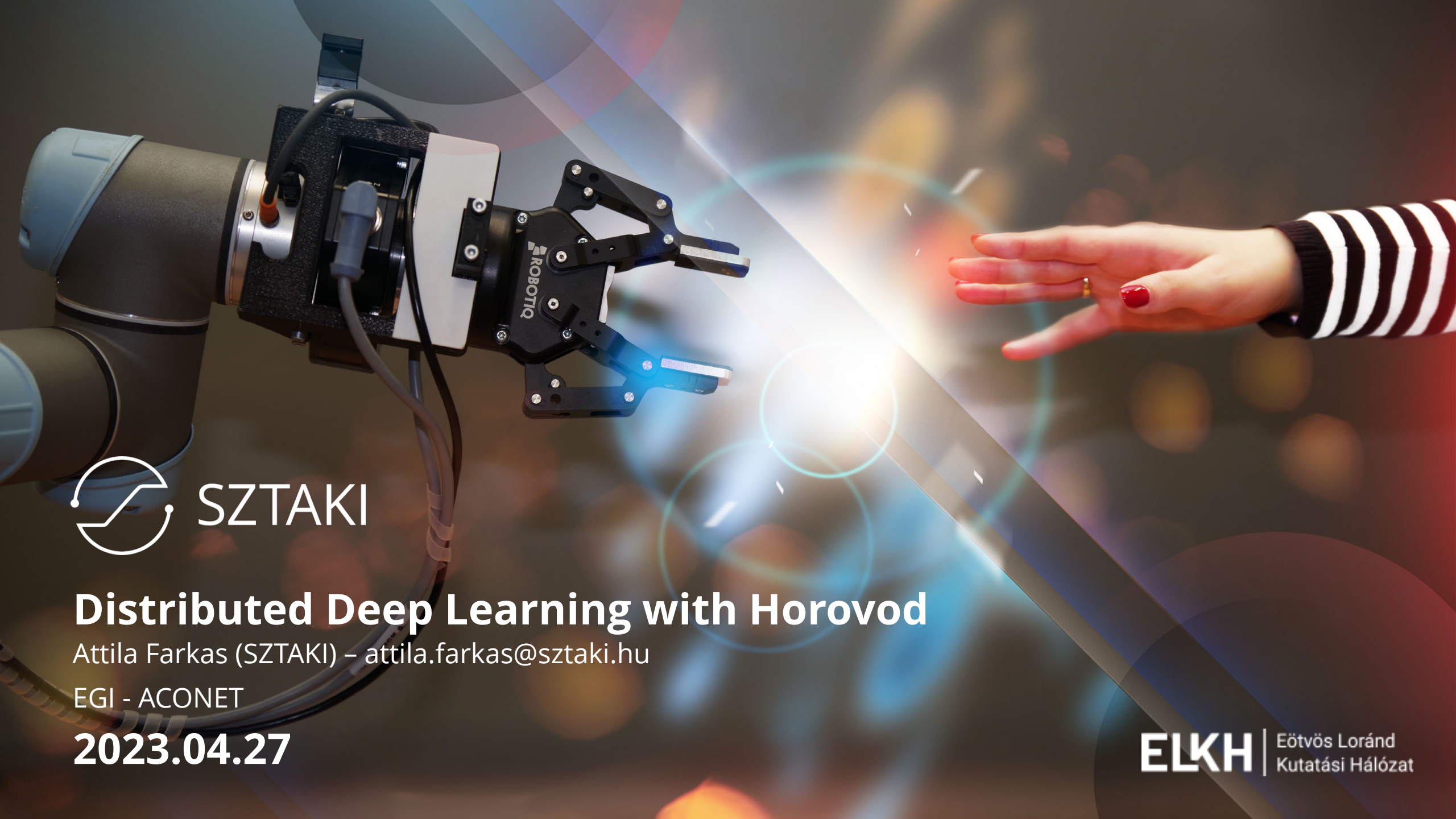


SZTAKI

Distributed Deep Learning with Horovod

Attila Farkas (SZTAKI) – attila.farkas@sztaki.hu

EGI - ACONET

2023.04.27

ELKH | Eötvös Loránd Kutatási Hálózat

Agenda

- **ELKH Cloud**
  - Introduction
  - Reference Architecture Concept
- **Distributed Deep Learning**
  - Motivation
  - Methodology
  - Uber’s Journey
  - Introduction to Horovod
- **Horovod Reference Architecture**
  - Introduction
  - Monitoring
  - Performance Evaluation
  - EOSC Service

ELKH Cloud

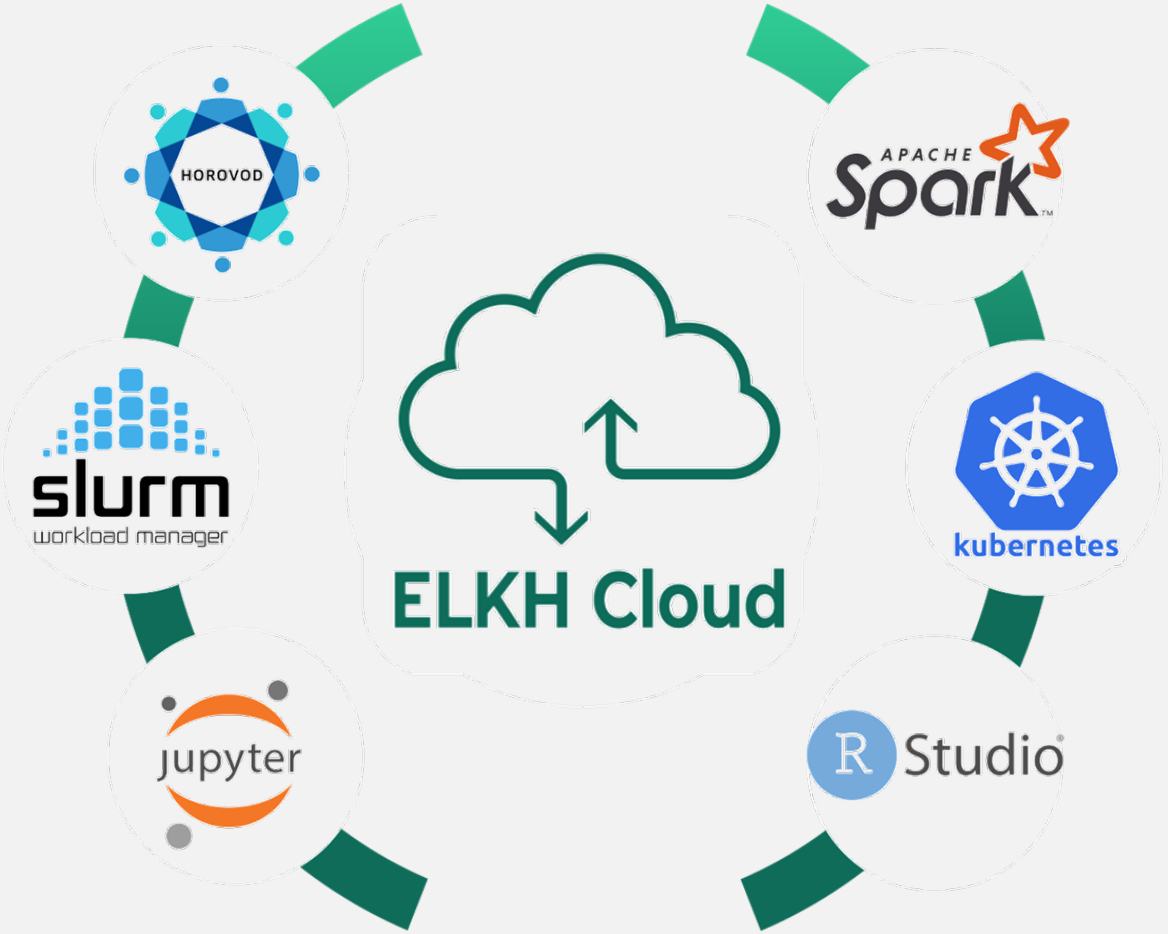
- OpenStack based federated cloud in Hungary
- Supporting the Hungarian scientific community with large scale computing capabilities
- Cloud services for more than 150 research projects
- Extensive support for AI-related research
  - On both hardware and software level
- Part of the European infrastructure landscape
  - EGI-ACE, SLICES-SC
- Pay-for-use provider in AI4PublicPolicy

science-cloud.hu/en

SZTAKI ELKH Eötvös Loránd Research Network

ELKH Cloud Reference Architectures

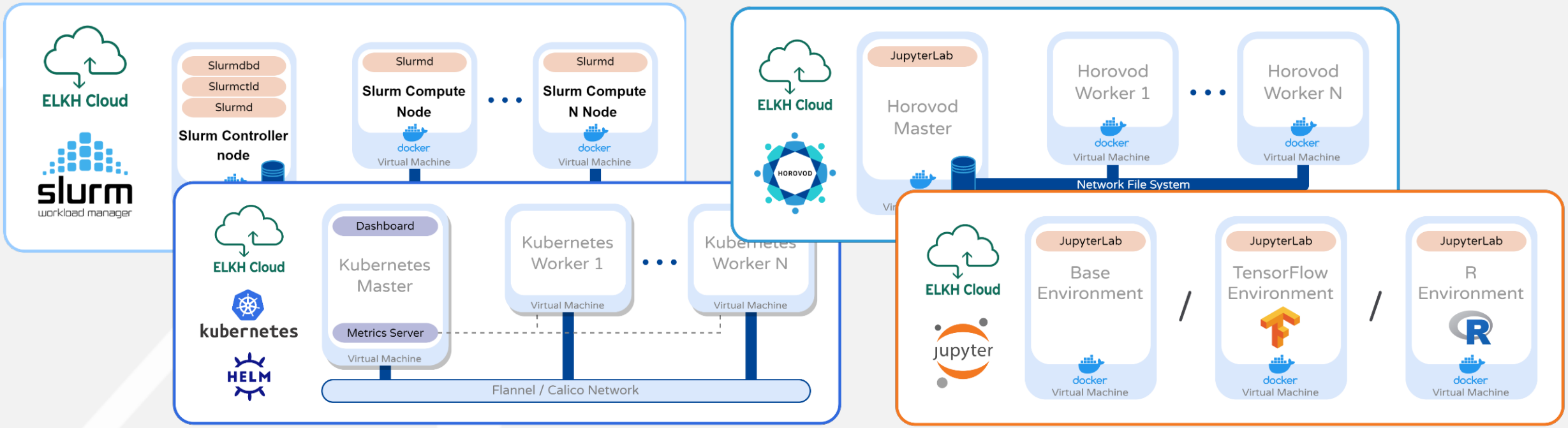
- Quickly and easily deployable digital research infrastructures
- Cover a variety of common use-cases
- Open-source and up-to-date
- Infrastructure as Code
- Based on widespread tools
  - Terraform
  - Ansible
  - Docker
- Long-term support



SZTAKI ELKH Eötvös Loránd Research Network

ELKH Cloud Reference Architectures

https://science-cloud.hu/en/reference-architectures



SZTAKI ELKH Eötvös Loránd Research Network

ELKH Cloud Reference Architectures - Deployment

HashiCorp Terraform

- **Provision infrastructure**
  - Virtual machines
  - Network settings
  - Firewall rules
  - Execute tasks
  - Invoke Ansible

ANSIBLE

- **Configure nodes**
  - Install packages
  - Start services
  - Run Docker containers

I. Set authentication information
II. Set resource parameters
III. Customize through variables
IV. Deploy

```
horovod_master_node = ({
    name = "horovod-master"
    flavor_name = "SET_YOUR_FLAVOR_NAME"
    image_id = "SET_YOUR_IMAGE_ID"
    key_pair = "SET_YOUR_KEY_PAIR_NAME"
    floating_ip = "SET_YOUR_
    volume_size = 32
})
```

```
user_config = ({
    jupyter_password = "elkhcloud"
    enable_gpu = true
    monitoring = false
    monitoring_password = "elkhcloud"
    nvidia_driver_install = false
    elkh_cloud_dedicated_network = false
})
```

ELKH Cloud Reference Architectures - Automatic Testing

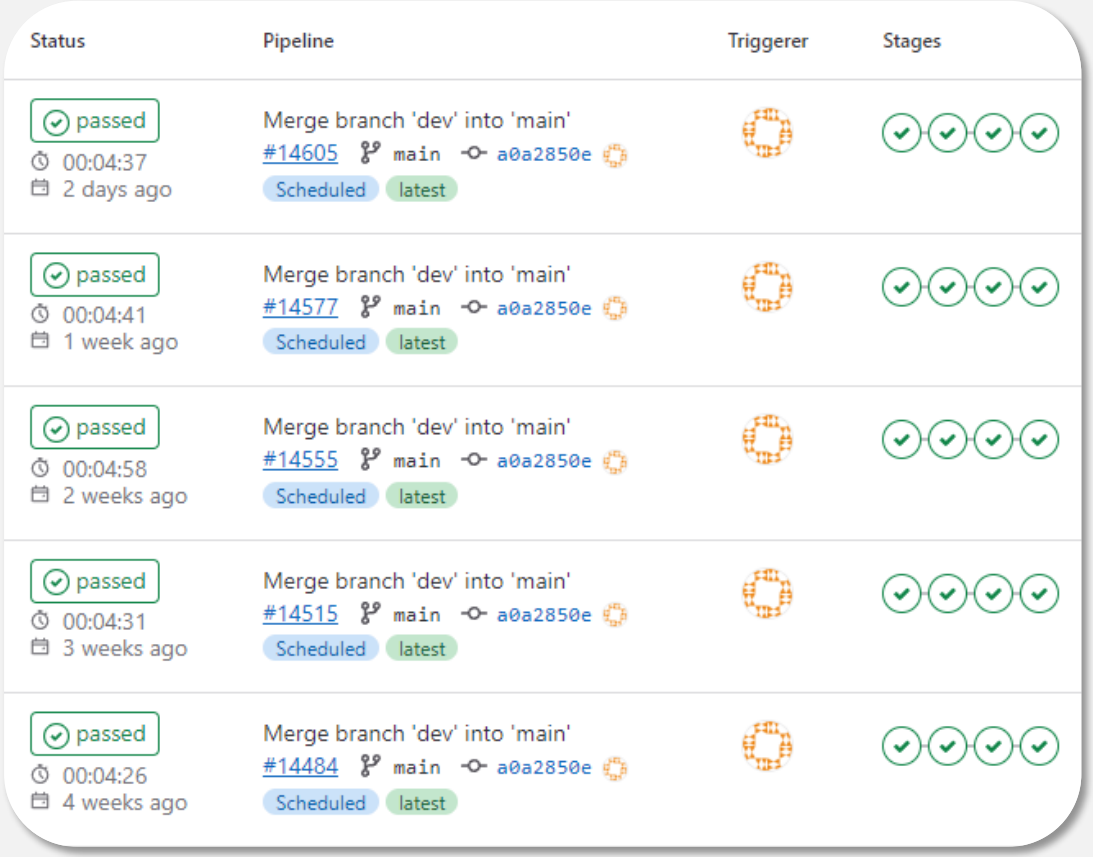
- Increased reliability
- Dedicated project in ELKH Cloud
- Weekly scheduled pipelines
- Email notifications
- Modular functional tests
- Easily expandable

Current

Automatic testing of reference architectures (SZTAKI)

Institution
Institute for Computer Science and Control

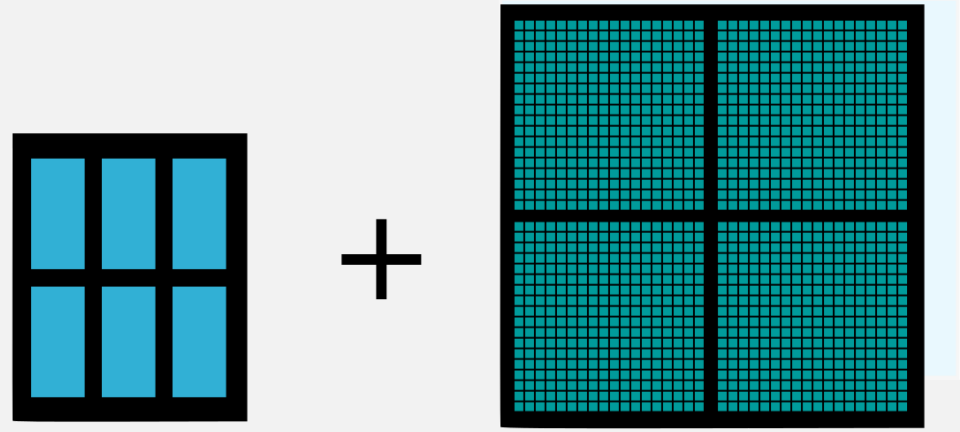
Project leader name
Emődi Márk Benjámin



SZTAKI ELKH Eötvös Loránd Research Network

Distributed Deep Learning - Motivation

- Deep Learning
  - Neural networks with a large number of layers
  - Significant increase in trainable parameters
  - Better suited for solving complex tasks
- Highly resource intensive
  - Large volume of training data
  - Prolonged training time
  - A limiting factor for a long time
- GPU acceleration
  - Simple operations
  - Massively parallel
  - Multiple cards can be used

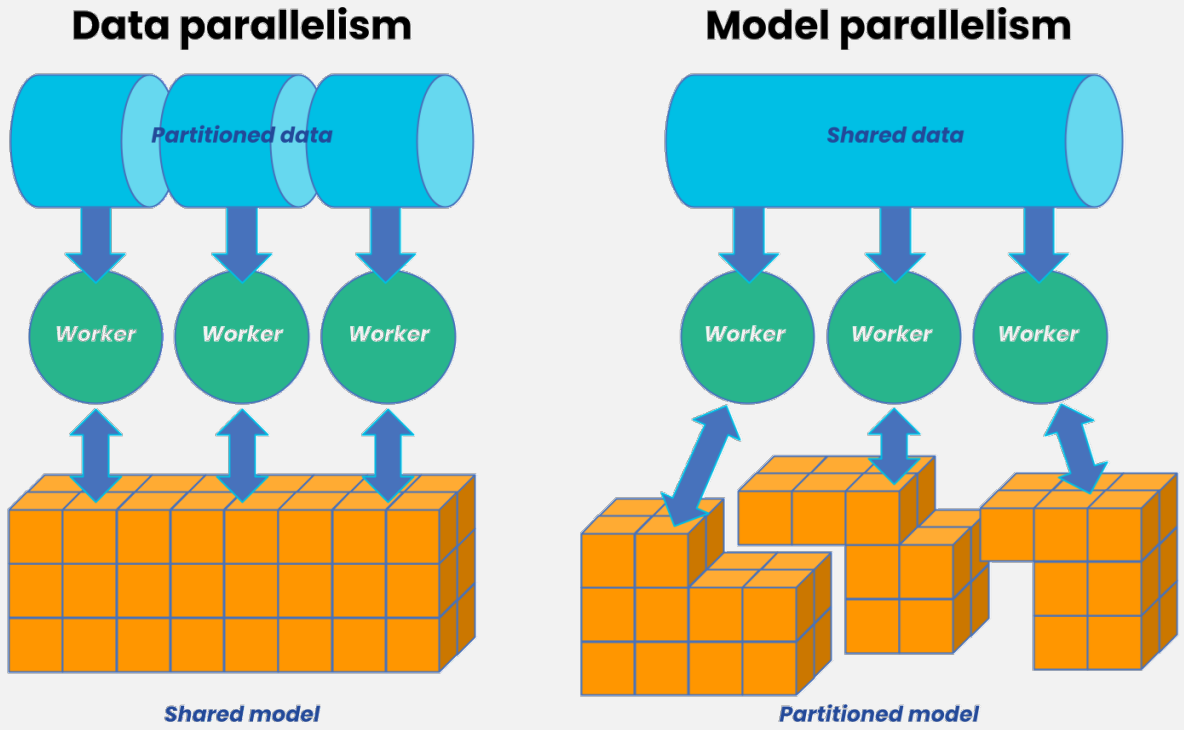


https://www.cherryservers.com/blog/gpu-vs-cpu-what-are-the-key-differences

SZTAKI ELKH Eötvös Loránd Research Network

Distributed Deep Learning - Methodology

- Vertical scaling
  - Increasing resources of a single node
  - Physical limitations
- Horizontal scaling
  - Connecting multiple compute nodes
  - Bandwidth is often a bottleneck
- Data Parallelism
  - Copy the model to each node
  - Compute on a different subset of data
  - Synchronize gradients
- Model Parallelism
  - Train different layers on different nodes
  - Applied due to memory limitations

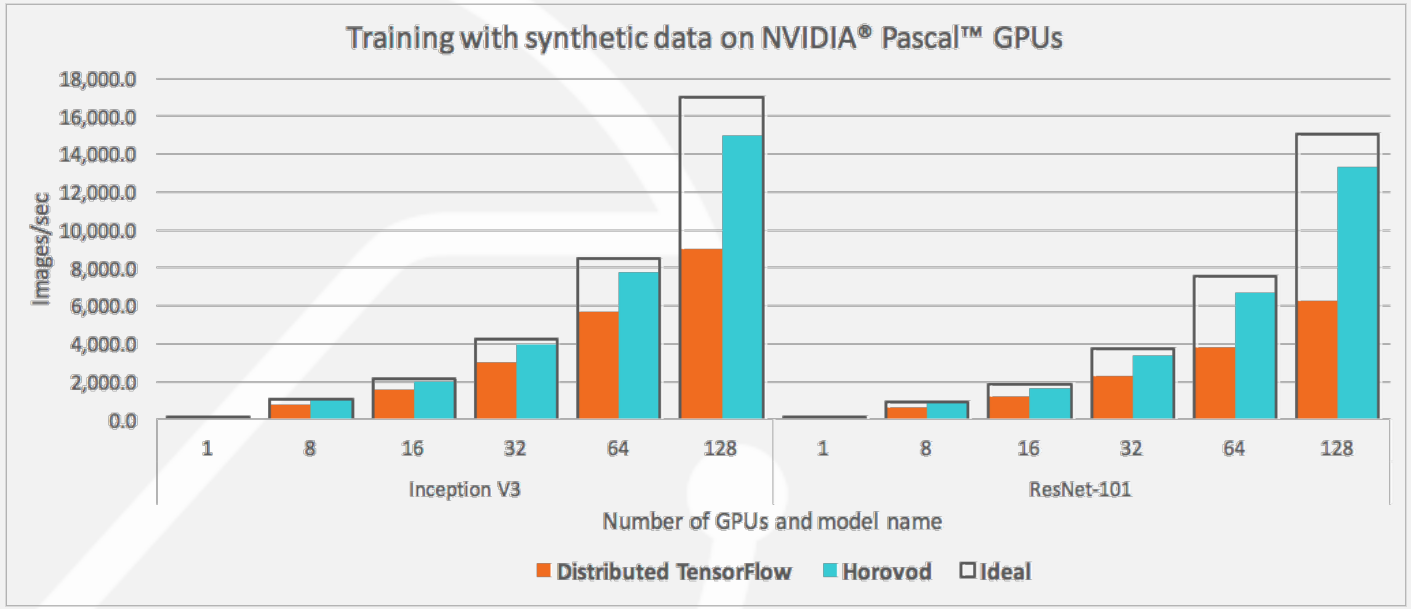


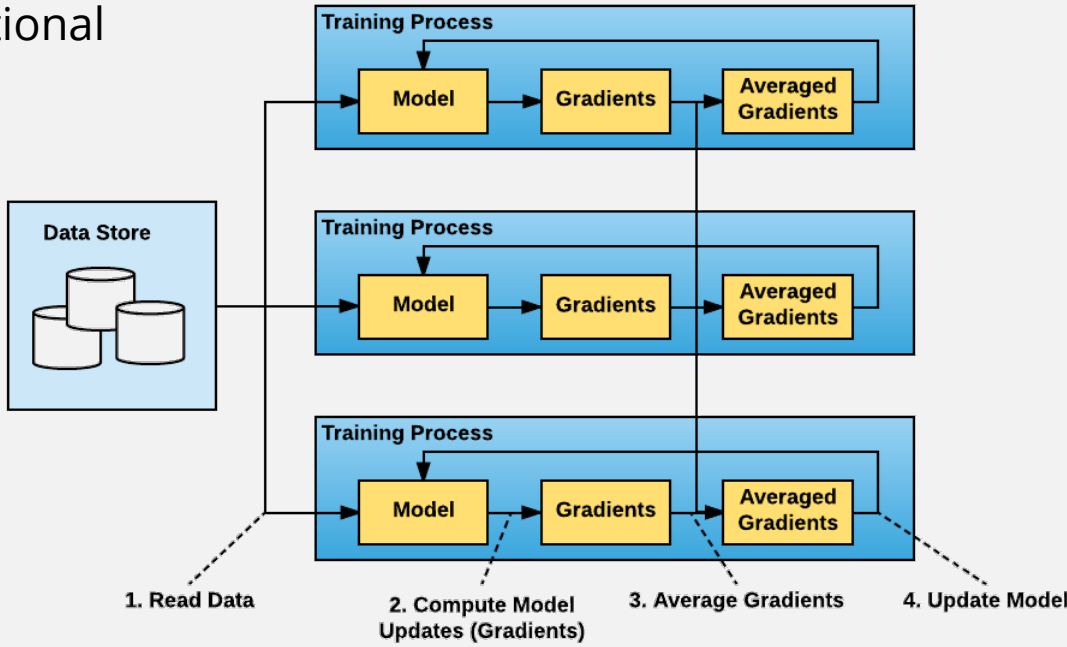
https://www.anyscale.com/blog/what-is-distributed-training

SZTAKI ELKH Eötvös Loránd Research Network

Introduction to Horovod

- Open-source distributed deep learning framework by Uber
- Supports TensorFlow, Keras, PyTorch, Apache MXNet and Spark
- Provide an easy-to-use framework for distributed training
  - Execute on hundreds of GPUs with just a few lines of additional code
- Data parallel execution



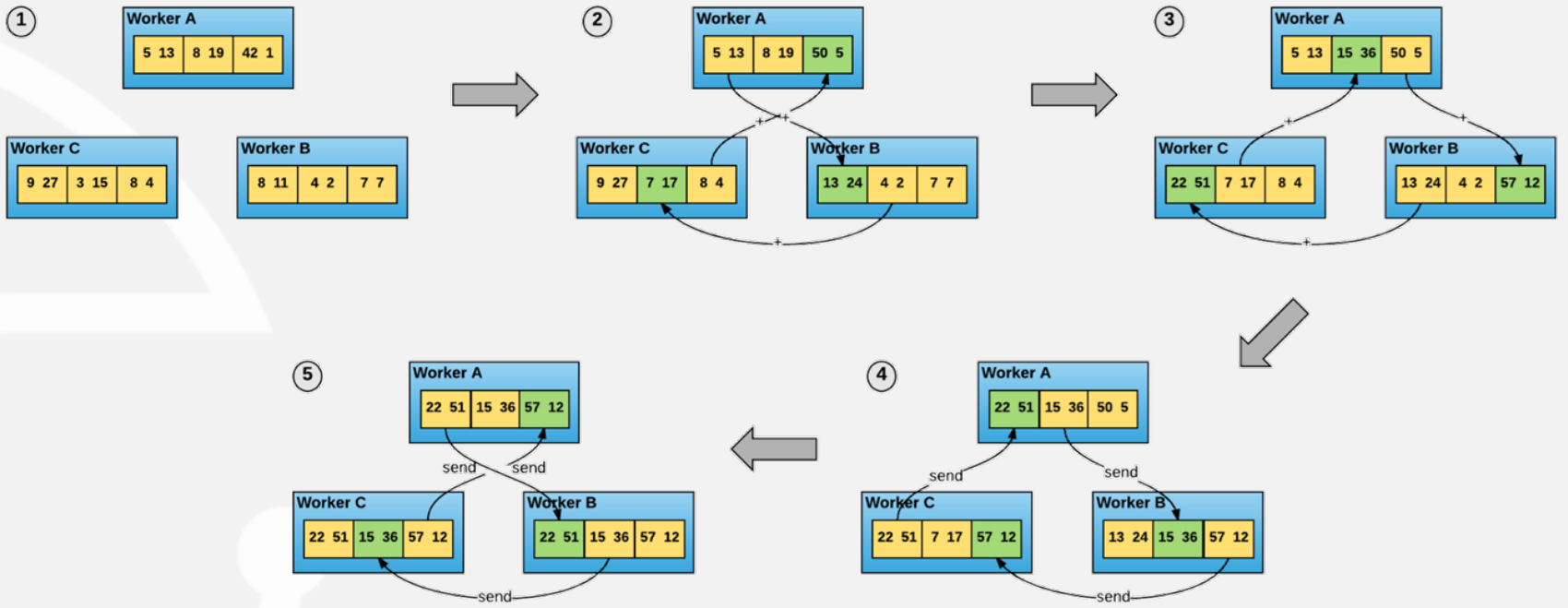


Sergeev, A., & Balso, M.D. (2018). Horovod: fast and easy distributed deep learning in TensorFlow. ArXiv, abs/1802.05799.

SZTAKI ELKH Eötvös Loránd Research Network

Introduction to Horovod

- Ring-Allreduce strategy
  - Each node communicates with two of its peers 2*(N-1) times
  - NVIDIA NCCL 2.0 for intra-node communication
  - Bandwidth optimal

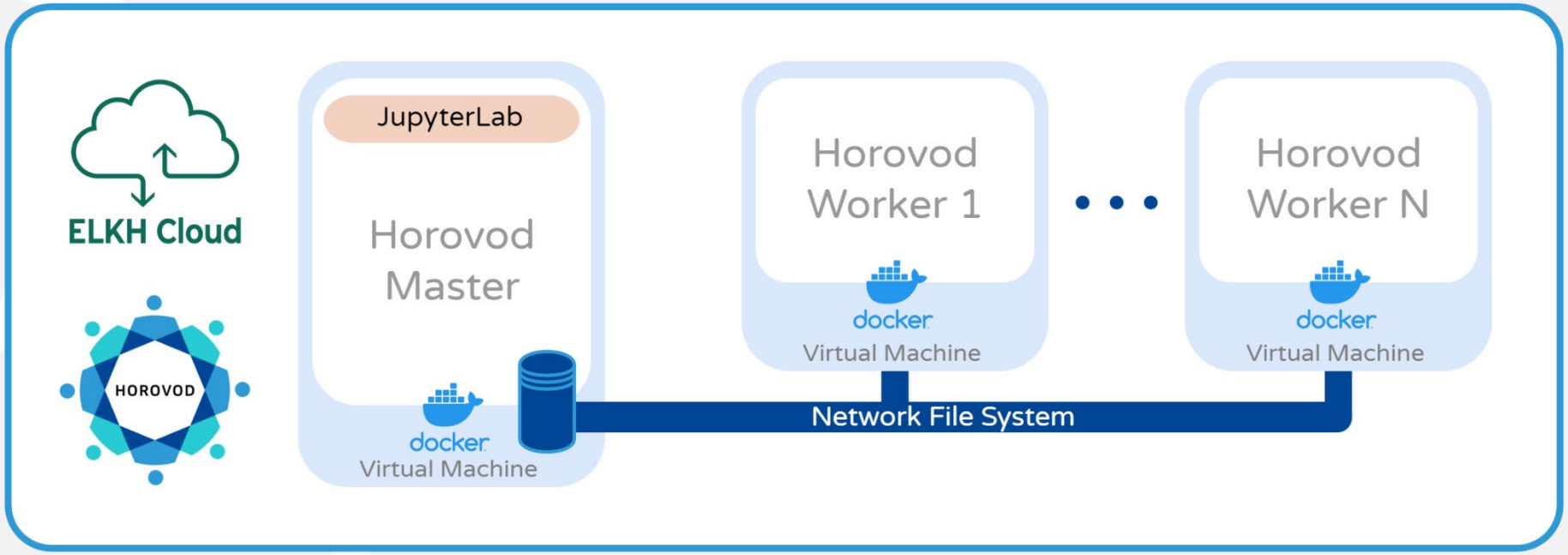


Sergeev, A., & Balso, M.D. (2018). Horovod: fast and easy distributed deep learning in TensorFlow. ArXiv, abs/1802.05799.

SZTAKI ELKH Eötvös Loránd Research Network

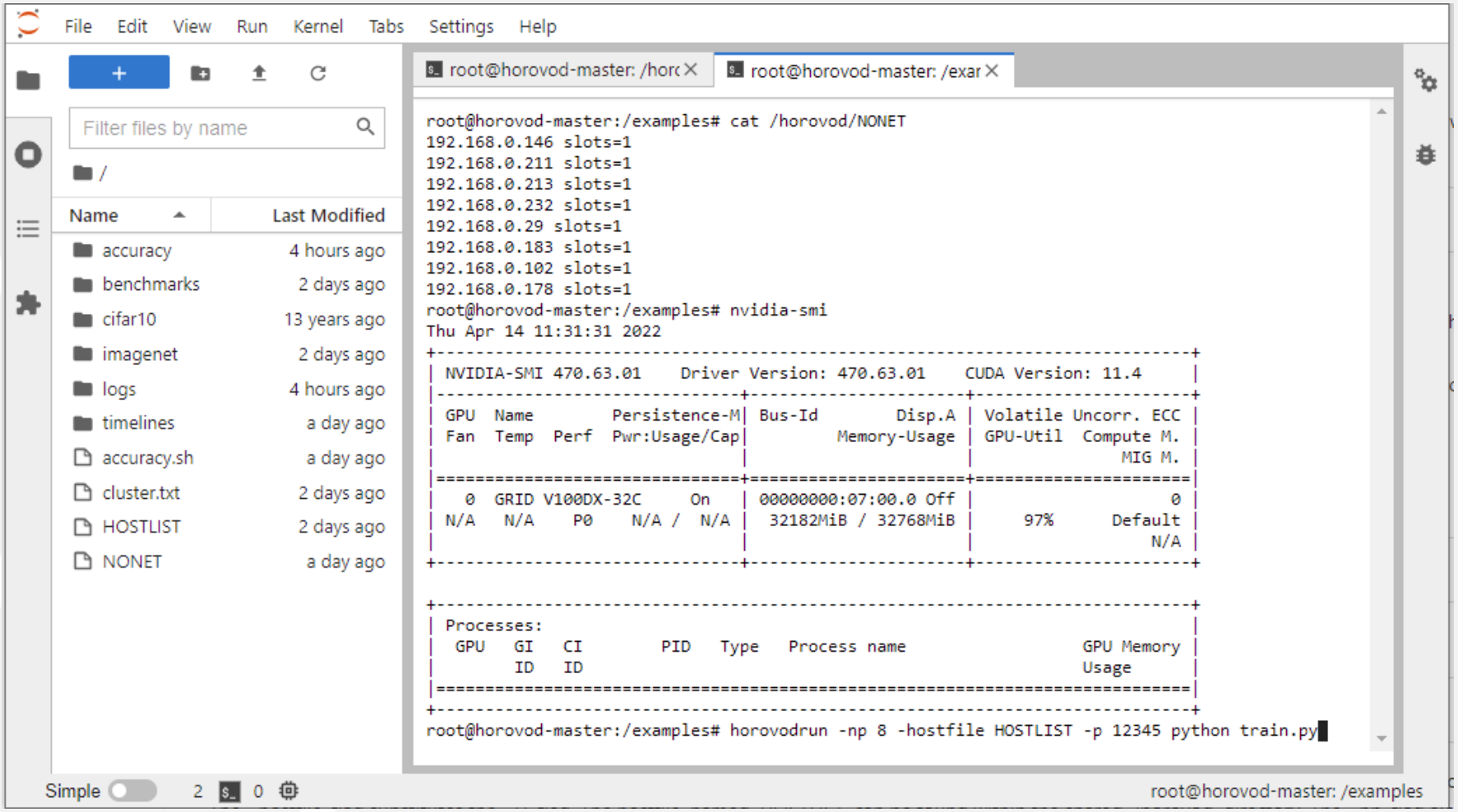
Horovod Reference Architecture

- Support for distributed deep learning applications
- JupyterLab development environment
- Network based file sharing between nodes
- Utilization of GPU resources



SZTAKI ELKH Eötvös Loránd Research Network

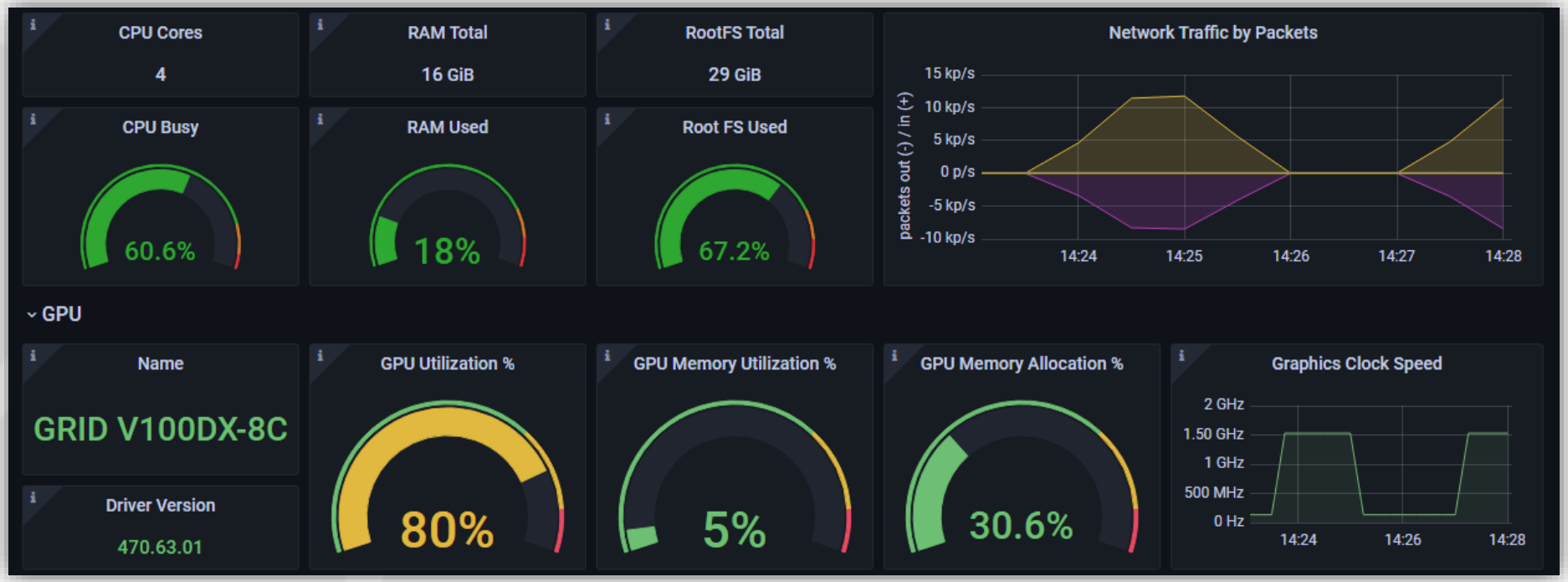
Horovod Reference Architecture – Usage



SZTAKI ELKH Eötvös Loránd Research Network

Horovod Reference Architecture - Monitoring

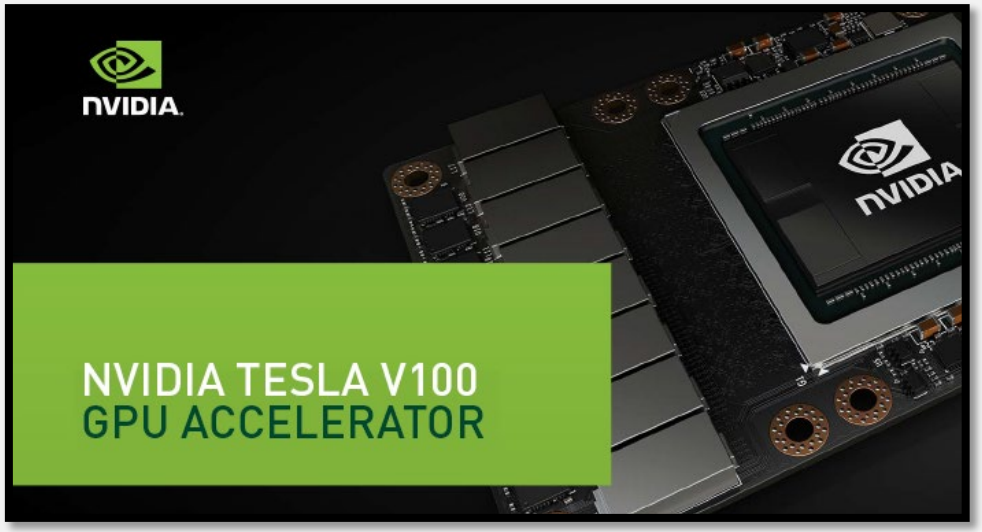
- Optional
- Containerized Prometheus & Grafana stack
- Covers metrics from GPU Utilization to Network Traffic



SZTAKI ELKH Eötvös Loránd Research Network

Horovod Performance Evaluation - Environment

- Executed on ELKH Cloud
- 1 to 8 Virtual machines, each containing:
  - 16 vCPU
  - 64GB RAM
  - NVIDIA Tesla V100 GPU (32GB VRAM)
  - 10Gbps network connection
- 1TB SSD-based storage, shared using NFS
- Official TensorFlow benchmark scripts
- ResNet models

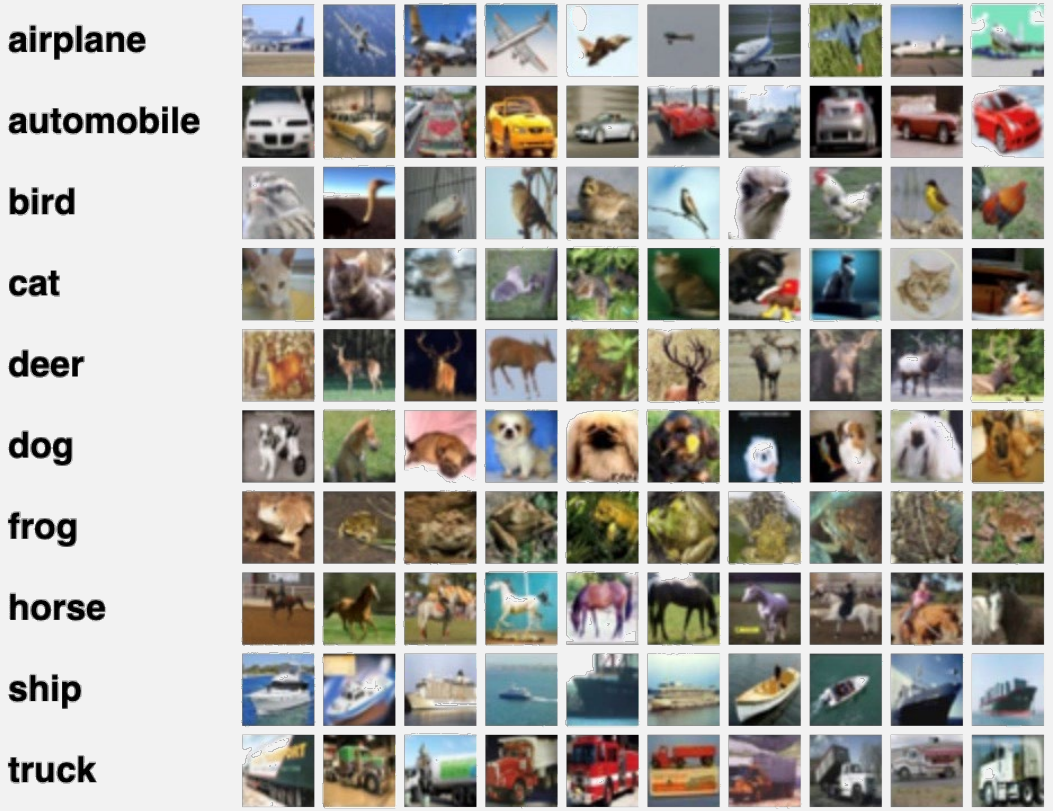


https://www.nvidia.com/en-us/data-center/v100/

SZTAKI ELKH Eötvös Loránd Research Network

Horovod Performance Evaluation - Parameters

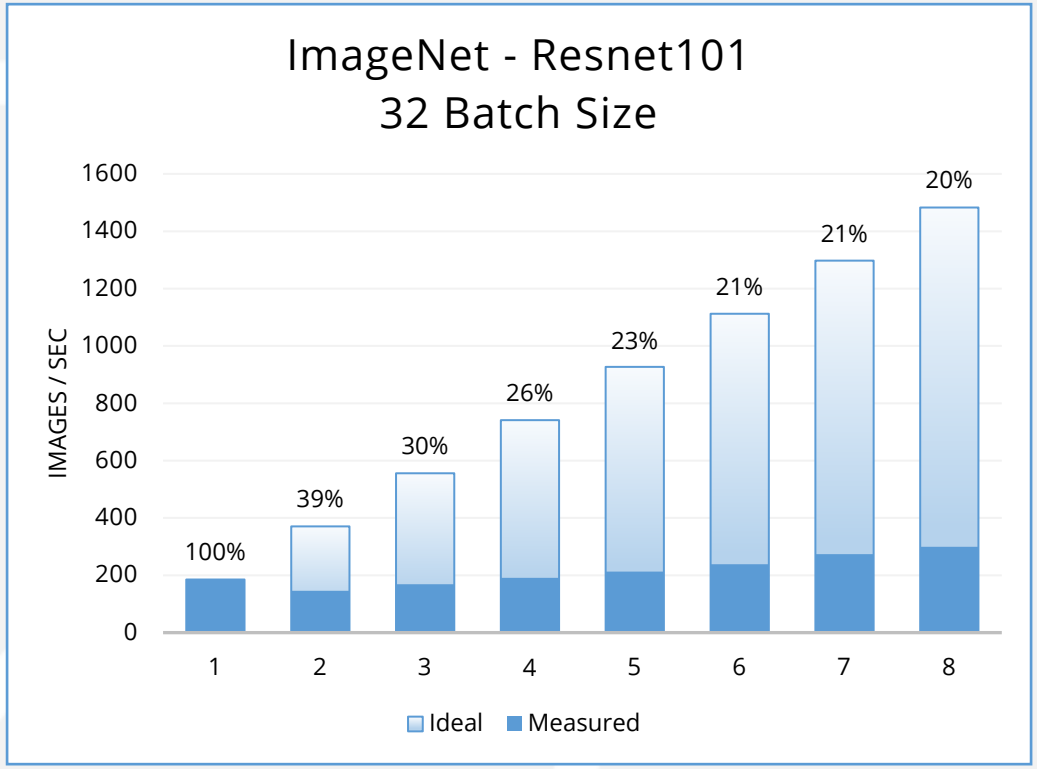
- Number of nodes (1-8)
- Batch size
  - 32 – default
  - 196 – maximum (physical limits)
- Training datasets
  - CIFAR-10 (178 MB)
  - ImageNet (141 GB)
- Measured parameters
  - Processing performance
  - Scaling efficiency
  - Model accuracy

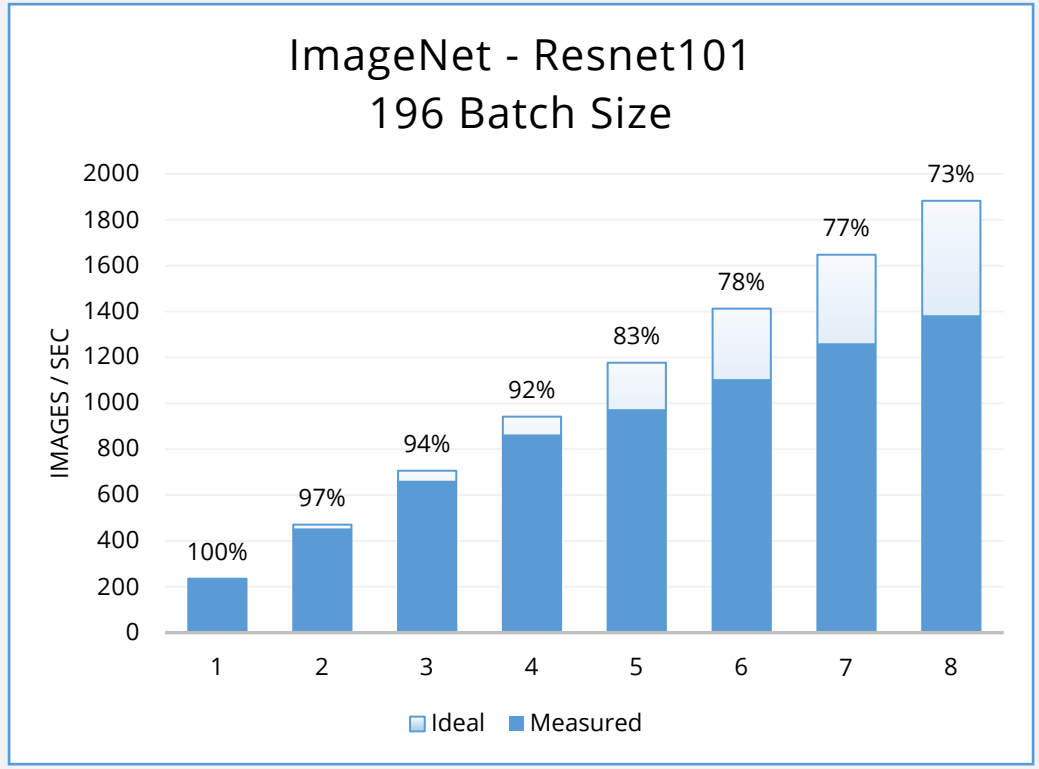


SZTAKI ELKH Eötvös Loránd Research Network

Horovod Performance Evaluation - Results

- Processing performance measured in images per second
- Communication frequency
- Optimization of training parameters

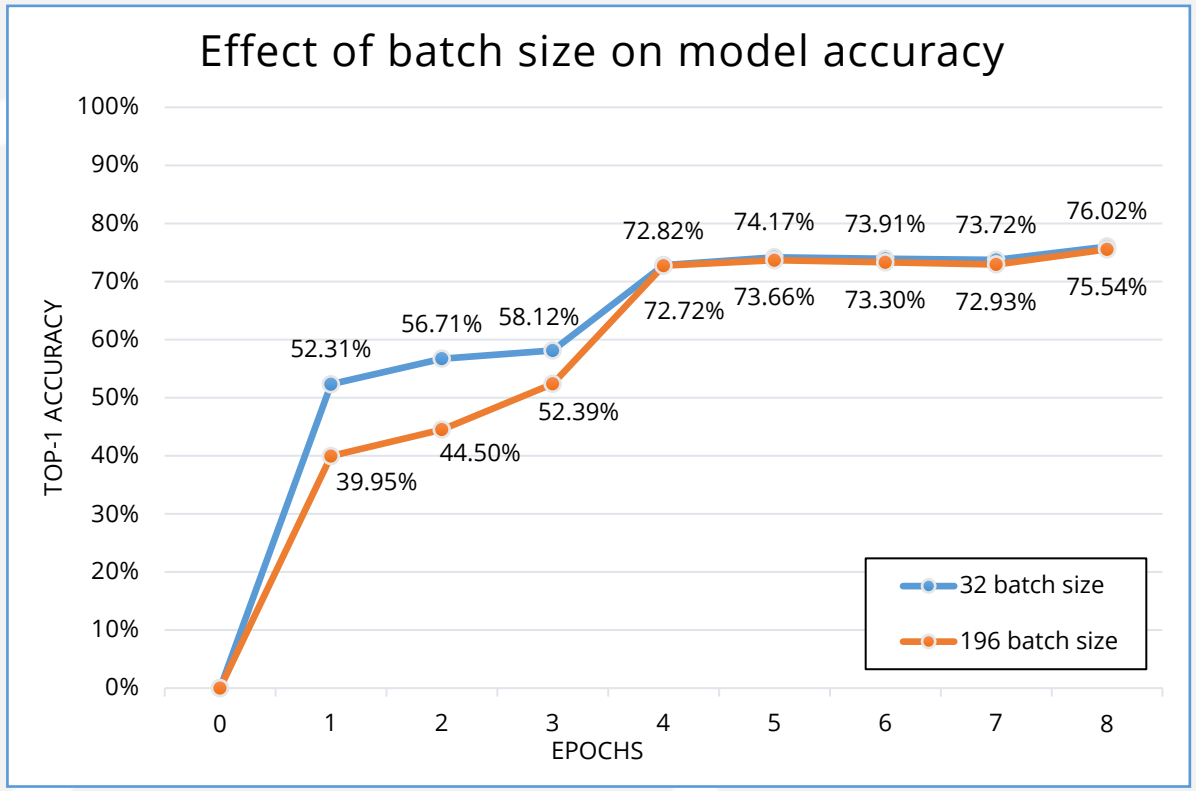


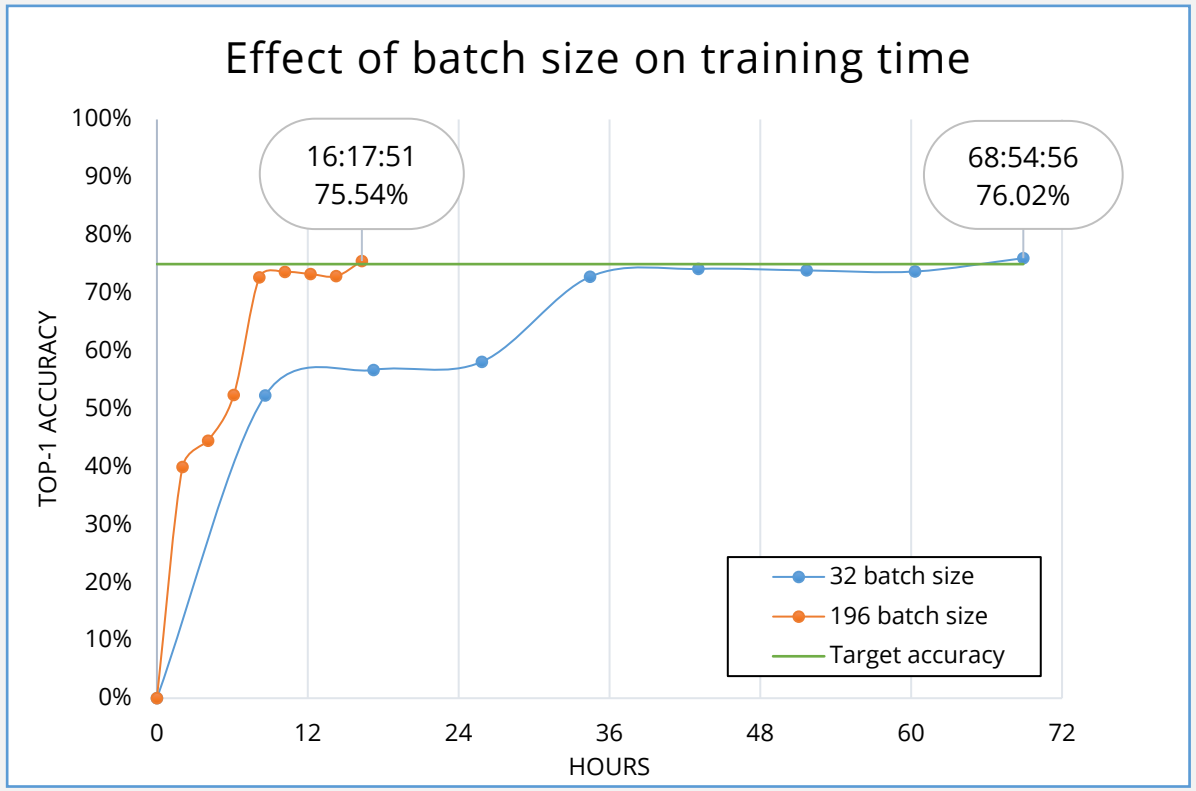


SZTAKI ELKH Eötvös Loránd Research Network

Horovod Performance Evaluation - Results

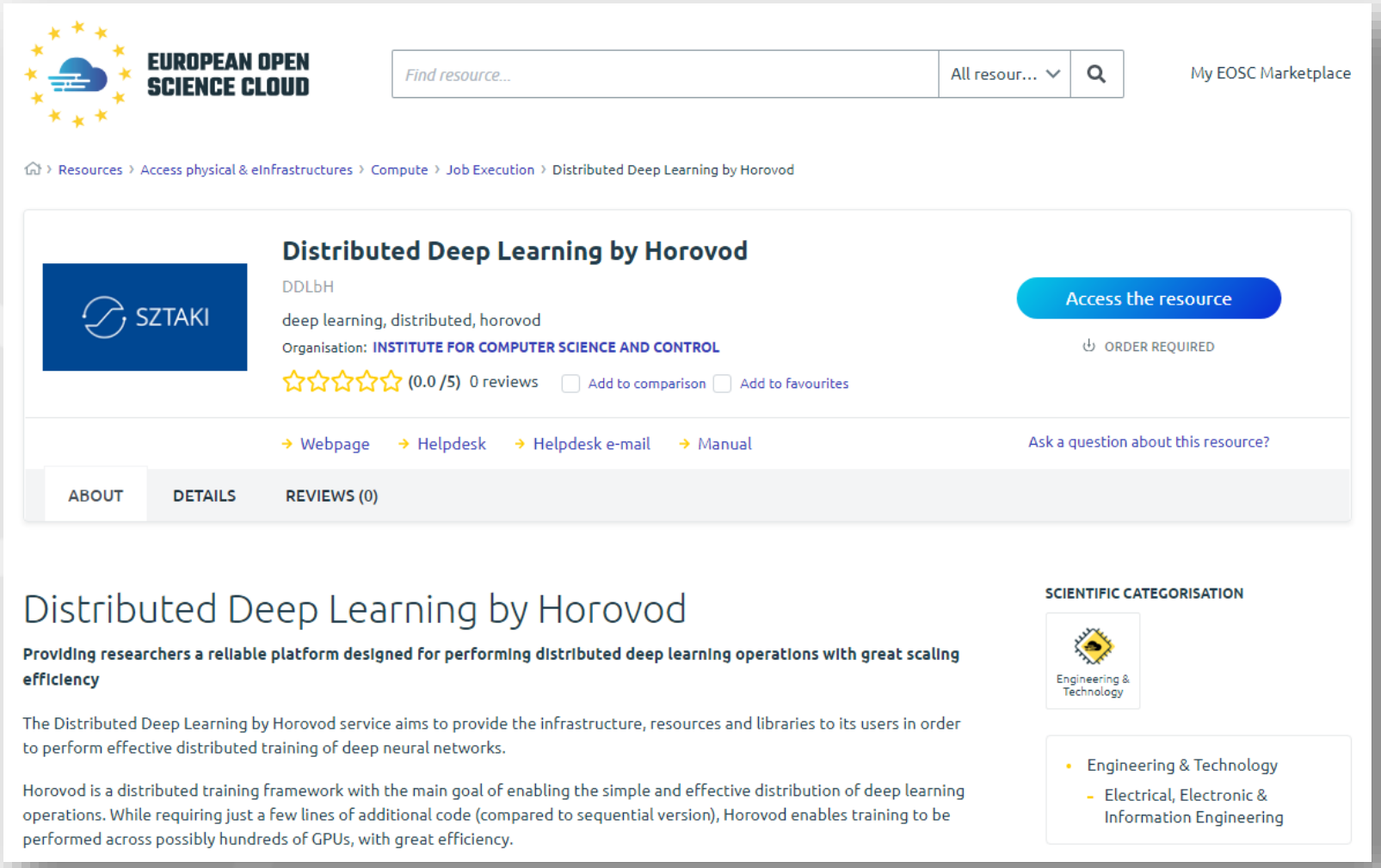
- 8 node cluster
- Training a ResNet101 model on ImageNet to up to 75% top-1 accuracy
- Evaluation after each epoch





SZTAKI ELKH Eötvös Loránd Research Network

Onboarding Horovod to the EOSC Marketplace



SZTAKI ELKH Eötvös Loránd Research Network

EOSC Horovod Service - Access Modes

Limited-time demo

- Gain short term access to a demo cluster
- Hosted on ELKH Cloud / EGI-ACE resources (CESNET)
- 4 GPU enabled nodes

ELKH Cloud

Request deployment on EOSC cloud resources

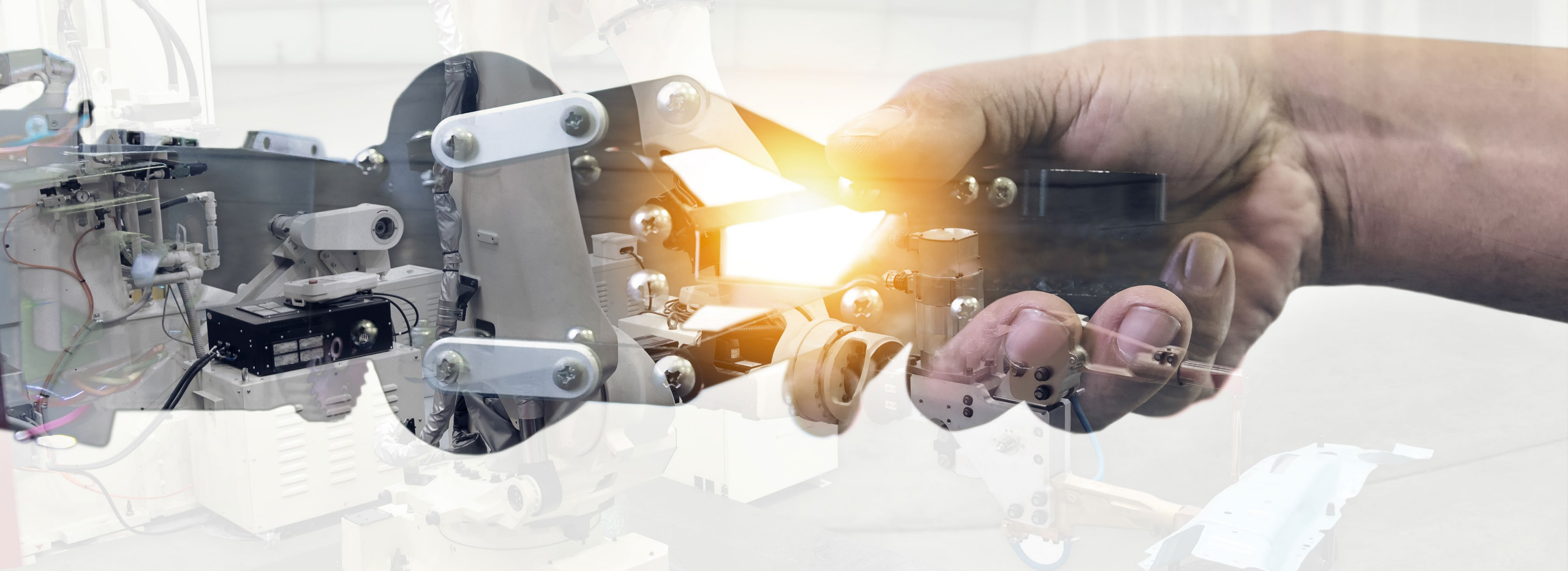
- Hosted on an OpenStack based cloud* provided by the user
- Long term access
- Exact period length and node count is up to negotiation

EUROPEAN OPEN SCIENCE CLOUD

Self-hosted

- User manual
- Technical consultation

*: Additional restrictions apply, exact requirements are listed in the service manual



THANK YOU FOR YOUR ATTENTION!

www.sztaki.hu

SZTAKI ELKH Eötvös Loránd Research Network