

Network Bandwidth and Learning Speed

Network Bandwidth: Critical for multi-node distributed training

Single Node

Fast

8 GPUs with NVLink/PCIe

Direct GPU communication

Multi-Node

Moderate

Ethernet: 10-100 Gbps

InfiniBand: 200-400 Gbps

Network dependent

Communication Bottleneck Example

Increases with more nodes and smaller batches

Gradient sync time > Computation time = Bottleneck

Gradient Compression

Reduce data to transfer

Reduce Frequency

Communicate less often

ZeRO (DeepSpeed)

Partition optimizer states

⚡ Design Rule

Network bandwidth should match GPU compute capability

NVLink

600 GB/s

GPU-to-GPU direct

High-end servers

PCIe 4.0 x16

32 GB/s

Most common setup

Standard

PCIe 5.0 x16

64 GB/s

Latest generation

Emerging