4.2 On-Premises to Cloud and Cloud-to-Cloud Replication

Tenants can completely self-serve and automate replication, migration, failover, and failback of VM and vApps and post failover operations from their vCenter plugin or from the VMware Cloud Director interface (via the existing Host Based Replication (HBR) ESXi service) using the solutions symmetric capabilities. The user interface provides better usability and efficiency of easier replication management and overview of the tasks by simplifying the management interface.

In 4.2 tenants can also migrate workloads to VMware Cloud Director service managed VMware Cloud on AWS resources (from on-premises or from VMware Cloud Director clouds). This new warm migration capability is using a new “Data Mover” due to the restrictions in VMware Cloud on AWS to underlying ESXi APIs. The new data mover uses a ‘H4DM’ (Host for Data Mover) service to the VMC data engine, that, if a tenant selects a VMware Cloud on AWS target, will automatically be selected for the operation, there is no difference to the tenant, it uses the same familiar UI’s and workflows.

For the VMware Cloud on AWS migration, the VMware Cloud VPC contains a management and a compute resource pool. The service provider must manage the deployment on behalf of the tenant into the compute pool, this will include management appliances, the Replicator, Client and Tunnel appliance. There are some specific requirements for static IP on appliances and routing to management zone containing the SDDC vCenter and customer vmcOrg-Net.

This migration to VMware Cloud on AWS is not a reversible migration in this version of the H4DM. The migration is a warm migration. With a warm migration the VM doesn’t have to be powered off at migration time. This reduces downtime significantly and is one of the main reasons it is ideal for mission-critical workloads. However, the migrated VM typically has new network settings that might require additional actions on the target side once the process is complete. This method is still suitable for self-service, even though providers who have access to their tenants’ on-premises environments can offer it as a managed service.
Warm Migration

A warm migration consists of the following steps (regardless of whether the target is VMware Cloud Director or VMware Cloud Director service managed):

Configure Migration

Once the migration is configured in VMware Cloud Director Availability through the DRaaS vCenter plugin, or directly, in the VMware Cloud Director Availability UI, the initial data sync will be complete (this will use the HDM or H4DM depending on the target cloud). It has no impact on the source workloads, and they continue running without any interruptions. When the initial sync is complete, the workloads are ready to be switched to the destination site. During this ‘wait period’, changes are synchronized once every 24h if the target is a VMware Cloud Director endpoint. If H4DM is being used and the target is Cloud Director service, then migrations are sync’ed once at the beginning (when replication is started or at the specified time for ‘the delayed initial sync’), and then during the ‘planned migration’ (a.k.a. switchover) workflow. There are no syncs in-between.

Start Migration.

After the preparation for the migration is done, it can be started by the user. At the time of switchover, online sync is executed first, followed by a graceful shutdown. If the graceful shutdown fails with a timeout, a forced power off will be triggered. When the machine is offline, a rapid sync is performed to capture any changes since the previous one.

Power on the migrated VM

Changing network settings and IP addresses can be negated using layer 2 stretch, this is popular with customers and providers alike as it mitigates risks. For this reason, support for layer 2 warm migration has been included in 4.2

NOTE that Migration to VMware Cloud on AWS using the H4DM Warm Migration function does not support layer 2 connectivity, only SSL VPN.
In terms of flow, it is the same as a warm migration but with an extra prerequisite – to set up Layer 2 network stretch. This is now included in 4.2 of VMware Cloud Director Availability. Using NSX-T Data Center and its’ Layer 2 VPN solution, Cloud Providers can now stretch on-premises network segments to the cloud from VMware Cloud Director Availability (note that this is not available in VMware Cloud Director). The on-premises VPN ‘client’ is support by VMware® NSX Edge™ appliance, called NSX Autonomous Edge. Essentially this new update means that Cloud Providers no longer have to worry about the configuration of setting up L2 VPN to customer premises and managing it, whether over the internet or dedicated links, this is now configurable in VMware Cloud Director Availability. The requirement for on-premises is only that the workloads are attached to a VLAN-based network of distributed switches, i.e. vSphere distributed port groups.

To ensure security is maintained, Internet Protocol Security (IPSec) is used between tunnel endpoints. For each L2 VPN session an NSX Route based IPSec VPN tunnel secures traffic flowing between two networks connected over a public network through IPSec gateways called endpoints, these endpoints are managed by VMware Cloud Director Availability as a part of the configuration. It is not permitted to stretch networks to more than one vSphere Distributed Switch (VDS) trunk, for this another NSX Autonomous Edge instance must be deployed.

For the purposes of migration, only workload replicated data will utilize the tunnel, it cannot be used for anything else, however once configured can remain active for as long as necessary. IP communication is managed using Network Address Translation (NAT) through IPSec L2 VPN to ensure IP addresses are masked for outside. When a source workload is communicating to another target workload, MTU discovery is used by default, meaning the source workload will learn the path MTU value for the destination host through the L2 VPN tunnel, this helps optimize performance across the VPN and avoids fragmentation.