Essential Component of a Continuous Integration Pipeline

Bill Houston, Eddie Shafaq



Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or plans 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 may differ materially. 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, it may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements made herein.

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 functionalities described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Turn Data Into Doing, 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. © 2019 Splunk Inc. All rights reserved.



Our Speakers

Our Speakers



EDDIE SHAFAQ

Splunk Infrastructure



BILL HOUSTON

Splunk Tooling



Our Speakers

Eddie Shafaq

Splunk Infrastructure

Joined Splunk in August 2011 as a Systems Administrator. Aided in expanding engineering support in "exotic operating system" (AIX, HPUX, S390X and PowerLinux). Served as a member of release engineering to address operational and infrastructure support for products team. Currently serving in an DevOps role around **Engineering Effectiveness and** Infrastructure Engineering services.

Bill Houston

Splunk Tooling

Bill started his career as an analog hardware engineer designing professional recording equipment. Currently he is a senior release engineer at Splunk working on improvements to the Jenkins Cl systems. Prior to Splunk he spent 16 years at Adobe working in various roles; the last four were spent using Jenkins to build and test Adobe Flash.



AgendaIntroduction

- The Challenge Give Devs the results of their CI faster
- Overview Splunk's CI environment
- Splunk/Jenkins Plug-in
- Investigation Using Splunk
- Using Splunk Collecting Build Data
- Build Parallelization
- DistCC, PIGZ
- Test Parallelization
- Multiple test instances
- Notifying Developers
- Jenkins App
- Monitoring Essential Service



Introduction

Why Splunk for CI Use Cases

Multiple critical systems working together

- Some built and hosted internally
- Others consumed as Cloud services.
- Diverse data sources

Need Health & Performance Data (standard troubleshooting)

CPU/Mem/IO

Need Application/Service Level Visibility

- Build process
- Test platform
- Storage

Engineering Productivity

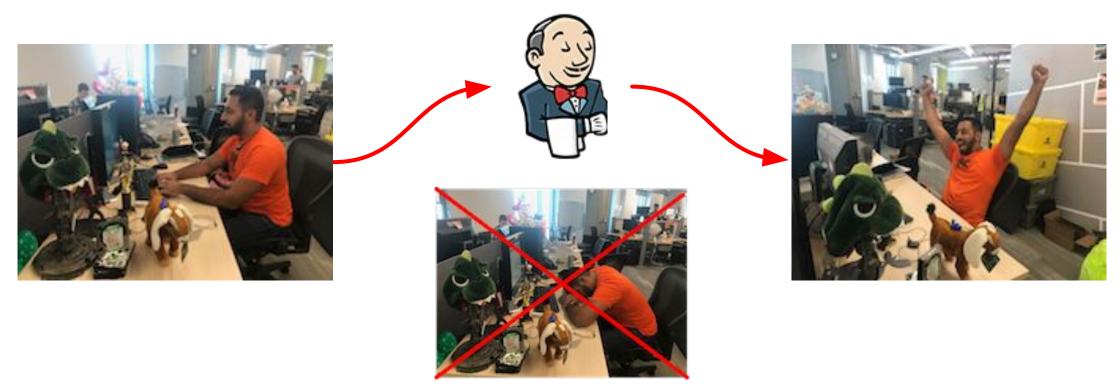
- Test results
- Open stories/tasks/bugs



The Challenge

Improve Engineering Productivity

Our contribution: Get CI test results to developers faster



System Before Optimization



The system has 30 dedicated Linux Docker agents to perform continuous integration testing

Each job ran for approximately **90 minutes**, performing a build of Splunk and running a set of validation tests

That meant it could perform an approximate average of **20 jobs per hour**



System Before Optimization

If more than **20 triggers** were received in a **one hour** period the excess **triggers were queued** waiting for a Docker agent to run on

Under "normal" circumstances the system operated with minimum delays, however during peak load periods when the pressure on developers was the highest...

We experienced **significant delays** resulting in frustration and phone calls as the **engineers waited** for results of the validation test jobs they were required to run before they could commit their work



Our Analysis

Our Analysis

Understanding the Situation

There were four factors that affected the delay developers experienced while waiting for test results

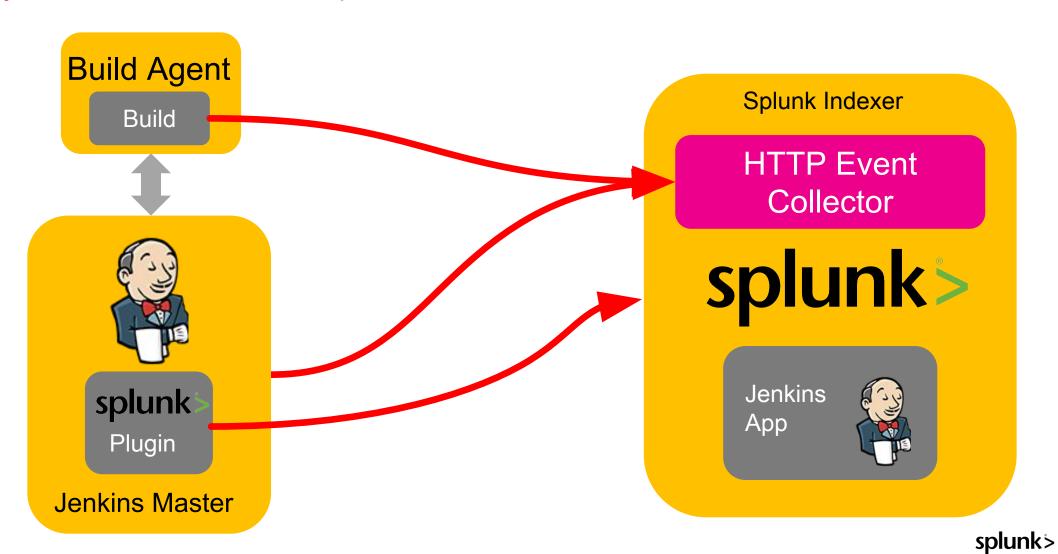
- Build time how long it takes to build the Splunk executables
- Test time how long it takes to perform the required set of tests
- Queue time how long before each phase of the builds start to run
- Notification how long before developers know the test results

We need to quantify each of those factors and determine what we could do to mitigate their effects on the overall time

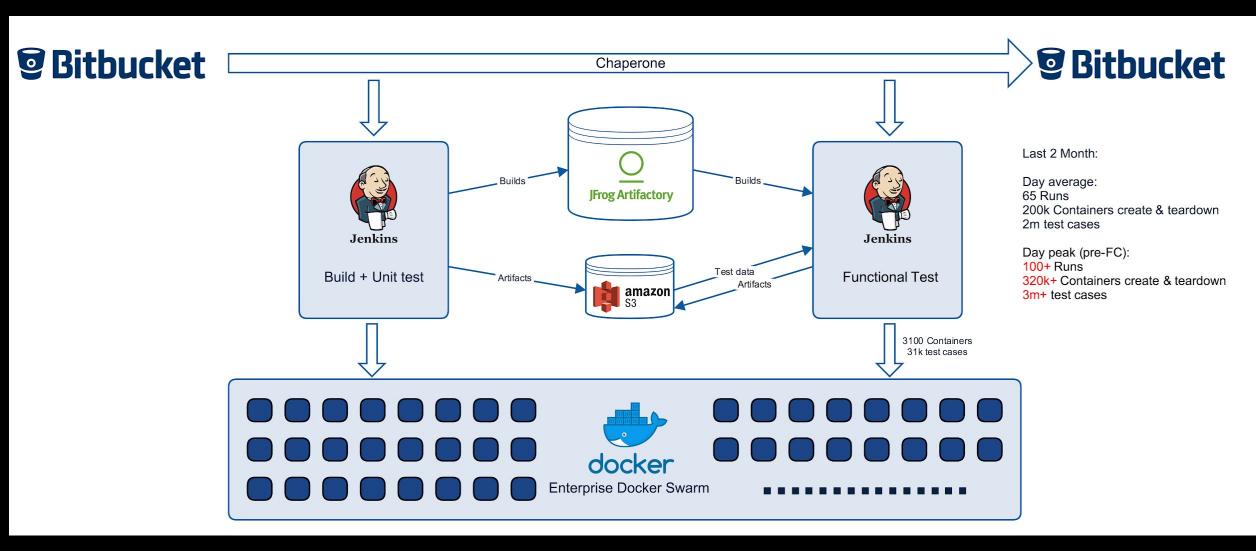


How We Collected The Data

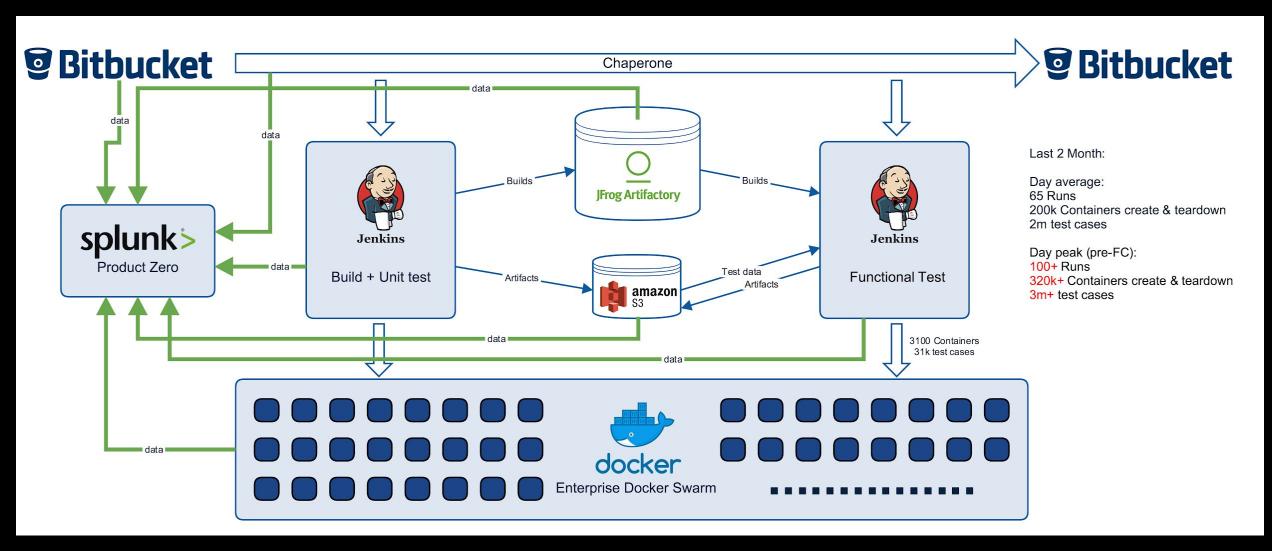
Our jobs are well connected to Splunk



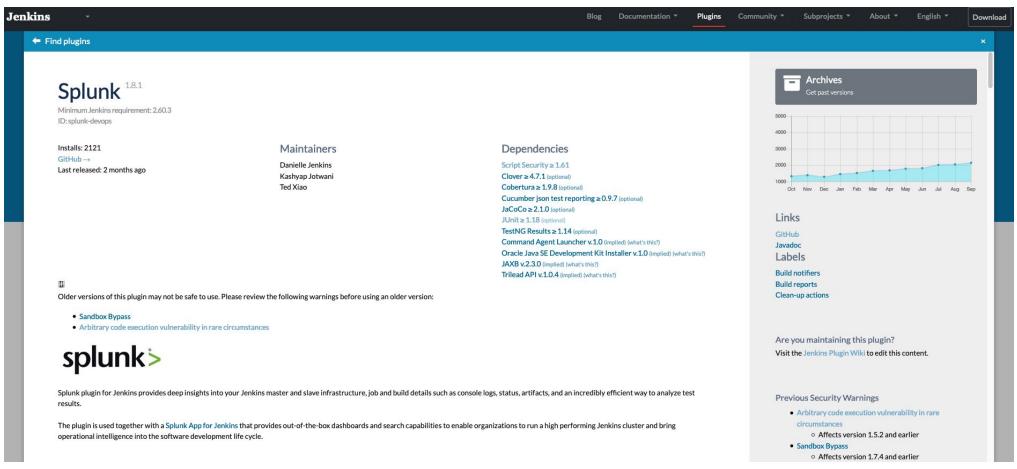
Internal CI Infrastructure - Overview



highlevel



Makes it Easy to Send Your CI/CD Data to Splunk

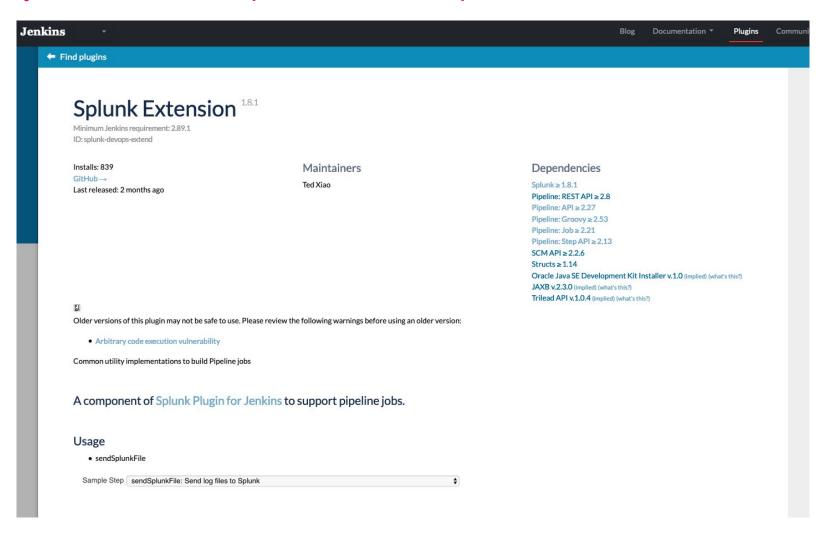


https://wiki.jenkins.io/display/JENKINS/Splunk+
plugin+for+Jenkins



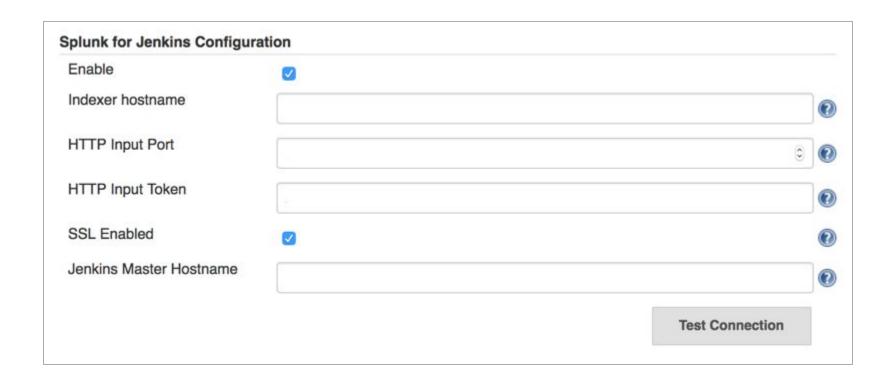
The Splunk Extension Plugin for Jenkins

Makes it Easy to Send Your Pipeline Data to Splunk

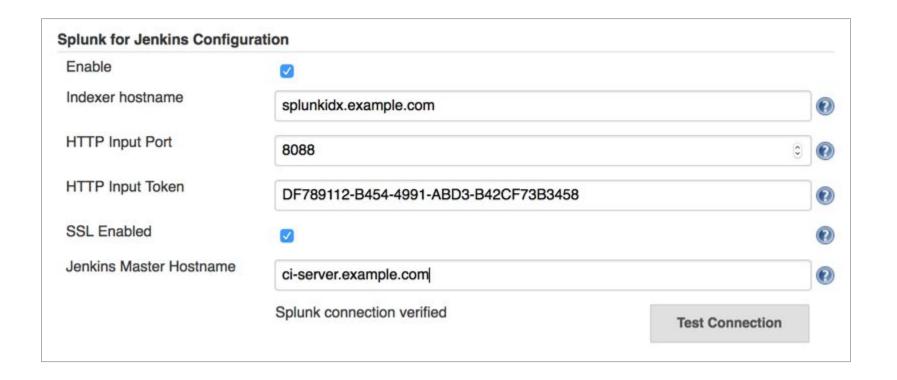




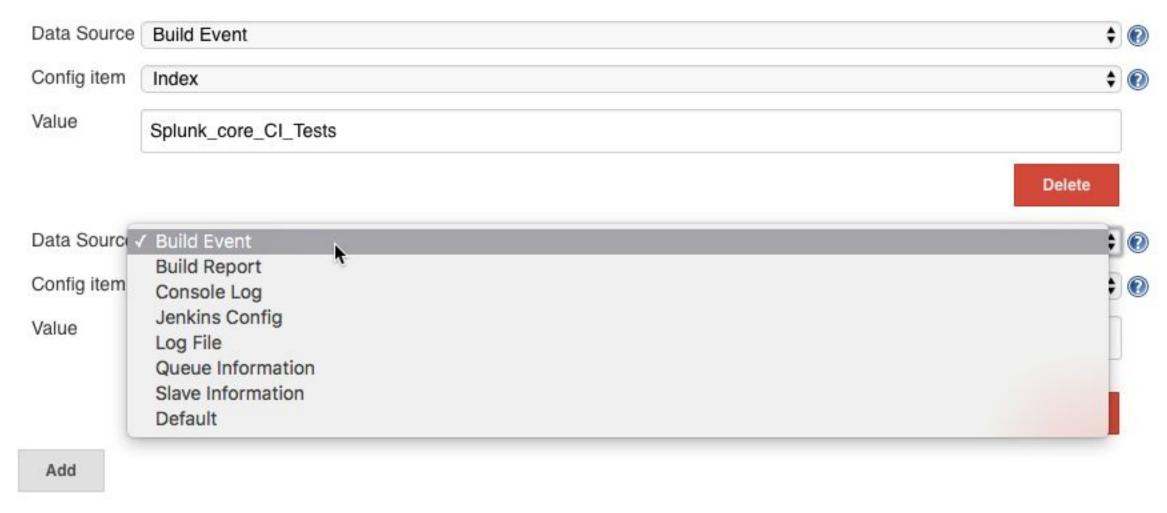
Easy to install



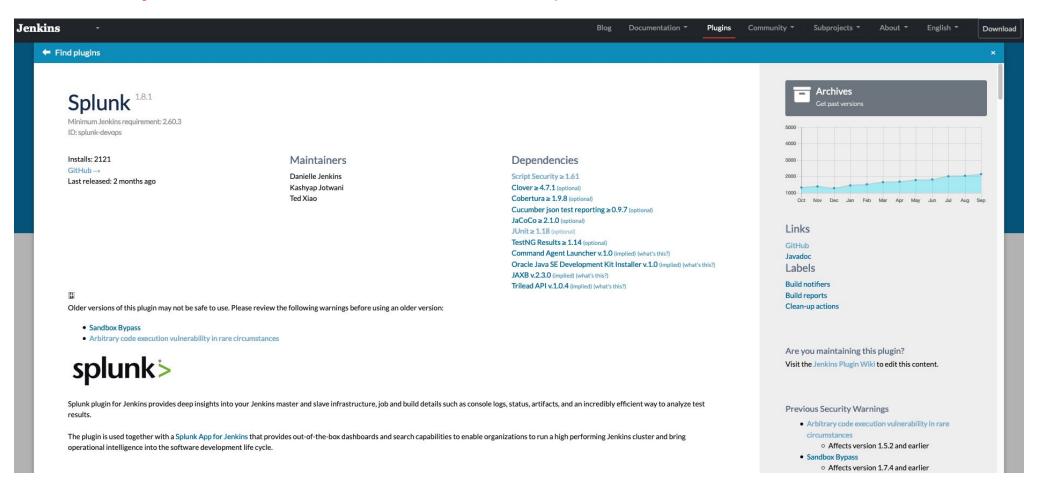
Easy to configure



Easy to customize



Makes it Easy to Send Your CI/CD Data to Splunk





The Splunk HTTP Event Collector

Simple to send custom data to your Splunk instance

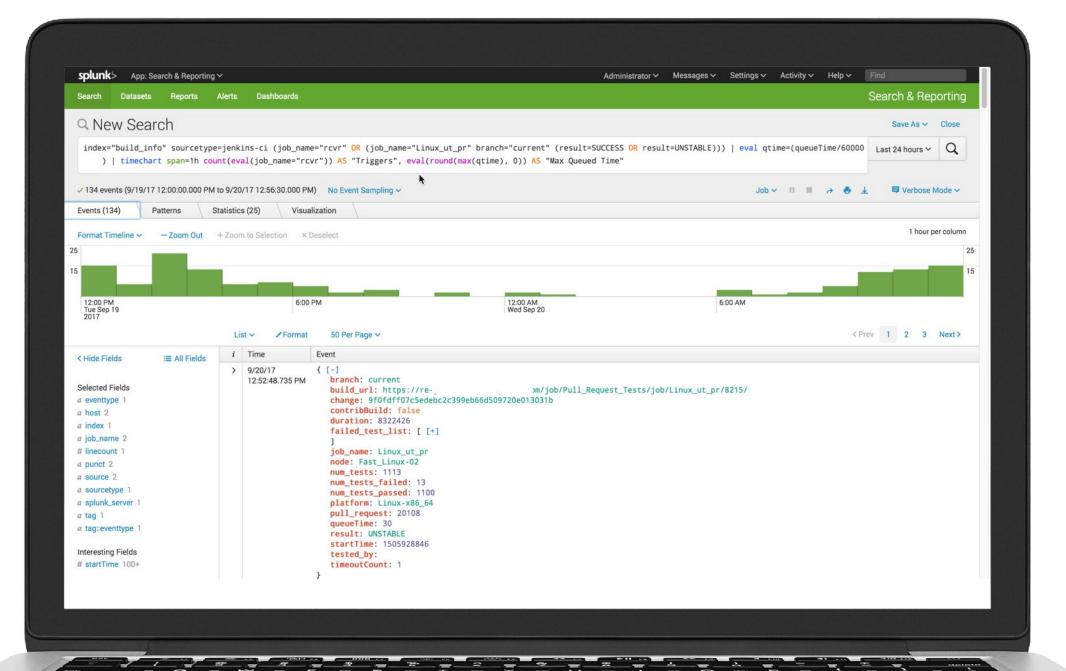
In Bash

```
Format your data as a JSON string:
jsonData="{\"time\": 12345, \"index\": \"YourIndex\", \"sourcetype\": \"YourSourceType\",
\"source\": \"YourSource\", \"event\": \{\"YourFieldName\": \"SomeData\", more json formatted data goes here}}"
Include as much json formatted information as you need in the event section

Then execute a curl call:
curl \
    --tlsv1.2 --header "Authorization: <Splunk_auth_token_goes_here>" \
    --header "Content-Type: application/json" \
    --request 'POST' \
    --data $jsondata \
    https://YourSplunkInstance/services/collector/event
```

Its that simple...





Of course this won't work for you, but...

```
index="jenkins console" host="aJenkins.ourco.com" source="*Linux ut pr*" ("make -j48 | |
"Install the project..." OR "Core build is done" OR "run the tests again" OR "Starting backend unit
tests" OR "Package and publish Splunk" OR "starting Linux 64 test" OR "fetch the jenkins scripts
directory" OR ("nodes run >>>> STARTING ACTION" AND "Write splunk-version.txt") OR
("STARTING COMMAND" AND "Running the contrib command") OR "Done all requested steps")
field=source "job/Pull Request Tests/job/Linux ut pr/(?<build number&gt;.*)/console" | eval
buildStep=case(searchmatch("fetch the jenkins scripts directory"), "start", searchmatch("starting Linux
64 test"), "clone", searchmatch("Running the contrib command"), "chroot", searchmatch("Write
splunk-version.txt"), "contrib", searchmatch("make -j48 | exit 0"), "setup", searchmatch("Core build is
done"), "build 1", searchmatch("Install the project..."), "build 2", searchmatch("Starting backend unit
         "package", searchmatch("run the tests again"), "tests 1", searchmatch("Package and publish
Splunk"), "tests 2", searchmatch("Done all requested steps"), "publish")
                                                                          chart values( time) by
build number, buildStep
                            eval gc = round(('clone' - 'start')/60) | eval cs = round((chroot -
'clone')/60) | eval "cb" = round((contrib - chroot)/60) | eval "bs" = round((setup - contrib)/60)
eval "cub" = round((build 1 - setup)/60) | eval "cbc" = round((build 2 - build 1)/60)
round((package - build 2)/60) | eval "pst" = round((tests 1 - package)/60) |
                                                                            eval "sst" =
round(('tests 2' - 'tests 1')/60) | eval "pub" = round(('publish' - 'tests 2')/60) | search cb < 5 |
search sst > 0 | search pst < 25 | chart values(pub) as Publishing, values(sst) as "Sequential
Smoke Tests", values(pst) as "Parallel Smoke Tests", values(ts) as "Test Setup", values(cbc) as "Core
Build Continues", values(cub) as "Core and UI build", values(bs) as "Build Setup", values(cb) as
"Contrib Build", values(cs) as "Chroot Setup", values(gc) as "Git Clone" by build number
```

Build the final table with user friendly names for display

```
index="jenkins statistics" job name="pr-corebld"
host="jenkins-core-delivery.splunkeng.com"
build number=* type=completed
"metadata.BRANCH"=current | eval fields =
mvzip('stages{}.name','stages{}.duration')
"(?<stagePlatName>[^,]+),(?<stageSeconds>[^,]+)"
rex field=stagePlatName
"(?<plat>[^\n-]+)[\n-](?<stageName>.+)$"
                                       eval
stageMinutes=stageSeconds/60 timechart span=1d
eval(round(avg(stageMinutes), 2)) by stageName
```

Build the final table with user friendly names for display

```
index="jenkins_statistics" job_name="pr-corebld"
host="jenkins-core-delivery.splunkeng.com"
build_number=* type=completed
"metadata.BRANCH"=current
```

Correctly associate each builds stage data

```
| eval fields =
mvzip(mvzip('stages{}.name','stages{}.duration'),
'stages{}.start_time')
```

```
2019-10-21 08:41:01.328
```

```
intel_linux_64_2.6-Sync,86.021,1571667908
intel_linux_64_2.6-Build contrib,70.814,1571667994
intel_linux_64_2.6-Build,996.183,1571668065
intel_linux_64_2.6-Show Warnings,9.532,1571669061
intel_linux_64_2.6-Package,1771.809,1571669071
intel_linux_64_2.6-Publish,1006.698,1571670843
intel_linux_64_2.6-Backend Fast Testing,606.131,1571671850
intel_linux_64_2.6-Backend Unit Results,1.13,1571672456
intel_linux_64_2.6-Postbuild,0.057,1571672457
```

rexpand into stage specific fields with associated time stamps

mvexpand fields

_time \$	fields \$
2019-10-21 08:41:01.328	intel_linux_64_2.6-Sync,86.021,1571667908
2019-10-21 08:41:01.328	intel_linux_64_2.6-Build contrib,70.814,1571667994
2019-10-21 08:41:01.328	intel_linux_64_2.6-Build,996.183,1571668065
2019-10-21 08:41:01.328	intel_linux_64_2.6-Show Warnings,9.532,1571669061
2019-10-21 08:41:01.328	intel_linux_64_2.6-Package,1771.809,1571669071
2019-10-21 08:41:01.328	intel_linux_64_2.6-Publish,1006.698,1571670843
2019-10-21 08:41:01.328	intel_linux_64_2.6-Backend Fast Testing,606.131,1571671850
2019-10-21 08:41:01.328	intel_linux_64_2.6-Backend Unit Results,1.13,1571672456
2019-10-21 08:41:01.328	intel_linux_64_2.6-Postbuild,0.057,1571672457

Build the final table with user friendly names for display

```
| rex field=fields
"(?<stagePlatName>[^,]+),(?<stageSeconds>[^,]+)"
| rex field=stagePlatName
"(?<plat>[^\n-]+)[\n-](?<stageName>.+)$" | eval
| stageMinutes=stageSeconds/60
```

Build the final table with user friendly names for display

_time \$	stageName \$	/	stageMinutes /
2019-10-21 08:41:01.328	Sync		1.4337
2019-10-21 08:41:01.328	Build contrib		1.1802
2019-10-21 08:41:01.328	Build		16.6031
2019-10-21 08:41:01.328	Show Warnings		0.1589
2019-10-21 08:41:01.328	Package		29.53015
2019-10-21 08:41:01.328	Publish		16.77830
2019-10-21 08:41:01.328	Backend Fast Testing		10.1022
2019-10-21 08:41:01.328	Backend Unit Results		0.0188
2019-10-21 08:41:01.328	Postbuild		0.00095

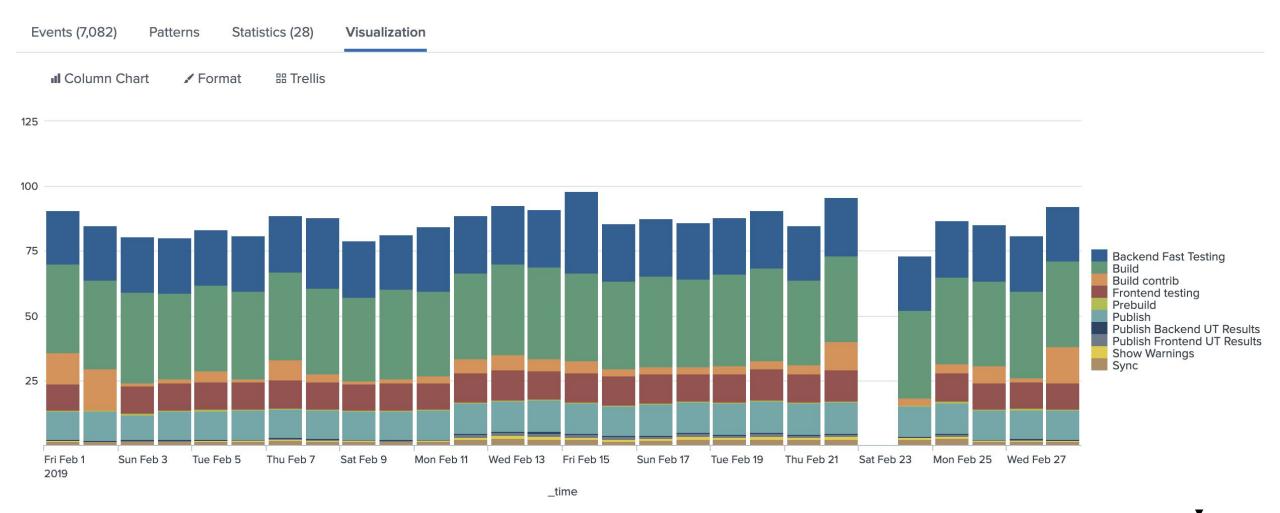
Build the final time chart showing each builds results

```
| timechart span=1d eval(round(avg(stageMinutes),
1)) by stageName
```

Our Analysis

How long does each step of a job take?

We used the previous search to chart the time each step took



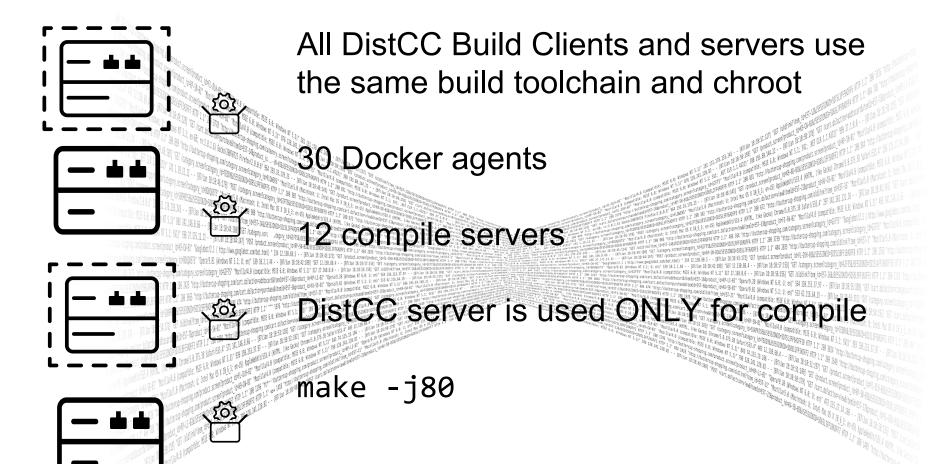


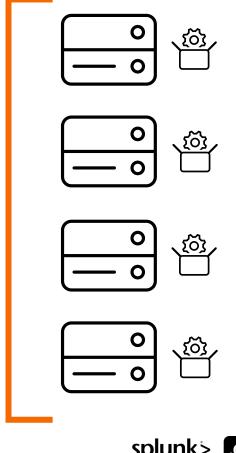
Speeding up the Splunk Build

Distcc Architecture

Containers:

12 x 20 Core VMs







DistCC VS Normal Build

- ▶ 24 Min Build
 - make -j 24
 - Web UI –j 1
 - Optimal 24 core VM agent

24m ----

- 19 Min Build
- make -j 48
- Web UI -<u>j</u> 1

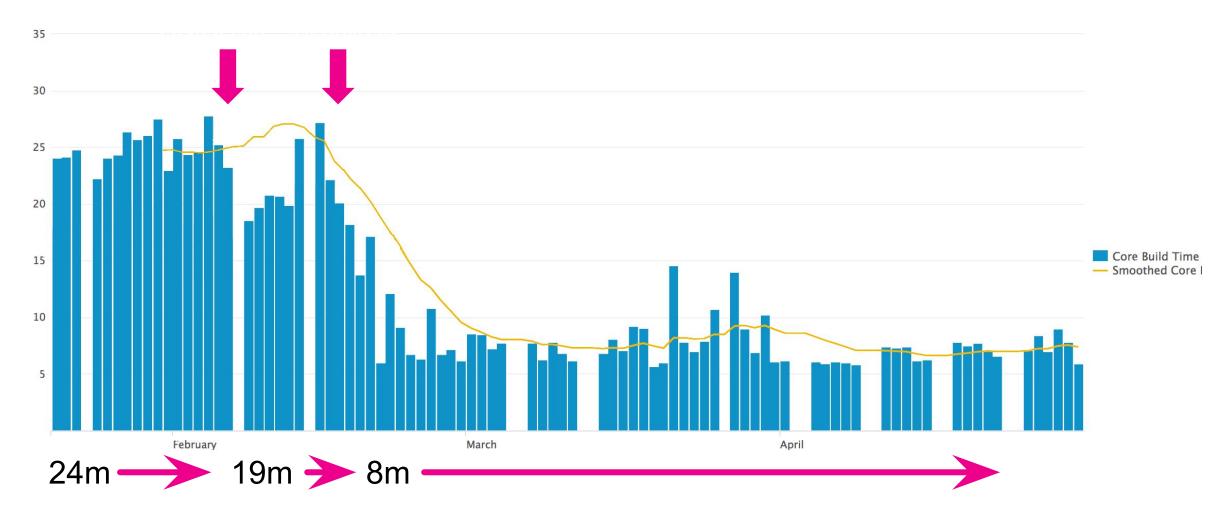
- Optimal 24 core VM agent
 - 12 DistCC hosts

19m ———

- ▶ 8 Min Build
 - make -j 48
 - Web UI -j 6
 - Optimal 24 core VM agent
 - 12 DistCC hosts

Build Time Improvement - Results

Dramatic reduction in the overall build time





Speeding up Testing

Our Analysis

Reducing the Build Testing Time

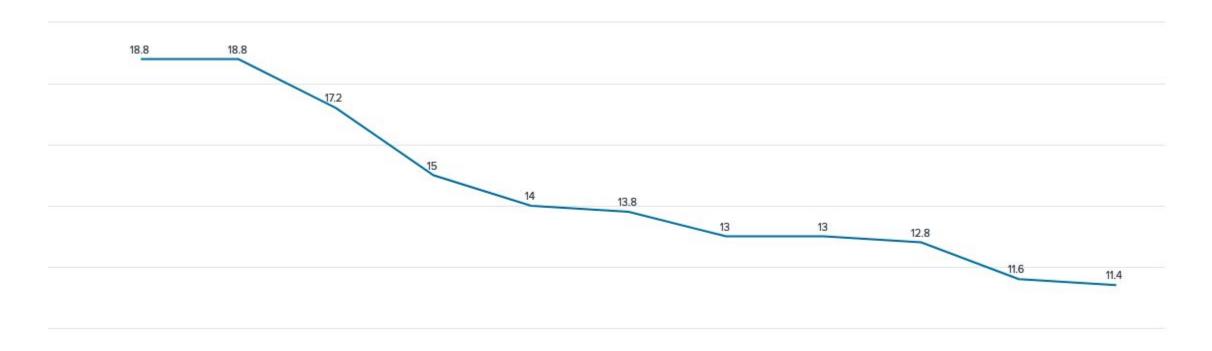
Analysis:

- Based on our previous analysis the backend unit tests were executed in parallel across 16 test instances on the build agents
- We used Splunk to measure the overall timing of each test configuration as well as the individual tests

Mitigation:

- Over time we increased the number of parallel test instances from 16 to 40 measuring the overall test time
- We stopped at 40 parallel tests because the build containers have 48 cores

Improvement in the overall test time as parallel instances were increased



Before Optimization

Peak load period – > significant delays

Fri Feb 1

2019

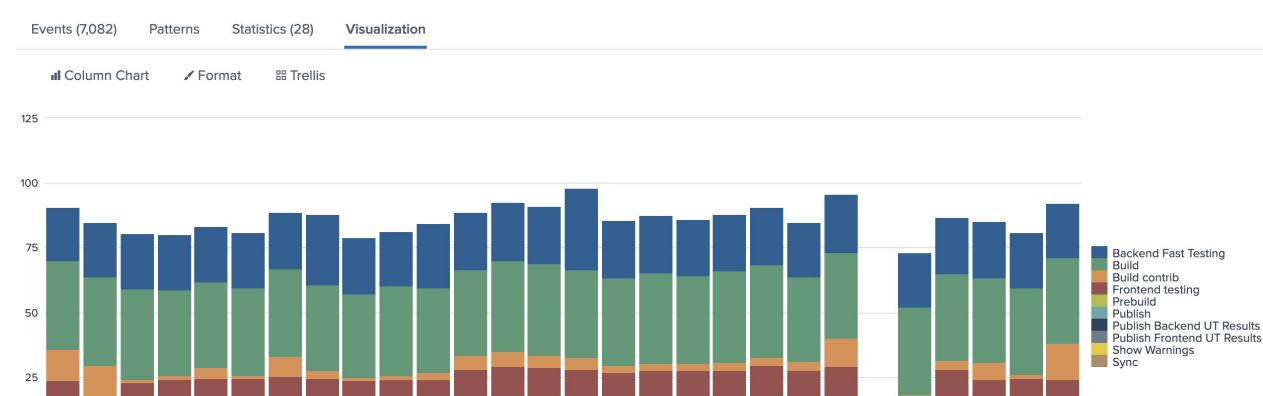
Sun Feb 3

Tue Feb 5

Thu Feb 7

Sat Feb 9

Mon Feb 11



Fri Feb 15

Sun Feb 17

Tue Feb 19

Thu Feb 21

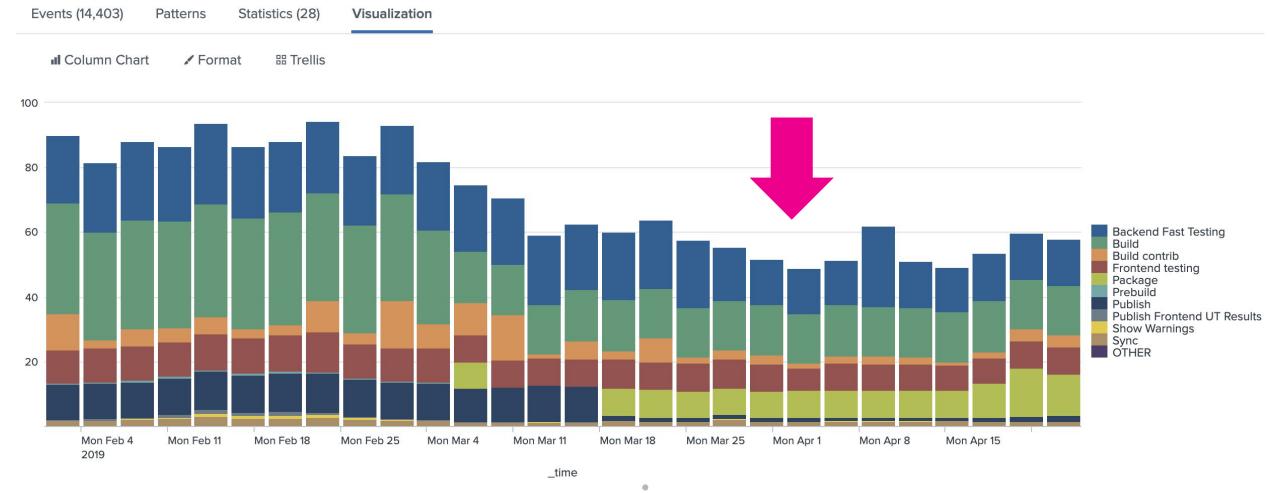
Sat Feb 23

Mon Feb 25 Wed Feb 27

Wed Feb 13

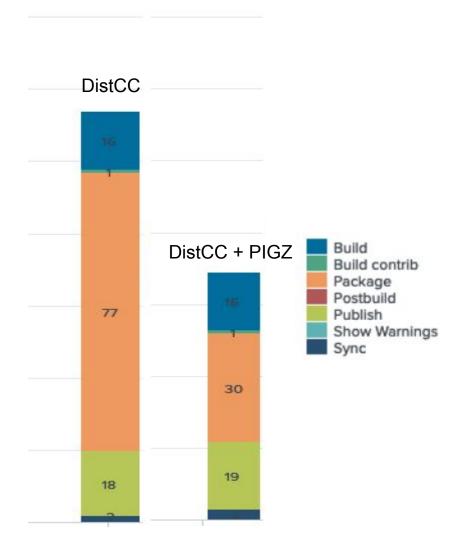
After Optimization

Recent similar trigger conditions





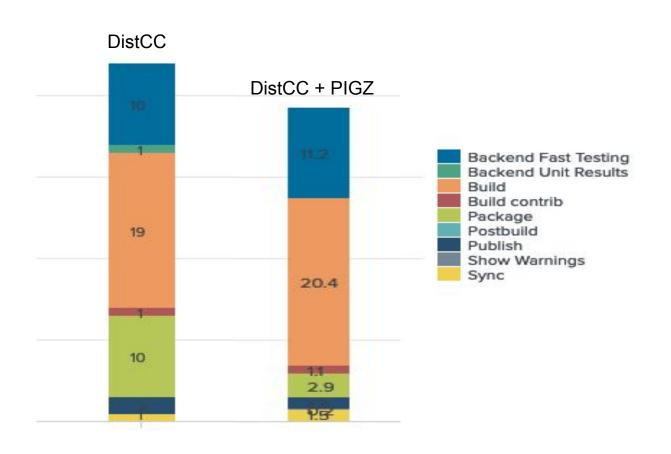
Production release build improvement with



- Production release build
 - 5 Components
 - 10 archives created
- Multicore archiving program
 - Utilizes multiple cores on the container

2x time reduction for packaging

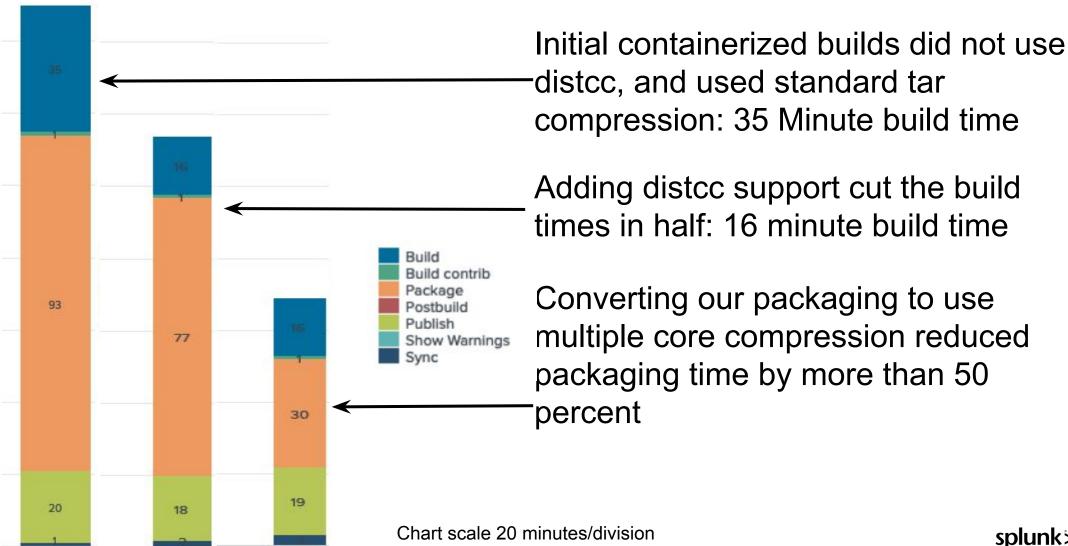
Pull Request Build Improvement with PIGZ



- Pull Request Build build
 - 3 Components
 - 3 archives created
- Multicore archiving program
 - Utilizes multiple cores on the container

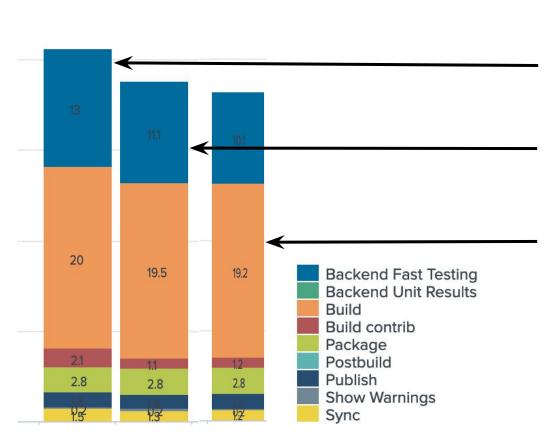
2x time reduction for packaging

Splunk Production Build Time reduced by 50 percent



Splunk Pull Request Time Improvements

Overall average running time reduced by 50 percent



Initial test configuration with 16 tests running in parallel takes 13 minutes

Mid change, 25 tests running in parallel takes 11.1 minutes

Final test configuration has 40 backend unit tests running in parallel, with distcc and pigz performance improvements takes 10.1 minutes, final total running time 36 minutes



Improving Notifications

Our Analysis

Speeding up Developer Notification

Analysis:

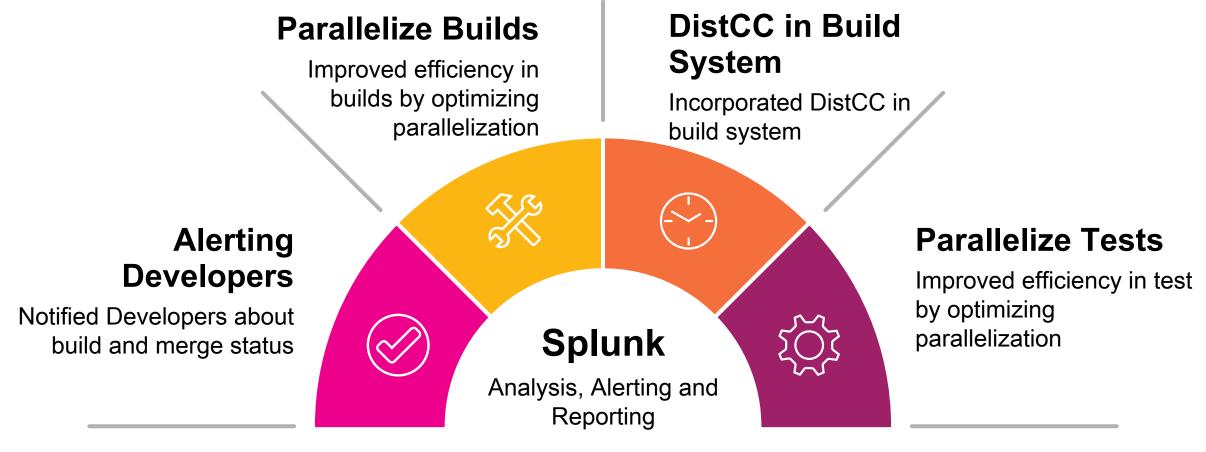
- Test results only available as comments in pull request UI
- Developers need to go to Bitbucket frequently to check for results
- All employees are connected via Slack

Mitigation:

- Added personal Slack messages
 - On receipt of trigger
- When job is completed -> includes test results

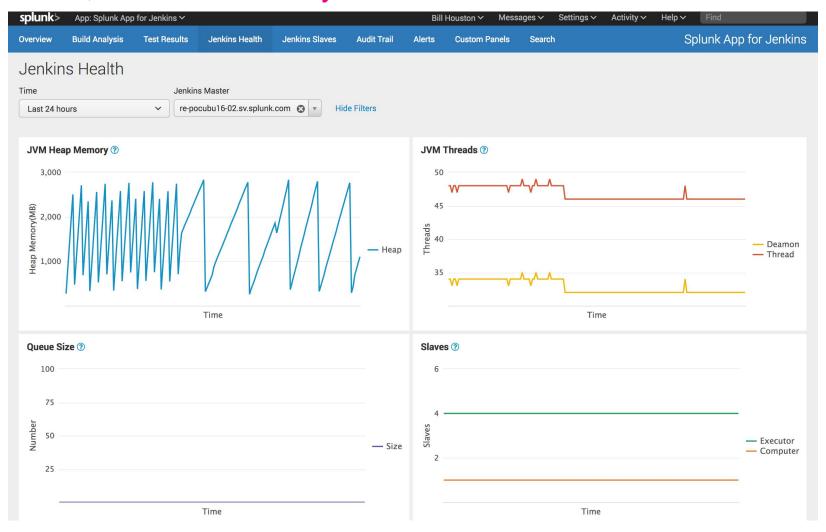
Overall Improvement

Key Takeaways



The Jenkins App for Splunk

Seamlessly collect, monitor and analyze Your Jenkins Data



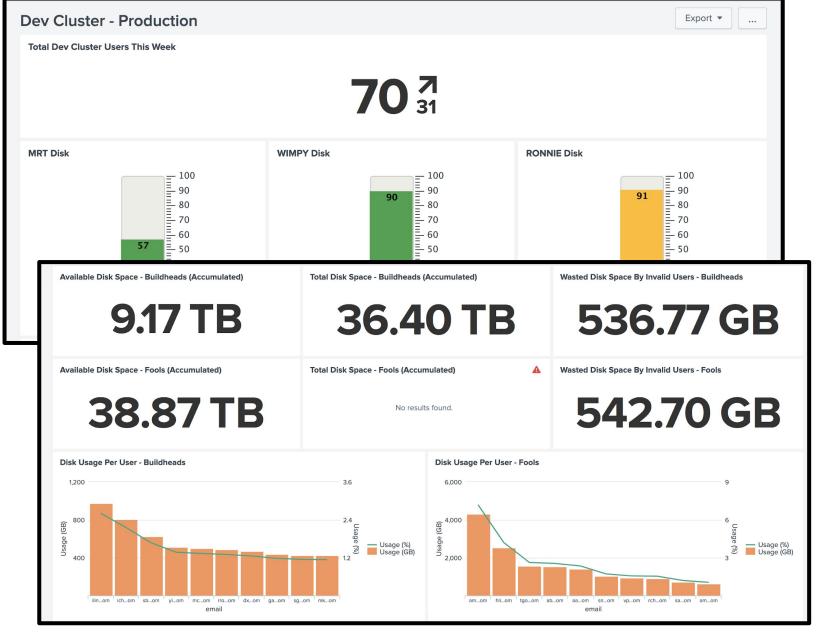


DevOps

DevCluster

☐ Artifactory

☐ Jenkins





DevOps

□ DevCluster

Artifactory

☐ Jenkins





DevOps DevCluster

☐ Artifactory

→ Jenkins

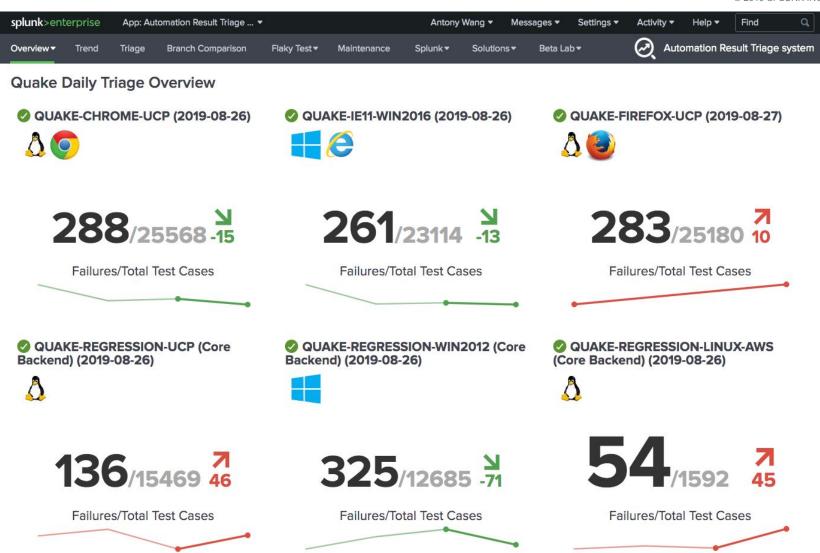




Golden Branch Health

→ Test Triage (ARTs)

Code Coverage



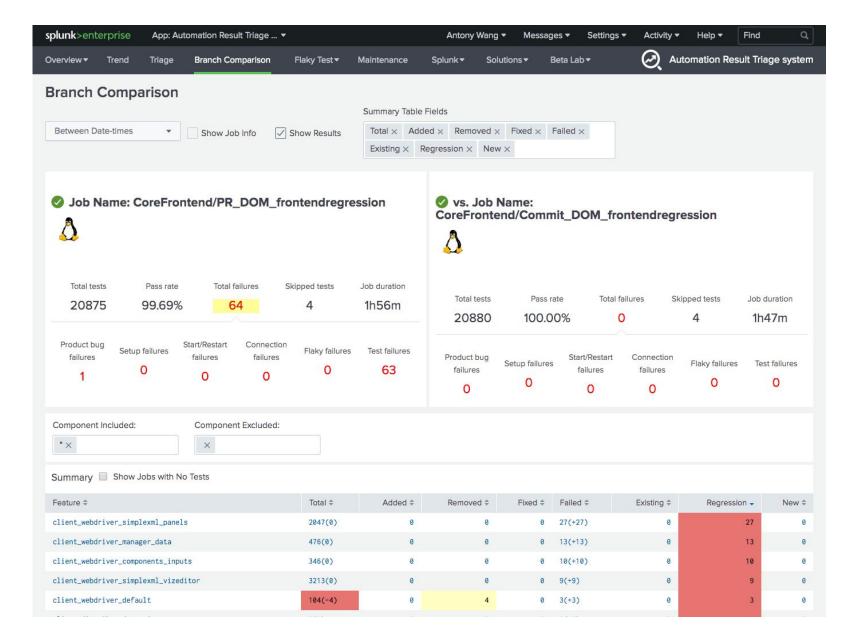


Test, CI & more

Golden Branch Health

Test Triage (ARTs)

Code Coverage



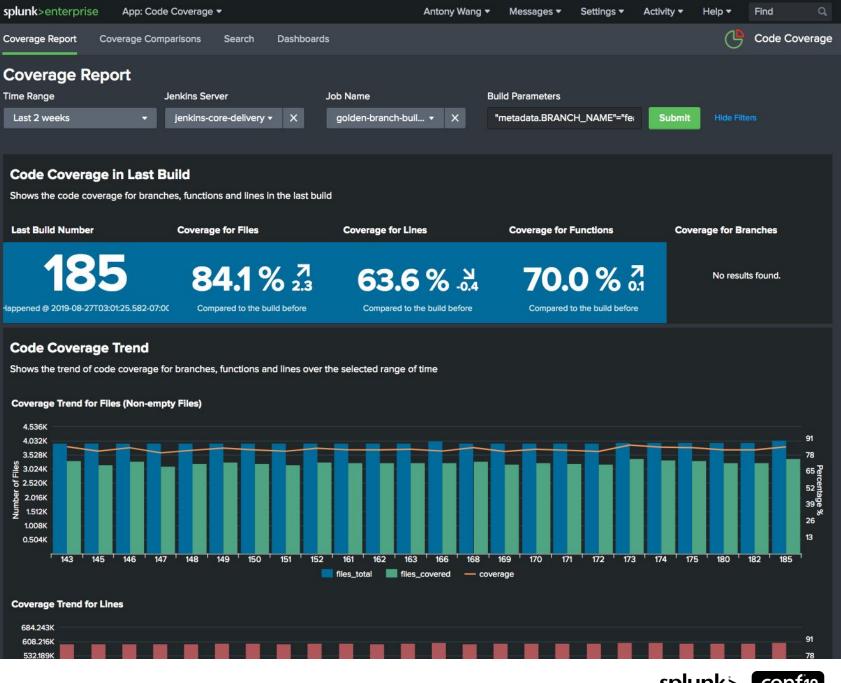


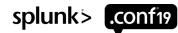
Test, CI & more

Golden Branch Health

Test Triage (ARTs)

Code Coverage



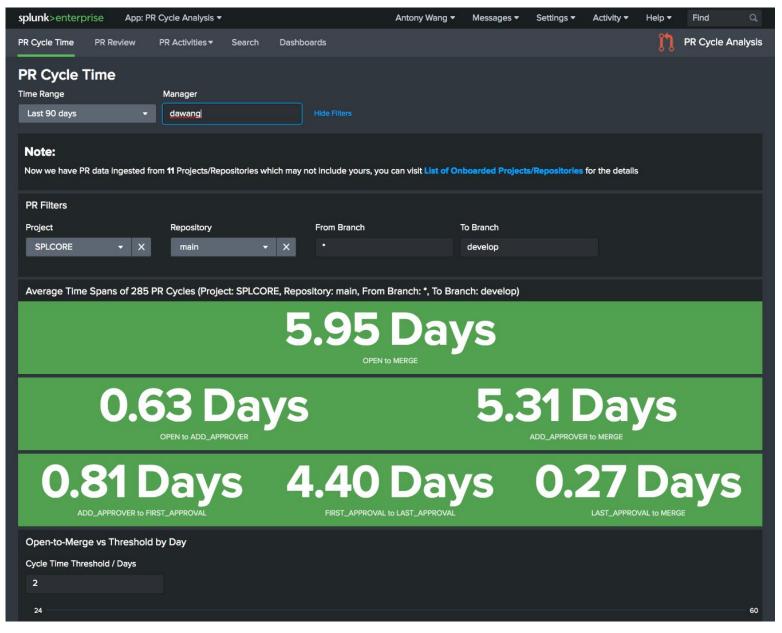


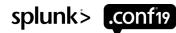
Test, CI & more

Golden Branch Health

Test Triage (ARTs)

Code Coverage





.Conf19
splunk>

Thank

You

Go to the .conf19 mobile app to

RATE THIS SESSION





Q&A