

大模型系列之AI集群

基本介绍



ZOMI



大模型 + AI系统全栈架构



提供大模型算法

提升 AI 集群整体利用率

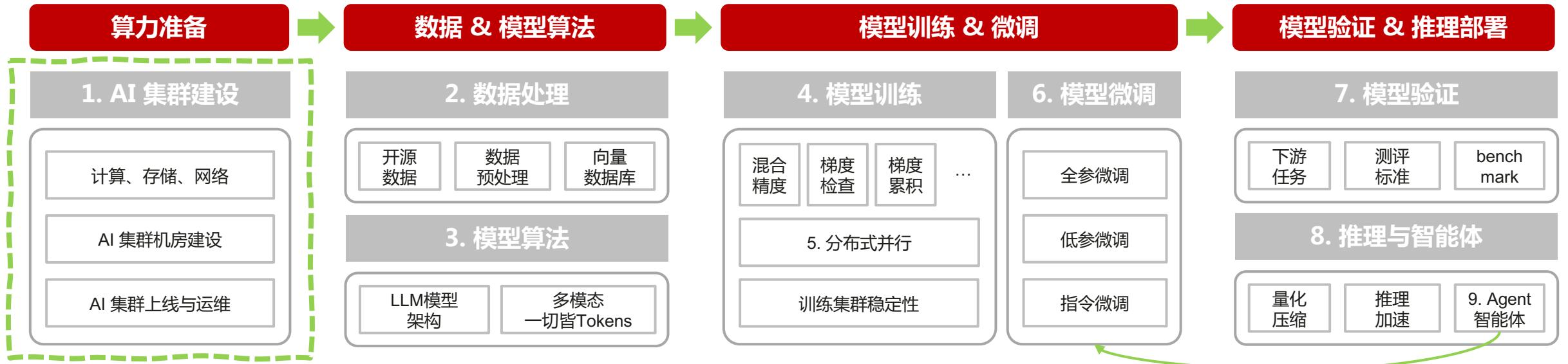
通过软硬件协同优化

基于硬件的编译、框架、使能层

硬件体系结构不仅需要算力

还包括网络、存储组成的超计节点

大模型业务全流程



大模型不仅需要 LLM 算法，同时需要提供
AI 集群、海量数据、分布式并行、推理部署等 AI 系统全栈软硬件协同优化

关于本内容

I. 内容背景

- **AI 集群 + 大模型** : AI集群服务器形态 – 集群训练的指标

2. 具体内容

- **AI 集群服务器架构** : 参数服务器模式 – 同步与异步并行
- **AI 集群通讯方式** : 通信硬件实现 - 集群组网 – 集群软件通信 - 通信实现方式
- **分布式通讯原语** : 通信源语 - 点对点通信 – 集合通信
- **分布式存储系统** : 大模型权重存储方式 – 多级存储系统
- **AI 集群回顾** : NVIDIA 与 TPU 超级计算节点POD

1. 大模型训练

需要AI集群

人工智能发展与大模型关系

Artificial Intelligence

Early artificial intelligence
stris excitement

Machine Learning

Machine learning begins
to flourish

Deep Learning

Deep learning
breakthroughs
drive AI boom

Foundation Models

General pre
training model

1950's

1960's

1970's

1980's

1990's

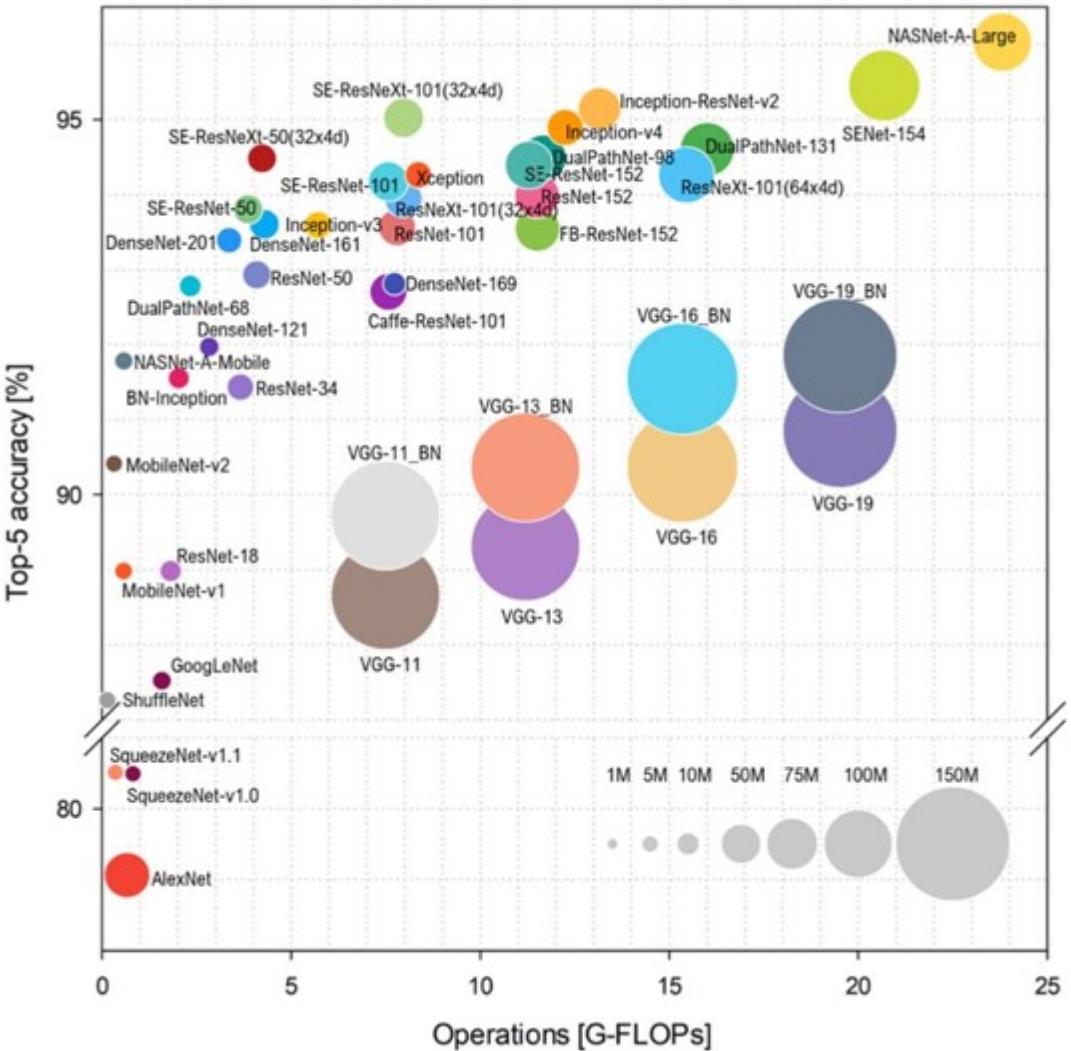
2000's

2010's

2020's

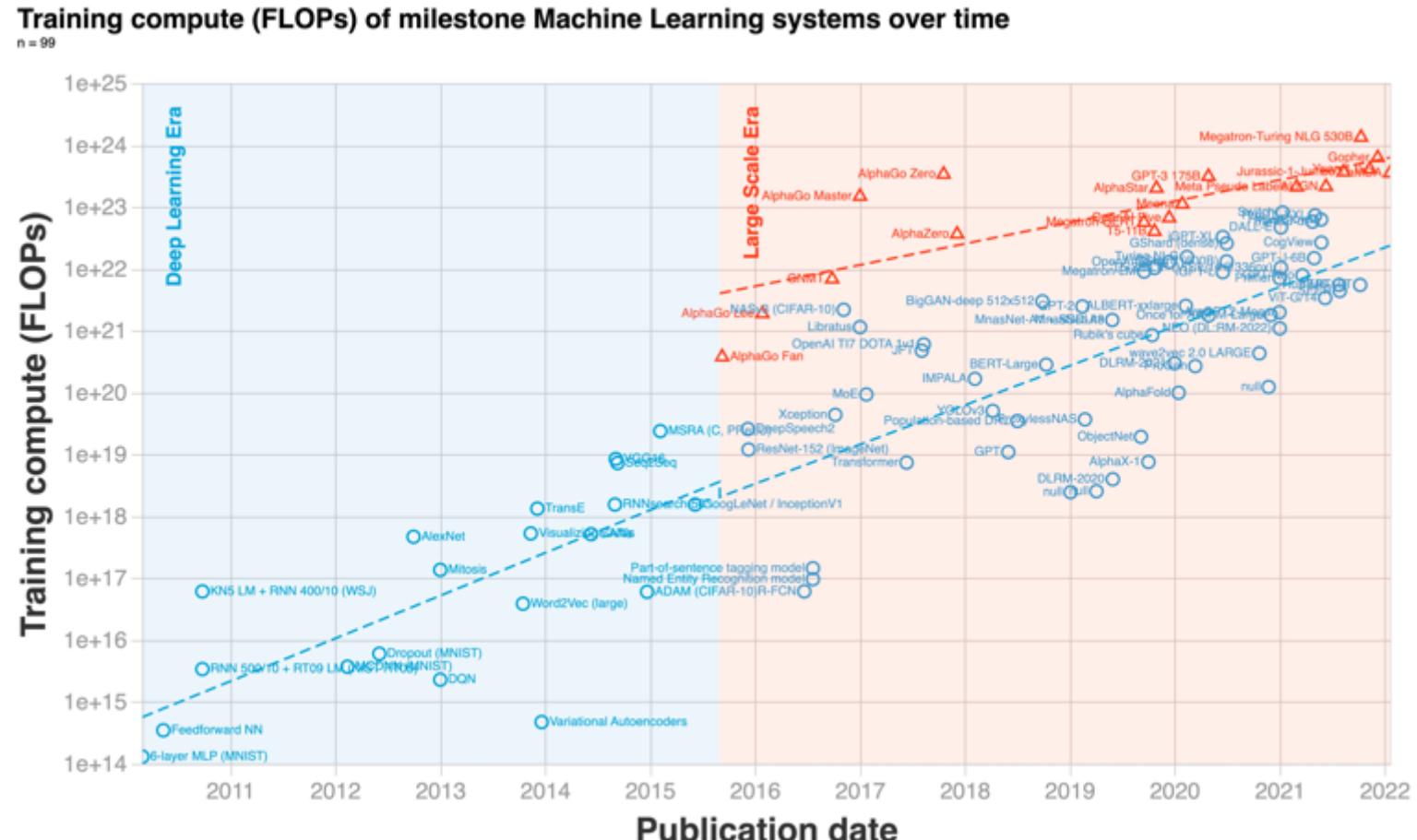
深度学习 + 小模型

1. 容易训练、快速部署应用
 2. 创新优化快，学术和工业界紧密结合
 3. 烟囱式发展，一个模型一个任务
 4. 需要投入大量算法人员进行优化
- e.g. 视觉模型：
 - MobileNet > 轻量化结构
 - YOLO/SSD > 检测与分割



深度学习 + 大模型 (Foundation Models)

1. 自监督学习方法，可以减少数据标注，降低训练研发成本
2. 模型参数规模越大，有望进一步突破现有模型结构的精度局限
3. 解决模型碎片化，提供预训练方案
 - e.g. 语言模型 GPT-3
 - 8 张 V100，训练时长 36 年
 - 512 张 V100，训练近 7 个月



大模型训练遇到的挑战

内存

- 175B参数模型，参数放内存占用700GB内存（HBM），训练过程会增加7~8倍内存开销~5600GB内存

性能

不仅要求AI芯片的计算性能足够强悍，同时也依赖于AI框架的大规模分布式训练的运行和调度效率

通讯

大模型参数量过多，需要对模型进行切分（模型并行/数据并行），因此引入了通信问题。

调优

在数千节点的集群上，要保证计算的正确性/性能/可用性

大模型训练考验的是算法、数据、框架、资源调度
全栈和全流程的综合能力

AI 集群：解决内存、通信、性能问题



2. AI集群

服务器形体



JEWELLERS
DEALS IN: GOLD & SILVER ORNAMENTS
91.98721-38531

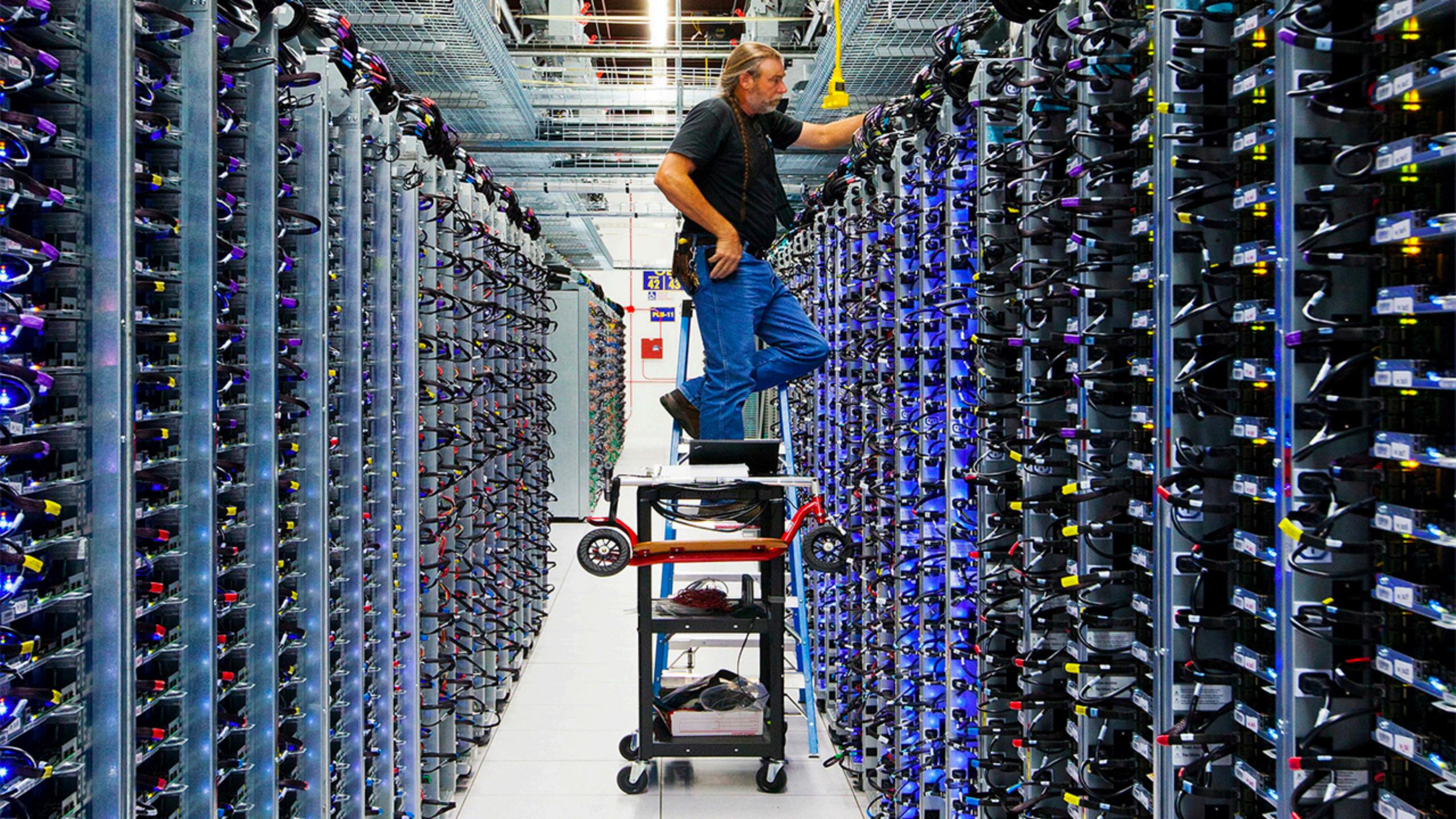
MALHOTRA
MARKET

मल्होत्रा
भारवीट

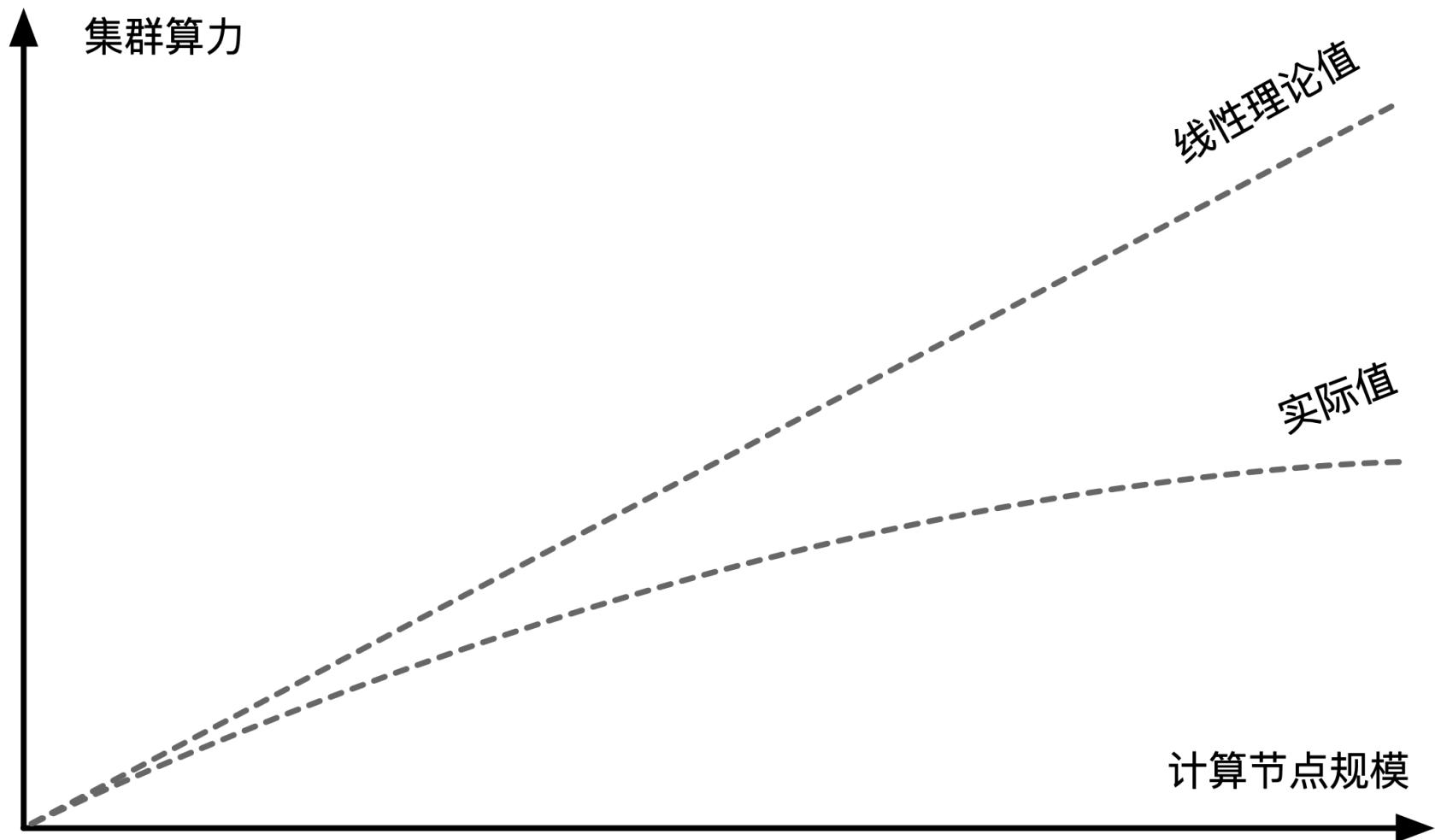
←
Mahavir
Packaging Co.
Pouches, Boxes,
Set Boxes, Silver
Set & Set Boxes etc.

JEWI

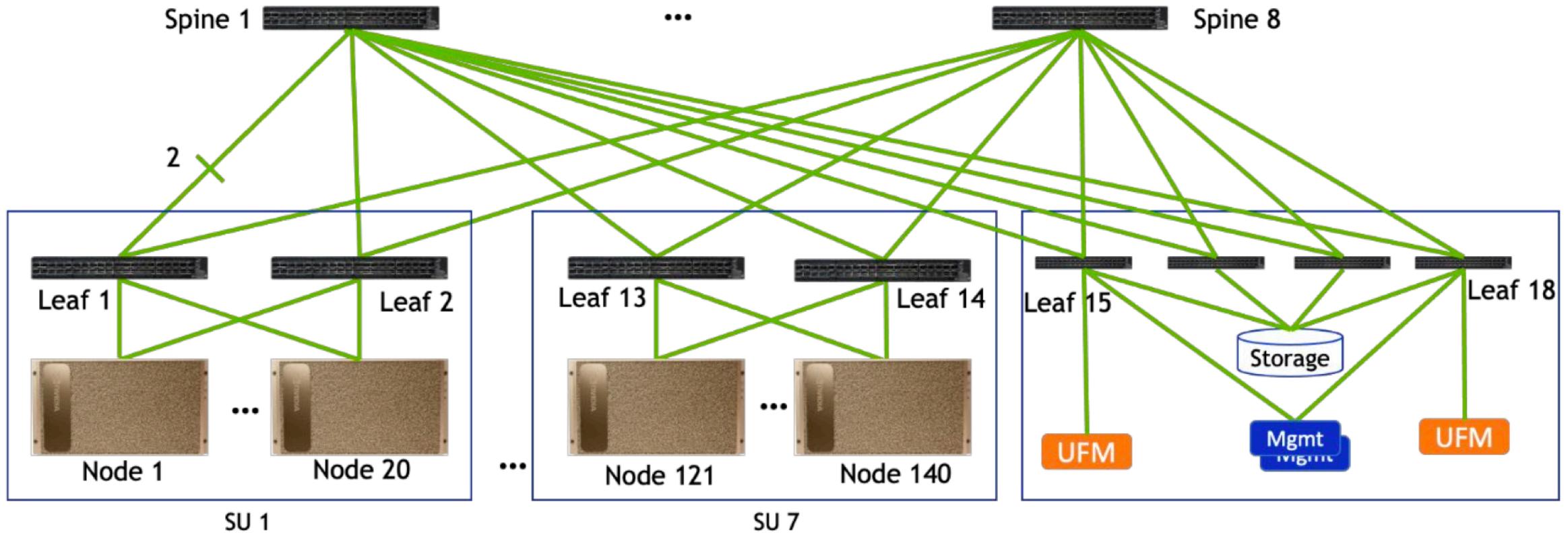




AI集群网络和算力协同

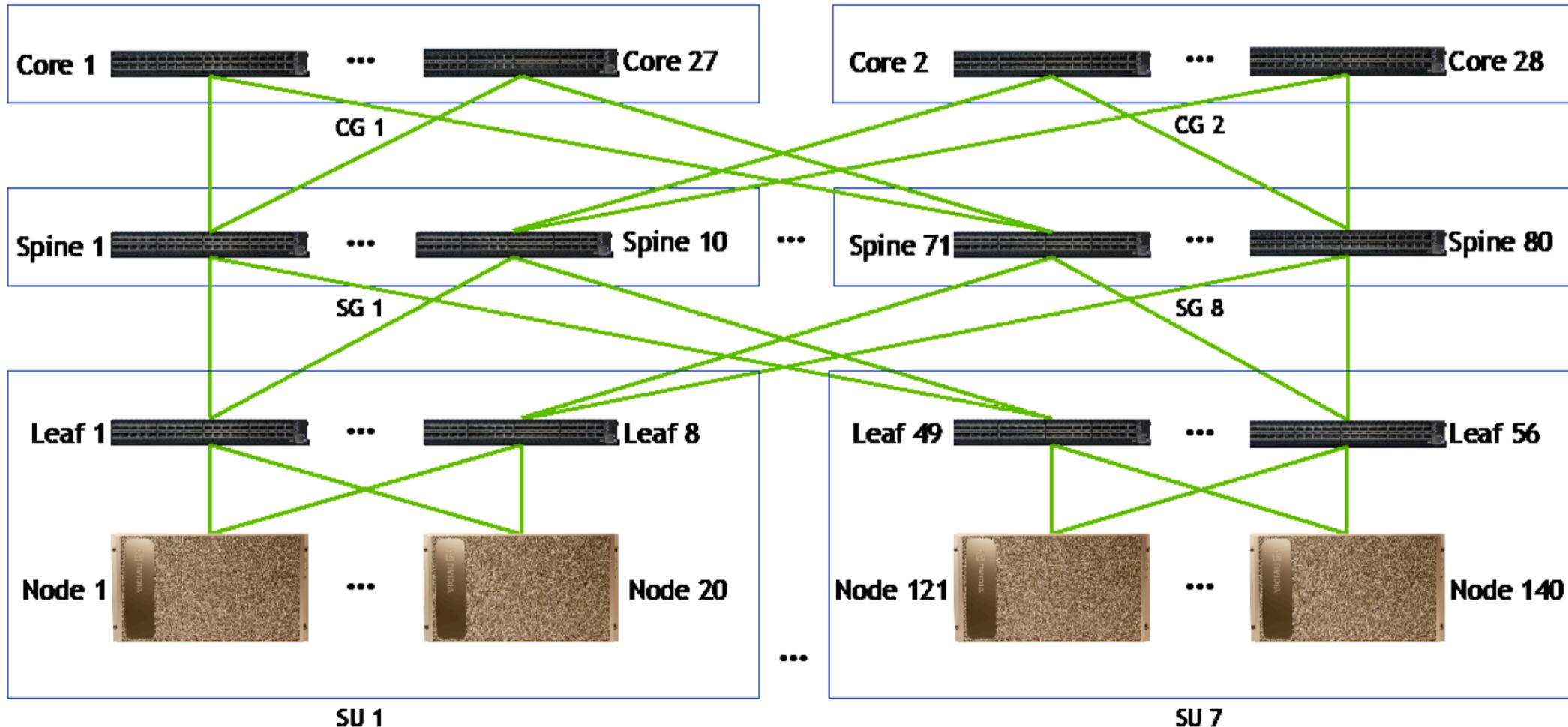


集群服务器架构



集群服务器架构

Figure 4. Compute fabric topology for a 140-node DGX SuperPOD



3. AI集群

训练关键指标

大模型训练遇到的挑战

内存

- 175B参数模型，参数放内存占用700GB内存（HBM），训练过程会增加7~8倍内存开销~5600GB内存

通讯

- 大模型参数量过多，需要对模型进行切分（模型并行/数据并行），因此引入了通信问题。

性能

- 不仅要求AI芯片的计算性能足够强悍，同时也依赖于AI框架的大规模分布式训练的运行和调度效率

• TTA (Time To Accuracy)

减少训练时间

• 计算速率 (Computation rate)

加快计算速率

大模型的训练指标

- **减少训练时间 TTA (Time To Accuracy) :**

$$\text{训练时间} = \text{训练数据规模} \times \text{单步计算量} / \text{计算速率}$$


模型相关，相对固定 可变因素

- **End-to-End Training Time (E2ETT) :**

$$E2ETT \approx 8 \times T \times P / n X$$

《Efficient Large-Scale Language Model Training on GPU Clusters Using Megatron-LM》

大模型的训练指标

- 减少训练时间 TTA (Time To Accuracy) :

$$\text{训练时间} = \text{训练数据规模} \times \frac{\text{单步计算量}}{\text{计算速率}}$$

模型相关，相对固定 可变因素

- 加快计算速率 (Computation rate) :

$$\text{计算速率} = \text{单设备计算速率} \times \text{设备数} \times \text{多设备并行效率 (加速比)}$$

Moore定律+算法优化 可变因素

加快大模型的训练速率

$$\text{训练时间} = \text{训练数据规模} \times \frac{\text{单步计算量}}{\text{计算速率}}$$

模型相关，相对固定

可变因素

- **加快计算速率 (Computation rate) :**

$$\text{计算速率} = \text{单设备计算速率} \times \text{设备数} \times \text{多设备并行效率 (加速比)}$$

混合精度

算子融合

梯度累积

服务器架构

通信拓扑优化

存储系统优化

数据并行

张量并行

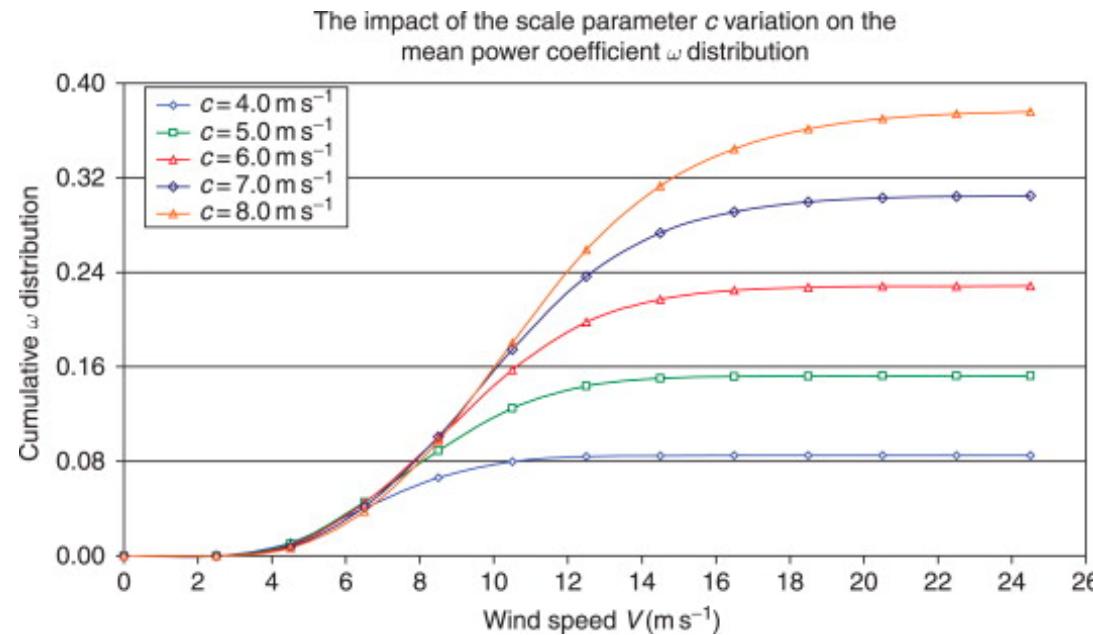
流水并行

AI 集群加速比

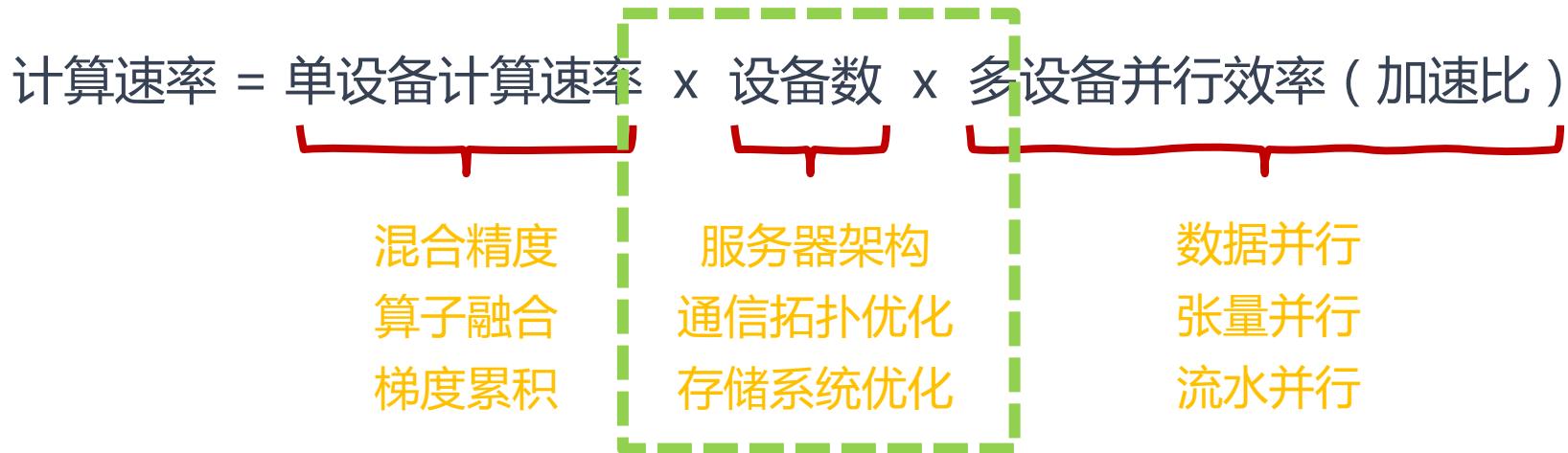
假设单设备吞吐量为 T ， n 个设备系统吞吐量为 $n \times T$ ，系统实际达到吞吐量为 T_n ，则加速比为：

$$Scale Factor = T_n / n \times T$$

边际效应受限



加快大模型的训练速率



集群加速比

- 尽可能提升 AI 集群加速比，减少边际效应带来的影响。

集群利用率

- 提升集群的利用率，降低因为通信、存储所引起的时延，让时间都在计算。

小结

1. 了解从小模型到大模型，大模型带来的优点
2. 了解了大模型训练需要AI集群，去降低模型训练时间
3. 了解通过集群架构、通信、存储来提升AI集群加速比和利用率





Thank you

把AI系统带入每个开发者、每个家庭、
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and
organization for a fully connected,
intelligent world.

Copyright © 2023 XXX Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. XXX may change the information at any time without notice.



Course chenzomi12.github.io

GitHub github.com/chenzomi12/DeepLearningSystem