CARMA: Collocation-aware Resource Management System for Deep Learning Training Tasks

Ehsan Yousefzadeh-Asl-Miandoab[1] (ehyo@itu.dk), Reza Karimzadeh[2], Bulat Ibragimov[2], Florina M. Ciorba[3], Pınar Tözün[1]
(1) IT University of Copenhagen, (2) University of Copenhagen, (3) University of Basel

1 GPU Underutilization: Causes and Opportunities

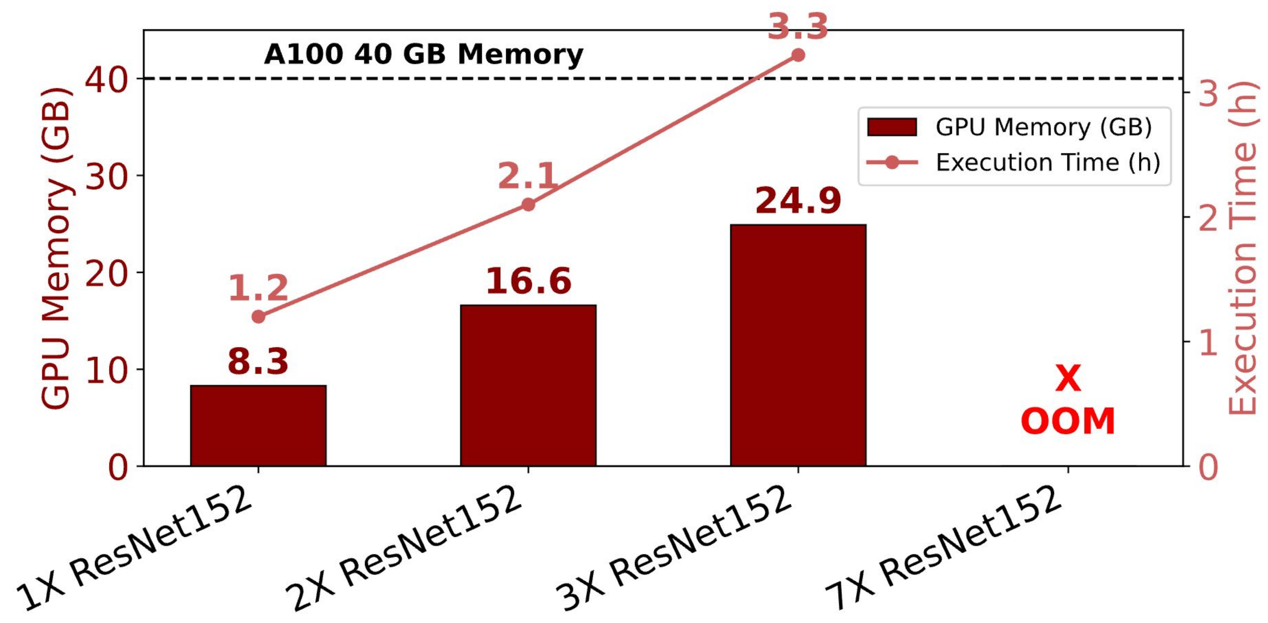
Real-world clusters exhibit only ~50% GPU utilization *

1- GPUs' lack of **fine-grain sharing** and **virtual memory**
2- **Exclusive** GPU assignment by resource managers
3- **Black box** view of tasks and GPUs

Collocating tasks together increase GPU utilization!

* Yanjie Gao et al. "**An Empirical Study on Low GPU Utilization of Deep Learning Jobs**," ICSE'24.

2 OOM Crashes & Interference!



GPU memory estimation is essential before robust collocating.

3 Estimating GPU memory: GPUMemNet

- Lightweight deep learning-based estimator.
- Synthetic, architecture-guided datasets.
- Rarely underestimates GPU memory need.

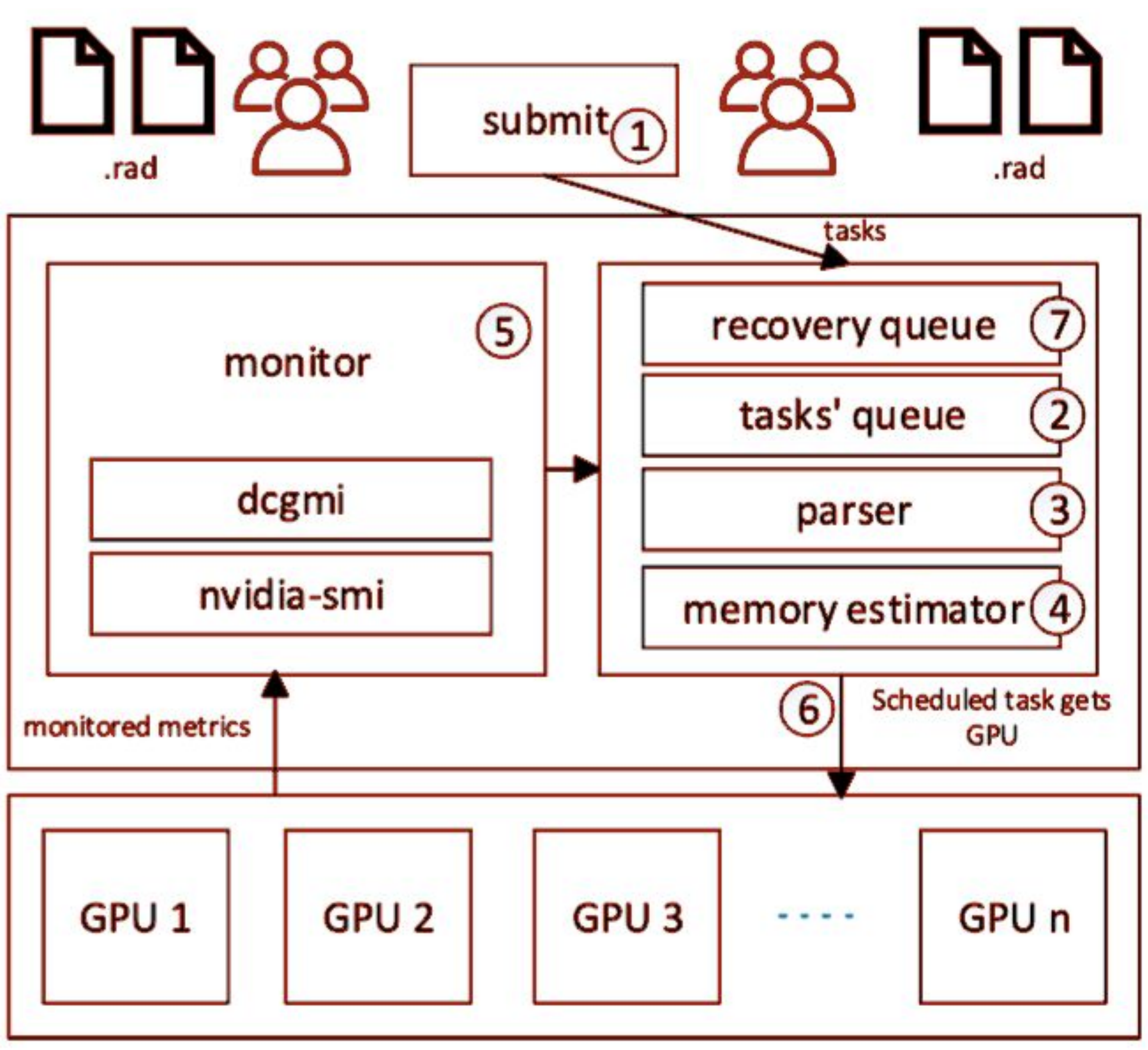
Machine Learning excels at pattern recognition.

4 Interference

- Resource interference can erase collocation throughput gains.
- No throughput gains when GPU is already more than 80% utilized.

5 CARMA Architecture

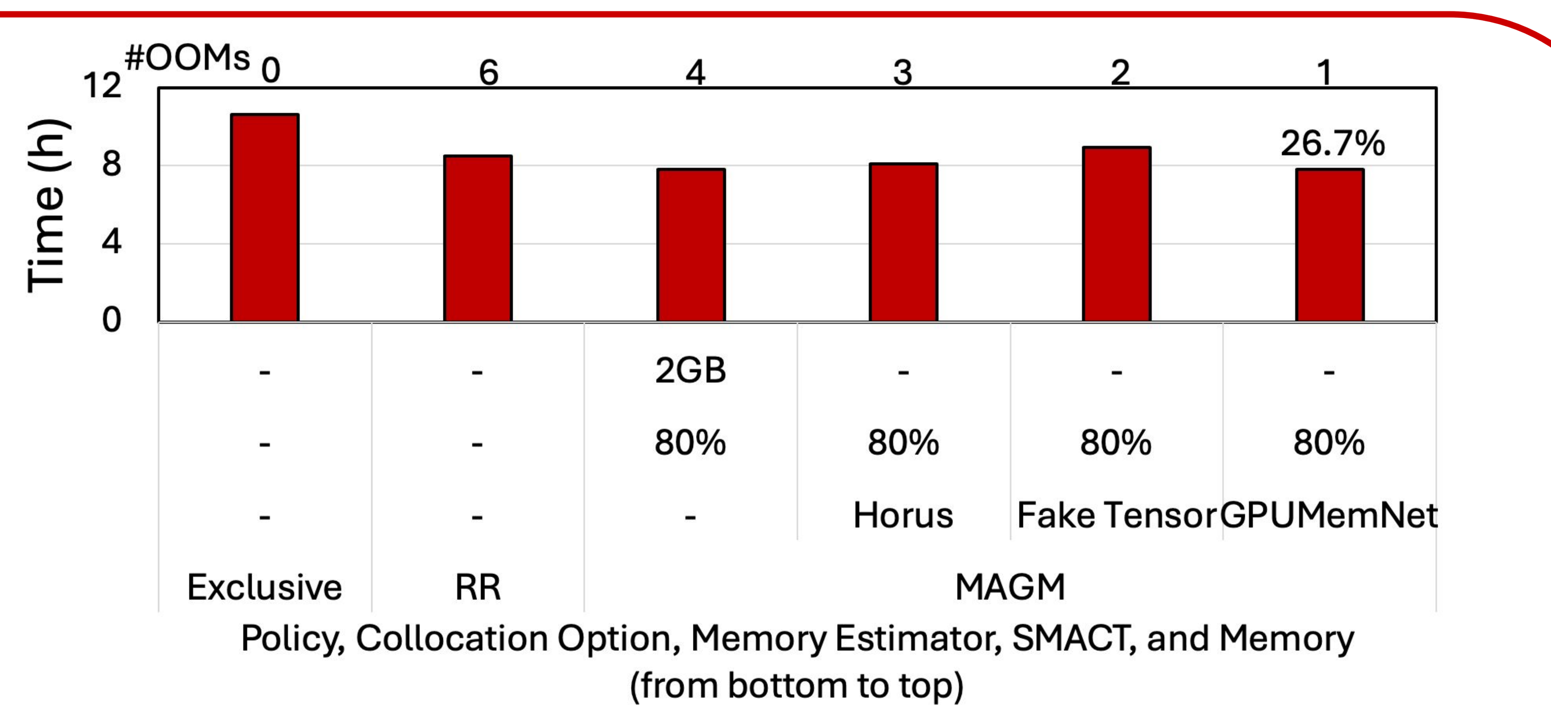
- Monitoring
  - **SMACT** (GPU Utilization): averaged over 1 min
  - GPU Memory: last observed value used
- Preconditions
  - **SMACT** <= 75%-80%
  - GPU Memory >= 2GB, 5GB
- Collocation Policies
  - Exclusive
  - Round Robin (**RR**)
  - Most Available GPU Memory (**MAGM**)
- Recovery when memory estimation falls short.

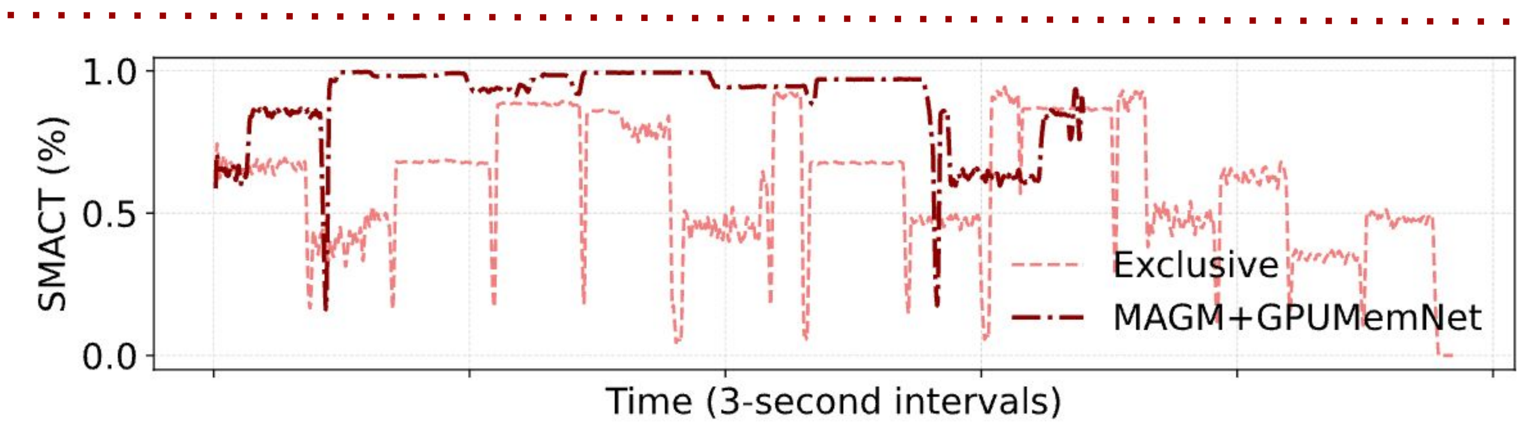


6 Evaluation

- **60-task Philly-based Trace**
- **A100 DGX Station (4X GPUs)**

- Collocation policies & preconditions affect #OOMs.
- Least interference promises higher performance via throughput gain.
- Collocation-aware resource management improves GPU utilization (**39.3%**) and energy efficiency (**14.2%**).





IT UNIVERSITY OF COPENHAGEN
itu.dk

Universität Basel

RAD
rad.itu.dk