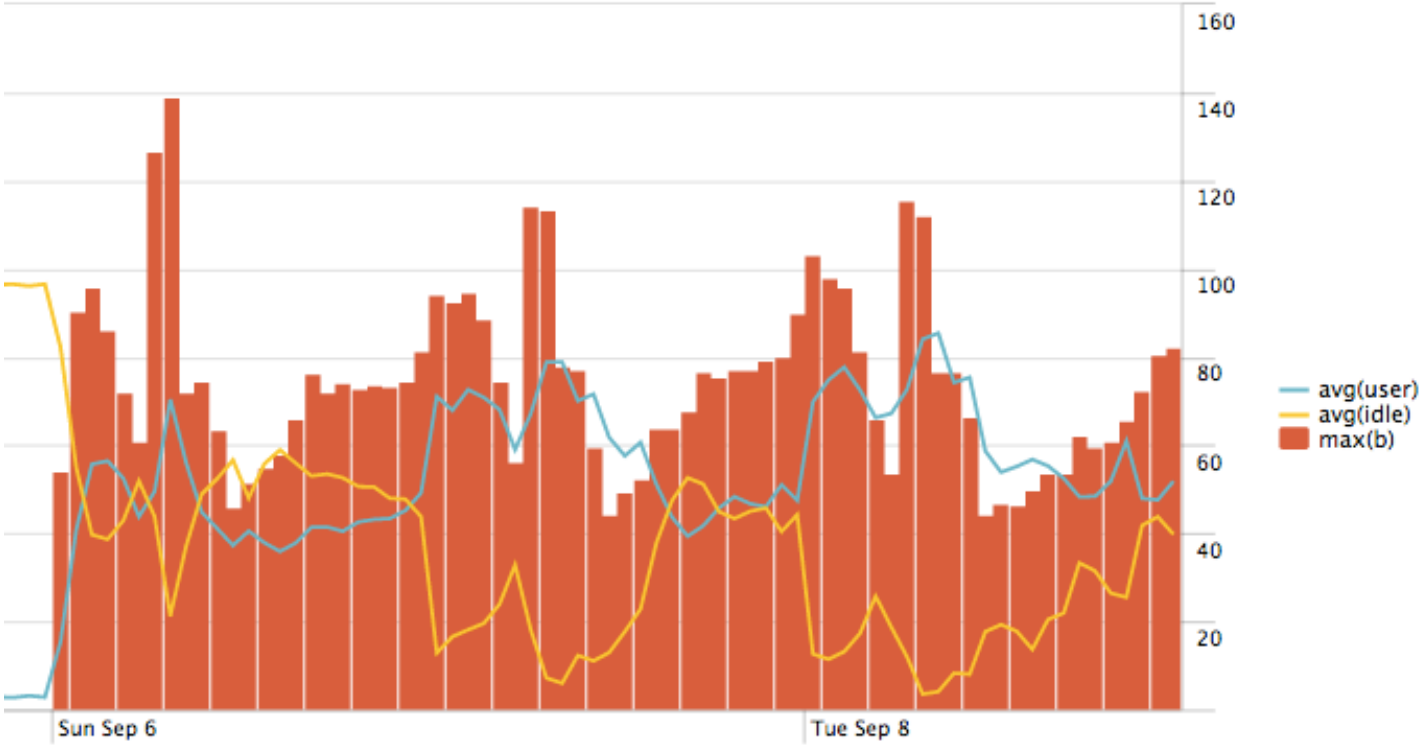
Planning for the Future

- With all the volume metrics information you now have, you can start to look at the trends & analyze the rates of growth that you are seeing.
- Questions you will be able to answer:
  - Which source types are growing the fastest?
  - Is your volume growth consistent with the numbers of hosts you have deployed?
  - Or is trending with the number of customers you're supporting?

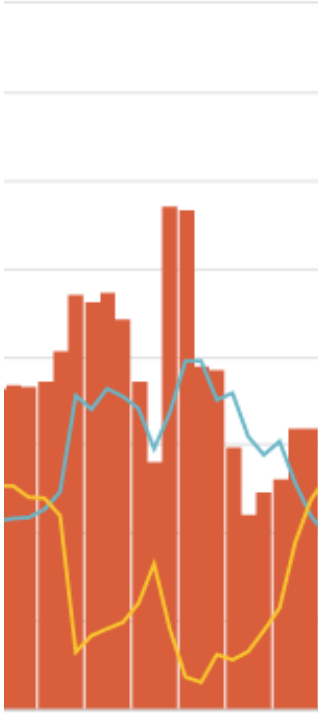
Bringing It Together - Capacity

- Splunk gives us the rough estimate that you can index 50-100GB a day on a bare-metal indexer.

Bringing It Together - Capacity



Bringing It Together - Capacity



%USER IS ABOVE 70%

%IDLE IS BELOW 10%

Splunk It – Capacity Query

- host=<license master> index=_internal source="/opt/splunk/var/log/splunk/license_usage.log" b>0
- [| rest /services/licenser/localslave splunk_server=<filter for a TYPE of indexer>| fields slave_id | rename slave_id AS i]
- | bucket _time span=1h | stats sum(b) AS b by i,_time
- |append [search source=sar host=<filter for a TYPE of indexer> perc_idle>60 AND perc_user>30
- | bucket _time span=1h | stats values(perc_idle) AS idle, values(perc_user) AS user by _time]
- | eventstats values(idle) AS idle, values(user) AS user, avg(b) AS b by _time | table _time idle user b
- | where isnotnull(idle) | eval gb=b/1024/1024/1024
- | stats avg(idle), avg(user) avg(gb) AS avg_h_gb by _time
- | where isnotnull(avg_h_gb)
- | stats min(avg_h_gb) as min, max(avg_h_gb) as max, stdev(avg_h_gb) as stdev
- | eval mid1=min+stdev
- | eval mid2=max-stdev
- | eval d1=mid1*24
- | eval d2=mid2*24 | table min max d1 d2

What is This?

- Healthy Total Daily Capacity for an indexer.

Job ⌄ ⏸ ■ ↗ ⬇ 🖨 ⚡ Fast Mode ⌄

d1	d2
104.0375	116.1677

Tips for Managing Your Splunk Installation

Topics

- Using Splunk's REST API For Troubleshooting and Reporting
- Detecting Latency Between Forwarders and Indexers
- Dealing With High Volume Sources

Splunks Rest API

- Every feature of SPLUNK is accessible from the REST API
- Indexers
- Forwarders
- Search Heads
- Deployment Servers

Check

- http://docs.splunk.com/Documentation/Splunk/6.2.5/RESTREF/RESTlist

REST API Reference Manual

Version

6.2.5 (latest release)

- Introduction
  - About the REST API Reference Manual
  - URI quick reference
- + Access endpoints
- + Application endpoints
- + Cluster endpoints
- + Configuration endpoints
- + Deployment endpoints
- + Input endpoints
- + Introspection endpoints
- + Knowledge endpoints
- + KV store endpoints

URI quick reference

Jump to: A - C - D - I - L - M - P - R - S

alerts/ URI	Summary	GET	PUT	POST	DEL
alerts/fired_alerts	Search Access all fired alerts	○			
alerts/fired_alerts/{name}	Search Access specific fired alert	○			○

apps/ URI	Summary	GET	PUT	POST	DEL
apps/appinstall	Applications Install app from URL or local file			○	
apps/apptemplates	Applications Access app templates for creating new apps	○			

Ways to Access the API

- Directly From Search UI on Management Search Head
  - Rest (SPL) Command
- Web Browser or Curl

Danger, Will Robinson!

- You can delete & modify using post requests to rest

The Rest (SPL) Command

| rest <ENDPOINT> <SPLUNK_SERVER> --optional (searches all peers by default)

```
| rest
```

Matching searches

/ rest /services/admin
/ rest /services/search
/ rest /servicesNS/admin
/ rest /services/admin
/ rest /services/data/inputs/script

Command history

... | rest /services/admin/inputstatus/TailingProcessor:FileStatus
... | rest /services/admin
... | rest /services/data/inputs/script
... | rest /services/search
... | rest /servicesNS/admin

rest Help More »

Access a `rest` endpoint and display the returned entities as search results.

Examples

Access saved search jobs.

| rest /services/search/jobs count=0 splunk_server=local | search isSaved=1

How to Search

Using Search Commands

More advanced searches use commands to transform, filter, and report on the events you retrieved.

- Use the vertical bar, or pipe character, to apply a command to the retrieved events:
  `sourcetype=access_* error | top 20 uri`
- Further refine or transform your search results with a additional commands:
  `sourcetype=access_* error | top 20 uri | search count>5`

Search assistant will suggest commands for you to use next and show you examples to help you build your search.

Real World Examples

- Reporting On Peered Indexers
- Viewing Recent Errors
- Reports of Configured Alerts
- Forwarder Monitor Troubleshooting

Reporting On Peered Indexers

Information You Can Get….

- Peer Name
- SPLUNK Version
- Status

Reporting On Peered Indexers

| rest /services/search/distributed/peers | search disabled=0 AND status=Up AND (splunk_server=search_head_1 OR splunk_server=search_head_2) | stats dc(peerName) AS indexerCount by splunk_server

```
| rest /services/search/distributed/peers | search disabled=0 AND status=Up (splunk_server=spl-[redacted]
[redacted]com OR splunk_server=spl-[redacted]com)› | stats dc(peerName) AS
indexerCount by splunk_server
```

Custom time

✓ 0 events (8/5/15 2:41:30.000 PM to 8/5/15 2:56:30.000 PM)

Job Smart Mode

Events | Statistics (2) | Visualization

100 Per Page Format Preview

	splunk_server	indexerCount
1	spl-[redacted]com	383
2	spl-[redacted]com	383

Viewing Recent Search Errors

- | rest /servicesNS/-/-/search/jobs

Information You Can Get….

- Search Query Used
- Time of Execution
- User
- Error & Warn Messages

Viewing Recent Search Errors

- | rest /servicesNS/-/-/search/jobs |fields published label author messages.error messages.warn label title | rename title AS search | search messages.warn="*" OR messages.error="*"

published	search	messages.error	messages.warn	author
2015-09-01T03:05:55.000+00:00	[redacted]		Unable to find tag [redacted] Unable to find tag [redacted]	[redacted]
2015-08-31T14:11:41.000+00:00	[redacted] scout server	Search results may be incomplete: the search process on peer's (http://spl-[redacted]8089) search ended prematurely. The remote search id (SID) was not received from peer so we don't know which search failed! Consult peer's logs, such as $SPLUNK_HOME/var/log/splunk/splunkd.log.	Unable to distribute to peer named [redacted] at uri http://1[redacted]:8089 because peer has status = "Down". Could not read remote messages	[redacted]

Reporting On Scheduled Alerts

- | rest /servicesNS/-/-/saved/searches/

Information You Can Get….

- Name Of Saved Search
- Email Distribution
- Run Frequency
- Search Time Frame
- Alert Condition

Reporting On Scheduled Alerts

- | rest /servicesNS/-/-/saved/searches/ | search is_scheduled=1 AND actions=email AND action.email.sendresults=1 AND disabled=0 | rename alert_condition AS AlertCondition | rename cron_schedule AS Schedule | rename dispatch.latest_time AS TimeFrameLatest | rename dispatch.earliest_time AS TimeFrameEarliest | rename eai:acl.owner AS CreatedBy| rename title AS AlertName | rename search AS SearchQuery | rename eai:acl.app AS SplunkAppLocation | fields AlertName SearchQuery action.email.to AlertCondition Schedule TimeFrameLatest TimeFrameEarliest SplunkAppLocation CreatedBy splunk_server | makemv delim="," action.email.to | rename action.email.to AS Recipients

Events | Statistics (1) | Visualization

100 Per Page | Format | Preview

	AlertName	SearchQuery	Recipients	AlertCondition	Schedule	TimeFrameLatest	TimeFrameEarliest	SplunkAppLocation	CreatedBy	splunk_server
1	Just A Test Alert	index=_internal just a test | stats count by host	joe_cramasta2@cable.comcast.com	| where count > 0	0 * * * *	now	-1h	search	jcramasta	s[redacted]1[redacted]e.co

Troubleshooting Monitors

Using The Web Browser *no admin:changeme

https://splunk-forwarder:8089/services/admin/inputstatus/TailingProcessor:FileStatus

/opt/splunkforwarder/var/log/splunk/audit.log	file position	628235
	file size	628235
	parent	$SPLUNK_HOME/var/log/splunk/splunkd.log
	percent	100.00
	type	finished reading
/opt/splunkforwarder/var/log/splunk/btool.log	file position	0
	file size	0
	parent	$SPLUNK_HOME/var/log/splunk/splunkd.log
	percent	100
	type	finished reading

Detecting Latency

Detecting Latecy – Method #1

```
index=main | stats latest(_time) AS _time by host source | reltime | sort 0 + _time
```

All time (real-time)

648,336 of 648,336 events matched

Job | Smart Mode

Events | Statistics (4,791) | Visualization

100 Per Page | Format | Preview

< Prev 1 2 3 4 5 6 7 8 9 ... Next >

	host	source	_time	reltime
1	[redacted]	[redacted]	2015-08-05 19:07:37	12 minutes ago
2			2015-08-05 19:07:46	12 minutes ago
3			2015-08-05 19:08:13	12 minutes ago
4			2015-08-05 19:15:44	4 minutes ago
5			2015-08-05 19:18:29	1 minute ago
6			2015-08-05 19:18:30	1 minute ago
7			2015-08-05 19:18:30	1 minute ago

Detecting Latecy – Method #2

- Using The Internal Fields that track WHEN The Event Was Indexed
- _index_earliest & _index_latest

```
_index_earliest=-5m@m _index_latest=now index=main | stats latest(_time) AS _time by host source | reltime sort 0 + _time
```

All time

✓ 26,761,176 events (Partial results for before 8/5/15 7:24:26.000 PM)

Job | Smart Mode

Events | Statistics (50,026) | Visualization

100 Per Page | Format | Preview

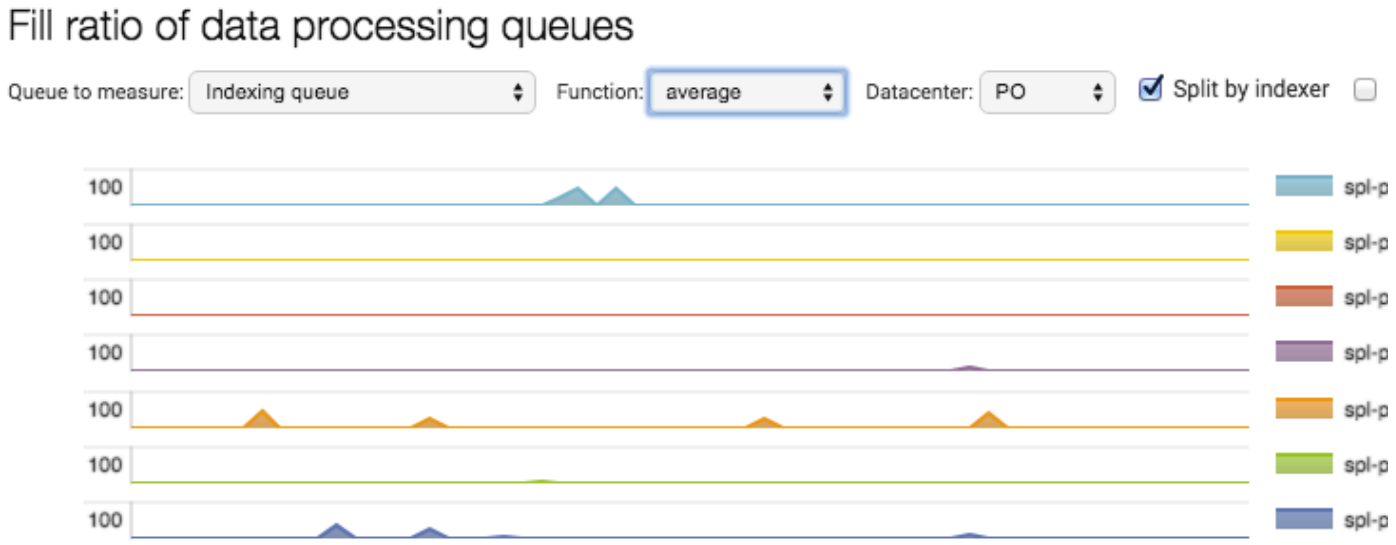
< Prev 1 2 3 4 5 6 7 8 9 ... Next >

	host	source	_time ^	reltime
1	[redacted]	[redacted]	2015-08-05 15:19:22	4 hours ago
2			2015-08-05 19:09:16	15 minutes ago
3			2015-08-05 19:09:27	15 minutes ago
4			2015-08-05 19:16:43	7 minutes ago
5			2015-08-05 19:19:02	5 minutes ago

Reasons Why You Might Have Latency

Overloaded Indexers

- SOS has some great dashboard for Distributed Indexing Performance which shows how full the different indexer queues are



Reasons Why You Might Have Latency

- Increasing maxKBps in the limits.conf on the forwarder.
- However this may just make ANOTHER situation worse.

Indexer Affinity!

ASSUMPTION

- Forwarders Automatically Distribute My Data To All Indexers

REALITY

- Forwarders Only Switch When When A Source Has 3 Seconds With No Activity

INDEXER AFFINITY

- Term Used To Describe When A Forwarder Is “Stuck” Sending A Source To A Single Indexer

Detecting Indexer Affinity

- index=* | stats count dc(splunk_server) AS indexerCount values(splunk_server) AS indexers by host source | where indexerCount=1 | sort 0 - count

host	source	count	indexerCount	indexers
[redacted]	[redacted]	200347	1	[redacted]
[redacted]	[redacted]	196733	1	[redacted]
[redacted]	[redacted]	191429	1	[redacted]
[redacted]	[redacted]	184216	1	[redacted]
[redacted]	[redacted]	182174	1	[redacted]
[redacted]	[redacted]	179332	1	[redacted]
[redacted]	[redacted]	174313	1	[redacted]

forceTimebasedAutoLB

WHAT DOES IT DO?

- Forces forwarders to switch indexers when it reaches the configured autoLBFrequency duration, even if the switch occurs in the middle of a single event that's being generated!
- Automatically reconciles any event that was in the middle of being written during the forced switch so there is no data loss OR half events!

Thanks!

- Kate_Lawrence-Gupta@cable.comcast.com
- Joe_Cramasta2@cable.comcast.com

Q & A

THANK YOU