Leveraging the Full Power of Oracle Cloud VMware Solution and VMware vSphere

Performance study featuring improved data performance

June 14, 2023

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

Consistent performance and control are critical when deploying or migrating mission-critical applications to a public cloud. VMware customers have run Oracle databases in their vSphere-based private clouds for over ten years. With Oracle Cloud VMware Solution (OCVS), enterprise workloads can be migrated to the public cloud without refactoring. In addition, applications already running in a vSphere-based private cloud can be directly migrated with a few mouse clicks to the vSphere-based OCVS solution that is hosted in the Oracle Cloud Infrastructure (OCI)-based public cloud.

Introduction

Over the years, vSphere private clouds have become mature platforms that provide excellent performance for enterprise applications, including mission-critical databases. With OCVS, public clouds can now be extended to provide vSphere private cloud capabilities. Additionally, OCVS allows complete control of the cloud environment—all existing best practices and performance tunings can continue to be used.

In this paper, we focus on how OCVS differentiates itself from other public cloud environments by:

- Migrating Oracle and SQL Server VMs from on-premises to OCVS with no refactoring and achieving the same performance.
- Showing OCVS AMD processors achieve the same database performance boosts as on-premises performance.
- Showing NSX Data Center can accelerate Oracle workloads, citing VMware and Deloitte Consulting studies.
- Showing the cost and technical benefits of customer-managed host remediations.

Database and enterprise workloads can run in OCVS with outstanding performance, just as they have been doing in private clouds for many years. OCVS is not limited to Oracle-based applications and is truly an open platform for any application. We ran SQL Server and Oracle databases to illustrate this point in our tests. Having vSphere as the underlying platform in public and private clouds lets customers decide the best place to run their applications.

Oracle Cloud Infrastructure (OCI)

Oracle Cloud Infrastructure (OCI) is a deep and broad platform of public cloud services that enables customers to build and run a wide range of applications in a scalable, secure, highly available, and high-performance environment.

VMware Cloud Provider Stack

The Oracle Cloud VMware Solution uses the VMware Cloud Provider Stack (VCPS) bundle that includes the necessary software-defined data center (SDDC) product components and its version details to deploy an SDDC and manage its lifecycle. The VCPS solution offers cloud providers an opportunity to offer their customers a subscription-based VMware SDDC solution on bare-metal elastic infrastructure service as part of the metal-as-a-service (MaaS) program.

The Cloud Providers use the VCPS bundle that includes the necessary SDDC product components (vSphere, vSAN, and NSX) and Hybrid Cloud Extension (HCX) to deploy an SDDC and manage its lifecycle.

vSphere

vSphere uses virtualization to transform individual data centers into aggregated computing infrastructures that include CPU, storage, and networking resources. VMware vSphere manages these infrastructures as a unified operating environment and provides the tools to administer the data centers participating in that environment.

The two core components of vSphere are ESXi and vCenter Server. ESXi is the virtualization platform where virtual machines and virtual appliances are run. vCenter Server is the service through which multiple hosts are connected and managed together as a pool of host resources.

vSphere with Tanzu transforms vSphere into a platform for running Kubernetes workloads natively on the hypervisor layer. When enabled on a vSphere cluster, vSphere with Tanzu provides the capability to run Kubernetes workloads directly on ESXi hosts and to create upstream Kubernetes clusters within dedicated resource pools.

vSAN

vSAN aggregates local or direct-attached data storage devices to create a single storage pool shared across all hosts in the vSAN cluster. Using vSAN removes the need for external shared storage and simplifies storage configuration and virtual machine provisioning. Built-in policies allow flexibility in data availability.

NSX Data Center

NSX Data Center is focused on providing networking, security, automation, and operational simplicity for emerging application frameworks and architectures that have heterogeneous endpoint environments and technology stacks. NSX Data Center supports cloud-native applications, bare-metal workloads, multi-hypervisor environments, public clouds, and multiple clouds.

Hybrid Cloud Extension (HCX)

VMware HCX, an application mobility platform, simplifies application migration, workload rebalancing, and business continuity across data centers and clouds. VMware HCX enables high-performance, large-scale app mobility across vSphere and non-vSphere cloud and on-premises environments to accelerate data center modernization and cloud transformation. A key benefit is the ability to create network extensions while seamlessly stretching IP address space across multiple clouds, allowing workloads to be relocated without a complex re-IP exercise.

Oracle Cloud VMware Solution

Oracle Cloud VMware Solution (OCVS) is a VMware Cloud–based environment that empowers customers to have full administrative control over their VMware environments while simultaneously leveraging the capabilities of Oracle Cloud Infrastructure. This solution includes the VMware products: vSphere, vSAN, NSX Data Center, and vCenter Server. It provides bare-metal compute, an advanced L2 network, and a storage-rich infrastructure that delivers a full-fidelity VMware service for predictable application performance. Moreover, OCVS is designed to cater to dynamic multi-cloud VMware environments.

This unified cloud infrastructure and operations platform enables enterprise IT personnel to migrate and modernize applications quickly while seamlessly moving workloads between on-premises and Oracle Cloud Infrastructure at scale. Enterprises can now migrate or extend VMware-based workloads without the need for rearchitecting applications or retooling operations. Additionally, IT teams can easily leverage services such as Oracle Autonomous Database, Oracle Exadata Cloud, and Oracle Database Cloud with consistent portal access and modernized APIs.

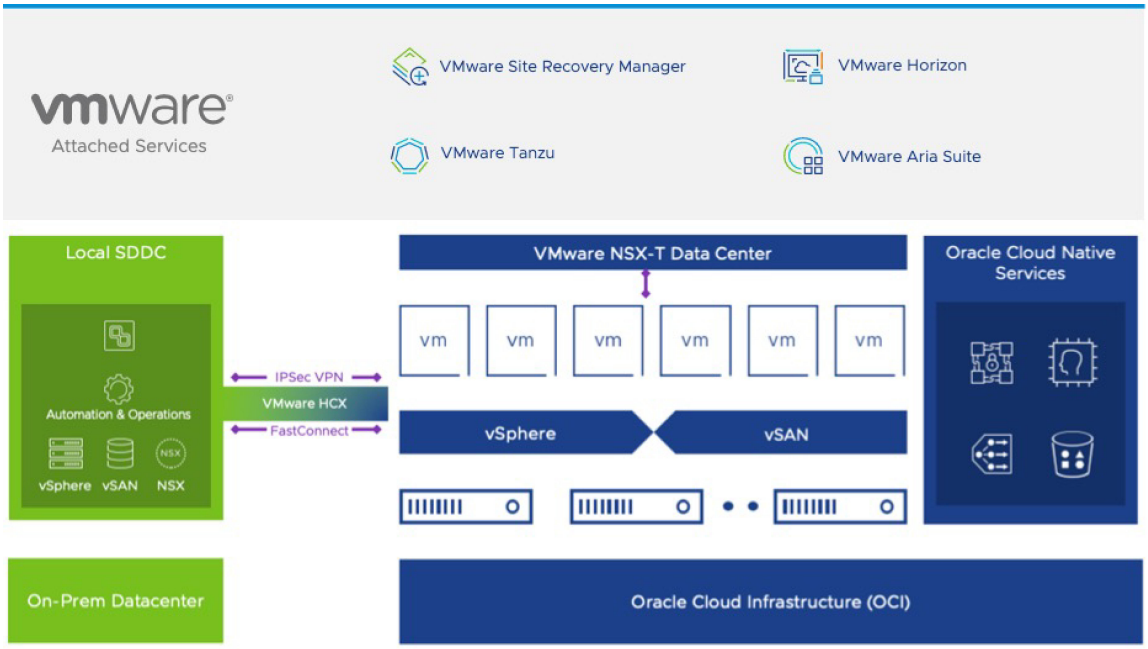

Figure 1. Diagram of how VMware products fit with Oracle Cloud Infrastructure and Oracle Cloud Native services

Detailed Benefits of OCVS

In this section, we provide an overview of the benefits of OCVS with respect to performance and enterprise workloads.

VMware Enterprise Cloud Platform

When deploying traditional enterprise applications in the cloud, there are many factors that must be considered by the application owners, architects, and the business. VMware and Oracle customers have been running Oracle databases successfully with high availability and stability in production for well over 10 years on-premises.

Any Workload

VMware has continually improved the vSphere platform to be able to run all workload types with good performance. VMware has worked with partners and customers to bring enhancements and tuning directly into the core product to address customer needs that have evolved over time. In addition, guides, best practices, and knowledge base articles show how to configure the Oracle stack on the vSphere platform. These are designed to further optimize performance, reliability, and availability when running on vSphere. The tunings range from BIOS settings, virtual machine settings, operating system settings, and host settings outlined in our documentation. The guide, *VMware Hybrid Cloud Best Practices Guide for Oracle Workloads*, simplifies creating Oracle database virtual machines. VMware's vast technical library is not limited to relational databases and includes analytic databases, in-memory databases, NoSQL databases, Telco workloads, media and entertainment workloads, and others.

Importance of Full Management

OVCS is a fully customer-managed offering that allows Oracle virtual machines in OCVS to be configured and optimized identically to those running on-premises.

Data Transfer – Outbound/Inbound Traffic

When deploying or migrating applications to the cloud, planning is crucial to avoid any cost or technical surprises such as egress costs, data transfer fees, or incomplete application dependency mapping. In a multi-cloud environment, when supporting enterprise applications, it is important not to treat each cloud as a silo. Enterprise applications and business processes typically depend on other applications or data feeds that can cross clouds and potentially lead to unexcepted data transfers.

Egress and Data Transfer Rates

The flexibility of Oracle Cloud and its networking features allows customers to minimize or eliminate these costs. With Oracle Cloud networking, there are no data transfer or egress charges associated with communications between OCVS and OCI because they are both services in the Oracle Cloud. Additionally, inbound traffic and network traffic between OCI availability domains within a region is free of charge.

For outbound traffic, the first 10 terabytes are free. This allowance allows customers to correct any surprise cloud-to-cloud dependencies missed during initial migrations. In the case of hybrid cloud deployments where egress costs are expected, the 10TB allowance may be sufficient and can be managed accordingly.

Once the network allowance has been consumed, there is a per-gigabyte charge for outbound traffic, which varies from region to region as outlined in table 1.

Description	Unit Price	Metric
Inbound data transfer	Free	
Outbound data transfer – first 10TB/month	Free	
Outbound data transfer – over 10TB/month, originating in North America and Europe	$0.0085 USD	Gigabyte outbound data transfer per month
Outbound data transfer – over 10TB/month, originating in APAC, Japan, and South America	$0.025 USD	Gigabyte outbound data transfer per month
Outbound data transfer – over 10TB/month, originating in Middle East and Africa	$0.050 USD	Gigabyte outbound data transfer per month

Table 1. Per-gigabyte charges for outbound network traffic

Oracle FastConnect

Finally, there are no charges for inbound or outbound data sent over the Oracle FastConnect service. FastConnect lets customers connect directly to their Oracle Cloud Infrastructure (OCI) virtual cloud network via dedicated, private, high-bandwidth connections. Customers choose an appropriate port speed plan based on their requirements and pay a consistent monthly price.

Poor or incomplete migration planning can contribute heavily to unexpected egress or data transfer costs. To minimize surprises, VMware Aria Operations for Networks can be used to perform application dependency mapping and discovery. This is discussed in the section, "VMware Aria Operations for Networks."

OCVS Customer-Managed Host Remediation

OCVS is a fully customer-managed environment that allows host remediation and the addition or removal of hosts. This aspect of OCVS provides potential Oracle licensing cost benefits.

Host Licensing

It's important to note that any discussion concerning Oracle licensing must ultimately be between Oracle and the customer. VMware can only provide guidance and perspective; licensing of Oracle is entirely between Oracle and its customers.

Whether the host is physical or virtual, all sockets and cores must be fully licensed on a server as required by Oracle. This is not specific to vSphere—this is the case for virtual or physical servers. In the event of a host failure using automated remediation, a host is monitored and then waits to determine if issues are transient; if they're not transient, then the host is replaced. While the new host is being added, the failing host has not been taken down to comply with Oracle licensing. All running hosts must be licensed at any point in time, so N+1 hosts in the cluster must be licensed.

Benefits of Customer-Managed Remediation

In a fully managed OCVS cluster, customers can create their own host replacement strategies using monitoring and alerts. Upon triggering an alert, a host can be scheduled for replacement as shown in figure 2.

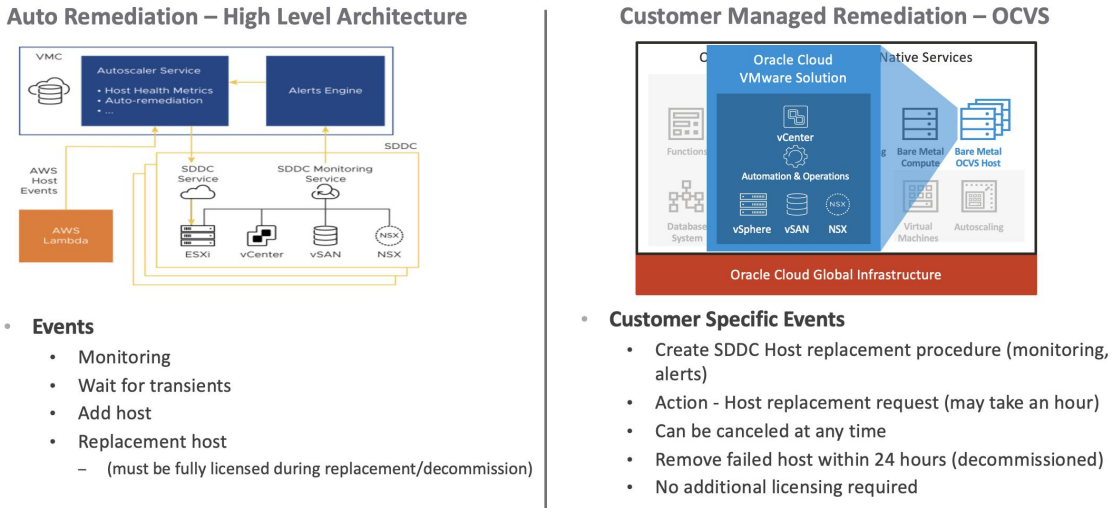

Figure 2. Architectural diagrams of auto remediation and customer managed remediation

In this scenario, there is no need to license any additional hosts, which significantly reduces licensing on an OCVS production cluster. An N+1 licensing strategy is not required; however, the downed host must be removed from the cluster within 24 hours.

Cloud Economics – Driving Down Costs

When deploying enterprise workloads on vSphere, there are many best practices associated with memory and CPU resources; however, a fundamental and universal best practice across all mission/business critical applications is not to over-commit memory or CPU resources. Specific to Oracle, VMware recommends memory reservations equal to the sum of the size of the Oracle SGA, the Oracle PGA, the Oracle background processes, and operating system used memory. Usually, this means a reservation for all of the memory assigned to the VM. We also recommend not to over-allocate vCPUs and to maintain a 1:1 ratio of physical cores to vCPUs such that virtual machines will not be sharing CPU resources.

The performance studies in “Enterprise Benchmark Performance” show that vSphere can leverage AMD EPYC processors' architecture to increase Oracle database performance. Increasing performance, in turn, increases the consolidation ratios in OCVS, essentially doing more work with less and driving down the total cost of ownership.

OCVS Compute Shapes & AMD EPYC Processors

Oracle Cloud VMware Solution offers flexible compute instances that are available as X7 dense 52-core Intel Xeon and E4 dense AMD EPYC processors. Customize shapes with 32-, 64-, and 128-core configurations with 2.7X to 3.5X1 more memory and 1.2X to 3.5X2 more storage per host than other offerings. These offerings allow customers to leverage the vSphere CPU scheduler, which is optimized for AMD EPYC processors.

Leveraging the Full Power of OCVS and vSphere

VMware vSphere 7.0 U2 and later versions include a CPU scheduler that is architecturally optimized for AMD EPYC processors. This scheduler is designed to take advantage of the multiple last-level caches (LLCs) per CPU socket offered by the AMD EPYC processors. This section shows that OCVS running vSphere 7.0 U2 can achieve optimal performance using out-of-the-box options and configurations. The VMware whitepaper "Performance Optimizations in VMware vSphere 7.0 U2 CPU Scheduler for AMD EPYC Processors," which was published in 2021 when the enhancement was included in the release of vSphere 7.0 U2, has extensive details about the EPYC architecture and the scheduler enhancement. Here, we present a brief overview of some of those details and refer the reader to the full paper for more information.

Background and Introduction

Prior to AMD EPYC processors, there was one L3 cache per NUMA node, meaning that all cores on a NUMA node shared a single L3 cache for that NUMA node. AMD "Zen" changed this by introducing multiple L3 caches within each NUMA node.

AMD EPYC processors are built using the multi-chip module (MCM) architecture where a set of core complex dies (CCDs) are connected to an I/O die via AMD Infinity Fabric.

In total, there are up to eight CCDs in the AMD EPYC 7002 Series processors shown in the simplified logical diagram in figure 3.

Two core complexes (CCXs) make up each CCD. A CCX is up to four cores sharing an L3 cache of 16MB in size, as shown in figure 4 for a 2nd generation CCD. As seen at the software layer, inter-CCX traffic still travels through the AMD Infinity Fabric. In AMD EPYC 7003 Series Processors, there is just one CCX/LLC per CCD and up to eight cores per LLCs.

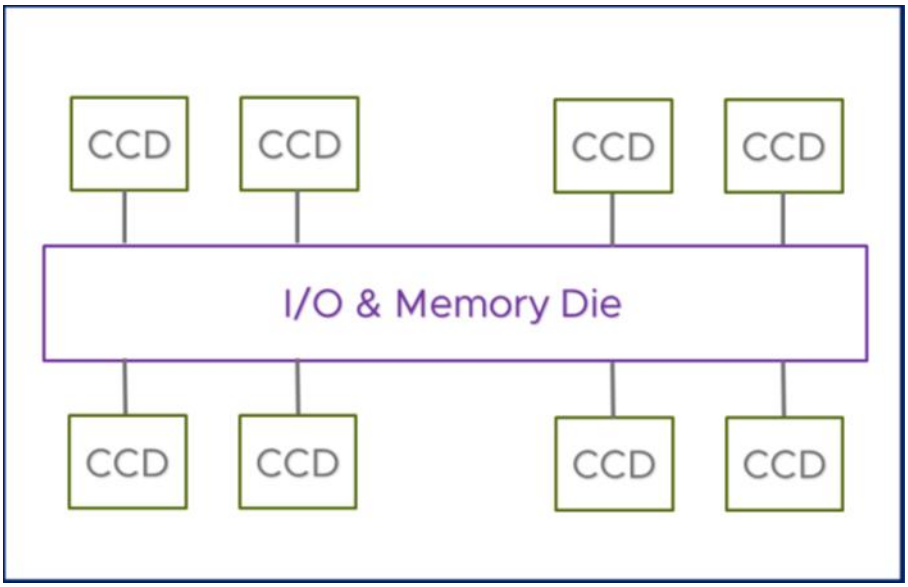

Figure 3. Simplified view of a single socket AMD EPYC 7702 (Rome) processor architecture Performance

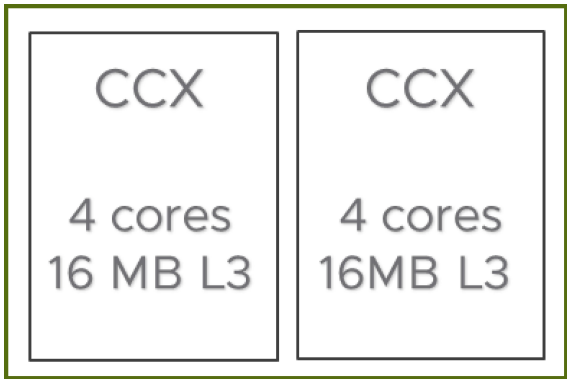

Figure 4. Compute/Core Complex Die (CCD) for AMD EPYC 7702

AMD EPYC 7702/7742 processors each contain 64 cores and 4 cores per LLC, with a total of 16 LLCs. AMD EPYC sockets have a similar CCX-based design across all generations: Gen1, Gen2, and Gen3. On AMD EPYC CPUs, the cost of cache-to-cache sharing may also vary greatly depending on the relative location of the caching/home agent and the target cache line, as well as processor generations. Such sharing cost tends to remain constant within a socket on other x86 processors. For example, as seen in software, the cache access latency can increase up to 3x for inter-CCX cache access compared to intra-CCX cache access latency. In short, one 2nd Gen socket has up to 256MB (16x16MB) system-wide L3 cache, but this large cache size comes at the cost of non-uniform cache-access (NUCA) latencies. Software vendors must use these architectural features to optimize their software stacks.

Design of vSphere CPU Scheduler Support for AMD EPYC Processors

VMware augmented the vSphere scheduler such that it is aware of the topology where each NUMA node may have multiple LLCs with distinct cache cost or non-uniform cache access latency (NUCA). The goal is to automatically explore the locality benefits offered by the hardware architecture by placing related contexts, such as vCPUs of the same VM, or multiple communicating contexts, onto as few LLCs as possible. Intuitively, a VM is divided into logical groups that can be individually placed on LLCs to benefit from LLC locality without over-packing.

Performance Testbed and Methodology

We set up a private cloud in our on-prem environment to match the OCVS cloud shape as closely as possible. Both were a single host with two AMD EPYC CPUs of the same generation with 64 cores per socket. The OCVS system had a slight advantage with NVMe-based local storage, while our on-prem system had SSDs.

Oracle 19c was installed on a Linux VM, and the popular DVD Store open-source database benchmark was set up. This is an online transaction processing (OLTP)–based benchmark that simulates an online store that sells DVDs. It simulates many aspects of an online store including browsing products, reading and writing product reviews, new customer signups, premium memberships, and purchasing products. The key metric for this test is the number of orders per minute (OPM), which reflects the throughput of the system under test.

Two sets of tests were run. The first directly compared the performance of the on-prem private cloud to the public cloud in OCVS. To determine whether there was a noticeable performance difference between the private and public clouds, the systems were set up to be almost identical.

The second test measured the performance of the scheduler enhancements to optimize the AMD EPYC architecture. These enhancements were originally implemented and measured with the release of vSphere 7.02 on-prem. With the availability of AMD-EPYC CPU-based public cloud options in OCVS, we wanted to again measure these enhancements in the new public cloud environment.

Enterprise Benchmark Performance

Figures 5 and 6 show that the performance of both Oracle database VMs and SQL Server VMs performed essentially the same in OCVS as they did in the on-prem private cloud environment. Due to the advantage of NVMe-based storage in the OCVS-based solution, it achieved a 4% and 1% gain in throughput for the Oracle and SQL Server tests respectively.

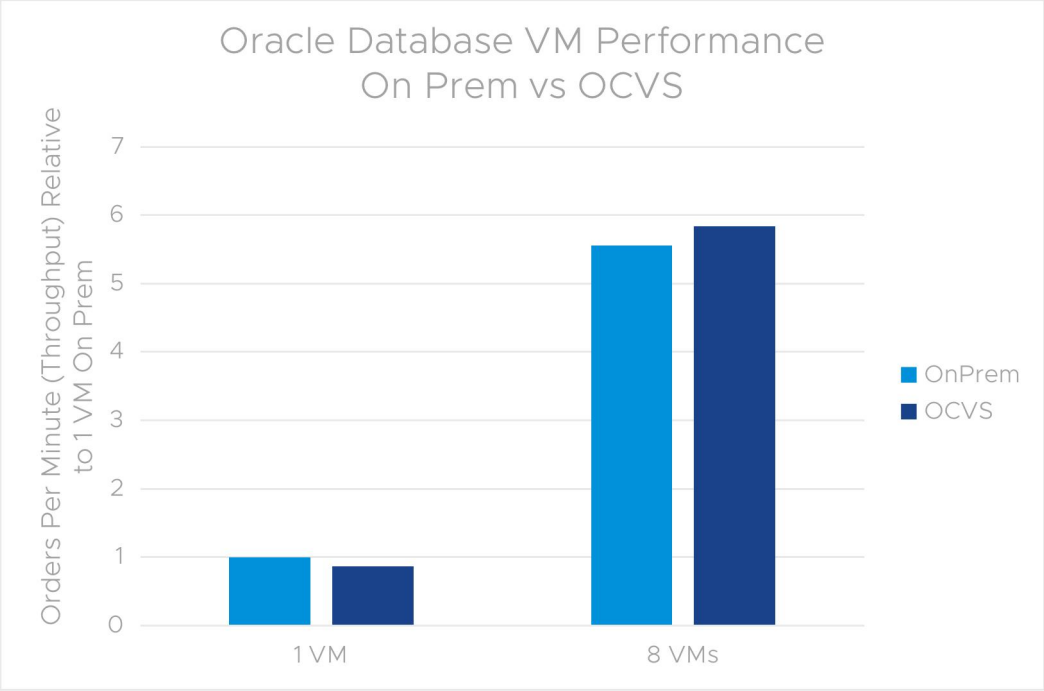

Figure 5. The number of orders per minute that **Oracle Database** can perform is about the same when comparing the database running on-premises to using OCVS

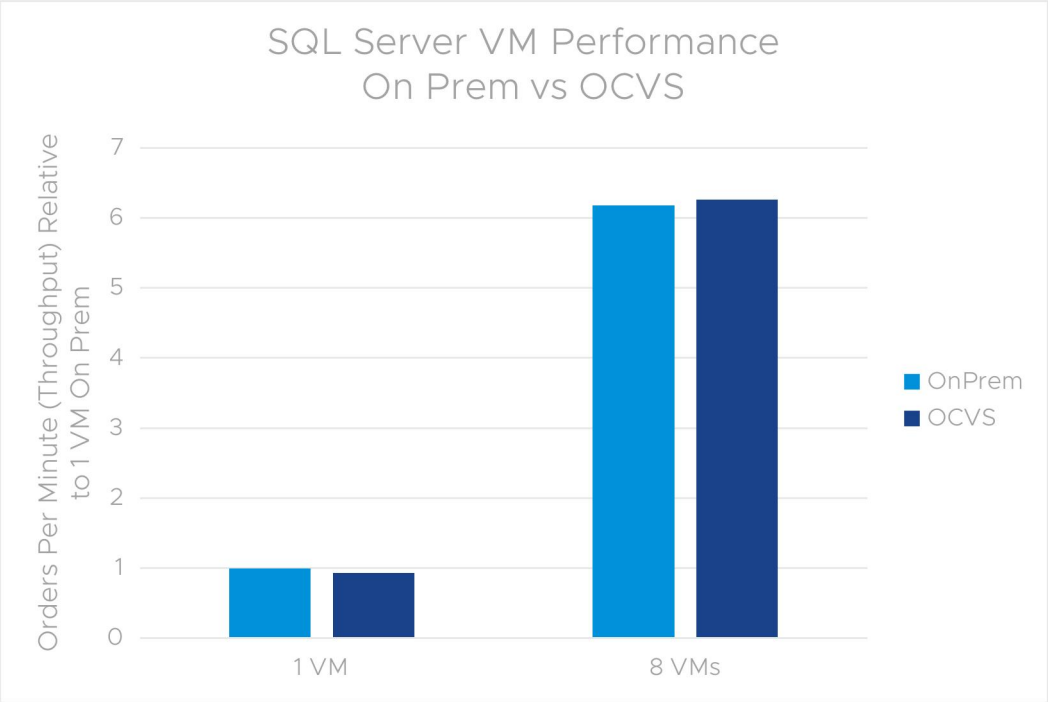

Figure 6. The number of orders per minute that **SQL Server** can perform is about the same when comparing the database running on-premises to using OCVS

In the tests to measure the performance gain from the EPYC AMD scheduler enhancement from VMware, a gain of 10% and 5% was found (figure 7). This shows that the performance gains built into the scheduler for the AMD EPYC processor continue to provide benefits in the OCVS environment.

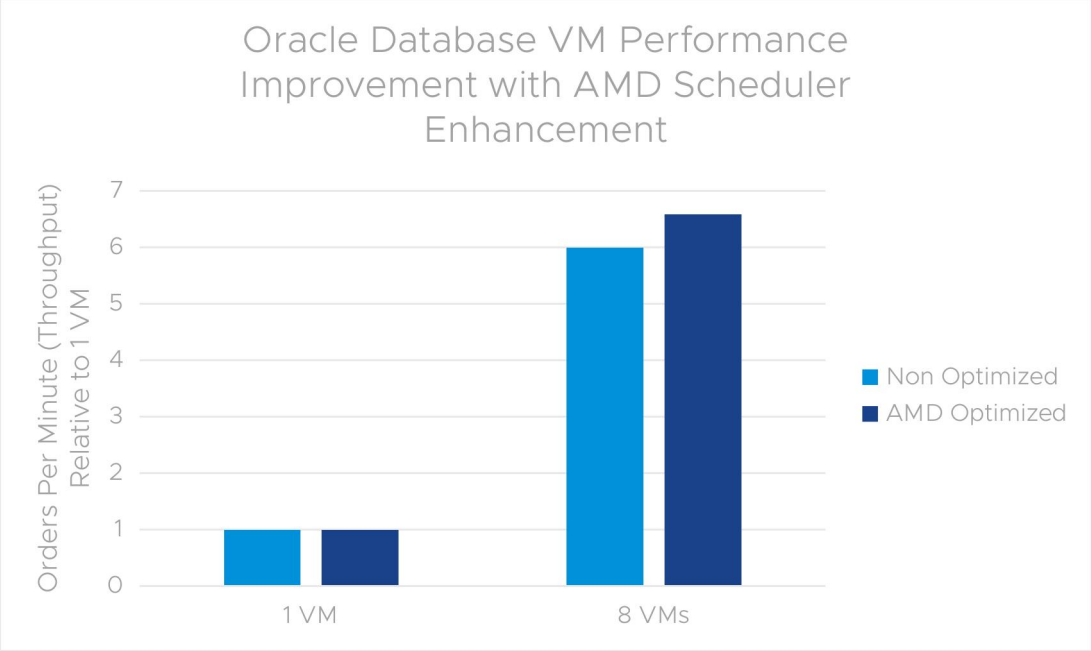

Figure 7. The **Oracle Database VMs** performed slightly more orders per minute with the AMD scheduler enhancement on-board than without it

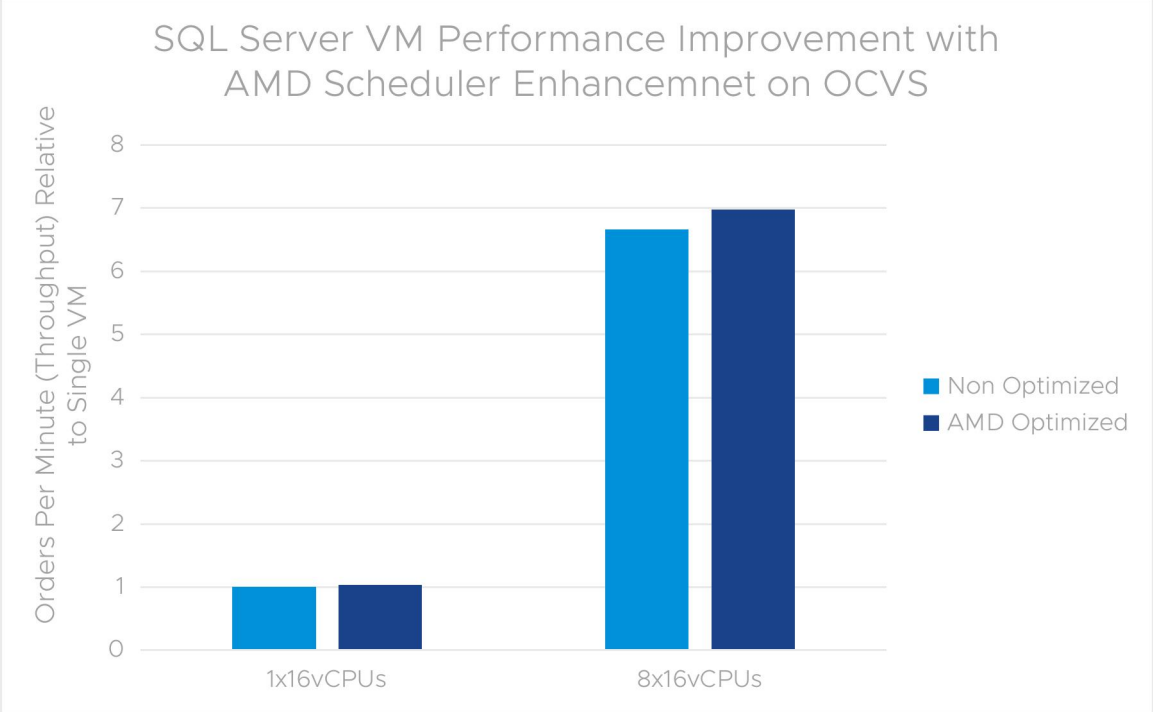

Figure 8. The **SQL Server VMs** performed slightly more orders per minute with the AMD scheduler enhancement on-board than without it

NSX: Accelerating Oracle Workloads in the Cloud

NSX can be used to connect and protect applications across on-prem and cloud infrastructure, including VMs, containers, and bare metal systems. VMware NSX Data Center is a complete Layer2-Layer7 networking and security virtualization platform that brings the public cloud experience to the private cloud.

An additional benefit of network virtualization is that by abstracting and virtualizing network services with NSX, customers can increase the overall performance of typical day 2 operational tasks that are network-intensive such as data loads, backup/recovery, batch processing, and others—essentially any management operation that involves the movement of data.

This performance analysis and study conducted by VMware are based on typical data movement operations related to an Oracle Business Intelligence landscape. The analysis showed that some operations in that environment could be accelerated by 2x to 3x with NSX and intelligent VM placement. This is accomplished through the ability of virtual networking to simplify network paths, often referred to as the elimination of network hairpinning. To validate VMware's result, Deloitte conducted their own independent performance study using Apache Bench HTTP-based tool, which showed gains ranging from 1.9x to 7.6x for multi-tier applications like Oracle.

Elimination of Network Hairpinning

The results of these studies clearly show the benefits that can be achieved in some scenarios with the elimination of network hairpinning on enterprise applications. The National Institute of Standards and Technology (NIST) defines *hairpinning* as "the performance penalty due to increased latency involved in routing the virtual network traffic to the physical network outside the virtualized host and then back to the virtual network inside the virtualized host." In simpler terms, virtualizing the network services like routing and firewalls simplifies the network paths, which can result in improved performance.

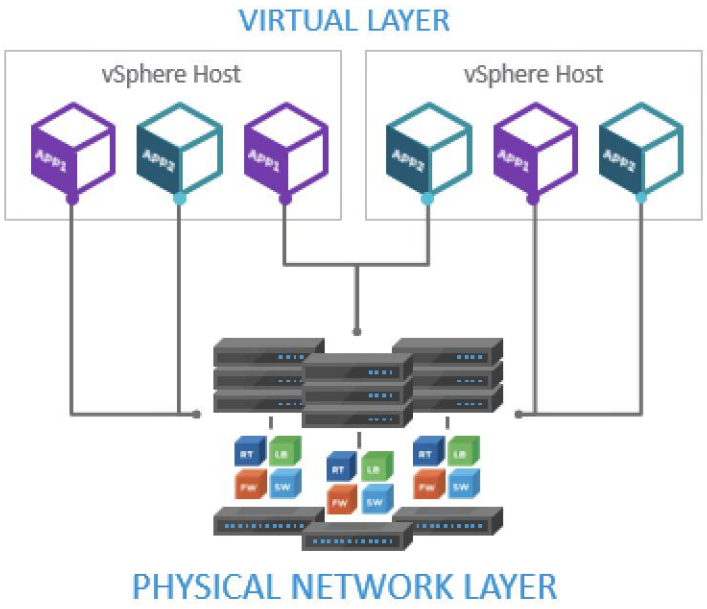

Figure 9. Without NSX, network services communicate between application tiers, and VMs must traverse the physical network layer

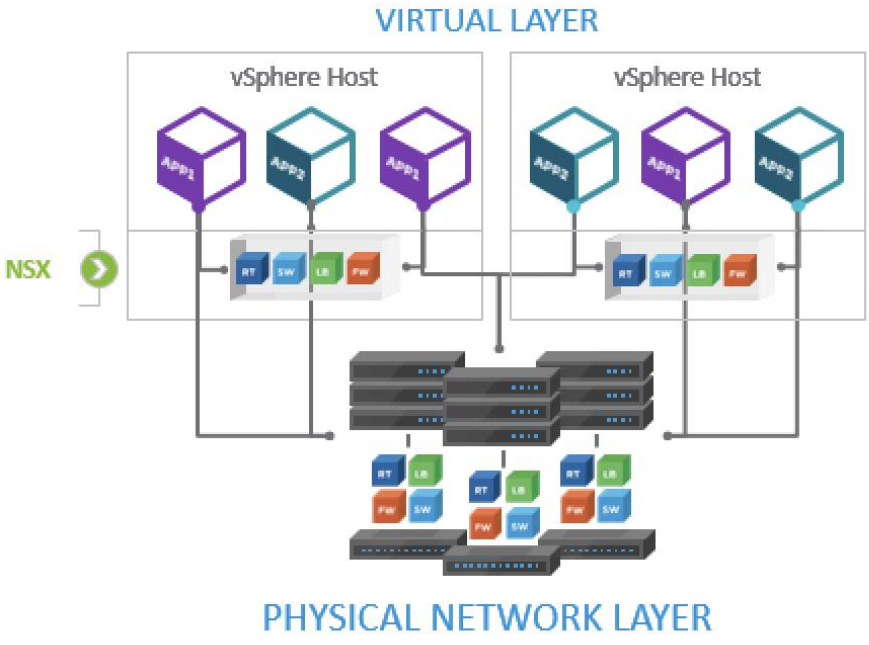

Figure 10. With NSX, eliminate network hairpinning and colocate VMs to enable efficient app communication at memory speed inside the virtualized network

Testing Methodology

There is a common misconception when using a standard or distributed switch that when VMs reside on the same host, all network communications never leave the host. This is an incorrect assumption: standard and virtual distributed switches are layer 2 constructs and have no routing capabilities; therefore, they must send traffic from different subnets or VLANs to the physical network to be routed. This can create a hairpin where traffic from VMs on the same host is sent out to the physical router and back in an inefficient path.

Using an Oracle Business Intelligence landscape configured with and without NSX, measurements were recorded when VM-to-VM communications were initiated from the same host and on different hosts. On the Oracle BI landscape without NSX, virtual distributed switches were used. The results showed performance gains with NSX and significant gains when both VMs were on the same host. This is due to VM-to-VM network communications being executed in the kernel space, essentially as memory references.

Test Results

Figure 11 represents our internal testing showing 1.3 to 2.7 Oracle application performance gains.

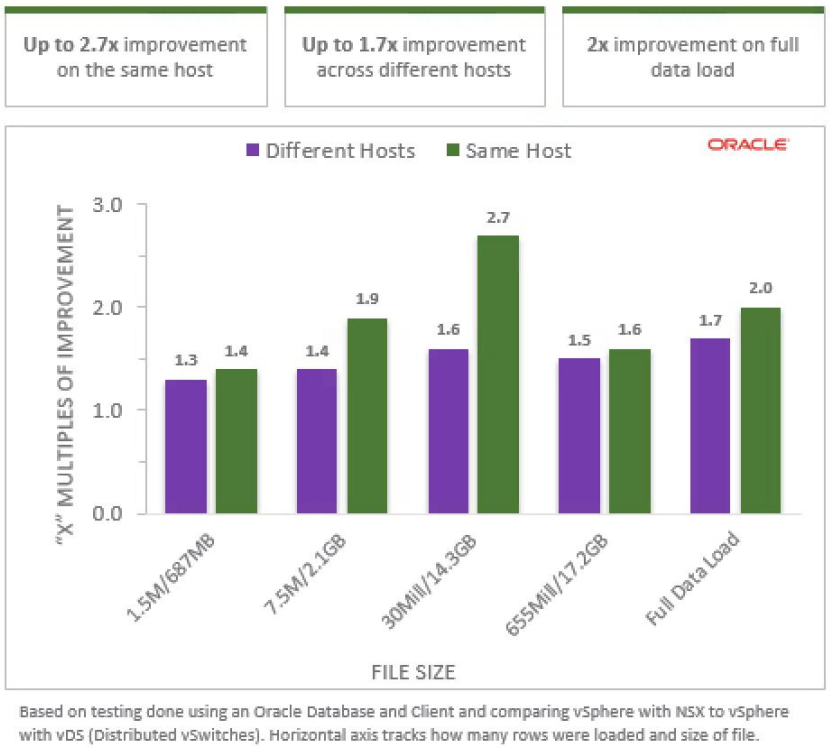

Figure 11. Our internal testing showis 1.3 to 2.7 gains in Oracle application performance

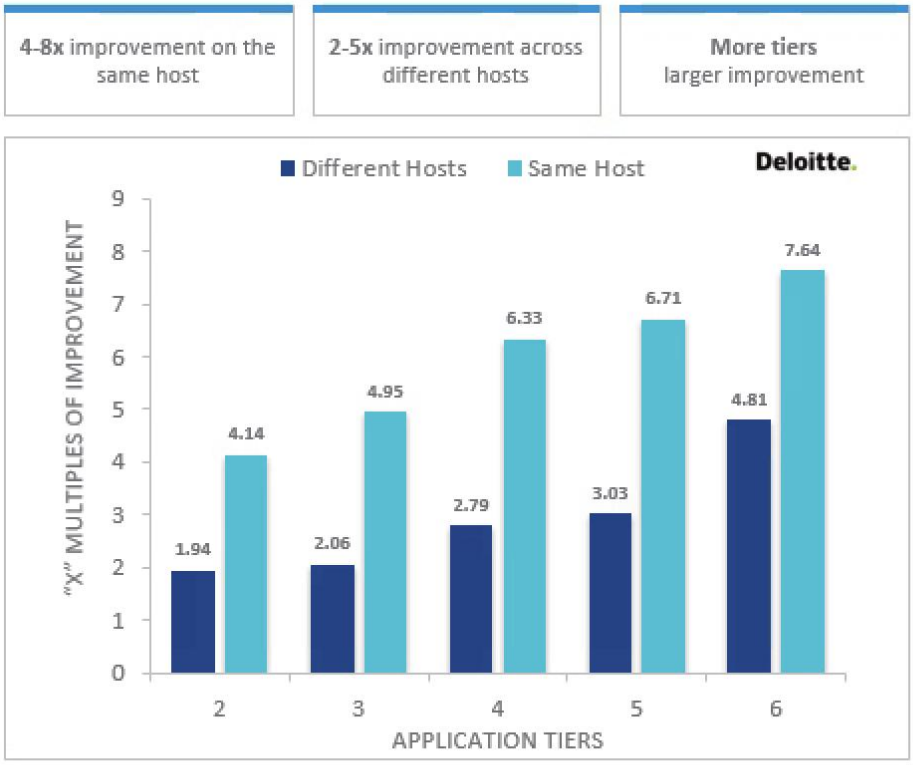

Figure 12. Testing showing performance gains from optimal placement of VMs with NSX using Apache Bench HTTP-based tool

VMware Aria Operations for Networks

VMware Aria Operations for Networks[1] has vast capabilities to monitor, discover, analyze, and build an optimized, highly available, and secure network infrastructure across clouds. Here, we focus on its application of discovery and visibility. For a more in-depth understating of VMware Aria Operations for Networks, see the link provided in the "References" section of this paper.

[1] Formally vRealize Network Insight

Application Discovery

For any successful migration, modernization, or optimization of a customer's cloud landscape, application discovery and visibility are essential. VMware Aria Operation for Networks provides visibility into applications and constructs meaningful insights using machine learning (ML) to discover which VMs are part of each application. VMware Aria Operations for Networks can provide the insights necessary to accelerate enterprise application performance by using ML to discover application boundaries automatically. Traffic flows are analyzed to determine application boundaries, group VM communications, and determine tiers within the application based on similar traffic patterns. The ML algorithms also detect shared services, identifying services like Active Directory or DNS.

Exposing Optimization Opportunities

There are many tools and standard dashboards which can be used to determine how application performance can be optimized and the degree of optimization which can be achieved. The "honeycomb widget" in figure 13 shows discovered applications in a way that lets admins filter and sort components visually and group them into user-defined applications. Analyzing the applications' traffic flows can help determine which workloads on the network communicate and over which protocols. Admins then group these components into applications and optionally mark components as shared between applications.

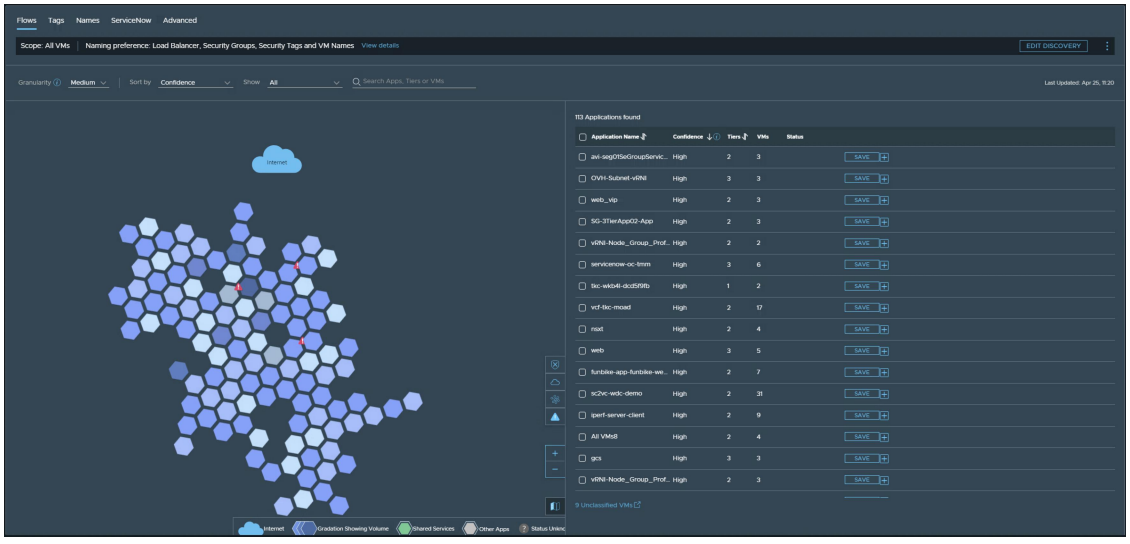

Figure 13. Applications widget

Dashboards and Visualizations

After determining the application boundaries and dependencies, understanding the traffic distribution can be used to determine the degree of optimization which can be achieved in a landscape. A key indicator is the Traffic Distribution Overview dashboard included with VMware Aria Operations for Networks as shown in figure 14. The **Routed** traffic indicates how much traffic is being routed externally to the physical layer from the virtual layer, detecting potential network hairpinning. The higher the percentage, the greater the opportunity there is to eliminate or reduce network hairpinning and improve application performance.

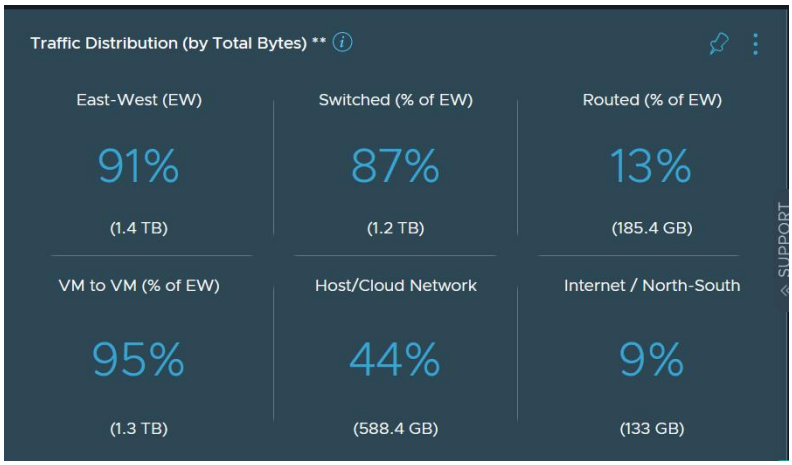

Figure 14. Traffic Distribution widget

Further drilling down on the Application Path Topology widget shown in figure 15 can identify which physical network devices and services the application traverses and then virtualize them with NSX. By minimizing network flows to the physical layer, application traffic can be isolated to the virtual layer, thereby eliminating network hairpinning.

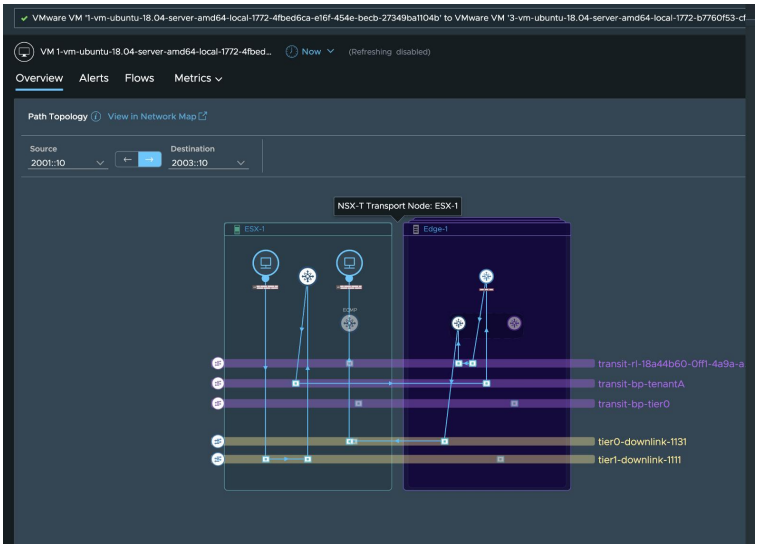

Figure 15. Application Path Topology widget

The Guided Network Troubleshooting feature in VMware Aria Operations for Networks helps users see the virtual and physical network resources that applications and VMs depend on, directly or indirectly, and narrow down potential root causes of application performance problems. See figure 16.

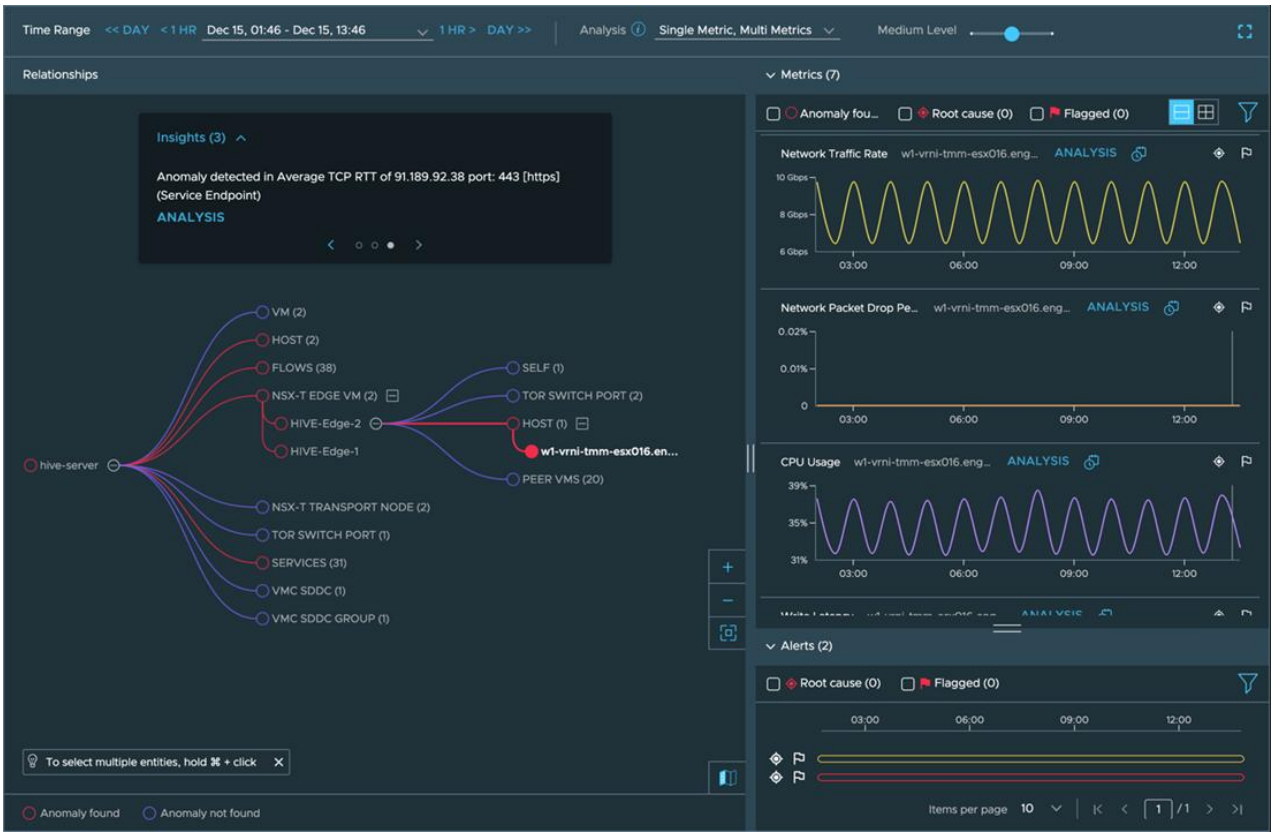

Figure 16. Troubleshoot latency using Guided Network Troubleshooting with VMware Aria Operations for Networks

There are many prepopulated dashboards and widgets within VMware Aria Operations for Networks that can be used to derive the information necessary to eliminate network hairpinning. This section illustrated just one technique to accomplish application acceleration.

Conclusion

Oracle Cloud VMware Solution extends VMware vSphere into the public cloud with high performance solution offerings. OCVS solutions allows customers to run their enterprise applications with the same performance and optimizations that are possible with vSphere on-premises.

References

- Performance Optimizations in VMware vSphere 7.0 U2 CPU Scheduler for AMD EPYC Processors
- Getting Started with Oracle Cloud VMware Solution e-book
- VMware Hybrid Cloud Best Practices Guide for Oracle Workloads
- VMware Cloud Provider Platform
- VMware Aria Operations for Networks
- Gaining Application Visibility with VMware Aria Operations for Networks
- Oracle Cloud VMware Solution Release

About the Authors

Bob Goldsand, Staff Global Solutions Architect | VMware Strategic Ecosystem and Industry Solutions

Todd Muirhead, Staff 2 Engineer R&D | VMware Performance Engineering

Acknowledgments

We thank Oracle, Deloitte, and AMD as well as our VMware colleagues for their support and collaboration:

Sheryl Sage, Senior Principal Product Manager, Oracle Cloud Infrastructure | Oracle

Praveen Kumar Pedda Vakkalam, Principal Solutions Architect, Oracle Cloud Infrastructure | Oracle

Adeel Amin, Principal Product Manager, Oracle Cloud Infrastructure | Oracle

Sankar Prabhukumar, Senior Manager, Hybrid Cloud Solutions | Deloitte Consulting

Joseph D'Alessandro, Alliance Marketing Manager | Deloitte Consulting

Andrew Ruffin, Senior Product Manager, Data Center Software & Systems | AMD

Qasim Ali, Senior Staff Engineer, Performance Engineering, Platform Services R&D | VMware

Brighten Godfrey, Technical Director, Modern Apps and Management Business Group | VMware

Sehjung Hah, Senior Product Line Marketing Manager, Modern Apps and Management Business Group | VMware

Sesh Sayani, Senior Director Product Management, Modern Apps and Management Business Group | VMware

Steve Dockar, WW OCVS Field Solution Architect Director, Hyperscalers | VMware