Efficient Neural Network Training for an AI Radiologist on Intel® Xeon® based Supercomputers

Vikram A. Saletore, Ph.D.
Principal Engineer, AI Products Group, Intel

Valeriu Codreanu, Ph.D. and Damian Podareanu, MSc,
Research & Data Scientist, SURFsara B.V.

Lucas A. Wilson, Ph.D. and Alex Filby
HPC and AI Engineering, Dell EMC

The AI Conference, Sept. 4-7, San Francisco
Agenda

- AI Usages & Performance Drivers
- Efficient Scaling of Neural Network Training on Supercomputers
- AI Radiologist Trained on Intel® Xeon® Scalable Processors
- Call To Action
AI Usage Growth

Consumer
- Smart Assistants
- Chatbots
- Search
- Personalization
- Augmented Reality
- Robots

Health
- Enhanced Diagnostics
- Drug Discovery
- Patient Care
- Research
- Sensory Aids

Finance
- Algorithmic Trading
- Fraud Detection
- Research
- Personal Finance
- Loyalty
- Supply Chain
- Security

Retail
- Support
- Experience
- Marketing
- Merchandising
- Safety & Security
- Resident Engagement
- Smarter Cities

Gov’t
- Defense
- Data Insights
- Smart Grid
- Operational Improvement
- Conservation

Energy
- Oil & Gas Exploration
- Automated Trucks
- Aerospace
- Shipping
- Search & Rescue

Transport
- Autonomous Cars
- Automated Trucks
- Aerospace
- Shipping
- Field Automation

Industrial
- Factory Automation
- Predictive Maintenance
- Precision Agriculture
- Field Automation

Other
- Advertising
- Education
- Gaming
- Professional & IT Services
- Telco/Media
- Space Exploration

Optimization notice
* Other names and brands may be claimed as the property of others.
Performance Drivers for AI Workloads

Compute

The FOUNDATION for AI

Fabric

Deep Learning Frameworks + M
KL-DNN

Intel® Omni-Path™ Architecture Fabric

SW Optimizations
Intel® Xeon® Scalable Processors
the foundation of data center innovation

BCS: https://www.bsc.es/
*TACC (Texas Advanced Computing Center): https://www.tacc.utexas.edu/
*DellEMC HPC and AI Innovation Lab

Architected for efficient, secure, and agile HPC Supercomputing center
Efficient scaling of Neural Network training on supercomputers

Valeriu Codreanu, Ph.D. and Damian Podareanu
Compute Services, SURFsara B.V.
Intel & SURFsara IPCC* Team

Valeriu Codoreanu, Ph.D. (PI), Senior HPC Consultant, SURFsara* B.V., The Netherlands, Intel Parallel Computing Center

Vikram Saletore, Ph.D. (Co-PI), Principal Engineer & Performance Architect, AI Products Group, Intel Corp.

Damian Podareanu, M.Sc., HPC Consultant, SURFsara* B.V., The Netherlands, Intel Parallel Computing Center

*IPCC: Intel® Parallel Computing Centers
IPCC@SURFsara: Scaling up Deep Learning

Research goals:
• Speeding up time-to-train for deep neural network models on large datasets
• Improve convergence accuracy
• Generalization of methodology across Intel® CPU architectures

Main Results
• Efficient scaling
• 512 Intel®2S Xeon® 8160 nodes, with a TTT of 44 minutes on ImageNet-1K
• Improved SOTA using a reduced number of epochs on ImageNet-1K
Accuracy vs Large Batch Size

Datasets

- ImageNet-1K | 1.2 million | 1000 categories => ~1200 examples / class
- Chest-Xray14 | 0.07 million | 14 categories => ~200-20000 examples / class

Training from scratch (< 2% accuracy degradation)

- ImageNet-1K | Batch size up to 32K | ~ 40 updates / epochs | 70-90 epochs

Fine tuning (< 2% accuracy degradation)

- Chest-Xray14 | Batch size up to 8K | ~ 10 updates / epoch | 70-90 epochs
Accuracy, Training Epochs, HW Scaling

- Achieving reasonably good to significantly better accuracy requires:
  - Increased Training time with a fixed level of HW scaling
  - Increased HW scaling for a desired Training Time

- We show results that trade-off accuracy with the number of training epochs
  - >74.0% Top-1 Accuracy
  - >75.5% Top-1 Accuracy
  - >76.5% Top-1 Accuracy

- Using several hardware architectures
  - Intel® Xeon® Platinum Processor Family with Intel® Omni-Path® Architecture (Intel® OPA) Fabric

BCS: https://www.bsc.es/
ResNet-50 Scaling on 2S Intel® Xeon® Platinum 8160 Processor Cluster

- Up to 90% Scaling Efficiency
- Top-1/Top-5 > 74%/92%
- Global BS=8192
- Throughput: 15170 Img/sec
- Time-To-Train: 70 minutes

MareNostrum4 Barcelona Supercomputing Center

Best Practices From SURFsara B.V: https://surfdrive.surf.nl/files/index.php/s/xrEFLPvo7IDRARs
ResNet-50 Training Time to 74% Top-1 Accuracy
Intel® Xeon® Platinum 8160 Processor Cluster MareNostrum4*

Intel® Distribution of Caffe* with ImageNet-1K dataset

<table>
<thead>
<tr>
<th>TRAINING TIME</th>
<th>TRAINING ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>400 Intel® Xeon® Processor Nodes</td>
<td>74.0% Top-1</td>
</tr>
<tr>
<td>56 Minutes</td>
<td>19047 Images/sec</td>
</tr>
<tr>
<td>LOWER IS BETTER</td>
<td>HIGHER IS BETTER</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TRAINING TIME</th>
<th>TRAINING ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>512 Intel® Xeon® Processor Nodes</td>
<td>74.0% Top-1</td>
</tr>
<tr>
<td>44 Minutes</td>
<td>24240 Images/sec</td>
</tr>
<tr>
<td>LOWER IS BETTER</td>
<td>HIGHER IS BETTER</td>
</tr>
</tbody>
</table>

Optimized DL Framework+M
KL-DNN

*MareNostrum4 (Barcelona Supercomputing Center): https://www.bsc.es/marenostrum

Intel® Xeon® Platinum 8160 Processor Cluster MareNostrum4:
- Global BS= 6400 & 50 Epochs
- 19047 Images/sec

Intel® Xeon® Platinum 8160 Processor Cluster MareNostrum4:
- Global BS= 8192 & 50 Epochs
- 24240 Images/sec

In the process of developing the Intel® distribution of Caffe, Intel® performed experiments using two different cluster configurations:
- **Configuration A**:
  - 400 Intel® Xeon® Processor Nodes
  - Global BS= 6400 & 50 Epochs
  - 19047 Images/sec
  - Training Time: 56 Minutes
- **Configuration B**: 512 Intel® Xeon® Processor Nodes
  - Global BS= 8192 & 50 Epochs
  - 24240 Images/sec
  - Training Time: 44 Minutes

Lower is better for training time, higher is better for training accuracy.
Extremely Large Batch Size convergence

Weight decay scaling throughout training eases the optimisation problem further

- 64K batch size: convergence in 2100 iterations to ~74% top-1 accuracy!

<table>
<thead>
<tr>
<th>Batch size</th>
<th>8K</th>
<th>16K</th>
<th>32K</th>
<th>48K</th>
<th>64K</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM [1]</td>
<td>75%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Facebook [2]</td>
<td>76.2%</td>
<td>75.2%</td>
<td>72.4%</td>
<td>-</td>
<td>66%</td>
</tr>
<tr>
<td>You et al. [3]</td>
<td>75.3%</td>
<td>75.3%</td>
<td>74.7%</td>
<td>-</td>
<td>72%</td>
</tr>
<tr>
<td>This work [4]</td>
<td>76.6%</td>
<td>76.3%</td>
<td>75.3%</td>
<td>74%</td>
<td>74%</td>
</tr>
</tbody>
</table>

Increasing Accuracy Using Collapsed Ensembles

Collapsed ensembles

Similar in fashion to the learning-rate collapses:

- However, after performing a partial collapse, LR is again increased
- Cycling the LR:
  - Improves single-model accuracy faster
  - Ensemble of the collapsed points leads to 77.5% accuracy using a ResNet-50 regular training budget

https://github.com/sara-nl/caffe/tree/master/models/intel_optimized_models/multinode/resnet50_custom_lr
Improving Hardware Efficiency

- Using 2 training processes per node increases HW efficiency significantly!
  - Each process has a local batch size of 16.
    At 448 nodes, global batch size is 14336, so no convergence issues.
  - Each process is pinned to a separate NUMA domain
  - Scaling efficiency is not negatively impacted (until 512 nodes)
- Caffe achieves good HW efficiency now!

Comparing efficiency of CPU to GPU-based training of ResNet50. GPU peak performance does not include the CPU hosts

<table>
<thead>
<tr>
<th>Work</th>
<th>HW type</th>
<th># nodes (devices)</th>
<th>Peak [FP32]</th>
<th>TTT</th>
<th>HW eff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>This work</td>
<td>SKX 8160</td>
<td>448 (896)</td>
<td>2682 TF</td>
<td>58 min</td>
<td>12.36</td>
</tr>
<tr>
<td>Facebook [5]</td>
<td>NVIDIA P100</td>
<td>32 (256)</td>
<td>2658 TF</td>
<td>60 min</td>
<td>12.03</td>
</tr>
<tr>
<td>You et al. [26]</td>
<td>SKX 8160</td>
<td>1024 (2014)</td>
<td>6144 TF</td>
<td>45 min</td>
<td>6.51</td>
</tr>
</tbody>
</table>

https://github.com/sara-nl/caffe/tree/master/models/intel_optimized_models/multinode/resnet50_448nodes
## Best Practices To Improve Accuracy of ResNet-50

<table>
<thead>
<tr>
<th>Technique</th>
<th>Approximate top-1 accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default augmentation</td>
<td>74,0%</td>
</tr>
<tr>
<td>Warm-up of LR</td>
<td>75,4%</td>
</tr>
<tr>
<td>Polynomial decay</td>
<td>75,7%</td>
</tr>
<tr>
<td><strong>Weight decay scaling</strong></td>
<td>76,2%</td>
</tr>
<tr>
<td><strong>Single collapse</strong></td>
<td>76,6%</td>
</tr>
<tr>
<td><strong>Collapsed ensembles</strong></td>
<td>77,5%</td>
</tr>
</tbody>
</table>
Summary of Caffe Work

• Extensively evaluated Intel® Xeon® Platinum Processors with Intel® OPA Fabric on training ResNet-50 with ImageNet-1K:
  • >90% scaling efficiency up to 256 nodes to achieve Top-1 >74% Accuracy
  • >85% scaling efficiency from 256 to 512 SKX nodes & achieve 76.5% Top-1
  • NUMA-awareness improves throughput significantly

• Introduced several techniques to improve accuracy:
  • Collapses
  • Weight decay scaling
  • Achieve SOTA at batch sizes of up to 64K

• Models achieve a SOTA of >76.5%+ Top-1 accuracy for ResNet-50 Benchmark

• Collapsed ensemble techniques lead to 77.5% accuracy using ResNet-50
Extending to Tensorflow and to scientific disciplines
Tensorflow Scalability on Intel® Xeon® Processors
ResNet-50 Scaling Efficiency With TensorFlow
Intel® Xeon® Platinum 8160 processor Cluster Stampede2 at TACC

ResNet-50: Training Performance
Intel(R) 2S Xeon(R) on Stampede2/TACC, Intel(R) OPA Fabric
TensorFlow 1.6+horovod, IMPI, ImageNet-1K, Core Aff. Intel BKM,s BS=64/Worker

- 81% scaling efficiency with TensorFlow+horovod
- Ideal Speedup

ResNet-50 with ImageNet-1K on 256 Nodes on Stampede2/TACC:

- Improved single-node perf with multi-workers/node
- 81% scaling efficiency
- Batch size of 64 per worker: Global BS=64K
- 16400 Images/sec on 256 nodes
- 26700 images/sec on 512 nodes
- Time-To-Train: ~2 Hrs on 256 Nodes

First to achieve convergence with state-of-the-art accuracy with TensorFlow on 256 node Intel® Xeon® cluster
Scaling up Training On ImageNet-1K

**Intel® Xeon® Gold 6148F processor Zenith* cluster at DellEMC**

<table>
<thead>
<tr>
<th>Configuration Details</th>
<th>ResNet-50: Time-To-Train (TTT) Perf.</th>
<th>DenseNet-121 Training at Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intel 25 Xeon(R) 6148F, 20C on Zenith/DellEMC, OPA (TM) Fabric</td>
<td></td>
</tr>
<tr>
<td>TensorFlow 1.6-horovod, IMPI, ImageNet-1K, Core Aff. AIPG BKMds, BS=32/Worker</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### DenseNet-121 Training at Scale

<table>
<thead>
<tr>
<th>Global batch size</th>
<th># nodes</th>
<th># epochs</th>
<th>Time/epoch (secs/epoch)</th>
<th>Time-To-Train</th>
<th>% Top1 Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>8192</td>
<td>64</td>
<td>90</td>
<td>346 s</td>
<td>8h40m</td>
<td>74.9</td>
</tr>
<tr>
<td>16384</td>
<td>128</td>
<td>64</td>
<td>187.5 s</td>
<td>3h20</td>
<td>74.5</td>
</tr>
</tbody>
</table>

These models are to be further fine-tuned on the real-world dataset:

**Chest-Xray14**
Transfer Learning Using Highly Accurate Benchmark for Real Use Case

Fine-tuned ResNet-50 that was pre-trained on ImageNet using the Zenith cluster.

To increase accuracy:
• When picking a pre-trained checkpoint do not pick the last one.
• Start with the learning rate at which the model was training when it was checkpointed.
• Perform gradual warmup of the learning rate, proportionally to the global batch size.

Comparative timings for 128-node fine-tuning run

<table>
<thead>
<tr>
<th>Global batch size</th>
<th>Framework</th>
<th># nodes</th>
<th>Time/epoch</th>
</tr>
</thead>
<tbody>
<tr>
<td>4096</td>
<td>Keras</td>
<td>128</td>
<td>85 s</td>
</tr>
<tr>
<td>4096</td>
<td>Tensorflow</td>
<td>128</td>
<td>18 s</td>
</tr>
</tbody>
</table>
An AI Radiologist Trained on Intel® Xeon® Scalable Processors

Automatically Identifying Thoracic Pathologies in Chest X-rays

Lucas A. Wilson, Ph.D. and Alex Filby
HPC and AI Engineering, DellEMC
DellEMC AI Engineering Team – Intel Projects

Onur Celebioglu
Director, HPC and AI Engineering

Quy Ta
Manager, AI Engineering

Lucas A. Wilson
Artificial Intelligence Research

Vineet Gundecha
AI Software Principal Engineer

Srinivas Varadharajan
AI Software Principal Engineer

Pei Yang
AI Software Principal Engineer

Alex Filby
Sr. Systems Development Engineer
The Importance of Early Detection

Emphysema is estimated to affect more than
1. 3 million people in the U.S.
2. 65 million people worldwide
   - Severe emphysema (types 3 / 4) are life threatening
     - Early detection is important to try to halt progression

Pneumonia affects more than 1 million people each year in the U.S.3, and more than 450 million4 each year worldwide.
   - 1.4 million deaths per year worldwide
     - Treatable with early detection

1. www.emphysemafoundation.org/index.php/the-lung/copd-emphysema
4. https://doi.org/10.1016%2FS0140-6736%2810%2961459-6
CheXNet

Developed at Stanford University, CheXNet is a model for identifying thoracic pathologies from the NIH ChestXray14 dataset

- DenseNet121 topology
  - Pretrained on ImageNet
- Dataset contains 112K images
  - Multicategory / Multilabel
  - Unbalanced

http://academictorrents.com/details/557481faacd824c83fbf57dcf7b6da9383b3235a
https://stanfordmlgroup.github.io/projects/chexnet/
Building CheXNet
Training CheXNet

AUROC – Single Process (BZ=8)

Dell EMC C6420 with 2S Intel® Xeon® Scalable Gold 6148

High-accuracy model

✓ 84% accuracy identifying pneumonia
✓ 89% accuracy identifying emphysema

Baseline performance on CPUs

- 4 images per second
- 1 epoch takes 5 hours!

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, including the limitations applicable to this product.
Parallelizing CheXNet
Faster Model Development with Distributed Deep Learning

1. Broadcast initial model
2. Distribute images
3. Shuffle images
4. Forward pass (inference)
5. Aggregate Gradients
6. Update model based on aggregated gradients
7. Repeat steps 3-6
CheXNet – Parallel Speedup

46x Speedup using 32 Nodes!
(64 processes - 2 processes/node)

Dell EMC PowerEdge C6420 with dual Intel® Xeon® Scalable Gold 6148 on Intel® Omni-Path fabric.

Training time reduced from 5 hours per epoch to 7 minutes!
Parallelizing CheXNet - Accuracy

![Graph showing accuracy of parallelizing CheXNet]

- **Cardiomegaly**
- **Emphysema**
- **Effusion**
- **Hernia**
- **Nodule**
- **Pneumonia**
- **Atelectasis**
- **PT**
- **Mass**
- **Edema**
- **Consolidation**
- **Infiltration**
- **Fibrosis**
- **Pneumothorax**

- **P=1, GBZ=8**
- **P=64, GBZ=4096, 100 Epochs**

Performance results are based on testing as of May 17, 2018 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, 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](http://www.intel.com/performance)
Parallelizing CheXNet – Accuracy Relative to single-process

<table>
<thead>
<tr>
<th>Condition</th>
<th>Relative Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumothorax</td>
<td>-6.21%</td>
</tr>
<tr>
<td>Fibrosis</td>
<td>-7.92%</td>
</tr>
<tr>
<td>Infiltration</td>
<td>-1.18%</td>
</tr>
<tr>
<td>Consolidation</td>
<td>-1.47%</td>
</tr>
<tr>
<td>Edema</td>
<td>-4.27%</td>
</tr>
<tr>
<td>Mass</td>
<td></td>
</tr>
<tr>
<td>PT</td>
<td></td>
</tr>
<tr>
<td>Atelectasis</td>
<td></td>
</tr>
<tr>
<td>Pneumonia</td>
<td>-12.33%</td>
</tr>
<tr>
<td>Nodule</td>
<td>-9.66%</td>
</tr>
<tr>
<td>Hernia</td>
<td>-6.53%</td>
</tr>
<tr>
<td>Effusion</td>
<td>-5.24%</td>
</tr>
<tr>
<td>Emphysema</td>
<td>-2.63%</td>
</tr>
<tr>
<td>Cardiomegaly</td>
<td></td>
</tr>
</tbody>
</table>

Configuration Details: Performance results are based on testing as of May 17, 2018 and may not reflect all publicly available security update. See configuration disclosure for details. No product can be absolutely secure. Optimization notices refer to common use or use not common to the product. 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, and functions. Any change to any of those factors may cause the results to vary. Software and workloads used in performance tests may have been optimized. You should consult the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. 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. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Performance tests may have been optimized for specific Intel microprocessors and may not be optimized for other microprocessors. These results may not be representative of all Intel microprocessor-based products and should not be considered for evaluation of all Intel microprocessors. These results may not be valid for all microprocessor-based products. These results are not a measurement of Intel microprocessor performance.
Can We Do Better?

DenseNet121 is a very deep topology with lots of batch normalization

- Batch normalization with large batches (thousands) can hinder convergence

VGG16 and ResNet50 are shallower topologies with less batch normalization

- ResNet50 contains less than half the batch normalization layers of DenseNet121
- VGG16 has no batch normalization

Why not try another topology?
Accuracy of VGG16 relative to DenseNet-121

Better accuracy in 4 categories with parallel trained VGG16

VGG16,P=128,GBZ=8192,260 Epochs
Accuracy of ResNet50 relative to DenseNet-121

Better accuracy in 10 categories with parallel trained ResNet50!

- Cardiomegaly
- Emphysema
- Effusion
- Hernia
- Nodule
- Pneumonia
- Atelectasis
- PT Mass
- Edema
- Consolidation
- Infiltration
- Fibrosis
- Pneumothorax

ResNet50, P=512, GBZ=8192, 103 Epochs
ResNet50, P=512, GBZ=4096, 83 Epochs

-3.00%  -2.00%  -1.00%  0.00%  1.00%  2.00%  3.00%  4.00%  5.00%
Categorical Accuracy of ResNet-50 based AI Radiologist

- **90% accuracy** identifying emphysema
- **88% accuracy** identifying pneumonia

**AUROC – Validation Set**

- **ResNet50, P=512, GBZ=4096, 83 Epochs**
- **ResNet50, P=512, GBZ=8192, 103 Epochs**
- **ResNet50, P=800, GBZ=8000, 72 Epochs**

Configuration Details at the end. Performance results are based on testing as of May 17, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. Optimization notices do not apply to older versions of BIOS or operating systems. For information about performance of Intel® products in specific applications, visit the Intel® website or contact Intel®. Performance results are based on testing as of May 17, 2018 and may not be reflective of all publicly available security updates. No product can be absolutely secure. Optimization notices do not apply to older versions of BIOS or operating systems. For information about performance of Intel® products in specific applications, visit the Intel® website or contact Intel®. Certain optimizations defined by Intel® are unique to Intel® microprocessors and may not be used in other processors. System administrator(s) may need to verify compatibility with other products. For more complete information visit: [http://www.intel.com/performance](http://www.intel.com/performance)
Training Throughput with VGG and ResNet

**Time to Solution DenseNet vs VGG vs ResNet**

<table>
<thead>
<tr>
<th>Model</th>
<th>P</th>
<th>GBZ</th>
<th>Time to Solution (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DenseNet121, P=1, BZ=8</td>
<td>386845</td>
<td></td>
<td>4.5 DAYS to reach a solution with DenseNet121 using 2 Intel® Xeon® Scalable Gold 6148 processors!</td>
</tr>
<tr>
<td>DenseNet121, P=64, BZ=64, GBZ=4096</td>
<td>8319</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VGG16, P=128, GBZ=8192</td>
<td>16532</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ResNet50, P=512, GBZ=4096</td>
<td>1742</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ResNet50, P=512, GBZ=8192</td>
<td>1362</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ResNet50, P=800, GBZ=8000</td>
<td>825</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ResNet50, P=1024, GBZ=8192</td>
<td>675</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**573x FASTER total time to solution going from 1 to 256 Dell EMC PowerEdge C6420 nodes**

**11.25 MINUTES to reach a solution with ResNet50 using 512 Intel® Xeon® Scalable Gold 6148 processors!**
Call To Action

More information at
ai.intel.com/framework-optimizations/

- Tensorflow: https://ai.intel.com/tensorflow/
- Dell EMC Ready Solutions for AI Blog: https://community.dell EMC.com/community/products/rs_for_ai

Use Intel’s performance-optimized libraries & frameworks
Contact us/Intel for help and collaboration opportunities
**Stampede2*/TACC* Configuration Details**

**Stampede2/TACC**: https://portal.tacc.utexas.edu/user-guides/stampede2

**Compute Nodes**: 2 sockets Intel® Xeon® Platinum 8160 CPU with 24 cores each @ 2.10GHz for a total of 48 cores per node, 2 Threads per core, L1d 32K; L1i cache 32K; L2 cache 1024K; L3 cache 33792K, 96 GB of DDR4, Intel® Omni-Path Host Fabric Interface, dual-rail. Software: Intel® MPI Library 2017 Update 4Intel® MPI Library 2019 Technical Preview OFI 1.5.0PSM2 w/ Multi-EP, 10 Gbit Ethernet, 200 GB local SSD, Red Hat* Enterprise Linux 6.7.

TensorFlow 1.6: Built & Installed from source: [https://www.tensorflow.org/install/install_sources](https://www.tensorflow.org/install/install_sources)

**Model**: Topology specs from [https://github.com/tensorflow/tpu/tree/master/models/official/resnet](https://github.com/tensorflow/tpu/tree/master/models/official/resnet) (ResNet-50); Batch size as stated in the performance chart


**Performance measured on 256 Nodes with**:  
OMP_NUM_THREADS=24 HOROVOD_FUSION_THRESHOLD=134217728 export I_MPI_FABRICS=tm1, export I_MPI_TMI_PROVIDER=psm2 \  
mpirun -np 512 -ppn 2 python resnet_main.py --train_batch_size 8192 --train_steps 14075 --num_intra_threads 24 --num_inter_threads 2 --  
mkl=True --data_dir=/scratch/04611/valeriuc/tf-1.6/tpu_rec/train --model_dir model_batch_8k_90ep --use_tpu=False --kmp_blocktime 1

**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/performanc](http://www.intel.com/performanc)
**DellEMC Zenith Cluster Configuration Details**

**DellEMC Internal Cluster:**

**Compute Nodes:** 2 sockets Intel® Xeon® Gold 6148F CPU with 20 cores each @ 2.40GHz for a total of 40 cores per node, 2 Threads per core, L1d 32K; L1i cache 32K; L2 cache 1024K; L3 cache 33792K, 96 GB of DDR4, Intel® Omni-Path Host Fabric Interface, dual-rail. Software: Intel® MPI Library 2017 Update 4Intel® MPI Library 2019 Technical Preview OFI 1.5.0PSM2 w/ Multi-EP, 10 Gbit Ethernet, 200 GB local SSD, Red Hat* Enterprise Linux 6.7.

TensorFlow 1.6: Built & Installed from source: https://www.tensorflow.org/install/install_sources


DenseNet-121 Model: Topology specs from https://github.com/liuzhuang13/DenseNet

Convergence & Performance Model: https://surfdrive.surf.nl/files/index.php/s/xrEFLPvo7IDRARs

Dataset:
ChexNet: https://stanfordmlgroup.github.io/projects/chexnet/

Performance measured with:
OMP_NUM_THREADS=24 HOROVOD_FUSION_THRESHOLD=134217728 export I_MPI_FABRICS=tmi, export I_MPI_TMI_PROVIDER=psm2 \
mpirun -np 512 -ppn 2 python resnet_main.py --train_batch_size 8192 --train_steps 14075 --num_intra_threads 24 --num_inter_threads 2 --
mkl=True --data_dir=/scratch/04611/valeriuc/tf-1.6/tpu_rec/train --model_dir model_batch_8k_90ep --use_tpu=False --kmp_blocktime 1

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/performace
MareNostrum4/BSC* Configuration Details

*MareNostrum4/Barcelona Supercomputing Center: [https://www.bsc.es/](https://www.bsc.es/)

**Compute Nodes:** 2 sockets Intel® Xeon® Platinum 8160 CPU with 24 cores each @ 2.10GHz for a total of 48 cores per node, 2 Threads per core, L1d 32K; L1i cache 32K; L2 cache 1024K; L3 cache 33792K, 96 GB of DDR4, Intel® Omni-Path Host Fabric Interface, dual-rail. Software: Intel® MPI Library 2017 Update 4Intel® MPI Library 2019 Technical Preview OFI 1.5.0PSM2 w/ Multi-EP, 10 Gbit Ethernet, 200 GB local SSD, Red Hat* Enterprise Linux 6.7.


**Model:** Topology specs from [https://github.com/intel/caffe/tree/master/models/intel_optimized_models](https://github.com/intel/caffe/tree/master/models/intel_optimized_models) (ResNet-50) and modified for wide-RedNet-50. Batch size as stated in the performance chart

**Time-To-Train:** measured using “train” command. Data copied to memory on all nodes in the cluster before training. No input image data transferred over the fabric while training; Performance measured for node count: 128, 192, 256, 400, 512 & Performance projected for node count: 1-64.

**Performance measured with:**

export OMP_NUM_THREADS=44 (the remaining 4 cores are used for driving communication), export I_MPI_FABRICS=tmi, export I_MPI_TMI_PROVIDER=psm2

OMP_NUM_THREADS=44 KMP_AFFINITY="=proclist=[0-87],granularity=thread,explicit" KMP_HW_SUBSET=1t MLSL_NUM_SERVERS=4 mpirun -maprov=pm2 -mca sml no -n $SLURM_JOB_NUM_NODES -ppn 1 -t hosts2 -mca mpi_genv OMP_NUM_THREADS 44 -env KMP_AFFINITY="proclist=[0-87],granularity=thread,explicit" -env KMP_HW_SUBSET 1t -mca mpi_genv L_MPI_FABRICS tmi -mca mpi_genv L_MPI_HYDRA_BRANCH_COUNT $SLURM_JOB_NUM_NODES -mca mpi_genv L_MPI_HYDRA_PMICONNECT alttoall sh -c 'cat /lsivrc12_train_lmdb_striped_64/data mdb > /dev/null ; cat /lsivrc12_val_lmdb_striped_64/data mdb > /dev/null ; ulimit -u 8192 ; ulimit -a ; numactl -H ; /caffe/build/tools/caffe train --solver=/caffe/models/intel_optimized_models/multinode/resnet_50_256_nodes_8k_batch/solver_poly_quick_large.prototxt -engine "MKL2017"

Optimizer 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 use by Intel. 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, operating systems, 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 evaluating your anticipated applications, including the performance of that product when combined with other products. For more complete information visit: [http://www.intel.com/performance](http://www.intel.com/performance)