Kubernetes fundamentals

(for sysadmins)

About me

- Sysadmin
- Nimium
- Working with Kubernetes for the past year
- @0x6976

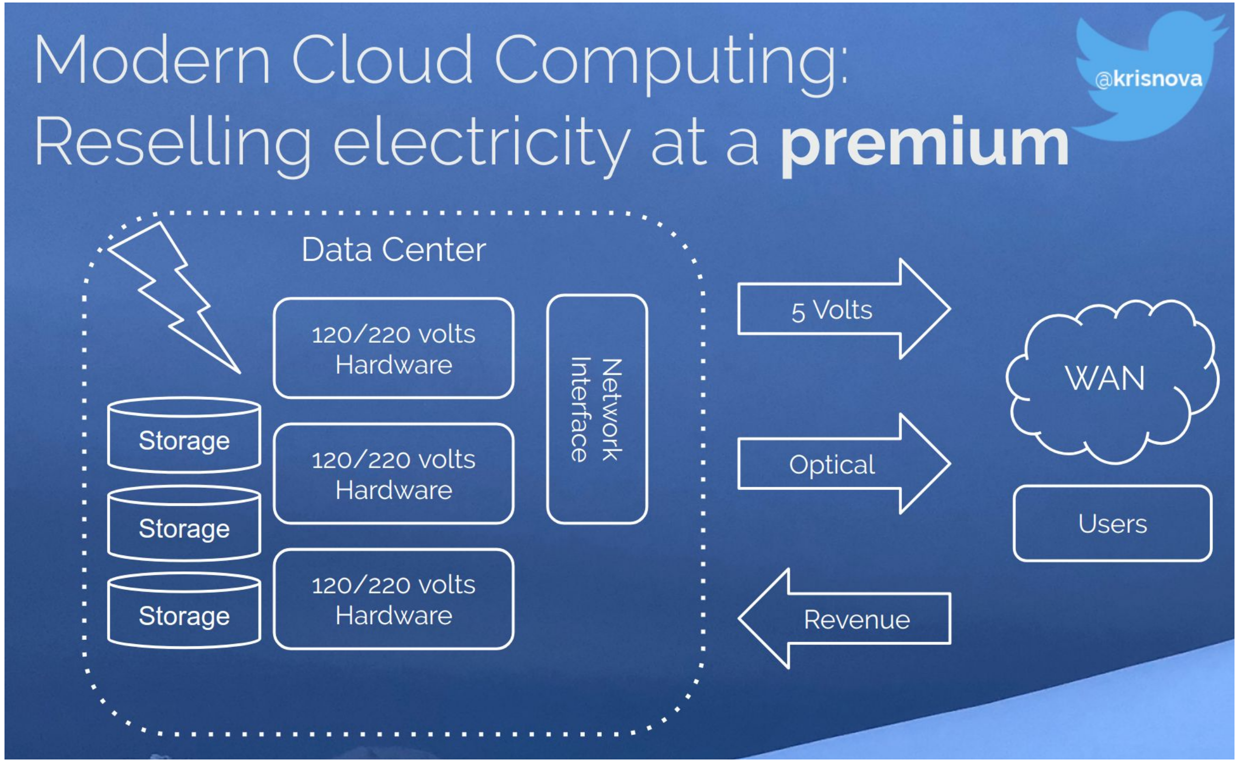
Overview

- What is Kubernetes
- The building blocks of Kubernetes
- Some hiccups you might run into on your way to Kubernetes
- Q&A

What is Kubernetes?

- “Kubernetes is named after the Greek god of spending money on cloud services”
    - Corey Quinn (@QuinnyPig)

What is the cloud?



- Kris Nova (@krisnova)

What is Kubernetes?

- Three main questions
- Where?
    - Public cloud
    - **On premises**
    - Some kind of a stretched configuration
- How?
    - **Self-supported**
    - Outsourcing
    - The cloud(TM)
- What actually is Kubernetes?

What is Kubernetes?

- A container orchestrator
- A set of software components that manages application lifecycle
- A scheduler
- A container runtime
- A set of pods (container groups)
- A set of replication and scaling rules for pods

“Not my monkey, not my circus” line

- Separation of concerns (KUAR book)
- Application developer
  - Uses the API
- API reliability engineer
  - Maintains the container orchestration API
- OS reliability engineer
  - Takes care of the operating system
- HW reliability engineer
  - Takes care of the hardware

- **What are the building blocks in Kubernetes?**

- etcd
- API server
- Storage
- DNS
- Ingress
- Container runtime
- Networking

- **What are the building blocks in Kubernetes?**
- Monitoring
- Deployment
- Service mesh
- Container registries
- Resource limits
- Upgrade story
- Software ecosystem

etcd

- Key-value store
- Make sure you have enough hardware
  - https://github.com/etcd-io/etcd/blob/master/Documentation/op-guide/hardware.md#hardware-recommendations
- **SSDs**
- Run a **multi-node cluster**
  - Five-member cluster recommended

etcd

- **Backup**
- Security
  - Firewall
  - PKI
- **Restrict access to etcd**
  - Having access to etcd == having root access to the cluster
- **Scaling an etcd cluster is done for reliability, not performance**

API server

- The front-end for the Kubernetes control plane
- Talks to etcd
- Can be scaled horizontally

Storage

- Well abstracted in Kubernetes
  - Volumes
  - Storage claims
- Numerous options
  - **Ceph**
  - NFS
  - Hardware-specific integrations

DNS

- Runs inside the cluster
- **Make sure your networking is set up properly**
  - Your worker nodes need to be able to talk to the DNS
  - Your master nodes need to be able to talk to the DNS
- **Make sure your pods and nodes are configured properly**
  - Queries which do not match the configured cluster domain suffix will be forwarded to the upstream DNS defined on the node

Ingress

- A lot of available options
  - **NGINX**
  - F5
  - Contour
  - HAProxy
  - Traefik
  - Istio

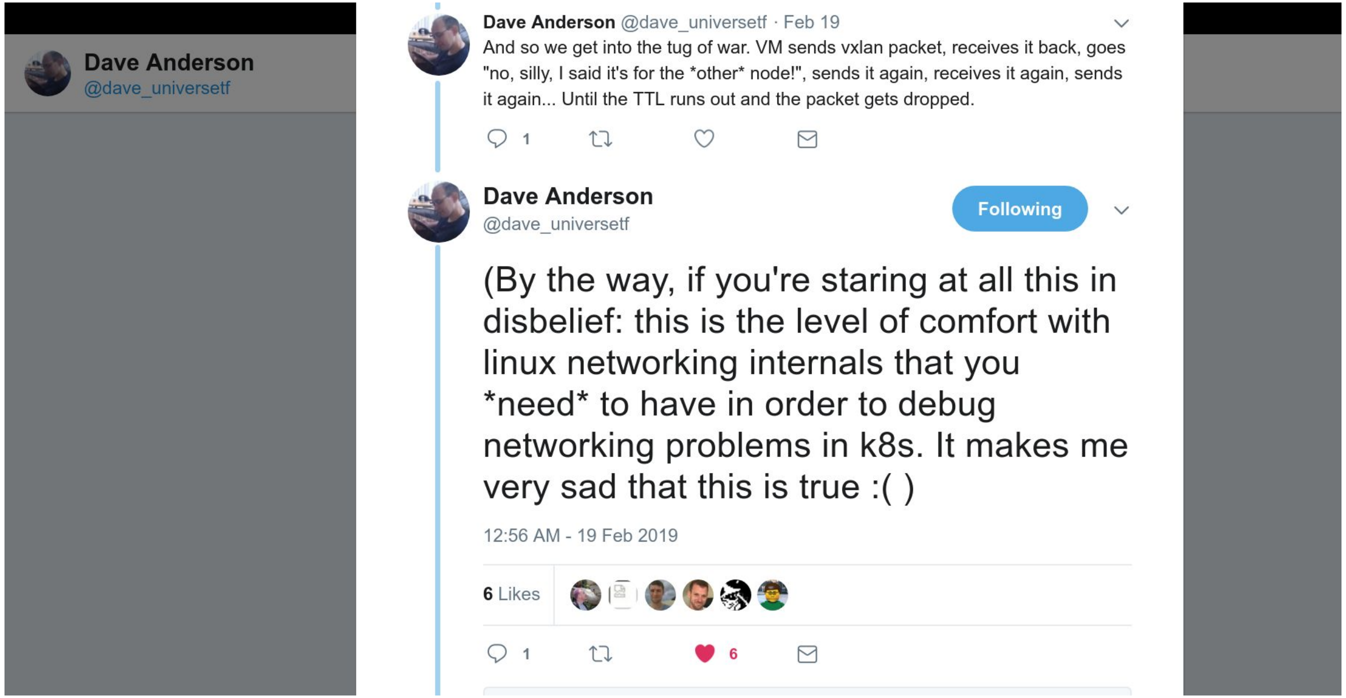
Container runtime

- Several options
  - Docker
  - rkt
  - CRI-O
  - frakti
- **Docker Engine is hard to avoid**

Networking

- Several options available
    - **Calico**
    - Flannel
- 2018 Public Cloud Performance Benchmark Report
    - https://www.thousandeyes.com/press-releases/2018-public-cloud-performance-benchmark-report

Networking



Dave Anderson
@dave_universetf

Dave Anderson @dave_universetf · Feb 19
And so we get into the tug of war. VM sends vxlan packet, receives it back, goes "no, silly, I said it's for the *other* node!", sends it again, receives it again, sends it again... Until the TTL runs out and the packet gets dropped.

1

Dave Anderson
@dave_universetf

(By the way, if you're staring at all this in disbelief: this is the level of comfort with linux networking internals that you *need* to have in order to debug networking problems in k8s. It makes me very sad that this is true :( )

12:56 AM - 19 Feb 2019

6 Likes

1 6

- Dave Anderson (@dave_universetf)

Monitoring

- Some great options here
  - Prometheus
  - TICK stack
- Absolutely crucial if you want to run a cluster by yourself
- **Prometheus + Alertmanager + Grafana + various exporters**

Monitoring

- Prometheus collects the data and stores it
  - Also provides a basic UI for your PromQL queries
- Alertmanager can route alerts
  - Also does deduplication and silencing of alerts
  - Supports various ways of sending out alerts
    - Slack, Opsgenie, PagerDuty, email
- **Try to figure out your monitoring high availability story early on**

Deployment

- **Helm**
- Roll out YAML by hand
- Terraform
  - Has a Kubernetes provider
- Pulumi

Container registries

- Nexus
- Artifactory
- Something that can handle both Helm charts and Docker images?
  - VMware Harbor
- **Handling images/charts is important**

Service mesh

- Service mesh brings control, security and observability of services, API calls and traffic for your Kubernetes clusters
- **Istio**

Resource limits

- Use limits and requests
  - CPU
  - Memory
- **Use reasonable values**
- Take care of your storage
- **Plan for failure**
  - OOM killer

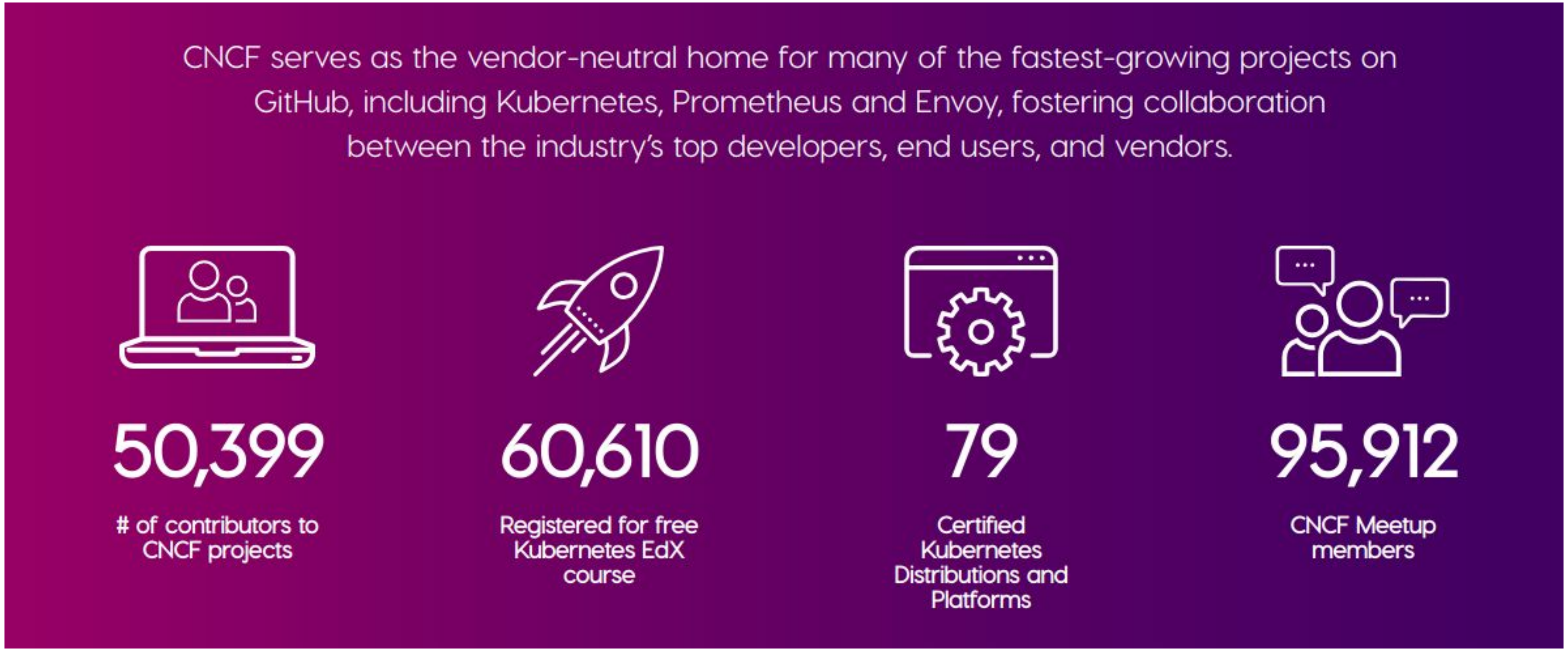
Upgrade story

- From the OS perspective
  - New OS versions
  - Security patching
  - Switching to another OS
- From the Kubernetes perspective
  - Kubernetes version upgrades

Software ecosystem

- **There’s a lot of stuff out there**
- Heptio (now VMware) things (Ark/Velero, Sonobuoy, Gimbal)
- kube-hunter
- kube-bench (CIS benchmark)
- Possible OpenStack ecosystem parallels in the future?

Software ecosystem


CNCF serves as the vendor-neutral home for many of the fastest-growing projects on GitHub, including Kubernetes, Prometheus and Envoy, fostering collaboration between the industry's top developers, end users, and vendors.

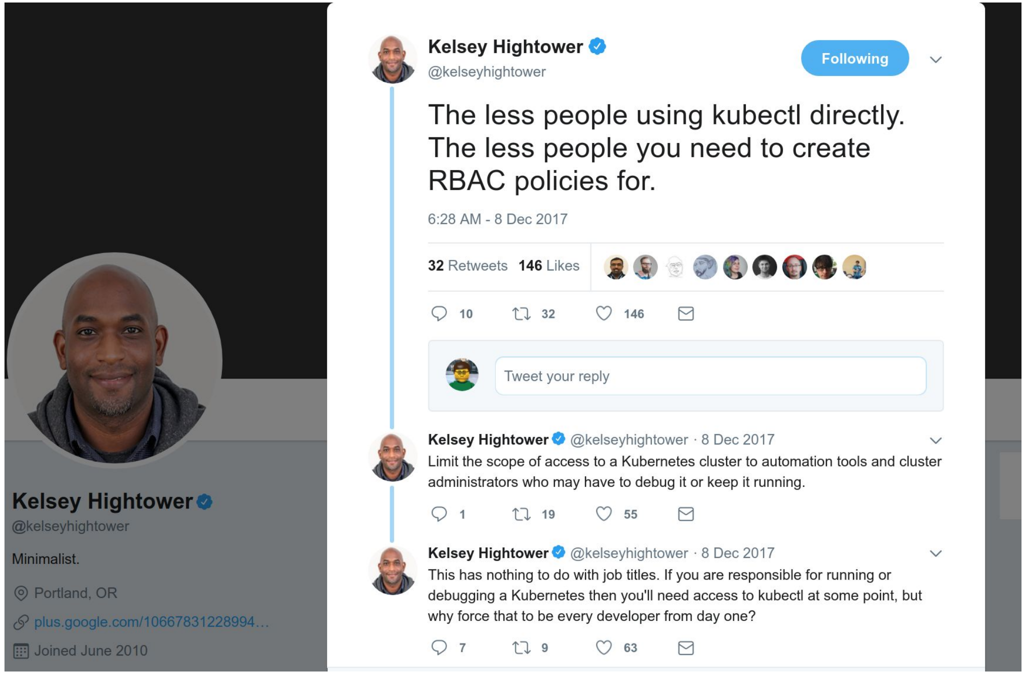
50,399	60,610	79	95,912
# of contributors to CNCF projects	Registered for free Kubernetes EdX course	Certified Kubernetes Distributions and Platforms	CNCF Meetup members

- CNCF (https://www.cncf.io)

RBAC

- If you don’t implement this you’re going to have a bad time
- Really hard to get right for your organization
- **Plan time to do this properly**
    - It can impact the processes

RBAC



Kelsey Hightower
@kelseyhightower

Following

The less people using kubectl directly. The less people you need to create RBAC policies for.

6:28 AM - 8 Dec 2017

32 Retweets 146 Likes

10 32 146

Tweet your reply

Kelsey Hightower @kelseyhightower · 8 Dec 2017
Limit the scope of access to a Kubernetes cluster to automation tools and cluster administrators who may have to debug it or keep it running.

1 19 55

Kelsey Hightower @kelseyhightower · 8 Dec 2017
This has nothing to do with job titles. If you are responsible for running or debugging a Kubernetes then you'll need access to kubectl at some point, but why force that to be every developer from day one?

7 9 63

Kelsey Hightower
@kelseyhightower

Minimalist.

Portland, OR

plus.google.com/10667831228994…

Joined June 2010

- Kelsey Hightower (@kelseyhightower)

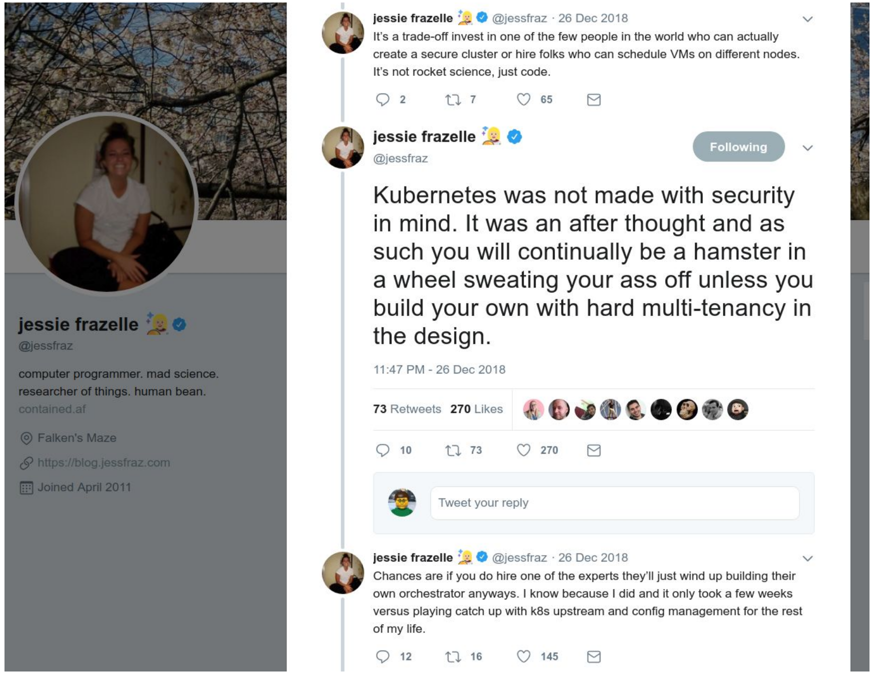
Cluster validation

- **Both when deploying and while it is in production**
- Serverspec/InSpec tests
- Running e2e tests from Kubernetes
- Sonobuoy
- OpenSCAP
- You may also want to think about performance testing

Security

- There are some vendors in the ecosystem
  - Aqua Security
  - NeuVector
  - Twistlock
  - Isovalent (Cilium)
- **Auditing**

Security



jessie frazelle
@jessfraz

computer programmer. mad science. researcher of things. human bean.
contained.af

Falken's Maze
https://blog.jessfraz.com
Joined April 2011

jessie frazelle @jessfraz · 26 Dec 2018
It's a trade-off invest in one of the few people in the world who can actually create a secure cluster or hire folks who can schedule VMs on different nodes. It's not rocket science, just code.
2 7 65

jessie frazelle @jessfraz
Following

Kubernetes was not made with security in mind. It was an after thought and as such you will continually be a hamster in a wheel sweating your ass off unless you build your own with hard multi-tenancy in the design.

11:47 PM - 26 Dec 2018

73 Retweets 270 Likes

10 73 270

Tweet your reply

jessie frazelle @jessfraz · 26 Dec 2018
Chances are if you do hire one of the experts they'll just wind up building their own orchestrator anyways. I know because I did and it only took a few weeks versus playing catch up with k8s upstream and config management for the rest of my life.
12 16 145

- Jessie Frazelle (@jessfraz)

Conclusion

- Kubernetes is good
- The ecosystem is rich
  - Vendors
  - Open source projects that are available
- **You need to make sure you're not just deploying to Kubernetes**
  - Make sure you have a complete infrastructure that you can leverage in day-to-day operations

Thank you!

Questions?