SQuaRE Upcoming Plans

Frossie Economou, SQuaRE

DMLT Nov 2018

Date long format

State of the Union

- Spurred by SQUASH-specific concerns as well as the need to kick off the work described in DMTN-082 we started a sprint to introduce InfluxDB+chronograf in our toolchain

- We now have a k8s (helm) deployable system (developer sandbox at chronograf-demo.lsst.codes - unstable) receiving lsst.verify payloads via the SQuaSH API (eg ap_verify metrics)

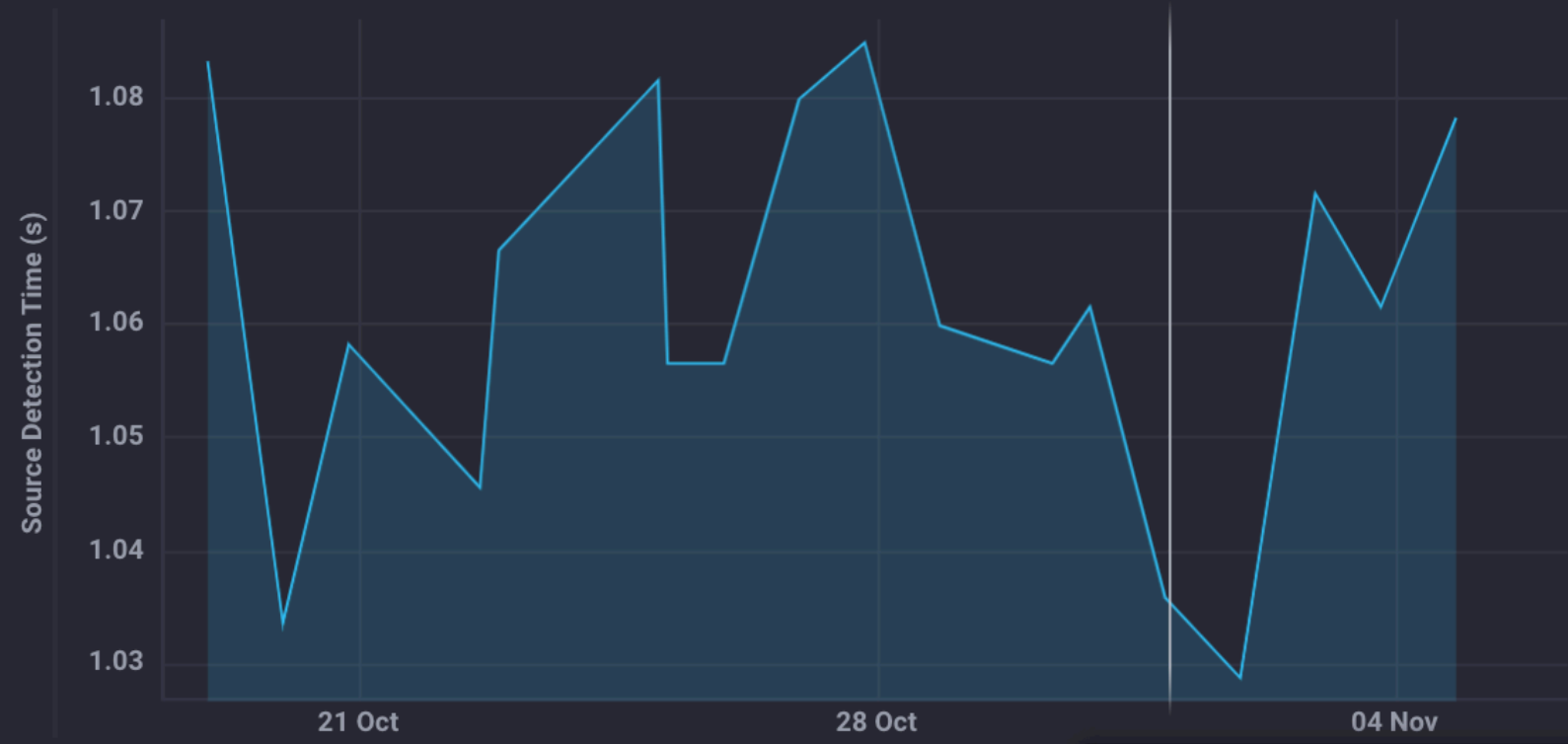- Initially this was going to start in December but was brought forward (project planning + EFD soak test)
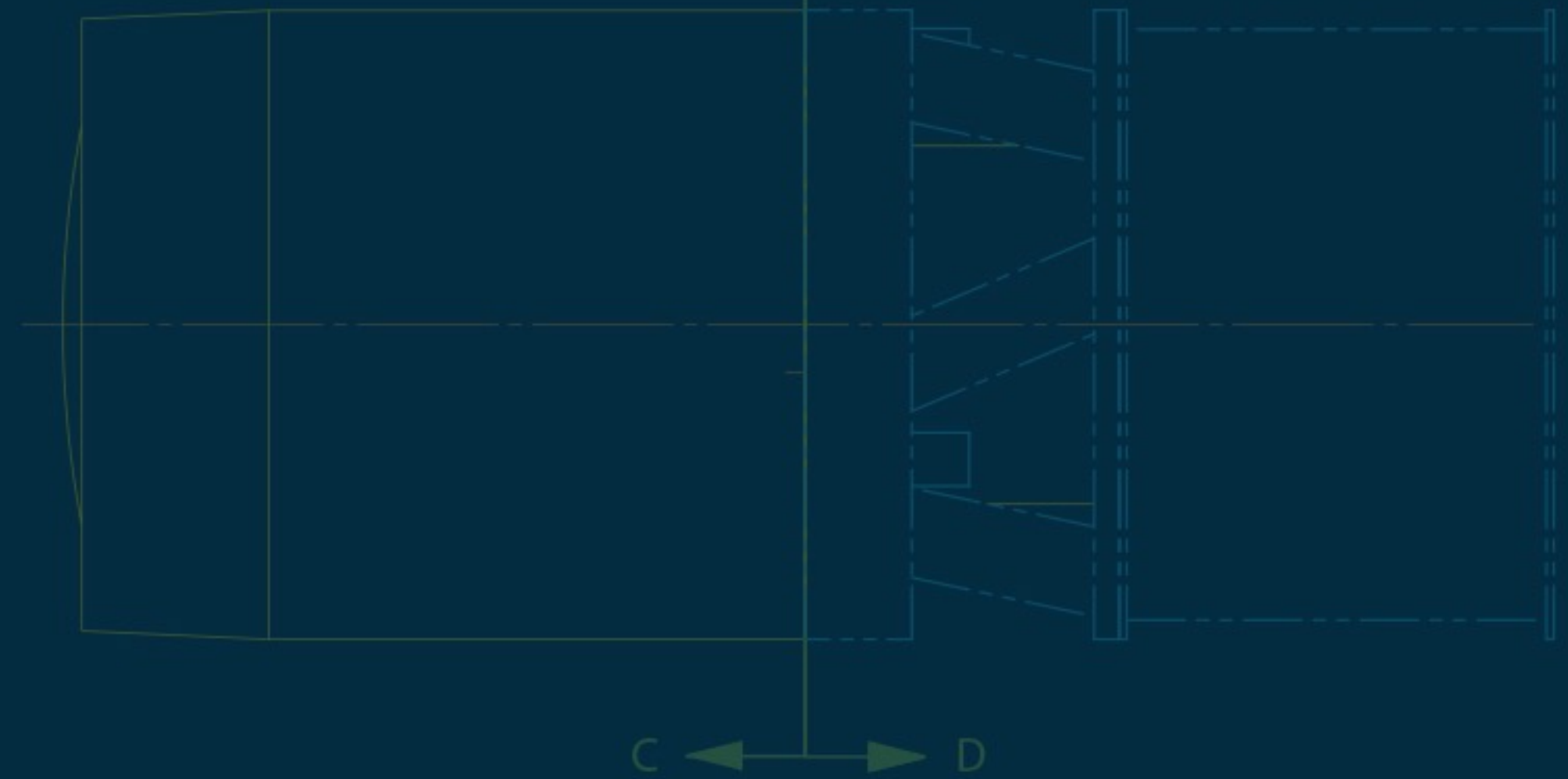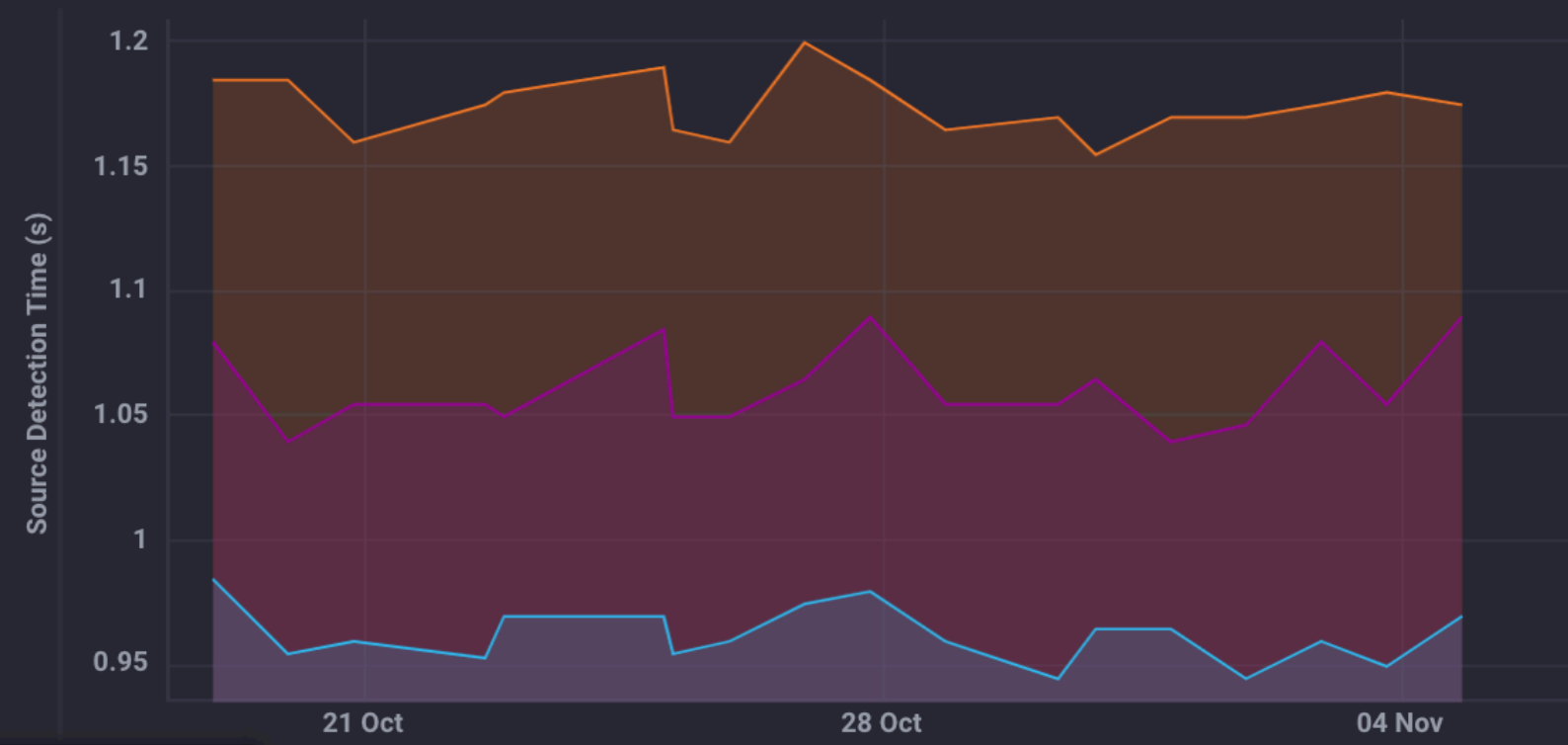
Introduction

Alert Production metrics from ap_verify

Data from the **ap_verify** nightly runs.

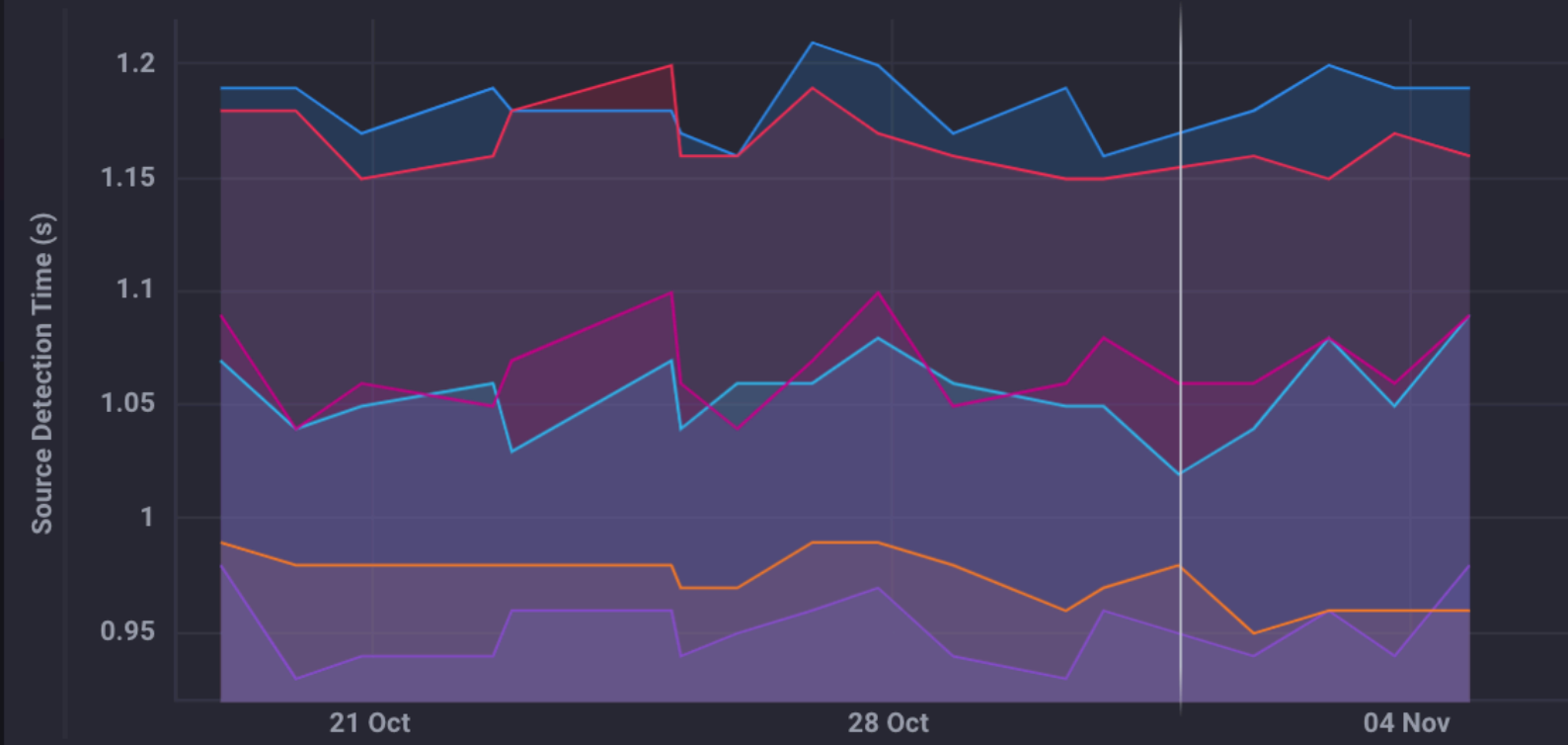Check also the **AP prototype pipeline documentation.**
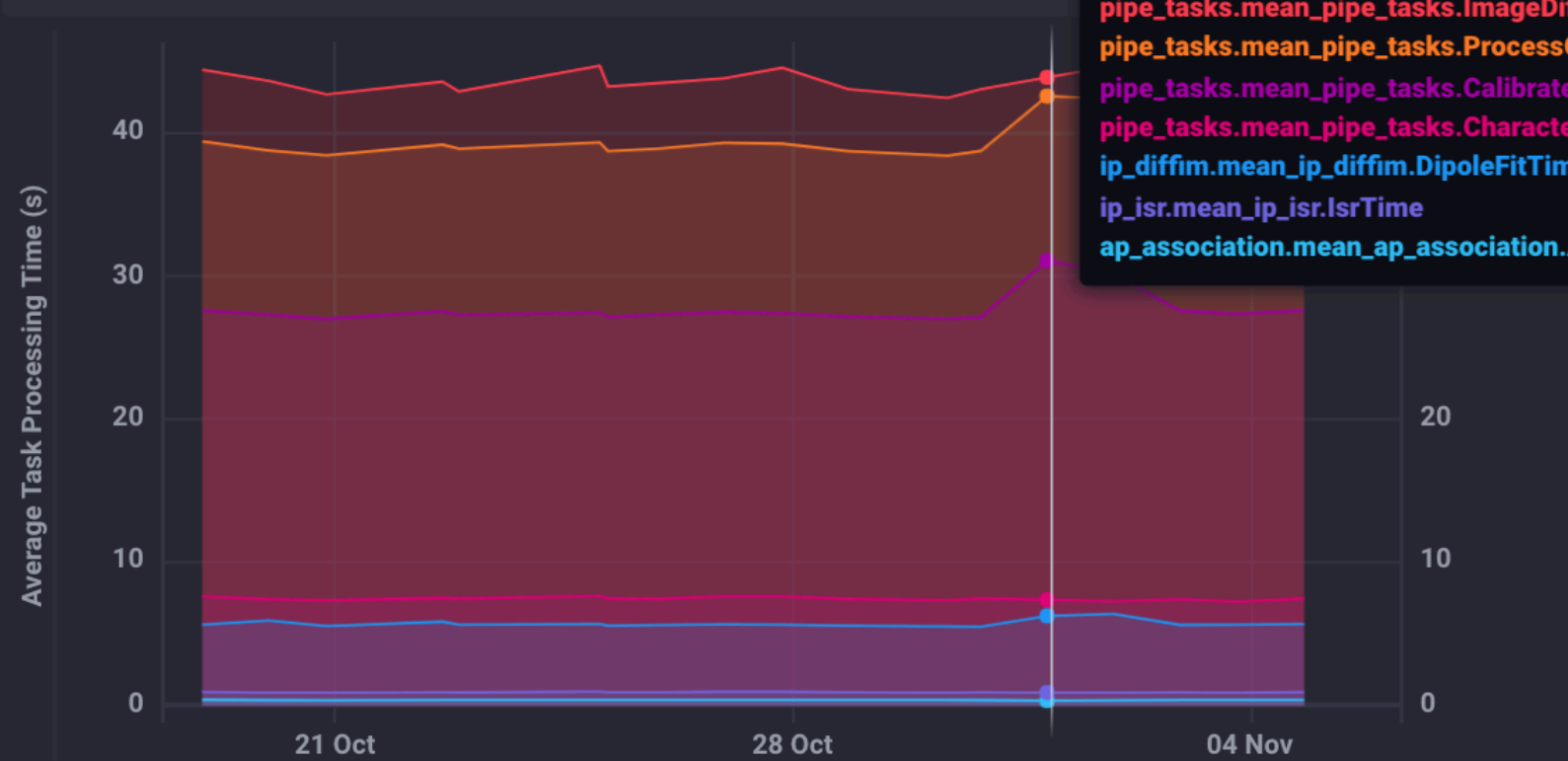
Average Source Detection Time

Average Source Detection Time per Visit

Average Source Detection Time per CCD

Average Task Processing Time

2018/10/31 21:26:27.188

pipe_tasks.mean_pipe_tasks.ImageDifferenceTime	43.92
pipe_tasks.mean_pipe_tasks.ProcessCcdTime	42.61
pipe_tasks.mean_pipe_tasks.CalibrateTime	31.12
pipe_tasks.mean_pipe_tasks.CharacterizeImageTime	7.4
ip_diffim.mean_ip_diffim.DipoleFitTime	6.28
ip_isr.mean_ip_isr.IsrTime	0.92
ap_association.mean_ap_association.AssociationTime	0.35

...urces per CCD
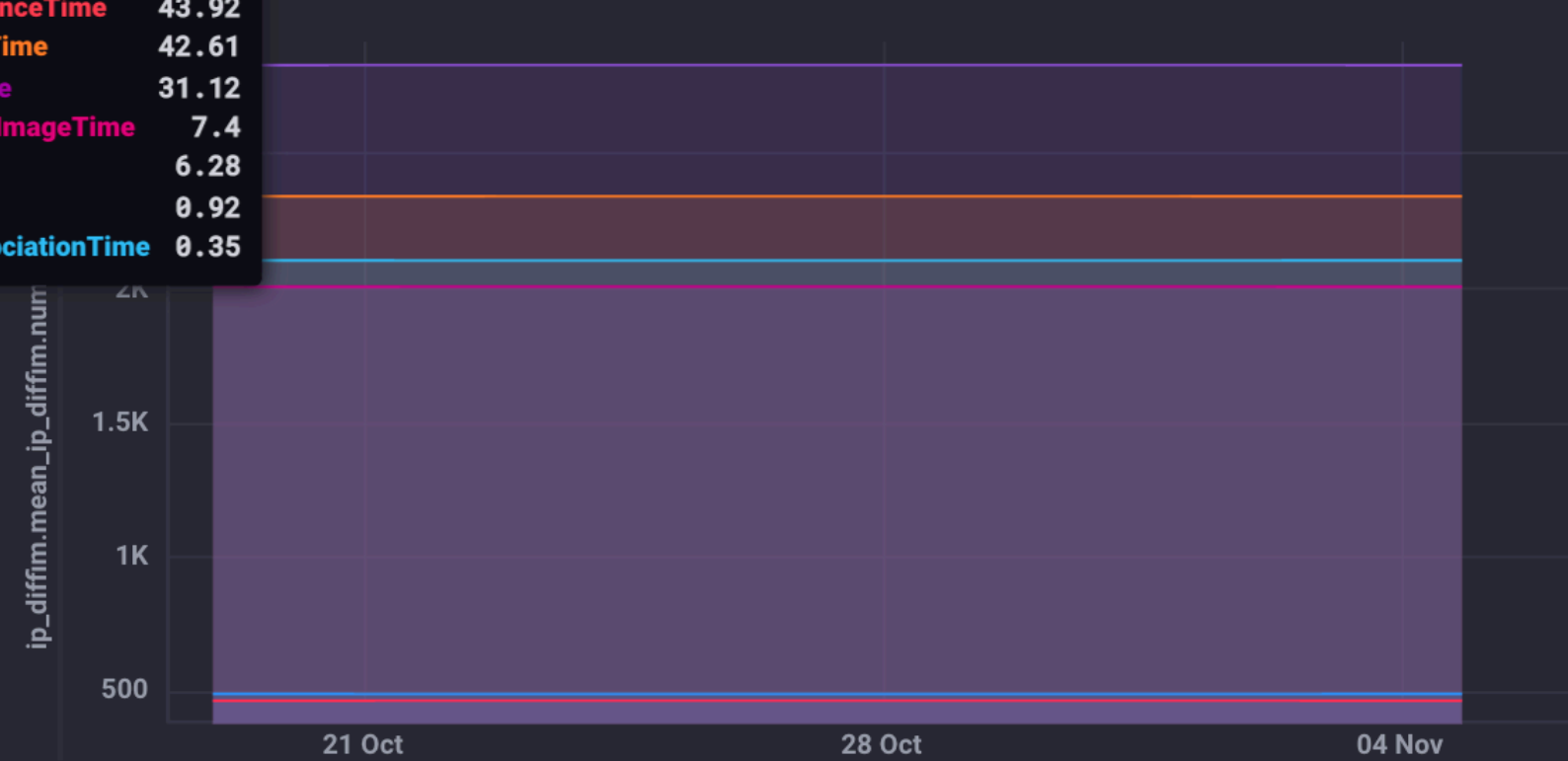
of DIAObjects per CCD
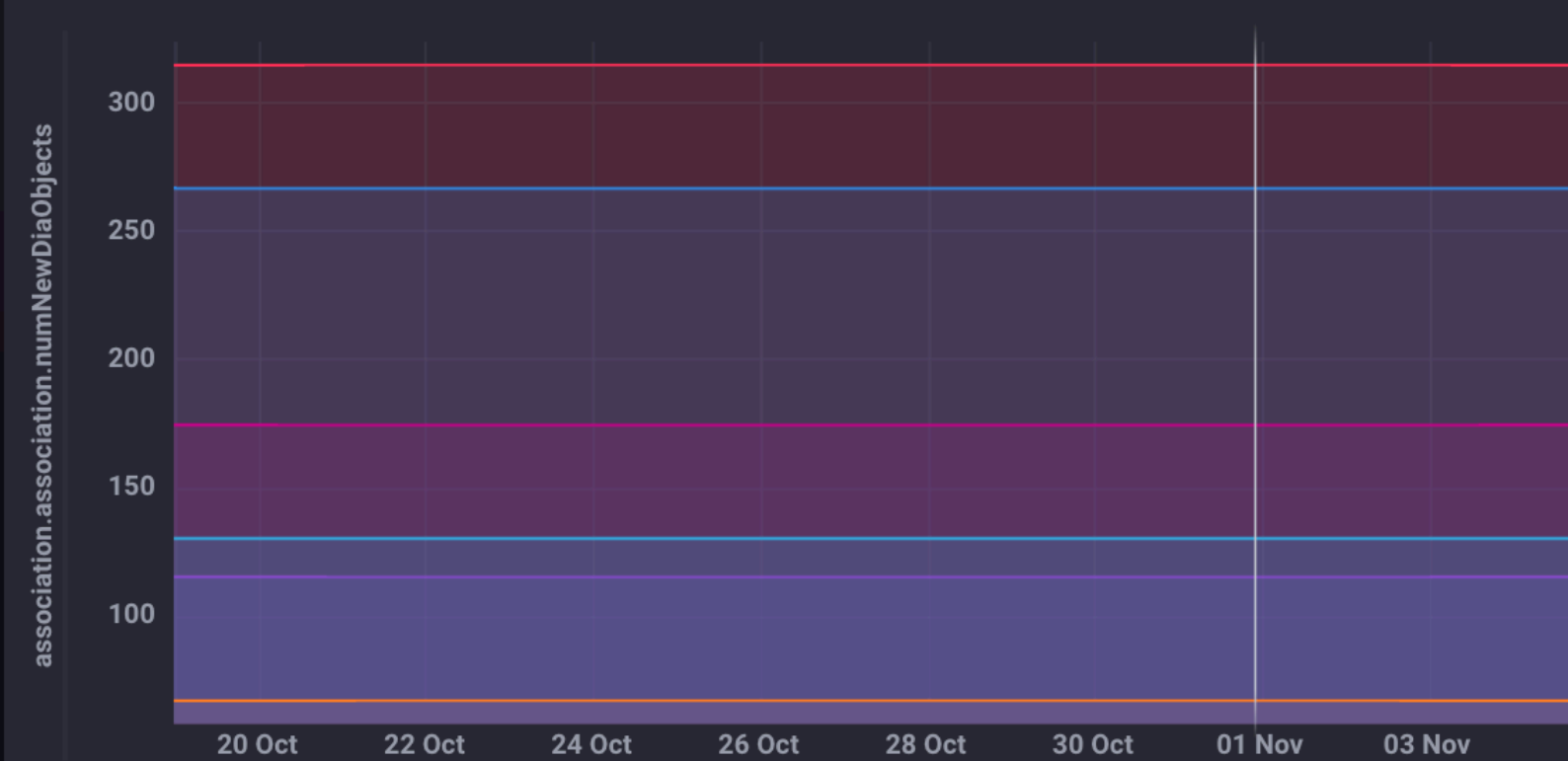
So about S19 (mostly S19A)

- Catch up on the work initially promised in November (sorry John!)

- Switch development focus from LSP back to metrics/QC

- DMTN-082 (what happens after proof of concept sprint?) - January lab work with AuxTel

- Stabilize jellybean and backlog major new features

  - punt production hardening (just did containers-not-as-root)

  - punt resource management issues (eg. dask) and LDF performance issues (file & forget)

  - given descope probably can't punt on firefly scaling / plugins / whatever IPAC needs

  - Vaikunth / AAS Demo

  - will keep up with upstream till it breaks, then might punt

  - not looking for new users (but would like to dogfood TAP service)

- Semantic search (for docs but are planning to re-use for notebooks)

- Leanne is claiming 50% of Simon & he just volunteered for processCCD conversion & bootcamps argh

- LSP review (LDM-542 section for jellybean not fit for purpose)

Support of Operational and Emerging services

- Priorities to support use of SQuaRE services:

  - Commissioning group (aka "Chuck" and "Patrick")

  - Verification & Validation Activities (aka "Leanne")

  - External users (eg science collaborations) (aka "Phil")
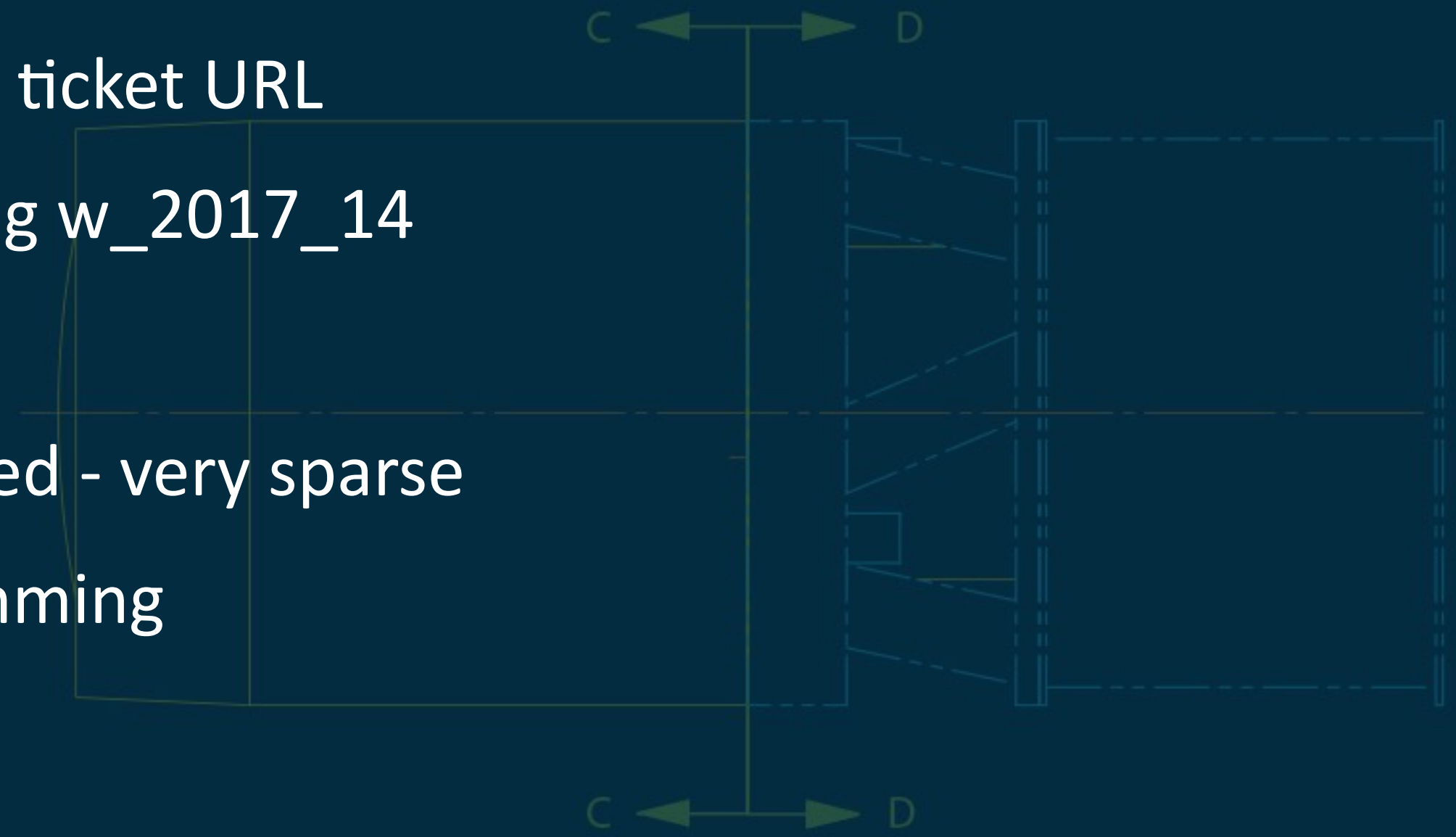
  - Everybody else

CI, and most services (including some of docs) at maintenance levels

The shims are creaking

Current Provenance And Workflow Shims For Metric System

- Provenance

    - Workflow-derived provenance

        - DATASET: name of the database being processed, e.g HSC RC2

        - DATASET_REPO_URL: do we have a git lfs repo for the dataset?

        - RUN_ID : can we use the associated jira ticket to identify this run?

        - RUN_ID_URL: could be the corresponding jira ticket URL

        - VERSION_TAG: the LSST stack version used, e.g w_2017_14

    - Data-derived provenance

        - dataIDs (visit, ccd, filter) of what was processed - very sparse

- Interface to workflow for notebook users not worth shimming

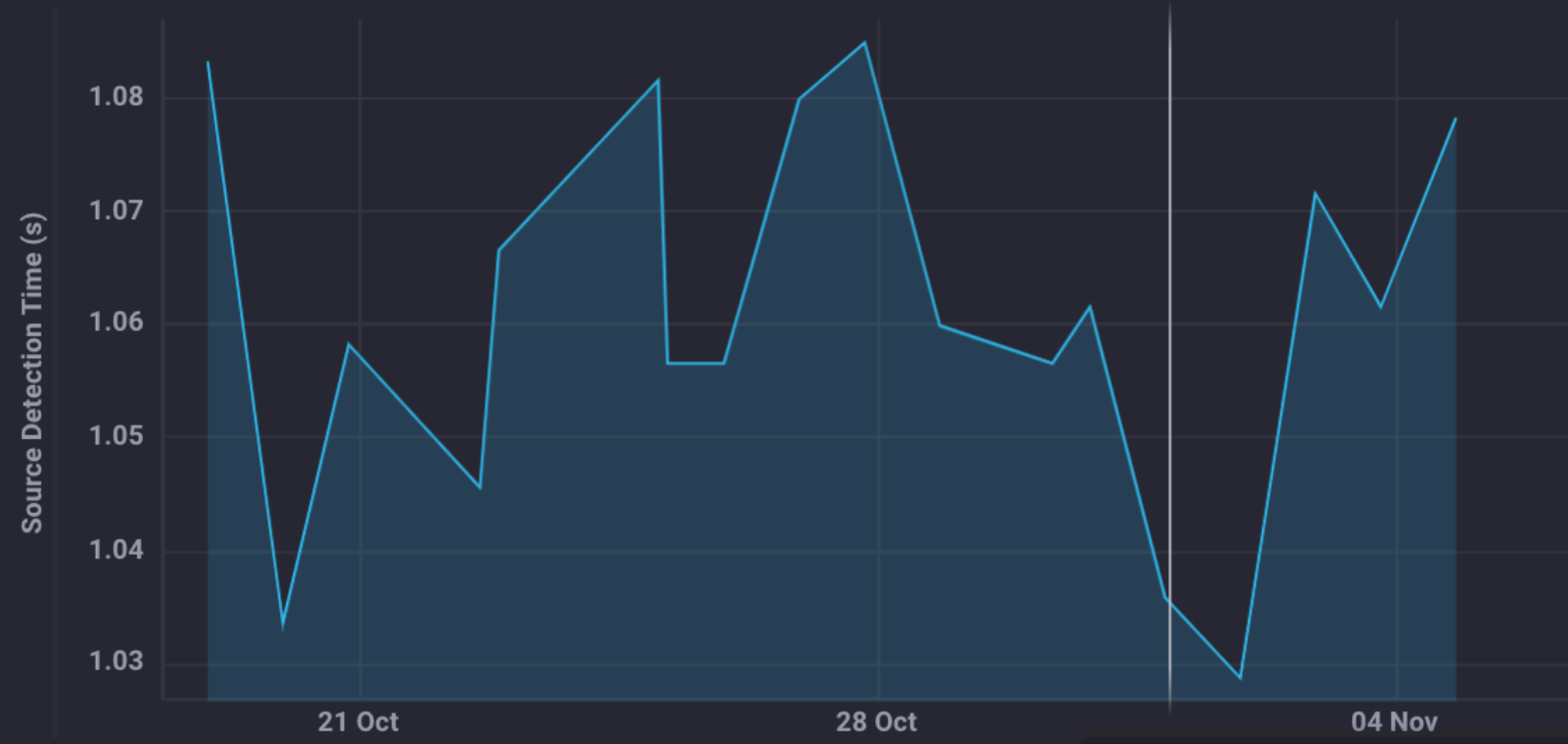- Jenkins no longer a suitable shim for Workflow
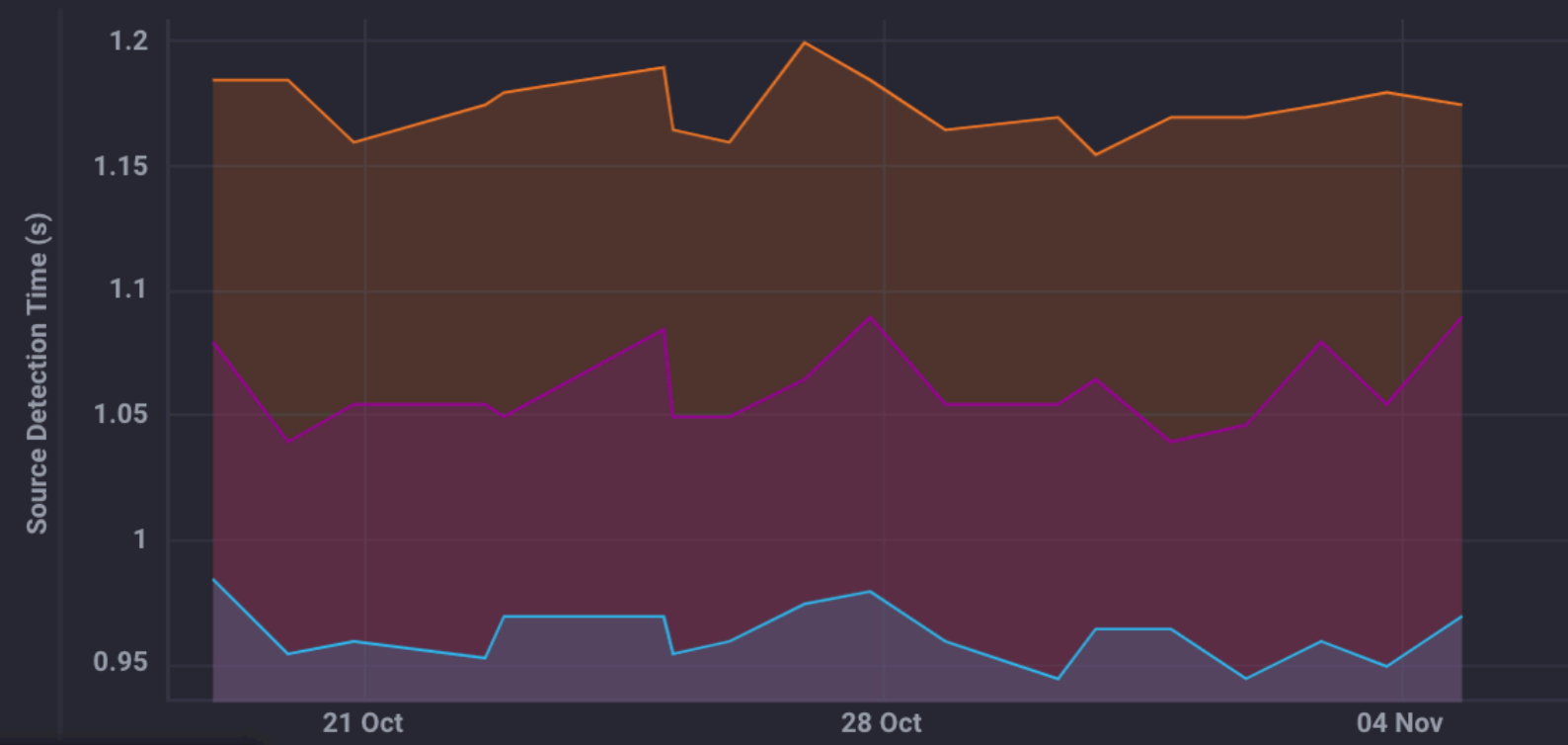
Introduction

Alert Production metrics from ap_verify

Data from the **ap_verify** nightly runs.
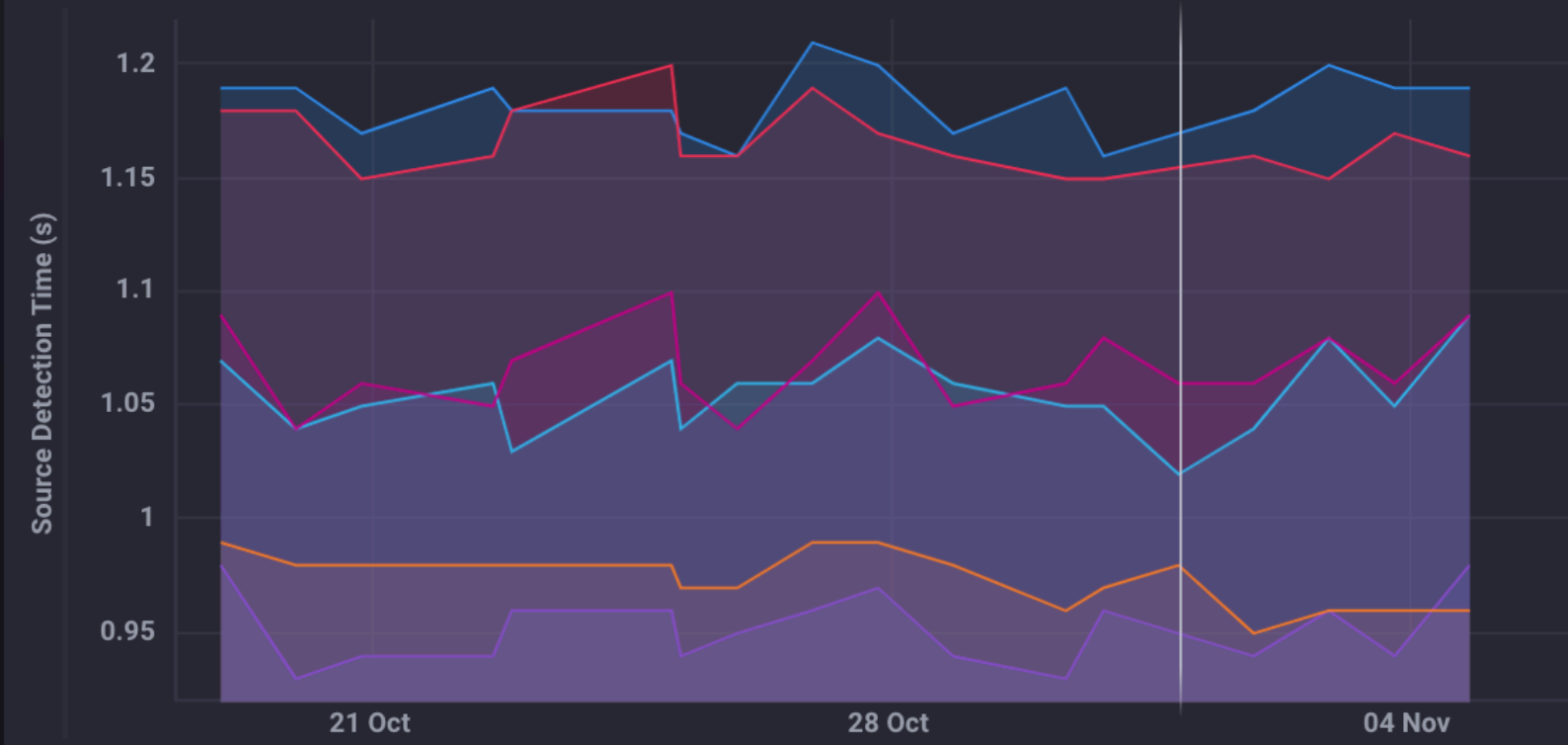Check also the **AP prototype pipeline documentation.**
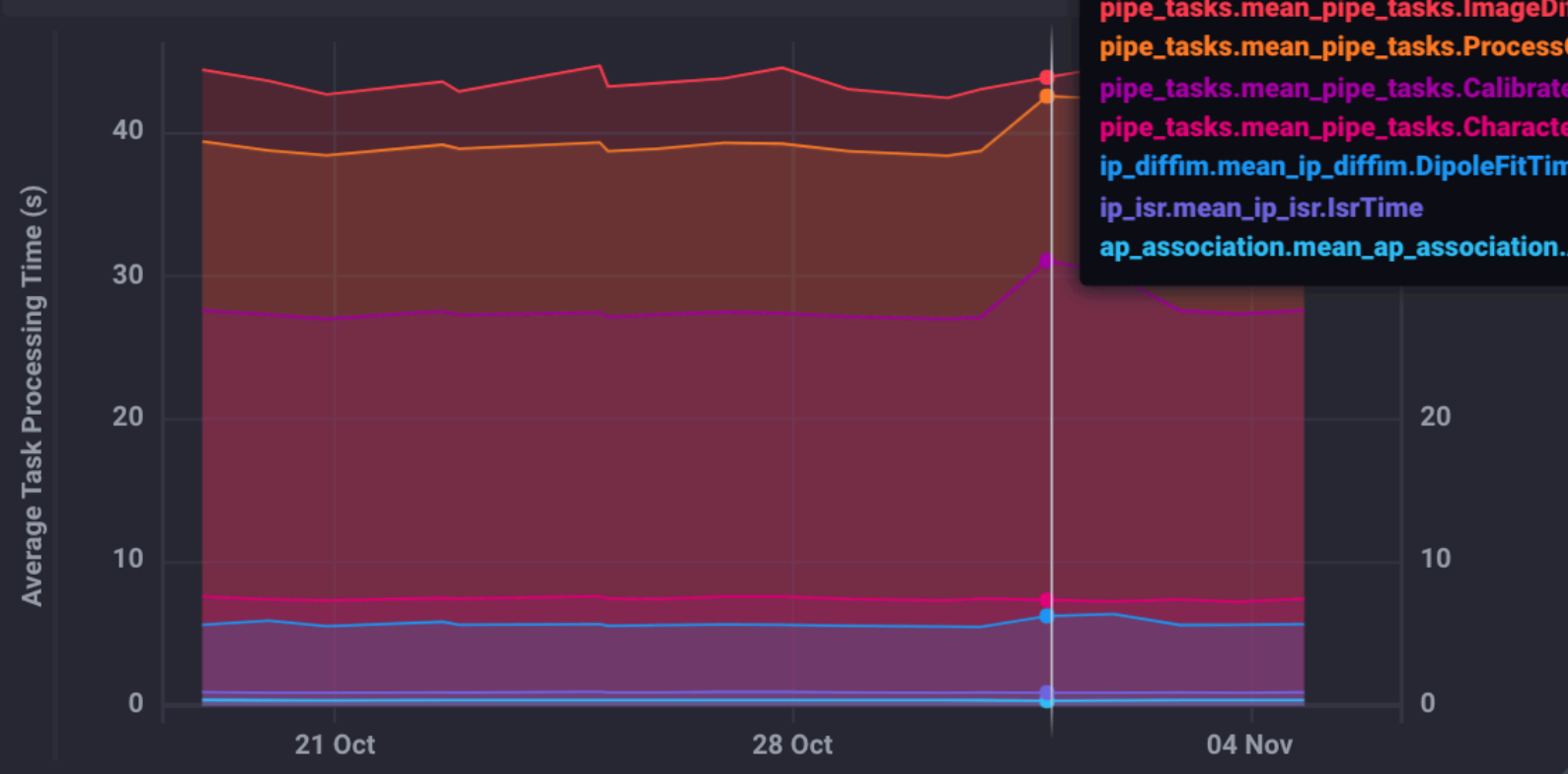
Average Source Detection Time

Source Detection Time (s)

1.08
1.07
1.06
1.05
1.04
1.03

21 Oct   28 Oct   04 Nov

Average Source Detection Time per Visit

Source Detection Time (s)

1.2
1.15
1.1
1.05
1
0.95

21 Oct   28 Oct   04 Nov

Average Source Detection Time per CCD

Source Detection Time (s)

1.2
1.15
1.1
1.05
1
0.95

21 Oct   28 Oct   04 Nov

Average Task Processing Time

Average Task Processing Time (s)

40
30
20
10
0

21 Oct   28 Oct   04 Nov

2018/10/31 21:26:27.188

pipe_tasks.mean_pipe_tasks.ImageDifferenceTime	43.92
pipe_tasks.mean_pipe_tasks.ProcessCcdTime	42.61
pipe_tasks.mean_pipe_tasks.CalibrateTime	31.12
pipe_tasks.mean_pipe_tasks.CharacterizeImageTime	7.4
ip_diffim.mean_ip_diffim.DipoleFitTime	6.28
ip_isr.mean_ip_isr.IsrTime	0.92
ap_association.mean_ap_association.AssociationTime	0.35

...urces per CCD

ip_diffim.mean_ip_diffim.num

2K
1.5K
1K
500

21 Oct   28 Oct   04 Nov

# of DIAObjects per CCD

association.association.numNewDiaObjects

300
250
200
150
100

20 Oct 22 Oct 24 Oct 26 Oct 28 Oct 30 Oct 01 Nov 03 Nov

Average Task Processing Time

Queries | Visualization

Dynamic Source ▾

Flux | *InfluxQL* | ?

⬇ CSV | ❚❚ ▾ | 🕐 2018-10-18 10:21 – 2018-11-06 10:21 ▾

40

20

Average Task Processing Time (s)

0

19 Oct 20 Oct 21 Oct 22 Oct 23 Oct 24 Oct 25 Oct 26 Oct 27 Oct 28 Oct 29 Oct 30 Oct 31 Oct 01 Nov 02 Nov 03 Nov 04 Nov 05 Nov 06 Nov

40

20

0
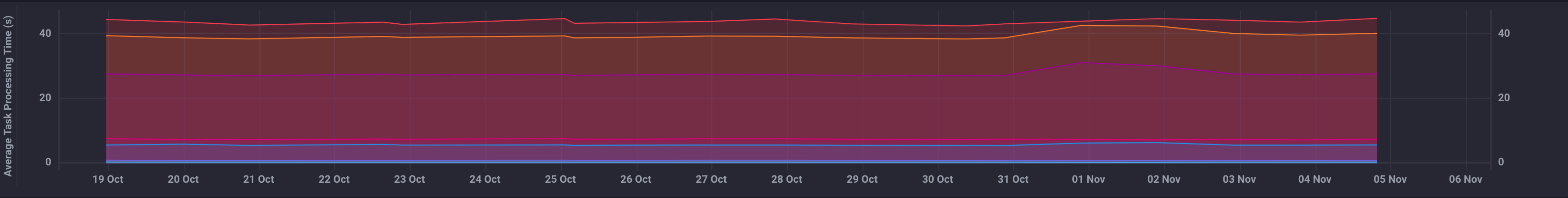
SELECT mean("pipe_tasks.ProcessCcdTime") AS ... ✕ | SELECT mean("ip_isr.IsrTime") AS "mean_ip_isr.Is... ✕ | SELECT mean("ip_diffim.DipoleFitTime") AS "mea... ✕ | SELECT mean("ap_association.AssociationTime")... ✕ | +

```
SELECT mean("pipe_tasks.ProcessCcdTime") AS "mean_pipe_tasks.ProcessCcdTime", mean("pipe_tasks.ImageDifferenceTime") AS "mean_pipe_tasks.ImageDifferenceTime", mean("pipe_tasks.CharacterizeImageTime") AS "mean_pipe_tasks.CharacterizeImageTime", mean("pipe_tasks.CalibrateTime") AS "mean_pipe_tasks.CalibrateTime" FROM "squash-demo"."autogen"."pipe_tasks" WHERE time > :dashboardTime: GROUP BY time(:interval:) FILL(null)
```

⚠ Unsubmitted

Show Template Values | Metaquery Templates ▾ | Submit Query

| DB.RetentionPolicy | Measurements & Tags | 🔍 Filter | Fields | Group by: auto ▾ | Compare: none ▾ | Fill: null ▾ |

DB.RetentionPolicy

_internal.monitor

chronograf.autogen

squash-demo.autogen

squash-prod.autogen

squash-sandbox.autogen

telegraf.autogen

Measurements & Tags

▶ association

▶ ip_diffim

▶ ip_isr

▶ meas_algorithms

▼ **pipe_tasks**

  ▶ ap_association.AssociationTime.estimator — 1

  ▶ ccdnum — 4

  ▶ ci_dataset — 1

Fields

● pipe_tasks.CalibrateTime   `1 Function`

● pipe_tasks.CharacterizeImageTime   `1 Function`

● pipe_tasks.ImageDifferenceTime   `1 Function`

● pipe_tasks.ProcessCcdTime   `1 Function`

Chronograf niceties

- - Display multiple time series plots

- - Use the query builder to define the time series we want to display

- - Combine more than one time series in a single plot (e.g. querying different databases)

- - Make interactive plots using template variables (e.g. select dataset)

- - Export/import dashboards to/from JSON functionality

- - Keep a copy of the Dashboard on GitHub for version control

- - Sharing URL

- - Presentation mode if you add ?presentation=true to the URL

- - Markdown/HTML content in cells (available in Chronograf 1.7)