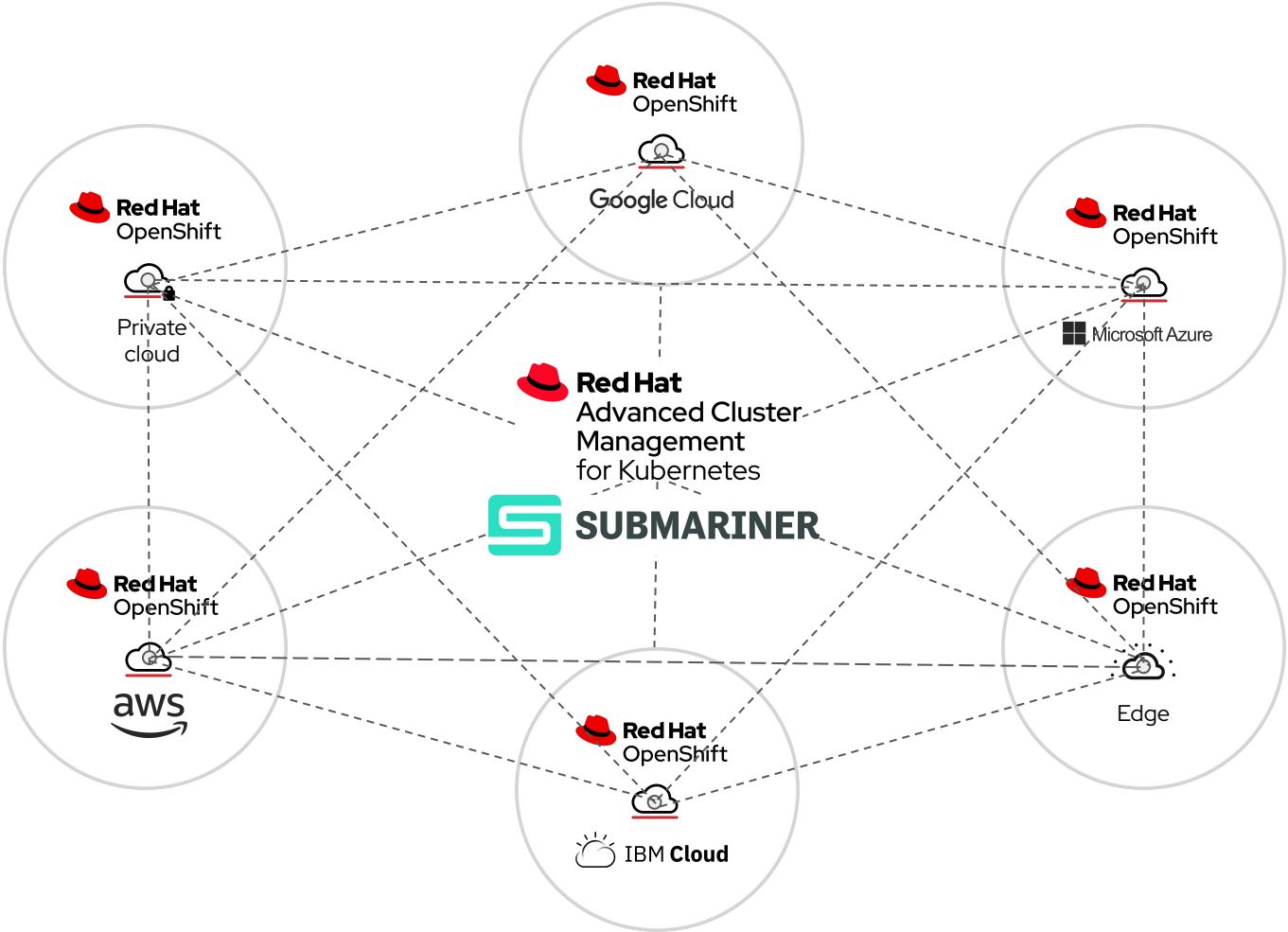
Submariner: Connecting Workloads Across Kubernetes Clusters

Red Hat

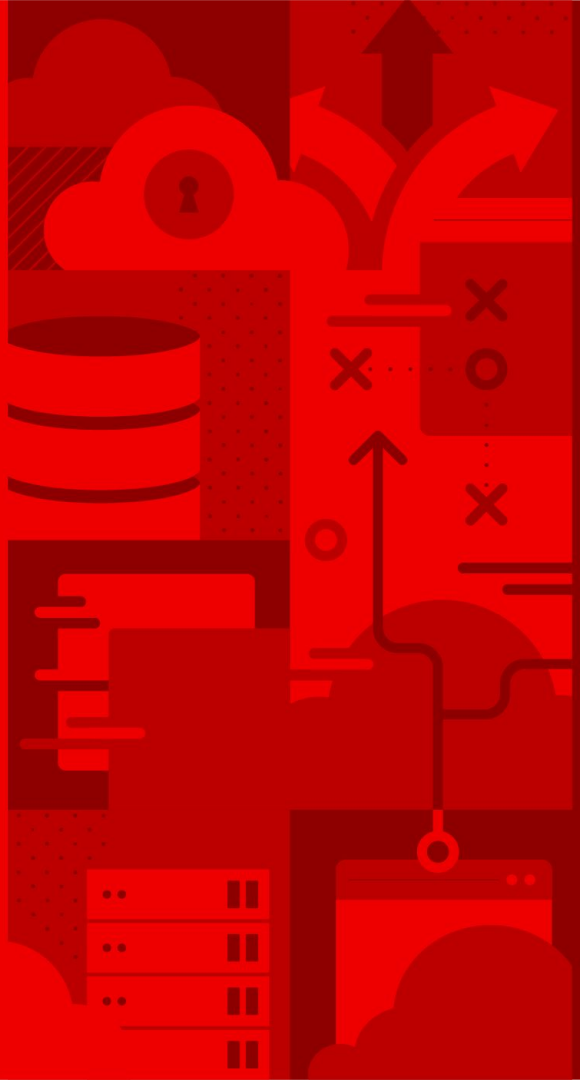

Red Hat OpenShift
Google Cloud
Red Hat OpenShift
Private cloud
Red Hat OpenShift
Microsoft Azure
Red Hat Advanced Cluster Management for Kubernetes
SUBMARINER
Red Hat OpenShift
aws
Red Hat OpenShift
Edge
Red Hat OpenShift
IBM Cloud

Red Hat

Agenda

- Market Trends and Challenges
- Multicluster Networking
- Introducing Submariner
- Architecture Overview
- Integration with Red Hat ACM

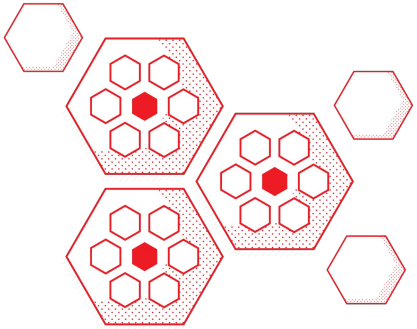
Red Hat



Market Trends and Challenges

Red Hat

Kubernetes adoption leads to multicluster



“As Kubernetes gains adoption across the industry, scenarios are arising in which teams are finding they must deploy and manage multiple clusters, either in a single region on-premises or in the cloud, or across multiple regions…. for a number of reasons, including multi-tenancy, disaster recovery, and with hybrid, multi-cloud, or edge deployments.”

Source:
Assessing Patterns for Deploying Distributed Kubernetes Clusters doc # G00465217, by Tony Iams

Red Hat

Where is the growth in cluster deployments?

Small Scale Dev Teams

Managing clusters across Dev/QA/Prod clusters

Medium Scale Organizations

Local retails with clusters across 100s of locations

Large Scale Organizations

Global organizations with 100s of clusters, hosting thousand of applications

Edge Scale Telco

100s of zones, 1000s of clusters and nodes across complex topologies

Red Hat

Reasons for deploying multiple clusters

Application availability

Reduced latency

Address industry standards

Geopolitical data residency guidelines

Disaster recovery

Edge deployments

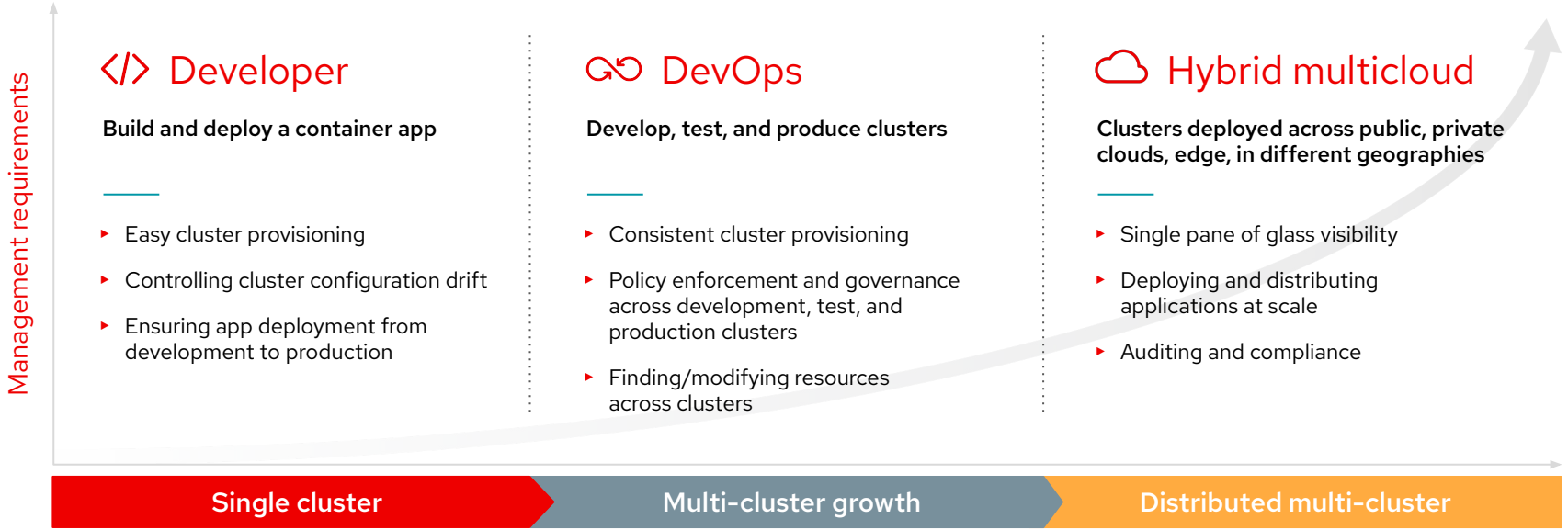
CapEx cost reduction

Avoid vendor lock-in

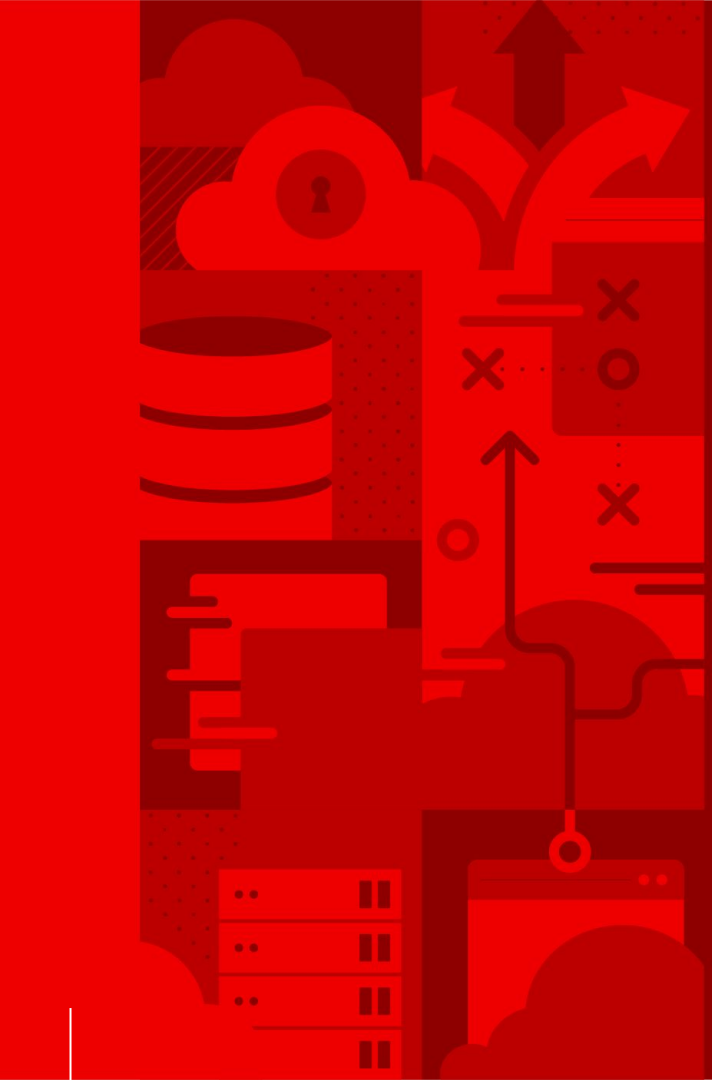
Red Hat

Multicluster management challenges

How do I normalize and centralize key functions across environments?



Red Hat



Multicluster Networking

Red Hat

Why is multi-cluster connectivity important?

- Multi-cluster connectivity is at the core of Red Hat's open hybrid cloud strategy and required for a wide range of use cases
- Our customers demand choice: a robust solution that works across different infrastructure providers and network (CNI) plug-ins
- Complement Red Hat's product portfolio:
  - OpenShift Container Platform
  - Advanced Cluster Management for Kubernetes (ACM)
  - Red Hat OpenShift Data Foundation (previously known as OCS)
  - Red Hat Service Mesh (Istio)

Red Hat

Multicluster networking

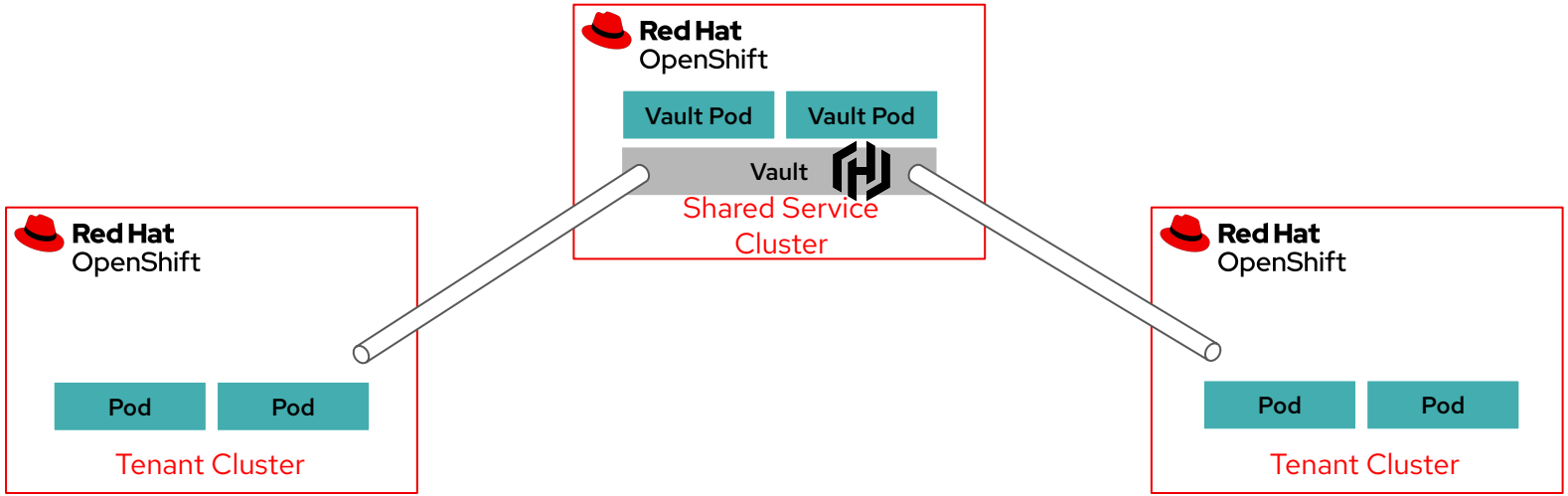
Use case for connecting multiple clusters: secure inter-service communication

- OpenShift clusters deployed on different infrastructure providers
- Some components of an app deployed in one cluster, others in another cluster
- Goal: secure service-to-service communication across clusters

Red Hat

Multicluster networking

Use case for connecting multiple clusters: shared services

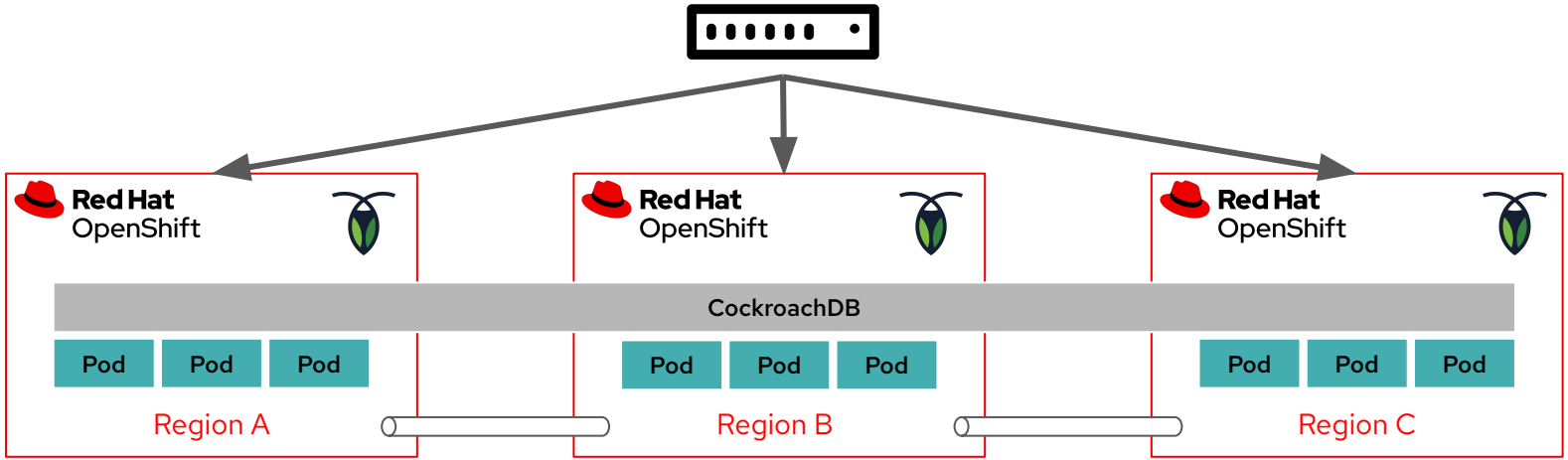


- Example: setting up Vault as a common source for secrets, certificates and credentials across clusters
- Also applicable for other common services like logging, monitoring, SSO, and metrics collection
- Goal: keep tenant clusters isolated, while avoiding operational overhead in maintaining shared services

Red Hat

Multicluster networking

Use case for connecting multiple clusters: Distributed Data



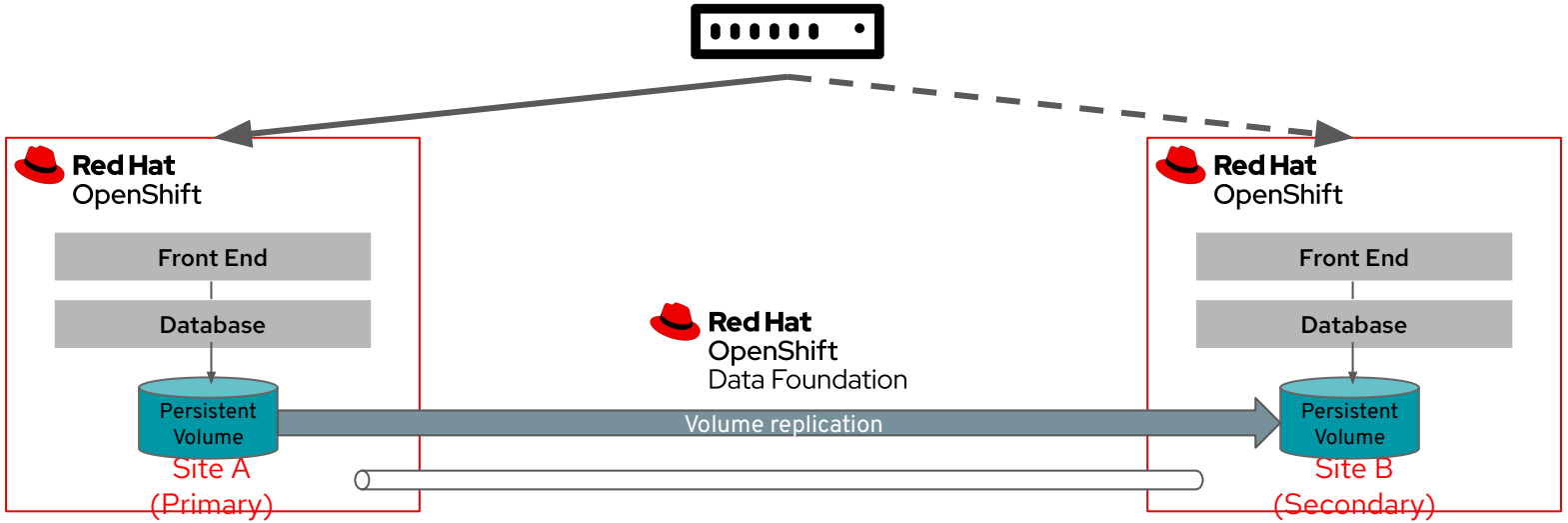
- Example: multi-region CockroachDB cluster
- OpenShift clusters in multiple regions, with replicas of the same service running in each cluster
- Goal: keep data close to the user, while reducing latency and fault tolerance to improve user experience
- Blog post | demo

Red Hat

GA with OCP 4.14

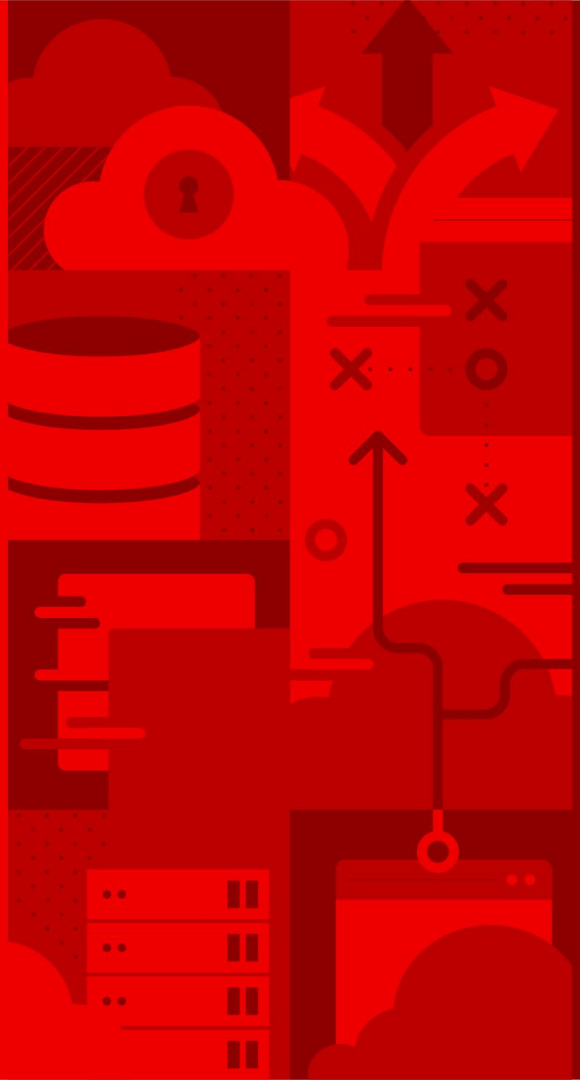
Multicluster networking

Use case for connecting multiple clusters: OpenShift Data Foundation Regional Disaster Recovery



- OpenShift clusters deployed across primary/secondary sites. ODF enables cross-cluster replication of data volumes
- Automated per application failover management through ACM
- Goal: protection against geographic disasters
- Documentation | demo

Red Hat



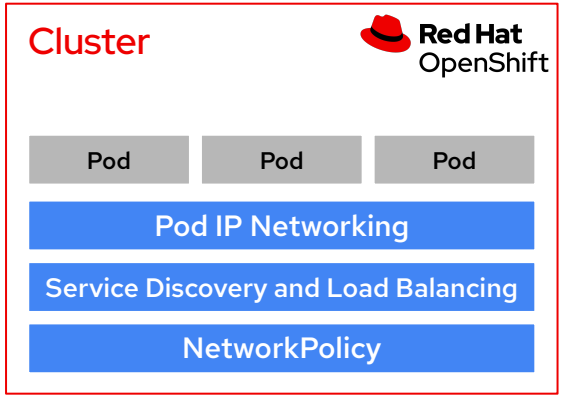
Introducing Submariner

Red Hat

Submariner project

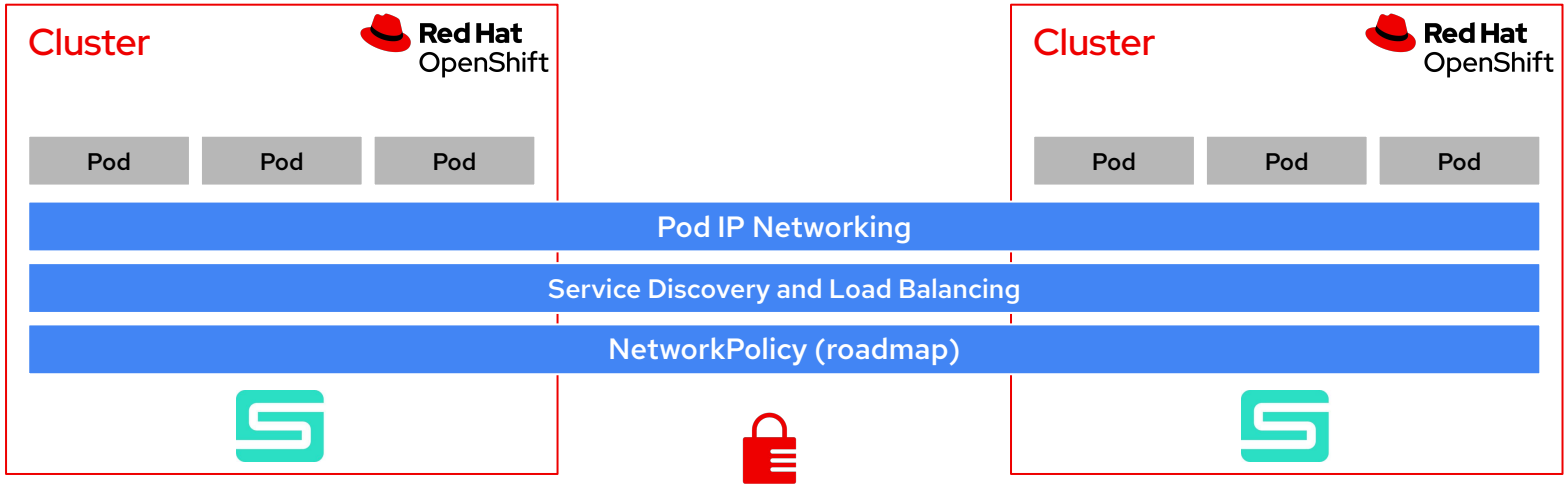
- “Connect all your Kubernetes clusters, no matter where they are in the world”
- Open source, vendor neutral project: https://submariner.io/
- Originally started by Rancher; now a CNCF Sandbox project

Kubernetes cluster networking



Red Hat

Kubernetes cluster networking with Submariner



- Different regions of the same public cloud provider
- Multiple public clouds
- Multiple on-prem sites
- Hybrid cloud, including a mix of on-prem and public cloud

Red Hat

Key personas

IT Operations

- Monitor usage across multiple clusters and cloud providers

SRE/NetOps

- Automate provisioning/deprovisioning of cluster interconnections
- Understand network infrastructure health and impact on cross-cluster application availability

SecOps

- Set consistent network policies across multiple clusters and ensure compliance

Application Developer

- Develop and deploy services to multiple clusters

Red Hat

Usage

1. Admin joins two or more clusters
    - Submariner provides full IP reachability between pods and services among the participating clusters, aka **ClusterSet**
2. Application developers then **export** selected **services** to expose them across the ClusterSet
    - Submariner automatically sets up DNS for the exported services

Step 1 (network setup) is done once to create inter-cluster L3 connectivity.

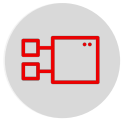
Step 2 (service export) can then be performed on-demand, leveraging the underlying connectivity.

Red Hat

Benefits

Pod-to-pod and pod-to-service routing with native performance

Direct network tunnel to support any application on top; eliminate the need for proxies, external load-balancers or ingress gateways



Deploy services across clusters

Beyond connectivity, also address the challenge of cross-cluster service discovery and network policy (roadmap)

Enhanced security

All traffic flow between clusters is encrypted using IPsec by default

Extend existing OpenShift deployments

Compatible with different infrastructure providers and network (CNI) plugins; benefit the wider OpenShift ecosystem

Red Hat

Key features

- Cross-cluster “east/west” L3 connectivity
  - Using encrypted or unencrypted connections
- Service Discovery across clusters
  - Implements the MCS API to facilitate multi-cluster DNS; *cluster.local* becomes *clusterset.local*
  - Support for ClusterIP as well as headless services
- Support for interconnecting clusters with overlapping CIDRs (Globalnet)
- `subctl` - a CLI utility that simplifies the deployment and management of Submariner
- Integration with Red Hat Advanced Cluster Management for Kubernetes (ACM)

Red Hat

Competitive landscape

Key differentiators for Submariner

- Rich and performant network connectivity
- Open source, standard-based
- Supported across a variety of infrastructure providers
- Compatible with most network (CNI) plug-ins

VMware/Tanzu

- Multi-cluster networking, load-balancing, and service mesh federation via their NSX portfolio
- Proprietary, forcing vendor lock-in

Public cloud providers

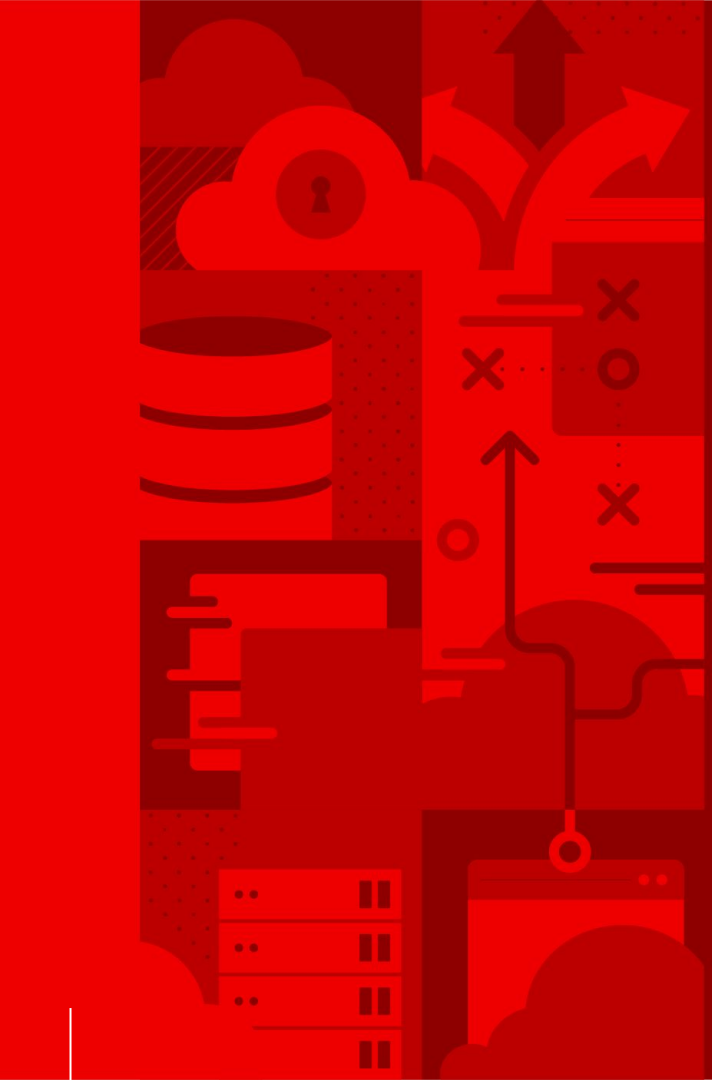
- Feature set tailored to their infrastructure or available in specific regions only (e.g Cloud VPN, VPC Peering)
- Proprietary, forcing infrastructure lock-in

Calico, Cilium

- Include multi-cluster networking capabilities as part of their Kubernetes CNI implementation
- Not available with OpenShift SDN/OVN; also limit customers who want to “mix and match” CNIs

Rancher/SUSE

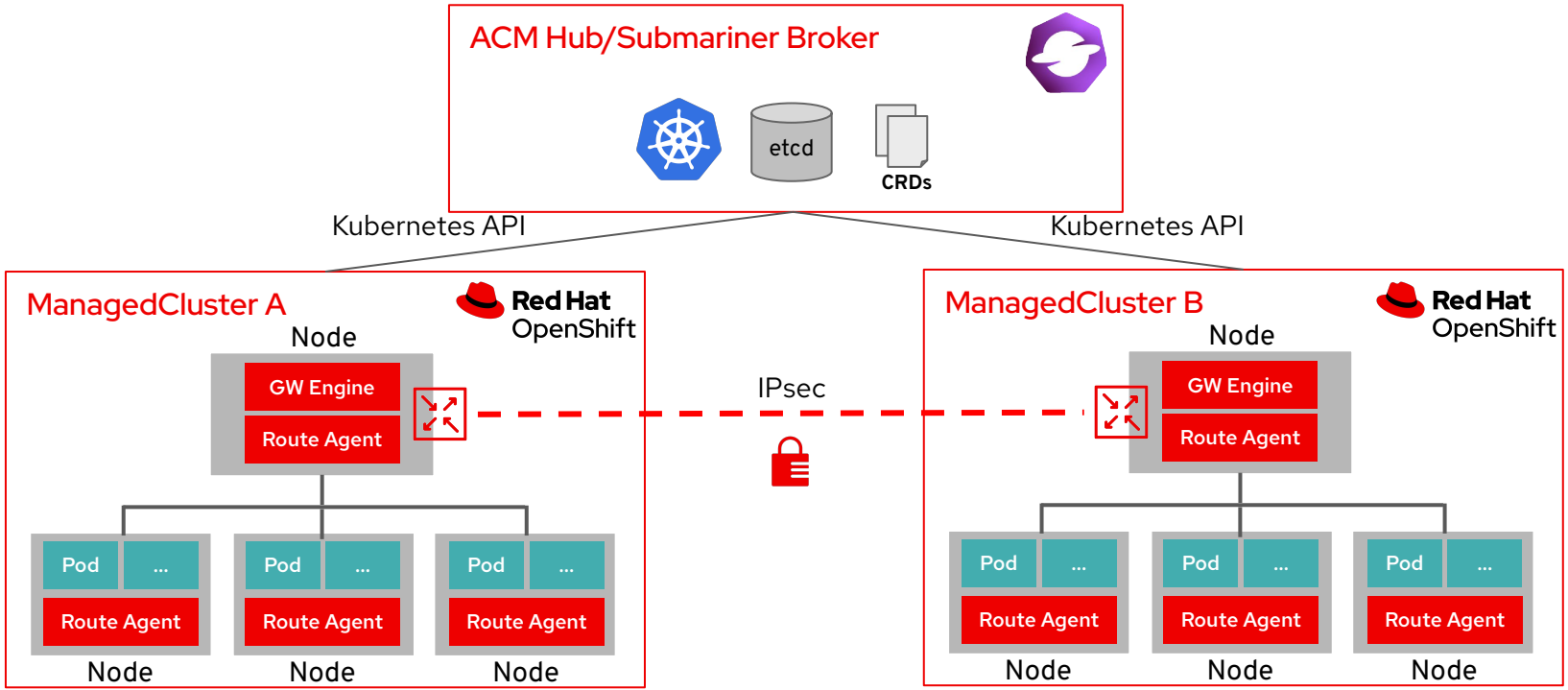
- Originator of the Submariner project, but Red Hat has much stronger community contribution and leadership
- Lacking enterprise support for Submariner



Architecture Overview

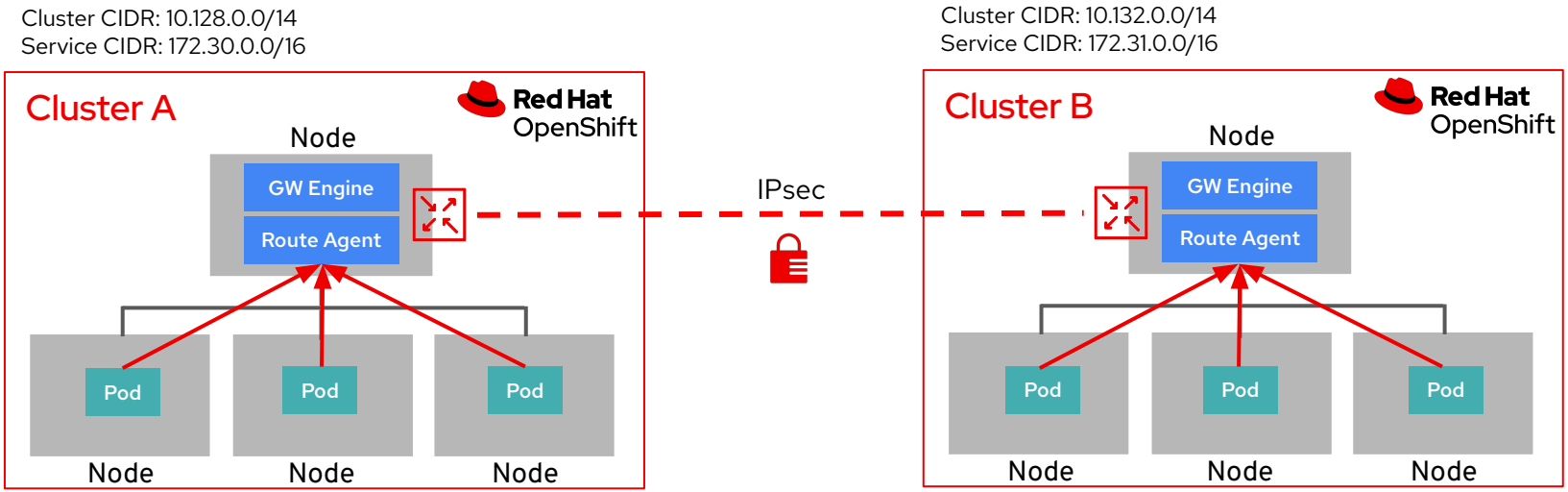
Red Hat

Architecture



Red Hat

Network connectivity



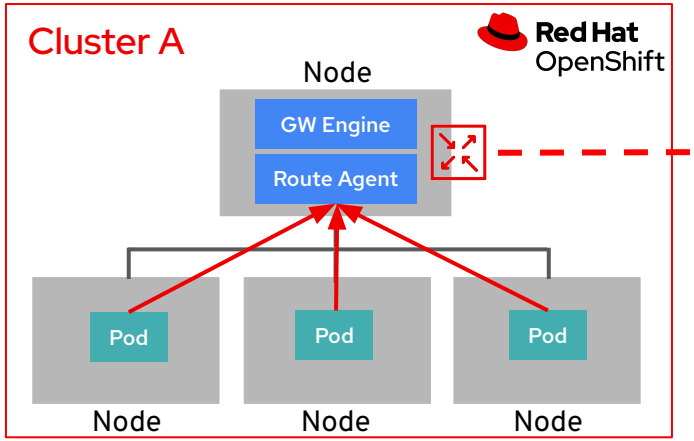
- No impact on intra-cluster traffic (handled by local network plugin)
- Traffic destined to remote clusters is tunneled to a gateway node; source IP is preserved
- Cross-cluster traffic is encrypted with IPsec by default

Red Hat

Network connectivity

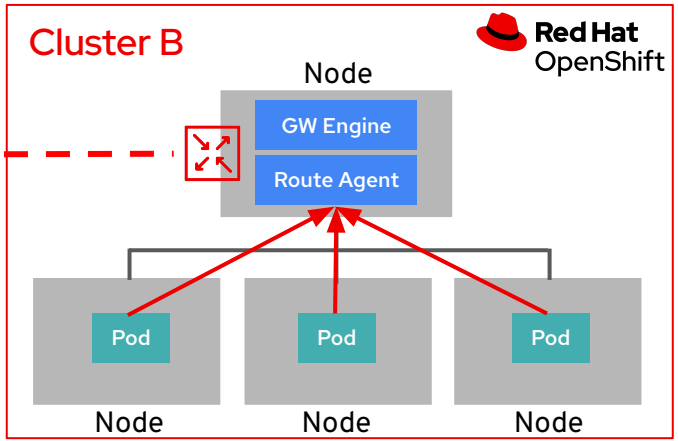
Globalnet

Cluster CIDR: 10.128.0.0/14
Service CIDR: 172.30.0.0/16
Global CIDR: 242.0.0.0/16



IPsec

Cluster CIDR: 10.128.0.0/14
Service CIDR: 172.30.0.0/16
Global CIDR: 242.1.0.0/16



- See next slides for details

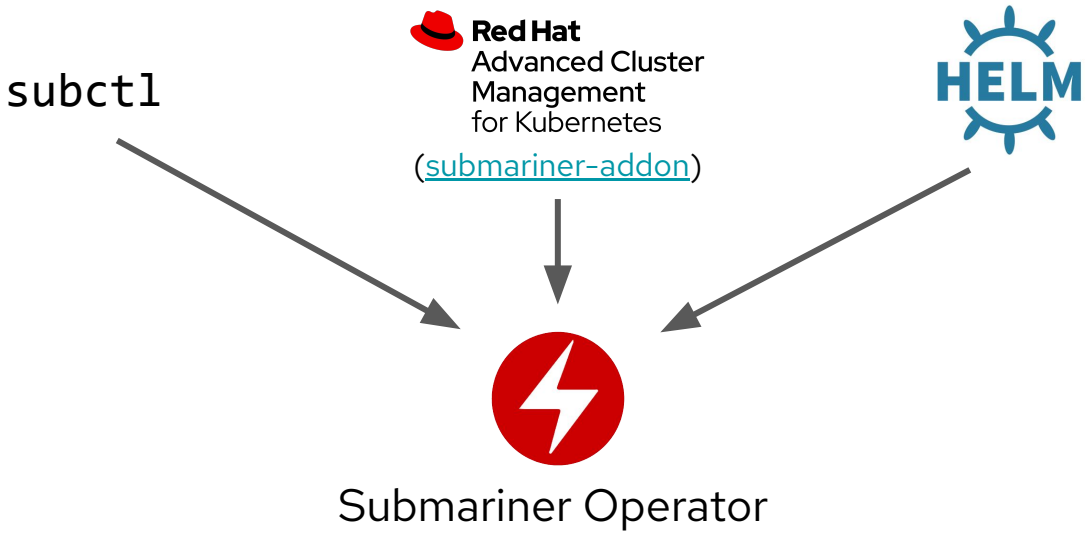
Red Hat

Globalnet

- Each cluster is given a unique subnet from a `GlobalCIDR` range (default 242.0.0.0/8)
- Cluster-scoped global egress IPs
  - Every cluster is assigned a configurable number of global IPs (default 8), represented by a `ClusterGlobalEgressIP` resource, which are used as egress IPs for cross-cluster communication
- Namespace-scoped global egress IPs
  - A user can assign a configurable number of global IPs per namespace by creating a `GlobalEgressIP` resource. These IPs are used as egress IPs for all or selected pods in the namespace and take precedence over the cluster-level global IPs
- Exported ClusterIP services are automatically allocated a `GlobalIngressIP` from the `GlobalCIDR`. For headless services, each backing pod is allocated a global IP that is used for both ingress and egress
- All address translations occur on the active gateway node of the cluster

Red Hat

Deployment and management



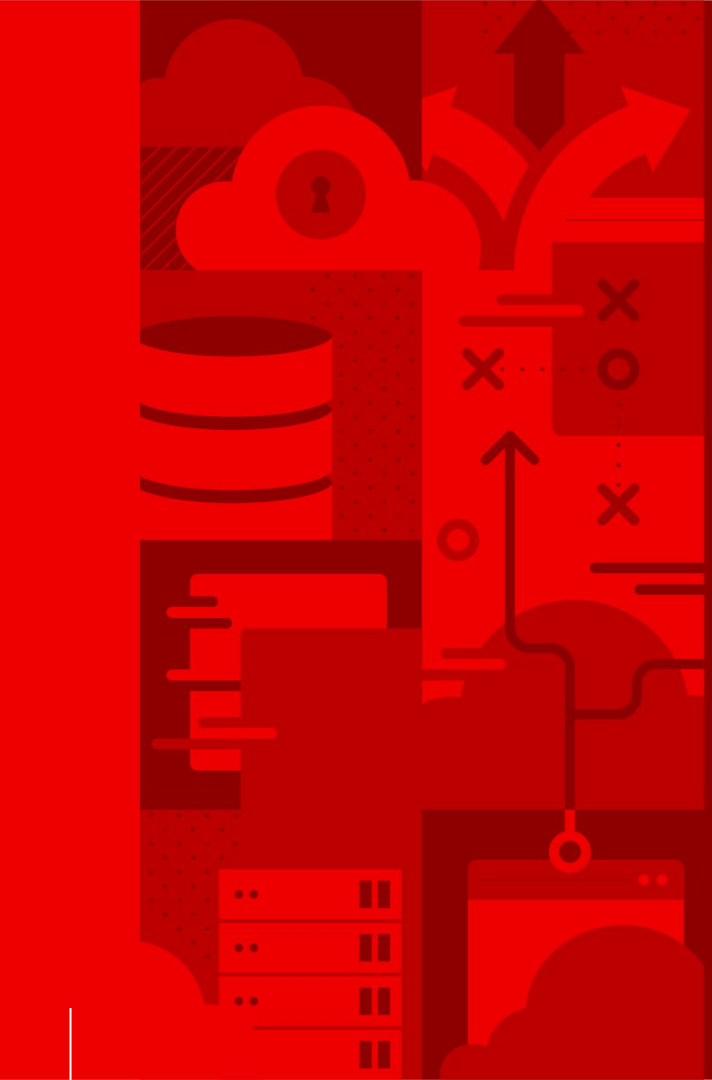
`subctl`

Submariner's CLI utility

- Can be used to easily deploy Submariner, but offers much more than that...
  - `subctl show` - reports various status information, including health of connections with other clusters
  - `subctl export` - creates a `ServiceExport` resource for a given service/namespace
  - `subctl benchmark` - runs a throughput/latency test between two specified clusters or within a single cluster
  - `subctl diagnose` - runs automated checks to help diagnose common issues in a Submariner deployment
  - `subctl verify` - verifies a Submariner deployment between two clusters is functioning properly
  - `subctl gather` - collects logs and other information to aid in troubleshooting

https://submariner.io/operations/deployment/subctl/

Red Hat



Service Discovery

Red Hat

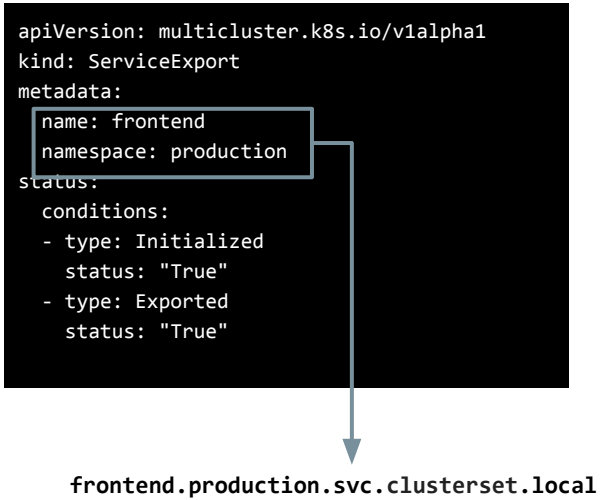
Multicluster service discovery and consumption

Terminology and workflow

- **`ClusterSet`** - a group of clusters with a high degree of mutual trust that share services. Namespaces present in multiple clusters are considered to be the same across the set

- **`ServiceExport`** (CRD) - used to specify which services should be exposed across all clusters (services aren't shared automatically)

- **`ServiceImport`** (CRD) - in-cluster representation of a multi-cluster service in each cluster. Also sets up DNS for the service



KEP-1645 Kubernetes Multi-Cluster Services API

Multicluster service discovery and consumption

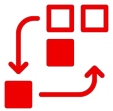
Terminology and workflow

Infra Admin

- **`ClusterSet`** - a group of clusters with a high degree of mutual trust that share services. Namespaces present in multiple clusters are considered to be the same across the set

App Dev

- **`ServiceExport`** (CRD) - used to specify which services should be exposed across all clusters (services aren't shared automatically)



- **`ServiceImport`** (CRD) - in-cluster representation of a multi-cluster service in each cluster. Also sets up DNS for the service

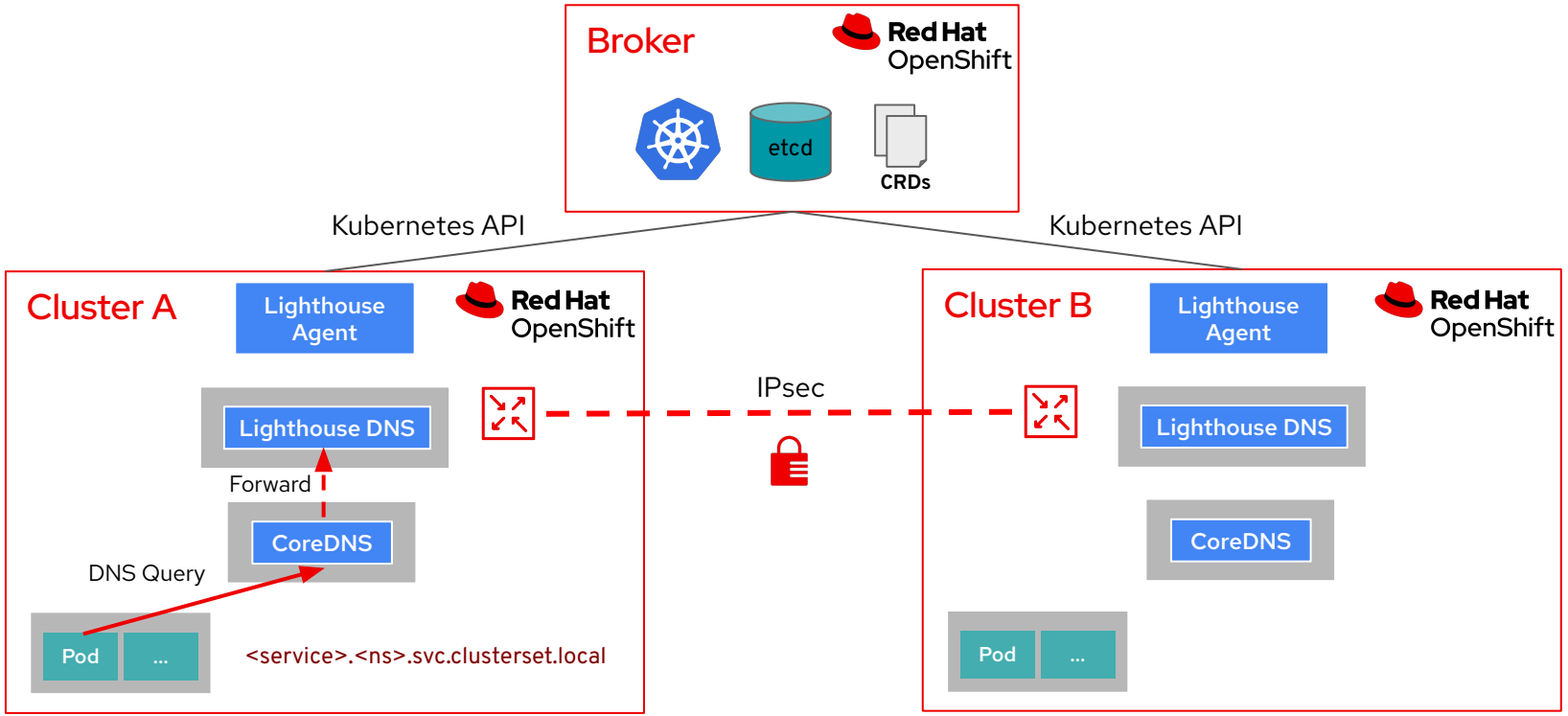
```
apiVersion: multicluster.k8s.io/v1alpha1
kind: ServiceExport
metadata:
  name: frontend
  namespace: production
status:
  conditions:
  - type: Initialized
    status: "True"
  - type: Exported
    status: "True"
```

`frontend.production.svc.clusterset.local`

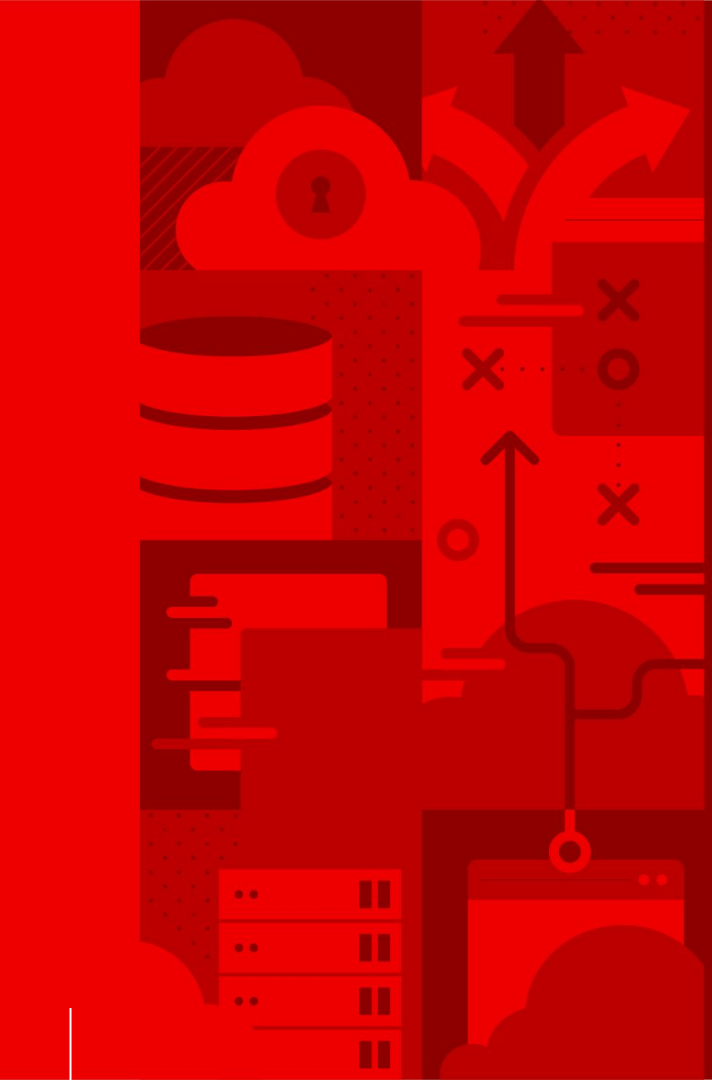
KEP-1645 Kubernetes Multi-Cluster Services API

Red Hat

Service Discovery



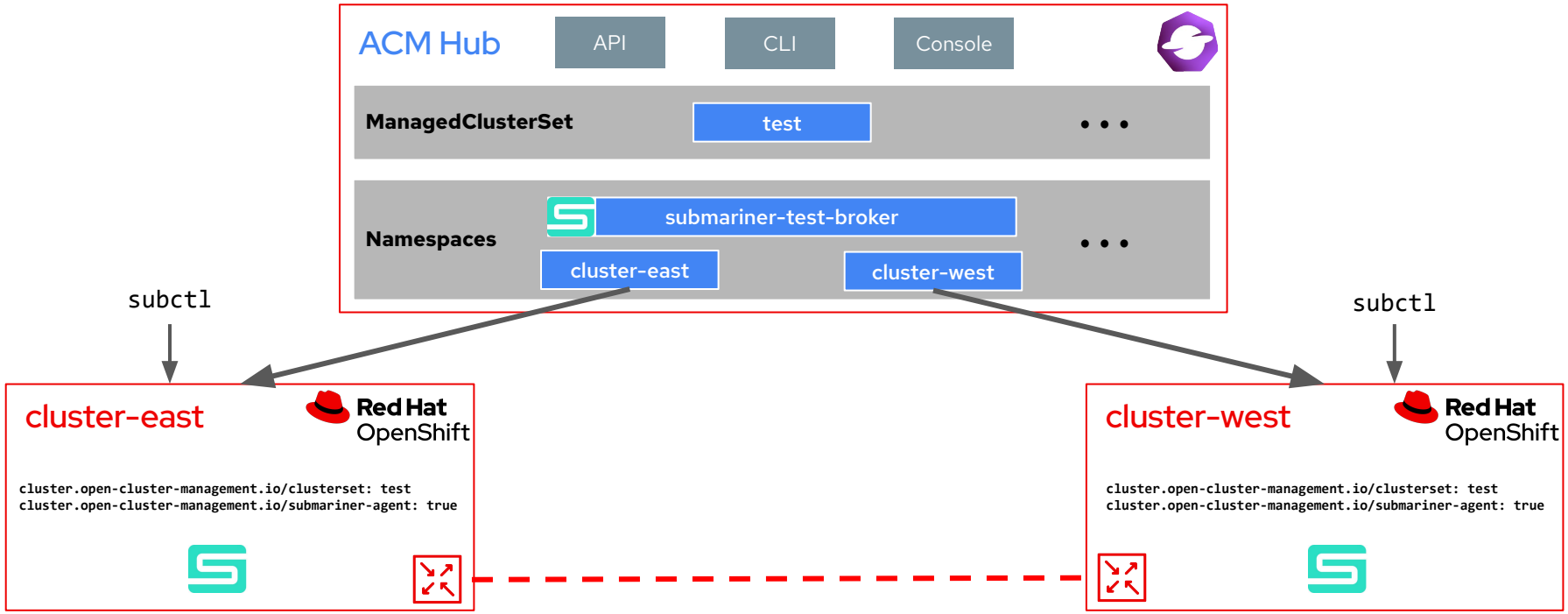
Red Hat



ACM Integration

Red Hat

Integration with ACM



An instance of the Submariner Broker is created on the Hub for each ClusterSet

Red Hat

Integration with ACM

Main APIs

- ManagedCluster
- ManagedClusterSet
- ManifestWork
- SubmarinerConfig

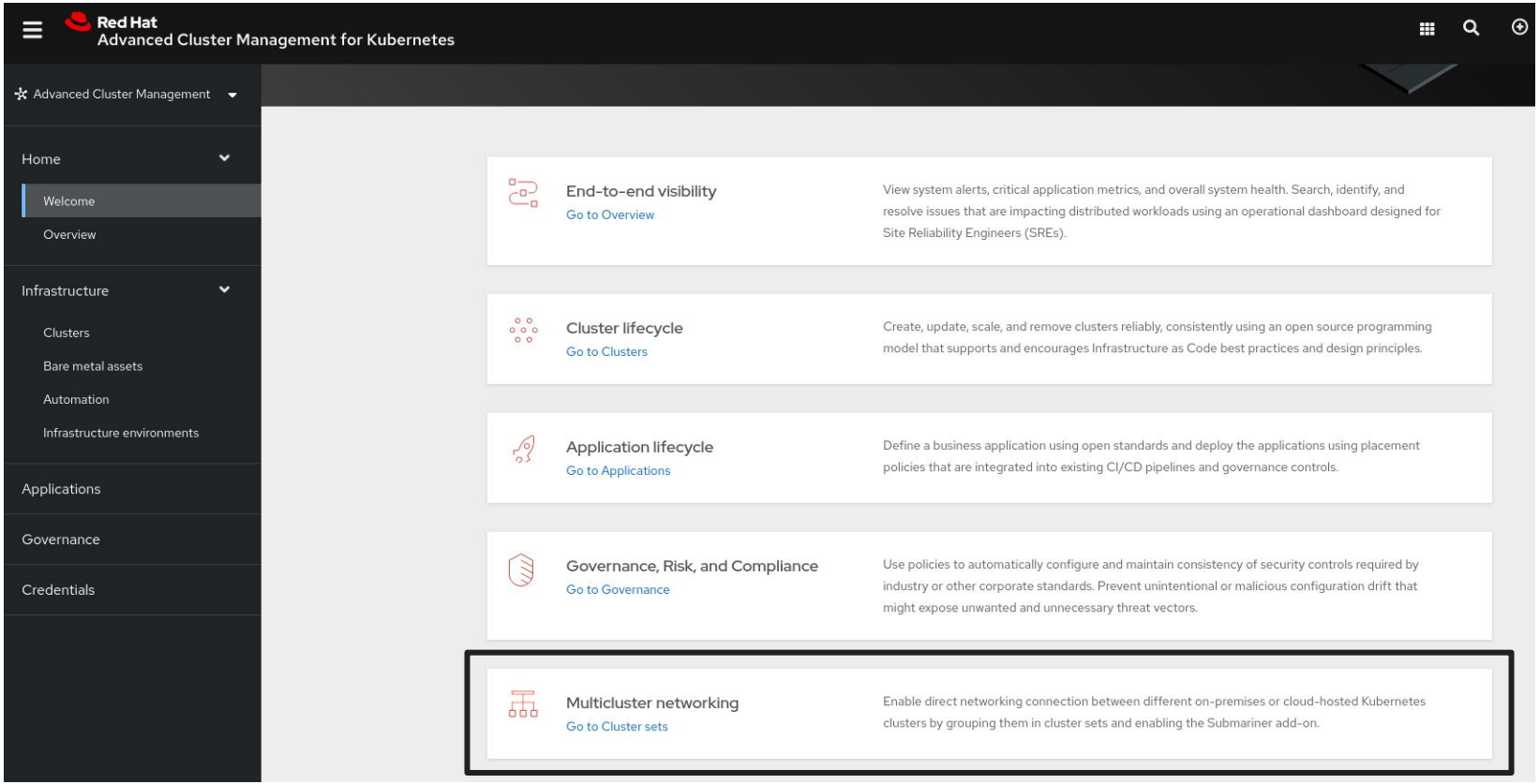
https://github.com/open-cluster-management-io/api
https://github.com/open-cluster-management/submariner-addon/blob/main/docs/submarinerConfig.md

Red Hat

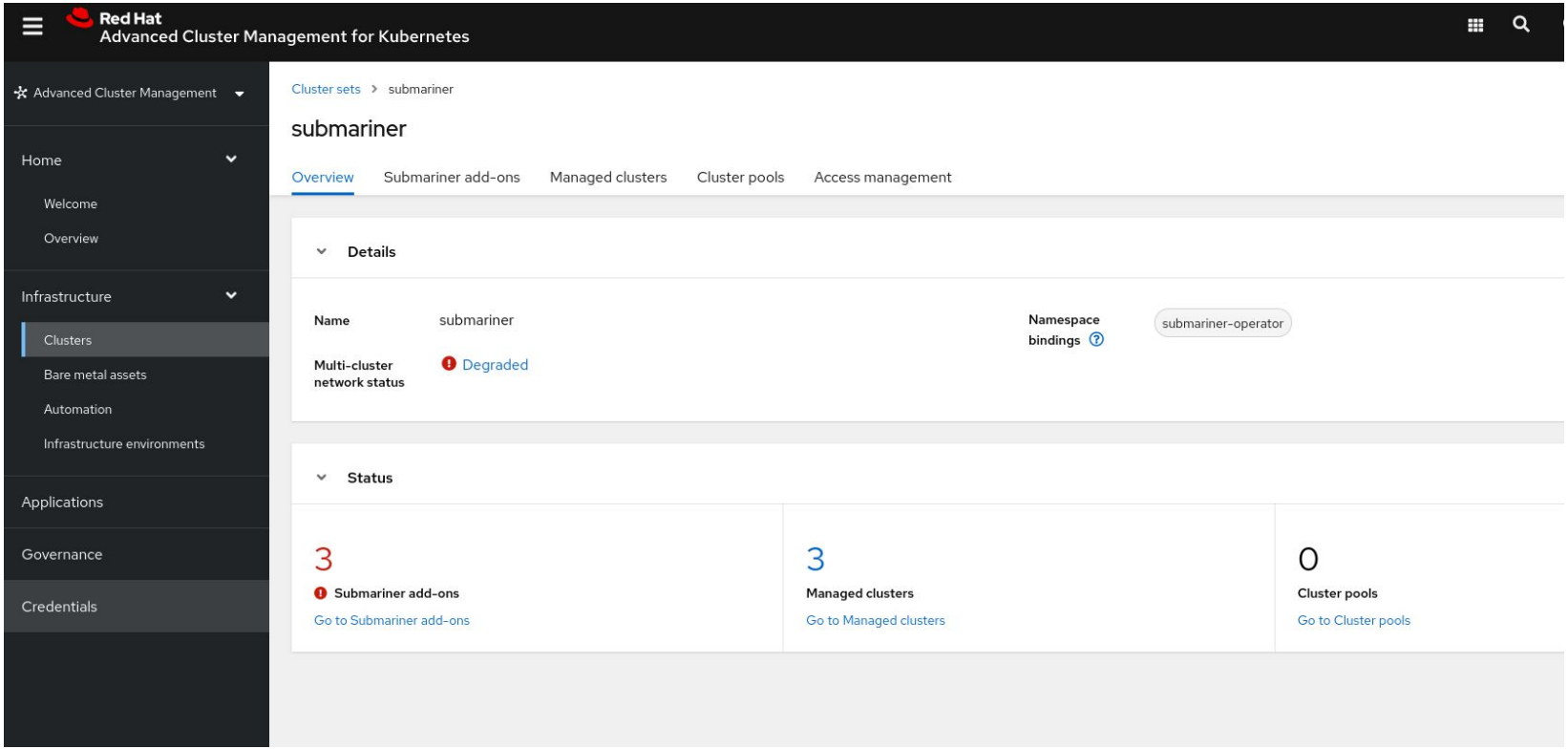
Deployment tips

- Globalnet (overlapping CIDRs) support was introduced in ACM 2.5
  - Prior to 2.5, interconnected clusters must have unique (non-overlapping) CIDRs
- Support for air-gapped (disconnected) environments was introduced in ACM 2.7
- ACM takes care of Submariner deployment and configuration
  - An instance of the Submariner Broker is created on the Hub for each ManagedClusterSet with the Submariner Add-on
- `subctl` is very useful
  - For e.g, for performance benchmarking or advanced troubleshooting
- Consider node types appropriate for gateways
  - Extra node(s) are required to be used as dedicated gateways
  - IPsec is CPU-bound on a single core. For e.g, on AWS, *c5d.large* would provide better performance comparing to *m5n.large*
  - Not all node types are available in all regions
- IPsec and NAT Traversal (NAT-T)

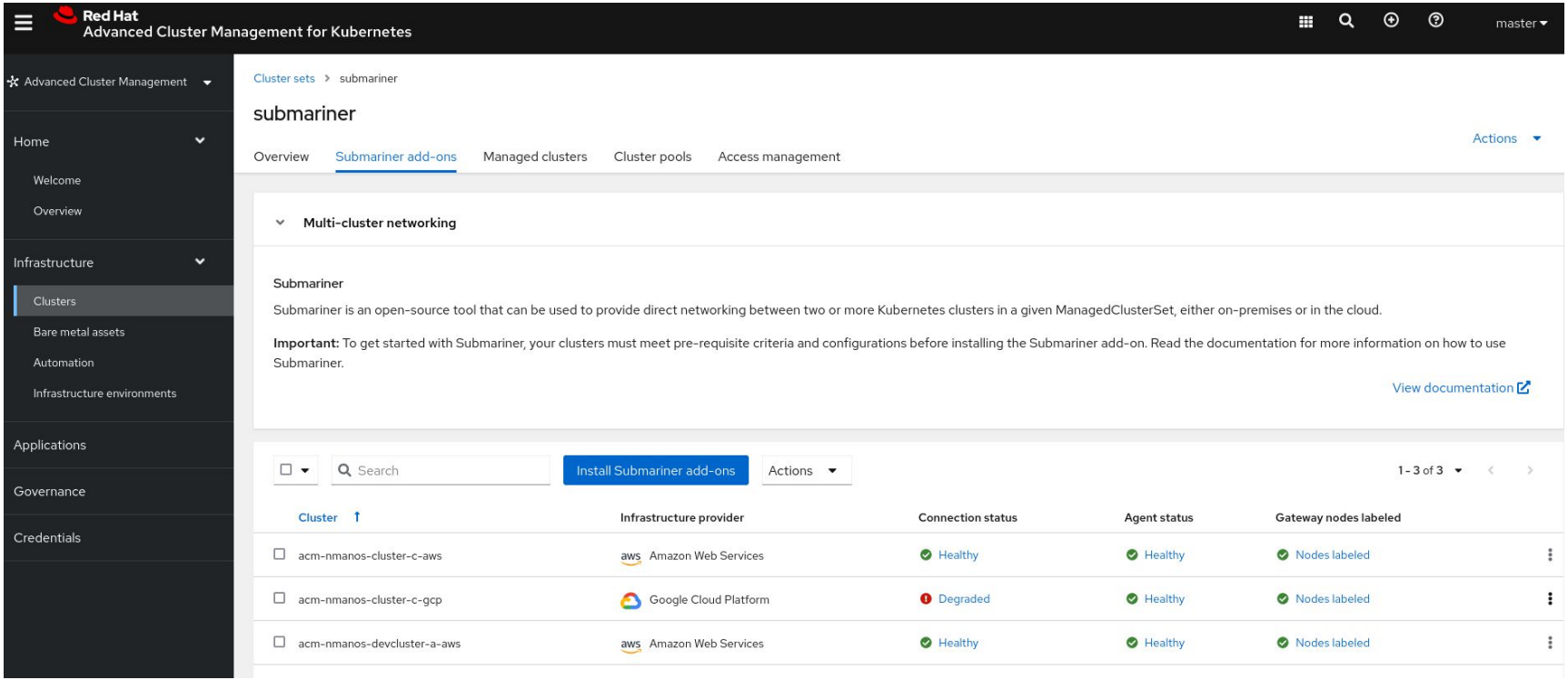
Red Hat


Red Hat
Advanced Cluster Management for Kubernetes
Advanced Cluster Management
Home
Welcome
Overview
Infrastructure
Clusters
Bare metal assets
Automation
Infrastructure environments
Applications
Governance
Credentials
End-to-end visibility
Go to Overview
View system alerts, critical application metrics, and overall system health. Search, identify, and resolve issues that are impacting distributed workloads using an operational dashboard designed for Site Reliability Engineers (SREs).
Cluster lifecycle
Go to Clusters
Create, update, scale, and remove clusters reliably, consistently using an open source programming model that supports and encourages Infrastructure as Code best practices and design principles.
Application lifecycle
Go to Applications
Define a business application using open standards and deploy the applications using placement policies that are integrated into existing CI/CD pipelines and governance controls.
Governance, Risk, and Compliance
Go to Governance
Use policies to automatically configure and maintain consistency of security controls required by industry or other corporate standards. Prevent unintentional or malicious configuration drift that might expose unwanted and unnecessary threat vectors.
Multicluster networking
Go to Cluster sets
Enable direct networking connection between different on-premises or cloud-hosted Kubernetes clusters by grouping them in cluster sets and enabling the Submariner add-on.

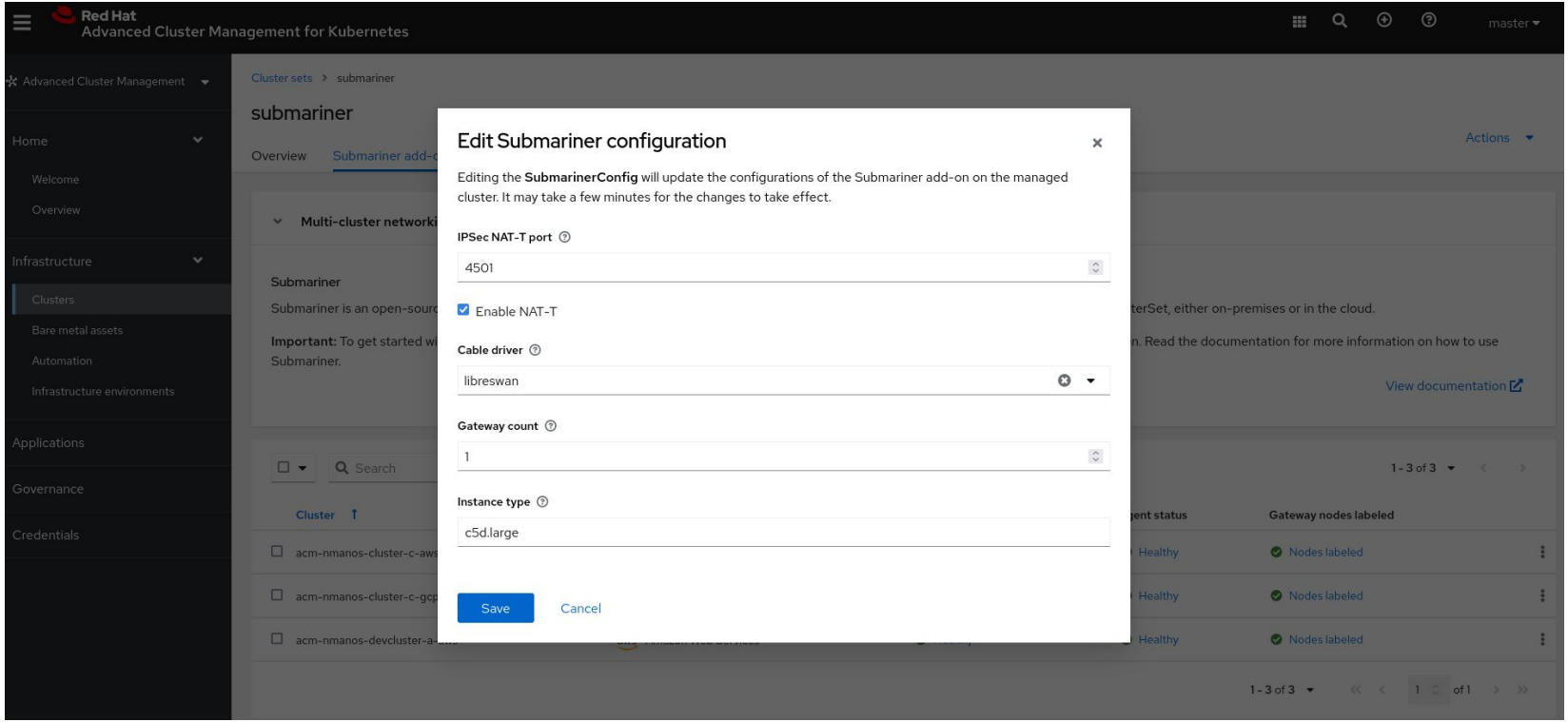
Red Hat


Red Hat
Advanced Cluster Management for Kubernetes
Advanced Cluster Management
Home
Welcome
Overview
Infrastructure
Clusters
Bare metal assets
Automation
Infrastructure environments
Applications
Governance
Credentials
Cluster sets > submariner
submariner
Overview
Submariner add-ons
Managed clusters
Cluster pools
Access management
Details
Name
submariner
Multi-cluster network status
Degraded
Namespace bindings
submariner-operator
Status
3
Submariner add-ons
Go to Submariner add-ons
3
Managed clusters
Go to Managed clusters
0
Cluster pools
Go to Cluster pools

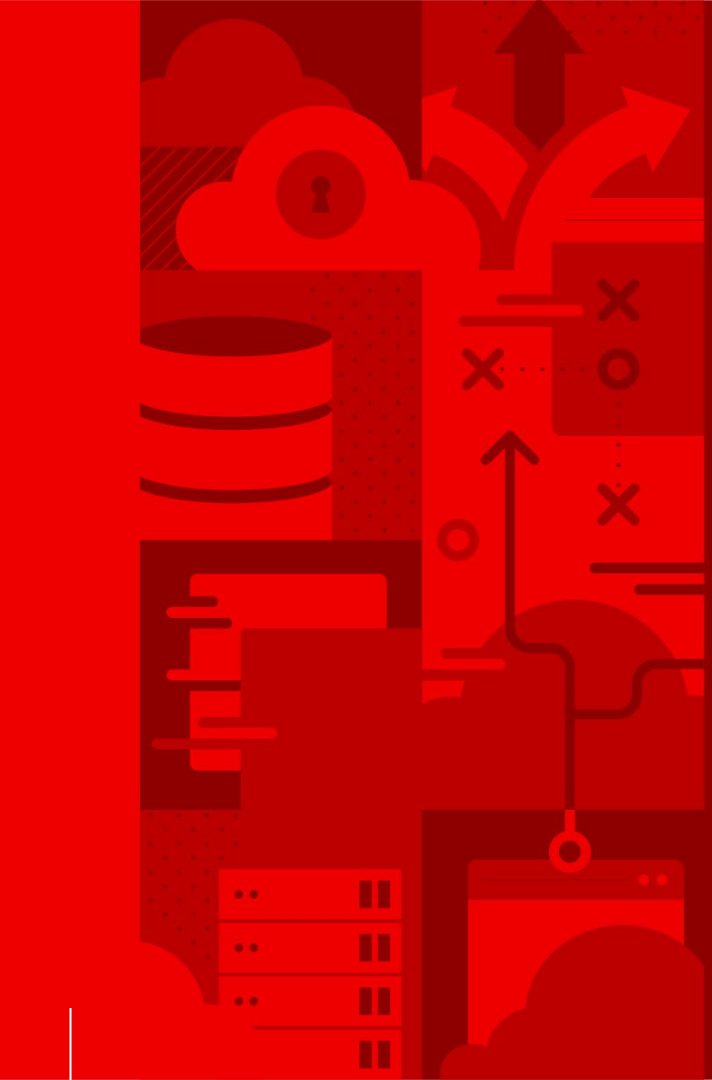
Red Hat


Red Hat
Advanced Cluster Management for Kubernetes
master
Advanced Cluster Management
Home
Welcome
Overview
Infrastructure
Clusters
Bare metal assets
Automation
Infrastructure environments
Applications
Governance
Credentials
Cluster sets
submariner
submariner
Actions
Overview
Submariner add-ons
Managed clusters
Cluster pools
Access management
Multi-cluster networking
Submariner
Submariner is an open-source tool that can be used to provide direct networking between two or more Kubernetes clusters in a given ManagedClusterSet, either on-premises or in the cloud.
Important: To get started with Submariner, your clusters must meet pre-requisite criteria and configurations before installing the Submariner add-on. Read the documentation for more information on how to use Submariner.
View documentation
Search
Install Submariner add-ons
Actions
1 - 3 of 3
Cluster
Infrastructure provider
Connection status
Agent status
Gateway nodes labeled
acm-nmanos-cluster-c-aws
Amazon Web Services
Healthy
Healthy
Nodes labeled
acm-nmanos-cluster-c-gcp
Google Cloud Platform
Degraded
Healthy
Nodes labeled
acm-nmanos-devcluster-a-aws
Amazon Web Services
Healthy
Healthy
Nodes labeled

Red Hat


Red Hat
Advanced Cluster Management for Kubernetes
master
Advanced Cluster Management
Home
Welcome
Overview
Infrastructure
Clusters
Bare metal assets
Automation
Infrastructure environments
Applications
Governance
Credentials
Cluster sets > submariner
submariner
Overview
Submariner add-
Actions
Multi-cluster networki
Submariner
Submariner is an open-sourc
terSet, either on-premises or in the cloud.
Important: To get started wi
n. Read the documentation for more information on how to use
Submariner.
View documentation
Search
1 - 3 of 3
Cluster
ent status
Gateway nodes labeled
acm-nmanos-cluster-c-aws
Healthy
Nodes labeled
acm-nmanos-cluster-c-gcp
Healthy
Nodes labeled
acm-nmanos-devcluster-a-
Healthy
Nodes labeled
1 - 3 of 3
1
of 1
Edit Submariner configuration
Editing the SubmarinerConfig will update the configurations of the Submariner add-on on the managed cluster. It may take a few minutes for the changes to take effect.
IPSec NAT-T port
4501
Enable NAT-T
Cable driver
libreswan
Gateway count
1
Instance type
c5d.large
Save
Cancel

Red Hat



Resources

Red Hat

Give it a try

- Community project
  - Website: https://submariner.io
    - https://submariner.io/getting-started/quickstart/
    - https://submariner.io/operations/usage/
  - GitHub: https://github.com/submariner-io
  - YouTube: https://tinyurl.com/submariner-youtube
  - Slack (Kubernetes space): #submariner

- Latest Red Hat product docs
  - ACM 2.9 Submariner product documentation
  - ACM 2.9 Release Notes
  - ACM 2.9 Support Matrix
  - OpenShift Data Foundation Disaster Recovery for OpenShift Workloads

Red Hat

Further reading and other resources

- https://submariner.io/other-resources/
- Red Hat blogs:
  - https://cloud.redhat.com/blog/geographically-distributed-stateful-workloads-part-one-cluster-preparation
  - https://cloud.redhat.com/blog/geographically-distributed-stateful-workloads-part-two-cockroachdb
  - https://cloud.redhat.com/blog/geographically-distributed-stateful-workloads-part-3-keycloak
  - https://cloud.redhat.com/blog/geographically-distributed-stateful-workloads-part-four-kafka
  - https://cloud.redhat.com/blog/geographically-distributed-stateful-workloads-part-five-yugabytedb

Red Hat

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

linkedin.com/company/red-hat

youtube.com/user/RedHatVideos

facebook.com/redhatinc

twitter.com/RedHat

Red Hat