

Kubernetes infrastructure

for CERN's Content Management Systems

Konstantinos Samaras-Tsakiris
IT-CDA-WF



Coalescing Web Frameworks on K8s

Unique infrastructure per use case

PaaS

Openshift 3 (Kubernetes)

Static sites

VMs with custom config

CMS

Physical machines with custom config

*Content
Management
Systems
(currently **Drupal**)*

Low reuse of components



Converging on a cloud native platform

PaaS

Openshift **4** (Kubernetes)

Static sites

Openshift 4

CMS

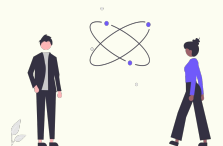
Openshift 4

Many **shared** components

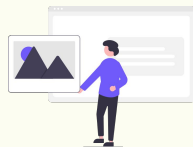
Drupal @ CERN



Drupal @ CERN



Physicists



Administration



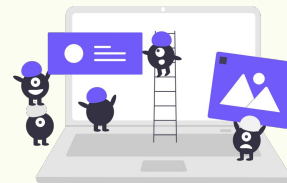
Drupal expert



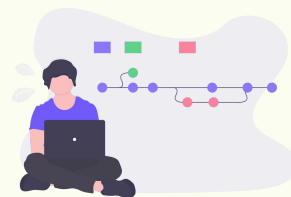
Comms expert



CERN
Drupal Distribution



Reliable
hosting



Custom
modules &
themes

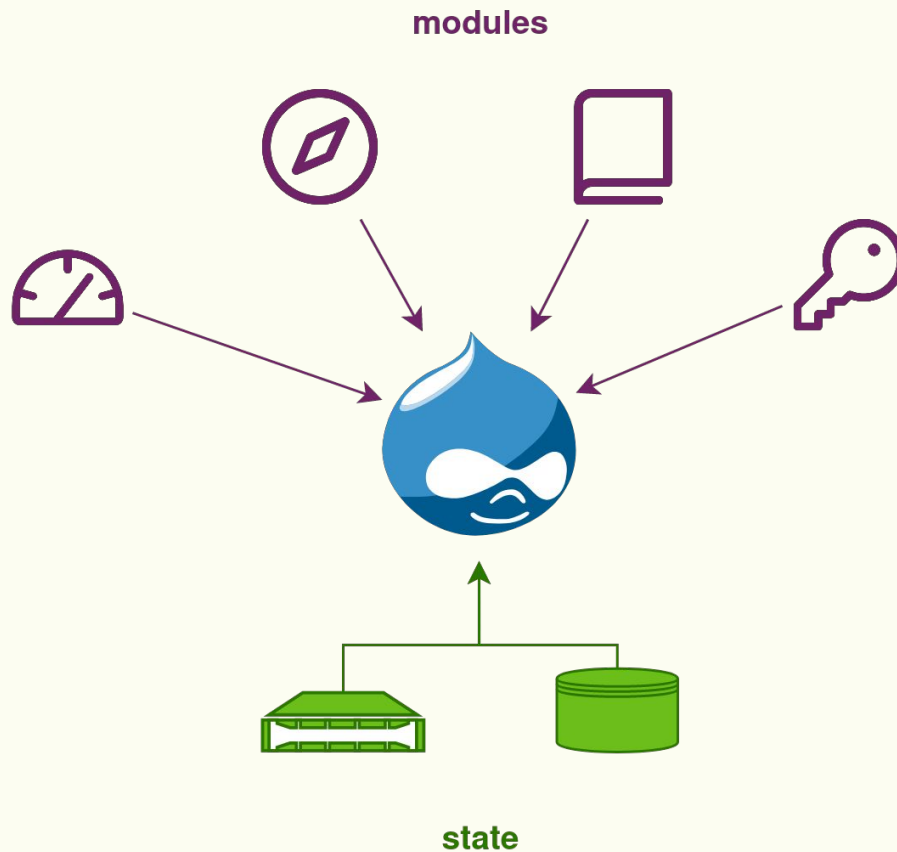


Upgrades
Failure recovery

Not just hosting, but fully managed

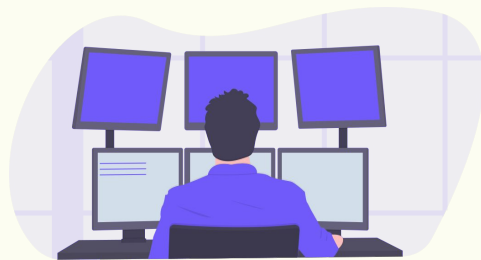
Software as a Service

Parts of a Drupal site



SaaS requirements

- Take **1500** instances of a complicated thing
- Automate business/operational logic
- Let users **self-provision** websites
- All this with a very small team!



Drupal sites on K8s



“Container orchestrator”

- not a workflow engine
- set of independent, **composable** control processes
- continuously drive the current state towards the provided desired state

API resources



OKD4 (OpenShift) cluster

website project

DrupalSite



OKD4 (OpenShift) cluster

components



argo



drupal site
operator



authorization
operator



external DB
operator

website project

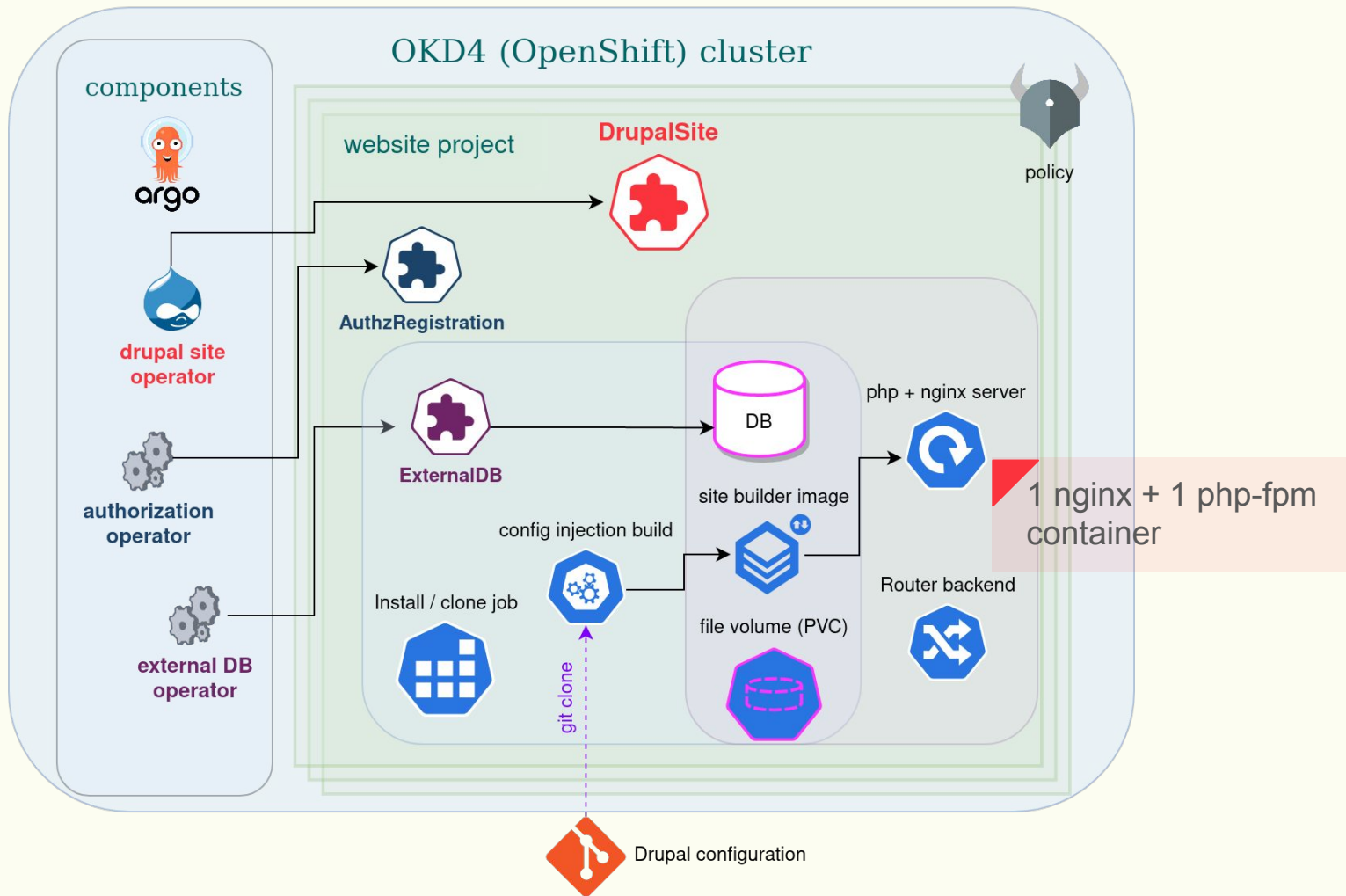


AuthzRegistration

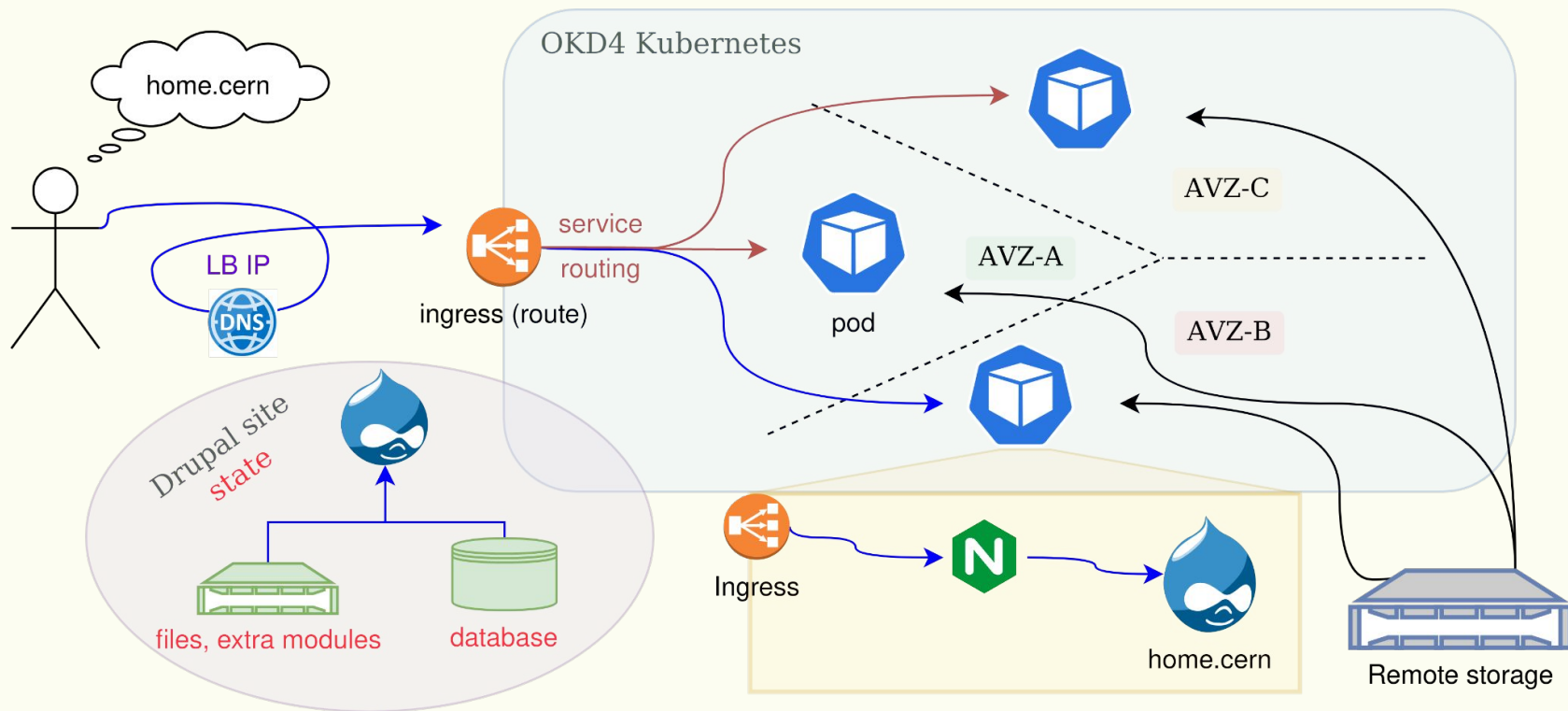
DrupalSite



policy



Request journey



Declarative API: desired / current state

```
apiVersion: drupal.cern.ch/v1alpha1
kind: DrupalSite
metadata:
  name: vcchep-2021
spec:
  publish: true
  drupalVersion: "8.9.14-72ac18b"
  environment:
    name: "production"
    qosClass: "standard"
    dbodClass: "test"
  diskSize: "1Gi"
```

```
status:
  conditions:
    - type: Initialized
      status: "false"
    - type: Ready
      status: "False"
    - type: DBUpdatesPending
      status: Unknown
      reason: K8sAPIClientError
      message: 'K8sAPIClientError:
Deployment.apps "vcchep-2021" not found'
      failsafeDrupalVersion: "8.9.13-37d1b8a"
```



Does it work?

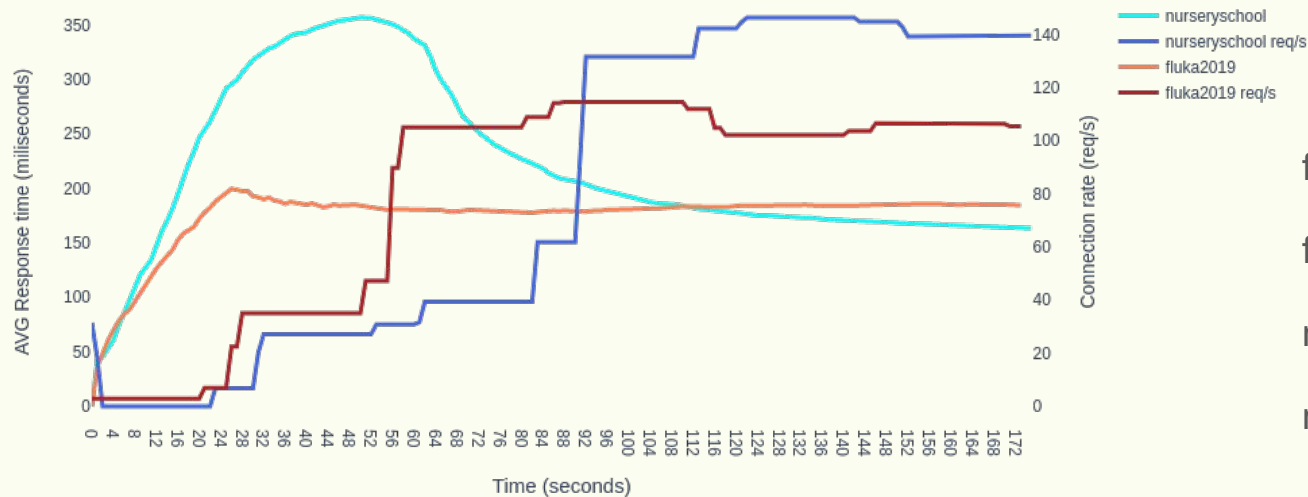


Pre-production

- Preparing to migrate all websites
- Automated provisioning / updates
- Flexible settings: tune critical websites, conserve resources

PaaS, static site infrastructures **already in production**

Experiment: web crawling



Memory consumption

fluka2019 idle	100MB
fluka2019 max	270MB
nurseryschool idle	100MB
nurseryschool max	390MB

Stress test: crawl a typical site with many clients simultaneously
(very large sites behave worse)

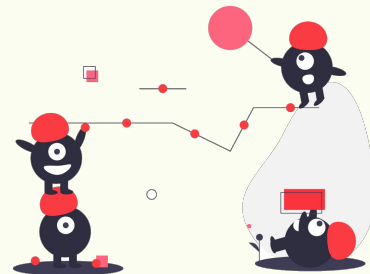
Estimate resources to sustain traffic according to site's QoS class

- Extrapolating to 300GB memory for the entire infrastructure
(2TB in the old infrastructure)

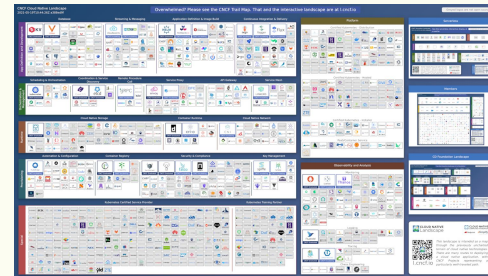
Containerization gives
opportunity to **conserve**
resources

Conclusions

- We can provision a highly **automated** infrastructure to solve a complex problem with a **very small team** →



- Kubernetes API: observability, isolation, high availability “for free”
- Take advantage of the CNCF Landscape →



gitlab.cern.ch/drupal/paas/drupalsite-operator

↙
Apache licence

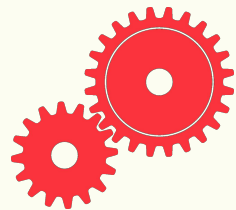


That's all Folks!

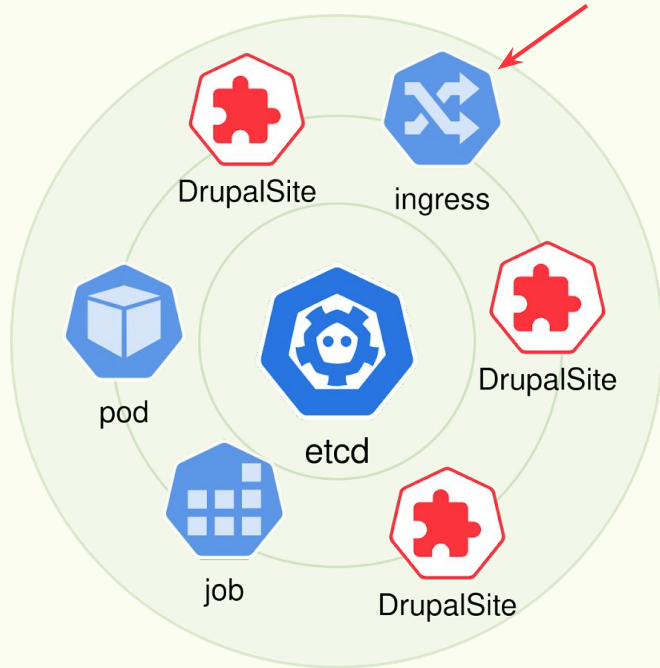
Check out our Kubecon presentation for more details:

<https://youtu.be/4O-pcSQR8Vw>

Appendix Operator Pattern



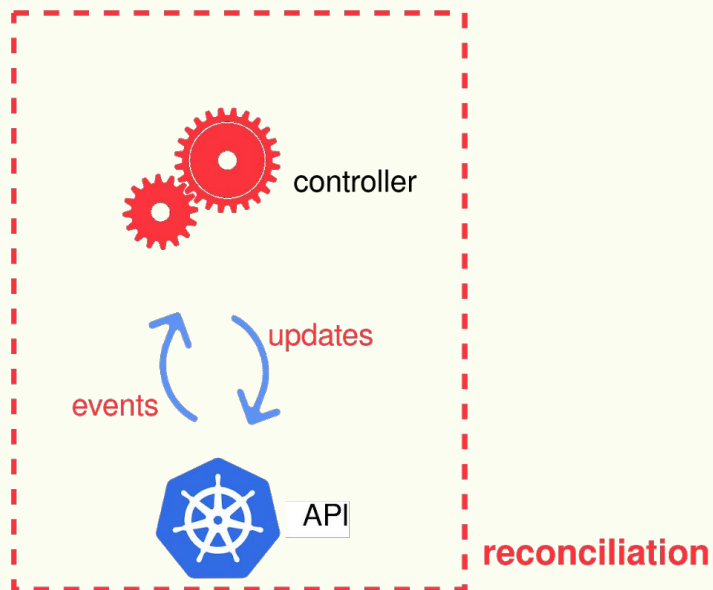
Standard API resource

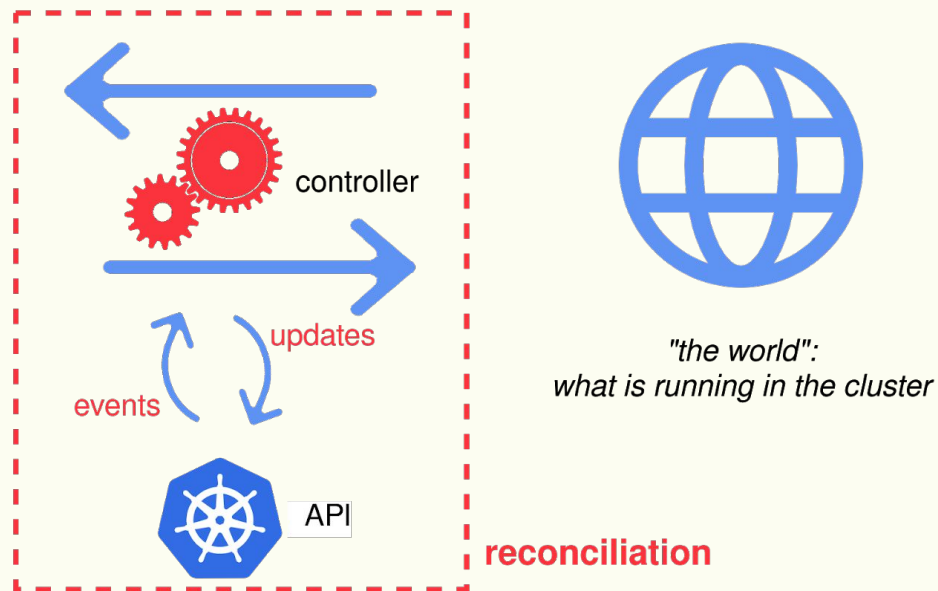
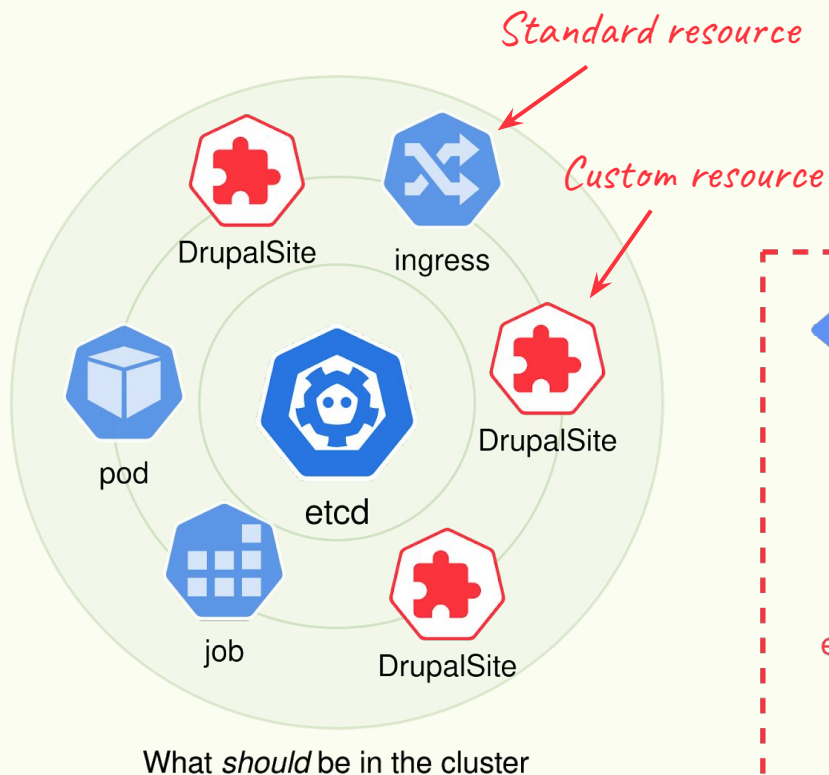


What *should* be in the cluster



What *should* be in the cluster





Operator a custom resource and controller, conceptually similar to an OOP **class** with data & methods