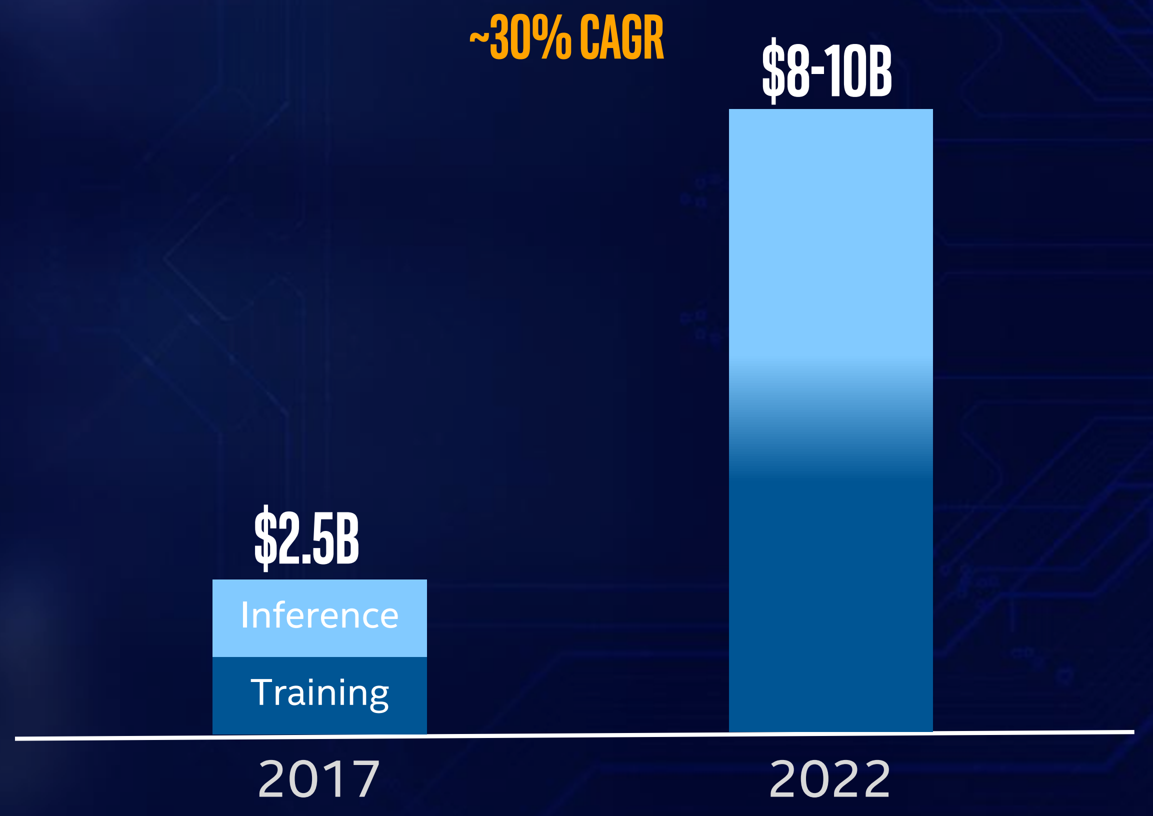
intel | DATA-CENTRIC INNOVATION SUMMIT

UNLOCKING DATA VALUE WITH AI

NAVEEN RAO
CORPORATE VICE PRESIDENT & GENERAL MANAGER
ARTIFICIAL INTELLIGENCE PRODUCTS GROUP

AI IS EXPLODING

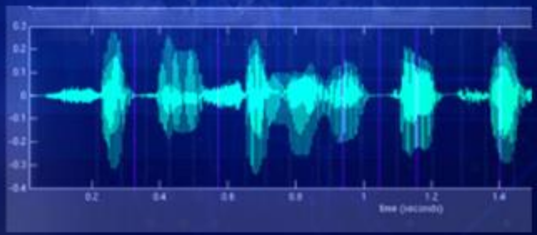
DATA CENTER LOGIC SILICON TAM

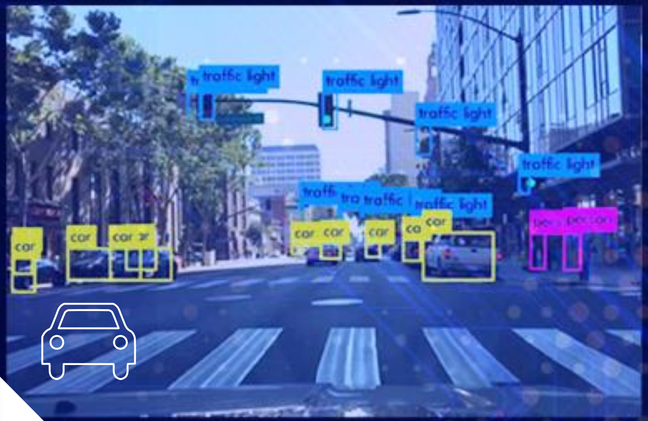


EMERGING AS A CRITICAL WORKLOAD

1. Source: AI Si Server TAM is based on amalgamation of analyst data and Intel analysis, based upon current expectations and available information and are subject to change without notice.

AI IS EVOLVING

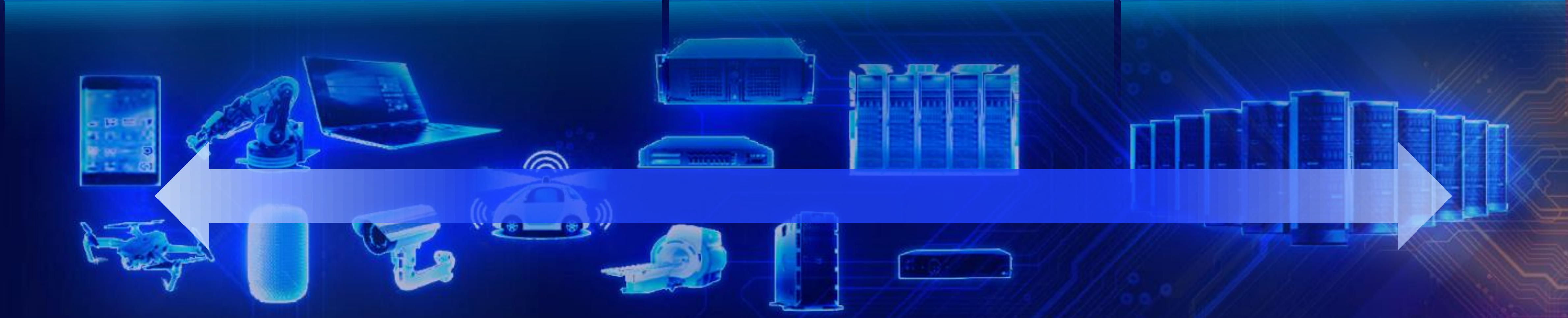




PROOFS OF CONCEPTS → UNLOCKING REAL VALUE

AI IS EXPANDING

END POINT	EDGE	DATA CENTER



intel MOVIDIUS inside

MOBILEYE An Intel Company

intel ATOM inside

intel FPGA inside™

intel CORE i7 inside

intel NERVANA inside

intel XEON PLATINUM inside

COMPREHENSIVE AI PORTFOLIO

ONE SIZE DOES NOT FIT ALL

END POINT

IOT SENSORS
(Security, home, retail, industrial...)

- intel MOVIDIUS inside — Vision & Inference
- intel ATOM inside + intel GNA inside — Speech

SELF-DRIVING VEHICLE

- Autonomous Driving

DESKTOP & MOBILE

- intel CORE i7 inside + intel IRIS + intel GNA inside + intel MOVIDIUS inside

Display, video, AR/VR, gestures

CONVERGED MOBILITY

- intel ATOM inside + intel MOVIDIUS inside

Vision, speech, AR/VR

EDGE

SERVERS, APPLIANCES & GATEWAYS

- intel XEON PLATINUM inside — Most use cases
- intel ARRIA 10 inside — Streaming latency-bound systems
- intel MOVIDIUS inside — Vision & Inference for various systems types

DATA CENTER

SERVERS & APPLIANCES

- intel XEON PLATINUM inside — Foundation for AI
- intel NERVANA inside — Built for Deep Learning
- intel STRATIX 10 inside — Flexible & memory bandwidth bound use cases

WINNING TOGETHER WITH INTEL AI

AAEON, ADLINK, ADVANTECH, AGENT, Alibaba.com, AllGoVision, ASUS, aws, Azure

axxonsoft, Baidu 百度, bluedata, cdhi, 中国电信 CHINA TELECOM, CRAY, alhua TECHNOLOGY, databricks, DELL EMC, dji

GE Healthcare, GeoVision, Getac, GIGASPACES, Google, Gorilla, Hewlett Packard Enterprise, HIKVISION, IEI Integration Corp.

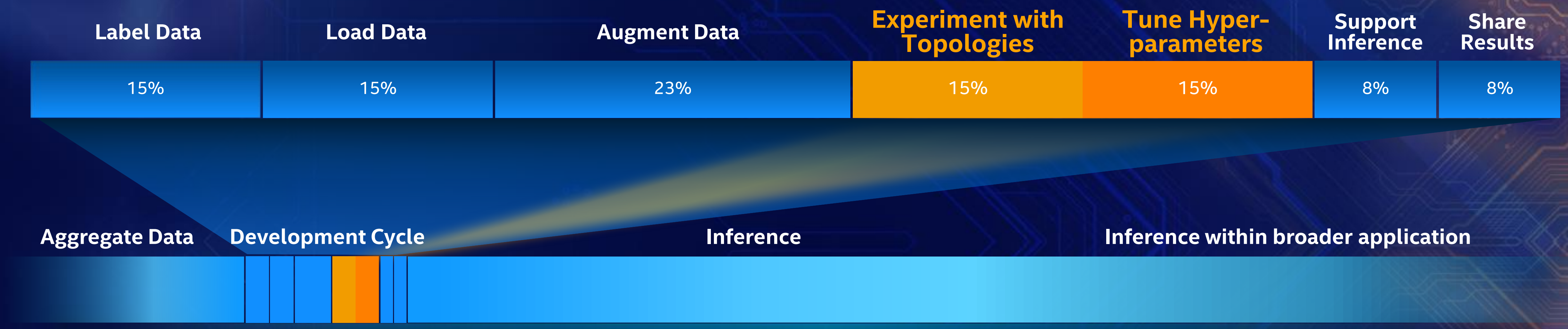
科大讯飞 iFLYTEK, iRONYUN, Kyoto University, JD.COM 京东, Lenovo, Microsoft, Midea, Mobiliya QuEST — Mobiliya is now a QuEST Global Company, Mindtree Welcome to possible, MTC, NEXCOM Building The Digital Infrastructure

NOVARTIS, QNAP, Taboola, Tencent Cloud, Telefónica, 头条 TOUTIAO, UCLOUD 专业云计算服务商, UnionPay 银联, V5 SYSTEMS, Yandex

Subset of full customer and partner list

$1B+ AI BUSINESS FOR INTEL TODAY

AI DEVELOPMENT LIFECYCLE



INTEL WORKS WITH CUSTOMERS ACROSS ENTIRE PROCESS FLOW

BROUGHT TO LIFE THROUGH DATA SCIENTISTS

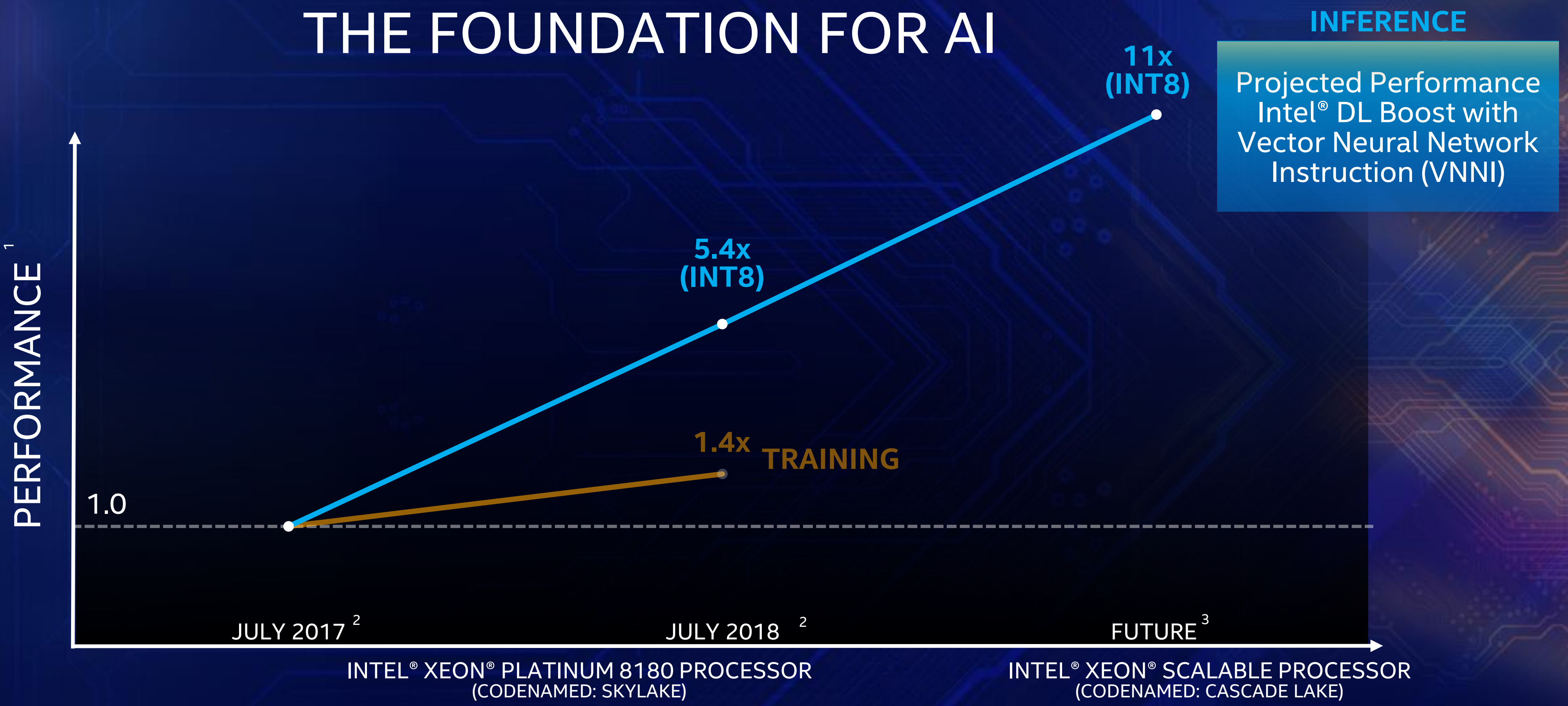
RESEARCH

CUSTOMIZE

DEPLOY

INTEL® XEON® SCALABLE PROCESSORS

THE FOUNDATION FOR AI



CONTINUED INVESTMENTS IN OPTIMIZATIONS TO DELIVER INCREASED PERFORMANCE

[1] Intel® Optimization for Caffe Resnet-50 performance does not necessarily represent other Framework performance.

[2] Based on Intel internal testing: 1X (7/11/2017), 2.8X (1/19/2018), 1.4x (8/2/2018) and 5.4X (7/26/2018) performance improvement based on Intel® Optimization for Café Resnet-50 inference throughput performance on Intel® Xeon® Scalable Processor. See Configuration Details Slide #19
Performance results are based on testing as of 7/11/2017(1x), 1/19/2018(2.8x), 8/2/2018 (1.4x) & 7/26/2018(5.4) and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. 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.
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: http://www.intel.com/performance.

INTEL® NERVANA™ NNP L-1000

PURPOSE-BUILT FOR REAL WORLD AI PERFORMANCE

Optimized across memory, bandwidth, utilization and power

3-4x training performance of first-generation NNP product

High-bandwidth, low-latency interconnects

bfloat16 numerics

FIRST COMMERCIAL NNP IN 2019

Source: Based on Intel measurements on limited distribution SDV (codenamed: Lake Crest) compared to Intel measurements on NNP-100 simulated product

SOFTWARE IS ESSENTIAL

ABSTRACTION

TOOLKITS
Application Developers

OpenVINO™ Toolkit

Intel® Movidius™ SDK

LIBRARIES
Data Scientists

MACHINE LEARNING LIBRARIES

Scikit-Learn NumPy MLlib

DEEP LEARNING FRAMEWORKS

TensorFlow, mxnet, Caffe, BigDL for Spark, Caffe2, PYTORCH, Microsoft CNTK, PaddlePaddle

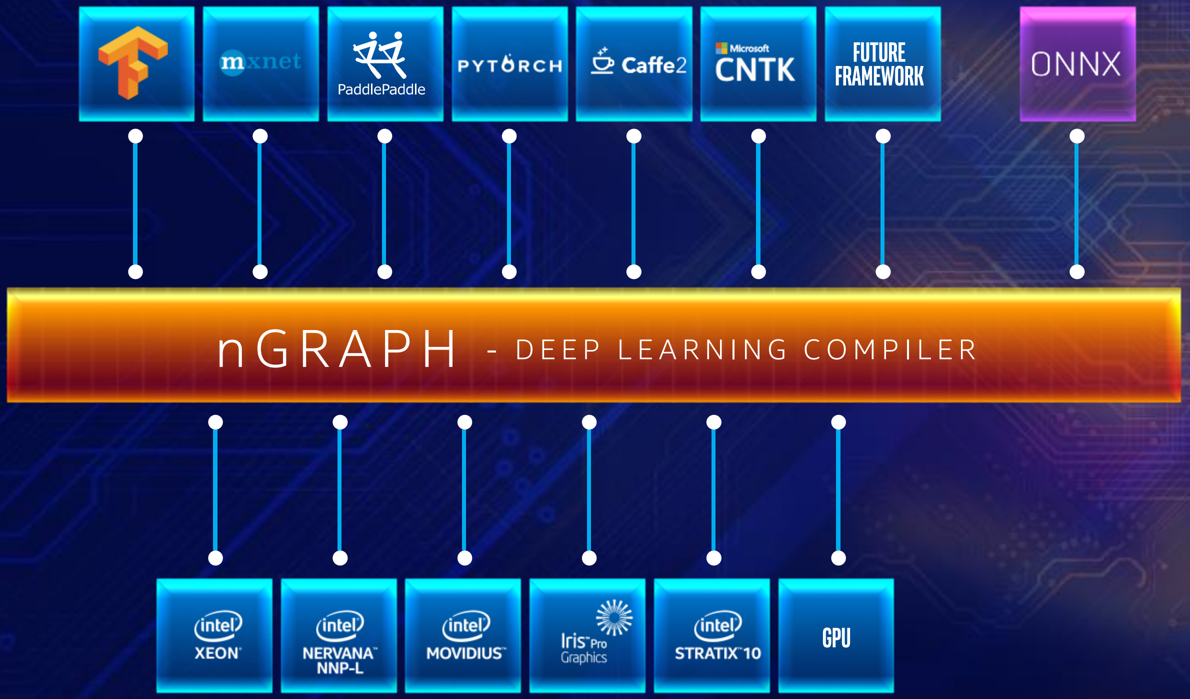
FOUNDATION
Library Developers

ANALYTICS, MACHINE & DEEP LEARNING PRIMITIVES

MKL-DNN clDNN Python DAAL

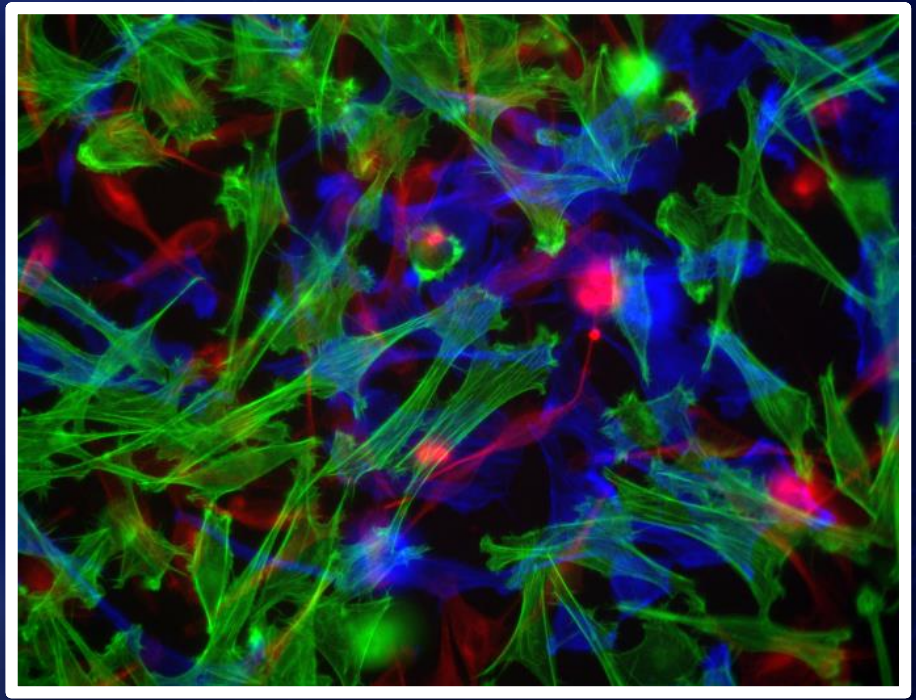
DEEP LEARNING GRAPH COMPILER

Intel® nGraph™ Compiler



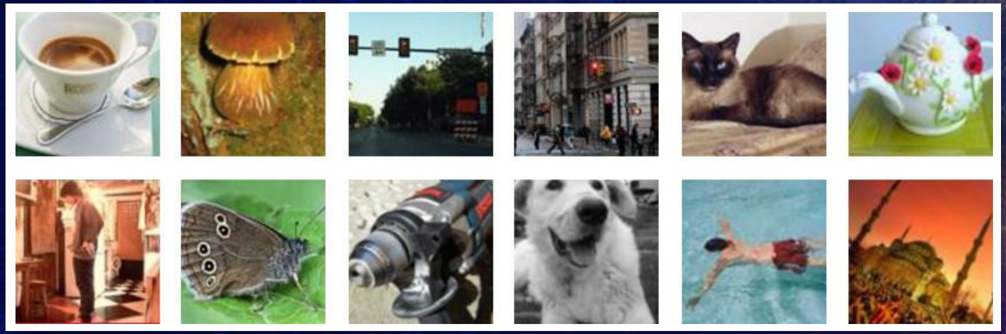
NOVARTIS DRUG DISCOVERY

26X LARGER



1024 X 1280 X 3

IMAGENET

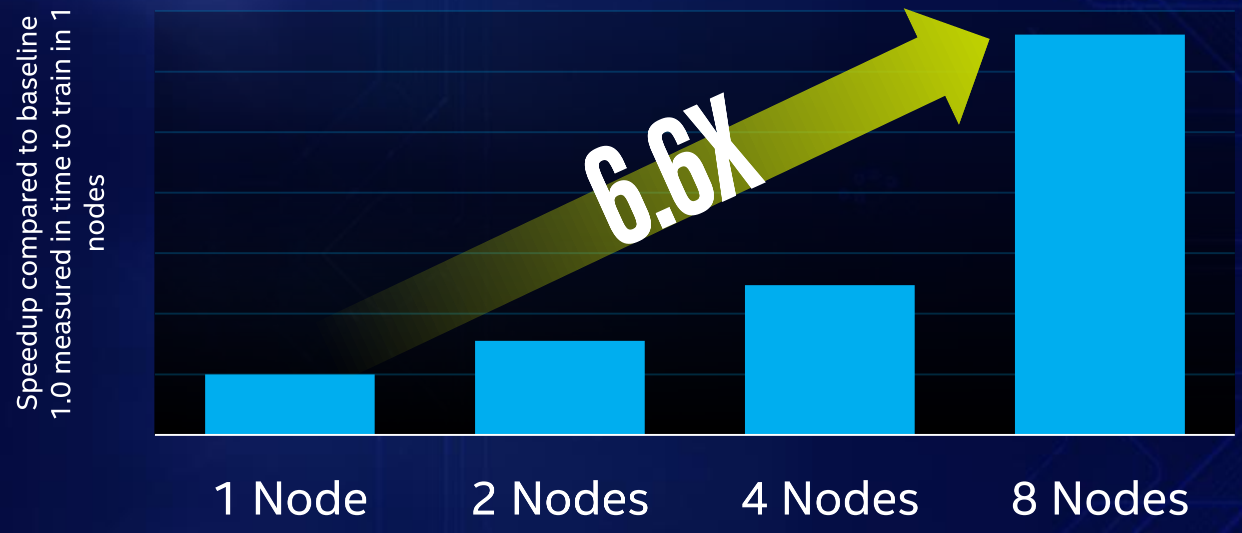


224 X 224 X 3

HIGH PERFORMANCE AT SCALE

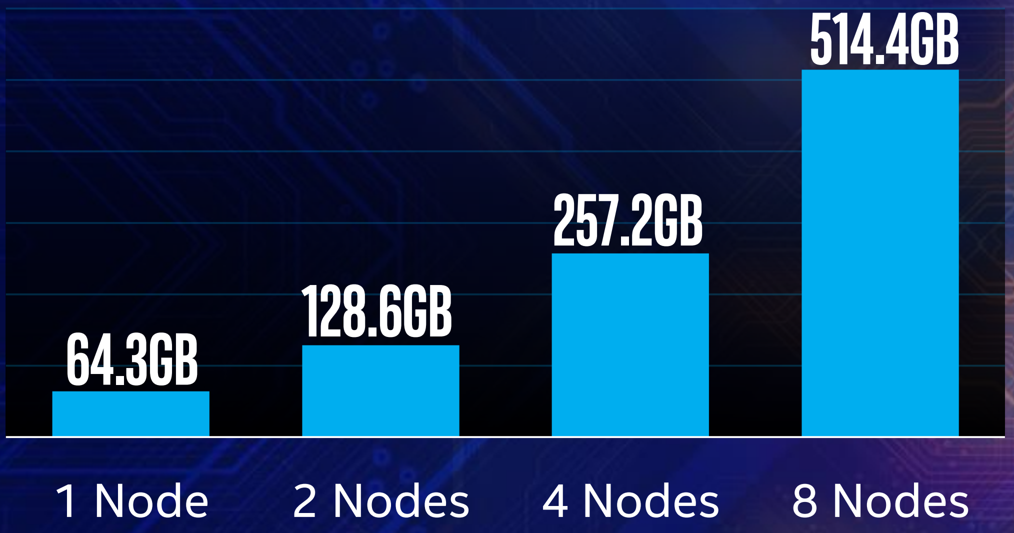
SCALING OF TIME TO TRAIN

INTEL® OMNI-PATH ARCHITECTURE, HOROVOD AND TENSORFLOW®



TOTAL MEMORY USED

192GB DDR4 PER INTEL® SP 2S XEON® 6148 PROCESSOR



MULTISCALE CONVOLUTION NEURAL NETWORK

TensorFlow

OPTIMIZED LIBRARIES

Intel® MKL/MKL-DNN, clDNN, DAAL

INTEL® OMNI-PATH ARCHITECTURE

intel XEON GOLD inside

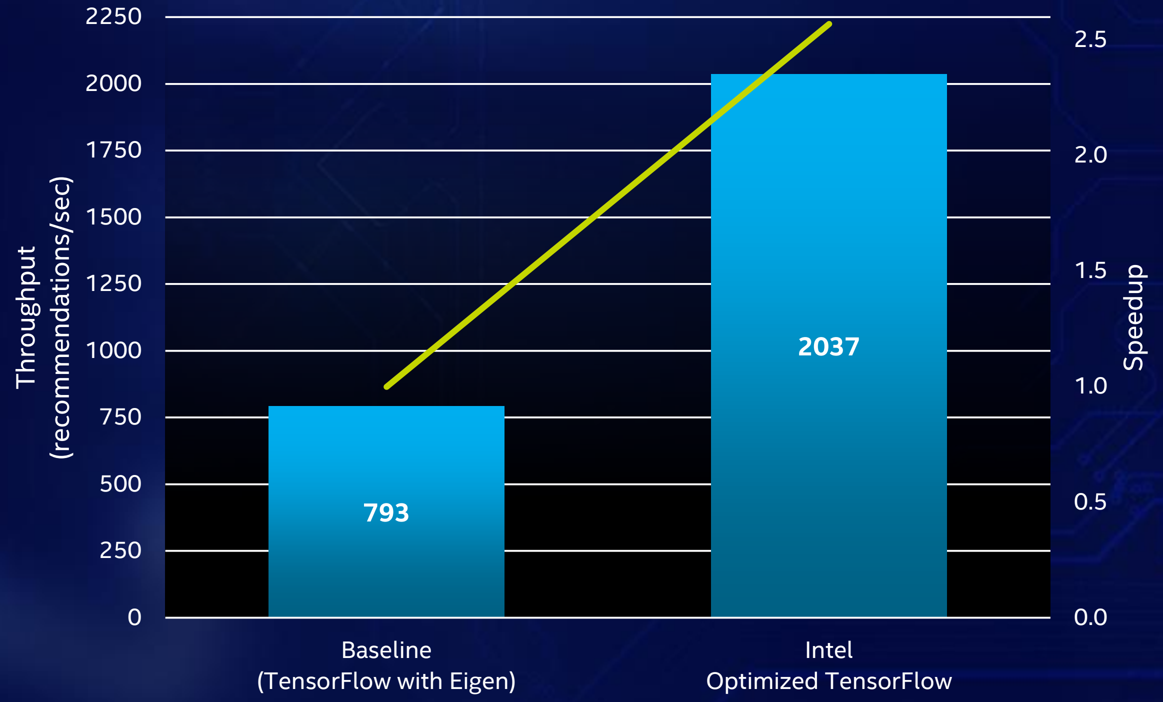
DATA-CENTRIC INNOVATION SUMMIT

§ Configuration: CPU: Intel Xeon 6148 processor @ 2.4GHz, Hyper-threading: Enabled. NIC: Intel® Omni-Path Host Fabric Interface, TensorFlow: v1.7.0, Horovod: 0.12.1, OpenMPI: 3.0.0. OS: CentOS 7.3, OpenMPU 23.0.0, Python 2.7.5
Time to Train to converge to 99% accuracy in model
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 http://www.intel.com/performance.
Performance results are based on testing as of 5/25/2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

intel

TABOOLA CHOOSES INTEL® XEON® SCALABLE PROCESSORS TO SCALE INFERENCE

2.5x INFERENCE IMPROVEMENT



"Serving from the CPUs helped us reduce costs, increase efficiency, and provide better content recommendations."
- Ariel Pisetzky, VP of Information Technology

Taboola

DATA-CENTRIC INNOVATION SUMMIT

Performance results are based on testing as of 8/6/2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product 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.
Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

VIBRANT AI ECOSYSTEM

CROSS VERTICAL

OEM
COLFAX Customized Solutions, DELL EMC, hp, Lenovo

SYSTEM INTEGRATORS
accenture High performance. Delivered., TATA CONSULTANCY SERVICES, Lambda, NCR, NTT DATA, Mobiliya, wipro

VERTICAL

HEALTHCARE
OrboGraph Intelligent Payment Automation, Carestream, SARADA, SIG{ }TUPLE, HEARTVISTA, lu

FINANCIAL SERVICES
WorkFusion, NEXT labs, vPhrase Insights from data, in words., Alpaca, nuueo, Allgoo Intelligence of Banking & A.I., Arya.ai

RETAIL
AiHello, NEXUS EDGE, locaweb, bigdatacorp., GIGASPACES innovate with confidence

TRANSPORTATION
GLOBAL EDGE Intelligence Of Things, Playment, iERA, WI-TRONIX

NEWS, MEDIA & ENTERTAINMENT
sensifai, Gramener A Data Science Company, ZIVA, keemotion, Taboola Content You May Like.

AGRICULTURE
LeBov, tbit

LEGAL & HR
reachr, finch soluções, LEGAL LABS, seedlink PEOPLE SCIENCE

ROBOTIC PROCESS AUTOMATION
UiPath

HORIZONTAL

BUSINESS INTELLIGENCE & ANALYTCS
DataRobot, SIGOPT

VISION
VISIOINGENII Being Different, Matroid, restb.ai, Linker Networks, imagga

CONVERSATIONAL BOTS
[24]7.ai, SET SAIL, nama, avaamo, gamalon

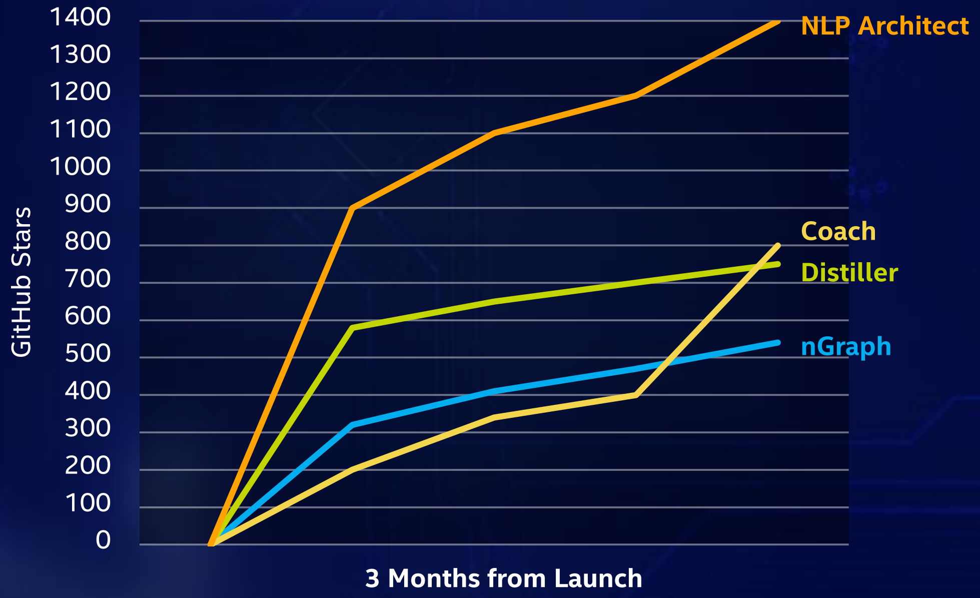
AI TOOLS & CONSULTING
Mighty AI, axondata, iMerit human-empowered computing, SET SAIL, LEAPMIND, Julia computing, cloudera, QuikFynd, CALLIGO TECHNOLOGIES High Performance & Beyond!, UiPath, minds.ai, nologin

AI PAAS
Arya.ai, DOMINO, Paperspace, G3 IoT, H_2O, RocketML, KuberLab

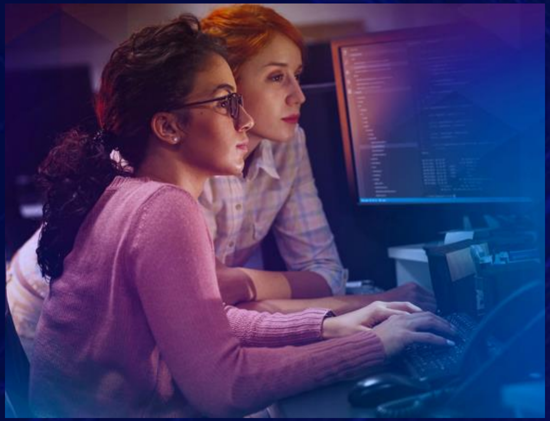
DESIGNED TO ACCELERATE CUSTOMER ADOPTION

ENGAGING WITH DEVELOPERS

OPEN SOURCE COMMUNITY



AI ACADEMY & AI DEVCLOUD



- Trained 110K developers
- Engaged with 90 universities
- 150k users each month, sharing 800+ AI projects

AI DEVELOPERS CONFERENCE

- 950 attendees
- 50+ sessions - 50% by customers, partners & academia
- 90% of sessions standing room only
- Global – US, India, Europe, China

SUMMARY

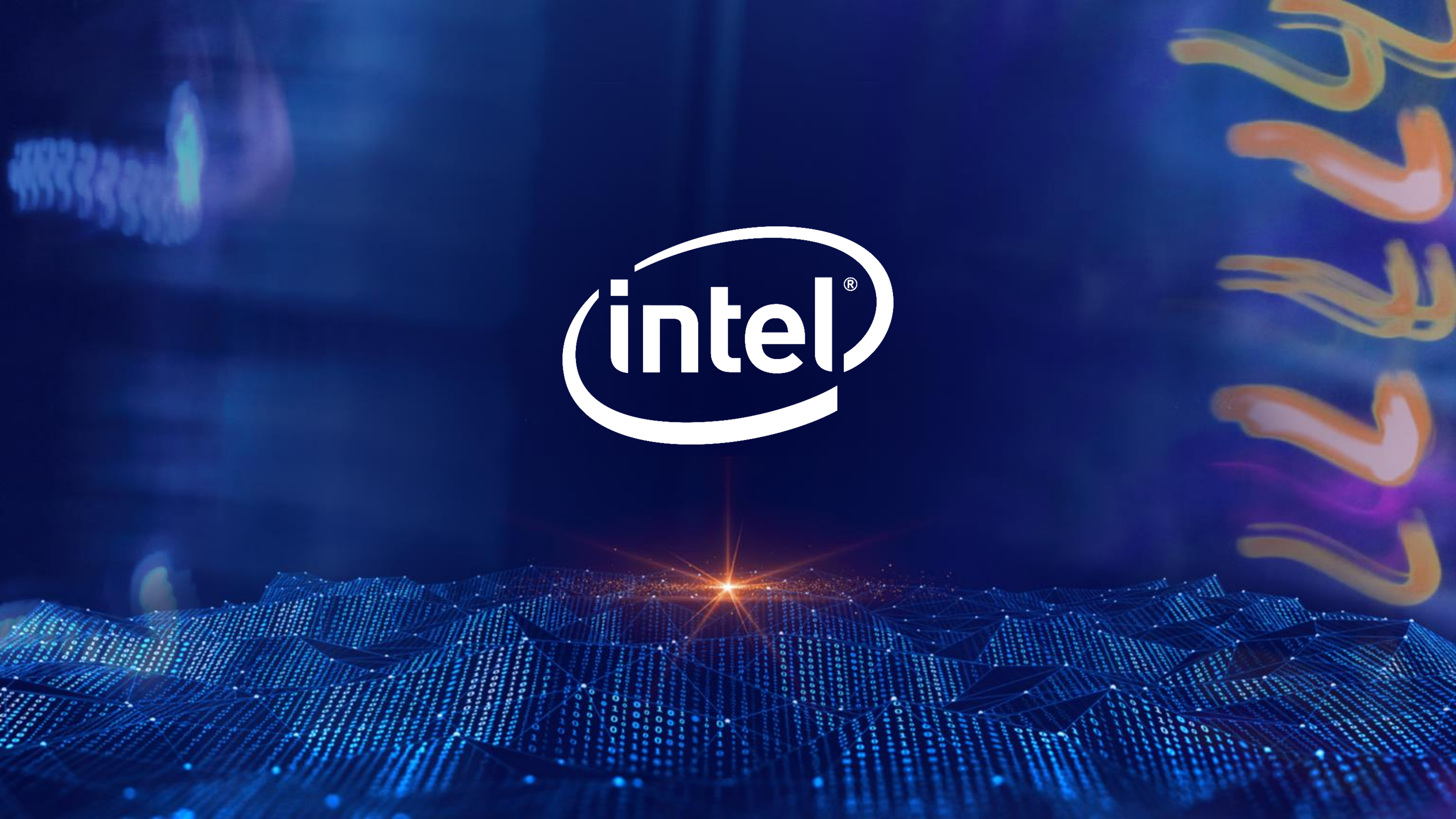
Intel® Xeon® Scalable processors are the foundations for AI, $1B+ business

Delivering tools and software that simplify the development of AI applications

Investing in cutting-edge, purpose-built silicon; engineered for the future of AI

MACHINE LEARNING IMAGE RENDERING

ZIVA


intel

CONFIGURATION DETAILS

1.4x training throughput improvement in August 2018:

Tested by Intel as of measured August 2nd 2018. Processor: 2 socket Intel(R) Xeon(R) Platinum 8180 CPU @ 2.50GHz / 28 cores HT ON , Turbo ON Total Memory 376.46GB (12slots / 32 GB / 2666 MHz). CentOS Linux-7.3.1611-Core kernel 3.10.0-693.11.6.el7.x86_64, SSD sda RS3WC080 HDD 744.1GB,sdb RS3WC080 HDD 1.5TB,sdc RS3WC080 HDD 5.5TB , Deep Learning Framework Intel® Optimizations for caffe version:a3d5b022fe026e9092fc7abc7654b1162ab9940d Topology::resnet_50 BIOS:SE5C620.86B.00.01.0013.030920180427 MKLDNN: version: 464c268e544bae26f9b85a2acb9122c766a4c396 NoDataLayer. Measured: 123 imgs/sec vs Intel tested July 11th 2017 Platform: Platform: 2S Intel® Xeon® Platinum 8180 CPU @ 2.50GHz (28 cores), HT disabled, turbo disabled, scaling governor set to “performance” via intel_pstate driver, 384GB 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 (800GB, 2.5in SATA 6Gb/s, 25nm, MLC).Performance measured with: Environment variables: KMP_AFFINITY='granularity=fine, compact‘, OMP_NUM_THREADS=56, CPU Freq set with cpupower frequency-set -d 2.5G -u 3.8G -g performance. Caffe: (http://github.com/intel/caffe/), revision f96b759f71b2281835f690af267158b82b150b5c. Inference measured with “caffe time --forward_only” command, 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. Topology specs from https://github.com/intel/caffe/tree/master/models/intel_optimized_models (GoogLeNet, AlexNet, and ResNet-50), https://github.com/intel/caffe/tree/master/models/default_vgg_19 (VGG-19), and https://github.com/soumith/convnet-benchmarks/tree/master/caffe/imagenet_winners (ConvNet benchmarks; files were updated to use newer Caffe prototxt format but are functionally equivalent). Intel C++ compiler ver. 17.0.2 20170213, Intel MKL small libraries version 2018.0.20170425. Caffe run with “numactl -l“.

5.4x inference throughput improvement in August 2018:

Tested by Intel as of measured July 26th 2018 :2 socket Intel(R) Xeon(R) Platinum 8180 CPU @ 2.50GHz / 28 cores HT ON , Turbo ON Total Memory 376.46GB (12slots / 32 GB / 2666 MHz). CentOS Linux-7.3.1611-Core, kernel: 3.10.0-862.3.3.el7.x86_64, SSD sda RS3WC080 HDD 744.1GB,sdb RS3WC080 HDD 1.5TB,sdc RS3WC080 HDD 5.5TB , Deep Learning Framework Intel® Optimized caffe version:a3d5b022fe026e9092fc7abc7654b1162ab9940d Topology::resnet_50_v1 BIOS:SE5C620.86B.00.01.0013.030920180427 MKLDNN: version:464c268e544bae26f9b85a2acb9122c766a4c396 instances: 2 instances socket:2 (Results on Intel® Xeon® Scalable Processor were measured running multiple instances of the framework. Methodology described here: https://software.intel.com/en-us/articles/boosting-deep-learning-training-inference-performance-on-xeon-and-xeon-phi) NoDataLayer. Datatype: INT8 Batchsize=64 Measured: 1233.39 imgs/sec vs Tested by Intel as of July 11th 2017:2S Intel® Xeon® Platinum 8180 CPU @ 2.50GHz (28 cores), HT disabled, turbo disabled, scaling governor set to “performance” via intel_pstate driver, 384GB 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 (800GB, 2.5in SATA 6Gb/s, 25nm, MLC).**Performance measured with**: Environment variables: KMP_AFFINITY='granularity=fine, compact‘, OMP_NUM_THREADS=56, CPU Freq set with cpupower frequency-set -d 2.5G -u 3.8G -g performance. Caffe: (http://github.com/intel/caffe/), revision f96b759f71b2281835f690af267158b82b150b5c. Inference measured with “caffe time --forward_only” command, 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. Topology specs from https://github.com/intel/caffe/tree/master/models/intel_optimized_models (ResNet-50). Intel C++ compiler ver. 17.0.2 20170213, Intel MKL small libraries version 2018.0.20170425. Caffe run with “numactl -l“.

11X inference thoughput improvement with CascadeLake:

Future Intel Xeon Scalable processor (codename Cascade Lake) results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance vs Tested by Intel as of July 11th 2017: 2S Intel® Xeon® Platinum 8180 CPU @ 2.50GHz (28 cores), HT disabled, turbo disabled, scaling governor set to “performance” via intel_pstate driver, 384GB 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 (800GB, 2.5in SATA 6Gb/s, 25nm, MLC).**Performance measured with**: Environment variables: KMP_AFFINITY='granularity=fine, compact‘, OMP_NUM_THREADS=56, CPU Freq set with cpupower frequency-set -d 2.5G -u 3.8G -g performance. Caffe: (http://github.com/intel/caffe/), revision f96b759f71b2281835f690af267158b82b150b5c. Inference measured with “caffe time --forward_only” command, 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. Topology specs from https://github.com/intel/caffe/tree/master/models/intel_optimized_models (ResNet-50),. Intel C++ compiler ver. 17.0.2 20170213, Intel MKL small libraries version 2018.0.20170425. Caffe run with “numactl -l“.

CONFIGURATION DETAILS

2.5x Taboola inference Improvement

Intel(R) Xeon(R) Platinum 8180 CPU @ 2.50GHz; 2 Sockets, 56 cores/socket, Hyper-threading ON, Turbo boost OFF, CPU Scaling governor "performance"; RAM: Samsung 192 GB DDR4@2666MHz. (16Gb DIMMS x 12); BIOS: Intel SE5C620.86B.0X.01.0007.062120172125; Hard Disk: INTEL SSDSC2BX01 1.5TB; OS: CentOS Linux release 7.5.1804 (Core) (3.10.0-862.9.1.el7.x86_64)

Baseline: TensorFlow-Serving r1.9 -- https://github.com/tensorflow/serving. Intel Optimized TensorFlow: TensorFlow-Serving r1.9 + Intel MKL-DNN + Optimizations.
MKL-DNN: https://mirror.bazel.build/github.com/intel/mkl-dnn/archive/0c1cf54b63732e5a723c5670f66f6dfb19b64d20.tar.gz
MKLML: https://mirror.bazel.build/github.com/intel/mkl-dnn/releases/download/v0.15/mklml_lnx_2018.0.3.20180406.tgz
Performance results are based on testing as of (08/06/2018) and may not reflect all publicly available security updates. No product can be absolutely secure.

DISCLOSURES

Statements in this presentation that refer to business outlook, future plans and expectations are forward-looking statements that involve a number of risks and uncertainties. Words such as "anticipates," "expects," "intends," "goals," "plans," "believes," "seeks," "estimates," "continues," "may," "will," “would,” "should," “could,” and variations of such words and similar expressions are intended to identify such forward-looking statements. Statements that refer to or are based on projections, uncertain events or assumptions also identify forward-looking statements. Such statements are based on management's current expectations, unless an earlier date is indicated, and involve many risks and uncertainties that could cause actual results to differ materially from those expressed or implied in these forward-looking statements. Important factors that could cause actual results to differ materially from the company's expectations are set forth in Intel's earnings release dated July 26, 2018, which is included as an exhibit to Intel’s Form 8-K furnished to the SEC on such date. Additional information regarding these and other factors that could affect Intel's results is included in Intel's SEC filings, including the company's most recent reports on Forms 10-K and 10-Q. Copies of Intel's Form 10-K, 10-Q and 8-K reports may be obtained by visiting our Investor Relations website at www.intc.com or the SEC's website at www.sec.gov.

All information in this presentation reflects management’s views as of the date of this presentation, unless an earlier date is indicated. Intel does not undertake, and expressly disclaims any duty, to update any statement made in this presentation, whether as a result of new information, new developments or otherwise, except to the extent that disclosure may be required by law.