If There Are No Metrics, Did It Even Happen? Tracking Open Source Contributions at Comcast

Shilla Saebi @shillasaebi David Grizzanti @dgrizzanti



Comcast Open Source Program Office

Opened beginning of 2017 Support 7k+ developers Support 141 project repos on GitHub 4 person team



How do we quantify the work?

We needed a way to measure





Project health metrics can help us determine which projects are thriving.



Project health metrics can help us determine which ones may need attention. 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?



It was time to cross-collaborate.











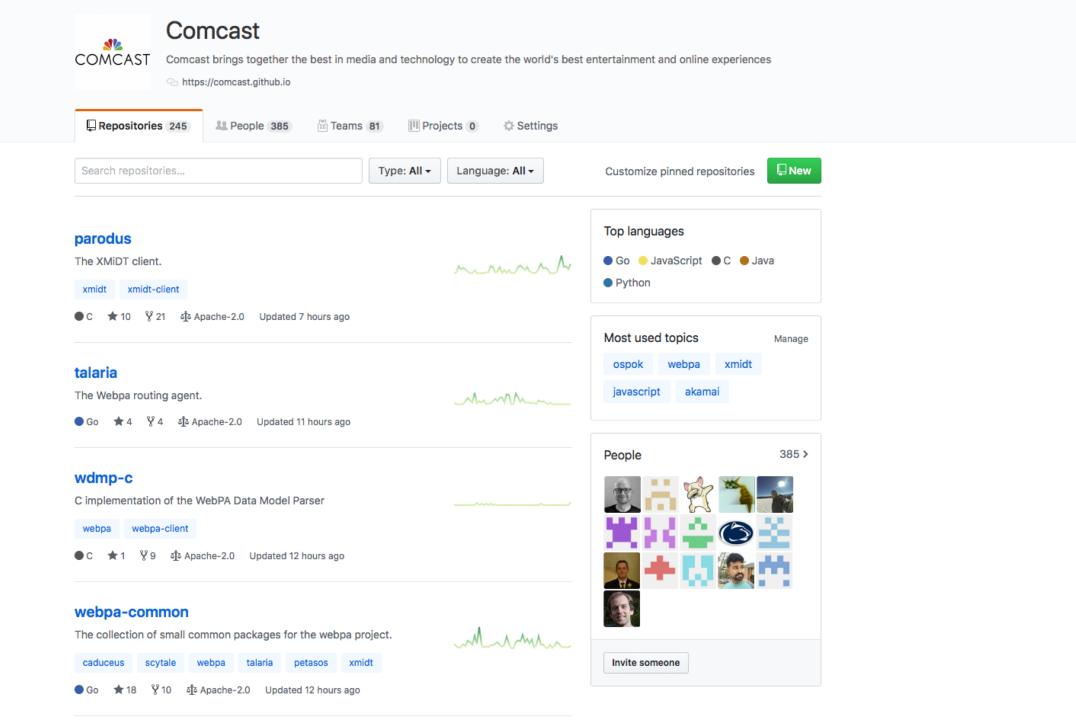
Which projects are we contributing to?

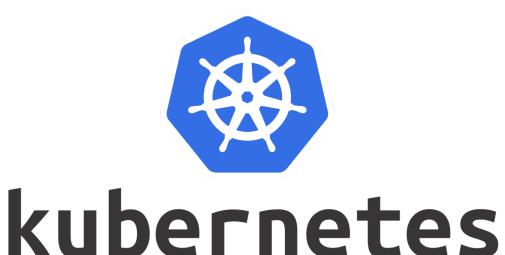
Are people outside of Comcast contributing to our projects?

Which developers are actively contributing to open source projects?

Which Comcast projects have an active community?









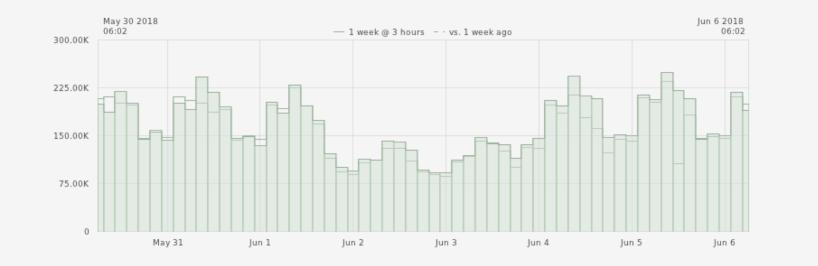












Open-source developers all over the world are working on millions of projects: writing code & documentation, fixing & submitting bugs, and so forth. GH Archive is a project to **record** the public GitHub timeline, **archive it**, and **make it easily accessible** for further analysis.

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:

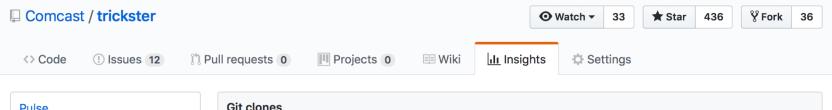
Query	Command					
Activity for 1/1/2015 @ 3PM UTC	wget http://data.gharchive.org/2015-01-01-15.json.gz					
Activity for 1/1/2015	<pre>wget http://data.gharchive.org/2015-01-01-{023}.json.gz</pre>					
Activity for all of January 2015	wget http://data.gharchive.org/2015-01-{0130}-{023}.json.gz					

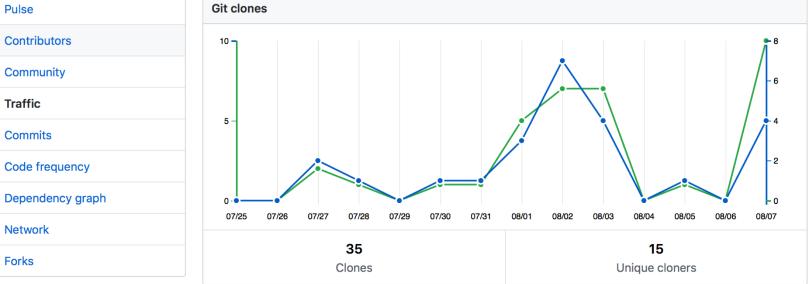


GitHub archives, git, GitHub API -> Postgres databases -> Grafana dashboards https://k8s.devstats.cncf.io/

githubarch	ive po	stgres	kuber	rnetes	prometheus	open	tracing	fluentd	linkerd	grpc	postgres-database	grafana-dashboard
metrics	apache	rkt	cni	envoy	jaeger	notary	cncf	statistics				

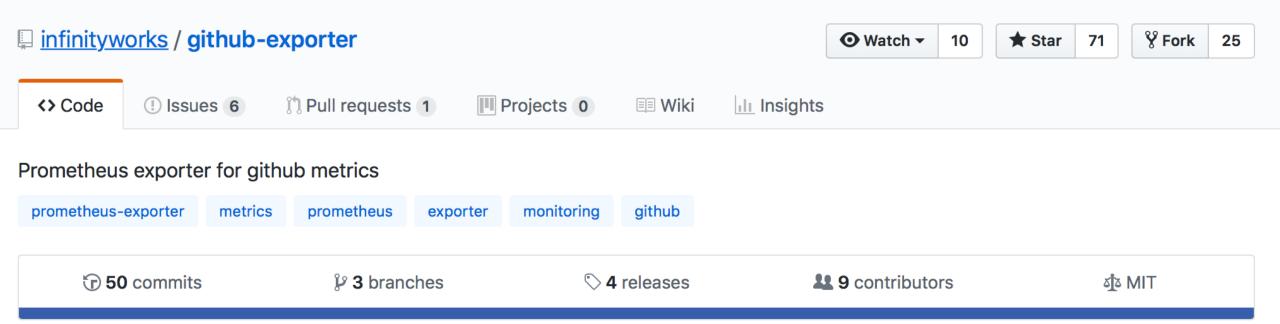




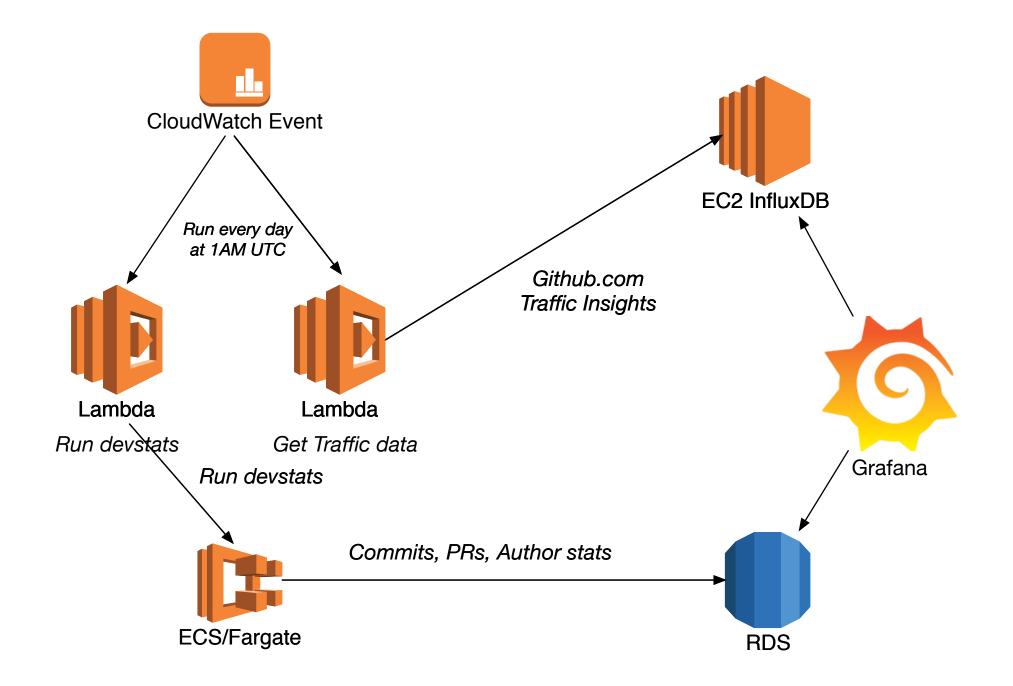


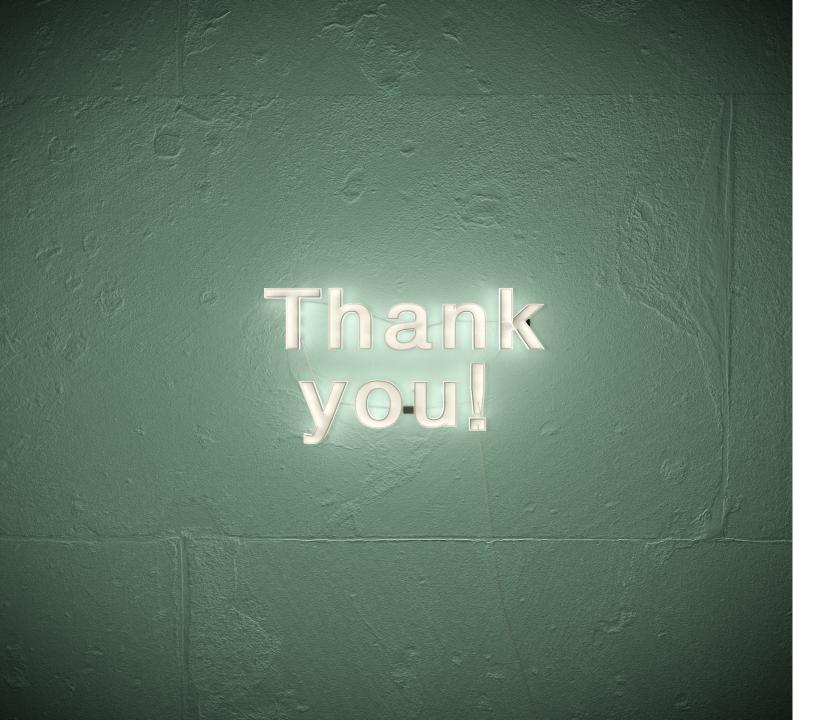












@shillasaebi @dgrizzanti

