From Request to ROI: Tracking your Company's Investment in Open Source & Compliance

Krista Khare
David Grizzanti

Open FinTech Forum
October 11, 2018

Comcast Open Source Program Office

Opened beginning of 2017

4 person team

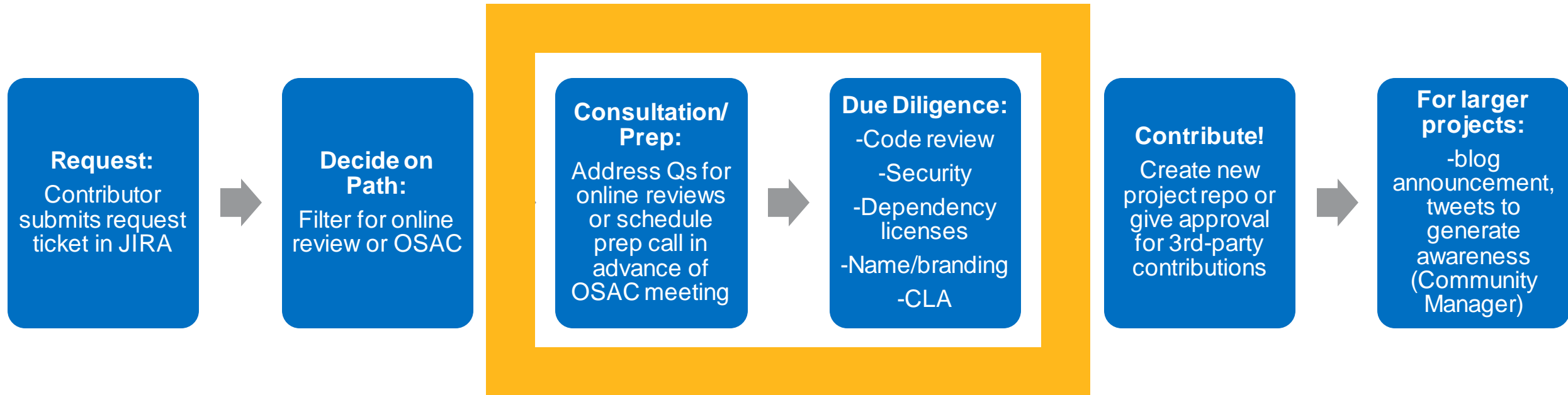Support 7k+ developers

Support ~150 project repos on GitHub

Members of Open Source Advisory Council

- OSPO
- Legal
- Strategic IP/Patents
- Engineering
- Security

Open Source Contribution Request Flow

Request:
Contributor submits request ticket in JIRA

Decide on Path:
Filter for online review or OSAC

Consultation/Prep:
Address Qs for online reviews or schedule prep call in advance of OSAC meeting

Due Diligence:
-Code review
-Security
-Dependency licenses
-Name/branding
-CLA

Contribute!
Create new project repo or give approval for 3rd-party contributions

For larger projects:
-blog announcement, tweets to generate awareness (Community Manager)

Pre-Approval Checkpoints

✓ Legal, Strategic IP – part of OSAC
✓ Security – code review part of OSAC; additional assessments as advised
✓ Compliance – dependency check and license scans
✓ Naming, branding/logo

Post-Approval Checkpoints

✓ Repo creation (Github management)
✓ Repo Linter from TODO Group

What's the ROI?

✓ Releasing secure code
✓ Complying with license obligations
✓ Fostering better ecosystem for scanning tools, machine readability for our projects
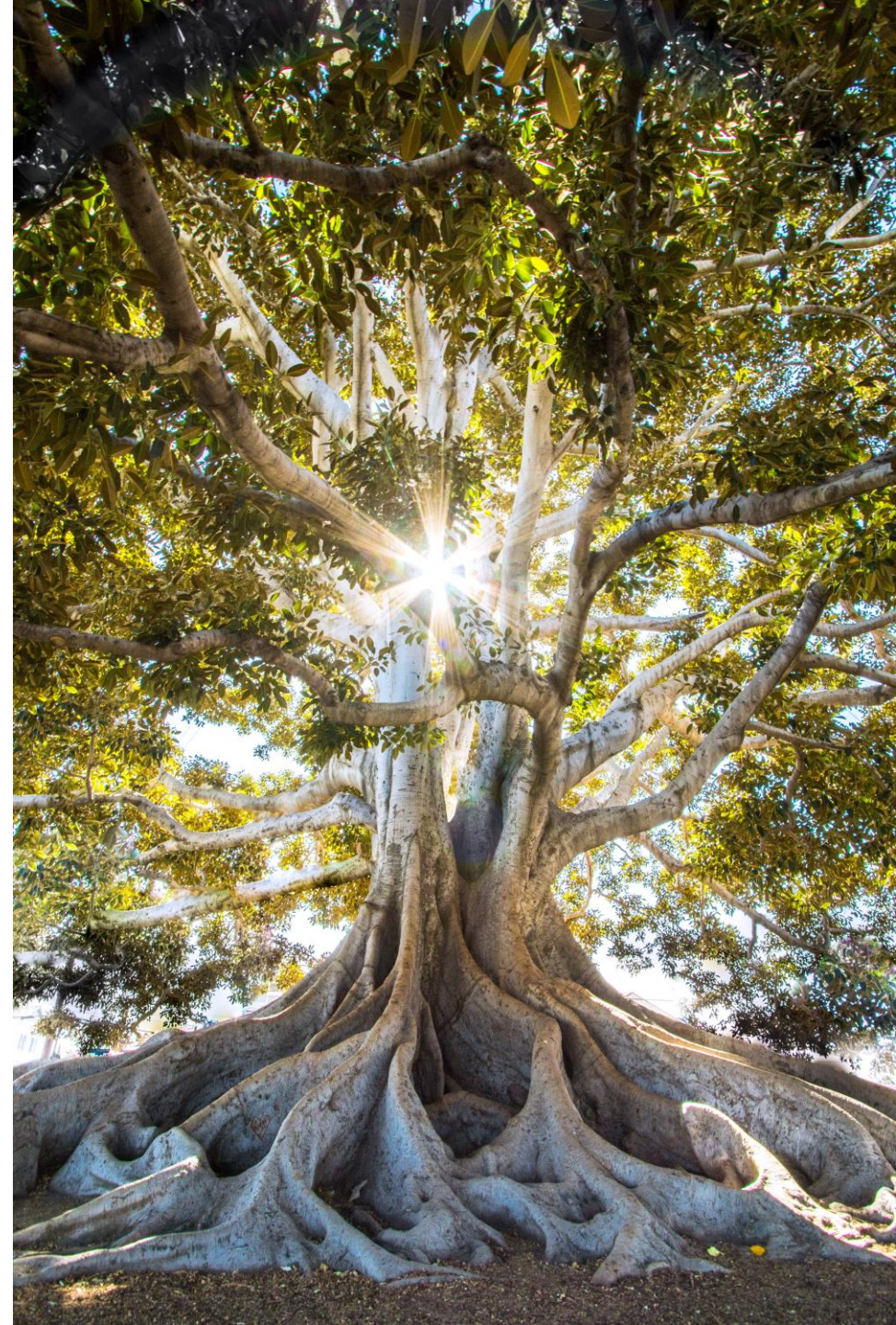
Where engineers can help…

- Catching projects at the beginning is great, but what about projects that have been out there for years?
- How do we know which projects need attention or need help with attracting contributors?
- Looking at compliance from multiple angles helps us play catch up
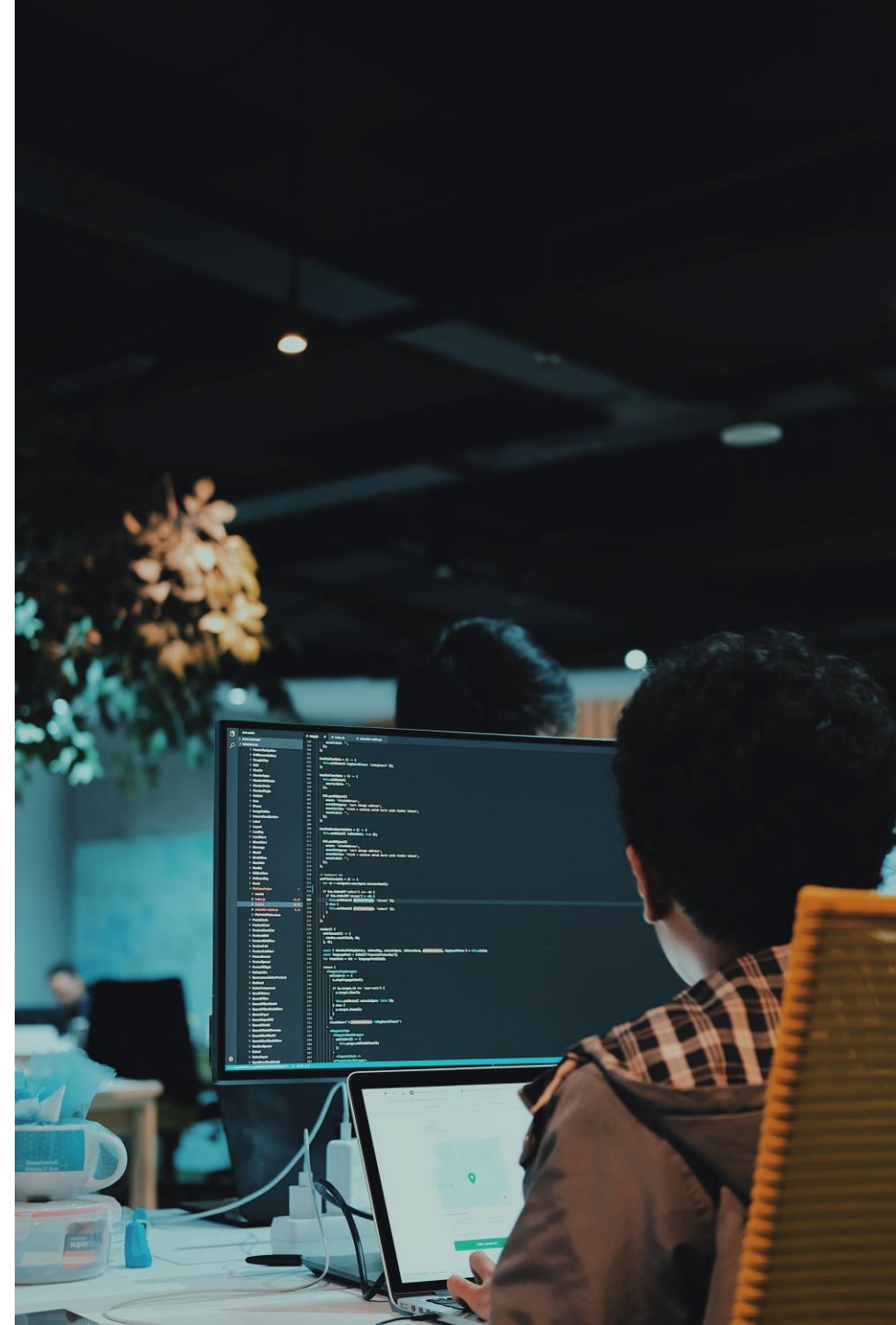
Invest in review

- Invest in review as a way to improve our relationship with the OS community and be a good corporate citizen
- Make sure we have the correct licenses on existing OS projects (code of conduct, etc.)
- Make sure are internal projects are using the complying with our open source license policies

Metrics

- Which projects are we contributing to?

- Are people outside of Comcast contributing to our projects?

- Which developers are actively contributing to open source projects?

- Which Comcast projects have an active community?
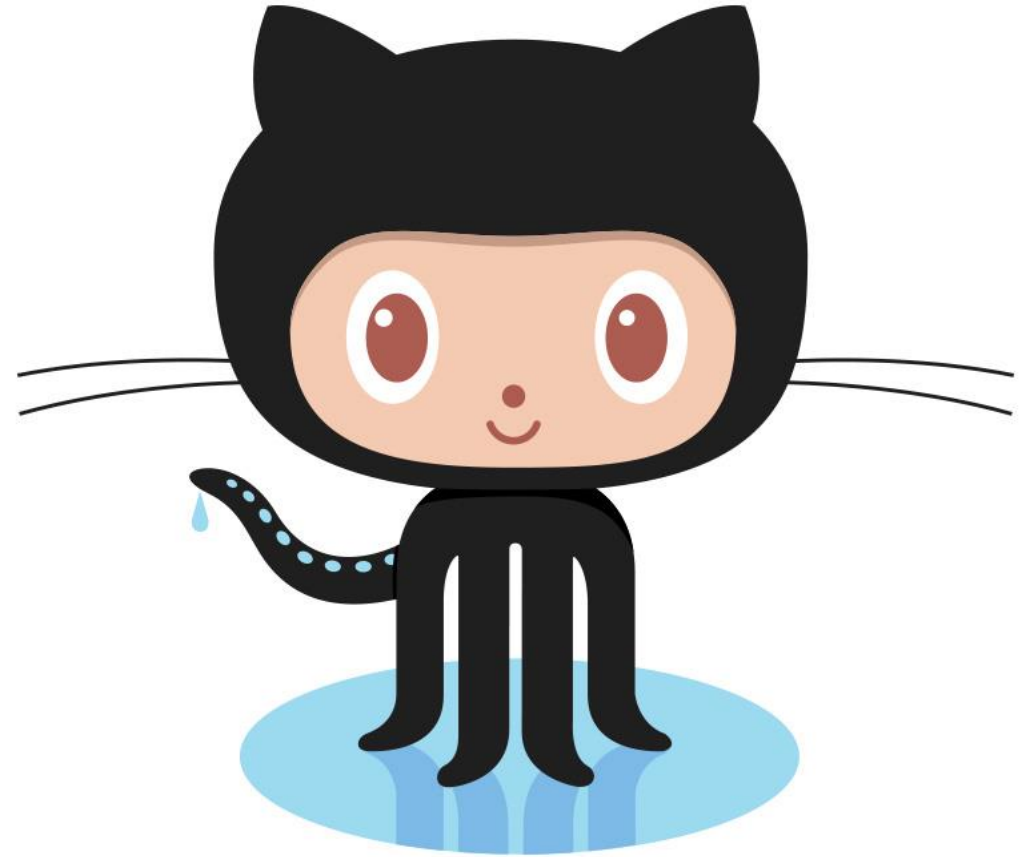
Metrics

- Use to showcase our work and activity in open source to leadership
- Involvement in CHAOSS community
  - Presented our work at 2nd annual conference in Vancouver (co-located with the Open Source Summit)

Comcast

Comcast brings together the best in media and technology to create the world's best entertainment and online experiences

🔗 https://comcast.github.io

Search repositories...   Type: All ▾   Language: All ▾   Customize pinned repositories   📖 New

parodus2ccsp

● C   ★ 1   ⑂ 13   ⚖ Apache-2.0   Updated 31 minutes ago

parodus

The XMiDT client.

xmidt   xmidt-client

● C   ★ 11   ⑂ 22   ⚖ Apache-2.0   Updated 7 hours ago

webpa-common

The collection of small common packages for the webpa project.

caduceus   scytale   webpa   talaria   petasos   xmidt

● Go   ★ 20   ⑂ 12   ⚖ Apache-2.0   Updated 3 days ago

Bynar

Server remediation as a service

rust   protobuf   zeromq   ceph

● Rust   ★ 27   ⑂ 4   ⚖ Apache-2.0   1 issue needs help   Updated 4 days ago

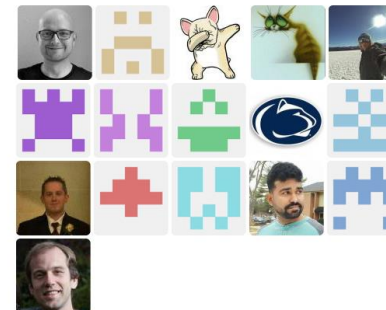### Top languages

● Go   ● Java   ● JavaScript   ● C
● Python

Most used topics   Manage

ospok   webpa   xmidt
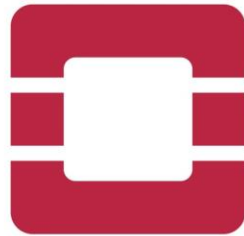
javascript   akamai

People   383 ›

Invite someone

Tools

git GH **Archive**

May 30 2018
06:02   — 1 week @ 3 hours   - - vs. 1 week ago   Jun 6 2018
                                                                                     06:02

300.00K

225.00K

150.00K

75.00K

0

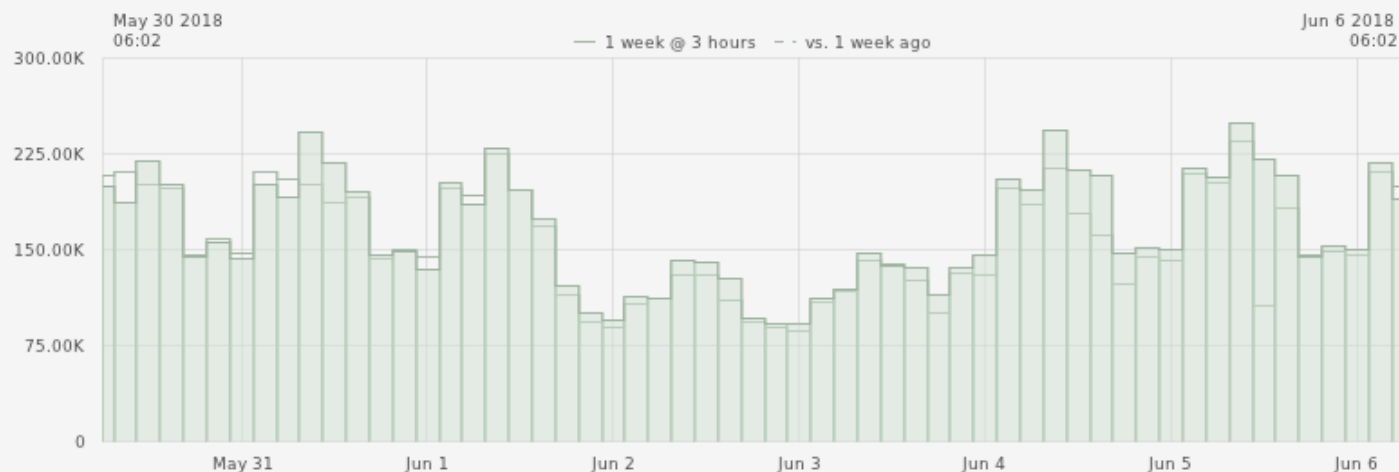   May 31   Jun 1   Jun 2   Jun 3   Jun 4   Jun 5   Jun 6

Open-source developers all over the world are working on millions of projects: writing code & documentation, fixing & submitting bugs, and so forth. GH Archive is a project to **record** the public GitHub timeline, **archive it**, and **make it easily accessible** for further analysis.

GitHub provides 20+ event types, which range from new commits and fork events, to opening new tickets, commenting, and adding members to a project. These events are aggregated into hourly archives, which you can access with any HTTP client:

Query	Command
Activity for 1/1/2015 @ 3PM UTC	`wget http://data.gharchive.org/2015-01-01-15.json.gz`
Activity for 1/1/2015	`wget http://data.gharchive.org/2015-01-01-{0..23}.json.gz`
Activity for all of January 2015	`wget http://data.gharchive.org/2015-01-{01..30}-{0..23}.json.gz`
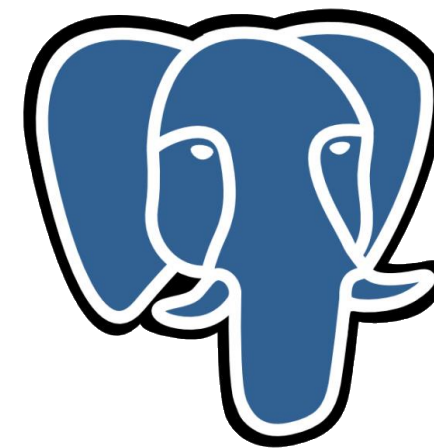
Pulse

Contributors

Community

Traffic

Commits

Code frequency

Dependency graph

Network

Forks

Git clones
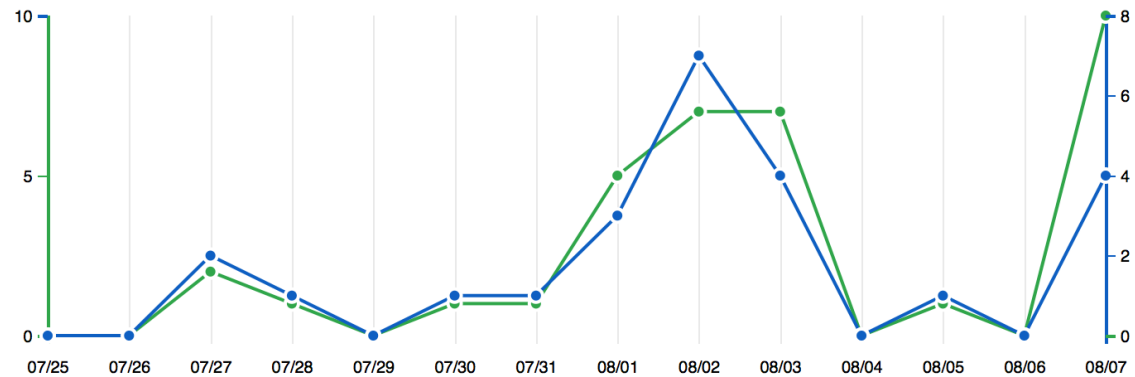
35
Clones

15
Unique cloners

Visitors
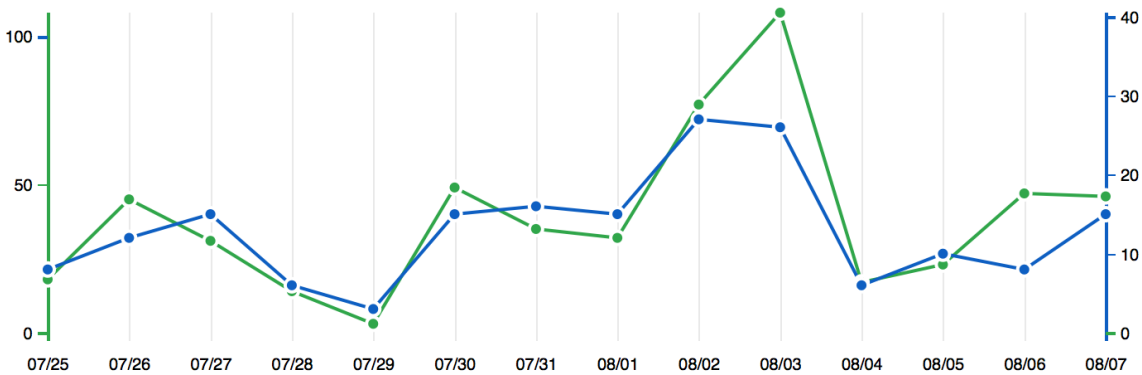
545
Views

139
Unique visitors

influxdb

<> Code   ⊘ Issues 6   ⑂ Pull requests 1   ▦ Projects 0   📖 Wiki   📊 Insights

Prometheus exporter for github metrics

prometheus-exporter   metrics   prometheus   exporter   monitoring   github

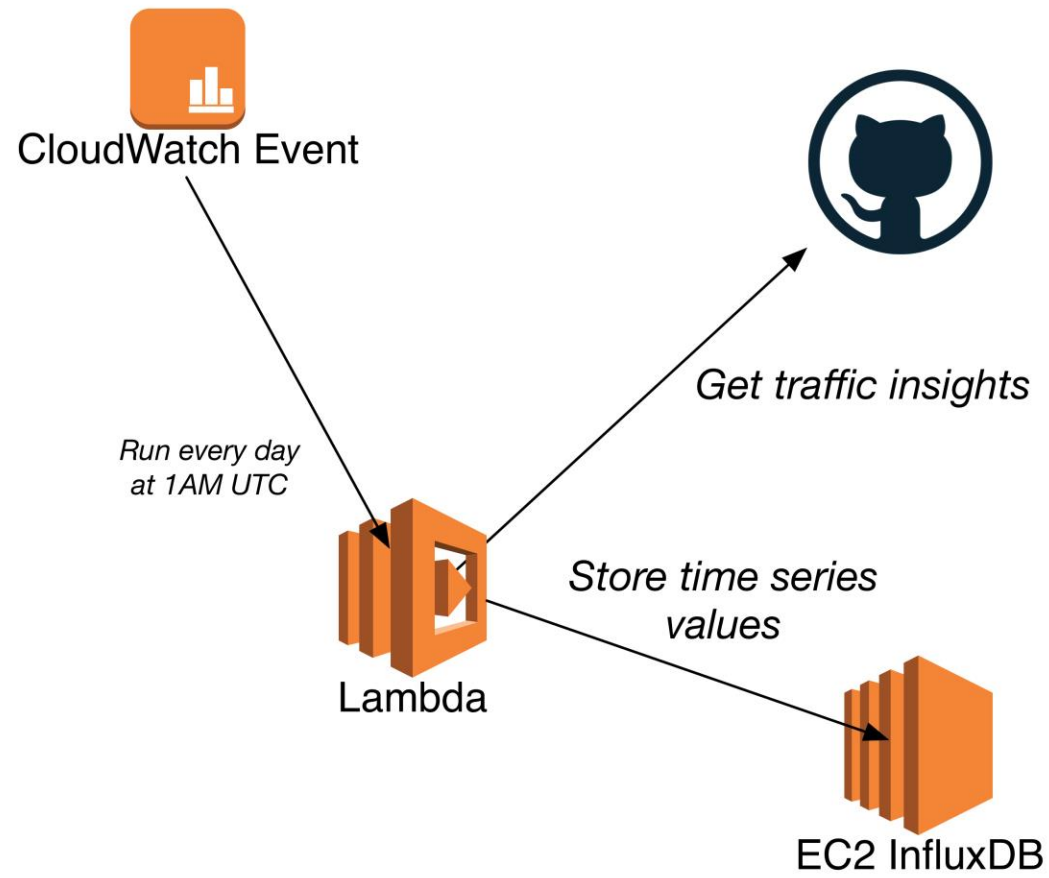⊙ **50** commits   ⑂ **3** branches   🏷 **4** releases   👥 **9** contributors   ⚖ MIT

CloudWatch Event

*Run every day
at 1AM UTC*

Get traffic insights

*Store time series
values*

Lambda

EC2 InfluxDB

CloudWatch Event

Run every day
at 1AM UTC

Lambda

Lambda

Get traffic insights

Store time series
values

EC2 InfluxDB

Commits, PRs, Author stats

RDS

ECS/Fargate

Run devstats

CloudWatch Event

github exporter

github exporter + Prometheus

Run every day at 1AM UTC

Lambda

Lambda

Get traffic insights

Store time series values

EC2 InfluxDB

ECS/Fargate

Commits, PRs, Author stats

RDS

Run devstats

CloudWatch Event

Run every day at 1AM UTC

Lambda

Lambda

github exporter

+

Prometheus

Get traffic insights

Store time series values

EC2 InfluxDB

Grafana

ECS/Fargate

Run devstats

Commits, PRs, Author stats

RDS

Resources

- TODOgroup.org

- Repo Linter: https://github.com/todogroup/repolinter
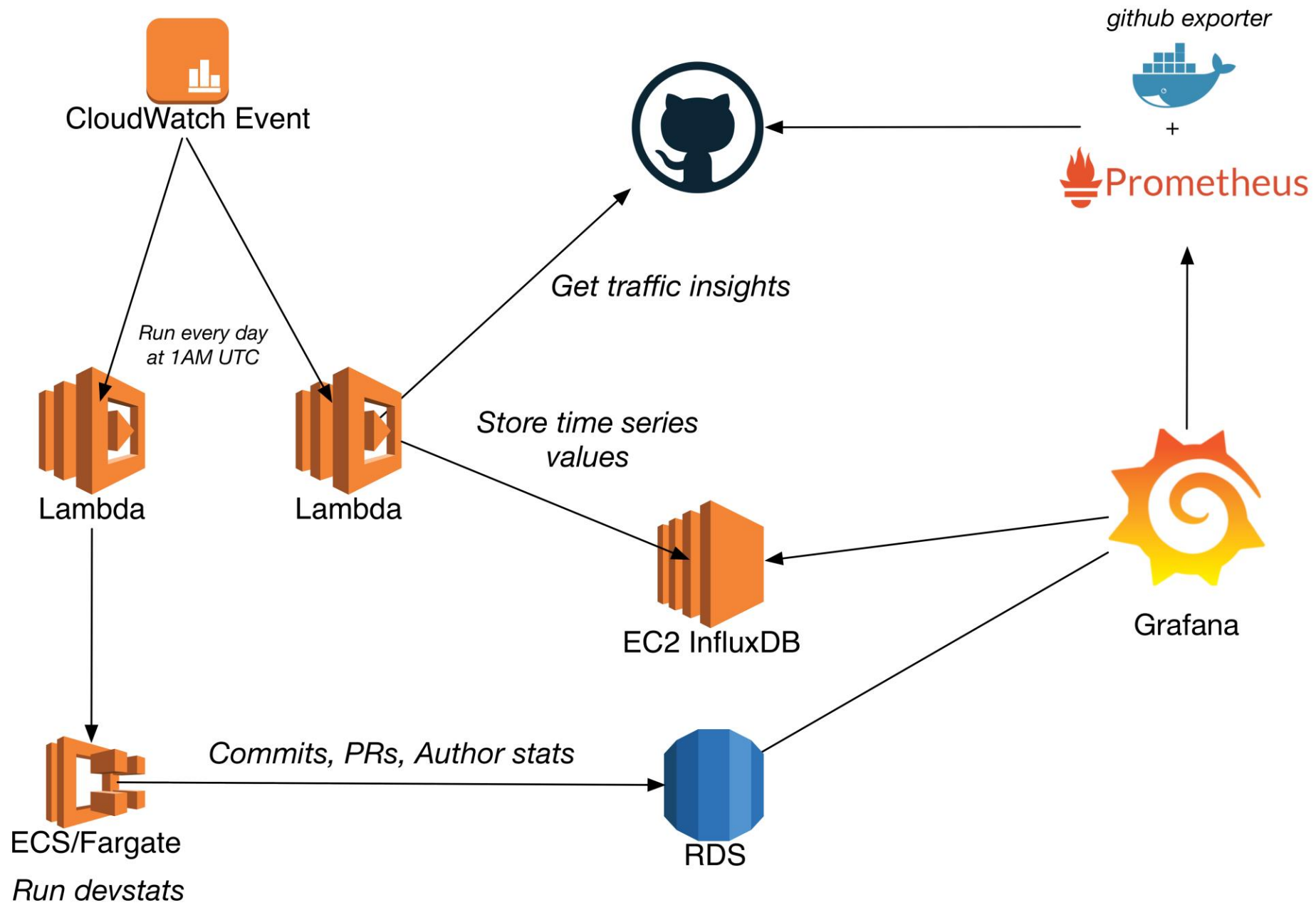
- CHAOSS Community (Community Health Analytics Open Source Software): https://chaoss.community