

# **Robot-shop Production-Grade DevSecOps, GitOps & FinOps Platform on AWS Cloud-Native Kubernetes Architecture**



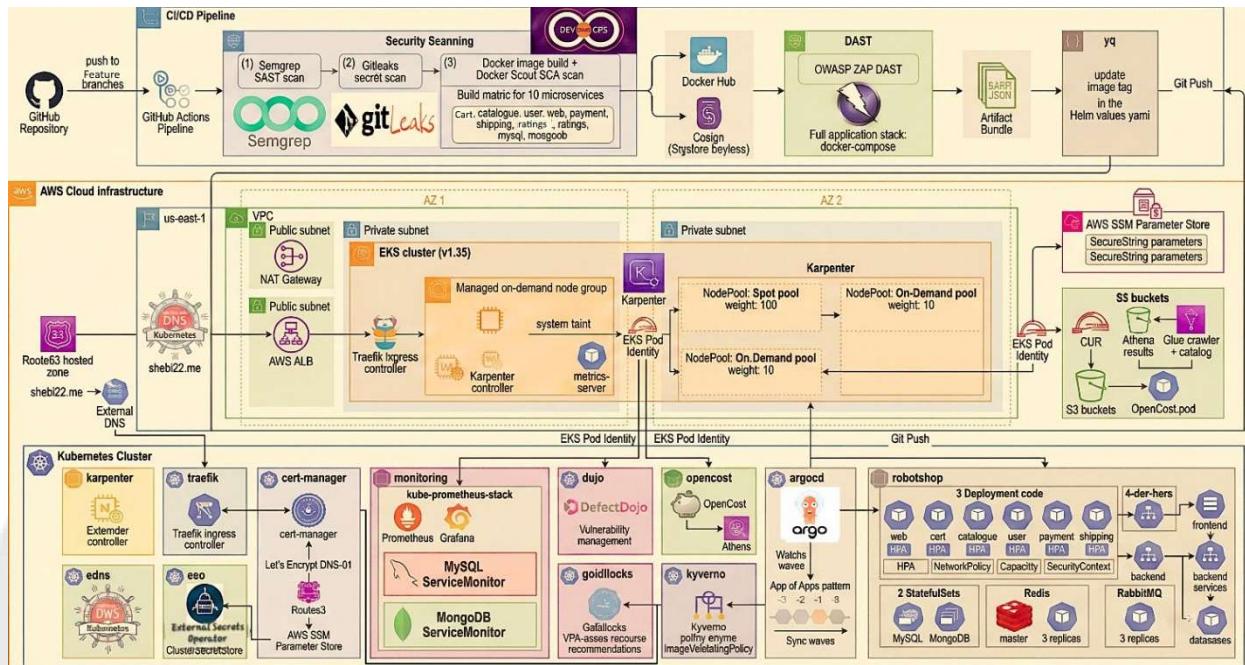
By Eng: Abdelrahman Mahmoud

Abdelrahman Ayman

## Contents

Introduction .....	4
App overview .....	5
Docker layer .....	6
Kubernetes layers .....	8
Umbrella Helm .....	10
Our Charts .....	12
Karpenter .....	12
Defectdojo .....	15
Kyverno .....	18
Opencost(FinOps) .....	20
Goldilocks .....	25
Prometheus stack and Grafana dashboards .....	32
Terraform & AWS .....	42
Infrastructure layer .....	42
Gitops & Application layer .....	47
Robot-shop application .....	51
CI-CD pipeline with DevSecOps automation .....	54
Final Infrastructure .....	56
Conclusion .....	57

**Warning In this Documentation shows how to reach to this diagram with out **hacking**:**



# LET'S GOOOOO! FEELS EOOOOOO!

## Robot Shop — Production DevSecOps Platform

### Enterprise-Grade Cloud-Native Reference Architecture on AWS

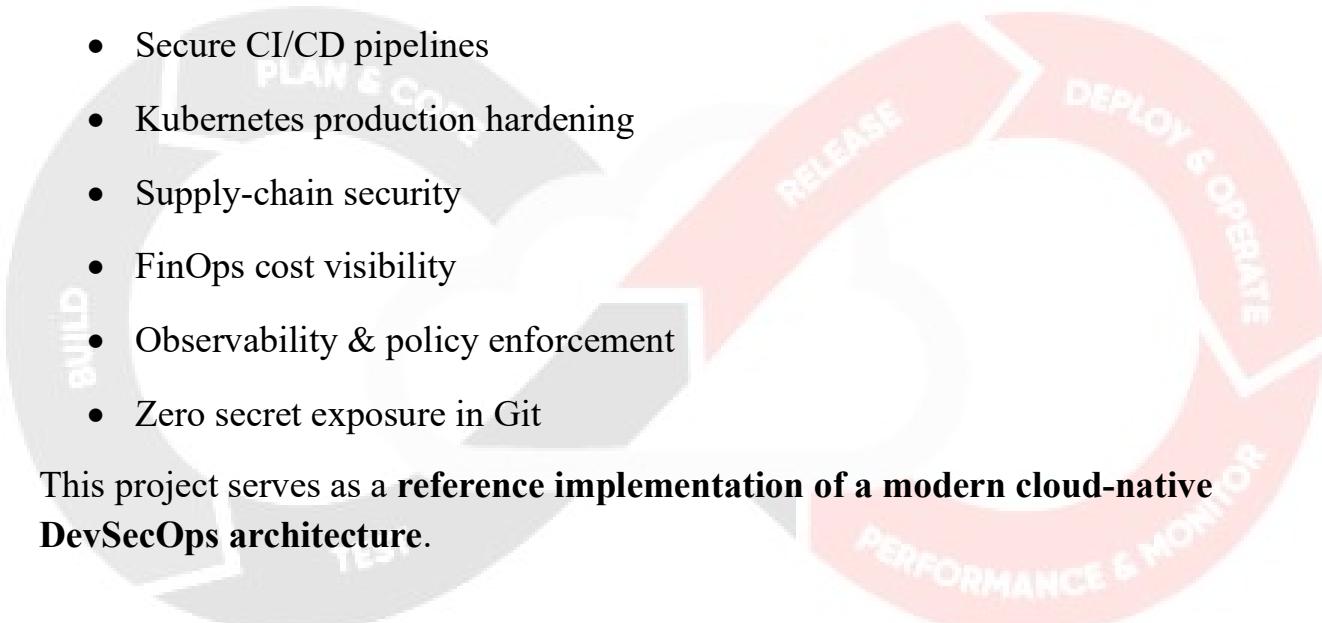
#### Introduction

Robot Shop is a **12-microservice** e-commerce application transformed into a **production-grade DevSecOps platform** running on AWS.

This project is completely demonstrating:

- Infrastructure as Code (Terraform)
- GitOps (ArgoCD)
- Secure CI/CD pipelines
- Kubernetes production hardening
- Supply-chain security
- FinOps cost visibility
- Observability & policy enforcement
- Zero secret exposure in Git

This project serves as a **reference implementation of a modern cloud-native DevSecOps architecture**.



# Application Overview

Robot Shop is a polyglot microservices application composed of 12 services.

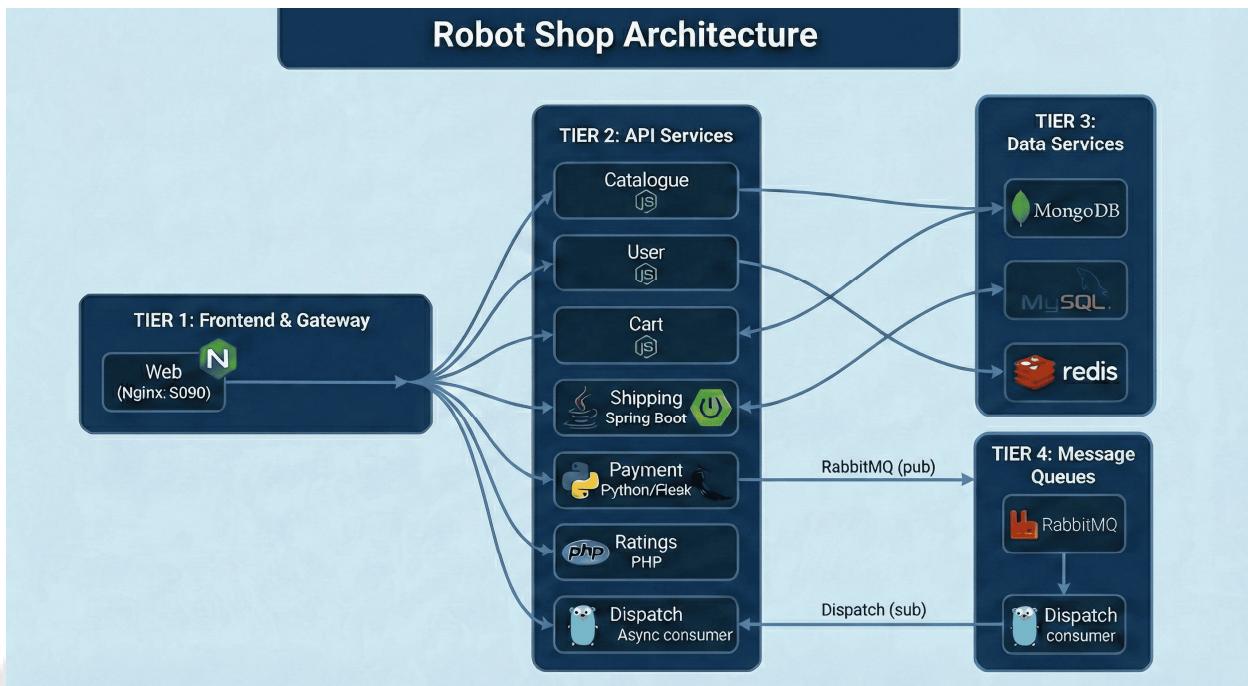
## Technology Stack

Service	Technology	Responsibility
web	Nginx	Reverse proxy & frontend gateway
catalogue	Node.js (Express)	Product catalog (MongoDB)
user	Node.js (Express)	User authentication (MongoDB + Redis)
cart	Node.js (Express)	Shopping cart (Redis)
shipping	Java (Spring Boot)	Shipping calculation (MySQL)
payment	Python (Flask)	Payment processing (RabbitMQ)
ratings	PHP (Apache)	Product ratings (MySQL)
dispatch	Go	Async order processing
mongodb	MongoDB	Document database
mysql	MySQL	Relational database
redis	Redis	Session cache
rabbitmq	RabbitMQ	Message broker

## Communication Model

- REST-based synchronous communication
- Asynchronous messaging via RabbitMQ
- Internal service isolation via Kubernetes Network Policies

## The structure of Robot shop



## Docker Layer Local Development Environment

### Purpose

The Docker layer provides a fast and isolated local development environment for engineers.

It enables:

- Rapid iteration
- Service debugging
- Security testing (DAST stage)
- Environment parity before Kubernetes deployment

## Architecture

All 12 microservices and dependencies run via:

- Docker Compose
- Internal Docker network
- Environment variables from .env

Services include:

Service	Technology	Role
Web	Nginx	Reverse proxy & frontend
Catalogue	Node.js	Product API
User	Node.js	Authentication
Cart	Node.js	Cart service
Shipping	Java (Spring Boot)	Shipping calculation
Payment	Python (Flask)	Payment processing
Ratings	PHP	Ratings service
Dispatch	Go	Async consumer
MongoDB	MongoDB	Document DB
MySQL	MySQL	Relational DB
Redis	Redis	Cache & sessions
RabbitMQ	RabbitMQ	Message broker

## Enhancement in Docker files

- Using light weight base images
- Using multi-stage in some micro services
- Making docker files for all micro services except ( Redis, RabbitMQ )

Also we made Docker compose file for the project

---

## Kubernetes Layer Production Workload Architecture

### Purpose

Kubernetes provides orchestration, scalability, security boundaries, and resilience.

The cluster runs on **Amazon Elastic Kubernetes Service**.

### Core Design Principles

#### 1. Security-First Pods

#### 2. Horizontal Pod Autoscaling

Each application service includes:

- CPU-based scaling
- Memory-based scaling
- Environment-specific replica tuning

#### 3. Network Policies (4-Tier Model)

Tier	Allowed Ingress	Allowed Egress
Frontend	Ingress Controller	Backend + DNS
Backend	Frontend	DB + Queue + DNS
Database	Backend only	DNS only
Queue	Backend only	DNS only

This ensures strict east-west traffic control.

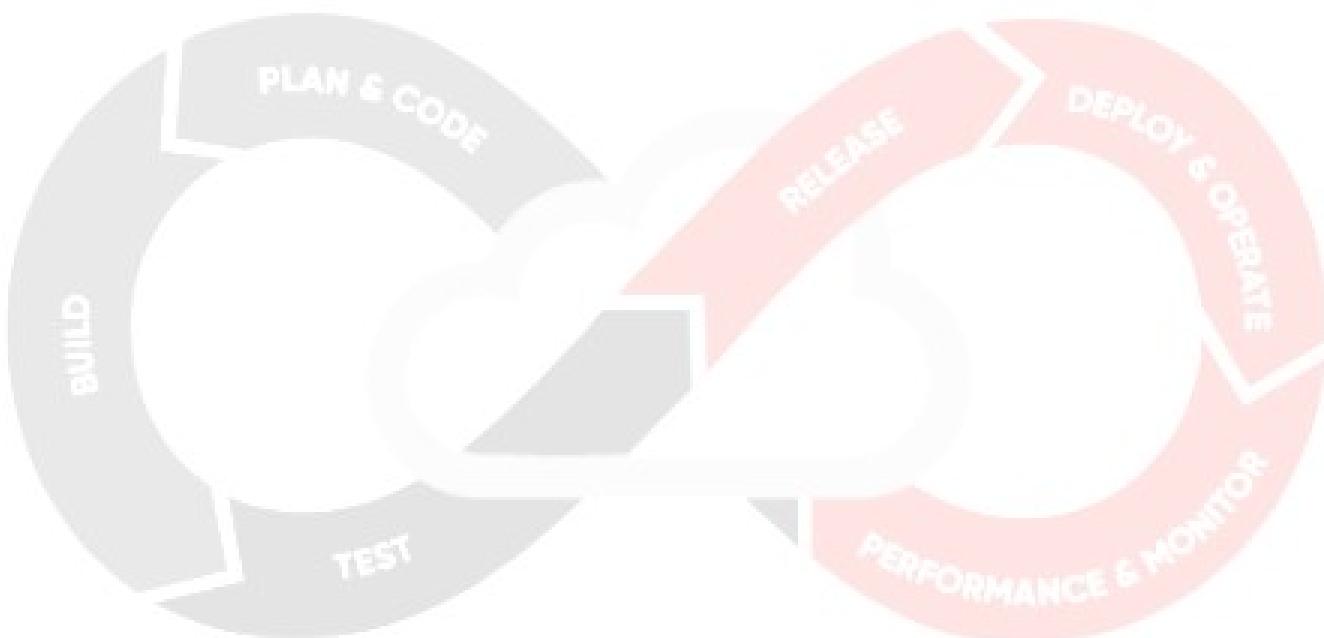
## 4. Zero Git Secrets

No credentials exist in Git.

Secrets flow:

Terraform → AWS SSM → ESO → Kubernetes Secrets → Pods

---



# Umbrella Helm Layer Deployment Abstraction

## Purpose

The umbrella Helm chart provides:

- Centralized deployment logic
- Environment-specific overrides
- Reusable security templates
- DRY chart architecture

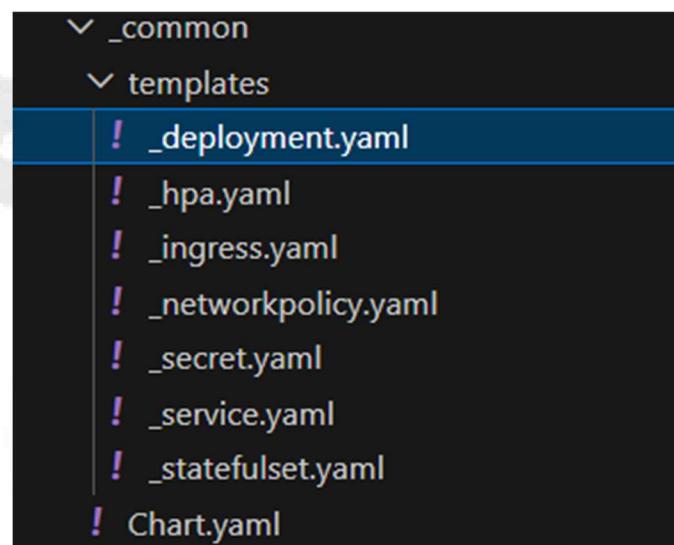
## Structure

```
└── helm
    └── robot-shop
        ├── charts
        │   ├── _common
        │   ├── cart
        │   ├── catalogue
        │   ├── dispatch
        │   ├── mongodb
        │   ├── mysql
        │   ├── payment
        │   ├── ratings
        │   ├── shipping
        │   ├── user
        │   └── web
        └── templates
            └── ! backend-networkpolicy.yaml
            └── ! database-networkpolicy.yaml
            └── ! frontend-networkpolicy.yaml
            └── ! message-queue-networkpolicy.yaml
            └── NOTES.txt
            └── .helmignore
            └── ! Chart.yaml
            └── ! values-dev.yaml
            └── ! values-prod.yaml
            └── ! values.yaml
```

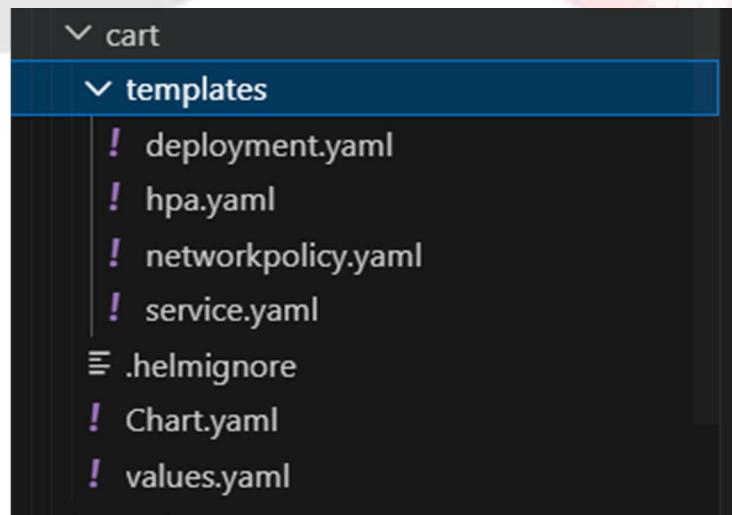
## Shared Library Chart

The \_common chart enforces:

- Security contexts
- HPA templates
- NetworkPolicy templates
- Standardized labels
- Resource limits



Using these templates in all microservices like:



Put values of cart svc and in the manifests define the common template.

## Dev vs Production

Setting	Dev	Production
Replicas	1	2–6
Resources	Minimal	Tuned
HPA	Optional	Required
TLS	Optional	Enforced

## Charts explanation and their uses:

### Karpenter Chart

#### Overview

Karpenter is a next-generation Kubernetes node provisioning system that dynamically launches compute capacity based on pending pods.

It replaces traditional Cluster Autoscaler with faster, more intelligent scaling.

#### Purpose in This Platform

- ❖ Dynamic node provisioning
- ❖ Spot-first cost optimization
- ❖ Instance type auto-selection
- ❖ Workload-aware scaling

## Architecture

Components deployed:

- Karpenter Controller (Deployment)
- ServiceAccount (Pod Identity attached)

- CRDs:
  - NodePool
  - EC2NodeClass
  - NodeClaim

## Scaling Behavior

1. Pod becomes unschedulable
2. Karpenter evaluates resource requirements
3. Selects cheapest matching instance
4. Launches EC2 instance
5. Joins EKS cluster automatically

## Security

- Uses EKS Pod Identity (no IRSA)
- Least-privilege IAM role
- Runs in dedicated namespace
- Tolerations for system nodes

## Cost Optimization Strategy

- Spot instances preferred
- On-demand fallback
- Consolidation enabled
- Unused nodes automatically terminated

## Results of this chart:

The screenshot shows the AWS EC2 Instances page. The left sidebar navigation includes: Dashboard, EC2 Global View, Events, Instances (selected), Instance Types, Launch Templates, Spot Requests, Savings Plans, Reserved Instances, Dedicated Hosts, Capacity Reservations, Capacity Manager (New), Images (AMIs, AMI Catalog), Elastic Block Store (Volumes, Snapshots, Lifecycle Manager), Network & Security (Security Groups, Elastic IPs, Placement Groups, Key Pairs, Network Interfaces). The main content area displays 'Instances (1/3) info' with a table showing three instances:

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 IP	Elastic IP
ip-10-0-2-119...	i-01faf4172094eb506	Running	c7i-flex.large	3/3 checks passed	View alarms +	us-east-1b	-	-	-
ip-10-0-2-119...	i-00dcfcb051bca4bf7c	Running	c7i-flex.large	3/3 checks passed	View alarms +	us-east-1b	-	-	-
karpenster-nod...	i-01ca41f56395a70da	Running	c7i-flex.large	3/3 checks passed	View alarms +	us-east-1a	-	-	-

Details for the selected instance (i-00dcfcb051bca4bf7c) are shown on the right:

- Stop protection:** Disabled
- Launch time:** Tue Feb 17 2026 07:15:47 GMT+0200 (Eastern European Standard Time) (about 11 hours)
- Instance auto-recovery:** Default
- AMI location:** amazon/amazon-eks-node-al2023-x86\_64-standard-1.35-v20260209
- Stop-hibernate behavior:** Disabled
- AMI Launch Index:** 0
- Lifecycle:** spot
- State transition reason:** -
- Credit specification:** Not supported by instance type
- Key pair assigned at launch:** -
- State transition message:** -
- Kernel ID:** -
- Usage operation:** RunInstances
- RAM disk ID:** -
- Boot mode:** -

The screenshot shows the AWS EC2 Spot Requests page. The left sidebar navigation is identical to the previous screenshot. The main content area displays 'Spot Requests (2)' with a table showing two active spot requests:

Request...	Request type	Instance t...	State	Capacity	Status	Persistence	Created
sir-wj7qefrk	instance	c7i-flex.large	active	i-00dcfcb0...	fulfilled	one-time	11 hours ago
sir-zsgfrnlj	instance	c7i-flex.large	active	i-01faf41...	fulfilled	one-time	10 hours ago

**Spot savings summary:**

A high-level summary of your savings across all of your running and recently terminated Spot Instances. For detailed reporting on your account-level Spot usage, visit Cost Explorer.

Total Spot Instances	Total Spot cost (USD)	Total usage	Average cost per usage (USD)
2 Instances	\$0.68 You saved 62%	42 vCPU hours 84 Memory(GB) hours	\$0.0165 per vCPU hour \$0.0081 per Memory hour

**Details by instance types:**

\* Spot savings are estimated savings and may differ from actual savings. This is because the savings shown on this page do not include the billing adjustments for your usage.

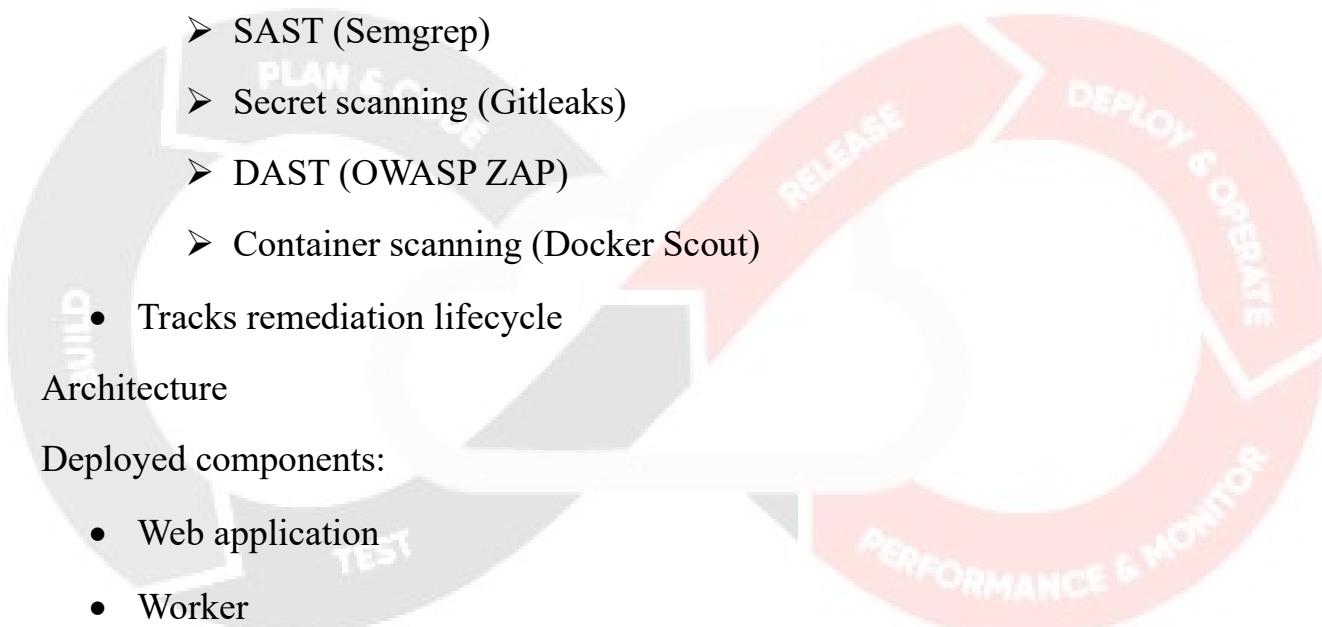
## DefectDojo Chart

### Overview

DefectDojo is a centralized vulnerability management platform used to aggregate and track security findings.

### Purpose in This Platform

- Central security dashboard
- Aggregates:
  - SAST (Semgrep)
  - Secret scanning (Gitleaks)
  - DAST (OWASP ZAP)
  - Container scanning (Docker Scout)
- Tracks remediation lifecycle



### Architecture

#### Deployed components:

- Web application
- Worker
- Redis
- PostgreSQL

### CI Integration

Pipeline uploads scan results automatically via API:

- SARIF
- JSON reports
- ZAP reports

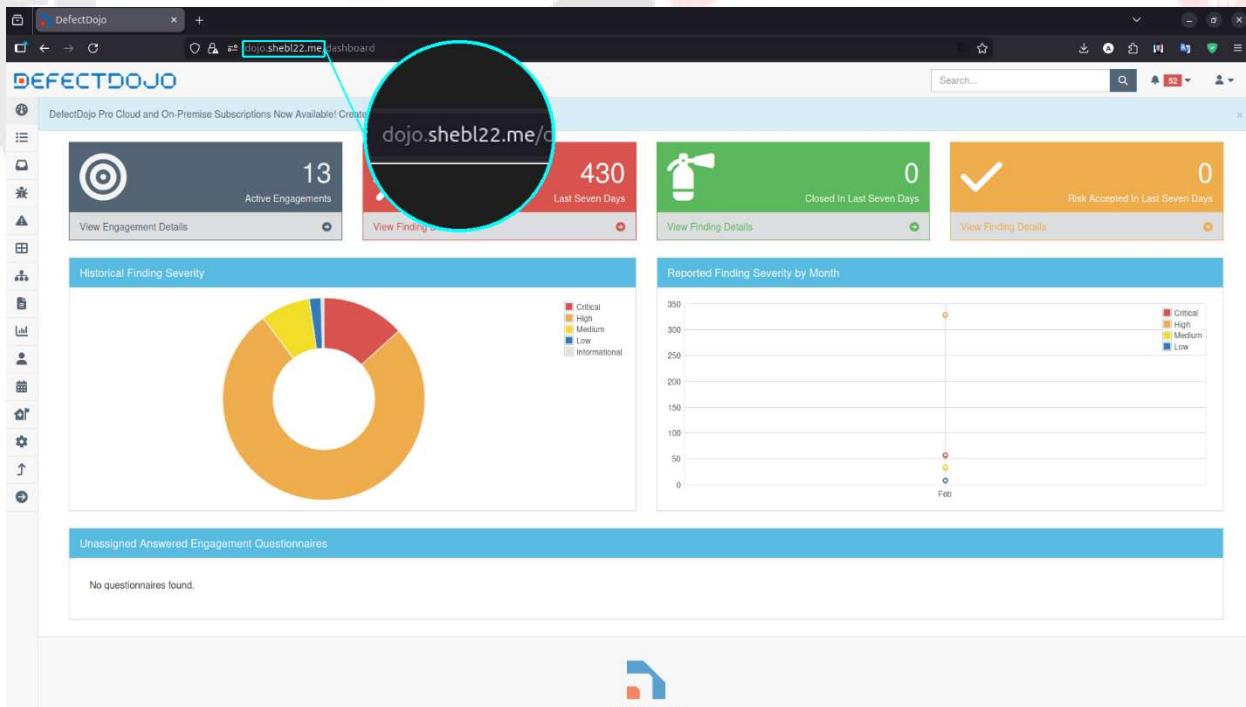
## Security

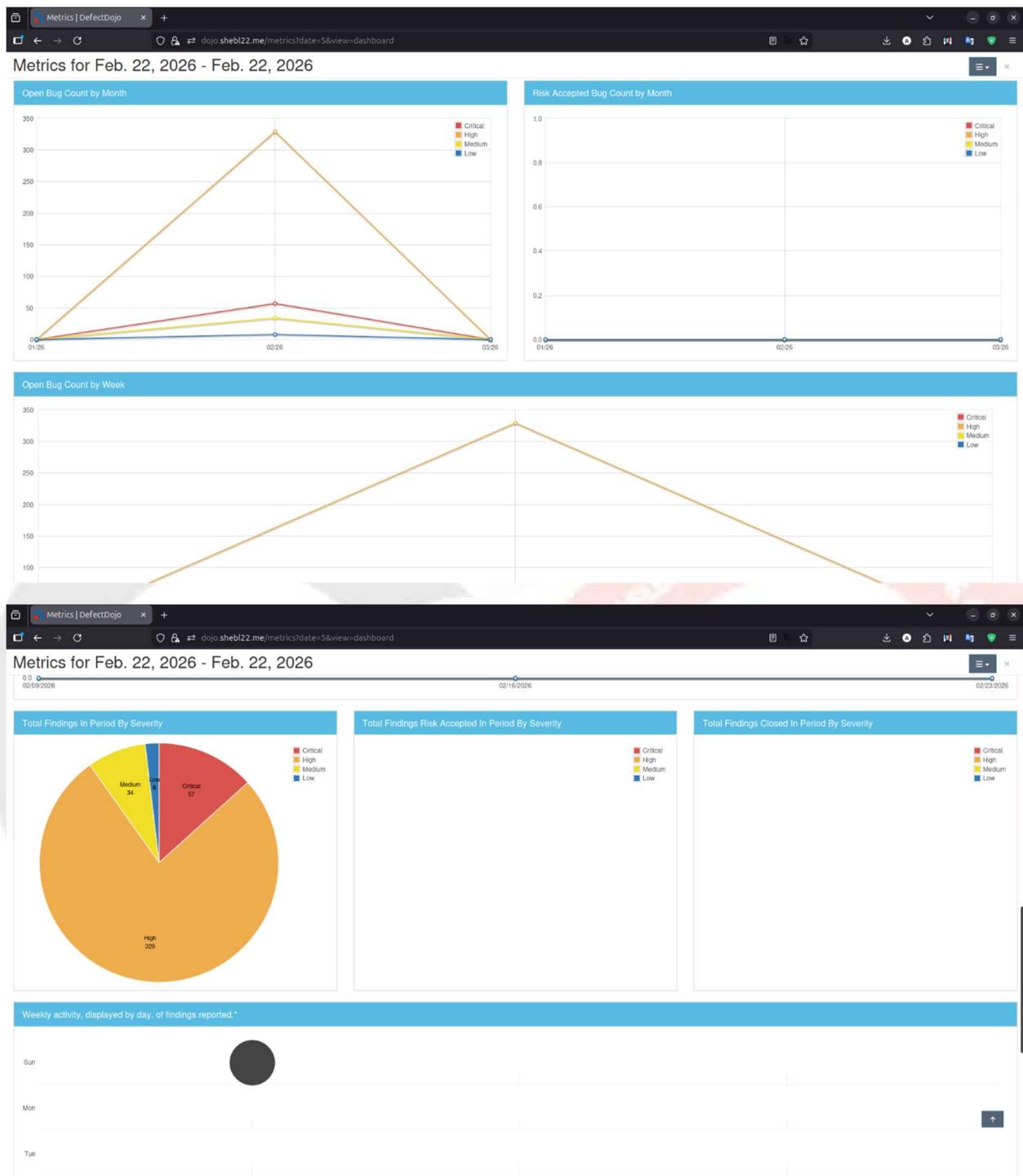
- Runs in isolated namespace
- Access restricted via Ingress + TLS
- Admin credentials stored in AWS SSM

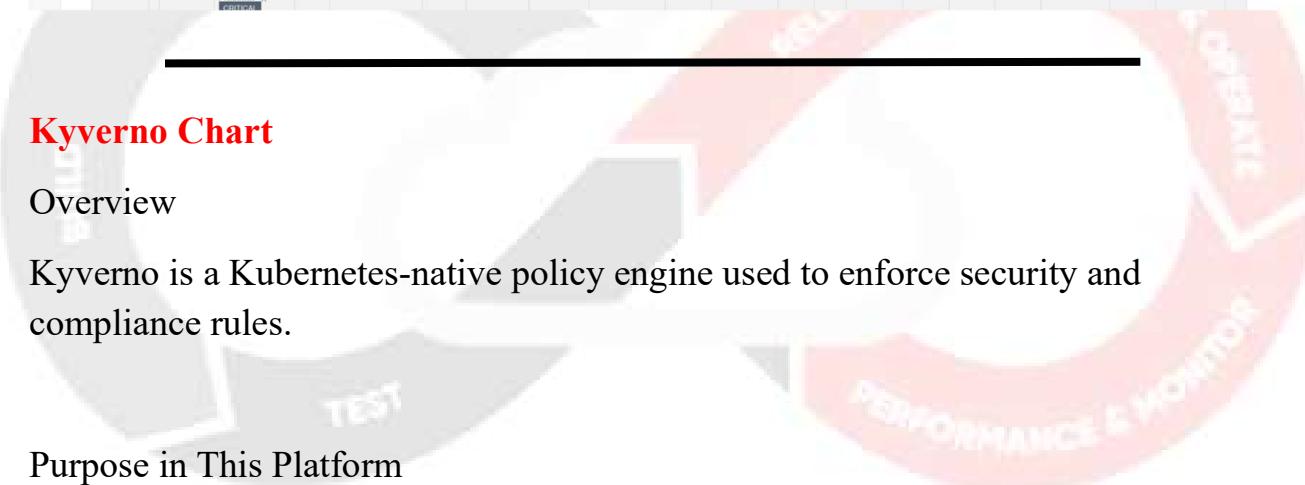
## Benefits

- Centralized vulnerability visibility
- SLA tracking
- Risk prioritization
- Audit readiness

## Results:







A screenshot of the DefectDojo web application showing a list of findings. The table has the following columns: Severity, Name, CWE, Vulnerability Id, EPSS Score, EPSS Percentile, Known Exploited, Used in Ransomware, Date Added to KEV, Date, Age, SLA, Reporter, Found By, Status, Group, Product, Service, Planned Remediation, and Rev. There are five rows of data, each with a red 'Critical' severity level and a 'Vulnerability : CVE-2025-15467' entry.

	Severity	Name	CWE	Vulnerability Id	EPSS Score	EPSS Percentile	Known Exploited	Used in Ransomware	Date Added to KEV	Date	Age	SLA	Reporter	Found By	Status	Group	Product	Service	Planned Remediation	Rev
	Critical	Vulnerability : CVE-2025-15467 Severity : CRITICAL Package : CRITICAL		CVE-2025-15467	N.A.	N.A.	No	No	Feb. 22, 2026	0	7	Admin User	docker-scout Scan (SARIF)	Active		Robot-Shop-Microservices-Image-catalogue				
	Critical	Vulnerability : CVE-2025-15467 Severity : CRITICAL Package : CRITICAL		CVE-2025-15467	N.A.	N.A.	No	No	Feb. 22, 2026	0	7	Admin User	docker-scout Scan (SARIF)	Active		Robot-Shop-Microservices-Image-catalogue				
	Critical	Vulnerability : CVE-2025-15467 Severity : CRITICAL Package : CRITICAL		CVE-2025-15467	N.A.	N.A.	No	No	Feb. 22, 2026	0	7	Admin User	docker-scout Scan (SARIF)	Active		Robot-Shop-Microservices-Image-cart				
	Critical	Vulnerability : CVE-2025-15467 Severity : CRITICAL Package : CRITICAL		CVE-2025-15467	N.A.	N.A.	No	No	Feb. 22, 2026	0	7	Admin User	docker-scout Scan (SARIF)	Active		Robot-Shop-Microservices-Image-catalogue				
	Critical	Vulnerability : CVE-2025-15467 Severity : CRITICAL Package : CRITICAL		CVE-2025-15467	N.A.	N.A.	No	No	Feb. 22, 2026	0	7	Admin User	docker-scout Scan (SARIF)	Active		Robot-Shop-Microservices-Image-catalogue				

## Kyverno Chart

### Overview

Kyverno is a Kubernetes-native policy engine used to enforce security and compliance rules.

### Purpose in This Platform

- Enforce signed container images
- Prevent privileged containers
- Block hostPath volumes
- Enforce security context standards

## Implemented Policies

### 1. Signed Image Verification

- Uses Cosign
- Rejects unsigned images

### 2. Pod Security Enforcement

- runAsNonRoot required
- Drop ALL capabilities
- Disallow privilege escalation

### 3. Resource Limits Enforcement

- All pods must define CPU and memory limits

## Why Kyverno Instead of OPA?

- Native Kubernetes CRDs
- YAML-based policies
- Easier maintainability

## Results:

```
@adelrahman-shebl + ~/Robot-Shop-Microservices/K8s/microservices/mongodb (feature/pipeline) $ kubectl describe polr 033214f1-27d8-47ea-89e8-8232aa3b1052
Annotations:  <none>
API Version:  wgpolicy.k8s.io/v1alpha2
Kind:         PolicyReport
Metadata:
  Creation Timestamp: 2026-02-23T05:58:50Z
  Generation: 48
  Owner References:
    API Version: v1
    Kind:       Pod
    Name:      dispatch-deployment-f48685b4f-hqrbz
    UID:       033214f1-27d8-47ea-89e8-8232aa3b1052
  Resource Version: 30575
  UID:       8740f980-8328-4d61-853f-eb0856321de75
Results:
  Message: success
  Policy: verify-robot-shop
  Properties:
    Process: background scan
    Result: pass
    Rule: verify-robot-shop
    Scored: true
    Source: KyvernoImageValidatingPolicy
    Timestamp:
      Nanos: 0
      Seconds: 1771827673
  Scope:
    API Version: v1
    Kind:       Pod
    Name:      dispatch-deployment-f48685b4f-hqrbz
    Namespace: default
    UID:       033214f1-27d8-47ea-89e8-8232aa3b1052
  Summary:
    Error: 0
    Fail: 0
    Pass: 1
    Skip: 0
    Warn: 0
  Events:  <none>
```

NAME	READY	STATUS	RESTARTS	AGE
cart-deployment-6bd6749cd4-qk5fj	1/1	Running	0	19m
cart-deployment-6bd6749cd4-xrlcd	1/1	Running	0	19m
catalogue-deployment-5f68fcc4b9-gtkcx	0/1	Pending	0	19m
catalogue-deployment-5f68fcc4b9-mxgxk	0/1	Pending	0	19m
dispatch-deployment-f48685b4f-hqrbz	1/1	Running	0	19m
mongo-ss-0	1/1	Running	0	19m
mongo-ss-1	1/1	Running	0	6m15s
mongo-ss-2	1/1	Running	0	6m7s
mysql-ss-0	1/1	Running	0	19m
mysql-ss-1	1/1	Running	1 (13m ago)	17m
mysql-ss-2	1/1	Running	0	17m
payment-deployment-5bbb688856-cnx5d	1/1	Running	0	19m
payment-deployment-5bbb688856-n8k9g	1/1	Running	0	19m
rabbitmq-0	0/1	Init:ImagePullBackOff	0	19m
ratings-deployment-6488f96bf4-zwkwg	1/1	Running	0	19m
redis-master-0	1/1	Running	0	19m
redis-replicas-0	1/1	Running	0	19m
redis-replicas-1	1/1	Running	0	8m17s
redis-replicas-2	1/1	Running	0	7m43s
shipping-deployment-5b7b9ddf59-9qjx2	1/1	Running	0	19m
shipping-deployment-5b7b9ddf59-xxlcj	1/1	Running	0	19m
user-deployment-7d9b7c49cb-znvz5	1/1	Running	0	19m
web-deployment-8b8965f45-qc972	1/1	Running	0	19m

## OpenCost Chart

### Overview

OpenCost provides real-time cost visibility inside Kubernetes clusters.

### Purpose in This Platform

- Namespace cost allocation
- Pod-level cost breakdown
- Spot vs On-Demand comparison
- FinOps visibility

### Architecture

- Prometheus integration
- Cost model engine
- Cost API endpoint

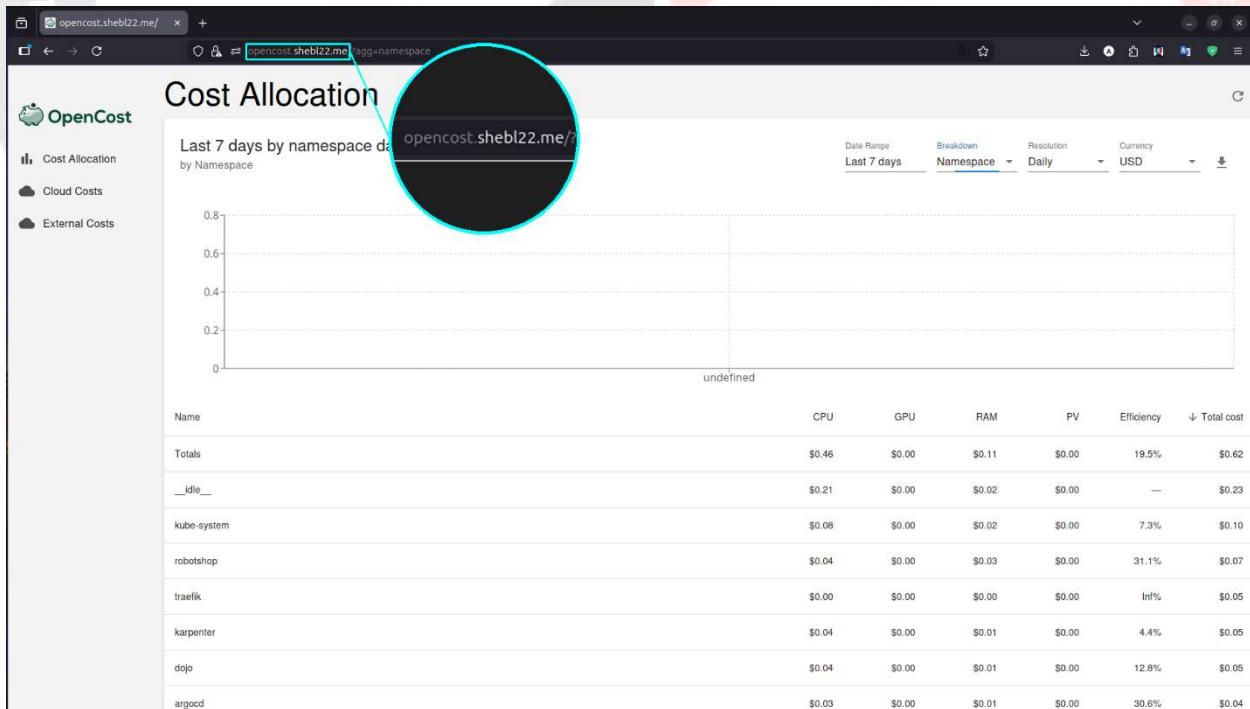
## What It Tracks

- CPU usage cost
- Memory usage cost
- Network egress cost
- Persistent volume cost

## Business Impact

- Identify waste
- Optimize node sizing
- Reduce cloud bill
- Justify scaling decisions

Results:



OpenCost version: 1.119.1 (0a1cccb)

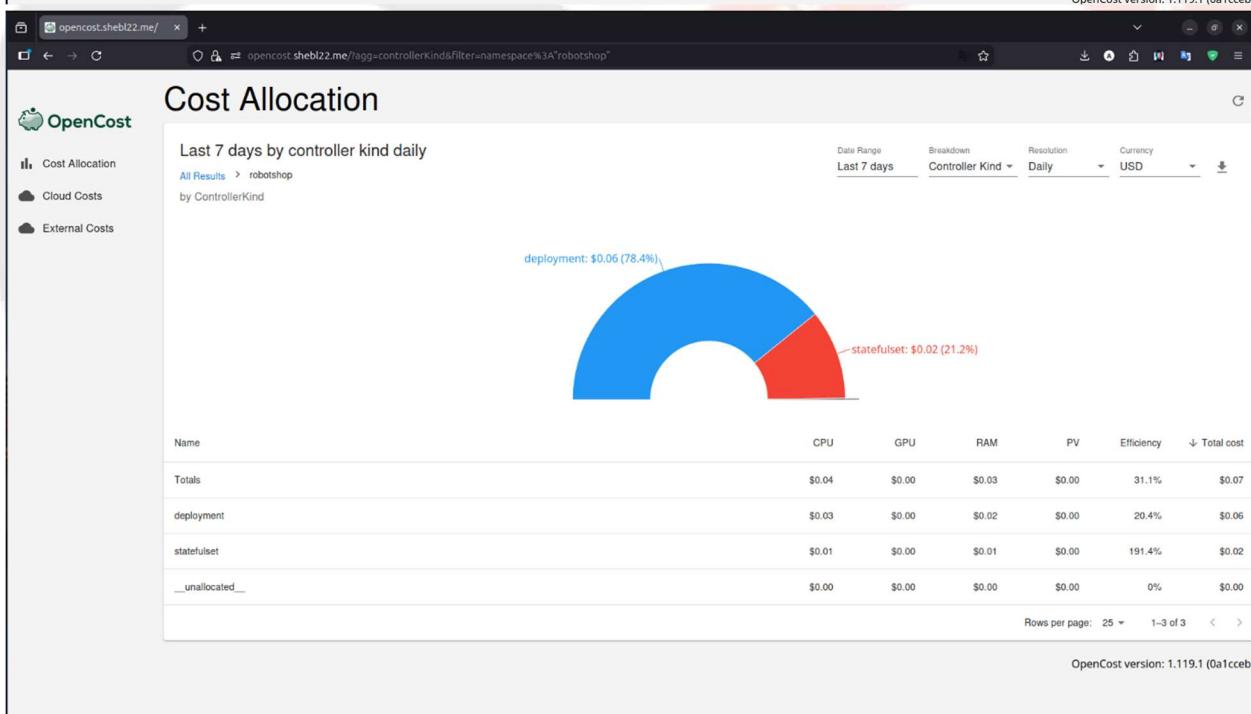
[opencost.shebl22.me/](#) + [opencost.shebl22.me/agg=namespace](#)

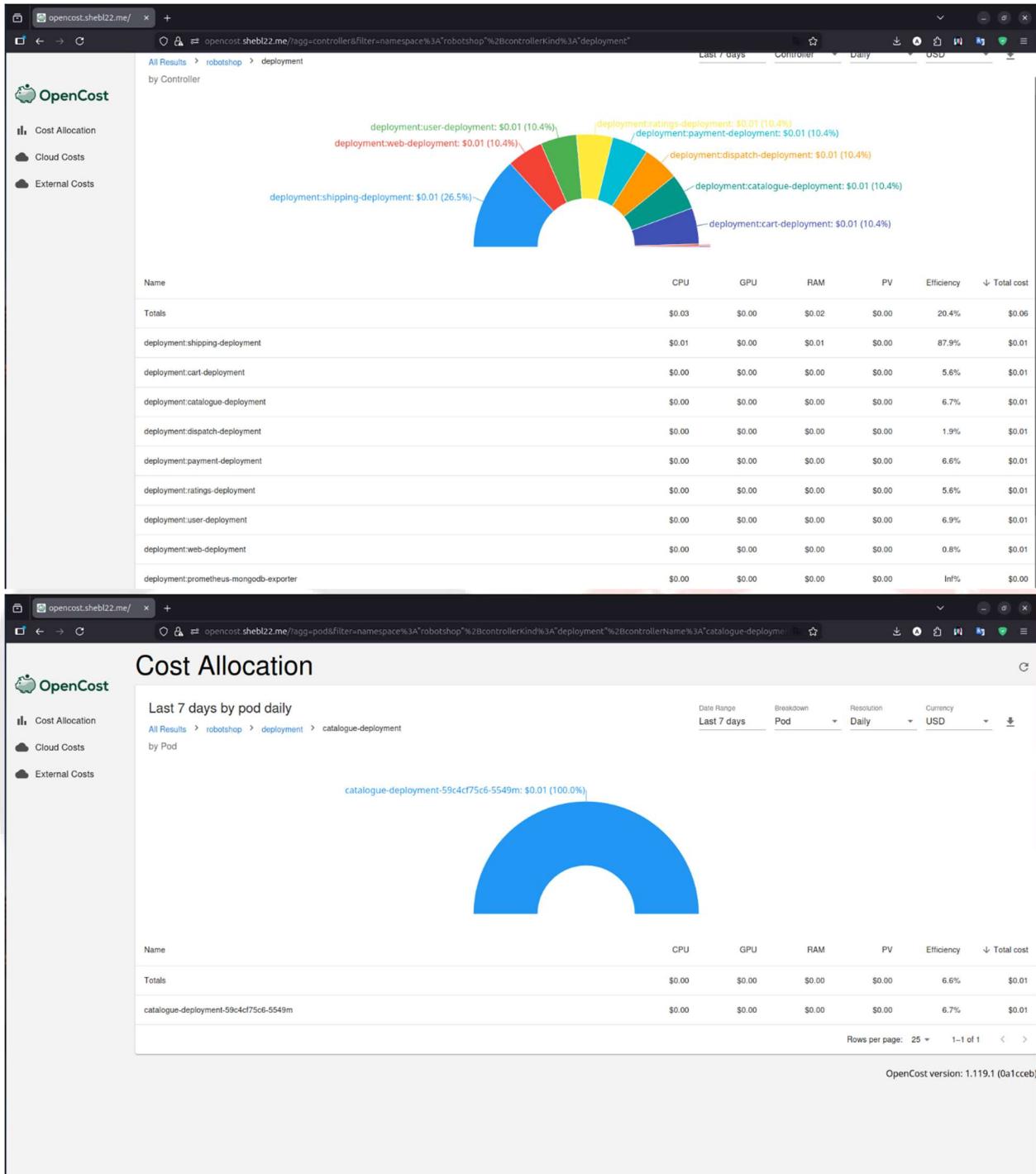
### OpenCost

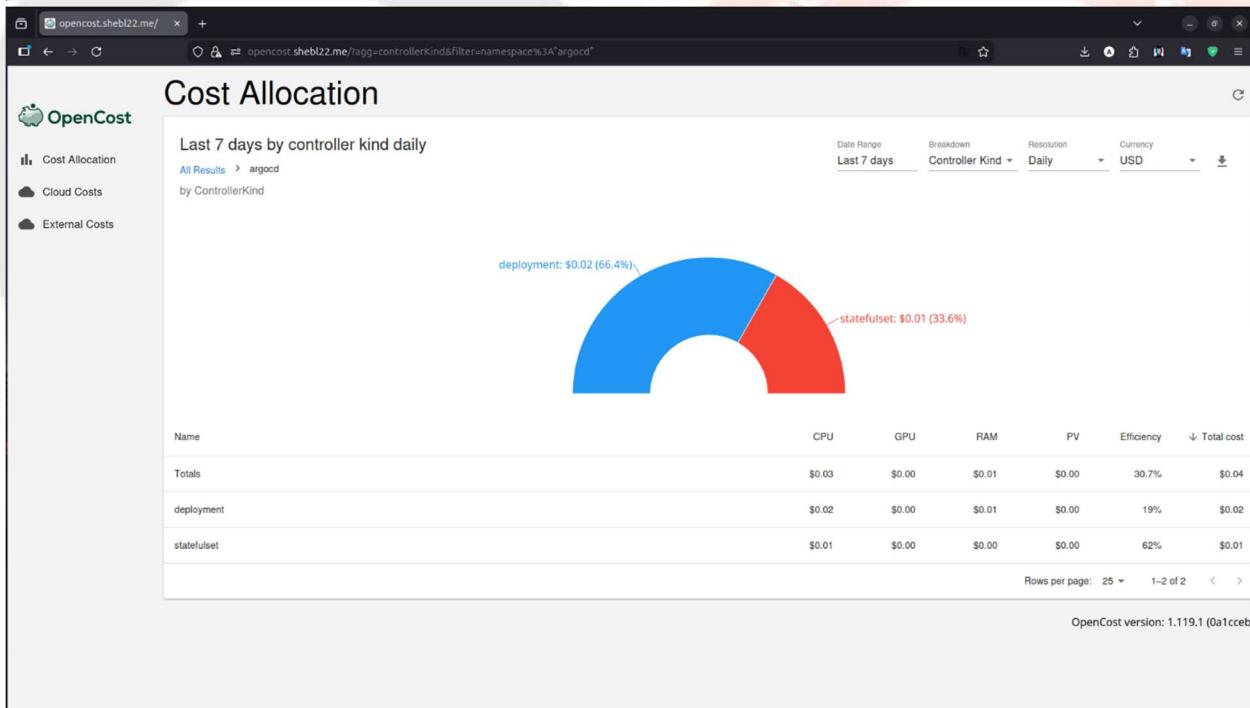
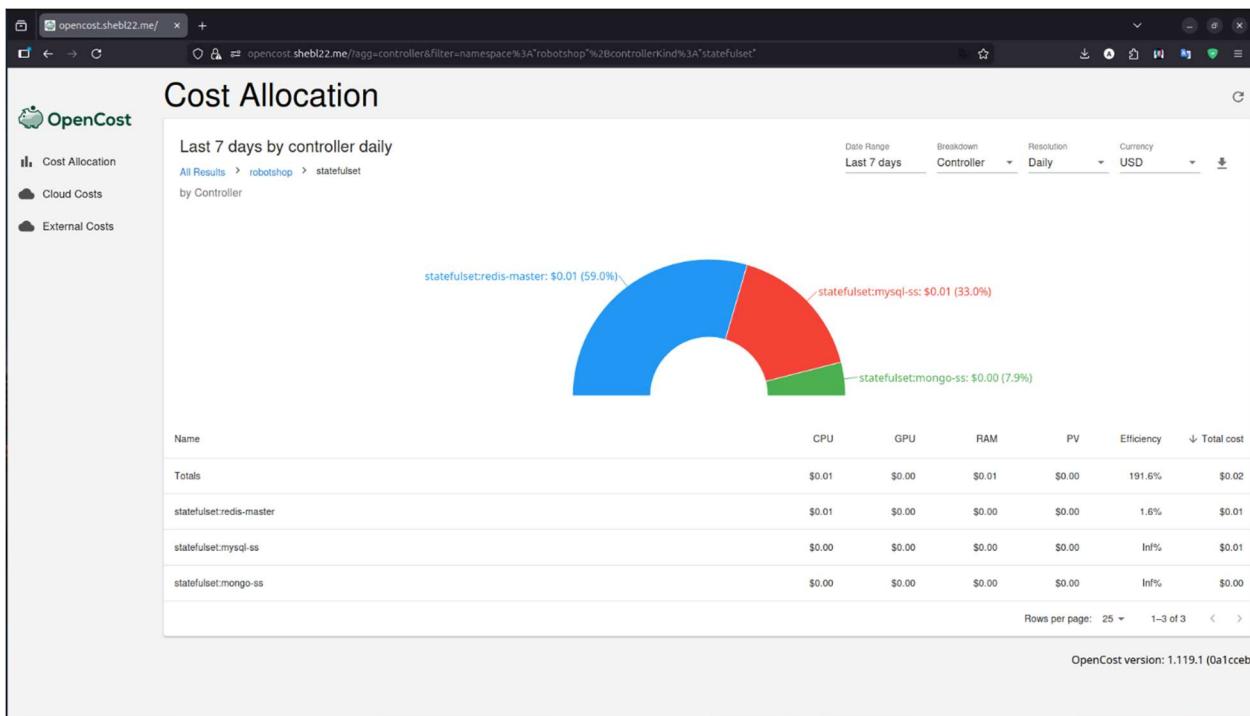
- Cost Allocation
- Cloud Costs
- External Costs

Name	CPU	GPU	RAM	PV	Efficiency	Total cost
Totals	\$0.46	\$0.00	\$0.11	\$0.00	19.5%	\$0.62
__idle__	\$0.21	\$0.00	\$0.02	\$0.00	—	\$0.23
kube-system	\$0.08	\$0.00	\$0.02	\$0.00	7.3%	\$0.10
robotshop	\$0.04	\$0.00	\$0.03	\$0.00	31.1%	\$0.07
traefik	\$0.00	\$0.00	\$0.00	\$0.00	Infl%	\$0.05
karpenter	\$0.04	\$0.00	\$0.01	\$0.00	4.4%	\$0.05
dojo	\$0.04	\$0.00	\$0.01	\$0.00	12.8%	\$0.05
argocd	\$0.03	\$0.00	\$0.01	\$0.00	30.6%	\$0.04
goldilocks	\$0.01	\$0.00	\$0.01	\$0.00	3.3%	\$0.03
monitoring	\$0.00	\$0.00	\$0.01	\$0.00	368.8%	\$0.01
opencost	\$0.00	\$0.00	\$0.00	\$0.00	26.5%	\$0.00
cert-manager	\$0.00	\$0.00	\$0.00	\$0.00	158.4%	\$0.00
eso	\$0.00	\$0.00	\$0.00	\$0.00	Infl%	\$0.00
edns	\$0.00	\$0.00	\$0.00	\$0.00	Infl%	\$0.00

Rows per page: 25 < 1–13 of 13 >







## Goldilocks Chart

### Overview

Goldilocks analyzes resource usage and recommends optimal CPU/memory requests and limits.

Built on top of Vertical Pod Autoscaler.

### Purpose in This Platform

- Right-size workloads
- Prevent overprovisioning
- Improve cluster efficiency
- Support FinOps goals

### How It Works

1. Collects Prometheus metrics
2. Uses VPA recommendation engine
3. Provides dashboard suggestions

### Output

- Recommended CPU requests
- Recommended memory limits
- Efficiency comparison

### Production Use

Recommendations are reviewed before updating Helm values.

## Results:

The screenshot displays two views of the Goldilocks by Fairwinds dashboard. The top view shows the main dashboard with a search bar and a 'Submit' button. A teal circle highlights the URL in the browser's address bar: `goldilocks.shebl22.me`. The bottom view shows the 'Namespace Details' page for the 'argocd' namespace. It lists workloads and provides detailed resource settings for the 'argo-cd-argocd-application-controller' statefulset's 'application-controller' container. The 'Guaranteed QoS' section shows current and guaranteed values for CPU and Memory requests and limits. The 'Burstable QoS' section shows current and burstable values for CPU and Memory requests and limits. Both sections include YAML snippets for recommended settings.

**Namespace Details**

Filter namespaces. Results update as you type.

12 namespaces found

**argocd**

Limit results to the argocd namespace

► Workloads

**cert-manager**

Limit results to the cert-manager namespace

► Workloads

**argocd**

Limit results to the argocd namespace

► Workloads

**STATEFULSET**

**argo-cd-argocd-application-controller**

► Containers

**application-controller**

► Details

**Guaranteed QoS**

	Current	Guaranteed
CPU Request	100m	> 93m
CPU Limit	1	> 93m
Memory Request	256Mi	< 814M
Memory Limit	2Gi	> 614M

**Burstable QoS**

	Current	Burstable
CPU Request	100m	! 15m
CPU Limit	1	> 456m
Memory Request	256Mi	< 811M
Memory Limit	2Gi	! 2598M

**YAML for Recommended Settings**

```
resources:
  requests:
    cpu: 93m
    memory: 814M
  limits:
    cpu: 93m
    memory: 814M
```

```
resources:
  requests:
    cpu: 15m
    memory: 811M
  limits:
    cpu: 456m
    memory: 2598M
```

The screenshot shows the Goldilocks by Fairwinds dashboard for the **robotshop** namespace. Under the **cart-deployment** section, the **cart** container is selected. The dashboard displays two QoS sections: **Guaranteed QoS** and **Burstable QoS**. In the **Guaranteed QoS** section, CPU Request is 50m > 15m, CPU Limit is 200m > 15m, Memory Request is 200Mi > 105M, and Memory Limit is 500Mi > 105M. In the **Burstable QoS** section, CPU Request is 50m > 15m, CPU Limit is 200m > 37m, Memory Request is 200Mi ! 105M (with a red exclamation mark), and Memory Limit is 500Mi > 219M.

The screenshot shows the Goldilocks by Fairwinds dashboard for the **catalogue** namespace. Under the **catalogue-deployment** section, the **catalogue** container is selected. The dashboard displays two QoS sections: **Guaranteed QoS** and **Burstable QoS**. In the **Guaranteed QoS** section, CPU Request is 50m > 15m, CPU Limit is 200m > 15m, Memory Request is 200Mi > 105M, and Memory Limit is 500Mi > 105M. In the **Burstable QoS** section, CPU Request is 50m > 15m, CPU Limit is 200m > 39m, Memory Request is 200Mi ! 105M (with a red exclamation mark), and Memory Limit is 500Mi > 228M. Below each QoS section, there is a link to **YAML for Recommended Settings** and a code block showing the recommended YAML configuration for each container.

**DEPLOYMENT**

## dispatch-deployment

▼ Containers

**CONTAINER**

### dispatch

▼ Details

**Guaranteed QoS**

	Current	Guaranteed
CPU Request	50m	> 15m
CPU Limit	200m	> 15m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 105M

**Burstable QoS**

	Current	Burstable
CPU Request	50m	> 15m
CPU Limit	200m	> 37m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 105M

▼ YAML for Recommended Settings

```
resources:  
  requests:  
    cpu: 15m  
    memory: 105M  
  limits:  
    cpu: 15m  
    memory: 105M
```

▼ YAML for Recommended Settings

```
resources:  
  requests:  
    cpu: 15m  
    memory: 105M  
  limits:  
    cpu: 37m  
    memory: 105M
```

**STATEFULSET**

## mongo-ss

▼ Containers

**CONTAINER**

### mongo

▼ Details

**Guaranteed QoS**

	Current	Guaranteed
CPU Request	Not Set	15m
CPU Limit	Not Set	15m
Memory Request	Not Set	226M
Memory Limit	Not Set	226M

**Burstable QoS**

	Current	Burstable
CPU Request	Not Set	15m
CPU Limit	Not Set	44m
Memory Request	Not Set	225M
Memory Limit	Not Set	921M

▼ YAML for Recommended Settings

```
resources:  
  requests:  
    cpu: 15m  
    memory: 226M  
  limits:  
    cpu: 15m  
    memory: 226M
```

▼ YAML for Recommended Settings

```
resources:  
  requests:  
    cpu: 15m  
    memory: 225M  
  limits:  
    cpu: 44m  
    memory: 921M
```

**mysql-ss**

**CONTAINER** mysql

**Guaranteed QoS**

	Current	Guaranteed
CPU Request	Not Set	15m
CPU Limit	Not Set	15m
Memory Request	Not Set	717M
Memory Limit	Not Set	717M

**Burstable QoS**

	Current	Burstable
CPU Request	Not Set	15m
CPU Limit	Not Set	43m
Memory Request	Not Set	713M
Memory Limit	Not Set	2866M

**YAML for Recommended Settings**

```
resources:
  requests:
    cpu: 15m
    memory: 717M
  limits:
    cpu: 15m
    memory: 717M
```

**DEPLOYMENT**

**payment-deployment**

**CONTAINER** payment

**Guaranteed QoS**

	Current	Guaranteed
CPU Request	50m	> 15m
CPU Limit	200m	> 15m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 105M

**Burstable QoS**

	Current	Burstable
CPU Request	50m	> 15m
CPU Limit	200m	> 37m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 171M

**YAML for Recommended Settings**

```
resources:
  requests:
    cpu: 15m
    memory: 105M
  limits:
    cpu: 15m
    memory: 105M
```

**YAML for Recommended Settings**

```
resources:
  requests:
    cpu: 15m
    memory: 105M
  limits:
    cpu: 37m
    memory: 171M
```

**Copy**

**DEPLOYMENT**

The screenshot displays three separate browser windows showing the Goldilocks by Fairwinds dashboard for different container configurations.

**Container: ratings**

**Deployment:** ratings-deployment

**Container:** ratings

**Guaranteed QoS**

	Current	Guaranteed
CPU Request	50m	> 15m
CPU Limit	200m	> 15m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 105M

**Burstable QoS**

	Current	Burstable
CPU Request	50m	> 15m
CPU Limit	200m	> 39m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 179M

**YAML for Recommended Settings**

```
resources:
  requests:
    cpu: 15m
    memory: 105M
  limits:
    cpu: 15m
    memory: 105M
```

**Container: redis**

**StatefulSet:** redis-master

**Container:** redis

**Guaranteed QoS**

	Current	Guaranteed
CPU Request	100m	> 23m
CPU Limit	150m	> 23m
Memory Request	128Mi	> 105M
Memory Limit	192Mi	> 105M

**Burstable QoS**

	Current	Burstable
CPU Request	100m	> 22m
CPU Limit	150m	> 91m
Memory Request	128Mi	> 105M
Memory Limit	192Mi	> 105M

**YAML for Recommended Settings**

```
resources:
  requests:
    cpu: 23m
    memory: 105M
  limits:
    cpu: 23m
    memory: 105M
```

**Container: shipping**

**Deployment:** shipping-deployment

The screenshot shows the Goldilocks by Fairwinds dashboard for a deployment named "user-deployment".

**Containers:** user

**Guaranteed QoS:**

	Current	Guaranteed
CPU Request	50m	> 15m
CPU Limit	200m	> 15m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 105M

**Burstable QoS:**

	Current	Burstable
CPU Request	50m	> 15m
CPU Limit	200m	> 39m
Memory Request	200Mi	! 105M
Memory Limit	500Mi	> 228M

**YAML for Recommended Settings:**

```
resources:
  requests:
    cpu: 15m
    memory: 105M
  limits:
    cpu: 15m
    memory: 105M
```

The screenshot shows the Goldilocks by Fairwinds dashboard for a deployment named "web-deployment".

**Containers:** web

**Guaranteed QoS:**

	Current	Guaranteed
CPU Request	50m	> 15m
CPU Limit	200m	> 15m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 105M

**Burstable QoS:**

	Current	Burstable
CPU Request	50m	> 15m
CPU Limit	200m	> 37m
Memory Request	200Mi	> 105M
Memory Limit	500Mi	> 105M

**YAML for Recommended Settings:**

```
resources:
  requests:
    cpu: 15m
    memory: 105M
  limits:
    cpu: 15m
    memory: 105M
```

The screenshot shows a web browser window with the title "Goldilocks by Fairwinds". The URL in the address bar is "goldilocks.shebl22.me/dashboard". The main content area displays the "traefik" namespace, which has been limited to workloads. Below this, there is a "Glossary" section with three entries:

- QoS:** "QoS (Quality of Service) class is a designation assigned to a pod based on its resource requests and limits. Kubernetes uses QoS classes to make decisions about scheduling and evicting Pods."  
*Kubernetes Documentation: Configure Quality of Service for Pods*
- Guaranteed QoS:** "For a Pod to be given a QoS class of Guaranteed:
  - Every Container, including init containers, in the Pod must have a memory limit and a memory request, and they must be the same.
  - Every Container, including init containers, in the Pod must have a CPU limit and a CPU request, and they must be the same."  
*Kubernetes Documentation: Create a Pod That Gets Assigned a QoS Class of Guaranteed*
- Burstable QoS:** "A Pod is given a QoS class of Burstable if:
  - The Pod does not meet the criteria for QoS class Guaranteed.
  - At least one Container in the Pod has a memory or CPU request."  
*Kubernetes Documentation: Create a Pod That Gets Assigned a QoS Class of Burstable*

At the bottom of the page, there are links for "Contact Us", "Feedback", social media icons (Facebook, Twitter, LinkedIn, Email), and the Fairwinds logo. Copyright information at the bottom includes "©2022-2023 Fairwinds Ops, Inc.", "Privacy Policy", and "Responsible Disclosure".

## Prometheus Stack Chart

### Overview

Prometheus +  
Grafana

Deployed via kube-prometheus-stack Helm chart.

### Components

- Prometheus
- Alertmanager
- Grafana
- Node Exporter
- kube-state-metrics

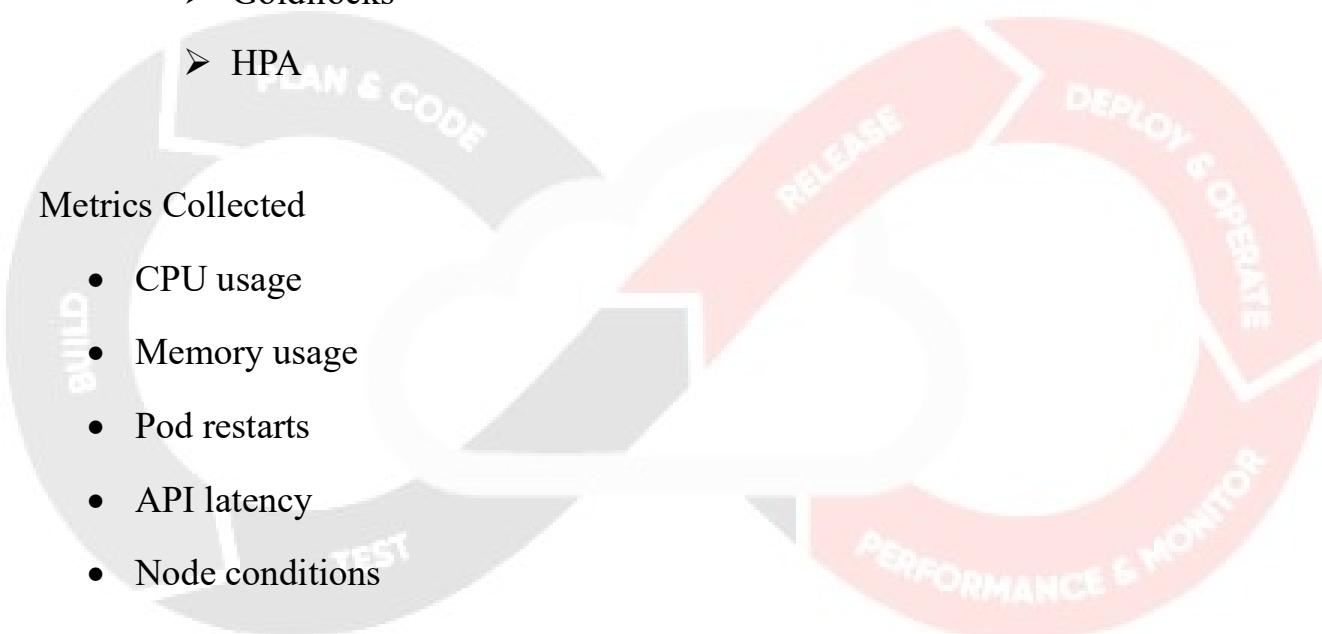
## Purpose in This Platform

- Cluster monitoring
- Application monitoring
- Node health tracking
- Alerting
- Metrics backend for:

- OpenCost
- Goldilocks
- HPA

## Metrics Collected

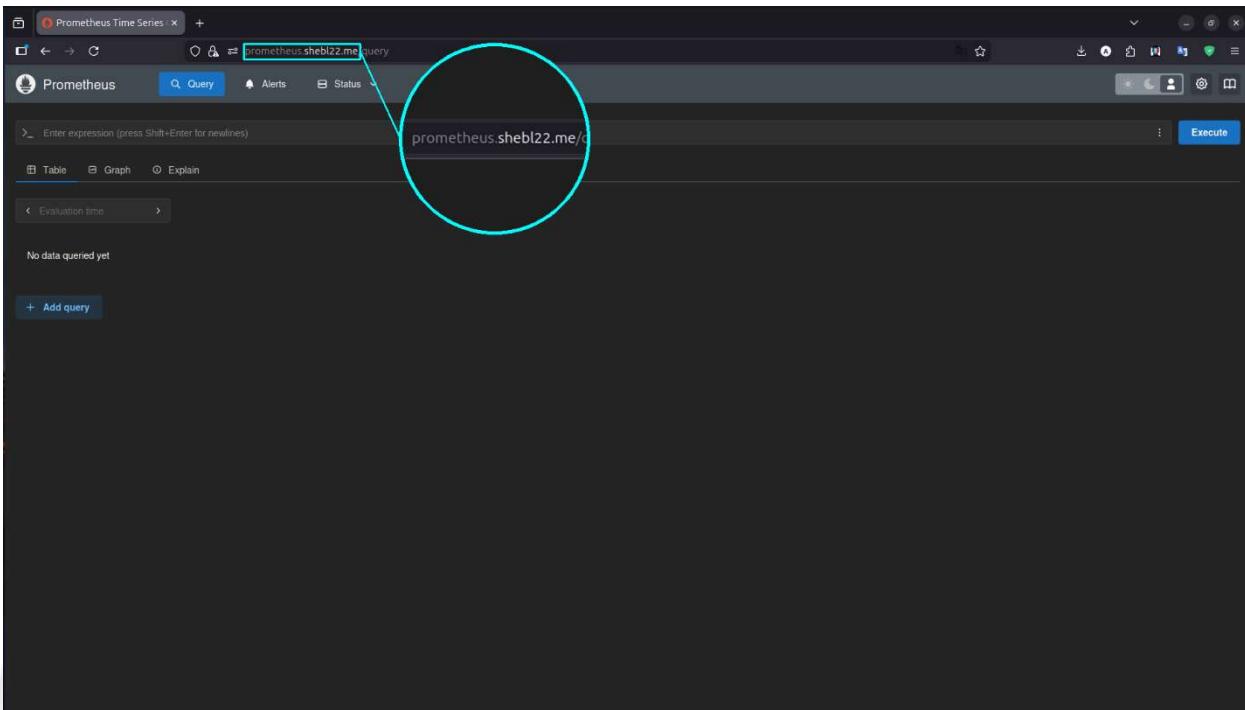
- CPU usage
- Memory usage
- Pod restarts
- API latency
- Node conditions



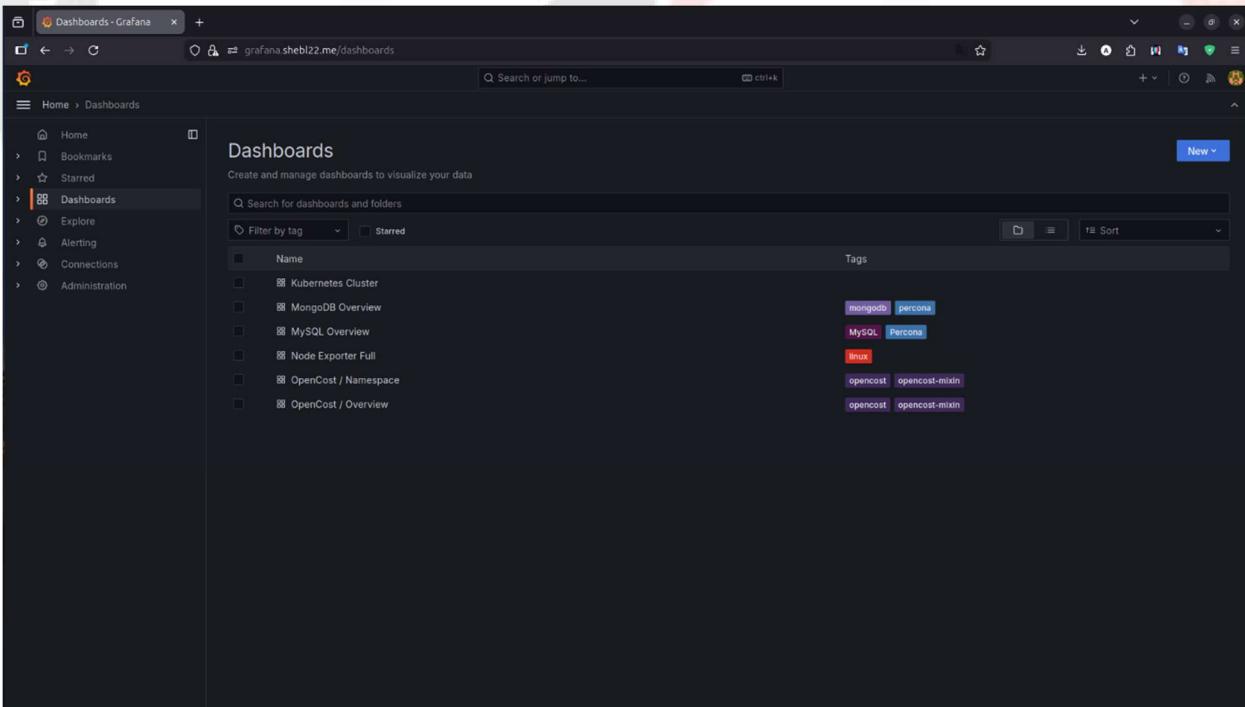
## Alerting

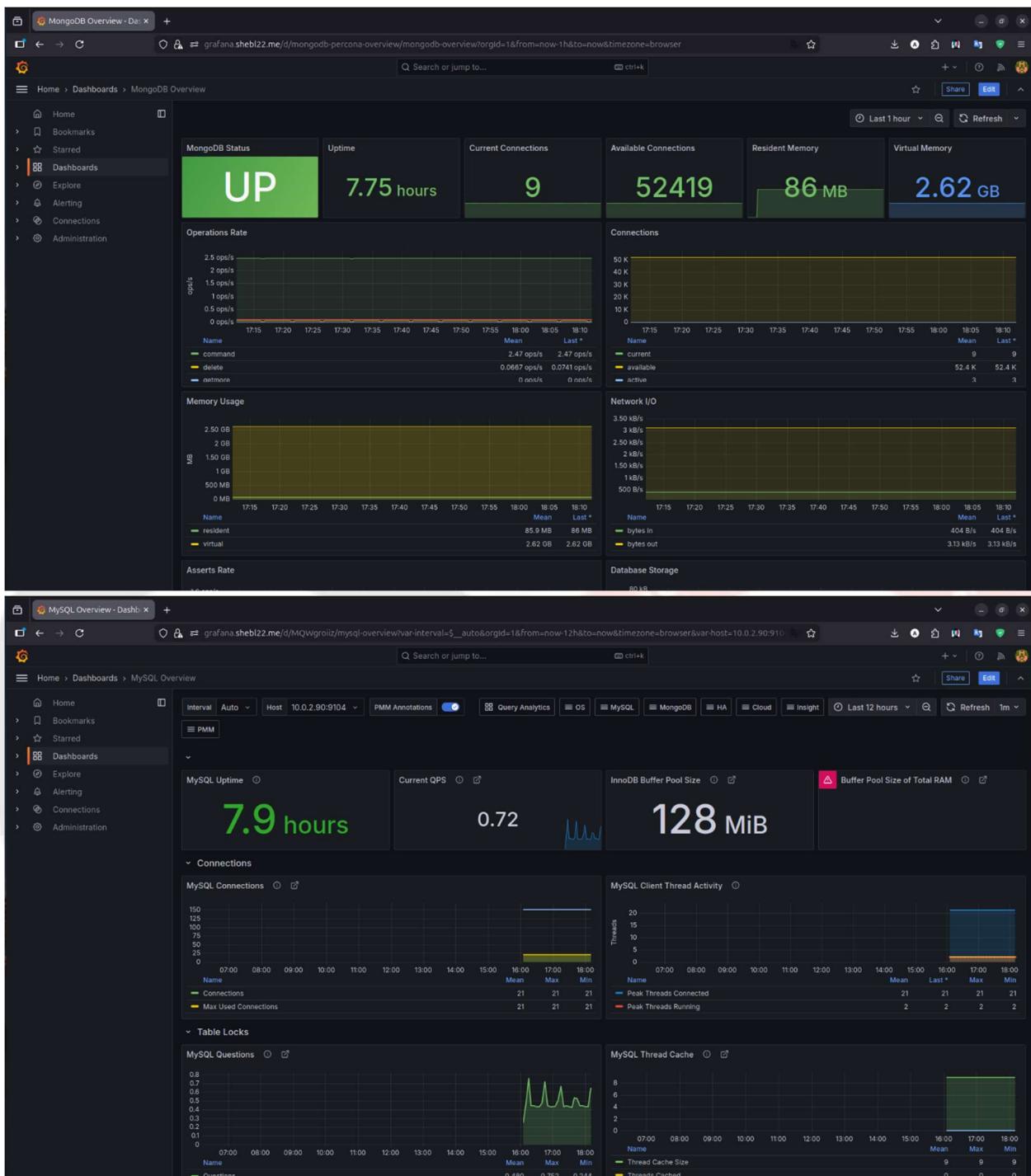
- High CPU usage
- Pod crash loops
- Node not ready
- Resource saturation

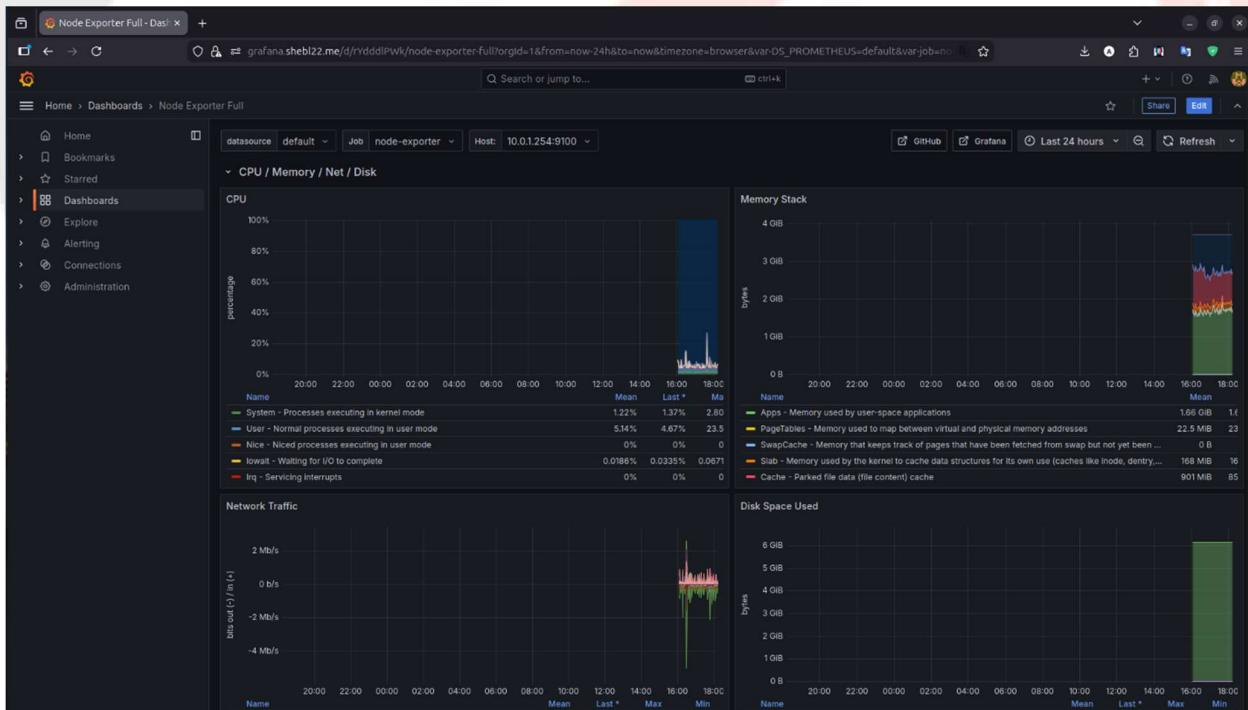
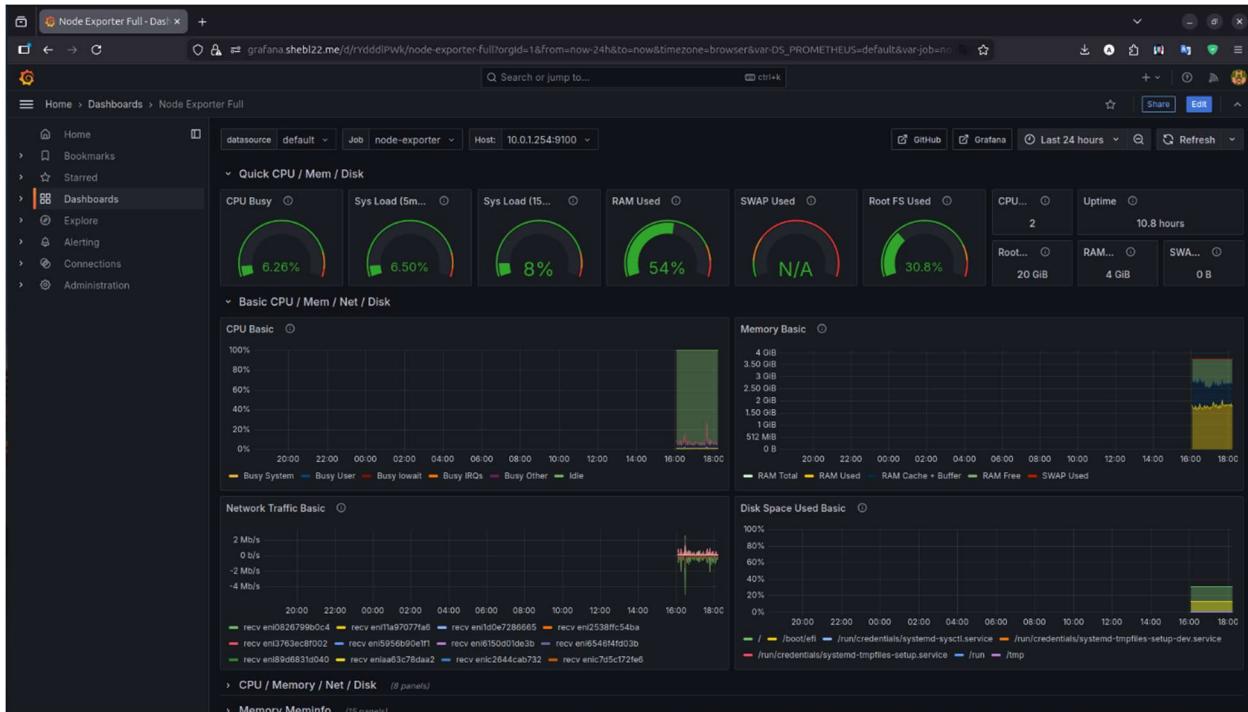
## Results:



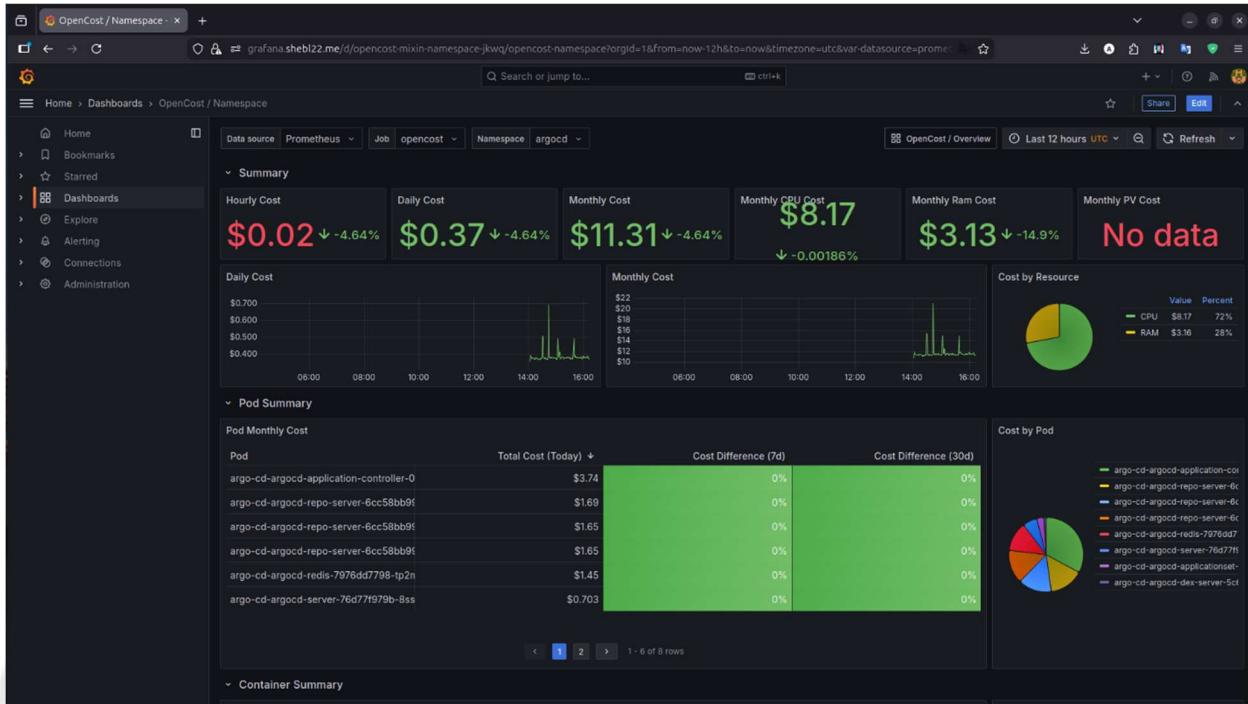
## Grafana:

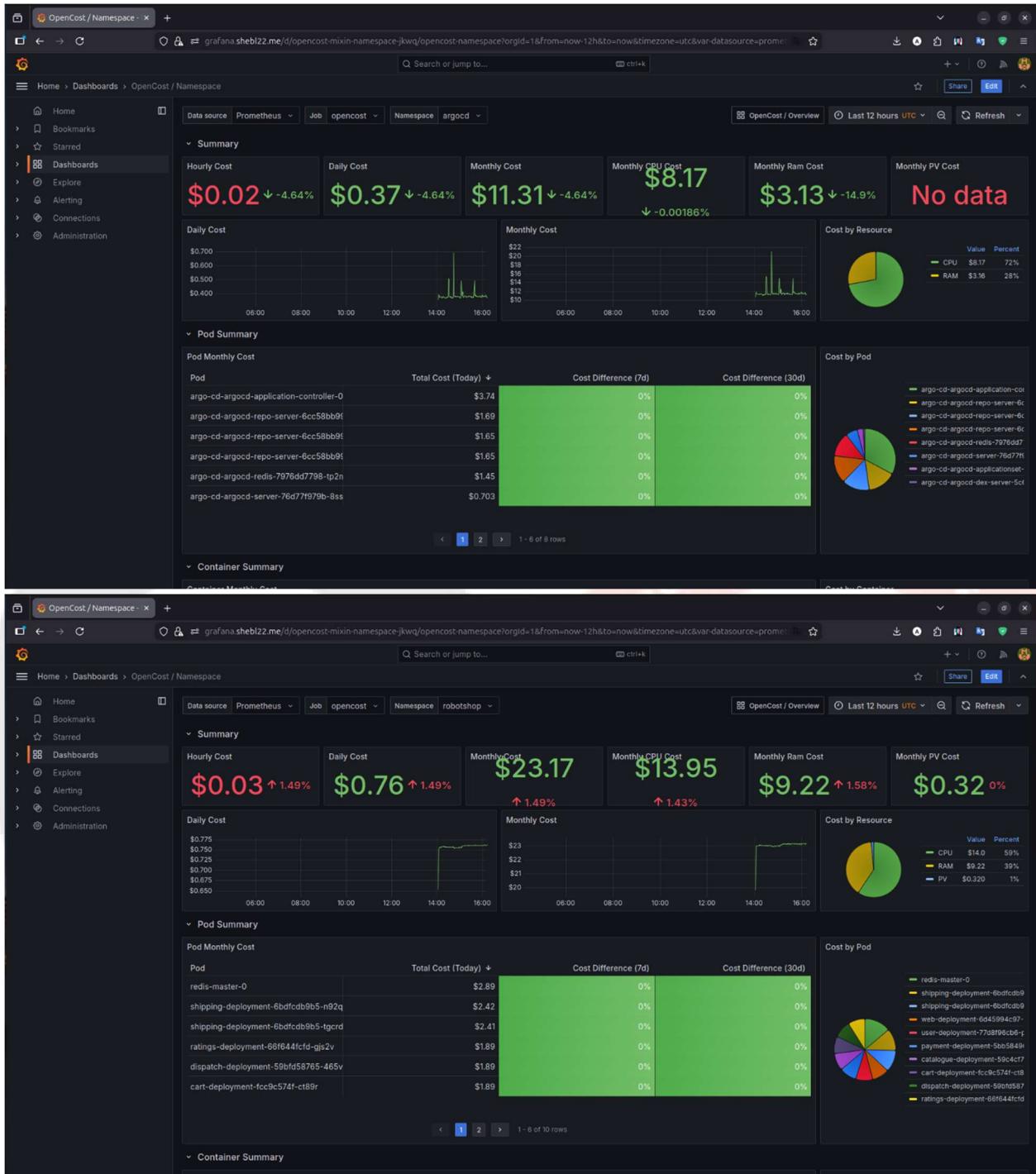


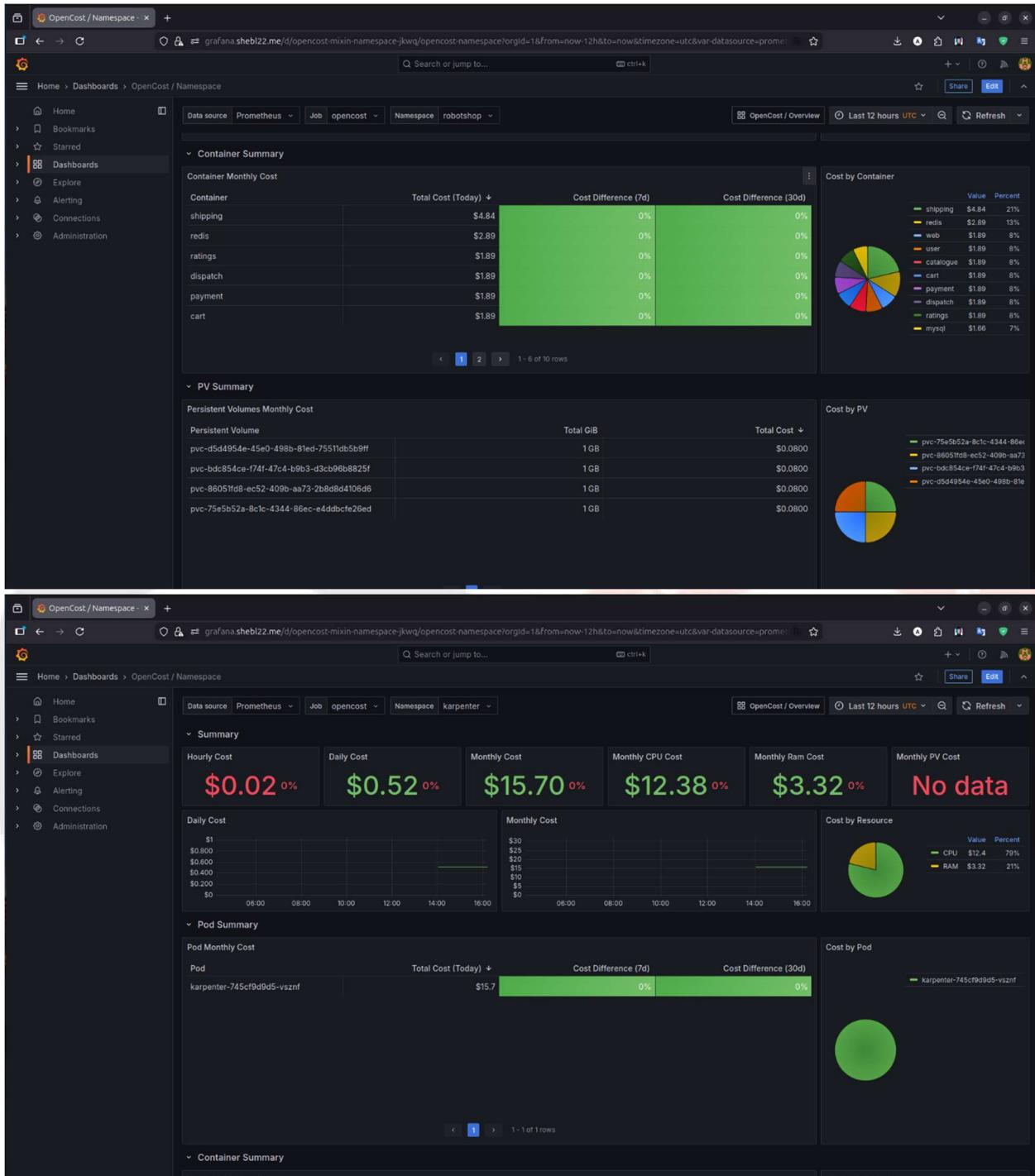


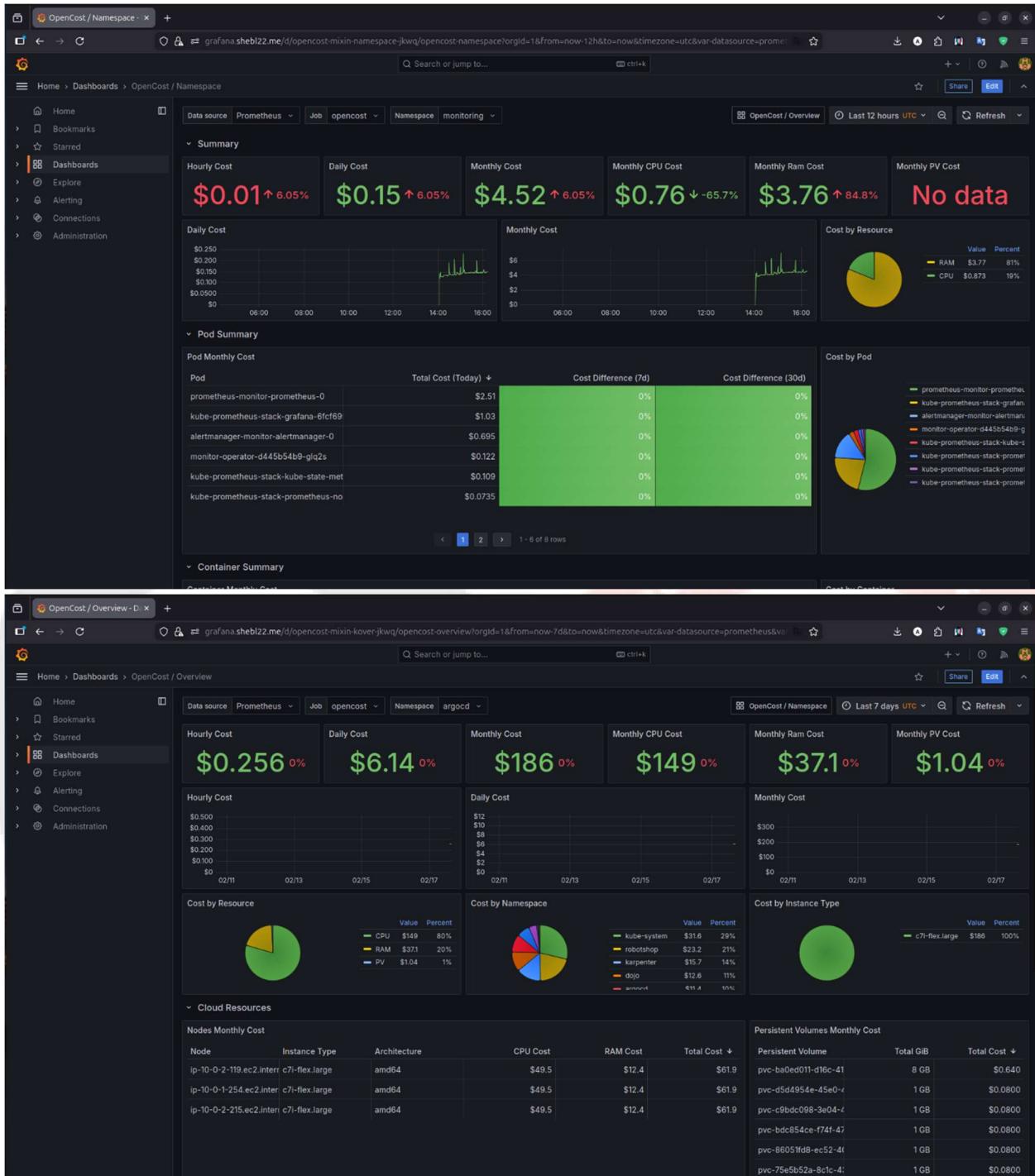


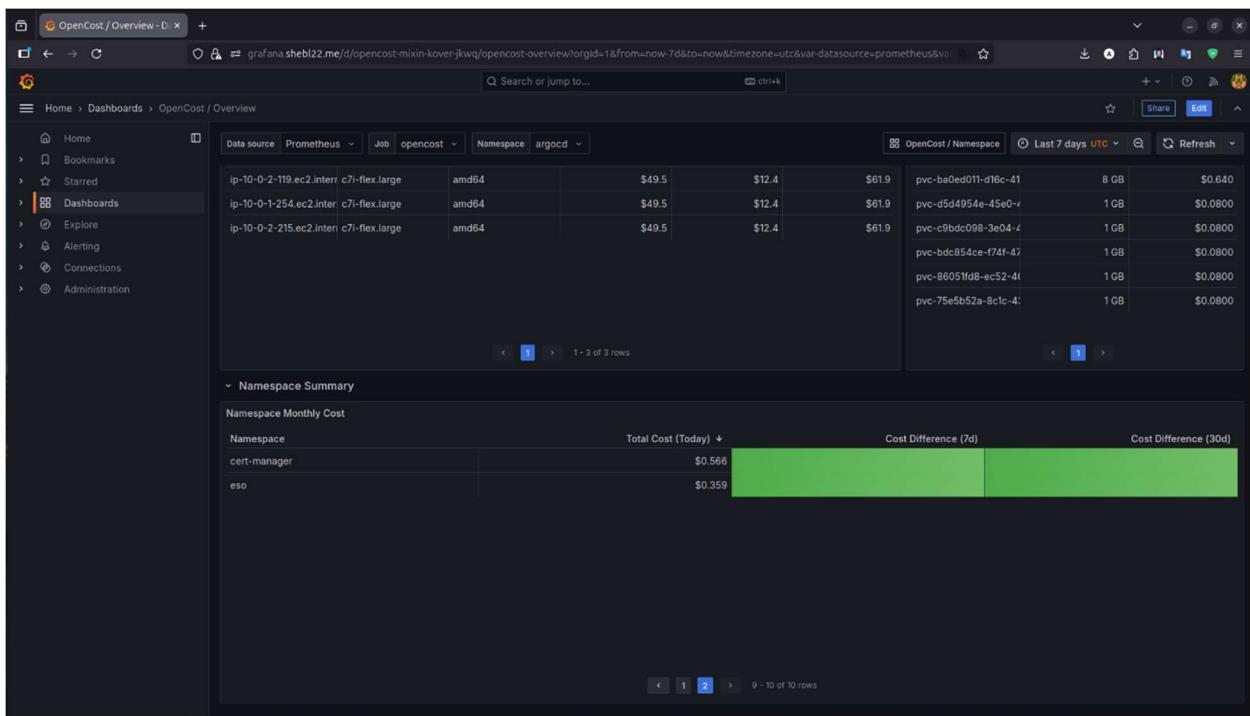
## Opencost dashboards:











## Platform-Level Integration Summary

Tool	Category	Role
Karpenter	Node Autoscaling	Dynamic infrastructure scaling
DefectDojo	Security	Vulnerability aggregation
Kyverno	Policy	Cluster security enforcement
OpenCost	FinOps	Cost visibility
Goldilocks	Optimization	Resource right-sizing
Prometheus Stack	Observability	Metrics & alerting backbone

## Cloud Infrastructure Layer Terraform + AWS

All infrastructure is provisioned using Terraform with official AWS modules.

Our infrastructure:

```
terraform/
├── main.tf          # EKS, VPC, Karpenter, Addons, IAM, Route53
├── variables.tf     # Cluster config (region, domain, EKS version)
├── outputs.tf        # Route53 nameservers
├── providers.tf      # AWS + Helm + Kubectl providers
└── modules/
    ├── addons/        # ArgoCD + ArgoCD Apps (deploys all charts)
    ├── karpenter/     # Karpenter Helm + NodePool + EC2NodeClass
    ├── ssm/           # AWS SSM Parameter Store (secrets)
    ├── eso/           # External Secrets Operator IAM
    ├── edns/          # External DNS IAM
    └── opencost/       # S3 + CUR + Glue + Athena for cost data
```

### Layer 1 Infrastructure Layer

(Terraform + AWS Foundation)

This layer provisions and configures all cloud infrastructure using Infrastructure as Code.

It is fully implemented using Terraform with official AWS modules.

The Kubernetes cluster runs on Amazon Elastic Kubernetes Service.

### Networking (VPC Layer)

Provisioned using the official VPC module.

Components:

- 2 Availability Zones
- Public subnets (Load Balancers)
- Private subnets (Worker nodes)
- NAT Gateway

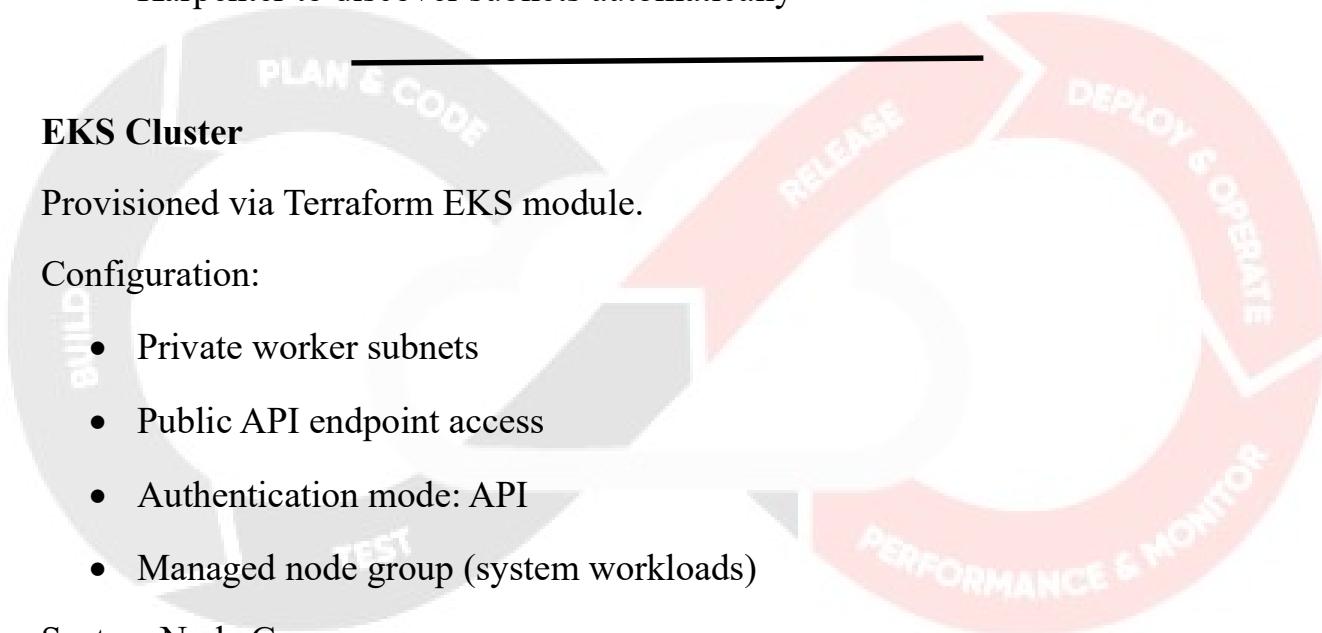
- Route tables

Kubernetes-Specific Tags:

- kubernetes.io/role/elb
- kubernetes.io/role/internal-elb
- karpenter.sh/discovery

These tags allow:

- LoadBalancer services to work correctly
- Karpenter to discover subnets automatically



System Node Group:

- Dedicated for core controllers
- Tainted: workload-type=system
- Hosts:
  - CoreDNS
  - metrics-server
  - Karpenter controller
  - AWS controllers

This prevents business workloads from running on system nodes.

---

## Karpenter Dynamic Node Provisioning

Node autoscaling is handled by  
Karpenter

Terraform Responsibilities:

- Create IAM role for nodes
- Create SQS interruption queue
- Attach required policies
- Enable Pod Identity for controller

Kubernetes Responsibilities:

- Karpenter controller runs in cluster
- Launches EC2 instances dynamically
- Prefers Spot instances
- Consolidates unused capacity

Result:

Cluster scales at node level automatically and cost-efficiently.

## Pod Identity (IAM Without IRSA)

Instead of IRSA, the platform uses EKS Pod Identity.

Terraform creates IAM roles and binds them to service accounts for:

- ExternalDNS
- External Secrets Operator
- cert-manager

- EBS CSI driver

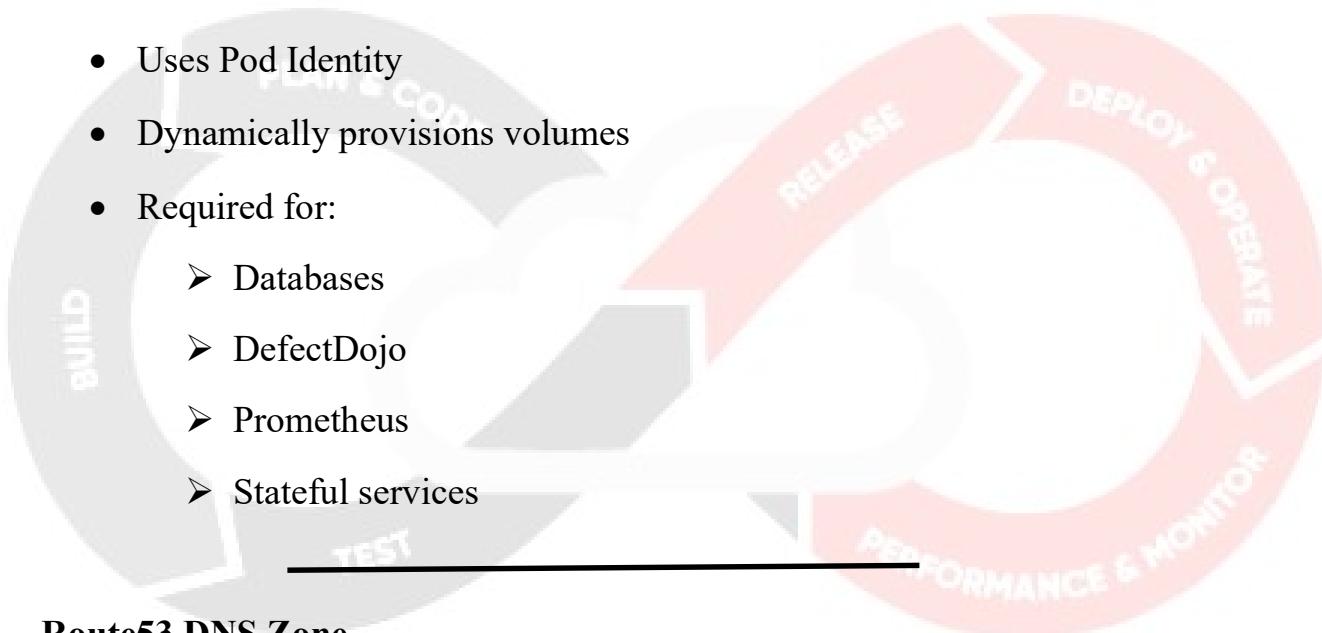
This ensures:

- No static AWS credentials
  - Least privilege access
  - Fine-grained AWS permissions per controller
- 

## Storage (EBS CSI Driver)

EBS CSI driver:

- Uses Pod Identity
  - Dynamically provisions volumes
  - Required for:
    - Databases
    - DefectDojo
    - Prometheus
    - Stateful services
- 



## Route53 DNS Zone

Terraform creates:

- Hosted zone
- DNS records managed dynamically

ExternalDNS uses this to automatically create domain records.

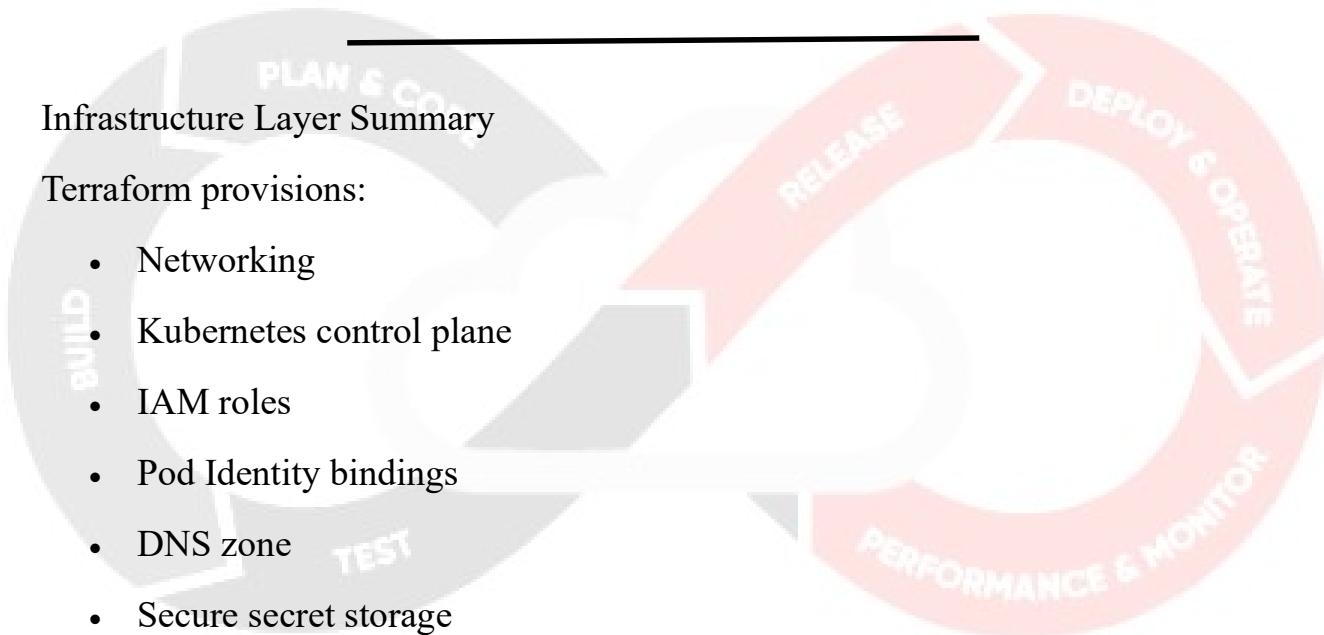
## SSM Parameter Store

Terraform provisions secure parameters in AWS SSM.

Used for:

- Application secrets
- OpenCost cloud integration
- Database credentials

Secrets **never** exist in Git.



This layer prepares AWS so Kubernetes controllers can operate securely.

## Layer 2 — GitOps & Application Layer

(ArgoCD + Helm)

This layer deploys and manages everything inside Kubernetes.

It is powered by

Argo CD

### GitOps Model

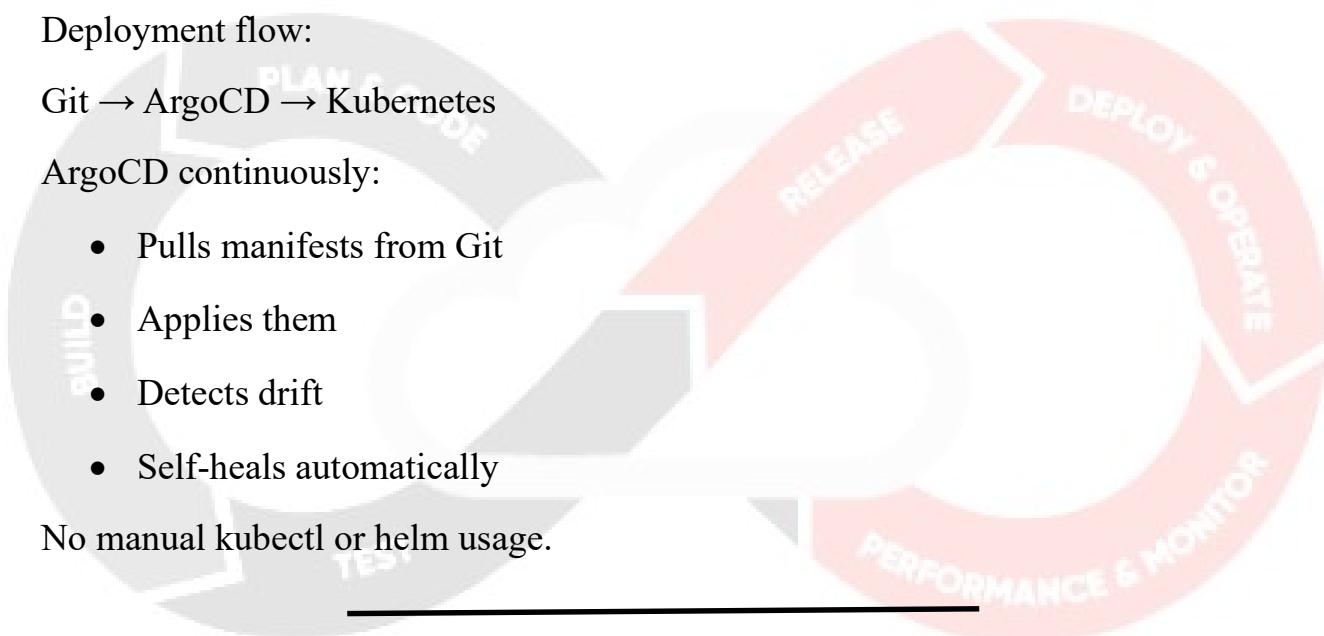
Deployment flow:

Git → ArgoCD → Kubernetes

ArgoCD continuously:

- Pulls manifests from Git
- Applies them
- Detects drift
- Self-heals automatically

No manual kubectl or helm usage.



### App-of-Apps Pattern

One root application manages all platform applications.

Each component is defined as an ArgoCD Application:

- cert-manager
- External Secrets
- ExternalDNS
- Traefik

- Prometheus stack
- OpenCost
- Goldilocks
- Kyverno
- DefectDojo
- Robot Shop application

Deletion uses cascade finalizers for clean teardown.

### Sync Waves (Ordered Deployment)

To ensure proper dependency ordering:

Wave	Component
-5	cert-manager
-4	ClusterIssuer
-3	ESO, ExternalDNS, Traefik, Prometheus
-2	Exporters, OpenCost, Goldilocks, Kyverno
-1	Kyverno policies
0	Robot Shop app

This guarantees:

- CRDs exist before CRs
- TLS before ingress
- Policies before workloads

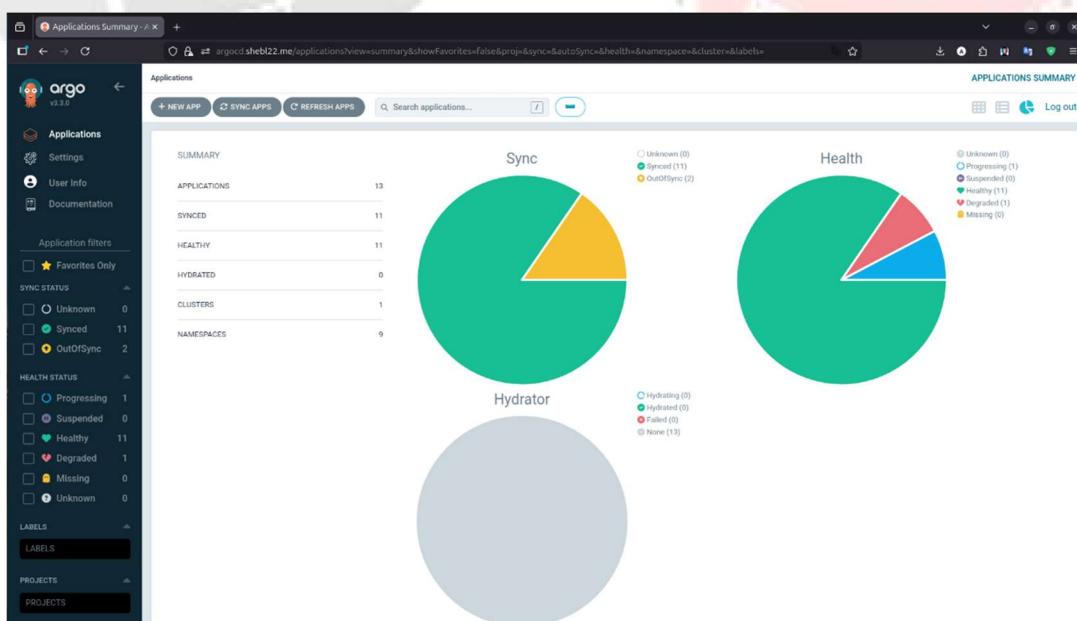
## Core Platform Controllers

- ❖ ExternalDNS
- ❖ External Secrets Operator
- ❖ cert-manager
- ❖ Kyverno
- ❖ OpenCost
- ❖ Goldilocks
- ❖ Prometheus Stack

Includes:

- Prometheus
  - Alertmanager
  - Grafana
  - Node exporters
  - kube-state-metrics
- ❖ DefectDojo

Results:



**Applications Tiles - Argo**

argocd.shebl22.me/applications?showFavorites=False&proj=&sync=&autoSync=&health=&namespace=&cluster=&labels=

**Argo v3.0**

**Applications**

+ NEW APP    SYNC APPS    REFRESH APPS    Search applications...    Log out

Sort: name ▾ Items per page: 15 ▾

Previous 1 2 Next

<b>cert-manager</b>	<b>cert-manager-manifests</b>	<b>defectdojo</b>	<b>external-dns</b>
Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Healthy Synced Repository: https://charts.jetstack.io Target: v1.17.1 Chart: cert-manager Destination: in-cluster Name: cert-manager Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 09:39:18 (an hour ago)	Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Healthy Synced Repository: https://github.com/abdelrahman-she... Target: feature/pipeline Path: K8s/cert-manager Destination: in-cluster Name: cert-manager Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 09:57:19 (37 minutes ago)	Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Progressing Synced Repository: https://raw.githubusercontent.com/D... Target: 1.9.12 Chart: defectdojo Destination: in-cluster Name: dojo Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 10:33:48 (a few seconds ago)	Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Healthy Synced Repository: https://kubernetes-sigs.github.io/extern... Target: external-dns Destination: in-cluster Name: edns Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 09:21:28 (an hour ago)
<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>	<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>	<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>	<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>

**external-secrets-manifests**

**external-secrets-operator**

**goldilocks**

**grafana-dashboards**

**LABELS**

**PROJECTS**

**Applications Tiles - Argo**

argocd.shebl22.me/applications?showFavorites=False&proj=&sync=&autoSync=&health=&namespace=&cluster=&labels=

**Argo v3.0**

**Applications**

+ NEW APP    SYNC APPS    REFRESH APPS    Search applications...    Log out

Sort: name ▾ Items per page: 15 ▾

Previous 1 2 Next

<b>kube-prometheus-stack</b>	<b>kyverno</b>	<b>kyverno-manifests</b>	<b>opencost</b>
Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Healthy Synced Repository: https://prometheus-community.github... Target: 6.8.2 Chart: kube-prometheus-stack Destination: in-cluster Name: monitoring Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 09:55:00 (40 minutes ago)	Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Healthy OutOfSync Repository: https://kyverno.github.io/kyverno/ Target: 3.7.0 Chart: kyverno Destination: in-cluster Name: kyverno Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 10:35:03 (a few seconds ago)	Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Healthy Synced Repository: https://github.com/abdelrahman-she... Target: feature/pipeline Path: K8s/kyverno Destination: in-cluster Name: kyverno Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 09:22:50 (an hour ago)	Project: default Labels: app.kubernetes.io/managed-by=Helm Status: Healthy Synced Repository: https://opencost.github.io/opencost-... Target: 2.5.5 Chart: opencost Destination: in-cluster Name: opencost Created: 02/22/2026 09:21:13 (an hour ago) Last Sync: 02/22/2026 09:22:48 (an hour ago)
<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>	<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>	<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>	<b>SYNC</b> <b>REFRESH</b> <b>DELETE</b>

**prometheus-mongodb-exporter**

**prometheus-mysql-exporter**

**robot-shop**

**LABELS**

**PROJECTS**

## Robot Shop Application

Deployed via umbrella Helm chart.

Features:

- HPA per service
- Network policies
- Strict security contexts
- Environment-based values
- Ingress with TLS

Observability & FinOps

Monitoring

- kube-prometheus-stack
- Custom Grafana dashboards

FinOps

OpenCost integrates with:

- AWS CUR
- S3
- Athena

Autoscaling Strategy

Karpenter provides:

- Spot-first provisioning
- On-demand fallback
- NodePool abstraction

Horizontal Pod Autscalers scale application services based on CPU and memory utilization.

## Supply Chain Security

The platform enforces image verification:

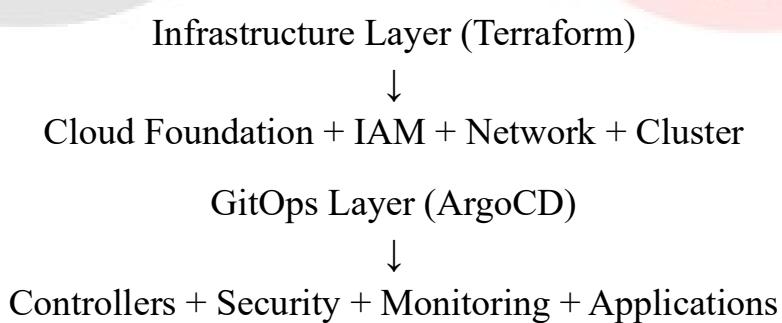
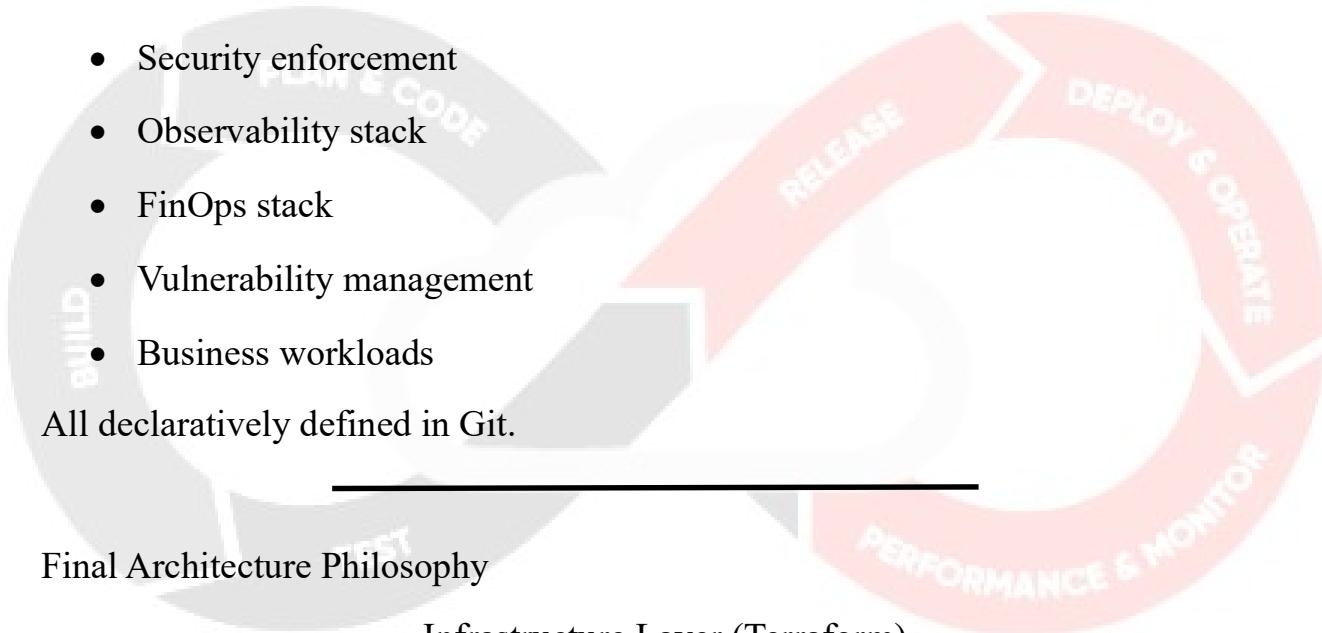
- All images signed via Cosign
- Kyverno verifies signatures before pod admission
- Unsigned images are rejected

## GitOps Layer Summary

ArgoCD manages:

- Platform controllers
- Security enforcement
- Observability stack
- FinOps stack
- Vulnerability management
- Business workloads

All declaratively defined in Git.



Together they create:

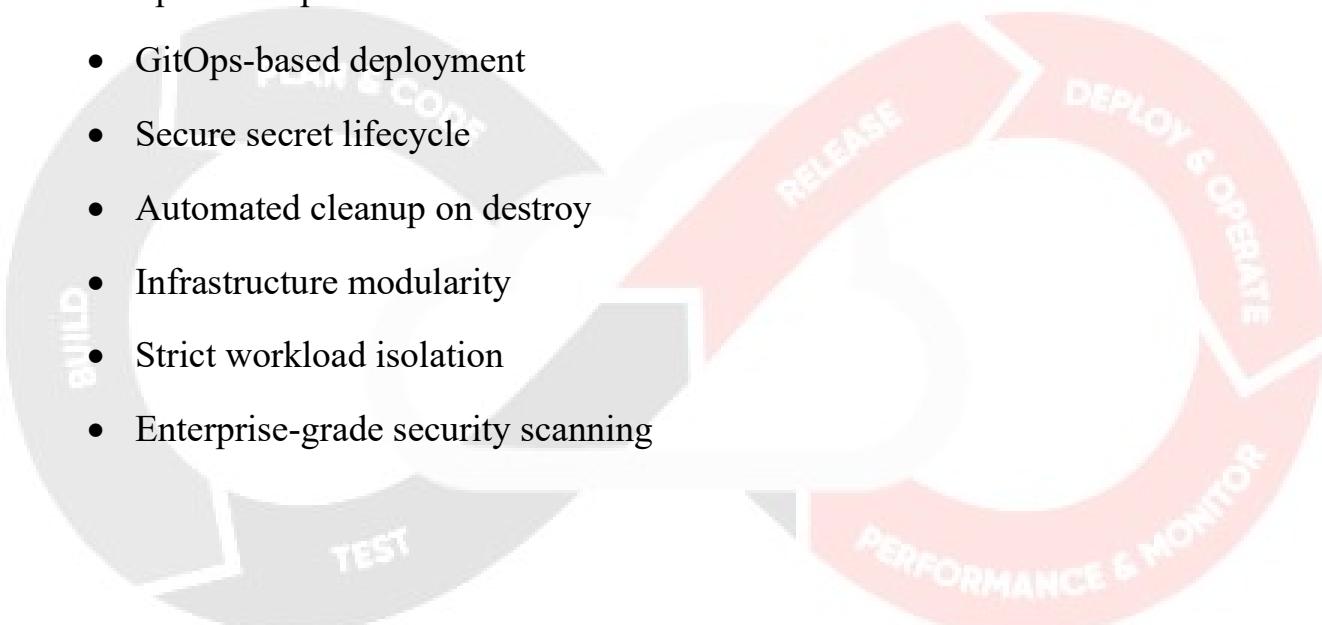
- Fully automated provisioning
- Secure workload isolation

- Cost-aware scaling
- Git-driven deployments
- Zero manual operations

## Production Characteristics

This platform includes:

- High availability
- Spot cost optimization
- GitOps-based deployment
- Secure secret lifecycle
- Automated cleanup on destroy
- Infrastructure modularity
- Strict workload isolation
- Enterprise-grade security scanning



# CI/CD Pipeline Full DevSecOps Automation

Pipeline implemented in GitHub Actions.

Triggered on:

- main
- feature/\*

Pipeline Flow

Push → SAST → Secret Scan → Image Build → Image Scan → Sign → DAST → Deploy

Security Layers

Tool	Type	Purpose
Semgrep	SAST	Code vulnerabilities
Gitleaks	Secret detection	Prevent leaked credentials
Docker Scout	SCA	CVE scanning
OWASP ZAP	DAST	Runtime scanning
Cosign	Supply chain	Image signing

Image Supply-Chain Security

- Images built with BuildKit
- Scanned for High/Critical CVEs
- Signed using Cosign (keyless via Sigstore)
- Verified in cluster by Kyverno

Unsigned images are rejected.

## GitOps Auto Deployment

Final stage:

1. Update Helm values.yaml image tag
2. Commit with [skip ci]
3. ArgoCD detects change
4. Automatic sync to cluster
5. Zero manual kubectl / helm usage

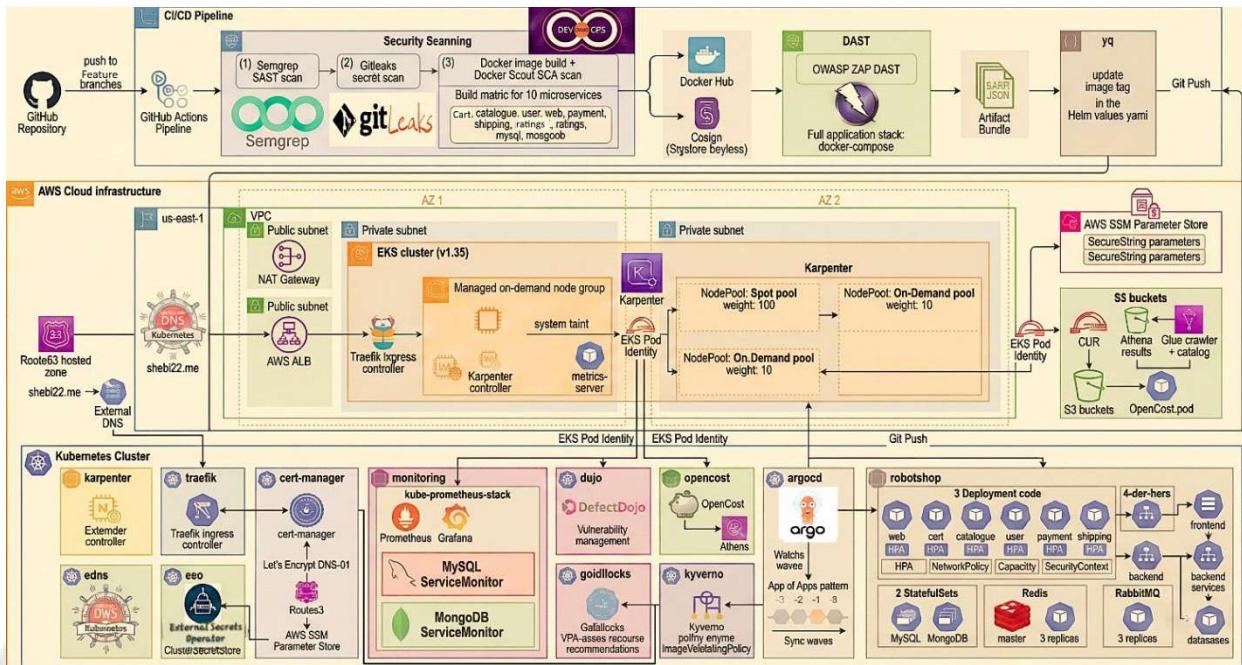
Flow Description

1. Secrets defined in variables.yaml (git-ignored)
2. Terraform stores them in AWS SSM as SecureStrings
3. EKS Pod Identity grants ESO IAM access
4. ESO syncs secrets into Kubernetes
5. Applications consume Kubernetes Secrets only

Benefits:

- No secret leakage
- No secret in Terraform state
- Centralized secret management

# Finally the infrastructure of the project:



## Final Project Structure

The project is organized into:

- Application services
- Helm umbrella chart
- Kubernetes manifests
- Terraform infrastructure
- CI/CD workflows
- Component documentation

Clear separation between:

- Infrastructure
- Platform services
- Application layer

## Conclusion

Robot Shop evolved from a demo microservices app into a full production-grade DevSecOps reference platform.

It demonstrates:

- Secure CI/CD pipelines
- GitOps automation
- Infrastructure as Code
- Cloud-native best practices
- Cost observability
- Kubernetes hardening
- Supply chain integrity

This architecture can serve as a blueprint for real-world enterprise Kubernetes deployments.

