



**KubeCon**



**CloudNativeCon**

North America 2018

# Monitor The World

Meaningful Metrics for Kubernetes  
Applications and Clusters



# Agenda



KubeCon



CloudNativeCon

North America 2018

- Kubernetes Monitoring Intro
- Applications
- Control Plane
- Monitoring at Planet Labs

# Who am I?



KubeCon



CloudNativeCon

North America 2018

- Currently working on Amazon EKS
- Formerly worked at two Seattle startups using Kubernetes, Porch and OfferUp

# Why



KubeCon



CloudNativeCon

North America 2018

- Problem Detection
- Outage Prevention
- Optimization
- I am Nosy

# Its Hard



KubeCon



CloudNativeCon

North America 2018

- Many Microservices
- Many Containers
- Many Perspectives

# A Method to the Madness



KubeCon



CloudNativeCon

North America 2018

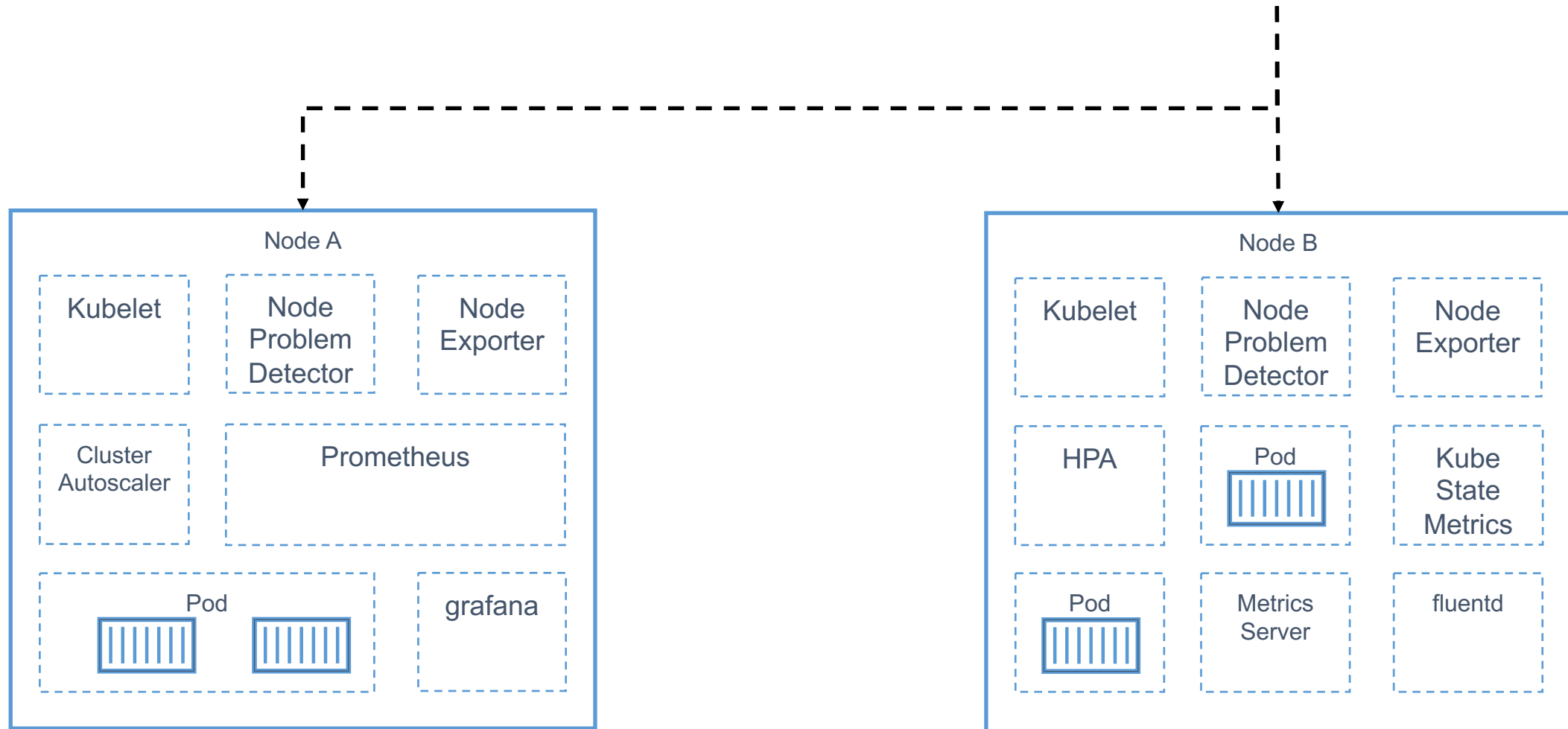
- Resources

- USE method by Brendan Gregg
- For every resource, check:
  - Utilization
  - Saturation
  - Errors

- Services

- RED method by Tom Wilkie
- For every service, monitor request:
  - Rate
  - Errors
  - Duration

# Metrics Environment





# Applications



KubeCon



CloudNativeCon

North America 2018





# Start with Your Users



KubeCon



CloudNativeCon

North America 2018

## Business Metrics

- Orders fulfilled successfully
- Rides completed
- Pictures taken and sent to earth

## Application Request Errors

- Tells you where to start
- Use tracing and logs to determine where to look next

## Application Latency

- Critical measurement of user experience

# A Complete Picture



KubeCon



CloudNativeCon

North America 2018

## Know Your Code and Configuration Version

- Know what version your code is, and where it has been deployed
- The same goes for configuration!
- Add a version label to your PodSpecs

```
sum(kube_pod_labels{label_version != "", label_app = "myapp"}) by (label_version)
```

## Request Rate & Saturation

# Resources



KubeCon



CloudNativeCon

North America 2018

- Consume from Kubelet (over Kube State Metrics)
- From Kubelet:
  - **container\_cpu\_usage\_seconds\_total**
  - **container\_memory\_working\_set\_bytes**
  - **kubelet\_volume\_stats\_available\_bytes**

# Kube State Metrics



KubeCon



CloudNativeCon

North America 2018

## Container restarts

- `kube_pod_container_status_restarts_total`

## Pods % available

- `kube_deployment_status_replicas_available / kube_deployment_status_replicas`
- `kube_poddisruptionbudget_status_current_healthy / kube_poddisruptionbudget_status_desired_healthy`



# The Control Plane



KubeCon



CloudNativeCon

North America 2018



# The Bare Minimum



KubeCon



CloudNativeCon

North America 2018

RED for API Server

Kube System Availability

Etcctl Availability

# As Your Cluster Scales



KubeCon



CloudNativeCon

North America 2018

## API Server Resource Usage

- Do you see dropped requests?
- Should you adjust **--target-ram-mb** and **--max-requests-inflight**?
- **requests** and **limits** based on object churn, node count, pod density.

## API Server clients

- **--kube-api-burst**
- **--kube-api-qps**

# As Your Cluster Scales



KubeCon



CloudNativeCon

North America 2018

## Scheduling Latency

- From Kube State Metrics:
  - **kube\_pod\_status\_scheduled\_time**
- From Kube Scheduler:
  - **scheduler\_e2e\_scheduling\_latency** histogram

## Controller Work Time

- From Controller Manager:
  - **\*\_work\_duration**
  - **\*\_queue\_latency**





## Leader Elections

- **etcd\_server\_has\_leader**
- **etcd\_server\_leader\_changes\_seen\_total**

## Disk Write Performance

- **etcd\_disk\_wal\_fsync\_duration\_seconds\_bucket**
- **etcd\_disk\_backend\_commit\_duration\_seconds\_bucket**



## Database Size

- When **etcd\_mvcc\_db\_total\_size\_in\_bytes** reaches the quota limit, etcd will trigger a **NOSPACE** alarm

## Corruption

- **--experimental-initial-corrupt-check**





# Nic Cope Kubernetes Infrastructure Lead

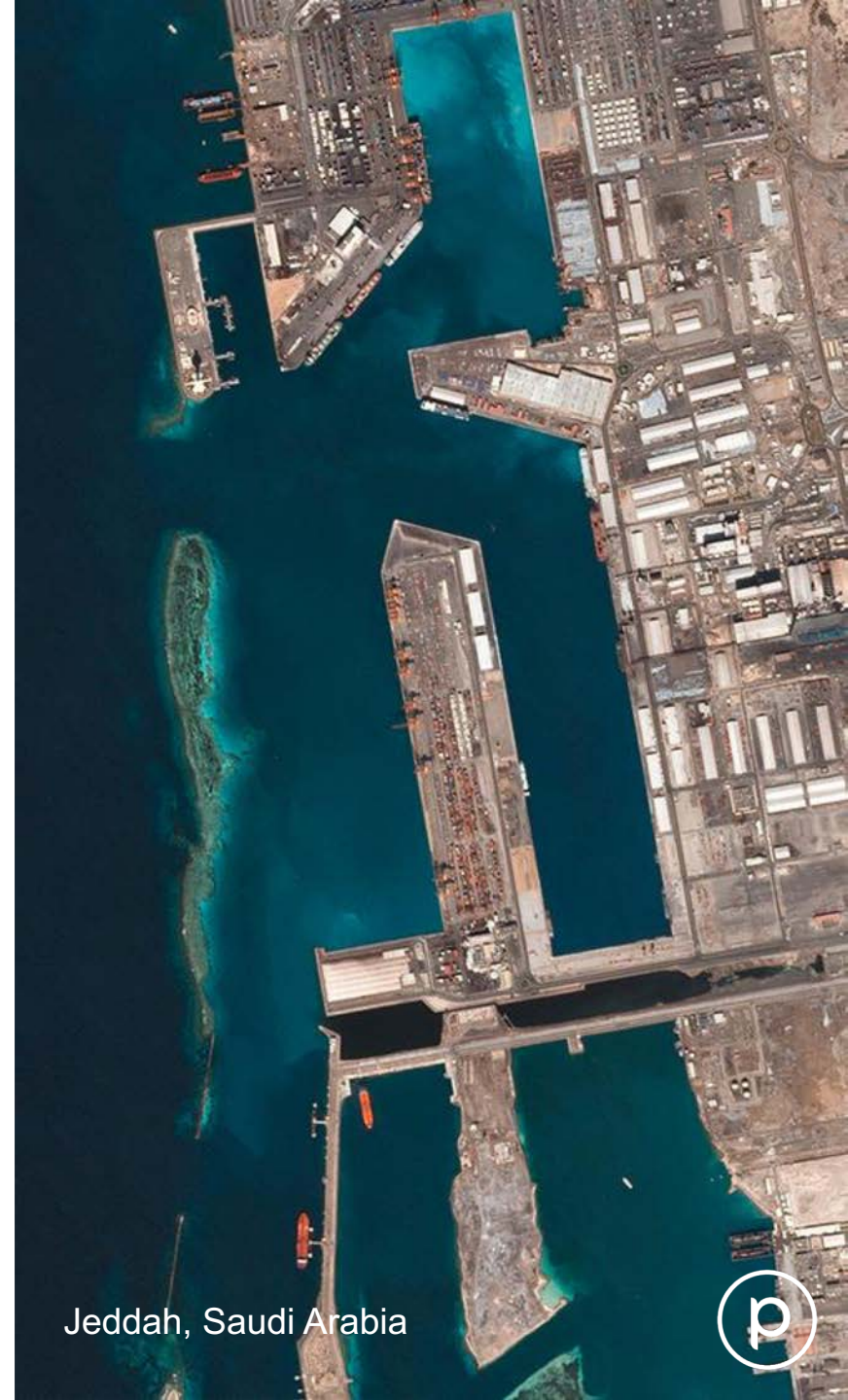
+

Mailiao Refinery, Taiwan – May 31, 2016





- ~70 engineers
- Tens of services
- Five person Kubernetes team, aka Hobbes



Jeddah, Saudi Arabia







**“Welcome to your first day at Planet! Can you have a ~5,000 node Kubernetes cluster running in a quarter?”**



An aerial photograph of the Great Barrier Reef, showing the intricate patterns of the coral and the surrounding ocean. The water is a deep blue, and the reef structures are a lighter, sandy color. A semi-transparent dark blue rectangular box is centered over the image, containing the title text.

# Our Monitoring Philosophy

Great Barrier Reef, Australia – July 8, 2016





Page only when **customers are affected.**



Page only when **software won't fix it.**



Page only **the team who can fix it.**



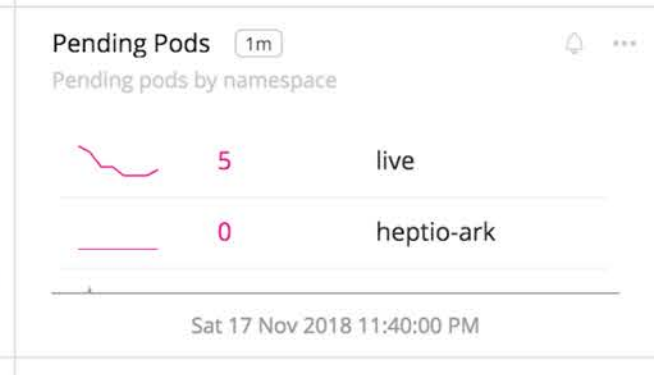
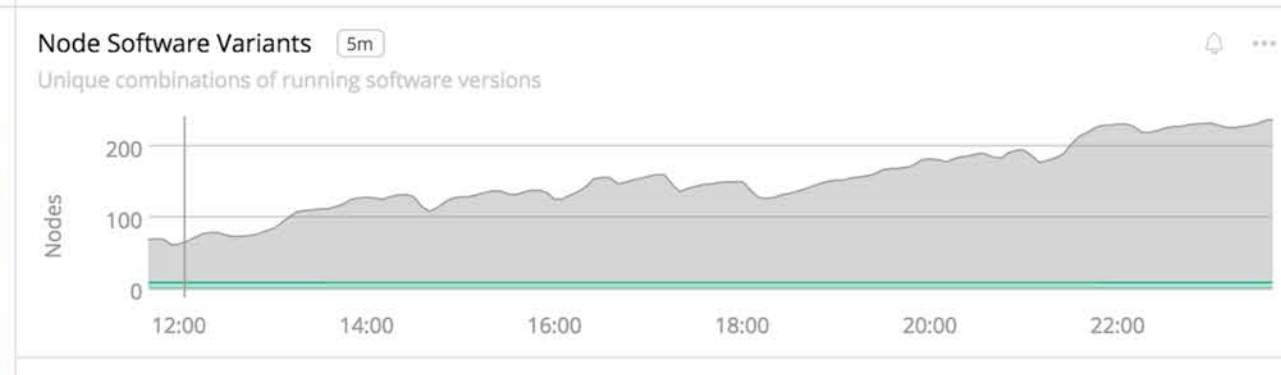
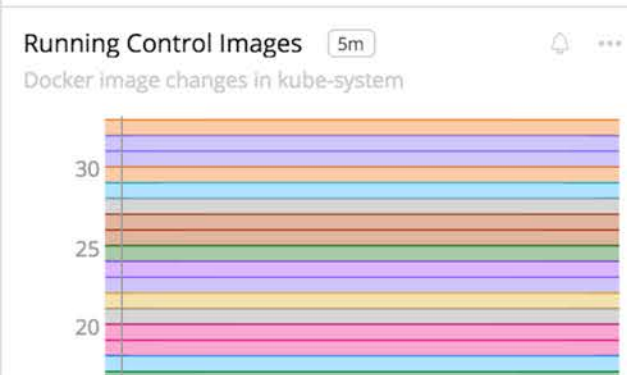
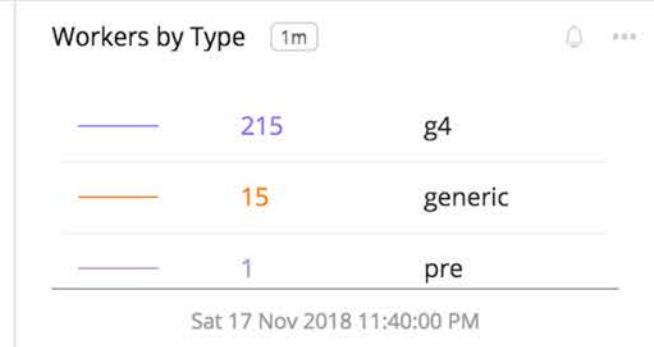
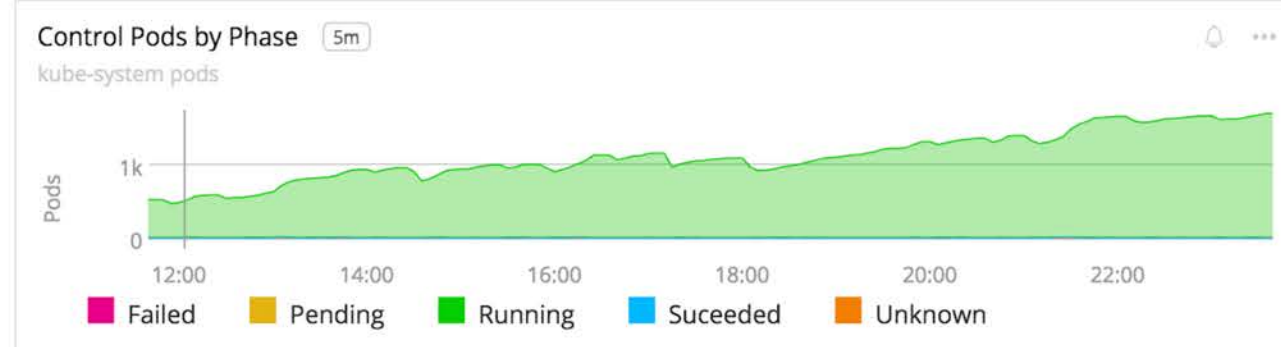
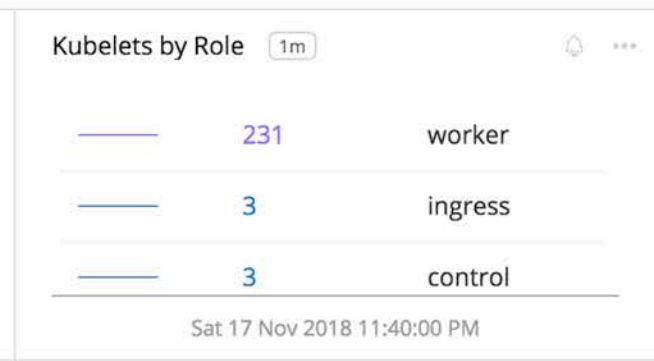
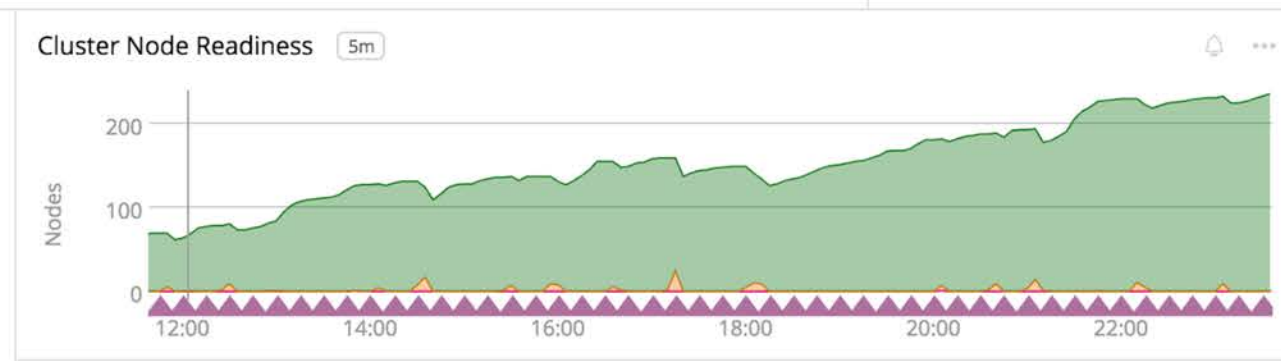
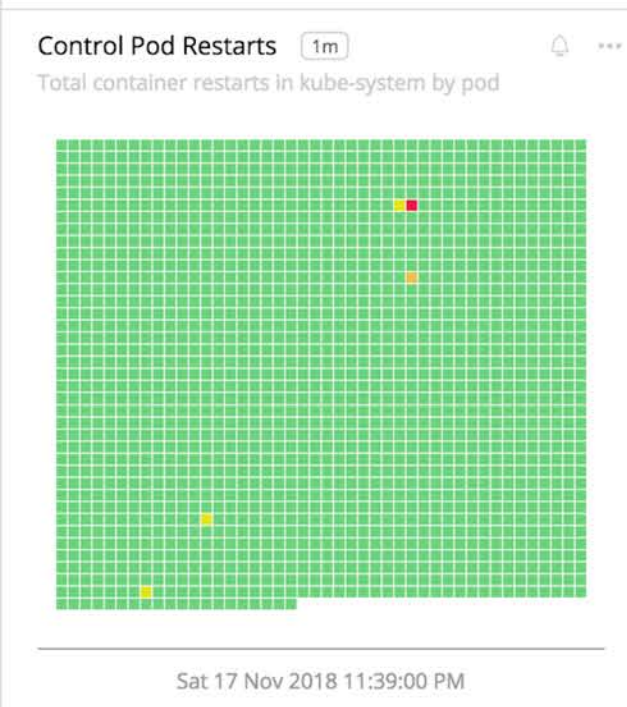
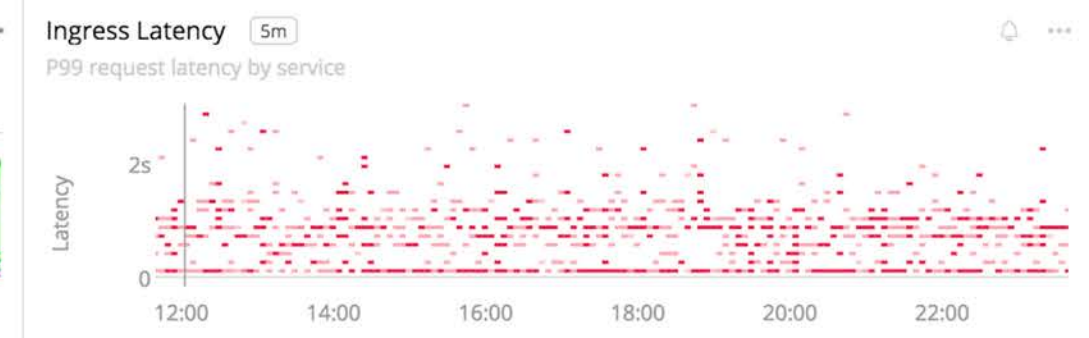
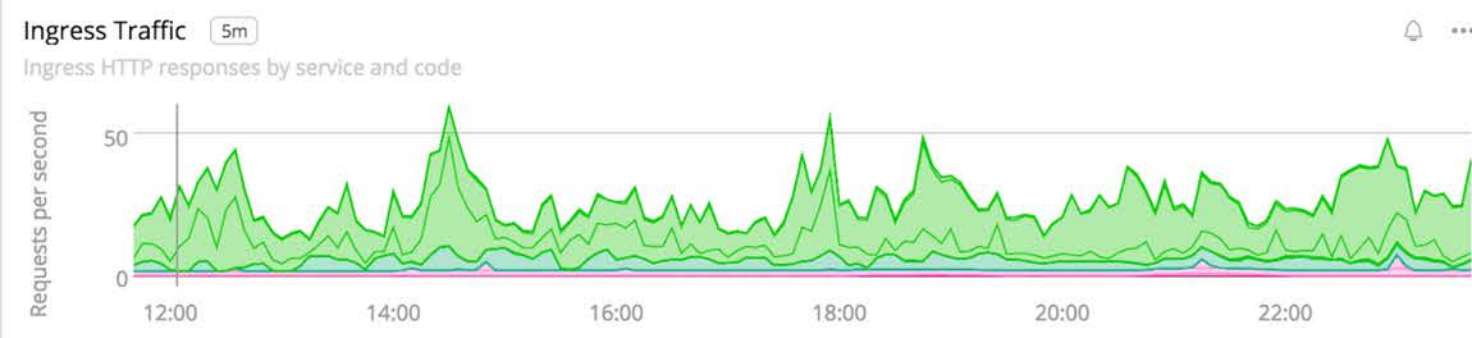


# What we monitor

Great Barrier Reef, Australia – July 8, 2016









- Customers instrument their own services.
- Hobbes provides standard metrics: service mesh health, state of Kubernetes, service costs.
- Customers define their own alerts.





An aerial photograph of the Great Barrier Reef, showing the intricate patterns of the coral and the surrounding ocean. A semi-transparent dark blue rectangular box is centered over the image, containing the text "But where is Prometheus?".

**But where is Prometheus?**



- It's difficult to balance under vs over alerting.
- Look deeper when customers ask for observability functionality.
- It's hard to beat Prometheus when it comes to monitoring cloud native software.





# Summary



KubeCon



CloudNativeCon

North America 2018

- Kubernetes monitoring environment is complex.
- Start with the metrics that effect your users.
- Scaling your cluster involves reactively tuning based on some important metrics.
- It helps to approach your metrics environment with a method and philosophy that works for you and your organization.

# Thank You!



KubeCon



CloudNativeCon

North America 2018

[nic@amazon.com](mailto:nic@amazon.com)

[https://twitter.com/Nck\\_T](https://twitter.com/Nck_T)

<https://github.com/nckturner>

[negz@planet.com](mailto:negz@planet.com)

<https://twitter.com/internegz>

<https://github.com/negz>

Draino <http://github.com/planetlabs/draino>

Kostanza <http://github.com/planetlabs/kostanza>

Grab us after for questions/feedback/beers/etc.



**KubeCon**

**CloudNativeCon**

————— **North America 2018** —————

