Reveal: Hardware-Centric Detection of Silent Inefficiencies in Machine Learning Systems

Funded by ARIA

Chen Ziji, Steven Chien, Peng Qian, Noa Zilberman
Department of Engineering Science, University of Oxford

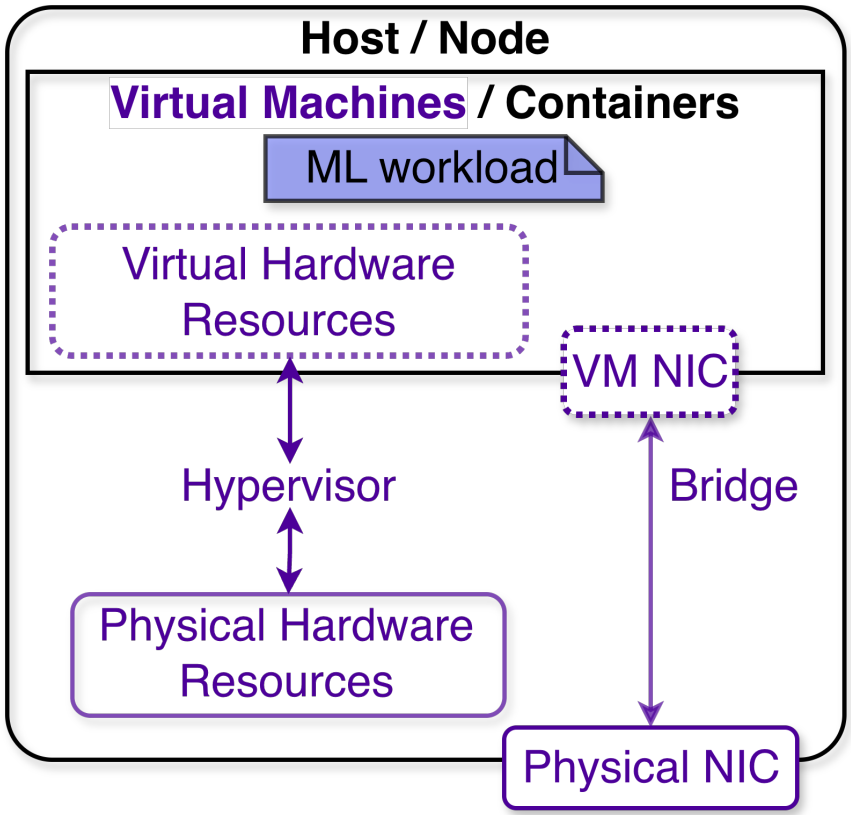
Tenth Annual UK System Research Challenges Workshop
April 22, 2026

Accelerate your ML model training or inference.

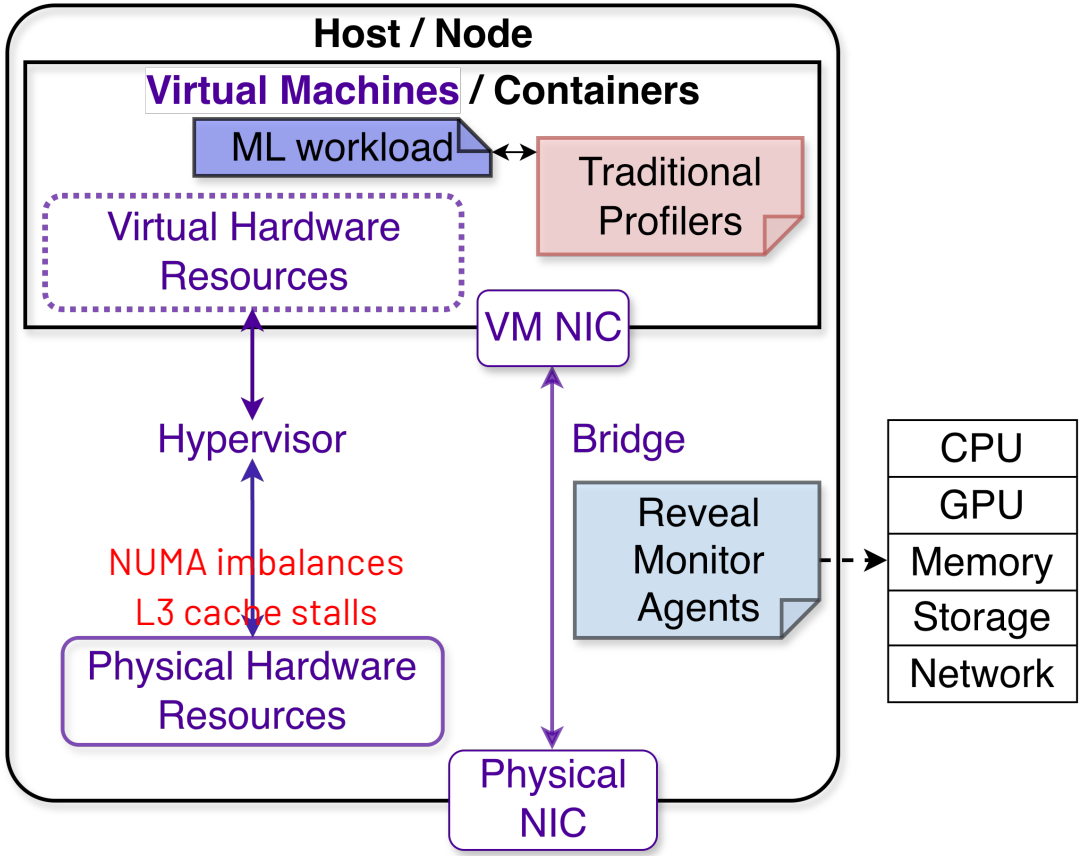
Is it the **CPU**? The **GPU**? **Network**? **Memory**? **Disk**?
Or… something you never thought about?

Background & Motivation

- Modern ML workloads **distributed** run in **virtualized** or **containerized** environments.
- Developers have **limited visibility** into host/node-level behavior.

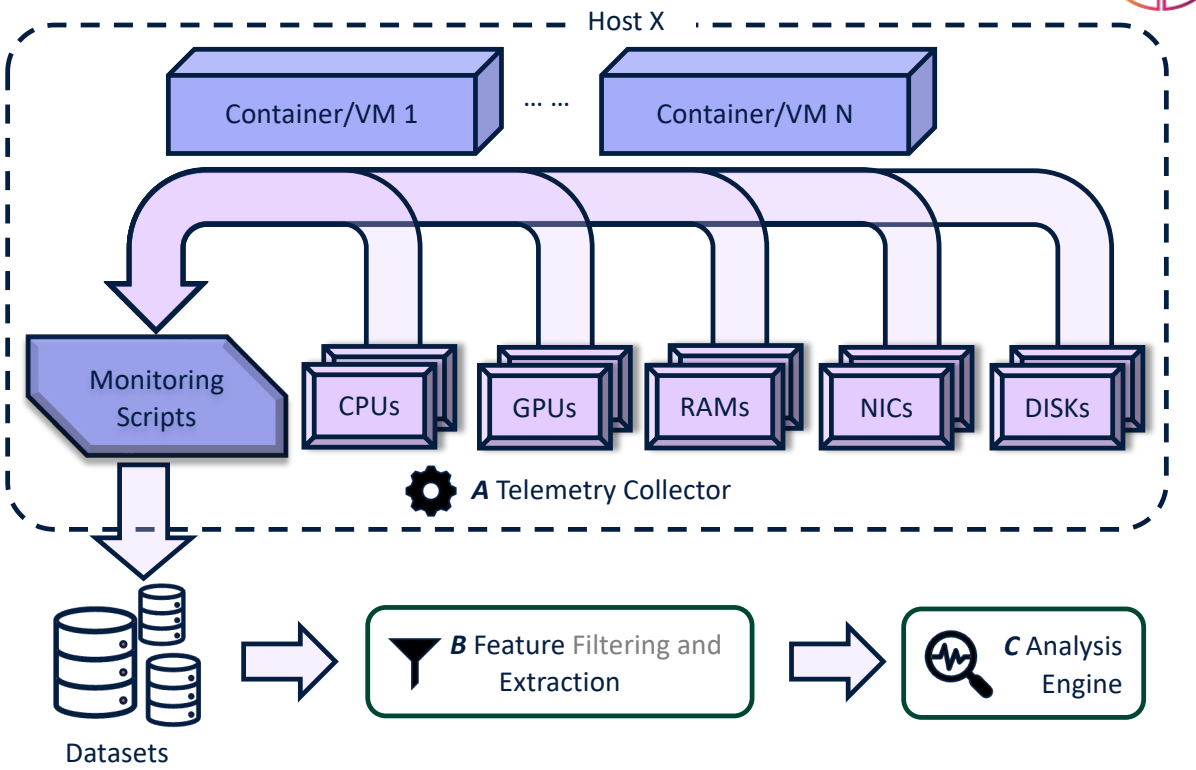


Why Existing Tools Aren't Enough



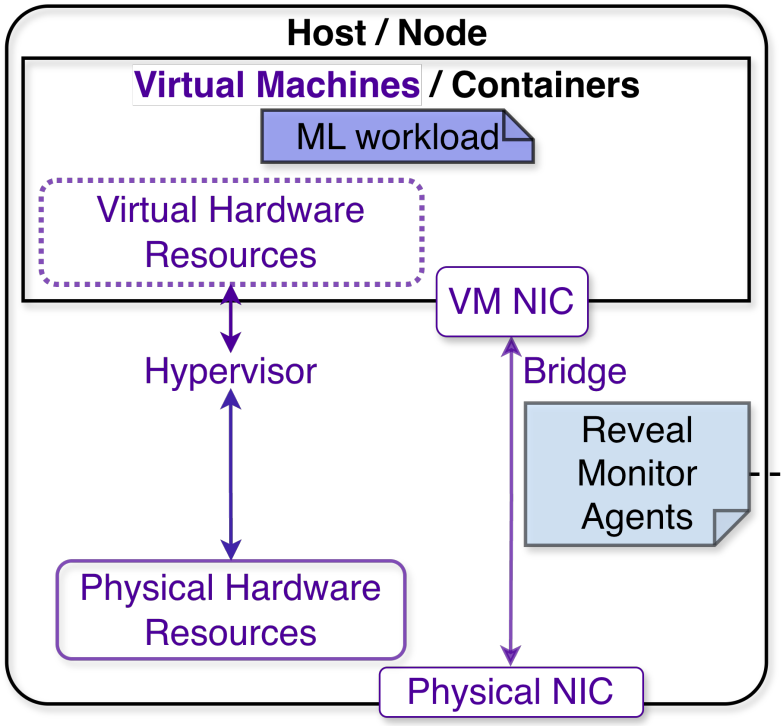
- High overhead
- Missing silent inefficiencies
- Requiring user code changing

System Overview



Goal: *Reveal* silent inefficiencies with minimal system overhead.

A Telemetry Collection



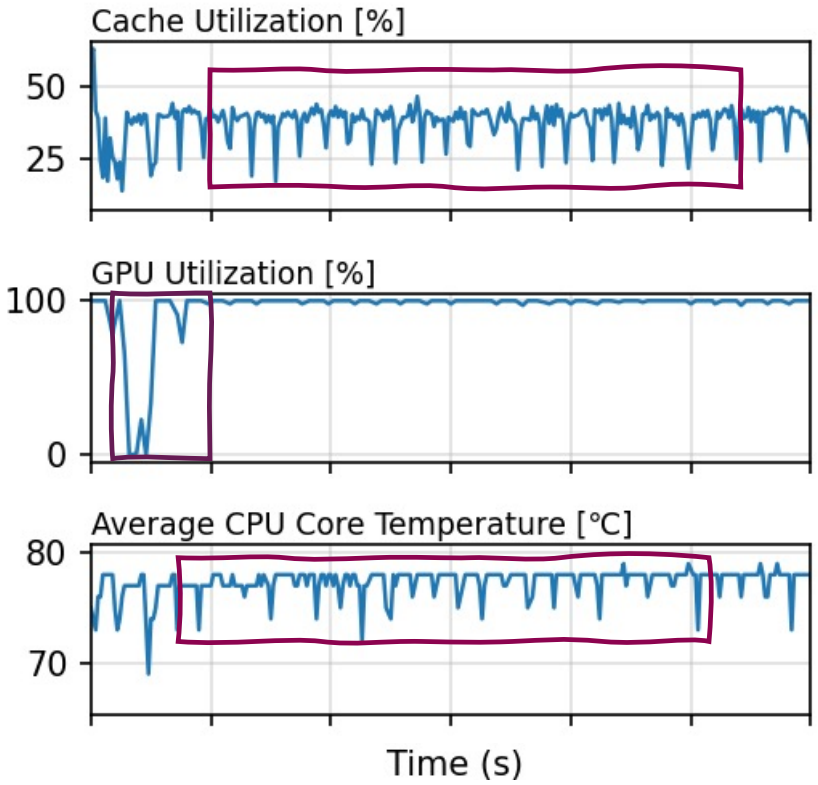
- 700+ system-level metrics
- CPU overhead < 1.5% at 100ms sampling
- Fed into the feature extraction pipeline

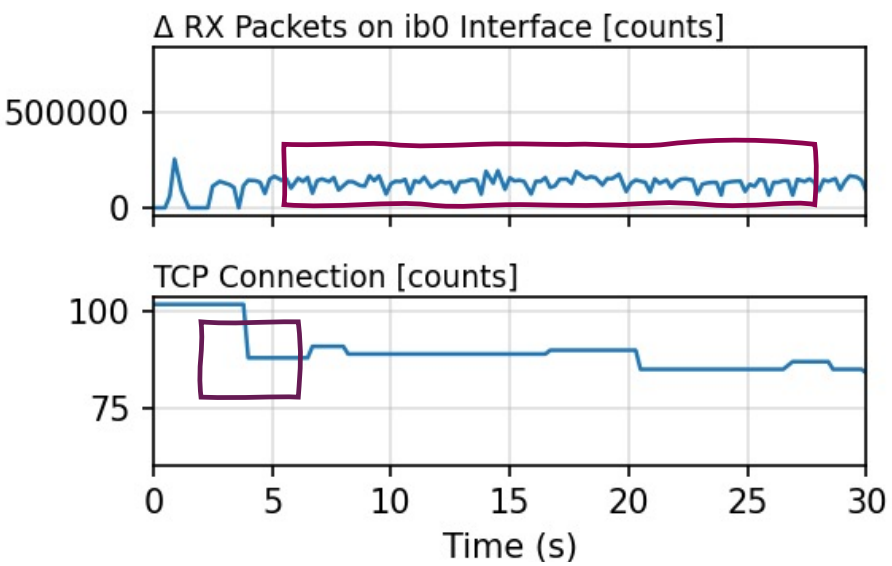
CPU	perf, /proc/stat
Memory	perf, /proc/meminfo
Storage	/proc/diskstat
Network	ss, sar, nstat
GPU	nvidia-smi

Timestamped

n-1	n-2	...	1	0	CPU Metrics
n-1	n-2	...	1	0	GPU Metrics
n-1	n-2	...	1	0	Memory Metrics
n-1	n-2	...	1	0	Network Metrics
n-1	n-2	...	1	0	Storage Metrics

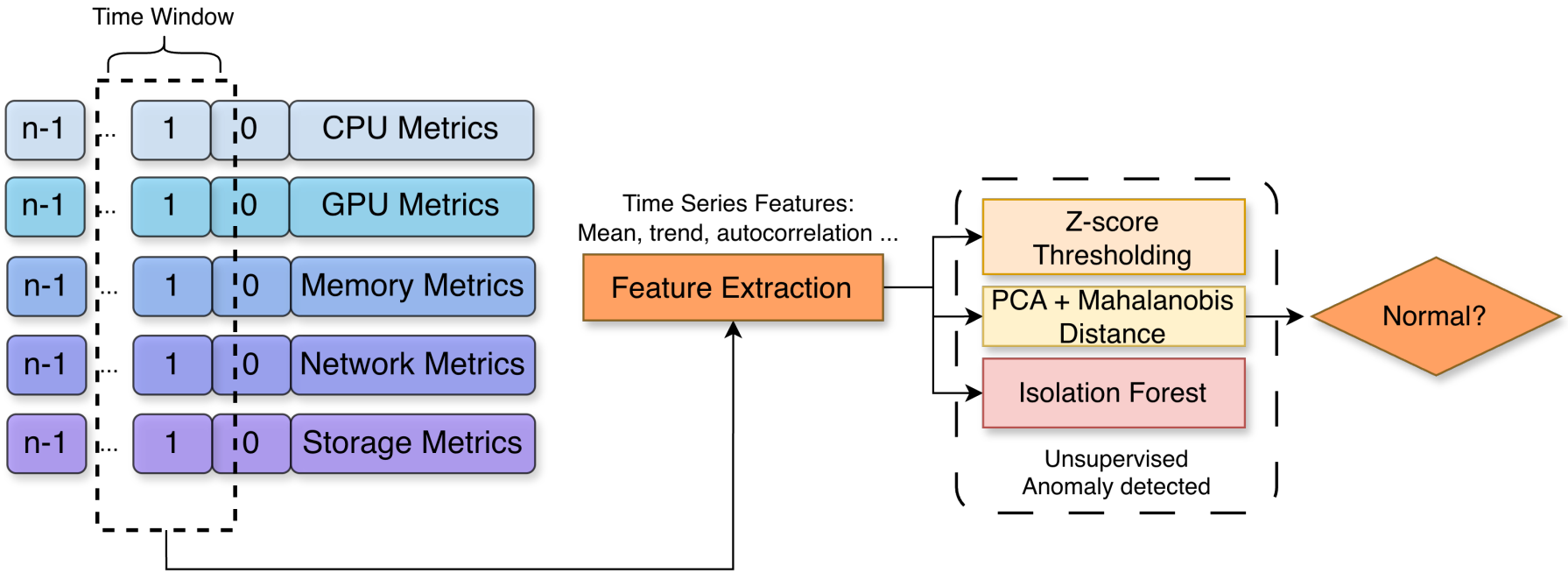
Observations



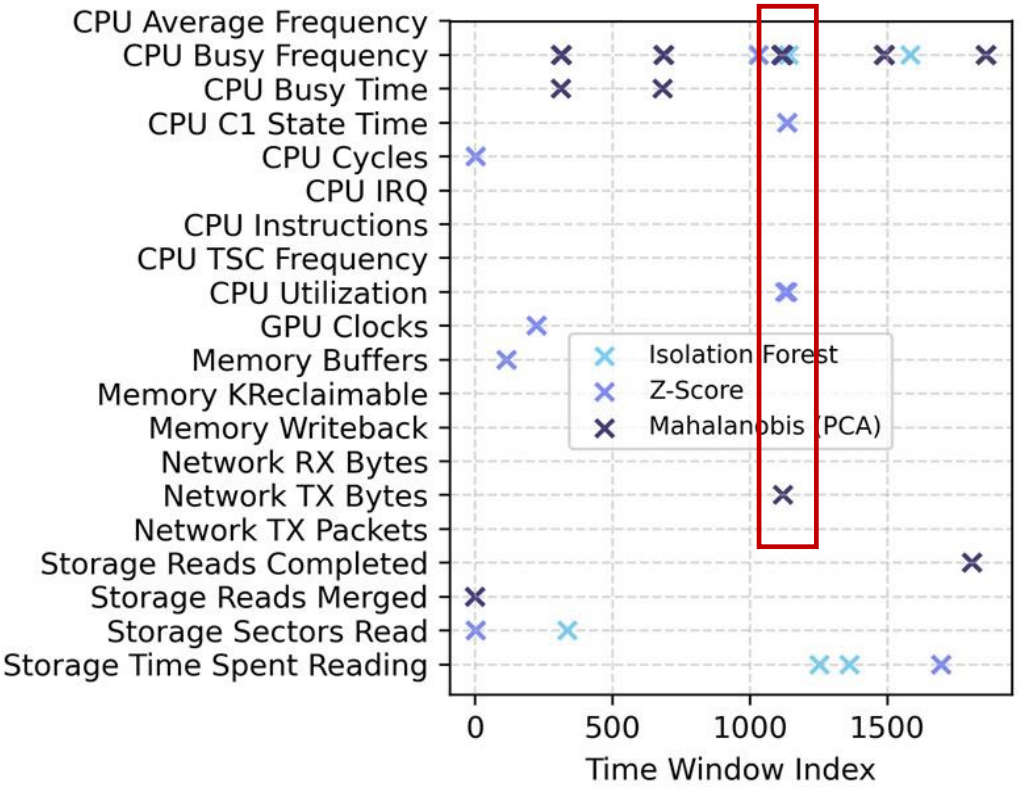


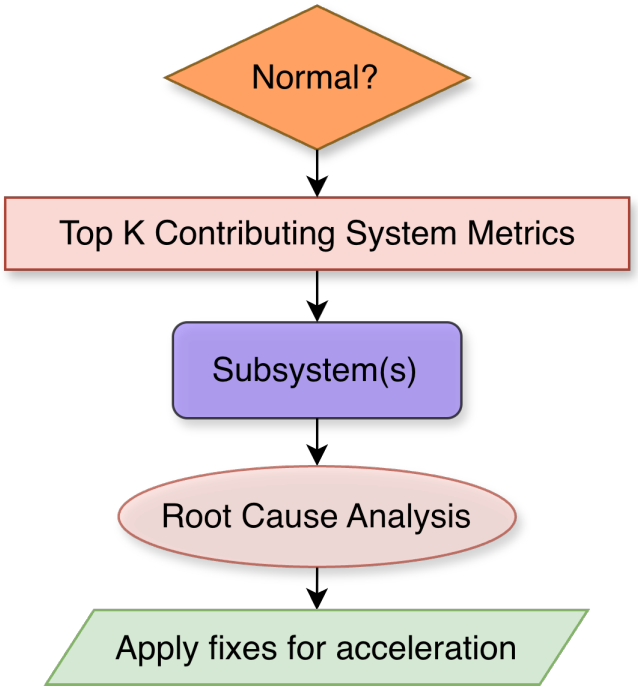
- (1) Periodic behavior
- (2) Large shifts

B Feature Filtering and Extraction & C.1 Analysis Engine (Unsupervised Anomaly detected)



C.2 Analysis Engine (Root Cause analysis and fixes)





Evaluation & Platform

DEPARTMENT OF ENGINEERING SCIENCE

UNIVERSITY OF OXFORD

1. Evaluated Models and Workloads

Model Type	Models	Model Type	Models
LLMs (MoE mode)	Llama-4-Scout-17B-16E/Llama-4-Maverick-17B-128E	**Encoder-Decoder**	BART (base/large)
	DeepSeek-MoE-16b-base, **DeepSeek-R1-Dstil-Qwen-7b**	**Encoders**	BERT (base/large), DistilBERT
LLMs (Dense mode)	LLama-3.1-8B/3.2-3B	**CNNs**	ResNet18/50, VGG16/19
	Deepseek-LLM-7B/67B-base	**Transformers (ViT)**	ViT (Base/Large)

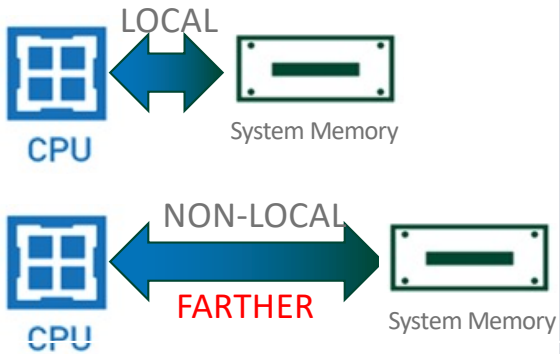
2. Hardware Configuration

Num. of GPUs	Num. of Nodes	CPU Type	Memory	Network
16 L40S	**4**	Intel Emerald Rapids (64 cores)	384 GB / node	NDR400 InfiniBand
4 H100	**2**	Intel Sapphire Rapids (96 cores)	1 TB / node	HDR100 InfiniBand
4 A100	**2**	Intel Cascade Lake (48 cores)	384 GB / node	HDR100 InfiniBand
4 V100	**2**	**Intel Cascade Lake (48 cores)**	**384 GB / node**	**HDR100 InfiniBand**
0	**10**	**AMD EPYC 7443P**	**256 GB / node**	**100GE Ethernet (Tofino)**

Detected Misconfigurations

NUMA misplacement (Memory not local)

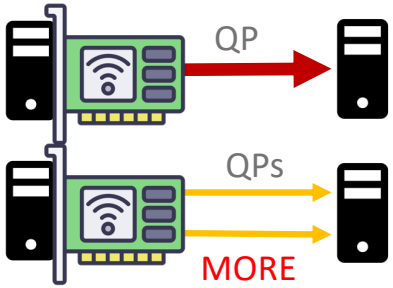
IPC ↓, L3 miss ↑, stall ratio ↑



+ 6% speed up.

Network misconfiguration (NCCL QP)

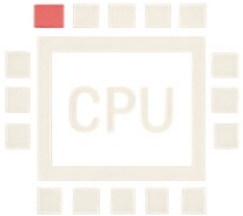
CPU Busy% ↑, ib0 TX/RX bursty, GPU power ↓



+ 3% speed up.

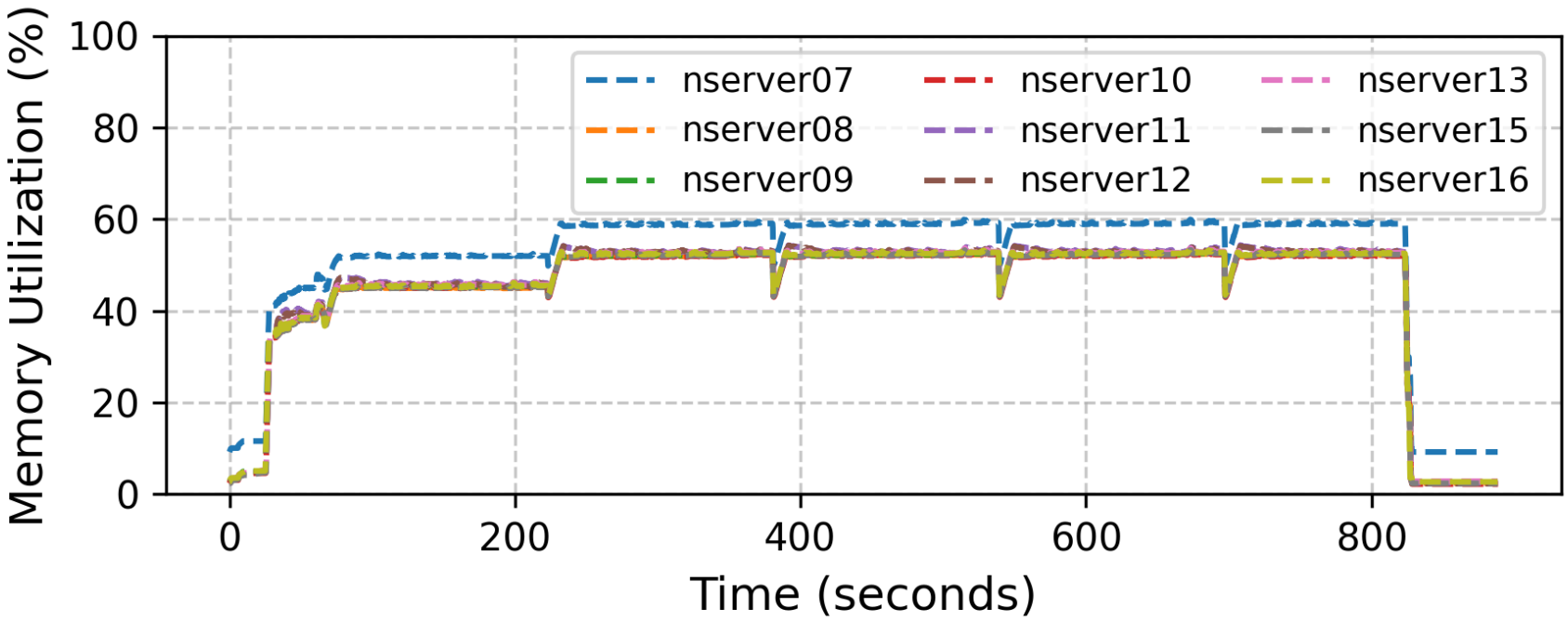
CPU IRQ imbalance

IRQ bursts on few cores, CPU Busy% ↑



Mitigated.

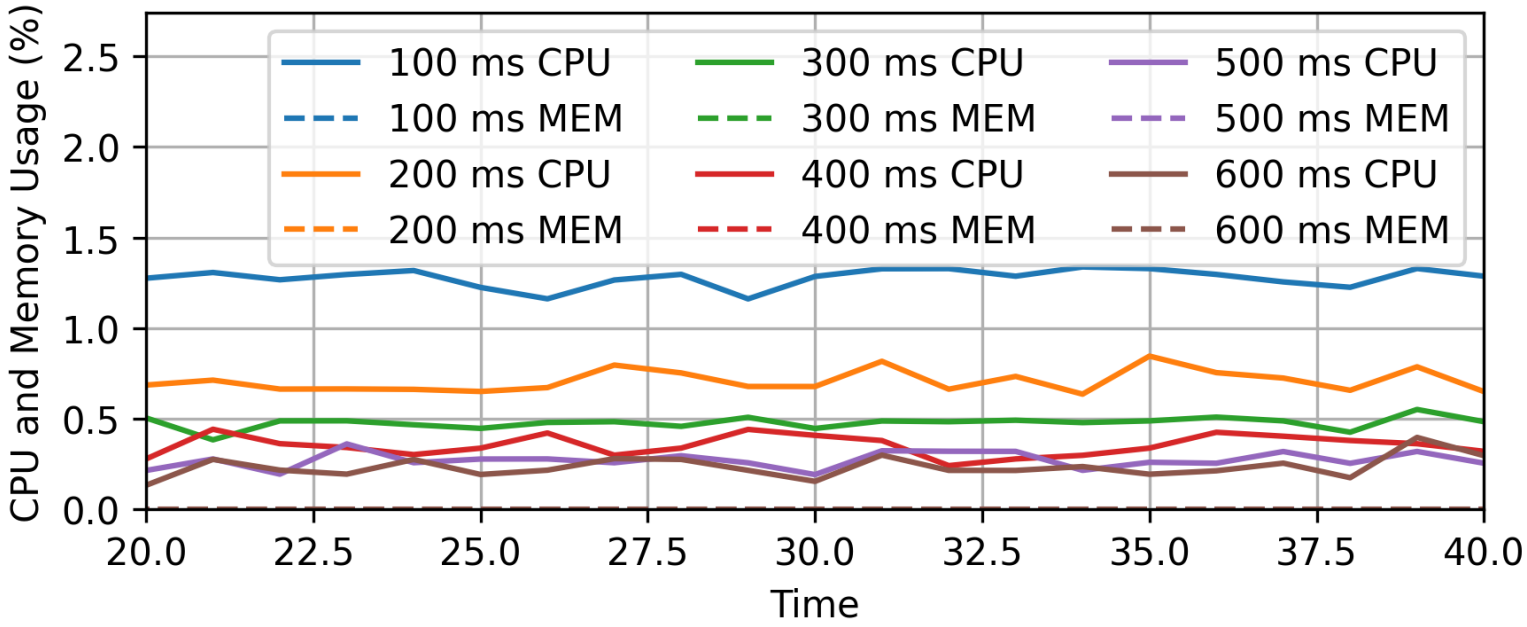
Memory Misconfiguration



- **Problem:** One node exhibited unusually high memory usage during distributed training; Distribution-aware workload expected balanced memory footprint.
- **Root Cause:** Inconsistent HugePages configuration across nodes.
- **Solution:** Unified HugePages settings across the cluster.
- **Impact:** Memory usage normalized;Improved consistency in resource scheduling and model performance.

Scalability & Portability

- CPU overhead **< 1.5% @ 100ms sampling**
- Memory overhead **negligible**
- Storage: **14–22 KB/s per node**
- Feature extraction + detection: **millisecond–second level latency**
- Won't scales with node count



Summary

- ML can be slow — and it's **not always the model's fault**
- **We need visibility** across the full stack
  → Our framework monitors CPU, GPU, memory, network, storage — from the host side, with <1.5% overhead at 100ms sampling.
- Unsupervised methods help surface anomalies
  → Even **without labels**, we can detect weird patterns in system behavior.
- **Real misconfiguration. Real fixes.**
  → Like HugePages misconfigurations that caused memory imbalance — found and resolved.

Thank you!

ziji.chen@eng.ox.ac.uk