

DESIGN AND PERFORMANCE EVALUATION OF A COMMODITY GPU CLUSTER FOR HPC AND DEEP LEARNING WORKLOADS

Valeriu Codreanu, Ph.D.

Senior HPC Consultant, SURFsara B.V.

Outline

- SURFsara and GPU usage
- Motivations
- Design decisions HW & SW infrastructure
- Performance evaluation Deep Learning
- Performance evaluation GROMACS
- Performance evaluation Cryo-EM (Relion)
- Successful applications

SURFsara: Dutch National Infrastructure/Service Provider

History:

- 1971: Founded by the VU, UvA, and CWI
- 2013: SARA (Stichting Academisch Rekencentrum A'dam) becomes part of SURF

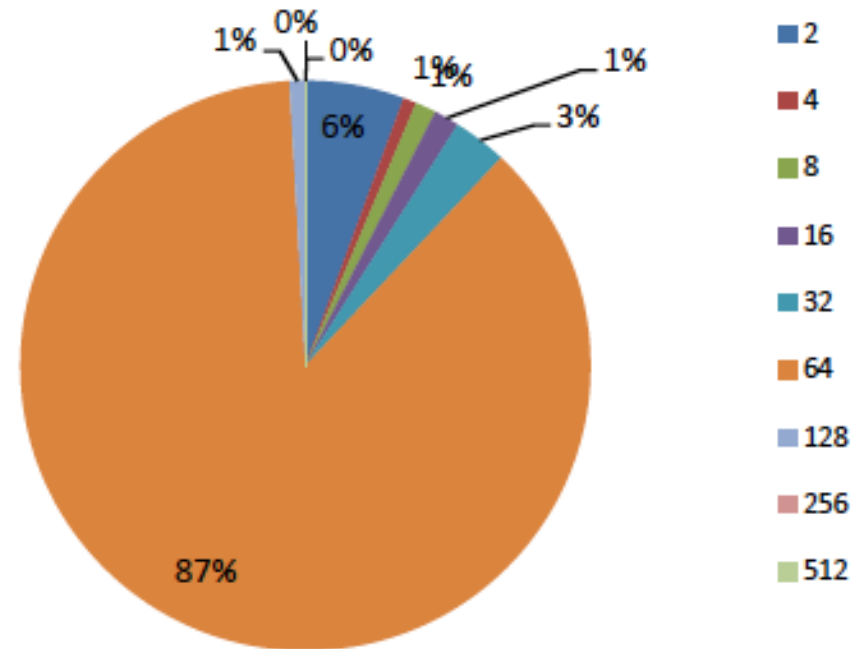
Cartesius (Bull supercomputer):

- 40.960 Ivy Bridge / Haswell cores: 1327 TFLOPS
- 56Gbit/s Infiniband
- 64 nodes with 2 x K40m GPUs each: 210 TFLOPS
- Broadwell & KNL extension (Nov 2016)
- 177 BDW and 18 KNL nodes: 284TFLOPS
 - Number 359 in latest Top500 (Nov. 2018) ☹️



Usage of GPUs on Cartesius - 2015

Distribution of SBU by gpu-counts 2015

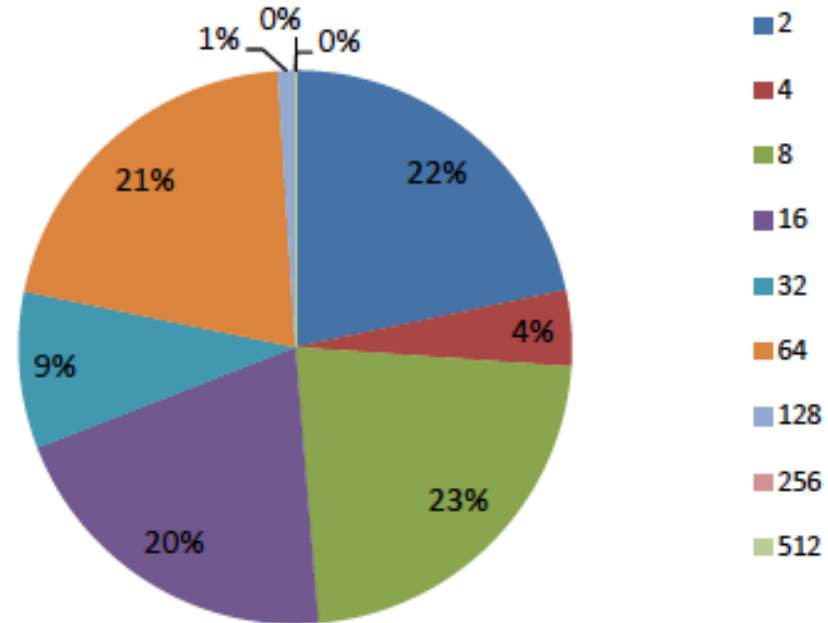


- Comes mainly from large jobs
 - From few large users
 - 32-node jobs dominate

Usage of GPUs on Cartesius - 2016

- Number of GPU users is growing
 - But still usage is dominated by large, experienced users
 - Only 22% of the GPU jobs are using a single-node

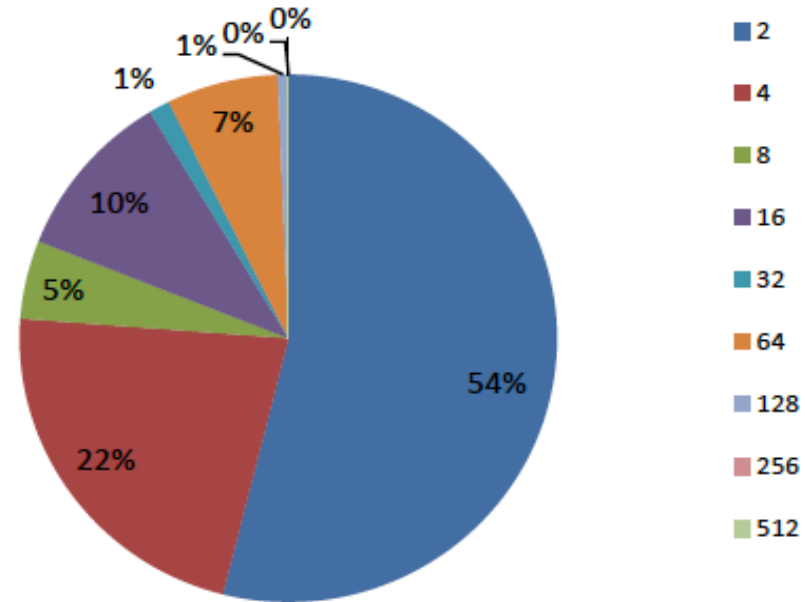
Distribution of SBU by gpu-counts 2016



Usage of GPUs on Cartesius - 2017

- But it's changing
 - So far almost 50% GPU SBUs are used on single-node jobs
 - Deep Learning, Cryo-EM,...
 - More GPU users, less GPU/CUDA developers
 - Libraries have succeeded!
 - **Particularly in Deep Learning**

Distribution of SBU by gpu-counts 2017



The LISA Compute Cluster

LISA (Dell cluster):

- 7484 cores (16 cores per node, Xeon E5-2650, Xeon Silver 4110)
 - Heterogeneous - multiple interconnect options (FDR Infiniband / 10Gbit Ethernet)
- Supports large community of users:
 - Genetics community is a very big user
- Peak performance: 149 TFLOPS
- Uses Torque as scheduler/resource manager

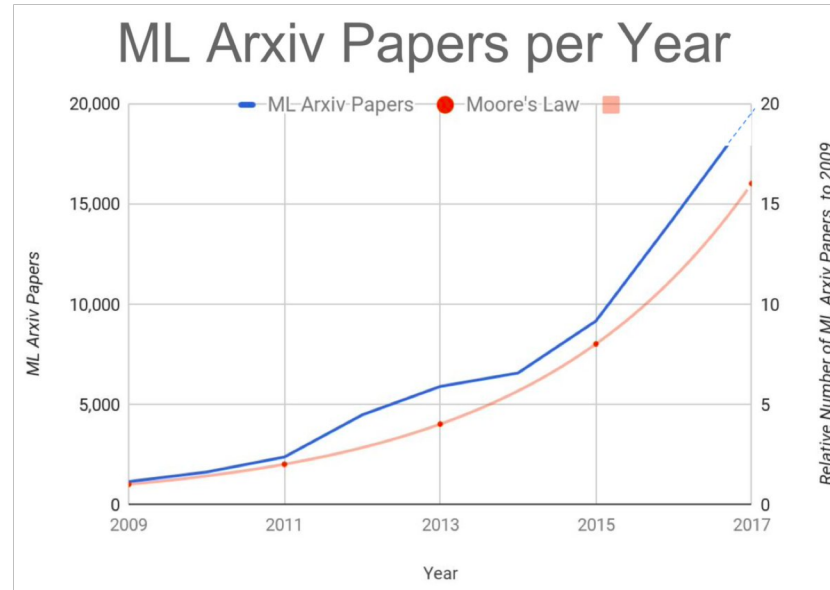
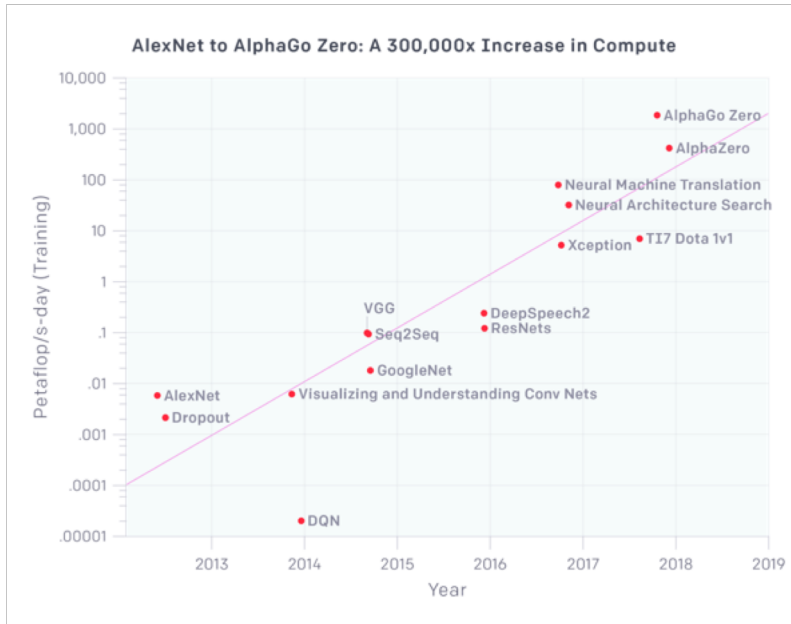
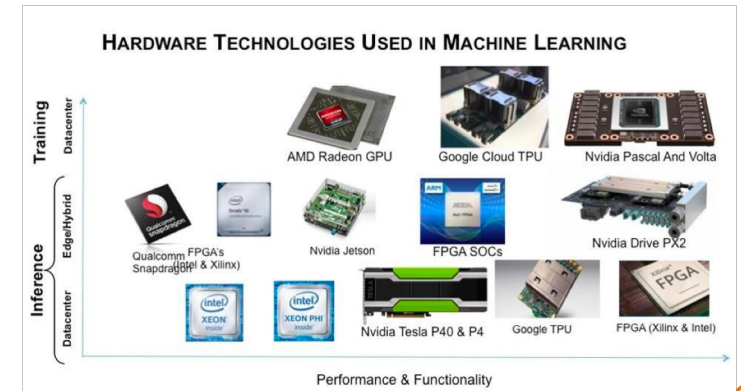


We noticed that the machine (deep) learning community had some special requests, not fitting neither LISA nor Cartesius!

dense GPU nodes

Isn't ML like other fields?

- Very dynamic field. Frameworks, programming languages, algorithm designs, hardware, and libraries change rapidly
- It is not trivial to achieve high-performance solutions



Isn't ML like other fields?

Complex hardware / (system) software integration!

INTEL AI PORTFOLIO

EXPERIENCES: Icons for various AI applications like autonomous vehicles, drones, and smart home devices.

TOOLS: Intel® Deep Learning Deployment Toolkit, Intel® Computer Vision SDK, Movidius Neural Compute Stick, Saffron Technology*.

FRAMEWORKS: Apache Spark*, MLlib, mxnet, Microsoft CNTK*, theano*, Caffe, E2E Tool, torch.

LIBRARIES: Intel Dist, Intel® DAAL, Intel® Nervana™ Graph*, Intel® MKL, MKL-DNN, Intel® MSL, Movidius™ MyTensor Library, Associative Memory Base.

HARDWARE: Intel CPUs, GPUs, and accelerators categorized into Compute, Memory and Storage Networking, and Visual Intelligence.

UNLEASH FULL POTENTIAL

IBM PowerAI Platform

PowerAI Software Distribution

Deep Learning Frameworks: Caffe, NVIDIA Caffe, IBM Caffe, torch, TensorFlow, theano, Chainer.

Supporting Libraries: DIGITS, OpenBLAS, Distributed Frameworks, Bazel, NCCL.

IBM Power System for HPC, with NVLink

Breakthrough performance for GPU accelerated applications, Including Deep Learning and Machine Learning

ROCmSoftware

Machine Learning Applications

Frameworks: Caffe, TensorFlow, Torch 7, MxNet, CNTK, Chainer, Theano.

Middleware & Libraries: MIOpen, BLAS, FFT, RNG, NCCL, C++ STL.

ROCm: ROC, HIP, OpenCL, Python.

ROCm Platform

AMD | RADEON

ACCELERATED DEEP LEARNING TRAINING STACK

COMPUTER VISION: Image Classification, Object Detection.

SPEECH AND AUDIO: Voice Recognition, Language Translation.

NATURAL LANGUAGE PROCESSING: Recommendation Engines, Sentiment Analysis.

UI / JOB MANAGEMENT / DATASET VERSIONING / VISUALIZATION: DIGITS, NVIDIA GPU Cloud, NVDock, Keras, Kubernetes.

NV OPTIMIZED: TensorFlow, PyTorch, Caffe2, mxnet, Caffe, theano, Bazel, Paddle, Chainer, MATLAB, KALDI, MINERVA, PyLearn2.

NV ACCELERATED: cuDNN, cuBLAS, cuSPARSE, cuFFT, NCCL.

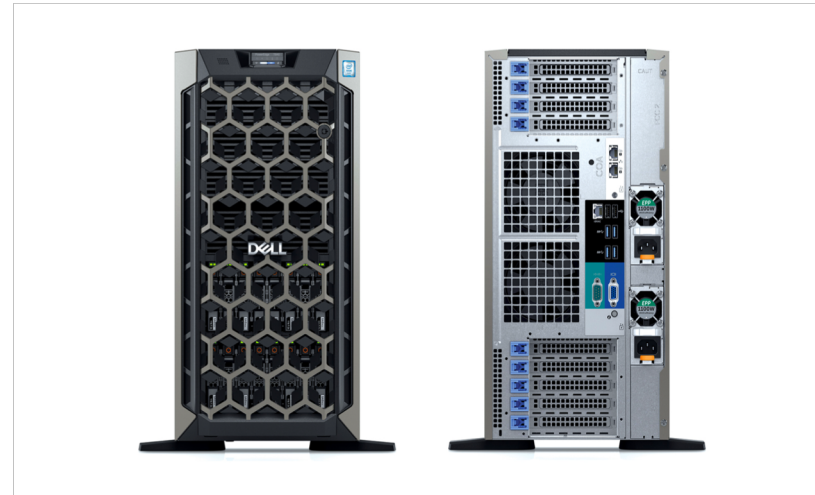
At Your Desk **On-Prem** **In-the-Cloud**

LISA (GPU) Machine Learning Extension – HW Design Requirements

- **GPUs** can provide high levels of performance in dense enclosures (4 modern GPUs -> ~50 TFLOPS of single precision in a single server)
- **(NVIDIA) GPUs** have a very high quality software ecosystem
 - CUDA gets continuous improvements since 2007
 - cuBLAS/cuDNN/NCCL libraries provide efficient (ML) kernels and multi-GPU communication primitives
- I/O is very important in distributed deep learning
 - **Local storage** alleviates the issues (cache the datasets)
- Communication fabric is stressed in distributed deep learning
 - Needs high-performance **communication fabric**
- CPU performance is not generally important, but sometimes it is 😞

LISA (GPU) Machine Learning Extension – HW Design Decisions

- 23 * 4 NVIDIA GTX 1080TI (92 GTX 1080Ti GPUs)
 - ~45 TFLOPs per node (FP32)
- 2 * 4 NVIDIA Titan V (8 Titan V GPUs)
 - ~55 TFLOPs per node (FP32)
 - **~27.5 TFLOPs per node (FP64)**
 - **440 TFLOPs per node (FP16)**



Each node features:

- 2S Intel® Xeon® Bronze 3104 1.7G, 6C/6T, 9.6GT/s UPI, 8M Cache
- 256GB DDR4 2667MT/s
- Intel 40Gb/s Ethernet
- Dell 1.6TB, NVMe, Mixed Use Express Flash, HHHL Card, PM1725

GPU extension – Software infrastructure

- Based on Debian Linux
- Torque is used for job scheduling / resource management
 - We are (currently) transitioning to SLURM.
 - SLURM supports single-GPU requests (effectively splits node using cgroups)
- Software builds are mostly done using **Easybuild**
- **Singularity** containers are supported and recommended
 - Can be downloaded from NVIDIA NGC: <https://ngc.nvidia.com>
 - Multi-node multi-GPU execution using containers is possible
- We regularly update CUDA toolkits for full Tensor Core support (Titan V)



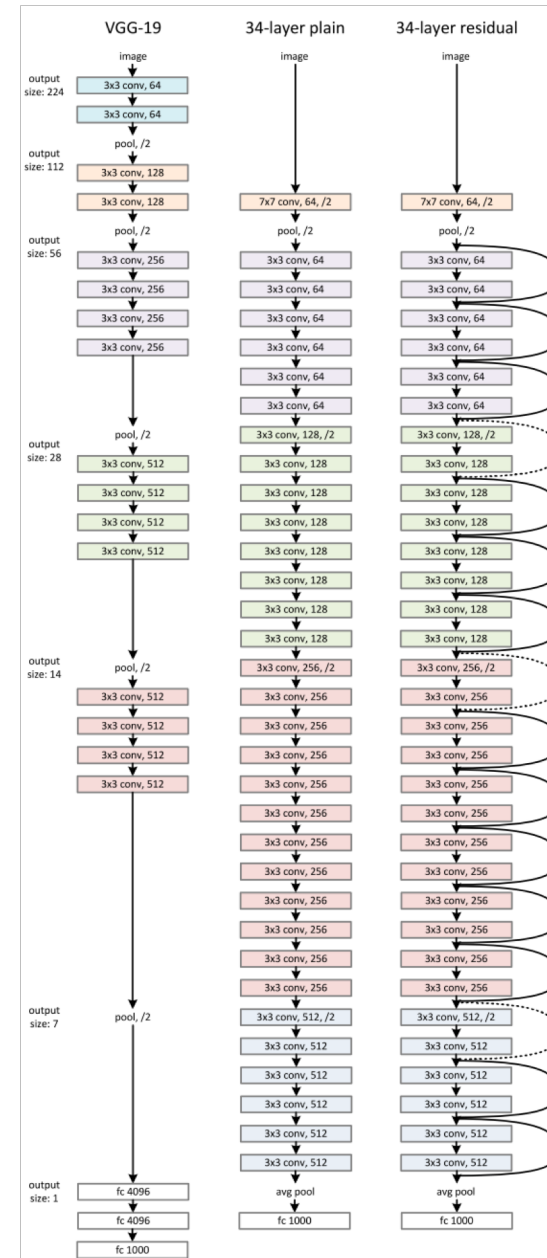
Easybuild advantages



- Automated software installations.
- Software packages and all their dependencies are built with a consistent set of compilers & libraries (known as toolchains).
- Modules automatically load all their dependencies. The compiler used to build the module and its dependencies also gets loaded.
- Tab completion for the module command.
- The same environment on Cartesius and Lisa. Note that not all modules are installed on both systems.
- Software is reinstalled when upgrading the OS. This allows for Lisa and Cartesius to have a mix of OS versions on the nodes, and the correct software build will be loaded automatically.
- A convenient way for you to install complex software locally (EasyBuild supports over 1200 packages) - ebllocalinstall

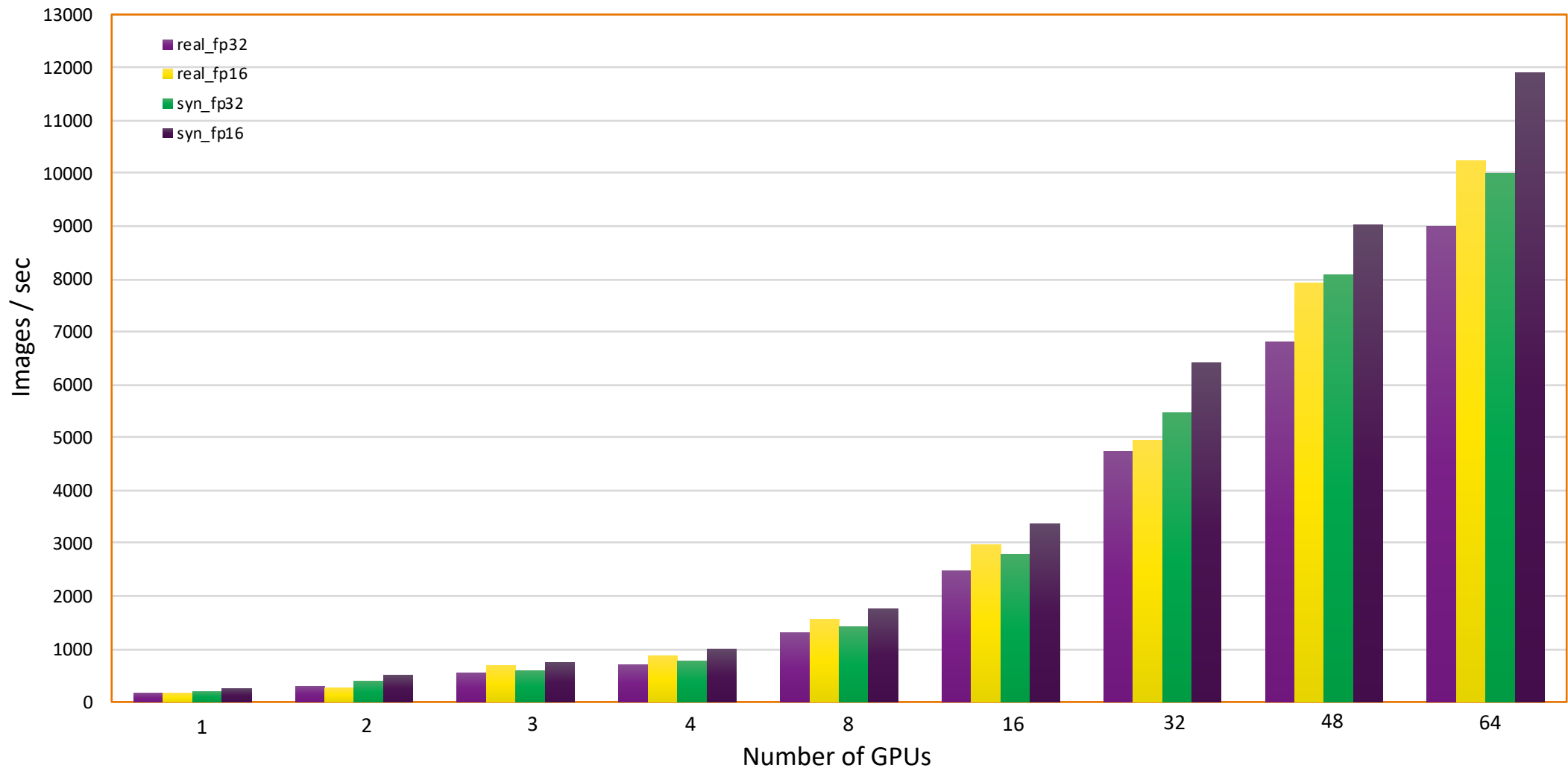
Performance evaluation – Deep Learning

- We selected ResNet-50 training on ImageNet-1K as benchmark and evaluated at scale:
 - TensorFlow
 - PyTorch
 - MXNet
- Benchmarks are performed both using the GPUs from a node, as well as in a distributed fashion across GPU nodes (using data parallelism)



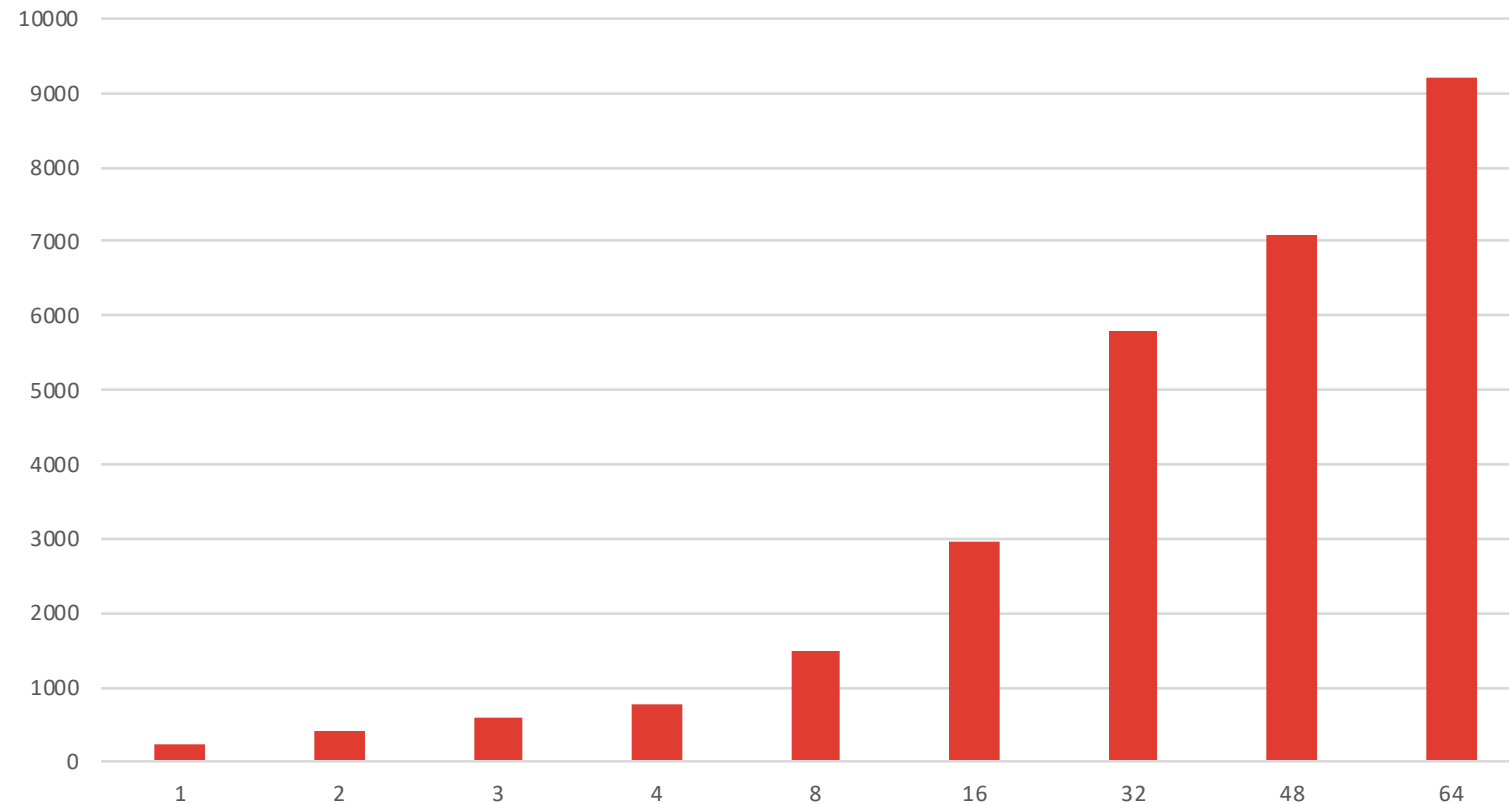
Performance evaluation – Deep Learning - Tensorflow

Tensorflow + Horovod - Scaling behavior of Resnet50 on 1080Ti GPUs with real and synthetic dataset

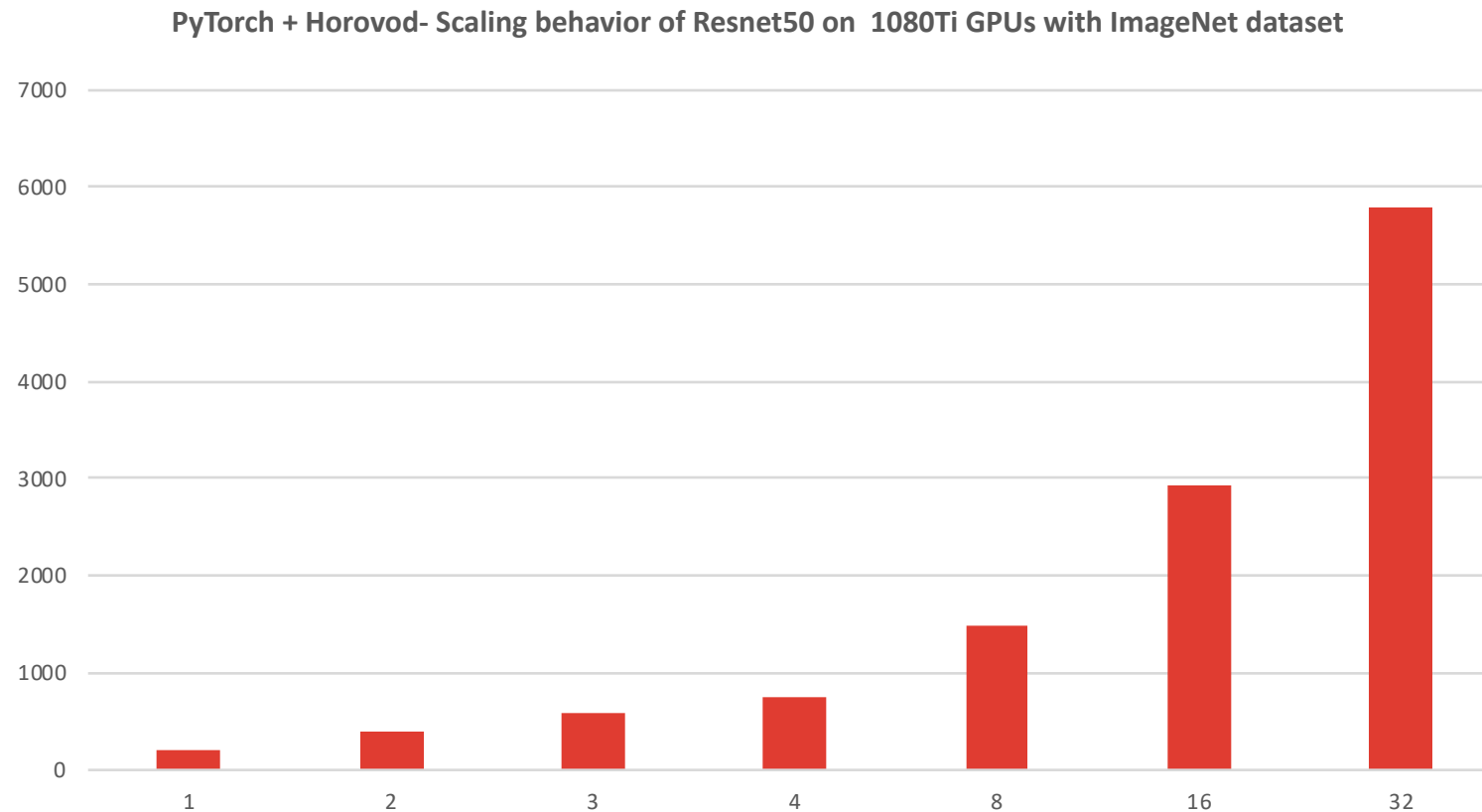


Performance evaluation – Deep Learning - MXNet

MXNet + ZeroMQ- Scaling behavior of Resnet50 on 1080Ti GPUs with ImageNet dataset



Performance evaluation – Deep Learning - PyTorch



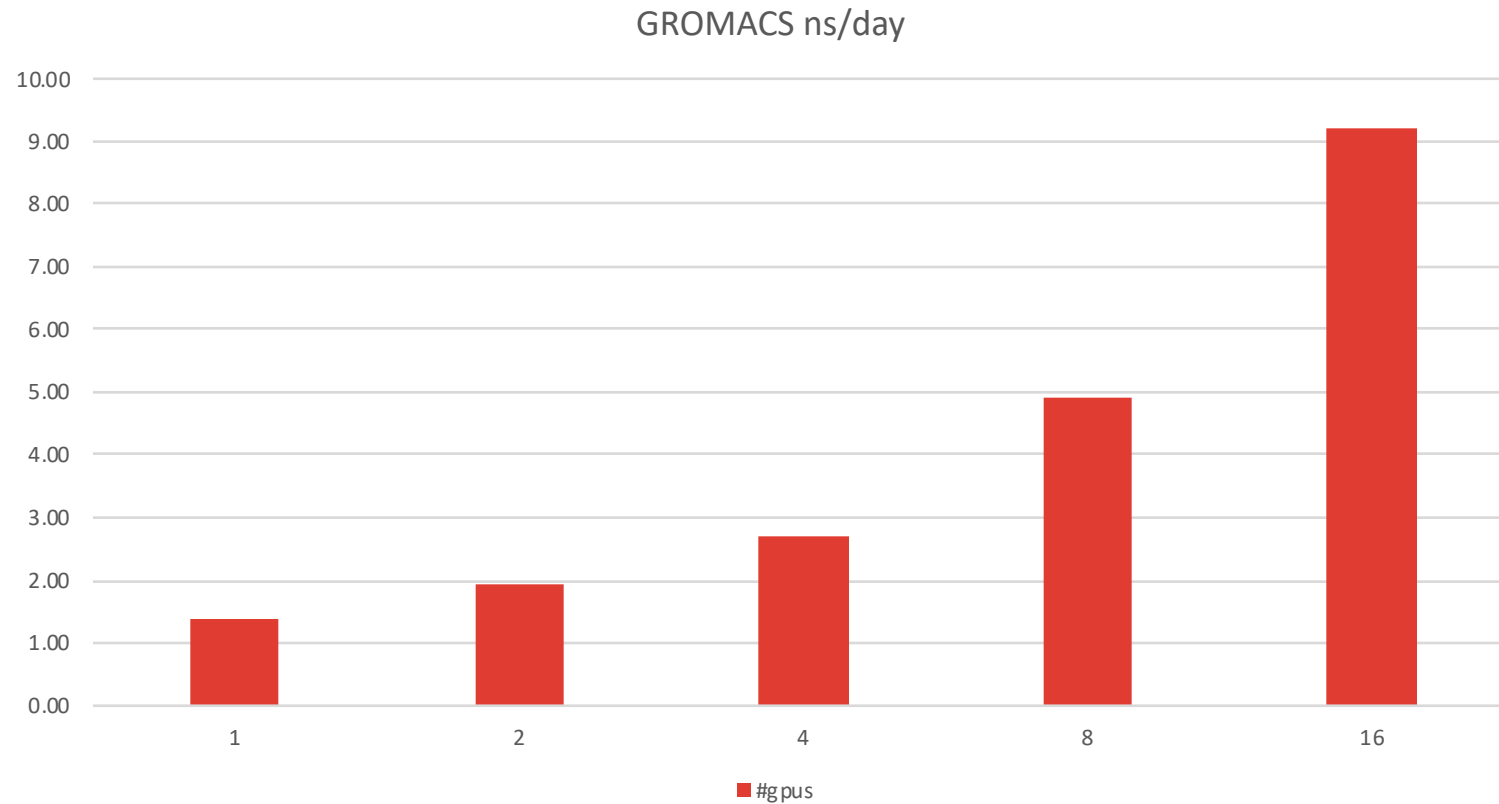
Performance evaluation – GROMACS

Gromacs Test Case B from the Unified European Applications Benchmark Suite.

Description: A model of cellulose and lignocellulosic biomass in an aqueous solution. This system of 3.3M atoms is inhomogeneous.

Example performance: 1 x Cartesius CPU node (Haswell) – 0.9 ns/day

1 x Cartesius GPU node (2x K40) – 1.4 ns/day

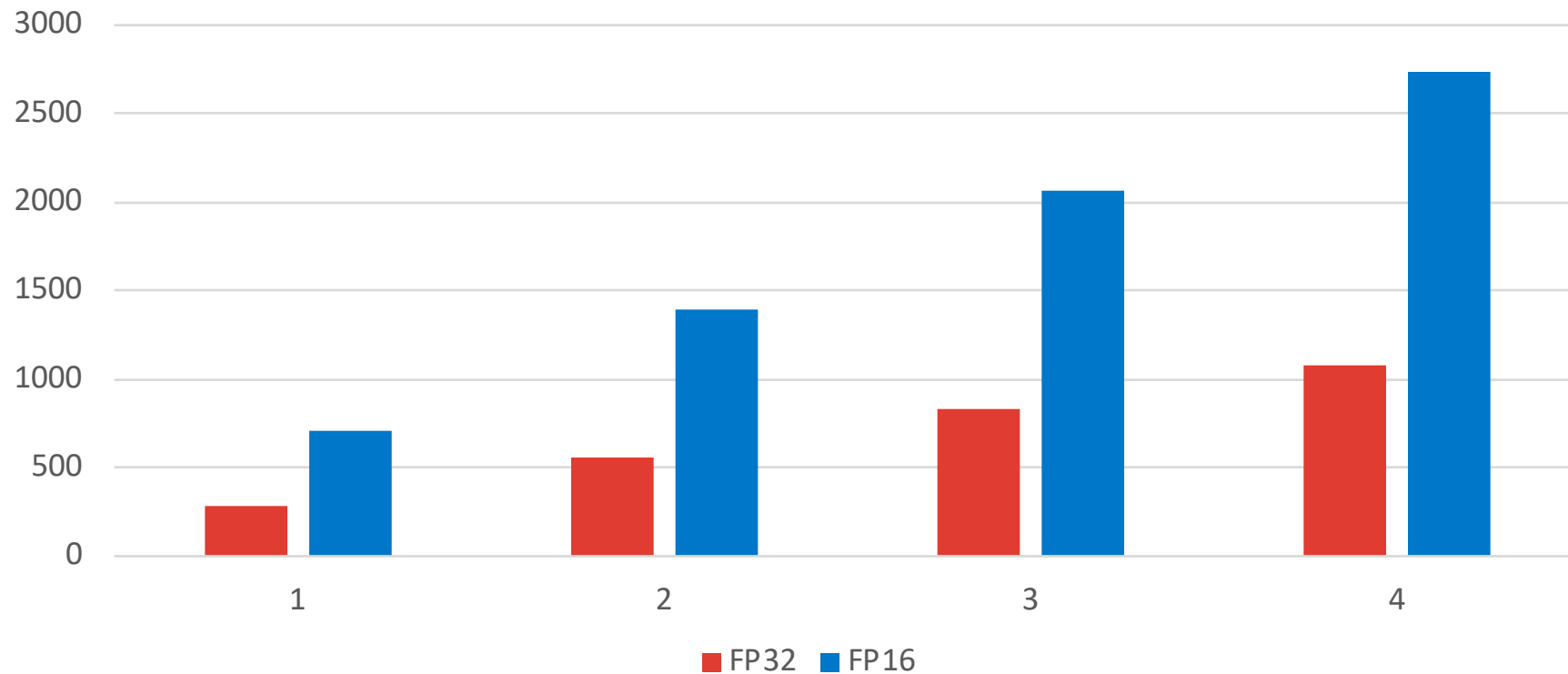


Performance evaluation – Titan V

- Titan V is offering V100-like performance with reduced costs.
- There are few downsides compared to V100, but not significant:
 - Less memory: 12 GB instead of 16/32GB
 - Lack of NVLink
- But is the only Titan card to offer double-precision performance, so quite fit for HPC
- It also offers Tensor Core capabilities, at up to 110TFLOPs per card

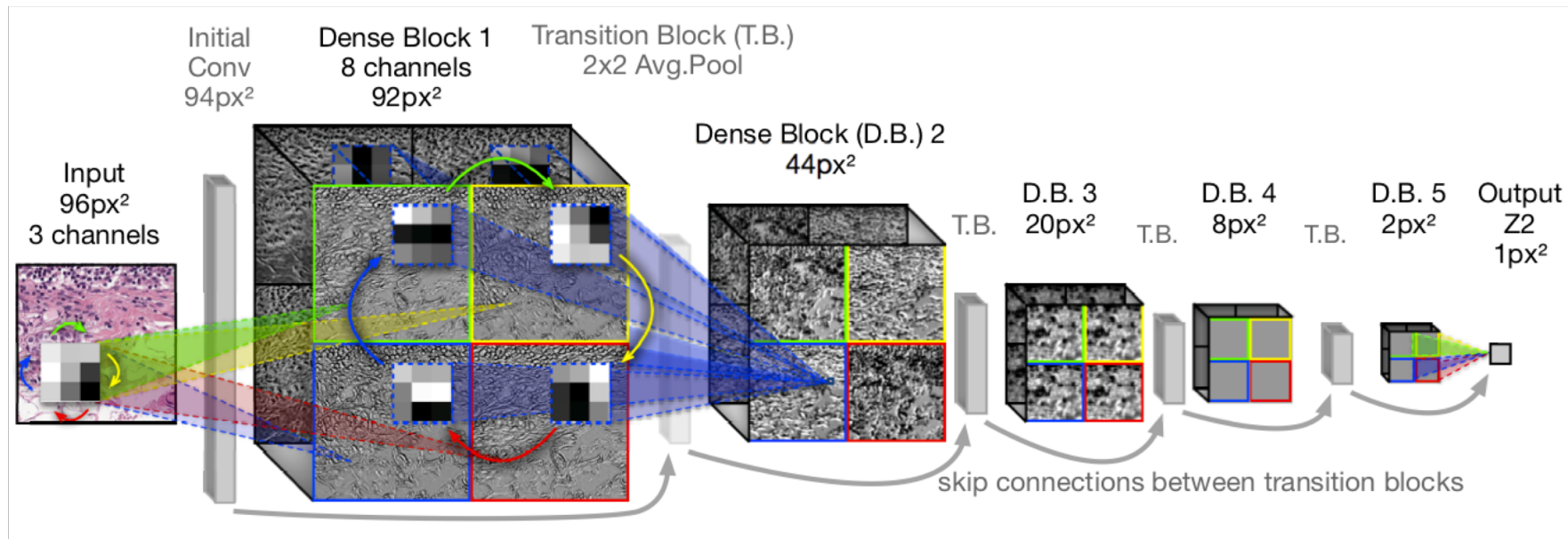
Performance evaluation – Titan V + CUDA10

Tensorflow + Horovod - Scaling behavior of Resnet50 on Titan V GPUs with Imagenet dataset (reduced precision)



Research on LISA GPU – Successful applications

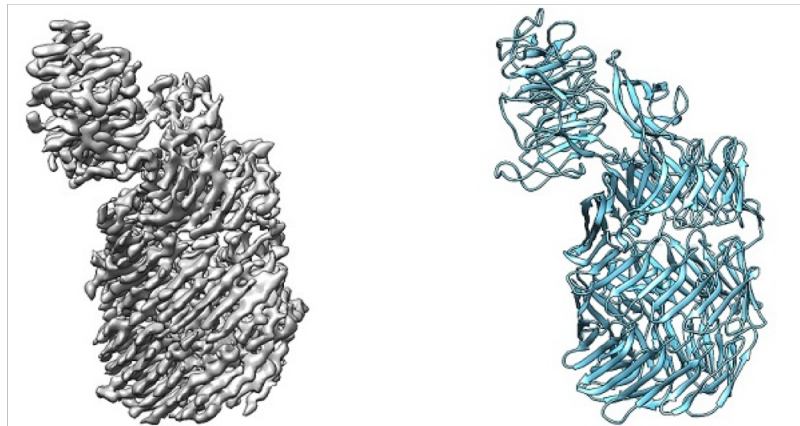
B. S. Veeling, J. Linmans, J. Winkens, T. Cohen, M. Welling. "Rotation Equivariant CNNs for Digital Pathology". arXiv [cs.CV] (2018), (available at <http://arxiv.org/abs/1806.03962>).



Research on LISA GPU – Successful applications

Structures of Teneurin adhesion receptors reveal an ancient fold for cell-cell interaction

Verity A. Jackson, Dimphna H. Meijer, Maria Carrasquero, Laura S. van Bezouwen, Edward D. Lowe, Colin Kleanthous, Bert J. C. Janssen & Elena Seiradake



**1 LISA node is ~3x faster than a Cartesius node for Cryo-EM!
Group has moved from Cartesius to LISA**

Research on LISA GPU – SOIL Deep Learning Enhanced HPC

- Generative networks for high-energy-physics
 - PI: Dr. Sascha Caron, Radboud University Nijmegen
- Machine-learned turbulence in next-generation weather models
 - PI: Dr. Ir. Chiel van Heerwaarden, Wageningen University
- 3DeepFace: Distinguishing biological interfaces from crystal artifacts in biomolecular complexes
 - PI: Prof. Dr. Alexandre Bonvin, Utrecht University
- Machine learning for accelerating planetary dynamics in stellar clusters
 - PI: Prof. Dr. Simon Portegies Zwart, Leiden Observatory

Education on LISA GPU

- Due to increased request from various AI-related courses, we have offered part of the LISA GPU capacity to education (mostly from UvA at the moment).
- Courses such as: Reinforcement Learning, Information Retrieval, Machine Translation, etc. (> 200 students in parallel)
- In order to allow this (and not monopolize the research part of LISA GPU):
 - Moved part of the cluster to a SLURM scheduler/resource manager
 - This allows scheduling single-GPU jobs (effectively splits node resources using cgroups)
 - Students submit max 4h jobs, max 1 job in run queue, max 1 GPU per job – special partition
 - So far, successful experiment!

Lessons learnt / Conclusions

- For further extensions of the cluster, stick to 4 GPUs per node
- Titan V GPUs are fit for multiple application domains (HPC/AI), although more expensive
- The SLURM scheduler makes single-GPU jobs possible
- All resources are shared (CPUs, main memory), including the NVME
- Easybuild and/or Singularity allows for SW reproducibility
- We should consider better performing CPUs to remove current CPU bottlenecks
- Maybe AMD?
- Particularly for GROMACS, that is very used by Dutch research
- The cluster can be also used effectively for education
- We should pursue this and increase acceptance of GPU/heterogeneous computing for future scientists

Thank you!
Q&A