Splunking the User Experience:

Going Beyond Application Logs with APM data

Doug Erkkila and Diviyesh Patel

CSAA Insurance Group

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.

Agenda

- Introductions
- How We Introduced Splunk – Logs
- Expanding Beyond Log Files with dynaTrace
- Feeding Knowledge back into Development
- Lessons Learned
- Questions?

Who are we?

splunk>

Who are we?

Diviyesh Patel

DevOps and Optimization Manager

- Extensive history in Enterprise ITOps (Fidelity and PepsiCo)

Doug Erkkila

Capacity Management Analyst

- 10 years as an application developer
  - Specialized in data related products

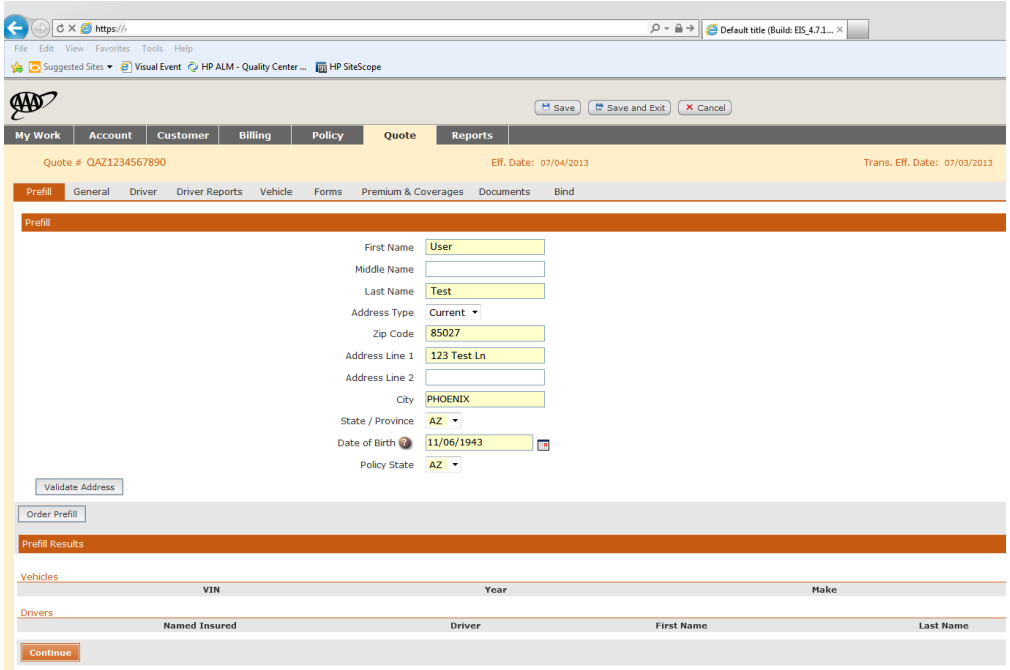
What is CSAA Insurance Group?

- Insurance company offering automobile, homeowners and other personal lines of insurance to AAA members through AAA clubs
- More than 3600 employees coast to coast
- Reaching nearly 17 million AAA members in 23 states and Washington DC

What is PAS?

PAS is the central Policy Administration System

- Consolidate policy administration from across different insurance categories and across different states into one central system
- A Team of over 300 developers and analysts
  - Optimization Team is 5 Analysts



In the Beginning: Logs

splunk>

What started it all?

Customer complaints about Application Errors (Error 500s)

- Was there an increase in exceptions?
- Difficult to understand error patterns across a cluster of servers
  - Was an exception only occurring on a single server?
  - Was an exception only occurring at a certain time of day?
- Is this a matter of perception or application quality?

What started it all?

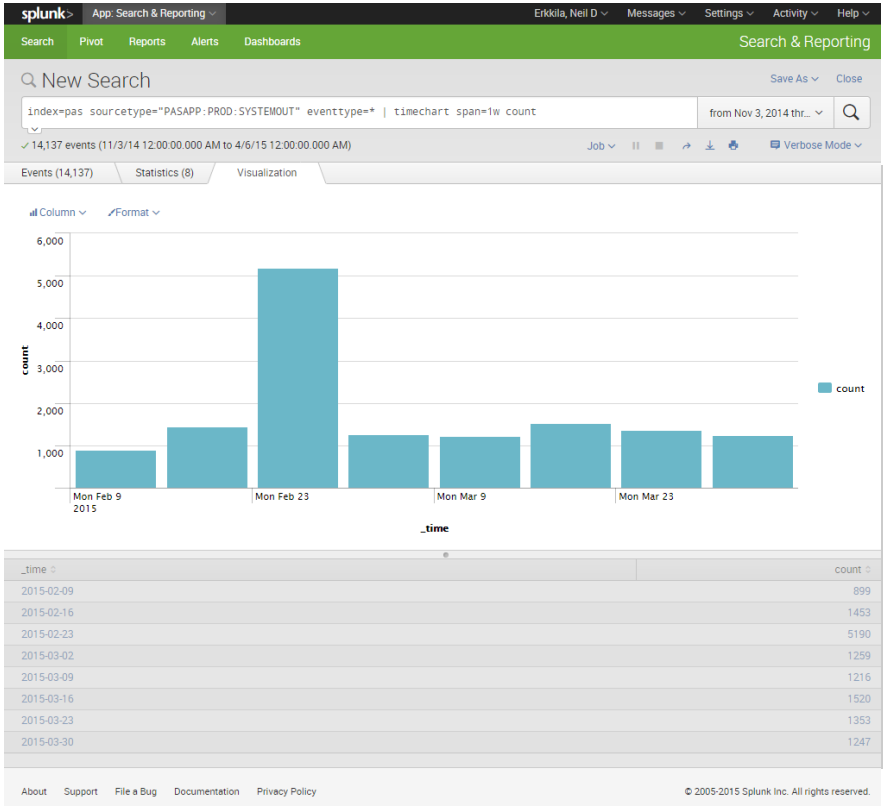
A lot of questions, but finding answers was difficult

- Logs were rolled off so exception history was hard to gather
- Difficult to understand error patterns across a cluster of servers
  - Any patterns had to come from intimate knowledge of logs
  - Wasted support time reading logs every day
- Customer feedback can be sporadic and you tend to hear from the same small group

Error 500s

A single measurable use case

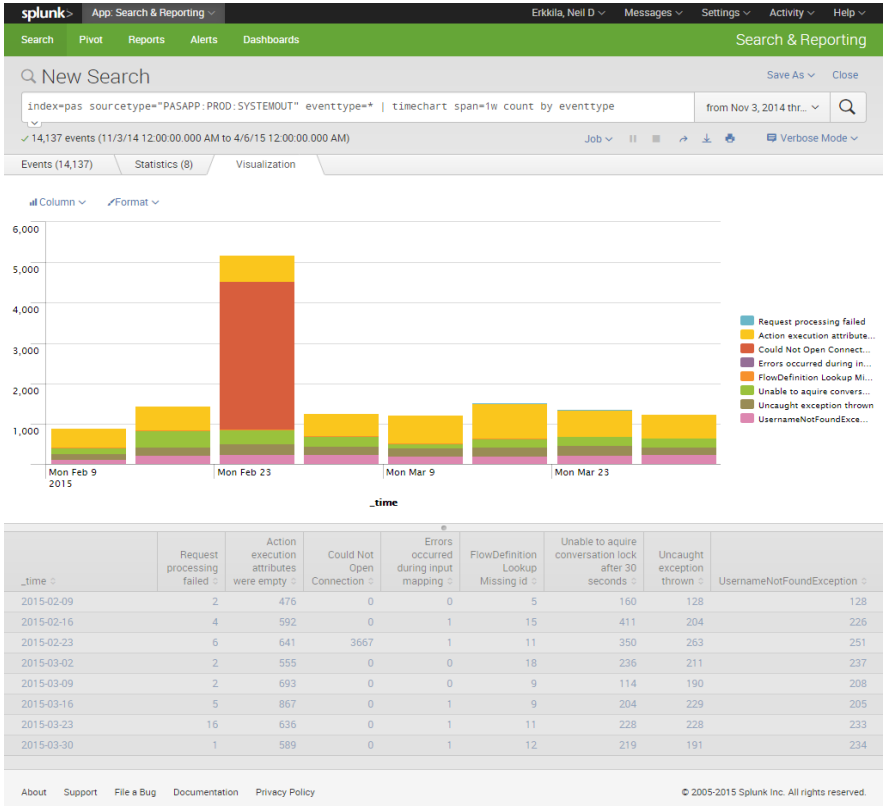
- Basic Aggregation and Trends



Error 500s

A single measurable use case

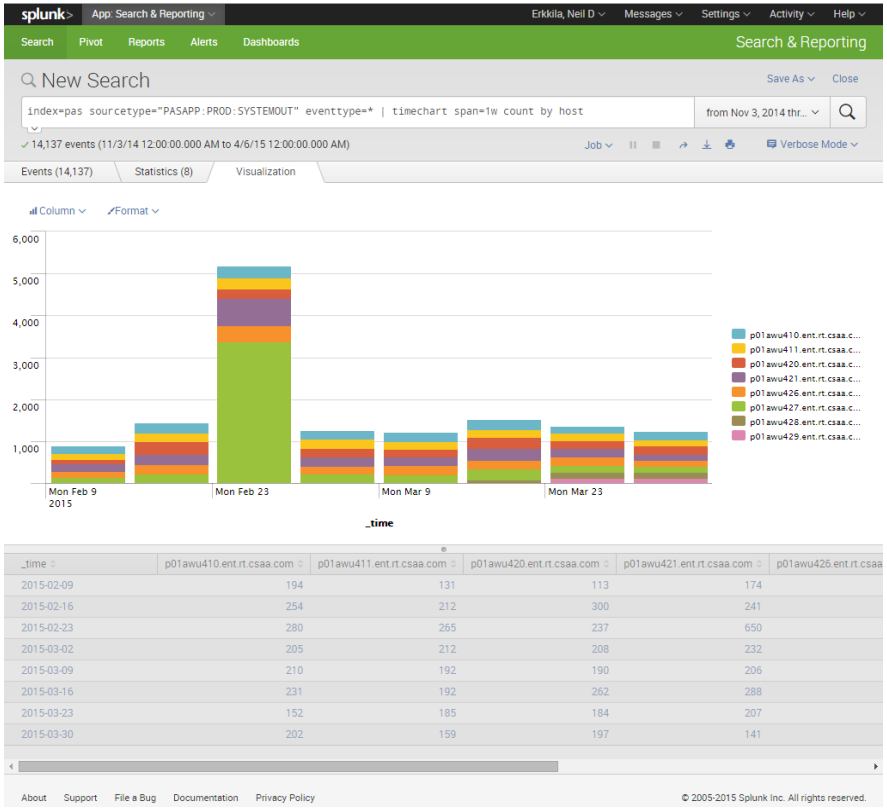
- Basic Aggregation and Trends
- Added categorization



Error 500s

A single measurable use case

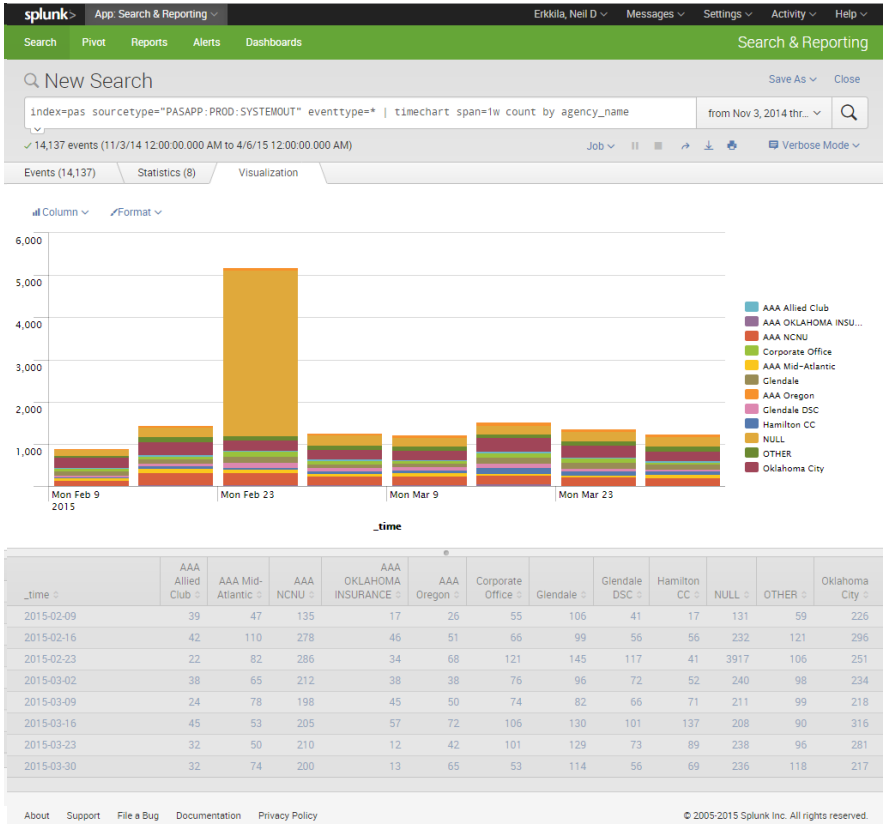
- Basic aggregation and trends
- Added categorization
- Segment by server



Error 500s

A single measurable use case

- Basic aggregation and trends
  - Helped us see the increase in exceptions
- Added categorization
  - Narrowed down timeframes for specific errors
- Segment by server
  - Easy to see if the issue was server specific
- Merge user information
  - Answered our perception vs quality question



What started it all?

Customer complaints about Application Errors (Error 500s)

- Was there an increase in exceptions?
- Difficult to understand error patterns across a cluster of servers
  - Was an exception only occurring on a single server?
  - Was an exception only occurring at a certain time of day?
- Is this a matter of perception or application quality?

What started it all?

Customer complaints about Application Errors (Error 500s)

- Was there an increase in exceptions?
- Difficult to understand error patterns across a cluster of servers
  - Was an exception only occurring on a single server?
  - Was an exception only occurring at a certain time of day?
- Is this a matter of perception or application quality?

What started it all?

Customer complaints about Application Errors (Error 500s)

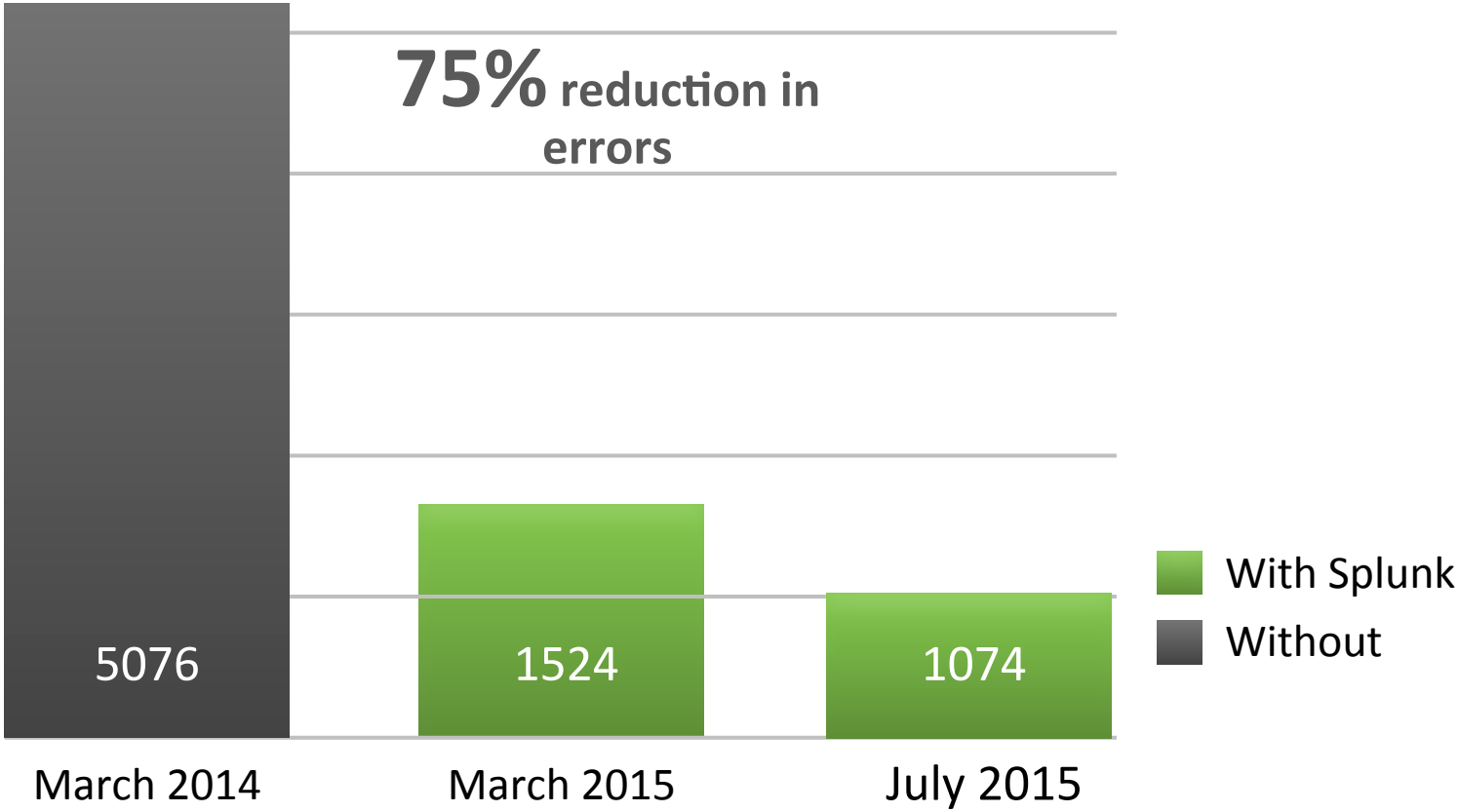
- Was there an increase in exceptions?
- Difficult to understand error patterns across a cluster of servers
  - Was an exception only occurring on a single server?
  - Was an exception only occurring at a certain time of day?
- Is this a matter of perception or application quality?

What started it all?

Customer complaints about Application Errors (Error 500s)

- Was there an increase in exceptions?
- Difficult to understand error patterns across a cluster of servers
  - Was an exception only occurring on a single server?
  - Was an exception only occurring at a certain time of day?
- Is this a matter of perception or application quality?

Changing the reality



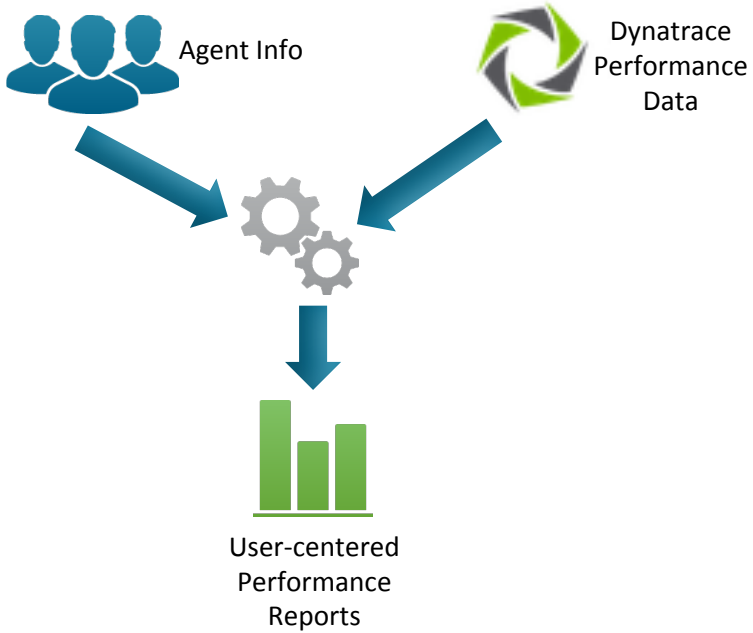
Moving Beyond Logs with dynaTrace APM

splunk>

The power of APM

We were already using Dynatrace for APM

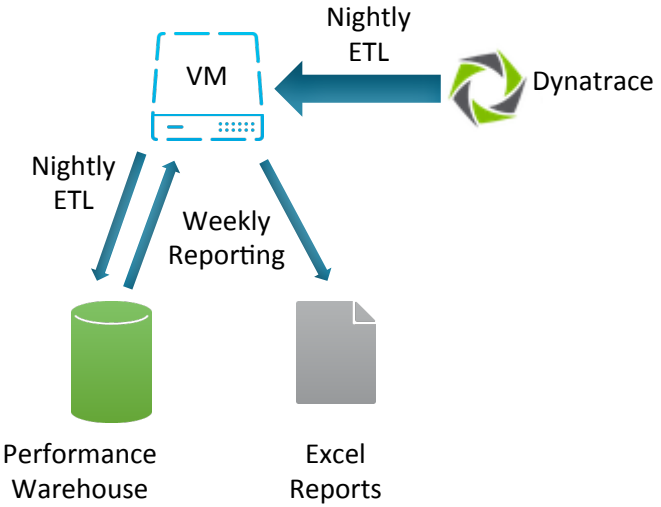
- Method level execution times
- User Experience Management
  - Server/Network/Client side load times
  - Browser data
- Database response times
- JVM Heap Stats



Exporting data out of Dynatrace

Before Splunk

- Using the Dynatrace REST API
- Nightly scheduled ETL
- Database storage
- Excel based reporting



Exporting data out of Dynatrace

But there were flaws…

- Not real time
- Limited user population
- Fragile
  - Scheduled tasks failed to trigger
  - User account credentials changed frequently
  - VM configuration issues
  - REST API calls timed out
- Hours spent cleaning up Excel reports every week

Exporting data out of Dynatrace

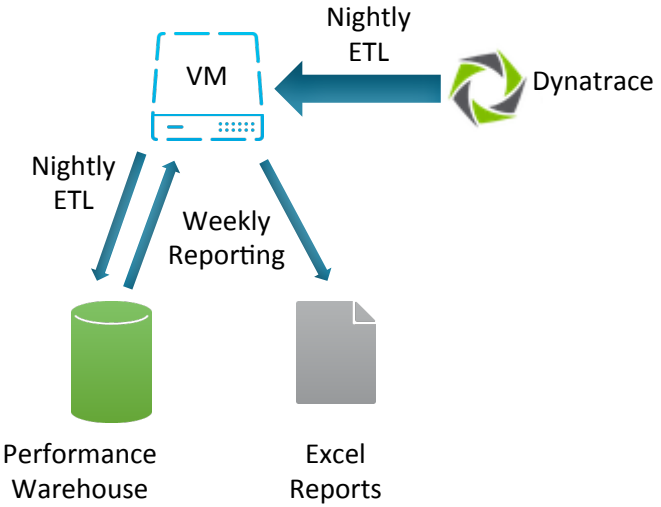
We addressed those flaws

- Real time push from Dynatrace
- Anyone with Splunk access could get performance data
- Limited areas for interruption
- No manual report modification
- Added additional report types

Eliminating those flaws

Dynatrace teams up with Splunk

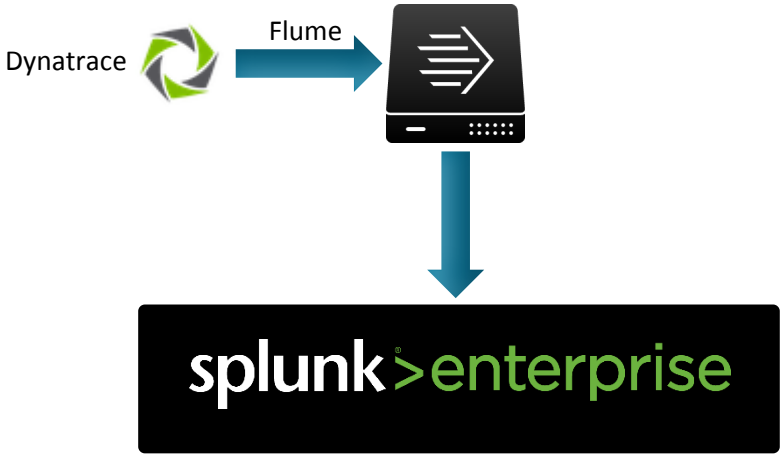
- Dynatrace's Big Data Business Transaction Bridge
- Flume
- Splunk Forwarder
- Pull vs. Push!



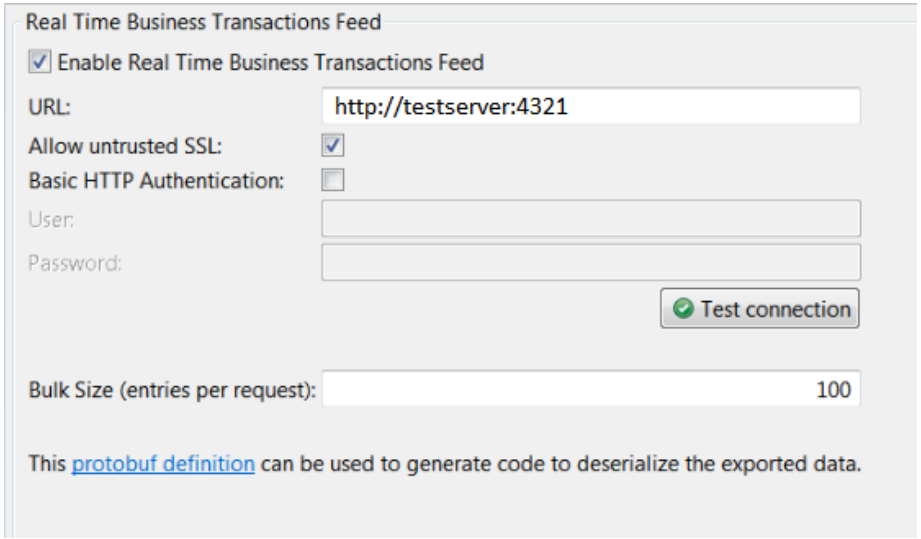
Eliminating those flaws

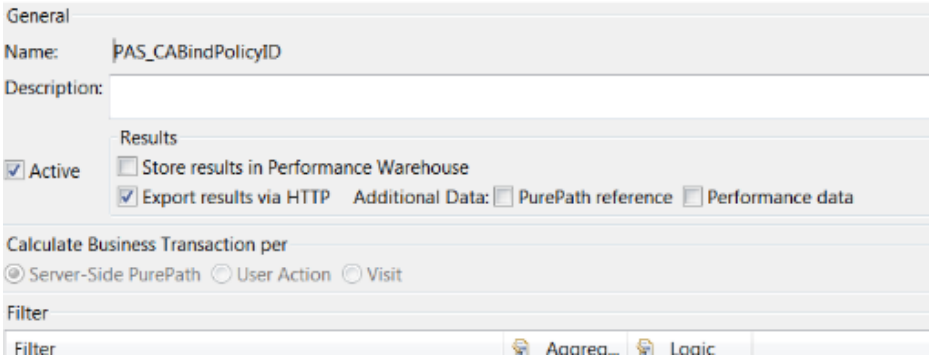
Dynatrace teams up with Splunk

- Dynatrace's Big Data Business Transaction Bridge
- Flume
- Splunk Forwarder
- Pull vs. Push!

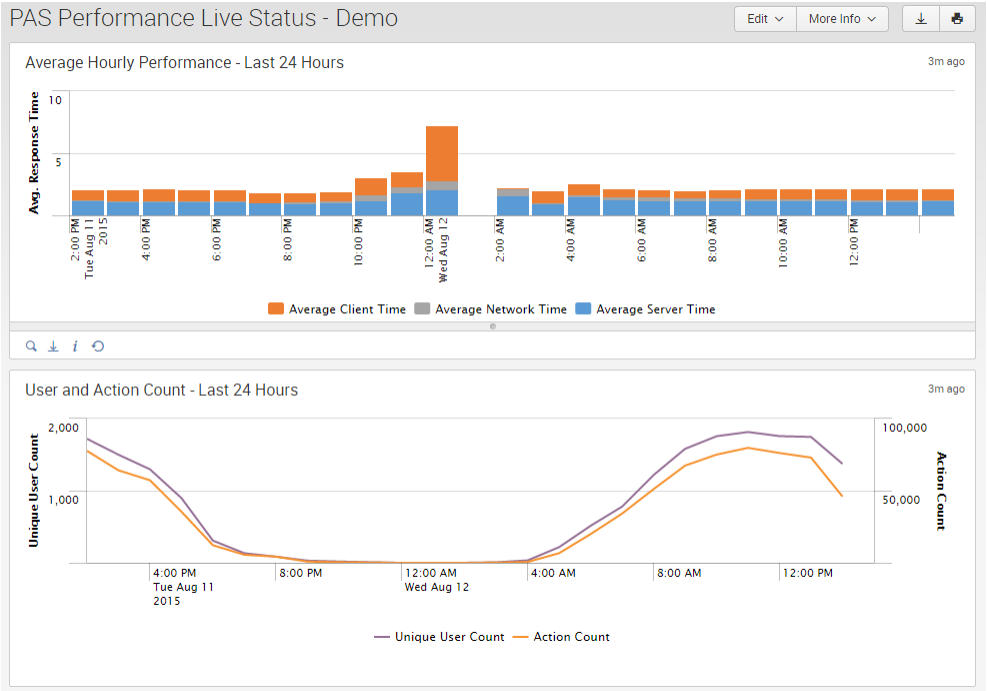


Configuring the big data bridge





Real time performance view



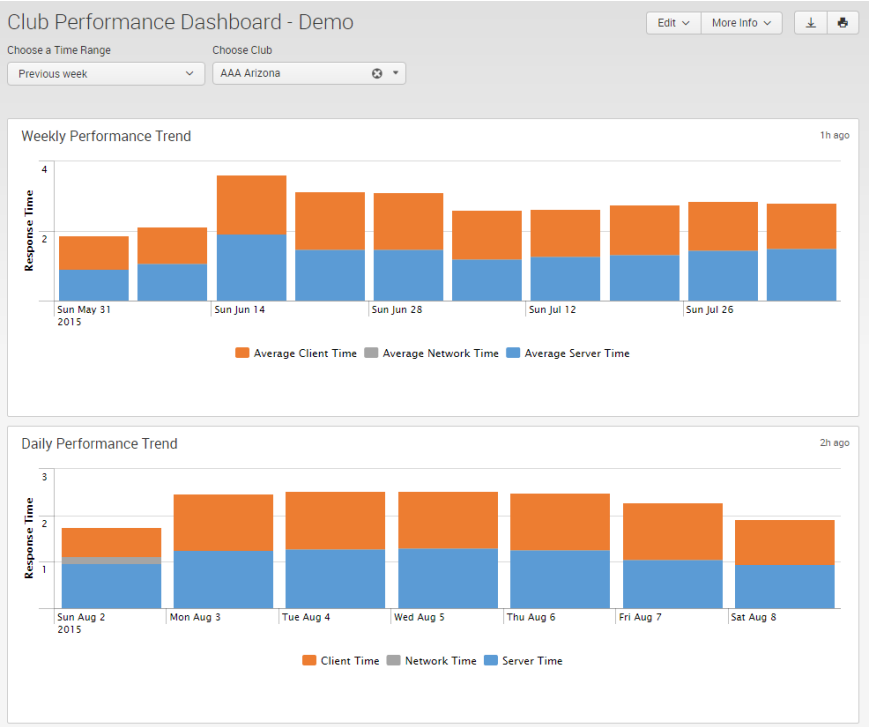
Worst Performers - Last 24 Hours

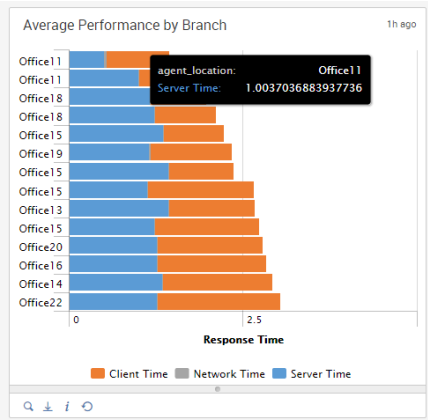
3m ago

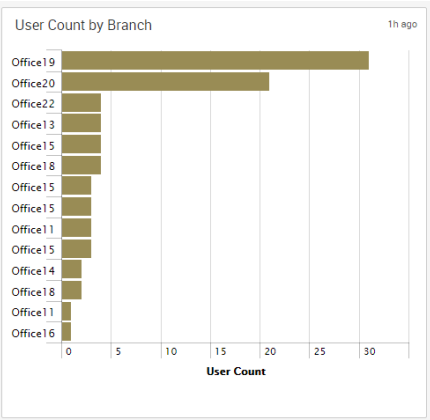
user_profile	username	Agency Name	agent_location	Action Count	Average Server Time	Average Network Time	Average Client Time	Average Response Time
user1234	John Doe	Agency 123	Mystery	37	1.601270410	0.384923211	7.282606042	9.305799663
user1234	John Doe	Agency 123	Mystery	103	1.511172853	0.365249843	7.090817309	9.070240005
user1234	John Doe	Agency 123	Mystery	115	1.507933790	0.047027918	5.977399716	7.647361424
user1234	John Doe	Agency 123	Mystery	44	1.772407568	3.317617502	2.419842452	7.553867522
user1234	John Doe	Agency 123	Mystery	35	2.037508619	3.946302354	1.406840695	7.425651668
user1234	John Doe	Agency 123	Mystery	171	5.179712686	0.018323435	1.928892979	7.297929100
user1234	John Doe	Agency 123	Mystery	136	1.851265685	0.144483190	5.130756176	7.262505051
user1234	John Doe	Agency 123	Mystery	63	4.566796363	0.346285079	2.178347851	7.154429293
user1234	John Doe	Agency 123	Mystery	123	4.487766939	0.226281785	2.246964097	7.084012821
user1234	John Doe	Agency 123	Mystery	201	5.518415205	0.093007131	0.821953687	6.634376023

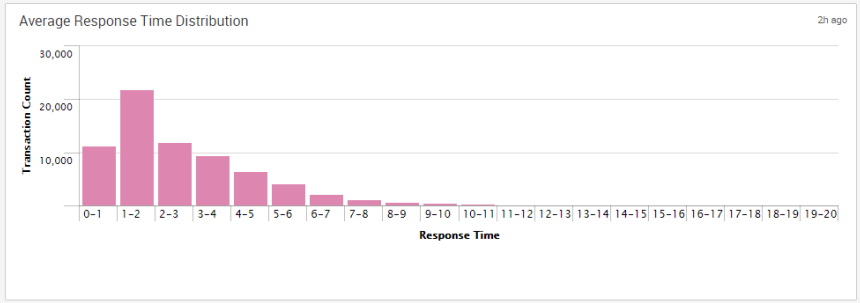
« prev 1 2 3 next »

Club Performance Reports

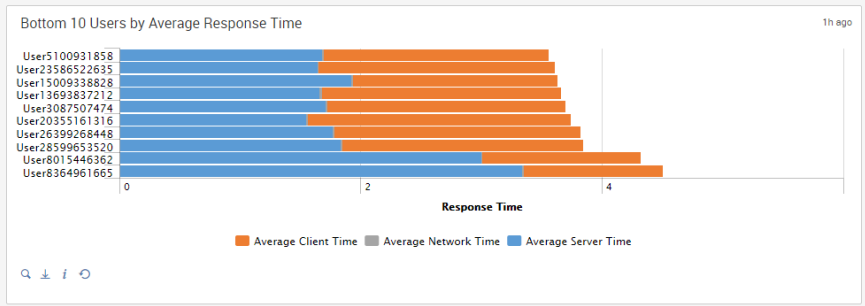


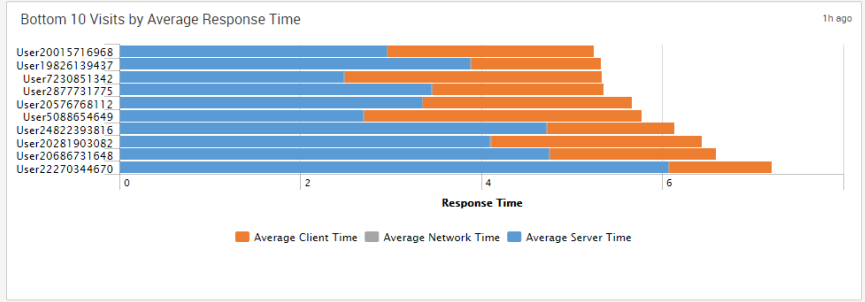


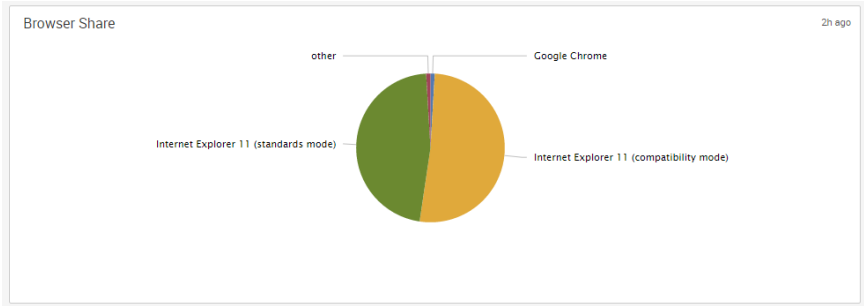




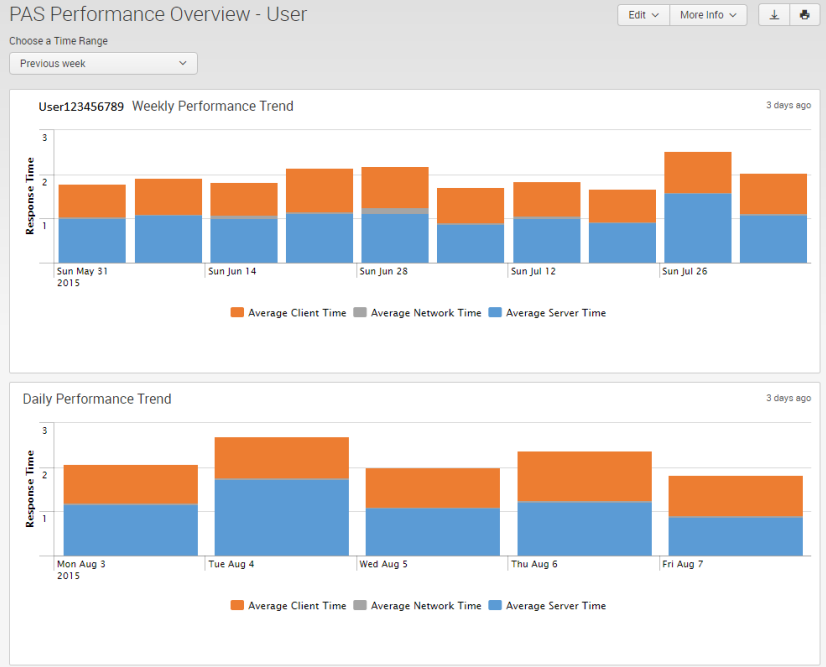
Club Performance Reports

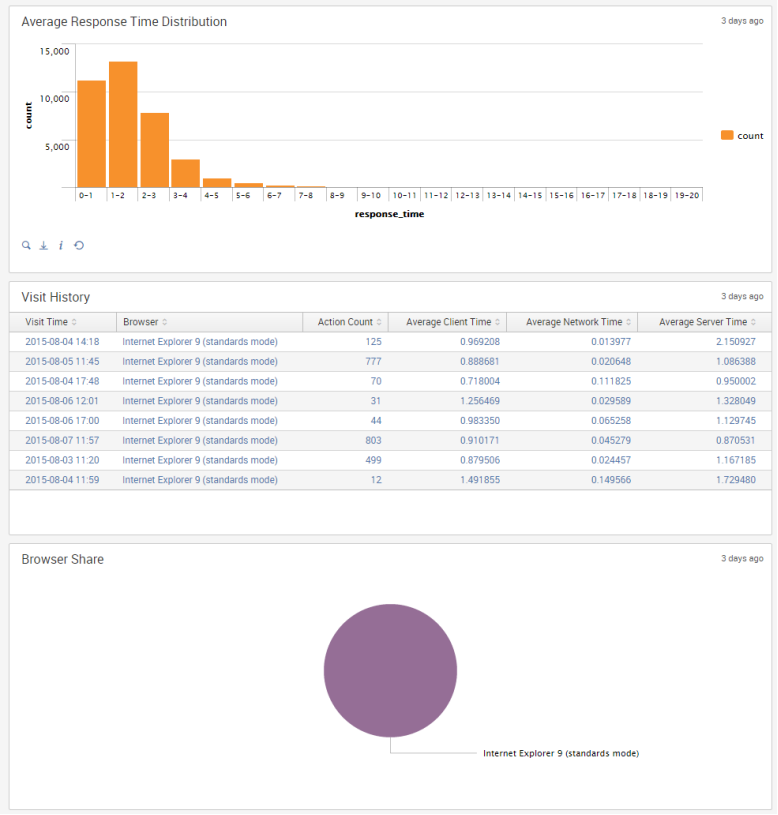






Individual User Performance Trends





Feeding Back Into Development

splunk>

Going Beyond Just Monitoring

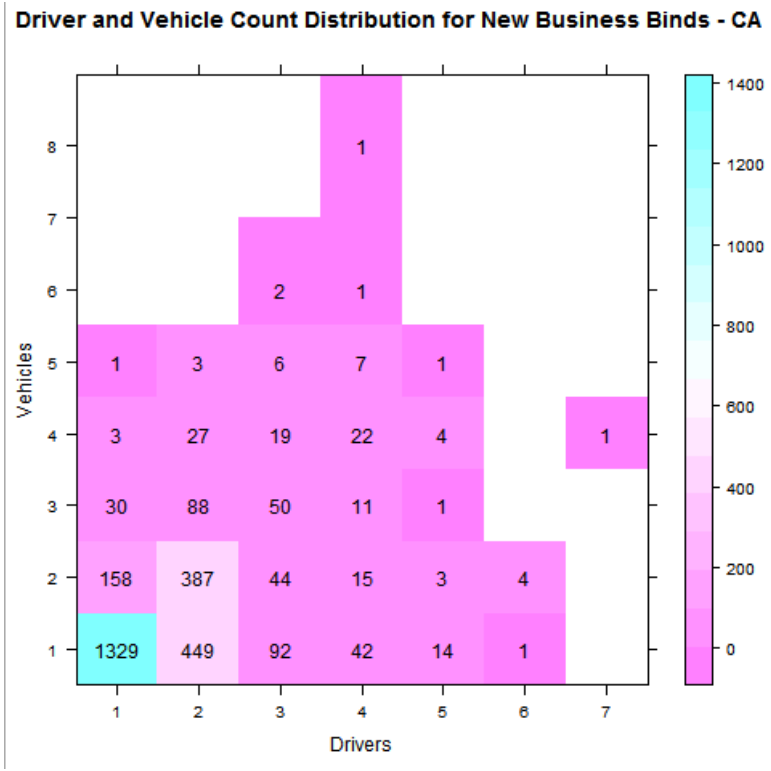
- Easily searchable data repository
- Quick long term trending
- Extensible…

Going Beyond Just Monitoring

- Easily searchable data repository
- Quick long term trending
- Extensible…

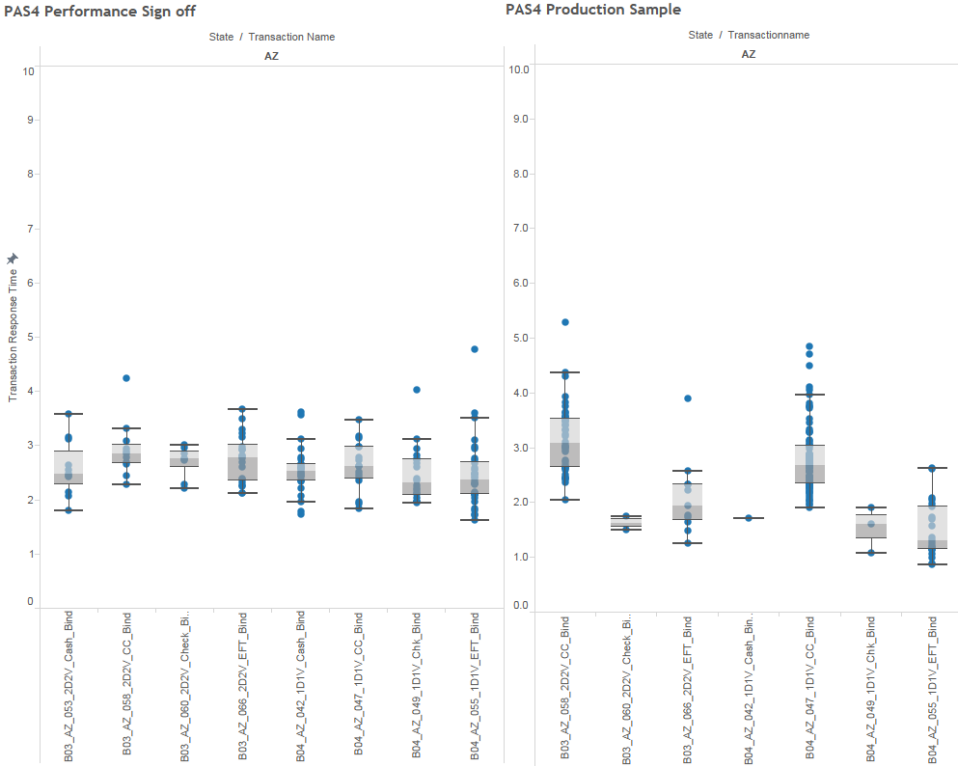
Updating Testing Plans

- Using production transaction frequency
- More realistic performance testing



Performance comparisons

- Typically baseline to baseline comparison in the save environment
- Differences between production and performance environments
- More accurate understanding forecasts



Lessons Learned

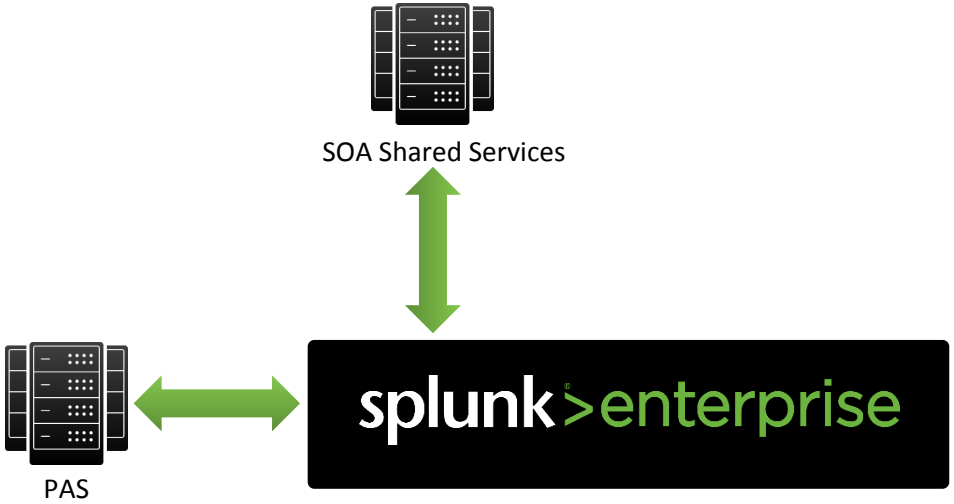
splunk>

Our Splunk Architecture

Splunk started out as a PAS only tool.

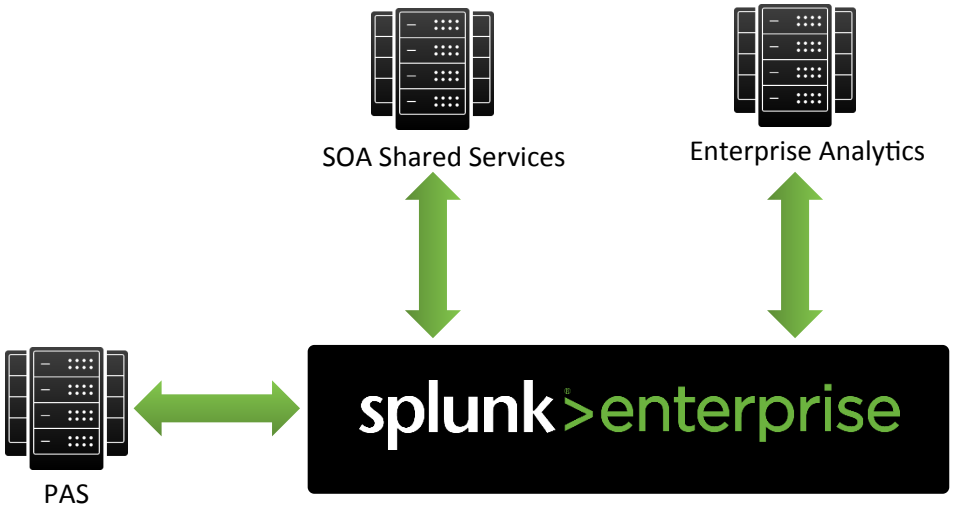
Our Splunk Architecture

But quickly other teams in the enterprise began using it too.



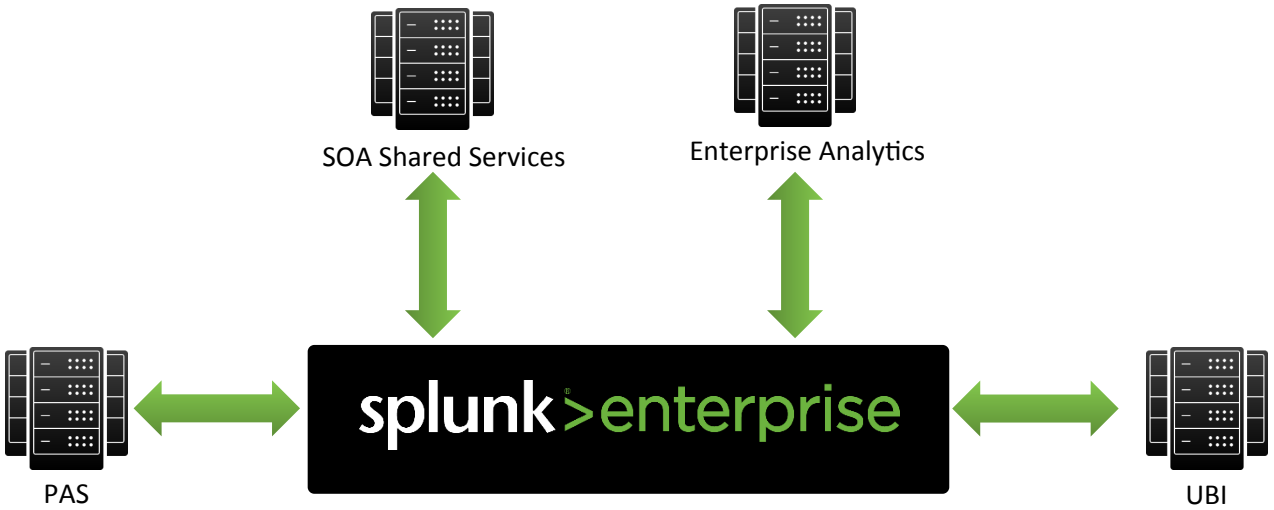
Our Splunk Architecture

But quickly other teams in the enterprise began using it too.



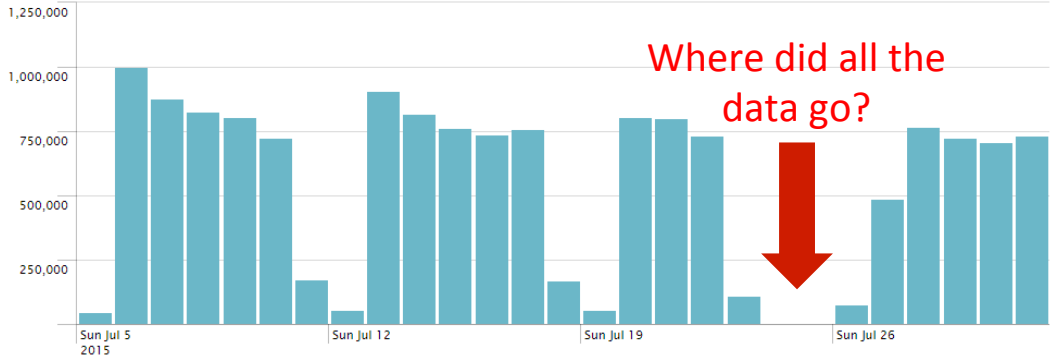
Our Splunk Architecture

Now, Splunk is an enterprise tool utilized by several groups



Capacity settings

- The default Flume capacities may not be enough
- Inadequate capacities can lead to missing or even duplicated data



- Watch for typos: It's "capacity" not "capactiy"

```
agent1.channels.UserActionChannel.type = memory
agent1.channels.UserActionChannel.capacity = 1000
agent1.channels.UserActionChannel.transactionCapactiy = 100
```

Questions?

splunk>

THANK YOU

splunk>