VMware Cloud: Planning an HCX Migration

VMware Cloud Migration

Table of contents

VMware Cloud: Planning an HCX Migration

Key Deliverables

The key deliverables that should be developed from HCX migration planning are as follows:

1. The migration “wave” plan that describes the workloads to be migrated and in what order.
2. A post-migration cutover plan that describes the final steps necessary to take the SDDC into full production.

Gathering Inventory

A necessary first step toward wave planning is to develop a deep understanding of the source environment. The key data points to collect are as follows:

1. An inventory of workloads to be migrated, including their compute/storage requirements.
2. An inventory of “anchor” workloads that cannot be migrated.
3. An IP addressing strategy for the migrated workloads.
4. A mapping of all network traffic flows and utilization levels.
5. An inventory of all on-premises networking gear that is in-path for the migration.
6. An inventory of all external storage devices that are used by workloads.

Identify Workloads to be Migrated

This process will involve an interview process whereas the architect will determine the scope of the migration project.

The key data points to gather are:

- An inventory of applications to be migrated.
- An inventory of workloads, and their compute/storage requirements, that are housing the applications. In addition to capacity requirements, be sure to gather requirements for storage I/O performance.
- An inventory of the networks (VLANs) providing connectivity to those workloads.

Be on the lookout for things that may complicate workload migration. Examples include:

- Workloads that have complex external dependencies, are large, or are intolerant of downtime.
- Workload that are difficult to migrate, including non-virtual workloads that must be virtualized, large or write-heavy databases, or workloads with multi-writer VMDK or Raw Device Mapping (RDM).

Identify Anchor Workloads

Anchor workloads are defined as those workloads (or appliances) that, for one reason or another, cannot be migrated. Examples include server infrastructure that cannot be virtualized (mainframes, non-x86 CPU, etc.) and network appliances such as firewalls and load balancers. Since anchor workloads may pose a major challenge to a migration project, it is critical to identify them as early as possible in the planning phase of a migration.

The key data points to gather are:

- An inventory of anchor workloads along with a description of their role in the current environment and reasons for why they cannot be migrated.
- An inventory of application integration points for these anchors. In other words, which applications are depending on them.
- An inventory of the networks (VLANs) providing connectivity to those anchors.

Determine an IP Addressing Strategy

A critical decision to make early-on is an IP addressing strategy for migrated workloads. This comes down to a binary decision between 2 possible scenarios.

Scenario 1: Workloads change their IP addresses

In case where workloads will change their IP addresses, then there are certain additional planning costs which you will incur.

- A new IP addressing scheme must be determined for the migrated workloads.
- Application owners must be made aware of the new IP addressing schemes and must be prepared to update their applications accordingly.
- System administrators must make plans for changing workload IP addresses post-migration.
- Updates to DNS must be coordinated to reflect changes in IP assignments.
- Firewalls may need to be updated to reflect changes in IP addressing.

Scenario 2: Workloads keep their existing IP addresses

A common goal of a migration project is to minimize disruption to application owners and users. Unfortunately, as seen above, IP address changes tend to be very disruptive. For this reason, most migration projects will require that workloads keep their IP addresses post-migration. In order to accomplish this, there are again 2 choices:

1. Migrate entire subnets worth of workloads at a time.
2. Utilize a layer-2 network extension.

Since the workloads housing applications aren't always arranged neatly within the bounds of a given subnet, it is often impractical to plan migrations around subnet boundaries. For this reason, the most common strategy is to utilize a layer-2 network extension to during a migration. Network extensions provide a great deal of flexibility in how a migration is performed and allow entire applications to be migrated regardless of the layout of the underlying network addressing scheme.

Understand Network Flows and Traffic Levels

In addition to the data gathering performed as part of the normal planning process for SDDC design, special consideration must be paid to the temporary traffic introduced as part of the migration. In particular, you must account for the following:

- Data replication – You must account for the additional north-south traffic resulting from data replication traffic for the migration. This includes both the initial replication as well as on-going delta synch for migration waves that are awaiting final migration.
- Network extension traffic – You must account for additional north-south traffic resulting from any network extensions that are in-place during the course of the migration. Having a detailed understanding of application network flows will help you to estimate this additional traffic load.

In-Path Networking Gear

Whenever possible, you should try to gather a basic inventory of the types of networking gear that is in-path for migration traffic. This data is useful in the event that you encounter legacy gear that may act as a bottleneck for migration traffic.

You should also establish basic monitoring for these devices (bits-per-second/packets-per-second for uplinks, CPU/memory for the device). This information is especially important in situations where there are performance issues with the migration. Basic monitoring data will help you to identify devices or uplinks that may be saturated.

External Storage

As with in-path networking gear, it is important to understand workload dependencies on external storage devices within the environment. It is not uncommon for legacy storage devices to become a bottleneck to the migration, so having an understanding of these dependencies will help in the event that a performance bottleneck is encountered during the migration.

Planning the Source Site Deployment

Migration projects will require that a certain number of HCX appliances be installed at the source site. It is important to ensure that these appliances are given sufficient resources.

The following are recommendations to consider when deploying the HCX components.

- Exclude the HCX appliances from DRS Automation.
- Configure High HA Restart Priority.
- When using HCX WAN Optimization service - assign SSD storage, or storage capable of 2500 IOPS.
- Ensure CPU/Memory resources are not constrained on the deployment cluster, reserve the resources if possible.
- If available, use dedicated mobility cluster (dedicated migration and network extension hosts) for maximum migration performance and network extension availability.

Wave Planning

Overview of Wave Planning

Wave planning is the process of sorting workloads into migration groups, or “waves”, that will be migrated concurrently as a single event. The general recommendations for planning HCX migration waves are as follows.

- Plan migration waves around applications whenever possible. In other words, attempt to migrate all workloads of a given application in as few waves as possible. This approach helps to keep intra-application traffic local to the SDDC.
- Plan to migrate interdependent applications in close sequence. This may help to reduce unnecessary north-south traffic on the SDDC.
- Plan to isolate larger/complex workloads to dedicated waves. Larger/complex workloads are typically databases or those VMs that have a high rate of change (lots of writes). These types of workloads tend to generate excessive amounts of delta data replication and will negatively impact other migrations.
- Choose the most appropriate migration option that fits the need of the individual workload. If a workload can afford to be powered off, then use a cold migration. Most workloads are suited well to bulk migration, as it is extremely rare that a workload cannot tolerate the small amount of cutover time imposed by a bulk migration. Additionally, bulk migration provides the opportunity to perform certain updates such as VM machine version and vmtools upgrades.
- Manage traffic flows for migrations carefully. Be conscious of the impact that a given migration wave will have on overall network utilization. In particular, pay attention to potentially sub-optimal network flows introduced by layer-2 extension.

Estimating Migration Times

There are a large number of variables to consider when estimating total time required to perform a migration. Some factors which may affect migration times include:

- WAN capacity (in terms of speed/throughput) between the source site and the target site.
- Network throughput in the source site LAN.
- Storage IOPS at the source and destination sites.
- Load (CPU/memory/storage) of individual hosts at the source site.
- Ability of WAN-opt appliance to dedup and compress data.
- Rate of change for workloads which are being migrated (i.e., how much delta data sync must regularly take place).
- Load at target site during the cutover phase of a migration (large migration waves will cause more resource contention).
- Time required to perform user validation on migrated workloads.

In order to make a reasonable estimate, it is important to understand which variables will represent the largest constraint. For example, if the migration is driven by a data center evacuation where the source site is running on older hardware, then storage or host constraints may easily become the limiting factor. It is also very common that an in-path network appliance is acting as a bottleneck on the source site LAN. There have been many cases where migrations were capped at artificially low rates due to an old in-path firewall or switch/router with a low backplane speed.

Understanding the complete end-to-end picture of the migration is vital when creating a time estimate. Performing pre-migration testing of the HCX installation will help to provide real-world data for this estimate.

Wave Planning Tips

The primary goals of wave planning are to facilitate speed, reliability, and minimal production impact. Keep the following in mind when performing wave planning:

- Seek to make the migration as transparent as possible. This will help minimize the need to coordinate between application owners, network engineers, and security administrators.
- Seek to minimize resource contention. Migration waves should be limited to 25-50 workloads in order to reduce network load during data replication and vCenter load during workload cutover.
- Overlap waves such that as one wave finishes its initial data replication the next wave is beginning its own. Overlapping in this manner will improve the data de-duplication of subsequent waves.

- Identify workloads with high rates of disk IO (specifically writes). This indicates a workload that will require large amounts of ongoing data replication. This type of workload should be migrated in a dedicated wave. Certain workloads, such as large write-intensive databases, may require services to be suspended prior to the migration.
- Create a vMotion Network Profile that connects to the cluster vMotion network (it is possible to route but may reduce migration performance). If clusters don't share vMotion networks, create unique Network Profiles (instead of routing to the vMotion networks).
- When the source site is at least vSphere 6, Replication traffic can use a dedicated vmkernel interface (similar to vmotion). If these exist, create a Replication Network Profile that connects to the cluster Replication network. If the customer wants to maximize migration performance, it may be beneficial to add this network (if it doesn't exist).
- Using a dedicated uplink network profile (instead of reusing the management network profile) may allow the network path being transmitted to the remote HCX target to use a less congested path. It may allow more efficient separation of the migration traffic (prioritize, segregate migration traffic on the physical fabric).
- Each Service Mesh (SM) includes one migration appliance (IX). It is possible to scale concurrent migration throughput to the same target SDDC by deploying additional IX appliances (this requires unique cluster/SDDC pairs).
- Scale using unique clusters at the source (CL1 -> SDDC1, CL2 -> SDDC1, etc.) or by adding SDDC targets (CL1 -> SDDC1, CL1 -> SDDC2, etc.)

Dealing with Anchors

Anchor workloads are especially problematic in situations where they reside within a network which is being migrated to an SDDC. Since long-term network extension is not recommended and a given network cannot exist in 2 places at once, there is really only 1 option: IP address change. Either the anchors must change their IP addresses or the to-be-migrated workloads that are sharing the same network with them must.

Note that in all cases there must be a plan for establishing some form of connectivity between the SDDC and the anchor workloads.

Migrating Over the Internet

In general, it is not recommended to use internet circuits for performing a migration. Private lines (such as AWS Direct Connect) are always preferable when available. In some cases, there may not be a choice and public internet may be the only available option. In these cases, you will need to keep the following in mind:

- You must manage contention for capacity on the internet circuit. Know that your normal user traffic will impact migration traffic (replication as well as layer-2 extension).
- You will need at least a 100Mbps internet circuit. This is the minimum required by HCX.

Considerations for Network Extension

It is important to give special consideration to planning around layer-2 network extensions. For anyone familiar with Ethernet, you will recall that there are no built-in mechanisms for loop detection or prevention. Layer-2 loops, when they occur, are often a disastrous event for a data center and were once a common problem for early LAN networks. Long ago, vendors began incorporating spanning-tree protocol into their products as a means of loop prevention in switched networks. Spanning-tree is not typically used with layer-2 extension over a WAN and is not compatible with modern virtualized network solutions such as NSX.

With layer-2 extension, loop prevention techniques vary between vendor and come with several complexities and caveats. HCX offers only minimal protection against layer-2 loops; mostly in the form of preventing the end user from deploying multiple L2C appliances in a given vSwitch. The purpose of this restriction is to prevent the user from extending the same VLAN multiple times and creating a loop. The danger comes from situations where multiple vCenter environments are connected to the same underlying switched infrastructure. In these scenarios, HCX cannot prevent multiple L2C appliances from extending the same set of VLANs multiple times and has no means of detecting this should it occur. In these situations, it becomes very possible to create layer-2 loops which can very easily ***bring down the entire network***. Awareness to this potential danger is critical when designing an HCX solution which involves network extension.

Recommendations

- Avoid using layer-2 extensions as a long-term solution. They are inherently risky and create sub-optimal network flows.
- Plan to tear down network extensions as soon as a given subnet has been fully evacuated. Traffic from extended networks will consume resources on the SDDC edge and WAN connections to the source site, possibly creating contention for available bandwidth. This contention may impact future migration waves. Additionally, layer-2 network extensions always introduce the slight risk of accidental creation of layer-2 network loops. Reducing concurrent extensions will help reduce this risk.
- Selecting a Deployment Cluster that is separate from the virtual machines on extended networks will reduce the packets forwarded by the L2C appliances and will result in improved performance.
- Network extension adds a 150byte encapsulation; end to end jumbo MTU will improve performance.
- Never extend cluster vmkernel networks using HCX (ESXi Management, vMotion, Replication, etc). Extending cluster networks is not supported.
- Never extend networks with the same VLAN backing to the same target SDDC. Layer-2 loops will ensue.
- Plan for 1 L2C per network being extended. Busy networks may require dedicated/semi-dedicated hosts in cluster-1 for the L2C appliance. Planning for ~1Gbps throughput per IX w/o WANOPT.

Planning for Post-Migration Validation

Post-migration validation provides a means of validating the success of a migration. You will want to develop plans for individual VMs/applications, migration waves, and final network cutover. Validation plans should include enough testing to provide confidence in the success of the migration and should include criteria that will trigger a backout of either a specific portion of the migration or of the entire migration.

Validation plans may include the following:

- Tests for individual VMs such as network reachability and performance checks. It may also include validation that all required processes are up and running.
- Application tests to ensure that applications running on migrated VMs are performing as expected.
- Wave checkpoint tests designed to spot-check the target site at the complete of each wave. This will help ensure that overall capacity (CPU/memory/storage) of the target site is in-line with projections.

Testing an HCX Deployment

In general, there are 3 categories of tests that should be performed prior to executing a migration plan. These are discussed below:

Basic Functional Testing

Basic functional tests will involve testing network extension as well as verifying that workloads may be migrated and reverse migrated. The takeaways from these tests will include ensuring that you have basic functional knowledge of HCX, as well as ensuring that you have a full understanding of how migration impacts the workloads.

Solution Design Testing

In addition to basic functional testing, it is important to validate the solution design itself. You must ensure that the migration plan meets the business requirements, that the technical aspects of the design are sound, and that the overall migration plan is feasible. In particular, you must understand how your plans for workload migration will impact the end-user, that your strategy for network extension and cutover behave as expected, and that your backout plans are workable.

Performance Testing

In order to develop a more accurate estimate of total migration times for a project, it will be important to gather baseline performance metrics for a real-world test migration. For testing, you will want to create dedicated test VMs which closely mirror your production workloads. Create small migration waves of a small handful of VMs and test each migration technique that you plan to utilize during the production migration. Gather metrics such as:

- Traffic amounts inbound to the IX appliance and outbound from the WAN-Opt appliance. This will help determine your dedup/compression rates. Low inbound traffic rates to the IX appliance may indicate storage or network bottlenecks in the source site.
- Data replication time for bulk migrations, and average cutover times for the various types of VMs which will be migrated. Understand how hardware and/or vmtool upgrades impact the cutover times.
- The effect of storage conversions on migration workloads (e.g., thick to thin).
- The effect of bulk migrations on applications. How do applications handle the “reboot to cloud” model of bulk migration?
- Migration time for cold migrations and vMotion.
- Network performance. Test latency and throughput for workloads that are utilizing network extension as well as those which are natively routed within the SDDC and reachable via IPSec VPN or other means.

In addition to gathering metrics for the migration from the source site to the SDDC, you will also want to test the functionality of a reverse migration; where a VM is migrated from the SDDC back to the original source. Understand how upgrades to migrated VMs affect reverse migrations. For bulk migrations, familiarize yourself with the archival locations of migrated VMs on the source site. These may be important for backout planning.

As an additional note, it is extremely common for legacy devices within the source site to create performance bottlenecks for HCX. If your migration testing is yielding poor results, then you may be hitting a bottleneck not related to HCX. Things to be on the lookout for will include:

- Storage - HCX can only migrate workloads as fast as it can read from storage. Are there any slow or over utilized storage devices to consider?
- Compute - Performance of the source vSphere deployment is a factor. Are these resources taxed?
- Network - It is surprisingly common for an over utilized or otherwise low-speed network device, which is in-path for HCX, to be the root cause of poor network performance.
- WAN - WAN performance is especially relevant if you are performing your migration over the internet. Basic troubleshooting such as bi-directional traceroutes will help identify your full path between the source and destination sites and help identify issues.

Backout Planning

Every good project plan includes backout planning in case something goes critically wrong. The quality of the backout plan will ultimately be determined by the level of testing that is performed prior to the actual migration. It will be critical to understand not only the time required to undo current migrations which may have already been performed, but also the potential impact of VM upgrades made to migrated workloads.

The HCX Migration Process

Once the HCX setup has been sufficiently tested, documented, and migration planning and backout plans have been developed, then the migration may commence. You will want to ensure that all required network extensions are in place prior to starting any migrations.

Migrations will take place in waves, per the wave planning developed as part of the migration plan. After each migration wave completes, you will want to perform validation of the given migration wave, per the post-migration validation plan. At the completion of the migration, a final validation should be performed prior to marking the migration completed.

See Migrating Virtual Machines with VMware HCX for additional information.

Cutover of Extended Networks

Workload migration typically involves workloads keeping their original IP addresses. Since network extension do not typically remain in place permanently, at some point a network cutover will be required.

Once a given network (VLAN) has been completely evacuated (i.e., all attached workloads have been migrated) then a network "cutover" should be performed for that network. A network cutover will involve removing the network extension to the SDDC and converting that network from a "disconnected" to a "routed" type within the SDDC. As a prerequisite for a cutover event, the source side of an extended network ***must*** have been fully evacuated (i.e., contains no devices other than the default gateway). If the network contains resources that will not be migrated to the SDDC, then they must be moved to another network that will remain native to the source site.

The process for a network cutover is roughly as follows:

1. Prepare the SDDC to act as the authority for the networks. This means reviewing the security policy of the SDDC to ensure that workloads will be reachable once the networks have been cut over.
2. Review the routing between the source site and the SDDC. Understand what will be required to ensure that the migrated networks are known via the SDDC. For example, this may mean adjusting prefix-lists to ensure that BGP routes propagate over IPSec or Direct Connect.
3. Schedule a maintenance.
4. At the start of the maintenance, unstretch the networks. You may choose to convert the networks to routed or leave as disconnected as part of the unstretch.
5. Shut down the networks on the source site (shut down default gateway interfaces).
6. If SDDC networks were left as disconnected, convert them to routed.
7. Adjust routing such that the migrated networks are reachable via the SDDC from the source site.
8. Verify connectivity to migrated workloads.

Authors and Contributors

Author: Dustin Spinhirne