.conf18

splunk>

A Deep Dive Into Splunk LDAP Authentication and How it Affects Cluster Performance

Jacky Chen - Developer – Atlassian

October 2018

Forward-Looking Statements

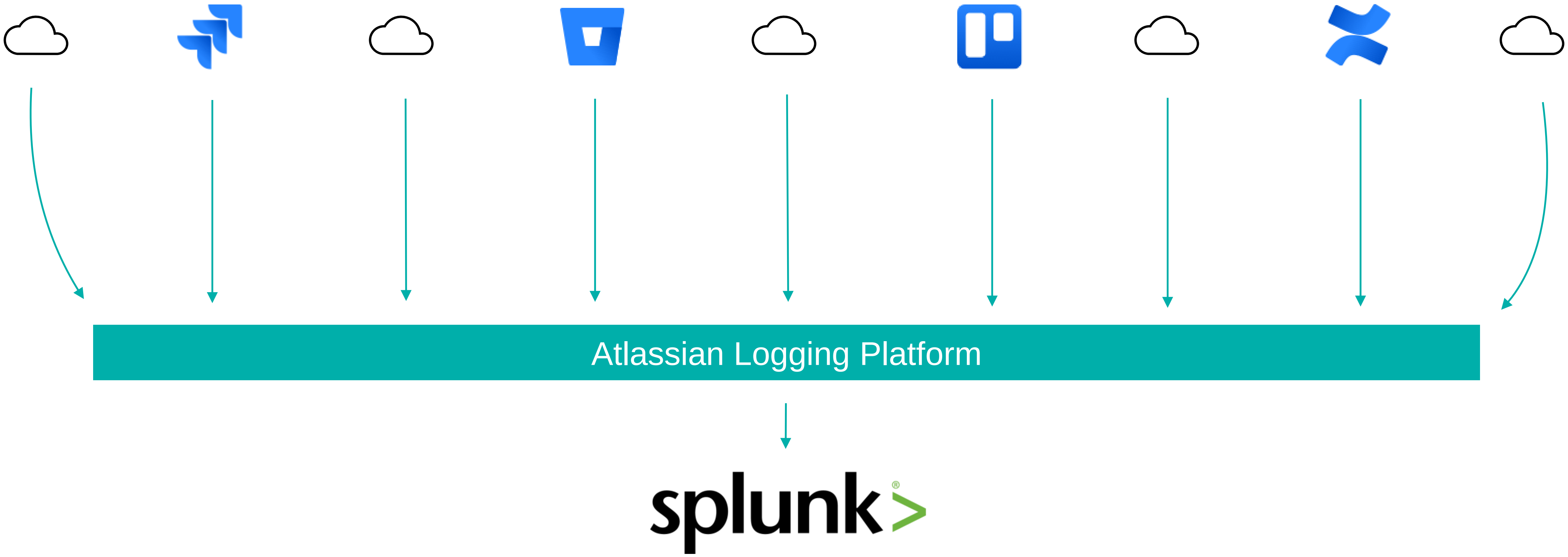
During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results 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.

How we scaled Splunk LDAP auth to handle hundreds of users and groups

Logging @ Atlassian

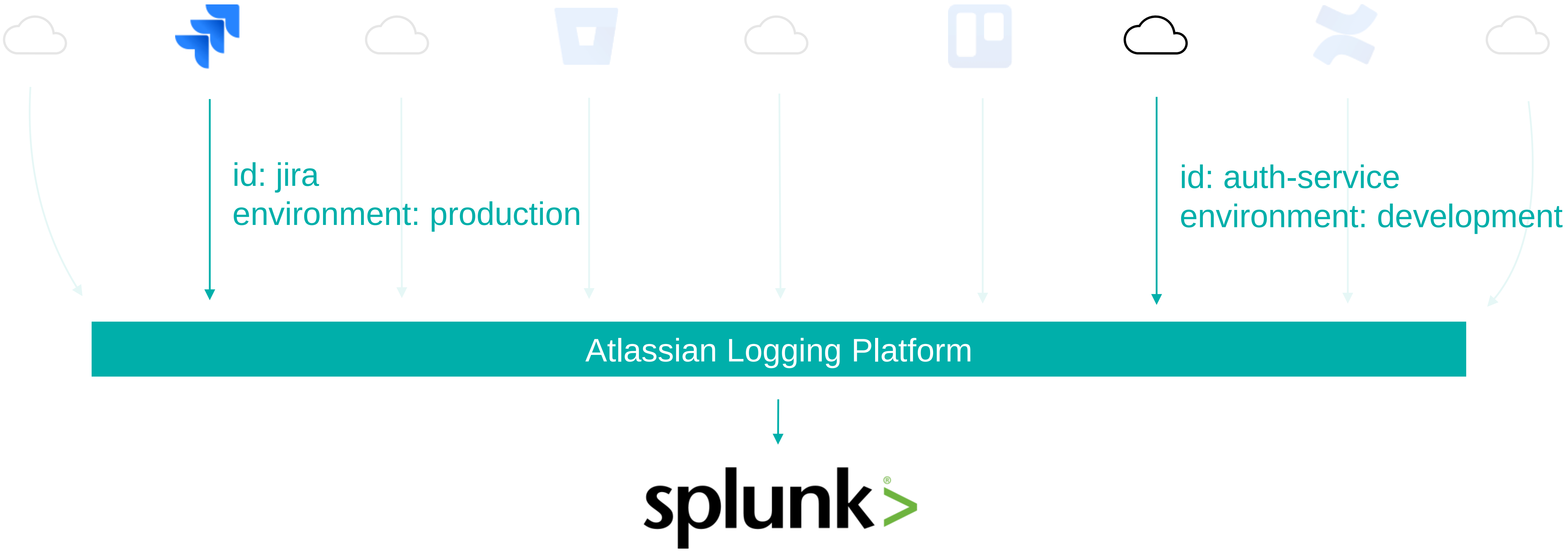
Logging @ Atlassian



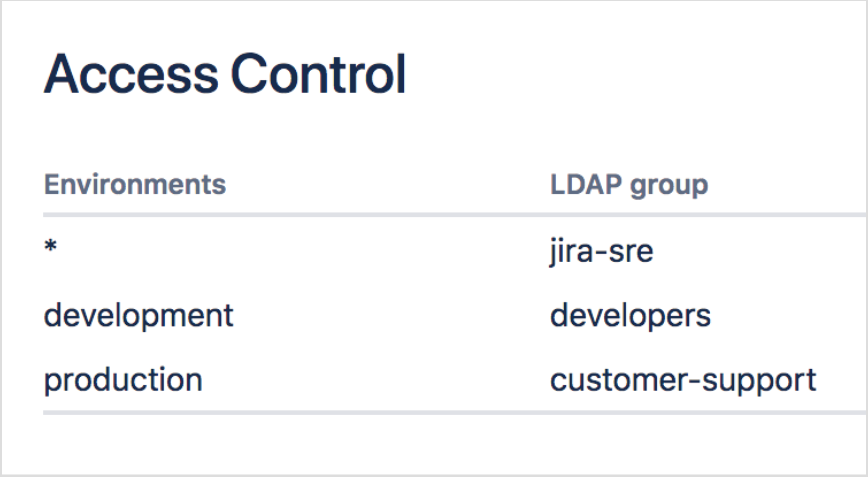
Atlassian Logging Platform

.CONF 2017 Talk: Traversing the Cloud: Atlassian’s Journey Building a Logging Pipeline with Splunk on AWS

Logging @ Atlassian



Access Control Model



Access Control

Environments	LDAP group
*	jira-sre
development	developers
production	customer-support

Generates

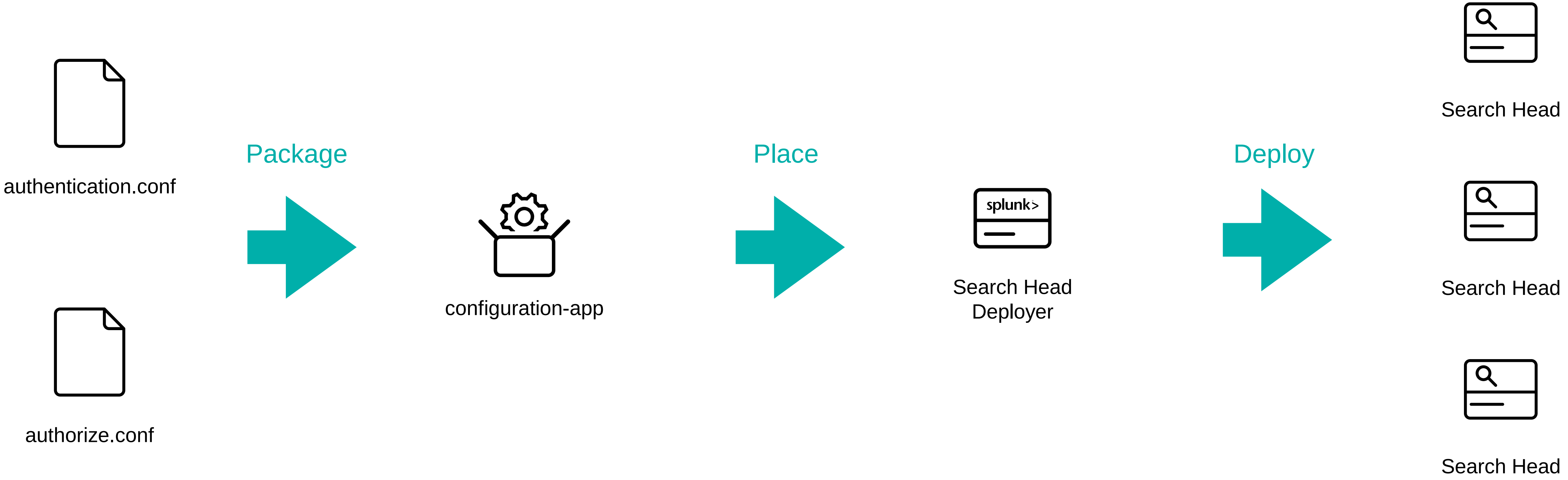
authorize.conf:

```
[role_jira_all_environments_read]
importRoles = user
srchIndexesAllowed = jira


[role_jira_development_read]
importRoles = user
srchIndexesAllowed = jira
srchFilter = environment = development


[role_jira_production_read]
importRoles = user
srchIndexesAllowed = jira
srchFilter = environment = production
```

authentication.conf:

```
[roleMap_ldap]
role_jira_all_environments_read = jira-sre
role_jira_development_read = developers
role_jira_production_read = customer-support
```

Config Deployment Process



Our Scale

392

Splunk Roles

685

LDAP Groups

1

SH Cluster

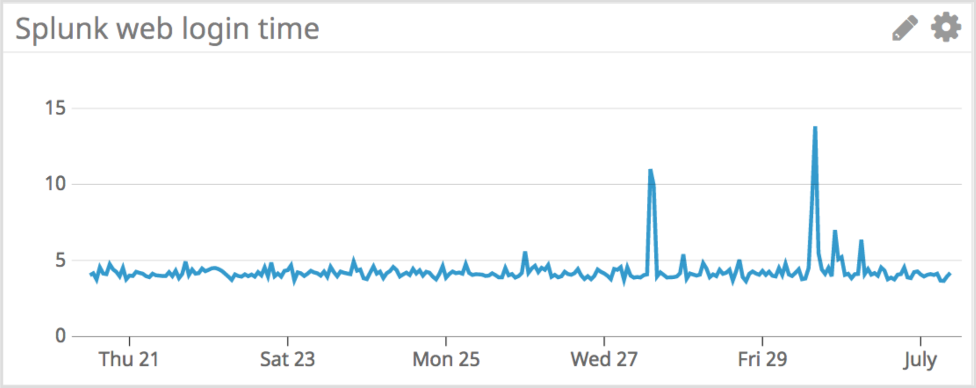
478

DAUs

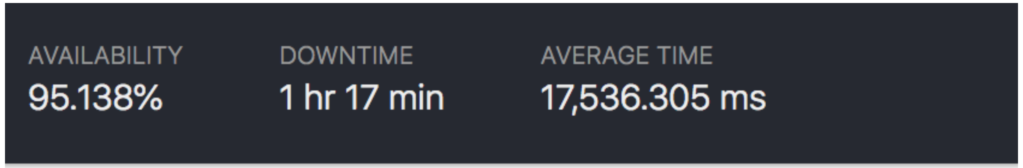
Problems

LDAP component slowness

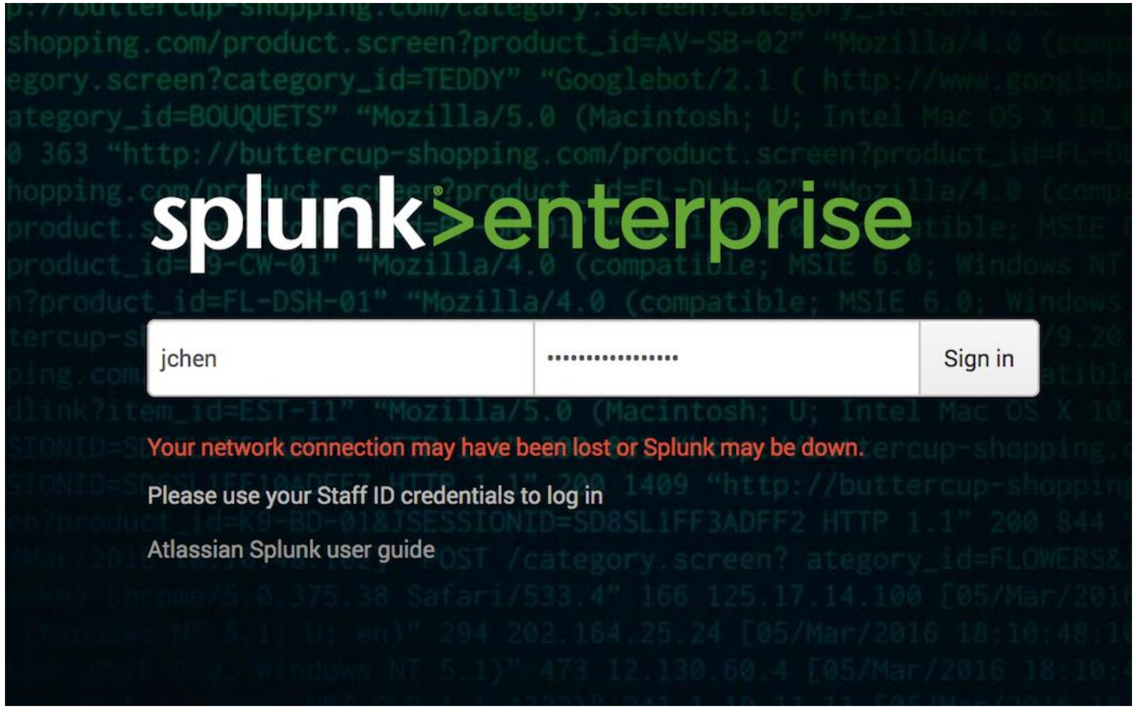
- ~4s login time on average



- ~17s login + page load



AVAILABILITY	DOWNTIME	AVERAGE TIME
95.138%	1 hr 17 min	17,536.305 ms



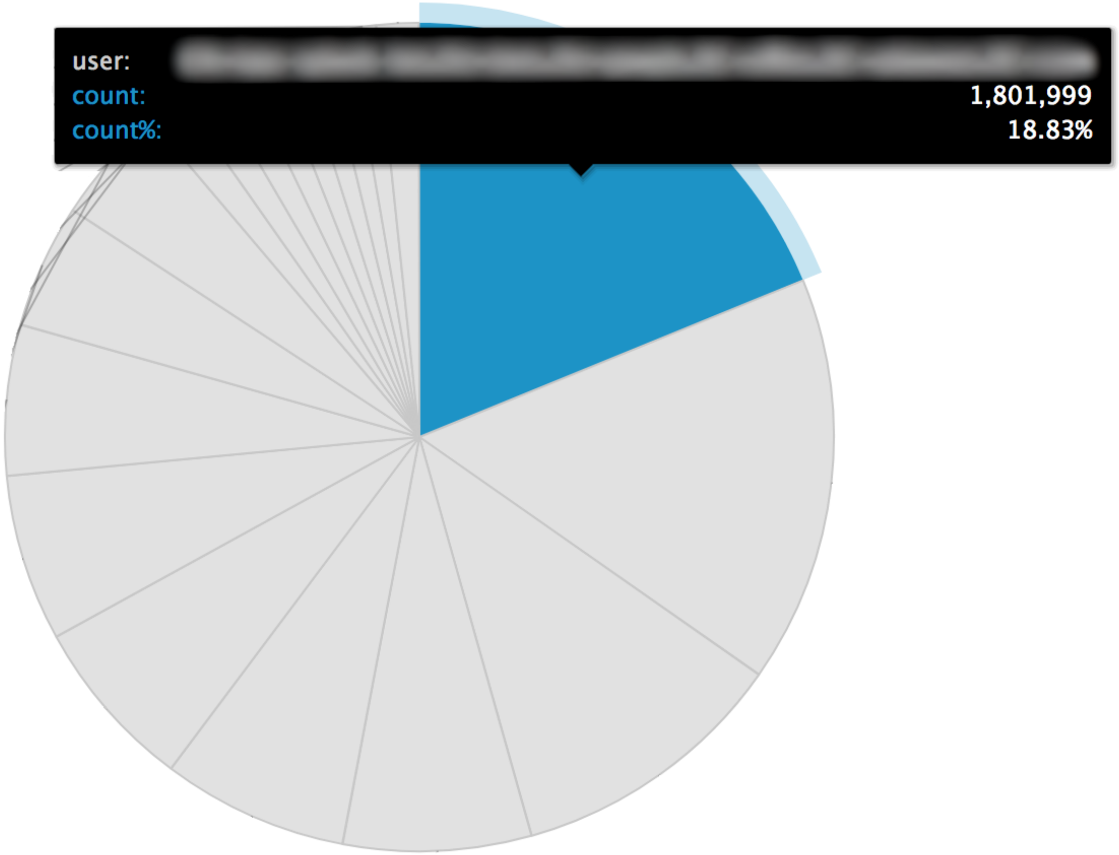
Your network connection may have been lost or Splunk may be down.

Please use your Staff ID credentials to log in

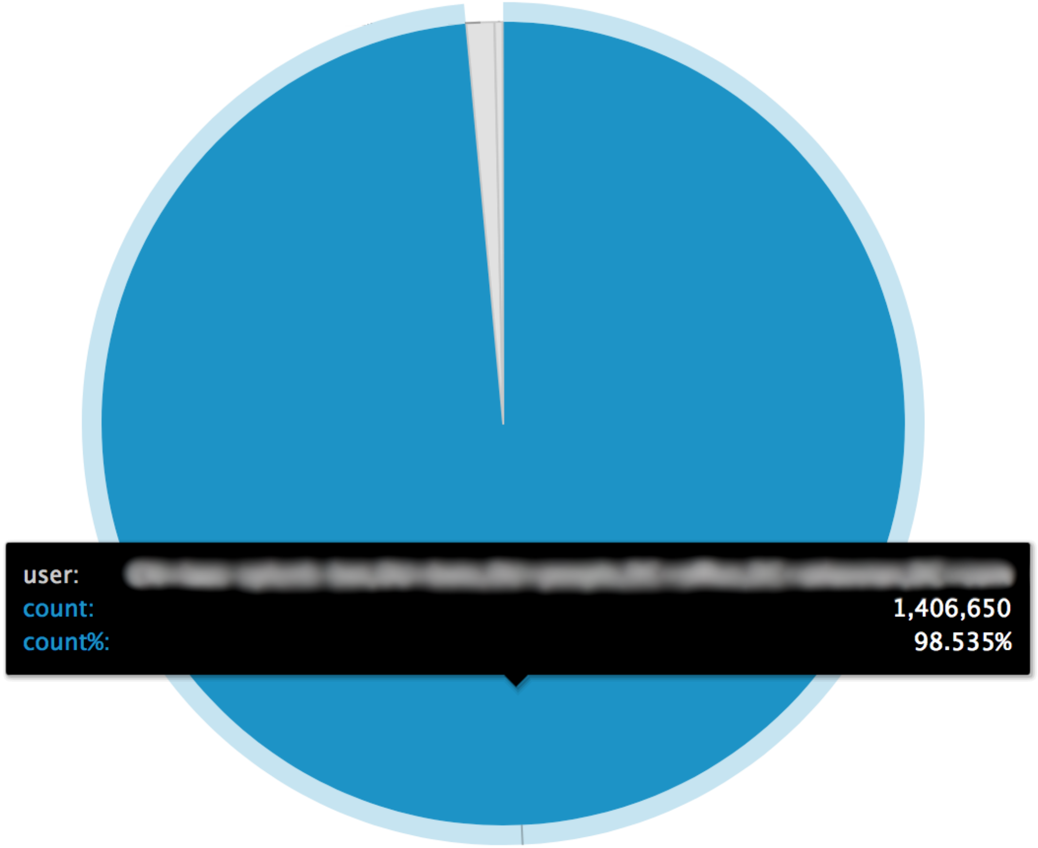
Atlassian Splunk user guide

Active Directory (AD) server load

Splunk instance does the most queries and 98% of the big queries



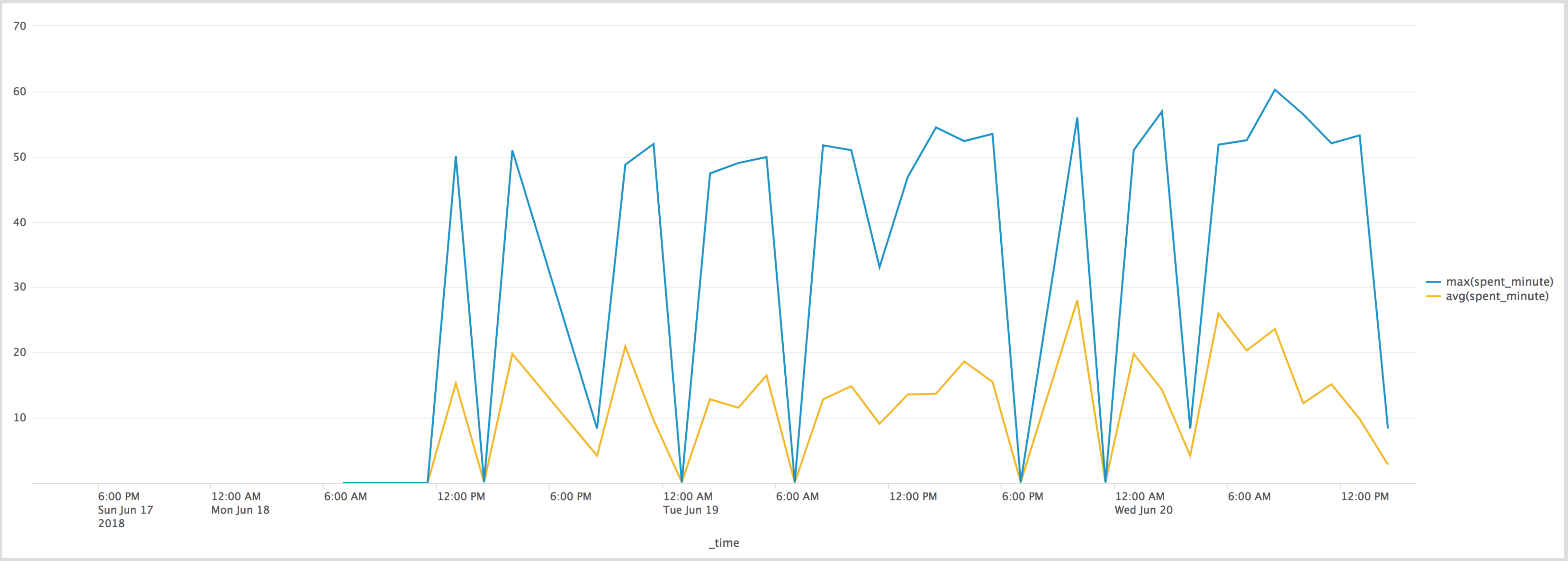
user: [redacted]
count: 1,801,999
count%: 18.83%



user: [redacted]
count: 1,406,650
count%: 98.535%

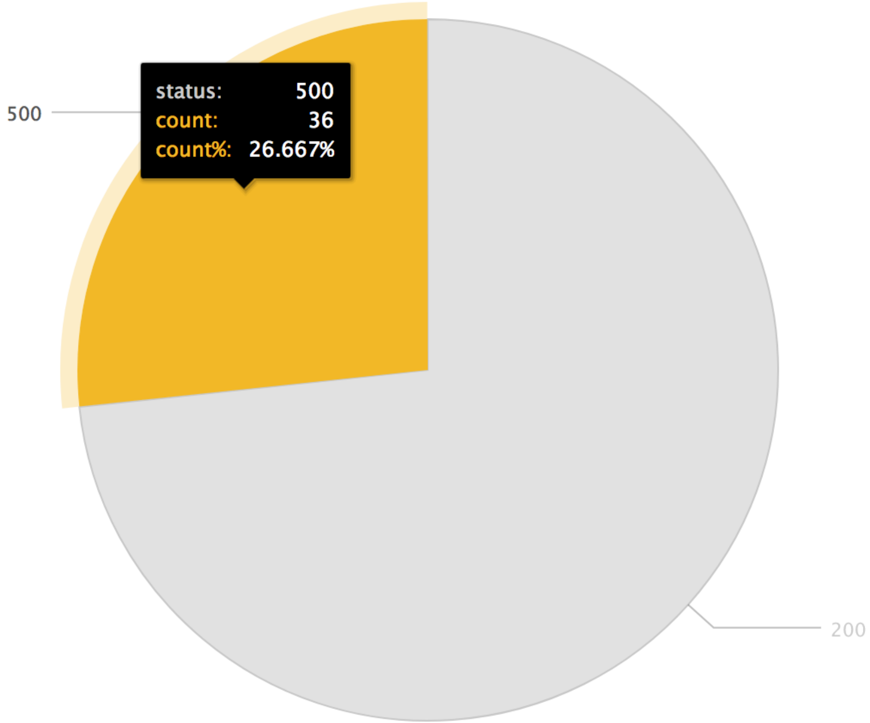
Slow auth config deployment

~1 hour deployment time



Auth config bundle deployment failures

1/4 of our bundle deployment failed



Questions to answer

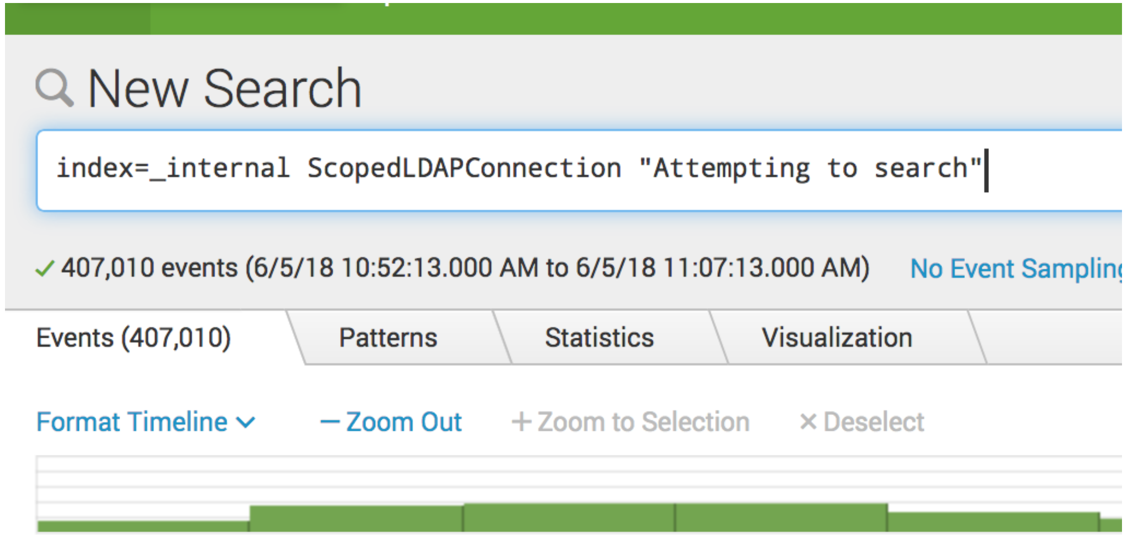
- Why is LDAP component (e.g. login) slow?
  - What LDAP queries does Splunk perform?
  - Why are they slow?
  - Why are they resource heavy?

- Why is bundle deployment slow and error-prone?
  - What does bundle deployment do?
  - Which part of the bundle deployment is slow?

Investigation

What LDAP queries are being done?

- ScopedLDAPConnection = DEBUG
  - 400k LDAP searches in 15 minute!



```
index=_internal ScopedLDAPConnection "Attempting to search"
```

✓ 407,010 events (6/5/18 10:52:13.000 AM to 6/5/18 11:07:13.000 AM) No Event Samplin

What is querying LDAP?

- Components that queries LDAP
  - 1 User login = 300+ LDAP searches
  - Any components that need user roles:
    - Search scheduler
    - Search dispatcher
    - Data model
    - …

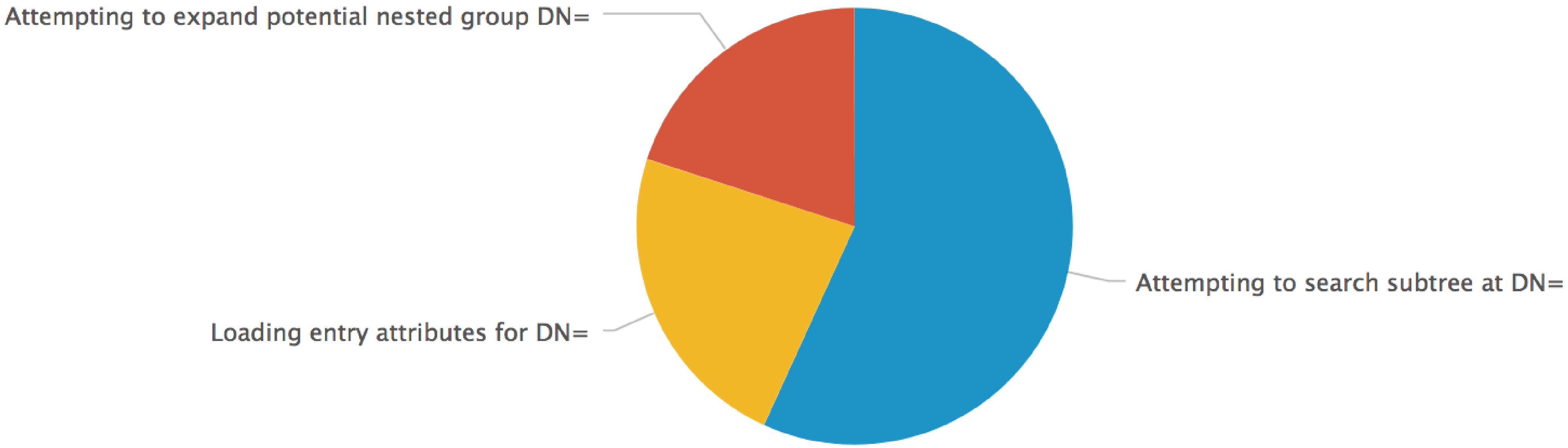
Why are queries expensive?

- Two key factors:
  - The number of LDAP groups
  - Depth of LDAP group nesting

E.g. a troublesome group:

- atlassian-staff:
  - member_groups (55):
    - exec-team
      - Scott
      - Mike
      - …
    - sydney-staff
      - observability-team
      - …
    - mountain-view-staff
    - customer-support
      - amsterdam-support
      - …
    - …

LDAP query breakdown



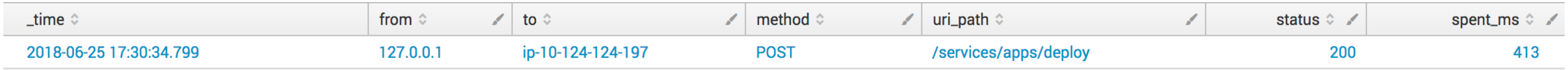
What does bundle deployment do?

Cli

?

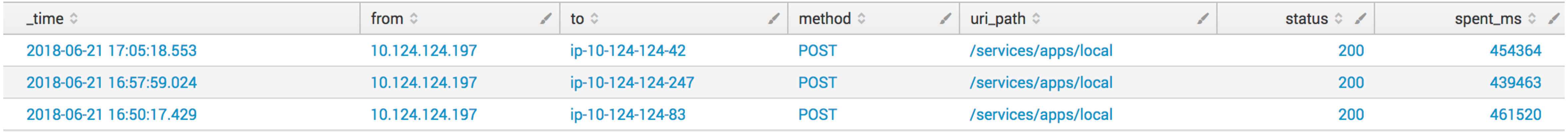
Access Logs

- Deployer: sourcetype=splunkd_access uri_path="/services/apps/deploy"



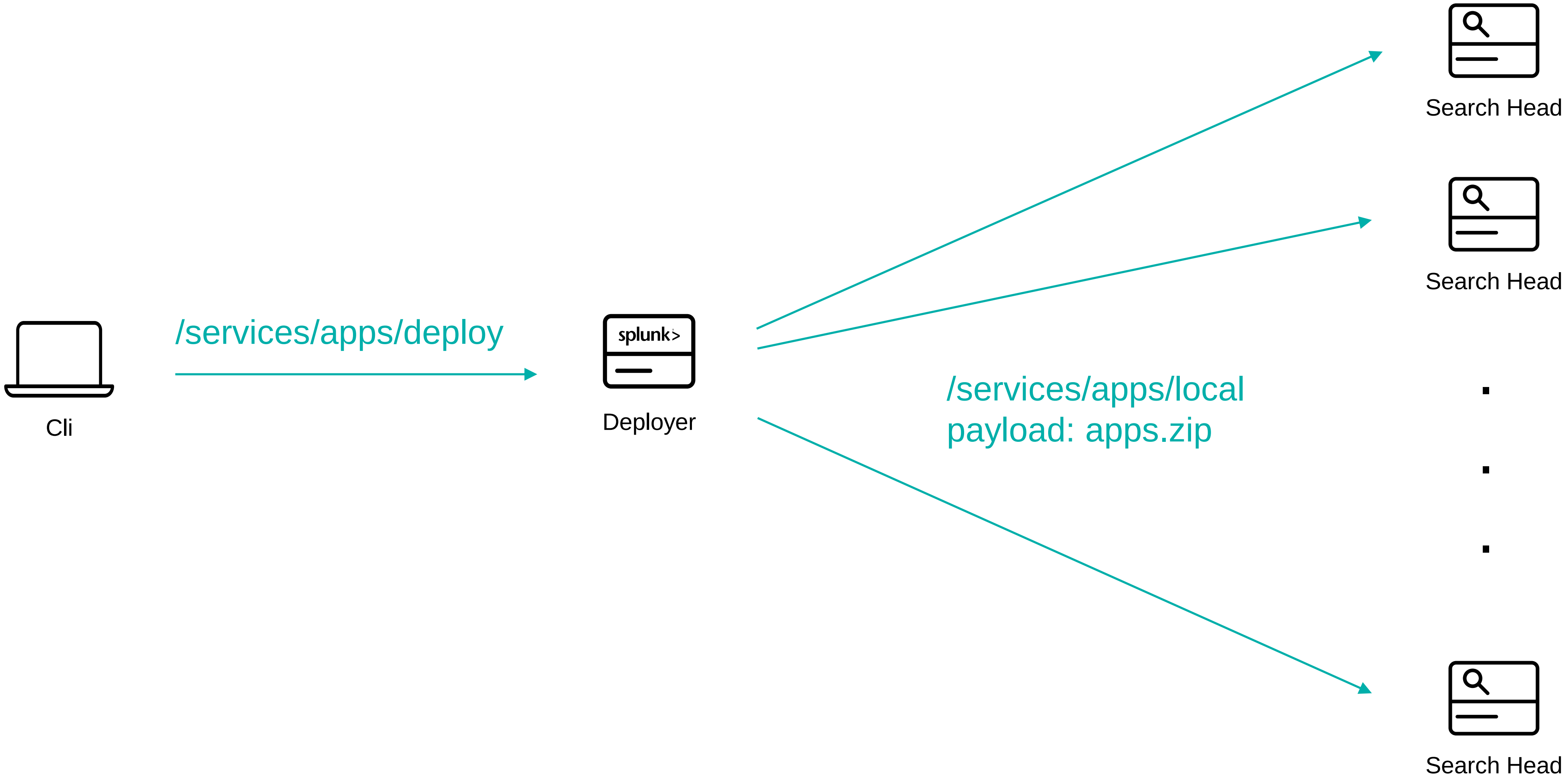
_time	from	to	method	uri_path	status	spent_ms
2018-06-25 17:30:34.799	127.0.0.1	ip-10-124-124-197	POST	/services/apps/deploy	200	413

- Searchhead: sourcetype=splunkd_access uri_path="/services/apps/local"



_time	from	to	method	uri_path	status	spent_ms
2018-06-21 17:05:18.553	10.124.124.197	ip-10-124-124-42	POST	/services/apps/local	200	454364
2018-06-21 16:57:59.024	10.124.124.197	ip-10-124-124-247	POST	/services/apps/local	200	439463
2018-06-21 16:50:17.429	10.124.124.197	ip-10-124-124-83	POST	/services/apps/local	200	461520

Bundle Deployment - Putting them together



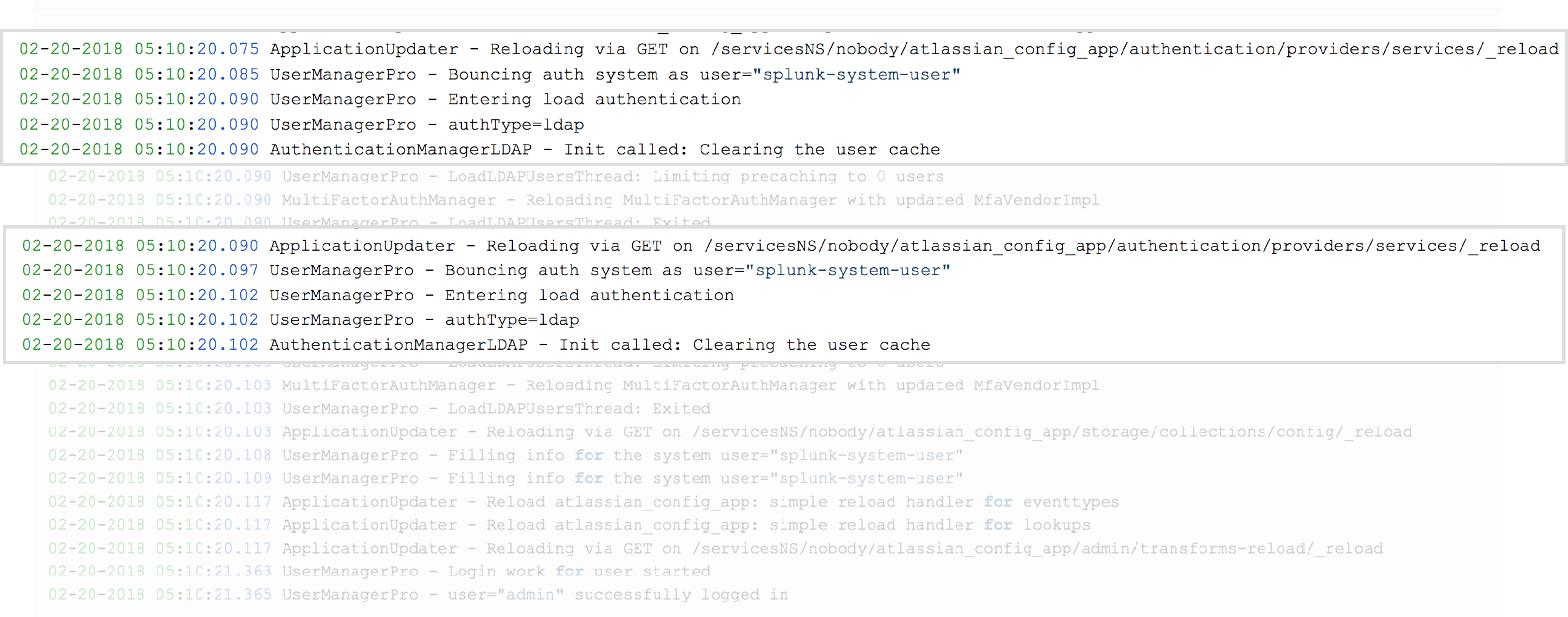
Bundle Installation - Loggers

- Again, we enabled some useful debug logs:
  - AppsDeployHandler
  - LocalAppsAdminHandler
  - ApplicationManager
  - ApplicationUpdater
  - UserManagerPro

Bundle Installation - Debug logs

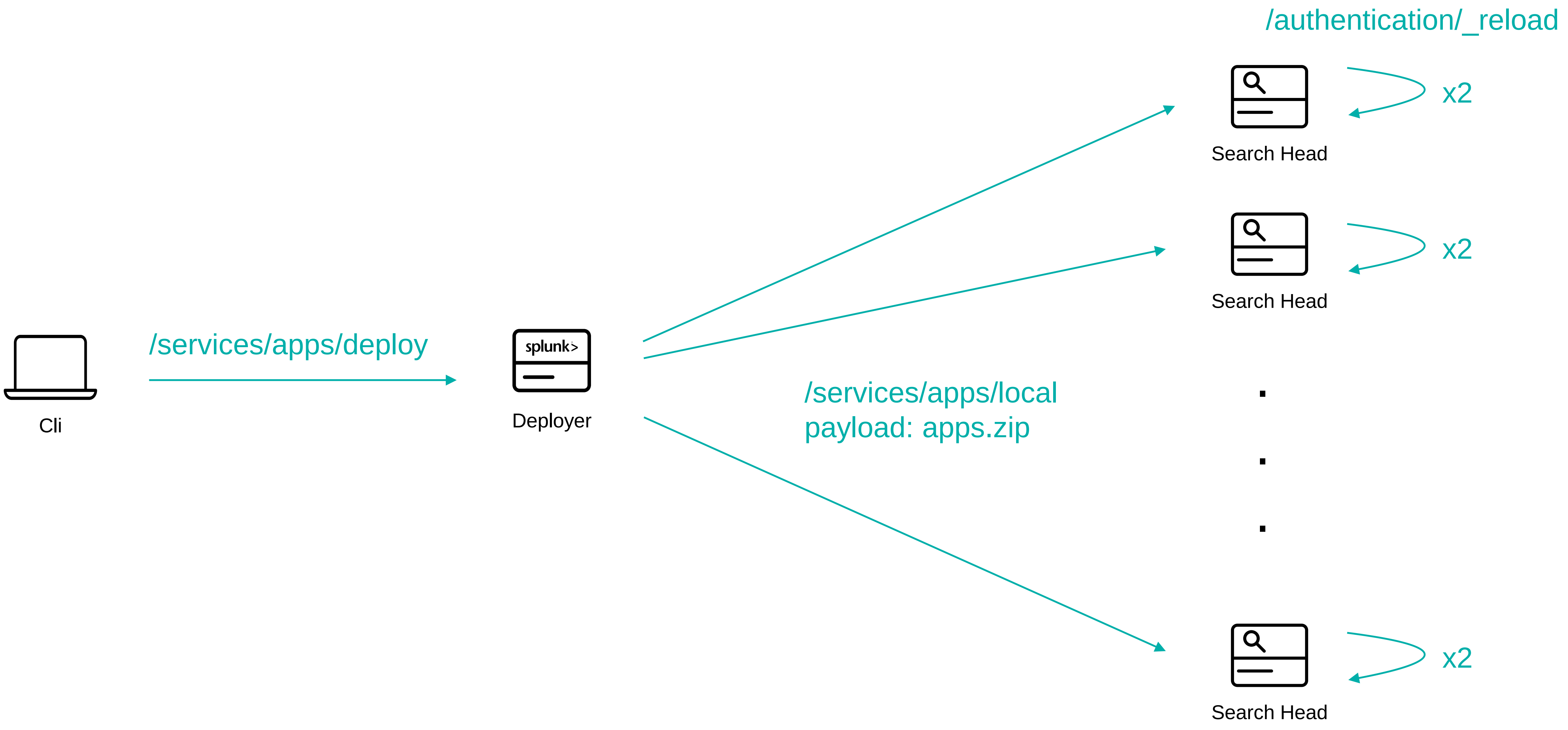
```
02-20-2018 05:10:20.075 ApplicationUpdater - Reload atlassian_config_app: simple reload handler for app
02-20-2018 05:10:20.075 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/authentication/providers/services/_reload
02-20-2018 05:10:20.085 UserManagerPro - Bouncing auth system as user="splunk-system-user"
02-20-2018 05:10:20.090 UserManagerPro - Entering load authentication
02-20-2018 05:10:20.090 UserManagerPro - authType=ldap
02-20-2018 05:10:20.090 AuthenticationManagerLDAP - Init called: Clearing the user cache
02-20-2018 05:10:20.090 UserManagerPro - LoadLDAPUsersThread: Limiting precaching to 0 users
02-20-2018 05:10:20.090 MultiFactorAuthManager - Reloading MultiFactorAuthManager with updated MfaVendorImpl
02-20-2018 05:10:20.090 UserManagerPro - LoadLDAPUsersThread: Exited
02-20-2018 05:10:20.090 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/authentication/providers/services/_reload
02-20-2018 05:10:20.097 UserManagerPro - Bouncing auth system as user="splunk-system-user"
02-20-2018 05:10:20.102 UserManagerPro - Entering load authentication
02-20-2018 05:10:20.102 UserManagerPro - authType=ldap
02-20-2018 05:10:20.102 AuthenticationManagerLDAP - Init called: Clearing the user cache
02-20-2018 05:10:20.103 UserManagerPro - LoadLDAPUsersThread: Limiting precaching to 0 users
02-20-2018 05:10:20.103 MultiFactorAuthManager - Reloading MultiFactorAuthManager with updated MfaVendorImpl
02-20-2018 05:10:20.103 UserManagerPro - LoadLDAPUsersThread: Exited
02-20-2018 05:10:20.103 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/storage/collections/config/_reload
02-20-2018 05:10:20.108 UserManagerPro - Filling info for the system user="splunk-system-user"
02-20-2018 05:10:20.109 UserManagerPro - Filling info for the system user="splunk-system-user"
02-20-2018 05:10:20.117 ApplicationUpdater - Reload atlassian_config_app: simple reload handler for eventtypes
02-20-2018 05:10:20.117 ApplicationUpdater - Reload atlassian_config_app: simple reload handler for lookups
02-20-2018 05:10:20.117 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/admin/transforms-reload/_reload
02-20-2018 05:10:21.363 UserManagerPro - Login work for user started
02-20-2018 05:10:21.365 UserManagerPro - user="admin" successfully logged in
```

Bundle Installation - Debug logs

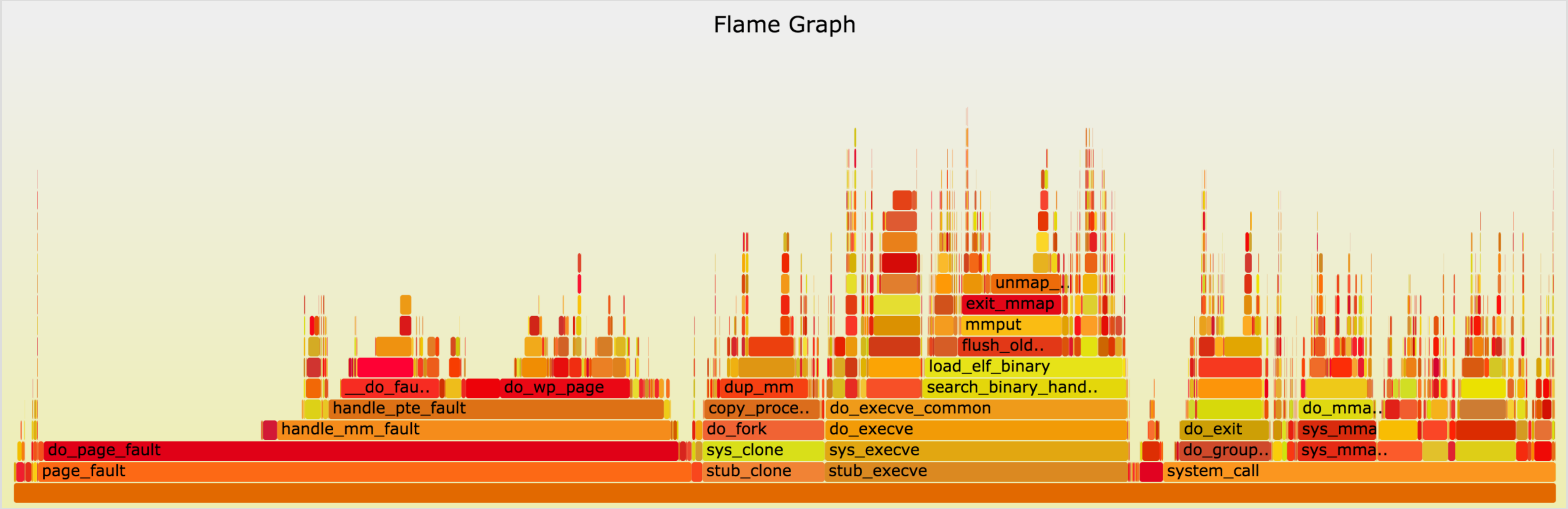


```
02-20-2018 05:10:20.075 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/authentication/providers/services/_reload
02-20-2018 05:10:20.085 UserManagerPro - Bouncing auth system as user="splunk-system-user"
02-20-2018 05:10:20.090 UserManagerPro - Entering load authentication
02-20-2018 05:10:20.090 UserManagerPro - authType=ldap
02-20-2018 05:10:20.090 AuthenticationManagerLDAP - Init called: Clearing the user cache
02-20-2018 05:10:20.090 UserManagerPro - LoadLDAPUsersThread: Limiting precaching to 0 users
02-20-2018 05:10:20.090 MultiFactorAuthManager - Reloading MultiFactorAuthManager with updated MfaVendorImpl
02-20-2018 05:10:20.090 UserManagerPro - LoadLDAPUsersThread: Exited
02-20-2018 05:10:20.090 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/authentication/providers/services/_reload
02-20-2018 05:10:20.097 UserManagerPro - Bouncing auth system as user="splunk-system-user"
02-20-2018 05:10:20.102 UserManagerPro - Entering load authentication
02-20-2018 05:10:20.102 UserManagerPro - authType=ldap
02-20-2018 05:10:20.102 AuthenticationManagerLDAP - Init called: Clearing the user cache
02-20-2018 05:10:20.103 UserManagerPro - LoadLDAPUsersThread: Limiting precaching to 0 users
02-20-2018 05:10:20.103 MultiFactorAuthManager - Reloading MultiFactorAuthManager with updated MfaVendorImpl
02-20-2018 05:10:20.103 UserManagerPro - LoadLDAPUsersThread: Exited
02-20-2018 05:10:20.103 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/storage/collections/config/_reload
02-20-2018 05:10:20.108 UserManagerPro - Filling info for the system user="splunk-system-user"
02-20-2018 05:10:20.109 UserManagerPro - Filling info for the system user="splunk-system-user"
02-20-2018 05:10:20.117 ApplicationUpdater - Reload atlassian_config_app: simple reload handler for eventtypes
02-20-2018 05:10:20.117 ApplicationUpdater - Reload atlassian_config_app: simple reload handler for lookups
02-20-2018 05:10:20.117 ApplicationUpdater - Reloading via GET on /servicesNS/nobody/atlassian_config_app/admin/transforms-reload/_reload
02-20-2018 05:10:21.363 UserManagerPro - Login work for user started
02-20-2018 05:10:21.365 UserManagerPro - user="admin" successfully logged in
```

Bundle Deployment - Putting them together

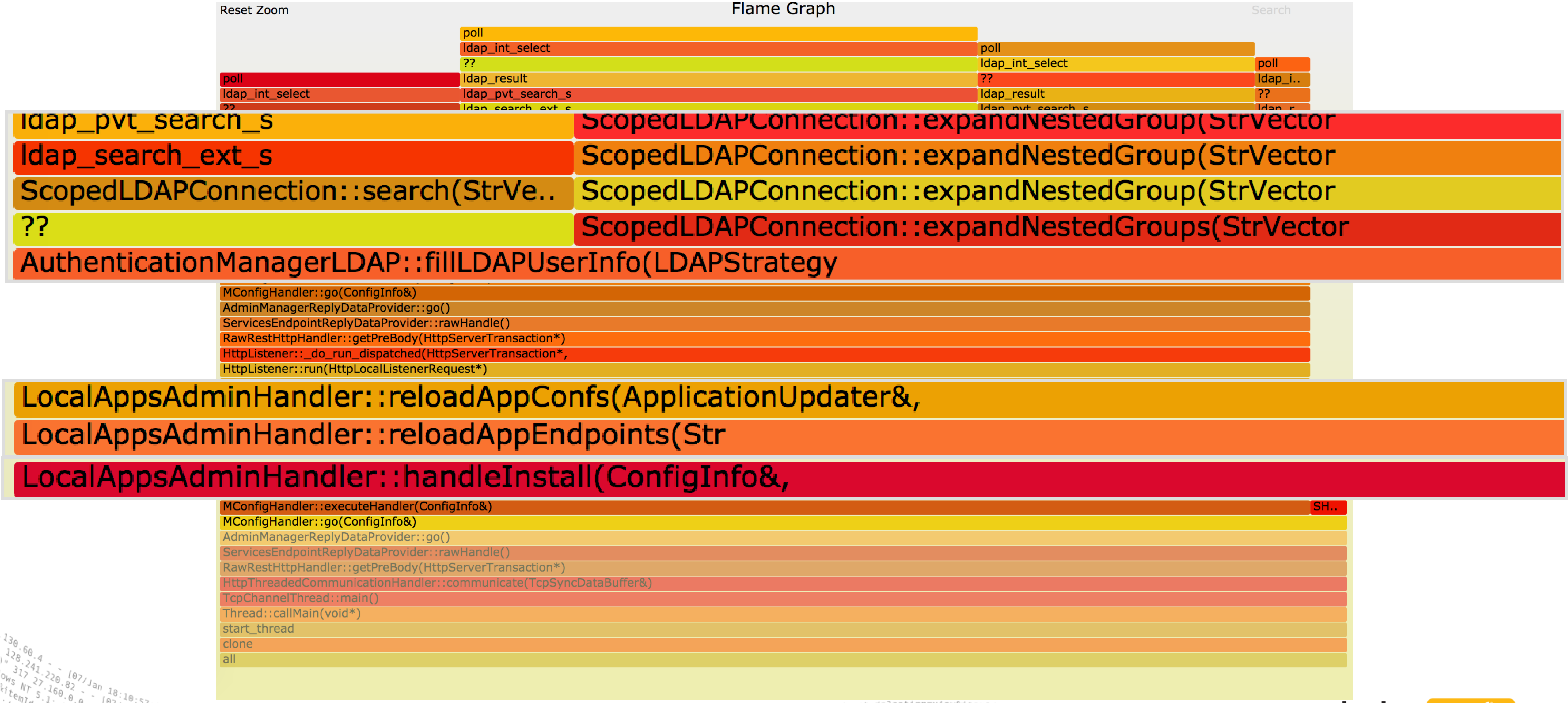


Flamegraphs



Source: *http://www.brendangregg.com/FlameGraphs/cpu-linux-execs.svg*

Bundle Installation Flamegraph



Recap

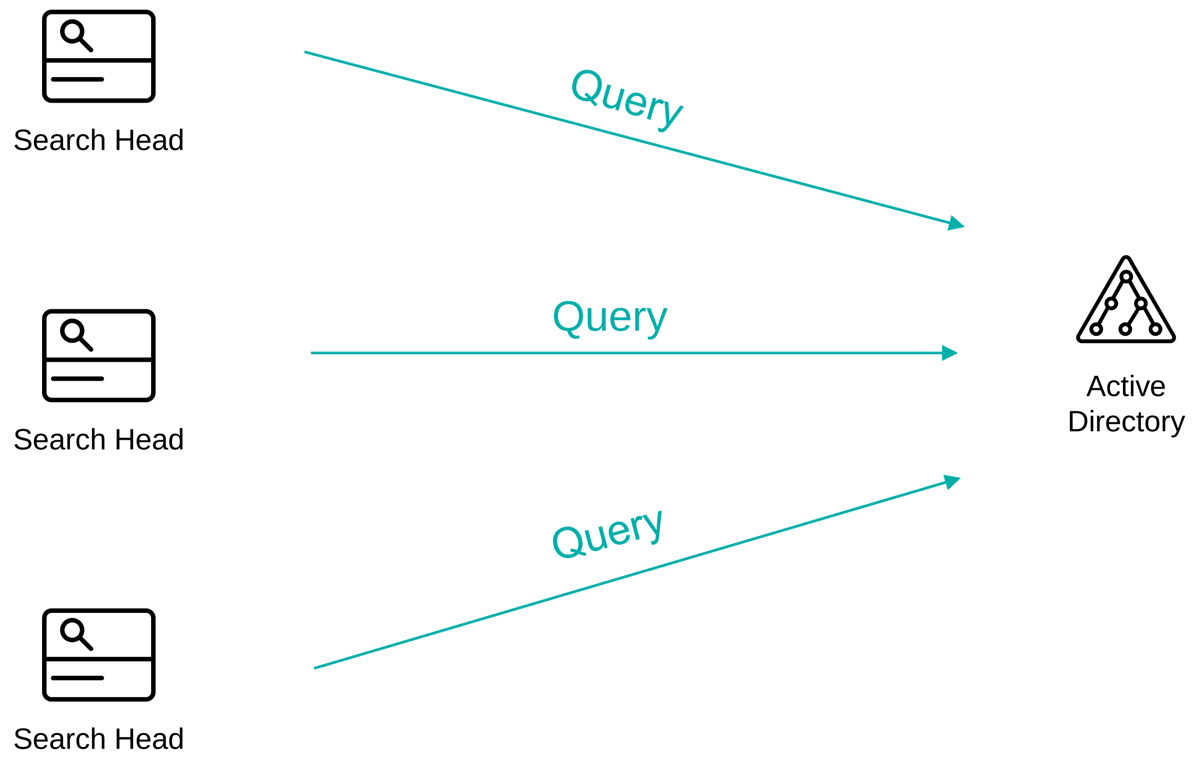
- LDAP component slowness
  - Splunk needs to query a big tree of LDAP groups for each user.

- AD server load
  - Compounding the problem above, a lot of components in Splunk also need to query user info.

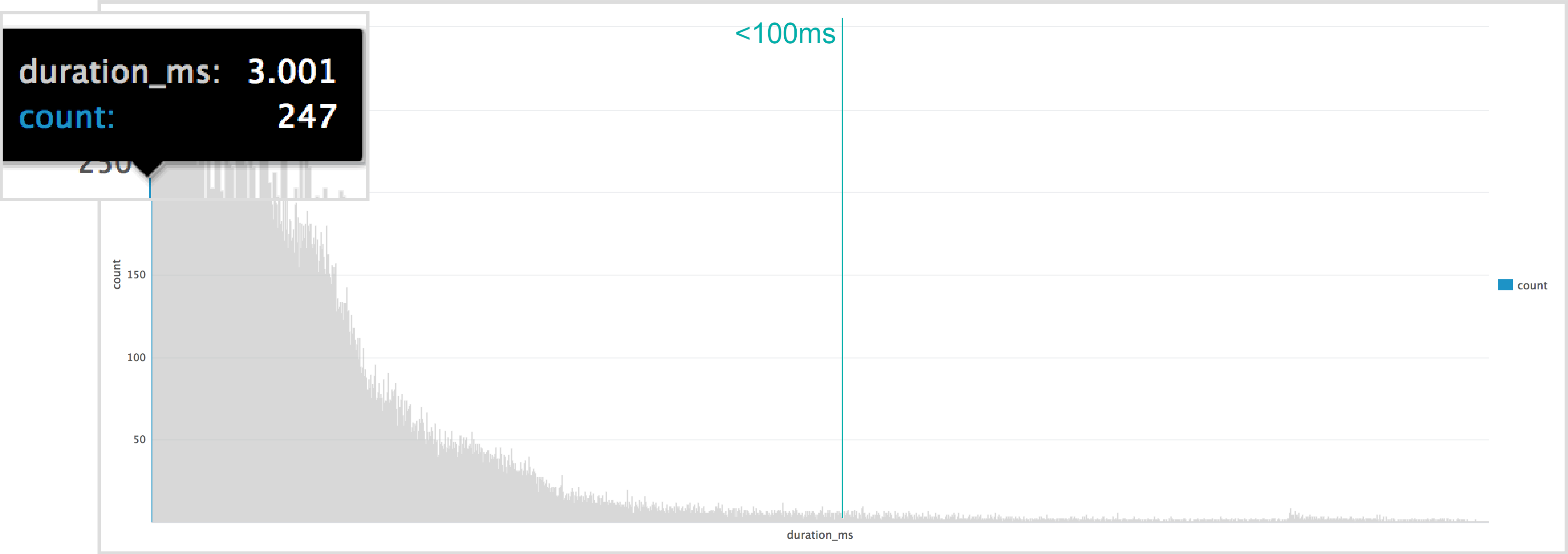
- Slow bundle deployment
  - Bundle installation includes reloading authentication information for everyone **twice**.

Solutions

Previous Architecture



LDAP Cache?


duration_ms: 3.001
count: 247
<100ms
250
150
100
50
count
duration_ms
count

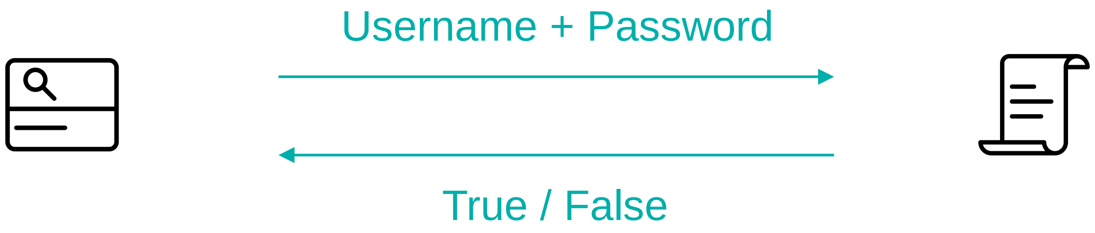
Possible Solutions

- Restructure LDAP groups
- Change authentication scheme

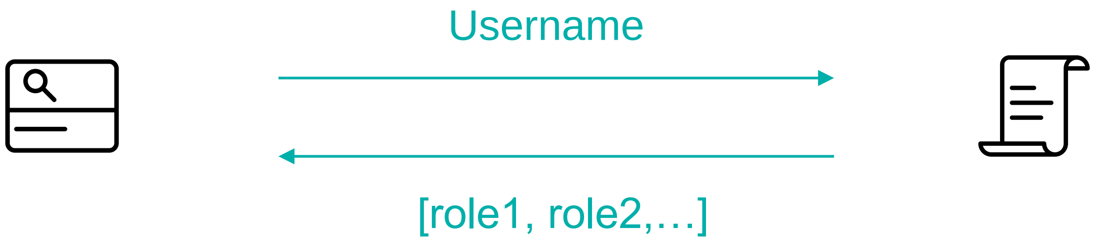
Scripted Auth

Script Functions

- userLogin
  - Can be a basic LDAP bind



- getUserInfo
  - Determines ACL
  - Gets called frequently and concurrently
  - Unrealistic to do LDAP query

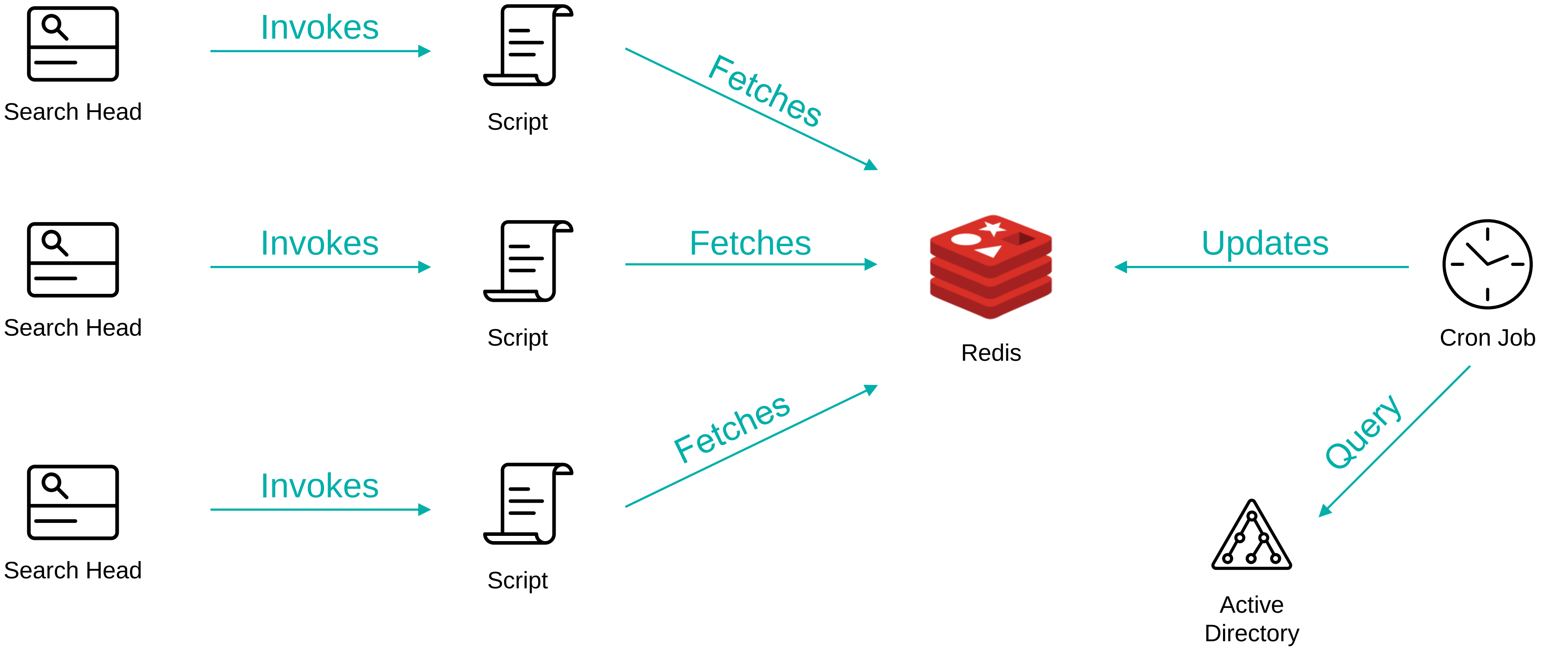


Script Functions

- getUsers
  - Roles are not used for ACL (7.0.4)
  - Determines if schedule search gets run

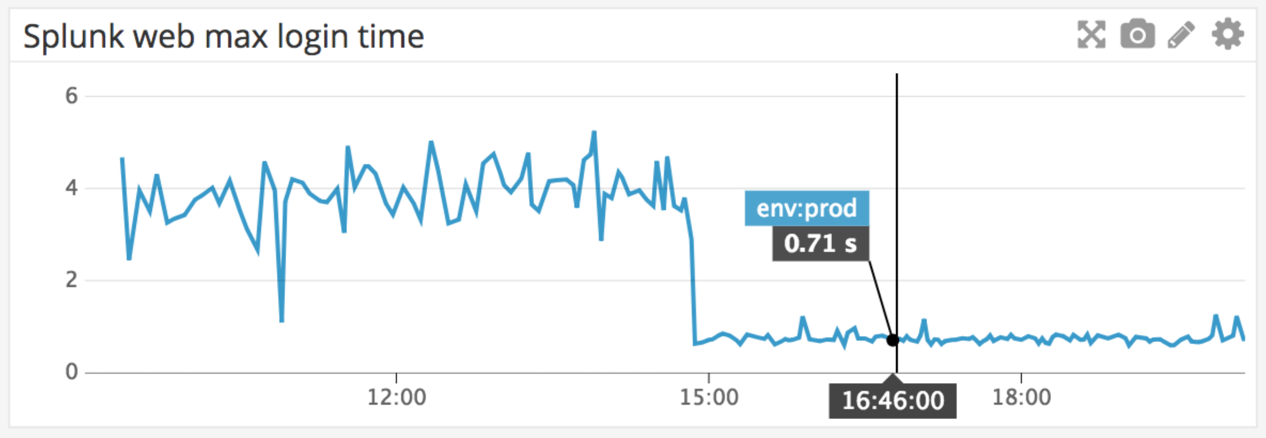
[[username,role1,role2,…],[username,role1,role2,…]]

New Architecture

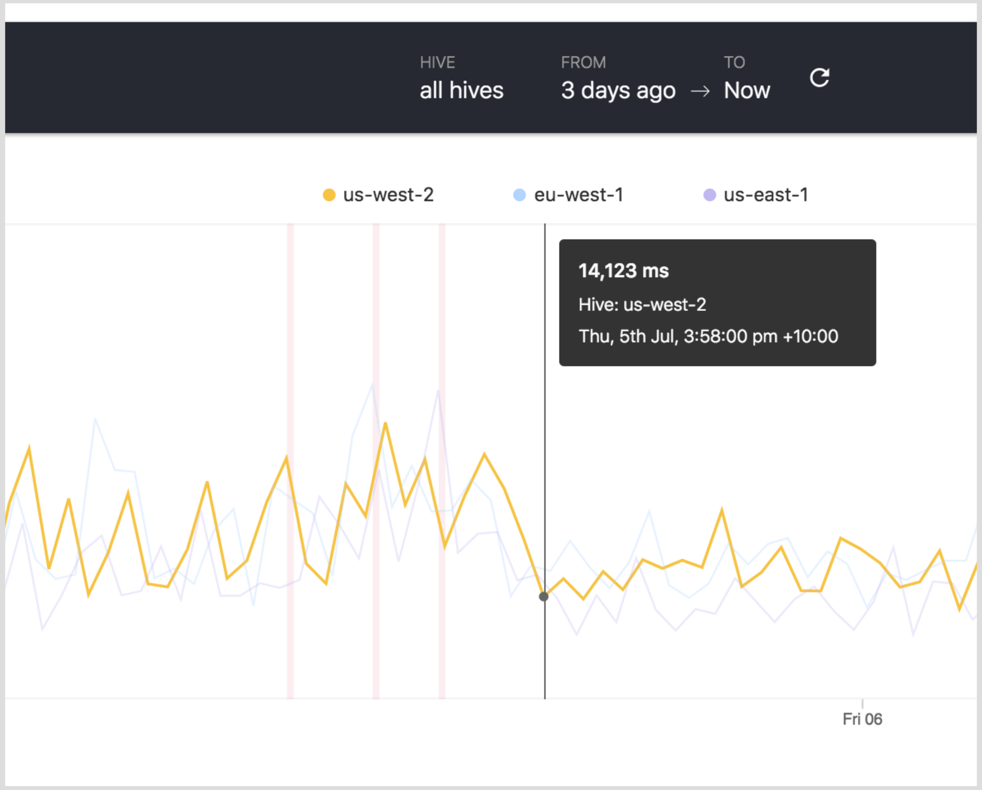


Results

- Login API (4s -> 0.7s)

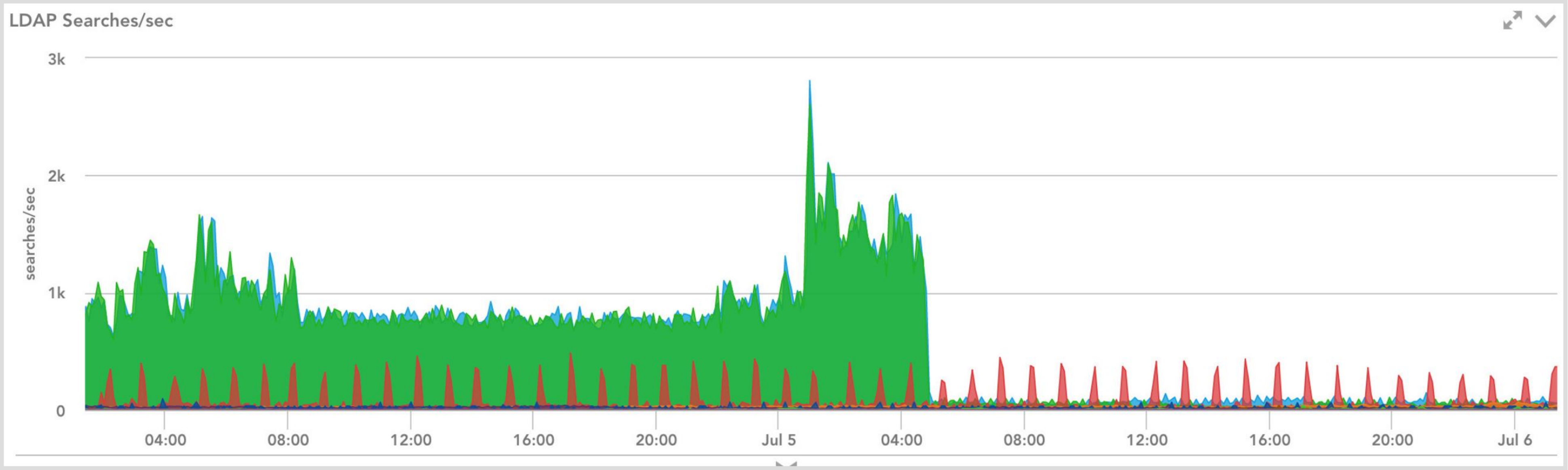


- Login API + Page load (17s -> 14s)



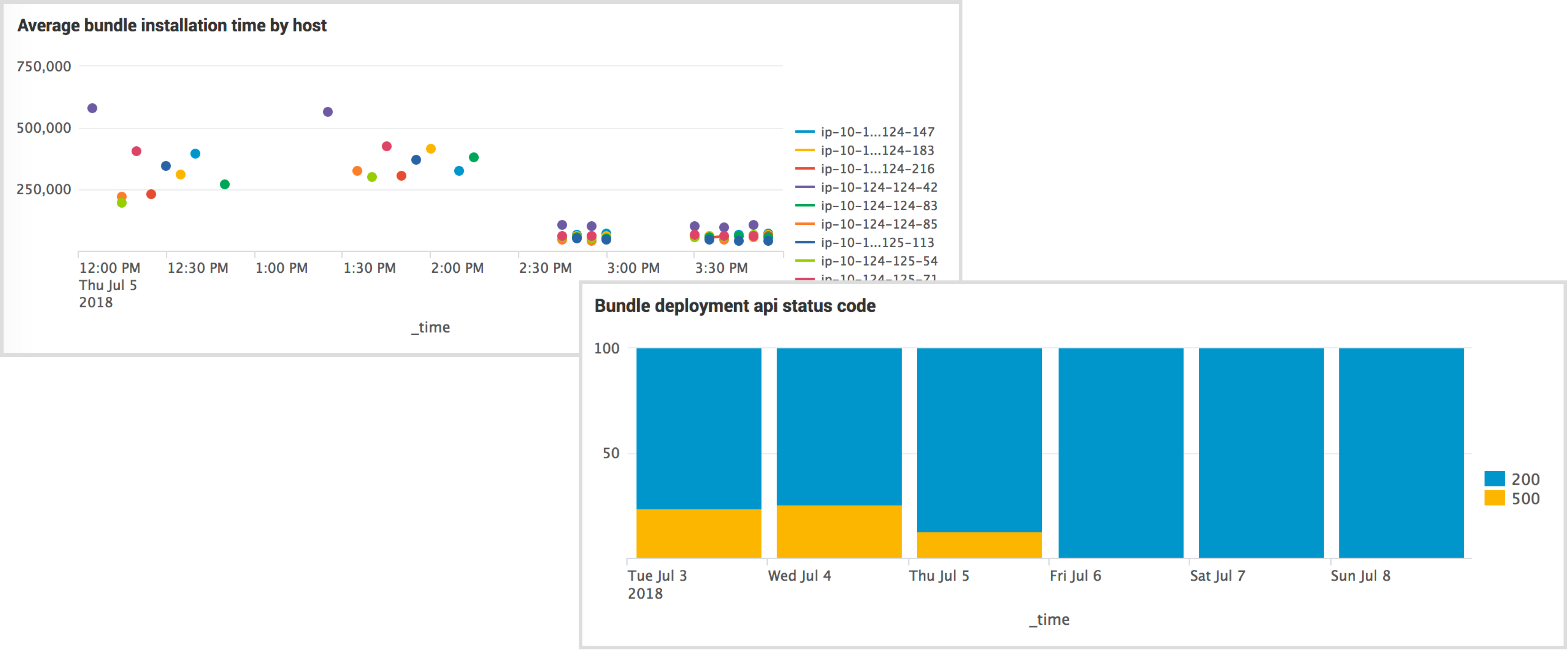
Results

- No more LDAP query spamming



Results

- Faster and less erroneous bundle deployment (8mins -> 1.5mins)

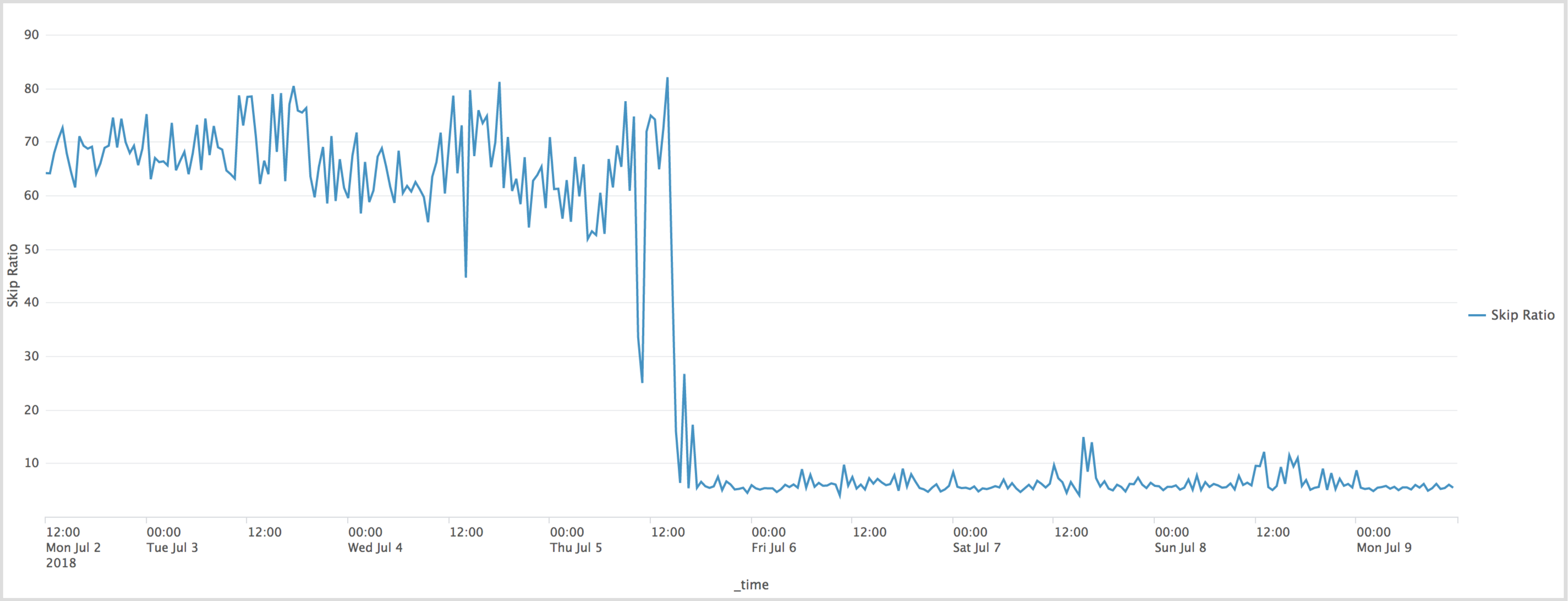


Average bundle installation time by host

Bundle deployment api status code

Results

- Lower saved search skip ratio (75% -> 8%)



Lessons

Key Takeaways

LDAP auth

1. Nested LDAP structure and the number of groups can negatively affect performance
2. Use ScopedLDAPConnection to profile your LDAP queries
3. Beware of auth reloads when deploying authorize.conf and authentication.conf via bundle

Key Takeaways

Splunk Debuging

1. Loggers are great, enable them for troubleshooting
2. Flamegraph can help with finding problematic components
3. Access log gives you a glimpse of the internal workings of Splunk

Key Takeaways

Scripted Auth

1. Instrument your auth script
2. Ingest script logs back into Splunk for debugging
3. The script need to be implemented efficiently

Thank You

Don't forget to rate this session in the .conf18 mobile app

.conf18

splunk>