
KubeCon
CloudNativeCon
North America 2022
BUILDING FOR THE ROAD AHEAD
DETROIT 2022

KubeCon CloudNativeCon North America 2022

BUILDING FOR THE ROAD AHEAD

DETROIT 2022

SIG Multicluster Intro & Deep Dive

KubeCon CloudNativeCon North America 2022

BUILDING FOR THE ROAD AHEAD

DETROIT 2022

October 24-28, 2021

Laura Lorenz
Google

Paul Morie
Apple

Jeremy Olmsted-Thompson
Google

BUILDING FOR THE ROAD AHEAD

DETROIT 2022

SIG Multicluster Intro & Deep Dive

Laura Lorenz (lauralorenz)
Paul Morie (pmorie)
Jeremy Olmsted-Thompson (jeremyot)

Welcome!

We'll cover:

- What this SIG is about
- Current activity
  - Kubefed Deprecation
  - ClusterSet / Namespace Sameness
  - About API for storing cluster properties such as Cluster ID / ClusterSet membership
  - Multicluster Services API / Multicluster DNS
  - StatefulSetSlices for migrating stateful sets between clusters
- Deep dive and Demo
  - About API and MCS API in action using AWS CloudMap MCS Controller for K8s
- How to contribute

What this SIG is about

- What should be the Kubernetes-native way to ….
  - Expose workloads from multiple clusters to each other
  - Share cluster metadata and its place relative to others
  - Generally break down the walls between clusters
- Touches many different functional areas, but we are still working to identify the best, most durable primitives
- We want AND need your input!
  - Real user stories and use cases are extremely valuable
  - Tell us what you're working on!

Our approach

- Avoid premature standardization
- Focus on APIs
- Avoid solving any optional problems
- Keep multicluster consistent with single cluster
- Work backwards from specific problems into something bigger, maybe

Current activity

Kubefed - on archival decision

- Kubefed will be archived
- Great learnings came from this project but is no longer under active development
- Archival is not deletion
- Thank you to everyone who contributed!

ClusterSet

- ClusterSet represents a pattern of use from the field:
  - A group of clusters governed by a single authority
  - High degree of trust within the set
  - Namespace Sameness applies to clusters in the set
    - Permissions and characteristics are consistent across clusters for a given namespace
    - Namespaces don't have to exist in every cluster, but behave the same across those in which they do
- A cluster's ClusterSet membership is stored in the about.k8s.io/ClusterProperty `clusterset.k8s.io`
  - See next slide …

About API: cluster metadata

- KEP-2149
- Now available at sigs.k8s.io/about-api
- Cluster scoped ClusterProperty CRD - name: value
- Uniquely identify clusters and identify their membership in a ClusterSet, for the lifetime of membership
- Provides a reference for multi-cluster tooling to build on within a cluster set (e.g. valid DNS label)
- Now a well-known place to store these or any other cluster properties that might otherwise be ad-hoc annotations on semantically adjacent objects

```
apiVersion: about.k8s.io/v1
kind: ClusterProperty
metadata:
  name: cluster.clusterset.k8s.io
spec:
   value: cluster-1
```

```
apiVersion: about.k8s.io/v1
kind: ClusterProperty
metadata:
  name: clusterset.k8s.io
spec:
   value: mycoolclusterset
```

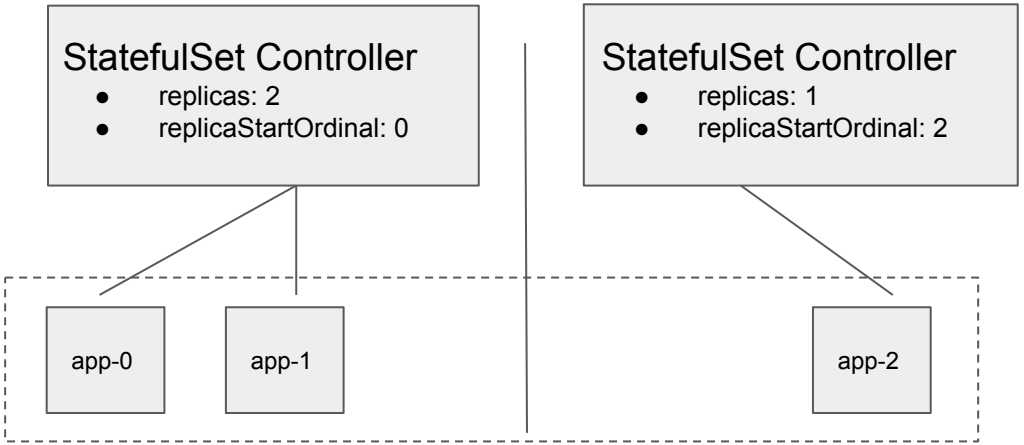
```
apiVersion: about.k8s.io/v1
kind: ClusterProperty
metadata:
  name: fingerprint.mycoolimplementation.com
spec:
  value:  '{"major": "1","minor":
"18","gitVersion": "v1.18.2","gitCommit":
"52c56ce7a8272c798dbc29846288d7cd9fbae032","git
TreeState": "clean","buildDate":
"2020-04-30T20:19:45Z","goVersion":
"go1.13.9","compiler": "gc","platform":
"linux/amd64"}'
```

MC Services API

- KEP-1645 and sigs.k8s.io/mcs-api
- Services are a multi-cluster building block
- Allows a single service to span and/or be consumed by multiple clusters
- Focused only on the API and common behavior, leaving room for various implementations
  - Submariner, GKE, Istio, AWS
- Consumers only ever rely on local data
- ClusterIP and headless services just work as expected across clusters.

```yaml
apiVersion: v1
kind: Service
metadata:
  name: foo
  namespace: bar
spec:
  ports:
  - port: 80
  selector:
    app: foo
---
apiVersion: multicluster.x-k8s.io/v1alpha1
kind: ServiceExport
metadata:
  name: foo
  namespace: bar
```

StatefulSetSlices



- KEP-3335, proposing a mechanism to support the “split brain” of individual StatefulSet controllers during migration of a StatefulSet across clusters
- Very cool POC leveraging both StatefulSetSlices and MCS
  - see @pwschuurman and @mattschallert showcasing this in their KubeconNA talk

Additional Topics

- More sophistication on MC networking
  - Network policy - applying policy uniformly across clusters
  - Multi-network - stitching together clusters on different networks
- Multicluster controllers / MC leader election
  - what SIG-MC should recommend or implement as a reference
- Work API
  - Spreading groups of resources to different cluster

KubeCon CloudNativeCon North America 2022
BUILDING FOR THE ROAD AHEAD
DETROIT 2022

About API and MCS API in action with AWS CloudMap MCS Controller for K8s

Major shoutout to @runakash and @astaticvoid

Get involved

We need your input

Share your use cases, problems, and ideas

- Home page: https://github.com/kubernetes/community/tree/master/sig-multicluster
- Slack channel: https://kubernetes.slack.com/messages/sig-multicluster
- List: https://groups.google.com/forum/#!forum/kubernetes-sig-multicluster
- Meetings are biweekly Tuesdays, 12:30 eastern, 9:30 pacific, 16:30 UTC

KubeCon | CloudNativeCon
North America 2022

BUILDING FOR THE ROAD AHEAD

DETROIT 2022

Please scan the QR Code above
to leave feedback on this session