Hybrid Cloud with LinuxONE

Matt Mondics
matt.mondics@ibm.com
Client Technical Specialist – Hybrid Cloud on IBM zSystems

Contents

What is Hybrid Cloud really?

What is Hybrid Cloud really?

“Hybrid cloud integrates **public cloud** services, **private cloud** services **and on-premises infrastructure** and provides orchestration, management and application portability across all three. The result is a single, unified and flexible distributed computing environment where an organization can run and scale its **traditional or cloud-native workloads** on the most appropriate computing model.” https://www.ibm.com/topics/hybrid-cloud

Role-Based Access Control (cont.)

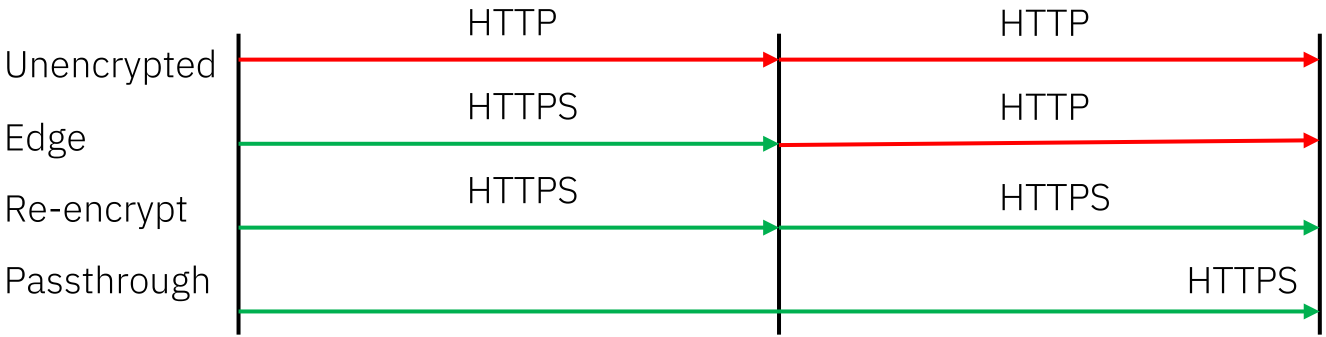
- Once an identity provider is configured, Red Hat recommends that you remove the default kubeadmin user.
- Kubeadmin is a special user that does not rely on an identity provider, but rather a token that is associated with a secret in the OpenShift cluster.
- If you remove kubeadmin before configuring an identity provider with at least one new cluster administrator, your OCP cluster will be **completely inaccessible**.
- If you remove kubeadmin and then your identity provider is then removed or goes down, your OCP cluster will be **completely inaccessible**.
- Therefore, it might be good idea to keep a backup cluster administrator that relies on a secret rather than the identity provider. This is, however, a less secure option.

Role-Based Access Control (cont.)

- Even though users now have access to the OpenShift cluster after configuring authentication, **by default they have access to nothing**.
- You must assign permissions to users or groups of users to allow them to do anything in the cluster.
- Permissions are granted by creating a RoleBinding between a specific role and a user or group.
- There are many default roles that come with OpenShift out of the box (cluster-administrator)
- Custom roles are supported with extremely granular permissions. All actions in OpenShift consist of a verb and a resource, and either verbs, resource types, or specific projects can be specified.
  - verbs: get, list, watch, create, update, delete, etc
  - resources: pod, service, deployment, secret, configmap, etc.
- Permissions are either namespace-scoped, or cluster-wide. Namespace-scoped provide access to resources in a specific project. Cluster-wide provide access to resources at the cluster scope, such as platform monitoring, projects managing OpenShift and Kubernetes services, etc.

Encrypting OpenShift Routes

- Routes are the OpenShift object responsible for getting external traffic into the cluster and its applications
- Various types of route encryption are supported:
  - Edge: traffic encrypted from the external user to the router
  - Re-encryption: traffic encrypted from external user to the router, and then again from router to the application
  - Passthrough: traffic encrypted once from the external user through the router and to the application



Red Hat Advanced Cluster Security

Vulnerability Management

Protect yourself against known vulnerabilities in images and running containers

Network Segmentation

Apply and manage network isolation and access controls for each application

Security Configuration Management

Compliance

Meet contractual and regulatory requirements and easily audit against them

Risk Profiling

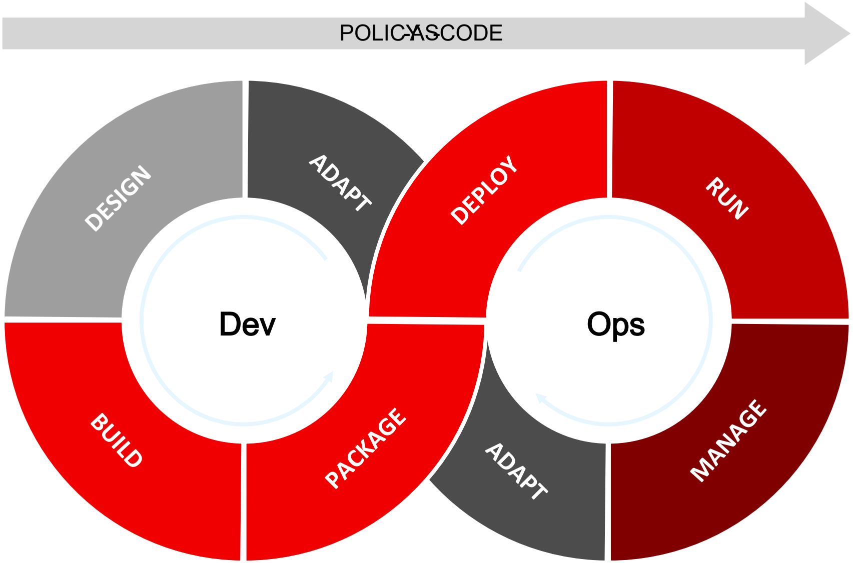
Gain context to prioritize security issues throughout OpenShift and Kubernetes clusters

Detection and Response

Carry out incident response to address active threats in your environment

Red Hat Advanced Cluster Security

- ➢ Image scanning
- ➢ Host scanning
- ➢ Serverless scanning
- ➢ Configuration scanning
- ➢ Compliance checks, auditing, reporting, remediation
- ➢ CI/CD integration and automation
- ➢ Artifact attestation



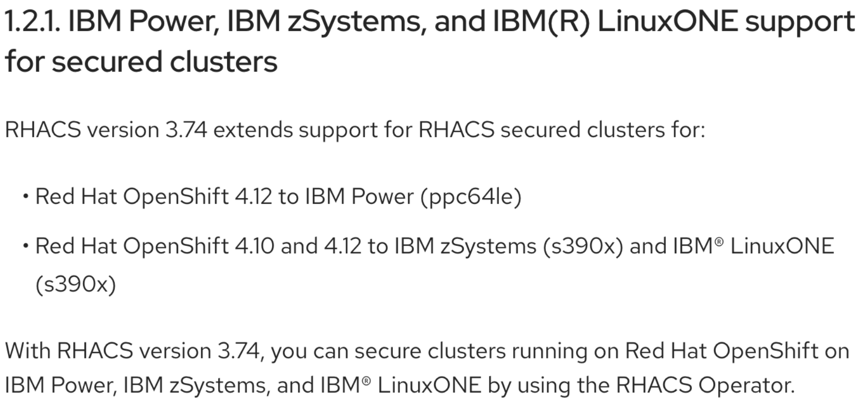
- ➢ Runtime threat detection
  - ✓ Process allowlisting
  - ✓ Anomaly detection
  - ✓ Policy-based detection
- ➢ Runtime vulnerability management
- ➢ Incident response
- ➢ Integrations
  - ✓ SIEM
  - ✓ Registries, CI/CD, runtimes, notification tools
- ➢ Feedback loop

VISIBILITY (images, deployments, network flows, processes, secrets use)

CONTAINERS AND K8S (on-premises, cloud/hybrid, edge)

Red Hat Advanced Cluster Security

RHACS version 3.74 now supports managing-to OpenShift on IBM zSystems and LinuxONE

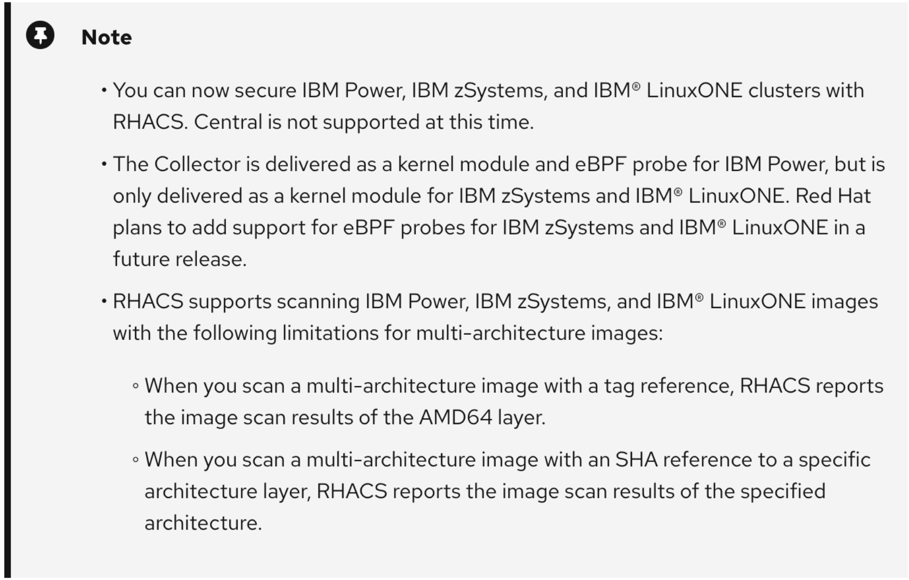


1.2.1. IBM Power, IBM zSystems, and IBM(R) LinuxONE support for secured clusters

RHACS version 3.74 extends support for RHACS secured clusters for:

- Red Hat OpenShift 4.12 to IBM Power (ppc64le)
- Red Hat OpenShift 4.10 and 4.12 to IBM zSystems (s390x) and IBM® LinuxONE (s390x)

With RHACS version 3.74, you can secure clusters running on Red Hat OpenShift on IBM Power, IBM zSystems, and IBM® LinuxONE by using the RHACS Operator.



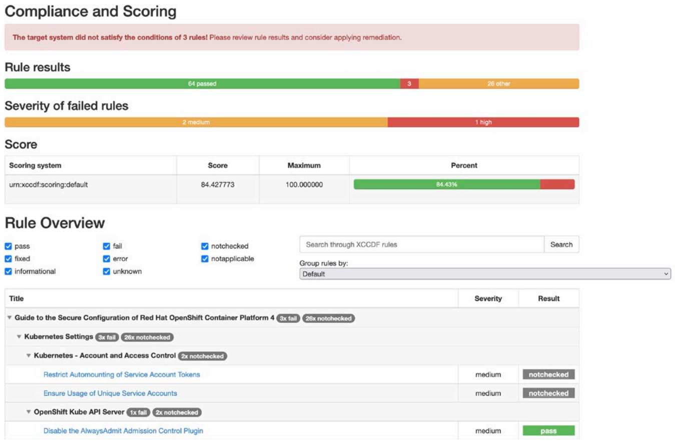
Note

- You can now secure IBM Power, IBM zSystems, and IBM® LinuxONE clusters with RHACS. Central is not supported at this time.
- The Collector is delivered as a kernel module and eBPF probe for IBM Power, but is only delivered as a kernel module for IBM zSystems and IBM® LinuxONE. Red Hat plans to add support for eBPF probes for IBM zSystems and IBM® LinuxONE in a future release.
- RHACS supports scanning IBM Power, IBM zSystems, and IBM® LinuxONE images with the following limitations for multi-architecture images:
  - When you scan a multi-architecture image with a tag reference, RHACS reports the image scan results of the AMD64 layer.
  - When you scan a multi-architecture image with an SHA reference to a specific architecture layer, RHACS reports the image scan results of the specified architecture.

OpenShift Compliance Operator

- The Compliance Operator lets OpenShift Container Platform administrators describe the required compliance state of a cluster and provides them with an overview of gaps and ways to remediate them.
- As of OpenShift version 4.10, the Compliance Operator is supported on IBM Z and LinuxONE clusters.
- Scan against compliance profiles such as NIST, CIS, etc. and generate a report of cluster compliance issues.
- Use the Compliance Operator to automatically remediate failed results or manually perform the remediation steps.

```
NAME                                                                        STATUS         SEVERITY
ocp4-cis-accounts-restrict-service-account-tokens                           FAIL           medium
ocp4-cis-accounts-unique-service-account                                    FAIL           medium
ocp4-cis-api-server-admission-control-plugin-alwaysadmit                    PASS           medium
ocp4-cis-api-server-admission-control-plugin-alwayspullimages               PASS           high
ocp4-cis-api-server-admission-control-plugin-namespacelifecycle             PASS           medium
ocp4-cis-api-server-admission-control-plugin-noderestriction                PASS           medium
ocp4-cis-api-server-admission-control-plugin-scc                            PASS           medium
ocp4-cis-api-server-admission-control-plugin-securitycontextdeny            PASS           medium
ocp4-cis-api-server-admission-control-plugin-serviceaccount                 PASS           medium
ocp4-cis-api-server-anonymous-auth                                          PASS           medium
ocp4-cis-api-server-api-priority-gate-enabled                               PASS           medium
ocp4-cis-api-server-audit-log-maxbackup                                     PASS           low
ocp4-cis-api-server-audit-log-maxsize                                       PASS           medium
ocp4-cis-api-server-audit-log-path                                          PASS           high
ocp4-cis-api-server-auth-mode-no-aa                                         PASS           medium
```



Compliance and Scoring

The target system did not satisfy the conditions of 3 rules! Please review rule results and consider applying remediation.

Rule results

Severity of failed rules

Score

Scoring system	Score	Maximum	Percent
urn:xccdf:scoring:default	84.427773	100.000000	84.43%

Rule Overview

☑ pass ☑ fixed ☑ informational
☑ fail ☑ error ☑ unknown
☑ notchecked ☑ notapplicable

Search through XCCDF rules

Group rules by: Default

Title	Severity	Result
▾ Guide to the Secure Configuration of Red Hat OpenShift Container Platform 4 3x fail 26x notchecked		
▾ Kubernetes Settings 3x fail 26x notchecked		
▾ Kubernetes - Account and Access Control 2x notchecked		
Restrict Automounting of Service Account Tokens	medium	notchecked
Ensure Usage of Unique Service Accounts	medium	notchecked
▾ OpenShift Kube API Server 1x fail 2x notchecked		
Disable the AlwaysAdmit Admission Control Plugin	medium	pass

OpenShift Service Mesh

- **Red Hat OpenShift Service Mesh** (based on Istio) addresses a variety of problems in a microservice architecture by creating a centralized point of control in an application. It adds a transparent layer on existing distributed applications without requiring any changes to the application code.

- OpenShift Service Mesh provides several capabilities, including:

  - **Traffic Management** - Control the flow of traffic and API calls between services, make calls more reliable, and make the network more robust in the face of adverse conditions.

  - **Service Identity and Security** - Provide services in the mesh with a verifiable identity and provide the ability to protect service traffic as it flows over networks of varying degrees of trustworthiness.

  - **Policy Enforcement** - Apply organizational policy to the interaction between services, ensure access policies are enforced and resources are fairly distributed among consumers. Policy changes are made by configuring the mesh, not by changing application code.

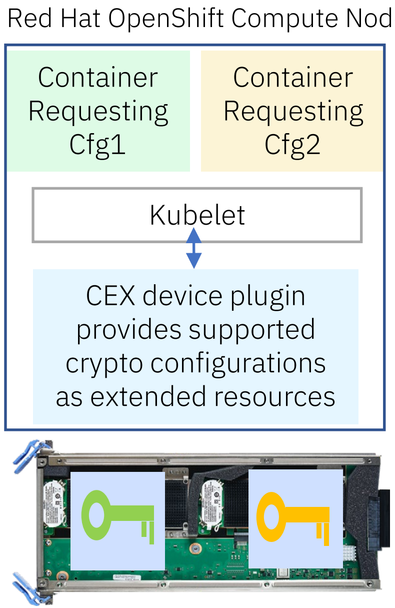
  - **Telemetry** - Gain understanding of the dependencies between services and the nature and flow of traffic between them, providing the ability to quickly identify issues.

  - **Federation** - Share services and workloads between separate meshes managed in distinct administrative domains (share traffic across service meshes in separate OCP clusters).

OpenShift Service Mesh

- Putting all the other functionality aside, OpenShift Service Mesh is an invaluable tool for security between complex microservice applications.
- Enable strict mTLS between microservices in the service mesh
  - Configure sidecars for incoming or outgoing connections with the microservices in the service mesh
  - Enforce minimum/maximum TLS versions used by the service mesh components
- Configure role-based access control
  - Intra-mesh communication (i.e. deny requests coming from any source other than a specific project)
  - Allow or deny access to the service mesh’s ingress gateway
  - Restrict access with JSON Web Tokens
- Configure specific cipher suites and ECDH curves to use within the service mesh

IBM Crypto Express Support for Containers on Red Hat OpenShift

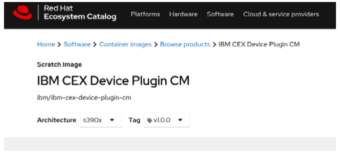

Red Hat OpenShift Compute Node
Container Requesting Cfg1
Container Requesting Cfg2
Kubelet
CEX device plugin provides supported crypto configurations as extended resources

Kubernetes device plug-in for IBM Crypto Express (CEX) cards

Enables Red Hat OpenShift containers to take advantage of crypto resources on IBM Crypto Express adapters

Generate up to 100,000 certificates per second using protected keys exploiting Crypto Express 8S adapters running application pods on Red Hat OpenShift Container Platform on a single IBM z16 drawer.[6]

Supported as Red Hat certified container


Red Hat Ecosystem Catalog
Platforms
Hardware
Software
Cloud & service providers
Home > Software > Container images > Browse products > IBM CEX Device Plugin CM

Scratch image

IBM CEX Device Plugin CM

ibm/ibm-cex-device-plugin-cm

Architecture s390x ▾ Tag v1.0.0 ▾

[6] see p18

OpenShift Resilience

OpenShift Availability Overview

- In the OpenShift world, availability means that the containerized applications running in the OpenShift cluster remain available regardless of any underlying failures.
- High availability is built into OpenShift at each level of the infrastructure stack, including the:
  - Pod level
  - Node level
  - Cluster level

Pod-level Availability

- In all Kubernetes platforms, terminating pods are an extremely common occurrence. Kubernetes is designed to restart pods if they go down (unless it was explicitly told to stop the pods).

- OpenShift includes many functions to maintain pod availability, including:
  - ReplicaSets and DaemonSets
  - Horizontal Pod Autoscaling
  - Liveliness and readiness probes
  - Pod Disruption Budgets

ReplicaSets and DaemonSets

- ReplicaSets are used to ensure a given number of pods are running at any given time
- DaemonSets are used to ensure that a replica of a pod is running on all nodes at any given time
- Services are used to load balance across the multiple pods running the same application

```
➜  ~ oc get pod -l app=nationalparks -owide
NAME                                READY   STATUS    RESTARTS   AGE      NODE
nationalparks-58f98949db-275br      1/1     Running   0          17m      compute-0.atsocpd1.dmz
nationalparks-58f98949db-s2lhx      1/1     Running   0          17m      compute-1.atsocpd1.dmz
nationalparks-58f98949db-xl5sv      1/1     Running   0          17m      compute-2.atsocpd1.dmz
```

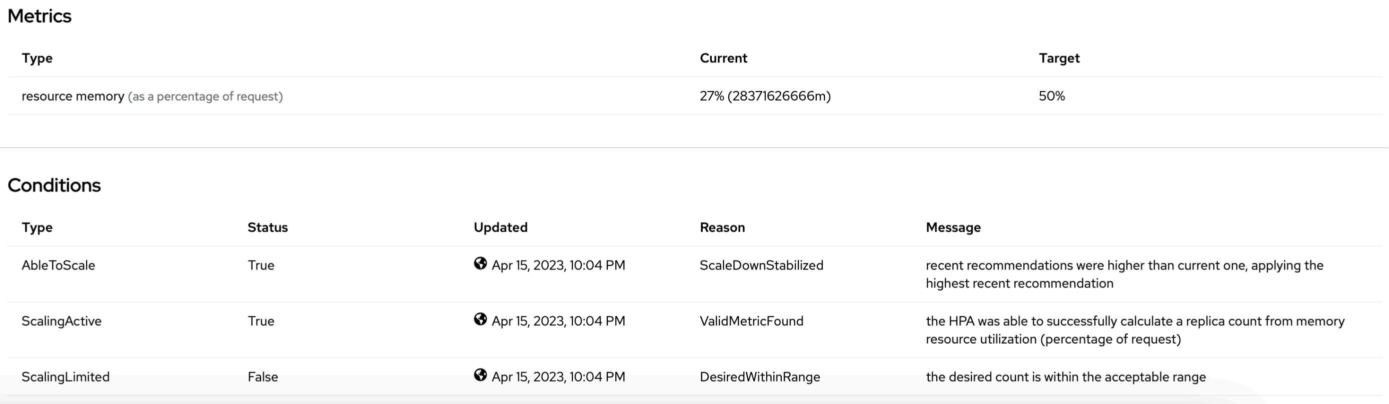
Horizontal Pod Autoscaling

- Specify the minimum and maximum number of pods you want to run, as well as the CPU utilization or memory utilization your pods should target

```
➜  ~ oc get hpa

NAME                REFERENCE                   TARGETS   MINPODS   MAXPODS   REPLICAS   AGE

nationalparks-hpa   Deployment/nationalparks    27%/50%   3         10        3          30s
```



Metrics

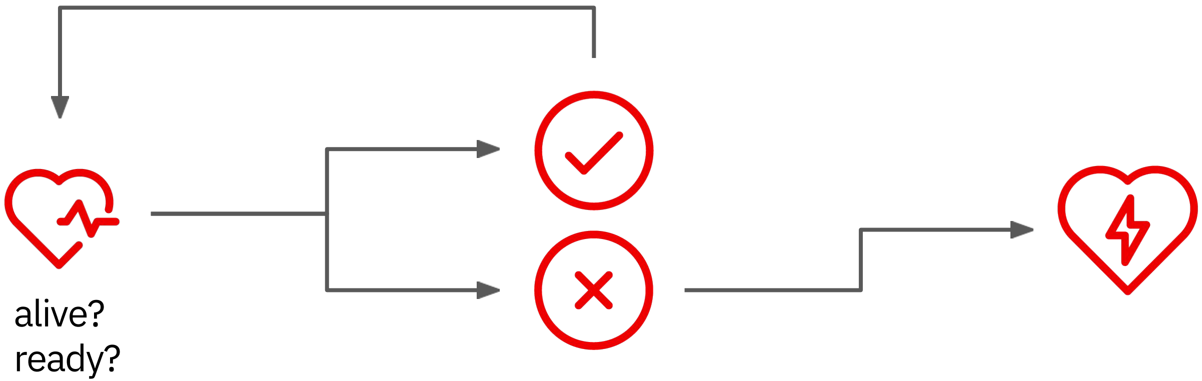
Type	Current	Target
resource memory (as a percentage of request)	27% (28371626666m)	50%

Conditions

Type	Status	Updated	Reason	Message
AbleToScale	True	Apr 15, 2023, 10:04 PM	ScaleDownStabilized	recent recommendations were higher than current one, applying the highest recent recommendation
ScalingActive	True	Apr 15, 2023, 10:04 PM	ValidMetricFound	the HPA was able to successfully calculate a replica count from memory resource utilization (percentage of request)
ScalingLimited	False	Apr 15, 2023, 10:04 PM	DesiredWithinRange	the desired count is within the acceptable range

Liveliness and readiness probes

- Automatically check if pods are running and if the application they contain is ready
- If not, the pod is restarted



Pod Disruption Budgets

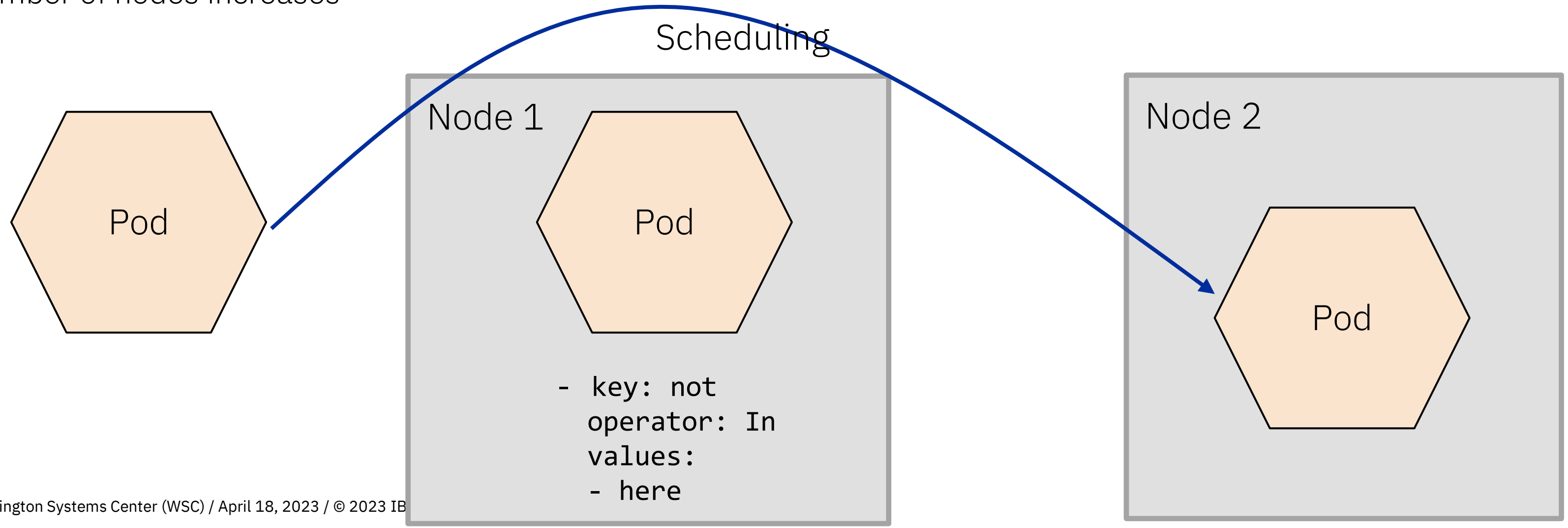
- Specify the minimum number or percentage of replicas that must be up at a time
- Only honored on voluntary evictions, not unplanned outages such as node failures
- For example,
  - A quorum-based application must have >50% of its replicas running at any given time
  - A business-critical application must have enough replicas to meet SLAs
- Useful during cluster upgrades or when performing maintenance on nodes

Node-level Availability

- Anti-affinity rules
- Rescheduling after node
- Descheduler
- Distribute nodes across multiple LPARs or physical servers.

Anti-affinity rules

- Force pods to be scheduled onto separate nodes so that if one node goes down, another replica of the workload pod remains available
- Checks the labels on each pod to ensure placement would not violate anti-affinity rule
- This does incur a performance hit at the time of pod scheduling with performance degrading more significantly as the number of nodes increases



```
- key: not
  operator: In
  values:
  - here
```

Rescheduling after Node Failure

- If an entire compute node is lost, OpenShift will reschedule the application pods to remaining compute nodes.
- The OpenShift scheduler is responsible, and there are a few scheduling profiles that can be used.
  - LowNodeUtilization - This profile attempts to spread pods evenly across nodes to get low resource usage per node. This profile provides the default scheduler behavior.
  - HighNodeUtilization - This profile attempts to place as many pods as possible on to as few nodes as possible. This minimizes node count and has high resource usage per node.
  - NoScoring - This is a low-latency profile that strives for the quickest scheduling cycle by disabling all score plugins. This might sacrifice better scheduling decisions for faster ones.

Descheduling

- The OpenShift descheduler can evict pods so they can be rescheduled onto more suitable nodes
- You can benefit from descheduling running pods in situations such as the following:
  - Nodes are underutilized or overutilized.
  - Pod and node affinity requirements, such as taints or labels, have changed and the original scheduling decisions are no longer appropriate for certain nodes.
  - New nodes are added to clusters.

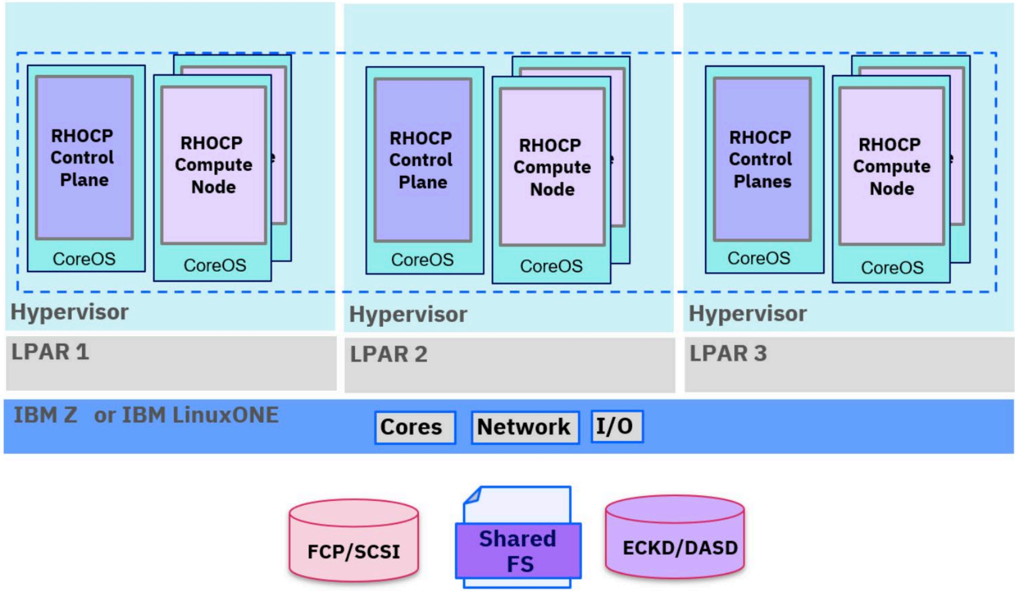
Separating Nodes by LPAR

- Control Planes should be distributed across multiple LPARs in the event of an LPAR failure.

- IBM recommends 3 LPARs with a Control Plane in each

- Various OpenShift components that run on Control Planes are quorum-based and need >50% available at any given time. Therefore, a cluster with three Control Planes across only two LPARs is not truly highly available.

- If you lose 2 out of 3 Control Planes, applications will become unresponsive and it becomes a Disaster Recovery situation.



What about cluster-level availability?

Red Hat Advanced Cluster Management for Kubernetes (RHACM)

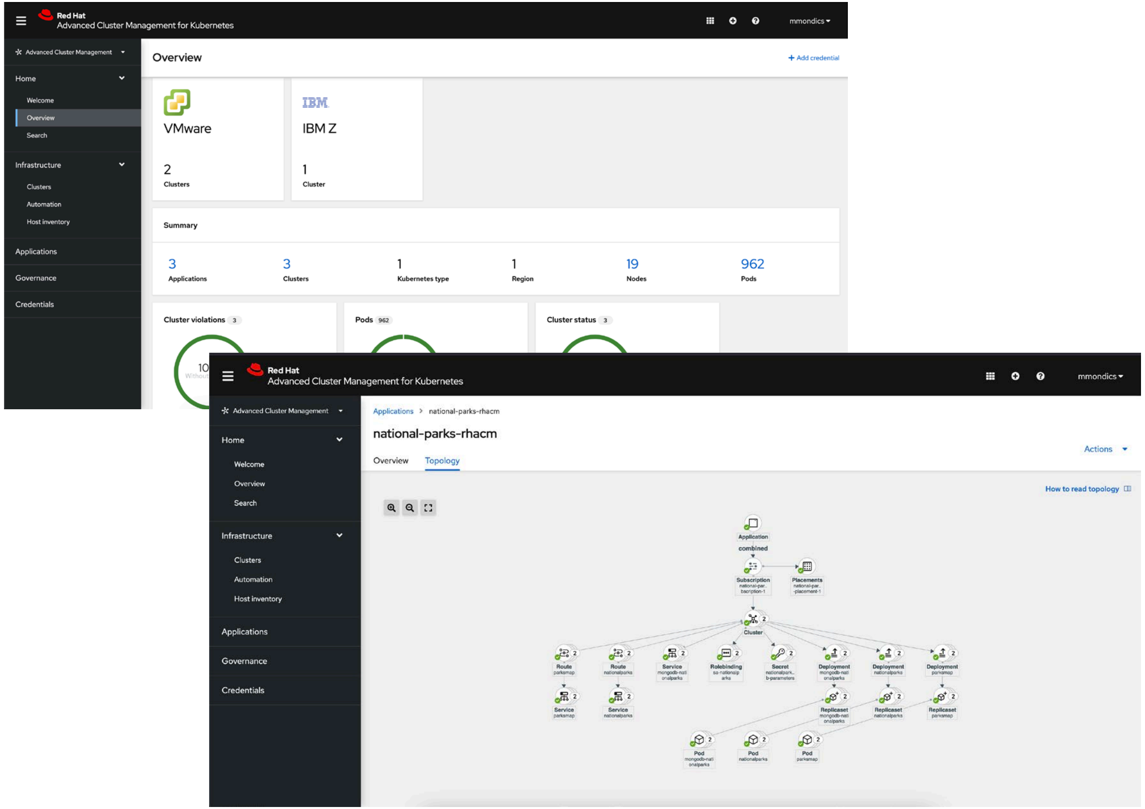
RHACM Overview

Multicluster lifecycle management

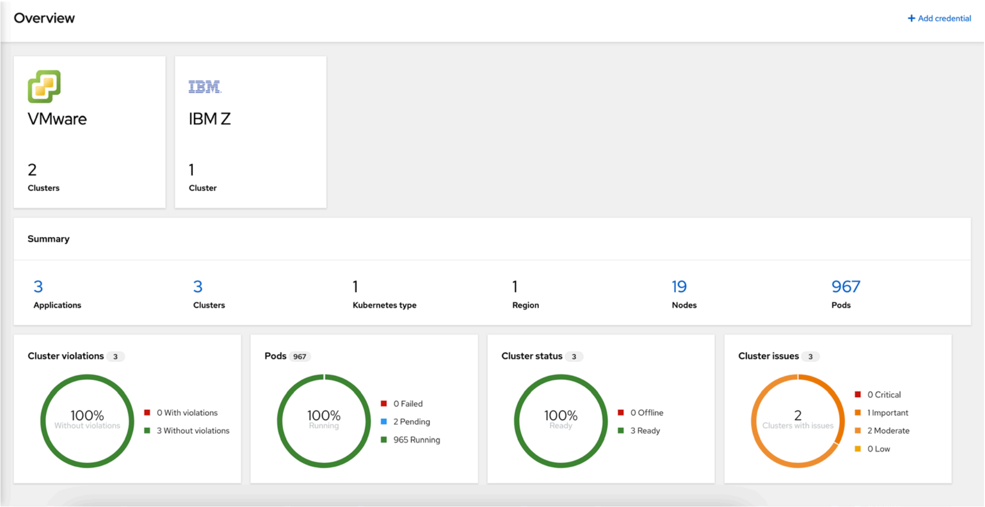
Multicluster observability for health and optimization

Advanced application lifecycle management

Policy driven governance, risk and compliance



Unified Multi-cluster Management



Overview

VMware
2
Clusters

IBM Z
1
Cluster

Summary

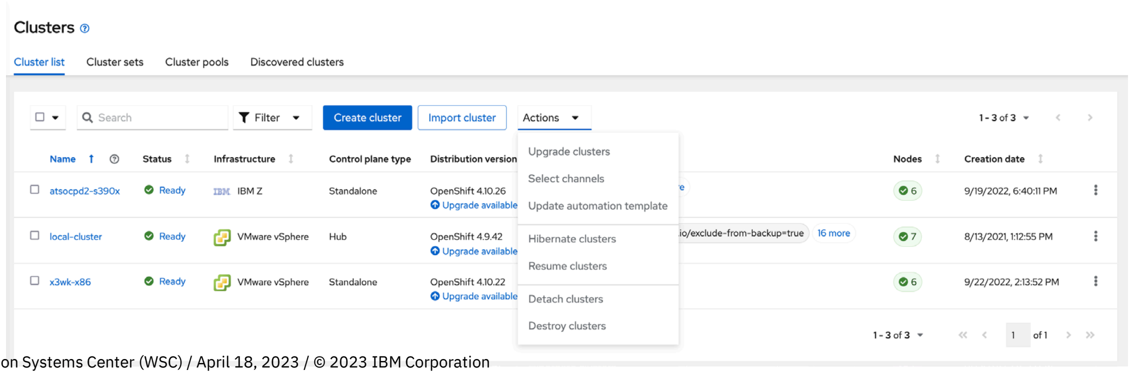
3	3	1	1	19	967
Applications	Clusters	Kubernetes type	Region	Nodes	Pods

Cluster violations 3
100% Without violations
- 0 With violations
- 3 Without violations

Pods 967
100% Running
- 0 Failed
- 2 Pending
- 965 Running

Cluster status 3
100% Ready
- 0 Offline
- 3 Ready

Cluster issues 3
2 Clusters with issues
- 0 Critical
- 1 Important
- 2 Moderate
- 0 Low



Clusters

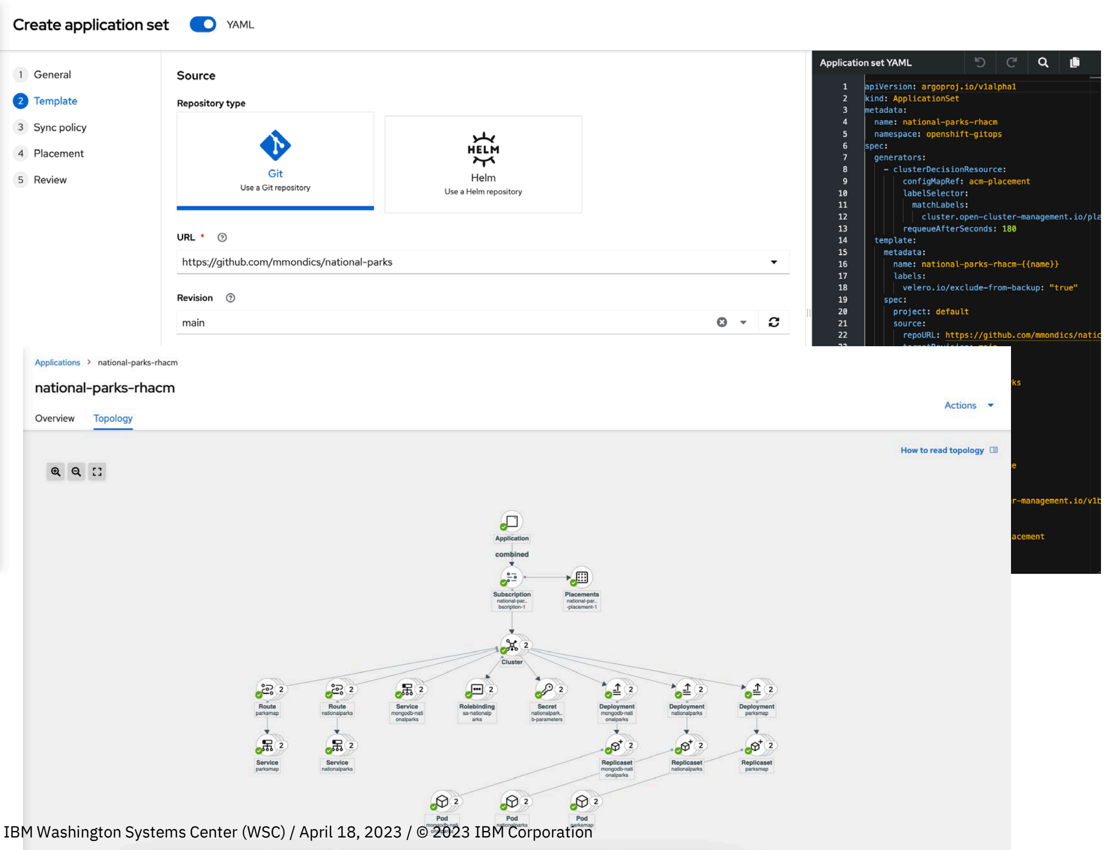
Name	Status	Infrastructure	Control plane type	Distribution version		Nodes	Creation date
atsocpd2-s390x	Ready	IBM Z	Standalone	OpenShift 4.10.26 Upgrade available	e	6	9/19/2022, 6:40:11 PM
local-cluster	Ready	VMware vSphere	Hub	OpenShift 4.9.42 Upgrade available	io/exclude-from-backup=true 16 more	7	8/13/2021, 1:12:55 PM
x3wk-x86	Ready	VMware vSphere	Standalone	OpenShift 4.10.22 Upgrade available		6	9/22/2022, 2:13:52 PM

- **Centrally** create, update and delete Kubernetes clusters across multiple private and public clouds

- Search, find and modify any Kubernetes resource across the **entire domain**

- Quickly troubleshoot and resolve issues across your **federated** domain

Multi-cluster Application Deployment and Lifecycle Management



Create application set YAML

1 General
2 Template
3 Sync policy
4 Placement
5 Review

Source

Repository type

Git
Use a Git repository

Helm
Use a Helm repository

URL *: https://github.com/mmondics/national-parks

Revision: main

Application set YAML

```
apiVersion: argoproj.io/v1alpha1
kind: ApplicationSet
metadata:
  name: national-parks-rhacm
  namespace: openshift-gitops
spec:
  generators:
    - clusterDecisionResource:
        configMapRef: acm-placement
        labelSelector:
          matchLabels:
            cluster.open-cluster-management.io/pla
        requeueAfterSeconds: 180
  template:
    metadata:
      name: national-parks-rhacm-{{name}}
      labels:
        velero.io/exclude-from-backup: "true"
    spec:
      project: default
      source:
        repoURL: https://github.com/mmondics/natic
```

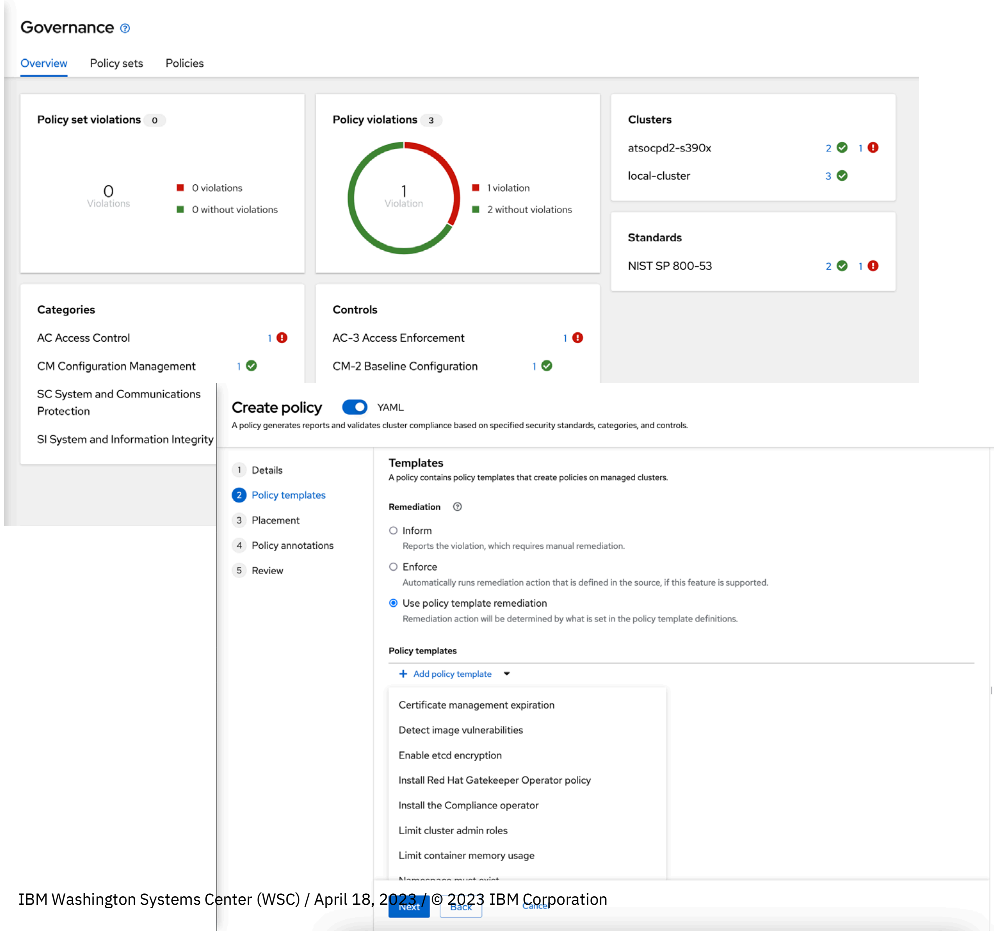
Applications > national-parks-rhacm

national-parks-rhacm

Overview Topology

- **Deploy applications and infrastructure with a GitOps approach** where Git is the single “source of truth”
- **Easily deploy** an Application using the Application Builder
- Deploy Applications from **multiple sources** (Git / Helm / Object Storage)
- Quickly **visualize** application relationships **across** clusters on **different infrastructure**, zone, or Kubernetes version

Policy-Based Governance, Risk, and Compliance



Governance

Overview | Policy sets | Policies

Policy set violations 0

0 Violations

🟥 0 violations
🟩 0 without violations

Policy violations 3

1 Violation

🟥 1 violation
🟩 2 without violations

Clusters

atsocpd2-s390x	2 ✓	1 !
local-cluster	3 ✓	

Standards

NIST SP 800-53	2 ✓	1 !

Categories

AC Access Control	1 !
CM Configuration Management	1 ✓
SC System and Communications Protection	
SI System and Information Integrity	

Controls

AC-3 Access Enforcement	1 !
CM-2 Baseline Configuration	1 ✓

Create policy YAML

A policy generates reports and validates cluster compliance based on specified security standards, categories, and controls.

1. Details
2. Policy templates
3. Placement
4. Policy annotations
5. Review

Templates

A policy contains policy templates that create policies on managed clusters.

Remediation

○ Inform
Reports the violation, which requires manual remediation.

○ Enforce
Automatically runs remediation action that is defined in the source, if this feature is supported.

● Use policy template remediation
Remediation action will be determined by what is set in the policy template definitions.

Policy templates

+ Add policy template ▾

- Certificate management expiration
- Detect image vulnerabilities
- Enable etcd encryption
- Install Red Hat Gatekeeper Operator policy
- Install the Compliance operator
- Limit cluster admin roles
- Limit container memory usage
- Namespace must exist

- **Centrally** set & enforce policies for security, applications and infrastructure
- Quickly **visualize** detailed **auditing** on configuration of apps and clusters
- Built-in compliance policies and audit checks
- **Immediate visibility** into your compliance posture based on your defined standards

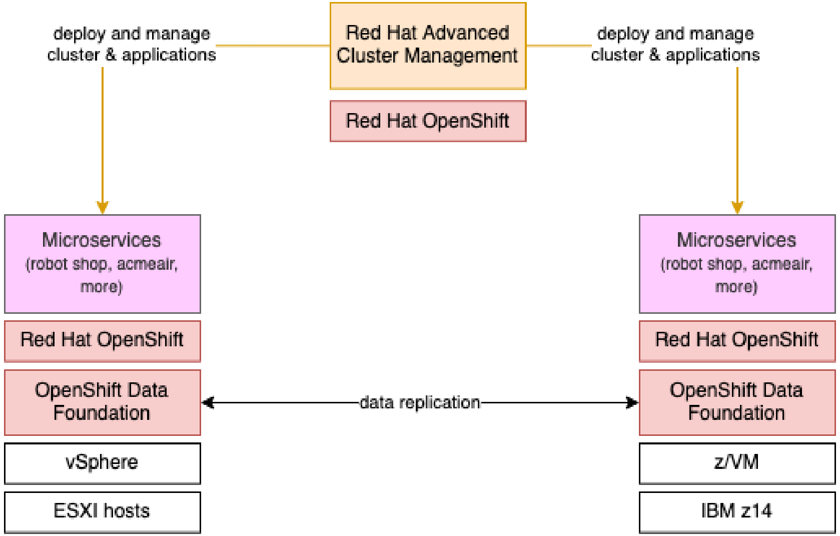
Policy-Based Governance, Risk, and Compliance (cont.)

- Use RHACM as an Infrastructure-as-Code platform
- Deploy infrastructure/security components from YAML files hosted in a Git solution in a standardized way
  - etcd encryption
  - LDAP configuration
  - Roles and RoleBindings
  - Any cluster customization

```
apiVersion: config.openshift.io/v1
kind: OAuth
metadata:
  name: cluster
spec:
  identityProviders:
  - ldap:
      attributes:
        email: []
        id:
        - dn
        name:
        - cn
        preferredUsername:
        - uid
      bindDN: uid=zz-rhosz1-prod-oc-console,ou=bind-ids,dc=wsc,dc=ibm
      bindPassword:
        name: ldap-bind-password-wscldap
      insecure: true
      url: ldap://192.168.176.6:389/dc=wsc,dc=ibm?uid?sub?(objectclass=top)
    mappingMethod: claim
    name: ldap-ats-wscdmz-wfwfsldapcl01
    type: LDAP
  tokenConfig:
    accessTokenMaxAgeSeconds: 172800
```

RHACM for High Availability

- Red Hat does not recommend “stretched” clusters – clusters with nodes in different datacenters
- Typically, multiple cluster in an active – passive configuration for Disaster Recovery
- Red Hat Advanced Cluster Management can be used to migrate workload applications from active to passive
- Or simply deploy the same application on multiple active clusters simultaneously
  - Same concept applies to dev/test/production clusters


Red Hat Advanced Cluster Management
Red Hat OpenShift
deploy and manage cluster & applications
deploy and manage cluster & applications
Microservices (robot shop, acmeair, more)
Red Hat OpenShift
OpenShift Data Foundation
vSphere
ESXI hosts
data replication
Microservices (robot shop, acmeair, more)
Red Hat OpenShift
OpenShift Data Foundation
z/VM
IBM z14

RHACM Live Demo

Thank you

Matt Mondics

Client Technical Specialist

Hybrid Cloud on IBM zSystems

—

matt.mondics@ibm.com