NEST: Network- and Memory- Aware Device Placement for Distributed Deep Learning

Irene Wang, Vishnu Venkata, Arvind Krishnamurthy, Divya Mahajan

SYSTEMS FOR SCALABLE INTELLIGENCE

Georgia Tech

UNIVERSITY of WASHINGTON

Problem: Efficient distributed training is bottlenecked by frameworks that rely on simplified, topology-agnostic search and handle memory constraints post-hoc, leading to excessive synchronization, over-sharding, and poor scalability on realistic, hierarchical datacenter networks.

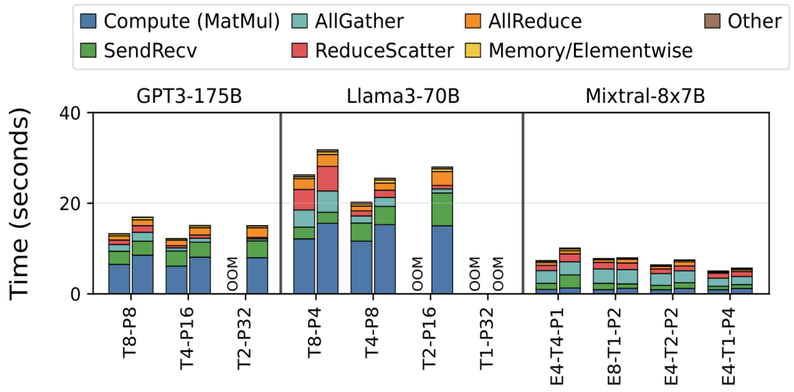
Our Approach: We introduce NEST, a network-, compute-, and memory-aware framework that incrementally models communication and memory within a structured dynamic program to jointly optimize hybrid parallelism and device placement with provable optimality on realistic datacenter topologies.

Findings: NEST demonstrates that explicitly modeling network hierarchy and memory constraints enables superior placement strategies, achieving up to **2.4×** higher throughput than state-of-the-art baselines and sustaining performance scalability on clusters with over 1,000 GPUs.

Paper Link

How to Train DNN Efficiently?

Today's AI models are trained at massive scales. Communication can become the bottleneck.

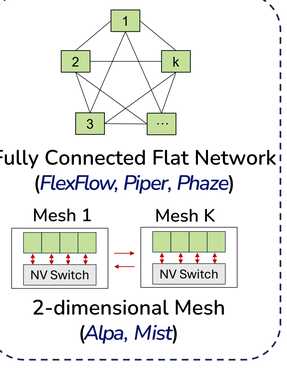


Optimal partitioning and placement requires being aware of the underlying network.

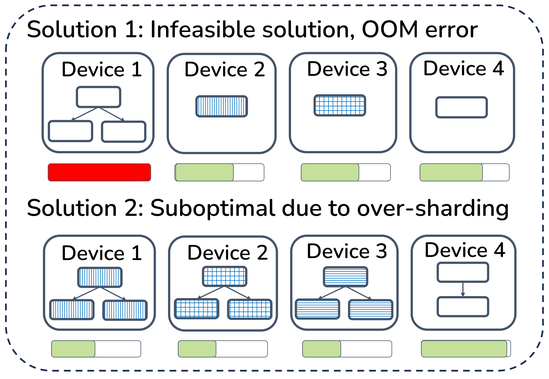
Network- & Memory- Awareness

Prior works fail to consider Network and Memory realities at the same time and lead to suboptimal placement plans

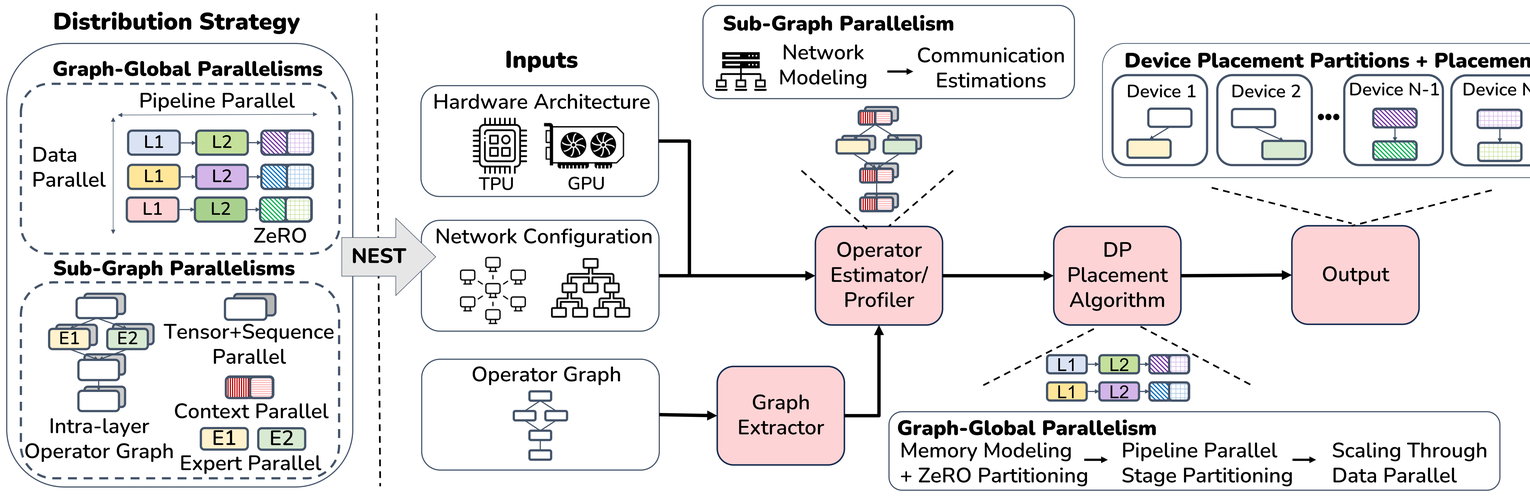
Oversimplify Networks



Post-Placement Memory Check



The NEST Framework



Key Insights

- Topology-adaptive partitioning **balances latencies** across uneven network links
- Explicit memory modeling enables **early bottleneck detection** and scalable training
- Template-based parallelism **matches fine-grained sharding** on repetitive Transformer architectures
- **Validated** to exceed profiling-based techniques **on real-world clusters**

Results

NEST 's distribution strategy achieves **2.4x** higher throughput than state-of-the-art baseline Alpa