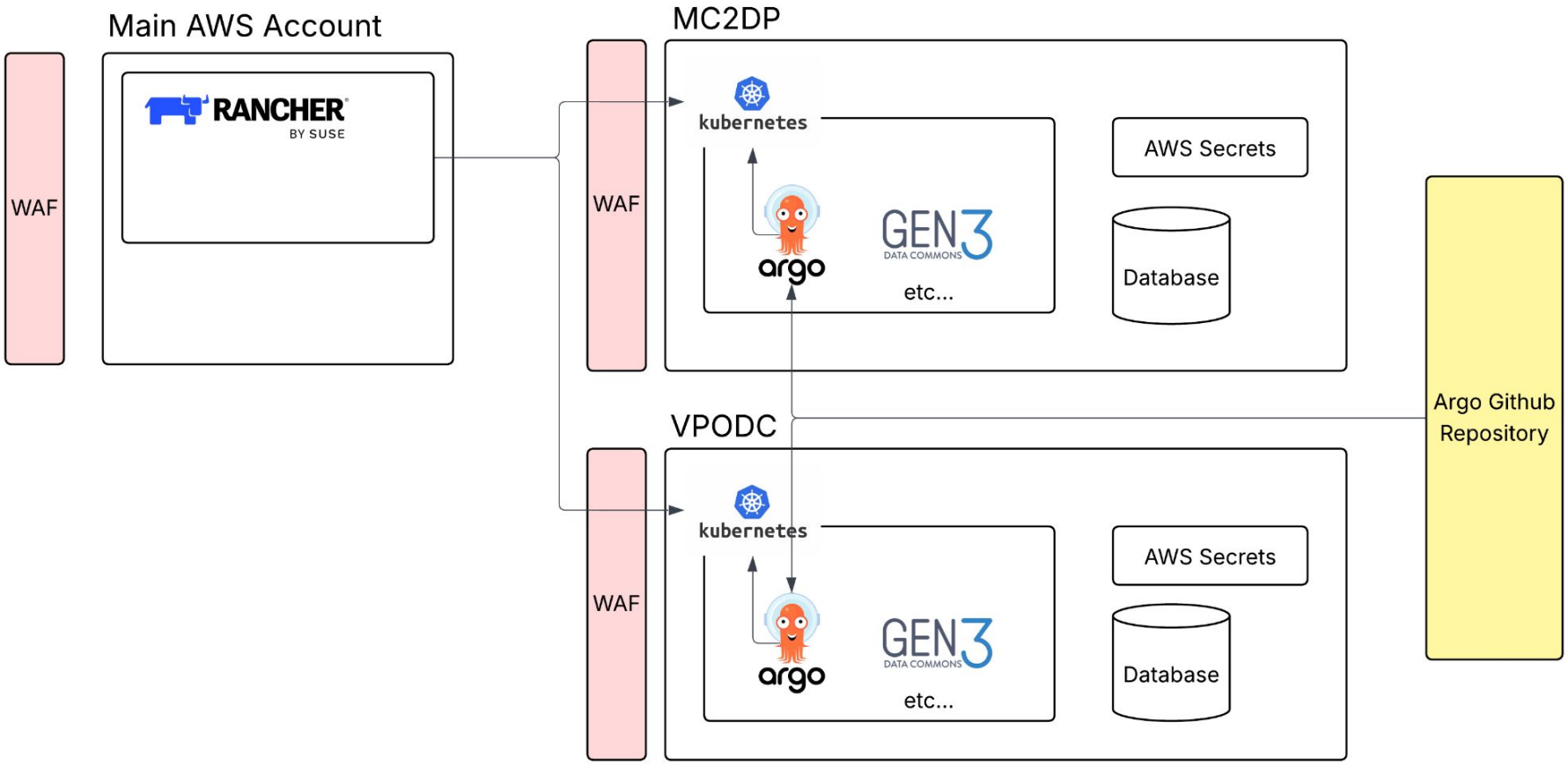
OCC/Krumware Partnership
CSOC Development


Main AWS Account
RANCHER
BY SUSE
WAF
WAF
MC2DP
kubernetes
argo
GEN3
DATA COMMONS
etc...
AWS Secrets
Database
VPODC
WAF
kubernetes
argo
GEN3
DATA COMMONS
etc...
AWS Secrets
Database
Argo Github Repository

Terraform Features (from github.com/krumIO/krum-gen3-aws-terraform)

- **Complete Infrastructure Deployment:** Creates all required AWS resources, Kubernetes clusters, and application configurations
- **Multi-Environment Support:** Deploy development, testing, and production environments with consistent configurations
- **Multi-Commons Support:** Host multiple data commons instances in a single infrastructure
- **Rancher Integration:** Manages Kubernetes clusters through Rancher for improved operations
- **Security-Focused:** Implements AWS and Kubernetes security best practices throughout
- **Secret Management:** Automatically rotates OpenSearch credentials and S3 bucket credentials
- **Database Management:** Provides secure and scalable PostgreSQL databases with automated user provisioning and access control
- **Certificate Management:** Supports both HTTP01 and DNS01 validation for Let's Encrypt certificates
- **Load Balancer Integration:** Configurable AWS Load Balancer Controller with ALB/NLB support
- **DNS Management:** External DNS integration for automatic DNS record creation

Terraform Setup Procedure

`occ-csoc-infra` Create new Terraform file for data commons (mostly boilerplate!)

`occ_gen3_dev.tf` (Example configuration changes)

Argo CD Git repository

Our local AWS profile

```
locals {
  # Github
  github_org = "occ-data"
  github_repo = "occ-csoc-deploy"

  # AWS profiles
  occ_gen3_dev_aws_profile = "occ-gen3-dev"

  # AWS region
  occ_gen3_dev_aws_region = "us-east-1"

  # Rancher configuration
  rancher_api_url = "https://rancher.occ-pla.net"
  rancher_token_key = var.rancher_token != "" ?
var.rancher_token :
module.occ_rancher.rancher_token_key
}
```

```
# Configure AWS provider for occ-gen3-dev
account
provider "aws" {
  alias   = "occ_gen3_dev"
  region  = local.occ_gen3_dev_aws_region
  profile = local.occ_gen3_dev_aws_profile
}
```

```
# Create IAM User, Roles, and Policies for
Rancher in Downstream Account
module "occ_gen3_dev_rancher_permissions" {
  source =
"git::https://github.com/krumio/krum-gen3-aws-
terraform.git//modules/aws_iam_rancher_managem
ent?ref=v2.0.2"
  providers = {
    aws = aws.occ_gen3_dev
  }
}
```

```
// Create Route53 NS record in Parent zone hosted in
occ_rancher module provider
resource "aws_route53_record"
"gen3_environment_dev_mc2dp_ns" {
  provider = aws.occ_rancher # must match the aws
provider used for the parent zone hosted in the
occ_rancher module provider
  zone_id  = data.aws_route53_zone.occ_pla_net.zone_id
  name     = "mc2dp.occ-pla.net"
  type     = "NS"
  ttl      = "300"
  records  =
data.aws_route53_zone.mc2dp_occ_pla_net.name_servers
  depends_on = [
    module.occ_gen3_dev # Make sure the zone is
created by the occ_gen3_dev module before adding the
NS record
  ]
}
```

Our desired domain

"`occ_gen3_dev`" is prefix to make module unique, can be anything; we decided to use this pattern

Continuous Deployment with Argo

Argo watches a Github repository for changes, then deploys as needed

occ-csoc-deploy

- `fence.yaml`
- `guppy.yaml`
- `indexd.yaml`
- `values.yaml`
- GEN3

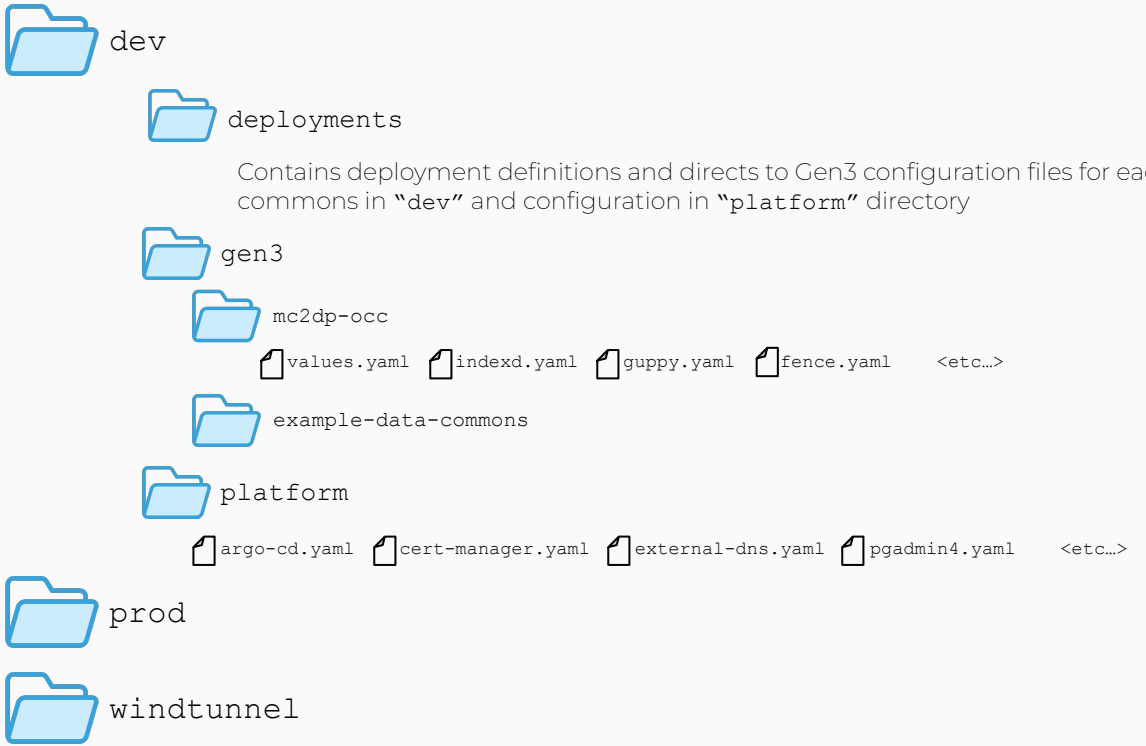
Other Services in Environment

- Argo CD
- Certificate Manager
- External DNS
- External Secrets
- Jaeger (monitoring)
- Loki & Promtail (logging)
- AWS LB Controller
- PGAdmin

Deploy Repository Folder Structure (for ArgoCD)

`occ-csoc-deploy`

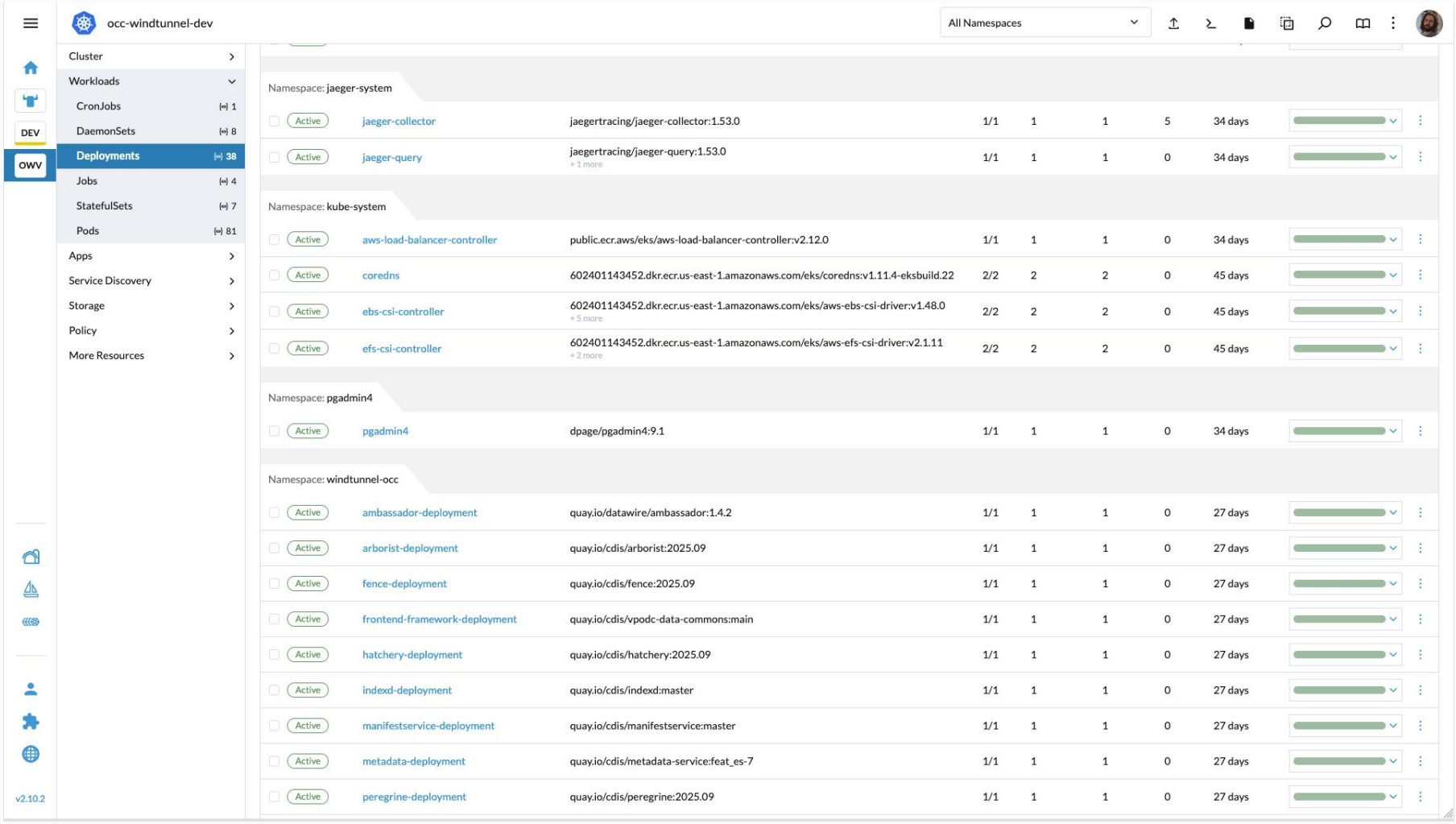
Each environment has its own path

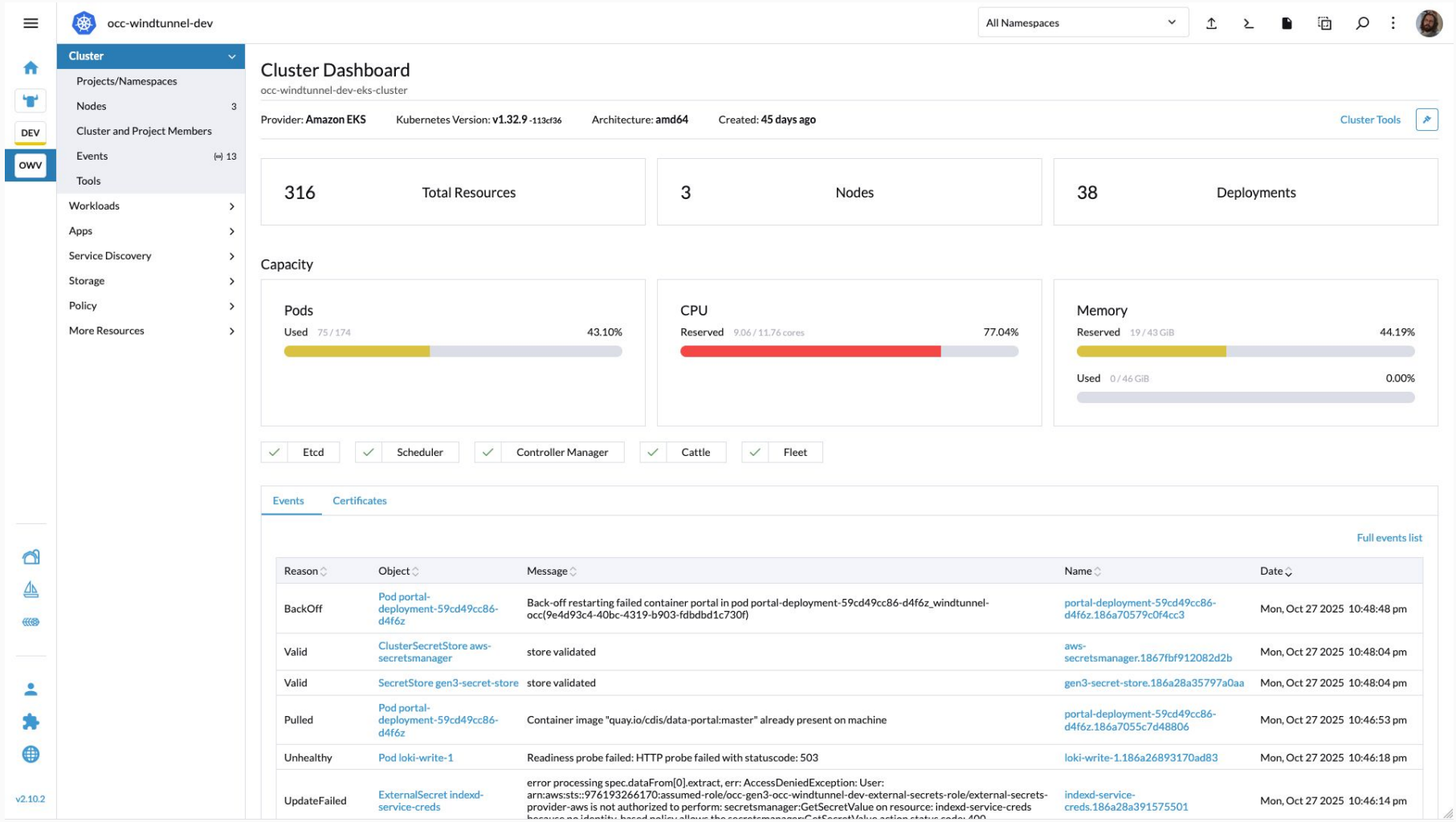


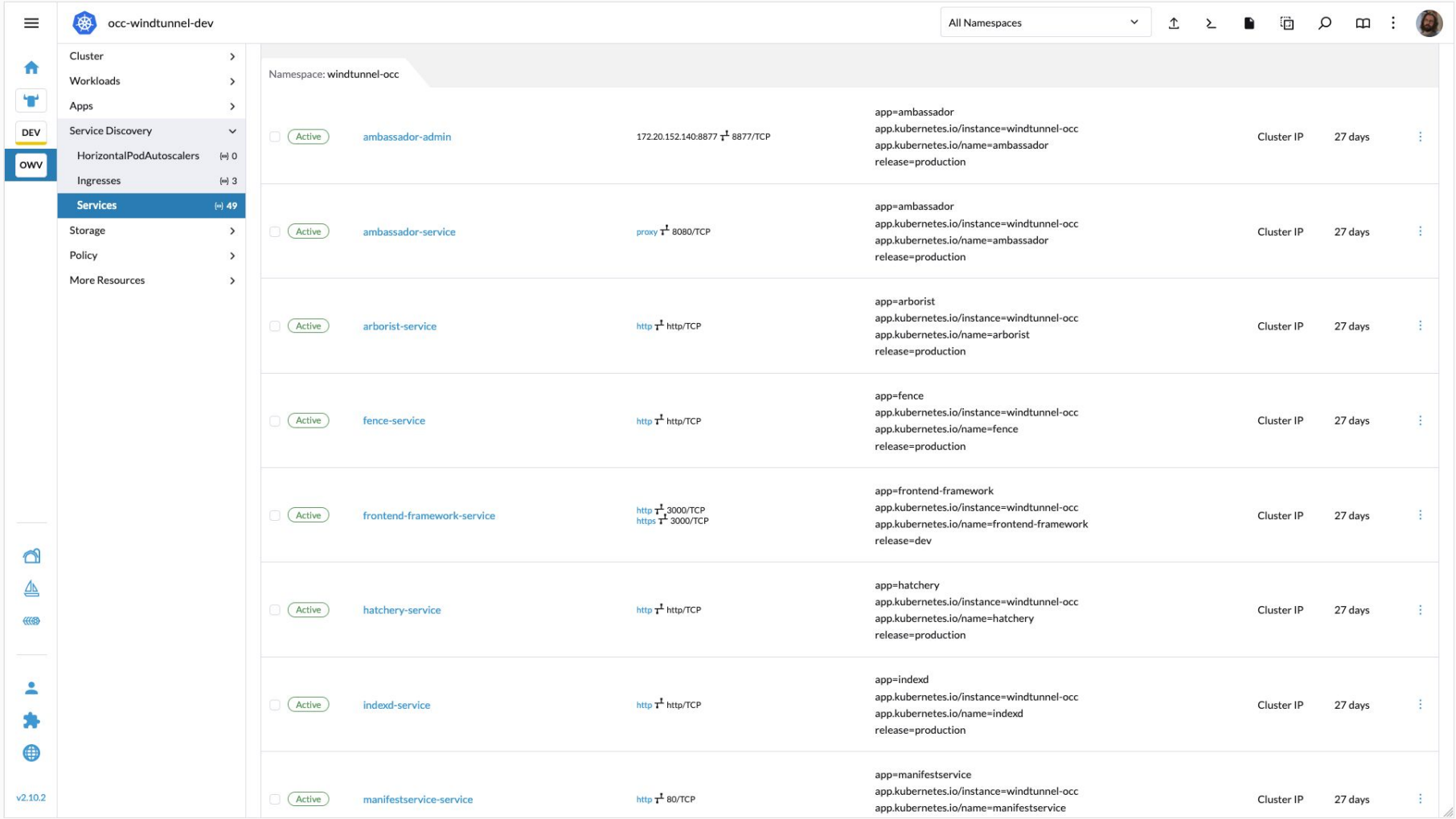
(these files are referenced in the Terraform config. For example…)

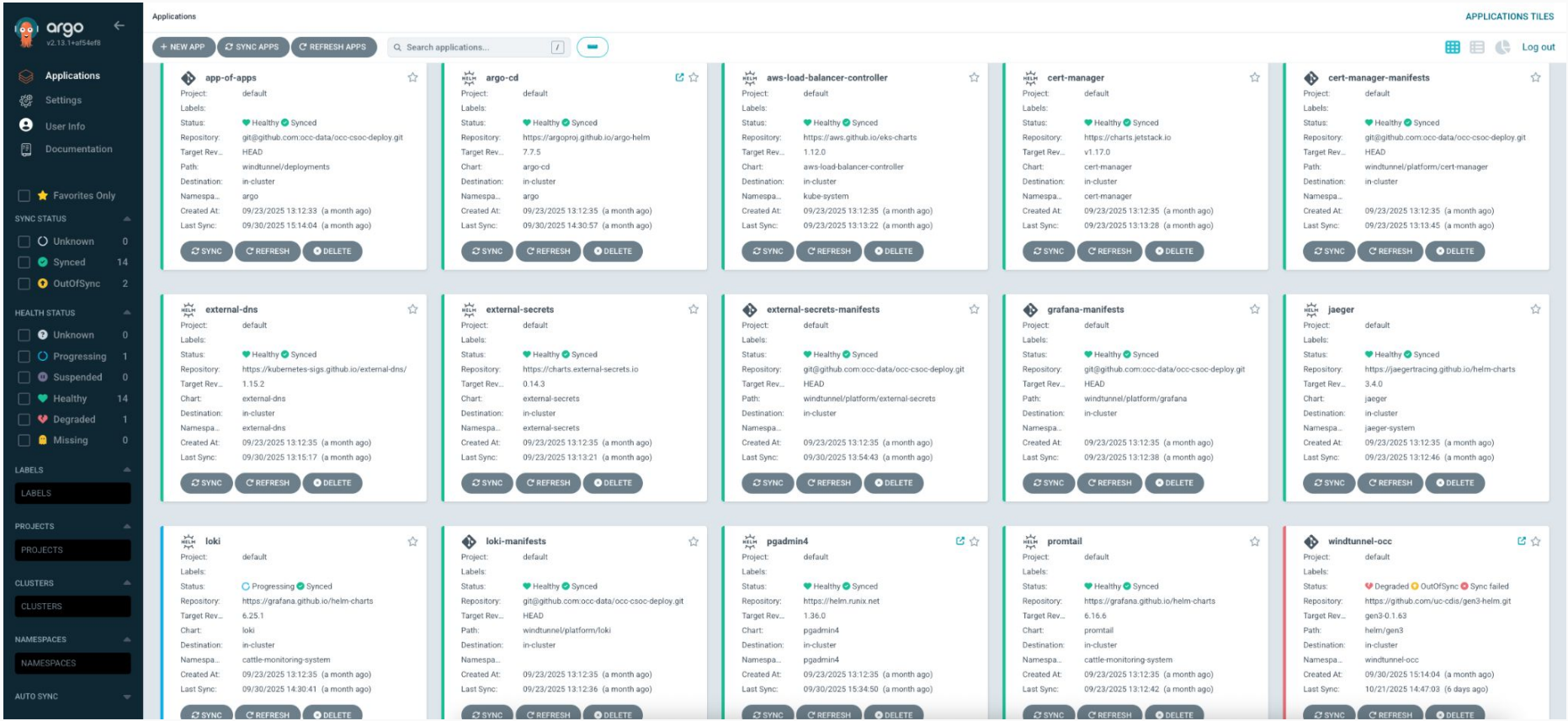
`occ_gen3_dev.tf`

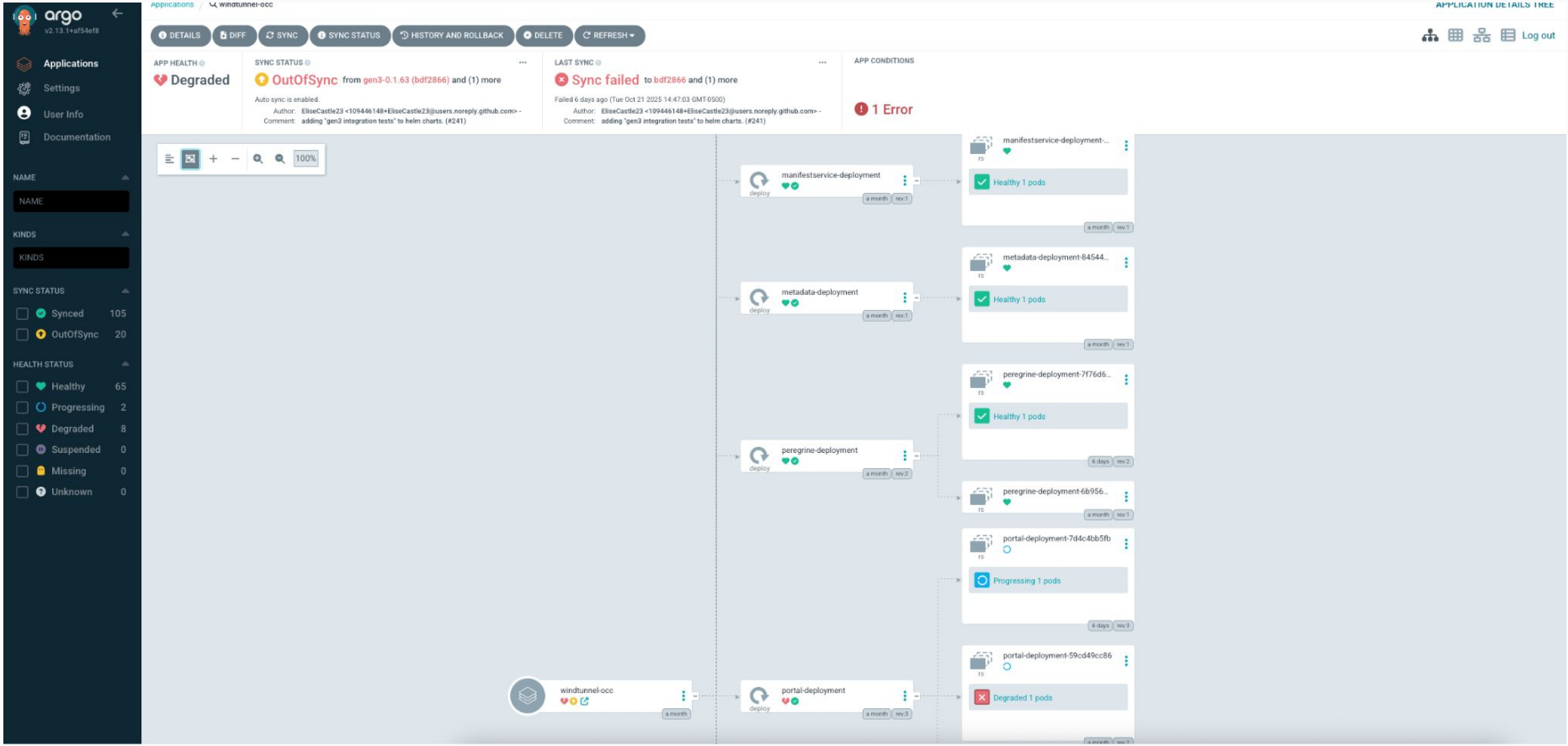
```
github_repo = local.github_repo
gitops_path = "dev/services"
gitops_branch = "main"
        gitops_external_secrets_manifest_path =
"dev/platform/external-secrets"
        gitops_external_dns_values_path =
"dev/platform"
        gitops_loki_manifest_path =
"dev/platform/loki"
        commons_env = {
          mc2dp = {
            secrets_namespace = "mc2dp-occ"
            commons_gitops_manifest_path =
"dev/gen3/mc2dp-occ/manifests"
```


occ-windtunnel-dev
All Namespaces
Cluster
Workloads
CronJobs 1
DaemonSets 8
Deployments 38
Jobs 4
StatefulSets 7
Pods 81
Apps
Service Discovery
Storage
Policy
More Resources
DEV
OWV
v2.10.2
Namespace: jaeger-system
Active jaeger-collector jaegertracing/jaeger-collector:1.53.0 1/1 1 1 5 34 days
Active jaeger-query jaegertracing/jaeger-query:1.53.0 + 1 more 1/1 1 1 0 34 days
Namespace: kube-system
Active aws-load-balancer-controller public.ecr.aws/eks/aws-load-balancer-controller:v2.12.0 1/1 1 1 0 34 days
Active coredns 602401143452.dkr.ecr.us-east-1.amazonaws.com/eks/coredns:v1.11.4-eksbuild.22 2/2 2 2 0 45 days
Active ebs-csi-controller 602401143452.dkr.ecr.us-east-1.amazonaws.com/eks/aws-ebs-csi-driver:v1.48.0 + 5 more 2/2 2 2 0 45 days
Active efs-csi-controller 602401143452.dkr.ecr.us-east-1.amazonaws.com/eks/aws-efs-csi-driver:v2.1.11 + 2 more 2/2 2 2 0 45 days
Namespace: pgadmin4
Active pgadmin4 dpage/pgadmin4:9.1 1/1 1 1 0 34 days
Namespace: windtunnel-occ
Active ambassador-deployment quay.io/datawire/ambassador:1.4.2 1/1 1 1 0 27 days
Active arborist-deployment quay.io/cdis/arborist:2025.09 1/1 1 1 0 27 days
Active fence-deployment quay.io/cdis/fence:2025.09 1/1 1 1 0 27 days
Active frontend-framework-deployment quay.io/cdis/vpodc-data-commons:main 1/1 1 1 0 27 days
Active hatchery-deployment quay.io/cdis/hatchery:2025.09 1/1 1 1 0 27 days
Active indexd-deployment quay.io/cdis/indexd:master 1/1 1 1 0 27 days
Active manifestservice-deployment quay.io/cdis/manifestservice:master 1/1 1 1 0 27 days
Active metadata-deployment quay.io/cdis/metadata-service:feat_es-7 1/1 1 1 0 27 days
Active peregrine-deployment quay.io/cdis/peregrine:2025.09 1/1 1 1 0 27 days


occ-windtunnel-dev
All Namespaces
DEV
OWV
Cluster
Projects/Namespaces
Nodes 3
Cluster and Project Members
Events (=) 13
Tools
Workloads
Apps
Service Discovery
Storage
Policy
More Resources
Cluster Dashboard
occ-windtunnel-dev-eks-cluster
Provider: Amazon EKS
Kubernetes Version: v1.32.9 -113cf36
Architecture: amd64
Created: 45 days ago
Cluster Tools
316 Total Resources
3 Nodes
38 Deployments
Capacity
Pods
Used 75 / 174 43.10%
CPU
Reserved 9.06 / 11.76 cores 77.04%
Memory
Reserved 19 / 43 GiB 44.19%
Used 0 / 46 GiB 0.00%
Etcd
Scheduler
Controller Manager
Cattle
Fleet
Events
Certificates
Full events list
Reason
Object
Message
Name
Date
BackOff
Pod portal-deployment-59cd49cc86-d4f6z
Back-off restarting failed container portal in pod portal-deployment-59cd49cc86-d4f6z_windtunnel-occ(9e4d93c4-40bc-4319-b903-fdbdbd1c730f)
portal-deployment-59cd49cc86-d4f6z.186a70579c0f4cc3
Mon, Oct 27 2025 10:48:48 pm
Valid
ClusterSecretStore aws-secretsmanager
store validated
aws-secretsmanager.1867fbf912082d2b
Mon, Oct 27 2025 10:48:04 pm
Valid
SecretStore gen3-secret-store
store validated
gen3-secret-store.186a28a35797a0aa
Mon, Oct 27 2025 10:48:04 pm
Pulled
Pod portal-deployment-59cd49cc86-d4f6z
Container image "quay.io/cdis/data-portal:master" already present on machine
portal-deployment-59cd49cc86-d4f6z.186a7055c7d48806
Mon, Oct 27 2025 10:46:53 pm
Unhealthy
Pod loki-write-1
Readiness probe failed: HTTP probe failed with statuscode: 503
loki-write-1.186a26893170ad83
Mon, Oct 27 2025 10:46:18 pm
UpdateFailed
ExternalSecret indexd-service-creds
error processing spec.dataFrom[0].extract, err: AccessDeniedException: User: arn:aws:sts::976193266170:assumed-role/occ-gen3-occ-windtunnel-dev-external-secrets-role/external-secrets-provider-aws is not authorized to perform: secretsmanager:GetSecretValue on resource: indexd-service-creds
indexd-service-creds.186a28a391575501
Mon, Oct 27 2025 10:46:14 pm
v2.10.2


occ-windtunnel-dev
All Namespaces
Cluster
Workloads
Apps
Service Discovery
HorizontalPodAutoscalers 0
Ingresses 3
Services 49
Storage
Policy
More Resources
DEV
OWV
v2.10.2
Namespace: windtunnel-occ
Active ambassador-admin 172.20.152.140:8877 8877/TCP app=ambassador app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=ambassador release=production Cluster IP 27 days
Active ambassador-service proxy 8080/TCP app=ambassador app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=ambassador release=production Cluster IP 27 days
Active arborist-service http http/TCP app=arborist app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=arborist release=production Cluster IP 27 days
Active fence-service http http/TCP app=fence app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=fence release=production Cluster IP 27 days
Active frontend-framework-service http 3000/TCP https 3000/TCP app=frontend-framework app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=frontend-framework release=dev Cluster IP 27 days
Active hatchery-service http http/TCP app=hatchery app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=hatchery release=production Cluster IP 27 days
Active indexd-service http http/TCP app=indexd app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=indexd release=production Cluster IP 27 days
Active manifestservice-service http 80/TCP app=manifestservice app.kubernetes.io/instance=windtunnel-occ app.kubernetes.io/name=manifestservice Cluster IP 27 days


argo
v2.13.1+af54ef8
Applications
Settings
User Info
Documentation
Favorites Only
SYNC STATUS
Unknown 0
Synced 14
OutOfSync 2
HEALTH STATUS
Unknown 0
Progressing 1
Suspended 0
Healthy 14
Degraded 1
Missing 0
LABELS
LABELS
PROJECTS
PROJECTS
CLUSTERS
CLUSTERS
NAMESPACES
NAMESPACES
AUTO SYNC
Applications
APPLICATIONS TILES
+ NEW APP
SYNC APPS
REFRESH APPS
Search applications...
Log out
app-of-apps
Project: default
Labels:
Status: Healthy Synced
Repository: git@github.com:occ-data/occ-csoc-deploy.git
Target Rev.. HEAD
Path: windtunnel/deployments
Destination: in-cluster
Namespa.. argo
Created At: 09/23/2025 13:12:33 (a month ago)
Last Sync: 09/30/2025 15:14:04 (a month ago)
SYNC REFRESH DELETE
argo-cd
Project: default
Labels:
Status: Healthy Synced
Repository: https://argoproj.github.io/argo-helm
Target Rev.. 7.7.5
Chart: argo-cd
Destination: in-cluster
Namespa.. argo
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/30/2025 14:30:57 (a month ago)
SYNC REFRESH DELETE
aws-load-balancer-controller
Project: default
Labels:
Status: Healthy Synced
Repository: https://aws.github.io/eks-charts
Target Rev.. 1.12.0
Chart: aws-load-balancer-controller
Destination: in-cluster
Namespa.. kube-system
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:13:22 (a month ago)
SYNC REFRESH DELETE
cert-manager
Project: default
Labels:
Status: Healthy Synced
Repository: https://charts.jetstack.io
Target Rev.. v1.17.0
Chart: cert-manager
Destination: in-cluster
Namespa.. cert-manager
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:13:28 (a month ago)
SYNC REFRESH DELETE
cert-manager-manifests
Project: default
Labels:
Status: Healthy Synced
Repository: git@github.com:occ-data/occ-csoc-deploy.git
Target Rev.. HEAD
Path: windtunnel/platform/cert-manager
Destination: in-cluster
Namespa..
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:13:45 (a month ago)
SYNC REFRESH DELETE
external-dns
Project: default
Labels:
Status: Healthy Synced
Repository: https://kubernetes-sigs.github.io/external-dns/
Target Rev.. 1.15.2
Chart: external-dns
Destination: in-cluster
Namespa.. external-dns
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/30/2025 13:15:17 (a month ago)
SYNC REFRESH DELETE
external-secrets
Project: default
Labels:
Status: Healthy Synced
Repository: https://charts.external-secrets.io
Target Rev.. 0.14.3
Chart: external-secrets
Destination: in-cluster
Namespa.. external-secrets
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:13:21 (a month ago)
SYNC REFRESH DELETE
external-secrets-manifests
Project: default
Labels:
Status: Healthy Synced
Repository: git@github.com:occ-data/occ-csoc-deploy.git
Target Rev.. HEAD
Path: windtunnel/platform/external-secrets
Destination: in-cluster
Namespa..
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/30/2025 13:54:43 (a month ago)
SYNC REFRESH DELETE
grafana-manifests
Project: default
Labels:
Status: Healthy Synced
Repository: git@github.com:occ-data/occ-csoc-deploy.git
Target Rev.. HEAD
Path: windtunnel/platform/grafana
Destination: in-cluster
Namespa..
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:12:38 (a month ago)
SYNC REFRESH DELETE
jaeger
Project: default
Labels:
Status: Healthy Synced
Repository: https://jaegertracing.github.io/helm-charts
Target Rev.. 3.4.0
Chart: jaeger
Destination: in-cluster
Namespa.. jaeger-system
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:12:46 (a month ago)
SYNC REFRESH DELETE
loki
Project: default
Labels:
Status: Progressing Synced
Repository: https://grafana.github.io/helm-charts
Target Rev.. 6.25.1
Chart: loki
Destination: in-cluster
Namespa.. cattle-monitoring-system
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/30/2025 14:30:41 (a month ago)
SYNC REFRESH DELETE
loki-manifests
Project: default
Labels:
Status: Healthy Synced
Repository: git@github.com:occ-data/occ-csoc-deploy.git
Target Rev.. HEAD
Path: windtunnel/platform/loki
Destination: in-cluster
Namespa..
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:12:36 (a month ago)
SYNC REFRESH DELETE
pgadmin4
Project: default
Labels:
Status: Healthy Synced
Repository: https://helm.runix.net
Target Rev.. 1.36.0
Chart: pgadmin4
Destination: in-cluster
Namespa.. pgadmin4
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/30/2025 15:34:50 (a month ago)
SYNC REFRESH DELETE
promtail
Project: default
Labels:
Status: Healthy Synced
Repository: https://grafana.github.io/helm-charts
Target Rev.. 6.16.6
Chart: promtail
Destination: in-cluster
Namespa.. cattle-monitoring-system
Created At: 09/23/2025 13:12:35 (a month ago)
Last Sync: 09/23/2025 13:12:42 (a month ago)
SYNC REFRESH DELETE
windtunnel-occ
Project: default
Labels:
Status: Degraded OutOfSync Sync failed
Repository: https://github.com/uc-cdis/gen3-helm.git
Target Rev.. gen3-0.1.63
Path: helm/gen3
Destination: in-cluster
Namespa.. windtunnel-occ
Created At: 09/30/2025 15:14:04 (a month ago)
Last Sync: 10/21/2025 14:47:03 (6 days ago)
SYNC REFRESH DELETE


argo
v2.13.1+af54ef8
Applications
Settings
User Info
Documentation
NAME
KINDS
SYNC STATUS
Synced 105
OutOfSync 20
HEALTH STATUS
Healthy 65
Progressing 2
Degraded 8
Suspended 0
Missing 0
Unknown 0
Applications / windtunnel-occ
APPLICATION DETAILS TREE
DETAILS
DIFF
SYNC
SYNC STATUS
HISTORY AND ROLLBACK
DELETE
REFRESH
Log out
APP HEALTH
Degraded
SYNC STATUS
OutOfSync from gen3-0.1.63 (bdf2866) and (1) more
Auto sync is enabled.
Author: EliseCastle23 <109446148+EliseCastle23@users.noreply.github.com> -
Comment: adding "gen3 integration tests" to helm charts. (#241)
LAST SYNC
Sync failed to bdf2866 and (1) more
Failed 6 days ago (Tue Oct 21 2025 14:47:03 GMT-0500)
Author: EliseCastle23 <109446148+EliseCastle23@users.noreply.github.com> -
Comment: adding "gen3 integration tests" to helm charts. (#241)
APP CONDITIONS
1 Error
100%
windtunnel-occ
manifestservice-deployment
manifestservice-deployment-...
Healthy 1 pods
metadata-deployment
metadata-deployment-84544...
Healthy 1 pods
peregrine-deployment
peregrine-deployment-7f76d6...
Healthy 1 pods
peregrine-deployment-6b956...
portal-deployment
portal-deployment-7d4c4bb5fb
Progressing 1 pods
portal-deployment-59cd49cc86
Degraded 1 pods

Key Benefits

- Works with OCC paradigm of separate AWS accounts
- Developed with best security practices; ensures each commons is properly secured
- WAF allows simple control over who can access server
- Reduces complexity; Rancher gives central view of all data commons
- Argo CD built-in, continuous deployment for all data commons
- Uses existing Gen3 Helm charts
- Logging/monitoring built into each commons environment (in development)

Potential Concerns

- Covers similar ground as gen3-terraform
  - Distinguishing Characteristics:
    - Single repo can deploy to multiple AWS repositories
    - Rancher very useful when dealing with several data commons
    - Argo CD built into deployment
- Currently only works with AWS
    - Terraform does not limit us to AWS, but development work is required to extend to Azure/GCP/etc.

Future Directions

- Terraform module will be made open source and public
- Users and system based logging and monitoring to be included soon
- Disaster Recovery for Gen3 commons
- OAuth/SSO for Rancher administration
- Enhanced Network Observability with OpenTelemetry
- Layer 7 Firewall and Container Security
- Addition of Researcher Authentication Services (RAS)