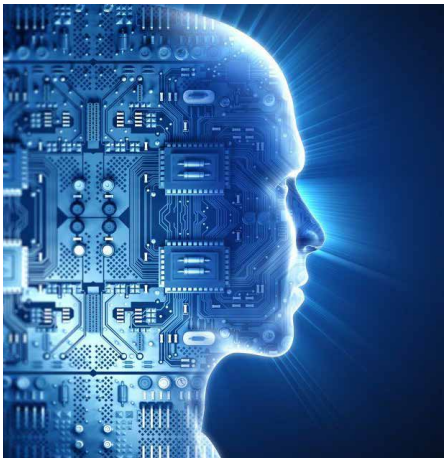
SOLUTION BRIEF

Enterprise
Artificial Intelligence

Accelerate AI Development with H2O.ai on Intel® Architecture

An integrated solution advances artificial intelligence (AI) development with H2O.ai's industry-leading software, validated and benchmarked on optimized Intel® technologies



H2O.ai

Executive Summary

While the promise of artificial intelligence (AI) is vast, enterprises working to develop AI solutions face talent shortages, complex development processes, and the need to ensure trusted results. Intel has collaborated with AI innovator H2O.ai to create an integrated solution that addresses these challenges. The solution combines H2O.ai's industry-leading AI software with hardware, libraries, and tools from Intel. These technologies help make AI innovation simple, fast, and explainable while enhancing total cost of ownership (TCO) and IT flexibility.

H2O.ai's software accelerates the tasks of training and deploying trusted machine-learning models. Key H2O technologies include a proprietary machine-learning solution that uses AI to automate AI, a fully open source platform for AI development, and tools to integrate H2O algorithms with the capabilities of Apache Spark*.

Intel® Xeon® Scalable processors, Intel® 3D NAND Solid State Drives (Intel® SSDs), and optimized Intel® software libraries provide the foundation for innovation. The solution is validated and benchmarked, enabling rapid deployment and robust performance.

Meeting the Challenges of AI Adoption

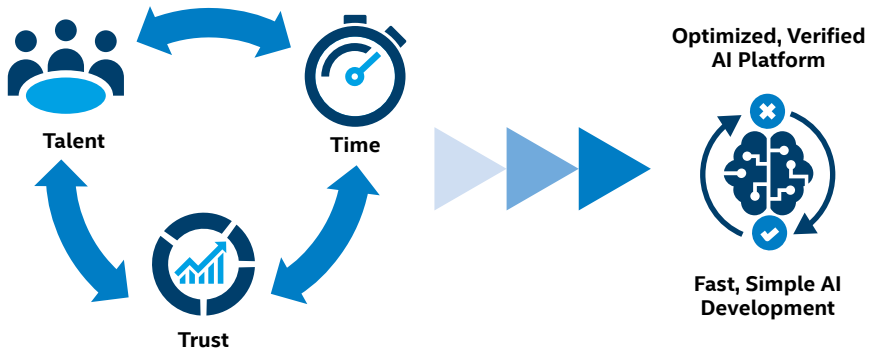


Figure 1. An integrated, validated solution based on technologies from H2O.ai and Intel empowers enterprises to simplify artificial intelligence (AI) development and speed innovation.

Authors

Vinod Iyengar
Head of Data Science Transformation
H2O.ai

Rick Cnossen
Solution Director
Intel Datacenter AI Solutions Development Group

Murali Madhanagopal
Lead Architect
Intel Datacenter AI Solutions Development Group

Business Challenge: Capture Value from AI

Artificial intelligence (AI) is rapidly becoming a critical differentiator—a revolutionary technology that organizations are using to maximize data value, augment human intelligence, and transform complex processes. AI adoption is rising, with 37 percent of respondents in a 2019 Gartner CIO survey saying they have implemented AI in some form.[1] PwC predicts AI's potential global economic impact will reach USD 15.7 trillion by 2030.[2]

However, three significant challenges are slowing AI development.[3]

- **Talent.** The demand for data science and AI experts far outstrips the supply. In Gartner's 2019 survey, 54 percent of the CIOs said skills shortages are their biggest AI challenge.[4]
- **Time.** AI and machine-learning workflows involve complex and iterative tasks. These workflows can delay results and impair the productivity of scarce talent.
- **Trust.** Many enterprises need more confidence in the validity of their AI models than a typical black-box environment can provide. Innovators must be able to explain what an AI solution is doing and how it reached its results.

Solution Overview: Addressing the Three Challenges

Addressing the challenge of talent, time, and trust requires easy-to-use tools and platforms that enable organizations to harness their data effectively and quickly develop AI algorithms.

H2O.ai and Intel, two companies with strong commitments to democratizing AI, have worked together to integrate H2O.ai's software portfolio with hardware and software technologies from Intel:

- **H2O Driverless AI*** is an automatic machine-learning platform that employs the techniques of expert data scientists in an easy-to-use application. Data science experts and beginners alike can streamline their AI development processes, accomplishing tasks in minutes that used to take months. The software provides information and insights about its operation and strategies, enabling organizations to explain their algorithms and results. **Impact: talent, time and trust.**
- **H2O Open Source Platform** offers a robust, high-performance machine-learning platform that supports widely used statistical and machine-learning algorithms. H2O includes an AutoML functionality that automatically evaluates the algorithms and produces a leaderboard of the best models for the task at hand. **Impact: time and trust.**
- **H2O Sparkling Water*** allows organizations employing the widely used Apache Spark* cluster computing framework to combine H2O.ai's machine-learning algorithms with their Spark technologies. **Impact: time.**
- **Intel's proven AI technology** includes optimized libraries, frameworks, and hardware infrastructure, including Intel® Xeon® Scalable processors, Intel® Solid State Drives (Intel® SSDs), and Intel® Ethernet Network Adapters. Together, they empower enterprises to optimize price/performance and speed deployment of the development infrastructure. **Impact: time.**

The overall result is a benchmarked solution for developing high-quality, trusted AI innovations, making the most of AI talent and accelerating time to market.

Solution Value: Accelerating AI Development

By addressing the challenges of talent, time, and trust, H2O.ai's software—integrated with the Intel® technologies—helps accelerate AI development and deliver the power of AI to a wide range of organizations. Enterprises gain capabilities to streamline the AI development workflow and increase productivity for data science experts as well as those with less advanced skills. Organizations can increase the business value of data and enjoy the competitive advantages of AI innovation. They can proceed with greater confidence in the workings of their algorithms and models—capabilities that are particularly important in regulated industries such as healthcare and financial services.

With libraries and frameworks optimized for Intel® architecture along with world-class Intel Xeon Scalable processors, the integrated solution enables enterprises to quickly deploy a reliable, powerful, comprehensive AI development platform. The validated solution helps improve price/performance while reducing the time needed to evaluate infrastructure elements and assemble the tools and services to advance AI innovation.

About H2O.ai

H2O.ai is a visionary, open source, artificial intelligence (AI) innovator with a mission to democratize AI for everyone. More than 14,000 companies use the H2O.ai's machine-learning platform in mission-critical use cases in finance, insurance, healthcare, retail, and other industries. H2O.ai is a leader in Gartner's Magic Quadrant for Data Science and Machine-Learning Platforms for 2018.[5]

Solution Architecture for AI Development

The integrated solution from H2O.ai and Intel offers a range of AI and data science platforms including H2O Driverless AI, H2O Open Source Platform, and H2O Sparkling Water, atop an environment of optimized libraries, frameworks, and software from Intel (Figure 2).

H2O.ai AI and Machine-Learning Software

H2O Driverless AI

H2O Driverless AI employs the techniques of expert data scientists in an easy-to-use application to automate complex tasks, scale an organization's data science efforts, and accomplish tasks in minutes that used to take months. The platform delivers unique and advanced functionality for data visualization, feature engineering, model interpretability, and low-latency deployment. Relevant capabilities include:

- **Feature engineering.** Feature engineering is the secret weapon that advanced data scientists use to extract the most accurate results from algorithms. Driverless AI employs a library of algorithms and feature transformations to automatically engineer new, high-value features for a given data set.
- **Explainability.** Driverless AI provides robust interpretability of machine-learning models to explain modeling results. The system generates charts including K-LIME*, Shapley*, Variable Importance, Decision Tree, Partial Dependence, and more.
- **Natural Language Processing (NLP).** Text data can contain critical information to inform better predictions. H2O Driverless AI automatically uses powerful NLP techniques to convert short text strings into features.
- **Scoring pipelines.** Driverless AI automatically generates Python* scoring pipelines and new ultra-low latency automatic scoring pipelines. This allows for highly optimized, low-latency, production-ready Java* code that can be deployed anywhere.
- **Time series.** Driverless AI delivers superior time-series capabilities to optimize for almost any prediction time window. The solution also incorporates data from numerous predictors; handles structured character data and high-cardinality categorical variables; and manages gaps in time-series data and other missing values.
- **Visualization.** To help users get a quick understanding of their data before model building begins, Driverless AI automatically generates visualizations and creates relevant data plots.
- **Data sets.** Driverless AI works across varied data sets, including Apache Hadoop* Distributed File System (HDFS) and Amazon S3*.

H2O Open Source Platform

H2O Open Source Platform is a fully open source, distributed, in-memory machine-learning platform with linear scalability. It supports the most widely used statistical and machine-learning algorithms, including gradient-boosted machines (GBM), generalized linear models (GLM), deep learning, and more. Key capabilities of the H2O Open Source Platform include:

- **Algorithm support.** H2O offers leading open source algorithms developed from the ground up for distributed computing and for both supervised and unsupervised approaches, including Random Forest, GLM, GBM, Deep Learning, XGBoost, Generalized Low Rank Models (GLRM), Word2Vec, and numerous others.
- **Language flexibility.** Developers can use familiar programing languages such as R*, Python, and others to build models in H2O. They can also use H2O Flow*, a graphical, notebook-based, interactive user interface that does not require coding.

AI Platform from H2O.ai and Intel

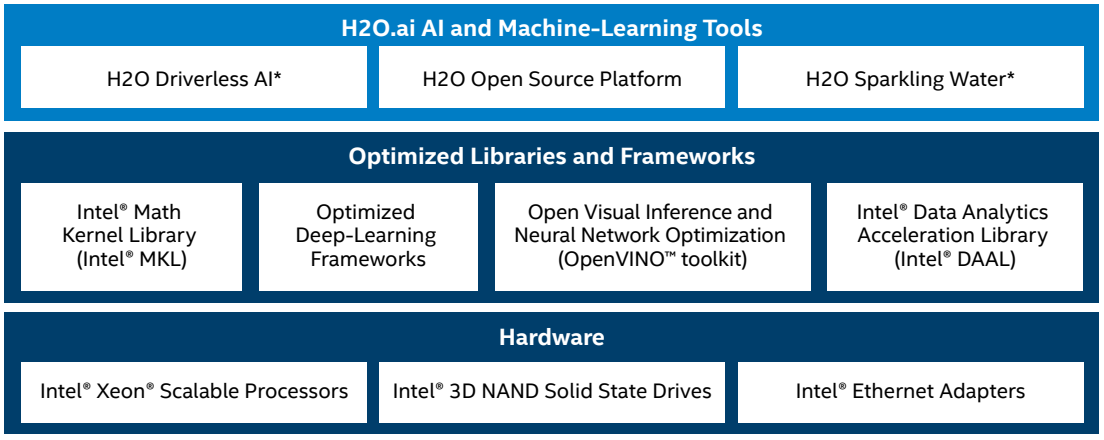


Figure 2. The artificial intelligence (AI) platform integrates industry-leading H2O.ai software on platform technologies, libraries, and frameworks from Intel, providing a validated, benchmarked environment for AI development.

- **Workflows automation.** Developers can automate the machine-learning workflow with AutoML, which includes automatic training and tuning of many models within a user-specified time. H2O's Stacked Ensembles feature identifies top-performing models.
- **High performance.** H2O uses in-memory processing with fast serialization between nodes and clusters to support massive data sets. Distributed processing on big data delivers dramatic speedups compared to fine-grain parallelism, enabling optimal efficiency without degrading computational accuracy.
- **Simple deployment.** Easy-to-deploy Java models allow for fast, accurate scoring in any environment, including very large models.

H2O Sparkling Water

Sparkling Water lets users combine H2O.ai's fast, scalable machine-learning algorithms with the capabilities of Apache Spark. Integrating these two open source environments provides a seamless experience for users who want to make a query using Apache Spark SQL*, feed the results into H2O to build a model and make predictions, and then use the results again in Spark. Significant capabilities of H2O Sparkling Water include:

- **Access.** Developers can add a broad range of H2O algorithms developed from the ground up for distributed computing. Algorithms are available for both supervised and unsupervised approaches, including Random Forest, GLM, GBM, XGBoost, GLRM, Word2Vec, and more.
- **Flexibility.** Developers can drive computation from Scala*, R, or Python and use the H2O Flow open source user interface, providing an ideal machine-learning platform for application developers.
- **Convenience.** Easy-to-deploy Java models allow for fast and accurate scoring in any environment, including very large models.

Intel® Technology-based Infrastructure

Intel offers an end-to-end, performance-tuned solution with H2O.ai to ensure outstanding performance and TCO on Intel technologies. This empowers enterprises to quickly harness a reliable, powerful, comprehensive solution that accelerates AI development as it optimizes TCO and simplifies the AI development environment. Organizations can:

- **Prepare for the future** with IT transformation investments on scalable, industry-leading, technologies for storage, memory, and compute.
- **Generate excellent TCO** with general-purpose hardware that IT organizations are used to managing, plus special-purpose AI accelerators that provide performance boosts as needed.
- **Accelerate time to market** by using a turnkey, validated solution with a rich development toolset and built-in optimizations.

Software Infrastructure

Optimized libraries, tools, and frameworks from Intel enhance performance while reducing programming complexity. These software optimizations, coupled with Intel Xeon Scalable processor-based hardware, deliver peak performance and compelling TCO across a broad range of analytics and AI workloads.

Intel® Math Kernel Library (Intel® MKL) optimizes code with minimal effort for current and future generations of Intel® processors. It is compatible with a range of compilers, languages, and operating systems, as well as linking and threading models.

Deep-learning frameworks such as TensorFlow* have been optimized using Intel MKL for Deep Neural Networks (Intel® MKL-DNN) primitives, a popular performance library for deep-learning applications.

Intel® Distribution of the Open Visual Inference and Neural Network Optimization (OpenVINO™) toolkit extends workloads across Intel® hardware (including accelerators) and maximizes performance.

Intel® Data Analytics Acceleration Library (Intel® DAAL) boosts machine-learning and data analytics performance by optimizing data ingestion and algorithm execution.

Hardware Infrastructure

Enterprises can gain a competitive advantage by using platforms based on Intel hardware building blocks—familiar, cost-effective technologies on a trusted platform.

Intel Xeon Scalable Processors

- World-class, highly integrated Intel Xeon Scalable processors provide exceptional performance for analytics, AI, and other data center workloads—including machine and deep learning.
- Deep-learning framework optimizations for the Intel Xeon Scalable platform deliver performance gains of up to 277 times in inference throughput[6] and up to 241 times training throughput.[7]
- Intel Xeon Scalable processors offer superb TCO. IT can replace four servers that are four-to-five years old with a single Intel Xeon Scalable processor-based server, lowering five-year TCO by up to 60 percent.[8]

Performance results are based on testing as of June 26, 2018, and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

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.

Intel® 3D NAND Solid State Drives (Intel® SSDs)

- Intel 3D NAND SSDs are transforming the economics of storage, delivering response times, capacity, and resilience that can significantly enhance the ability to access data faster for speedier analysis.
- Intel 3D NAND SSDs are on a path to approach cost parity with hard disk storage, enabling organizations to replace their data centers' spinning disks with 3D NAND flash storage.

Verified Performance and Scalability

To ensure outstanding performance and scalability, members of the AI Solutions Group at Intel verified the integrated solution from H2O.ai and Intel by conducting industry benchmark tests. The tests used the platform configurations in Table 1, along with H2O.ai's widely used GBM and XGBoost supervised learning algorithms.[9] For many problems, XGBoost is considered one of the best GBM frameworks today.

- **Performance.** The performance benchmark used Intel® Xeon® Platinum processors, optimized software libraries, and the popular Higgs Dataset.[10] Tests ran 500,000 particle samples, executing a classification problem to identify signal processes that produce Higgs bosons. A single node of the Intel Xeon Platinum 8180 processor-based configuration needed just over 3.5 minutes (214 seconds) to train 16 XGBoost algorithms. This result demonstrates the ability to achieve excellent training and classification performance with flexible, general-purpose CPU infrastructure.
- **Scalability.** To validate multi-node performance improvements with large data sets, the team used three-node and five-node configurations based on the Intel® Xeon® Gold 6148 processor and GBM algorithm. Tests used the Public Airline Dataset, a collection of 20 years of public airline data, to predict potential flight delays.[11] The five-node configuration showed a 38.5 percent improvement over the three-node configuration. These results demonstrate an impressive increase in speed as more compute resources are added, highlighting with the solution's ability to scale for long-running training jobs.

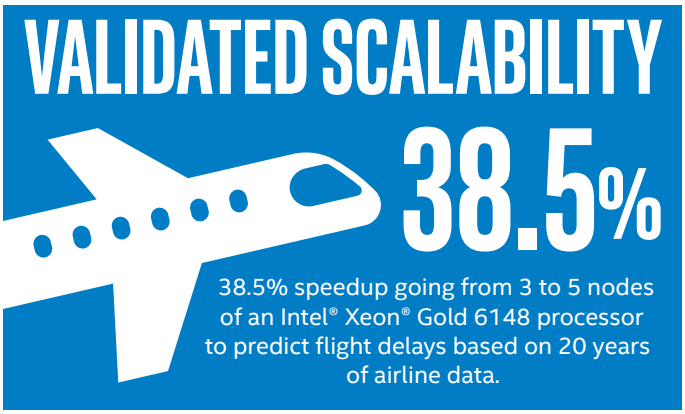


Table 1. Benchmark Configuration.

	CONFIGURATION		
	Performance Benchmark	3-Node Scaling Benchmark	5-Node Scaling Benchmark
Processor	• 2x Intel® Xeon® Platinum 8180 processors • 26 cores, 2.1 GHz	• 2x Intel® Xeon® Gold 6148 processors • 20 cores, 2.4 GHz	• 2x Intel Xeon Gold 6148 processors • 20 cores, 2.4 GHz
Nodes	1	3	5
Memory	• 384 GB (12x 32 GB 2666 MT/s DDR4 DIMM) per node	• 384 GB (12x 32 GB 2666 MT/s DDR4 DIMM) per node	• 384 GB (12x 32 GB 2666 MT/s DDR4 DIMM) per node
Storage	• Boot: 1x 1.2 TB Intel® SSD DC S4500 • Application: 1x 2 TB Intel® SSD DC P4600 (PCIe* NVMe)	• Boot: 1x 1.2 TB Intel SSD DC S4500 • Application: 1x 2 TB Intel SSD DC P4600 (PCIe NVMe)	• Boot: 1x 1.2 TB Intel SSD DC S4500 • Application: 1x 2 TB Intel SSD DC P4600 (PCIe NVMe)
Data network	• 2x Intel® 82599ES dual-port 10 Gb Ethernet	• 2x Intel 82599ES dual-port 10 Gb Ethernet	• 2x Intel 82599ES dual-port 10 Gb Ethernet
Firmware and software optimizations	• Intel® Turbo Boost Technology enabled • Intel® Hyper-Threading Technology (Intel® HT Technology) enabled	• Intel Turbo Boost Technology enabled • Intel HT Technology enabled	• Intel Turbo Boost Technology enabled • Intel HT Technology enabled
OS	• CentOS* 7.4+	• CentOS 7.4+	• CentOS 7.4+
Software setup	• Cloudera* 5.15 distribution with Apache Spark* 2.3.1 and Hadoop* Distributed File System (HDFS) 2.6	• Cloudera 5.15 distribution with Apache Spark 2.3.1 and HDFS 2.6	• Cloudera 5.15 distribution with Apache Spark 2.3.1 and HDFS 2.6

Configurations used the Intel® Server System platform, ucode: 0x004d, and Centos 7.4 kernel: 3.10.0-862.11.6.el7.x86_64. Performance results are based on testing as of December 10, 2018 and may not reflect all publicly available security updates. No product can be absolutely secure.

Intel anticipates additional performance breakthroughs with the next generation of Intel Xeon Scalable processor. This forthcoming processor, based on 14nm technology, includes a new AI extension called Intel® Deep Learning Boost (Intel® DL Boost). This feature extends Intel® Advanced Vector Extensions 512 (Intel® AVX-512) by adding a new vector neural network instruction (VNNI) designed to increase inferencing performance. The next generation of Intel Xeon Scalable processor will also debut Intel® Optane™ DC persistent memory, which extends accessible memory for currently constrained jobs. Several software optimizations, including Intel DAAL are expected to be integrated, further enhancing performance.

Conclusion

H2O.ai and Intel are collaborating to democratize AI with an integrated, validated, high-performance solution that empowers enterprises to optimize their AI development environment and meet the challenges of talent, time, and trust. Organizations can take advantage of the solution's easy-to-use tools and general-purpose Intel technologies to position themselves to accelerate time to market and gain the competitive advantages of AI innovation.

H2O.ai is a member of the Intel® AI Builders Program, an ecosystem of industry-leading independent software vendors (ISVs), system integrators (SIs), original equipment manufacturers (OEMs), and enterprise end users who have a shared mission to accelerate the adoption of AI across Intel® platforms.

Intel's proven AI technology is driving enterprise-wide AI adoption with industry-leading performance and security based on scalable storage, memory, and compute resources. Intel's broad product portfolio offers excellent TCO and flexibility aligned to customer needs, whether the need is general-purpose hardware that the IT organization is already managing or purpose-built AI accelerators for the most demanding performance requirements. The result is a significant step toward AI for everyone.

Find the solution that is right for your organization. Contact your Intel representative or visit **intel.com/ai**.

Solution Provided By:

Learn More

You may also find the following resources useful:

- Download Driverless AI for a free 21-day trial
- H2O.ai and its AI product suite
- Intel and AI
- Intel® AI Builders Program
- Intel® Xeon® Scalable processors
- Intel® SSD Data Center Family
- Intel® Ethernet Products
- Intel® Data Analytics Acceleration Library
- Intel® Math Kernel Library
- Intel® Distribution of OpenVino™ Toolkit
- Intel® Deep Learning Boost
- Intel® Optane™ DC Persistent Memory

[1] Gartner Press Release, 2019. gartner.com/en/newsroom/press-releases/2019-01-21-gartner-survey-shows-37-percent-of-organizations-have

[2] PwC, "Sizing the Prize: PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution," 2017. pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html

[3] Ingrid Burton, "What Business Leaders Need to Know about AI," H20.ai Blog, January 11, 2019. h2o.ai/blog/what-business-leaders-need-to-know-about-ai

[4] See endnote 1.

[5] Gartner, "Magic Quadrant for Data Science and Machine-Learning Platforms," gartner.com/doc/3860063/magic-quadrant-data-science-machinelearning

[6] Inference throughput testing by Intel as of June 26, 2018, using FP32 Batch Size Caffe GoogleNet v1 128 AlexNet 256. Configurations for Inference throughput Platform: 2 socket Intel® Xeon® Platinum 8180 CPU @ 2.50GHz / 28 cores HT ON , Turbo ON Total Memory 376.28GB (12slots/32GB/2666MHz), 4 instances of the framework, CentOS Linux-7.3.1611-Core, SSD sda RS3WC080 HDD 744.1GB, sdb RS3WC080 HDD 1.5TB, sdc RS3WC080 HDD 5.5TB , Deep Learning Framework caffe version: a3d5b022fe026e9092fc7abc7654b1162ab9940d Topology: GoogleNet v1 BIOS: SE5C620.8 6B.00.01.0004.071220170215 MKLDNN: version: 464c268e544bae26f9b85a2acb9122c766a4c396 NoDataLayer. Measured: 1449.9 imgs/sec vs Platform: 2S Intel® Xeon® CPU E5-2699 v3 @ 2.30GHz (18 cores), HT enabled, turbo disabled, scaling governor set to "performance" via intel_pstate driver, 64GB DDR4-2133 ECC RAM. BIOS: SE5C610.86B.01.01.0024.021320181901, CentOS Linux-7.5.1804(Core) kernel 3.10.0-862.3.2.el7.x86_64, SSD sdb INTEL SSDSC2BW24 SSD 223.6GB. Framework BVLC-Caffe: github.com/BVLC/caffe, Inference and Training measured with "caffe time" command. For "ConvNet" topologies, dummy dataset was used. For other topologies, data was stored on local storage and cached in memory before training. BVLC Caffe (github.com/BVLC/caffe), revision 2a1c552b66f026c7508d390b526f2495ed3be594

[7] Training throughput testing by Intel as of June 26, 2018, using FP32 Batch Size Caffe GoogleNet v1 128 AlexNet 256. Configuration for training throughput Platform: 2 socket Intel® Xeon® Platinum 8180 CPU @ 2.50GHz / 28 cores HT ON , Turbo ON Total Memory 376.28GB (12slots/32GB/2666MHz), 4 instances of the framework, CentOS Linux-7.3.1611-Core, SSD sda RS3WC080 HDD 744.1GB, sdb RS3WC080 HDD 1.5TB, sdc RS3WC080 HDD 5.5TB , Deep Learning Framework caffe version: a3d5b022fe026e9092fc7abc765b1162ab9940d Topology: alexnet BIOS:SE5C620.8 6B.00.01.0004.071220170215 MKLDNN: version: 464c268e544bae26f9b85a2acb9122c766a4c396 NoDataLayer. Measured: 1257 imgs/sec vs Platform: 2S Intel® Xeon® CPU E5-2699 v3 @ 2.30GHz (18 cores), HT enabled, turbo disabled, scaling governor set to "performance" via intel_pstate driver, 64GB DDR4-2133 ECC RAM. BIOS: SE 5C610.86B.01.01.0024.021320181901, CentOS Linux-7.5.1804(Core) kernel 3.10.0-862.3.2.el7.x86_64, SSD sdb INTEL SSDSC2BW24 SSD 223.6GB. Framework BVLC-Caffe: github.com/BVLC/caffe, Inference and Training measured with "caffe time" command. For "ConvNet" topologies, dummy dataset was used. For other topologies, data was stored on local storage and cached in memory before training. BVLC Caffe (github.com/BVLC/caffe), revision 2a1c552b66f026c7508d390b526f2495ed3be594

[8] Up to 60 percent TCO savings compared to five-year-old system: Tests performed by Intel as of June 26, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. Example based on estimates as of June 26, 2018 of equivalent rack performance over four-year operation on integer throughput workload (estimate based on SPECrate*2017_int_base on Intel internal platforms) running VMware vSphere* Enterprise Plus on Red Hat Enterprise Linux Server* and comparing 20 installed 2-socket servers with Intel® Xeon® processor E5-2690 (formerly "Sandy Bridge-EP") at a total cost of USD 737,460 [Per server cost USD 36.8K: acquisition=12,5K, infrastructure and utility=4.5K, OS and software=10.2K, maintenance=9.7K] vs. 5 new Intel® Xeon® Platinum 8180 (Skylake) at a total cost of USD 294,540 [Per server cost USD 58.9K: acquisition=12,5K, infrastructure and utility=10.1K, OS and software=10.1K, maintenance=9.7K]. Assumptions based on https://xeonprocessoradvisor.intel.com assumptions as of June 6, 2018. Per node 4X higher integer throughput performance: estimate based on SPECrate*2017_int_base on Intel internal platforms: 1x node, 2x Intel® Xeon® Processor E5-2690, 128 GB total memory, 16 slots/8 GB/1600 MT/s DDR3 RDIMM, Benchmark: SPEC CPU2017 V1.2, Compiler: Intel® Compiler IC17 update 2, Optimized libraries/versions: IC18.0_20170901, Other Software: MicroQuill SMART HEAP. uCode: 713, OS: Red Hat Enterprise Linux* 7.4, Kernel: 3.10.0-693.11.6.el7.x86_64 x86_64, Score 65.5 vs. 1x Node, 2x Intel® Xeon® Platinum 8180 Processor, 384 GB total memory, 12 slots/32 GB/2666 MT/s DDR4, Benchmark software: SPEC CPU* 2017, Compiler: Intel® Compiler IC18 OEM, Optimized libraries: AVX512, ucode:0x043, Red Hat Enterprise Linux* 7.4, 3.10.0-693.11.6.el7.x86_64, Score: 281. 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.

[9] H2O.ai, XGBoost, docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/xgboost.html

[10] HIGGS Dataset, archive.ics.uci.edu/ml/datasets/HIGGS

[11] Public Airline Dataset, stat-computing.org/dataexpo/2009/the-data.html

0219/JSTA/KC/PDF 338570-001US