VMware vCloud® Architecture Toolkit™
for Service Providers

Architecting a VMware vCloud® Availability for vCloud Director® Solution

Version 2.9

January 2018

Tomas Fojta
Principal Architect, VCDX

VMware, Inc.
3401 Hillview Ave
Palo Alto, CA 94304
www.vmware.com

Contents

List of Tables

List of Figures

Introduction

The VMware vCloud® Availability for vCloud Director® solution extends existing hybrid cloud offerings of VMware Cloud Providers™ on top of VMware vCloud Director with disaster recovery and migration services. vCloud Availability for vCloud Director allows vCloud Director to act as cloud replication endpoint for customer on-premises or co-located VMware vSphere® infrastructures and leverages hardware independent VMware vSphere Replication™ technology.

The purpose of this document is to provide guidance on how to properly architect and design vCloud Availability for vCloud Director on top of a vCloud Director infrastructure.

Note Refer to the *Architecting a VMware vCloud Director Solution for VMware Cloud Providers* white paper[1] for vCloud Director specific architectural considerations.

[1] https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/vcat/vmware-architecting-a-vcloud-director-solution.pdf

Use Cases

The vCloud Availability for vCloud Director solution enables usage of a public vCloud Director endpoint as a target for vSphere Replication. vSphere Replication is provided by the hypervisor and supports essentially any virtualized workload without any dependence on application or operating system disaster recovery (DR) capability. vSphere 5.5, 6.0 and 6.5 environments are supported, while only vSphere 6.*x* provides replication in both directions (to and from the cloud). Only running virtual machines are replicated which means templates or powered-off VMs cannot be protected and must be replicated with a different solution (for example, VMware vCloud Connector® or VMware vCenter® Content Library synchronization).

While the main use case is disaster recovery, the solution can be also leveraged for migration services where minimal workload downtime is required.

2.1 Disaster Recovery

Description:

Customer with on-premises or co-located vSphere environment is looking for a simple DR solution that would allow the replication of virtual workloads to a multitenant public cloud service. In the case of a complete loss of the vSphere infrastructure, the tenant can power on all replicated workloads in the public cloud with minimal loss of data (RPO = 15 mins).

Typical Use Case:

1. Customer purchases cloud Organization Virtual Data Center (Org VDC) with DR option enabled. This can be augmented with additional Org VDCs that are not enabled for DR to provide capacity to run permanent workloads that are protected by solutions other than storage-level replication (for example, Active Directory servers or databases).
2. Customer sets up cloud networking infrastructure with Organization VDC networks that use the same subnets as the on-premises or co-located environment. Optionally, additional networks can be created for test failover purposes (isolated networks).
3. Customer configures a new vCloud replication endpoint in the VMware vSphere Web Client with its vCloud Director credentials.
4. Customer maps the vSphere networks in vSphere Web Client with vCloud Org VDC networks for failover and fail back purposes.
5. Customer enables protection in vSphere Web Client for each workload that must be protected. Recovery Point Objective (RPO) and number of point-in-time snapshots (vSphere 6 feature only) must be specified. RPO can vary from workload to workload. Guest quiescing and replicated data compression are supported as well.
6. Initial sync for each protected workload must be done before the workload becomes protected. This could be sped up with pre-seeding of cold clones (for example, through vCloud Connector with Offline Data Transfer).
7. Customer monitors vSphere Web Client replication dashboard to see if all protected workloads satisfy the required RPO and are in compliance.

Expected Result:

- Customer can at any time initiate in vSphere Web Client test failover for given subset of protected workloads. These will be connected to vCloud Org VDC networks designated as test networks and powered on.
- In case of loss of on premises or co-located vSphere environment, the customer can initiate failover. Replicated workloads will be connected to the vCloud Org VDC networks designated as

recovery networks. Failover can be initiated from within vSphere Web Client or alternatively via vCloud API.

- After the rebuild of the customer on-premises vSphere environment, the customer can reverse replication to prepare the workloads for a failback (only supported with vSphere 6.x on-premises).

Other comments:

- Metering for VMware licensing is based on number of protected virtual machines.
- Customer billing is based on resource usage (CPU, memory, disk space).
- The service provider can optionally offer DR as a managed service.

2.2 Migration

Description:

Customer wants to migrate a large number of workloads to the public cloud with minimum downtime.

Typical Use Case:

1. Customer requests DR option for existing or new Organization Virtual Data.
2. Customer sets up cloud networking infrastructure with Organization VDC networks that use the same subnets as the on-premises or co-located environment.
3. Customer configures the new vCloud replication endpoint in vSphere Web Client with its vCloud Director credentials.
4. Customer configures the vSphere networks in vSphere Web Client to map to the vCloud Org VDC networks for failover purposes.
5. Customer enables replication for each workload that will be migrated in the vSphere Web Client. RPO, point-in-time snapshots, and guest quiescing are not critical for the migration use case. Data compression can be optionally enabled.
6. After the initial synchronization of each protected workload has completed, the migration can be performed on each virtual machine. To secure minimal workload downtime, migration is performed in following way:
    - New delta sync is performed
    - vSphere workload is gracefully shut down.
    - Last data sync is performed
    - Workload is powered on in the cloud

Expected Result:

- Workload is migrated to the public cloud with no data loss and with minimal downtime (minutes)

Other comments:

- Metering for VMware licensing is based on number of protected virtual machines
- Customer billing is based on resource usage (CPU, memory, disk space)
- The service provider can optionally offer migration as a managed service

vCloud Availability Architecture Design Overview

3.1 vCloud Availability Architecture

VMware vSphere Replication is the underlying replication technology that is used by vCloud Availability for vCloud Director. vSphere Replication provides a hypervisor-based, asynchronous replication of virtual machines, independent from the underlying storage hardware, which is essential for replications to and from the public cloud with disparate hardware used.

The vSphere Replication Agent is built into the VMware ESXi™ hypervisor, and with the help of the vSCSI filter, tracks block changes to virtual disks (VMDKs) of running virtual machines. Depending on the required RPO, the hypervisor then transmits the changed blocks in regular intervals to the target environment.

The following figure shows the components that comprise vCloud Availability for vCloud Director.

Figure 1. vCloud Availability for vCloud Director Component Diagram

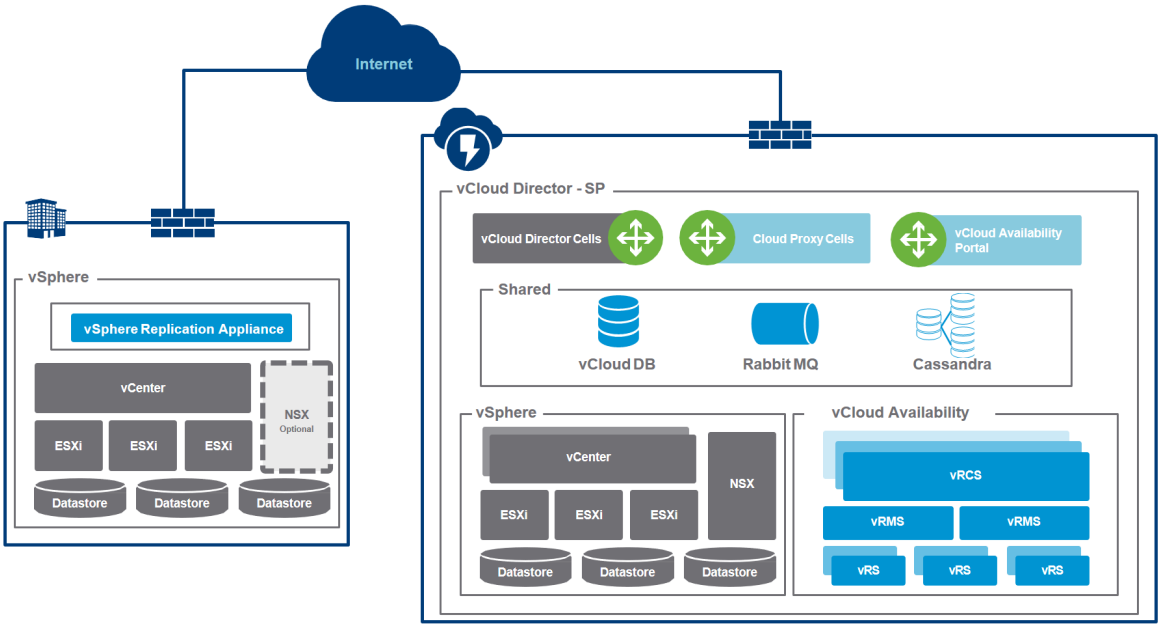


Table 1. Components

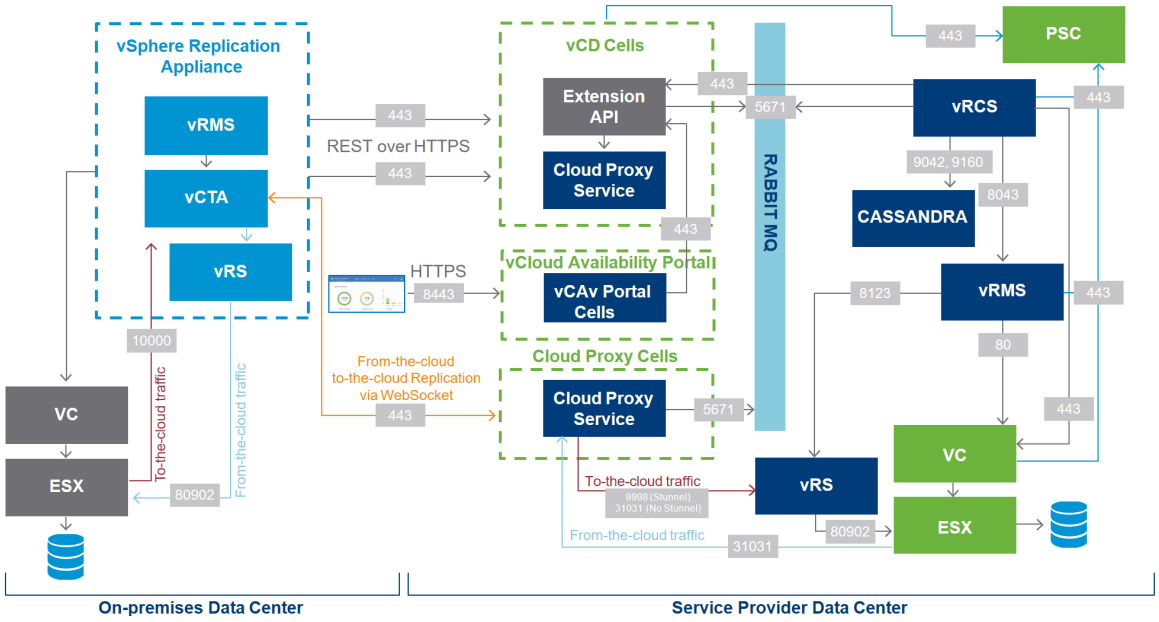
Component	Description
VMware vCloud Director	Existing vCloud Director components consisting of multiple vCloud Director cells, vCloud Director database (VCD-SP DB), and vSphere (vCenter Server, ESXi hosts, datastores) and VMware NSX® resources.
RabbitMQ	AMQP message broker that provides communication between vCloud Director and the other components that extend vCloud Director functionality.
vSphere Replication Cloud Service	Tenant-aware replication manager that extends vCloud API with vCloud Availability for vCloud Director APIs and manages replications and failovers.

Component	Description
Cassandra	Cassandra noSQL database cluster used for VM performance metrics and for replication metadata, and for configuration of vCloud Availability for vCloud Director infrastructure components.
Cloud Proxy Cells	vCloud Director cell-like components whose purpose is to tunnel replicated traffic to and from the cloud and establish control connections with on-premises vSphere Replication Appliance.
vCloud Availability Portal	Appliances that provide a tenant-level UI portal to manage replications in the cloud.
vSphere Replication Manager Server	Management component that manages the replication process of tenant virtual machines. It has a 1:1 relationship with the cloud resource VMware vCenter Server® instances.
vSphere Replication Server	Replication server that receives the replication data and records the changes for each replicated virtual machine. It has a 1:many relationship with cloud resource vCenter Server instances.
vSphere Replication Appliance	On-premises vSphere Replication component that combines both manager and replication server functionality. It also contains the vCloud Tunneling Agent that terminates the to-cloud and from-cloud replication tunnels. vSphere Replication Appliance can be extended with deployment of additional vSphere Replication Server instances for scale.
VMware Platform Services Controller™	An external Platform Services Controller runs the lookup service necessary for discovery of vCloud Availability components (Cassandra nodes, vSphere Replication Manager Server, and vSphere Replication Cloud Service) and provides secure token service to authenticate replication solution users to vCloud Director.

3.2 Network Flows

The following diagram shows network flows between all components of the vCloud Availability for vCloud Director solution.

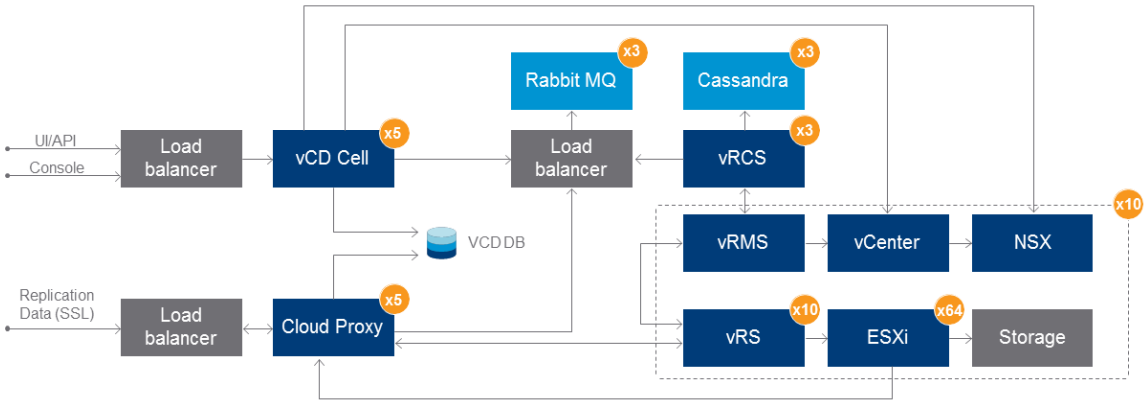
Figure 2. Network Flows



3.3 Conceptual Architecture

The following diagram shows all vCloud Director and vCloud Availability for vCloud Director components in a production environment using a highly available and scalable configuration.

Figure 3. vCloud Conceptual Architecture Overview



vCloud Availability Management Components

4.1 Logical Architecture

There are two types of vCloud Availability for vCloud Director components—vCloud Director management components that are coupled with each vCloud Director instance and vSphere management components needed for each vCloud Resource vCenter Server instance.

4.1.1 vCloud Director Management Components

- The vCloud Availability Portal is a virtual appliance that provides a graphical user interface for tenants to facilitate the management of vCloud Availability for vCloud Director operations
- Existing vCloud Director cells are extended with Cloud Proxies. These use the same Red Hat Enterprise Linux / CentOS guest OS and vCloud Director binaries as the cells, but have most of the vCloud services disabled. They still need to have access to the vCloud database as well as to the NFS transfer share. They can be load balanced with different public virtual IP addresses (VIPs) and have different SSL certificates than the vCloud cells. Deployment of Cloud Proxies scales out with the number of concurrent tunnels.
- RabbitMQ is used for vCloud API extension to provide communication between vCloud Director cells and vSphere Replication Cloud Service. For high availability, RabbitMQ usually consists of at least two clustered nodes with mirrored queues. The nodes must be behind a load balancer and presented under one virtual IP address.
- Cassandra noSQL database cluster is used for vSphere Replication Cloud Service replication metadata and configuration storage. For high availability, at least three nodes must be deployed to provide quorum consistency level, with replication factor 3. A network load balancer is not required to access the Cassandra database because vSphere Replication Cloud Service automatically chooses an available node to communicate with. The node discovery is performed through the Platform Services Controller lookup service where each Cassandra node is registered.
- vSphere Replication Cloud Service is a virtual appliance that provides the management functionality and APIs for vCloud Availability for vCloud Director. For high availability, at least two appliances must be deployed without a need for a network load balancer. The appliances use a common Cassandra database and communicate with vCloud Director through the RabbitMQ message bus. The appliances discover other service components through Platform Services Controller lookup service.
- VMware Platform Service Controller runs the lookup service that is required for discovery of vCloud Availability for vCloud Director components (Cassandra nodes, vSphere Replication Manager Server, vSphere Replication Cloud Service), and provides secure token service to authenticate replication solution users to vCloud Director. For high availability, an external Platform Services Controller is required and can be deployed in form of VMware vCenter Server Appliance with all but Platform Services Controller services disabled or as Windows installable. This Platform Services Controller can be separate from Platform Services Controller instances used by Resource Group vCenter Server instances but must be within the same vSphere domain. While a two-node load balanced Platform Services Controller set up in a single site is recommended, a single node Platform Services Controller protected with VMware vSphere High Availability might be sufficient, depending on required availability SLAs.
- vCloud Availability for vCloud Director Installer is a virtual appliance that simplifies the deployment, configuration, and management of all vCloud Availability for vCloud Director components.

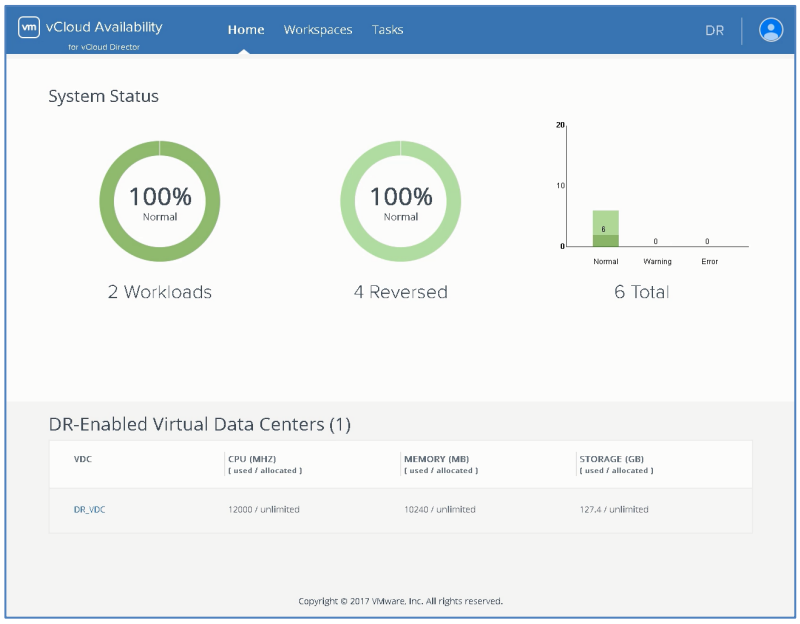
4.1.2 vSphere Management Components

- Each Resource Group vCenter Server is coupled with vSphere Replication Manager, which is provided as a virtual appliance. vSphere Replication Manager must be registered in the same vSphere domain as the vCloud Director Platform Services Controller. It is protected with vSphere HA.
- Multiple vSphere Replication servers are deployed for each vSphere Replication Manager Server. They receive the replication streams from the tunnel provided by cloud proxies and upload the data to ESXi hosts. They are deployed in scale-out fashion based on number of replicated virtual machines.

4.2 vCloud Availability Portal

The vCloud Availability for vCloud Director portal is a virtual appliance that provides a graphical user interface for tenants to manage their replications when they have no access to their on-premises vSphere Web Client user interface. The portal provides web UI on TCP port 8443 and requires a vCloud API connection.

Figure 4. vCloud Availability for vCloud Director Portal UI



The appliance is stateless, and with a network load balancer supporting sticky sessions, it can be deployed in highly available and horizontally scalable configuration. Additionally, it can be scaled up vertically in three different virtual hardware configurations depending on the number of concurrent sessions each appliance must serve.

Table 2. vCloud Availability for vCloud Director Portal Appliance Scaling

Deployment Type	vCPU	RAM	Java VM	Concurrent Sessions
Small	2	2 GB	512 MB	150
Medium	2	4 GB	1.5 GB	400

Large	4	6 GB	3 GB	800

4.3 Cloud Proxy

Cloud proxies are similar to vCloud Director cells. They are installed from the same binary on RHEL/CentOS VM using the `responses.properties` file, and they need access to the vCloud database and vCloud Director cell NFS transfer share. Cloud proxies, however, have all but one vCloud service disabled which is accomplished by editing the `global.properties` configuration file.

Figure 5. Cloud Proxy Application

```
[root@gcp-atx-sandbox-proxy1 ~]# /opt/vmware/vcloud-director/bin/cell-management-tool cell -a -i 1582
+-----------------------------------------------------------+---------+
| application                                               | state   |
+-----------------------------------------------------------+---------+
| com.vmware.vcloud.cloudproxy.server.CloudProxyApplication | STARTED |
+-----------------------------------------------------------+---------+
```

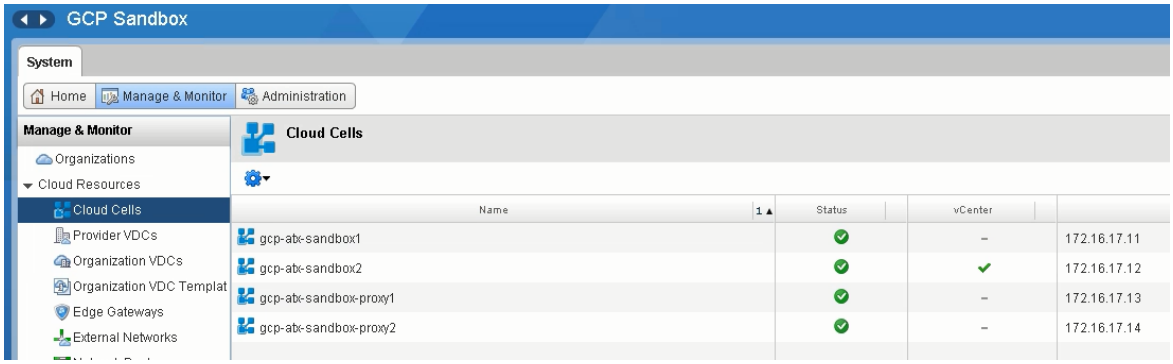
While in theory a single cell can provide both vCloud Director (UI, API, remote console) and cloud proxy services, such a setup should be used only for test and lab purposes. For the production environments, multiple cloud proxies should be deployed based on expected virtual machine replication count.

Table 3. Number of Cloud Proxies

Cloud Proxies	Replicated Virtual Machines
2	500
3	5000
5	10000

Note Although cloud proxies are displayed as regular cells in the vCloud Director Administrator UI, they do not count against vCloud Director maximums (see Appendix C – Maximums).

Figure 6. vCloud Director Cells and Cloud Proxies



Time keeping of the cloud proxies is critical for correct vCloud Availability for vCloud Director functionality. The same NTP-based internal time source must be used for the whole solution.

4.3.1 From-the-Cloud Tunnel

The from-the-cloud replication is a complex process that involves a *control connection*, initiated by the on-premises vSphere Replication vCloud Tunneling Agent, and replication data coming from the vSphere Replication agent (on the cloud ESXi host) to internally load balance cloud proxies. To properly "stitch" external and internal parts of the tunnel, the following mechanism is used:

1. The vSphere Replication vCloud Tunneling Agent opens a *control connection* (based on CloudProxyBaseUri set in vCloud API `/hybrid/settings`. (See Appendix F – Undocumented HybridSettings vCloud API.) The connection is forwarded by a cloud proxy load balancer to a cloud proxy (in Figure 7, it is Cloud Proxy 2).
2. The ESXi host starts a vSphere Replication session for a given replication *group-id*.
3. An internal load balancer forwards the TCP connection to an arbitrary cloud proxy (based on CloudProxyFromCloudTunnelHost configured in vCloud API `/hybrid/settings`). In Figure 7, it is Cloud Proxy 1.
4. The cloud proxy (through the vSphere Replication content-aware plug-in) decodes the *group-id* and requests the vSphere Replication Cloud Service to resolve it to a *destination-id* and *tenant-id.* (The cloud proxy acts as the vSphere Replication destination and "fabricates" the replication reply frames expected by the ESXi host until the real replication connection from the on-premises vSphere Replication Agent, relayed by the vSphere Replication vCloud Tunneling Agent joins.)
5. The cloud proxy broadcasts a reverse connection request over the internal message bus for the resolved tunnel (*destination-id*, *tenant-id*). It provides reverse Fully Qualified Domain Name (FQDN) based on *global.properties* override. See Figure 8.
6. Cloud Proxy 2 detects that it has a *control connection* corresponding to the given *destination-id*.
7. Cloud Proxy 2 sends a reverse connection request to the specific vSphere Replication vCloud Tunneling Agent over the corresponding *control connection* and announces the FQDN of Cloud Proxy 1 to connect.
8. The vSphere Replication vCloud Tunneling Agent opens the requested reverse connection (to the specific Cloud Proxy 1 FQDN).
9. Firewall forwards (Destination NAT) to specific Cloud Proxy 1.
10. Cloud Proxy 1 "stitches" the incoming vSphere Replication vCloud Tunneling Agent reverse connection with the pending context and will check that the corresponding stitched replication sessions are consistent with each other.

Figure 7. From-the-Cloud Tunnel Workflow

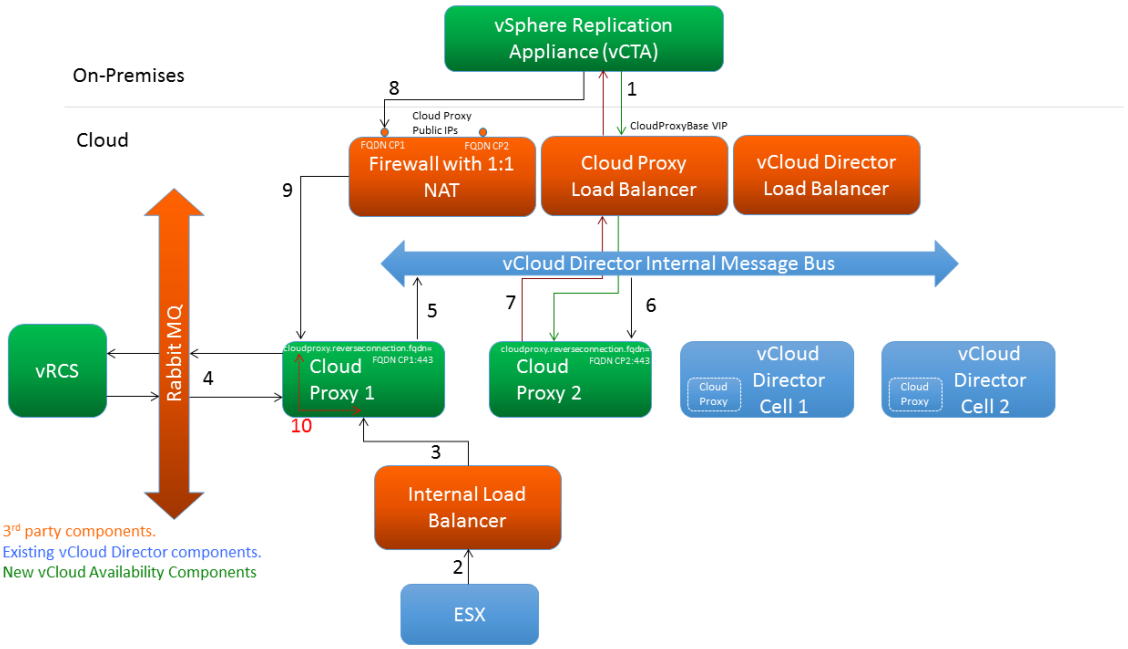


Figure 8. Cloud Proxy global.properties FQDN Override

```
[root@gcp-atx-sandbox-proxy1 ~]# cat /opt/vmware/vcloud-director/etc/global.properties
# Database connection settings
database.jdbcUrl = jdbc:jtds:sqlserver://172.16.17.10:1433/vcdmgmt;socketTimeout=90
database.username = vcdmgmt
database.password = lipGRrob77+MtxFFaBR7LQ==

# Product display name
product.display_name = VMware vCloud Director

# Maximum number blocking tasks (per cell) to be resumed when expired.
# This property is taken into account only when blocking task's default timeout action is "Resume".
# Uncomment this line and set the desired value. Otherwise, the default value (1000) will be used.
# blockingTasks.timeoutResumeRate = 1000
product.version = 8.10.0.3879706
product.build_date = 2016-05-12T20:32:07-0700
consoleproxy.host.https = 172.16.17.13
vcloud.cell.ip.primary = 172.16.17.13
consoleproxy.port.https = 8443
vcloud.http.port.standard = 80
vcloud.http.port.ssl = 443
vcloud.ssl.password = fxSk4RcI9zCdwx1g4TcxTvtPnTYDsqgF9vTP2rb2A5c=
vcloud.ssl.key.password = KqlbJHqYLNdIMBxm8RAaiSGBfk5gnPJlsy6AtkbLhh4=
vcloud.ssl.truststore.password = sw4h8FGFDVhKd66UDQ3KUsSCibbWbSx6k7oX49yyy98=
consoleproxy.keystore.password = bBf95CEeErz+tKcfbmCJLM2OoTydYXGOC7tYmV2Y+k8=
consoleproxy.keystore.path = /opt/vmware/vcloud-director/etc/proxycertificates
audit.syslog.host = 10.25.185.57
audit.syslog.port = 514
vcloud.cell.uuid = 7e7b70ee-d116-4795-8603-950cc6a54dec
vcloud.cell.ips = [10.10.13.13,172.16.17.13,127.0.0.1,fe80:0:0:0:250:56ff:fe88:aad0%eno16780032,fe80
]
system.info = rc6Px9h1IShFH8DRFZRojj/IJhJmm5ZLXHg5+i4lVek=
system.version = 1
com.vmware.cell.runtime.application = com.vmware.vcloud.cloud-proxy-server.cloudProxyApplication
user.keystore.path = /opt/vmware/vcloud-director/etc/certificates.ks
user.keystore.password = E3XqJ8+UWlV56Lw+uDQ9HQ==
cloudproxy.reverseconnection.fqdn = gcp-atx-sandbox-cloudproxy1.gcp.local:443
[root@gcp-atx-sandbox-proxy1 ~]#
```

4.3.2 Cloud Proxy Load Balancing

Configurations with multiple cloud proxies require an external network load balancer to direct traffic among the clustered proxies. The replicated traffic can come from the internet as well as internally from the cloud. Due to the amount of traffic, VMware recommends having dedicated load balancers for each replication direction.

Both the public cloud proxy VIP endpoint (URI) for to-the-cloud tunnel termination and the internal IP address VIP for from-the-cloud traffic (used by ESXi host-based replication) must be specifically configured in vCloud Director either with the vcav CLI or with a vCloud API call (see Appendix F – Undocumented HybridSettings vCloud API for more detail):

- vcav CLI, provided by the vCloud Availability for vCloud Director installer appliance:

```
vcav vcd set-cloud-proxy
\\ --to-the-cloud-address=<cloud-proxy-fqdn>
\\ --from-the-cloud-address=<from-cloud-IP>
\\ --vcd=<vcd instance>
```

- vCloud API:

```
PUT /api/admin/hybrid/settings

Headers:

Accept: application/*+xml;version=6.0
Content-Type: application/vnd.vmware.vcloud.hybridSettings+xml

Body:

<HybridSettings xmlns="http://www.vmware.com/vcloud/v1.5">
    <CloudProxyBaseUriOverride>wss://<cloud-proxy-
fqdn>:443/socket/cloudProxy</CloudProxyBaseUriOverride>
<CloudProxyFromCloudTunnelHostOverride><from-cloud-
IP></CloudProxyFromCloudTunnelHostOverride>
</HybridSettings>
```

Table 4. Example of Load Balancer Configuration

Attribute	Specification
Cloud Proxy FQDN (TCP 443)	<public VIP>
From Cloud IP (TCP 31031)	<internal VIP>
Application type	HTTPS
SSL mode	SSL passthrough
Persistence	SSL Session ID
Pool members	<CP1 internal IP> <CP2 internal IP>
Pool health check	TCP
Pool algorithm	LEASTCONN

Attribute	Specification
Pool TCP ports	443 and 31031

Table 5. From-the-Cloud Specific Cloud Proxy Configuration

Attribute	Cloud Proxy 1	Cloud Proxy 2
Public IP	<CP1 public IP>	<CP2 public IP>
Internal IP	<CP1 internal IP>	<CP2 internal IP>
Fully Qualified Domain Name	<FQDN CP1>	<FQDN CP2>
DNAT	DNAT: CP#**N** public IP:TCP 443 > CP#**N** internal IP:TCP 443	
global.properties: *cloudproxy.reverseconnection.fqdn*=	<FQDN CP1>:443	<FQDN CP2>:443

4.3.3 Cloud Proxy Certificates

A publicly trusted certificate for the fully qualified domain name of cloud proxy VIP must be imported to the cloud proxy cell with the *http* alias. Another certificate must be present with the alias *consoleproxy* even though it will be unused. This is because the cloud proxies should not be used for VM console proxy sessions. Both certificates are imported with the same process as vCloud Director cell certificates.

Because each cloud proxy is accessed from the internet under two different FQDNs (cloud proxy VIP in Table 4 and specific cloud proxy FQDN in Table 5), the *http* certificate must match both FQDNs. The easiest approach is to use a wild card certificate. If that is not possible, certificates with Subject Alternate Name (listing both FQDNs) can be used instead.

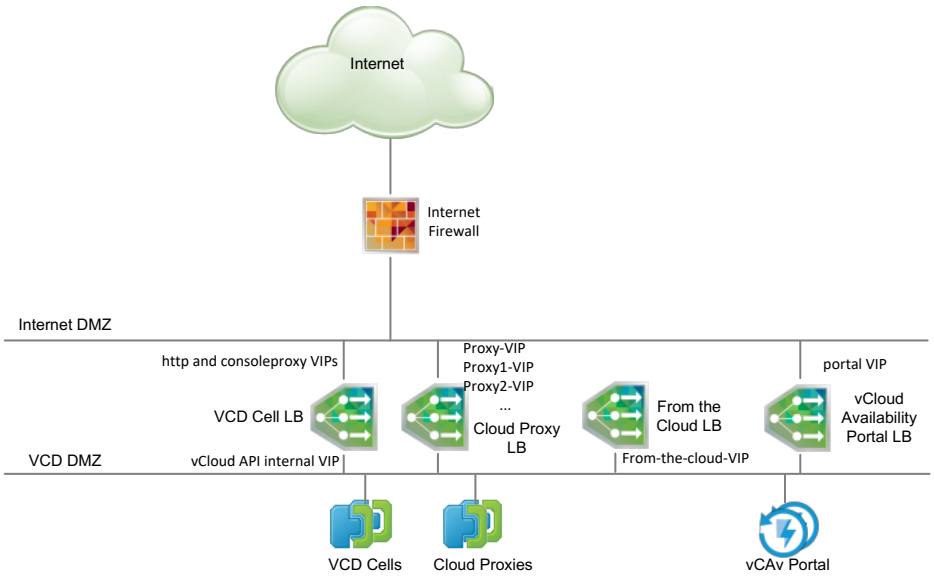
Table 6. FQDN Example

Attribute	Specification
vCloud Director UI/API	vcloud.example.com
vCloud Remote Console Proxy	console.example.com
vCloud Availability Portal	availability.example.com
Cloud Proxy VIP	vip.proxy.example.com
Cloud Proxy 1	cp1.proxy.example.com
Cloud Proxy 2	cp2.proxy.example.com
Certificate CN on Cloud Proxy Nodes	*.proxy.example.com

4.3.4 vCloud API Load Balancing

Multiple internal management components (vSphere Replication Cloud Service, Portal, Usage Meter, VMware vRealize® Orchestrator™) require access to the vCloud API. If the public vCloud API endpoint is not routable from the internal management networks, a separate internal VIP can be configured on the VCD DMZ network. Split-brain DNS or local hosts entries can then be utilized to provide correct internal VIP resolution of the vCloud API FQDN.

Figure 9. Load Balancing Design Example



4.4 RabbitMQ

vCloud extensibility provides the capability to connect vCloud Director with external systems through the AMQP message bus provided by the RabbitMQ highly available cluster.

For high availability, at least two RabbitMQ load-balanced nodes running with RabbitMQ clustering enabled and mirrored queues must be configured.

RabbitMQ must have SSL/TLS communication enabled because vSphere Replication Cloud Service uses an encrypted connection. Optionally, the vCloud Director connection can be encrypted as well.

Table 7. Example of RabbitMQ Load Balancer Configuration

Attribute	Specification
Virtual IP	…
Port	5671
Protocol	TCP
Pool Members	…

Attribute	Specification
Persistence	None
Application Type	TCP
LB Algorithm	LEASTCONN
Health check	Default TCP Monitor

The RabbitMQ SSL configuration (`/etc/rabbitmq/rabbitmq.conf`) is as follows:

```
[
  {ssl, [{versions, ['tlsv1.2', 'tlsv1.1', tlsv1]}]},
  {rabbit, [
     {ssl_listeners, [5671]},
     {ssl_options, [{cacertfile,"/etc/rabbitmq/server/cacert.pem"},
                    {certfile,"/etc/rabbitmq/server/cert.pem"},
                    {keyfile,"/etc/rabbitmq/server/key.pem"},
                    {versions, ['tlsv1.2', 'tlsv1.1', tlsv1]},
                    {ciphers,  ["ECDHE-ECDSA-AES256-GCM-SHA384",…]}
                     {verify,verify_none},
                     {fail_if_no_peer_cert,false}]}]}
].
```

All nodes must have identical certificate files: `cacert.pem` (SSL certificate of the signing Certificate Authority), `cert.pem` (RabbitMQ node certificate), and `key.pem` (private key of the RabbitMQ node certificate). The common name of `cert.pem` must be the FQDN of the load balancer VIP or a wild card certificate can be used instead.

4.5 Cassandra Database

The Cassandra noSQL database cluster is used for vSphere Replication Cloud Service replication metadata and configuration storage. vSphere Replication Cloud Service requires client to node encryption contrary to KairosDB, which is used for the optional vCloud Director VM historic performance data. Therefore, a separate Cassandra cluster must be deployed next to the VM Metric Cassandra cluster (if used). For high availability, at least three nodes must be deployed to provide quorum consistency, with a replication factor of 3. A load balancer is not required for accessing the Cassandra database, because vSphere Replication Cloud Service automatically chooses an available node to communicate with. The node discovery is performed through the Platform Services Controller lookup service where each Cassandra node must be registered.

Cassandra is available as Apache Software Foundation project or commercially by Datastax who provides supported as well as community editions.

Table 8. Example of Cassandra Configuration

Attribute	Specification
Nodes	3
Seeds	node 1 IP address, node 2 IP address
Endpoint Snitch	GossipingPropertyFileSnitch
Internode encryption	All
Client encryption	Enabled = True Optional = True
Required client authentication	False
Certificates	Self-signed certificates

4.6 VMware Platform Services Controller

vSphere Replication Cloud Service relies on the Platform Services Controller infrastructure to discover other services and for vCloud Director authentication. The following requirements must be met:

- All resource vCenter Server instances must be registered in a single Platform Services Controller vSphere domain (for example, `vsphere.local`).
- All Cassandra nodes must be registered in the Platform Services Controller lookup service.
- The vSphere Replication Manager Server must be registered in the Platform Services Controller lookup service.
- vCloud Director must be federated with the Platform Services Controller.
- The Platform Services Controller solution user for each vSphere Replication Cloud Service node is added to vCloud Director system administrator users.

As soon as more than one resource vCenter Server is deployed, you must have an external Platform Services Controller because you cannot have multiple embedded Platform Services Controller instances in a single Platform Services Controller domain (https://kb.vmware.com/kb/2108548).

An existing resource vCenter Server with an embedded Platform Services Controller can be repointed to an external Platform Services Controller if needed. To do this, use the instructions documented at http://vmw.re/emb2ext.

High availability for Platform Services Controller instances can be achieved with VMware vSphere High Availability or by deploying a load balancer in front of two nodes. For more information, see https://kb.vmware.com/kb/2113315.

The Platform Services Controller (vSphere 6.0) has the following scalability constraints:

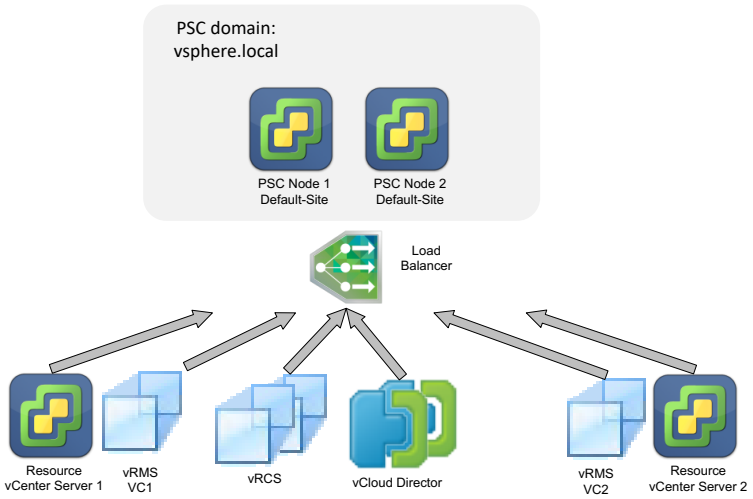
- 8 Platform Services Controller nodes per domain
- 4 vCenter Server nodes per single Platform Services Controller node
- 4 Platform Services Controller nodes per Platform Services Controller site
- 10 vCenter Server nodes per Platform Services Controller domain

To fulfill these requirements and constraints while keeping the complexity of the solution low, you can use the Platform Services Controller design options described in the following sections.

4.6.1 Single-Site

vCloud Director management components are co-located at the same site as the resource infrastructure. A load balanced, highly available pair of Platform Services Controller nodes is used for all components. If you are using Platform Services Controller 6.0, only four resource vCenter Server nodes can be used.

Figure 10. Platform Service Controller Nodes in Single-Site Configuration

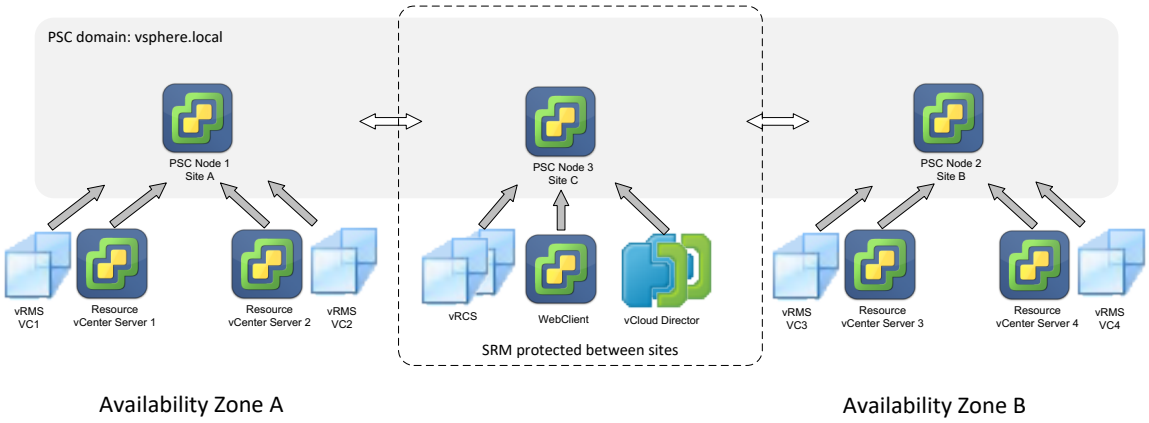


4.6.2 Multi-Site

Each availability zone (site) has a shared Platform Services Controller node with a dedicated Platform Services Controller site name (Site A, Site B). Each Platform Services Controller node can be used for up to four resource vCenter Server nodes that are installed in the same site and for related resource group management components (VMware NSX Manager™, vSphere Replication Manager).

A third Platform Services Controller node is deployed with Platform Services Controller Site C. This Platform Services Controller node is then used solely by non-resource group components, such as vCloud Director, vSphere Replication Cloud Service servers, and Cassandra., The Platform Services Controller is not site-specific—it can run on any site and fail over to another site (for example with Site Recovery Manager protection).

Figure 11. Platform Service Controller Nodes in Multi-Site Configuration



vCloud Director requires a vSphere Web Client endpoint to exist in the Platform Services Controller Site C LookupService service registration (otherwise the vCloud Director federation with Platform Services Controller node 3 fails). Therefore, an additional VMware vCenter Server Appliance™ must be deployed and registered against Platform Services Controller node 3.

Note All Platform Services Controller nodes must be in the same domain (for example, `vsphere.local`). It is also possible to deploy each node in a highly available configuration with two nodes in the same site behind a load balancer (as in the single-site case).

4.7 vSphere Replication Cloud Service

vSphere Replication Cloud Service virtual appliances provide management functionality and APIs for vCloud Availability for vCloud Director.

For high availability, a pair of vSphere Replication Cloud Service nodes must be deployed in active - active configuration without a load balancer because they share a common Cassandra database and RabbitMQ message bus for communication with vCloud Director.

Table 9. Number of vSphere Replication Cloud Service Nodes

vSphere Replication Cloud Service Nodes	Replicated Virtual Machines
1	Only for PoCs
2	0-3000
3	3000-10000

4.8 vSphere Replication Manager

There must be one vSphere Replication Manager appliance for each resource group vCenter Server. For vSphere Replication Manager Server nodes, high availability is achieved with the vSphere HA mechanism. VMware recommends deploying vSphere Replication Manager into the resource vSphere cluster managed by the vCenter Server it is associated with and not to the management cluster. This

is because the vSphere Replication Manager automatically registers itself as an extension to the vCenter Server where the appliance is deployed.

vSphere Replication Manager can use the internal vPostgreSQL database or an external Microsoft SQL or Oracle databases. All replication configurations in a paired vCenter Server are stored in the database. From a recoverability perspective, it is therefore important to provide continuous backup of the database. This is more easily accomplished with the external database because only a VMware vSphere Storage APIs – Data Protection based backup mechanism can be used for protection of vPostgreSQL.

4.9 vSphere Replication Servers

Multiple vSphere Replication Servers can be deployed for each vSphere Replication Manager Server. They receive the replication data from the tunnel provided by cloud proxies and upload them to ESXi hosts. A high-bandwidth uplink to ESXi vmknic ports is recommended. One vSphere Replication Server can be deployed for every 250 replicated virtual machines with a minimum configuration of two vSphere Replication Servers.

Table 10. Number of vSphere Replication Server Nodes

vSphere Replication Server Nodes	Replicated Virtual Machines
2	250
2	500
4	1000
…	…

Each replication is associated with one vSphere Replication Server. However, a replication can be manually moved to another vSphere Replication Server with an API call. When the vSphere Replication Server fails, recovery and reconfiguration of associated replications is not possible. Therefore, it is essential to protect all vSphere Replication Servers.

`GET /api/admin/extension/vr/vrServers` … returns vSphere Replication Server Node references (including UUID)

`GET /api/admin/extension/vr/vrs/<vrs-id>/statistics` … returns replication count, disk count and data transfer statistics.

The following API call moves a specific replication to another vSphere Replication Server:

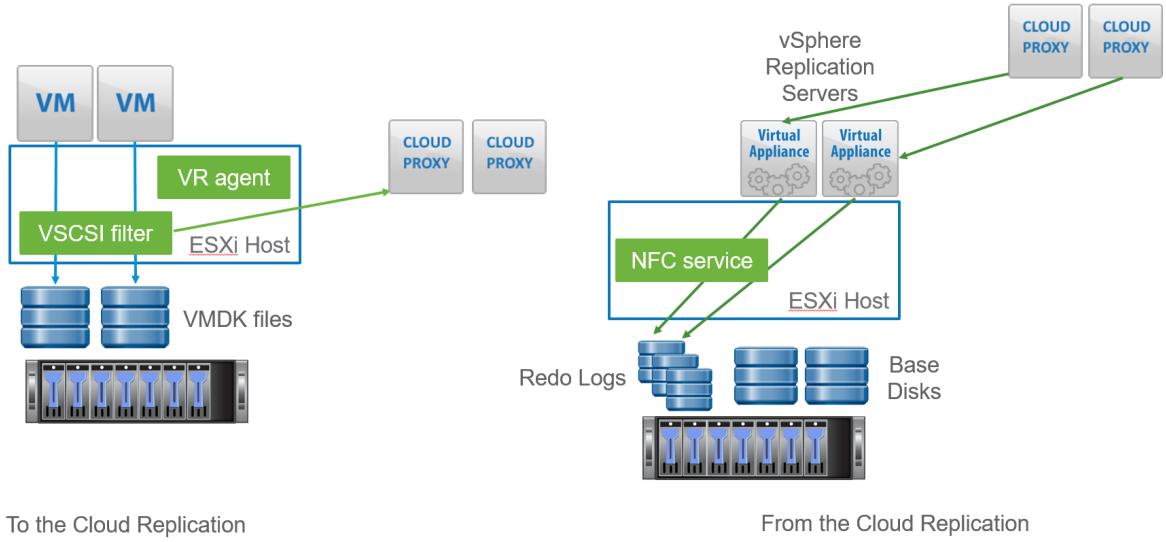
```
POST /api/vr/replications/<replication-id>/action/switchVrServer
Headers:
Accept: application/*+xml;version=20.0;vr-version=4.0
Content-Type: application/vnd.vmware.hcs.switchVrServerParams+xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:SwitchVrServerParams xmlns="http://www.vmware.com/vcloud/v1.5"
xmlns:ns2="http://www.vmware.com/vr/v6.0">
 <ns2:vrServerUuid><vrs-uuid></ns2:vrServerUuid>
```

```
</ns2:SwitchVrServerParams>
```

4.10 ESXi Hosts

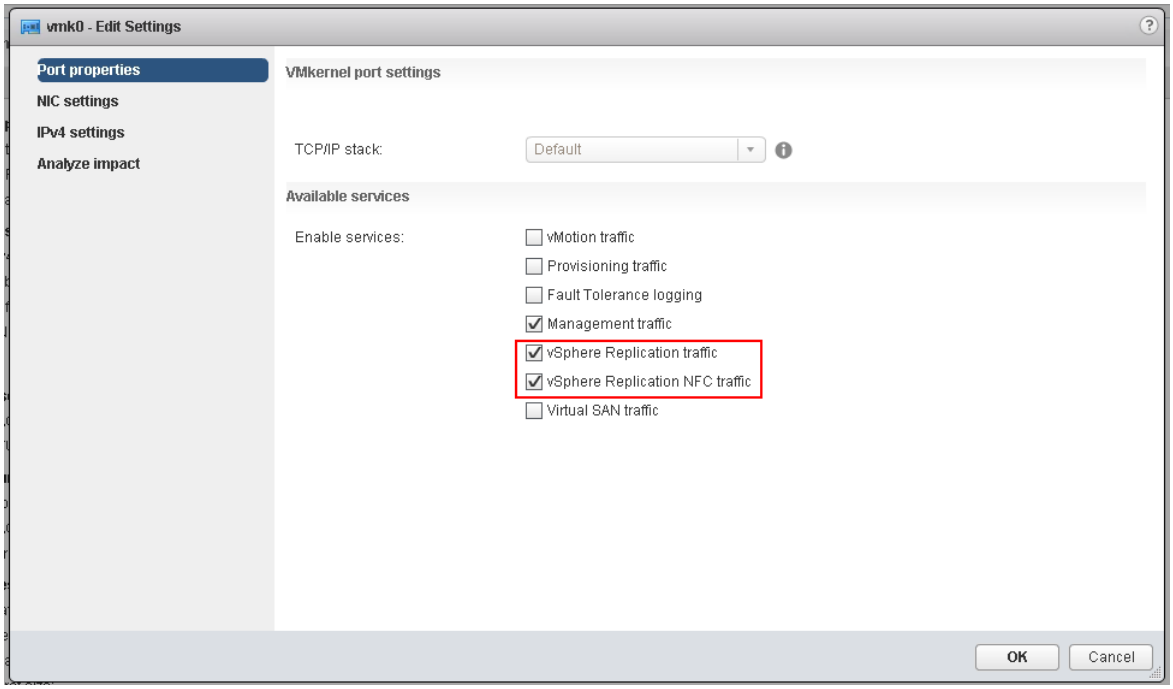
vSphere Replication Manager automatically opens necessary ports on each ESXi host to enable incoming replications from vSphere Replication Servers over NFC service and outgoing replication data from vSphere Replication Agent and vSCSI filter already present in the VMkernel to Cloud Proxies.

Figure 12. Provider vSphere Replication Traffic



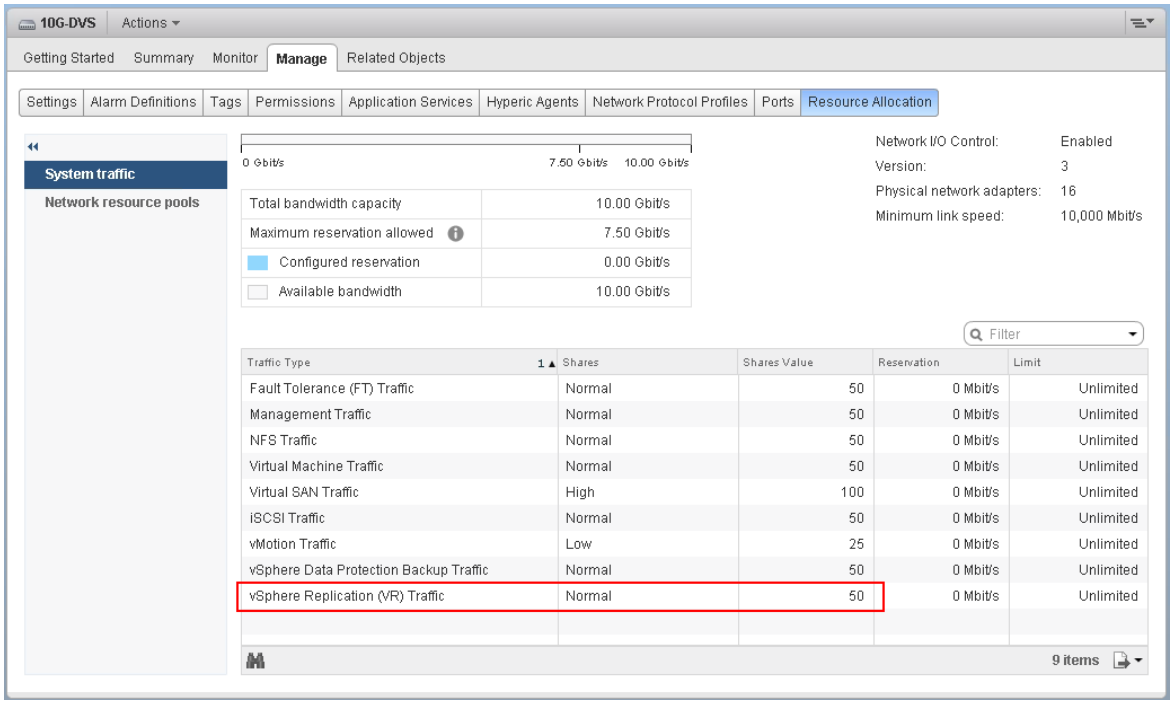
vSphere allows specifying which VMkernel ports should be used for both traffic types, so it is possible to dedicate specific network uplinks.

Figure 13. VMkernel vSphere Replication Traffic Types



In most cases vSphere Replication traffic will share the same uplink with other traffic types (Management, vSphere vMotion, and so on). In such cases, VMware vSphere Network I/O Control on VMware vSphere Distributed Switch™ can be used to define quality of service by setting shares, reservations, or limits.

Figure 14. Network I/O Control Traffic Shaping of vSphere Replication Traffic Type



4.11 vCloud Availability Metering

Currently VMware vCloud Usage Meter (3.5.x) does not meter vCloud Availability for vCloud Director consumption. While it still needs to be deployed for vCloud Director licensing, it is out of scope for this document.

Instead, a Python usage report script is provided (https://github.com/vmware/vcloud-availability-examples) that enables collection of necessary vCloud Availability for vCloud Director metrics through vCloud API.

The script can be installed on any machine that supports Python and has access to the vCloud API with provider level credentials.

The script provides point in time vCloud organization-level statistics about number of replicated virtual machines and storage capacity the virtual machines consume (including multiple point-in-time snapshots). The data can be exported to *json* or *csv* formats for additional processing.

4.12 vRealize Orchestrator

vRealize Orchestrator provides service orchestration. It can automate tasks across VMware products leveraging vCloud API, VIM API, NSX API, and so on. Using generic plug-ins (SSH, SOAP, HTTP REST, SQL, PowerShell, Active Directory) or third-party specific plug-ins, VMware vRealize® Orchestrator™ can orchestrate other systems as well.

vRealize Orchestrator has two distinct roles in vCloud Director environments:

- It acts as an extension that is subscribed to RabbitMQ and consumes vCloud Director messages. In this role, vRealize Orchestrator extends vCloud Director by providing additional services (backup, additional controls, CMDB integration, and so on).
- It acts as an orchestrator for common onboarding or tenant lifecycle tasks. The tasks are triggered by an external portal (VMware vRealize Automation™ Advanced Service Designer, or through the vRealize Orchestrator REST API). By utilizing vRealize Orchestrator plug-ins, the tenant service can be configured end-to-end.

For vCloud Availability use cases, the second role can be used to provide tenant onboarding and managing disaster recovery failover for fully provider managed services.

To provide high availability, two vRealize Orchestrator appliances should be deployed in active – standby configuration, with a network load balancer providing access to the active instance behind common virtual IP address. Both appliances must be configured exactly the same way and share a common database. Failover detection is provided through database heartbeats.

Table 11. Load Balancer Configuration Example

Attribute	Specification
Virtual IP	…
Port	8281
Protocol	HTTPS
Pool Members	vRealize Orchestrator node1, vRealize Orchestrator node2
Persistence	None
Application Type	HTTPS

Attribute	Specification
LB Algorithm	ROUND-ROBIN
Health check	Interval: 3s Timeout: 9s Max Retries: 3 Type: HTTPS Expected: 200 Method: GET URL: /vco/api/healthstatus Receive: RUNNING

4.13 Management Component Resiliency Considerations

The following table summarizes the high availability concept for each solution component that can be used to fulfill the requirements of SLAs.

Table 12. Management Component Resiliency

Management Component	HA Enabled?	Additional Availability
vCloud Director Cells	Yes	Load balancing, Active - Active
Cloud Proxies	Yes	Load balancing, Active - Active
vCloud Availability Portal	Yes	Load balancing, Active - Active
Platform Services Controller nodes	Yes	Distributed
vCloud Director Database	Yes	AlwaysOn SQL, Oracle RAC
vCloud NFS Transfer Share	Yes	Storage array provided
vSphere Replication Cloud Service	Yes	Active – Active
RabbitMQ	Yes	Application clustering, queue mirroring, load balancing
Cassandra	Yes	Application clustering, Active - Active
Resource vCenter Server nodes	Yes	
Resource NSX Manager instances	Yes	
vSphere Replication Manager instances	Yes	

Management Component	HA Enabled?	Additional Availability
vSphere Replication Manager databases	Yes	AlwaysOn SQL, Oracle RAC
vSphere Replication Servers	Yes	Active – Active (manual failover)
Active Directory servers	Yes	Application replication
vCloud Usage Meter	Yes	
vCloud Availability for vCloud Director Metering VM	Yes	
vRealize Orchestrator Server nodes	Yes	Load balancing, Active - Passive
vRealize Orchestrator database	Yes	AlwaysOn SQL, Oracle RAC
VMware vRealize Operations Manager™ servers	Yes	Application clustering

vCloud Director Configuration

vCloud Availability for vCloud Director is extending vCloud Director functionality, which is used as a platform for its services.

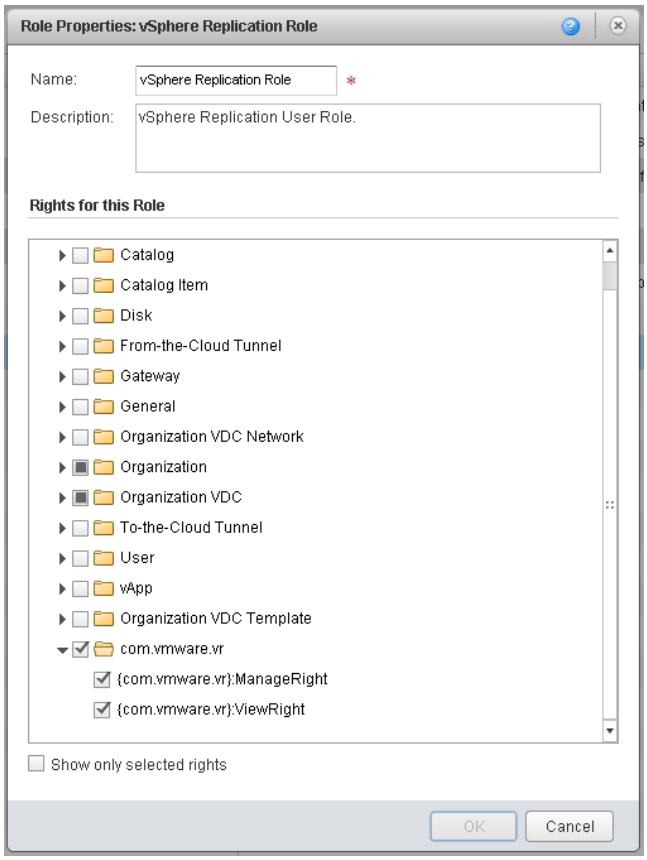
The following points summarize vCloud Availability for vCloud Director interaction with vCloud Director:

- vCloud Availability for vCloud Director registers itself as a vCloud API extension, which is provided by vSphere Replication Cloud Service appliances.
- RabbitMQ messaging bus and vCloud API are used for communication with vCloud Director.
- vSphere Replication Cloud Service uses the VMware vCenter Single-Sign On solution user to access the vCloud API as system administrator.
- Cloud Proxy is provided as part of vCloud Director cell binaries and its address is configured with vCloud API (see 4.3.1, From-the-Cloud Tunnel)
- vCloud Availability Portal relies on tenant vCloud API to collect and configure replications

5.1 User Roles

When vCloud Availability extension is registered, a new vSphere Replication role and rights are created in vCloud Director:

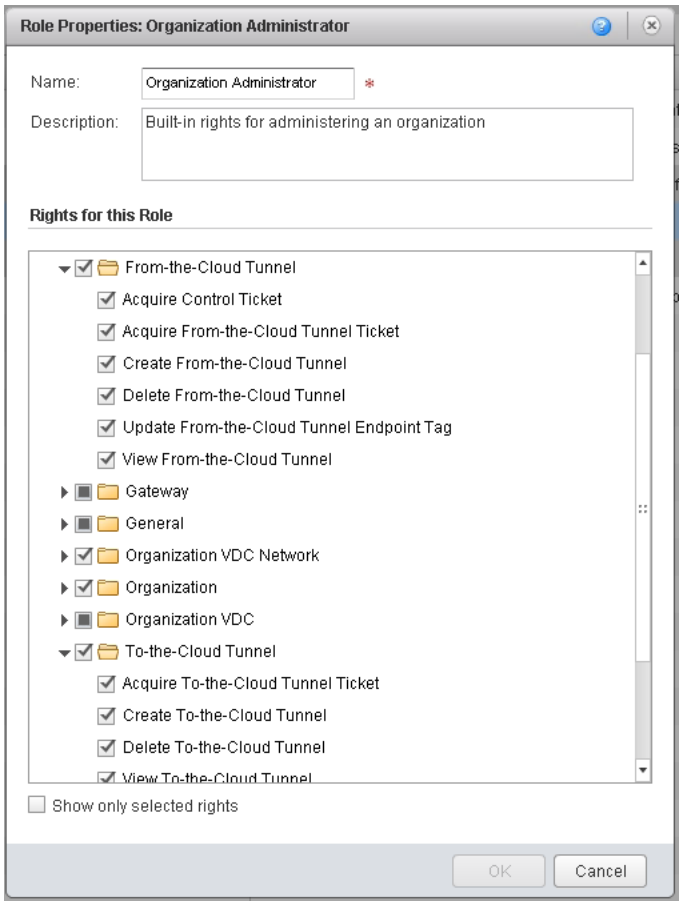
Figure 15. vCloud Director vSphere Replication Role



The Organization Administrator does not have these rights assigned by default.

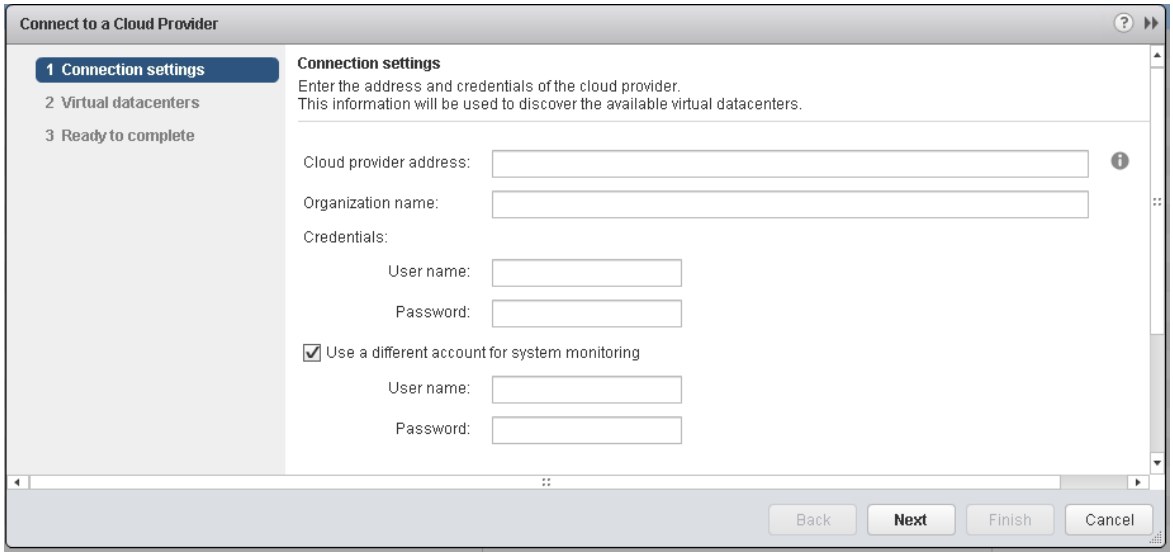
Cloud proxy-related rights (To-the-Cloud Tunnel and From-the-Cloud Tunnel) are also present in vCloud Director and are already assigned to the Organization Administrator.

Figure 16. Cloud Proxy-Related Rights



When the tenant is configuring its vCloud Availability for vCloud Director replication endpoint, he can enter two different credentials—one for administration and another for monitoring.

Figure 17. Tenant Connection Setting Dialog



To support two different roles the changes in the following table should be made to default vCloud Director roles.

Table 13. vCloud Director Role Adjustments

Role	Changes
Organization Administrator (existing role)	Add: {com.vmware.vr}.ManageRight {com.vmware.vr}:ViewRight
Replication Monitoring (new role)	Organizations – View Organization Networks Organizations – View Organizations Organization VDC – View Organization VDCs {com.vmware.vr}:ViewRight {com.vmware.vr}:ManageRight

5.1.1 vCloud Director 8.20 Changes

vCloud Director 8.20 enables creation of organization-specific roles and rights assignments.

- Service provider can still create global roles
- Service provider can selectively grant rights to specific tenants
- Organization administrators can create tenant specific roles from a subset of granted rights

The procedure to add new vCloud Availability for vCloud Director rights to a predefined Organization Administrator role is different and must be partially done with vCloud API:

1. Find new vCloud Availability for vCloud Director rights references with the following vCloud API call:

   ```
   GET /api/admin
   ```

```
Response:
</RightReferences>
…
  <RightReference href="https://.../api/admin/right/08401a7e-9898-4afe-
b07a-4e6e9a84c872" name="{com.vmware.vr}:ManageRight"
type="application/vnd.vmware.admin.right+xml"/>
  <RightReference href="https://.../api/admin/right/c6d72052-a986-4566-
9f08-4ee27938fd50" name="{com.vmware.vr}:ViewRight"
type="application/vnd.vmware.admin.right+xml"/>
</RightReferences>
```

2. For an Organization to retrieve its current rights with vCloud API GET call and from the response form a new payload for PUT call by including new right references retrieved from the step 1:

```
GET /api/<org-id>/rights
PUT /api/<org-id>/rights
```

3. In vCloud *Roles and Rights* user interface, find Organization Administrator role for the Organization used in step 2 and add the new rights. This will modify predefined Organization Administrator role which will be applied to all newly created organizations.

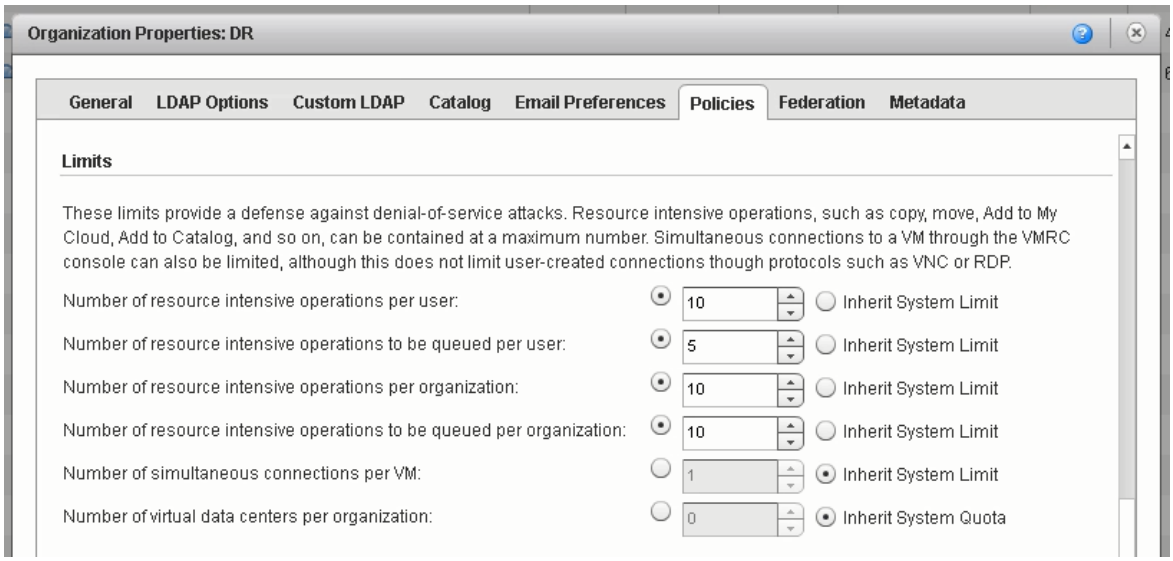
Note Existing organizations will still have the Organization Administrator role based on the unmodified predefined role, because they were not granted the new rights. Therefore step 2 will have to be repeated for all existing organizations that should have access to vCloud Availability for vCloud Director.

5.2 Tenant Limits and Leases

5.2.1 Limits

vCloud Director enables configuration of limits for resource-intensive operations that each tenant can run at the same time. This can be set at global, organization, and user level.

Figure 18. Organization Limits



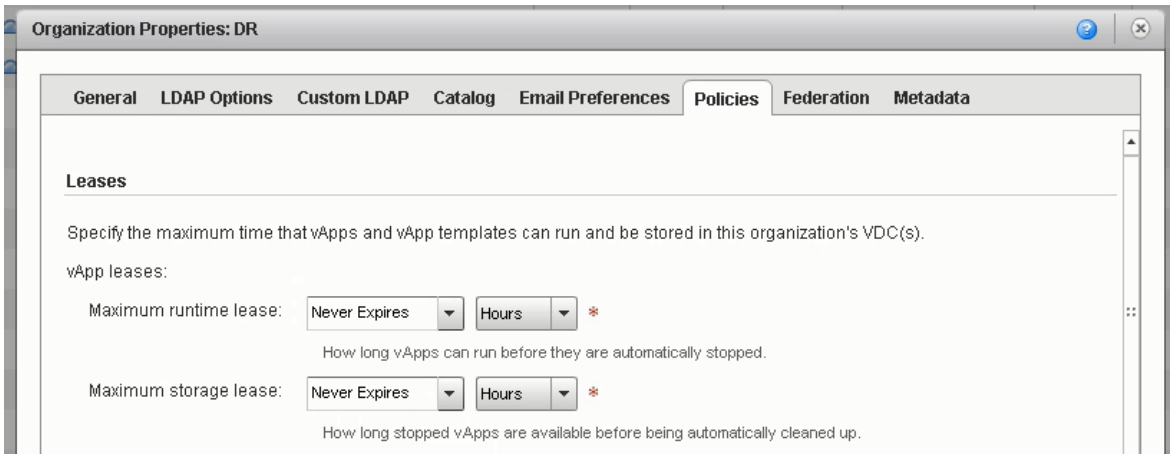
If the limit is set too low at the user level, it can negatively impact configuration of replications because these are all done under one Organization Administrator user which is defined in the VMware vSphere Web Client cloud endpoint configuration (Figure 17).

Note It might take up to 24 hours for the initial VM configuration task to expire. VMware recommends increasing these limits substantially during tenant onboarding when many VM replications are configured at once.

5.2.2 Leases

vApp runtime and storage leases must be set to "Never Expires" to avoid expiration or deletion of the placeholder VM.

Figure 19. Organization Leases



5.3 Organization Virtual Data Center

An organization virtual data center is a subgrouping of compute and storage resources allocated from a provider virtual data center and mapped to an organization. Organization virtual data centers are a deployment environment where vApps can be instantiated, deployed, and powered on.

5.3.1 Org VDC with Disaster Recovery Capabilities

vCloud Availability for vCloud Director is enabled at Organization VDC level with an API call:

`POST` /api/vdc/{orgvdc-uuid}/action/enableReplication

```
Headers:

Accept: application/*+xml;version=20.0;vr-version=4.0
Content Type: application/vnd.vmware.hcs.enableReplicationParams+xml
```

or via CLI from the installer appliance:

```
vcav org-vdc enable-replication --vcd=<vcd instance> --org=<org name> --
vdc=<org vdc name>
```

The provider can further view and limit vCloud Availability for vCloud Director features that will be available to tenants at the Org VDC limit with the following API calls:

`GET` /api/vdc/{orgvdc-uuid}/recoveryDetails

`POST` /api/admin/vdc/{orgvdc-uuid}/action/editRecoveryDetails

```
Headers:

Accept: application/*+xml;version=20.0;vr-version=4.0
Content-Type: application/vnd.vmware.hcs.editRecoveryDetailsParams+xml

Body:

EditRecoveryDetailsParams xmlns="http://www.vmware.com/vr/v6.0"
xmlns:vcloud_v1.5="http://www.vmware.com/vcloud/v1.5" name="xs:string">
    <vcloud_v1.5:VCloudExtension required="xs:boolean"/>
    <Description> xs:string </Description>
    <IsTestFailoverEnabled> xs:boolean </IsTestFailoverEnabled>
    <IsFailoverEnabled> xs:boolean </IsFailoverEnabled>
    <IsPlannedMigrationEnabled> xs:boolean </IsPlannedMigrationEnabled>
</EditRecoveryDetailsParams>
```

Table 14. Org VDC Replication Features

Element	Description
IsPlannedMigrationEnabled	Can run planned migrations? (Boolean)
IsFailoverEnabled	Can configure failover? (Boolean)
IsTestFailoverEnabled	Can configure test failover (Boolean)

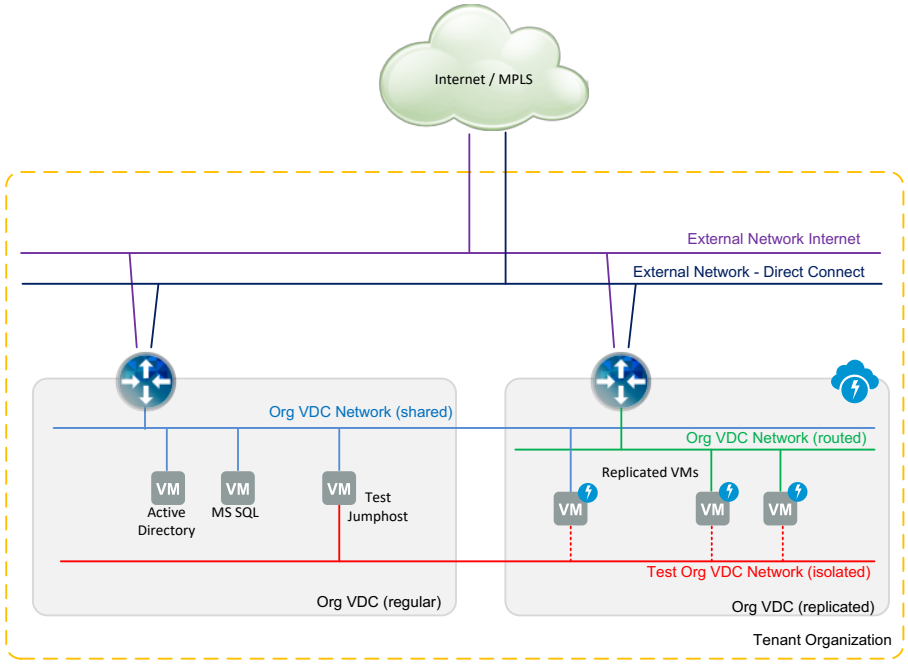
Limiting replication features at the Org VDC level might be useful in scenarios where vCloud Availability for Director is used only for onboarding of tenants' virtual machines and only Planned Migration must enabled.

Note Only Org VDCs with thin storage provisioning can be enabled for replication.

5.3.2 Org VDC Architecture

vCloud Availability for Director services can be offered on top of existing vCloud Director environments. In such cases, VMware recommends provisioning a separate Org VDC with enabled replication next to existing regular Org VDCs. Tenants can run regular workloads or workloads that use application-level replication (for example Active Directory servers or SQL AlwaysOn Database cluster nodes) in the separate Org VDC. Org VDC networks can be shared among regular and replication-enabled Org VDCs.

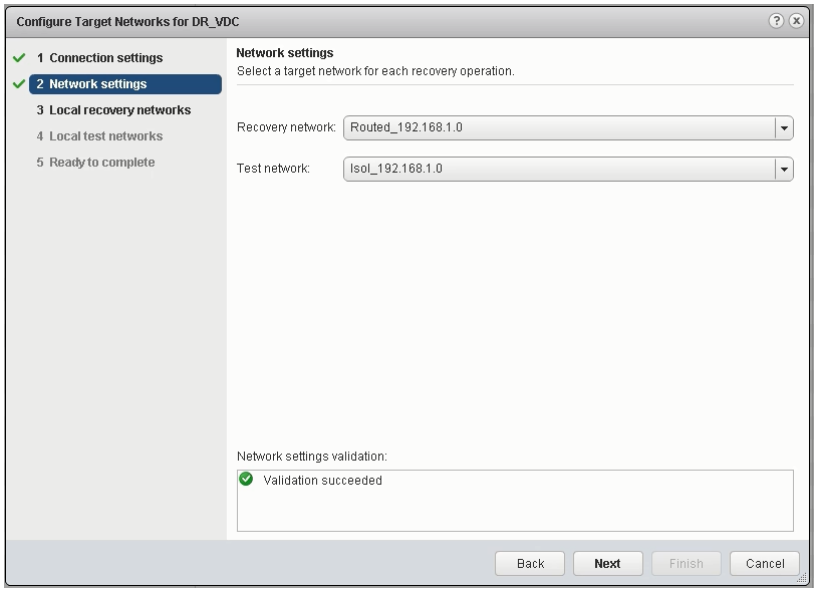
Figure 20. Organization VDC Layout (Example)



5.4 Network Management

During vSphere Replication configuration, the tenant selects one recovery network and one test Org VDC network for each Org VDC in the Target Networks Configuration dialog. All subsequently created replications will use the Recovery network setting. If the Recovery network is changed later, existing replications will not be affected. However, changing the Test network is applicable to all replications for the given Org VDC.

Figure 21. vSphere Replication Target Network Configuration



vCloud Availability for vCloud Director does not provide any mechanism for network extension from on-premises environments to the cloud. This can be achieved with Layer 2 or Layer 3 VPNs, direct connects, and similar technologies, but this is out of scope of this document.

5.5 Storage Management

Storage management is crucial for vCloud Availability for Director because storage is consumed by replicated base disks, change deltas, and point-in-time snapshots.

In vCloud Director, the tenant consumes cloud resources through Organization Virtual Data Centers (Org VDCs) that represent compute, networking, and storage resources. Storage is presented in terms of storage policies (with different performance or availability SLAs).

vCloud Availability for vCloud Director is consuming replication enabled Org VDCs, and Org VDC storage policies are presented as targets for replicated VMs. The files that represent replicated disks and multiple point-in-time (MPIT) storage snapshots of protected virtual machines, however, are not real vSphere objects, and therefore, vCloud Director has no visibility as to how much space they consume until the VM is failed over. Therefore, Org VDC storage quota management has no effect on the cold data.

Note If VMware vSAN™ is used as target storage policy, the replicated disk will be stored on *Default Virtual SAN Storage Policy* which might be different than the VM vSAN storage policy assigned to the Org VDC. In such cases, the provider needs to orchestrate remediation at the vSphere level after failover.

When a VM is configured for replication, vCloud Director creates a placeholder VM initially with full disk size, which is placed by the vCloud Director placement engine to a particular datastore and checked against Org VDC storage quota. Virtual machine disks are then replaced by 4 MB dummy disks so no storage consumption is reported to Org VDC. The virtual machine folder location is used for replication data. vSphere sees the replication files as independent data (*hbrdisk.RDID*) that is not coupled with the VM. VMware vSphere Storage vMotion® and vCloud API move replica disk call must be used to move the replication data (see Section 5.5.2, Moving Replica Disks).

VMware vSphere Storage DRS™ is supported only for placement purposes but not for data movement for load balancing or maintenance purposes.

5.5.1 MPIT Storage Snapshots

A tenant can create up to 24 multiple point-in-time storage snapshots. Currently, the service provider cannot limit this number. In theory, a storage snapshot can grow to the full size of the base disk if all disk block data has changed. This is very rare, however, and as a general practice, two to three times the size of each virtual machine disk should be reserved.

Upon failover (migration), the MPIT storage snapshots are not automatically consolidated because the process is time consuming and could impact virtual machine Recovery Time Objective (RTO). Service providers must therefore regularly consolidate these VMs because MPIT files take unaccounted for storage space and its delta disks negatively impact virtual machine storage performance. Consolidation can be scheduled automatically against the resource vCenter Server with the following VMware PowerCLI™ script:

```
Get-VM |
Where-Object {$_.Extensiondata.Runtime.ConsolidationNeeded} |
ForEach-Object {
  $_.ExtensionData.ConsolidateVMDisks()
}
```

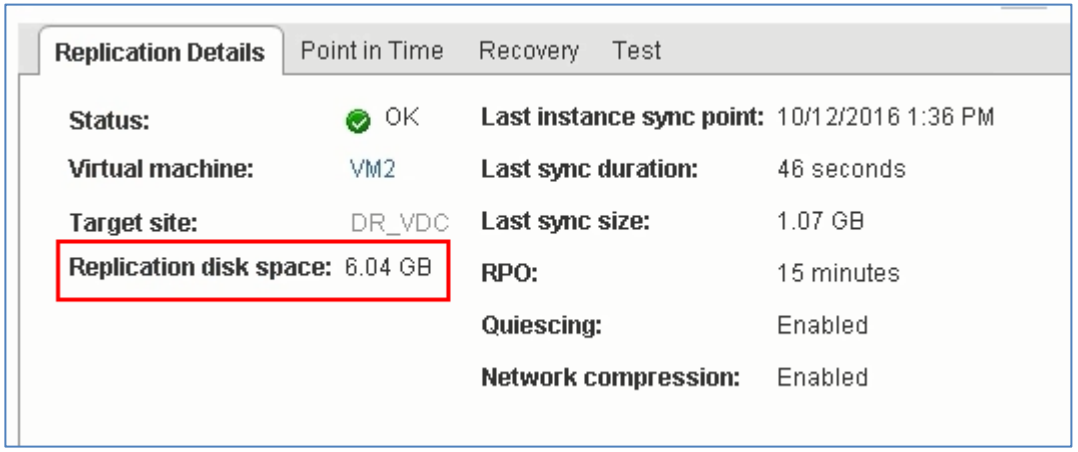
Note This consolidation does not impact VMs/disks that are in replication state.

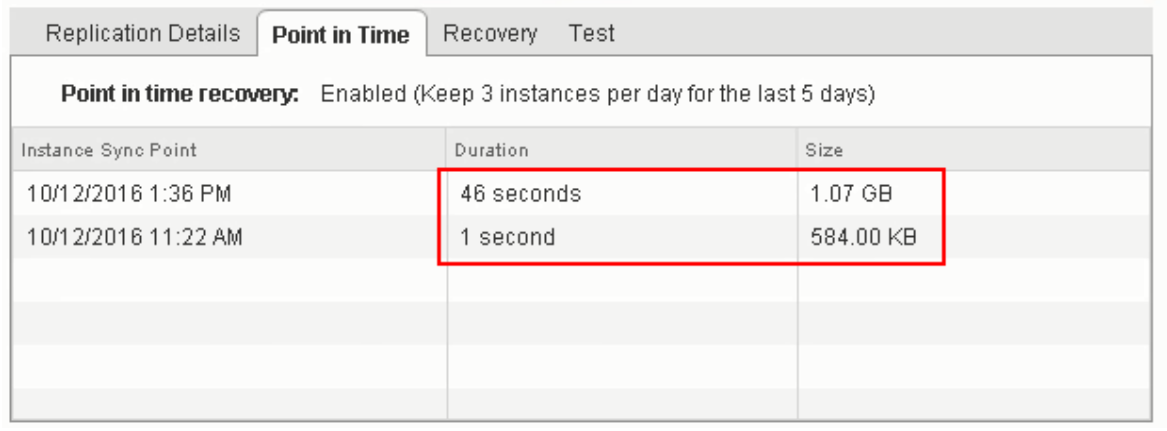
Both the provider and tenant can track disk usage of each replication and MPIT with vCloud API:

- List all replications at Org VDC level
  `GET /api/org/<org-id>/replications`
- List space requirements for replicated VM (per disk)
  `GET /api/admin/vr/replications/<replication-id>/details`
- List total space requirements and each transfer session size and duration
  `GET /api/vr/replications/<replication-id>/instances`

```
<ns2:ReplicaSpaceRequirements>6486491136</ns2:ReplicaSpaceRequirements>
    <ns2:Instance>
        <ns2:Id>GINST-Hbr.Replica.GroupInstance.000000000000014c/CGID-5345cefe-
3ad9-457a-8452-93759e8e54d6</ns2:Id>
        <ns2:TransferStartTime>2016-10-12T12:36:48.000Z</ns2:TransferStartTime>
        <ns2:TransferSeconds>46</ns2:TransferSeconds>
        <ns2:TransferBytes>1151172608</ns2:TransferBytes>
    </ns2:Instance>
    <ns2:Instance>
        <ns2:Id>GINST-Hbr.Replica.GroupInstance.000000000000013a/CGID-5345cefe-
3ad9-457a-8452-93759e8e54d6</ns2:Id>
        <ns2:TransferStartTime>2016-10-12T10:22:22.000Z</ns2:TransferStartTime>
        <ns2:TransferSeconds>1</ns2:TransferSeconds>
        <ns2:TransferBytes>598016</ns2:TransferBytes>
    </ns2:Instance>
```

Figure 22. Tenant View of Storage Consumption in vSphere Web Client





5.5.2 Moving Replica Disks

In case the service provider needs to migrate or free space on a datastore that is used for replication, the replica disks can be move with the following API call:

```
POST /api/admin/vr/replications/{id}/action/moveReplication


Accept: application/*+xml;version=6.0;vr-version=2.0

Content-Type:
application/vnd.vmware.hcs.adminMoveReplicationGroupParams+xml

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>


<MoveReplicationGroupParams xmlns="http://www.vmware.com/vr/v6.0"
xmlns:vmext="http://www.vmware.com/vcloud/extension/v1.5">


    <NewDatastore>
    <vmext:VimServerRef
    href="https://10.10.10.10/api/admin/extension/vimServer/c1abda33-81a8-
    4cf1-aa89-5af79658f84a"
    type="application/vnd.vmware.admin.vmwvirtualcenter+xml"/>
         <vmext:MoRef>datastore-26</vmext:MoRef>
         <vmext:VimObjectType>DATASTORE</vmext:VimObjectType>
    </NewDatastore>

</MoveReplicationGroupParams>
```

A virtual machine that is in test or recovery process cannot be relocated. Placeholder VM storage is moved by is vSphere Storage vMotion after its replica disks are moved.

5.5.3 Storage Management Recommendations

- Dedicate datastores or storage policies to Org VDCs used for replication.
- Use conservative datastore thresholds in vCloud Director (keep red threshold to 33%-50% of total datastore size).
- Charge tenants for cold and hot disk data.

5.6 vApps and Virtual Machines

A vApp is created in the target Org VDC for each replicated virtual machine. Inside the vApp is a placeholder virtual machine with the same compute (vCPU, RAM) size and operating system as the protected VM. The placeholder VM has certain operations disabled and cannot be moved or powered on by vCloud Director—these operations can be done only through vCloud Availability for vCloud Director APIs.

A vApp is created with the recovery network, and in the case of test failover, an additional test network is added so the recovered VM can be connected to it.

The vApp can be renamed and additional VMs and vApp networks can be added. However, there can be only one recovery VM in a vApp.

The recovery VM has guest customization disabled and its network IP Allocation Mode is set to DHCP which guarantees it will have the same IP and MAC address as the protected VM.

It is possible to change IP address of the recovered VM with guest customization but this must be orchestrated during the failover (see Section 7.4, Failover Orchestration).

Billing

The service provider can use various mechanisms to charge tenants for vCloud Availability for vCloud Director usage.

Pay per protected VM

- The simplest model aligns with VMware licensing. Protected VMs can be collected with the vCloud API and usage metering script. (See Section 4.11, vCloud Availability Metering.)

Pay per used storage

- Cold storage–used for replicated data (base disks, MPIT snapshots) can be metered with vCloud API. (See Section 5.5.1, MPIT Storage Snapshots.)
- Hot storage–used by (recovered) running VMs is reported as Org VDC allocated storage and can be metered via standard vCloud Director mechanisms.

Pay per used compute

- A replication enabled Org VDC should be set to Pay-As-You-Go allocation model. After failover, allocated compute of running VMs is reported through standard vCloud Director metering mechanisms.

Pay per replicated network traffic

- vCloud API (see Section 5.5.1, MPIT Storage Snapshots) can be used to collect how much data was transferred for each replication transfer. Note that this number does not include compression overhead and encryption efficiencies.

vRealize Orchestrator Configuration

The vSphere Replication UI is provided within the vSphere Web Client of the on-premises vCenter Server. This is the primary interface for configuration of the replications in both to-the-cloud and from-the cloud directions.

There is, however, potential need for an additional interface that can be provided by vRealize Orchestrator for two main use cases:

- Provider managed DR service
- Tenant-initiated failover when the on-premises environment (including vSphere Web Client) is down and not accessible

vSphere Replication includes the vRealize Orchestrator plug-in with out-of-the-box basic workflows for replication configuration and recovery, migration, or test actions at the VM level. These can be combined into elaborate recovery plans thanks to vRealize Orchestrator workflow and orchestration capabilities.

The placement of vRealize Orchestrator impacts which workflows can be used.

7.1 On-Premises Deployment

A tenant deploys vRealize Orchestrator within its own data center. The vSphere Replication plug-in communicates with the protected vCenter Server and the cloud provider endpoint. All plug-in workflows are available. In summary:

- Tenant self-service use case
- Simple networking setup
- Can be used as an alternative interface to vSphere Web Client UI
- Can be used to configure replications to and from the cloud
- Can be used to manage migrations and test failovers
- In case of complete loss of tenant data center, the tenant loses both the vSphere Replication user interface and vRealize Orchestrator.

7.2 In-the-Cloud Deployment

Tenant deploys vRealize Orchestrator in the provider cloud (Organization VDC). This can be simplified by the provider who can offer a pre-installed vRealize Orchestrator template with an already included vSphere Replication plug-in in the public catalog. The tenant must decide if they will connect the cloud-based vRealize Orchestrator only to cloud endpoint or to its own vCenter Server as well. Providing connectivity to the tenant's on-premises vCenter Server requires either usage of a direct connect network or an IPsec VPN.

If vRealize Orchestrator does not have connectivity to the on-premises vCenter Server, it can be used only for failover purposes during disaster recovery situation. The following workflows are then available:

- Run Real Recovery to Cloud
- Run Test Recovery at the Cloud Site
- Run Test Cleanup at the Cloud Site

Summary:

- Tenant self-service use case

- Complex networking setup if vCenter Server access is required (for full configuration and management functionality)
- Can be used during complete loss of tenant data center

7.3 Provider Deployment

The provider deploys vRealize Orchestrator in its own management infrastructure in a highly available and scalable configuration. vRealize Orchestrator is then used for tenant on-boarding activities (with vCloud Director plug-in) as well as for managed DR services for multiple tenants.

The provider vRealize Orchestrator does not have access to the tenant's on-premises vCenter Server nodes. Therefore, it can be used only for real and test recovery to the cloud.

Summary:

- Provider managed use case
- Scalable multi-tenant vRealize Orchestrator usage
- No connectivity to tenant on-premises vCenter Servers (uses *Register Standalone Org* configuration workflow)
- Used only for recovery to the cloud.

The vSphere Replication plug-in currently supports only tenant-level credentials. Therefore, the provider must have an Organization Administrator account for each managed organization and configure separate endpoints within vRealize Orchestrator.

7.4 Failover Orchestration

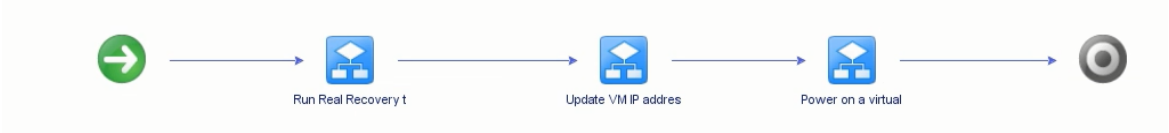
Combining the vCloud Availability for vCloud Director vRealize Orchestrator plug-in with vCloud Director (and optionally other) plug-ins allows orchestration of complex workflows that require additional tasks during VM failover.

A simple example is the need to change the IP address of the recovered VM. This can be accomplished with the following workflow:

1. Run real/test failover to the cloud without powering on the VM.
2. Enable guest customization of the recovered VM.
3. Change recovered VM IP address and IP address allocation mode to manual.
4. Power on the VM.

The vCloud Director vRealize Orchestrator plug-in currently does not have native workflow that would accomplish steps 2 and 3. Therefore, a new workflow called *Update VM IP Address* can be created.

Figure 23. Run Real Recovery to Cloud with Re-IP



The Update VM IP Address Workflow consists of two scripted tasks with *Wait for a task* workflow in between.

Figure 24. Update VM Address Workflow

Enable Guest Customization script:

Input: (vCloud:VM) **vm**
Output: (vCloud:Task) **task**

```
vm.updateInternalState();

var guestCustomizationSection = vm.getGuestCustomizationSection();
if (guestCustomizationSection.Enabled != "true") {
      guestCustomizationSection.Enabled = "true";
      System.log("Enabling guest customization");
}

task = vm.updateSection(guestCustomizationSection);
```

Update IP Address script:

Input: (vCloud:VM) **vm,** (Array/string) **ipAddress**
Output: (vCloud:Task) **task**

```
vm.updateInternalState();

var networkConnectionSection = vm.getNetworkConnectionSection();
var networkConnections =
networkConnectionSection.networkConnection.enumerate();

var poolAllocationModeNetworkConnectionIndexProp = new Properties();

var i = 0;
for each (var networkConnection in networkConnections) {

      networkConnection.IpAddress = ipAddress[i];
   if (networkConnection.ipAddressAllocationMode != "MANUAL") {
      networkConnection.ipAddressAllocationMode = "MANUAL";
   poolAllocationModeNetworkConnectionIndexProp.put(networkConnection.netwo
rkConnectionIndex,"");
   }
   i++;
}

task = vm.updateSection(networkConnectionSection);
```

Monitoring

8.1 Component Monitoring

vCloud Availability for vCloud Director components can be monitored with an API, agent-based monitoring solution (VMware vRealize Hyperic® / Endpoint Operations Management for VMware vRealize Operations™) and a syslog-based solution (VMware vRealize Log Insight™).

Table 15. Syslog Monitoring

Component	Syslog Configuration	Log Files
vSphere Replication Cloud Service	/etc/syslog-ng/syslog-ng.conf	/opt/vmware/hms/logs/hcs.log
vSphere Replication Server	/etc/syslog-ng/syslog-ng.conf	/var/log/vmware/hbrsrv.log
vSphere Replication Manager Server	/etc/syslog-ng/syslog-ng.conf	/opt/vmware/hms/logs/hms.log
Cloud Proxy	http://kb.vmware.com/kb/2004564	
Portal	/opt/vmware/conf/vcav-ui/log4j2.xmls	/opt/vmware/logs/vcav-ui/access.log /opt/vmware/logs/vcav-ui/dr2c.log /opt/vmware/logs/vcav-ui/error.log
Platform Services Controller		/var/log/vmware/sso/*
vRealize Orchestrator		/var/log/vco/app-server/server.log

Syslog-ng example for vSphere Replication Server:

Attach the following text at the end of `syslog-ng.conf`:

source *hbrsrv* {

file*("/var/log/vmware/hbrsrv.log*" follow_freq(1) flags(no-parse));

};

destination loginsight { udp("*gcp-atx-syslog.gcp.local*"); };

log { source(*hbrsrv*); destination(loginsight); };

And restart the syslog service:

> `service syslog restart`

Table 16. Monitoring via vCloud API

Component	vCloud API Call	Response
vSphere Replication Cloud Service	`GET /api/admin/vr/nodes`	`<ns2:Node id="{vrcs-node-id}"> <ns2:MaxSupportedApiVersion>4.0</ ns2 MaxSupportedApiVersion> <ns2:LastHeartbeatTimestamp>{epoc htime}</ns2:LastHeartbeatTimestam p> </ns2:Node>`
vSphere Replication Server	`GET /api/admin/extension/vr/ vrs/{vrs_id}`	`<ns2:IsConnected>true</ns2:IsConn ected>`
vSphere Replication Manager Server	`GET /api/admin/extension/ vimServer/{vim-server- id}/vrmServer`	`<ns2:IsConnected>true</ns2:IsConn ected>`

Cloud Proxy	`GET /api/query?type=cell`
	or alternatively
	`GET <proxy-ip>/ api/server_status`
		`<CellRecord … isActive="1" … name="{cell-hostname}" …/>`
		`Service is up.`

Portal	`GET <portal-ip>:8443/ api/vcd/server_status`	`Service is up.`

8.2 VM Replication Monitoring

Tenants and the provider (in managed service use case) monitor the VM replication state through the vCloud Availability for vCloud Director API.

A list of replications can be obtained either on the Organization object or the Peer (vSphere endpoint) object, which is associated with a particular Org VDC (see Appendix E – Tenant API Structure).

Detailed information about replication is provided through `GET /api/vr/replications` or `/api/vr/failbackreplications` calls.

Table 17. Replication Details

Element	Description
Name	Replication name, same as protected VM name
RPO	Recovery Point Objective (in minutes)
QuiesceGuestEnabled	Guest OS quiescing enabled? (Boolean)
NetworkCompressionEnabled	Network compression enabled? (Boolean)
PlaceholderVappId	vCloud vApp URN

Element	Description
EventPartitionId	vSphere Replication Manager Server UUID who manages replication
ReplicationState	OK or replication error
VrServerInfo	vSphere Replication Server who stages replicated data
NextInstanceSequenceNumber	Replication data sequence number
Paused	Is replication paused (Boolean)
CurrentRpoViolation	The time in minutes that has elapsed since the last RPO violation. It has a value of 0 if there is no RPO violation and a positive value otherwise.
TestRecoveryState	Unknown, none, test in progress, complete, clean up in progress, test error, cleanup error
RecoveryState	Unknown, not started, in progress, error, OK
ReplicationGroupInstance	MPIT snapshots (date, duration of transfer, amount of bytes)
LastTestFailoverOperationId	Test failover operation ID
Vm	vCloud VM URN and replication tag (original replication ID)

8.3 Backup Strategy

The backup and recovery is performed with combination of VMware vSphere Storage APIs - Data Protection compatible backup solution at VM level with database-level backups. The following table summarizes the backup strategy for each solution component:

Table 18. Backup Strategy

Component	Backup Strategy	Notes
vSphere Replication Cloud Service	VM level	stateless
vSphere Replication Server	VM level	
vSphere Replication Manager Server	VM level	+ external database backups
Cloud Proxy	VM level	stateless
vCloud Availability for vCloud Director Portal	VM Level	stateless
Platform Services Controller	VM level	http://pubs.vmware.com/vsphere-60/index.jsp?topic=%2Fcom.vmware.vsphere.install.doc%2FGUID-7FB2DA06-3737-41F0-AE53-12FBD0D1E6B7.html
vRealize Orchestrator appliance	VM level	+ external database backups
RabbitMQ	VM level	Messages are in durable HA queues
Cassandra	VM level	https://docs.datastax.com/en/cassandra/2.0/cassandra/operations/ops_backup_restore_c.html

Appendix A – Port Requirements / Firewall Rules

The following table provides the required TCP and UDP ports for vCloud Availability for vCloud Director.

Table 19. VMware vCloud Director Port Requirements

Description	Ports	Protocol
Cloud Proxy to MS SQL database	1433	TCP
Cloud Proxy to Oracle database	1521	TCP
Cloud Proxy to DNS Server	53	TCP, UDP
Cloud Proxy to RabbitMQ	5671, 5672	TCP, UDP
Cloud Proxy to/from NFS share	111,920	TCP, UDP
Cloud Proxy to Syslog Server	514	UDP
Cloud Proxy to/from vCloud Director cell / Cloud Proxies (Active MQ)	61611, 61616	TCP
Remote Manager to Cloud Proxy (JMX)	8999	TCP
Cloud Proxy to vSphere Replication	31031 9998 (stunnel)	TCP
External to Cloud Proxy	443	TCP
ESXi to Cloud Proxy	31031	TCP
Administrator to vSphere Replication Manager	5480	TCP
Administrator to vSphere Replication	5480	TCP
Platform Services Controller to Active Directory	389	TCP
vCloud Director to RabbitMQ	5671, 5672	TCP
vCloud Director to Platform Services Controller	443, 7444	TCP
vSphere Replication Manager to Platform Services Controller	443, 7444	TCP
vSphere Replication Cloud Service to Platform Services Controller	443, 7444	TCP
vSphere Replication Cloud Service to vCenter Server	80, 443	TCP
vSphere Replication Cloud Service to vCloud Director	443	TCP
vSphere Replication Cloud Service to RabbitMQ	5671	TCP

Description	Ports	Protocol
vSphere Replication Cloud Service to vSphere Replication Manager	8043	TCP
vSphere Replication Cloud Service to Cassandra	9042, 9160	TCP
vSphere Replication Manager to vCenter Server	80, 443	TCP
vSphere Replication Manager to vSphere Replication	8123	TCP
vSphere Replication to ESXi	80, 902	TCP, UDP
vCloud Availability for vCloud Director Portal to vCloud Director	443	TCP
External to vCloud Availability for vCloud Director Portal	8443	TCP

Appendix B – Glossary

Table 20. Glossary

Term	Definition
Cassandra	An open source distributed highly scalable and available database.
CBT	Changed Block Tracking mechanism vSphere Storage API - Data Protection use. Not used by vSphere Replication
Cloud proxy	Multitenant tunneling component running in the service provider environment
FastLZ	Compression library used by vSphere Replication. Provides good CPU : compression ratio
FQDN	Fully Qualified Domain Name (for example, *cloudproxy.example.com*)
HBR	Host-based replication (another name for vSphere Replication)
HCS	HBR Cloud Service (another name for vSphere Replication Cloud Service)
HMS	HBR Management Service (another name for vSphere Replication Manager Server)
hbrdisk.RDID	Redo log file in the target datastore location
MPIT	Multiple point-in-time (snapshots)
NFC	Network File Copy. Protocol used to move replicated data from a vSphere Replication Server to a vSphere datastore
RabbitMQ	An open source Message Queue Service used to route messages between vCloud Director and vCloud Availability for vCloud Director components
Replication group	Group of disks that belong to the same virtual machine and are replicated together
RPO	Recovery Point Objective. Policy that defines the maximum tolerable amount of data loss measured as a period of time
RTO	Recovery Time Objective. Policy that defines the maximum tolerable downtime of an application
vCloud Tunneling Agent	Software component running as part of vSphere Replication Manager Server on premises. Terminates to and from the cloud tunnels with Cloud Proxy

Term	Definition
vSphere Replication	Hypervisor-based replication technology used by vCloud Availability for vCloud Director
vSphere Replication Agent	vSphere component that manages replication at ESXi host level
vSphere Replication Cloud Service	Provides vSphere Replication APIs through vCloud extension and multitenancy
vSphere Replication Manager Server	Appliance that manages the environment and receives replicated data
vRealize Orchestrator	Extensible workflow solution used for orchestrating complex disaster recovery procedures
vSphere Replication Server	Appliance that receives replicated data
vSCSI Filter	vSphere component that capture writes for replication
VSS	Microsoft Volume Shadow Copy Service. Feature of Microsoft products enabling app-level and file-level consistency during backup or replication

Appendix C – Maximums

Table 21. vCloud Availability Maximums

Constraint	Limit	Explanation
Resource vCenter Server Instances	10	
vSphere Replication Cloud Service	3	
vSphere Replication Manager Servers	10	One per each resource vCenter Server
vSphere Replication Servers per vSphere Replication Manager Server	10	
Total vSphere Replication Server Nodes	100	
Replications per vCenter Server	2000	
Replication per vSphere Replication Server	200	
Multiple Point-in-Time Snapshots	24	
Recovery Point Objective	15 min – 24 hours	
Recovery Point Objective (VSAN-VSAN)	5 min – 24 hours	
Cloud Proxy Nodes	5	
Connections per Cloud Proxy	2000	
vCloud Director Tenants	500	
vCenter Server Instances per Platform Services Controller Node	4	
Platform Services Controller Nodes in vSphere Domain	8	
vCenter Server instances in vSphere Domain	10	

Appendix D – Reference Documents

Table 22. vCloud Availability for vCloud Director Reference Documents

Item	URL
Product Documentation	http://pubs.vmware.com/vcloud-availability-for-vcd-101/index.jsp
vCloud Director Documentation	https://www.vmware.com/support/pubs/vcd_sp_pubs.html
VMware vCloud Architecture Toolkit™ for Service Providers	http://www.vmware.com/solutions/cloud-computing/vcat-sp.html
vSphere Replication Documentation	https://www.vmware.com/support/pubs/vsphere-replication-pubs.html
vSphere Replication Technical Overview	https://www.vmware.com/files/pdf/vsphere/vmw-vsphere-replication-6-1.pdf

Appendix E – Tenant API Structure

Figure 25. To-the-Cloud Tenant APIs

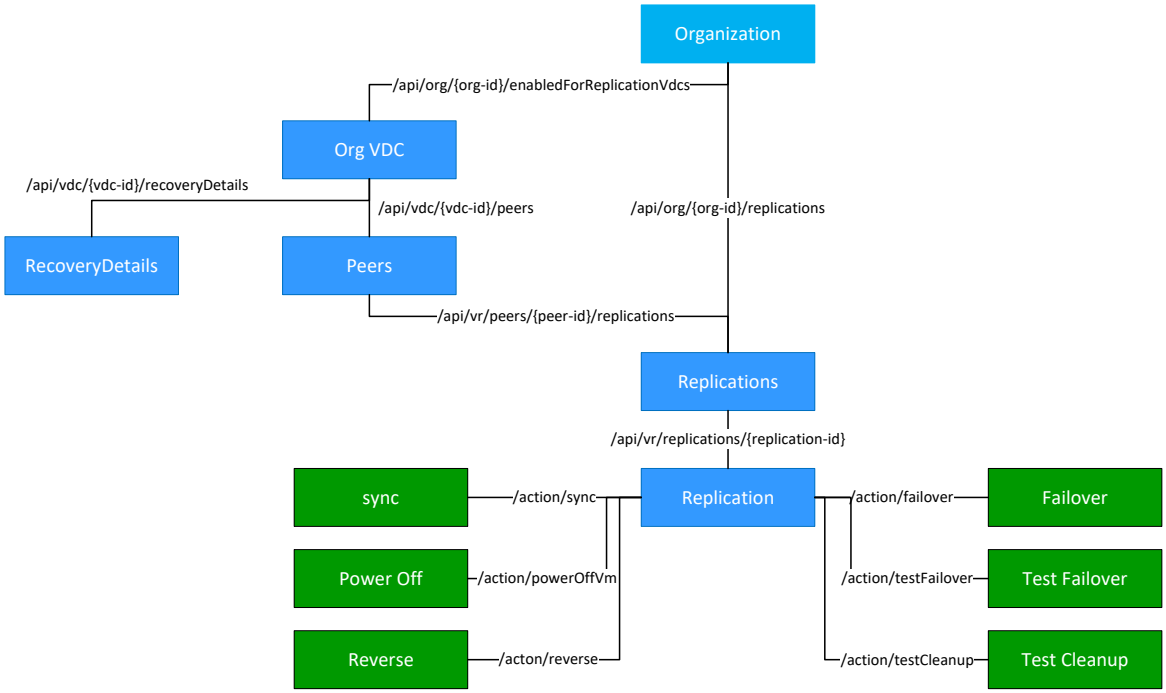
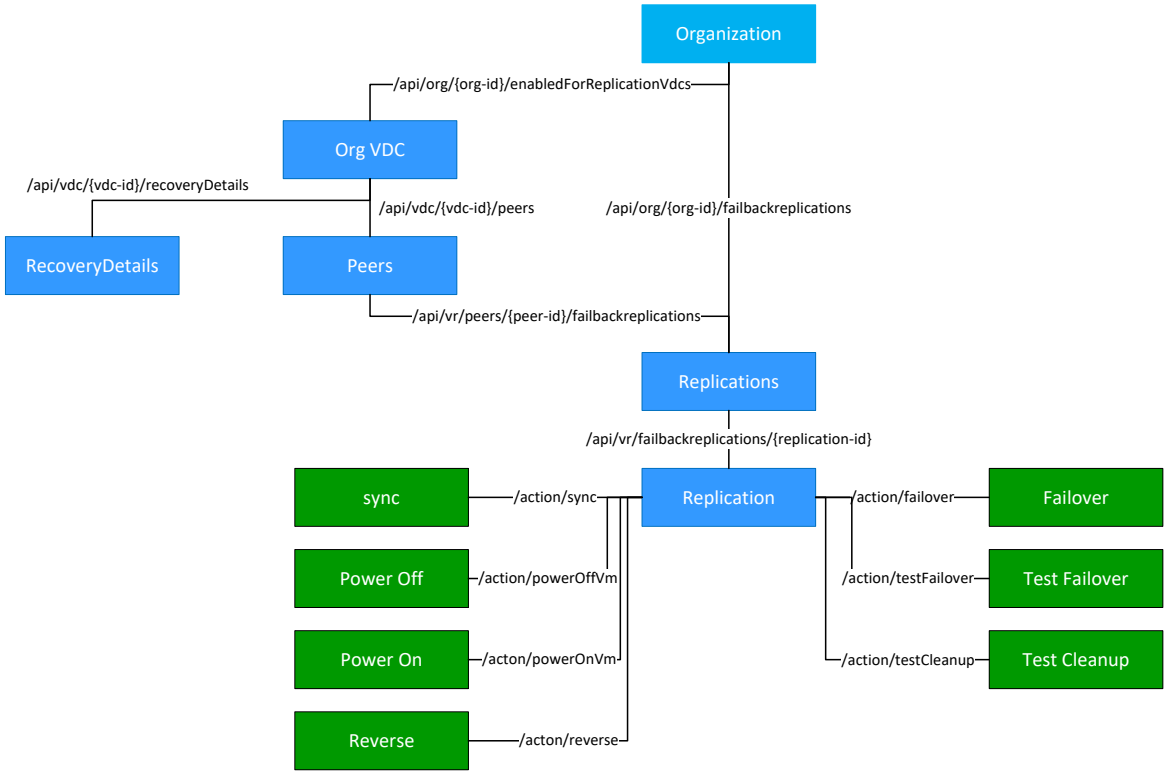


Figure 26. From-the-Cloud Tenant APIs



Appendix F – Undocumented HybridSettings vCloud API

HybridSettingsType

Element: HybridSettings

Type: HybridSettingsType

Namespace: http://www.vmware.com/vcloud/v1.5

Description: Public URL and certificate of the Cloud Proxy endpoint.

Since: 5.6

Internal: Yes

Schema: hybrid.xsd

Media type(s): application/vnd.vmware.vcloud.hybridSettings+xml

Extends: ResourceType

XML Representation:

```
<HybridSettings xmlns="http://www.vmware.com/vcloud/v1.5" href="xs:anyURI" type="xs:string">
    <Link href="xs:anyURI" id="xs:string" type="xs:string" name="xs:string"
            rel="xs:string"/>
    <CloudProxyBaseUri> xs:string </CloudProxyBaseUri>
    <CloudProxyBaseUriPublicCertChain> xs:string </CloudProxyBaseUriPublicCertChain>
    <CloudProxyBaseUriOverride> xs:string </CloudProxyBaseUriOverride>
    <CloudProxyBaseUriPublicCertChainOverride> xs:string </CloudProxyBaseUriPublicCertChainOverride>
    <CloudProxyFromCloudTunnelHost> xs:string </CloudProxyFromCloudTunnelHost>
    <CloudProxyFromCloudTunnelHostOverride> xs:string </CloudProxyFromCloudTunnelHostOverride>
</HybridSettings>
```

Attribute	Type	Required	Modifiable	Since	Description
href	anyURI	No	Always	5.6	The URI of the entity.

Attribute	Type	Required	Modifiable	Since	Description
type	string	No	Always	5.6	The MIME of the entity.

Element	Type	Required	Modifiable	Since	Description
CloudProxyBaseUri	string	No	none	5.6	Effective base URI for Cloud Proxy (wss) endpoint. By default, this value is the same as the base URI for the API endpoint. This value can be overridden by specifying CloudProxyBaseUriOverride.
CloudProxyBaseUriOverride	string	No	always	5.6	Base URI for Cloud Proxy (wss) endpoint. Leave empty to use the default.
CloudProxyBaseUriPublicCertChain	string	No	none	5.6	Effective SSL public certificate chain for the effective base URI for cloud proxy (wss) endpoint. The certificate chain is PEM encoded X.509 certificates. By default, this value is the same as the public certificate chain for the API endpoint. This value can be overridden by specifying CloudProxyBaseUriPublicCertChainOverride. This value will be empty if URI is overridden but the certificate is not.
CloudProxyBaseUriPublicCertChainOverride	string	No	always	5.6	SSL public certificate chain for the base URI for cloud proxy (wss) endpoint. The certificate chain must be PEM encoded X.509 certificates. A Base URI override must be specified if this value is included in a request.

Element	Type	Required	Modifiable	Since	Description
CloudProxyFromCloudTunnelHost	string	No	none	6	IP address or hostname of the effective host used for from-the-cloud tunnels.
CloudProxyFromCloudTunnelHostOverride	string	No	none	6	IP address or hostname of the host used for from-the-cloud tunnels. Leave empty to use default.
Link	LinkType	No	none	5.6	A reference to an entity or operation associated with this object.
VCloudExtension	VCloudExtensionType	No	always	5.6	An optional extension element that can contain an arbitrary number of elements and attributes. Not related to extension services.

GET /admin/hybrid/settings

Operation: GET /admin/hybrid/settings

Internal: Yes

Input parameters

Consume media type(s): None

Input type: None

Output parameters

current settings

Produce media type(s): application/vnd.vmware.vcloud.hybridSettings+xml

Output type: HybridSettingsType

PUT /admin/hybrid/settings

Operation: PUT /admin/hybrid/settings

Description: Updates settings

Internal: Yes

Input parameters

New settings

Consume media type(s): application/vnd.vmware.vcloud.hybridSettings+xml

Input type: HybridSettingsType

Output parameters

updated settings

Produce media type(s): application/vnd.vmware.vcloud.hybridSettings+xml

Output type: HybridSettingsType

Appendix G – Monitoring

Table 23. Monitoring with Endpoint Operations Management for vRealize Operations

Component	Process	Process Query
vCloud Director cells	vcd-watchdog	State.Name.eq=vmware-vcd-watc
	vCloud Director	State.Name.eq=java,Args.*.ct=vcloud*
CloudProxies	vcd-watchdog	State.Name.eq=vmware-vcd-watc
	vCloud Tunneling Agent	State.Name.eq=java,Args.*.ct=vcloud*
vSphere Replication Cloud Service	hcs	State.Name.eq=java,Args.*.eq=libs/hcs-6.1.0.jar
vSphere Replication Manager Server	hms	State.Name.eq=java,Args.*.eq=libs/hms.jar
vSphere Replication Server	hbr	State.Name.eq=hbrsrv-worker
Portal	nginx	State.Name.eq=nginx-linux-x64