Mobility Optimized Networking with HCX and VMware Cloud on AWS

HCX Technical Paper - April 5, 2021

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com

Table of Contents

Executive Summary 3

Introduction 3

The Case for Mobility Optimized Networking 5

How it Works 8

- Requirements 9

Policy Dependent Optimization 10

- About HCX MON Policy Routes 10
- Default MON Policy Configuration 10
- Best Practice MON Policy Configuration for VMC 11
- Policy Configuration for Amazon S3 Object Storage 11
- Policy Configuration with Route Based VPN 12

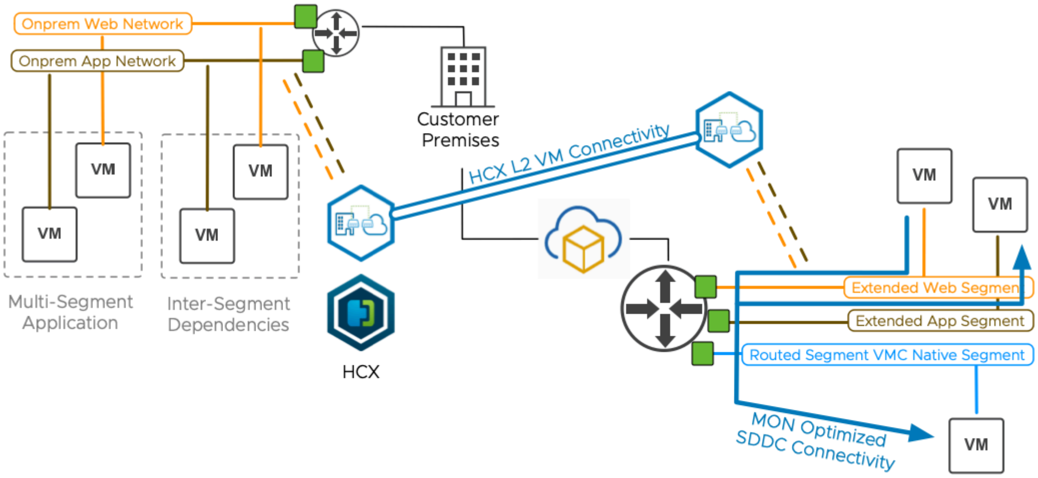
General Best Practices & Considerations 12

- Day 2 Operations & Outcomes 13
- Unsupported Configurations and Scenarios 14

About the Author 15

Executive Summary

HCX Mobility Optimized Networking (MON) is an enterprise capability of the VMware HCX Network Extension (HCX-NE) feature. MON enables optimized application mobility for virtual machine application groups that span multiple segmented networks or for virtual machines with inter-VLAN dependencies, as well as for hybrid applications, throughout the migration cycle. Migrated virtual machines can be configured to access the internet and AWS S3 storage buckets optimally, without experiencing the network tromboning effect.

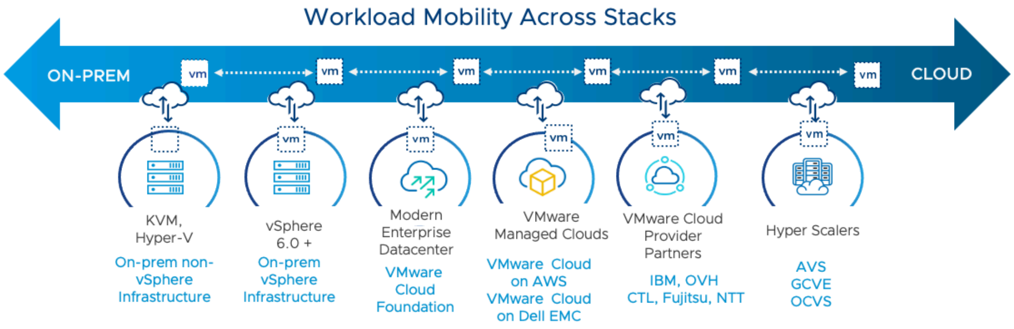


Introduction

This document describes the HCX Mobility Optimized Networking technology, the use cases and out of scope scenarios. It explores multiple scenarios specific to the MON feature in VMware Cloud on AWS.

About VMware HCX

VMware HCX delivers secure and seamless application mobility and infrastructure hybridity across environments running vSphere 6.0 and above, both on-premises and in the cloud. HCX abstracts on-premises and cloud resources and presents them as one continuous hybrid environment, enabling users to connect infrastructure and consume adopt a hybrid cloud vision, or a full migration to cloud as a consistent experience.



About VMware Cloud on AWS

VMware Cloud on AWS is an integrated cloud offering jointly developed by Amazon Web Services (AWS) and VMware. You can deliver a highly scalable and secure service by migrating and extending your on-premises VMware vSphere-based environments to the AWS Cloud running on Amazon Elastic Compute Cloud (Amazon EC2).

Intended Audience

This document is intended for hybrid cloud architects, technologists and system administrators, and any person planning workload migrations to VMware Cloud on AWS with HCX. The reader is expected to understand network patterns for their application network patterns for virtual machines being migrated.

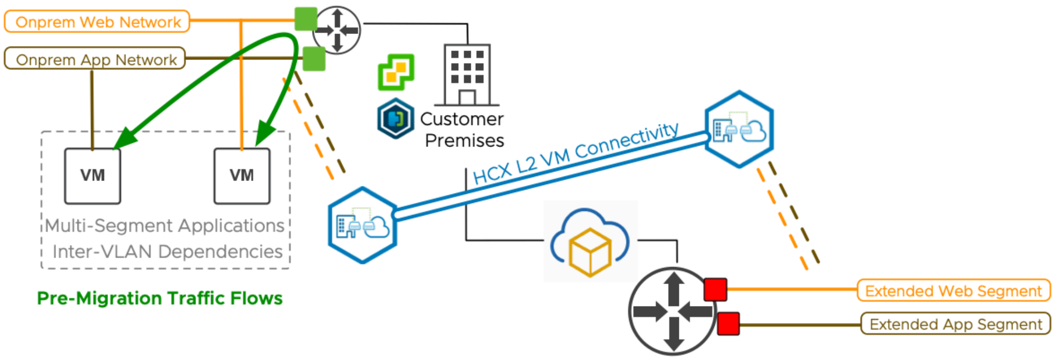
Definitions

Terminology	Definition
Flat Network	A network design approach where the topology is flattened to simplify configuration and administration. Flat networks with large broadcast domains are contrary in principle to segmented networks, which restrict broadcast domains and relies on VLANs, subnets and routers.
Segmented (Hierarchical) Network	A network design approach where the topology is segmented using variable length subnetting and VLANs to create a hierarchical routing configuration. Segmented, or hierarchical networks have controlled broadcast domains and are contrary in principle to flat networks, which rely on large broadcast domains and ARP discovery.
Network Latency	It is a measure of delay for data to go from the source server to the destination over a network.
Round-Trip Time / Delay	The length time it takes for a signal to be sent to a destination, plus the time it takes for the acknowledgement to be received.

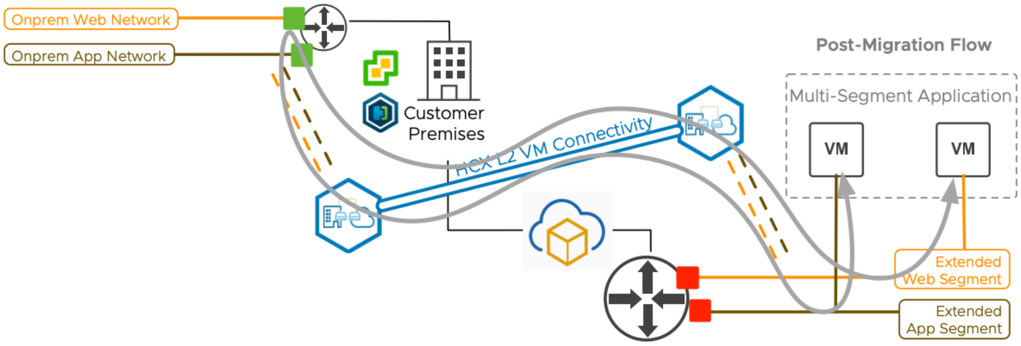
Network Tromboning	A network effect where traffic between topologically close virtual machines take a long detour, significantly increasing the total round-trip time. In the context of HCX NE and migrations to VMware Cloud on AWS: When SDDC virtual machines (connected to an HCX L2 network extensions) send traffic back to the premises router before returning to the same SDDC via HCX L2 service, or a routed path.

The Case for Mobility Optimized Networking

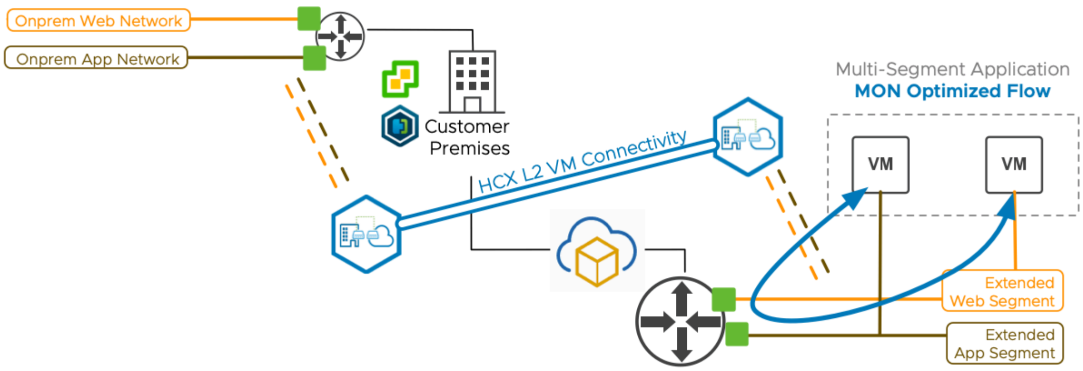
Mobility Optimized Networking improves routed connectivity patterns for multi-segment applications and virtual machines with inter-VLAN dependencies as those virtual machines are migrated into the cloud.



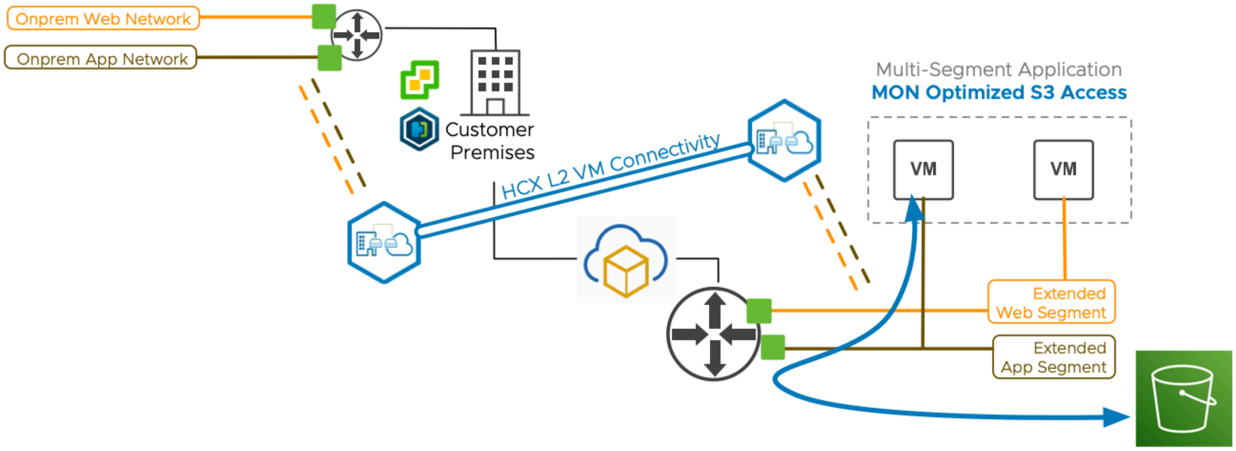
Without MON, HCX Network Extension expands the on-premises broadcast domain to the VMC SDDC while the first hop routing function remains at the source. The network tromboning effect is introduced observed when virtual machines connected to different extended segments communicate.



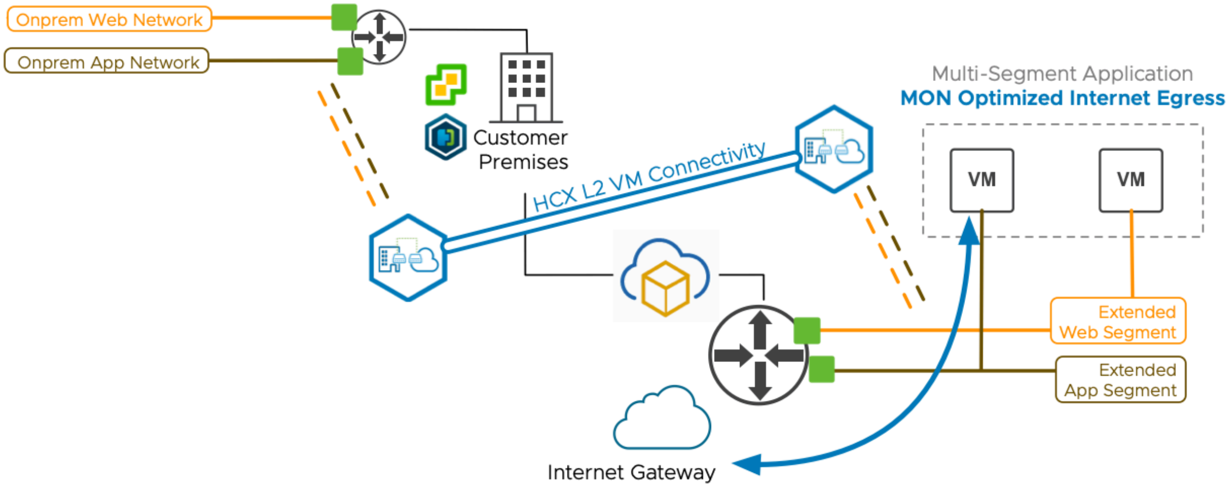
MON optimization enables migrated virtual machines to reach segments within the SDDC.



Mobility Optimized Networking can be configured to allow migrated virtual machines to reach S3 storage endpoints.



Mobility Optimized Networking enables migrated virtual machines to use the SDDC Internet interface (with SNAT).



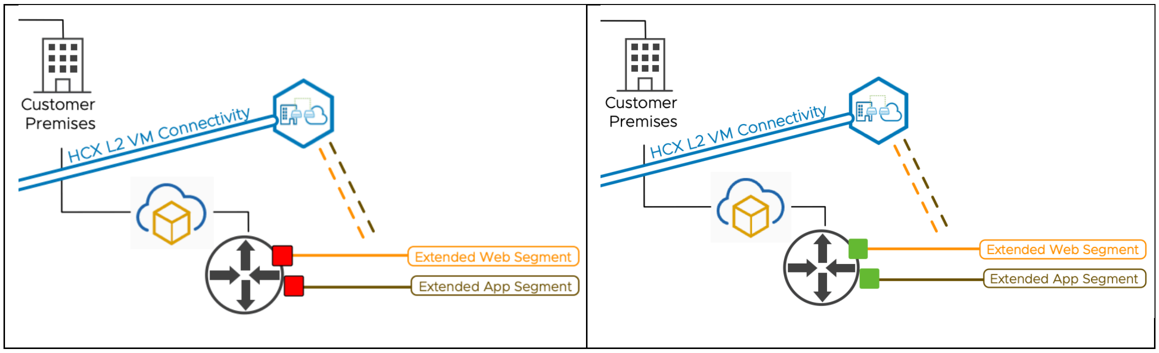
Beyond the Scope of Mobility Optimized Networking

- **MON does not optimize intra-VLAN traffic**
  Deployments with large flat networks on-premises may not benefit from the MON feature. Network extension (non-MON) forwarding functionality is used to achieve connectivity within a broadcast domain.
- **MON does not optimize within the on-premises network** or source environment. In SDDC-to-SDDC migrations, the optimization functions happen within the destination SDDC.
- **MON does not provide global ingress optimization**
  Network talkers outside of the SDDC must route to the MON-enabled targets via the source gateway.
- In multi-SDDC network extensions from on premises, **MON does not optimize SDDC to SDDC traffic**.

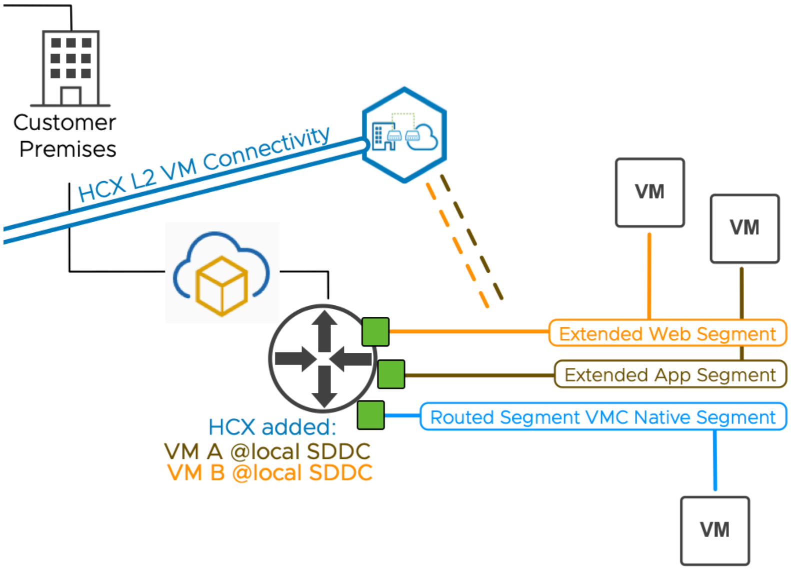
How it Works

This section explains what happens during the various phases of Mobility Optimized Networking:

1. MON is enabled for an HCX extended segment
   HCX enables the network ID (gateway IP) in the SDDC Compute Gateway. It is enabled with a limited /32 255.255.255.255 network mask.



2. When MON enabled on a virtual machine
   HCX adds reachability information for the migrated virtual machine (in the form of a virtual machine specific static route) to the SDDC Compute Gateway, allowing reachability within the SDDC. <u>This vm static route is not advertised to the on-premises environment</u>. The HCX L2 path is used to reach subnets not in the SDDC.
3. Using SDDC forwarding technology, the virtual machine able to use the SDDC Compute Gateway to reach the SDDC networks.



For reachability outside of the SDDC, the MON policy configuration is evaluated. According to the MON policy configuration. Matching subnets are sent to the original premises router. Non-matched subnets are sent to the SDDC Tier 0 router.

Note: This

With MON enabled, the following operations become available:

- Enable MON feature during the HCX Network Extension operation.
- Enable MON feature on existing HCX Network Extension.
- Disable MON feature on HCX Network Extension.
- Apply MON functionality during a VM migration. *Details in the next section.*
- Enable MON functionality for existing virtual machines.
- Disable MON functionality on virtual machines.
- Configure MON Route Policy to define on-premises (non-SDDC) subnets or exception/deny subnets for local egress.

MON Outcomes by Migration Type

In a MON-enabled segment the following behaviors are observed:

- **HCX Bulk** migrated virtual machines are automatically MON optimized in the SDDC.
- **HCX vMotion** migrated virtual use the on-premises gateway until they are transitioned in the MON interface.
- **HCX RAV** migrated virtual use the on-premises gateway until they are transitioned in the MON interface.
- Virtual machines created in the cloud, or created in the segment prior to enabling the MON feature will use the on-premises gateway until they are transitioned to use the SDDC cloud gateway.

Requirements

The following requirements apply:

- VMware HCX software with access to the HCX Enterprise feature set (HCX Enterprise features are bundled with VMware Cloud on AWS).
- Virtual Machines must be located in the MON enabled SDDC segment (the MON interface denotes virtual machines on the related on-prem network as ineligible).
- MON functionality requires VMware Tools IP Address detection. Virtual Machines must be running a supported version of VMware Tools.
- The cloud environment running MON compatible NSX-T 3.0+ software (this requirement is met automatically in a VMC SDDC)
- MON provides point to point functionality:

- It can be used to provide on-prem to SDDC functionality.
- It can be used in SDDC to SDDC migrations, where one SDDC is the source environment.

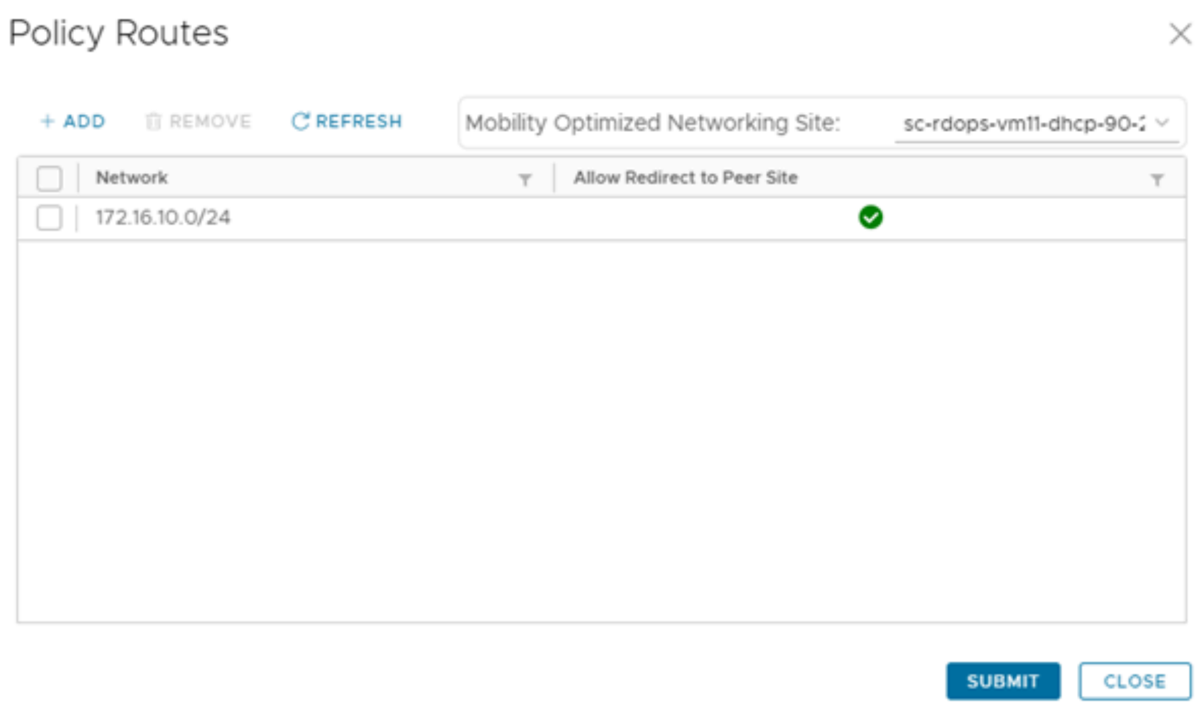
Policy Dependent Optimization

About HCX MON Policy Routes

When the destination network for a traffic flow is not within the SDDC, the MON policy is evaluated:

- If the destination IP is matched and configured as **allow** in the MON policy configuration, the packet is forwarded to the premises gateway using the HCX Network Extension appliance.
- If the destination IP is not matched, or configured to **deny** in the MON policy, the packet is forwarded to the SDDC Tier-0 to be routed according to the SDDC routing policy.

The HCX MON Policy configuration is not evaluated for virtual machines configured to use the on-premises gateway



Policy Routes

Mobility Optimized Networking Site: sc-rdops-vm11-dhcp-90-:

□	Network	Allow Redirect to Peer Site
□	172.16.10.0/24	✓

Default MON Policy Configuration

The default MON policy includes all RFC-1918 networks.

- This policy configuration forwards private subnet traffic (not destined to segments within the SDDC) to the on-premises router and sends internet egress traffic to the SDDC Compute Gateway.
- The default configuration should not be used when the SDDC has a BGP configuration that is learning a default route from the premises router, as it will result in asymmetric traffic patterns. See the **Best Practice MON Policy Configuration** in the next section.

Best Practice MON Policy Configuration for VMware Cloud on AWS

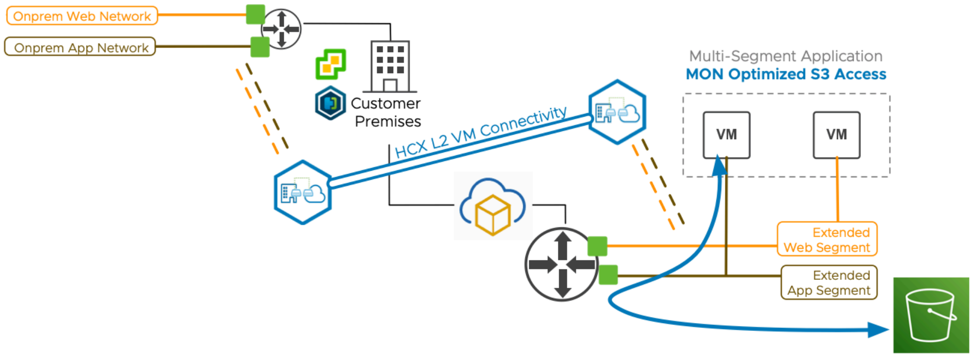
For MON deployments on VMware Cloud on AWS, it is recommended to replace the **Default MON Policy Configuration**.

- Remove the default RFC-1918 entries from the Policy Routes interface.
- Add a single entry for network 0.0.0.0/0 configured to **allow**.
- This policy configuration forwards private subnet traffic (not destined to segments within the SDDC) to the on-premises router, as well as internet egress traffic.

Policy Configuration for Amazon S3 Object Storage

The **best practice MON policy** can be revised to achieve S3 reachability.

- Locate the S3 IP Prefixes in the SDDC Compute Gateway firewall configuration.
- Configure the located ranges as deny entries (exclusions) to the **Best Practice MON Policy**.
- Deny entries are sent to the SDDC Compute gateway to be routed.
- The SDDC Compute gateway is able to use **Connected VPC** to reach S3 buckets.



⚠ S3 IP Prefixes may change. The user is responsible for adjustments to the MON policy.

Policy Configuration with Route Based VPN

In deployments with Route-based VPN (RBVPN), the SDDC router will advertise MON configured vm routes (prefix length 32) to the premises BGP peer.

⚠ VM static route advertisements only apply to SDDCs with RBVPN. connections.

Best Practice MON Policy approach (Recommended)

- Apply the **best practice MON policy** (replace the RFC-1918 entries with one 0.0.0.0/0 entry)
- In the premises BGP peer, filter routes with 32 prefix length advertised by the SDDC BGP peer.
- This configuration provides configuration consistency with DX deployments.

Propagated /32 vm routes approach (advanced)

- Remove the **default MON policy** (remove the RFC-1918 entries). Leave the policy blank.
- This configuration causes all policy evaluated public and private traffic to be sent to the SDDC for routing.
- The connectivity model is similar to the legacy Proximity Routing feature for NSXv:
  - This configuration requires the first hop router on premises be able to learn the SDDC advertised virtual machine routes.
  - Summarization and network design restrictions can prevent achieving proper connectivity with this approach.

General Best Practices & Considerations

Qualify the feature

- Does it make sense to enable it in your deployment? While enabling **Mobility Optimized Networking** can drastically reduce latency, it also introduces the complexity of additional network patterns.
- **Understand traffic patterns prior to enabling the feature:**
  - Quantify the Inter-VLAN packets and total traffic.
    vRealize Network Insight can help characterize the VM-to-VM network flows.
  - Intra-VLAN packets and traffic flows are outside of the scope of the MON functionality.

- **Understand your site-to-site latency**

  - MON eliminates round-trip latency between VMs in the SDDC.
  - MON does not eliminate point to point latency in post-migration inter-site traffic flows (across the WAN connection).

- Understand your application latency tolerance
  - Can the multi-VLAN applications to tolerate round-trip latency (2X site-to-site latency) after being migrated? If yes, then it may be preferrable to use the simpler network extension model.
  - If the multi-VLAN applications do not tolerate the increased latency, consider enabling MON on all HCX extended networks.

- Is cloud-based internet egress a requirement?
  - A MON migrated virtual machine can be configured with SNAT to access destinations.

Day 2 Operations & Outcomes

Unextending a Network Extension with MON and “Enable Cloud Gateway” checked

- The virtual machine static routes are removed.
- The Compute Gateway will be enabled
- The segment behaves like an SDDC native routed network.

Unextending a Network Extension with MON and “Enable Cloud Gateway” unchecked

- The virtual machine static routes are removed
- The Compute Gateway assignments remain, in a disabled state.
- The segment will be changed to disconnected.

Disabling MON on an existing network extension.

- The VMC network type changes from Routed to Isolated
- The Compute Gateway interface for the segment is set to disconnected, and it is changed from the /32 assignment to the original subnet.
- Virtual Machine static routes are removed from the Compute Gateway.
- Routed traffic is forwarded to the on-premises/source gateway (like NE without MON).

NE path becomes degraded on MON enabled extension

- MON enabled virtual machines are able to reach SDDC destinations.
- L2 traffic and Routed traffic targets on-prem are unreachable until the NE is restored.

Unsupported Configurations and Scenarios

Migrated VM & Extended Networks in the SDDC with MON Disabled

Traffic between MON-enabled migrated virtual machines to virtual machines on Extensions without MON is not optimized.

Migrated VM & SDDC Management Networks

Traffic between MON-enabled migrated virtual machines and the SDDC management networks is not optimized

Migrated VM & xENI Connected VPC Private IP Addresses

Traffic between MON-enabled migrated virtual machines and Connected VPC Private IP addresses.

Migrated VM & other SDDC VMs (over vTGW / Transit Connect)

Traffic between MON-enabled migrated virtual machines and virtual machines in other SDDCs (traffic over private Transit Connect)

Migrated VM with Multiple vNIC (MON functionality on secondary vNICs)

MON functionality is supported for one vNIC with one IP address.

Migrated VM & Multi-Tenant / CDS Networks (VMs on Tier 1 Compute Gateways in the Same SDDC)

Traffic between MON-enabled migrated virtual machines across Multi-Tenancy Cloud Director Service boundaries.

VMC SDDCs with Route-Based VPN connections to AWS VPCs

MON-enabled migrated virtual machines are advertised to Route-Based VPN connections. If the connected environment a VPC, it is subject to an inbound 100 route limit. Using Mobility Optimized Networking in this scenario is not unsupported.

⚠ Reaching the 100 route limit in AWS will transition the VPN to a down state.

About the Author

Gabe Rosas is a Staff Technical Product Manager for VMware HCX in the Networking and Security Business Unit at VMware. He is experienced in designing and operating traditional and software-defined datacenter infrastructure.

Twitter: gabe_rosas | **Blog:** hcx.design

Jesse Schachter is a Staff Engineer in the VMware HCX team. He has worked on products such as vCloud Director and NSX Edge. He has expertise in internet network security and virtualized routing and switching.

Acknowledgements

The author would like to thank **Michael Kolos** and **Aniket Deole** for their contributions to this paper.