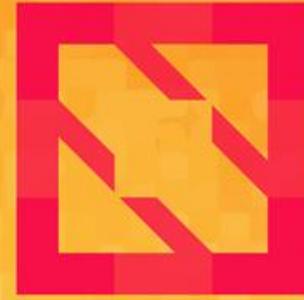




KubeCon



CloudNativeCon

North America 2019





KubeCon



CloudNativeCon

North America 2019

KubeDirector - Deploying Complex Stateful Applications on Kubernetes

Joel Baxter – Distinguished Engineer/HPE

Tom Phelan – Fellow/HPE



Agenda



KubeCon



CloudNativeCon

North America 2019

Agenda



KubeCon



CloudNativeCon

North America 2019

- Complex Stateful Applications on Kubernetes

Agenda



KubeCon



CloudNativeCon

North America 2019

- Complex Stateful Applications on Kubernetes
- Existing Kubernetes Tooling for Complex Stateful Applications

Agenda



KubeCon



CloudNativeCon

North America 2019

- Complex Stateful Applications on Kubernetes
- Existing Kubernetes Tooling for Complex Stateful Applications
- BlueK8s and KubeDirector

Agenda



KubeCon



CloudNativeCon

North America 2019

- Complex Stateful Applications on Kubernetes
- Existing Kubernetes Tooling for Complex Stateful Applications
- BlueK8s and KubeDirector
- KubeDirector Architecture

Agenda



KubeCon



CloudNativeCon

North America 2019

- Complex Stateful Applications on Kubernetes
- Existing Kubernetes Tooling for Complex Stateful Applications
- BlueK8s and KubeDirector
- KubeDirector Architecture
- KubeDirector Architecture Benefits

Agenda



KubeCon



CloudNativeCon

North America 2019

- Complex Stateful Applications on Kubernetes
- Existing Kubernetes Tooling for Complex Stateful Applications
- BlueK8s and KubeDirector
- KubeDirector Architecture
- KubeDirector Architecture Benefits
- Takeaways

Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based

Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next

Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration

Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications



KubeCon



CloudNativeCon

North America 2019

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



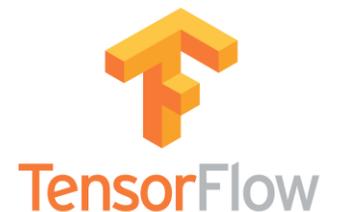
Complex Stateful Applications

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Complex Stateful Applications

- Complex
 - Distributed processing using multiple tools and services
 - Application architecture is not micro services based
- Stateful
 - Applications that save data from the activities of one session for use in the next
- Require persistence of “root” storage
 - Per application instance specific configuration
 - Per application node instance specific configuration



Existing K8S Stateful Support



KubeCon



CloudNativeCon

North America 2019

Existing K8S Stateful Support



KubeCon



CloudNativeCon

North America 2019

- Stateful Sets
 - Stable, unique network ids
 - Ordered deployment and scaling
 - Ordered automated updates

Existing K8S Stateful Support



KubeCon



CloudNativeCon

North America 2019

- Stateful Sets
 - Stable, unique network ids
 - Ordered deployment and scaling
 - Ordered automated updates
- Persistent Volumes
 - Life cycle of storage is independent from life cycle of container or pods

Existing K8S Stateful Support

- Stateful Sets
 - Stable, unique network ids
 - Ordered deployment and scaling
 - Ordered automated updates
- Persistent Volumes
 - Life cycle of storage is independent from life cycle of container or pods

Works great for applications that do not store their persistent data in /
(the root) file system

Existing K8S Complex Support



KubeCon



CloudNativeCon

North America 2019

Existing K8S Complex Support



KubeCon



CloudNativeCon

North America 2019

- Helm and helm charts, terraform, Kubeflow (ML focused)
 - Essentially package and deployment tools

Existing K8S Complex Support



KubeCon



CloudNativeCon

North America 2019

- Helm and helm charts, terraform, Kubeflow (ML focused)
 - Essentially package and deployment tools

No “server side” agent or daemon to respond to application/node events

Existing K8S Complex Support



- Helm and helm charts, terraform, Kubeflow (ML focused)
 - Essentially package and deployment tools

No “server side” agent or daemon to respond to application/node events

- Cleanup on exit for example

- Operators Framework and API
 - An operator needs to be written for each application

Existing K8S Complex Support



- Helm and helm charts, terraform, Kubeflow (ML focused)
 - Essentially package and deployment tools

No “server side” agent or daemon to respond to application/node events

- Cleanup on exit for example

- Operators Framework and API
 - An operator needs to be written for each application

“Day 2” challenges with user permissions when deploying new Custom Resources

BlueK8s and KubeDirector



KubeCon



CloudNativeCon

North America 2019

Source: www.bluedata.com/blog/2018/07/operation-stateful-bluek8s-and-kubernetes-director

BlueK8s and KubeDirector



KubeCon



CloudNativeCon

North America 2019

- BlueK8s is an Apache open source initiative focused on bringing enterprise support for complex stateful applications to Kubernetes

Source: www.bluedata.com/blog/2018/07/operation-stateful-bluek8s-and-kubernetes-director

BlueK8s and KubeDirector



KubeCon



CloudNativeCon

North America 2019

- BlueK8s is an Apache open source initiative focused on bringing enterprise support for complex stateful applications to Kubernetes
- A series of open source projects will be rolled out under the BlueK8s umbrella
 - The first major project is “KubeDirector”:

Source: www.bluedata.com/blog/2018/07/operation-stateful-bluek8s-and-kubernetes-director

BlueK8s and KubeDirector



KubeCon



CloudNativeCon

North America 2019

- BlueK8s is an Apache open source initiative focused on bringing enterprise support for complex stateful applications to Kubernetes
- A series of open source projects will be rolled out under the BlueK8s umbrella
 - The first major project is “KubeDirector”:

<https://github.com/bluek8s/kubedirector>

Source: www.bluedata.com/blog/2018/07/operation-stateful-bluek8s-and-kubernetes-director



KubeCon



CloudNativeCon

North America 2019

KubeDirector Architecture



A Non-App-Specific Operator



KubeCon



CloudNativeCon

North America 2019

- Model to support a *domain* of applications
- KubeDirectorApp CR: describe an "app type"
 - Images, service endpoints, directories to persist, etc.
 - Anything constant across app instances (some data-driven variation)
- KubeDirectorCluster CR: an "instantiation" of a KubeDirectorApp
- Lifecycle events drive config/maintenance of app instance
 - Events from CR create/edit; events from system state or metrics
 - E.g. create, resize, autoscale, node-down
 - Inside container: KD updates metadata file, optionally runs script hook

Example KubeDirectorApp 1/3



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: kubedirector.bluedata.io/v1alpha1
kind: KubeDirectorApp
metadata:
  name: cdh5142cm
spec:
  label:
    name: CDH 5.14.2 with Cloudera Manager
    description: blah blah more details
  version: "1.4"
  systemdRequired: true
```

```
defaultImageRepoTag: bluedata/cdh5142cm:3.0
defaultConfigPackage:
  packageURL: https://foo.com/cdhsetup.tgz
defaultPersistDirs:
- "/usr"
- "/opt"
- "/var"
- "/etc"
- "/data"
```

Example KubeDirectorApp 2/3



roles:

- cardinality: "1"
id: controller
- cardinality: "1+"
id: worker

services:

- id: cdh-scm-srvr
label:
name: Cloudera Manager
endpoint:
urlScheme: http
port: 7180
path: "/"
isDashboard: true
- id: ssh
endpoint:
port: 22
isDashboard: false
- *etc.*
- *etc.* (lots more services)

Example KubeDirectorApp 3/3



KubeCon



CloudNativeCon

North America 2019

config:

roleServices:

- roleID: controller

serviceIDs:

- cdh-scm-srvr

- cdh-scm-srvr-db

- cdh-scm-agent

- hdfs-dn

- ssh

- *etc.*

- *etc.* (lots more services)

- roleID: worker

serviceIDs:

- cdh-scm-agent

- yarn-nm

- ssh

- hdfs-dn

Example KubeDirectorCluster



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: kubedirector.bluedata.io/v1alpha1
kind: KubeDirectorCluster
metadata:
  name: cdh5142cm-persistent
spec:
  app: cdh5142cm
  roles:
  - id: controller
    resources:
      requests:
        memory: "16Gi"
        cpu: "2"
    storage:
      size: "200Gi"
```

```
- id: worker
  members: 10
  resources:
    requests:
      memory: "8Gi"
      cpu: "2"
  storage:
    size: "100Gi"
```

KubeDirectorApp Deploy



KubeCon



CloudNativeCon

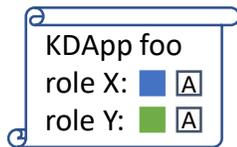
North America 2019

K8s Resources

container images

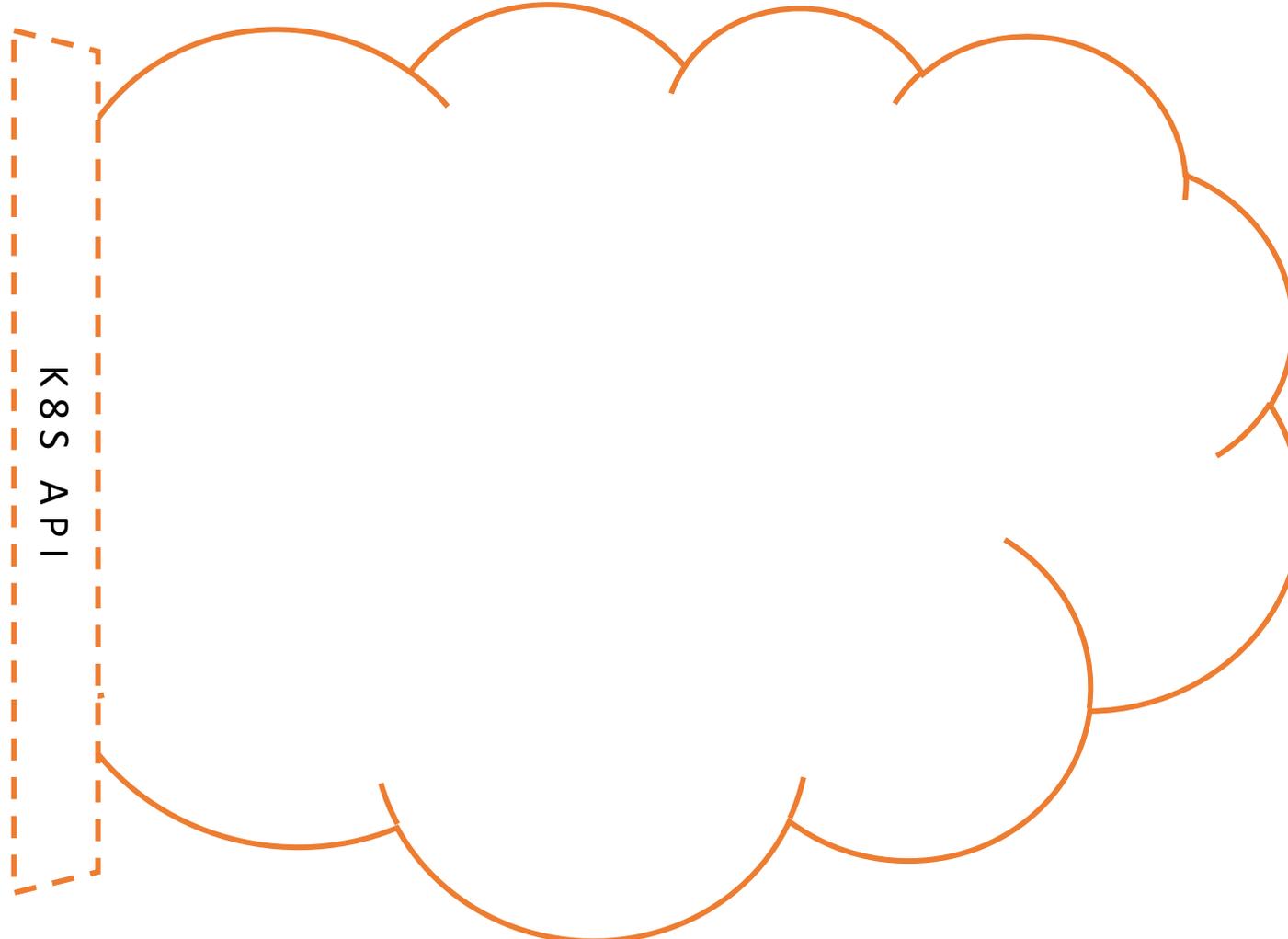


app config artifacts



K8S API

KubeDirector



KubeDirectorApp Deploy

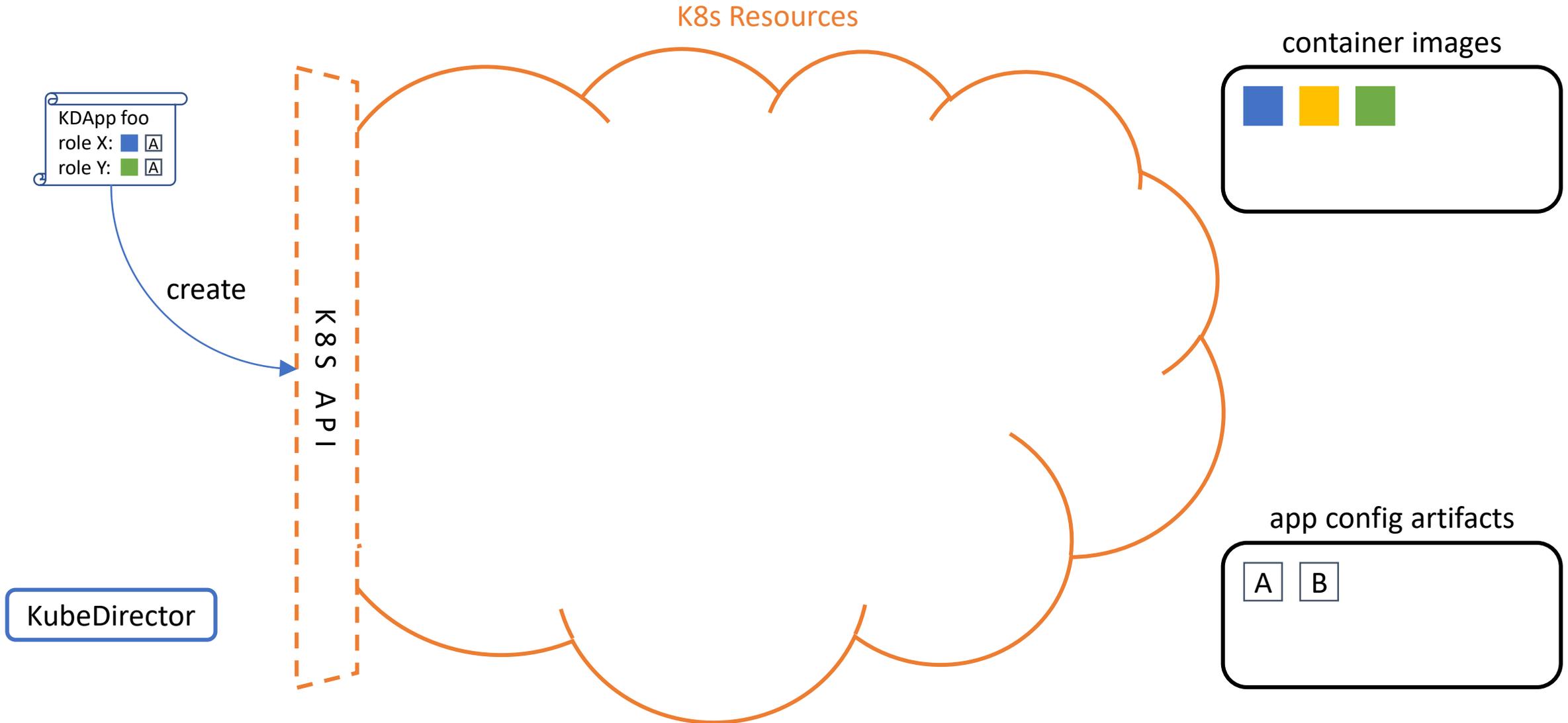


KubeCon



CloudNativeCon

North America 2019



KubeDirectorApp Deploy

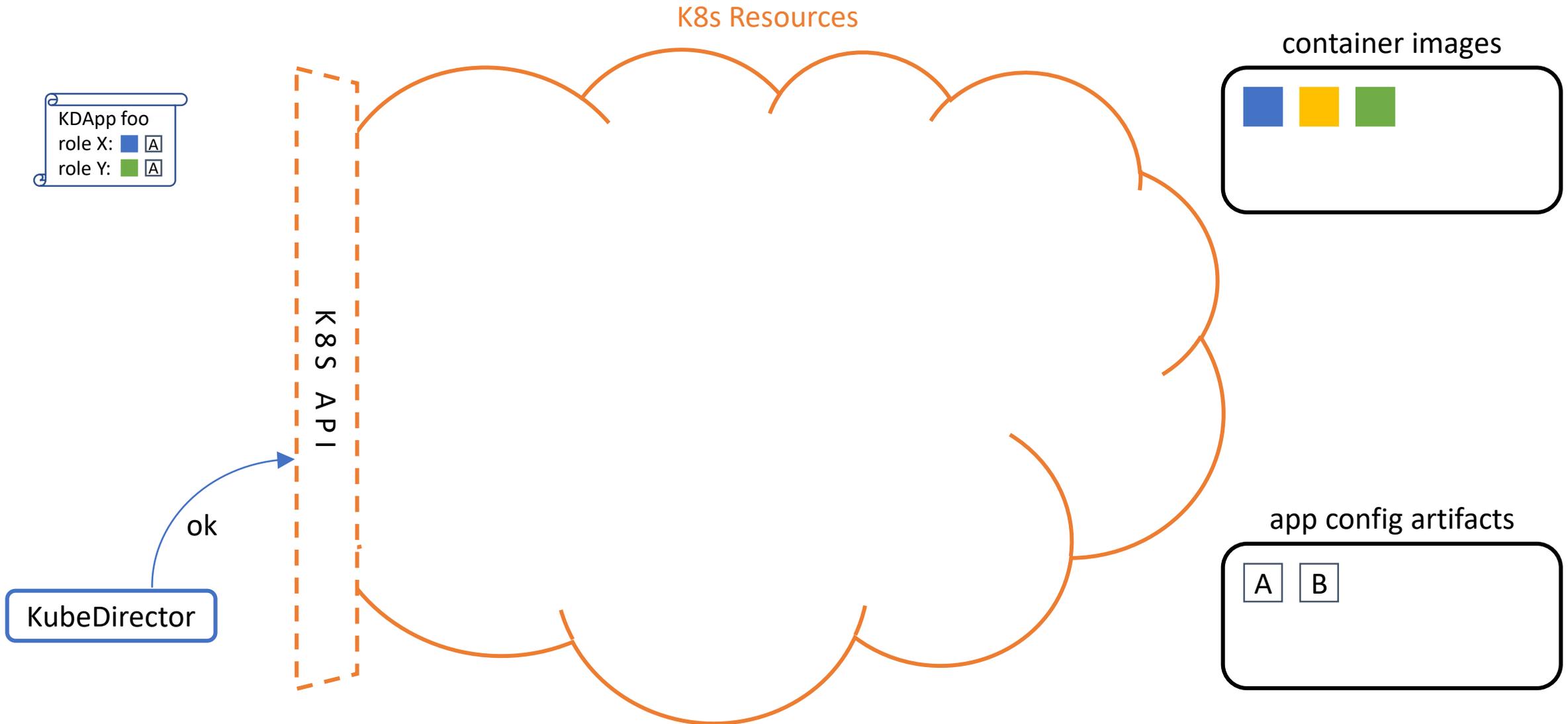


KubeCon



CloudNativeCon

North America 2019



KDApp foo
role X: ■ A
role Y: ■ A

KubeDirector

ok

K8S API

K8s Resources

container images

app config artifacts

AB

KubeDirectorApp Deploy

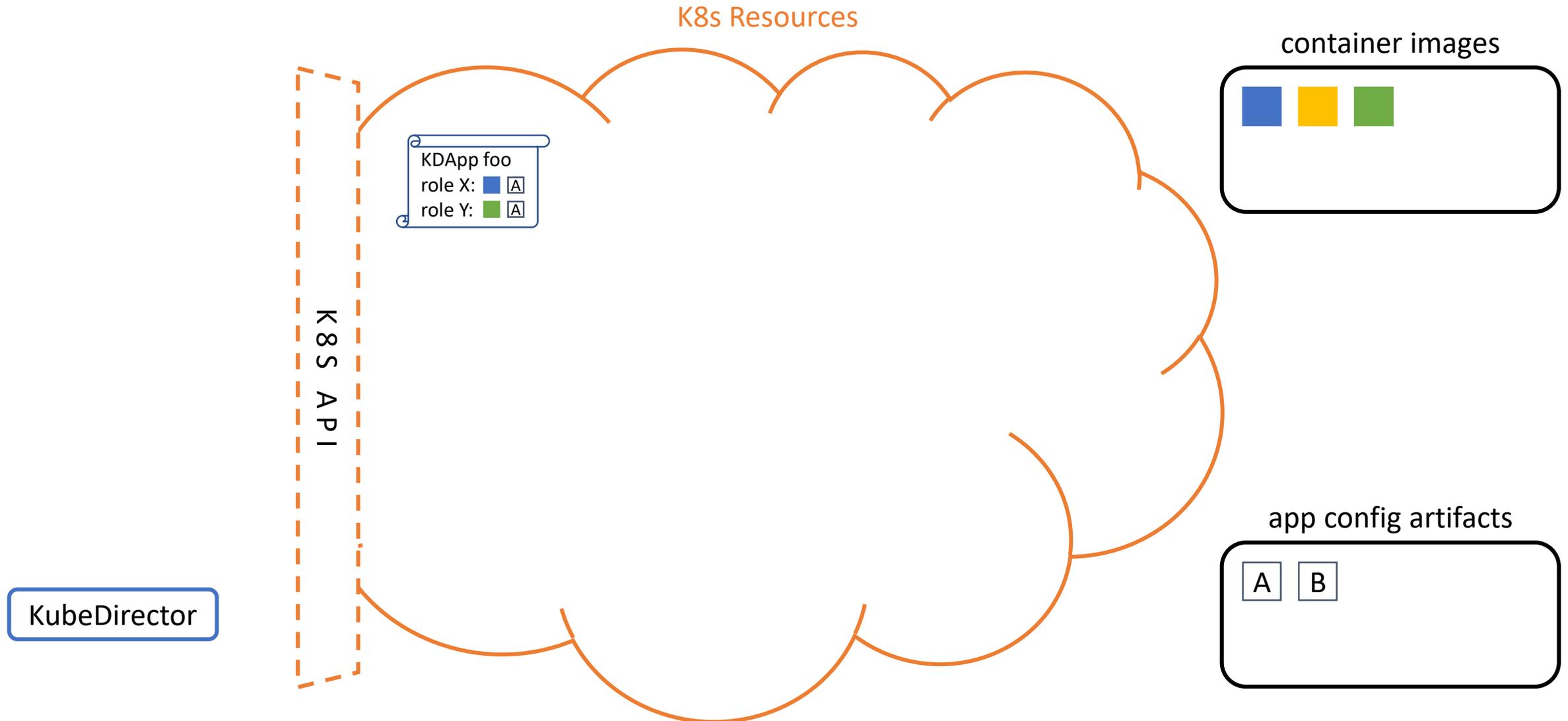


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

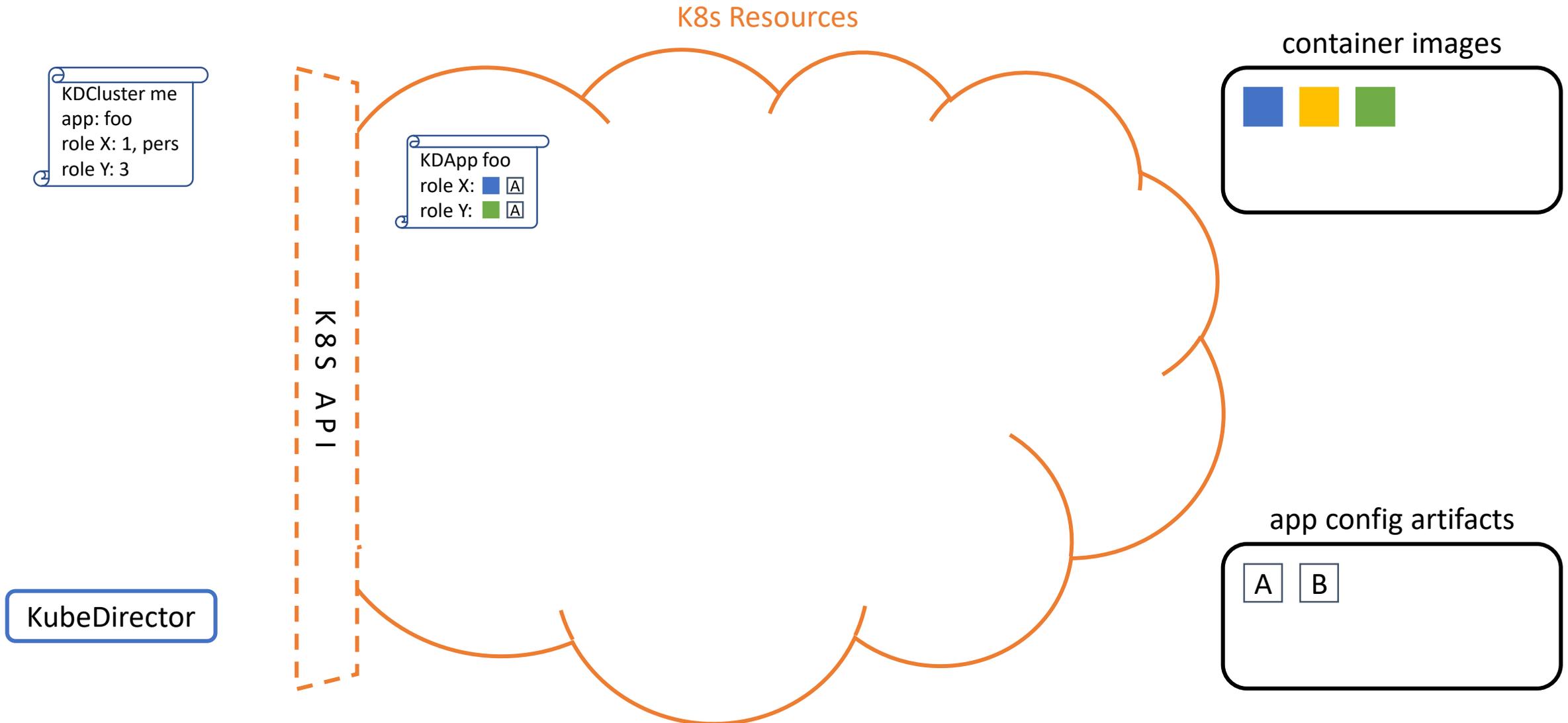


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

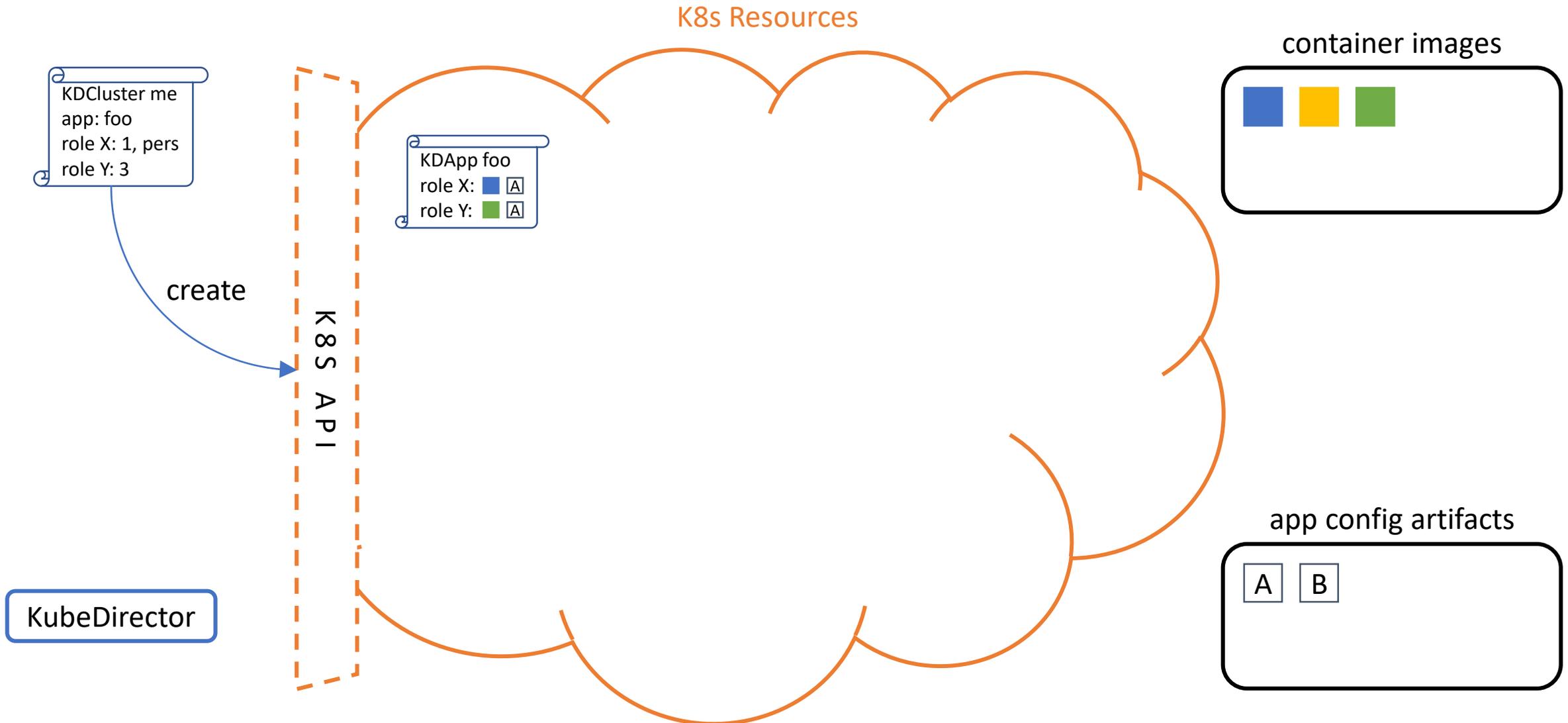


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

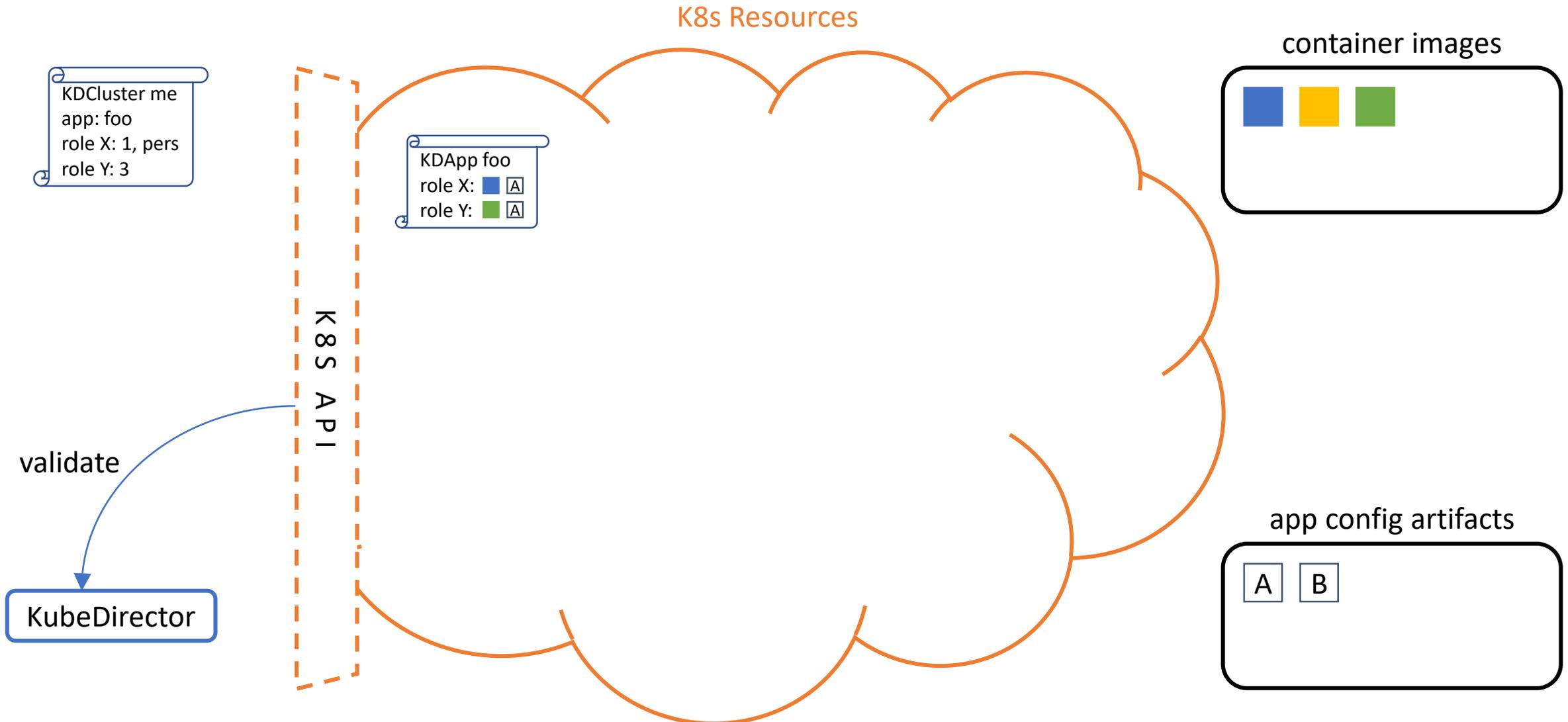


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy



KubeCon



CloudNativeCon

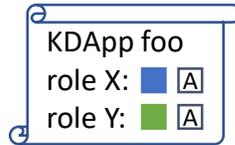
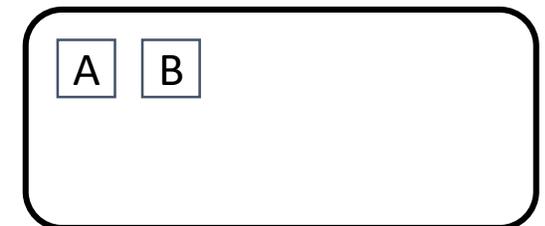
North America 2019

K8s Resources

container images

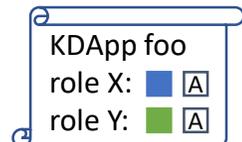


app config artifacts



K8S API

KubeDirector



KubeDirectorCluster Deploy

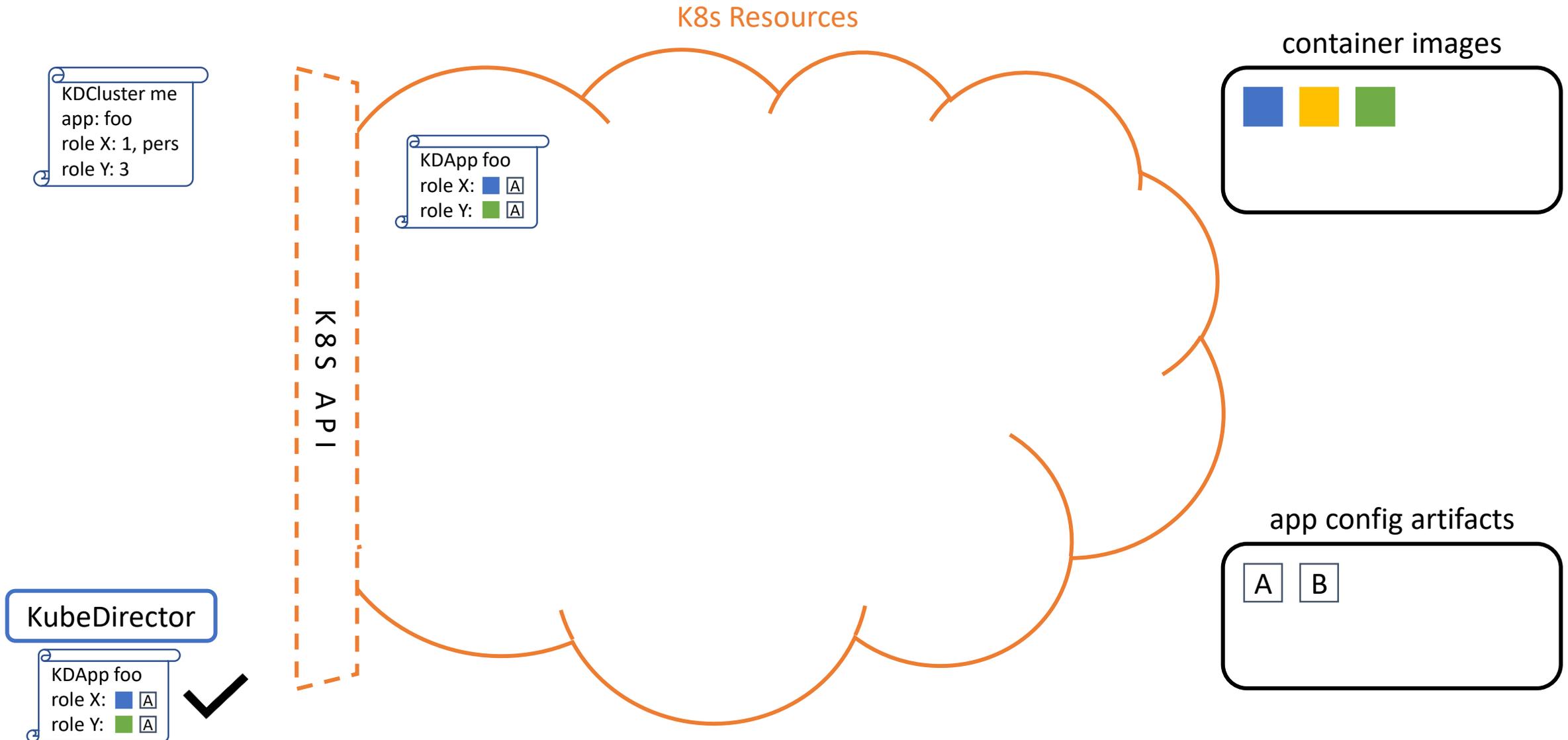


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

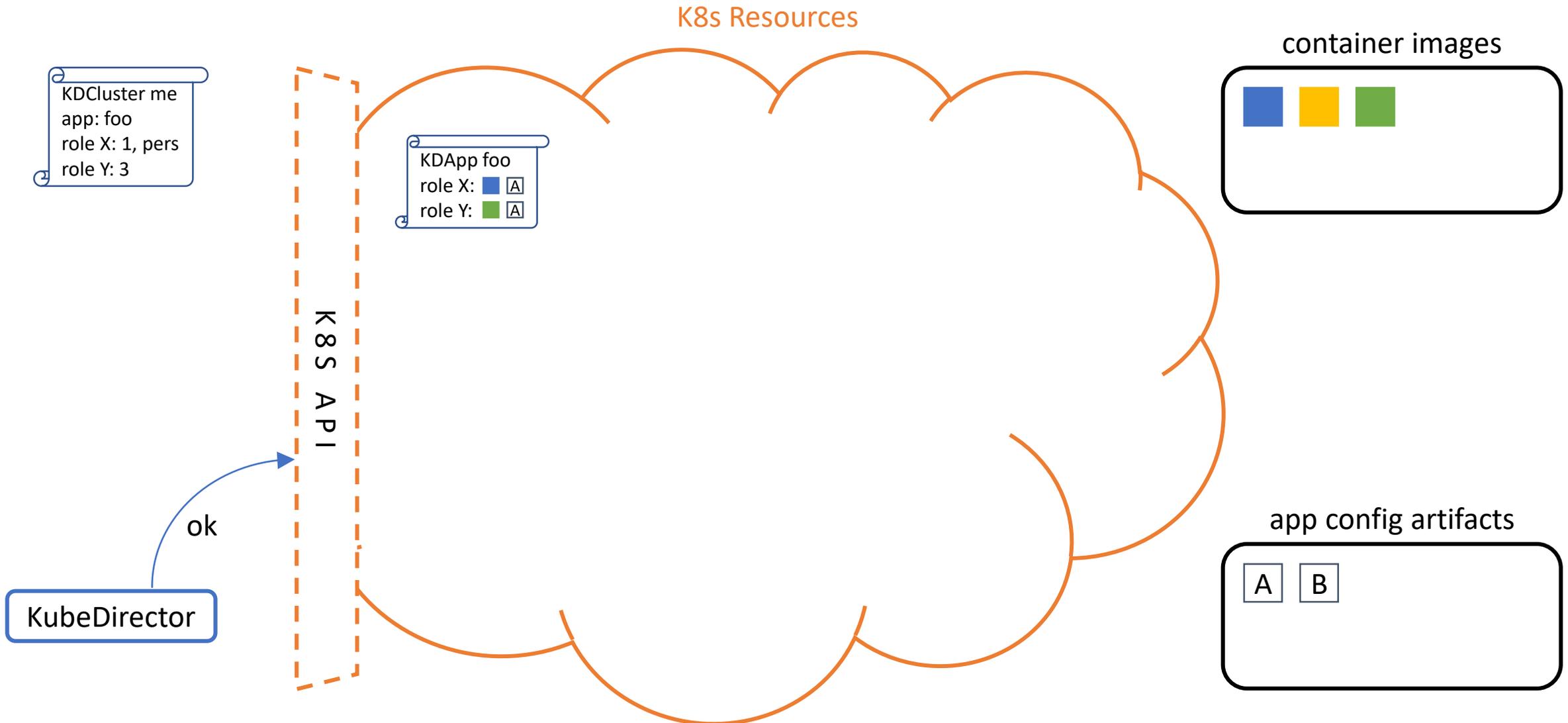


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

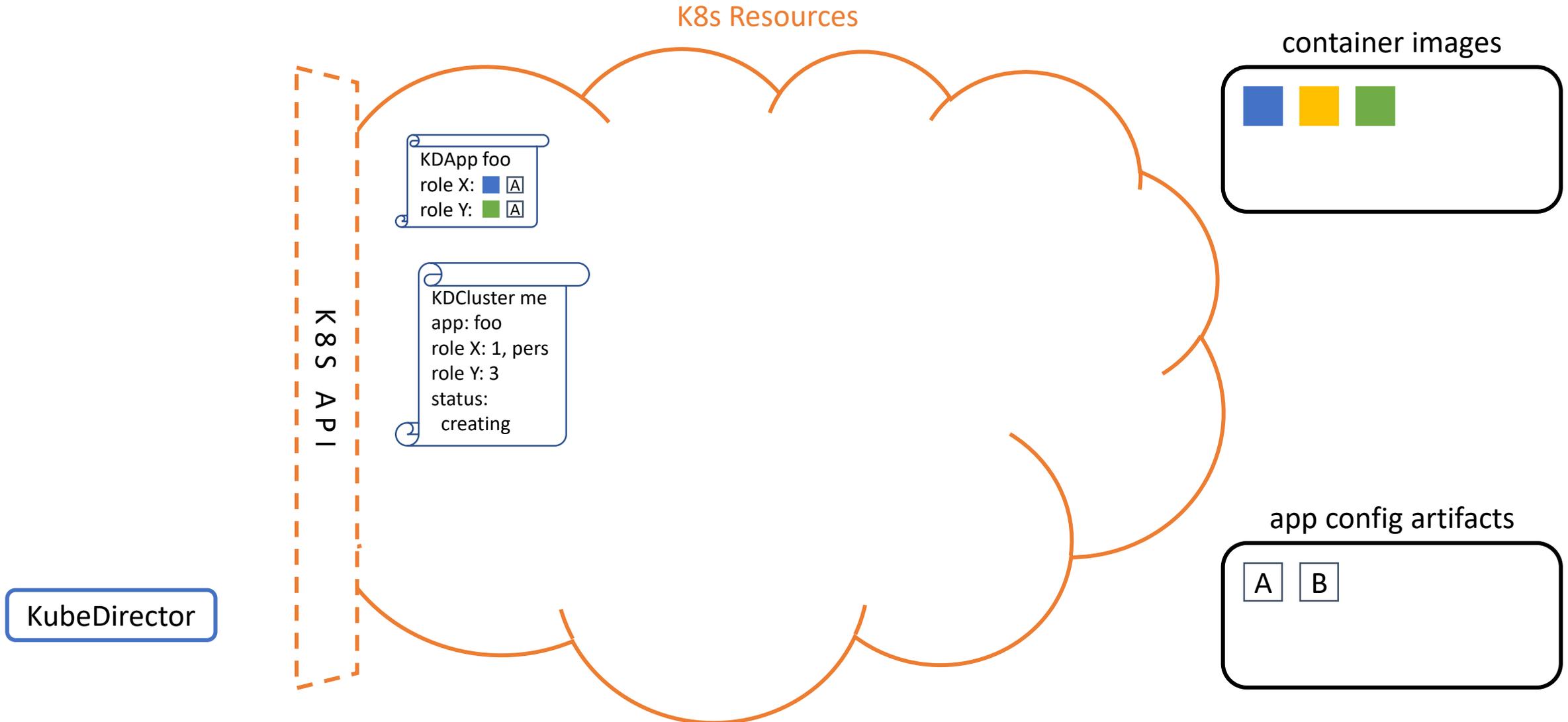


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

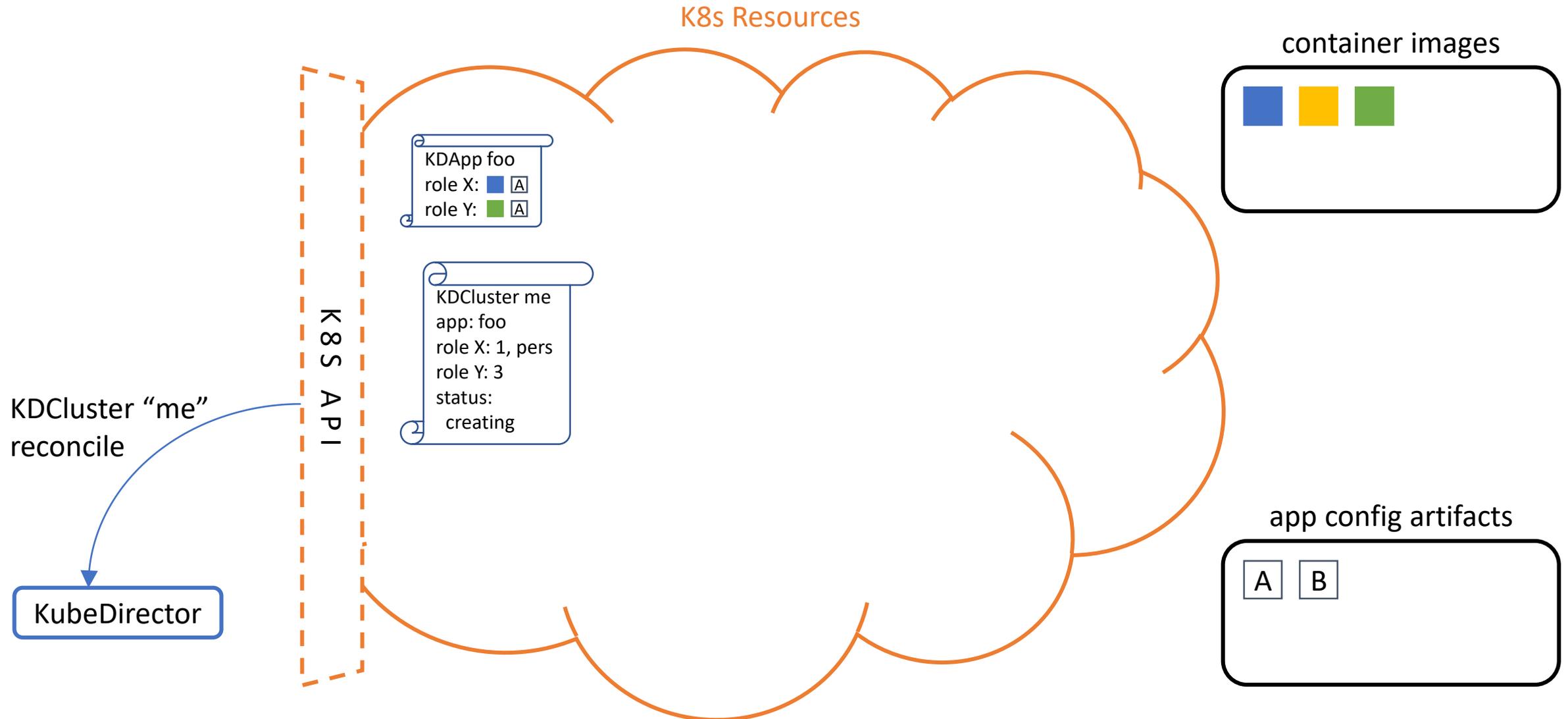


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

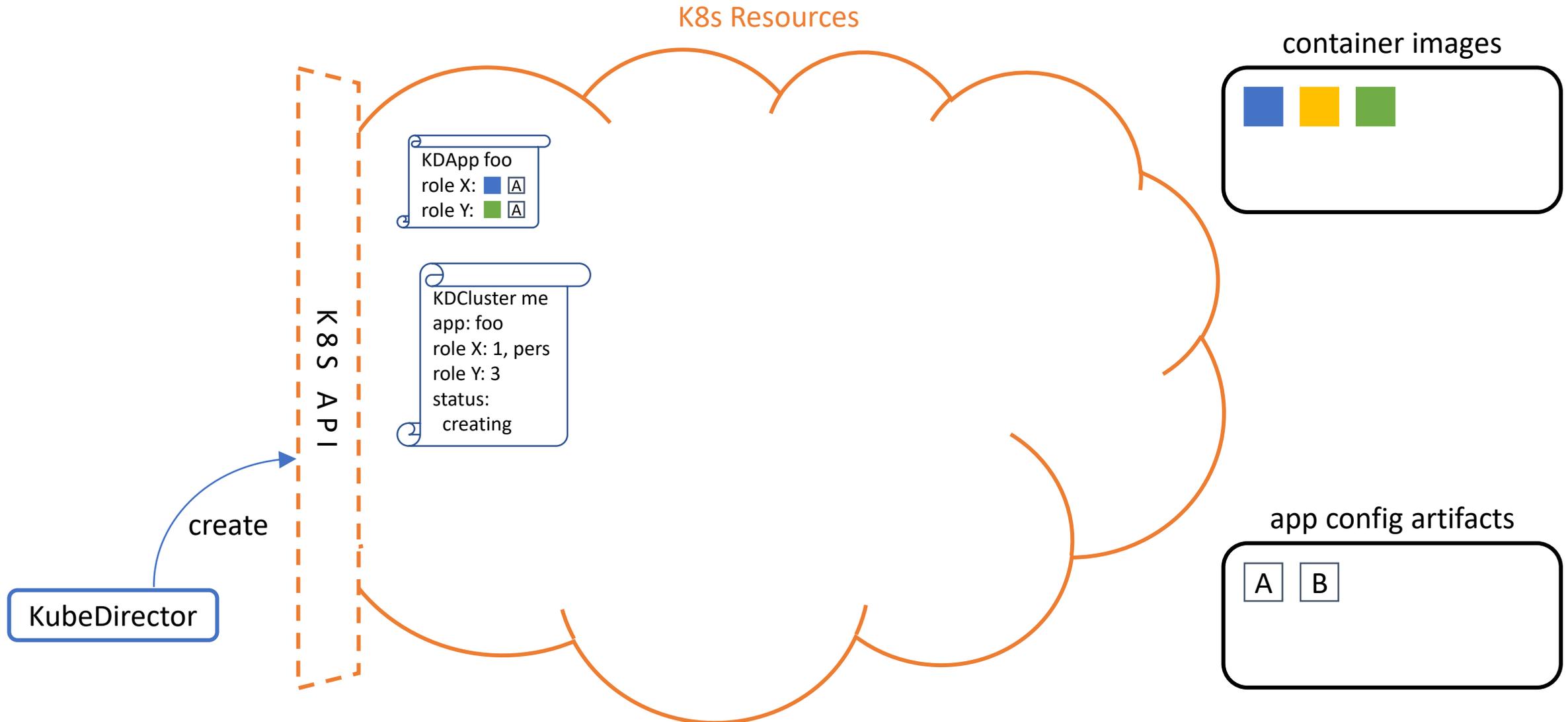


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

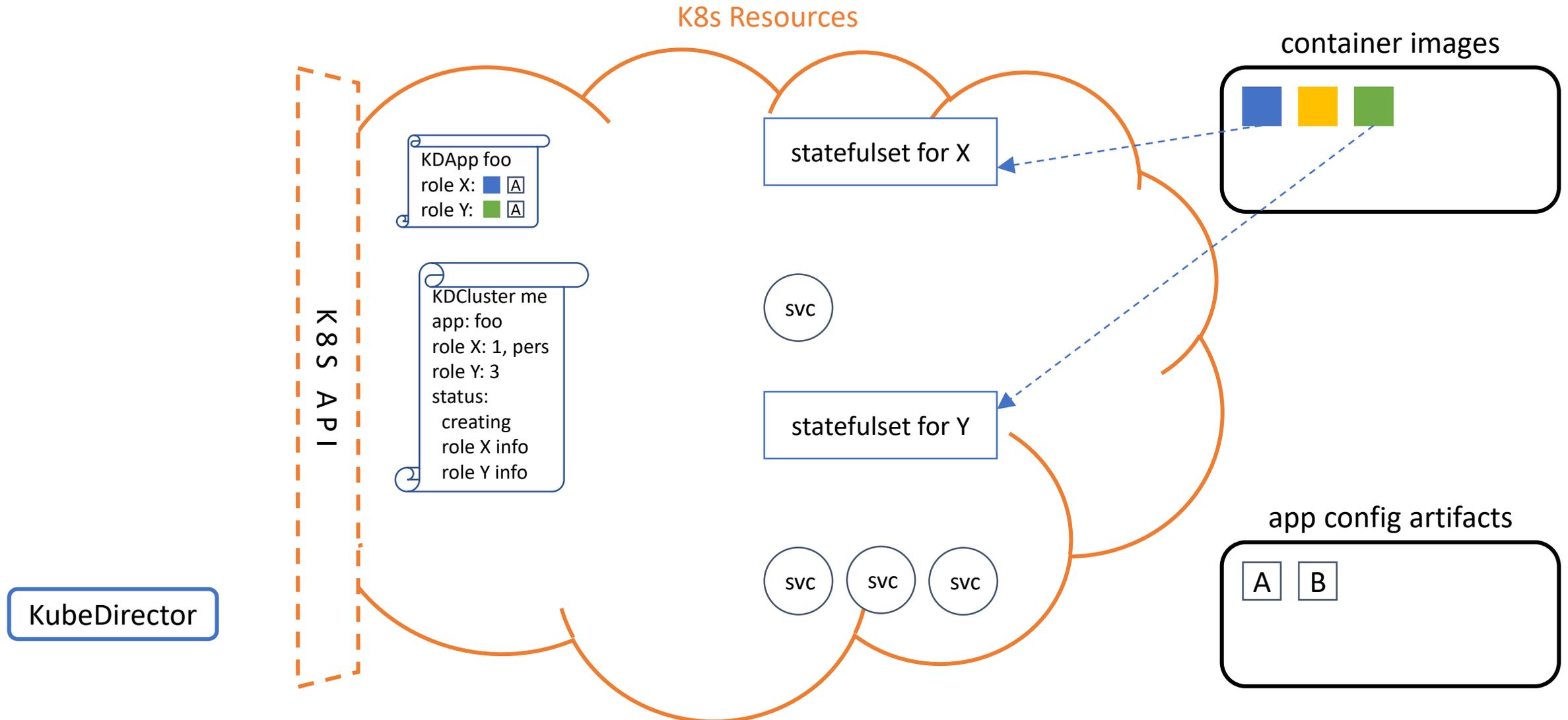


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

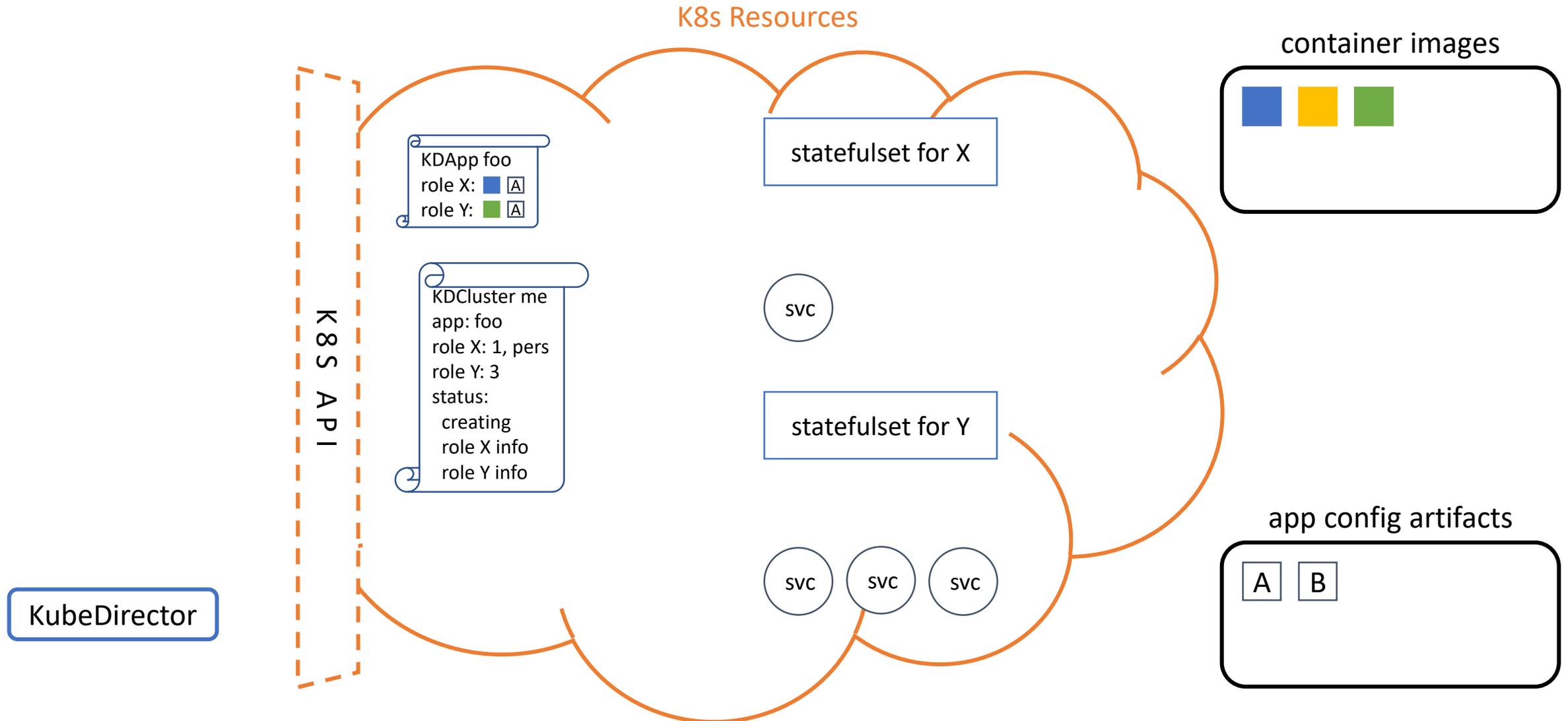


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

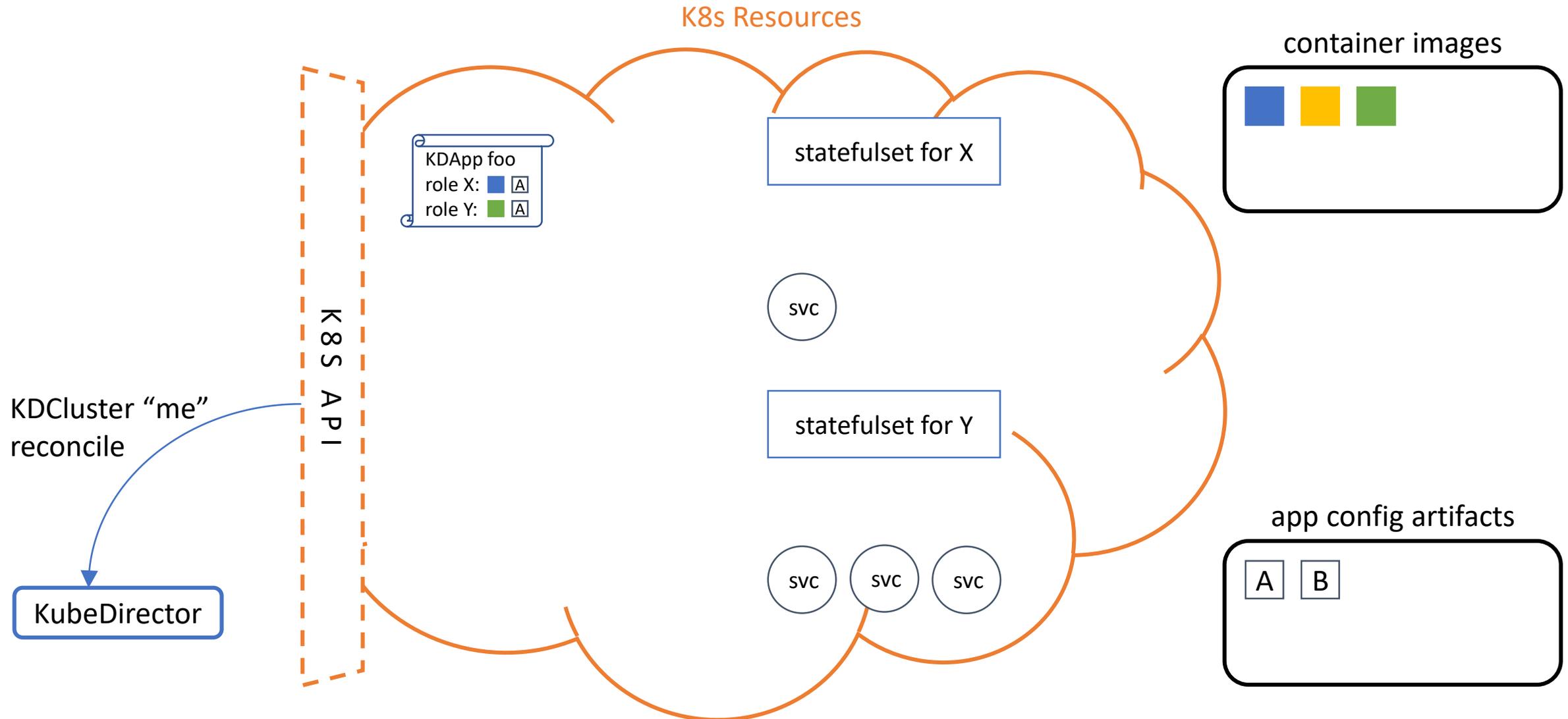


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

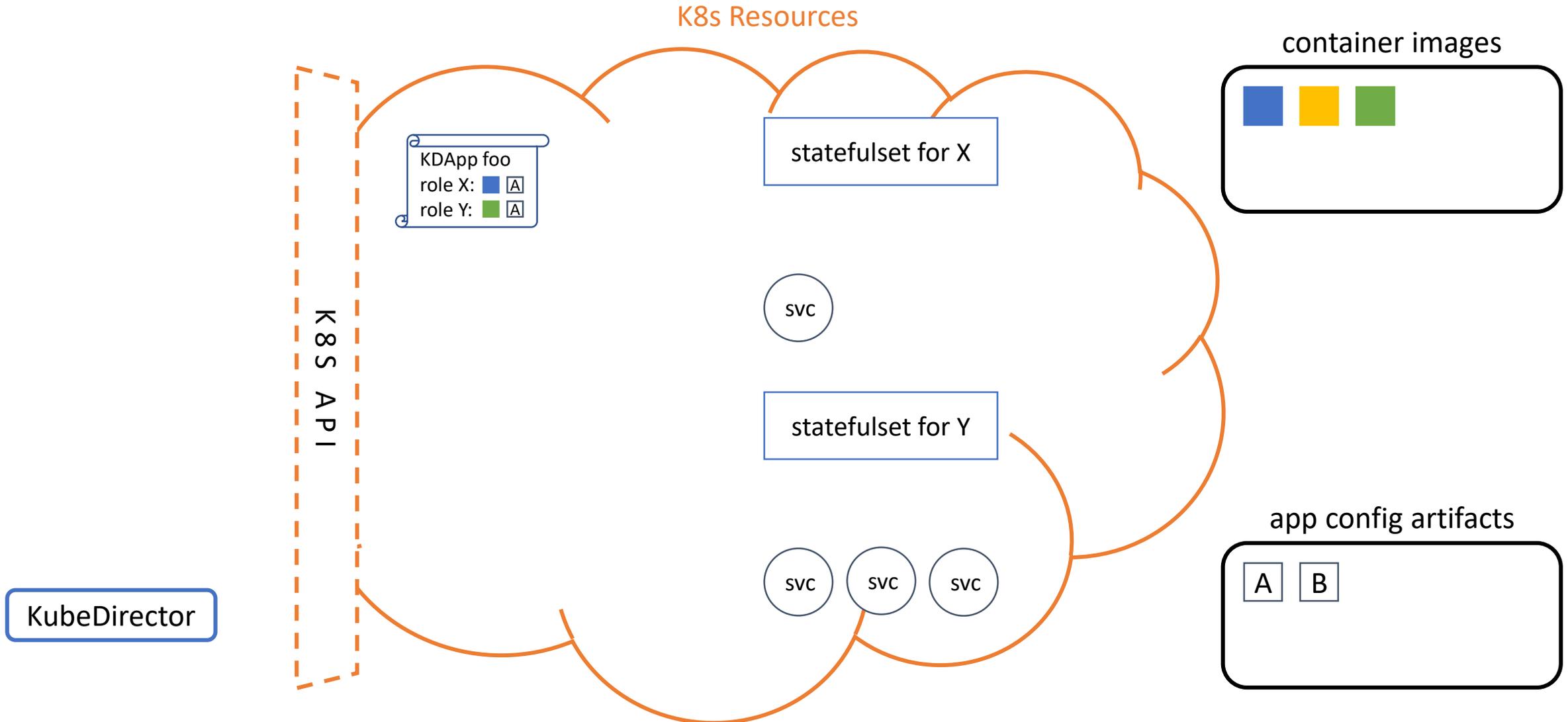


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

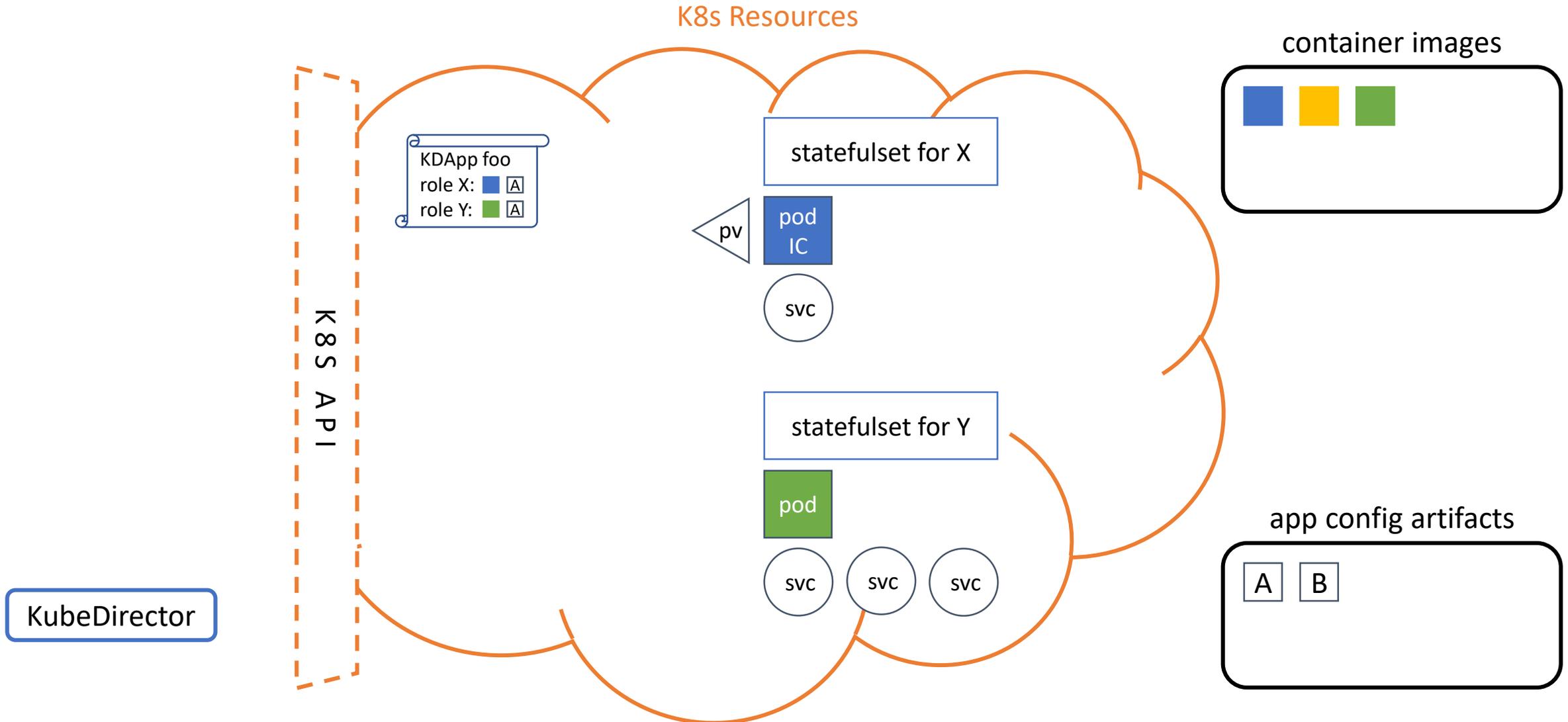


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

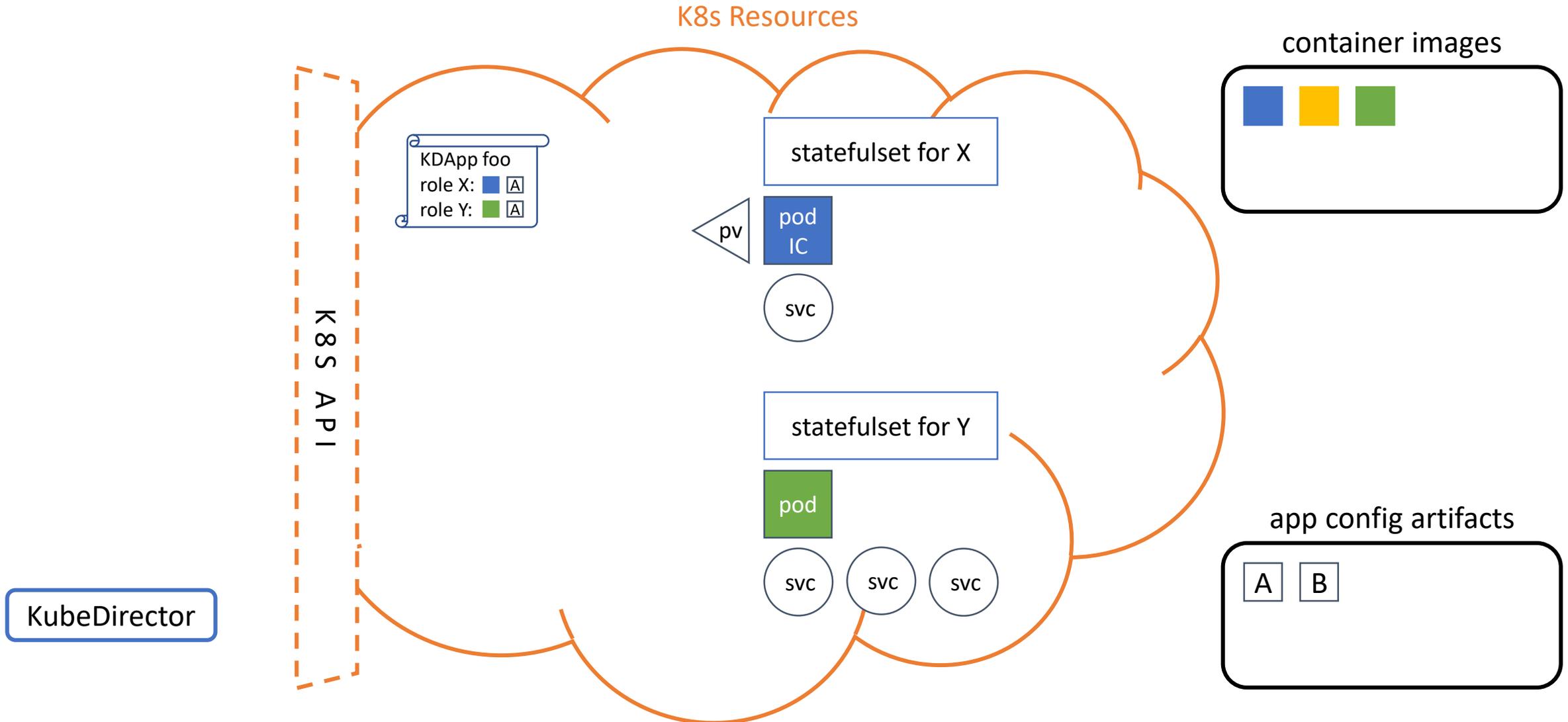


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

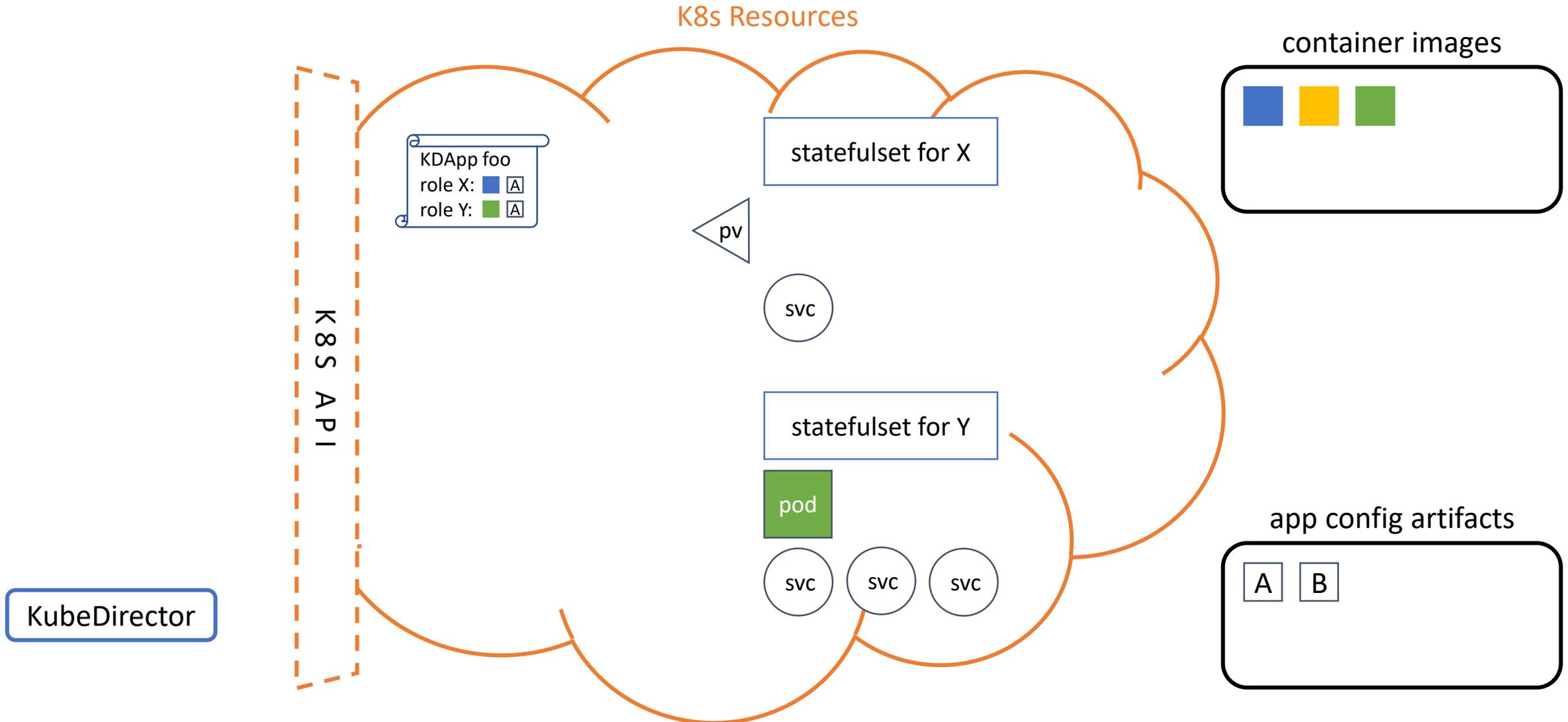


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

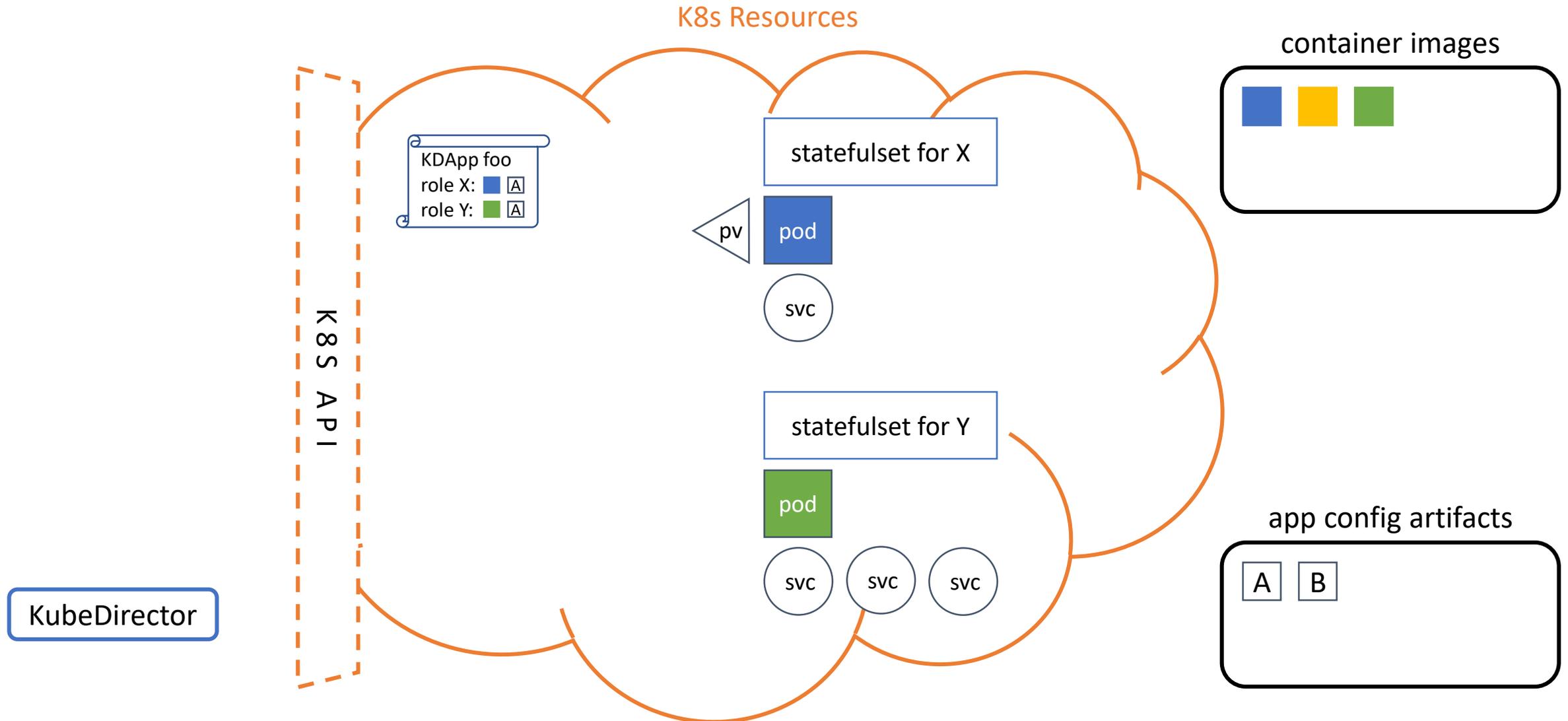


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

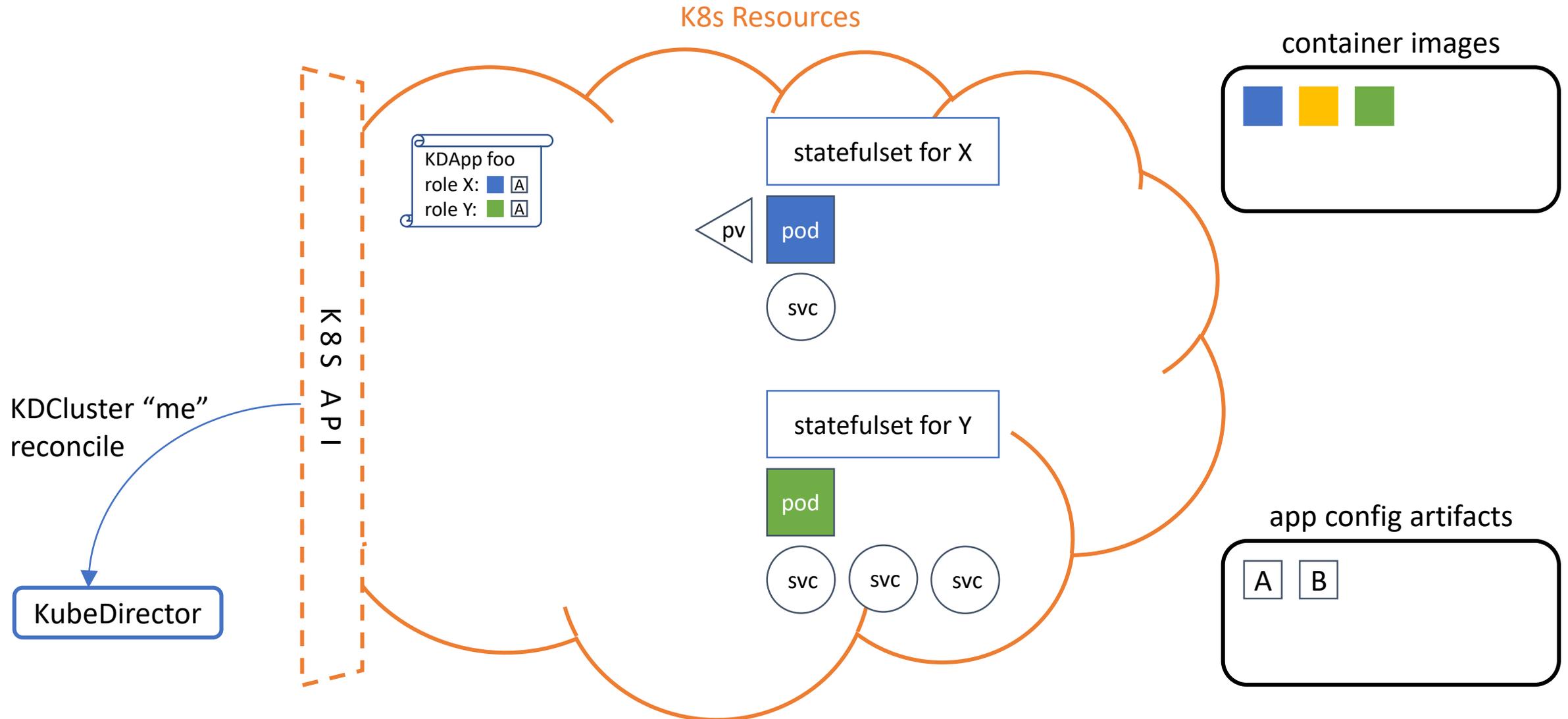


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

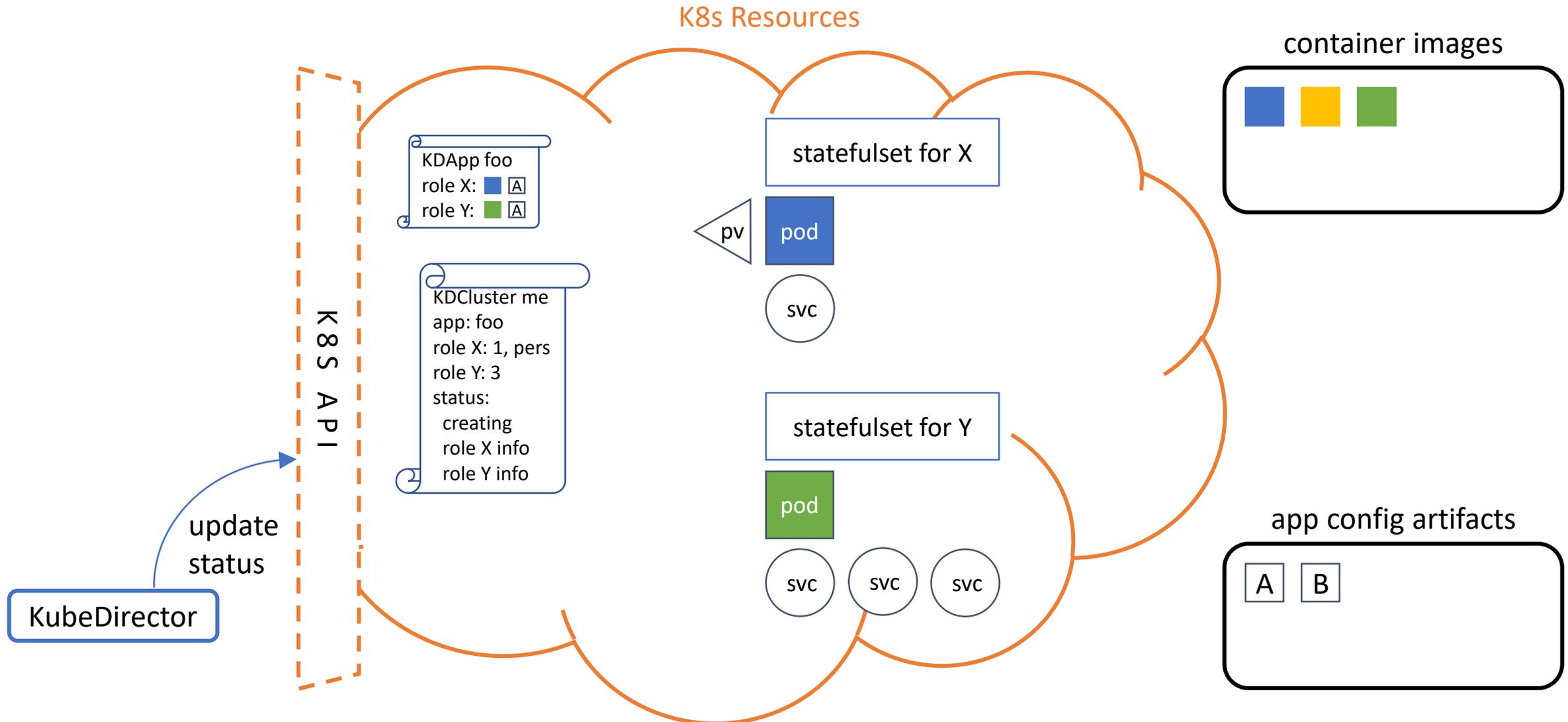


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

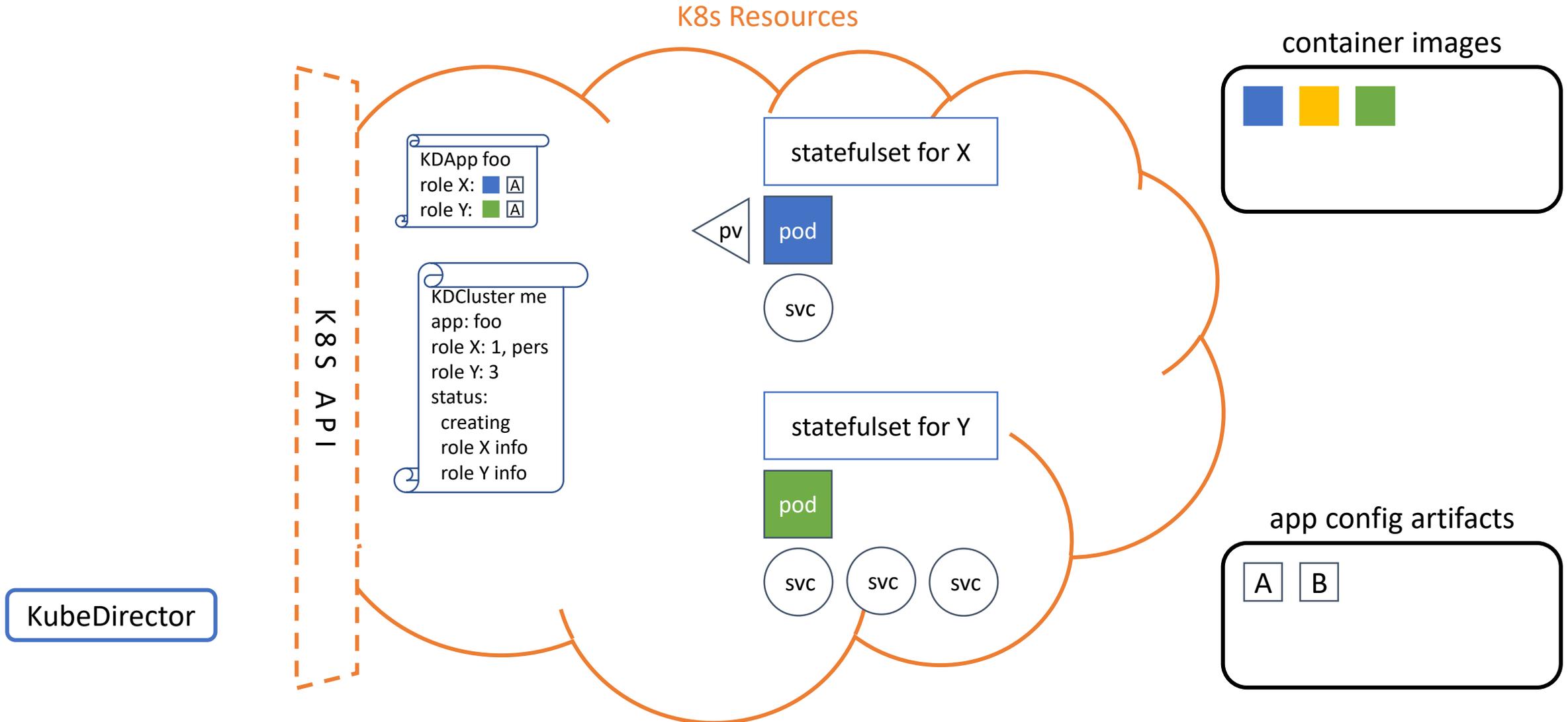


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

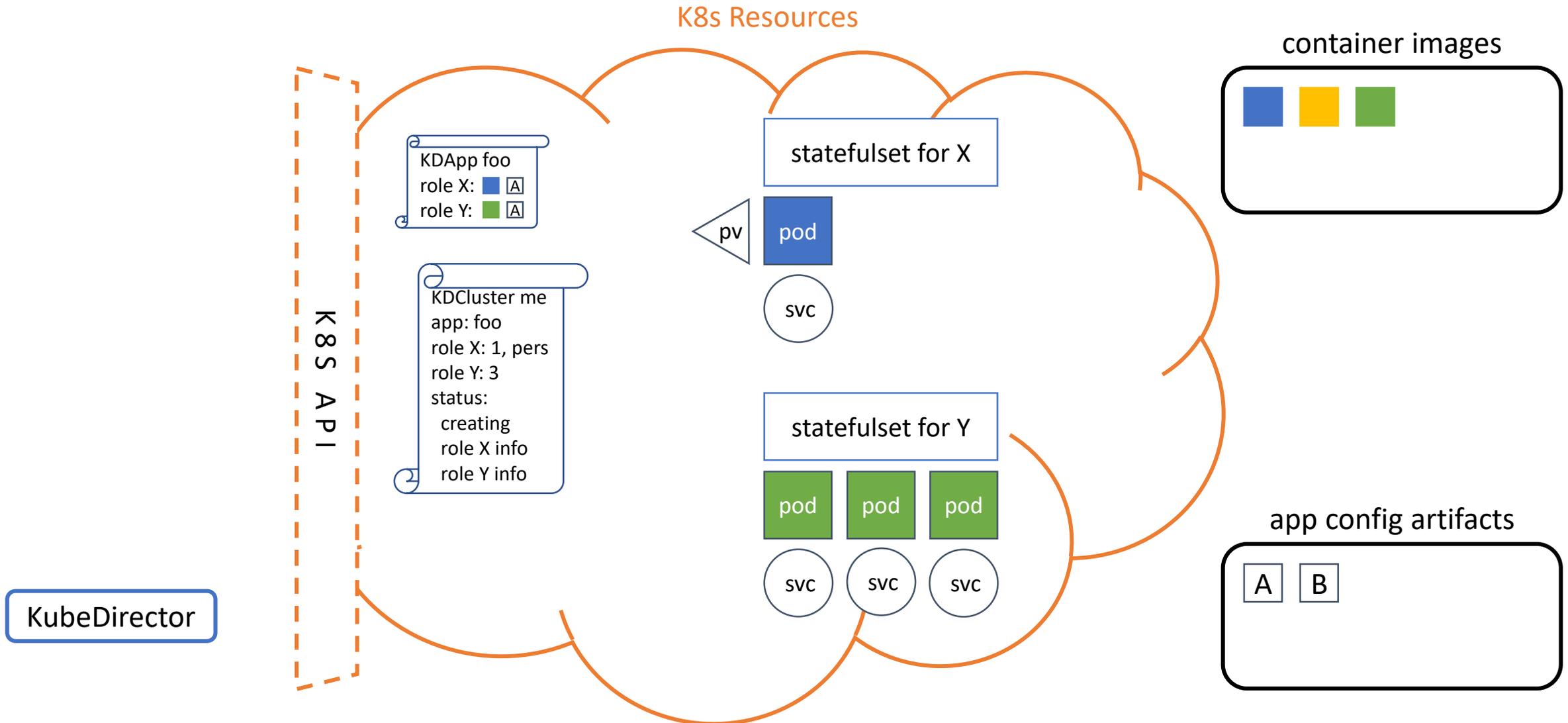


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

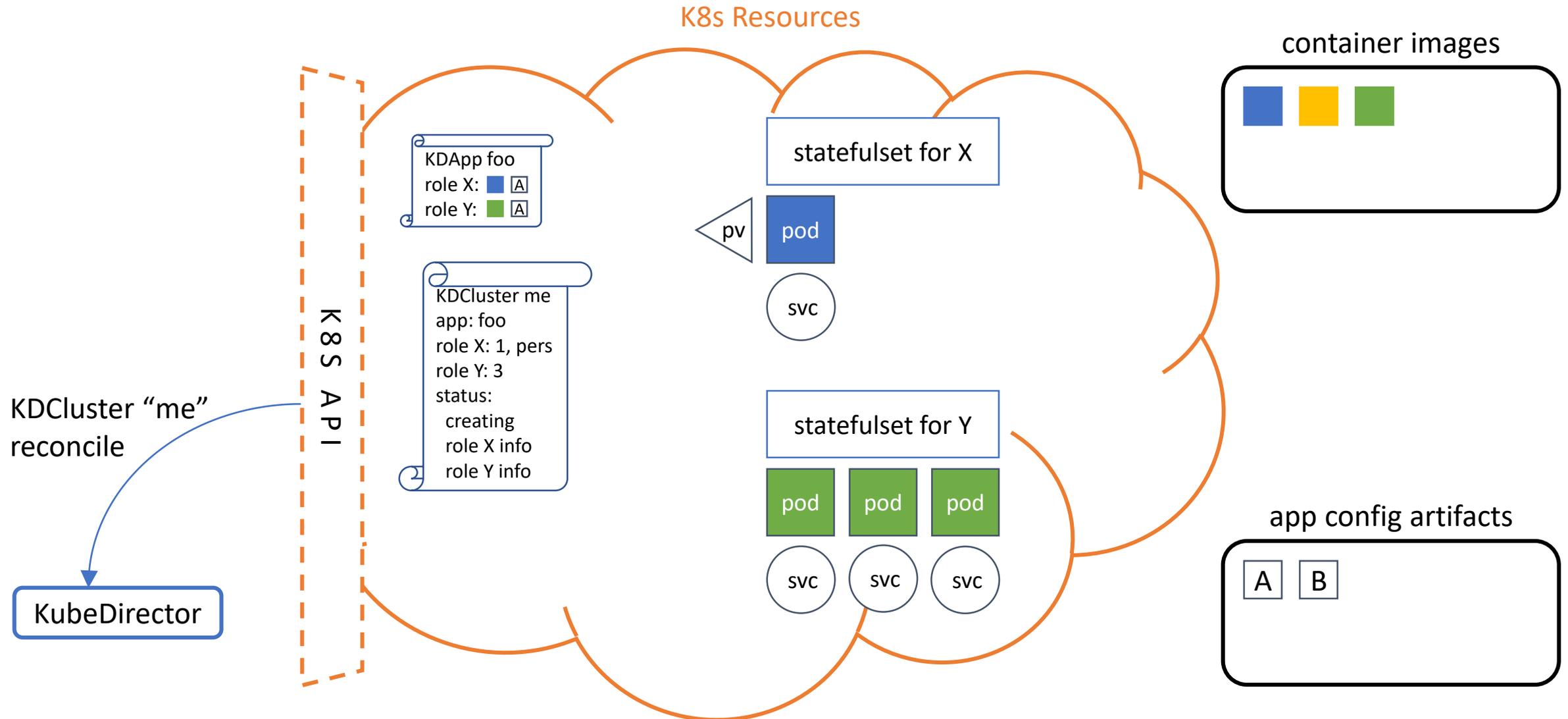


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

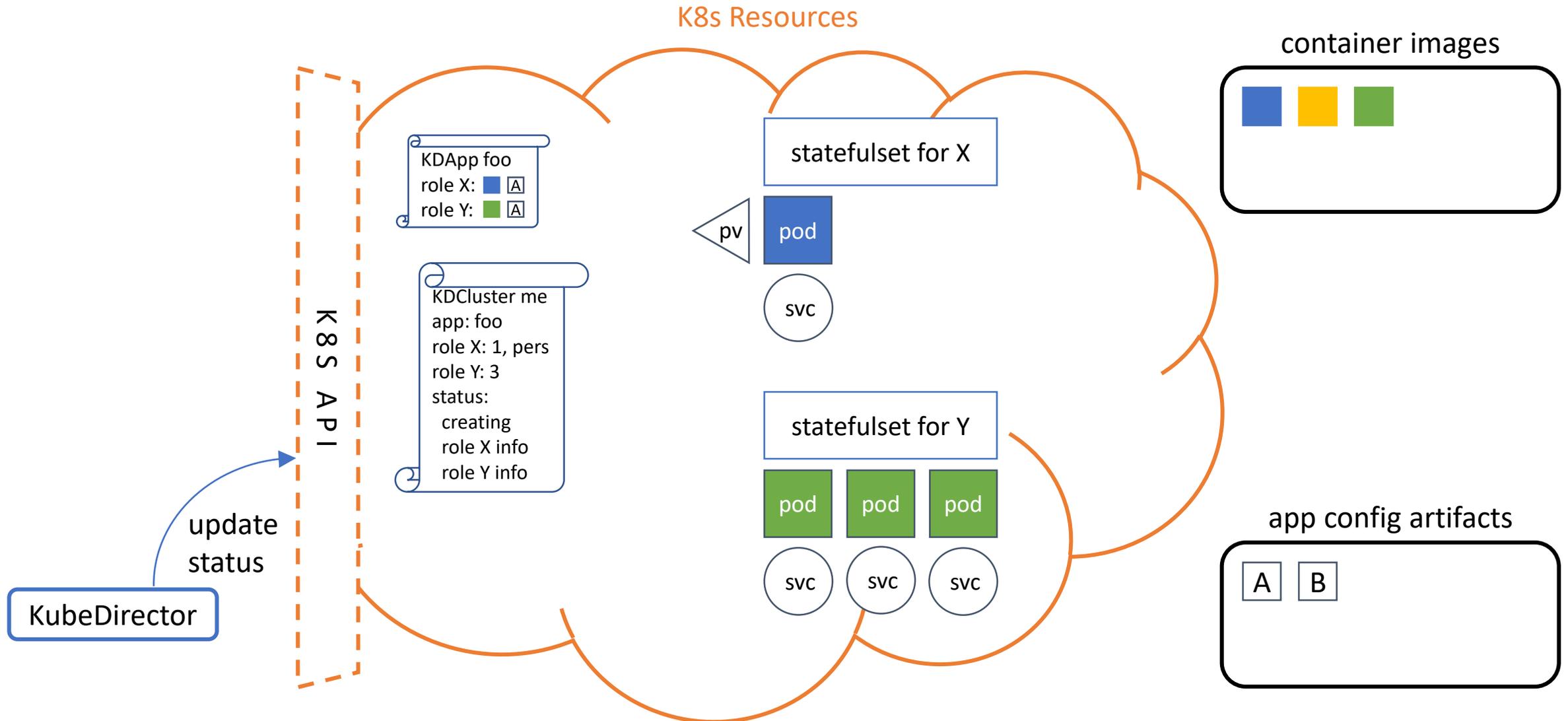


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

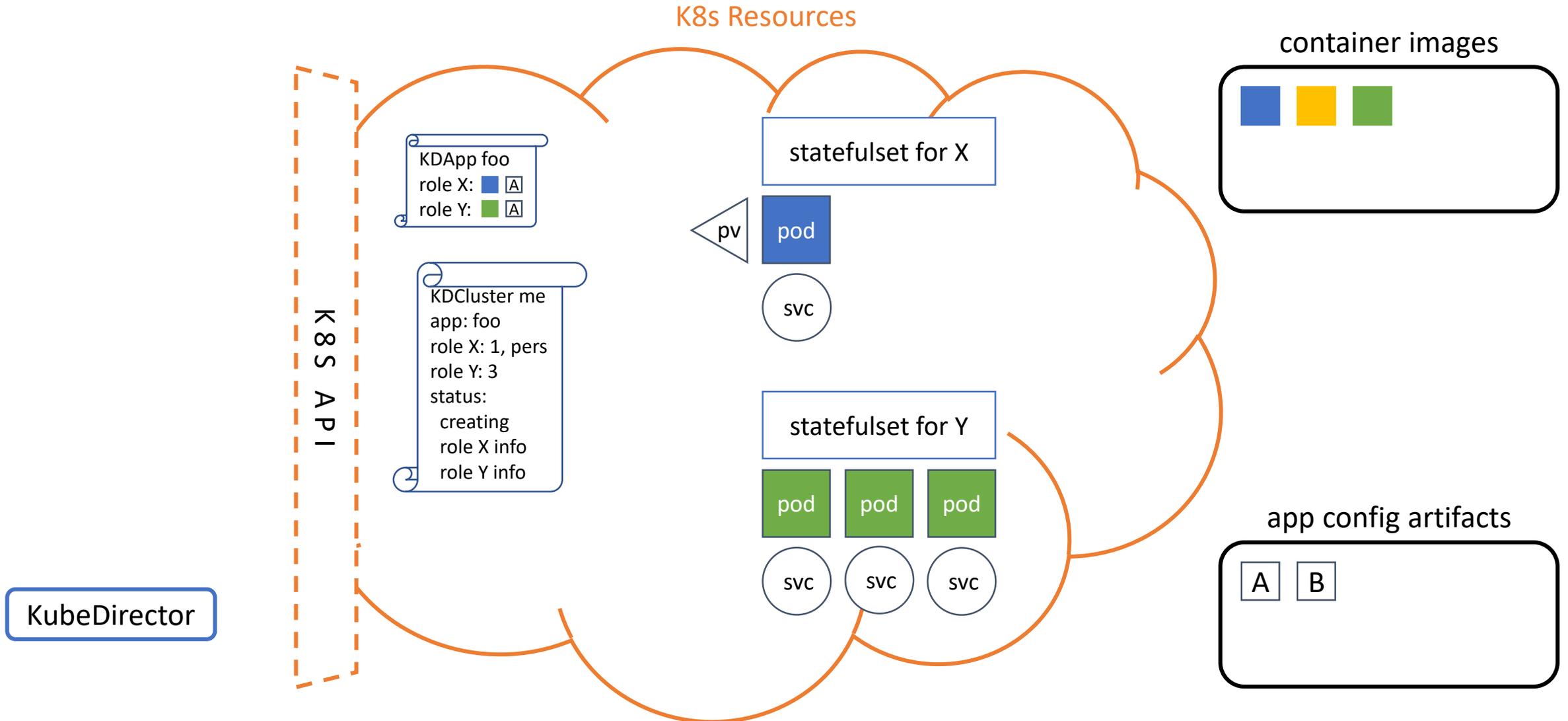


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

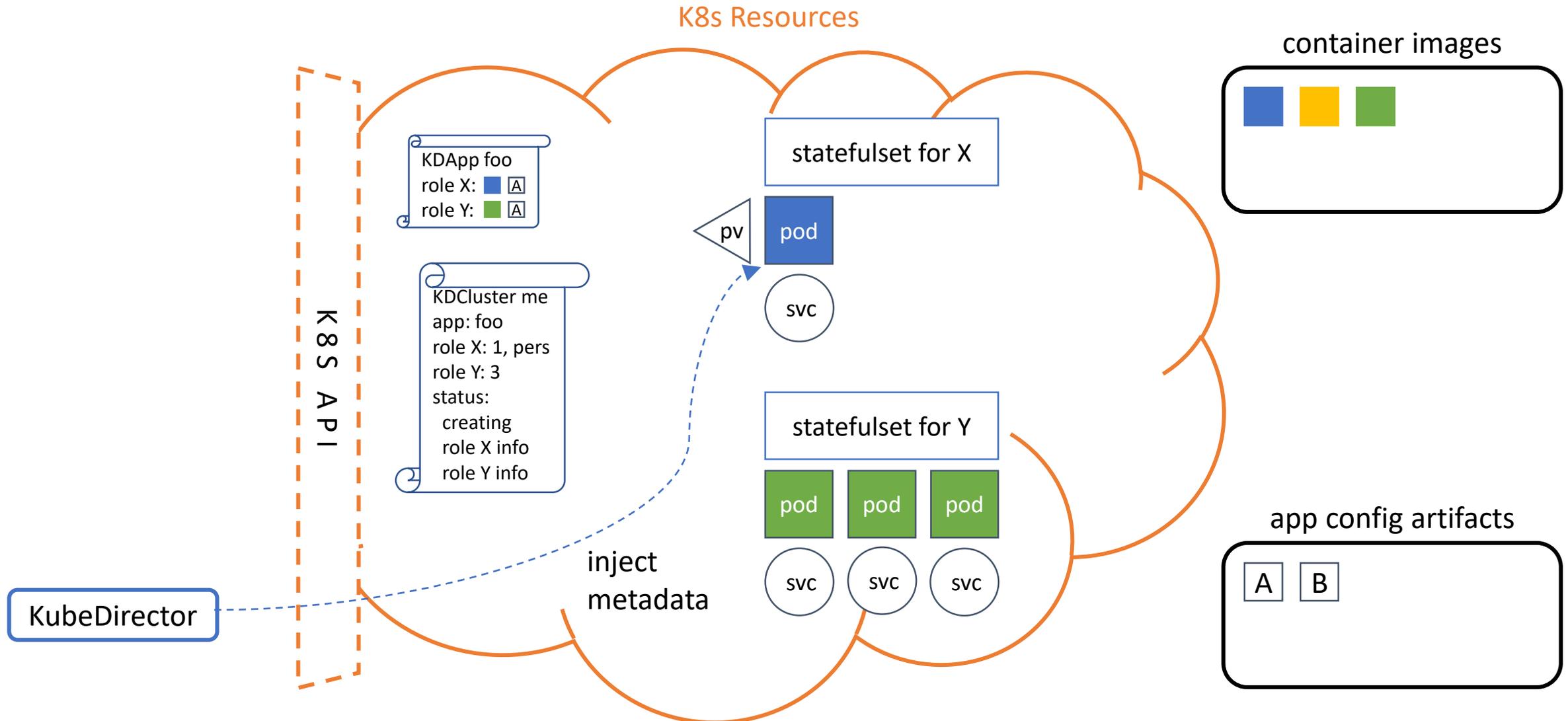


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

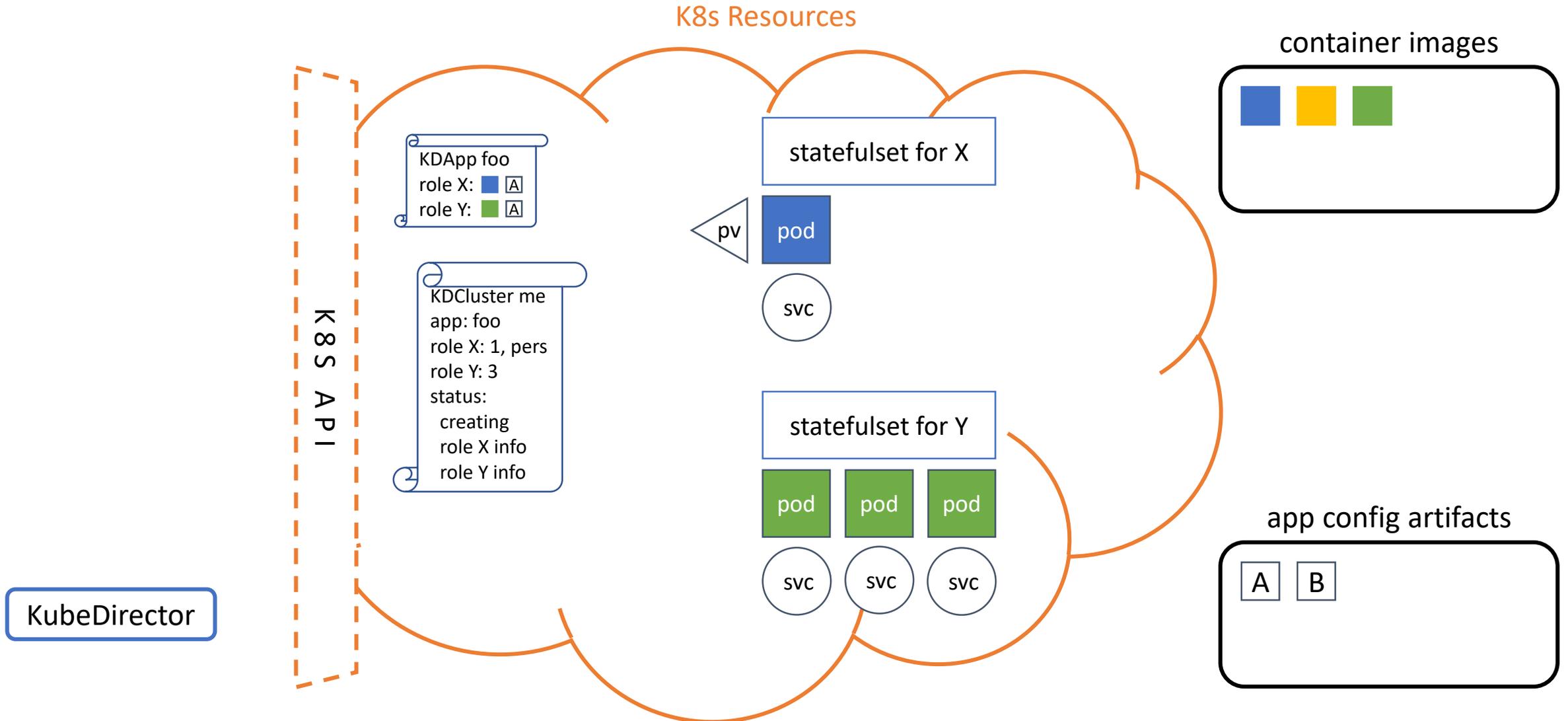


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

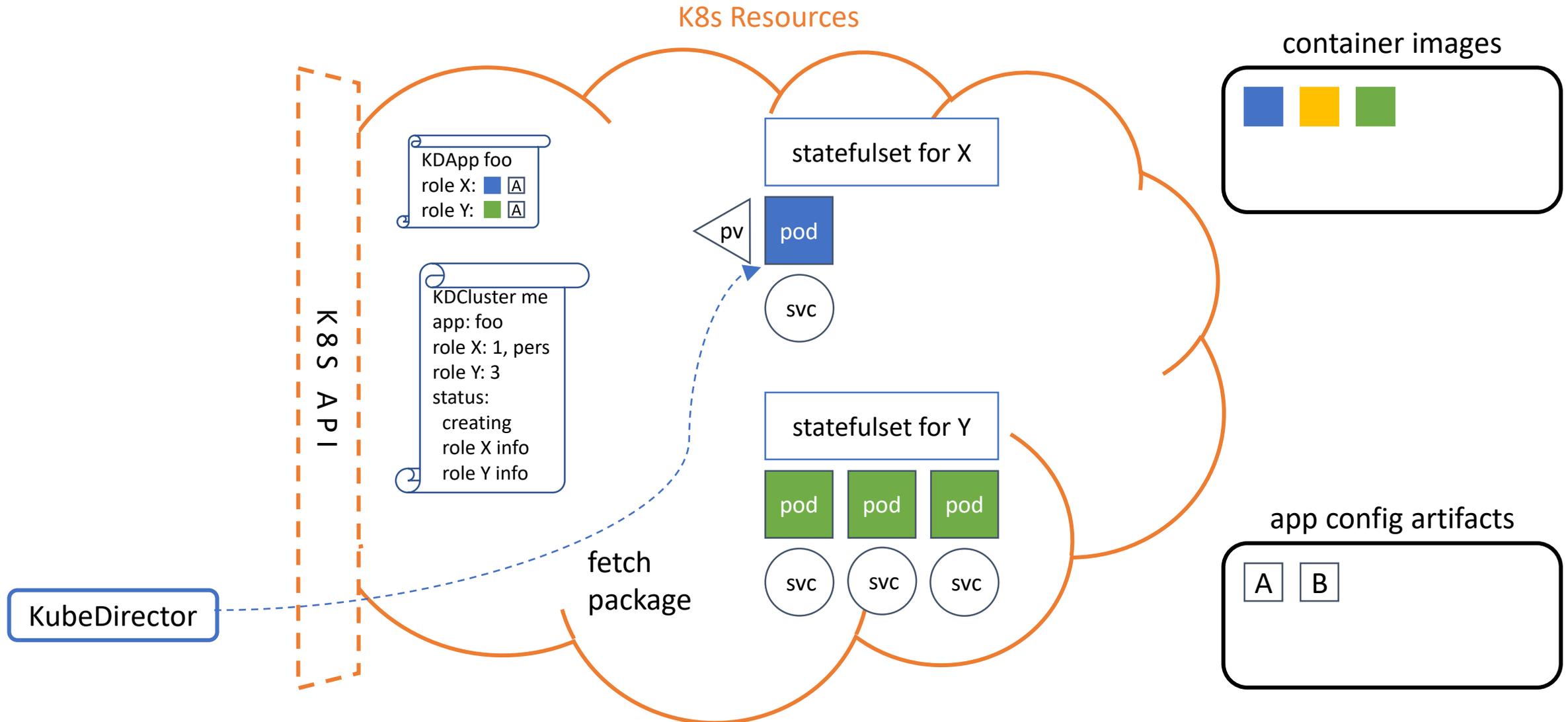


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

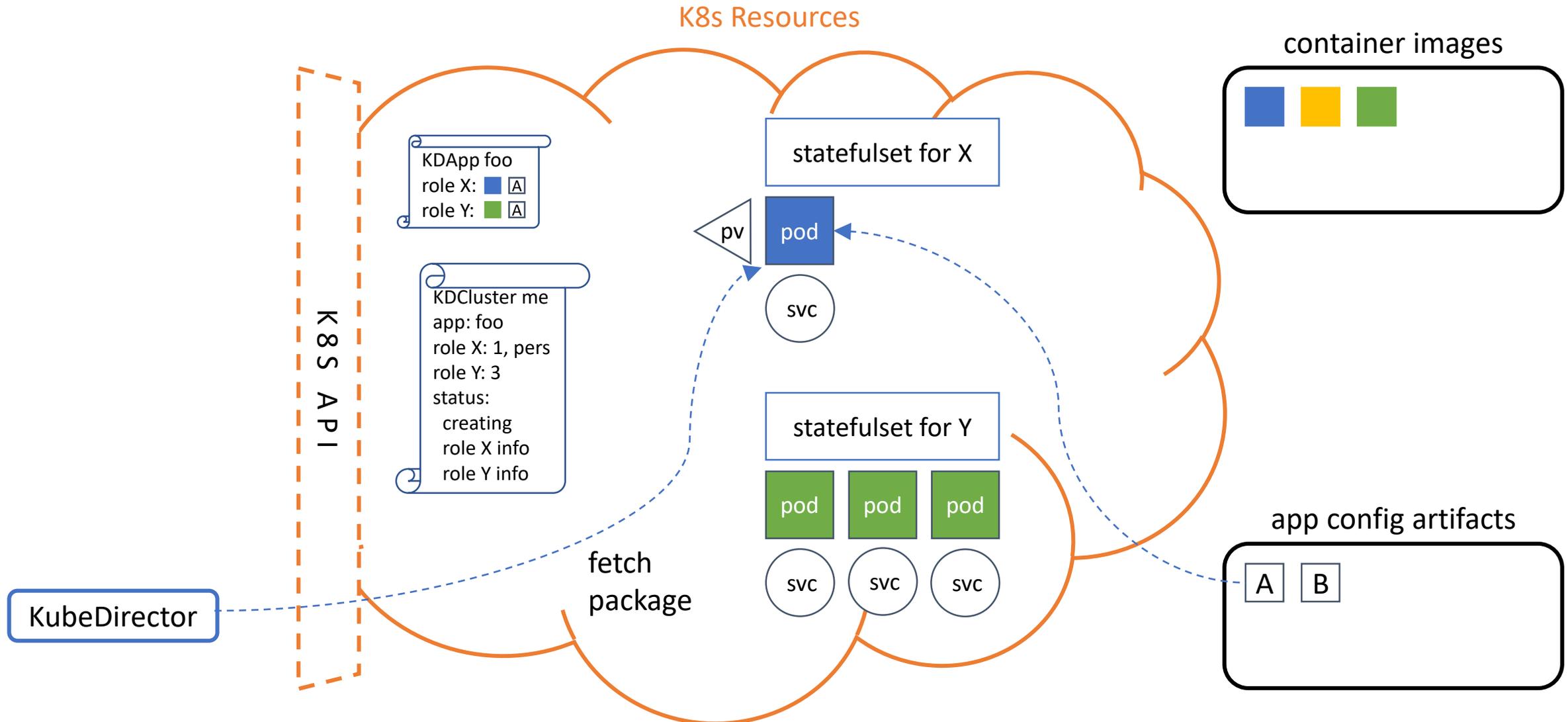


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

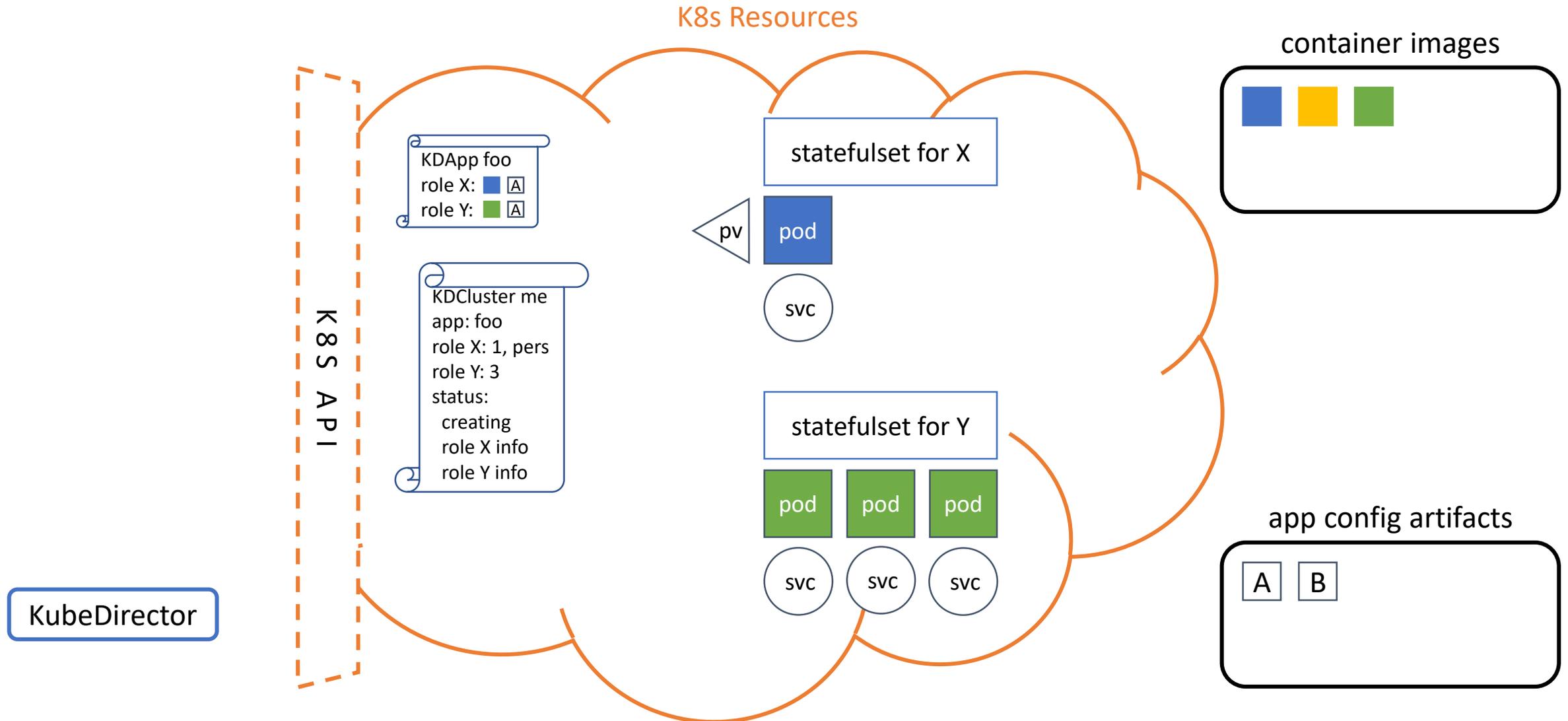


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

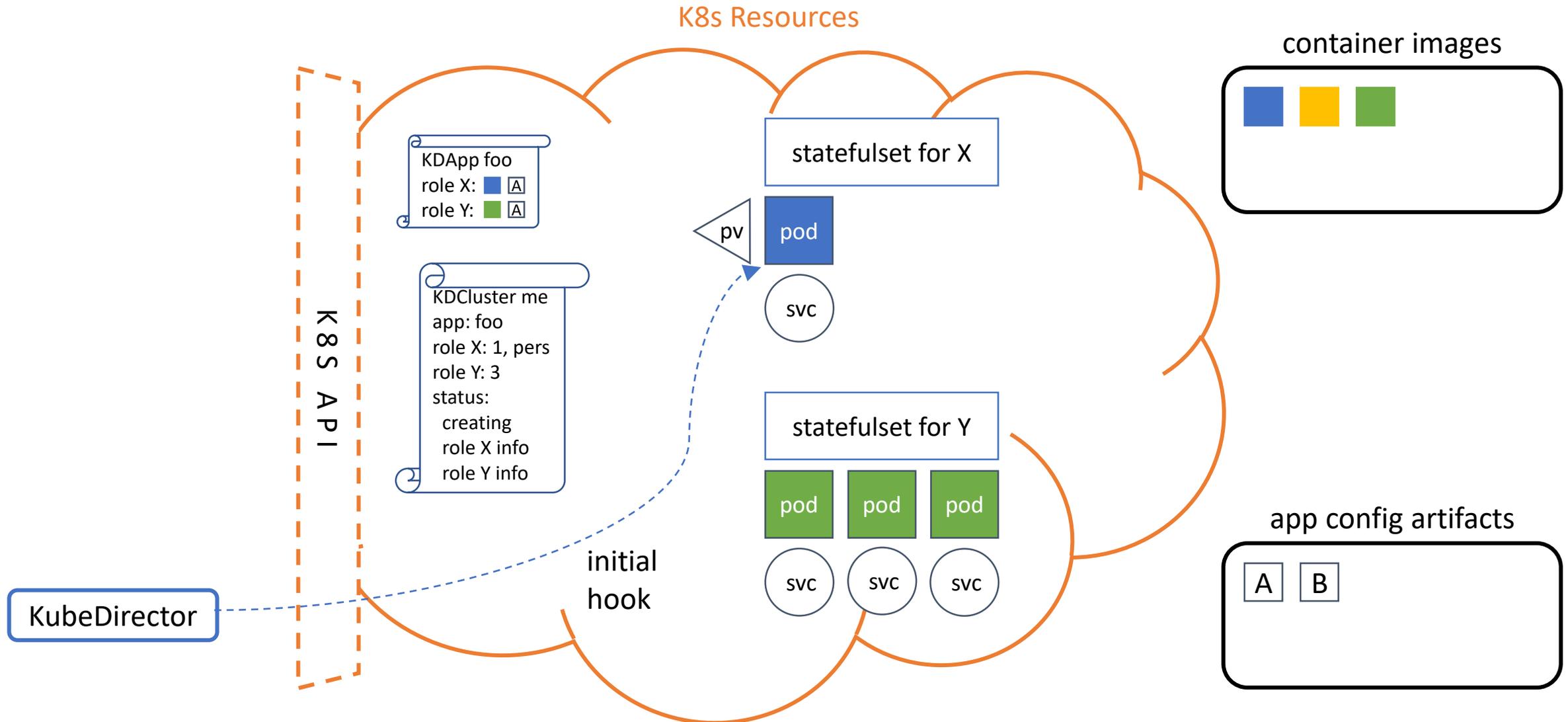


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

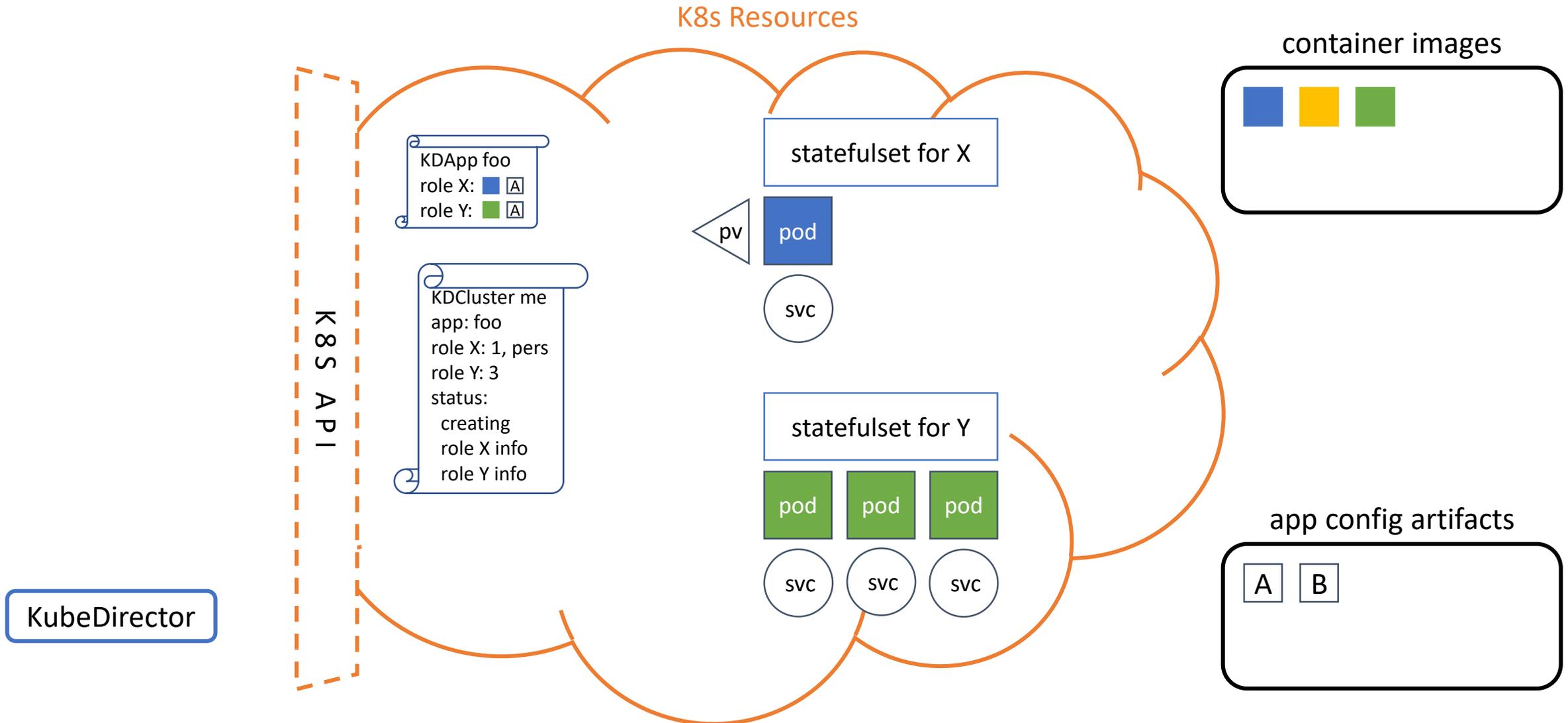


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

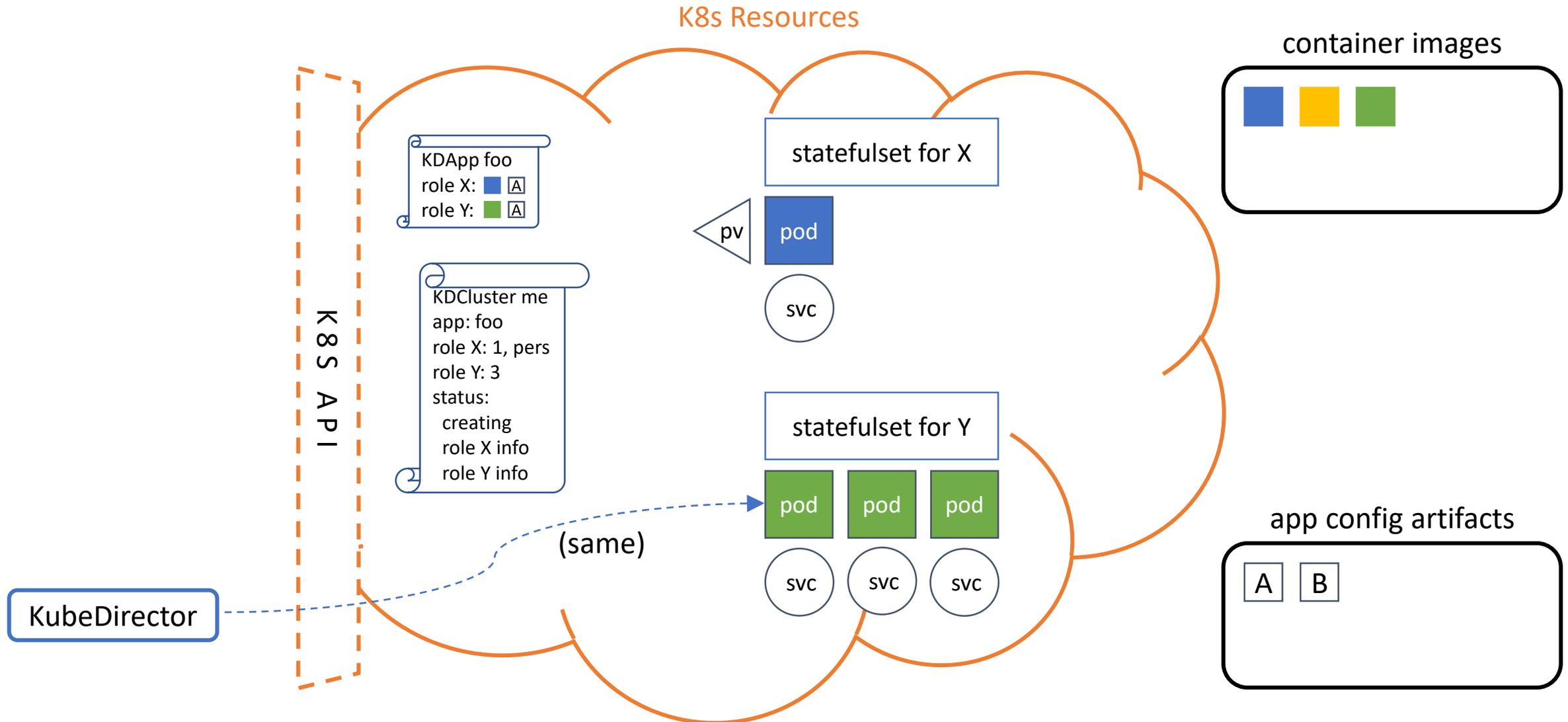


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

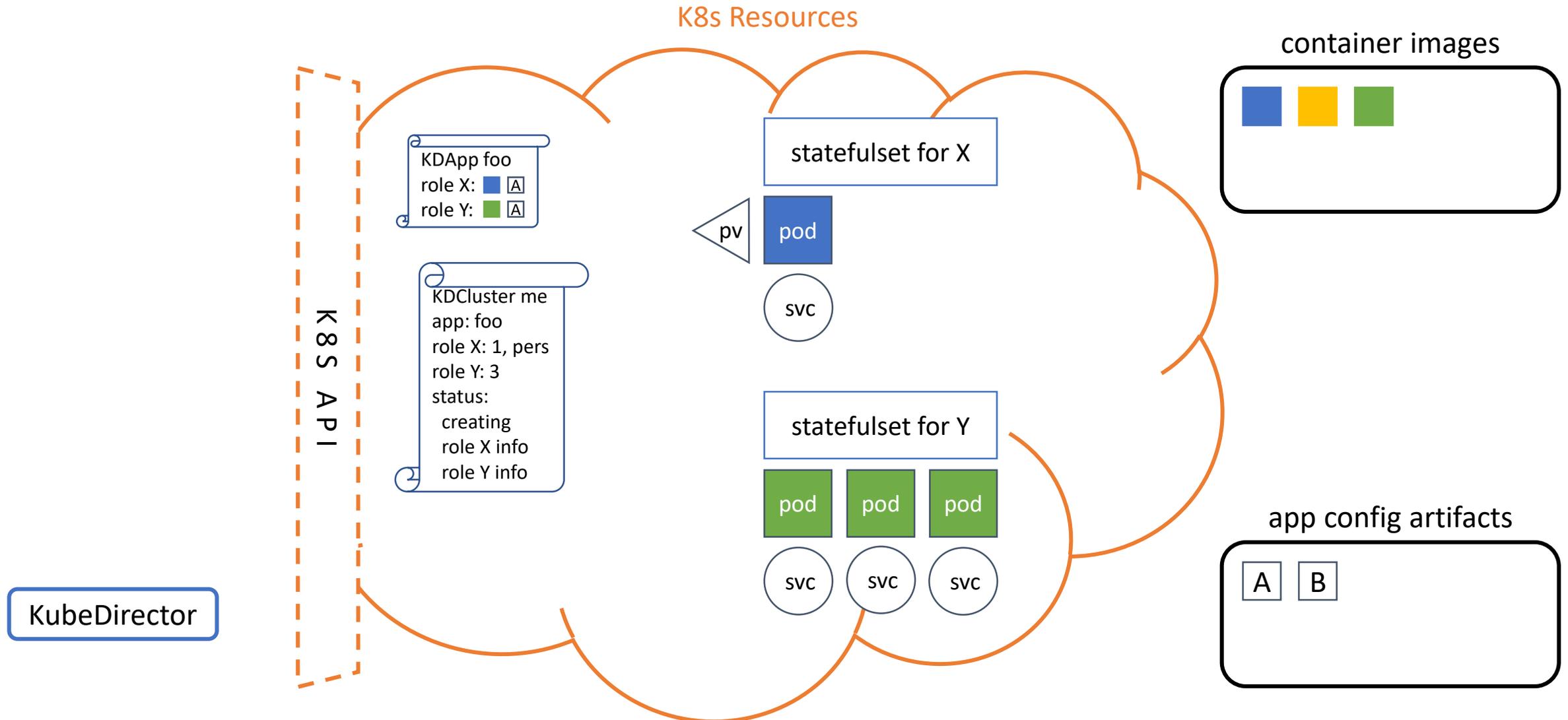


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

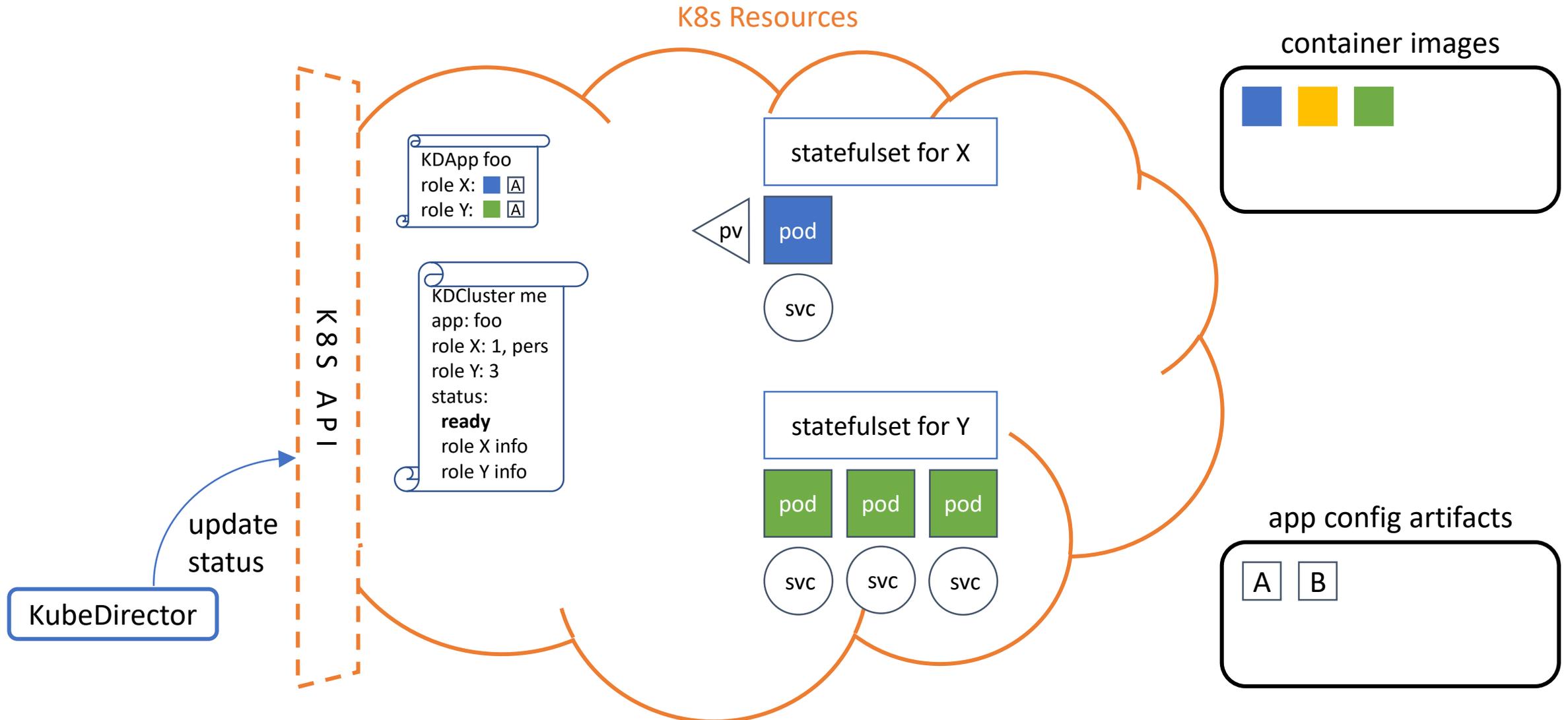


KubeCon



CloudNativeCon

North America 2019



KubeDirectorCluster Deploy

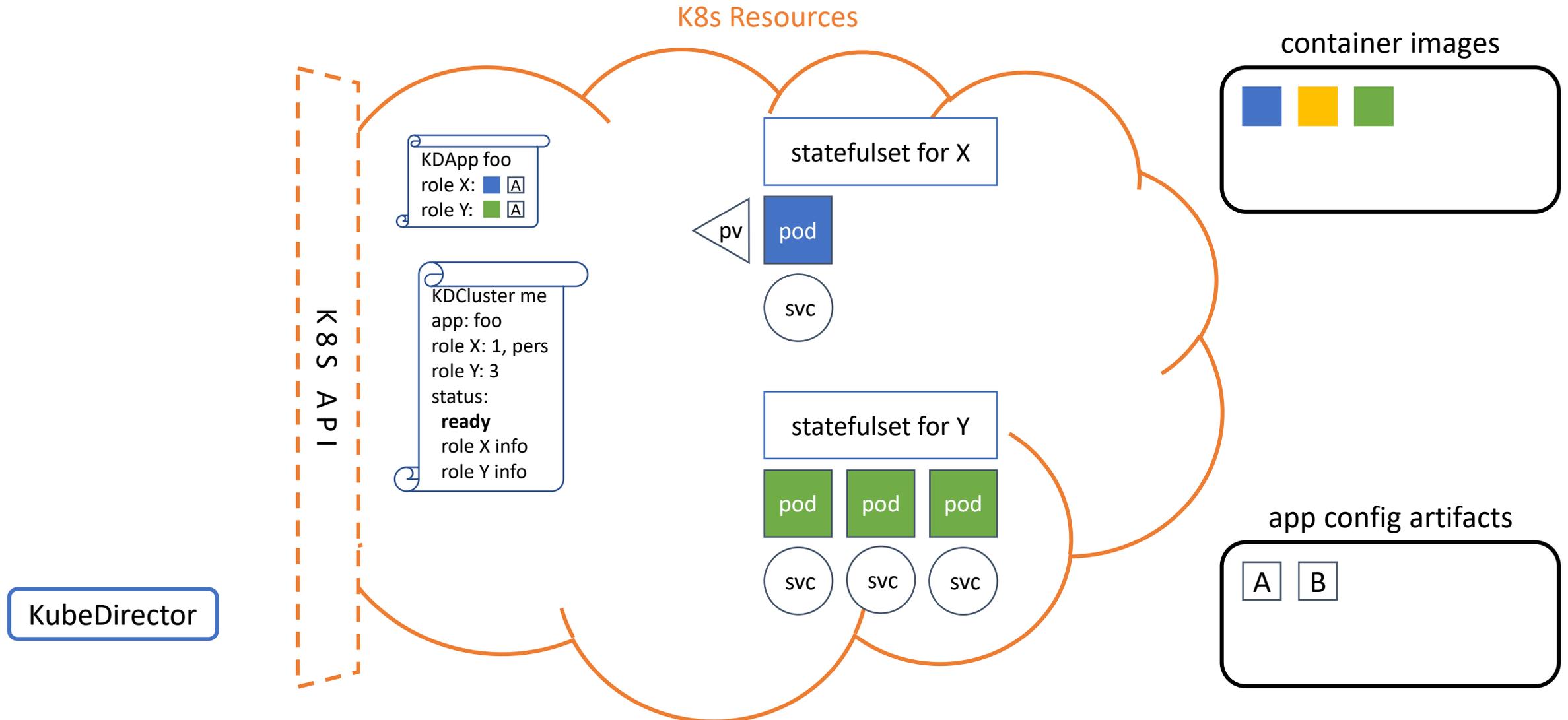


KubeCon



CloudNativeCon

North America 2019





KubeCon



CloudNativeCon

North America 2019

KD Architecture Benefits

...aka, “Why Do It Like That?”



Better Processes

- Less new things needed to add support for new/modified app
 - Don't need to know how to create/modify an operator
 - Don't need to rely on or certify a new operator from elsewhere
 - Don't need to register new CRDs and update all user/group ACLs
 - Don't need to (radically) change clients to deal with new schemas
- Advantages from separate app-type/app-instance CRs
 - KubeDirectorApp enshrines app-specific validation rules, e.g. resource minimums
 - Can optionally enforce “separation of concerns” for different user groups
 - Cluster admins, app experts, project administrators, data scientists
 - Some should be able to create/modify KubeDirectorApp; some just use what's there
- Obviously only works if our model captures the interesting bits
 - One reason we focus on a specific app domain/ecosystem

Handling Multiple App Types



KubeCon



CloudNativeCon

North America 2019

- Common-format app characteristics helps build pipelines
 - Specify “connected” app instances in KubeDirectorCluster
 - In-container metadata file also has metadata of connected app instances
 - App type, IPs, service endpoints, etc.
 - Connections metadata updated on lifecycle events, + optional script hook
 - App instance learns about its connections w/o needing K8s credentials
- Cross-app-type features “just work”
 - Specify a new app type, get access to existing KD features
 - Even if feature requires certain things in resource object creation specs
 - If label-driven, don't need to give clients write access to underlying resources



KubeCon



CloudNativeCon

North America 2019

Takeaways



State of the Project



KubeCon



CloudNativeCon

North America 2019

State of the Project



KubeCon



CloudNativeCon

North America 2019

- 0.3.0 release recently at <https://github.com/bluek8s/kubedirector/>
 - Based on controller-runtime now
 - App resource requirements, GPUs, file injection...

State of the Project



KubeCon



CloudNativeCon

North America 2019

- 0.3.0 release recently at <https://github.com/bluek8s/kubedirector/>
 - Based on controller-runtime now
 - App resource requirements, GPUs, file injection...
- Some upcoming features
 - Config choices (app config variations picked at instantiation)
 - Error recovery policies
 - Encrypted secrets

State of the Project

- 0.3.0 release recently at <https://github.com/bluek8s/kubedirector/>
 - Based on controller-runtime now
 - App resource requirements, GPUs, file injection...
- Some upcoming features
 - Config choices (app config variations picked at instantiation)
 - Error recovery policies
 - Encrypted secrets
- Some upcoming developer support
 - Factoring out common useful CR-processing utilities
 - Revision & documentation of tool used to access in-container metadata

State of the Project



KubeCon



CloudNativeCon

North America 2019

- 0.3.0 release recently at <https://github.com/bluek8s/kubedirector/>
 - Based on controller-runtime now
 - App resource requirements, GPUs, file injection...
- Some upcoming features
 - Config choices (app config variations picked at instantiation)
 - Error recovery policies
 - Encrypted secrets
- Some upcoming developer support
 - Factoring out common useful CR-processing utilities
 - Revision & documentation of tool used to access in-container metadata
- Expanding the example app catalog

State of the Project



KubeCon



CloudNativeCon

North America 2019

- 0.3.0 release recently at <https://github.com/bluek8s/kubedirector/>
 - Based on controller-runtime now
 - App resource requirements, GPUs, file injection...
- Some upcoming features
 - Config choices (app config variations picked at instantiation)
 - Error recovery policies
 - Encrypted secrets
- Some upcoming developer support
 - Factoring out common useful CR-processing utilities
 - Revision & documentation of tool used to access in-container metadata
- Expanding the example app catalog

In Summary...



KubeCon



CloudNativeCon

North America 2019

In Summary...



KubeCon



CloudNativeCon

North America 2019

- Running complex stateful applications on Kubernetes is challenging today

In Summary...



KubeCon



CloudNativeCon

North America 2019

- Running complex stateful applications on Kubernetes is challenging today
- The goal of BlueK8s and KubeDirector is to make it easier to run such applications on Kubernetes

In Summary...



KubeCon



CloudNativeCon

North America 2019

- Running complex stateful applications on Kubernetes is challenging today
- The goal of BlueK8s and KubeDirector is to make it easier to run such applications on Kubernetes
- KubeDirector is a key component of the HPE Container Platform
 - hpe.com/info/container-platform

In Summary...



KubeCon



CloudNativeCon

North America 2019

- Running complex stateful applications on Kubernetes is challenging today
- The goal of BlueK8s and KubeDirector is to make it easier to run such applications on Kubernetes
- KubeDirector is a key component of the HPE Container Platform
 - hpe.com/info/container-platform
- Learn more about KubeDirector:
 - <https://github.com/bluek8s/kubedirector/wiki>

In Summary...



KubeCon



CloudNativeCon

North America 2019

- Running complex stateful applications on Kubernetes is challenging today
- The goal of BlueK8s and KubeDirector is to make it easier to run such applications on Kubernetes
- KubeDirector is a key component of the HPE Container Platform
 - hpe.com/info/container-platform
- Learn more about KubeDirector:
 - <https://github.com/bluek8s/kubedirector/wiki>

Tom Phelan

 @tapbluedata

Joel Baxter

 @joel_k_baxter