Prometheus

Introduction and Production-ready Setup

Kubernetes and Cloud Native Group Dresden May 2019

Axel Köhler - axel.koehler@kiwigrid.com

Rico Pahlisch - rico.pahlisch@kiwigrid.com

KIWIGRID

Agenda

- 20’ Introduction
- 20’ Production-ready Setup
- 5’ Outlook

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Introduction

KIWIGRID

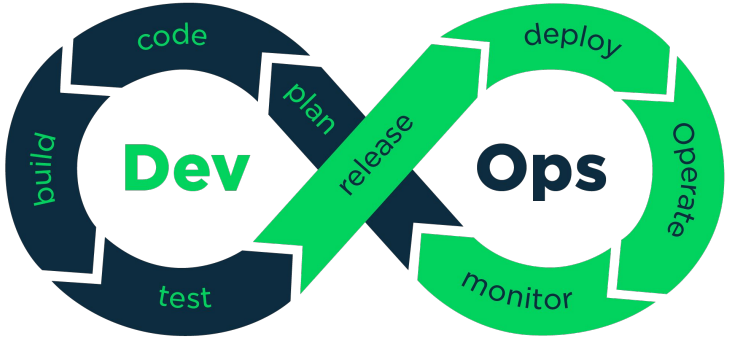
Prometheus

- Open-Source monitoring and alerting system
- Starts in 2012 by SoundCloud - v1.0 released in July 2016
- 24k stars on GitHub
- 330 contributors
- Latest Release 2.9.2 (24. April 2019)
- One of six CNCF Graduated projects

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Monitoring - Why?

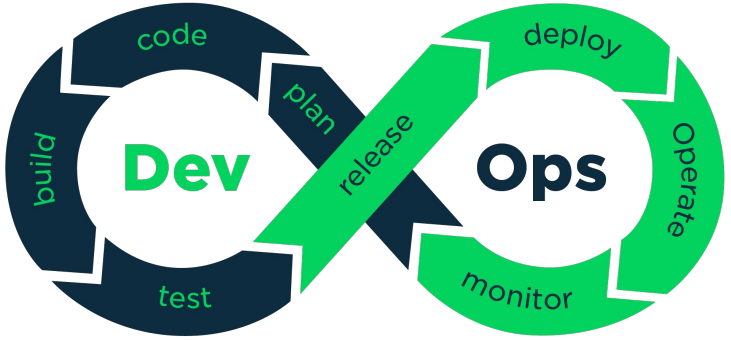
- To identify how system behaves in production
  - Bottlenecks
  - Prevention
  - Results of experiments
- “Monitoring enhances communication and trust.”
  - “CIO Guide to DevOps” (by Puppet)
- Metrics as input for scaling



https://cdn-images-1.medium.com/max/2600/1*EBXc9eJ1YRFLtkNI_djaAw.png

Metrics - What?

- Metrics represent the raw measurements of resource usage or behavior.
  - CPU / RAM / Disc usage, Throughput, ...
- Usual types:
  - Counter
    - cumulative metric that represents a single value (e.g. error count)
    - do not use for values that can decrease
  - Gauge
    - represents a single numerical value that can arbitrarily go up and down (e.g. temperature, no. of requests)
- Collected and processed by a monitoring system.
  - e.g. Nagios, Graphite, Prometheus



https://cdn-images-1.medium.com/max/2600/1*EBXc9eJ1YRFLtkNI_djaAw.png

Prometheus - Data Model

- Everything is a time-series
- Every time-series has a unique name and a set of key-value pairs called *labels*
- Notation
  - `<metric name>{<label name>=<label value>, ...}`
  - should have a single-word (application) prefix
  - `traefik_backend_requests_total{protocol="http"}`

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Prometheus - Jobs and Instances

- An endpoint you can scrape is called an instance
- A collection of instances (replicas) is called a job
- Prometheus attaches `job` and `instance` labels automatically to the scraped time series
  - `traefik_backend_requests_total{protocol="http", job="kubernetes-pods", instance="10.1.2.3:12345"}`

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Prometheus - PromQL

- Prometheus provides a functional query language called PromQL

```
sum by (backend) (rate(
  traefik_backend_requests_total{
    protocol=”http”, code=~"2.."
  }[5m]
))
```

- https://timber.io/blog/promql-for-humans/

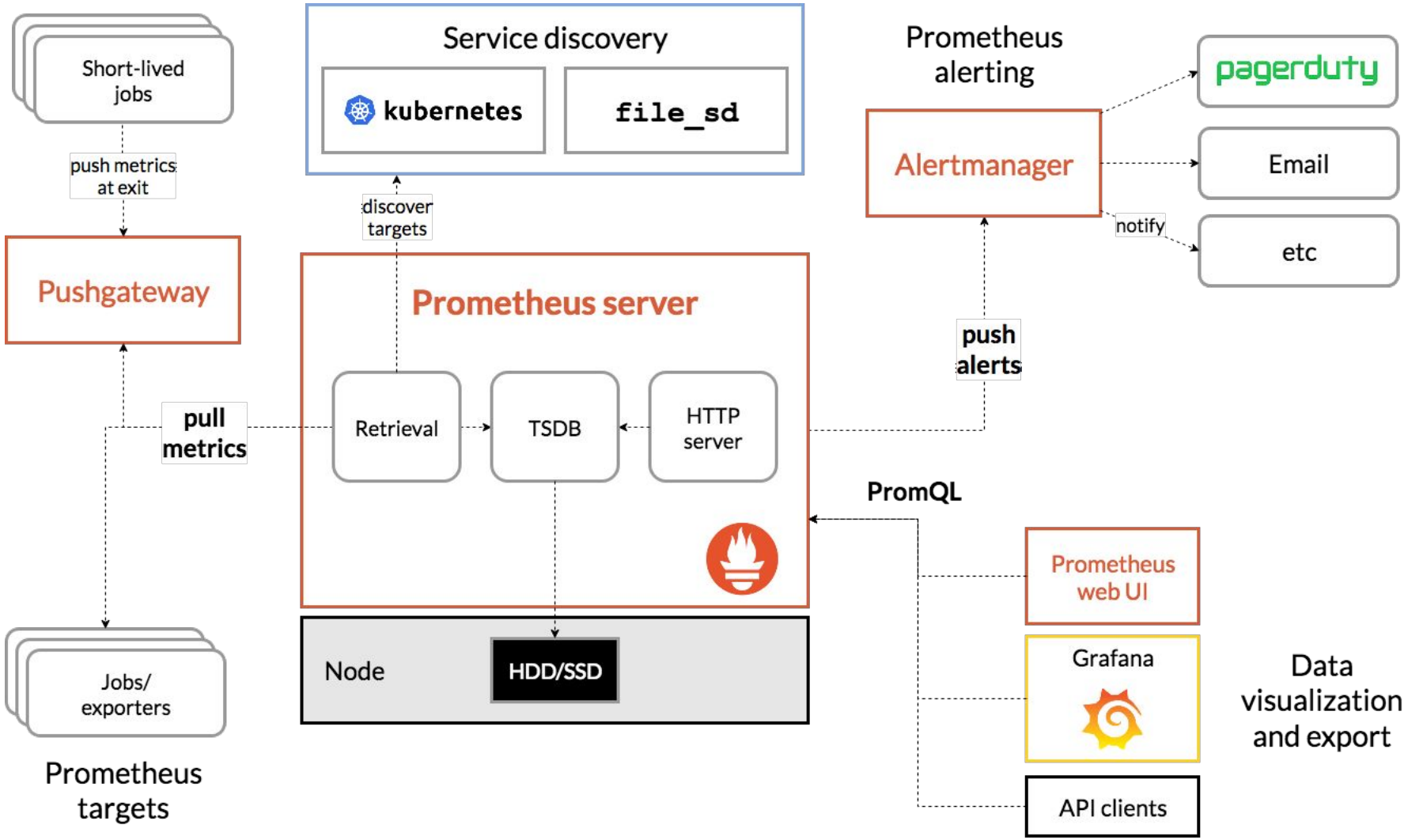
https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Prometheus - Scrape Config

- Control what and how to scrape → *Targets*
- Targets may be statically or dynamically discovered
- `relabel_configs` allow advanced modifications to any target and its labels before scraping, e.g.
  - Drop unnecessary metrics or labels
  - Adjust labels
  - Extract label values and create new labels

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Prometheus - Architecture



https://prometheus.io/assets/architecture.png

Prometheus - Pushgateway

- Allows ephemeral and batch jobs to expose their metrics
- For all cases where scraping is not possible (e.g. due short process lifespan, services behind firewalls)

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Prometheus - Alertmanager

- Handels alerts sent by clients (e.g. Prometheus, Thanos)
- Grouping, routing
- Lot of existing integrations
  - Webhook, Mail, Slack, PagerDuty, OpsGenie, ...

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Prometheus - Exporters

- Exports metrics from third-party systems as Prometheus metrics
  - *kube-state-metrics* - listens to the Kubernetes API server and generates metrics about the state of the objects
  - *node-exporter* - exports hardware and OS metrics exposed by *NIX kernels
  - *stackdriver-exporter* - requests Stackdriver API for the metrics

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Production-ready Setup

KIWIGRID

Demo Time

- Kubernetes single cluster setup for multiple teams
- Prometheus Server + Alertmanager
- Grafana for visualization
- Alerts per team
- Dashboards per team

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Prometheus Helpers

- https://github.com/kiwigrid/k8s-sidecar
  - watches kubernetes configmaps
  - notify Prometheus server / Grafana on changes

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Scrape Kubernetes Pods

- `prometheus.io/scrape: “true”`
- `prometheus.io/port: “9000”`
- `prometheus.io/path: /actuator/prometheus`
  - If path is not /metrics, define it with this annotation.

https://upload.wikimedia.org/wikipedia/en/thumb/3/38/Prometheus_software_logo.svg/1033px-Prometheus_software_logo.svg.png

Resources

- https://prometheus.io
- https://timber.io/blog/promql-for-humans/
- https://thanos.io

Go out and monitor!

KIWIGRID

Backup Slides

KIWIGRID