Executive Summary

A modern compute environment is key to remaining competitive. The traditional approach for deploying applications and services cannot deliver innovation at the pace that today’s businesses require. In addition, as data volumes grow, enterprises struggle to get more value out of their data. Data silos and cumbersome data management and analytics processes hinder discovering business insights that can drive competitive advantage. What’s more, as applications move to the edge in industries such as retail, establishing secure connectivity between the core data center, the cloud, and the edge becomes crucial to success.

Addressing these challenges involves replacing legacy hardware and software with modern, multicloud-capable solutions that can accelerate and streamline the entire software and hardware provisioning, deployment, and maintenance lifecycle. The multicloud analytics solution reference architecture described here deploys VMware Cloud Foundation on Dell VxRail, which is based on Intel® technologies. It offers an easily deployable platform for robust multicloud analytics, as illustrated in Figure 1. This solution helps eliminate data silos and provides security-enabled infrastructure, operations, and connectivity across private clouds, public clouds, and the edge. It offers businesses greater agility, operational efficiency, and future-ready scalability, with best-of-breed hardware and software technologies delivered as a highly integrated full stack.

This paper introduces the solution architecture and its business applicability to analytics in a multicloud environment. It begins with discussions of the business challenges and corresponding benefits of the solution. It then explores each of the solution building blocks in turn, including VxRail, the Intel technologies it is based on, and VMware Cloud Foundation 4.4. Next, it outlines a series of use cases to show how the solution applies to real-world business challenges. Finally, this paper presents test results that demonstrate performance on sample deep-learning inference workloads. The appendixes document the hardware and software used in the testing as a reference for use by implementation teams.

Deploy and manage data-intensive workloads from edge to cloud, taking advantage of high-performance 3rd Generation Intel® Xeon® Scalable processors and software optimized for Intel architecture

Figure 1. Multi-cloud analytics solution.
Business Challenges

Today’s enterprises want the flexibility to run analytics workloads where they make most sense—in the on-premises core data center, in one or more public clouds (multicloud), and at the edge. But to make this flexibility operationally feasible, there must be a way to efficiently manage all the analytics workloads, wherever they reside. Without a single pane of glass, management costs rapidly spiral out of control, application development becomes inconsistent, and performance may suffer.

Enterprises seek analytics infrastructure that is characterized by reduced downtime, less setup time, easier maintenance, and lower overhead costs—without sacrificing performance. Legacy data centers cannot take advantage of the cost efficiencies and new technologies available in a multicloud analytics environment. Nor can such data centers adapt to changing workload requirements quickly and nimbly.

For companies with outdated data center technologies, meeting these challenges requires replacing legacy hardware and software with modern, hybrid-cloud-capable analytics solutions. These solutions can accelerate the entire software and hardware provisioning, deployment, and maintenance lifecycle along with application development, testing, and delivery. But, whether it is an on-premises machine-learning cluster or a remote branch office analytics cluster, companies may find assembling and maintaining multicloud infrastructure daunting.

Solution Value

Intel, VMware, and Dell have teamed up to engineer a multicloud analytics solution to help take the guesswork out of building multicloud analytics. It provides a simple, security-enabled, and agile cloud infrastructure for on-premises, as-a-service public cloud, and edge analytics workloads. The foundation of the solution architecture is VMware Cloud Foundation on Dell VxRail—which includes innovative Intel technology—to provide a unified Software-Defined Data Center (SDDC) platform for running and managing private cloud, multicloud, and edge containerized analytics workloads.

Deploying VMware Cloud Foundation on Dell VxRail delivers full-stack integration across the infrastructure hardware layer and VMware cloud software that is not possible with other infrastructure. For example, VMware Cloud Foundation has been co-engineered to integrate with VxRail API to automate hardware and software lifecycle management and deployment, with flexible topologies that range from clusters of nodes to integrated rack offerings. Other VxRail integrations (such as vCenter plug in, SDDC Manager and VxRail Manager integration, and VxRail architecture awareness built into Cloud Builder) deliver a turnkey hybrid cloud user experience and simplify operations.

Enterprises can use the end-to-end multicloud analytics solution to quickly launch database processing and AI, and scale workloads to accommodate future needs. The unified cloud solution presented here can run containerized applications and traditional VMs that are located in an on-premises data center as well as in the public cloud, such as on Amazon Web Services (AWS) and Microsoft Azure.

Solution Benefits: VMware Cloud Foundation on VxRail

Running VMware Cloud Foundation on Dell VxRail infrastructure provides a joint value proposition compared to conventional servers that is greater than the sum of its parts:

- **Unified platform with intrinsic security** for running, managing, and seamlessly connecting traditional and cloud-native applications across private cloud, multicloud, and edge environments.

- **Full-stack automated lifecycle management (LCM)** for seamless infrastructure evolution with curated and automated end-to-end full stack non-disruptive hardware and software patching and upgrades.

- **Simplified support**, with a single point of global support 24x7 for both the hardware and software throughout the entire solution lifecycle.

- **Deep integration between SDDC Manager and VxRail Manager** to unify deployment, configuration, and lifecycle management of ESXi, vSAN, and hyperconverged infrastructure software and hardware.

- **Flexible storage extensibility** using VxRail dynamic nodes to integrate and support configuration of Dell PowerMax, PowerStore-T, and Unity XT external storage as cluster principal storage.

- **Optional data protection solutions** that unify the backup and recovery environment for Cloud Foundation on VxRail infrastructure along with VM and container-based workloads, eliminating silos and compatibility issues while enabling quick and easy deployment and scaling.

- **High performance across workloads** through the inclusion of accelerated analytics and increased throughput with Intel architecture-optimized deep-learning frameworks and flexible Dell VxRail hardware configurations powered by Dell PowerEdge server platforms.
Solution Architecture Building Blocks

The hybrid/multicloud structure of the solution allows enterprises to extend available resources and easily distribute analytics workloads between on-premises, public cloud, and the edge. VMware SD-WAN is used to provide reliable and secure network connectivity over public internet from any to any location (on-premises to the edge and to public cloud and vice versa). The building blocks of the solution are illustrated in Figure 2.

In addition to VMware Cloud Foundation on VxRail, the multicloud analytics solution incorporates VMware components that include VMware Secure Access Service Edge (SASE) with VMware Software-Defined WAN (SD-WAN), VMware Tanzu Mission Control, and VMware vRealize Suite to provide infrastructure-as-a-service capabilities. It also includes VMware services on public clouds—VMware Cloud on AWS (VMC) and Azure VMware Solution (AVS). Container provisioning and lifecycle management are provided by VMware Tanzu Kubernetes Grid (TKG).

Dell VxRail

VxRail, powered by Dell PowerEdge server platforms and VxRail HCI System Software, features next generation technology to future proof your infrastructure and enable deep integration across the VMware ecosystem. VxRail HCI System Software, a suite of integrated software elements that sits between infrastructure components such as vSAN and VMware Cloud Foundation, delivers a seamless and automated operational experience. VxRail supports Intel’s latest processor, memory, and storage technologies. Flexibility across compute, memory, storage, network, and graphics options delivers performance across a wide variety of applications and workloads.

VxRail contributes to optimizing costs by running business-critical workloads on a consolidated, high-performing platform with high reliability, functionality, and performance. It helps IT teams free up resources with productivity-enhancing capabilities such as one-touch deployments and automated patching and updates. By reducing the incidence and duration of service outages, VxRail protects against revenue loss, safeguards the user experience, and improves the ability to back up, protect, and recover data. It also generates value for the business by ensuring that IT can readily support expansion and support more effective and productive development activity.

Enterprises get greater control of their workloads by deploying VxRail across on-premises data centers, the edge, and co-locations. Customers reported business results such as the following:

- **More business agility:** 71 percent faster deployment of new storage
- **Faster business development:** 114 percent more new applications per year
- **Higher performance:** 32 percent faster execution of business transactions


---

Figure 2. Dell, Intel, and VMware provide the building blocks for the multicloud analytics solution.
The placement of Intel® Optane™ SSDs and Intel® Optane™ persistent memory within the architecture.

Intel Processor, Memory, and Storage Technologies

Underlying the software components of VMware Cloud Foundation in the on-premises core data center are 3rd Generation Intel Xeon Scalable processors, Intel Optane™ persistent memory (PMem), Intel Optane SSDs, Intel SSD D7 and D5 Series, and Intel Ethernet products.

Enterprises can use Intel Optane technology to boost their VMware Cloud Foundation workload performance by placing data closer to the CPU. This technology is a class of non-volatile memory and storage media that fills the gap between high-performing volatile memory and lower-performing NAND storage and HDDs. By placing data closer to the CPU, Intel Optane technology helps architects confidently deploy an agile, high-performing infrastructure so organizations can create innovative analytics services and optimize their infrastructure investments. Intel Optane technology can be deployed in two different ways (see Figure 3):

- **Intel Optane PMem**: gives enterprises the ability to extract more value from larger datasets by combining higher capacity and native persistence in a DIMM form factor. Data can be accessed, processed, and analyzed in near real time to deliver deep insights, improve operations, and create new revenue streams. Compared to DRAM-only implementations, VxRail with Intel PMem plus DRAM in Memory Mode offers 14 percent lower node cost per VxRail node while in App Direct Mode 5,3x more IOPS and 93 percent faster OLTP response time than VxRail with all-flash vSAN.  

- **Intel Optane SSDs**: help remove data bottlenecks to accelerate transactions and time to insights, so users get what they need, when they need it. VxRail with Intel Optane SSDs sustained nearly 61 percent more transactions per minute and new orders per minute than with SAS SSDs. For tiered storage like vSAN, it is recommended to use Intel Optane SSDs in the cache tier and Intel SSD D7 or D5 Series in the capacity tier.

VMware Cloud Foundation 4.4

VMware Cloud Foundation is an integrated software stack that bundles compute virtualization (VMware vSphere), storage virtualization (VMware vSAN), network virtualization (VMware NSX), and automated lifecycle management (VMware SDDC). Integrating these capabilities into a single solution helps break down administrative silos to facilitate end-to-end support for application deployment.

The solution provides cloud management and operations with VMware vRealize Suite and Kubernetes-based container services with VMware Tanzu. It can be deployed on premises as a private cloud or run as a service within a public cloud to provide a common platform for traditional and containerized applications.

VMware Cloud Foundation 4.4 introduces several new features and enhancements that help customers deploy scalable, flexible infrastructure:

- **Enhanced workload domain deployment and lifecycle management** to support large-scale VM and container architectures
- **Integration with VMware vSphere 7 Update 3c** to deliver AI- and developer-ready infrastructure, boost data security, and help simplify operations
- **Integration with VMware vSAN 7 Update 3c**, which provides enhancements to the vSAN Data Persistence Platform for improved cloud-native storage and persistent services support
- **Enhanced management flexibility** to better accommodate asynchronous vRealize Suite component upgrades
- **Platform and security operations enhancements** include stronger security mechanisms to improve the management and administration of security settings within VMware Cloud Foundation

For more details about what’s new in VMware Cloud Foundation 4.4, visit the release announcement. More information about the VMware Cloud Foundation 4.4 on Dell VxRail 7.0.320 release of the co-engineered turnkey platform can be found here.

VMware Cloud Foundation includes access to the Tanzu Application Catalog, which contains more than 70 Kubernetes applications and components from the Bitnami collection that are maintained and verified for use in production environments. Among these applications are popular analytics tools like TensorFlow, MxNet, PyTorch, and many others.

Dell APEX Cloud Services

APEX Cloud services is a breakthrough portfolio of as-a-Service offerings that simplify digital transformation while increasing IT agility and control. APEX makes it easier than ever to leverage Dell Technologies innovation when and where you want it, while easing or eliminating infrastructure management. APEX helps you react quickly to capture new opportunities and ensures that your technology stays aligned with your business requirements. With APEX, you are always in full control, minimizing risk and maximizing performance, all on your terms.
Benefit from the Dell Technologies innovation you know and trust delivered to you with the combined benefits of public and private cloud—however you choose to consume it and wherever you want to deploy it. You decide what is right for you: on-premises in your data center, out to your edge environments, in co-location facilities adjacent to your cloud providers, or natively in a public cloud. Wherever APEX goes, you get a more consistent cloud experience across your enterprise—throughout your multi-cloud, multi-edge, and multi-data center IT landscape—giving you more flexibility and choice to achieve strategic objectives.

The APEX portfolio also includes a managed solution that gives you the ability to combine the simple operations of public cloud and the control, security, and performance of private cloud with APEX Cloud Services with VMware Cloud. Designed for both virtualized and containerized VMware workloads, the infrastructure is owned and managed by Dell Technologies so you can focus on driving innovation and not have to worry about lifecycle management of firmware and software updates, patching, upgrades, and remediation.

**Use Cases**

The combination of VMware Cloud Foundation and Intel technology running on VMs or in containers can support a wide variety of use cases, as discussed in this section.

**DevOps**

VxRail is the only jointly engineered system with deep VMware Cloud Foundation integration, making it ideal for existing vSphere customers who want to create and operate Kubernetes on-premises. VxRail comes fully integrated and ready to deploy, helping accelerate adoption of Kubernetes infrastructure such as VMware Tanzu. Cloud-like resource pooling, elasticity, agility/speed, and programmability accelerate Kubernetes infrastructure delivery, ensuring that developers always have the resources they need to deliver modern applications. New nodes can be non-disruptively added to a cluster in as little as 15 minutes, whether to provide additional compute and storage to support application development or to roll out a new app into production.

**Deep-Learning Inference**

Inference is compute-intensive and can benefit from innovations such as Intel Deep Learning Boost (Intel DL Boost) with Vector Neural Network Instructions (VNNI)—a special instruction set that speeds up inference—available starting with vSphere 7 and ESXi 7.0, which are foundational components of the VMware Cloud Foundation 4.4 platform.

Enterprises need high-performance data analytics and AI to remain competitive. They require flexible solutions that can run traditional data analytics and AI applications. The VMware multicloud platform includes components that take advantage of performance optimizations for Intel hardware. Intel supports developing machine-learning workloads at multiple layers in the solution stack.

These building blocks enable enterprises to quickly operationalize analytics, AI, and machine-learning workloads because they are already optimized for Intel architecture and have been verified in wide production use. The multicloud analytics solution reference architecture significantly improves deep-learning inference throughput, as discussed below.

**Retail at the Edge**

For retail stores, healthcare, and smart industry, running workloads closer to customers and closer to the sources of the data can improve performance, which can lead to increased customer satisfaction. VMware Cloud Foundation makes it easy to deploy and manage remote workloads, using the same technology that is used for public and private cloud workloads.

The proposed use case illustrates a solution that can increase retail customer engagement and improve the shopping experience. We include three scenarios:

- **Product recommendations.** When the client shows interest in a specific area or department, we can use a machine-learning algorithm to send personalized product recommendations. Based on people’s similar choices and the customer’s position in the store, the algorithm creates a list of the most relevant products. The customer is notified and can check the personalized recommendations using a mobile application. The process occurs every time the system discovers a new customer interest.

- **Presence detection.** We use deep-learning techniques and image recognition algorithms to detect customers in the Customer Service area. Cameras installed in the store send images to the deep-learning pipeline. When such an event occurs, the store staff is informed.

- **Hesitation detection.** When a customer is wandering around the store without stopping, the business rules engine assumes the customer is looking for something, is lost, or may need assistance. A notification—personalized with any known data about the individual customer—is sent to store staff so they can quickly offer assistance.

**Data Warehousing and Analytics**

Data warehouses are considered one of the core components of business intelligence. They are a central location to store data from one or more disparate sources as well as current and historical data. The VMware hybrid/multicloud platform supports data warehousing, including industry-proven solutions based on Microsoft SQL Server 2019 or Oracle Database 19c.
Results: Deep-Learning Inference

Image classification is a common use case for deep learning. Our tests benchmarked the ResNet50 v1.5 topology with int8 and fp32 precision, using the TensorFlow distribution from the Intel architecture-optimized container stack with Intel’s Model Zoo pretrained models.

We ran two tests (see Appendixes for hardware and software used in testing):

- Performance comparison of default TensorFlow container versus the Intel architecture-optimized TensorFlow container
- Performance comparison of fp32 precision versus int8 precision (both using Intel DL Boost with VNNI and the Intel architecture-optimized TensorFlow container)

As the following results illustrate, the hardware and software optimizations for inference have a substantial impact on improving the performance of inference. In this use case, the optimizations enabled a significant increase in throughput (frames per second). The VMware Cloud Foundation 4.4 platform is an excellent example of how software can take advantage of hardware innovations such as Intel DL Boost and VNNI to deliver faster insights.

Up to 6.78x Better Throughput by Optimizing TensorFlow for Intel Architecture

This testing compares throughput performance of the default TensorFlow container against a container using the Intel Optimization for TensorFlow, which is optimized to take advantage of Intel DL Boost and VNNI. Both containers used fp32 precision. As Figure 4 shows, framework optimizations from the Intel Optimization for TensorFlow can provide up to a 5.78x throughput improvement for the Base design and up to a 6.78x throughput improvement for the Plus design.¹

Up to 3.85x Better Throughput with int8 Precision

In this benchmark, we compared throughput performance of Intel DL Boost with VNNI using int8 precision against fp32 precision. Both containers used the Intel Optimization for TensorFlow. As shown in Figure 5, a small reduction in precision enabled up to a 3.58x throughput improvement for the Base design and up to a 3.85x throughput improvement for the Plus design.⁵

Learn More

- 3rd Generation Intel Xeon Scalable processors
- Intel Ethernet products
- Intel Optane persistent memory
- Intel Optane SSDs
- Dell VxRail
- VMware Cloud Foundation
- White paper: VMware Cloud Foundation on Dell VxRail
## Appendix A: Testing Software

### Table A1. Software versions used to test deep-learning inference.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Plus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Guest OS</strong></td>
<td>Ubuntu Server 20.04.4 LTS</td>
<td></td>
</tr>
<tr>
<td><strong>Guest OS Kernel</strong></td>
<td>5.4.0-100-generic</td>
<td></td>
</tr>
<tr>
<td><strong>Containers</strong></td>
<td>intel/intel-optimized-tensorflow:2.7.0-ubuntu-20.04</td>
<td>tensorflow/tensorflow:2.7.0</td>
</tr>
<tr>
<td><strong>AI Precision</strong></td>
<td></td>
<td>int8, fp32</td>
</tr>
<tr>
<td><strong>Other Software</strong></td>
<td>VMware Cloud Foundation 4.4; VMware vSAN 7.0 U3c; VMware vCenter Server 7.0 U3c; VMware NSX-T 3.1.3.56</td>
<td>VMware ESXi 7.0 U3c (build 19193900)</td>
</tr>
<tr>
<td><strong>VM vCPU</strong></td>
<td>42</td>
<td>56</td>
</tr>
<tr>
<td><strong>VM vRAM</strong></td>
<td>256 GB</td>
<td>256 GB</td>
</tr>
<tr>
<td><strong>Framework/Toolkit included version</strong></td>
<td></td>
<td>TensorFlow</td>
</tr>
<tr>
<td><strong>Framework URL</strong></td>
<td>TensorFlow Docker images used: intel/intel-optimized-tensorflow:2.7.0-ubuntu-20.04 and tensorflow/tensorflow:2.5.0</td>
<td></td>
</tr>
<tr>
<td><strong>Topology or ML algorithm</strong></td>
<td></td>
<td>ResNet50v1.5</td>
</tr>
<tr>
<td><strong>Compiler</strong></td>
<td>Not compiled, used Docker images</td>
<td></td>
</tr>
<tr>
<td><strong>Libraries</strong></td>
<td>Container with TensorFlow optimized with oneAPI Deep Neural Network Library (oneDNN)</td>
<td></td>
</tr>
<tr>
<td><strong>Dataset</strong></td>
<td>Synthetic data (autogenerated, --benchmark-only parameter)</td>
<td></td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td></td>
<td>int8, fp32</td>
</tr>
<tr>
<td><strong>Batch Size</strong></td>
<td></td>
<td>128</td>
</tr>
<tr>
<td><strong>Build Flags</strong></td>
<td>Not compiled, used Docker images</td>
<td></td>
</tr>
<tr>
<td><strong>KMP AFFINITY</strong></td>
<td>granularity=fine,verbose,compact,1,0/*p/'verbose,warnings,respect,granularity=fine,compact,1,0'</td>
<td></td>
</tr>
<tr>
<td><strong>NUMACTL</strong></td>
<td>Not used</td>
<td></td>
</tr>
<tr>
<td><strong>OMP_NUM_THREADS</strong></td>
<td></td>
<td>42/56</td>
</tr>
<tr>
<td><strong>Command Line Used</strong></td>
<td>python3 /tf/intel-models/benchmarks/launch_benchmark.py --in-graph $(IN_GRAPH) --model-name $(MODEL_NAME) --framework tensorflow --precision $(PRECISION) --mode inference --batch-size $(BATCH_SIZE) --benchmark-only</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix B: Recommended Hardware

### Table B1. Recommended hardware bill of materials.

<table>
<thead>
<tr>
<th></th>
<th>MANAGEMENT</th>
<th>BASE DOMAIN</th>
<th>PLUS DOMAIN</th>
<th>EDGE DOMAIN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Platform</strong></td>
<td>4x VxRail E660N</td>
<td>4x VxRail E660N</td>
<td>4x VxRail E660N</td>
<td>4x VxRail E660F</td>
</tr>
<tr>
<td><strong>CPU</strong></td>
<td>2x Intel® Xeon® Gold 5318Y processor (2.1 GHz, 24 cores)</td>
<td>2x Intel® Xeon® Gold 6342 processor (2.8 GHz, 24 cores)</td>
<td>2x Intel® Xeon® Gold 8362 processor (2.8 GHz, 32 cores)</td>
<td>2x Intel® Xeon® Silver 4310 processor (2.1 GHz, 12 cores)</td>
</tr>
<tr>
<td><strong>Memory (DRAM)</strong></td>
<td>512 GB (16x 32 GB DDR4 DRAM 3200 MHz)</td>
<td>512 GB (16x 32 GB DDR4 DRAM 3200 MHz)</td>
<td>256 GB (16x 16 GB DDR4 DRAM 3200 MHz)</td>
<td>256 GB (8x 32 GB DDR4 DRAM 3200 MHz)</td>
</tr>
<tr>
<td><strong>Memory (PMem)</strong></td>
<td>N/A</td>
<td>N/A</td>
<td>1024 GB (8x 128 GB, Intel® Optane™ PMem 200 series 3200 MHz)</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Boot Drive</strong></td>
<td>2x Dell BOSS-S2 M.2 480 GB</td>
<td>2x Dell BOSS-S2 M.2 480 GB</td>
<td>2x Dell BOSS-S2 M.2 480 GB</td>
<td>2x Dell BOSS-S2 M.2 480 GB</td>
</tr>
<tr>
<td><strong>Storage (Cache)</strong></td>
<td>2x Intel® Optane™ SSD DC P5800X 400 GB</td>
<td>2x Intel® Optane™ SSD DC P5800X 400 GB</td>
<td>2x Intel® Optane™ SSD DC P5800X 400 GB</td>
<td>1x 1.6 TB NVMe MU</td>
</tr>
<tr>
<td><strong>Storage (Capacity)</strong></td>
<td>4x SSD – 3.84 TB RI NVMe</td>
<td>4x SSD – 3.84 TB RI NVMe</td>
<td>6x SSD – 3.84 TB RI NVMe</td>
<td>4x SSD – 1.92 TB SATA RI</td>
</tr>
<tr>
<td><strong>Network Adapter</strong></td>
<td>1x Intel E810-XXV 25 GbE SFP28 Dual Port OCP 3.0</td>
<td>1x Intel E810-XXV 25 GbE SFP28 Dual Port OCP 3.0</td>
<td>2x Intel E810-XXV 25 GbE SFP28 Dual Port OCP 3.0</td>
<td>1x Intel E810-XXV 25 GbE SFP28 Dual Port OCP 3.0</td>
</tr>
<tr>
<td><strong>Top of the Rack (ToR) Switch</strong></td>
<td></td>
<td></td>
<td></td>
<td>25/40/50/100 GbE per port Switch capabilities: Jumbo Frames, BGP</td>
</tr>
</tbody>
</table>
Base Configuration: Single-node, 2x Intel Xeon Gold 6342 processor, 1x Intel Server Board MSGCPY2UR, total memory = 512 GB (16 slots/32 GB/3200 MHz), Intel Hyper-Threading Technology = OFF, Intel Turbo Boost Technology = ON, BIOS: SE5C6200.868.0022. D64.2105220049 (0x:00d0002b1), Storage (boot): 1x Intel Optane SSD P5800X 400 GB, Storage (cache): 4x Intel SSD D7-P5510 3.84 TB, Network devices: 1x Intel Ethernet Adapter E810-CQDA2 (100 GbE), Hypervisor: VMware ESXi 7.0 U3c (build 19193900) and VMware Cloud Foundation 4.4, OS/Software: Ubuntu Server 20.04.4 LTS, 5.4.0-100-generic, 42 vCPU, 256 GB vRAM, Containers: intel/intel-optimized-tensorflow:2.7.0-ubuntu-20.04, tensorflow/tensorflow:2.7.0, ResNet50 v1.5, int8, fp32, batch size = 128

Plus Configuration: Single-node, 2x Intel Xeon Platinum 8362 processor, 1x Intel Server Board MSGCPY2UR, total memory = 1280 GB (AD mode) = 256 GB (16 slots/16 GB/3200 MHz) + 1024 GB Intel Optane persistent memory 200 series (8 slots/128 GB/3200 MHz), Intel Hyper-Threading Technology = OFF, Intel Turbo Boost Technology = ON, BIOS: SE5C6200.868.0022. D64.2105220049 (0x:00d0002b1), Storage (boot): 1x Intel Optane P1600X 118 GB, Storage (cache): 2x Intel Optane SSD P5800X 800 GB, Storage (capacity): 6x Intel SSD D7-P5510 3.84 TB, Network devices: 2x Intel Ethernet Adapter E810-CQDA2 (100 GbE), Hypervisor: VMware ESXi 7.0 U3c (build 19193900) and VMware Cloud Foundation 4.4, OS/Software: Ubuntu Server 20.04.4 LTS, 5.4.0-100-generic, 56 vCPU, 256 GB vRAM, Containers: intel/intel-optimized-tensorflow:2.7.0-ubuntu-20.04, tensorflow/tensorflow:2.7.0, ResNet50 v1.5, int8, fp32, batch size = 128


* Testing by Intel as of April 04, 2022. Results may vary. See Appendix A for details on the software used for testing.