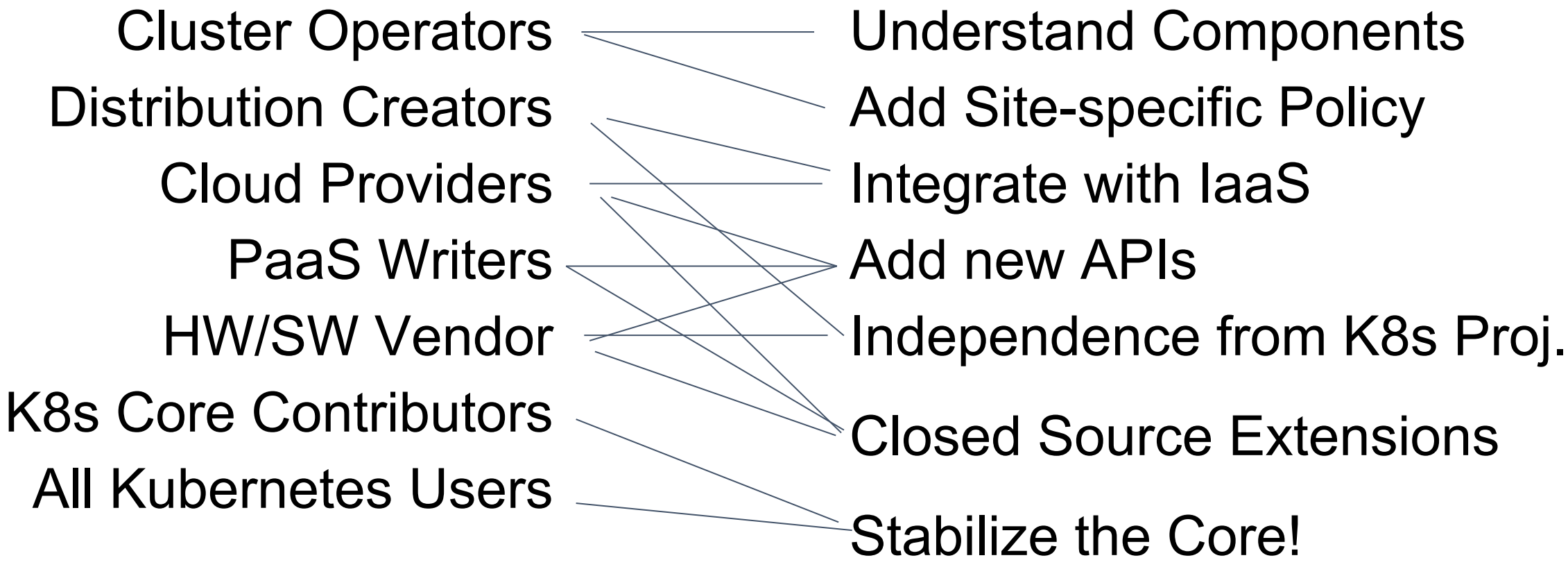
KubeCon

North America 2017

Extensibility

Daniel Smith – Staff Software Engineer, *Google*
Eric Tune – Sr. Staff Software Engineer, *Google*

Who should care and why



Extensibility

Kubernetes is

...Open Source

...Automatable

...Extensible

but

...Forking is hard
- Fast
- Big
- Hosted

…
- Asynchronous
- Cannot add APIs
- Cannot change APIs

… So *many* ways to extend

Kubernetes is...

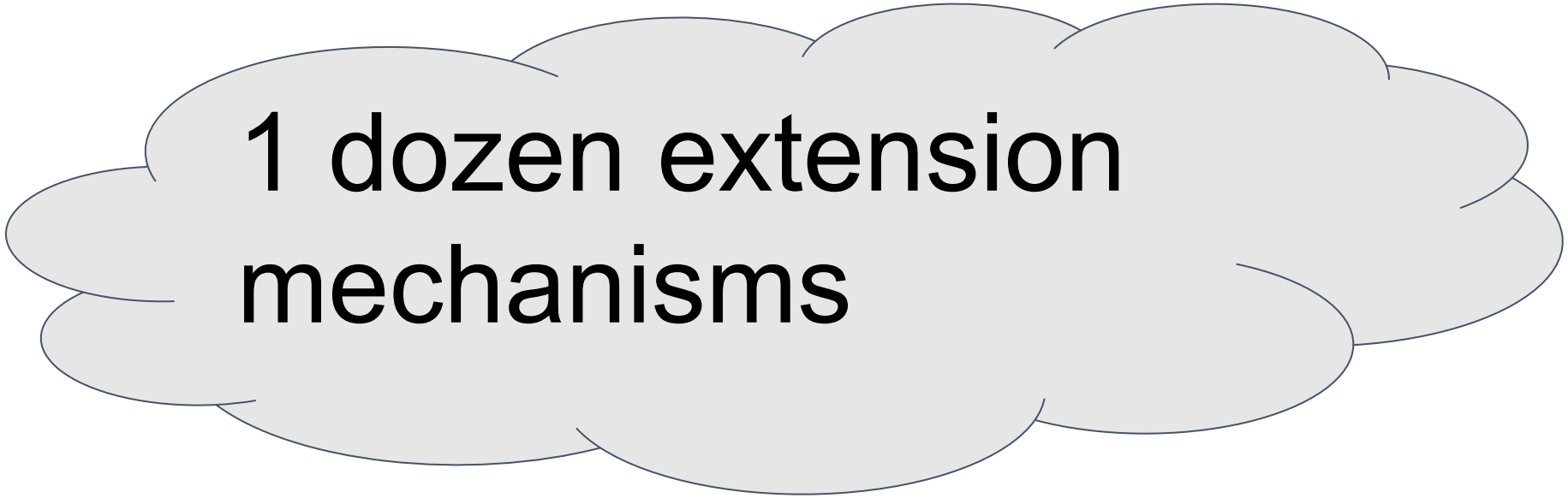
...an abstraction over infrastructure.

...a framework for declarative APIs and distributed control.

Kubernetes is...

...an abstraction over infrastructure.

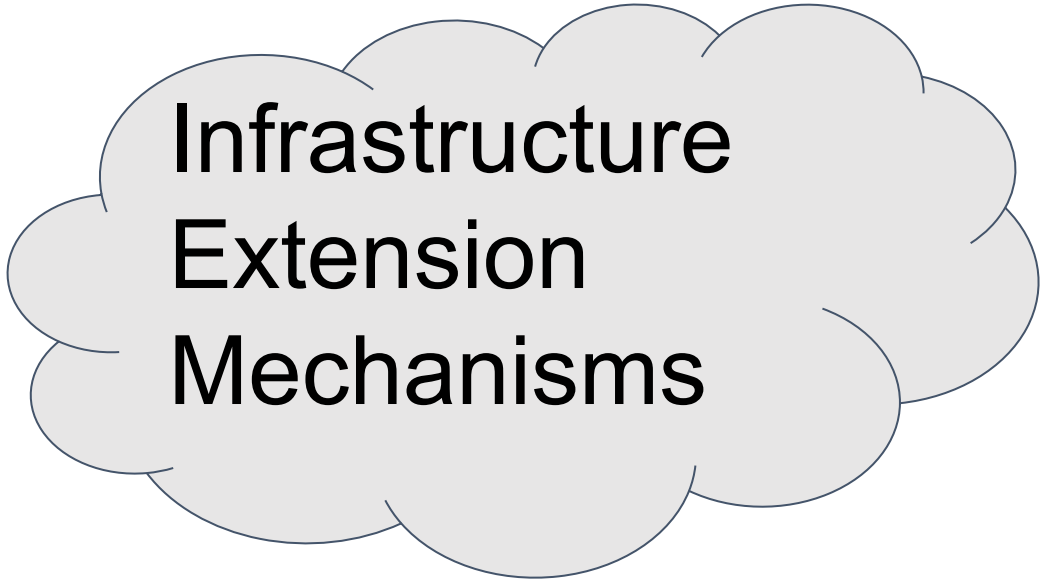
...a framework for declarative APIs and distributed control.

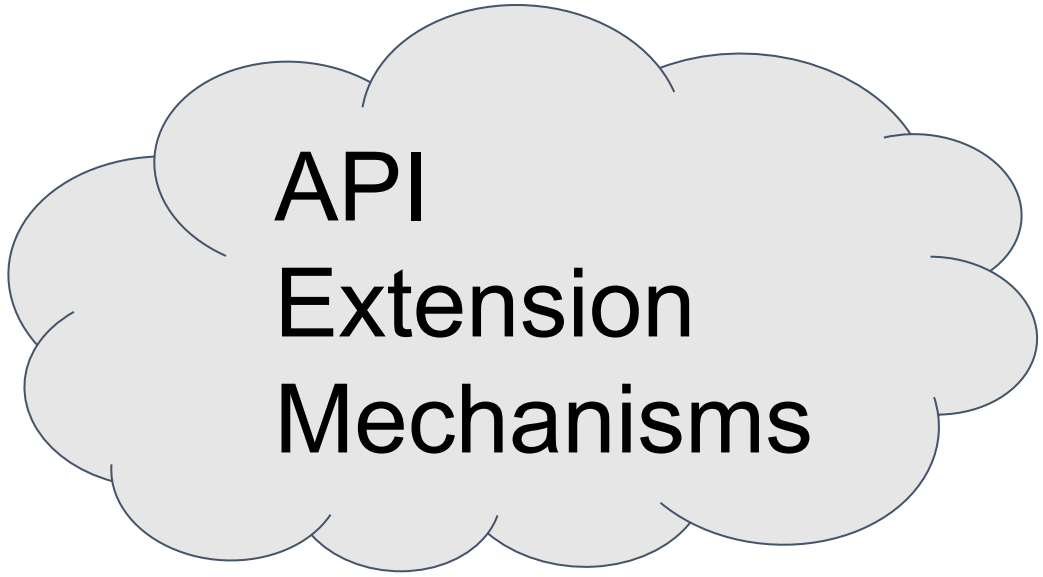


Kubernetes is...

...a abstraction over infrastructure.

...a framework for declarative APIs and distributed control.





Ready to hit the slopes?

Easiest and Stable

Moderately Hard and/or Less Stable

More Coding, Less Stable

Likely to crash and break a ~~leg~~ cluster

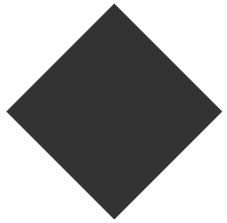
Infrastructure Extensibility

https://goo.gl/2qz8jW

- Storage ExtensionS – Allow new kinds of Volumes for Pods
  - Flex Volumes
    - easiest to write: binary plugin, bash scripts
    - expect it to stick around but not get better
  - CSI "Container Storage Interface"
    - Open: Docker, Kubernetes, Mesos, etc
    - easier to deploy/upgrade on top of K8s.
    - expect it to stick around to grow
    - alpha in 1.9

Infrastructure Extensibility

https://goo.gl/2qz8jW



- Cloud Controller Manager
  - "Cloud provider" now a separate binary.
    - Manages instance lifecycle, service IPs, load balancing, etc
  - Support your own cloud without forking the code, even use private code.
  - expected Alpha in 1.10.
  - 2018: expecting beta and shifting most/all providers out of main release.

Infrastructure Extensibility

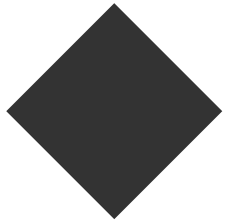
https://goo.gl/2qz8jW

- Device Plugins
  - Add discrete hardware resources such as:
    - "GPUs"
    - "FPGAs"
    - "QRNGs"
  - Name and number of devices Reported by kubelet on Node object
  - Considered by the scheduler
  - Kubelet Allocates # to a pod.
  - Alpha in v1.8

Infrastructure Extensibility

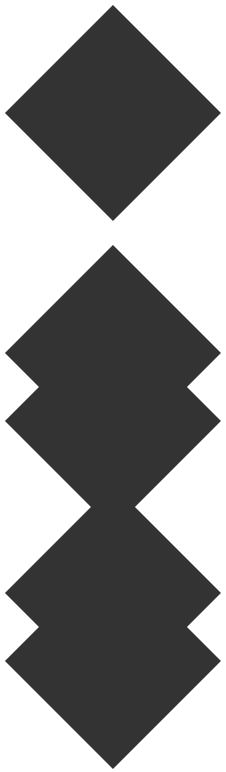
https://goo.gl/2qz8jW

- Network Plugins
  - Connect Pods to the network
  - Support for them is alpha in K8s
  - Open standard: CNI
    - two dozen or so available (not all support K8s hostport)
    - they work on 6 or so orchestrators



Infrastructure Extensibility

https://goo.gl/2qz8jW



Replace the Scheduler

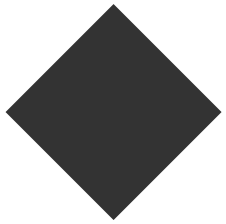
Multiple Schedulers

- Scheduler Extender
  - Plugin model

Infrastructure Extensibility

https://goo.gl/2qz8jW

- Secrets
  - Encryption at Rest – alpha in 1.9
  - Store the KEKs in a Key Store
    - e.g. Vault, Google KMS, Azure KMS, etc
    - Alpha planned for 1.10
    - Expected to GA by end 2018.

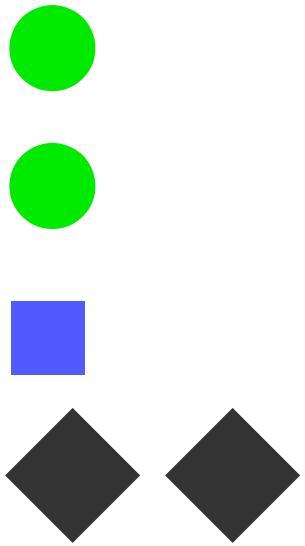


API Extensibility

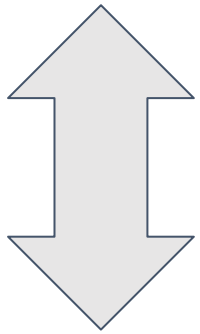
goo.gl/AJf3PU

A spectrum of API Extensions

goo.gl/AJf3PU



- CRD
- CRD + schema
- CRD + schema + validation webhook
- Aggregated apiserver



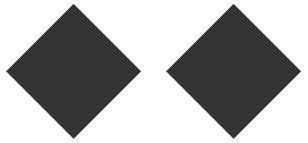
How to choose: goo.gl/zb2ssj

Custom Resources

goo.gl/AJf3PU

- CRDs to GA in 2018
- Completeness
  - Schema
  - Validation Webhooks
  - ClusterRole
- API-wide Consistency
  - Sub-resources (/status, /scale)

Aggregated API Servers



- Customize all the things!
  - Storage backend
    - Time-series data: metrics apiserver
  - Admission chain / business logic
  - Version conversion
- Who?
  - Kubernetes Developers

Admission Extensions
1.9 and beyond

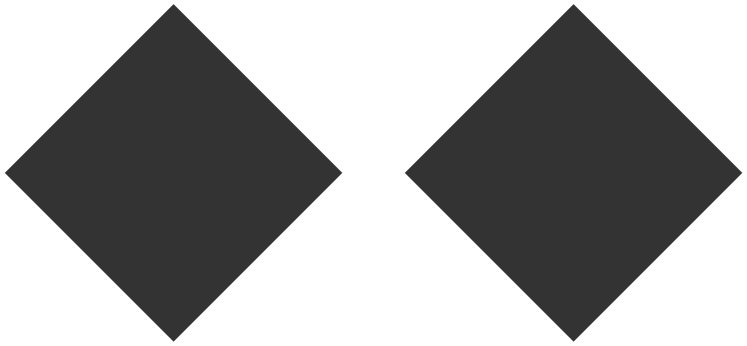
- What is the “admission stack”?

Admission Extensions
1.9 and beyond

- What is the “admission stack”?
- Everything on the request path...
  - ...that’s after the permissions check
  - ...and before the final storage operation.
- Ideal place for policy enforcement.

Admission Extensions
1.9 and beyond

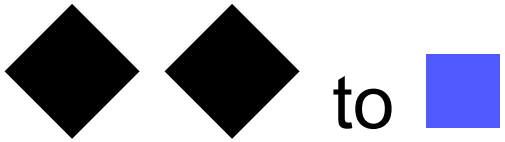
- Problem: admission plugins are all compiled-in.



Admission Extensions
1.9 and beyond

- Admission webhooks!
  - Beta in 1.9
  - GA sometime in 2018
  - Dynamic configuration
- Initializers!
  - Alpha

Future API extension work...

- kube-apiserver
  - flags to config files
  - config files to APIs
- Permissions (“authz”) webhook:  to
- End goal: fully portable extensions!

Combining Extensions Mechanisms

CRD + Control Loop	= etcd-operator
CRD + Control Loop + Volume Plugin	= Rook
CRD + Control Loop + Network Plugin	= Calico Canal
CRD + Validating Admission Webhook	= *better validation*

2018 Aspirations for API

- Automatic Rich CLI/GUI for Custom Resources
  - App Definition
  - Show status and children
- Scale for Custom Resources
  - Use HPA and PDB with Operators
- Version Conversion for CRDs
- Cluster introspection API
  - Garbage Collector has a Resource Graph

Conclusion

- Commitment to making and keeping Kubernetes extensible
  - Stabilizing extension mechanisms
  - Improving documentations
  - Using existing open standards where suitable
  - Offer multiple choices with graded difficulty where needed