SOLUTION BRIEF

Data Center
Hybrid Cloud

Next-Generation Hybrid Cloud Data Analytics Solution

Imagine if you could quickly operationalize analytics and AI on a hybrid cloud solution that is also optimized to run traditional enterprise workloads

Executive Summary

Enterprise architects are seeking a hybrid cloud solution that can handle traditional enterprise SQL/NoSQL workloads but can also extend to include compute and memory-hungry artificial intelligence (AI) jobs. Enterprises require both portability and performance. Portability to migrate and scale VMs across private and public cloud boundaries, and performance to support compute-heavy analytics and AI workloads. To deliver on these customer needs, Intel and VMware collaborated to create the Hybrid Cloud Analytics Solution, a platform that delivers support for advanced machine learning, while also providing a consistent cloud operating model regardless of where the workloads reside.

This solution is a unique combination of the latest hardware innovations from Intel, VMware Cloud Foundation*, container orchestration, and AI tools optimized to run on 2nd Generation Intel® Xeon® Scalable processors. These processors feature built-in inferencing acceleration through Intel® Deep Learning Boost (Intel® DL Boost) with Vector Neural Network Instructions (VNNI).

With this easy-to-deploy and comprehensive solution, enterprises can quickly operationalize database processing and AI to unlock the insights hidden in their data—and scale the solution as future needs dictate. The hybrid cloud capability provides flexibility in workload placement as well as business agility.

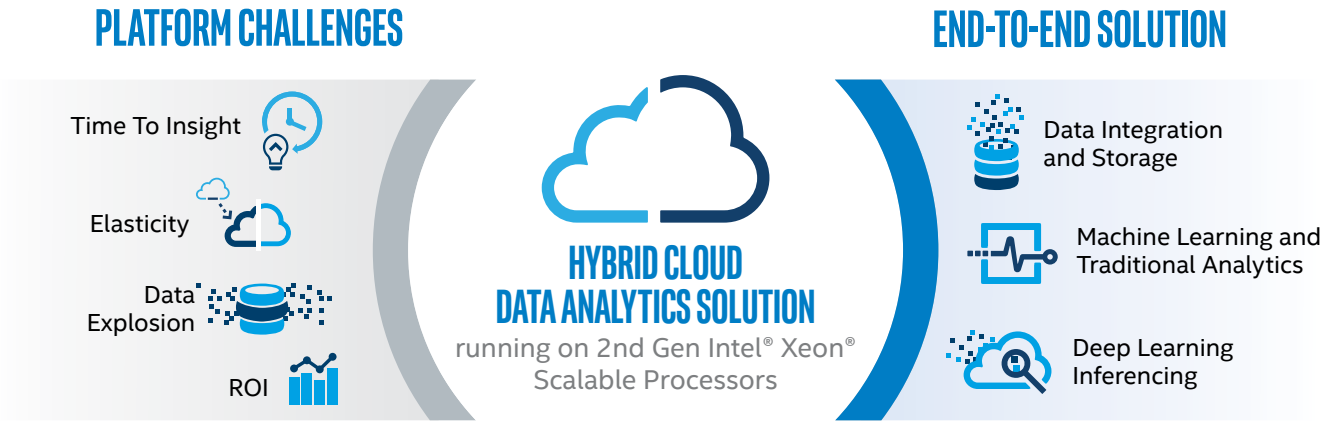


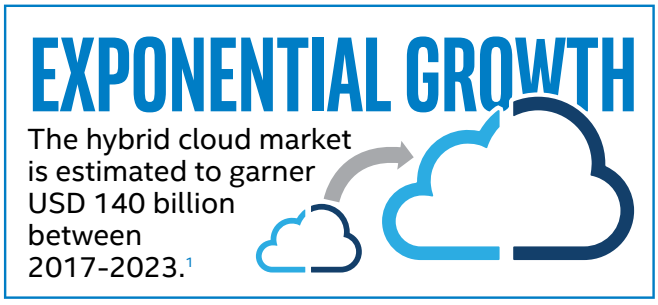
Figure 1. Meet your data analytics and AI challenges with the Hybrid Cloud Data Analytics Solution, which takes advantage of leading virtualization and manageability technologies from VMware, innovations in hardware from Intel, and optimized software from Intel.

Business Challenge: Extend Existing Workloads with AI at Scale

While many enterprises choose to run certain workloads in the public cloud, other workloads are often better suited to staying on-premises. Increasingly, enterprises want a hybrid cloud option for flexibility and business agility. In fact, the hybrid cloud market is estimated to garner USD 140 billion between 2017-2023.[1] Finding a solution that supports both traditional enterprise database workloads and artificial intelligence (AI) and machine-learning workloads–especially in a hybrid cloud–poses some significant challenges:

- **Data explosion.** With the amount of data expected to double every 12 hours by 2025,[2] businesses seek solutions that are both scalable and affordable. They can't risk investing in infrastructure that will be inadequate in just a few months or years.
- **Elasticity.** The current anything-as-a-service (XaaS) mindset requires applications that can land and run anywhere—private cloud, public cloud, or both.
- **Time to insight.** Insight-driven companies grow eight times as fast as the global gross domestic product (GDP), and customer-centric businesses that focus on digital insight are 60 percent more profitable than those that are not.[3] High-performance data analytics solutions can help organizations maintain a competitive advantage.
- **Return on investment (ROI).** Balancing performance and cost requires infrastructure utilization and efficient administration tools. The right approach to virtualization and automation can help drive ROI for data analytics in the hybrid cloud.

Solving these challenges with an end-to-end, easy-to-deploy, and high-performance data analytics solution is every enterprise's dream. And now it's possible.



Solution Benefits

- **Easy deployment.** Container solution deployed using the Deep Learning Reference Stack and VMware Enterprise Pivotal Container Service* (PKS*) lets you operationalize AI quickly and easily.
- **Future-ready.** Compute and storage resources on VMware vSAN* hyperconverged infrastructure are ready to scale up and out as needed.
- **Hybrid cloud-friendly.** As business needs change, workloads can be deployed on-premises or in a public cloud—or a combination of both using VMware Cloud Foundation*.
- **Accelerated inferencing.** While supporting traditional enterprise analytics workloads using SQL and NoSQL databases, powerful compute plus integrated, optimized artificial intelligence building blocks also speed inferencing workloads and hence shorten time to insight.

Flexible Solution Addresses the Data Life Cycle and Keeps up with Data Growth

An effective solution must address multiple phases of the data life cycle:

- **Data ingestion and storage.** Data warehousing (such as using Apache Hadoop*); a range of SQL and NoSQL databases; and extract, transform, and load capabilities are the starting point for enterprise data analytics and AI. With today's increasing amounts of data, storage needs can scale to multiple petabytes.
- **Traditional analytics and machine learning.** A good solution should support a wide variety of applications.
- **Deep learning inferencing.** This is a significant area of growth for AI, and includes such activities as image processing, face recognition, video analytics, and natural language processing (NLP).
- **Act.** This is where actionable insights are garnered from visualization and dashboard tools.

These phases can be applied across a wide variety of industries. For example, financial services institutions use AI for fraud detection, and rely heavily on efficient data storage and analysis. In healthcare, hospitals are modernizing their storage and analytics capabilities to accommodate data that is growing at a 36 percent compound annual growth rate (CAGR) through 2025.[4] Government entities are using AI to process and analyze large volumes of streaming video (which also requires vast storage capacities).

Solution Value: Fast Analytics and Inference, Easy Deployment

Big data analytics using SQL still accounts for a large portion of enterprise workloads in the data center. IDC expects worldwide revenue for big data and business analytics solutions to reach USD 260 billion in 2022, compared to a market value of USD 166 billion in 2018.[5] But AI inferencing is quickly becoming one of the top enterprise workloads. Experts predict that by 2020, the ratio between cycles of training and inference will shift from the current 1:1 ratio to well over 1:5.[6] In other words, AI inferencing is quickly becoming one of the top enterprise workloads. And the more quickly inferencing runs, the faster business can act.

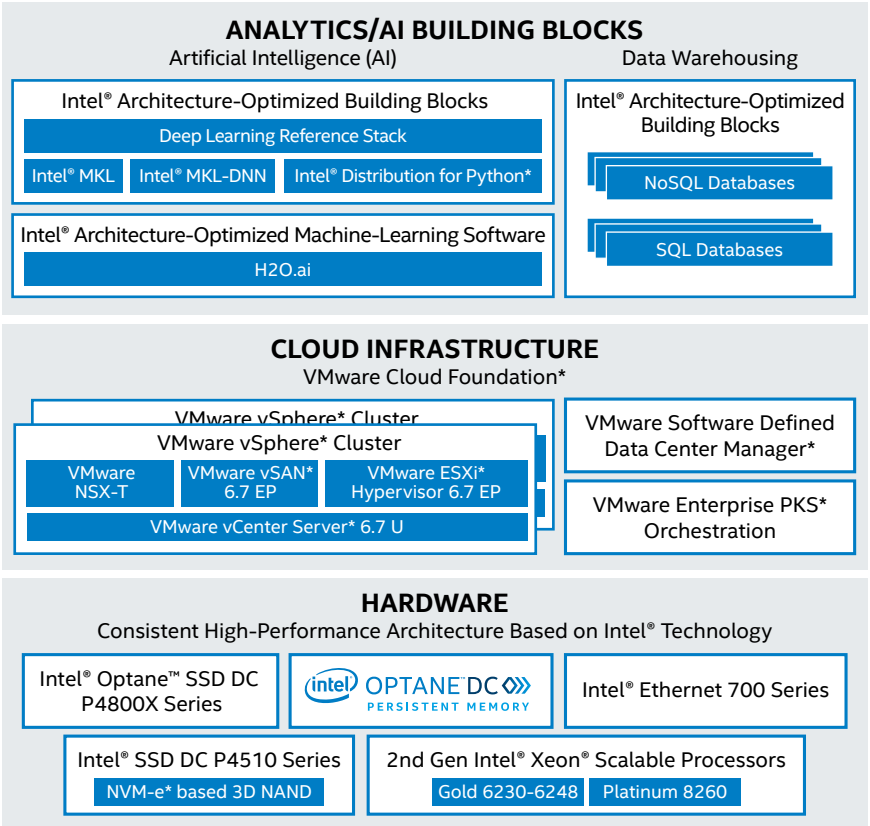
VMware and Intel have collaborated to create the Hybrid Cloud Data Analytics Solution, which combines Intel's hardware innovation and AI software expertise with VMware's deep experience with virtualization. This solution offers many benefits to enterprises that want to speed big data analytics and AI workloads:

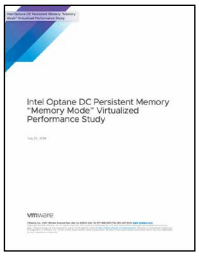
- Continue to run your traditional enterprise workloads while quickly operationalizing analytics and AI on familiar infrastructure with a turn-key solution.
- Scale from a few nodes to hundreds with a fully integrated, validated, and scalable hyperconverged infrastructure (HCI).
- Simplify the path to AI with the Deep Learning Reference Stack, which bundles many Intel® architecture optimizations for AI frameworks and libraries.
- Take advantage of VMware Cloud Foundation* for hybrid cloud deployments using a software-defined infrastructure.
- Improve total cost of ownership (TCO) with high capacity, high density, and high utilization.

You can continue to run your traditional enterprise workloads and accelerate your path to AI at scale with a solution that uses familiar and trusted technologies from both Intel and VMware—including Intel® Select Solutions for VMware vSAN* ReadyNodes.

Solution Architecture: Future-Proof Your Data Center for Modern AI Workloads

The Hybrid Cloud Data Analytics Solution (see Figure 2) consists of essentially three layers: hardware foundation from Intel, cloud/virtualization technology from VMware, and application building blocks optimized for Intel architecture. For a complete bill of materials, refer to Appendix A.





Read More about Intel® Optane® DC Persistent Memory and VMware: Intel Optane DC Persistent Memory "Memory Mode" Virtualized Performance Study

Figure 2. The Hybrid Cloud Data Analytics Solution is a complete, validated AI package—high-performance hardware from Intel, cloud infrastructure based on VMware Cloud Foundation*, and integrated optimized AI building blocks that speed inference workloads.

For fast data analytics and inferencing, the solution combines the following high-performance Intel® hardware components:

- Workload-optimized 2nd Generation Intel Xeon Scalable processors with support for Intel® Deep Learning Boost (Intel® DL Boost) with Vector Neural Network Instructions (VNNI)
- High-capacity Intel® DC 3D NAND SSDS for the vSAN capacity tier, which can scale to petabytes if necessary
- Low-latency Intel® Optane™ DC SSDs for the vSAN caching tier
- Reliable, fast Intel® Ethernet networking components

2nd Generation Intel® Xeon® Scalable processors with Intel DL Boost are a natural choice for inferencing workloads. These innovative processors are enhanced specifically to run performance-hungry AI applications, right alongside more traditional data center and cloud applications.

Intel DL Boost features VNNI,[7] which speeds up inferencing because it uses only a single instruction, instead of three separate instructions, to process deep learning convolution operations. What does this mean in the real world? As shown in Figure 3, AI workloads powered by 2nd Generation Intel Xeon Scalable processors, running on the Hybrid Cloud Data Analytics Solution, can complete inferencing up to 14x faster than a previous generation Intel Xeon Scalable processor.[8]

VMware Cloud Foundation is VMware's unified Software-Defined Data Center (SDDC) platform for the private and public cloud. Cloud Foundation brings together VMware's vSphere*, vSAN*, and NSX* into a natively integrated stack, delivering enterprise-ready cloud infrastructure with automation, self-driving operations, and management capabilities for a consistent experience across clouds. VMware Cloud Foundation can help simplify the path to the hybrid cloud while increasing admin productivity.

Simplified virtualization and container management is provided through VMware SDDC Manager* and VMware Enterprise PKS*.

An integrated AI inferencing solution is layered on this foundation, which includes the Deep Learning Reference Stack, the Intel® Distribution for Python*, and the Intel® Math Kernel Library (Intel® MKL). The Deep Learning Reference Stack includes additional optimized building blocks such as the Intel® Distribution of OpenVINO™ toolkit, which can help accelerate the development of high-performance deep-learning inference algorithms, and Intel Optimized for TensorFlow*.

Machine-learning software (like H2O.ai*) that has been optimized for Intel architecture can be included, depending on business needs, and will efficiently run on this solution.

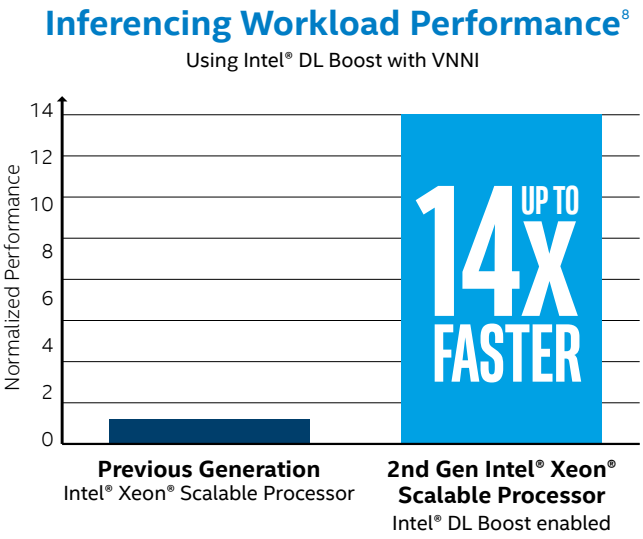


Figure 3. Intel® Deep Learning Boost (Intel® DL Boost) with Vector Neural Network Instructions (VNNI)—available only with 2nd Generation Intel® Xeon® Scalable processors—can speed up inferencing workloads by up to 14x.

Need More Memory?

For memory-centric workloads—including in-memory databases and machine-learning inferencing—you can add Intel® Optane™ DC persistent memory to the Hybrid Cloud Data Analytics Solution.[9] This disruptive technology is an additional level of memory to bridge the gap between DRAM and NAND SSD, in DIMM form factor. Intel Optane DC persistent memory can be configured at up to 3 TB per CPU socket. That means fewer I/O operations and less latency. Intel Optane DC persistent memory provides near-DRAM performance at up to 39 percent lower cost than a DRAM-only system.[10]

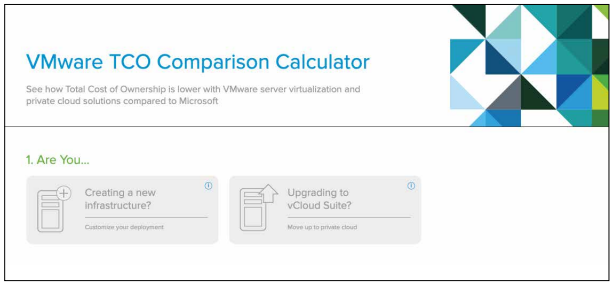
In tests performed by Intel and VMware, Intel Optane DC persistent memory resulted in 33 percent more memory than a server with DRAM only and a 22 percent increase in VM density.[11] For more information about how Intel Optane DC persistent memory can benefit in-memory databases and AI applications, read the "Intel Optane DC Persistent Memory 'Memory Mode' Virtualized Performance Study."

HCI Offers Many Benefits

VMware's hyperconverged infrastructure (HCI) provides a mix of compute, storage, and management resources that run on industry-standard servers. HCI supports a software-defined environment that can help solve today's IT challenges. The benefits of HCI include:

- **Simplicity.** Tightly-integrated hardware and software can eliminate the complexities associated with fiber channel networks, purpose-built storage, and complex storage array planning—thereby lowering operational costs.
- **Total cost of ownership.** A scale-up/scale-out architecture with industry-standard servers means IT can buy the compute and storage they need now, and increase capacity as needed, either by adding storage drives to existing nodes (scale-up) or nodes to a cluster (scale-out).
- **Agility.** HCI makes it easy to add the latest technologies to your infrastructure to improve compute, network, or storage performance.

Learn More at VMware TCO Comparison Calculator



Conclusion

With the need for high-performance data analytics and AI on the rise in an increasingly digital world, enterprises seek flexible solutions that can run traditional data analytics and AI applications on-premises or in the public cloud. The Hybrid Cloud Data Analytics Solution provides exactly that—a hybrid-cloud-friendly solution that can run workloads such as online analytics processing as well as machine-learning inferencing. And it doesn't just run them—it excels at them. Inferencing workloads running on 2nd Generation Intel Xeon Scalable processors can run up to 14x faster than on older technologies.[12] And faster inferencing means faster insights, which can translate to competitive advantage.

Based on the latest hardware and software innovations from Intel, combined with familiar and trusted virtualization technologies from VMware, the Hybrid Cloud Data Analytics Solution is an end-to-end, ready to deploy solution that can meet enterprise's current analytics needs and can easily scale to handle the needs of the future.

Find the solution that is right for your organization. Contact your Intel representative or visit **Intel® Select Solutions for VMware vSAN**.

Learn More

You may find the following resources helpful:

- Deep Learning Reference Stack
- 2nd Generation Intel® Xeon® Scalable processors
- Simplifying Access to SDS Solutions and Private Cloud
- Intel® Solid State Drives Data Center Family
- Intel® Distribution of OpenVINO™ Toolkit
- Intel® Deep Learning Boost
- Intel® Framework Optimizations
- Intel® Optane™ DC persistent memory
- Intel® Select Solutions for VMware vSAN*
- VMware vSAN
- VMware Cloud Foundation

Appendix A: Bill of Materials

Table A1 provides the bill of materials for the Hybrid Cloud Data Analytics Solution. Table A2 provides the version numbers for the software that is part of the solution.

The Base configuration consists of four management nodes and four worker nodes. For higher performance you can replace the Base worker nodes with the Plus worker nodes. The foundational building block for all nodes is Intel® Select Solutions for VMware vSAN* ReadyNodes, with replaced CPU (2nd Generation Intel® Xeon® Scalable processor instead of previous generation).

Table A1. Bill of Materials for the Hybrid Cloud Data Analytics Solution

COMPONENT	DESCRIPTION	QUANTITY
4+ Management Nodes – VMware NSX-T Virtual Machines (VMs)		
Base SKU	Intel® Server System VRN2208WFAF82R	1
Mainboard	Intel® Server Board S2600WF0R	1
CPU	Intel® Xeon® Gold 6230 processor	2
Memory	RDIMM 32 GB DDR4-2933	12
Caching Tier	Intel® Optane™ SSD DC P4800X Series: 375 GB PCIe* x4 U.2	2
Capacity Tier	Intel® SSD DC P4510 Series: 4 TB 2.5” NVMe* U.2	6
Boot Device	Intel® SSD D3-S4510 Series: 480 GB (M.2, 80mm)	1
Network Interface Card (NIC)	Intel® Ethernet Converged Network Adapter XXV710-DA2	1
4+ Base Worker Nodes		
Base SKU	Intel® Server System VRN2208WFAF82R	1
Mainboard	Intel® Server Board S2600WF0R	1
CPU	Intel® Xeon® Gold 6248 processor	2
Memory	RDIMM 32 GB DDR4-2933	12
Caching Tier	Intel® Optane™ SSD DC P4800X Series: 375 GB PCIe* x4 U.2	2
Capacity Tier	Intel® SSD DC P4510 Series: 4 TB 2.5” NVMe* U.2	6
Boot Device	Intel® SSD D3-S4510 Series: 480 GB (M.2, 80mm)	1
Network Interface Card (NIC)	Intel® Ethernet Converged Network Adapter XXV710-DA2	1
4+ Plus Worker Nodes		
Base SKU	Intel® Server System VRN2208WFAF83R	1
Mainboard	Intel® Server Board S2600WF0R	1
CPU	Intel® Xeon® Platinum 8260 processor	2
Memory	RDIMM 32 GB DDR4-2666	24
Caching Tier	Intel® Optane™ SSD DC P4800X Series: 375 GB PCIe* x4 U.2	4
Capacity Tier	Intel® SSD DC P4510 Series: 4 TB 2.5” NVMe* U.2	12
Boot Device	Intel® SSD D3-S4510 Series: 480 GB (M.2, 80mm)	1
Network Interface Card (NIC)	Intel® Ethernet Converged Network Adapter XXV710-DA2	1

Table A2. Software Components of the Hybrid Cloud Data Analytics Solution

Component	Version
VMware Cloud Foundation*	3.8
Cloud Builder VM	2.1.0.0
VMware ESXi* hypervisor	ESXi670-201906002
VMware vSAN*	6.7 Express Patch 10
VMware NSX Data Center* for vSphere*	6.4.5
VMware NSX-T Data Center*	2.4.1
VMware vCenter* Server Appliance	6.7 Update 2c
VMware SDDC Manager*	3.8
VMware vRealize* Suite Lifecycle Manager	2.1 Patch 1
VMware Enterprise PKS*	1.4.1

[1] MarketWatch, February 2019, "Hybrid Cloud Market 2019."
[2] Forbes, March 2019, "Leaders Need AI to Keep Pace with the Data Explosion."
[3] VisionCritical, February 2019, "13 stunning stats on insight-driven businesses."
[4] Health IT Analytics, December 2018, "Big Data to See Explosive Growth, Challenging Healthcare Organizations."
[5] IDC, August 2018, "IDC Predicts Strong Growth for Big Data and Business Analytics."
[6] The Next Platform, October 2018, "Deep Learning Is Coming of Age."
[7] Vector Neural Network Instructions (VNNI) is planned for general availability from VMware in Q1 2020.
[8] No product or component can be absolutely secure. **Baseline configuration:** Testing performed by Intel as of July 11, 2017 and may not reflect all publicly available security updates. Two-socket Intel® Xeon® Platinum 8180 CPU @ 2.50 GHz (28 cores), Intel® Hyper-Threading Technology OFF, Intel® Turbo Boost Technology OFF, scaling governor set to "performance" via intel_pstate driver, 384 GB DDR4-2666 ECC RAM. CentOS Linux* release 7.3.1611 (Core), Linux* kernel 3.10.0-514.10.2.el7.x86_64. SSD: Intel® SSD DC S3700 Series (800 GB, 2.5in SATA 6 Gb/s, 25nm, MLC). Performance measured with: environment variables: KMP_AFFINITY = 'granularity = fine, compact', OMP_NUM_THREADS = 56, CPU Freq set with CPU power frequency-set -d 2.5G -u 3.8G -g performance. Caffe*: (github.com/intel/caffe), revision f96b759f71b2281835f690af267158b82b150b5c. Inference measured with "caffe time --forward_only" command, training measured with "caffe time" command. For "ConvNet" topologies, synthetic dataset was used. For other topologies, data was stored on local storage and cached in memory before training. Topology specs from github.com/intel/caffe/tree/master/models/intel_optimized_models (ResNet-50). Intel® C++ compiler version 17.0.2 20170213, Intel® Math Kernel Library small libraries version 2018.0.20170425. Caffe* run with "numactl -l". **Test configuration:** Testing performed by Intel as of February 20, 2019 and may not reflect all publicly available security updates. Two-socket Intel® Xeon® Platinum 8280 processor, 28 cores, Intel® Hyper-Threading Technology ON, Intel® Turbo Boost Technology ON; total memory: 384 GB (12 slots/32 GB/2933 MHz), BIOS: SE5C620.86B.0D.01.0271.120720180605 (ucode: 0x200004d), Ubuntu* 18.04.1 LTS, kernel 4.15.0-45-generic, SSD 1x sda Intel® SSDSC2BA80 SSD 745.2 GB, nvme1n1 Intel® SSDPE2KX040T7 SSD 3.7 TB, Deep Learning Framework: Intel® Optimization for Caffe* version: 1.1.3 (commit hash: 7010334f159da247db3fe3a9d96a3116ca06b09a), ICC version 18.0.1, Intel® MKL-DNN version: v0.17 (commit hash: 830a10059a018cd2634d94195140cf2d8790a75a, model: github.com/intel/caffe/blob/master/models/intel_optimized_models/int8/resnet50_int8_full_conv.prototxt, BS = 64, synthetic data, 4 instance/2 socket, Datatype: INT8.
[9] Intel® Optane™ DC persistent memory is not officially part of the validated Hybrid Cloud Data Analytics Solution, but is supported.
[10] Testing performed by Intel as of January 30, 2019 and may not reflect all publicly available security updates. No product or component can be absolutely secure. **Baseline Configuration (DRAM only):** # of systems: 5; Memory Sub System Per Socket: DRAM – 1536 GB (12 x 128 GB); CPU SKU/# per system: 8280M (CLX, Plat, 28 core)/4; Storage Description/Total Storage Cost: # of HDD/SDDs/$36,000; Software License Description/Cost per System: Software Cost (per/core or per system)/$0; Relevant Value Metric:15.00; CPU Cost: 4x8280M (CLX, Plat 28core): $52,048; Memory subsystem: Total Capacity (DRAM only): 6144 GB (1536 GB/Socket): $91,834; Storage: # of HDD/SDDs: $36,000; RBOM: Chassis; PSUs; Boot drive, etc: $7,603; Software Costs: $0; Total System Cost: $187,485; Total Cost: 5 Systems x $187,485: $937,424; Cost/TB: $62,494.95. **Intel® Optane™ DC persistent memory and DRAM configuration:** # of systems: 5; Memory Sub System Per Socket: 2304 GB (6 x 256 GB DCPMM + 6x128 GB DRAM, 2-2-2 App Direct Mode; CPU SKU/# per system: 8280L (CLX, Plat, 28 core)/4; Storage Description/Total Storage Cost: # of HDD/SDDs|$54,000; Software License Description/Cost per System: Software Cost (per/core or per system)/$0; Relevant Value Metric: 30.00; CPU Cost: 2x8280L (CLX, Plat 28 core): $71,624; Memory subsystem: Total Capacity (Intel Optane DC persistent memory+DRAM): 9216 GB (2304 GB/Socket): $96,917; Storage: # of HDD/SDDs: $54,000; RBOM: Chassis; PSUs; Boot drive, etc: $7,603; Software Costs: $0; Total System Cost: $230,144; 5 Systems x $230,144: $1,150,720; Cost/TB: $38,357.32.
[11] Testing performed by Intel and VMware as of March 21, 2019 and may not reflect all publicly available security updates. No product or component can be absolutely secure. **DRAM-only configuration:** 2 x Intel® Xeon® Platinum 8280L processor (2.7 GHz, 28 cores, Intel® Hyper-Threading Technology ON, Intel® Turbo Boost Technology ON); 384 GB addressable memory (192 GB/socket) - 12 x 32 GB DDR4 2666 MHz; 1x Intel® SSD D5-P4320 Series (7.6 TB, 2.5 in.); 10 GbE Intel® X520 network interface card. **Intel® Optane™ DC persistent memory and DRAM configuration:** 2x Intel® Xeon® Platinum 8280L processor (2.7 GHz, 28 cores, Intel® Hyper-Threading Technology ON, Intel® Turbo Boost Technology ON); 512 GB addressable memory (256 GB/socket) - 12 x 16 GB DDR4 2666 MHz plus 4x 128 GB Intel Optane DC persistent memory modules (T1 interleaved); 1x Intel® SSD D5-P4320 Series (7.6 TB, 2.5 in.); 10 GbE Intel® X520 network interface card. **VDI VM configuration (for both DRAM-only and Intel Optane DC persistent memory configurations):** OS = Microsoft Windows* 10 Enterprise (version 1803, OS build 17134.1; number of vCPU cores = 4; memory capacity = 12 GB; VMDK size = 32 GB (thick-provisioned lazy zeroed); paging file size = 12 GB; software installed = Google Chrome* (version 73.0.3683.103, Official Build, 64-bit and Microsoft Office* Professional Plus 2013.
[12] See endnote 8.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software, or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer, or learn more at intel.com.

Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor family, not across different processor families: Learn About Intel® Processor Numbers.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information, visit intel.com/benchmarks.

Performance results are based on testing as of the date noted in the configuration details and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

Optimization Notice: Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations.

Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel.

Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

Notice Revision #20110804

 0819/RCNO/KC/PDF 340880-001US