Red Hat
Advanced Cluster Management
for Kubernetes

Business Continuity

ACM Deep Dive

Red Hat

Agenda

- Hub Backup and Restore
- VolSync
- Metro and Regional Disaster Recovery

Hub Backup and Restore

Hub Backup and Restore - Overview

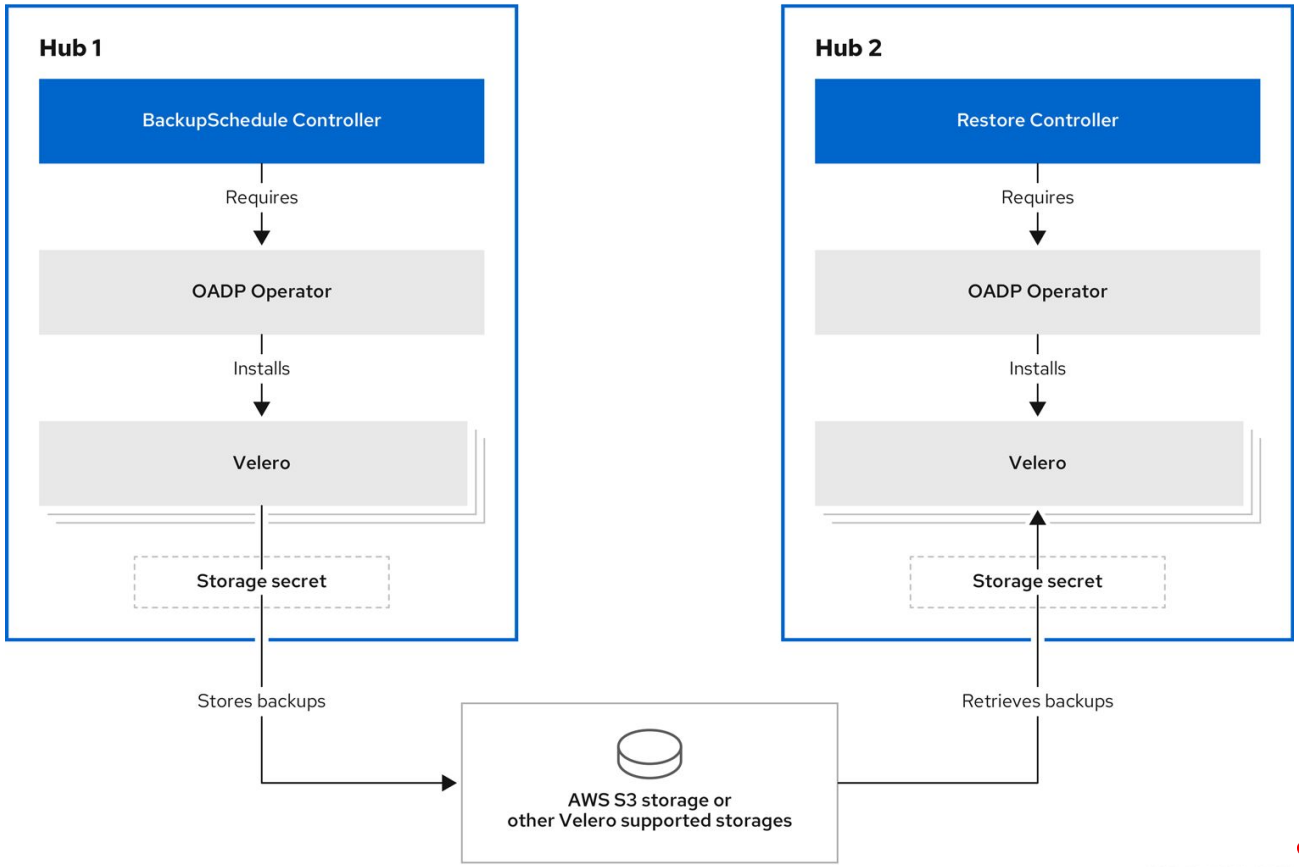
- **Provides backup and restore support for ACM HUB resources**
  - Policies configuration
  - Application placement configuration and definition
  - Managed cluster configuration
  - other resources created on the hub
- **In a disaster scenario, when the hub goes down, these resources can be restored on another hub using the backed up data**
- **Scenarios outside the scope of the Hub backup**
  - disaster recovery scenarios for applications running on managed clusters or scenarios where the managed clusters go down

Components/Architecture

The ACM Cluster Backup and Restore Operator consists of Kubernetes Custom Controllers that run on the hub cluster and depends on the OADP Operator to create a connection to a backup storage location on the hub. The OADP Operator also installs Velero which is the component used to backup and restore hub resources.

Cluster Backup and Restore flow:

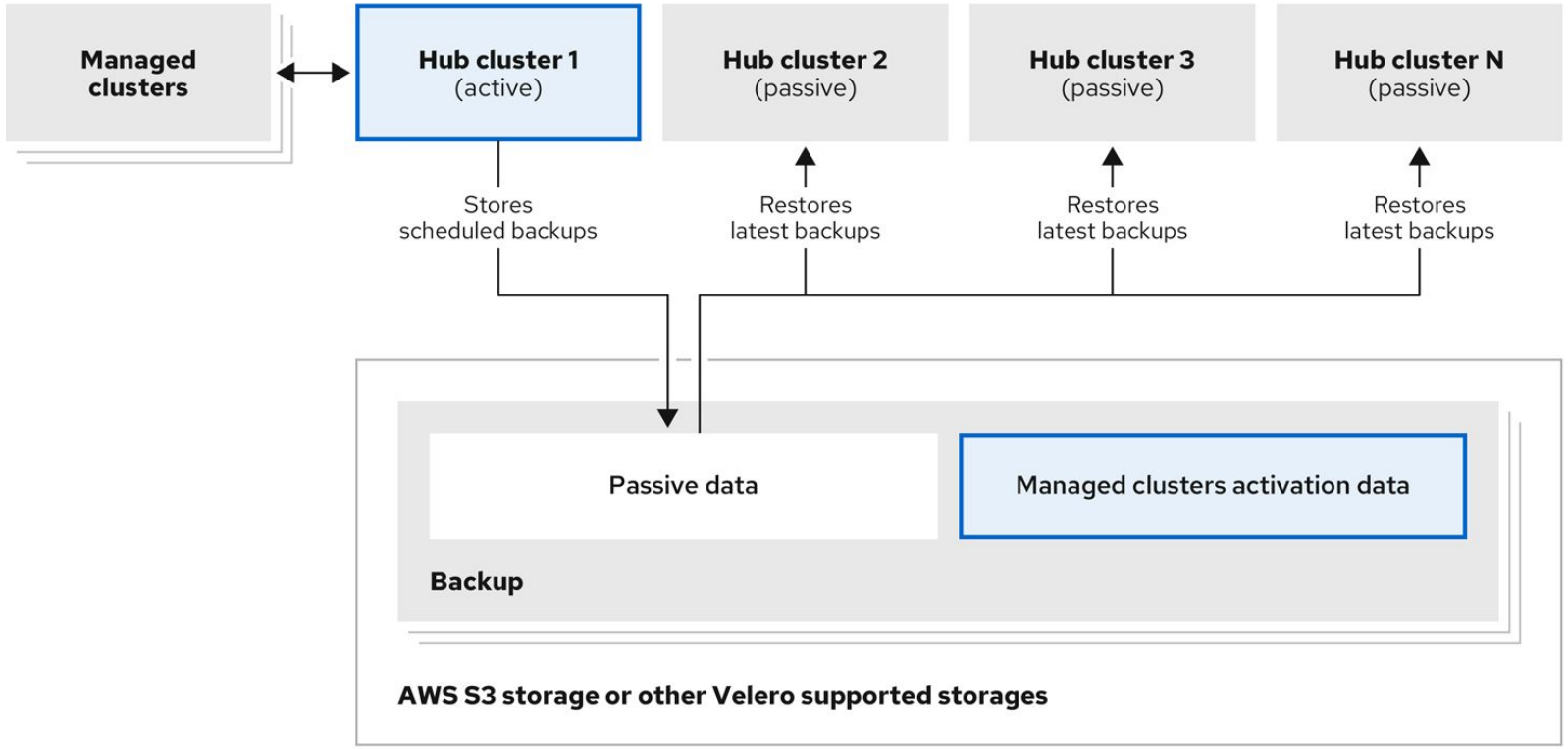
- ACM Backup and Restore operator provides 2 controllers
  - BackupSchedule controller to create backup schedules
  - Restore controller to support restoring resources from a backup



Red Hat

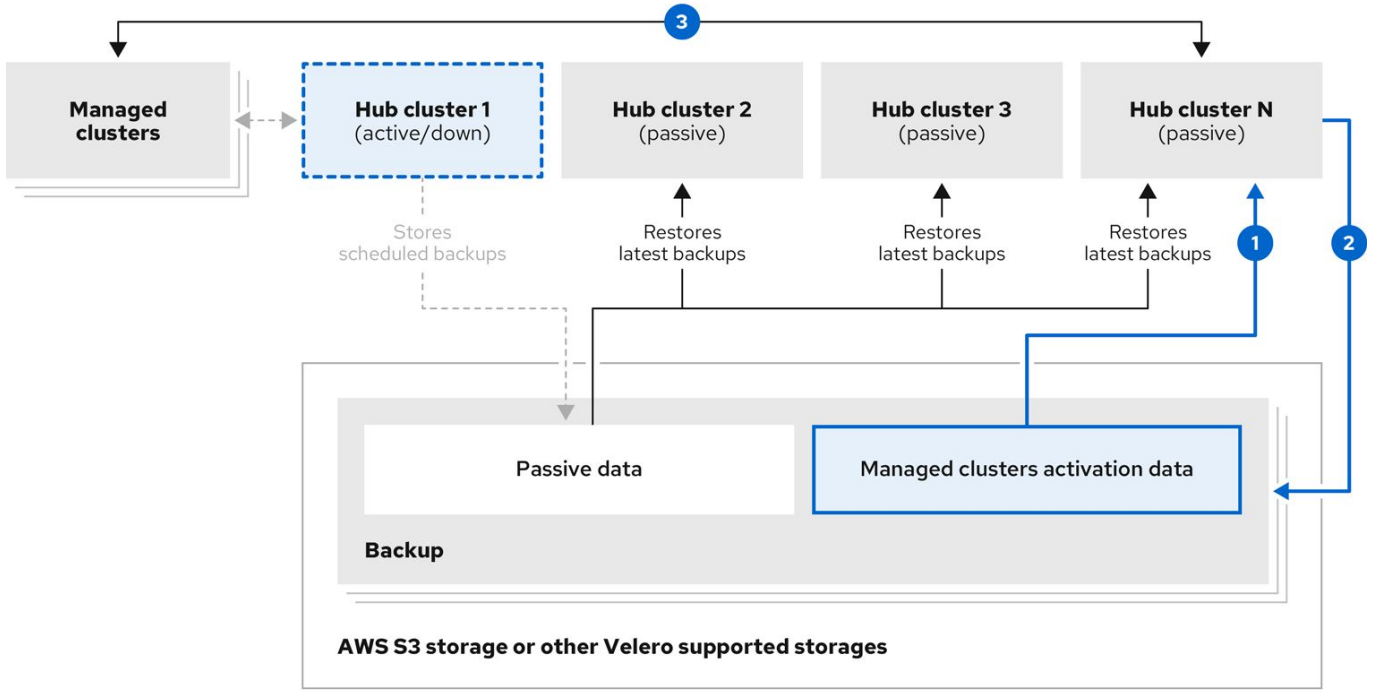
Active Passive Configuration Design

- The active hub manages the remote clusters and backs up hub data at regular intervals.
- The passive hubs restore this data, except for the managed clusters activation data, which would move the managed clusters to the passive hub. The passive hubs can restore the passive data continuously, or as a one time operation.



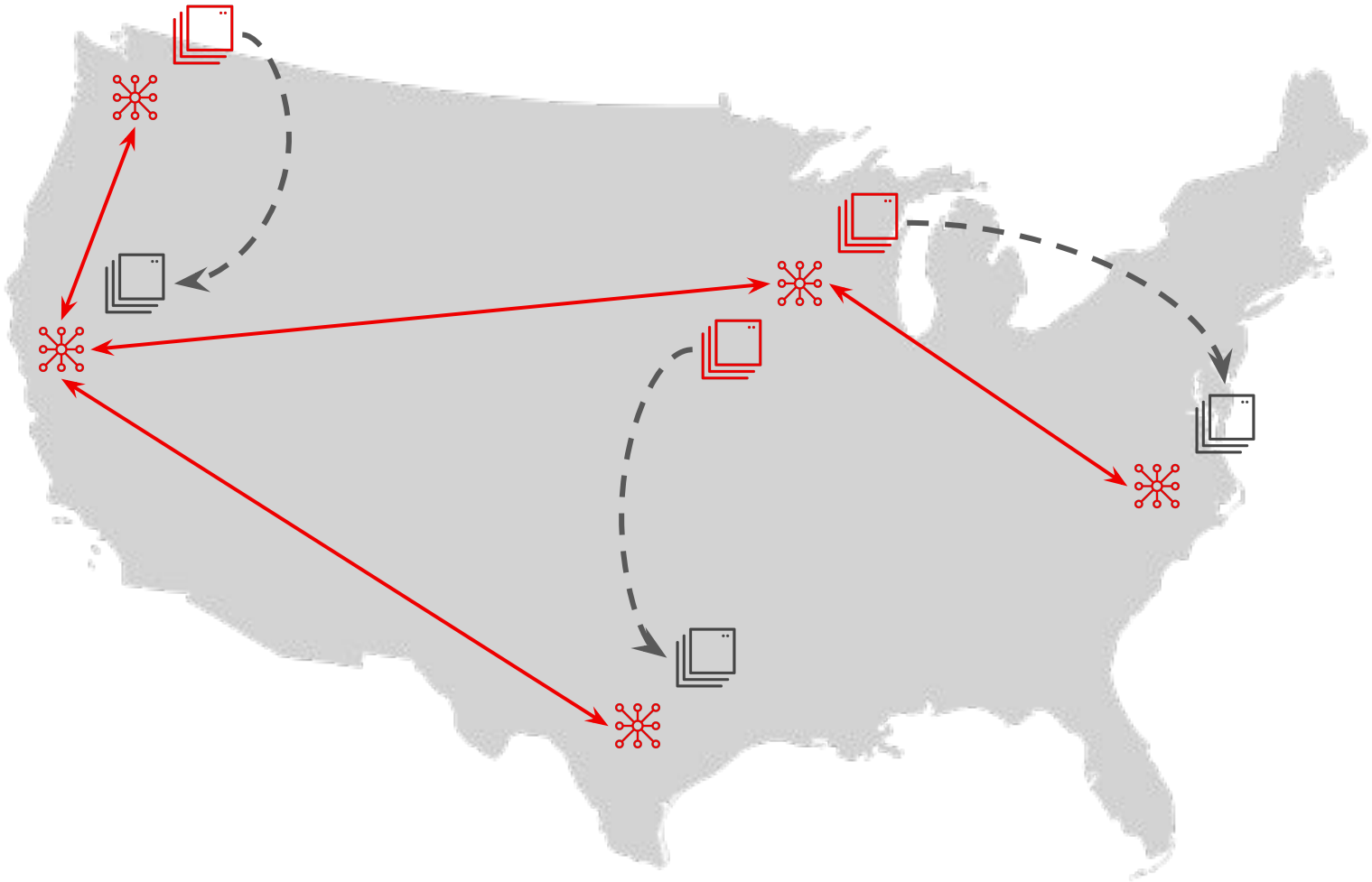
Disaster Recovery

- When the primary hub goes down, one of the passive hubs is chosen by the admin to take over the managed clusters. In the image below, the admin decides to use Hub N as the new primary hub.
- Hub N restores the Managed Cluster activation data. At this point, the managed clusters connect with Hub N.
- The admin starts a backup on the new primary Hub N by creating a BackupSchedule resource.



VolSync Operator

Application migration



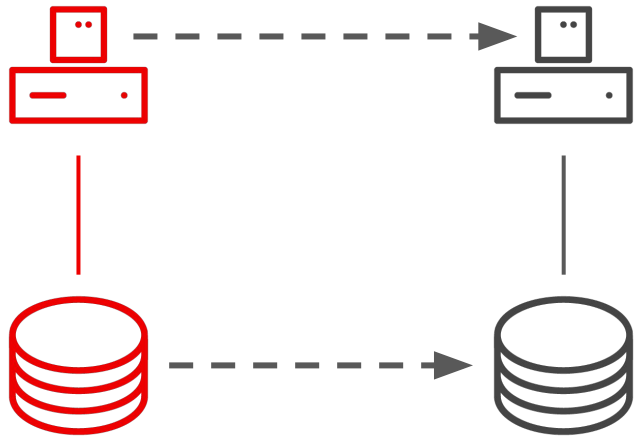
Why multi-cluster:

- Minimize cost
- Optimize delivery
- Disaster recovery

Moving applications:

- Stateless apps move easily
- **Data complicates movement**

How is data moved today?

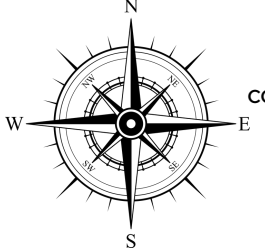


Traditional approach relies on the storage system:

- Vendor must support mobility
- Requires the same storage everywhere

The old way is insufficient:

- Cloud deployments are heterogeneous
- Expanded set of use-cases



Overview

VolSync is an operator/controller that can be installed on managed clusters to perform persistent volume replication within or across clusters. The replication provided by VolSync is independent of the storage system. This allows replication to and from storage types that don't normally support remote replication. Additionally, it can replicate across different types (and vendors) of storage.

- VolSync within ACM is managed as part of the DR4Hub squad. VolSync is GA as of ACM 2.6
- VolSync is intended to be installed on managed clusters and can be deployed from the ACM hub via a ManagedClusterAddOn. The ACM hub deploys the volsync-addon-controller which watches ManagedClusterAddOns and deploys the VolSync operator accordingly.

VolSync

Kubernetes operator for data mobility

Features

- Relies only on Kubernetes functionality
- Storage system independent
- Disparate sources & destinations
- Extensible set of data movers

Use cases

- Disaster recovery
- Data distribution
- Application migration
- Data backup
- Wide-area data sharing

Extensible data movers

ReplicationSource ⇒ ReplicationDestination

rsync - 1:1 replication

Provides efficient delta transfers directly from source to destination

rclone - 1:n data distribution

Uses intermediate cloud storage to support wide fan-out

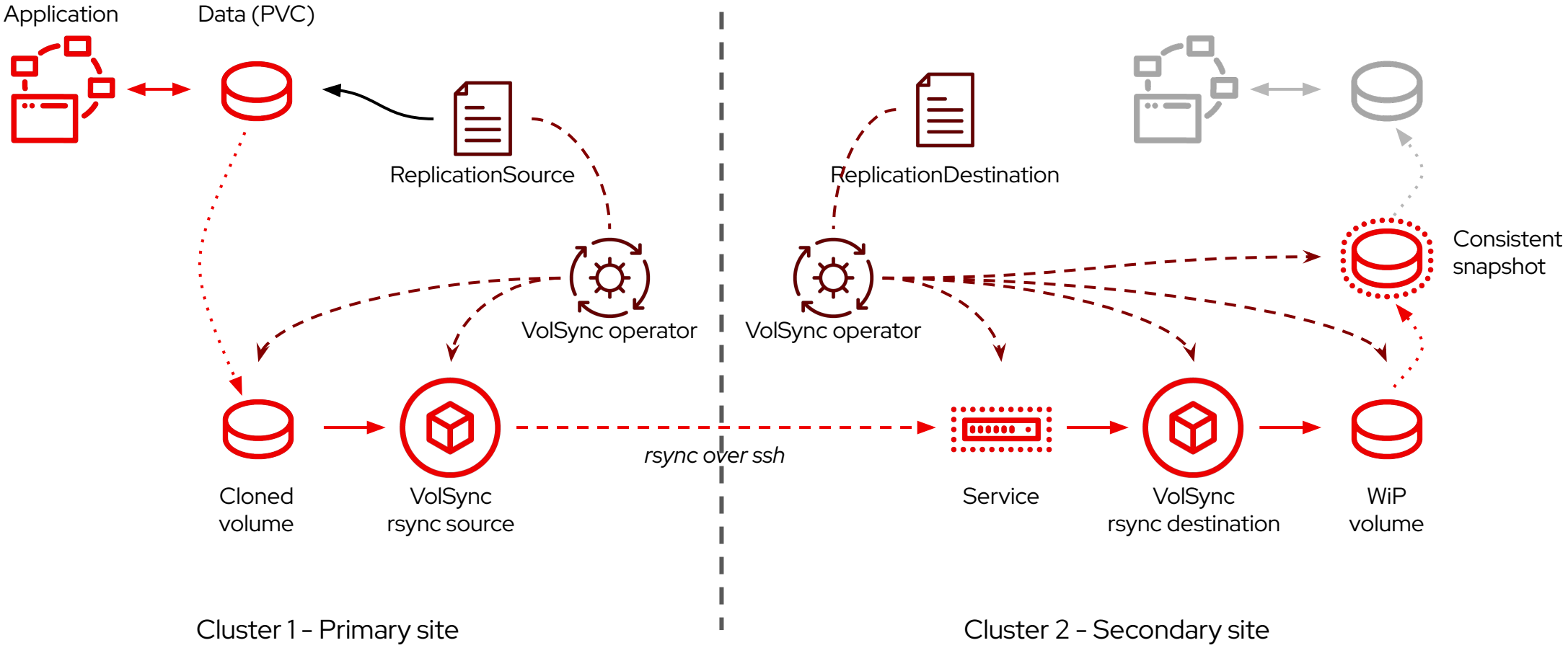
restic - PV data backup

Dedup-ed/compressed volume backups to object storage

syncthing - data sharing

Permits live, multi-site access to shared data with eventual consistency

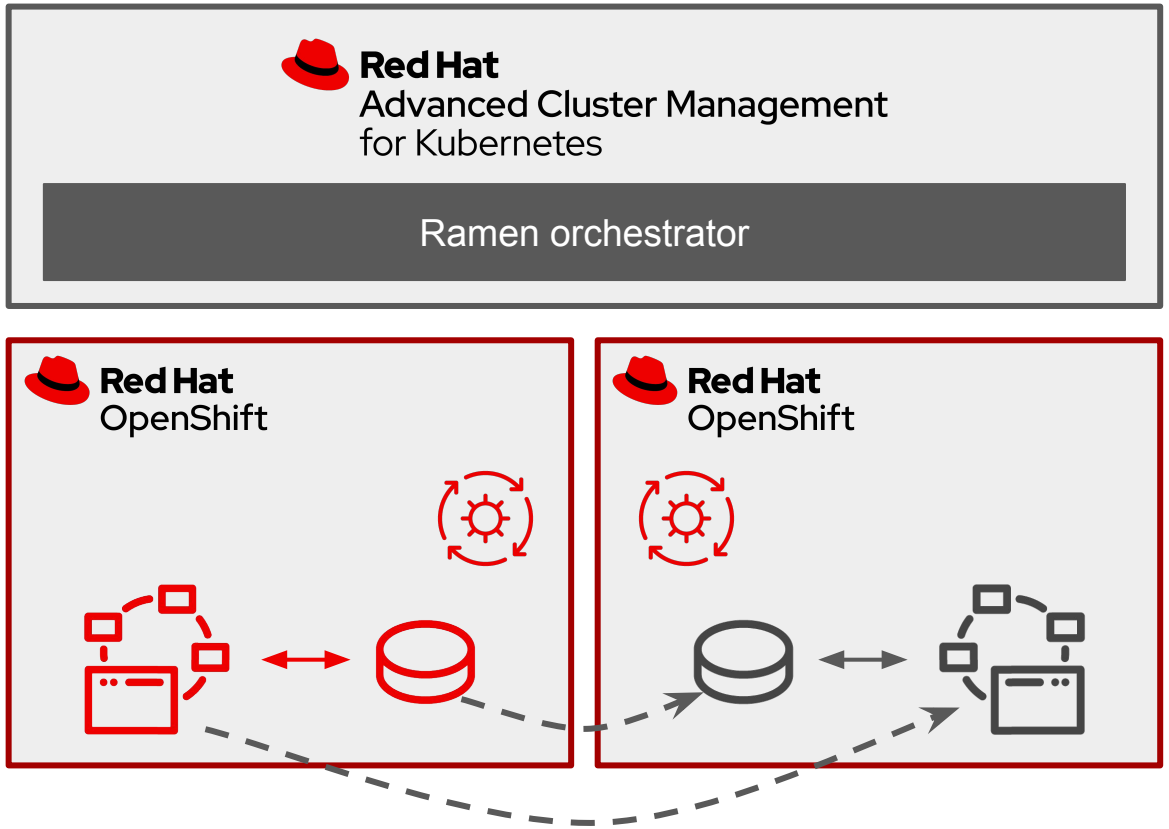
Application migration (rsync)



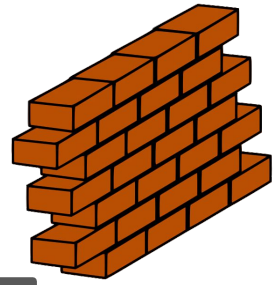
Red Hat

Application-level cross-cluster orchestration

- RHACM provides multi-cluster view
- Ramen orchestrator manages migration
  - Application setup
    - Initialize replication relationship
  - Application relocation
    - Shut down/Quiesce application
    - Issue final sync to VolSync
    - Start application at new site



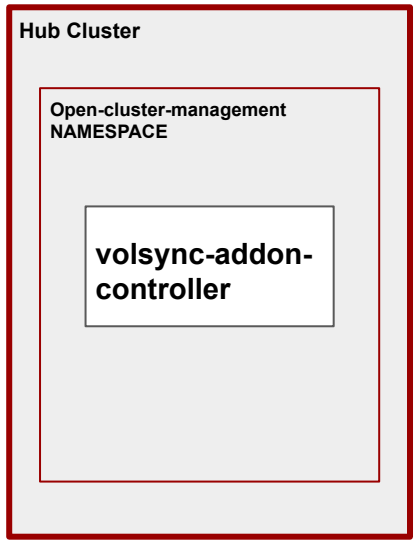
Components/Architecture

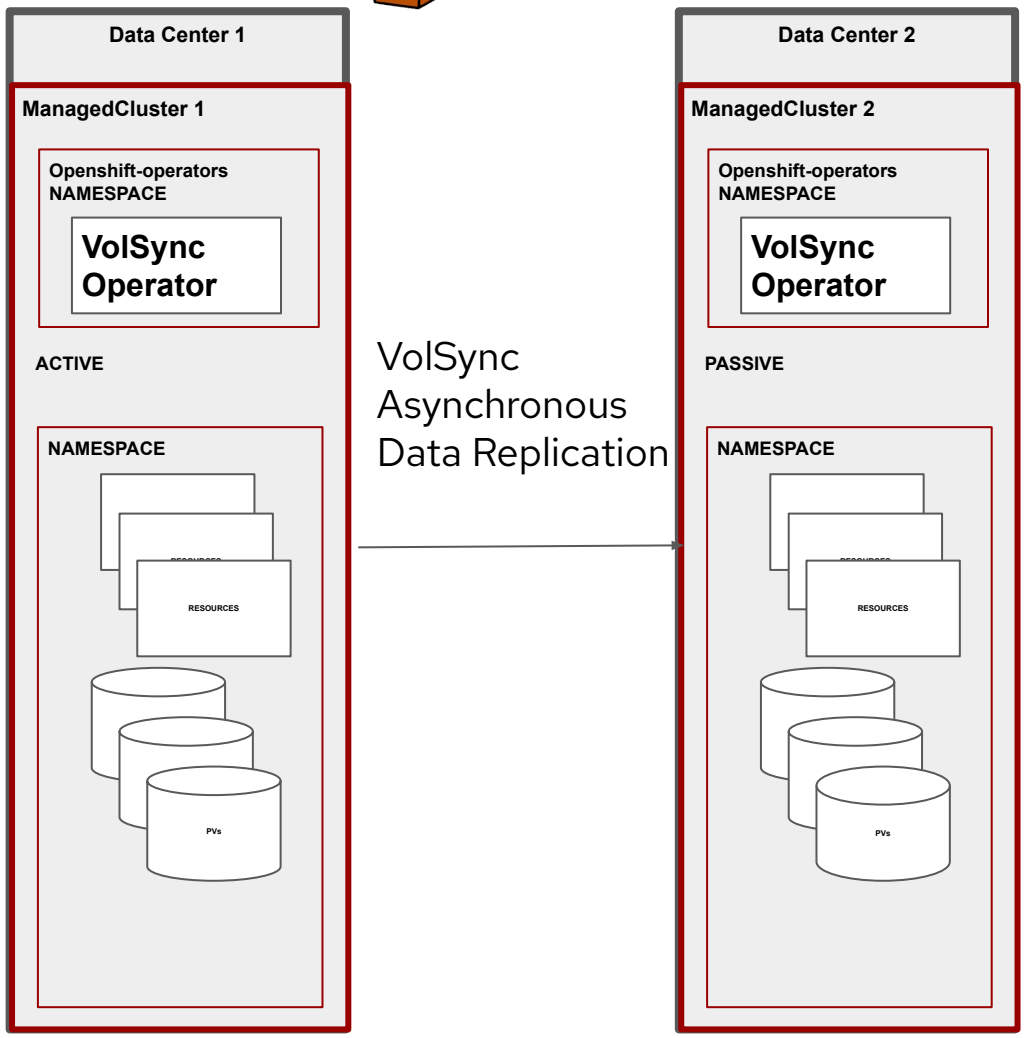


Components:

- Volsync-addon-controller (on ACM hub cluster)
- VolSync Operator (on ACM managed clusters)

Red Hat
Advanced Cluster Management
for Kubernetes

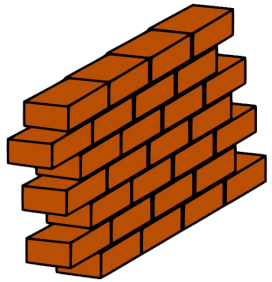




Red Hat

Components/Architecture

VolSync Model



Components:

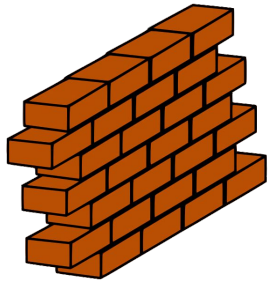
- VolSync-addon-controller is deployed to the “open-cluster-management” namespace in an ACM hub cluster.

```
❯ oc -n open-cluster-management get deploy volsync-addon-controller-c65ec-deploy
NAME                                    READY   UP-TO-DATE   AVAILABLE   AGE
volsync-addon-controller-c65ec-deploy   1/1     1            1           7d5h
```

- VolSync can be deployed to managed clusters via a ManagedClusterAddOn resource.

Components/Architecture

VolSync Model

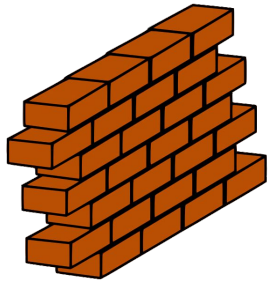


Components:

- VolSync is deployed as an operator to the “openshift-operators” namespace in an ACM managed cluster.
- Custom Resource Definitions:
  - ReplicationSource
  - ReplicationDestination

Components/Architecture

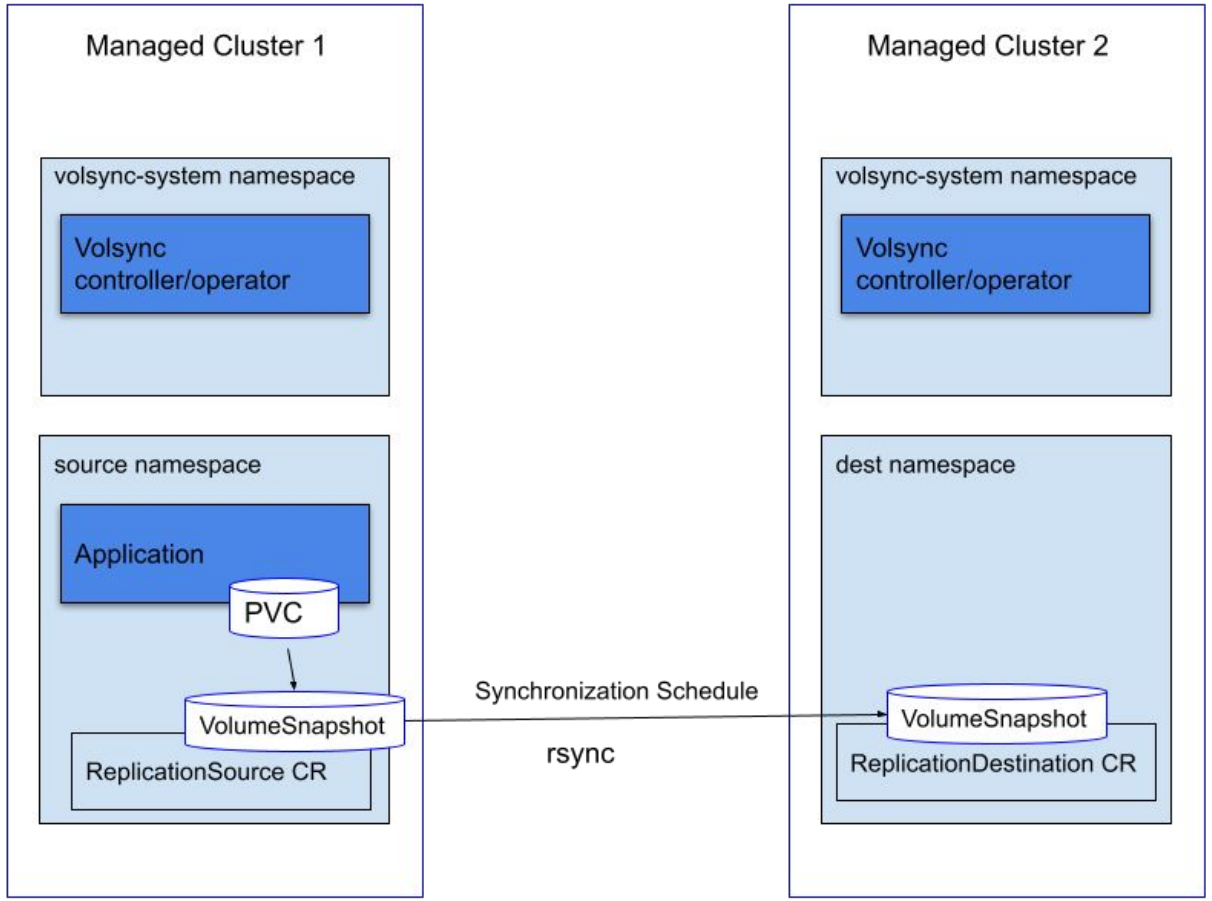
VolSync Model



- Replication types:
  - Rsync replication
    - Use Rsync-based replication for 1:1 replication of volumes in scenarios such as disaster recovery, mirroring to a test environment, or sending data to a remote site for processing.
  - Rclone replication
    - Use Rclone-based replication for multi-way (1:many) scenarios such as distributing data to edge clusters from a central site.
  - Restic backup
    - Create a Restic-based backup of the data in a PersistentVolume.
  - Syncthing live data synchronization (dev preview)
    - Synchronize data between multiple different locations

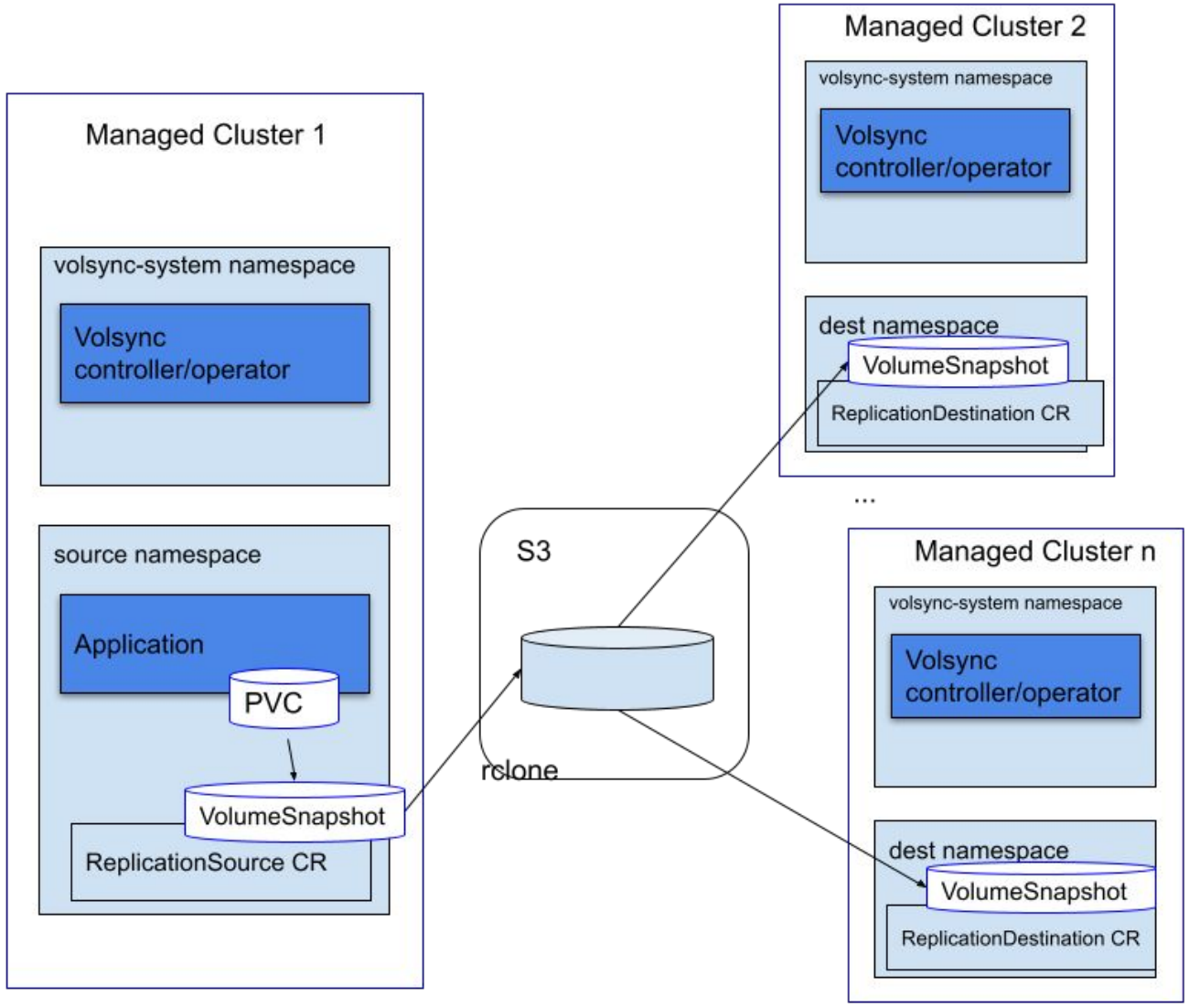
Rsync Scenario

Rsync Sample scenario with 2 managed clusters



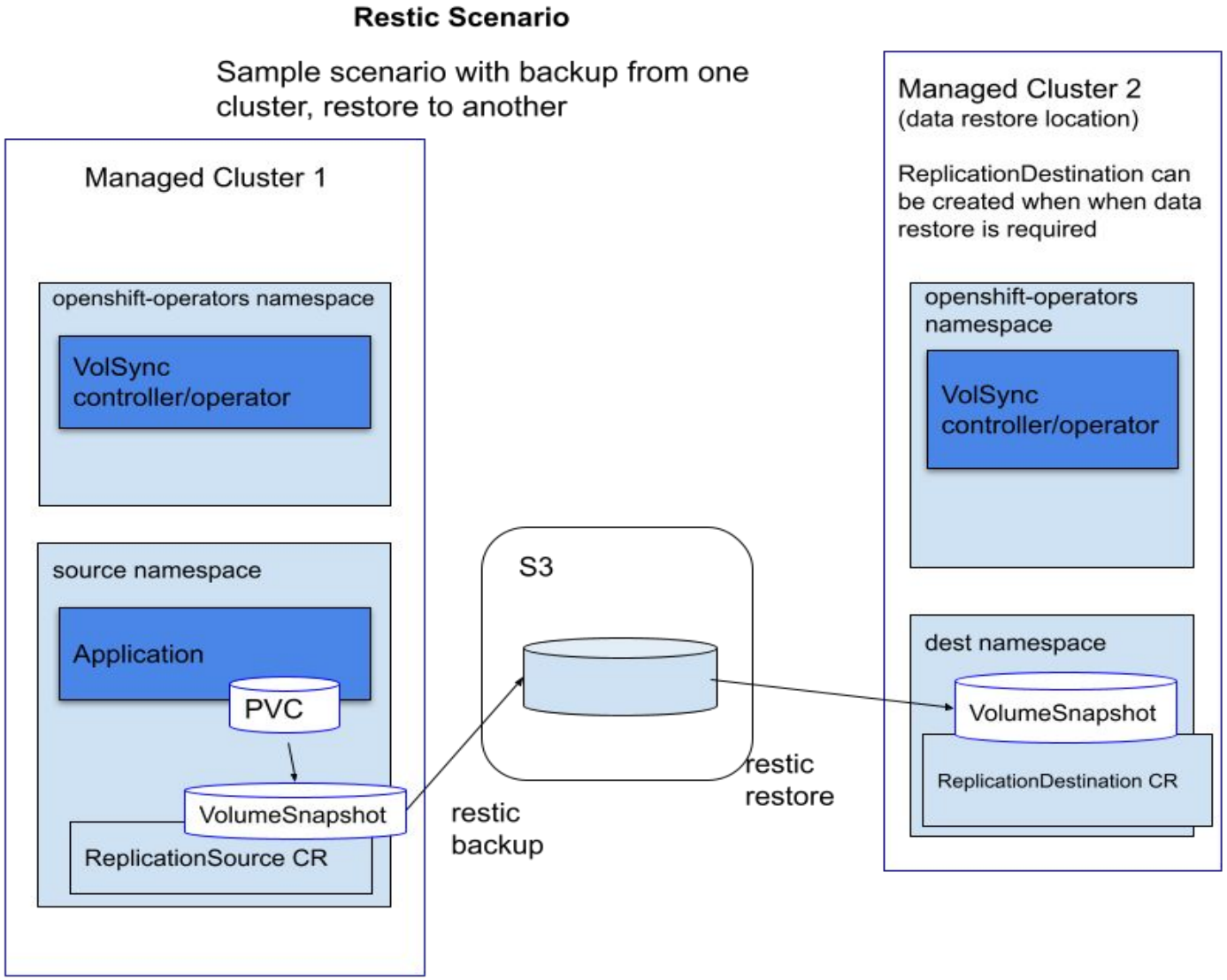
Rclone Scenario

Rclone Sample scenario with multiple managed clusters

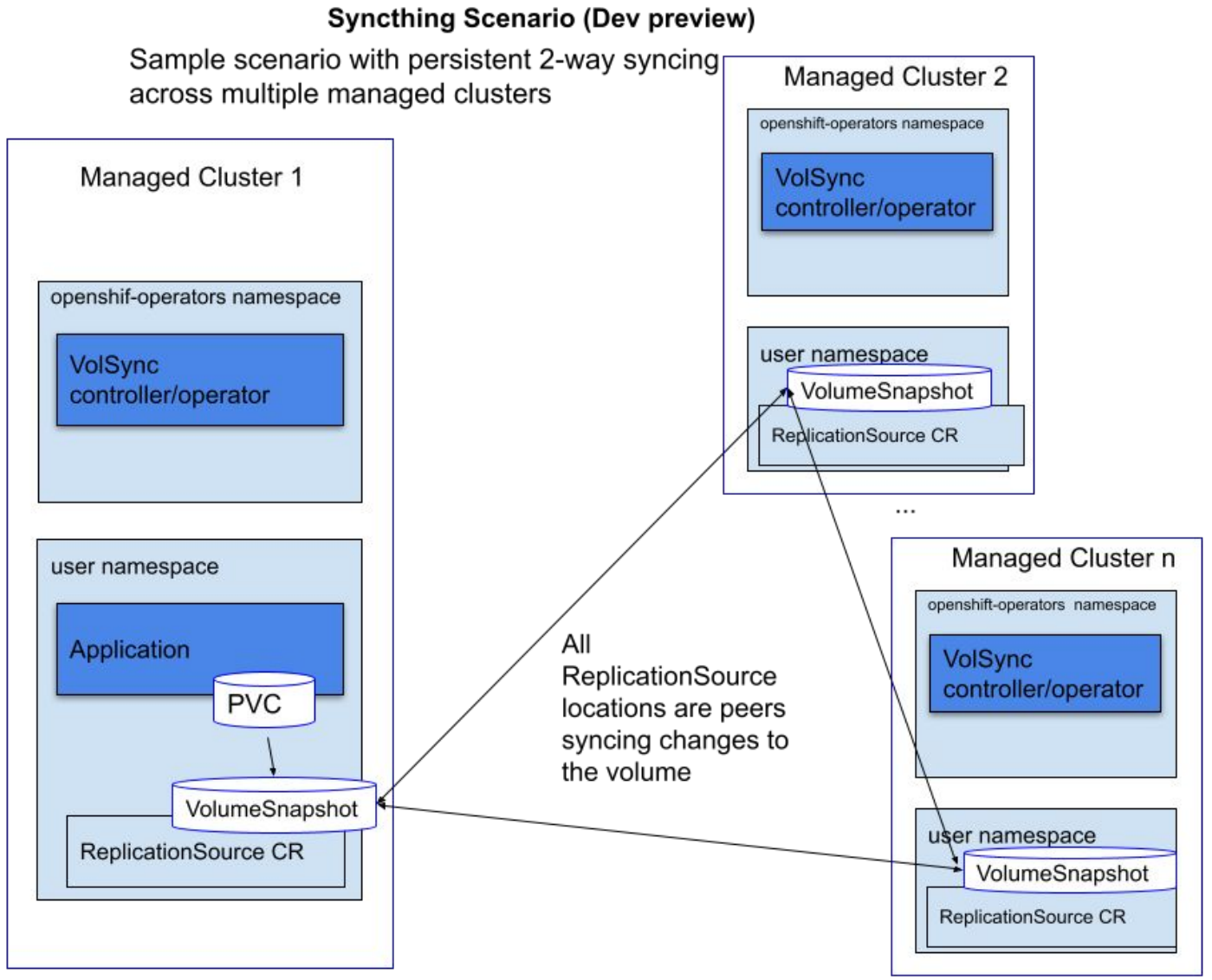


Red Hat

Restic Scenario (backup)

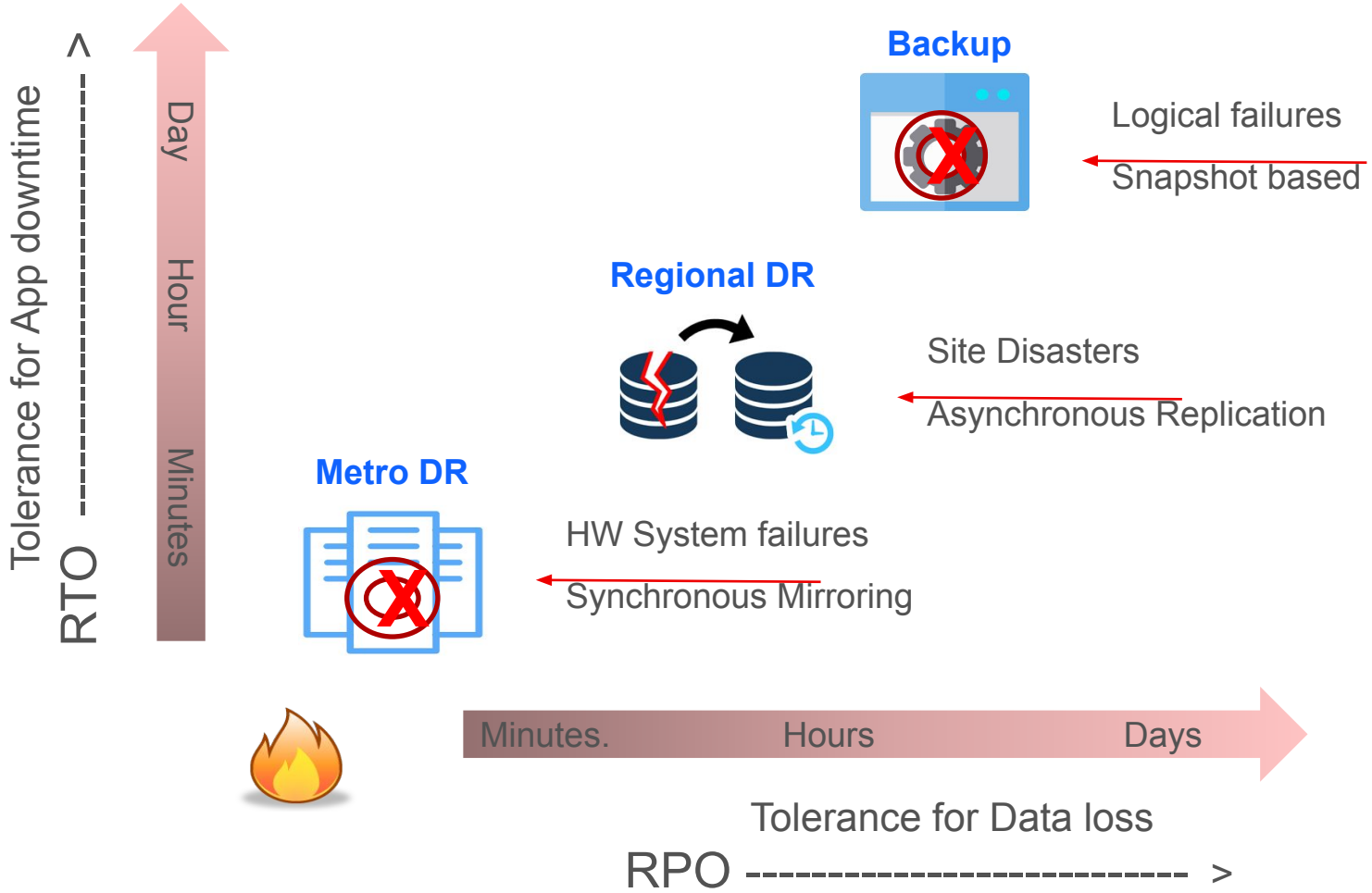


Syncthing Scenario



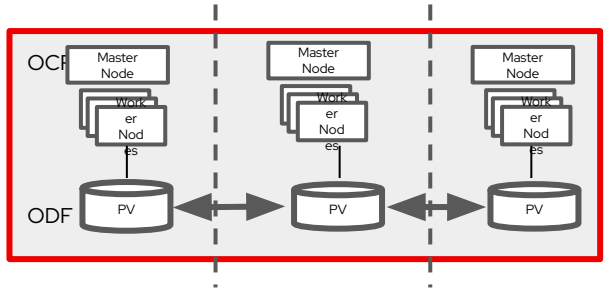
Metro and Regional DR

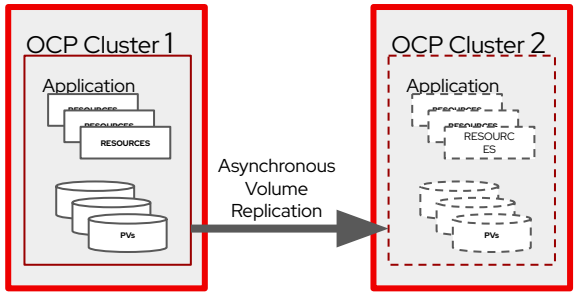
OpenShift Resiliency Solutions for different service level objectives

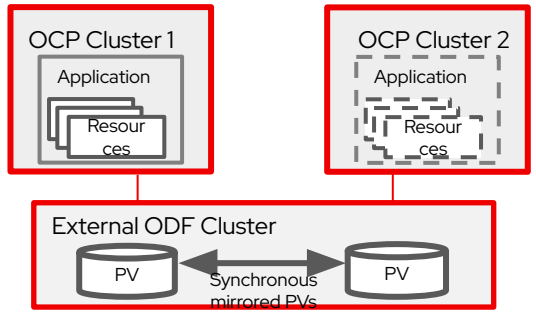


- Comprehensive protection solutions against wide spectrum of failures
- Beyond Data Protection --> Full Application protection
- Resiliency built into the platform – Available to all stateful and stateless applications on OpenShift
- OpenShift integrated stack lends towards Automated and Simplified Application granular protection

Start with HA configuration and add site DR







	Cluster HA	Regional-DR	Metro-DR
Topology	Single OCP+ODF clusters deployed over multiple AZs in a single region	Multi OCP + ODF clusters spread over multiple regions	Multi OCP clusters + single external ODF stretched cluster deployed over low latency networks
RTO (Downtime)	RTO=0 (Continuous)*	RTO = minutes DR Automation from ACM+ODF reduces RTO	RTO = minutes DR Automation from ACM+ODF reduces RTO
RPO (Data loss exposure)	RPO=0 No Data loss due due to Synchronous mirroring of ODF data	RPO > 0; Usually 5 min or higher Depends upon network bandwidth & change rate	RPO=0 No Data loss due due to Synchronous mirroring of ODF data
Infra Requirements	Multi-AZ supported public clouds (vSphere support in OCP 4.10)	All ODF supported platforms No network latency limits	On-prem only (vSphere, bare metal) <10ms network latency between sites

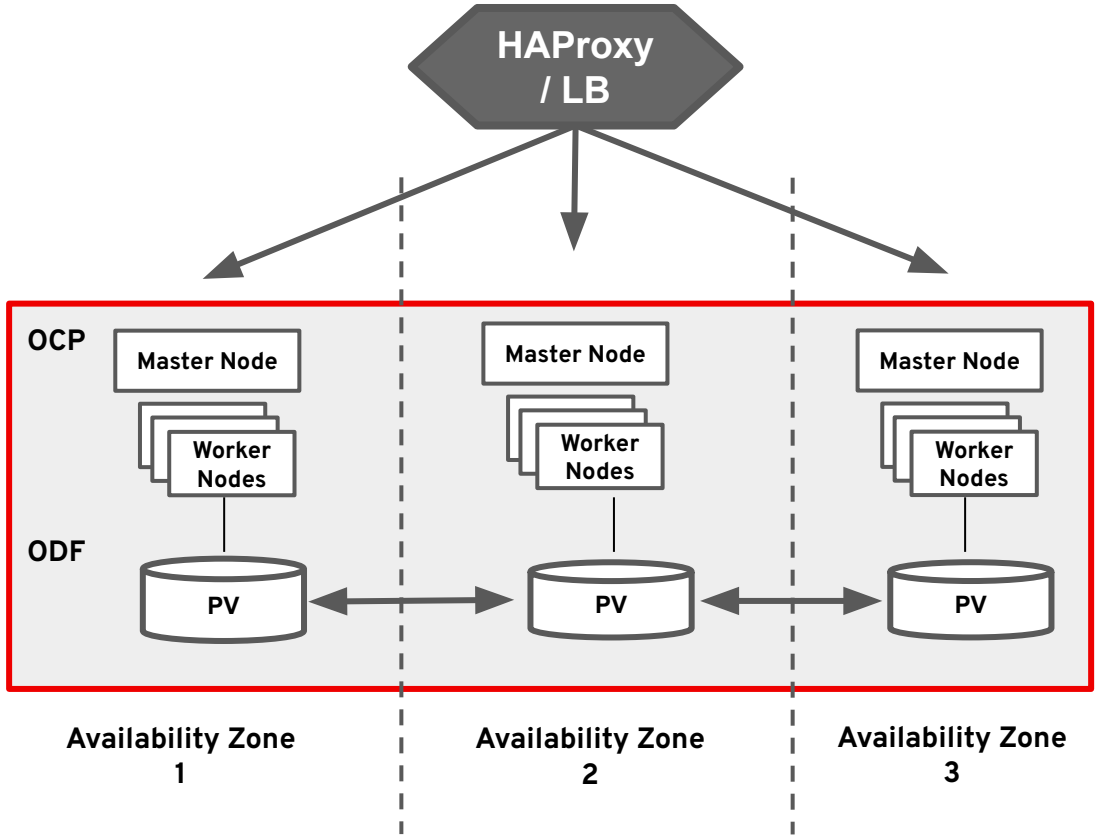
* Subject to RWO PV limitations

HA-DR Solutions Overview

Red Hat

Application HA

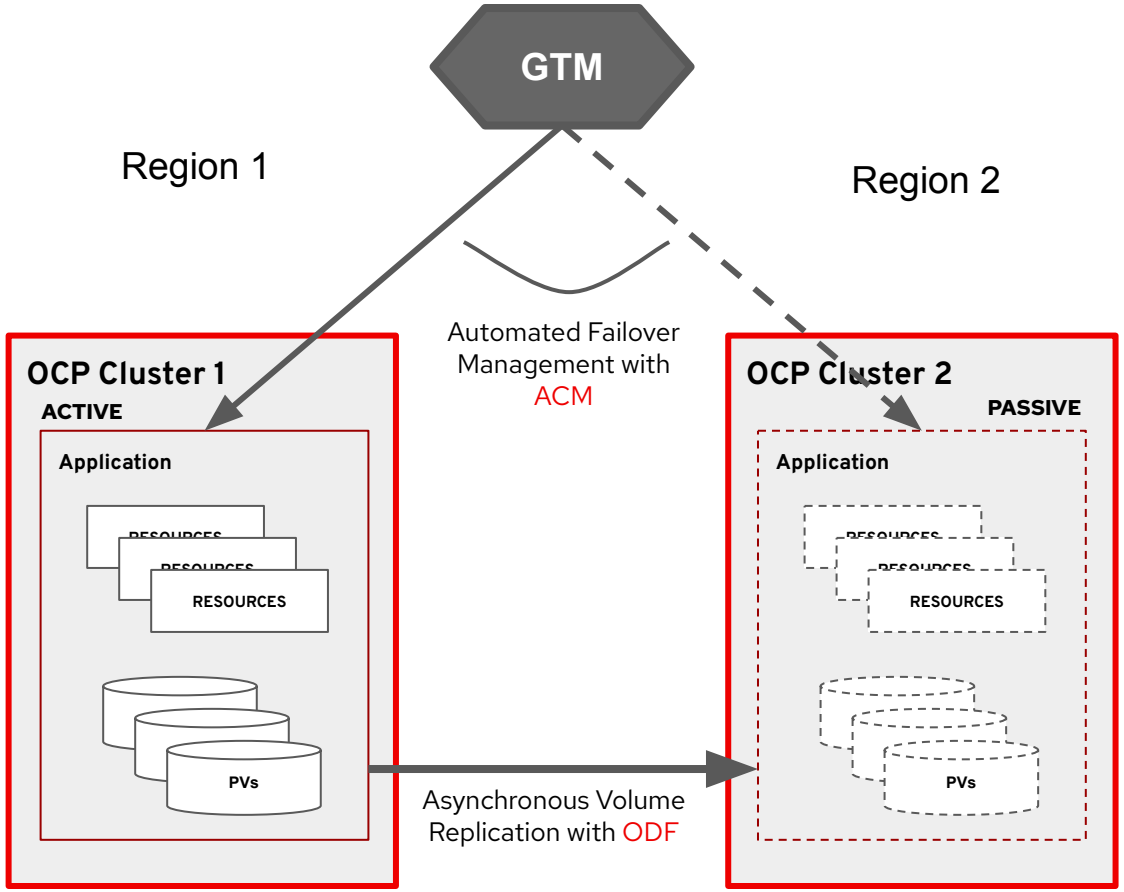
Multi-Zone spanning Cluster for local HA



- HA for Stateful Applications deployed on cluster that is stretched across Availability Zones within a region
- Installer ensures that resources are deployed across all AZs making the cluster resilient against failures of any single AZ
- ODF provides synchronous consistent copies in all AZs ensuring no data loss during zone failures
- Suitable for public cloud platforms with Regions supporting 3 or more AZs
    - Can be deployed on-prem when AZs are connected by networks with <10ms latency

Regional-DR with Failover Automation

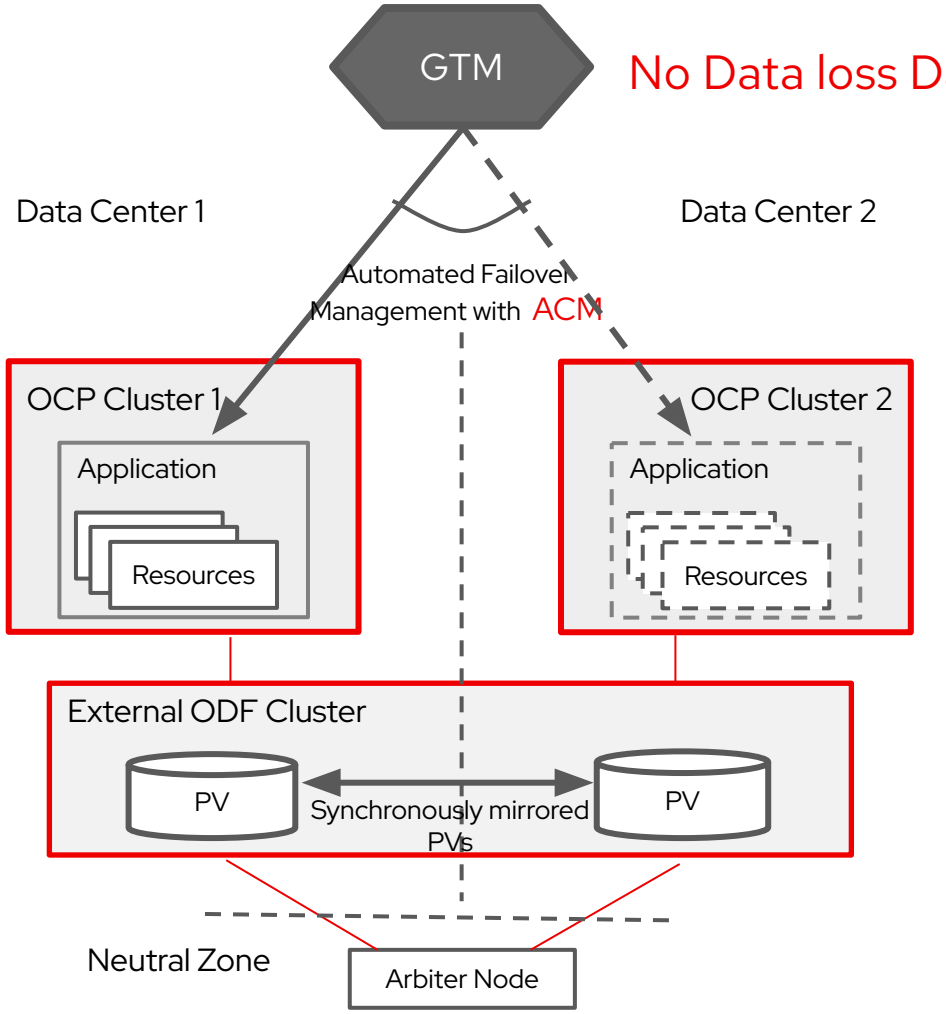
Protection against Geographic Scale Disasters



- Asynchronous Volume Replication => low RPO
  - ODF enables cross cluster replication of data volumes with replication intervals as low as 1 min
  - ODF Storage operators synchronizes both App data PVs and Cluster metadata
- Automated Failover Management => low RTO
  - ACM Multi-Cluster manager enables failover and failback automation at application granularity
- Both clusters remain active with Apps distributed and protected among them

RPO – 0
RTO – Mins

Metro-DR

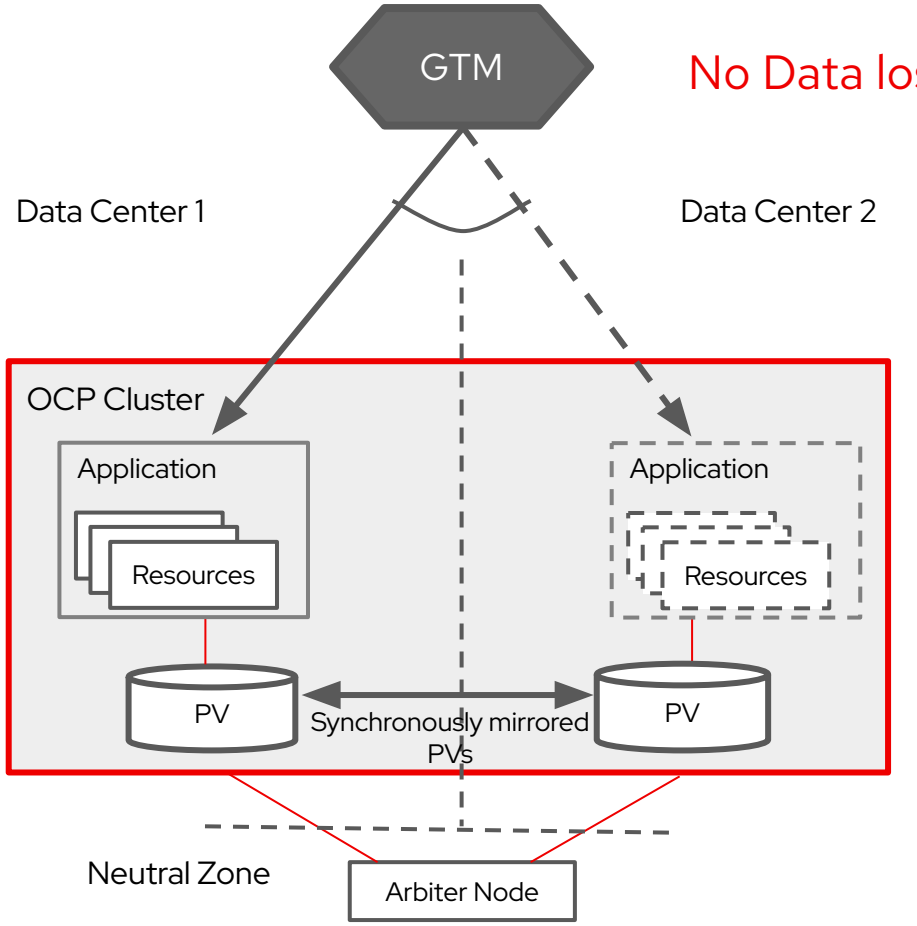


No Data loss Data Mirroring, across multiple OCP clusters

- Multiple OCP clusters deployed in different AZs provide a complete fault isolated configuration
- External Ceph storage cluster provides persistent synchronous mirrored volumes across multiple OCP clusters enabling zero RPO
- ACM managed automated Application failover across clusters reduces RTO
- Requires Arbiter node in a third site for storage cluster
    - Arbiter node can be deployed over higher latency networks provided by public clouds

RPO – 0
RTO – Secs

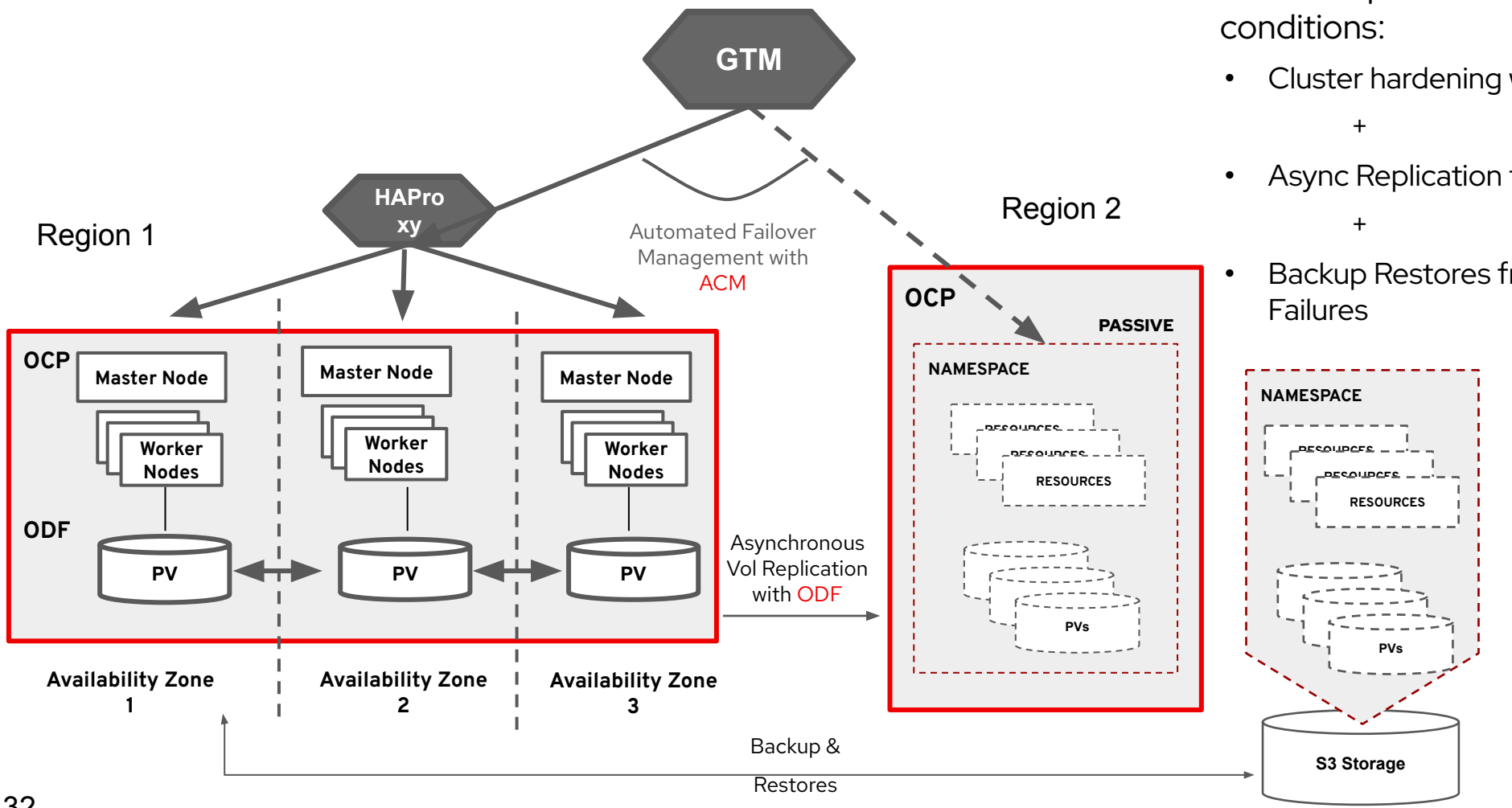
Stretch Cluster-DR



No Data loss Data Mirroring, with single stretched OCP cluster

- A single OCP and ODF Cluster is deployed in different AZs or data centers to provide a complete fault isolated configuration
- Ceph storage cluster provides persistent synchronous mirrored volumes across two data centers enabling zero RPO
- Applications recover automatically on surviving nodes in the other data center ensuring very low RTO
- Requires Arbiter node in a third site for storage cluster and OCP Master node
    - Follow RedHat guidelines for stretch OCP clusters

Comprehensive & Flexible Data Protection for the desired SLO (RPO+RTO)



Multi-tier protection against various failure conditions:

- Cluster hardening with Multi-Zone spanning OCP cluster.

  +

- Async Replication for HW and Data Center Failures

  +

- Backup Restores from Snapshots for Software & Logical Failures

Planning for OpenShift DR

DR solution Design Choices

Application RPO and RTO requirements are the primary factors guiding the DR solution choice.

The following Infrastructure and Operational factors influence DR solution choice:

Key Factors	Choices*
Distance between DR sites	• < 60 Kms - Stretch Cluster DR • < 100 Kms - Metro DR • > 100 Kms - Regional DR
Network Quality between the DR sites	• High Bandwidth, Low Latency (10ms RTT) – Metro DR • High Latency – Regional DR
Availability of DR site to run an active OCP Cluster	• Passive DR site with no active OCP cluster – OpenShift Backup • Active DR site with running OCP cluster – Metro or Regional DR
ODF is used as external mode or internal mode	• Internal Mode – Regional DR or Stretch Cluster DR • External Mode – Metro DR

* For Planning guidance only. RedHat support requirements may vary.

Metro-DR Considerations

- Metro-DR is the primary choice for no data loss (RPO=0) solution. But still has these limitations :
  - ODF external mode only – Need to be comfortable with Ceph administration of storage
  - Supported on On-Prem (Baremetal and VMWare) only – No Public Clouds.
    - Requires Arbiter on a 3rd site and can be on a public cloud (<100ms RTT latency)
  - You can choose what Apps are DR protected, but the failover granularity for MDR is per OCP cluster, unlike per-application granularity available in RDR
- Potential for Automatic Failover (RTO=0) – Roadmap
- Metro-DR avoids stretching OCP cluster across sites.
  - Stretching OCP clusters can complicate day 2 ops and can lead to cluster instability if the RedHat guidance is not followed

Stretch Cluster-DR Considerations

- Stretch Cluster DR can be considered when Metro-DR solution is not suitable
  - Supported with ODF internal mode – Hence no heavy reliance on Ceph administration of storage
  - Supported on On-Prem (Baremetal and VMWare) only – No Public Clouds.
    - Requires Arbiter node on a 3rd site with OCP Master and ODF monitors running
- ACM is not required as this is a single cluster solution
- Quick RTO with Applications are brought online quickly by OCP
- Stretching OCP cluster across sites can complicate day 2 ops and can lead to cluster instability if the RedHat guidance is not followed

Regional-DR Considerations

- Regional-DR is the most flexible and non-restrictive DR solution. Regional-DR should be considered as the first choice of DR solutions:
  - Protects against wide range of Disaster scenarios and large blast radius failures
  - With dual independent clusters offers the most robust DR solution
  - No Network restrictions, latency limits, Application performance impact …
  - Offered on ODF internal mode and will be support on external mode (Roadmap)
  - Supported on all platforms supported by ODF - both On-Prem and Public Clouds.
  - Application granular failover and DR management – does not force entire cluster failover.
- Potential data loss (RPO > 0) is the main limitation of this solution

Guidance on DR choices

- Backup vs. DR Solutions
- Both OpenShift enabled backup and DR solutions provide site DR. Key Difference: Cost and Protection levels
  - Backup solution TCO is lower as there is no standby site/cluster requirement. Backup can function with a cheaper remote Object storage.
  - Lower protection level as the data recovery starts after the failure.
- Users should be focused on Application protection and not cluster protection
  - All of our Backup and DR solutions are targeting (stateful) Application protection and not necessarily cluster protection
    - OpenShift makes cluster redeployment easy

Solution Overview

- Description
- Components
- Architecture

OCP Integrated, Full Stack DR Protection

Multi-Site Multi Cluster Manager

ACM

- ODR Hub Operator – Orchestrates & Automates DR operations across clusters

Platform

Red Hat OpenShift

- ODR Cluster Operator - Manages and synchronizes cluster meta data and application data

Persistence Layer

Red Hat OpenShift Data Foundation

- Asynchronous replication of Application volumes – or-
- Synchronous Mirroring of Application volumes

- Easy configuration DR across cluster and sites as part of application deployment
- Automated DR Failover and Failback operations reduces RTO
- Manage and Monitor DR across clusters and Apps
- Same consistent DR operations for both Metro-DR and Regional-DR
- Both Application Data and State is protected and used for Application granular protection
- Consistent Data replication or mirroring or both based on infrastructure and desired protection.

Solution Components

- ***Red Hat Advanced Cluster Management for Kubernetes (RHACM)***
  - RHACM provides the ability to manage multiple clusters and application lifecycles. Hence, it serves as a control plane in a multi-cluster environment.
  - RHACM is split into two parts:
    - RHACM Hub: components that run on the multi-cluster control plane
    - Managed clusters: components that run on the clusters that are managed
  - RHACM submariner addon: provides OpenShift private network connectivity between clusters
- ***OpenShift Data Foundation (ODF)***
  - ODF provides the ability to provision and manage storage for stateful applications in an OpenShift Container Platform cluster. It is backed by Ceph as the storage provider, whose lifecycle is managed by Rook in the OpenShift Data Foundation component stack. Ceph-CSI (Container Storage Interface) provides the provisioning and management of Persistent Volumes for stateful applications.
- ***Ceph Storage (Ceph)***
  - Ceph is a massively scalable, open, software-defined storage platform combined with a Ceph management platform, deployment utilities, and support services.
  - Metro-DR uses an external Ceph architecture called “stretch cluster mode with arbiter”.
  - ODF configured to connect to Ceph cluster that provides storage for OCP applications.

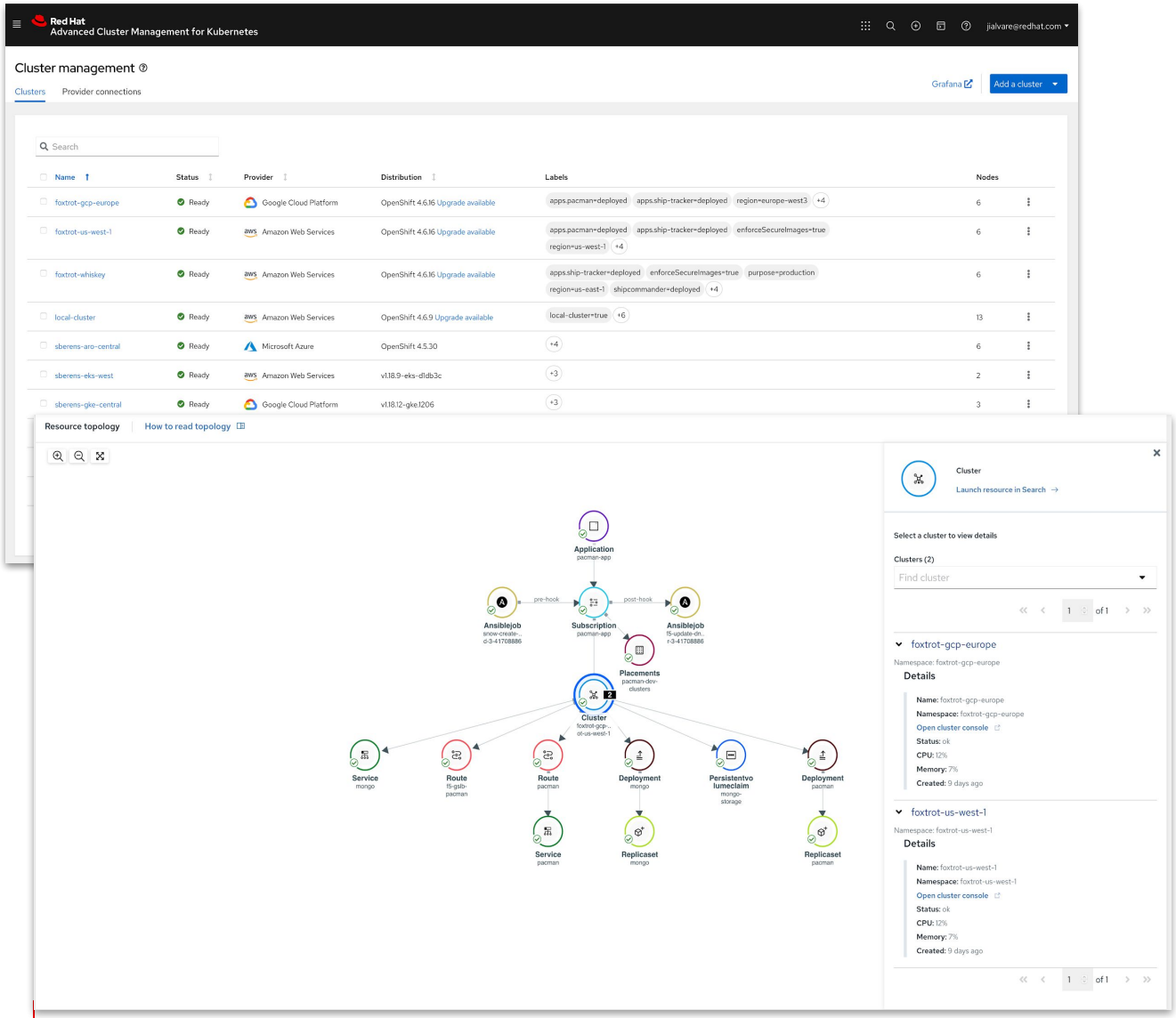
Components (continued)

- ***OpenShift DR (ODR)***
  - OpenShift DR is a disaster recovery capability for stateful applications across a set of peer OpenShift clusters which are deployed and managed using RHACM. Provides cloud-native interfaces to orchestrate the life-cycle of an application's state on Persistent Volumes. These include:
    - Protecting an application state relationship across OpenShift clusters
    - Failing over an application's state to a peer cluster
    - Relocate an application's state to the previously deployed cluster
  - OpenShift DR is comprised of three operators:
    - OpenShift DR Hub Operator
      - Installed on the hub cluster to manage failover and relocation for applications.
    - OpenShift DR Cluster Operator
      - Installed on each managed cluster to manage the lifecycle of all PVCs of an application.
    - OpenShift Data Foundation Multicluster Orchestrator
      - Installed on the hub cluster to automate numerous configuration tasks.
    - Requires the OpenShift Data Foundation advanced entitlement.

RH-ACM Based DR Automation

DR Automation enables quick and error free Application recovery enabling lowering RTO

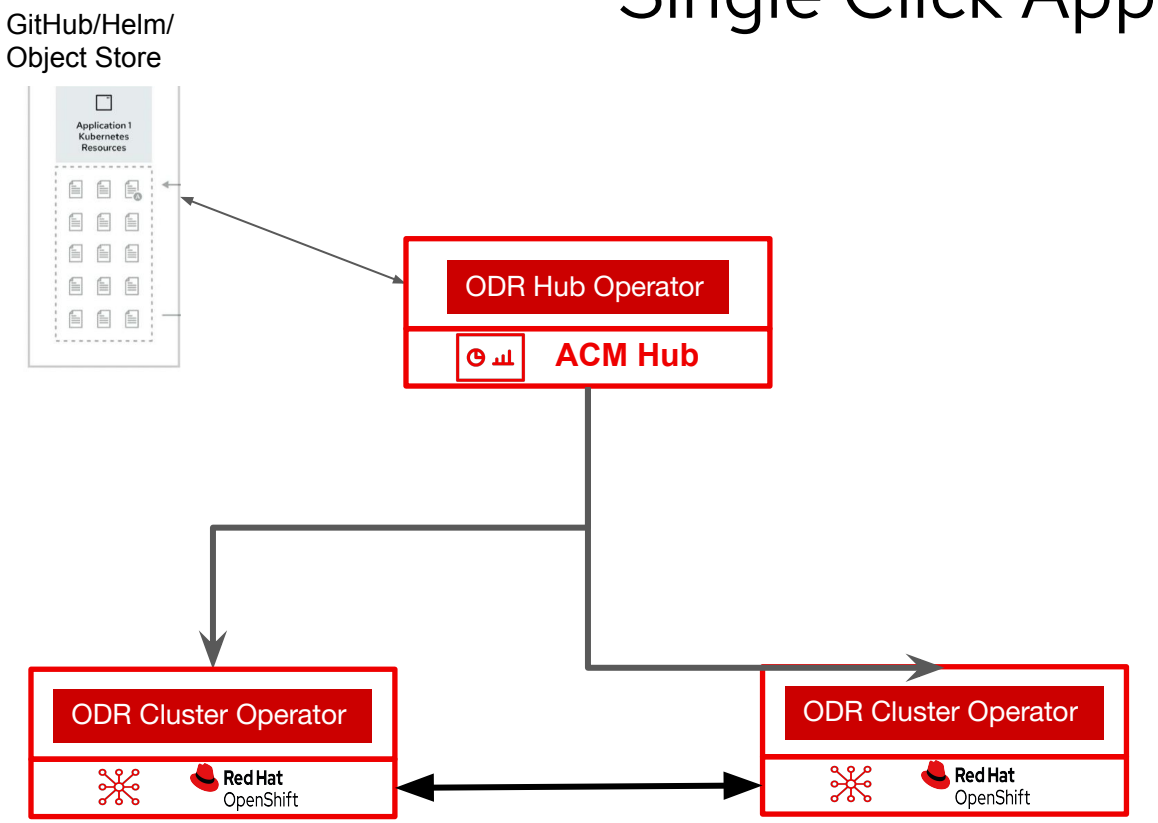
Simplified DR Management with ACM



- Centralized DR management across clusters, across sites, across cloud platforms
- Policy driven DR Configuration, governance, and compliance
- Better availability control with Application granular DR Management and operations
- Multicluster observability and alerting enables quick DR recovery and reduces downtime
- Multicluster networking simplifies cross-cluster networking for DR configuration

Operator based DR Automation

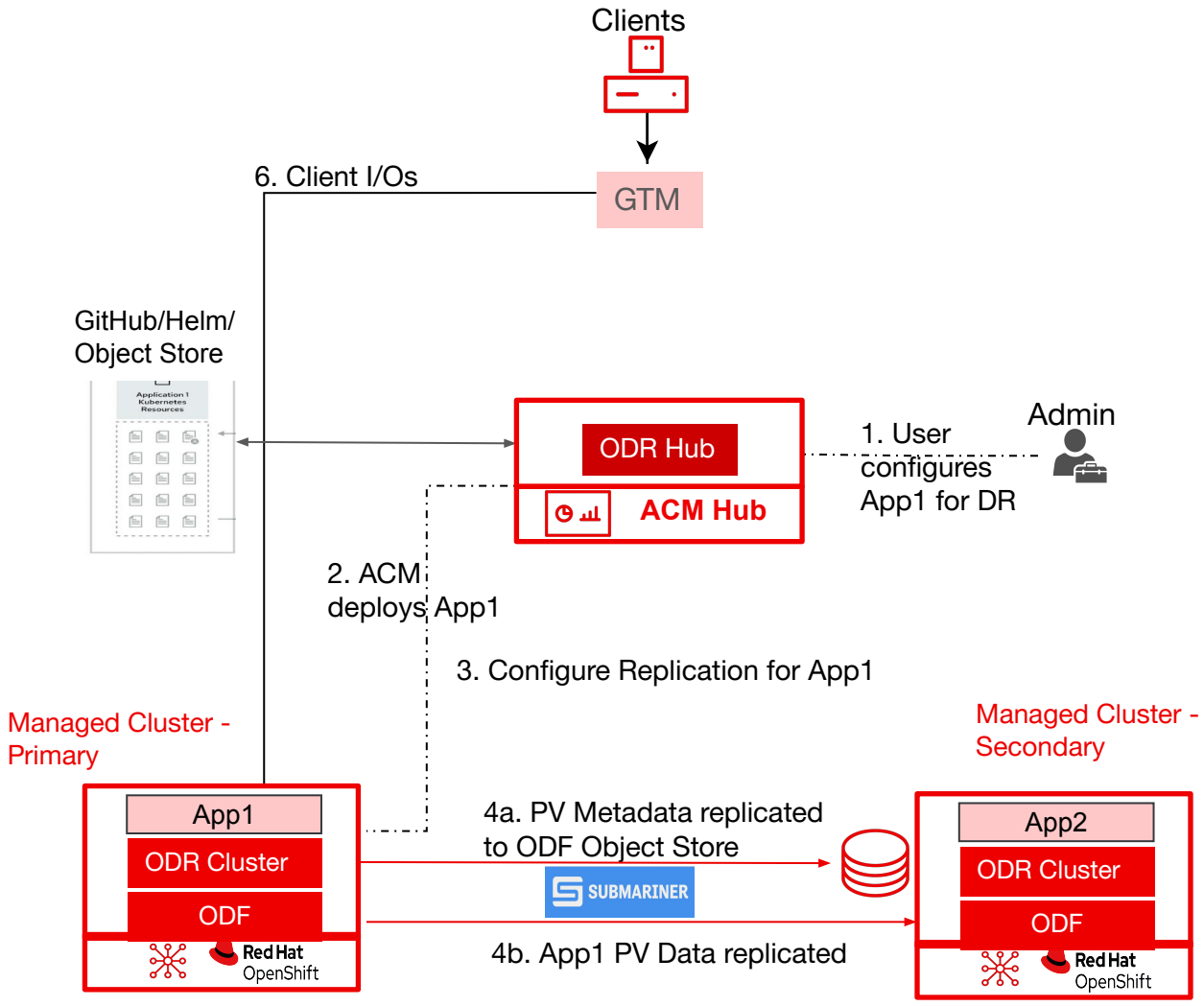
Single Click Application Recovery



A Centrally deployed ACM Hub managing DR between a pair of Managed clusters deployed at different regions

- OpenShift DR (ODR) operators reduce RTO and risk with automation of DR operations
- Simplified DR Orchestration via centrally managed ACM Hub via ODR Hub operator
- ODR Cluster Operator on each managed (workload) cluster manages data replication and synchronization
- ACM Hub recovery in case of site disasters ensures its availability

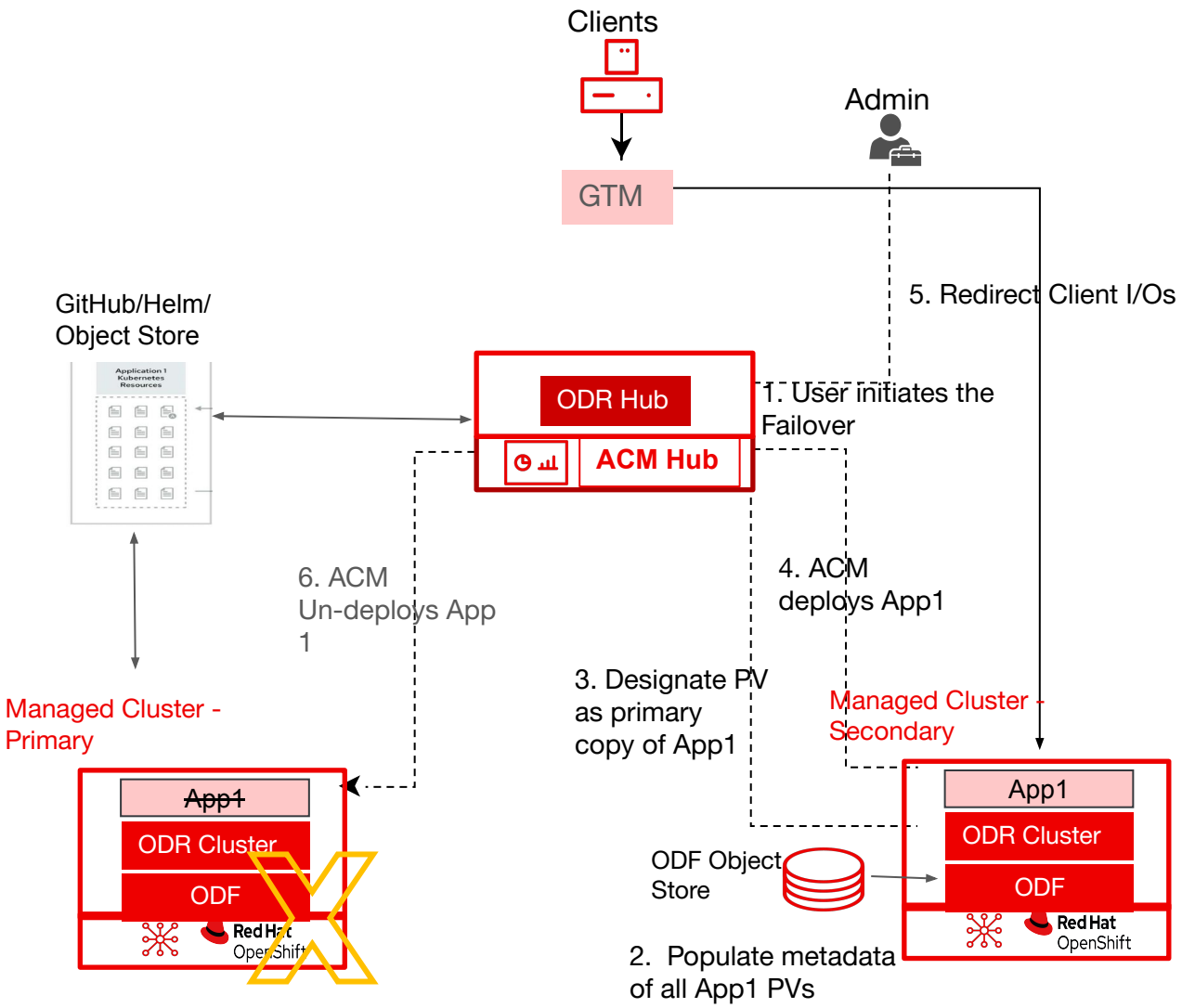
Simplified DR Orchestration



Regional-DR

- DR Orchestration in few easy steps with ODR Operators
  - Choose primary and secondary clusters for your Applications
  - Choose predefined DR Policy to apply for your application
  - ACM & ODR Operators deploy Application, configures it for DR and initiate data replication
- Both Application state and Data are replicated and protected
  - Uses ODF Object Store to capture App meta data
- Provide active-active use of both clusters, with different applications deployed at each site protected by each other
- Flexibility for replica copy to have different storage configuration than primary

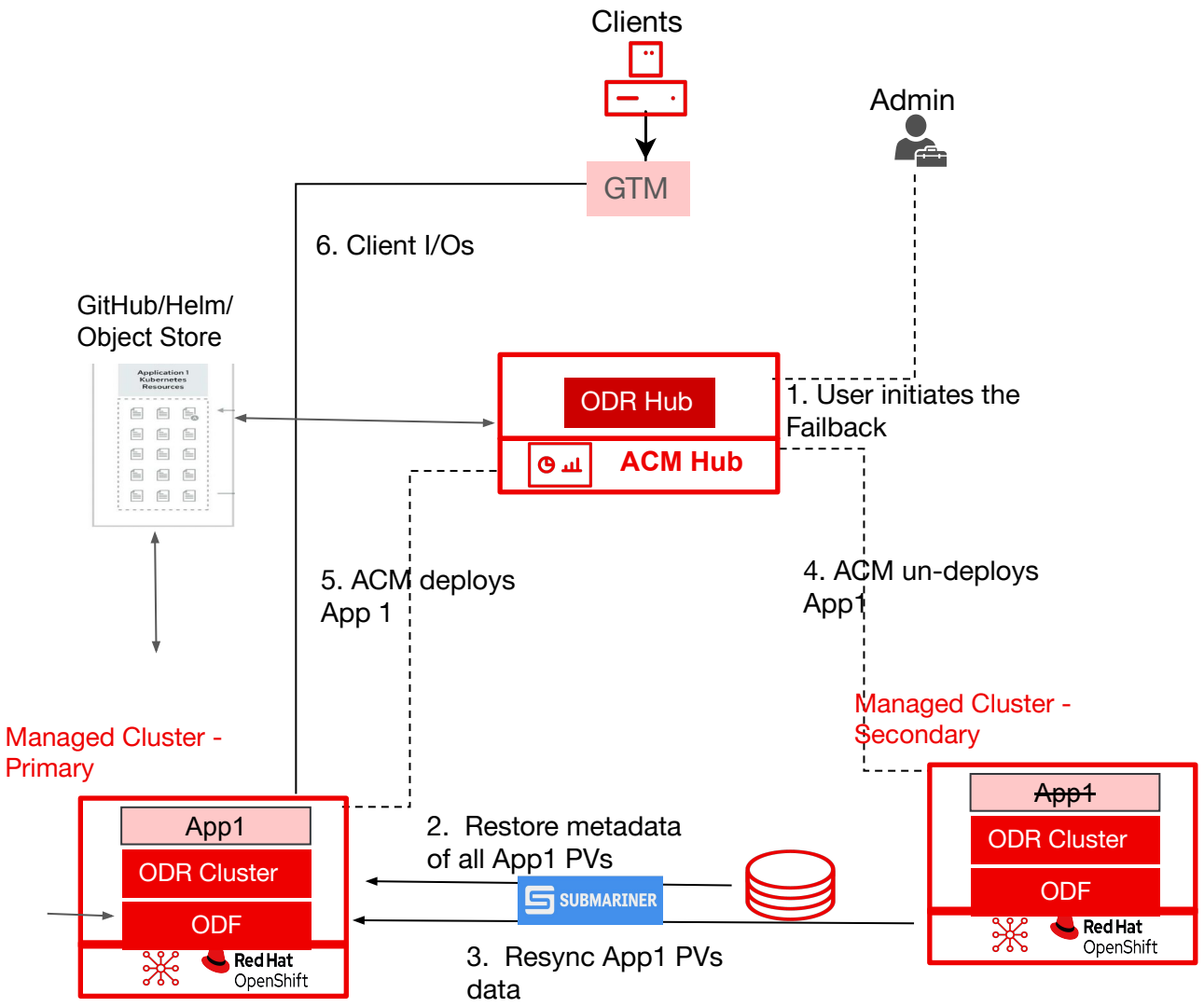
Automated DR Failover reduces RTO



Regional-DR

- ODR Operators automate the sequence of operations required for Application to recover on the secondary cluster and accomplished with a single command by the user
- Failover process automation – Increases application availability and reduces user errors
- Application(s) granular Failovers provides better priority based control
- Failovers are always user initiated and controlled, eliminates un-intended switch and data loss.

DR Failback ensures smooth recovery to Primary



Regional-DR

- DR Failbacks are planned, controlled and with no-data loss
- PV Data changes and meta data changes are restored to primary cluster before the fail-back is complete
- Fail-back operations are user initiated and controlled
- Workload migrations across hybrid platforms are also enabled with the same process

MetroDR Solution

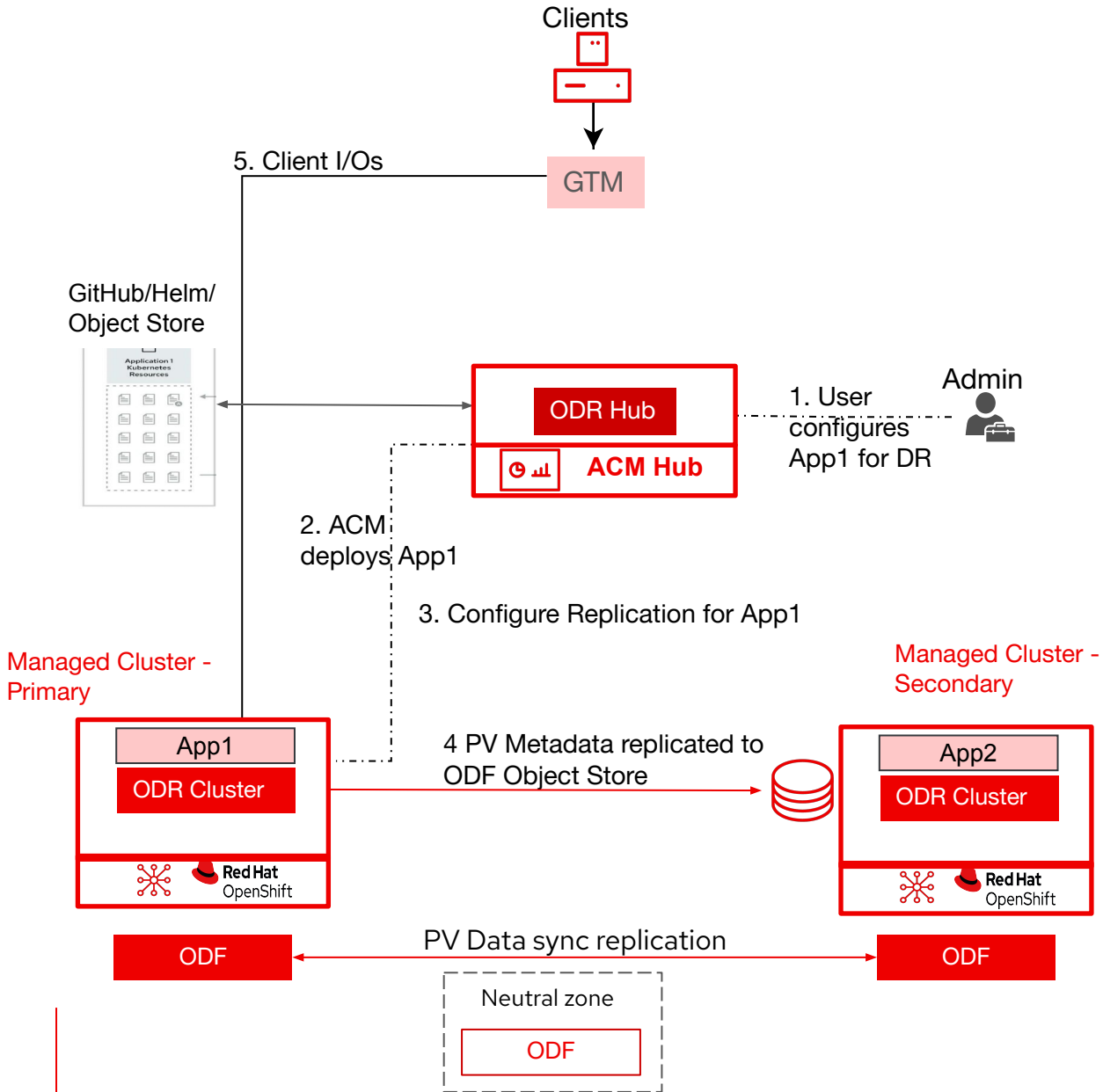
DR Automation enables

quick and error free

Application recovery

enabling lowering RTO

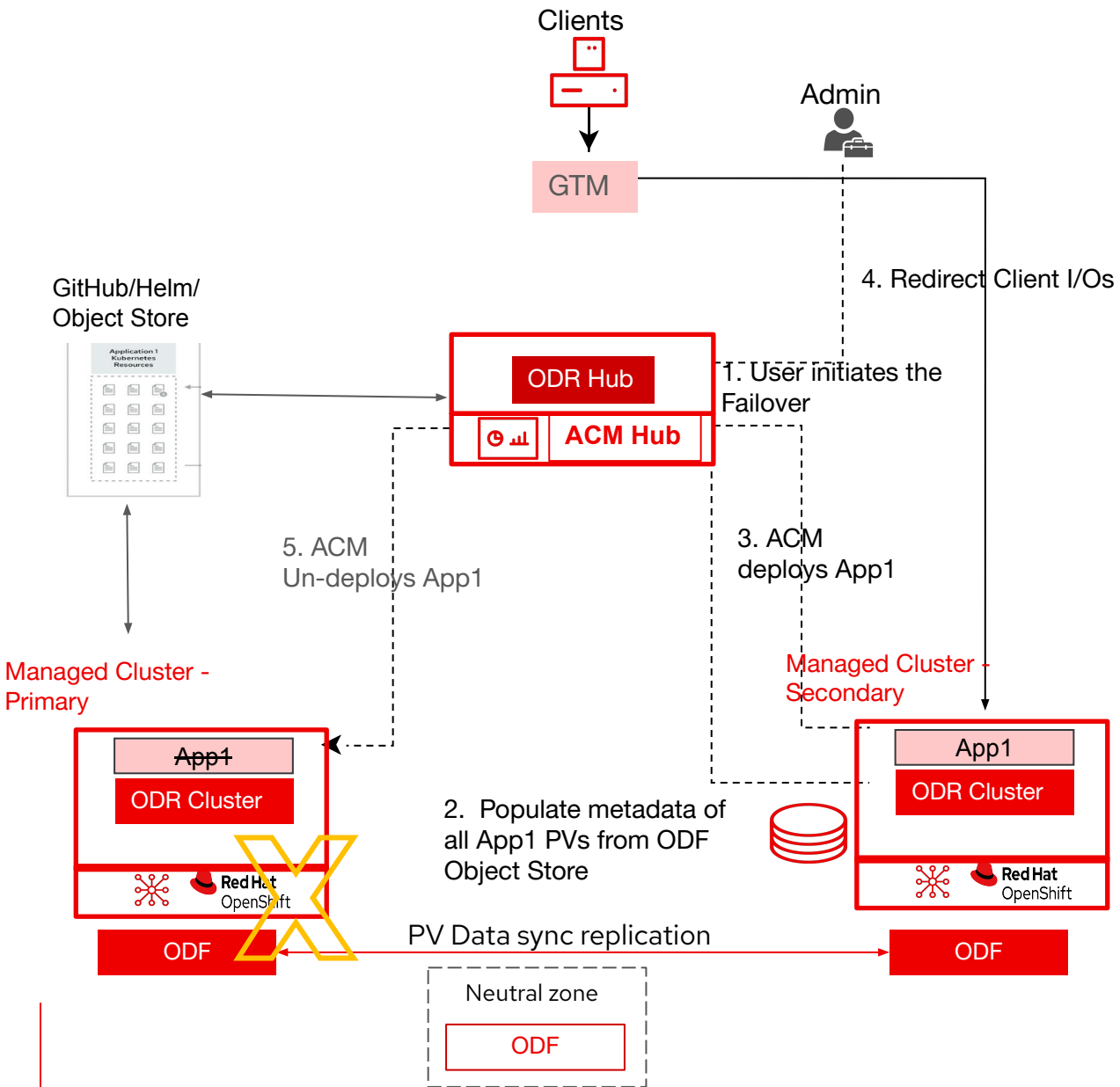
Simplified DR Orchestration



Metro-DR

- DR Orchestration in few easy steps with ODR Operators
  - Choose primary and secondary clusters for your Applications
  - Choose predefined DR Policy to apply for your application
  - ACM & ODR Operators deploy Application, configures it for DR and initiate data replication
- Both Application state and Data are replicated and protected
  - Uses ODF Object Store to capture PV/PVC App metadata
- Provide active-active use of both clusters, with different applications deployed at each site protected by each other

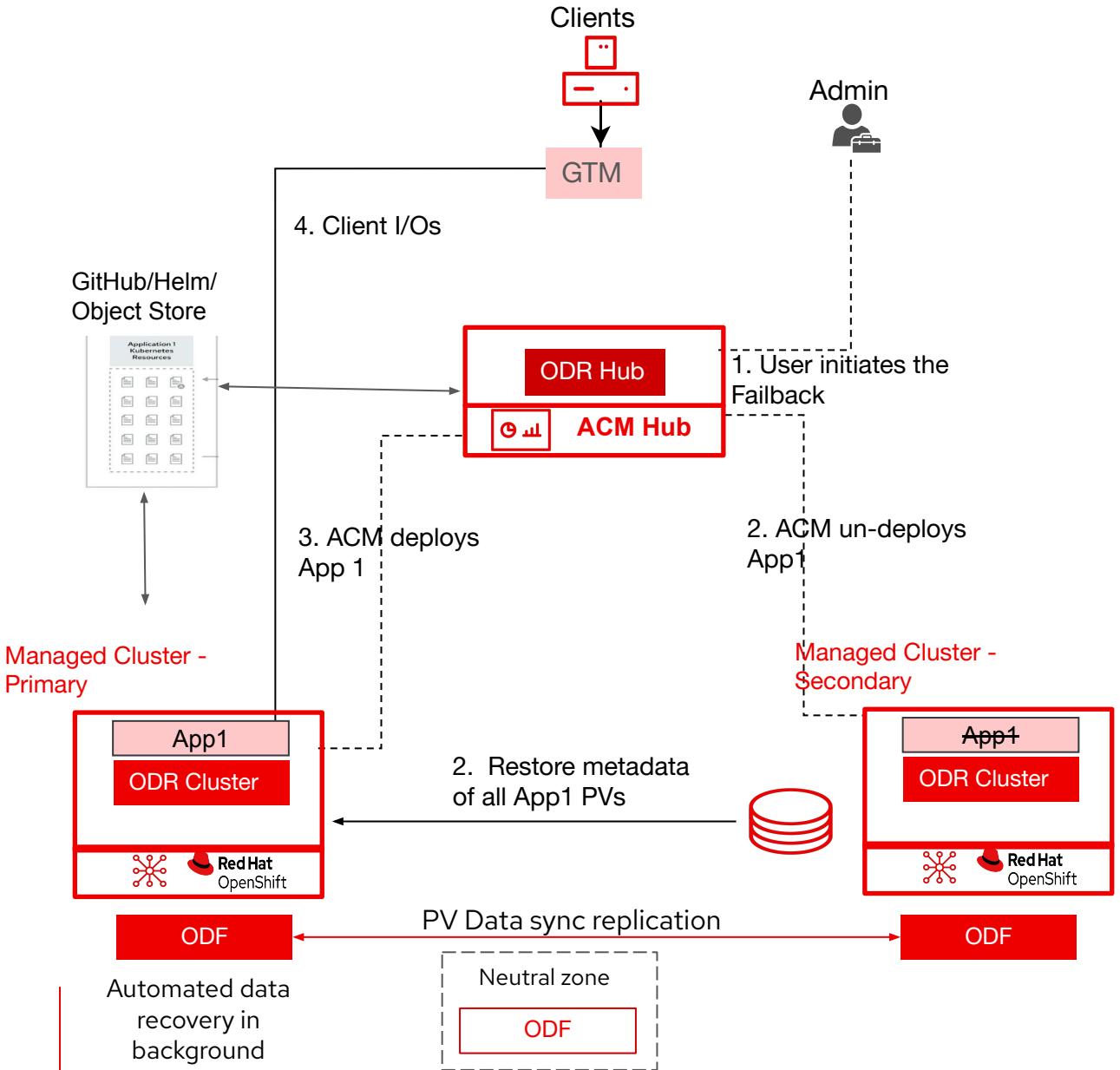
Automated DR Failover reduces RTO



Metro-DR

- ODR Operators automate the sequence of operations required for Application to recover on the secondary cluster and accomplished with a single command by the user
- Failover process automation – Increases application availability and reduces user errors
- Application(s) granular Failovers provides better priority based control
- Failovers are always user initiated and controlled, eliminates un-intended switch and data loss.
- OCP worker nodes on failed cluster are IO Fenced to prevent data corruption

DR Failback ensures smooth recovery to Primary



Metro-DR

- DR Failbacks are planned, controlled and with no-data loss
- PV Data changes and metadata changes are restored to primary cluster before the fail-back is complete
- Fail-back operations are user initiated and controlled

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