

SZTAKI

## Distributed Deep Learning with Horovod

Attila Farkas (SZTAKI) – [attila.farkas@sztaki.hu](mailto: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

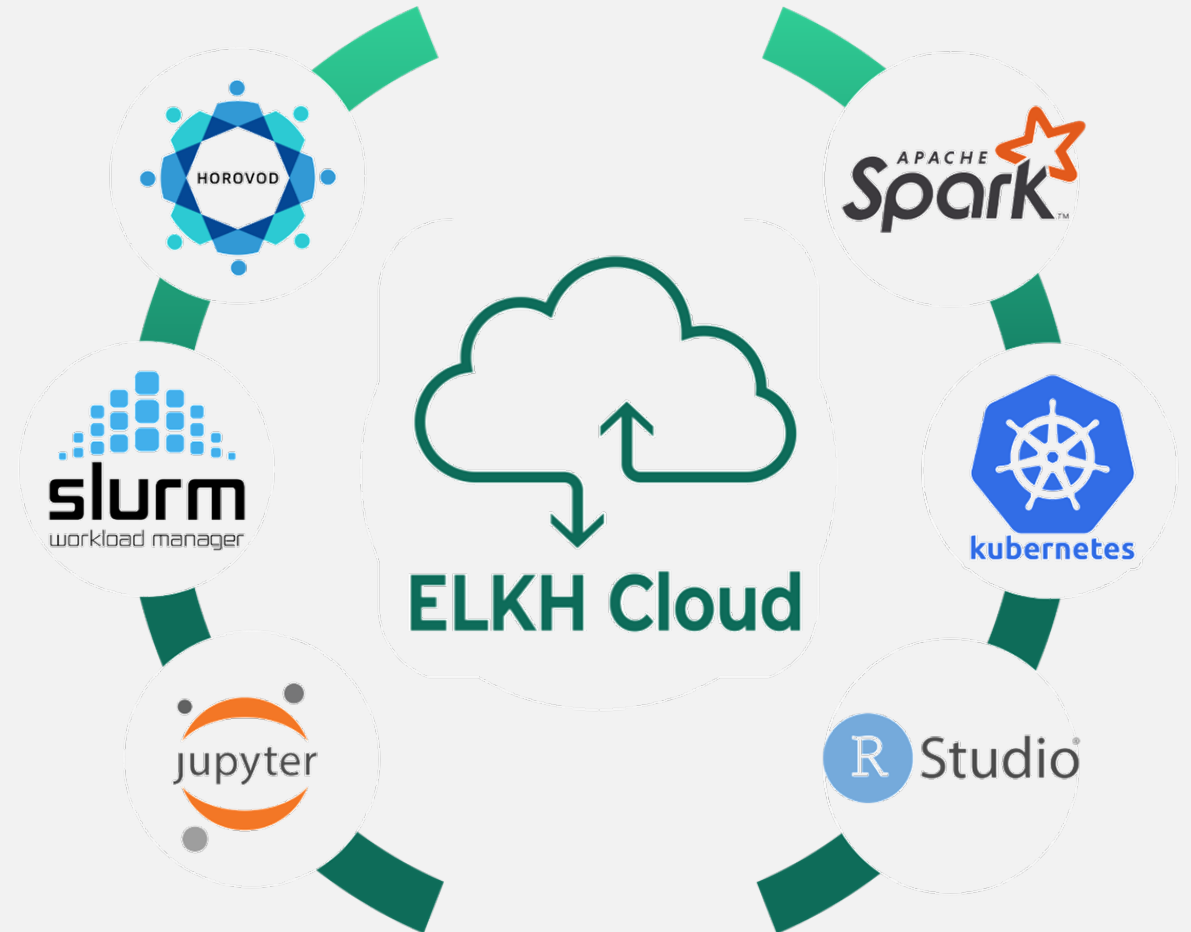
- OpenStack based federated cloud in Hungary
- Supporting the Hungarian scientific community with large scale computing capabilities
- Cloud services for more than 150 research projects
- Extensive support for AI-related research
  - On both hardware and software level
- Part of the European infrastructure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



[science-cloud.hu/en](https://science-cloud.hu/en)

# ELKH Cloud Reference Architectures

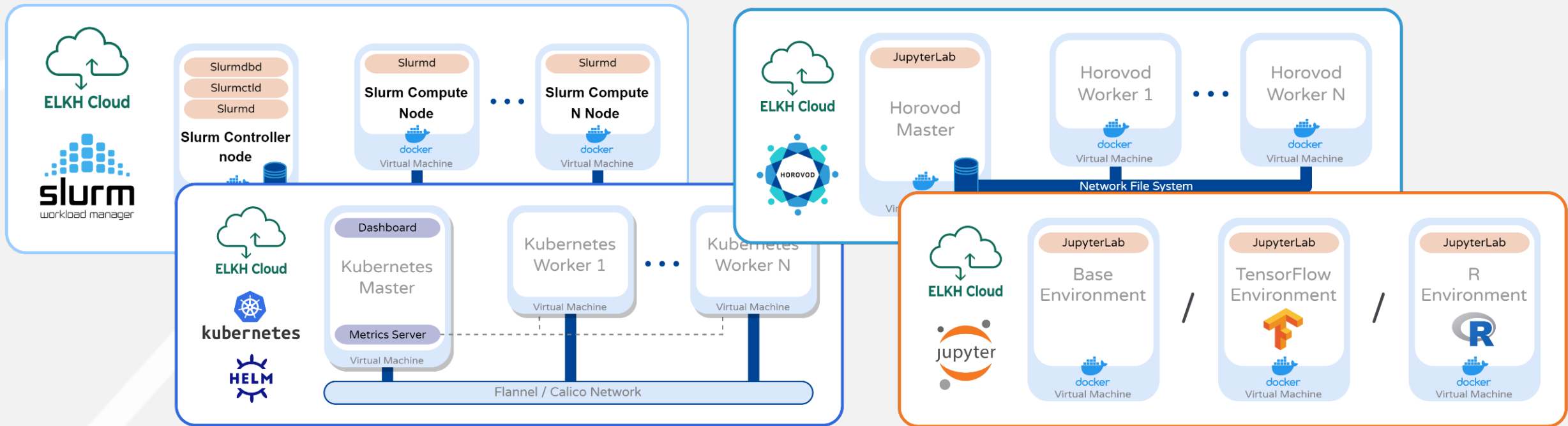
- Quickly and easily deployable digital research infrastructures
- Cover a variety of common use-cases
- Open-source and up-to-date
- Infrastructure as Code
- Based on widespread tools
  - Terraform
  - Ansible
  - Docker
- Long-term support





# ELKH Cloud Reference Architectures

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



# 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

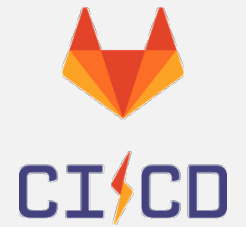
- I. Set authentication information
- II. Set resource parameters
- III. Customize through variables
- IV. Deploy

```
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

Status	Pipeline	Triggerer	Stages
<span>passed</span> ⌚ 00:04:37 📅 2 days ago	Merge branch 'dev' into 'main' <a href="#">#14605</a> main  a0a2850e Scheduled latest		<span>✓</span> <span>✓</span> <span>✓</span> <span>✓</span>
<span>passed</span> ⌚ 00:04:41 📅 1 week ago	Merge branch 'dev' into 'main' <a href="#">#14577</a> main  a0a2850e Scheduled latest		<span>✓</span> <span>✓</span> <span>✓</span> <span>✓</span>
<span>passed</span> ⌚ 00:04:58 📅 2 weeks ago	Merge branch 'dev' into 'main' <a href="#">#14555</a> main  a0a2850e Scheduled latest		<span>✓</span> <span>✓</span> <span>✓</span> <span>✓</span>
<span>passed</span> ⌚ 00:04:31 📅 3 weeks ago	Merge branch 'dev' into 'main' <a href="#">#14515</a> main  a0a2850e Scheduled latest		<span>✓</span> <span>✓</span> <span>✓</span> <span>✓</span>
<span>passed</span> ⌚ 00:04:26 📅 4 weeks ago	Merge branch 'dev' into 'main' <a href="#">#14484</a> main  a0a2850e Scheduled latest		<span>✓</span> <span>✓</span> <span>✓</span> <span>✓</span>

Current

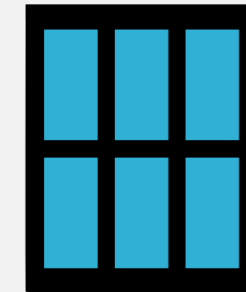
## Automatic testing of reference architectures (SZTAKI)

Institution  
Institute for Computer Science and Control

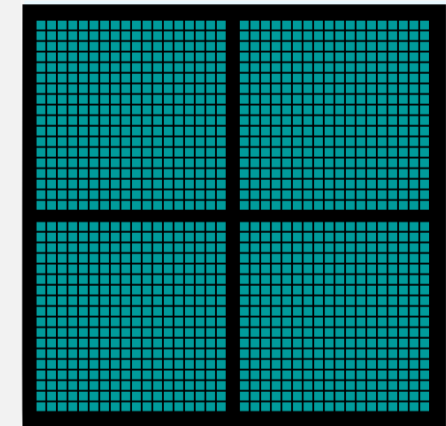
Project leader name  
Emödi Márk Benjámín

# 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



CPU  
Multiple Cores



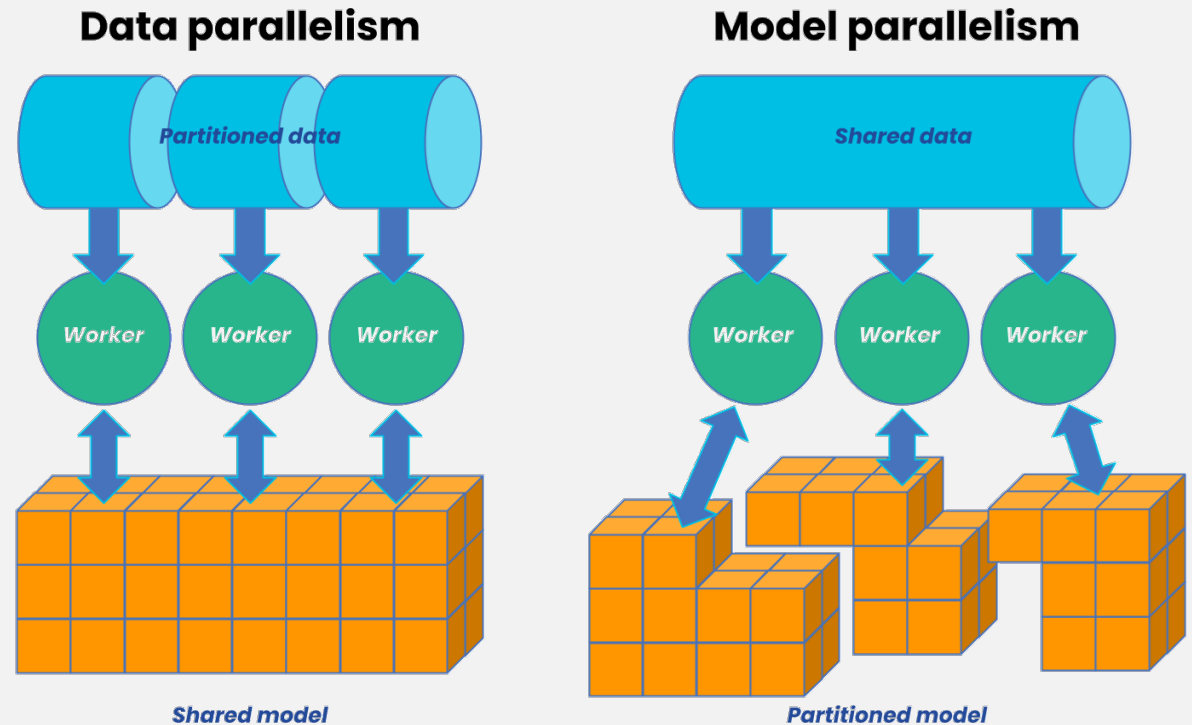
GPU  
Thousands of Cores

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



# Distributed Deep Learning - Methodology

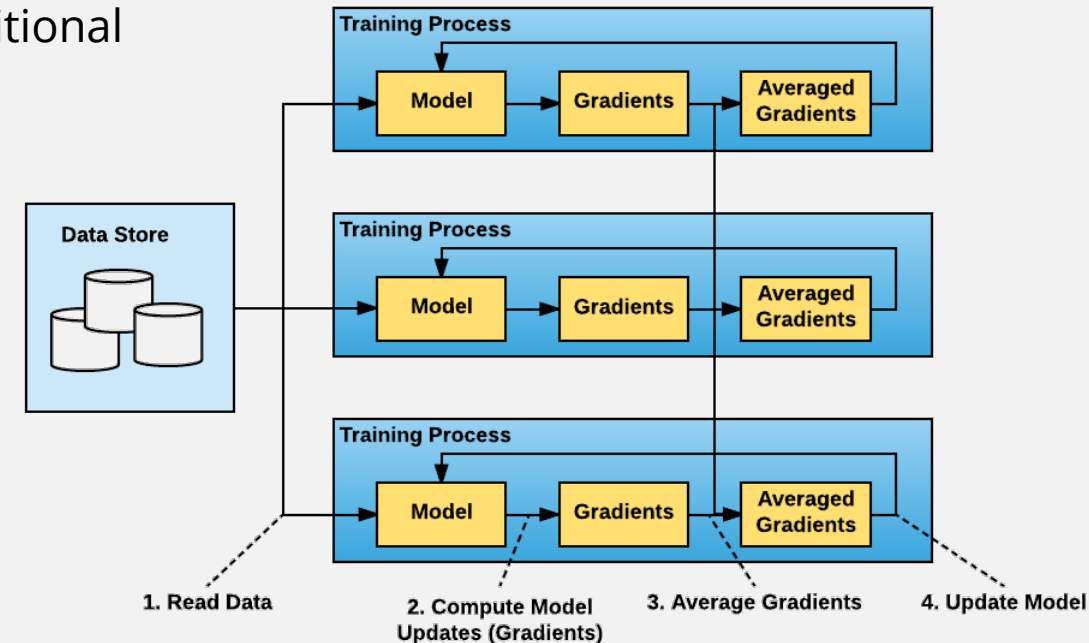
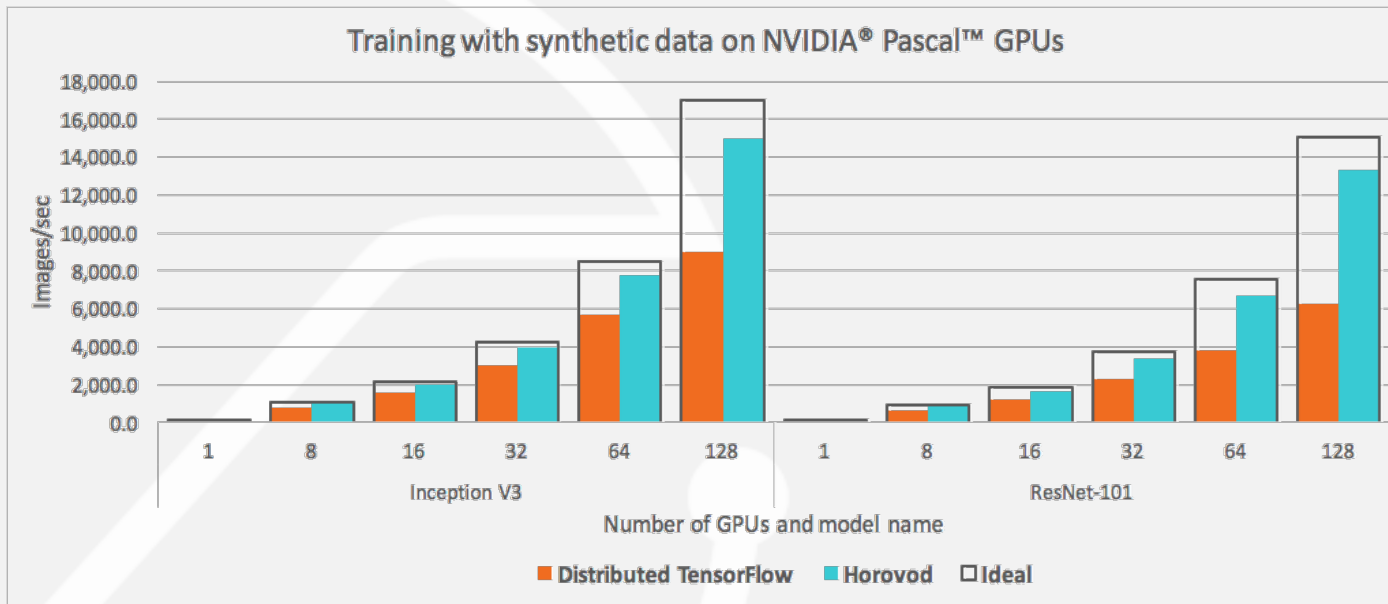
- Vertical scaling
  - Increasing resources of a single node
  - Physical limitations
- Horizontal scaling
  - Connecting multiple compute nodes
  - Bandwidth is often a bottleneck
- Data Parallelism
  - Copy the model to each node
  - Compute on a different subset of data
  - Synchronize gradients
-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



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

# 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

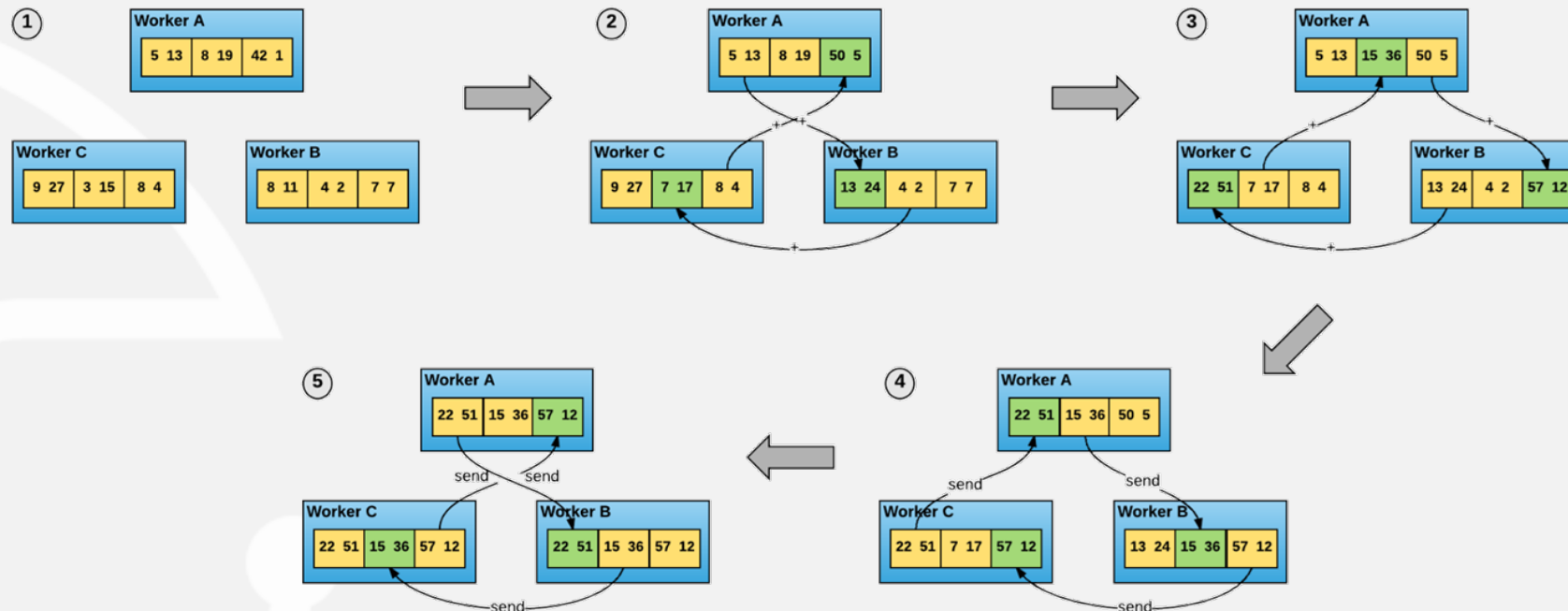


Sergeev, A., & Balso, M.D. (2018). Horovod: fast and easy distributed deep learning in TensorFlow. ArXiv, abs/1802.05799.

# Introduction to Horovod



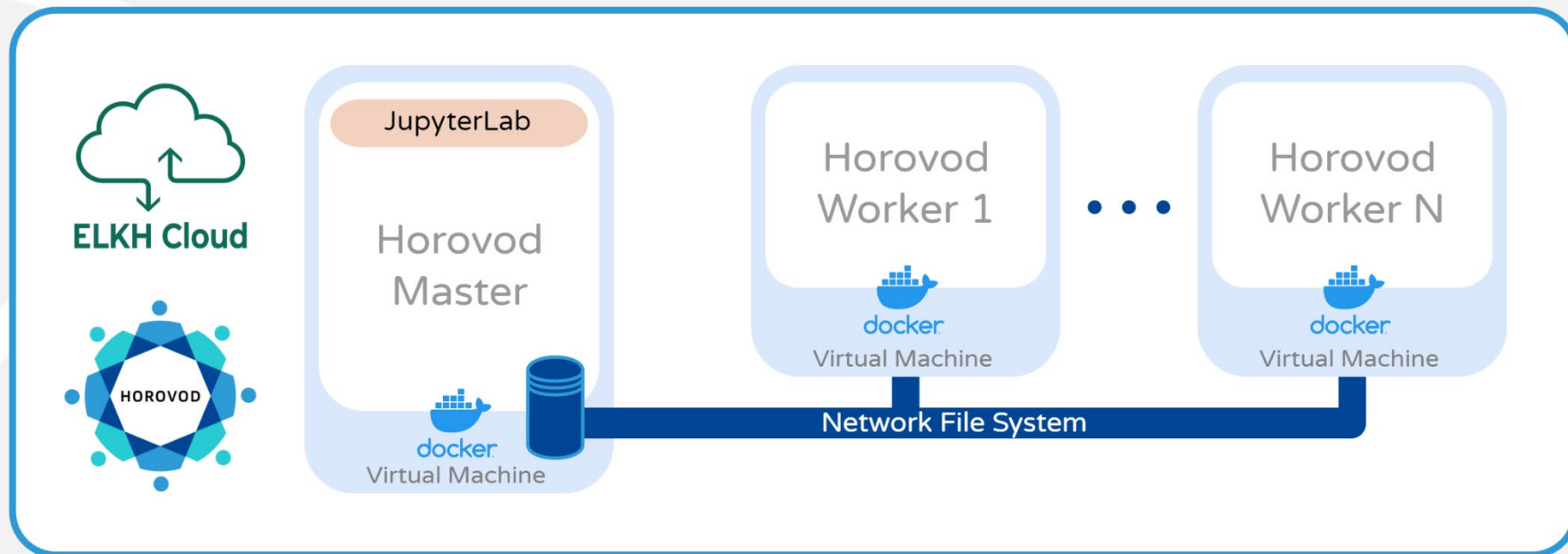
- Ring-Allreduce strategy
  - Each node communicates with two of its peers  $2*(N-1)$  times
  - NVIDIA NCCL 2.0 for intra-node communication
  - Bandwidth optimal



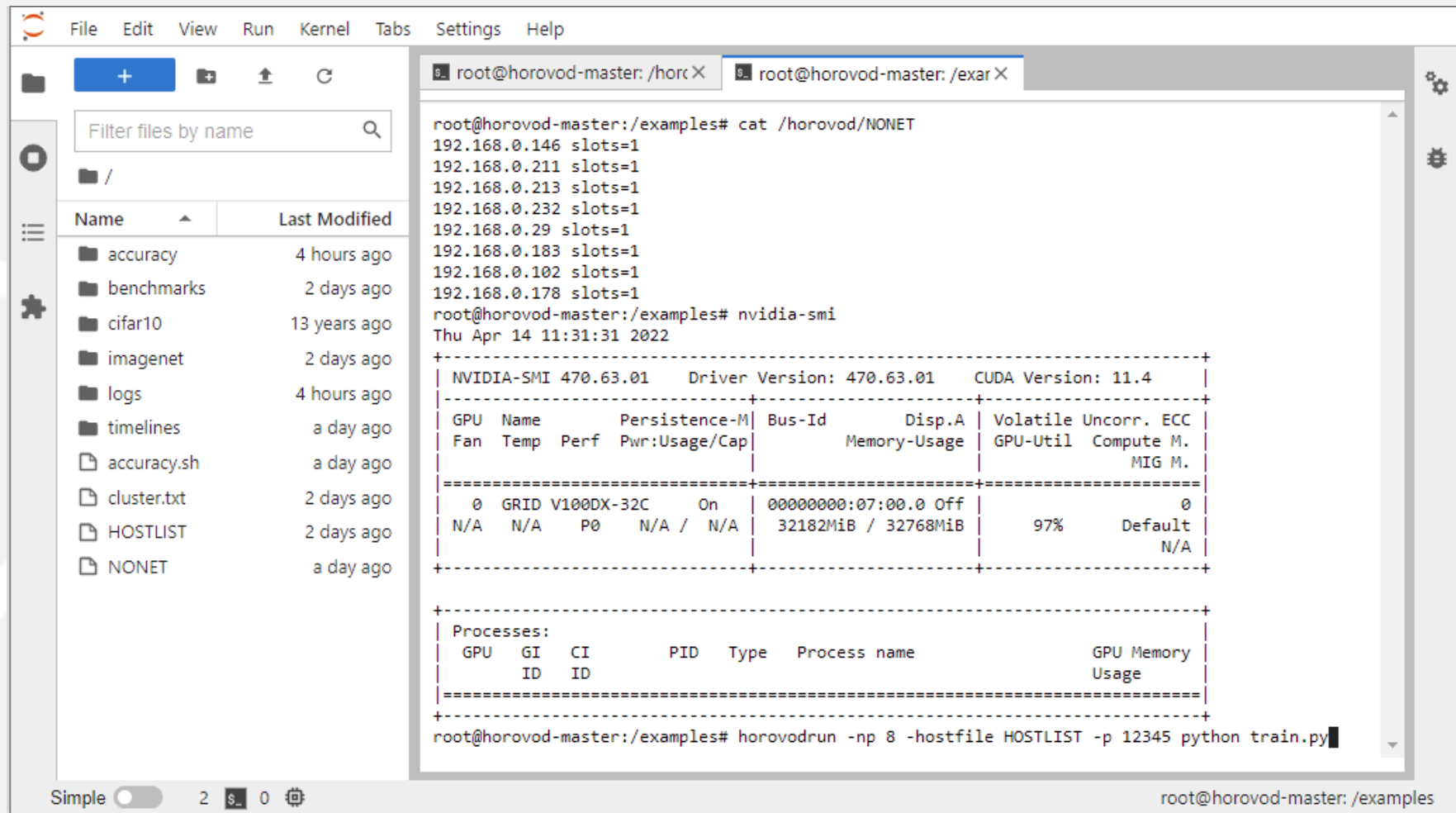
Sergeev, A., & Balso, M.D. (2018). Horovod: fast and easy distributed deep learning in TensorFlow. ArXiv, abs/1802.05799.

# Horovod Reference Architecture

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



# Horovod Reference Architecture - Usage



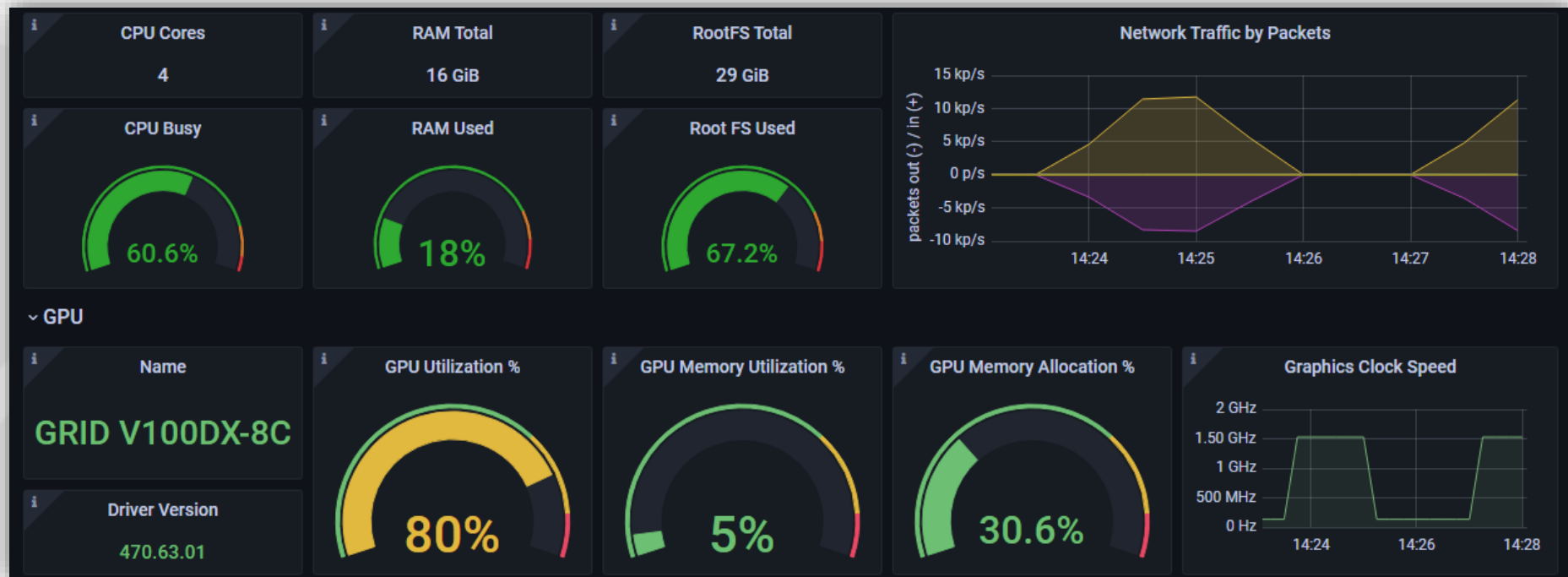
The screenshot shows a terminal window with a file explorer on the left and a terminal on the right. The file explorer shows a directory structure with folders like 'accuracy', 'benchmarks', 'cifar10', 'imagenet', 'logs', 'timelines' and files like 'accuracy.sh', 'cluster.txt', 'HOSTLIST', and 'NONET'. The terminal shows the following commands and output:

```
root@horovod-master:/examples# cat /horovod/NONET
192.168.0.146 slots=1
192.168.0.211 slots=1
192.168.0.213 slots=1
192.168.0.232 slots=1
192.168.0.29 slots=1
192.168.0.183 slots=1
192.168.0.102 slots=1
192.168.0.178 slots=1
root@horovod-master:/examples# nvidia-smi
Thu Apr 14 11:31:31 2022
+-----+
| NVIDIA-SMI 470.63.01    Driver Version: 470.63.01    CUDA Version: 11.4    |
+-----+
| GPU   Name           Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf    Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
|                                           MIG M.         |
+-----+-----+
| 0   GRID V100DX-32C   On          | 00000000:07:00.0 Off |                    |
| N/A   N/A    P0      N/A /  N/A | 32182MiB / 32768MiB | 97%      Default  |
|                                           N/A           |
+-----+-----+
+-----+
| Processes: |
| GPU   GI    CI          PID    Type    Process name                        GPU Memory |
| ID   ID   ID              |   |               | Usage |
+-----+-----+
root@horovod-master:/examples# horovodrun -np 8 -hostfile HOSTLIST -p 12345 python train.py
```



# Horovod Reference Architecture - Monitoring

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



# 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

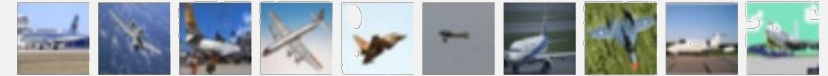


<https://www.nvidia.com/en-us/data-center/v100/>

# 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

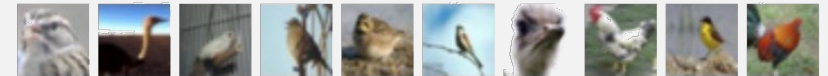
**airplane**



**automobile**



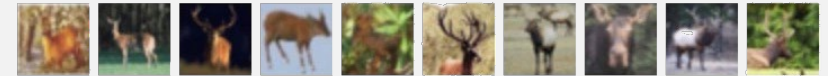
**bird**



**cat**



**deer**



**dog**



**frog**



**horse**



**ship**

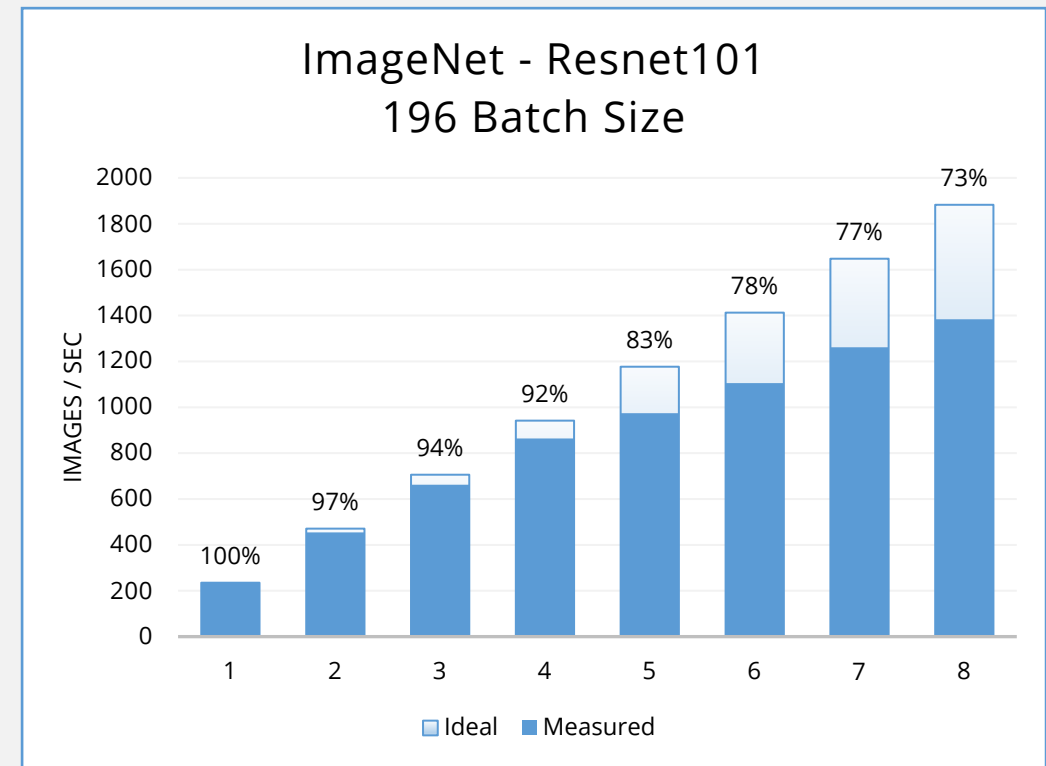
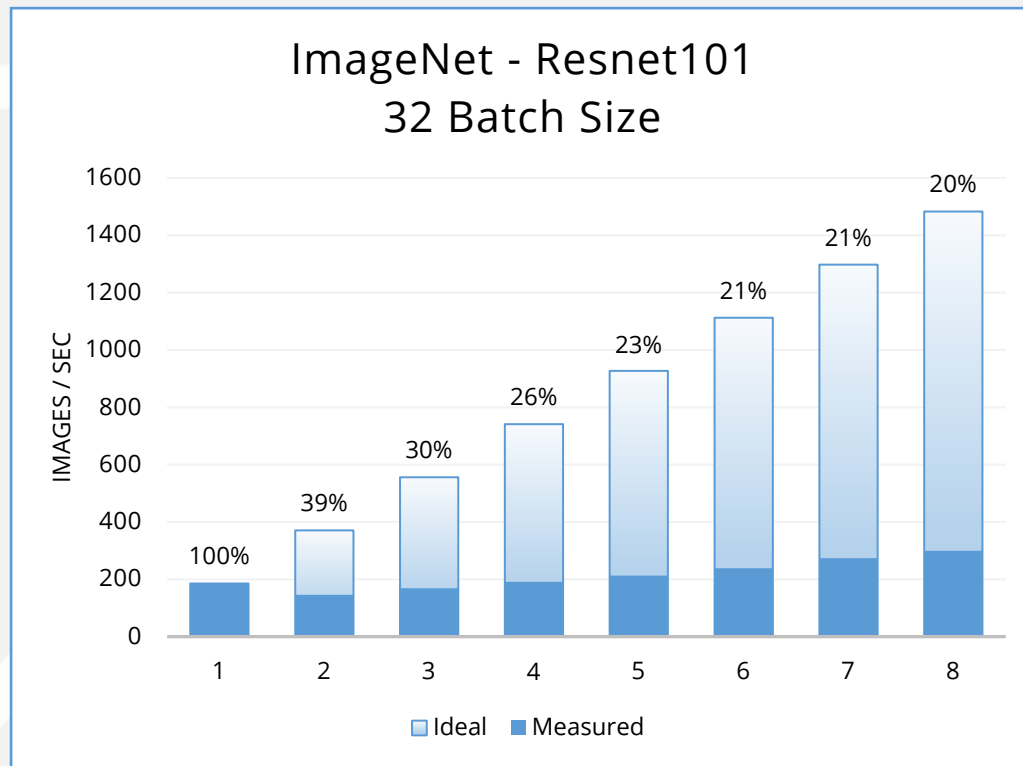


**truck**



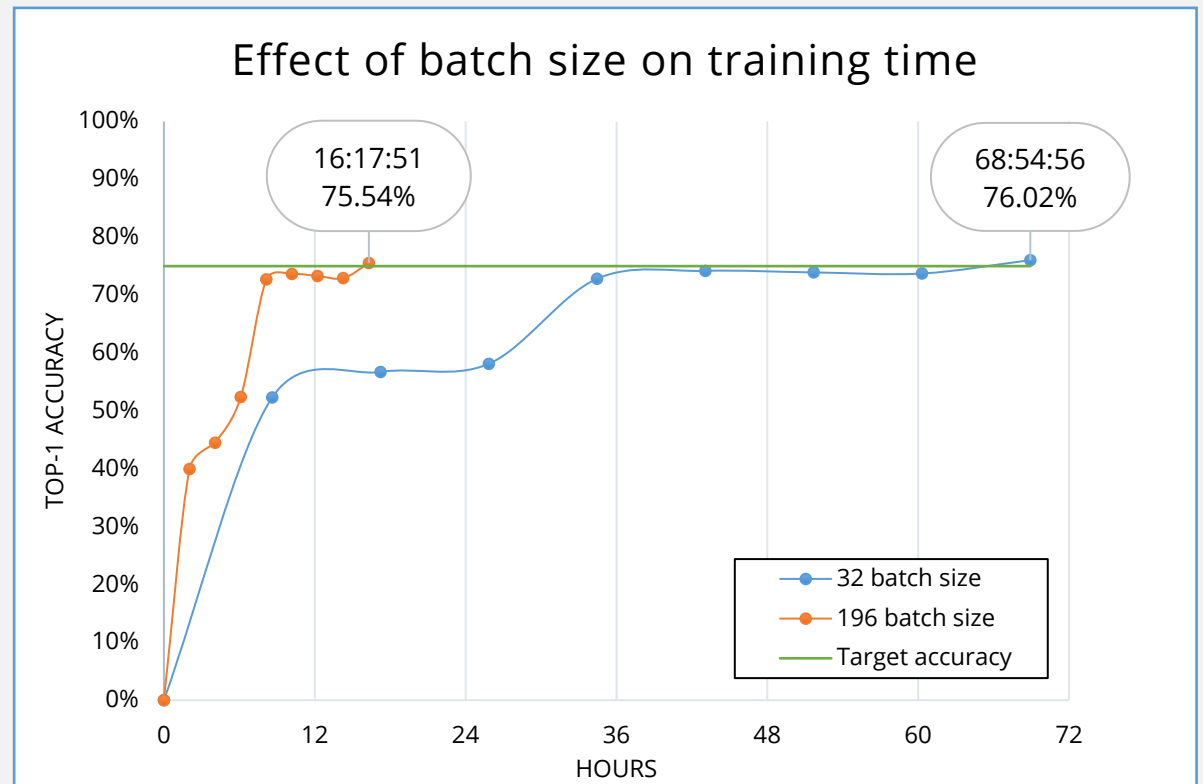
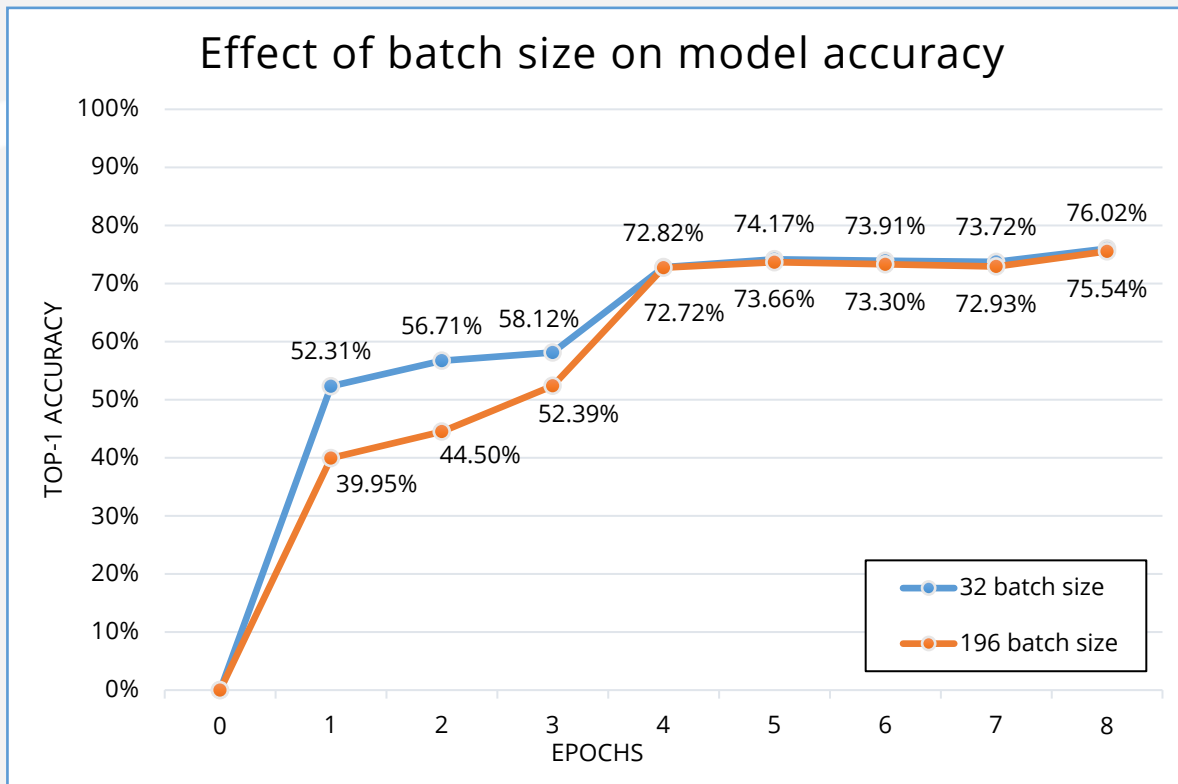
# Horovod Performance Evaluation - Results

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



# Horovod Performance Evaluation - Results

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





# Onboarding Horovod to the EOSC Marketplace

The screenshot shows the EOSC Marketplace interface. At the top left is the European Open Science Cloud logo. A search bar contains the text 'Find resource...'. To the right of the search bar are filters for 'All resour...' and a search icon. The user's name 'My EOSC Marketplace' is in the top right corner. Below the search bar is a breadcrumb trail: 'Resources > Access physical & infrastructures > Compute > Job Execution > Distributed Deep Learning by Horovod'. The main content area features a blue box with the SZTAKI logo. To the right of the logo is the title 'Distributed Deep Learning by Horovod', the acronym 'DDLbH', and a description 'deep learning, distributed, horovod'. The organization is listed as 'INSTITUTE FOR COMPUTER SCIENCE AND CONTROL'. There is a blue 'Access the resource' button with a lock icon and the text 'ORDER REQUIRED'. Below this are five empty star icons, a rating '(0.0 / 5)', and '0 reviews'. There are also checkboxes for 'Add to comparison' and 'Add to favourites'. A navigation bar includes links for 'Webpage', 'Helpdesk', 'Helpdesk e-mail', and 'Manual', along with a link to 'Ask a question about this resource?'. Below the navigation bar are tabs for 'ABOUT', 'DETAILS', and 'REVIEWS (0)'. The main text area contains the title 'Distributed Deep Learning by Horovod' and a bolded summary: 'Providing researchers a reliable platform designed for performing distributed deep learning operations with great scaling efficiency'. The text describes the service's aim to provide infrastructure and resources for distributed training of deep neural networks. It also mentions that Horovod is a distributed training framework designed for enabling the simple and effective distribution of deep learning operations across hundreds of GPUs with great efficiency. On the right side, there is a 'SCIENTIFIC CATEGORISATION' section with a yellow gear icon and the text 'Engineering & Technology'. Below this is a list of categories: 'Engineering & Technology' (indicated by a blue dot) and 'Electrical, Electronic & Information Engineering' (indicated by a blue dash).

# 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



## 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



## 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](http://www.sztaki.hu)