

# High Availability Alarm System Deployed With Kubernetes

---

Jesse Bellister  
12 October 2023

# Introduction

---

## Main Goals for the Project

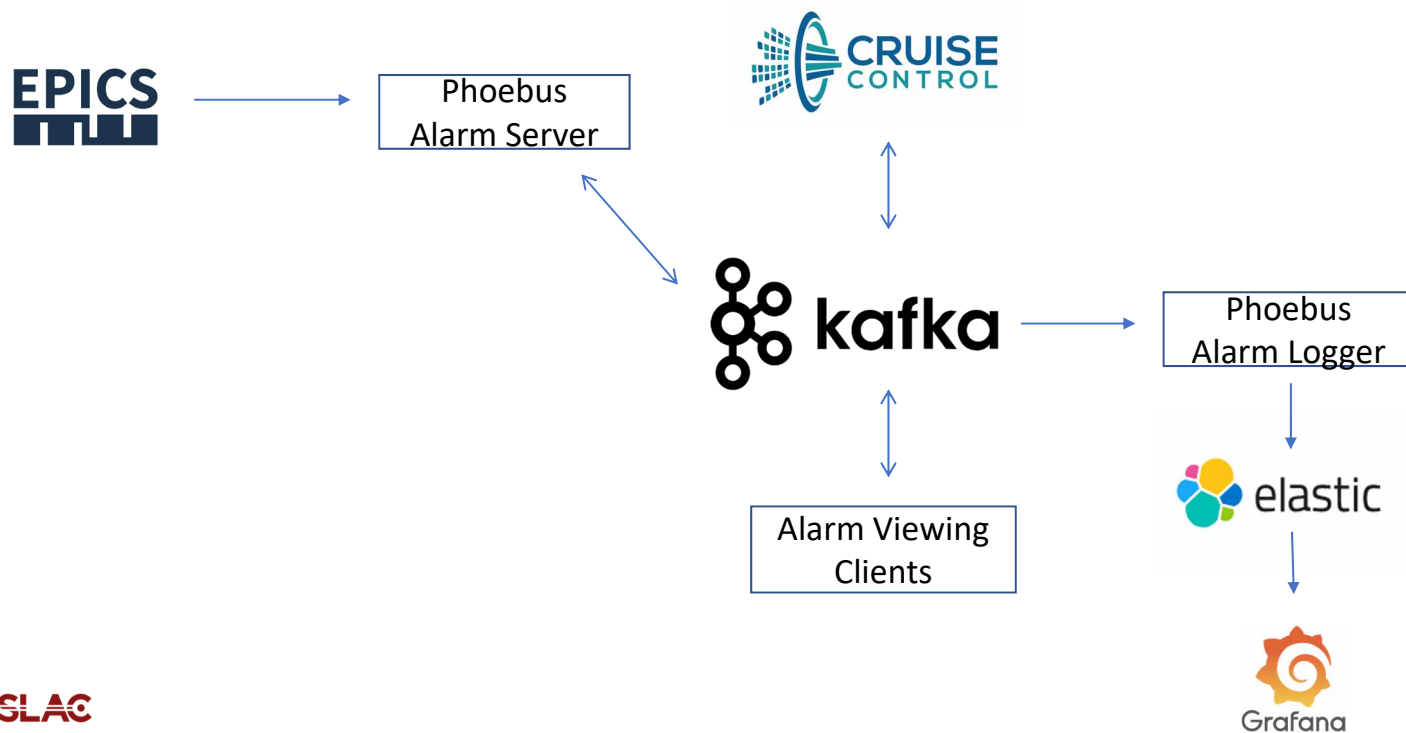
- Increased operator engagement with the alarm system
- Easy to keep the system updated as devices are added or removed
- Robust deployment with near constant uptime
- Keep in line with current best practices for technologies used

## Stakeholders

- Accelerator Operators, Scientists, End-Users
- IOC Engineers
- System Administrators

# Phoebus Alarm Server

---

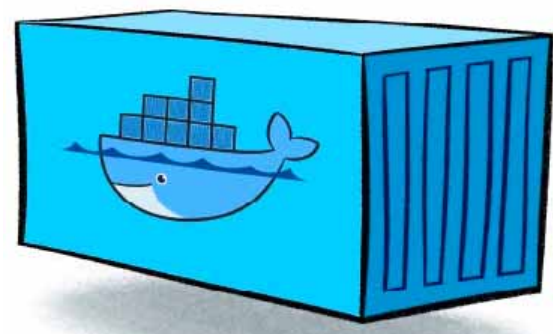


# Containerization

---

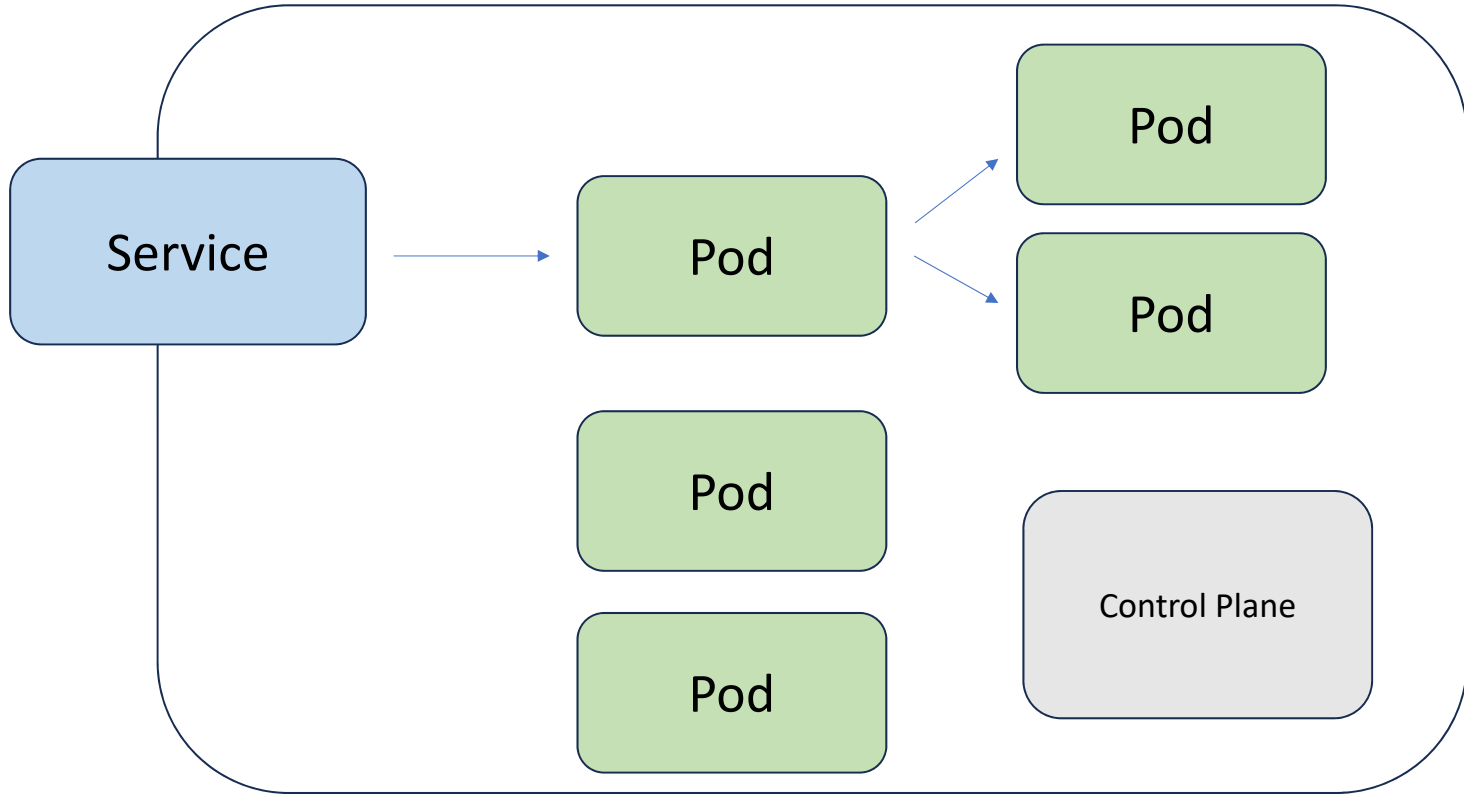
## Why?

- All dependencies are easy to track
- Build anywhere
  - Don't like yum? Can use apt instead even if the host OS is CentOS
- Consistent environment
  - Self-documenting, file for generating container is version controlled with git
  - Shareable
- Reasonable learning curve
- But how do we want to deploy and manage these containers?



# Kubernetes

---



# Kubernetes

---

## What are the benefits?

- High Availability
  - Kubernetes control plane is constantly monitoring all components of the deployment
  - If a node goes down, pods running on it are automatically moved to a different node
  - Set a number of replicas for each component
- Declarative Configuration
  - Tell the cluster what you want, not how to do it
  - “Give me 2 alarm servers and 3 kafka brokers, at least one of each must always be running”
- Deployments are easily replicated
  - Manifest files are defined with yaml, version controlled with git
  - Deployment of entire system can be done with a one line command “kubectl apply -k base”



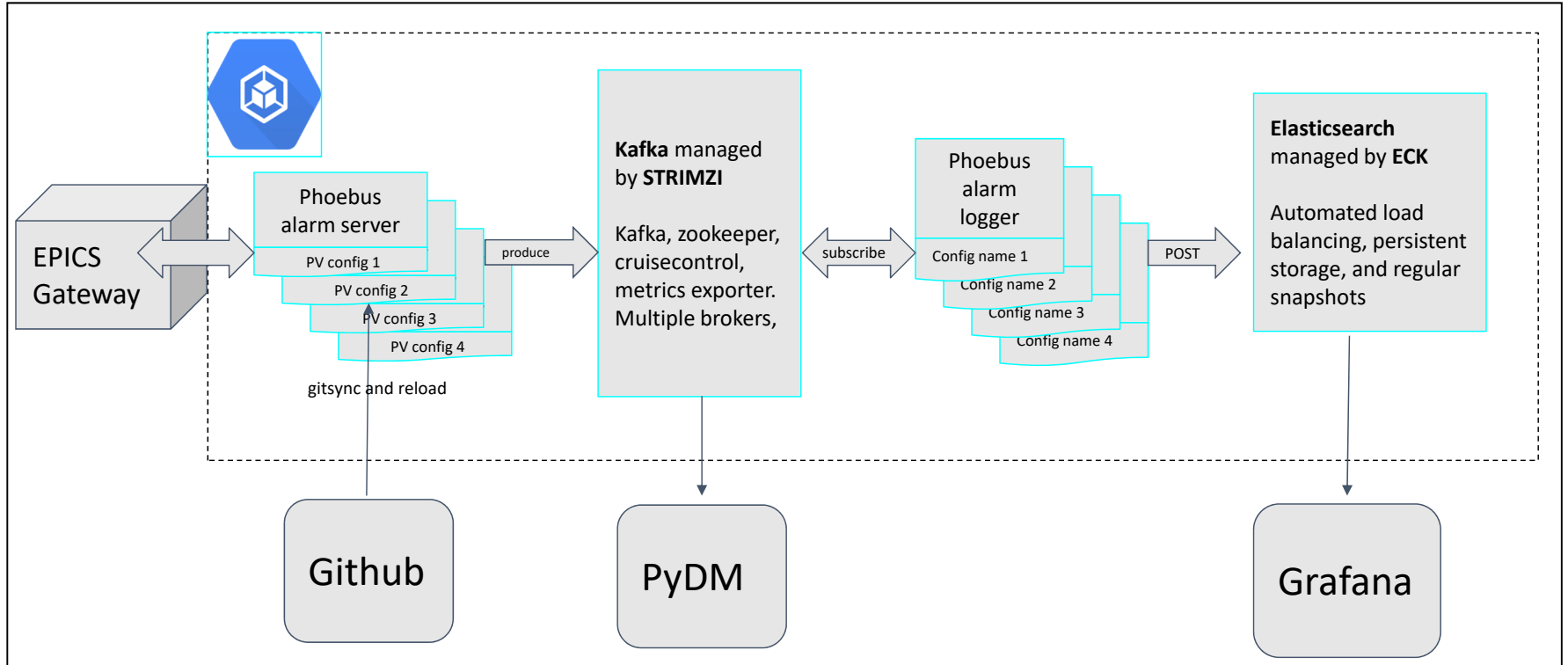
# Alarm System Deployment

---

## Focus on High Availability and Ease of Upgrade

- Kubernetes cluster is running in the SLAC Shared Scientific Data Facility (S3DF)
  - Virtual cluster (vcluster) isolates our system from everything else in the cluster
- Multiple replicas of each Phoebus alarm server are running
  - Pod anti-affinity ensures each copy runs on a separate node
  - Pod disruption budget ensures at least one copy is always running
  - Set a number of replicas for each component
- Operators are used to manage Kafka and Elasticsearch
  - Deployment pattern that attempts to automate repeatable tasks that would usually be handled by system operators
  - Make updates to new versions easy

# Alarm System Deployment





# Deployment Updates

---

## Make Editing Alarm Hierarchies as Simple as Possible

- Continuous Integration with GitHub Actions
  - User edits a csv file to reflect their changes and submits a pull request
  - GitHub action validates the change, updates the associated xml file, and lints that file
  - PR still requires approval before being merged
- Continuous Deployment with Kubernetes Sidecar
  - A sidecar is a deployment pattern in which the sidecar container enhances the functionality of the main container by running alongside it in the same pod
  - Our sidecar uses gitsync to monitor the git repository with the xml files
  - When a change is detected, it pulls in the new xml file and uses kcat to issue the alarm server restart command to kafka

# Python Alarm Manager

SLAC Alarm Manager

File Tools

GMDXGMD

- ▼ GMD
  - ▶ GMD ATCA system
  - ▶ GMD Keithley
  - ▶ GMD Machine input data
  - ▶ GMD chamber temperature
  - ▼ GMD high voltage
    - EM2K0:YGMD:SHV:01:M0:C0:VoltageMeasure
    - EM2K0:YGMD:SHV:01:M0:C0:isOn
    - EM2K0:YGMD:SHV:01:M0:C1:VoltageMeasure
    - EM2K0:YGMD:SHV:01:M0:C1:isOn
    - EM2K0:YGMD:SHV:01:M0:C14:VoltageMeasure
    - EM2K0:YGMD:SHV:01:M0:C14:isOn
    - EM2K0:YGMD:SHV:01:M0:C15:VoltageMeasu...
    - EM2K0:YGMD:SHV:01:M0:C15:isOn - MINOR...
  - ▶ GMD spinning rotor gauge
- ▼ XGMD
  - ▶ XGMD ATCA system
  - ▶ XGMD Keithley
  - ▶ XGMD chamber temperature
  - ▶ XGMD high voltage
  - ▶ XGMD machine input data
  - ▶ XGMD spinning rotor gauge

Active Alarms: 13

Filter

PV	Latched Severity	Current Severity	Description	Time	Value	
EM1K0:GMD:HPS:Nu...	INVALID	INVALID	Number of samples for pulse		12	U
EM1K0:GMD:ETM:...	INVALID	INVALID	Reading to average		0	C
EM2K0:YGMD:SHV:...	MINOR	MINOR	Electron multiplier voltage		0.20038128	H
EM2K0:YGMD:SHV:...	MINOR	MINOR	Electron multiplier powered on	2023-07-07 21:12:52	Off	S
EM2K0:YGMD:HPS:Nu...	INVALID	INVALID	Number of samples for pulse	2023-09-28 04:44:30	13	U
EM2K0:YGMD:ETM:...	INVALID	INVALID	Measurement function Keithley 1	2023-09-28 08:40:50	Current	S
EM2K0:YGMD:ETM:...	INVALID	INVALID	Measurement function Keithley 2	2023-09-28 08:40:50	Current	S
EM2K0:YGMD:SHV:...	MINOR	MINOR	Electron multiplier voltage	2023-09-29 10:03:43	0.22091386	H
EM2K0:YGMD:SHV:...	MINOR	MINOR	Electron multiplier powered on	2023-07-07 21:12:52	Off	S
EM1K0:GMD:GSR:...	INVALID	INVALID	Rotation speed	2023-09-28 16:36:07	435.0	T
EM1K0:GMD:GSR:...	INVALID	INVALID	Pressure	2023-09-19 13:50:12	1....	L
TPR:EM2K0:YGMD:...	MAJOR	OK	Timing status	2023-10-02 18:04:25	Link Down	S
TPR:EM1K0:GMD:...	MAJOR	OK	Timing status	2023-10-02 18:04:25	Link Down	S

Acknowledged Alarms: 0

Filter

PV	Latched Severity	Current Severity	Description	Time	Value	
----	------------------	------------------	-------------	------	-------	--

# Slack Integration

---

Yesterday ▾

 **Cryo Alarms** APP 4:50 PM  
CANL:CP12:ANL2:ALM Severity: MAJOR Message: STATE\_ALARM Description: CP0 CANL - C1-ANL-23822 High

 **Cryo Alarms** APP 7:30 PM  
CPT:CP08:LN1:ALM Severity: MAJOR Message: STATE\_ALARM Description: CP0 LN2D - CPT-88101 Press. Low

 **Cryo Alarms** APP 8:36 PM  
CLT:CP08:LN6:ALM Severity: MAJOR Message: STATE\_ALARM Description: CP0 LN2D - CLT-88101 Level High

Today ▾

 **Cryo Alarms** APP 10:31 AM  
WCMP:CP12:14:C100\_ALARM Severity: MAJOR Message: STATE\_ALARM Description: CP1 WCMP4 - Warm compressor 4  
CPT:CP12:14:ALM Severity: MAJOR Message: STATE\_ALARM Description: CP1 WCMP4 - Suction Pressure Low  
CPT:CP12:G15:ALM Severity: MAJOR Message: STATE\_ALARM Description: CP1 GMGT - MP Line Pressure Low

# Conclusion

---

- Kubernetes learning curve is fairly steep
  - Having dedicated personnel resources helps!
- End user response has been largely positive
- Future projects at SLAC are already slated to use Kubernetes based deployments