VMware Cloud Director Availability 4.3

Introduction to VMware Cloud Director Availability

Every business in every sector needs to protect their investment in their digital estate. Typically, companies employ backup to be able to restore in case of a disaster. However, with exponential growth in data, backups are rarely tested and do not meet the recovery point needed for data that changes frequently. VMware Cloud Director Availability whilst able to cover frequency for replication, goes much further in granularity to 1-minute low recovery point objectives and high frequency of point in time instances and advanced retention options - helping businesses keep recovery as close to real time as possible.

Lots of companies have an expectation of either a managed service or a self-service depending on this size and budget. Self-service Disaster Recovery has become a simple expectation that has been popular in Hyperscale clouds, but most of these clouds require disk conversion to the target hypervisor, making testing difficult if not impossible and Disaster Recovery becomes a migration. Having a vSphere source and vSphere target cloud means there is no disk conversion, no different architecture or security to worry about and makes true Disaster Recovery as a (self-)service possible. Indeed, VMware Cloud Director Availability is the only Disaster Recovery & Migration solution integrated to VMware Cloud Director that supports self-service Disaster Recovery and Migration at a click of a button.

Of course, recovery is only as good as the testing to ensure recovery will work, and whilst many solutions cannot provide self-service or simplistic testing, none-impacting testing is out of the box for VMware Cloud Director Availability. Ensuring that customers can test when they wish to reduce risk of issues in a real failover Disaster Recovery event.

Whilst Disaster Recovery solutions are targeting applications, not all requirements are equal; some applications are mission critical and some are business critical, others are non-critical. It is important that a customer understands their application portfolio and has the option to choose the appropriate coverage for the application. Like insurance policies, it is not one size fits all, but the ability to differentiate is key to aligning cost to capability and criticality. VMware Cloud Director Availability features capabilities to tier offerings in repeatable, sustainable way – meeting all workload types and criticality.

Using a Disaster Recovery solution can also be a great way of onboarding into a cloud environment as many companies wish to move out of their data centers to a Cloud Provider, replicating workloads and cutting over from on-premises to cloud is exactly the same for Disaster Recovery as it is for migration. VMware Cloud Director Availability supports both Disaster Recovery and Migration allowing for complex networking and security and self-service failover use cases.

“As the market grows more and more aware of how sensitive corporate data is and how the availability demand from the line of business grows rapidly, we are happy to be using VMware Cloud Director Availability to fully satisfy those demands and needs for our customers.”

JONAS EMILSON
DIRECTOR PRODUCT MARKETING
ATEA

NEW FEATURES IN 4.3 RELEASE

- 1 Min Recovery Point Objective
- DR & Migration Recovery Plans
- Advanced retention policies
- VCD placement policy integration
- Certificate Update and Management
- Backup and Restore improvement
- Tunnel Endpoint Health
- NIC selection for local component communications
- DR access via customer’s identity provider authentication to DR Plugin and Cloud Director
- Improved Backup and Restore
- Operational improvements
Main themes for 4.3 release

Increased control and granularity
With 4.3 release we have introduced a 1-minute recovery point objective (RPO), advanced retention policies and recovery plans:

1 min RPO
If granularity is important to your customers and you have critical VMs in which data changes regularly, then you want the smallest distance between changes, i.e., a more granular recovery. Now providers can enable a 1-minute RPO using replication policies and SLA policies, if chosen this is cycled every 24 minutes with the Multiple Point In Time (MPIT) rotation, with the last MPIT being consolidated (full recovery vmdk created) before starting the next cycle. Of course, this is a desirable feature for critical workloads, but it comes at a price – network, storage, and compute must be able to handle the additional read and write tasks, which will be intensive. For a 1 min RPO to be available VMware recommends the use of all flash storage and a low current utilization and cache to capacity ratio for disk groups. Please check the official documentation to ensure your infrastructure is suitable before offering this option to tenants.

What is the actual use case for a 1 min RPO? As already explained such a capability will require significant resource, and this will in turn mean cost, that is why a 1 min RPO should be used for mission-critical applications, those the business cannot live without and/or have significant financial impact with downtime. It should be noted that modern applications including database systems, clustering technologies and other solutions can help protect these applications with active-active and standby setups when zero downtime is required and disaster recovery, really means disaster recovery.

Advanced Retention Policies
All replications within VMware Cloud Director Availability are restricted to the vSphere replication standard of 24 MPIT. Before 4.3 Cloud Director Availability these MPIT were spread evenly across the chosen duration and there was no way to change this. Now retention policies allow a maximum of 5 rules to be defined to keep an instance of an MPIT (a change delta) in each schedule, max 24 – note that the source disk will still be required for recovery consolidation. For example:

Recovery and Migration plans
Recovery of an application can be a complex task; it is rarely as simple as just starting up a VM. Cloud Director Availability has always recovered the VM and allowed some networking changes on recover to assist with customized recovery. In 4.3 we have taken this one step further with the ability to group VMs and vApps according to their recovery sequence priority. This can be used in a test failover, real failover and will clean up in reverse order too. The UI will allow a tenant to configure the sequence, timing between the execution steps and even scheduling the initial sync for migrations in the plan, all with options for error handling, retry and rollback.
The initial sync feature is particularly important for migration tasks, where you may be doing failover in a change window and want to make sure the prior the window you have a full sync of all the VMs and vApps that you will be migrating. Before 4.3 without the ability to schedule the initial full sync, there would have been a delay whilst this was done in the change window, rather than scheduling it to execute prior.

### Simple Authentication

**Login with API token**

Cloud Director Availability has improved on premise customers using the Cloud Director Availability vSphere plugin capability to do replication management operations without having to login to the provider Cloud Director instance using local Cloud Director credentials. Previously credentials for on premises and at Cloud Director were both local, and would only work if they matched, or the tenant would be re-challenged to authenticate with Cloud Director credentials, also, critically, there was no option to use the tenants own identity provider for both. Now tenants are able to authenticate with their own identity provider on premise and also using an API token instead of a local username and password, the authentication can be relayed to their identity provider and reused for accessing VMware Cloud Director Availability at the provider site, eliminating the need to have local users matching.

### Multisite authentication

This feature, originally introduced in 4.1 expects that in a VMware Cloud Director multi-site configuration using site association in a cloud-to-cloud DR scenario. Some service providers have multiple instances of VMware Cloud Director Availability using the same VMware Cloud Director instance and tenants replicating between org Virtual Data Centers. Some of those tenants use their own SAML server and wish to use it for authentication. The problem is that those tenants cannot extend their session to a remote site. When opening the “extend session” dialog in the UI, you can see which remote orgs your organization is associated to. Since the organization is not associated to itself, the dropdown is empty (as expected). In 4.3 json web tokens are used and using a new API these sessions are extended, supporting VMware Cloud Director multi-site capabilities.
**Assessable provider Virtual Data Centers**
Many providers configure their environment with a single SSO Domain with multiple vCenters. Some of them are workload vCenter and provide Provider VDC (PVDC) resources and some are management vCenters which host all management components for the cloud. In the case where each PVDC has its own VMware Cloud Director instance there maybe network connectivity limitations to management vCenters or other workload vCenters (perhaps in other geos and would not be replication targets). Prior 4.3, this configuration resulted in connection errors in the UI as VMware Cloud Director Availability attempted to connect to vCenters for which it had no responsibility. Now with 4.3 a configuration ‘permit’ list capability means that “Accessible Provider VDCs” can be configured so the UI will return only the vCenters that are backing the corresponding PVDCs and not the other vCenters in the SSO domain.

**Operational Improvements**
There are many operational improvements in 4.3, the ones below are some of the important and most requested enhancements.

**Backup and Restore**
In previous releases we have provided the capability to back up the configuration and services of an appliance and restore in situ to a ‘vanilla’ appliance that was required to be deployed before the restore could be done. In this release we are removing the need for an existing appliance to be deployed to restore the back up to and also encrypting backups for additional security.

**Change Cloud Certificate**
For DRaaS Cloud Providers will have many on-premises customer peers which they replicate from. Before 4.3 changing the VMware Cloud Director Availability appliance certificate forces all the peers to re-pair which means there was outage with tenant replication, and this usually needed to be co-ordinated in a change window. This was due to the on-premise components assessed the centralized component trust by comparing certificate thumbprints with what was persisted in the local database (which is a manual record accepted by the user).

In 4.3 certificate details are now stored on-premises using the default java CA trust store and associated verification procedure to ascertain whether the central site certificate is valid or not, this provides a full certificate chain verification without manual intervention needed in a certificate change.

**Local NIC routing**
Whilst we have provided the ability to configure a tunnel appliance and replication appliance with multiple network settings where there are more than one NIC, this was via a private API and CLI only to define static routes. This was to take care of problems when any other NIC than first NIC is planned to be used by:

- Tunnel for communication with other local components
- Replicator for communication with other local components

In 4.3 we are pleased to say that it is now possible to configure traffic types and tunnel endpoints for the tunnel/replicator interfaces and traffic to other local components (for tunnel perhaps via DMZ) or to hosts (for replication) all in the user interface.

**Tunnel endpoint health**
In an environment with many customers replicating to cloud, there are many tunnel connections. These tunnels will typically traverse firewalls and other network components, which can make troubleshooting hard. In 4.3 we have introduced a tunnel API to query the health of all local endpoints (components) and all remote tunnel endpoints (customer side or other cloud site). The API will query DNS resolution, connection status, route to host, and endpoint availability.
This is now also shown in the system health tab for convenience:

Simple, Capable Disaster Recovery as a Service

From the installation in the provider cloud to implementation on premise, VMware Cloud Director Availability is a simplified very capable architecture making it easy for customers and providers to implement. Customers can now find and setup DRaaS with a partner with our new vSphere plugin for DRaaS and Migration. Qualified DRaaS validated and Cloud Verified partners are listed in distance priority to the customer vSphere console, with integrated lead gen a customer can click on the partner and through form fill out request more about their service.

Once a customer has an agreement with partner and the destination details, they can self-serve deploy a replication appliance into their vCenter and connect to the provider Virtual Data Center via an encrypted tunnel, then start protecting their workloads directly from vCenter or from the Provider UI using the symmetric nature of the solution. VMware Cloud Director Availability allows customers to configure and manage both incoming and outgoing replication from the source and recovery site.

Importantly there are no agents to deploy on ESXi hosts and starting replication is a quick activity, equally the networking is vastly simplified to make it straightforward to deploy and use. Providers who enable VMware Cloud Director Availability for customers, enable customers to understand their protected status and run DR workflows directly in VMware Cloud Director UI, thereby driving more consumption and better user experience for customers.

VMware Cloud Director Availability provides coverage for 2 main use cases; on-premise to cloud Disaster Recovery and / or migration and Cloud to Cloud Disaster Recovery and / or migration.

Migration capability is cold and warm and easily scheduled into maintenance windows to suit your customers. Cold Migration is the complete sync of an offline workload before cutover and warm migration just syncs the differential at the time of cutover and is faster to implement. Many providers use VMware Cloud Director Availability for migration as it is simple, at no charge to VMware, but importantly can be driven by customers and allowing a customer to self-migrate when it suits them is a great experience and selling point.
Usability in this release has been significantly improved and many UI improvements make the solution far more intuitive for customers to enjoy Disaster Recovery as a Service and to manage navigation with new collapsible sections. Having intuitive usage is preferential for customers to be able simply use the solution and drives better consumption.

Around the themes explained in the previous section, VMware Cloud Director Availability is really helping customers drive better protection and testing. In fact, one big aspect of the solution is the ability to test, i.e. the ability to ensure that you have limited any uncertainly in your capability to recovery in the event or a disaster.

Testing frequently is the key to decreasing risk and protecting against a disaster, unfortunately it is perhaps the least used feature in DRaaS. Typically, this is because Disaster Recovery is provided by products that do not suit self-service or because the provider needs to ensure resource availability at the target is managed between multiple customers.

VMware Cloud Director Availability is self-service and can also be a managed service; self-service, and this means a customer can test their failover, non-impacting, at any time on any frequency. Managed service would mean a provider does this testing for the customer and this could be complimented with additional application testing services. As a self-service capability it is important that there are adequate resources at the target end to manage all customers compute requirements as potentially all could choose to failover or test failover at the same time. The recommendation is to promote testing as feature to decrease risk of recovery uncertainty.

**Workload distribution**

It is important to realize that not all workloads are equal in requirements, some may require much higher replication frequency and granular recovery due to the nature of the speed and criticality of the changing data, others may be non-critical and have longer cycles with less granularity. When considering Disaster Recovery, you need to have cost and functionality allocated correctly to cost, i.e. the higher the importance of a workload the more cost it is likely to take to cover it as it will consume more data, more replicants more frequently.

Mission Critical characteristics are defined as affecting the entire business, business will stop quickly in the result of outage. These are applications that have serious impacts on a broad part of the business can be deemed mission critical. E.g. financial systems which transact millions of transactions per minute are critical to the business success. Customers for DRaaS need to think about what applications are in their business that the business cannot survive without for even the shortest duration?

Business Critical characteristics are different and affect Line of Business, but overall business can operate and survive. These Line of Business applications can be viewed as business critical. E.g., HR payroll system, although without it, payroll will be interrupted, the business will carry on.

Lastly there are non-critical workloads and applications, they affect people personally and may delay deliverables, but ultimately, they are not affecting the business, nor teams in the business for a short duration. Items such as personal file systems and possibly email could be viewed as non-critical, it all depends on how you run your business.
It is easy to see how the recovery characteristics can be composed for different workload types, the following graphic indicates how customers look at recovery point and time objectives by workload type, although this data is from 2019, it is unlikely to have changed much, if at all:

![Diagram showing recovery point and time objectives by workload type]

**Mission - Critical** example: Finance / Billing
- RPO Average 1 min
- RTO Average <15min
- Suggested method – Synchronous Replication

**Business - Critical** example: eMail / SharePoint
- RPO Average 15 min
- RTO Average <2hrs
- Suggested method – Async Replication

**Non-Critical** example: Personal File Folders
- RPO Average >24hrs
- RTO Average <48hrs
- Suggested method - Backup

For a customer to be able to match a workload to a tier of service for Disaster Recovery is important as it will be more cost effective to have the appropriate resource capacity aligned to the workload. Having a single tier DRaaS portfolio does not provide the flexibility to cover mission critical workloads vs noncritical – there will be underused functionality/capacity which may cost the customer more overall. From a partner perspective, consumption will be much higher and better aligned with a tiered offering to customers.

VMware Cloud Director Availability SLA profiles offer out of the box classes of service to offer, the defaults are detailed below and can be added/modified or changed to meet you overall or per customer DRaaS cost to performance needs. With a simple nomenclature, customers understand what they are getting; from a Gold service with a low RPO and long retention time to a Bronze service with a longer RPO for less critical workloads and a short limited retention time.

<table>
<thead>
<tr>
<th>Tier</th>
<th>RPO</th>
<th>Retention</th>
<th>Quiescing</th>
<th>Compression</th>
<th>Initial Synchronizing</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold</td>
<td>3h</td>
<td>Keep latest instance only</td>
<td>Enabled</td>
<td>Enabled</td>
<td>no-diff</td>
<td>1 Orgs, 3 Replications</td>
</tr>
<tr>
<td>Silver</td>
<td>2h</td>
<td>7 instances every 1 day</td>
<td>no</td>
<td>enabled</td>
<td>no-diff</td>
<td>1 Orgs, 0 Replications</td>
</tr>
<tr>
<td>Bronze</td>
<td>4h</td>
<td>14 instances every 1 day</td>
<td>no</td>
<td>no-diff</td>
<td>0</td>
<td>0 Orgs, 0 Replications</td>
</tr>
</tbody>
</table>

As already noted the retention time in these SLA profiles is the ‘span’ of the Multiple points in time instances, with 4.0 these are flexible and in 4.3 these are extended with Advanced Retention Policies, permitting even more granularity over the MPIT cycle.

Resources are not unlimited so having the right option choices and taking advantage of the Advanced Retention Policies will mean better coverage overall and more revenue ultimately. SLA profiles in is an important feature that allows providers to start tiering services in this way to tenants, making the decisions for them on the DR capability and functionality at each tier and if required allowing customers to have their own custom profiles.
Market opportunity

As more customers move to cloud or find themselves in Cloud Provider VMware clouds, the need to protect their workloads becomes more and more important, not only from disasters, but also from malicious intent as more and more hackers threaten company’s intellectual property.

For these reasons, the growth in the Disaster Recovery market continues to grow at a CAGR of 36% from 2018-2022\(^1\) and expected to increase to 41.8% from 2022 to 2025\(^2\). Migration, security and Backup and Disaster Recovery are also some of the highest in demand hosted and cloud managed services organizations are planning to introduce in 2021/2022:

With migration to cloud being a primary use case customers are looking for, VMware Cloud Director Availability provides inclusive cold and warm migration at no additional cost, with a simple vSphere plugin or via the VMware Cloud Director user interface, customers can self-manage their own migration or providers can deliver migration as a managed service.

Considering security, malware attacks are prevalent, the ability to restore quickly is key to business recovery. With Cloud Director Availability 4.3, a 1-min RPO and recovery plans are front line to delivering the granularity of restore points and fastest time to recover (RTO), ensuring businesses can get back working as quick as possible.

The market is neither fragmented nor, at this time, consolidated from a provider selling DRaaS perspective so there is plenty of opportunity for all VMware Cloud Providers. Hybrid (on-premise and cloud based) configurations account for much of the current market share and represents an opportunity this today is provided by several global and regional providers as well as hyperscale providers like AWS and Microsoft Azure. However, solutions to Hyperscale or different target hypervisors are really viewed as migration solutions and not true Disaster Recovery solutions due to disk conversions making failing back very complex and manual. VMware Cloud Providers therefore have a great opportunity to sell DRaaS from a hybrid on-premises customer to their cloud solution with the benefit that it is not a migration (although it could easily be used for this), it is a true self service Disaster Recovery as a service capability.

\(^2\) https://www.mordorintelligence.com/industry-reports/disaster-recovery-service-market
For more information on cloud computing and VMware vCloud Powered services, please visit https://cloud.vmware.com/ or contact your VMware representative.

For more information about VMware Cloud Director Availability 4.x please see https://www.vmware.com/products/cloud-director-availability.html

If you would like to understand what your opportunity could look like using VMware Cloud Director Availability, please use our online calculators https://cpscalculator.vmware.com/

Access the VMware Learning Zone for Cloud Providers to learn more about cloud technology you as a provider can use http://bit.ly/VCPPSolutionEnablementLearningPath