LGNNIC: Acceleration of Large-Scale GNN Training using SmartNICs

Liad Gerstman / 30 June 2024
Aditya Dhakal
Sai Rahul Chalamalasetti
Chaim Baskin
Dejan Milojicic

Agenda

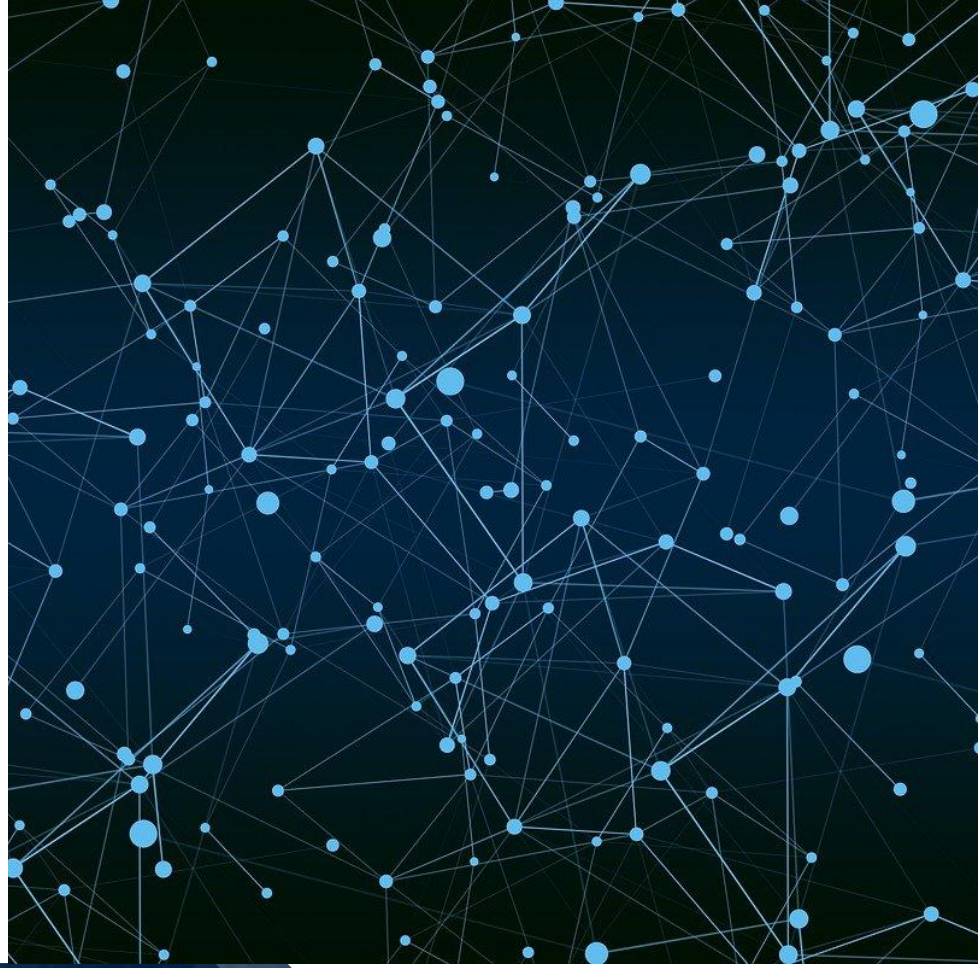
1 | Large-Scale GNN Challenge

2 | Our LGNNIC Architecture Solution

3 | Baseline Local Node Evaluation

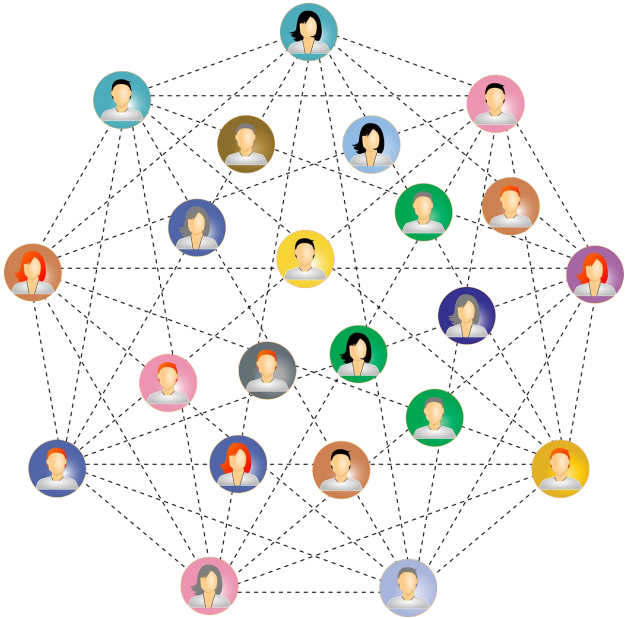
4 | Remote Node Evaluation

5 | Next Steps and Conclusions



Graph Neural Networks (GNNs)

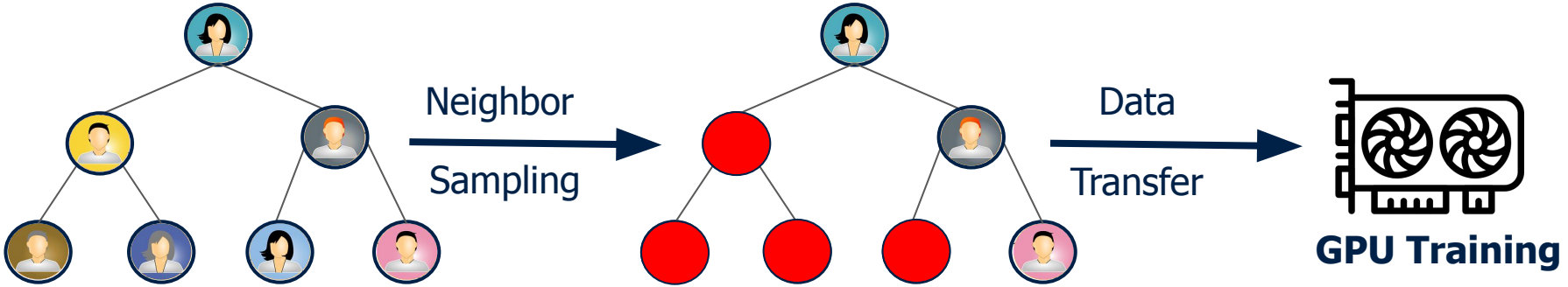
- Graphs are composed of relations (edges) between different components (nodes).
- For each node in the graph, a computational graph is constructed to calculate the node embeddings.
- Node and edge features in a node's neighborhood are aggregated using learned parameters -> GNN
- Several GNN types exist: GCN, GAT, **GraphSAGE**, GIN.



Large-Scale GNNs Challenge

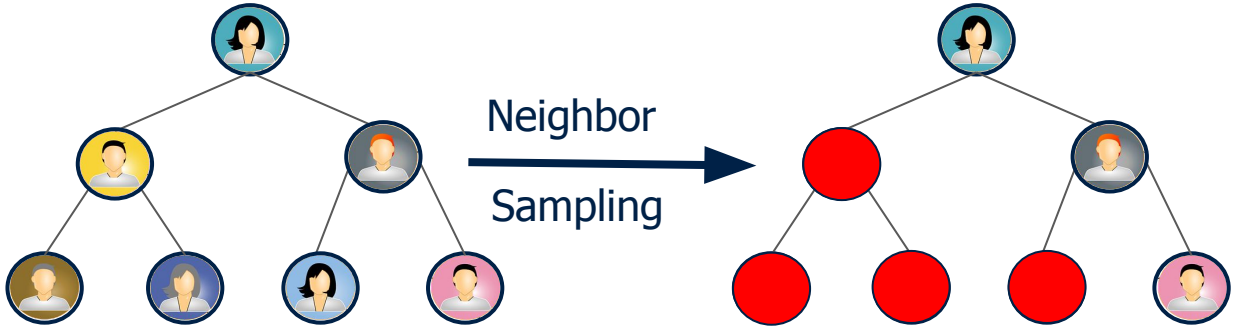
- Large-scale GNNs improve algorithmic performance but exceed GPU memory capacity.
- CPUs, with larger DRAM capacities, can store larger graphs in memory.

Mini-batch sampling methods can be executed on CPU to reduce the graph size so it can fit into GPU memory.



Mini-Batch Neighbor Sampling

- Recursively choosing a subset of neighbors for each mini-batch to prevent explosion.
- Random sampling vs Random Walk with Restarts.
- GraphSAGE model's neighbor sampling: Limiting number of layers & nodes per layer.
- Potential Sampling drawbacks:
  - execution time
  - reduced accuracy



What if the large graph does not fit CPU memory?

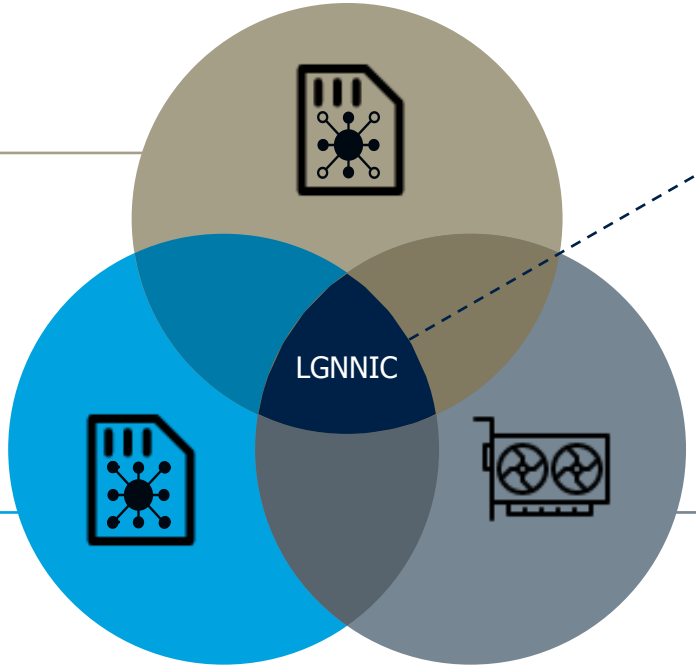
Our Solution - LGNNIC: Acceleration of Large-Scale GNN Training using SmartNICs

Remote Sampling

Neighborhood **sampling** (and more) on **remote** memory / storage nodes to reduce **network bottlenecks**.

Remote Memory

Using **remote** memory / storage nodes to **store** the large datasets can alleviate memory capacity bottlenecks



LGNNIC

Acceleration of Large-Scale GNN Training using SmartNICs

Local Training

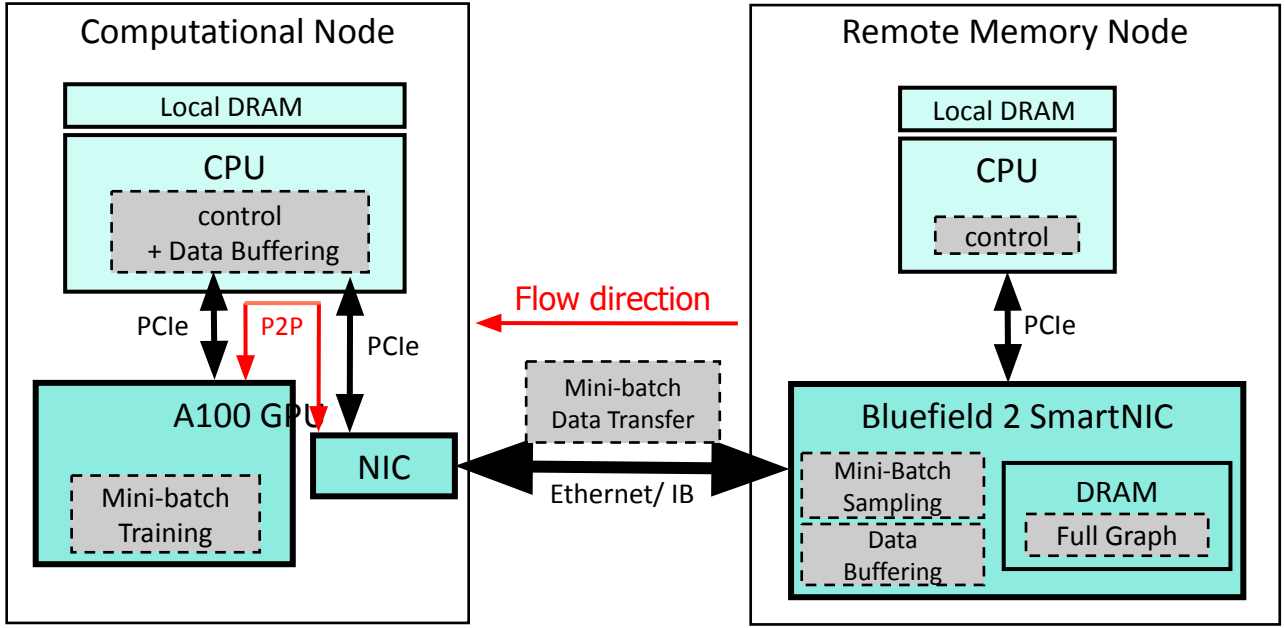
Transferring sampled graph to **local** node's GPU for mini-batch **training**

Hewlett Packard Labs

TECHNION Israel Institute of Technology

Our Solution - LGNNIC System Architecture

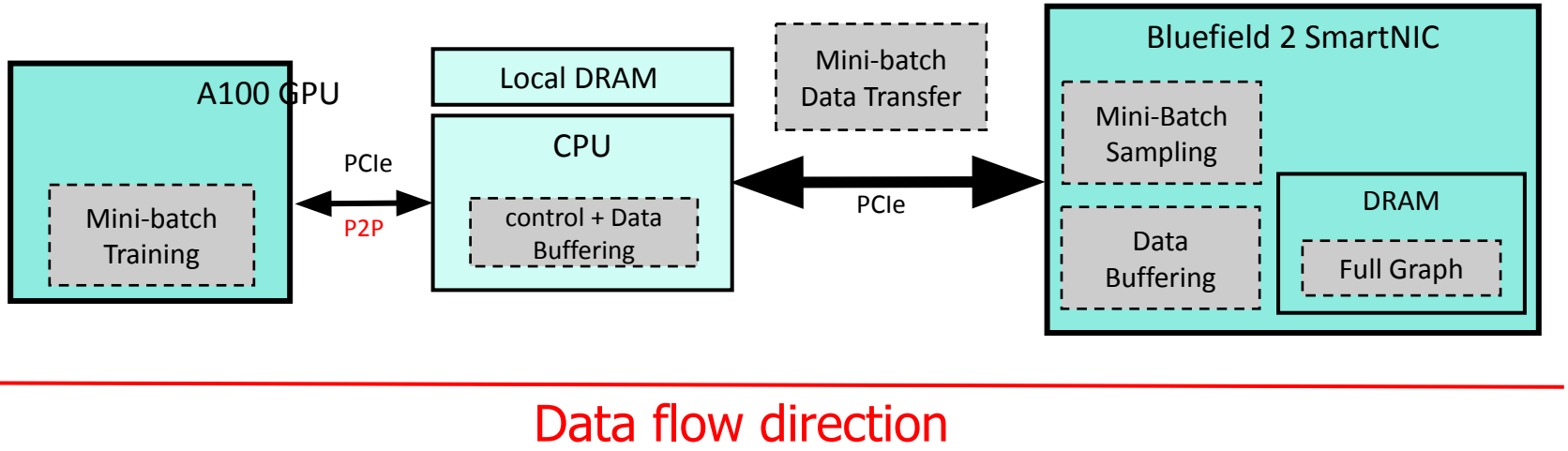
Neighborhood sampling (and more) on remote memory / storage nodes to reduce network bottlenecks.



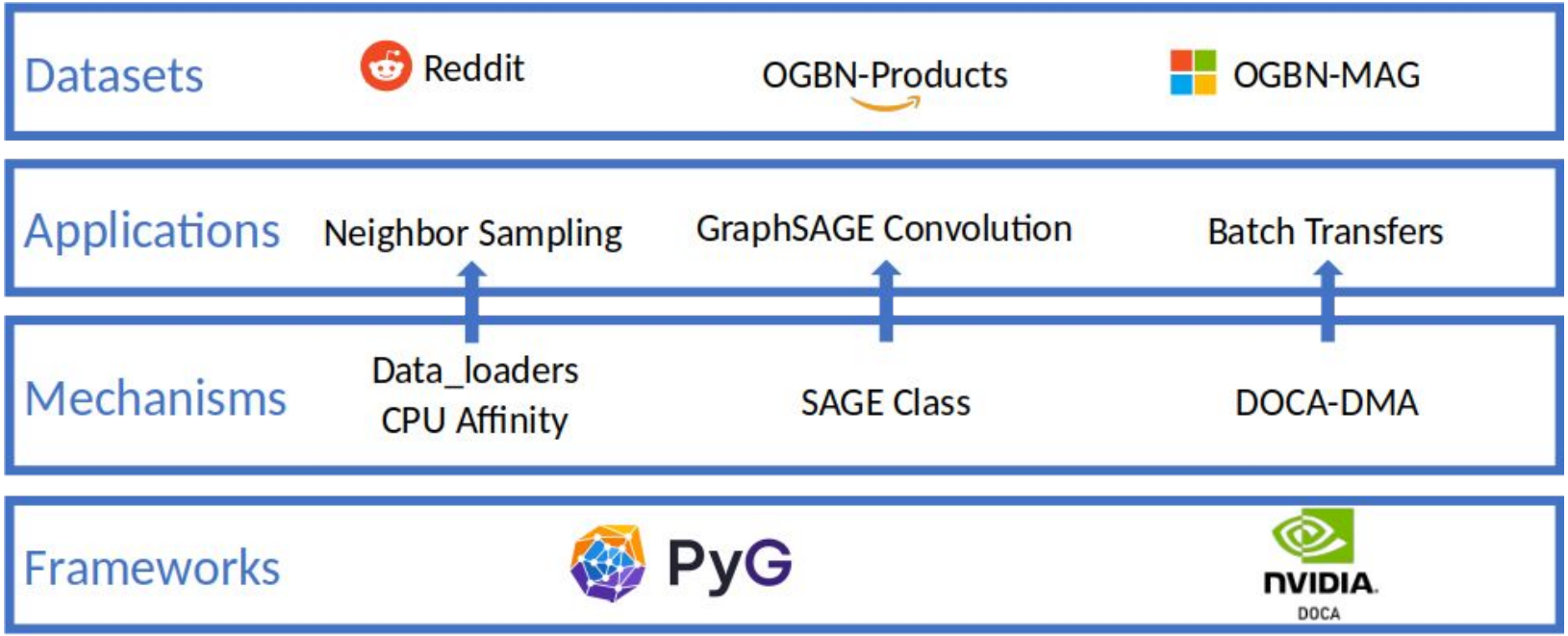
- We exemplify our concept focusing on memory nodes only.

POC System

Neighborhood sampling (and more) on remote memory / storage nodes to reduce network bottlenecks.

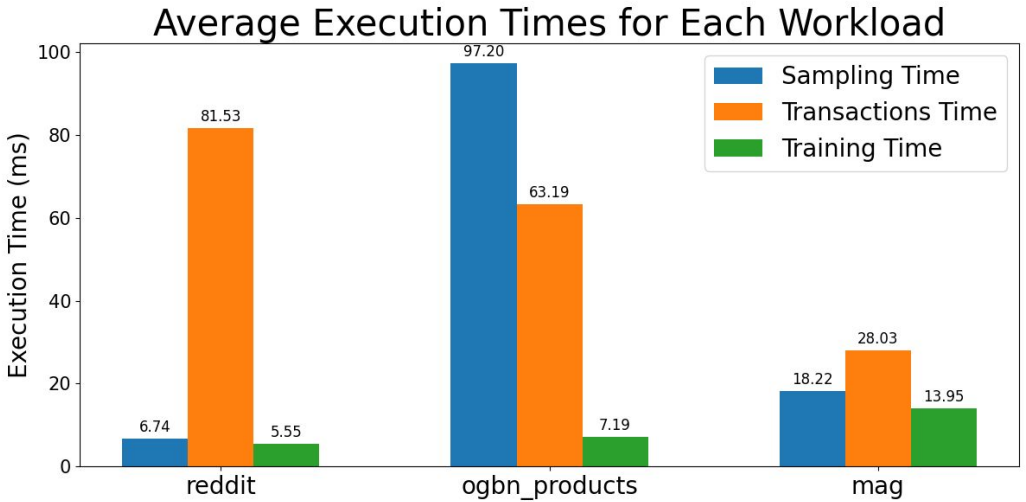


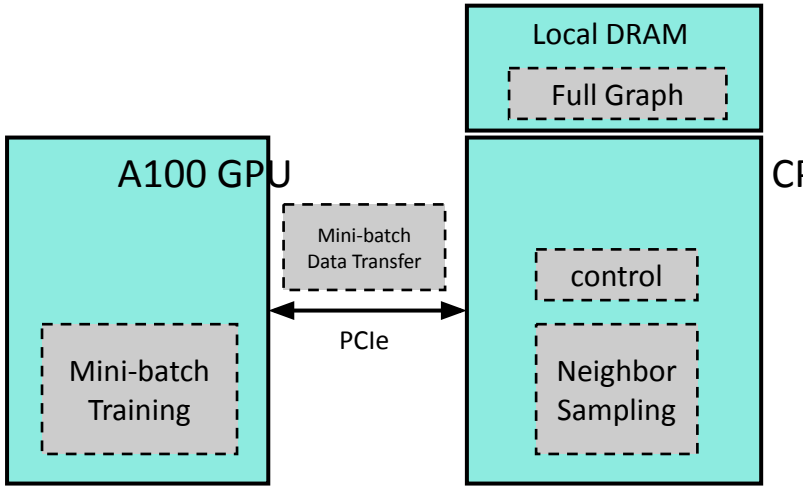
LGNNIC Synchronization Mechanism Baseline



Baseline Single Node Mini-Batch Sampling

- Setup: AMD EPYC 7513 32-Core processor and NVIDIA A100 GPU
  - 2TB CPU DRAM & 80GB GPU HBM
- Sampling is done on CPU.
- Transaction times range from 28ms to 81.53 ms.





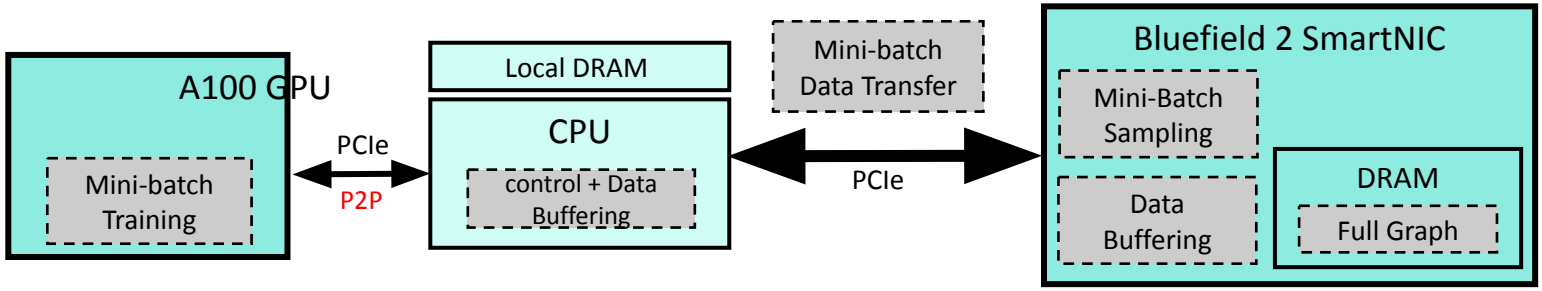
Hewlett Packard Labs

TECHNION Israel Institute of Technology

Mini-Batch Neighbor Sampling on SmartNIC

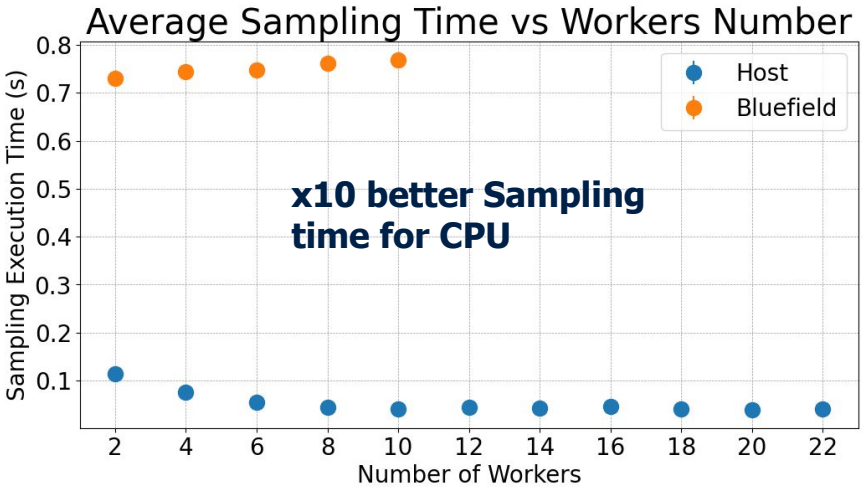
POC System - Reddit

Using remote memory / storage nodes to store the large datasets can alleviate memory capacity bottlenecks.



Mini-Batch Neighbor Sampling on SmartNIC Challenge

- What are the computational disparities between our SmartNIC and our local node CPU?
  - Bluefield-2 with 8 ARMv8 A72 Cores vs. EPYC 7513 32-Core processor
- GraphSAGE mini-batch Neighbor Sampling time execution measurements for Reddit

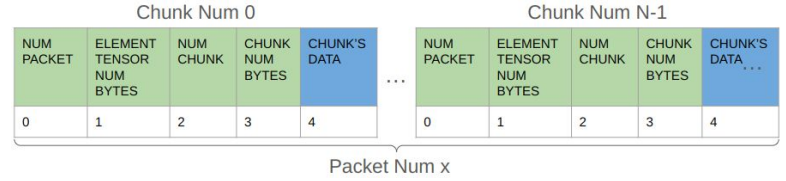


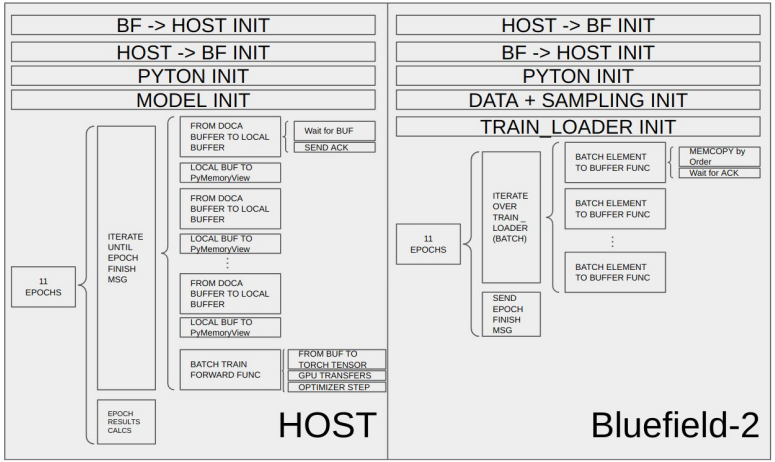
Hewlett Packard Labs

TECHNION Israel Institute of Technology

LGNNIC Synchronization Mechanism

- Ensures data **coherency** between local and remote buffers.
- DOCA DMA integration with PyG.
- Bluefield send mini-batches to host in chunks.
  - Inherent 1 MB DOCA buffer limitation.
  - ACK based mechanism.
- While host trains the current batch - the next one is already being transferred.



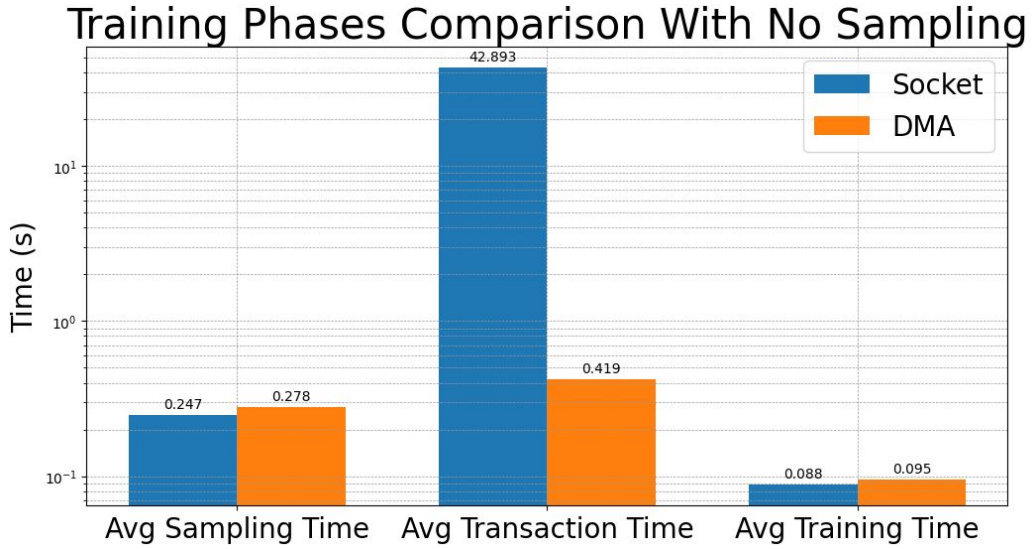


Hewlett Packard Labs

TECHNION Israel Institute of Technology

SmartNIC - CPU Transfers - Network Bottleneck

- Transferring mini-batches from remote node to local node is a network bottleneck.
- Our synchronization mechanism is much faster then the socket-based benchmark.

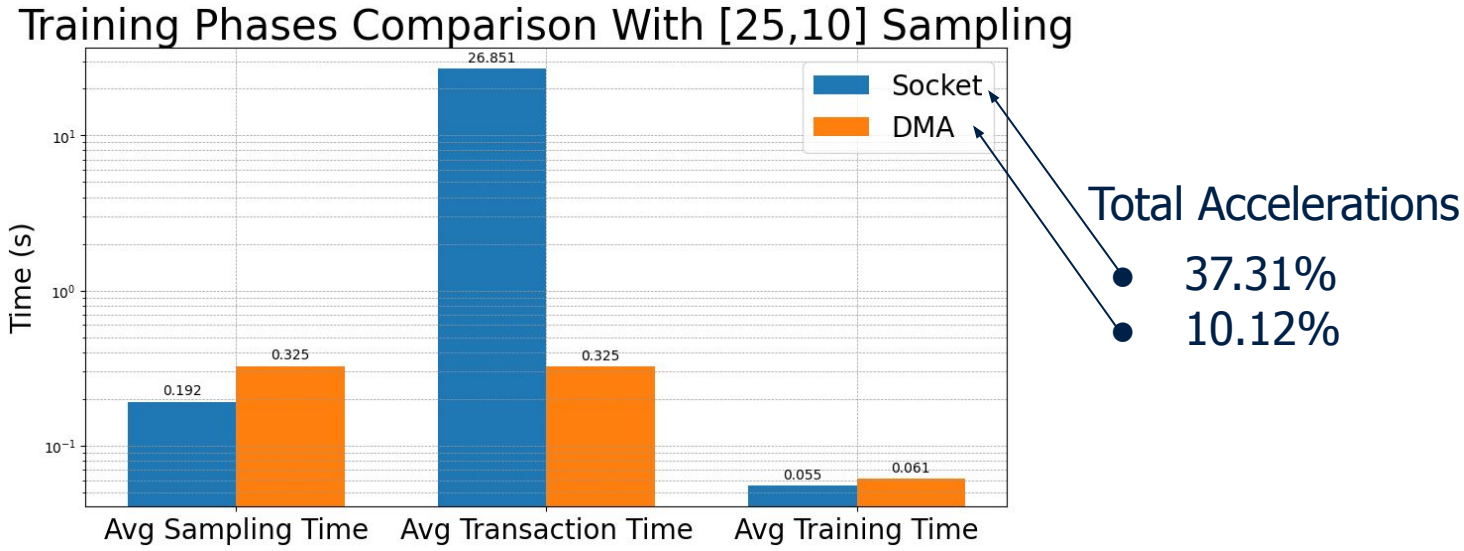


Hewlett Packard Labs

TECHNION Israel Institute of Technology

SmartNIC Sampling - Network Bottleneck Alleviation

- Sampling the mini-batches before the transfer can significantly reduce transfer time.

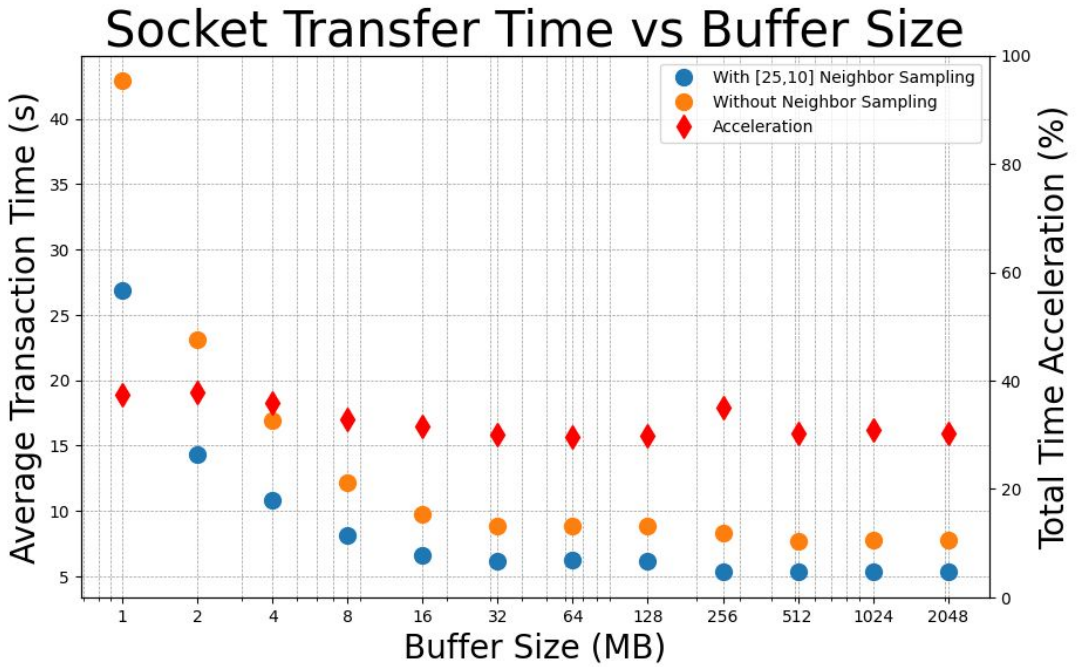


Hewlett Packard Labs

TECHNION Israel Institute of Technology

Large Buffer Sizes

- The problem occurs for large buffer sizes as well.



Hewlett Packard Labs

TECHNION Israel Institute of Technology

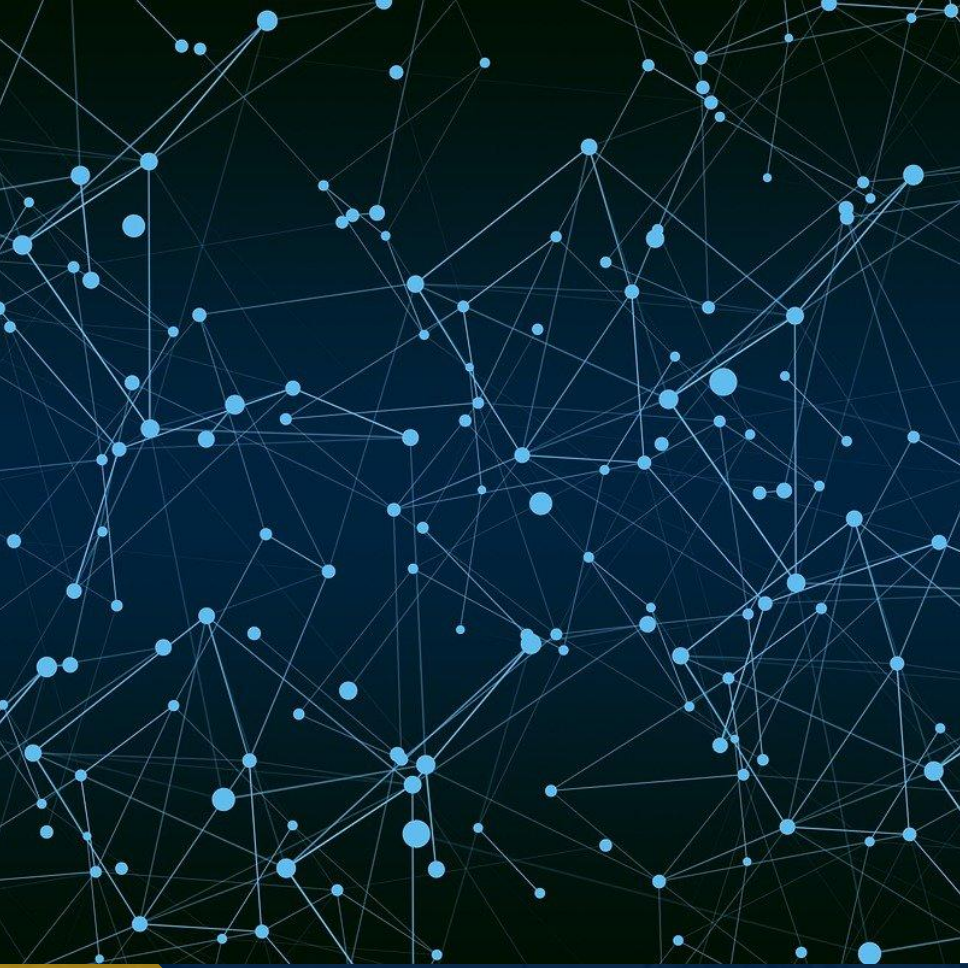
Conclusions and Next Steps

Conclusions:

- Total training time acceleration with our architecture & synchronization mechanism.
- Sampling time could be a bottleneck as well and has to be accelerated too.

Next Steps:

- Measuring different workloads.
- Investigating the characteristics of worth-accelerating workloads.
- Measuring the results using DOCA-RDMA.



Thank you!

liadgerstman@campus.technion.ac.il

www.linkedin.com/in/liad-gerstman-421643221

Technion, Haifa, Israel

Hewlett Packard Labs

TECHNION Israel Institute of Technology