You can’t do that with Kubernetes…
or can you?

Ricardo Rocha, CERN

Kubernetes
Community
Days
Denmark


MESOSPHERE
KubeCon
CloudNativeCon
Europe 2018
MESOSPHERE


ogle
ity Member,
KubeCon
CloudNativeCon
Europe 2018
Prometheus
fluentd
CNI


Europe 2018


Crossing the River by
Feeling the Stones
Simon Wardley
Researcher,
Leading Edge Forum
GRPC
OPENTRACING
kubernetes
CloudNativeCon
Europe 2018
Prometheus
fluentd
CNI


Keynotes
Auditoriums 10-12 + 15


CLOUD NATIVE
COMPUTING
FOUNDATION

2016 2017 2018 2019 2020 2021 2022

1. You can’t do scale?

One million requests per second: Dependable and dynamic distributed systems at scale

Wednesday, November 11, 2015

Recently, I've gotten in the habit of telling people that building a reliable service isn't that hard. If you give me two Compute Engine virtual machines, a Cloud Load balancer, supervisord and nginx, I can create you a static web service that will serve a static web page, effectively forever.

Brendan Burns, Senior Staff Software Engineer, Google, Inc.

2017

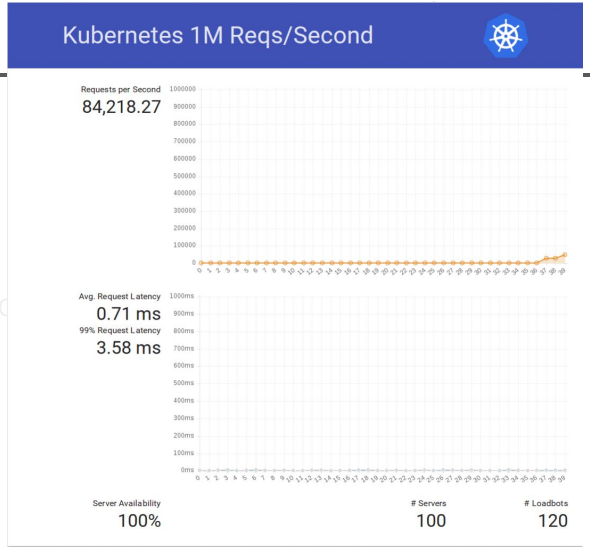
2016

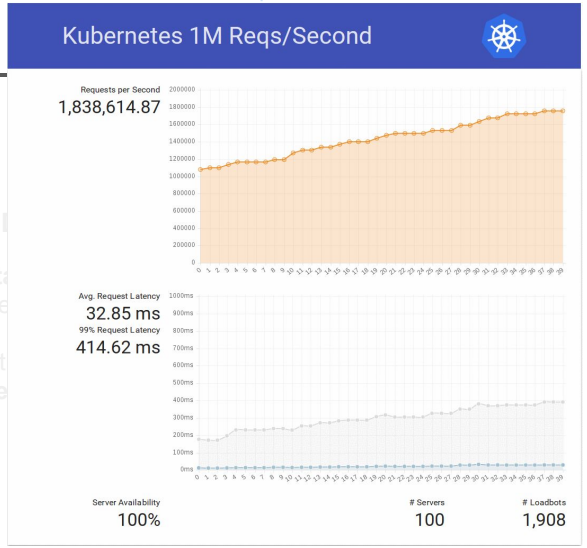
Pilot Service

Kubernetes, Swarm

Not obvious around this time which orchestrator would win. **Scalability and performance** tests.

Cluster Size (Nodes)	Deployment Time (min)
2	2.5
32	4
128	5.5
512	14
1000	23





One million requests per second: Dependable and dynamic distributed systems at scale

Wednesday, November 11, 2015

Recently, I've gotten in the habit of telling people that building a reliable service isn't that hard. If you give me two Compute Engine virtual machines, a Cloud Load balancer, supervisord and nginx, I can create you a static web service that will serve a static web page, effectively forever.

Brendan Burns, Senior Staff Software Engineer, Google, Inc.

Cluster Size (Nodes)	Deployment Time (min)
2	2.5
32	4
128	5.5
512	14
1000	23

The results of the first evaluation are presented in Table 1. Numbers for clusters up to 128 nodes are very good, with an almost flat deployment time. Above that there seems to be a close to linear increase following the cluster size, even if 23 minutes is a perfectly acceptable value for deploying a 1000 node cluster.

Kubernetes, Swarm

Not obvious around this time which orchestrator would win. **Scalability and performance** tests.

One million requests per second: Dependable and dynamic distributed systems at scale

Wednesday, November 11, 2015

Recently, I've gotten in the habit of telling people that building a reliable service isn't that hard. If you give me two Compute Engine virtual machines, a Cloud Load balancer, supervisord and nginx, I can create you a static web service that will serve a static web page, effectively forever.

Brendan Burns, Senior Staff Software Engineer, Google, Inc.

Cluster Size (Nodes)	Deployment Time (min)
2	2.5
32	4
128	5.5
512	14
1000	23

2016

2017

2018

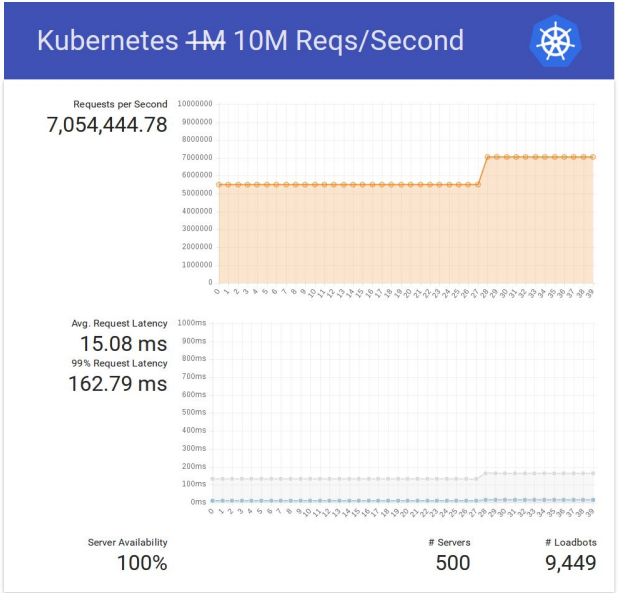
2022

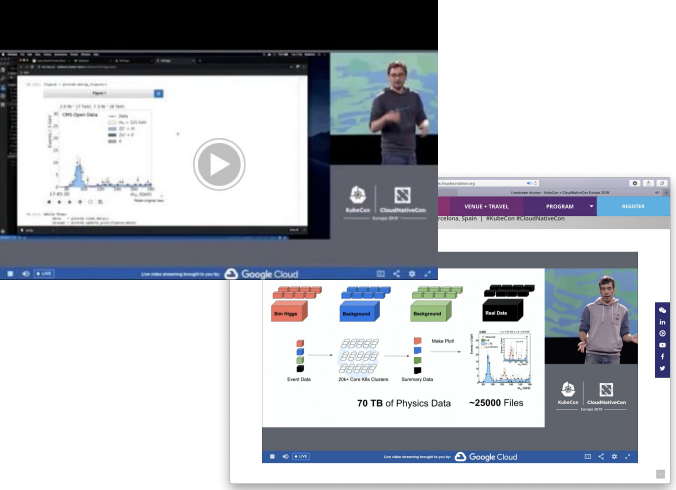
Pilot Service

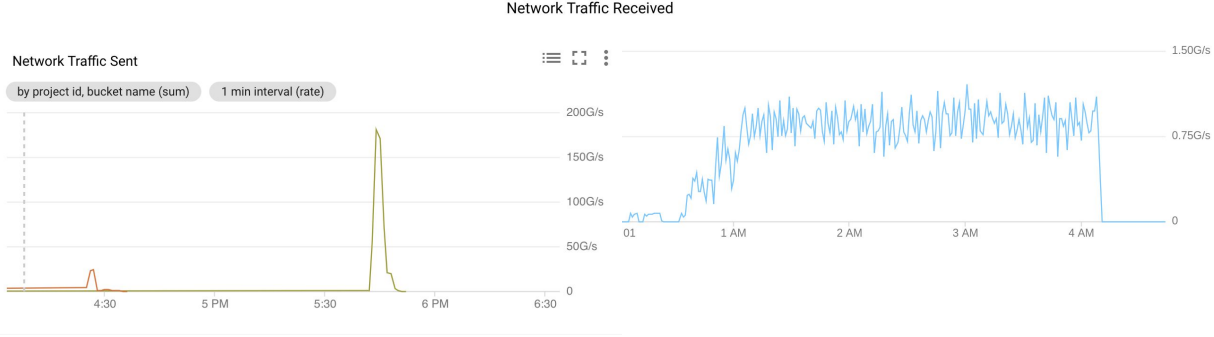
Kubernetes, Swarm

Not obvious around this time which orchestrator would win. **Scalability and performance** tests.

Kubernetes ~~1M~~ 10M Reqs/Second

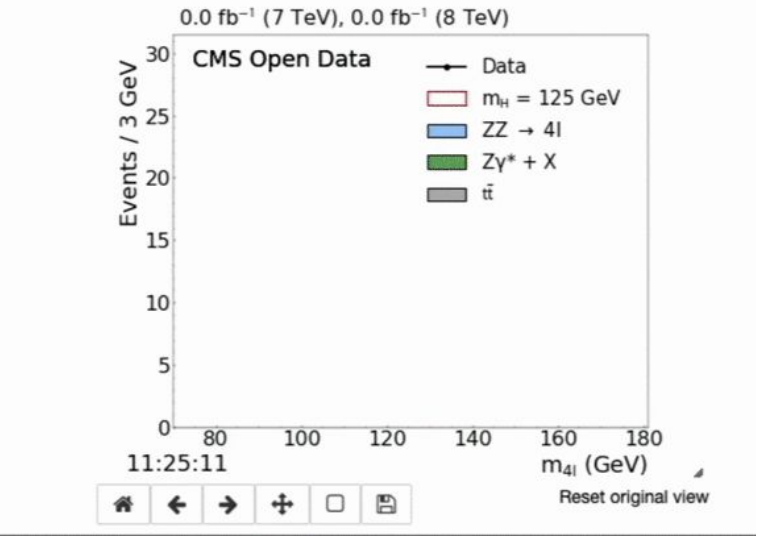






CERN openlab

Google Cloud



Google Cloud Storage → Cluster on GKE → Job Results → jupyter

70 TB Dataset

Cluster on GKE

Max **25000 Cores**

Single Region, 3 Zones

25000 **Kubernetes Jobs**

Job Results

Interactive Visualization

Aggregation

Containers & Kubernetes

65,000 nodes and counting: Google Kubernetes Engine is ready for trillion-parameter AI models

November 14, 2024

2. You can do containers

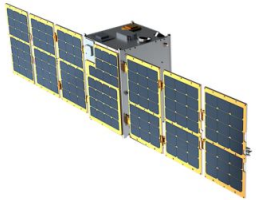
2. You can *only* do containers


Crossplane

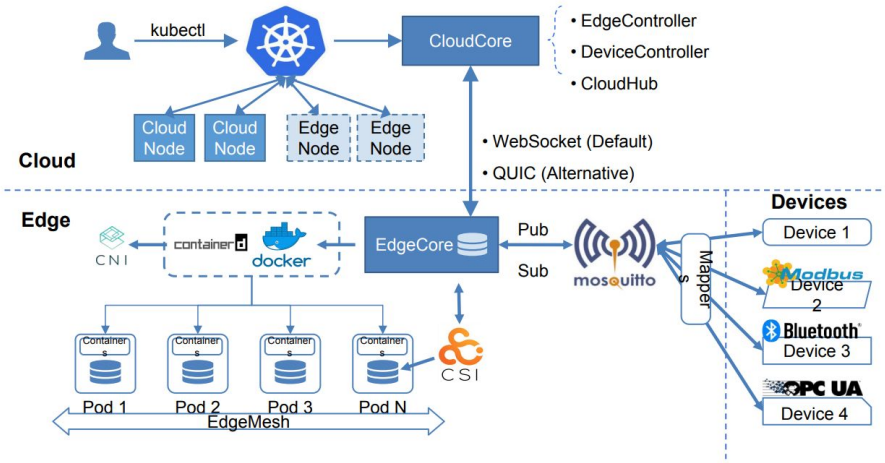

KubeVirt

3. You can’t do critical things?

Incremental Deep Learning for satellite with KubeEdge and MindSpore, Huawei



Number	Orbital Altitude	Mass	Battery Capacity	Spectrum	Uplink Rate	Downlink Rate	ISLs	Processors
1	500±50km	≤ 30kg	118Wh – 236Wh	X-band	0.1Mbps – 1Mbps	100Mbps – 600Mbps	NO	CPU/NPU
2	500±50km	≤ 30kg	118Wh – 236Wh	X-band	0.1Mbps – 1Mbps	100Mbps – 600Mbps	NO	CPU/NPU
3	500±50km	≤ 30kg	118Wh – 236Wh	X-band	0.1Mbps – 1Mbps	100Mbps – 600Mbps	NO	CPU/NPU
4	> 500km	> 50kg	> 360Wh	X, Ku, Ka bands	≥ 200Mbps	≥ 1Gbps	YES	CPU/NPU/GPU
5	> 500km	> 50kg	> 360Wh	X, Ku, Ka bands	≥ 200Mbps	≥ 1Gbps	YES	CPU/NPU/GPU
6	> 500km	> 50kg	> 360Wh	X, Ku, Ka bands	≥ 200Mbps	≥ 1Gbps	YES	CPU/NPU/GPU



ns were selected to build Minimum Viable Products (MVPs) using cloud native best
mple applications to demonstrate that it could be done. But Chaillan took a different
stems so General Officer and Senior Executives will pay attention, and that's where you
ing. So if you get people excited and show you can do it, then you can demonstrate
there is something there."

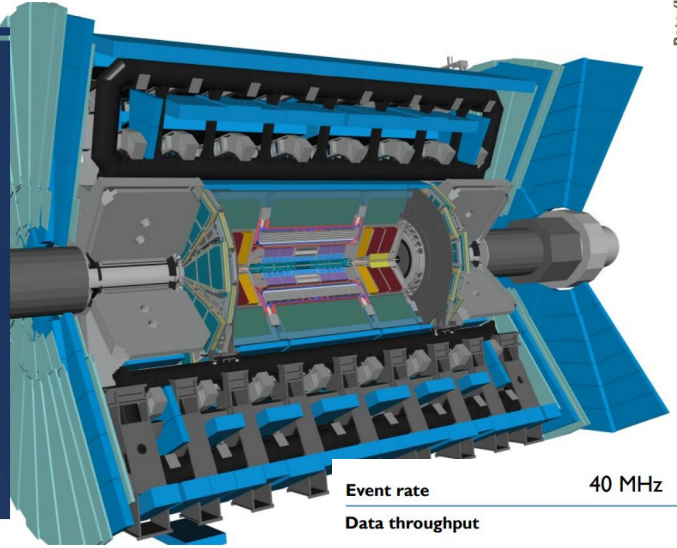
So in the fall of 2019, the SoniKube team based at Hill Air Force Base in Utah set out to get Kubernetes running on an F-16 jet. Members of the DoD's Platform One team, led by Jeff McCoy, were embedded with the group to teach them how to put Kubernetes on the jet's legacy hardware. "We had to be able to boot Kubernetes with Istio on the jet within two minutes, because that's a requirement for the jet if something goes wrong, and it has to be able to spin back up within two minutes," says Chaillan. "That was the biggest challenge."

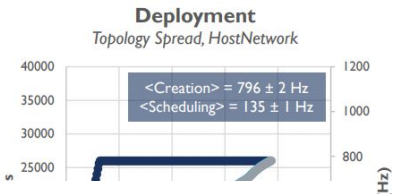
Within 45 days, the team accomplished that goal, and were able to do a demo on the jet for Dr. Will Roper, Assistant Secretary of the Air Force for Acquisition, Technology and Logistics. "We got the cluster on Istio running and then we launched five or six microservices," says Chaillan. "A lot of the jet runs in older programming languages, and so being able to run Go, Python, and Java was pretty exciting."

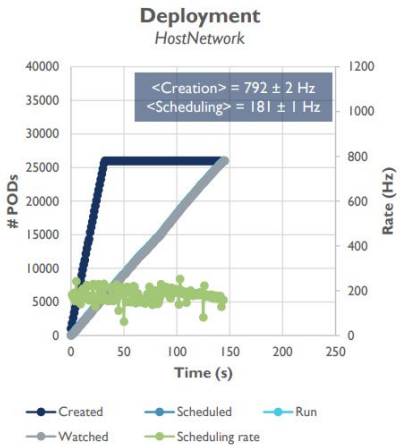
A total of 37 teams are currently working on building applications on top of Kubernetes: "We have teams doing this at every side of the weapons systems, from the space systems to the nuclear systems to the jets," he says.

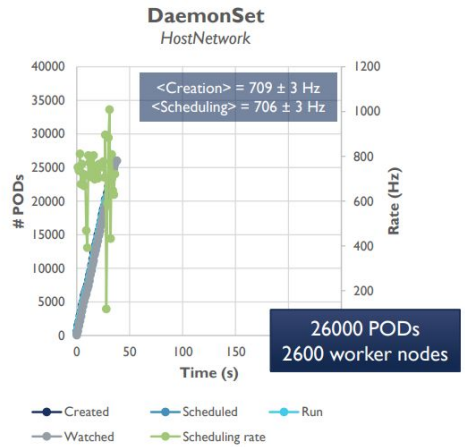
ATLAS

- **One of the four main particle detectors at the Large Hadron Collider (LHC)**
  - 44 m length
  - 25 m height
- **Collect data from proton collisions provided by the LHC**
  - 40 MHz collision rate
  - 14 TeV proton-proton energy

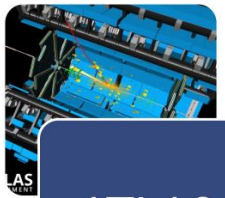








Event rate	40 MHz	100 kHz (**1 MHz**)	3 kHz (**10 kHz**)
Data throughput		300 GB/s (**5 TB/s**)	8 GB/s (**50 GB/s**)

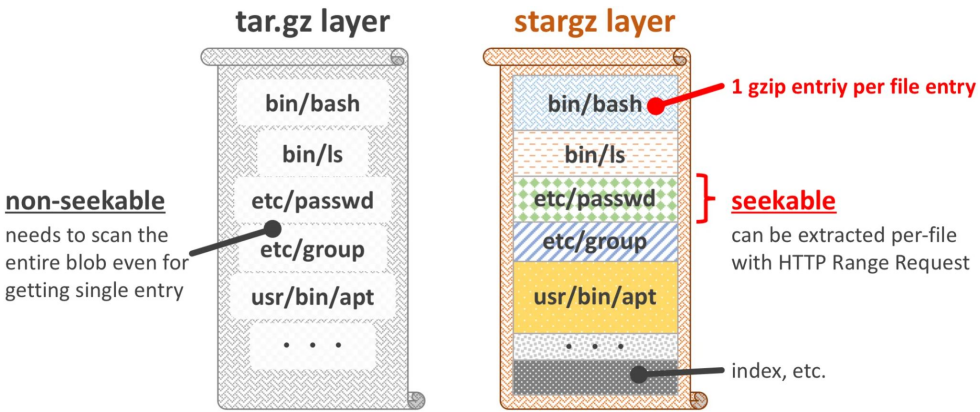


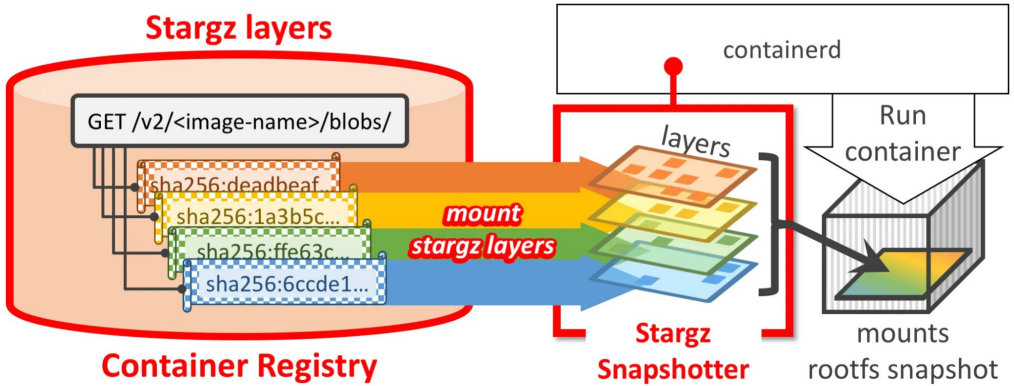
4. You can’t do batch?


VOLCANO


ARMADA

5. You need well behaved users?


tar.gz layer
stargz layer
bin/bash
bin/ls
etc/passwd
etc/group
usr/bin/apt
. . .
1 gzip entriy per file entry
non-seekable
needs to scan the entire blob even for getting single entry
seekable
can be extracted per-file with HTTP Range Request
index, etc.


Stargz layers
GET /v2/<image-name>/blobs/
sha256:deadbeaf...
sha256:1a3b5c...
sha256:ffe63c...
sha256:6ccde1...
mount
stargz layers
Container Registry
containerd
layers
Run container
Stargz Snapshotter
mounts rootfs snapshot

6. You can do single cluster

6. You can *only* do single cluster

cilium

Istio

linkerd

7. You can’t do… ?

Thank you!