

DOK DAY

NORTH AMERICA

DOK DAY
NORTH AMERICA

The State of Stateful on Kubernetes

Michelle Au, Google
Kaslin Fields, Google

Intros

DOK DAY
NORTH AMERICA

November 6, 2023 | Chicago, Illinois



Kaslin Fields
GKE & OSS K8s
Developer Advocate
Google



Michelle Au
GKE & OSS K8s
Software Engineer
Google

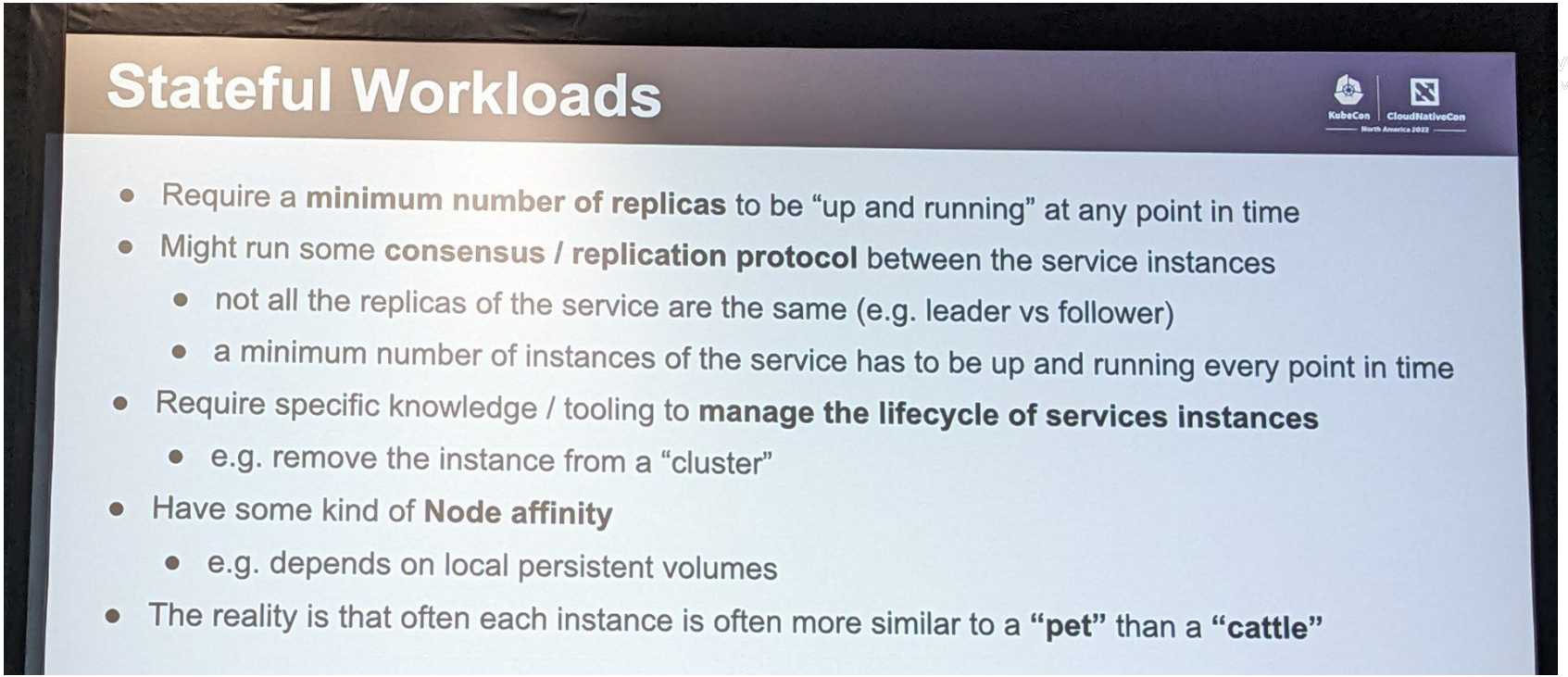
Stateful

Everything Has State

The difference is whether anyone cares or not.

Stateful Workloads in Kubernetes

DOK DAY
NORTH AMERICA



Andrea Tosatto
Kubernetes Contributor Summit NA 2022

Categorizing Workloads in Kubernetes

DOK DAY
NORTH AMERICA
KnativeCon
NORTH AMERICA

- Deployments
  - Long-running workloads, state is shared across replicas
- DaemonSets
  - Workloads that run on each node in the cluster
- Jobs
  - A workload that needs to run to completion
- CronJobs
  - Workloads that need to run to completion on a time-based schedule
- StatefulSets
  - Volume per replica, more sticky/persistent identity

StatefulSet

Manages the deployment and scaling of a set of Pods, *and provides guarantees about the ordering and uniqueness* of these Pods.

Unlike a Deployment, a StatefulSet maintains a **sticky identity** for each of its Pods.

Useful for workloads that require:

- Stable, unique network identifiers.
- Stable, persistent storage.
- Ordered, graceful deployment and scaling.
- Ordered, automated rolling updates.



KnativeCon
NORTH AMERICA

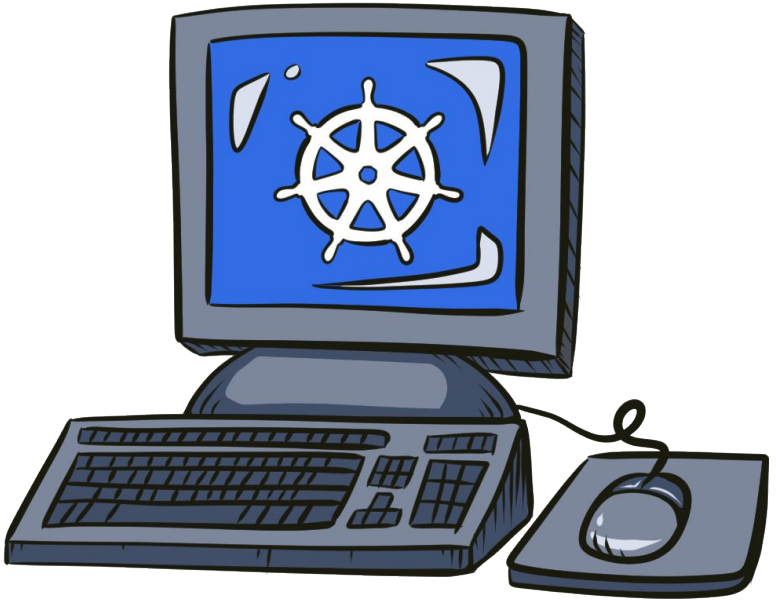
What kinds of workloads count as stateful?

What kinds of workloads count as stateful?

DOK DAY
NORTH AMERICA

- Pre-container style architectures
  - Wordpress (Usually *Deployment*)
- Game Servers
  - https://github.com/saulmaldonado/agones-minecraft (*CRD*)
- Things that deal intricately with data
  - Databases (Usually *StatefulSet/CRD*)
- AI/ML
  - Training datasets, models, checkpoints (Usually *Jobs*)

KnativeCon
NORTH AMERICA



What are the challenges stateful workloads face?

DOK DAY
NORTH AMERICA

KnativeCon
NORTH AMERICA

- Maintaining a consistent identity
  - Often for connection to other services
- High & Consistent Availability
  - Upgrades must be handled gracefully and carefully
  - *This* needs to be up and ready before *that*
  - Stateful workloads often have complex start and end processes

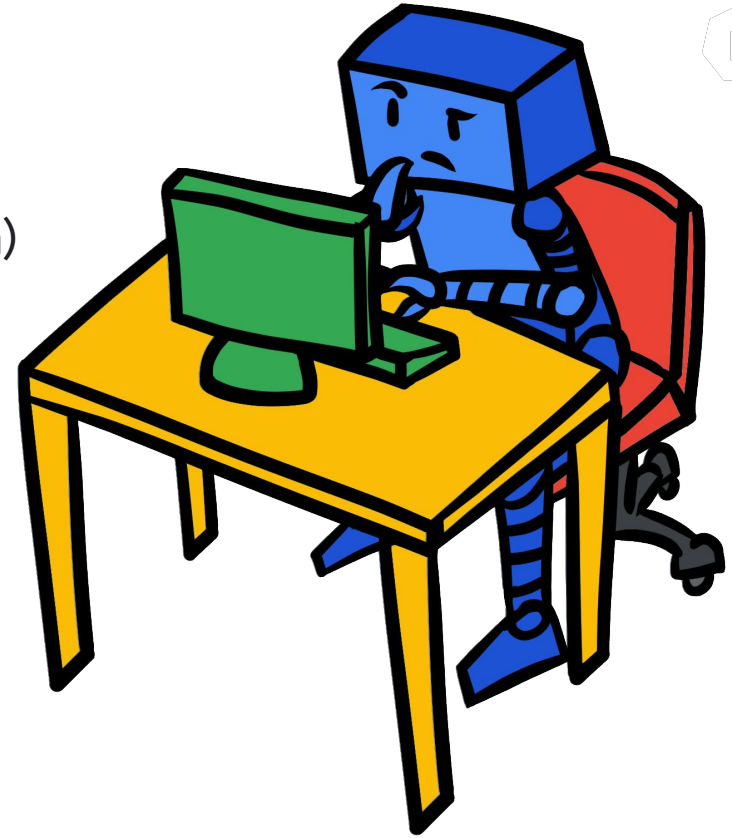
How does Kubernetes help?

What are we doing to address the challenges of Stateful workloads?

DOK DAY
NORTH AMERICA

Lifecycle and Day 2 Management

- StatefulSet
  - Ie. PVC deletion policies (beta)
- Custom Resources
  - Custom Resource Definitions
  - Operators (How Kubernetes runs CRDs)



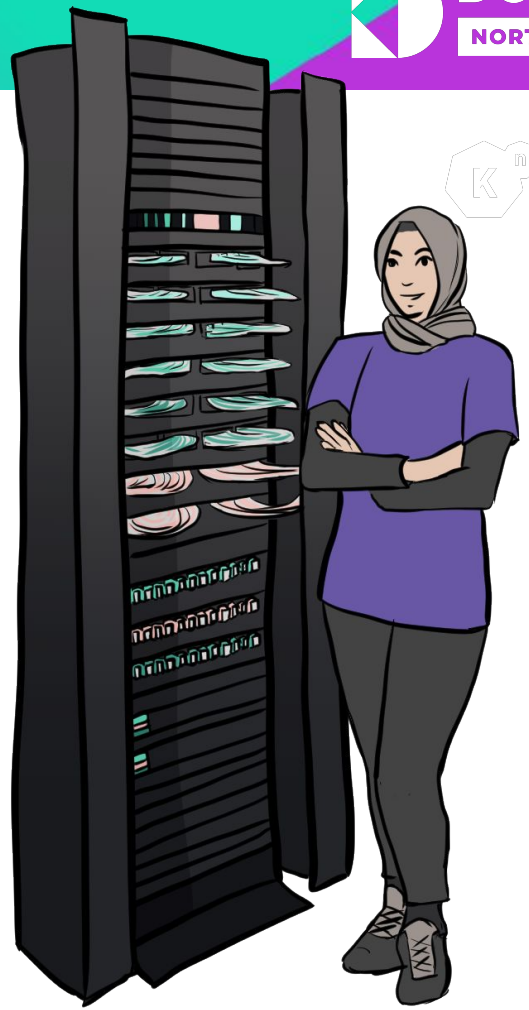
KnativeCon
NORTH AMERICA

What are we doing to address the challenges of Stateful workloads?

DOK DAY
NORTH AMERICA

[Persistent Volumes]

- Container Storage Interface (CSI) Ecosystem
  - Over [100 drivers!] (Out of tree!)
- Dynamic provisioning, resizing
- Snapshots, cloning, custom data sources (beta)



KnativeCon
NORTH AMERICA

Addressing challenges cont'd: Upgrades & Disruption

DOK DAY
NORTH AMERICA

- Fault tolerance
  - Pod topology spreading
- Workload isolation for critical workloads
  - Node Affinity, Taints/Tolerations
  - Pod Priority and Preemption
  - Pod Resources and QoS

KnativeCon
NORTH AMERICA



Addressing challenges cont'd: Upgrades & Disruption

DOK DAY
NORTH AMERICA

- Managing Pod eviction
  - Pod Disruption Budgets
  - Pod readiness probes
  - Graceful termination, pre-stop hooks
- Not doing upgrades is not an option! DO YOUR UPGRADES!



Future / Upcoming k8s and DoK features

DOK DAY NORTH AMERICA

KnativeCon NORTH AMERICA

k8s 1.29 alpha features:

Modify volumes - use cases like updating IOPS/throughput

Beyond:

STS volume expansion

Group volume snapshots

Cross-namespace snapshots (and other data sources)

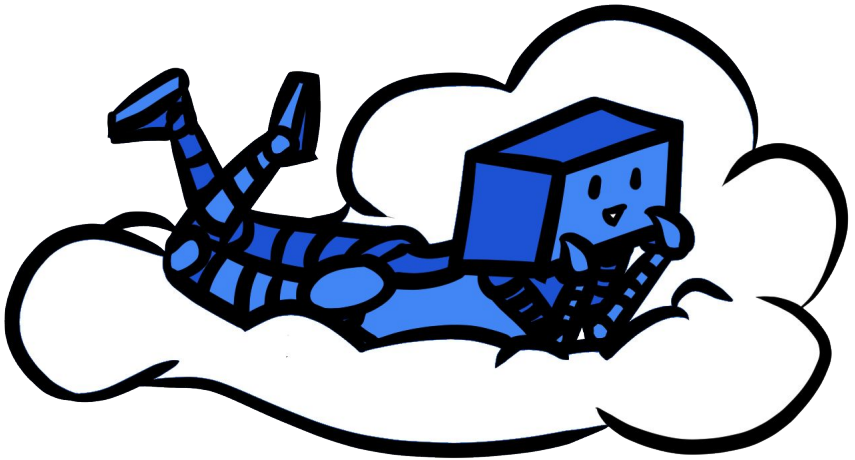
Declarative node maintenance

Topology-aware disruptions

DoK community developments:

Operator feature matrix

Security hardening guide

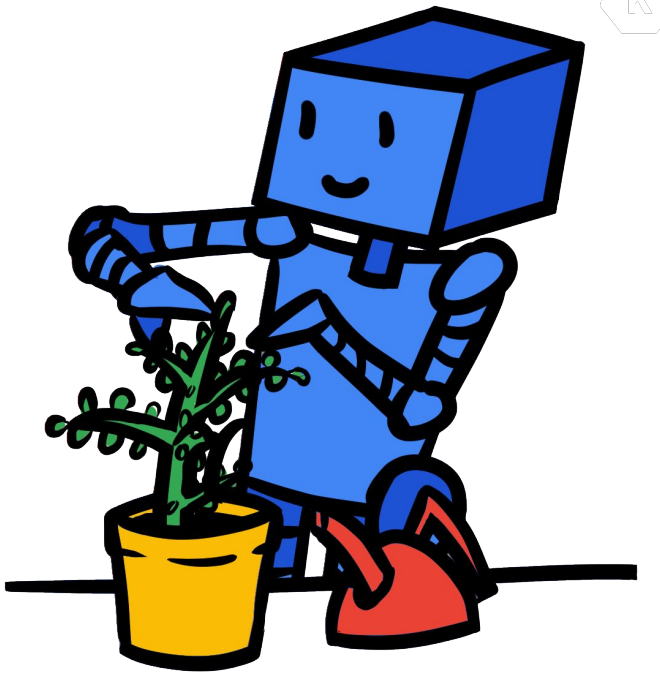


Best Practices

Best Practices for Stateful Workloads on Kubernetes

- Use the aforementioned features!
- Blue/green strategies for upgrades
- Chaos testing
- Take regular backups
  - Backups of the data
  - Backups of the config
- Actually test your recovery procedures!
- CI/CD best practices apply
- General Kubernetes best practices around security and networking apply

KnativeCon
NORTH AMERICA



Key Takeaways

DOK DAY
NORTH AMERICA

KnativeCon
NORTH AMERICA

Stateful is more than just databases

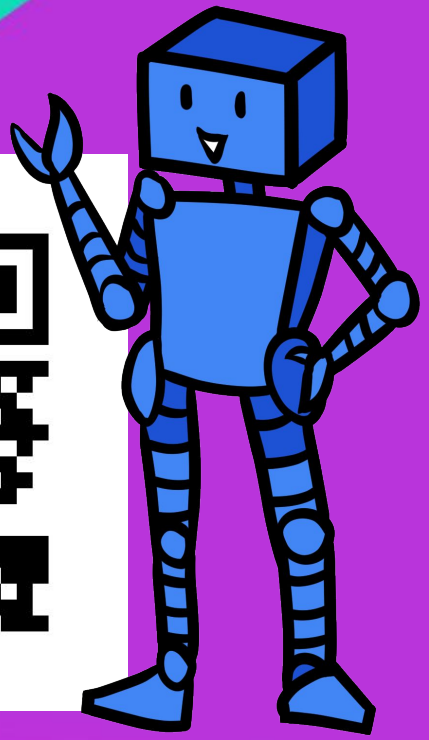
Kubernetes sees a workload as stateful if something cares about its state in some form (not just data!)

Kubernetes provides primitives for app lifecycle, storage, scheduling, and graceful disruption management. Look for these types of features for your stateful needs!

A good quality operator can simplify and manage complex day 2 workflows

Design your application with modern best practices

Thanks!
Q&A



DOK DAY
NORTH AMERICA

November 6, 2023 | Chicago, Illinois