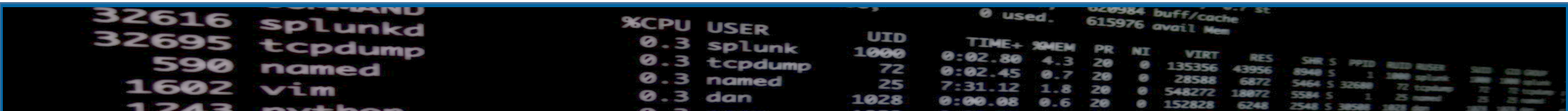
Splunk Index Automation in GitHub Repositories

Presented by Tom Kreiner

to Baltimore Splunk User Group on

August 23, 2021

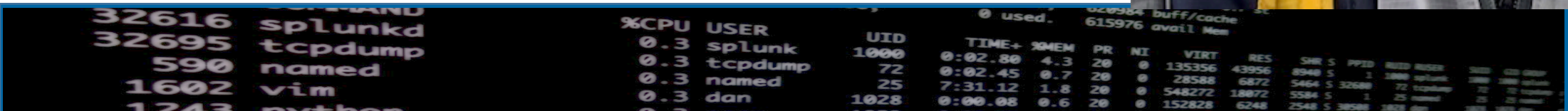


APLURA

Many Solutions, One Goal.

Who Am I?

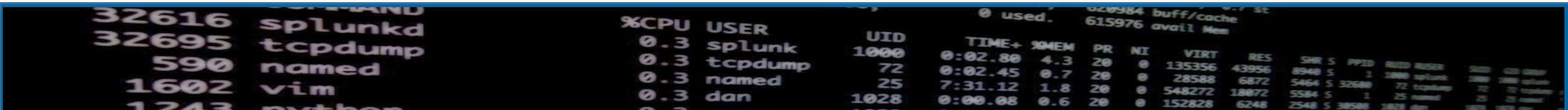
- Splunk professional since 2012
- Splunk Enterprise Certified Architect and Core Certified Consultant
- Joined Aplura in April 2021
- When not Splunking:
  - Husband, and father of two kids
  - Sailing
  - BBQ/grilling/smoking
  - Amateur radio



Many Solutions, One Goal.

Customer Profile

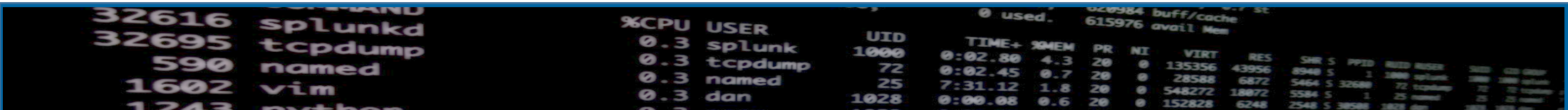
- Customer is a large dev ops focused environment.
- Splunk environment is handling 120+ TB/day.
- Indexer clustering used for storing indexes.
- Developers need to request new indexes as new applications/environments are built.



Many Solutions, One Goal.

Problem

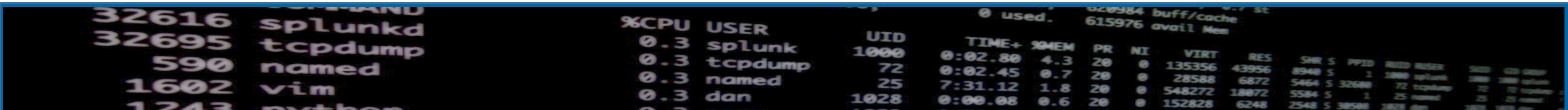
- Customer needed a better way to build documentation about their indexes.
- Documentation needed to include an index description and data owner.
- Users shouldn’t be required to understand Splunk indexes.conf structure.
- All requests will need approval before being deployed to Splunk.



APLURA
Many Solutions, One Goal.

WARNING – Lots of Information Ahead!!!

We are going to cover A LOT of information in the next hour. Use this session as a guide and the public repo as a resource.

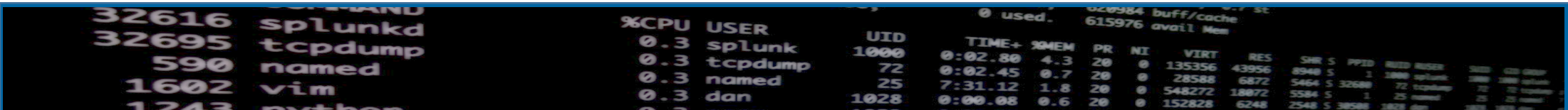


GitHub Repositories and Workflow Actions

Customer was already a user of GitHub to store their configurations. We could use Workflow Actions to automate the process.

https://docs.github.com/en/actions/creating-actions/about-actions

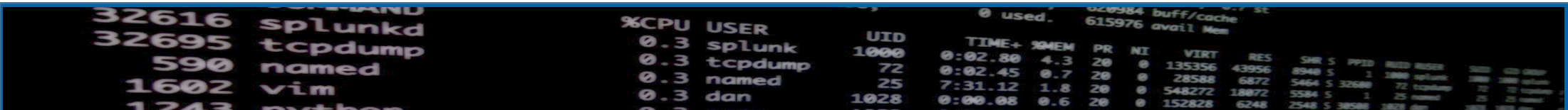
https://docs.github.com/en/actions/using-github-hosted-runners/about-github-hosted-runners



Index YAML Files

We defined a simple YAML file to define our indexes.

```
---
indexes:
- name: testing01
  description: Index for Splunk Admins to test data ingests before ingesting to production.
  owner: Splunk admins
  retention_days: 7
```

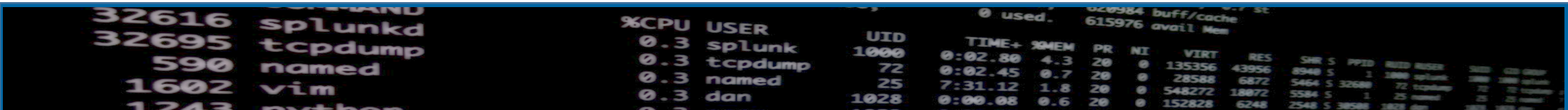


Many Solutions, One Goal.

Desired Output - Splunk Indexes.conf

```
# Description: Metrics index for Splunk admins to store performance data of Splunk
# Owner: Splunk admins
[splunk_metrics]
homePath   = volume:primary/splunk_metrics/db
coldPath   = volume:primary/splunk_metrics/colddb
thawedPath = $SPLUNK_DB/splunk_metrics/thaweddb
datatype   = metric

# Description: Index for Splunk Admins to test data ingests before ingesting to production.
# Owner: Splunk admins
[testing01]
homePath   = volume:primary/testing01/db
coldPath   = volume:primary/testing01/colddb
thawedPath = $SPLUNK_DB/testing01/thaweddb
frozenTimePeriodInSecs = 604800
```



Desired Output - Index Readme

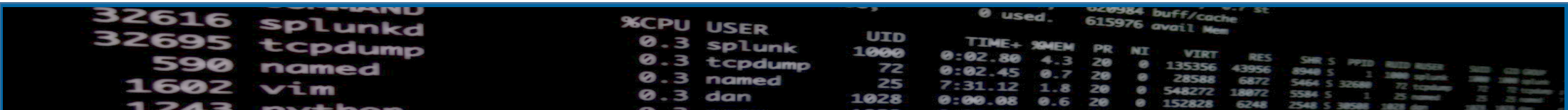
Index Definitions

This document provides a summary of all the indexes defined in Splunk.

Index	Owner	Description	Retention (Days)
splunk_metrics	Splunk admins	Metrics index for Splunk admins to store performance data of Splunk	365*
testing01	Splunk admins	Index for Splunk Admins to test data ingests before ingesting to production.	7

* - *These indexes do not have a specified retention time and are using the system wide setting of 365 retention days.*

This file is auto generated by the build_indexes.py Actions script. To edit any index information, please update the YAML file in the 'index_yaml' directory of the repository through a pull request.



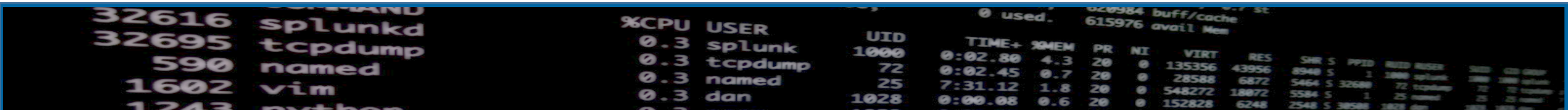
APLURA

Many Solutions, One Goal.

Python Script to Process YAML

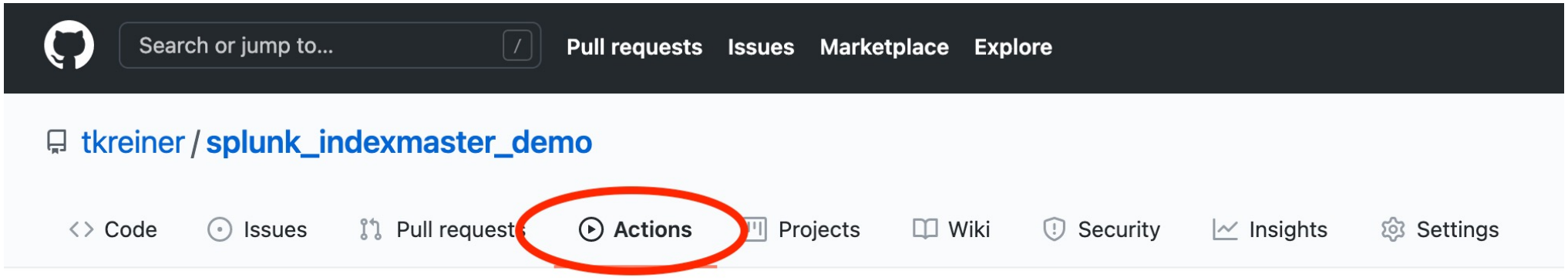
Once the YAML files are defined, we use a Python script to walk those files and build a Splunk indexes.conf and a user document in GitHub markdown.

Script - https://github.com/tkreiner/splunk_indexmaster_demo/blob/main/.github/scripts/build_indexes.py



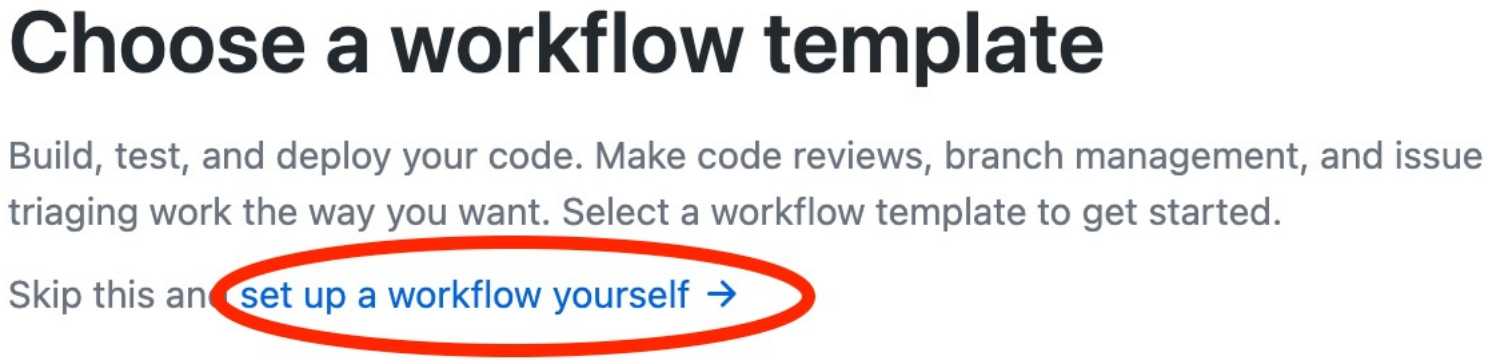
Workflow Action – Bringing It All Together

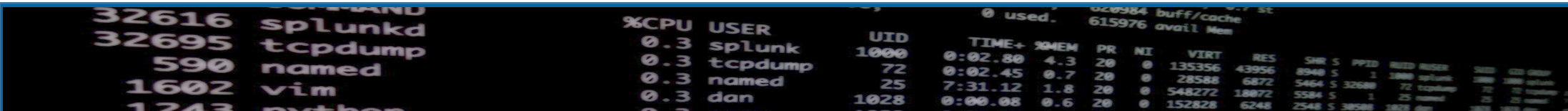
With everything ready, we build a workflow action in GitHub to bring it all together. In GitHub, click on the Actions menu for your repo.



Workflow Action – Bringing It All Together

Select the option to set up a custom workflow yourself.



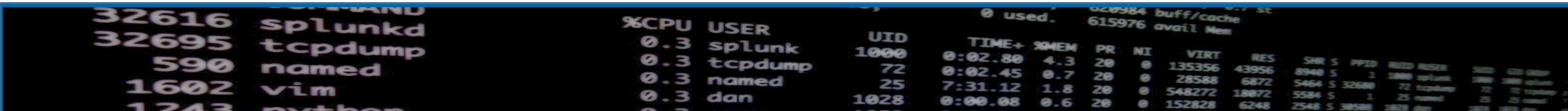


Many Solutions, One Goal.

Workflow Action – Bringing It All Together

Define your workflow action.

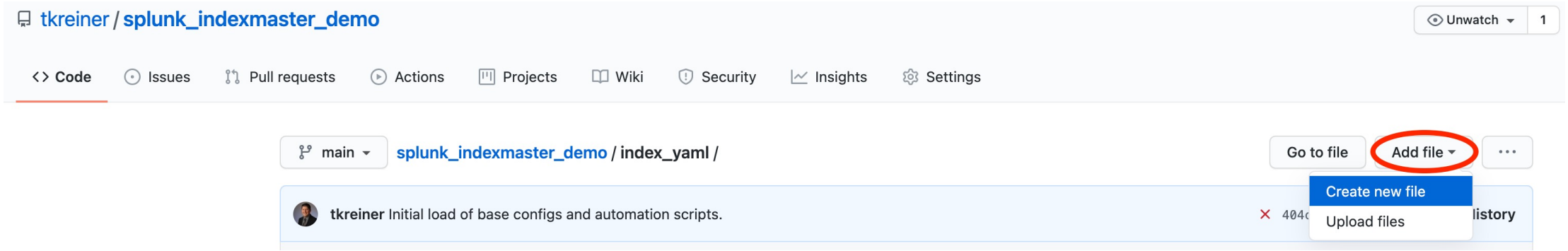
Action definition - https://github.com/tkreiner/splunk_indexmaster_demo/blob/main/.github/workflows/generate_index.yml



Many Solutions, One Goal.

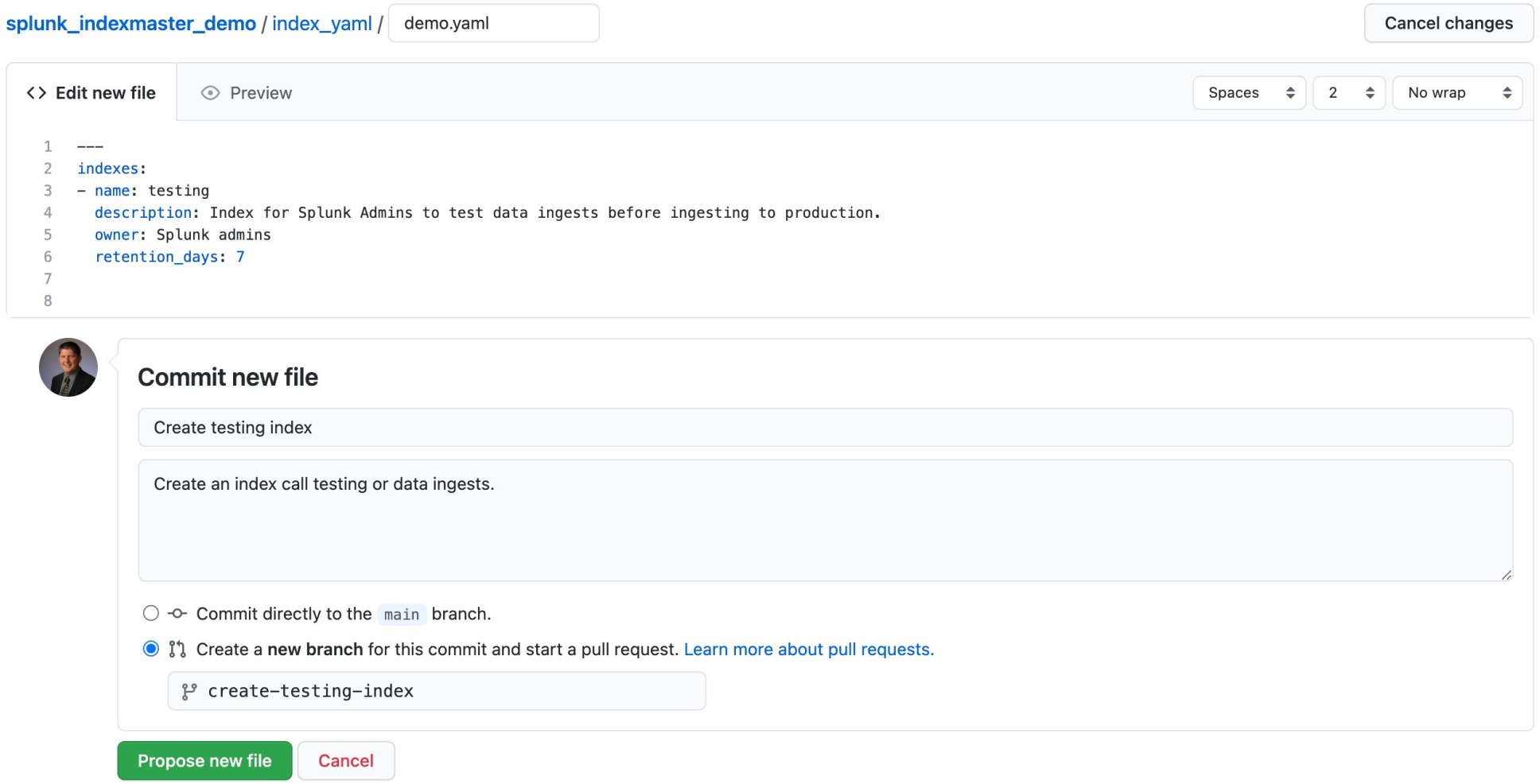
Creating the Index – Define the YAML

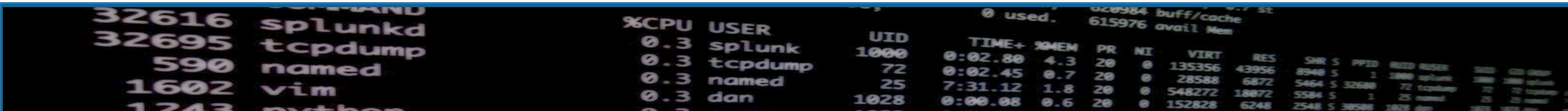
Let’s go into the index_yaml directory and create a new YAML file.



Creating the Index – Submit Branch

We name the file and create the YAML contents. Then provide a branch name and commit description.

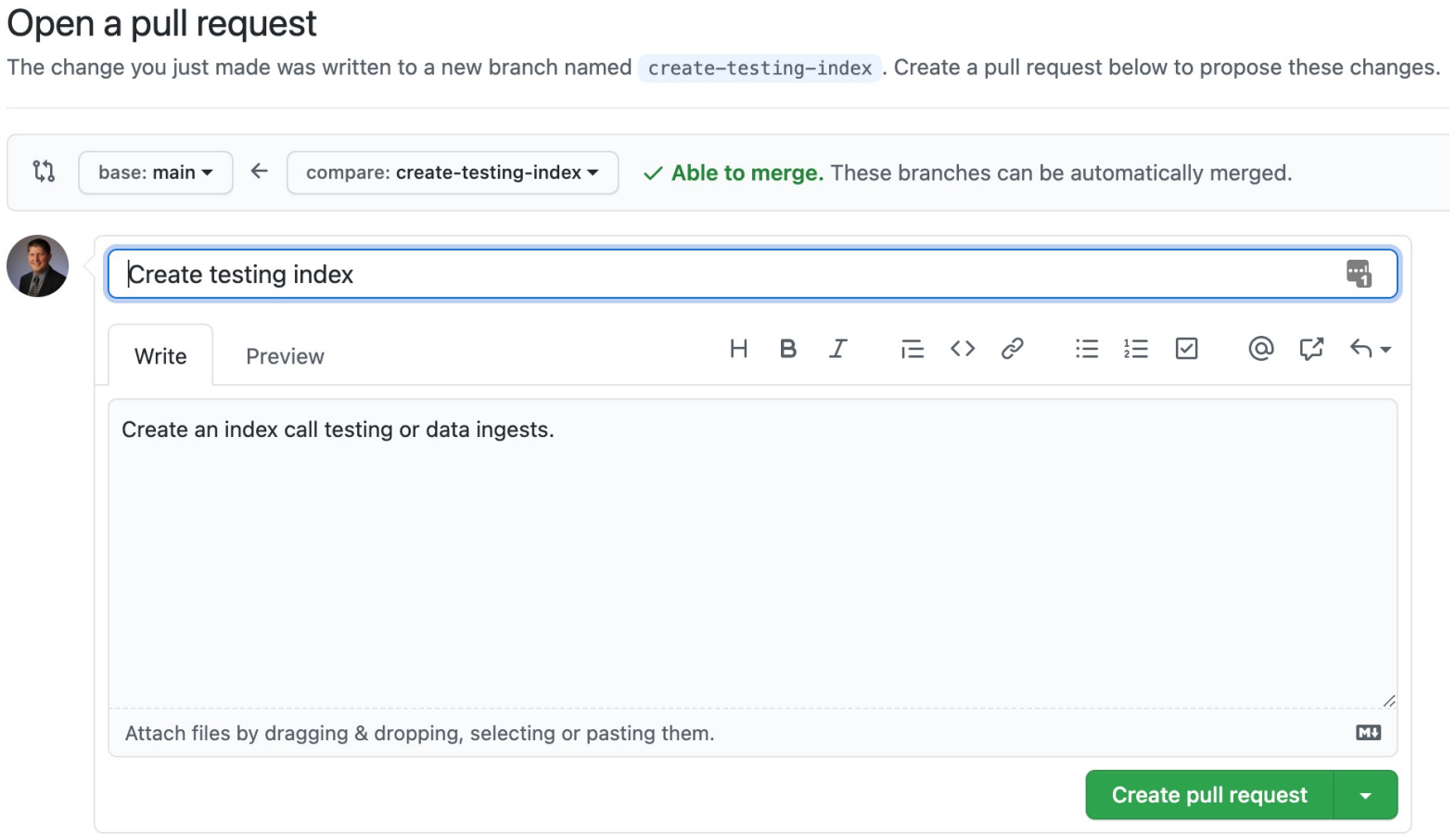


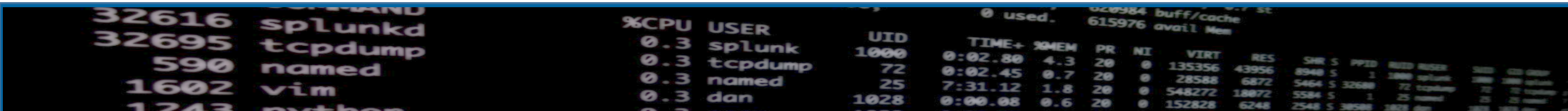


APLURA
Many Solutions, One Goal.

Creating the Index – Submit Pull Request

We submit a pull request to notify the repository administrators that we want to submit a change.

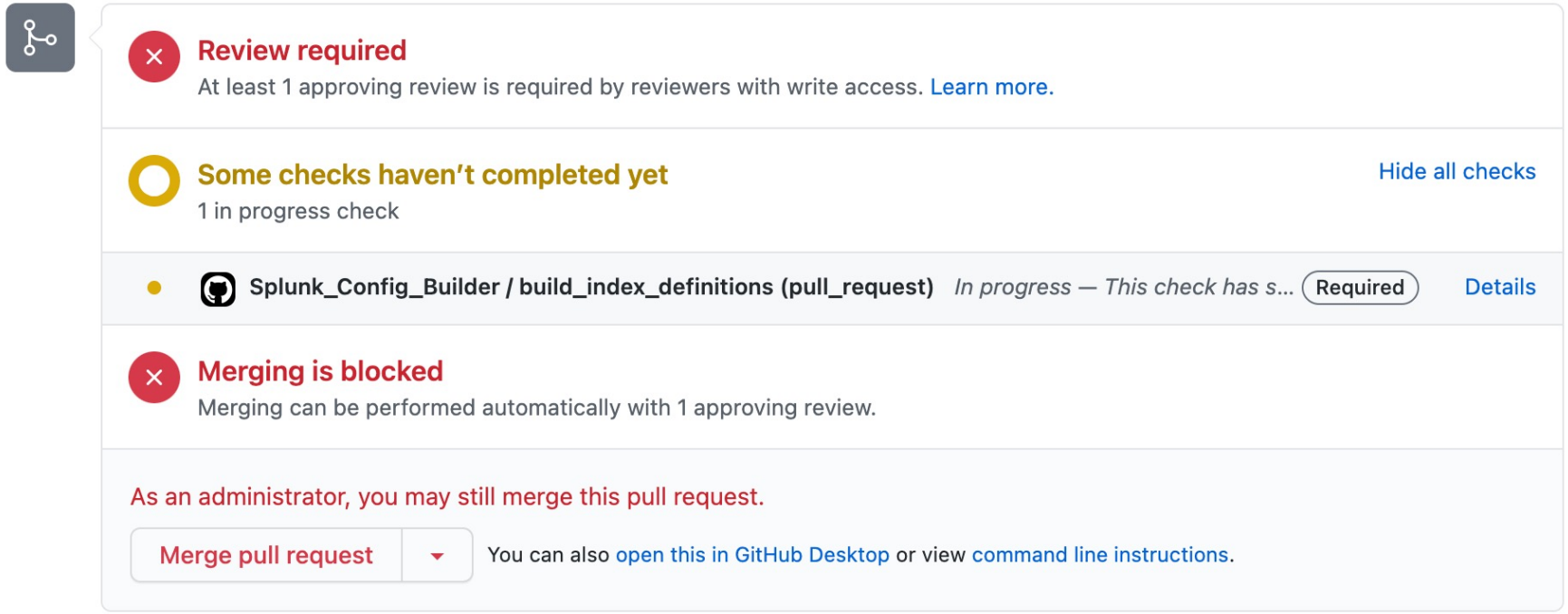


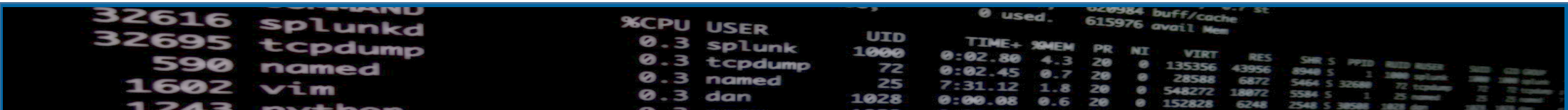


APLURA
Many Solutions, One Goal.

Creating the Index – Action Workflow

From the Pull Request page, we can see the workflow in Action.



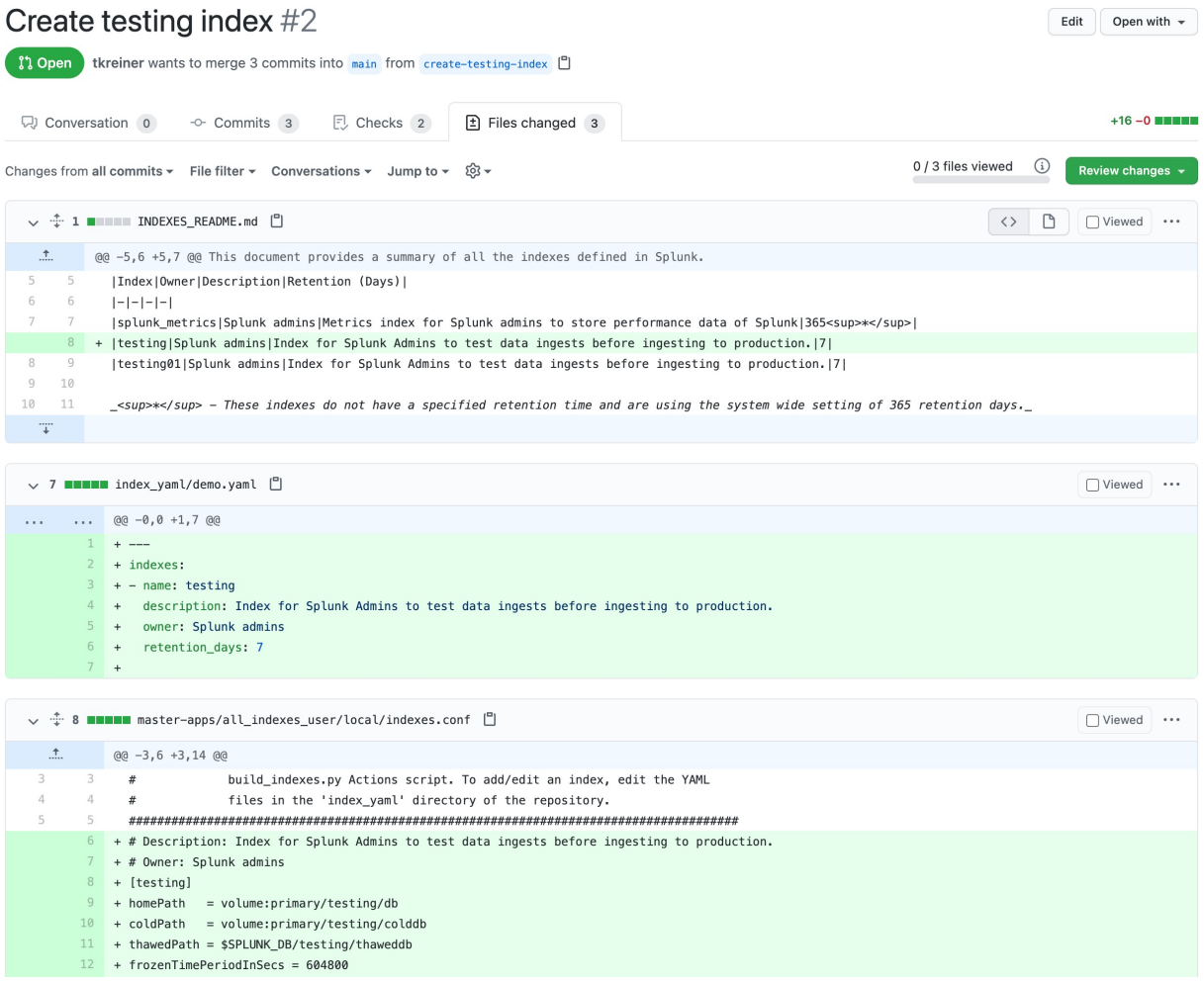


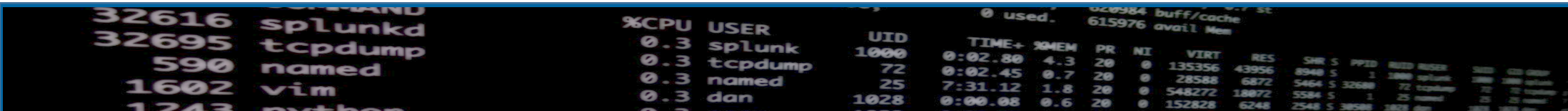
Many Solutions, One Goal.

Creating the Index – Changed Files

When the workflow is complete, we can see that our pull request now has 3 changed files.

- Our new YAML file
- INDEXES_README.md
- indexes.conf

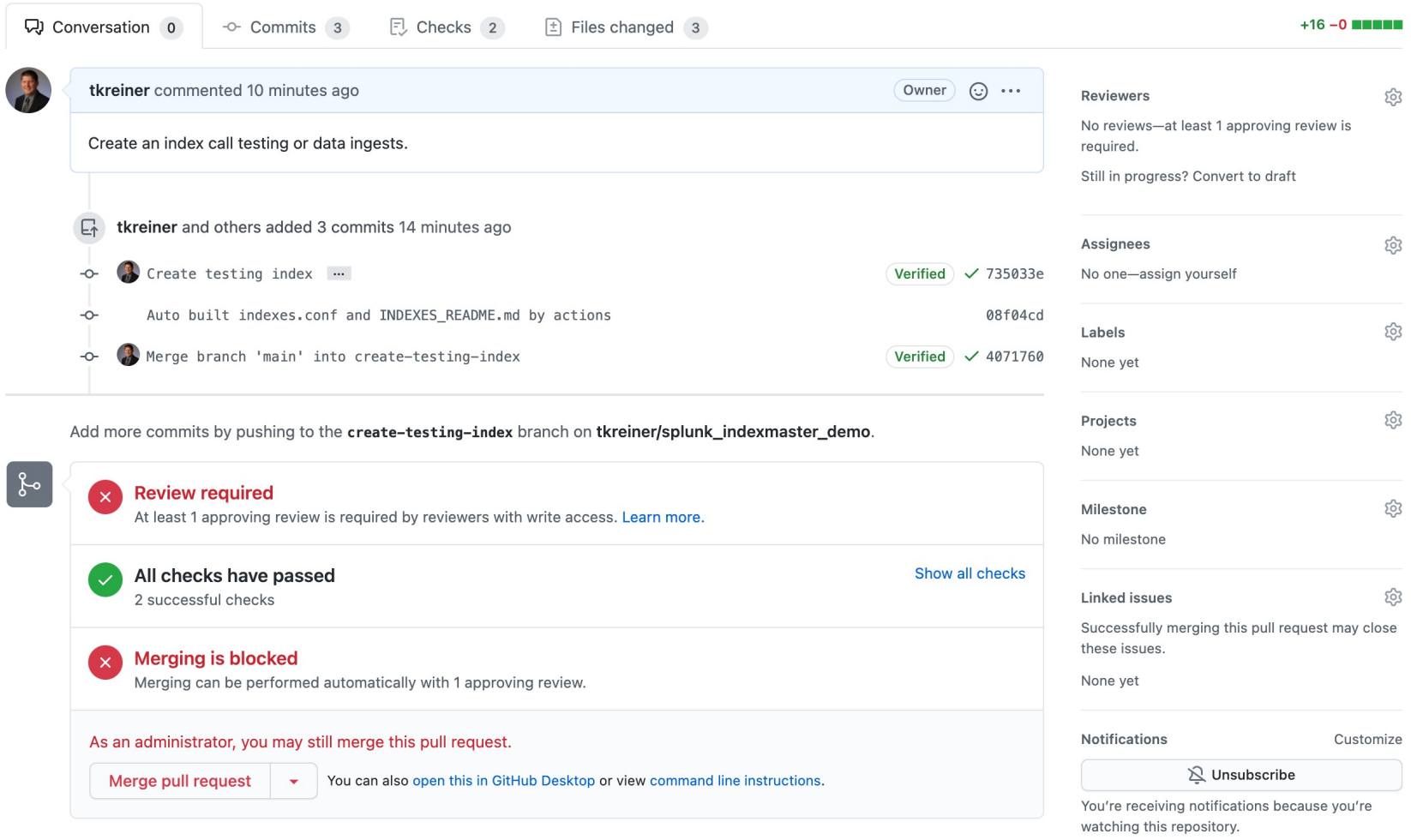


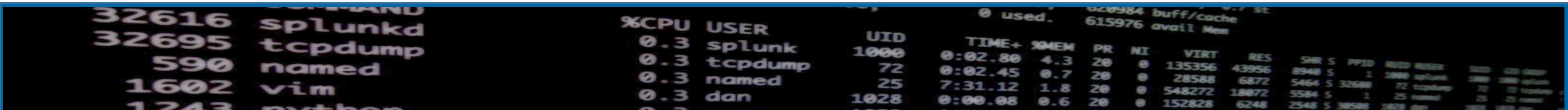


APLURA
Many Solutions, One Goal.

Creating the Index – Ready for Review

Our repo admins can now review our change and integrate it to Splunk.





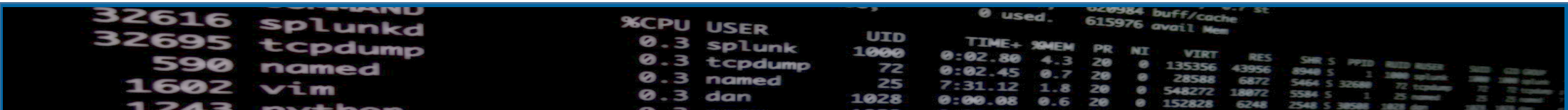
APLURA
Many Solutions, One Goal.

Fun Next Steps

- Customer heavily uses chatops features of Slack. Users can request an index through Slack with a command like the one below. This creates the YAML file and submits a pull request:

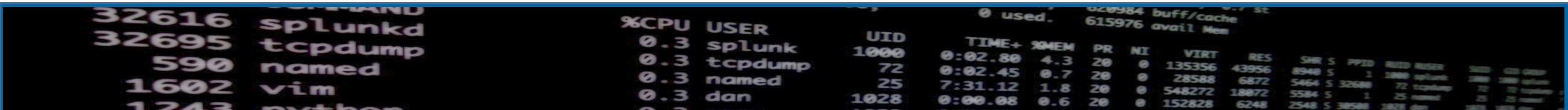
  .splunk createindex name=testing01 owner="Splunk Administrators" description="Index for testing ingest of data." retention_days=7

- Consider building Splunk authorization.conf for your search heads.



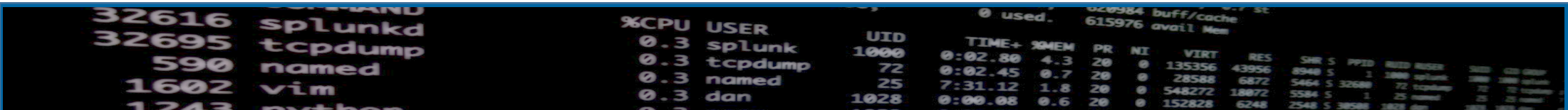
Gotchas

- This does not work with repos that require signed commits. The action that updates the INDEXES_README and indexes.conf does not support signed commits.



Problem – Solved!

- ✓Customer needed a better way to build documentation about their indexes.
- ✓Documentation needed to include an index description and data owner.
- ✓Users shouldn’t be required to understand Splunk indexes.conf structure.
- ✓All requests will need approval before being deployed to Splunk.

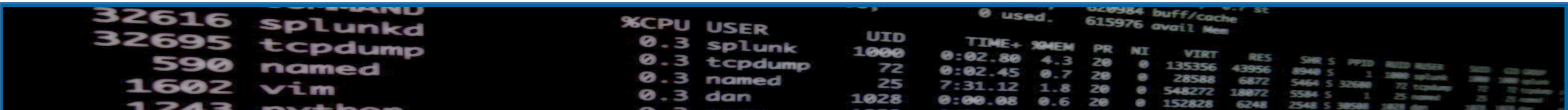


Many Solutions, One Goal.

Capistrano Demo – If Time Allows

Consider the Capistrano project to deploy your repo to Splunk.

https://capistranorb.com/

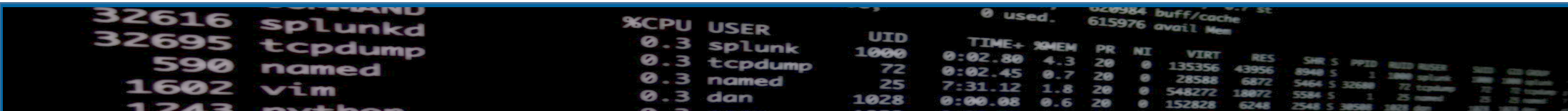


Many Solutions, One Goal.

Demo Repository Available

The repository from today's demo is available at:

https://github.com/tkreiner/splunk_indexmaster_demo



Many Solutions, One Goal.