KubeCon
CloudNativeCon
Europe 2023
TiKV

Distributing and Decentralizing Pod Disruption Budget (PDB)

Illya Chekrygin

Distributed PDB

Disruption* Sucks

* disturbance or problems which interrupt an event, activity, or process

Types of Disruption

Involuntary

- Hardware, VM, Network Failures
  - Machine failure
  - VM failure
    - accedental deletion
    - Kernel panic
  - Network partition
  - Kubelet Eviction (out of resources)
  - Taint based Eviction

Voluntary

- Caused or consented by the workload owner
  - Deletion
  - Update
  - Automation
- Caused by Infrastructure or Admin
  - Scheduler Eviction

Disruption Protection - PDB


```
apiVersion: policy/v1
kind: PodDisruptionBudget
metadata:
 name: test
  namespace: default
spec:
  minAvailable: 3
  selector:
    matchLabels:
      app: test
status:
  currentHealthy: 6
  desiredHealthy: 3
  disruptionsAllowed: 3
  expectedPods: 6
```

Per Namespace

“Healthy” pod requirements (absolute or %-age)

Pod Selector

How many “healthy” pods are there at this moment?

How many “healthy” pods are required?

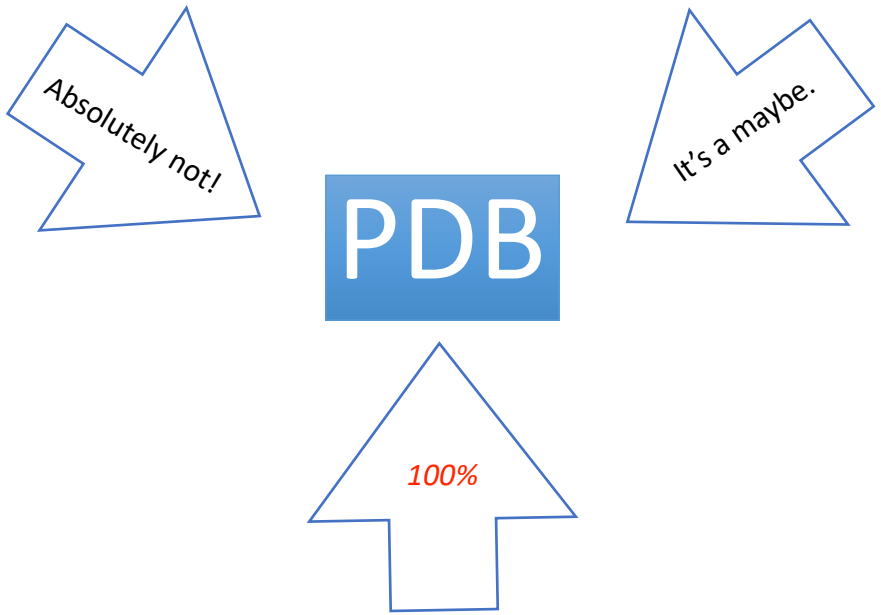
How many can be disrupted at this moment?

How many “healthy” pods are expected?

PDB - Support in K8s

Kubelet Eviction
(out of resources)

Tain based Eviction



Scheduler Eviction

Eviction API
(kubectl drain)

PDB

The Good

Simple

Eviction Protection

The Not So Good

Rudimentary (too simple)

Namespace Scoped

Not Universal

Does not support complex identity pods

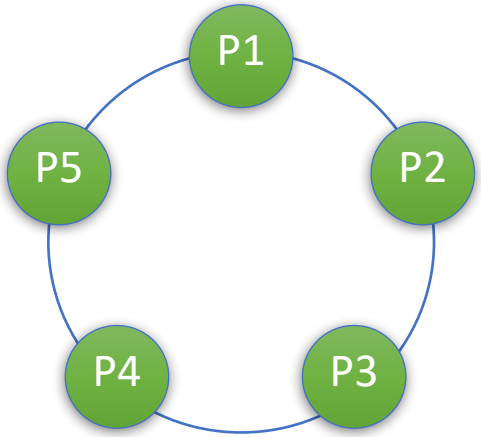
Not extensible

The Very Not So Good

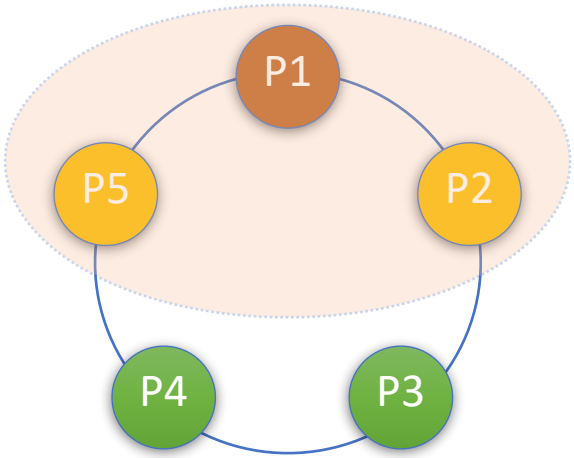
Err500 on multiple PDBs per Pod

Not extensible

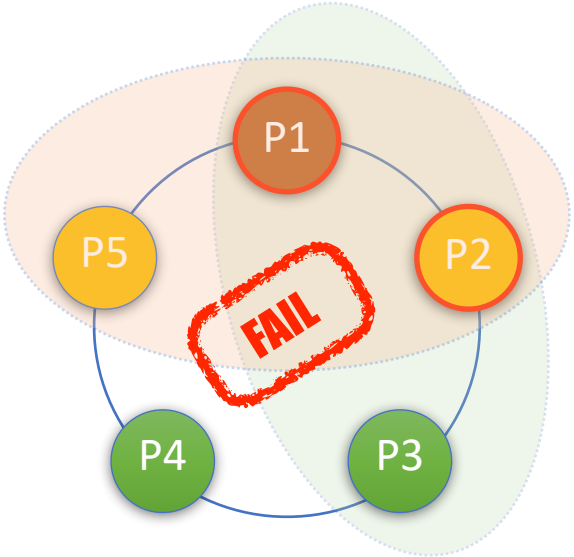
Distributed NoSQL DB Use case



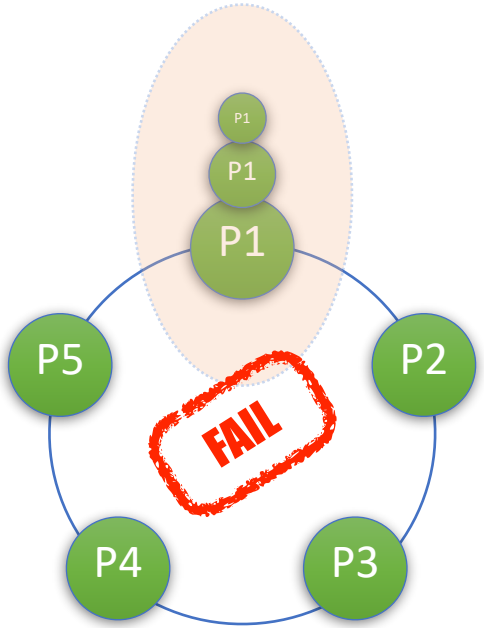
P1: P5,P2
P2: P1,P3
P3: P2,P4
P4: P3,P5
P5: P4,P1



P1: P5,P2
P2: P1,P3
P3: P2,P4
P4: P3,P5
P5: P4,P1



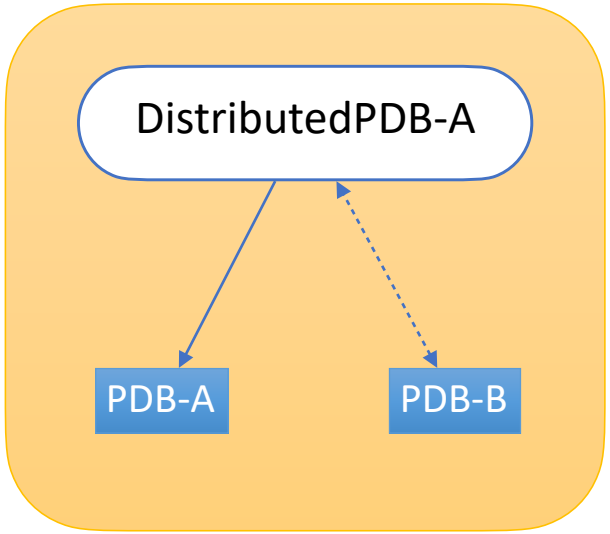
PDB1[P1,P5,P2]
PDB2[P2,P1,P3]
…



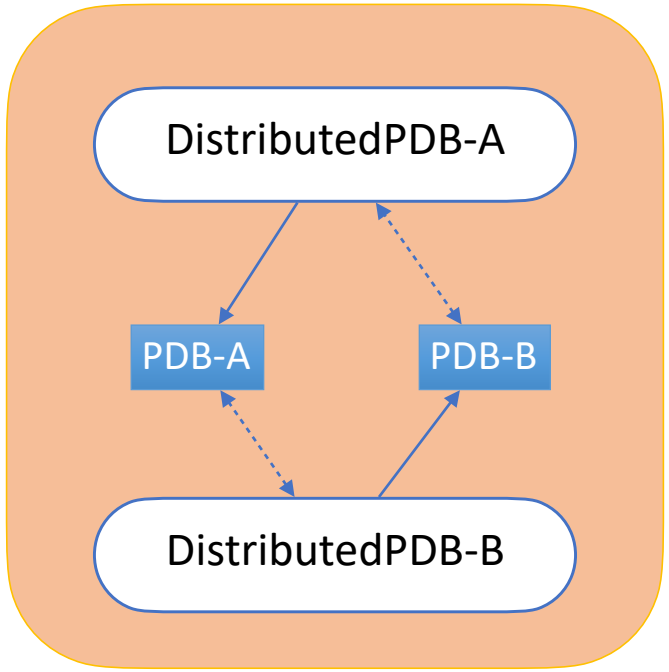
[P1,P5,P2] x N
…

Federated PDB

Unidirectional

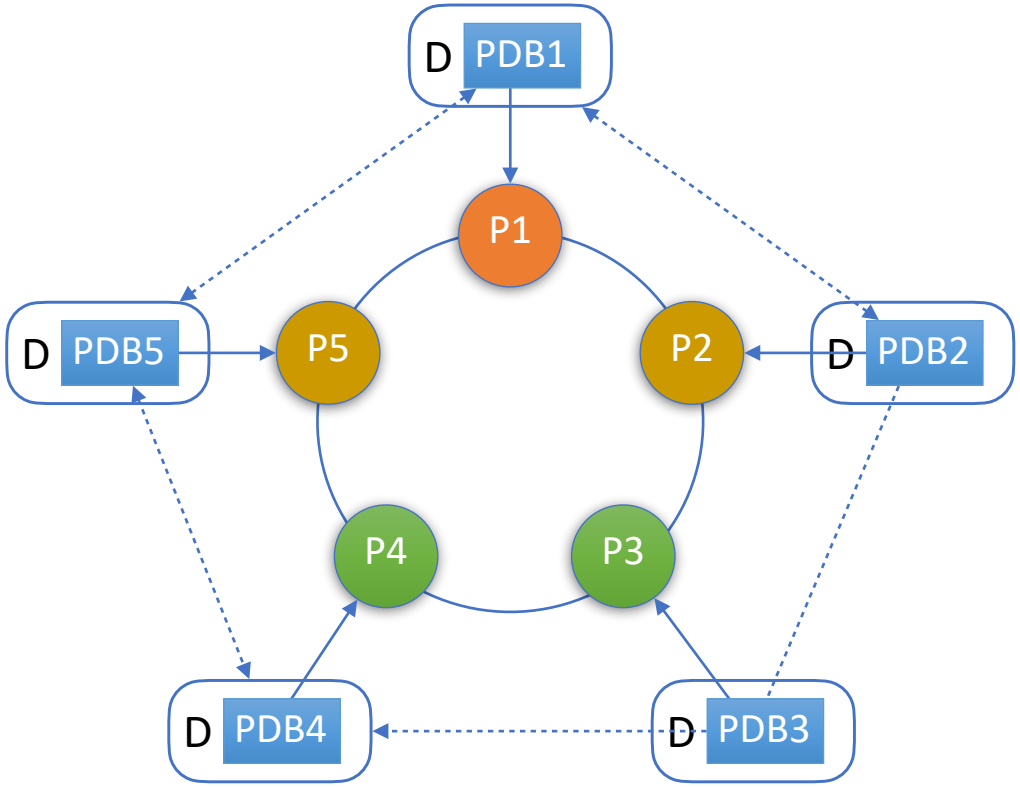


Bidirectional

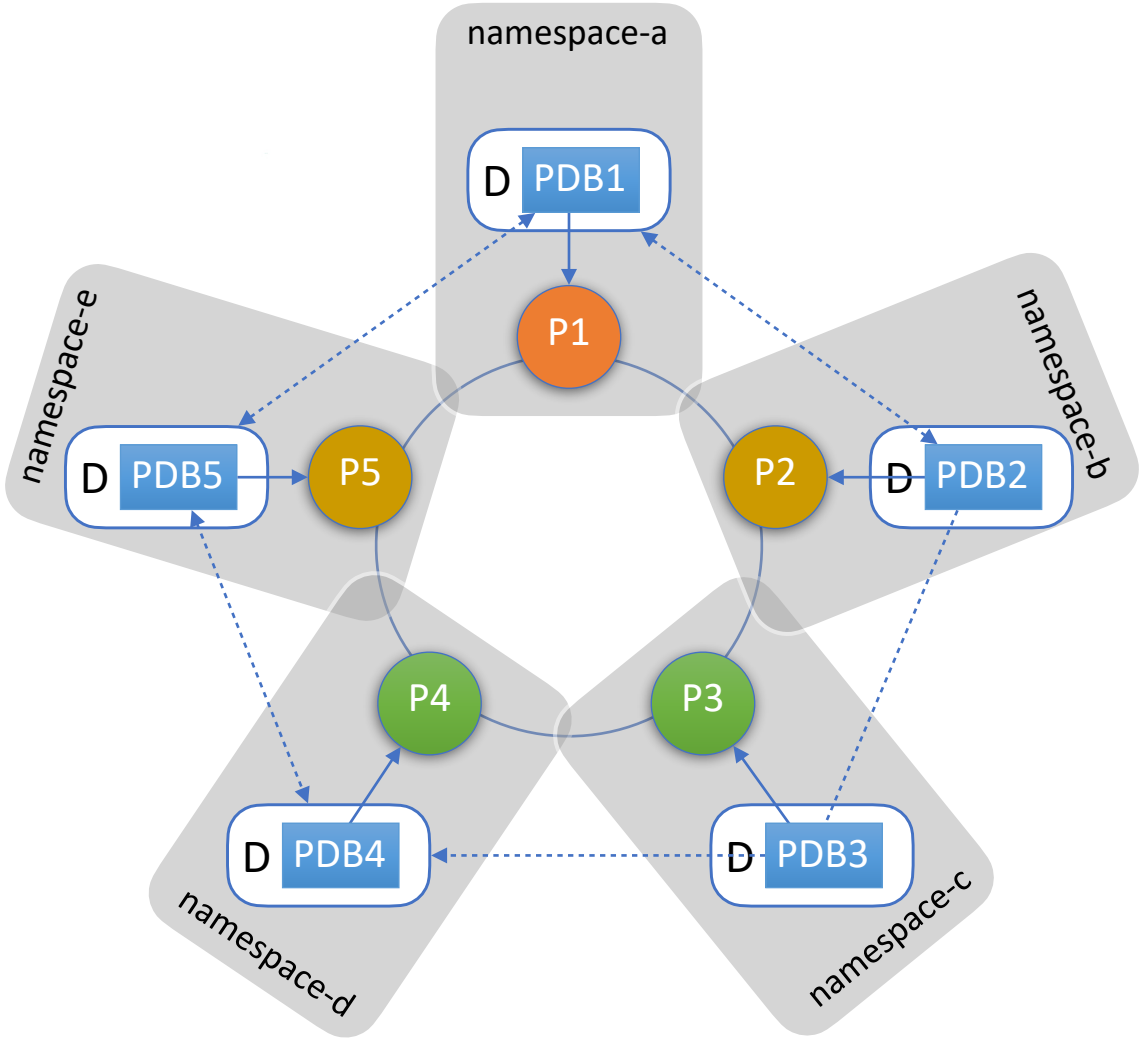


Distributed DB Federated PDB

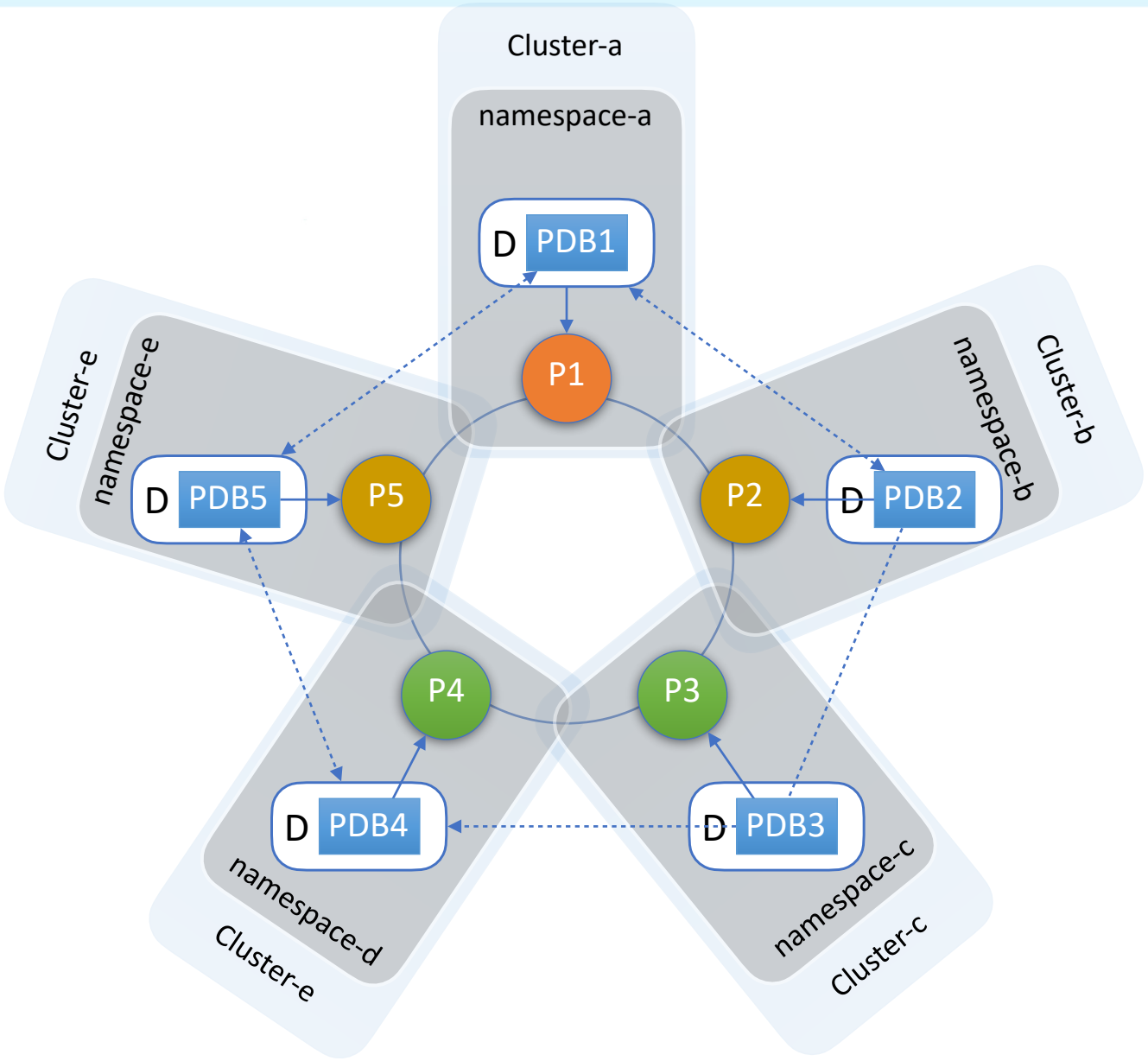

```yaml
apiVersion: demo.illya.com/v1alpha1
kind: DistributedPodDisruptionBudget
metadata:
  name: instance-1
spec:
  minAvailable: 2
  selector:
    matchLabels:
      app: instance-1
  federation:
    - name: instance-5
    - name: instance-2
```



Multi-Namespace PDB



Multi-Cluster PDB



KubeCon
CloudNativeCon
Europe 2023
DEMO TIME

Challenges

Extensibility

Or the lack thereof for built-in types.

Path Forward

KubeCon CloudNativeCon Europe 2023

Your Story

Thank you!

https://github.com/apple/distributed-disruption

https://github.com/ichekrygin

Slack:

Kubernetes - ichekrygin

CNCF: Illya Chekrygin

https://github.com/apple/distributed-disruption

https://github.com/ichekrygin

Slack:

Kubernetes - ichekrygin

CNCF: Illya Chekrygin

to leave feedback on this session