大模型系列 - 集合通信

ZOMI

集合通信原语

大模型业务全流程

数据 & 模型算法

2. 数据处理

- 开源数据
- 数据预处理
- 向量数据库

3. 模型算法

- LLM模型架构
- 多模态一切皆Tokens

模型训练 & 微调

4. 模型训练

- 混合精度
- 梯度检查
- 梯度累积
- …

5. 分布式并行

训练集群稳定性

6. 模型微调

- 全参微调
- 低参微调
- 指令微调

模型验证 & 推理部署

7. 模型验证

- 下游任务
- 测评标准
- bench mark

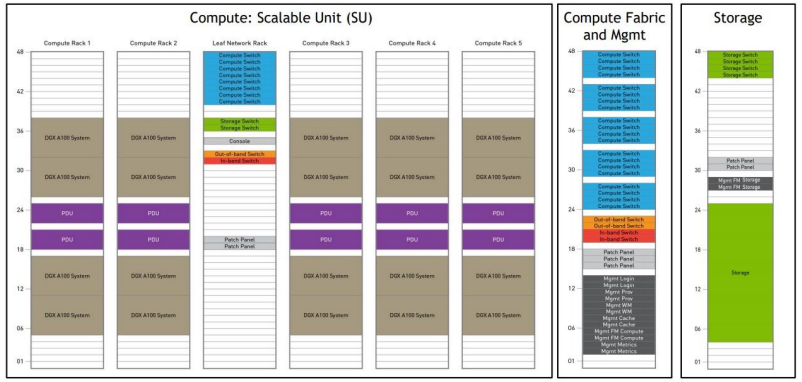
8. 推理与智能体

- 量化压缩
- 推理加速
- 9. Agent 智能体

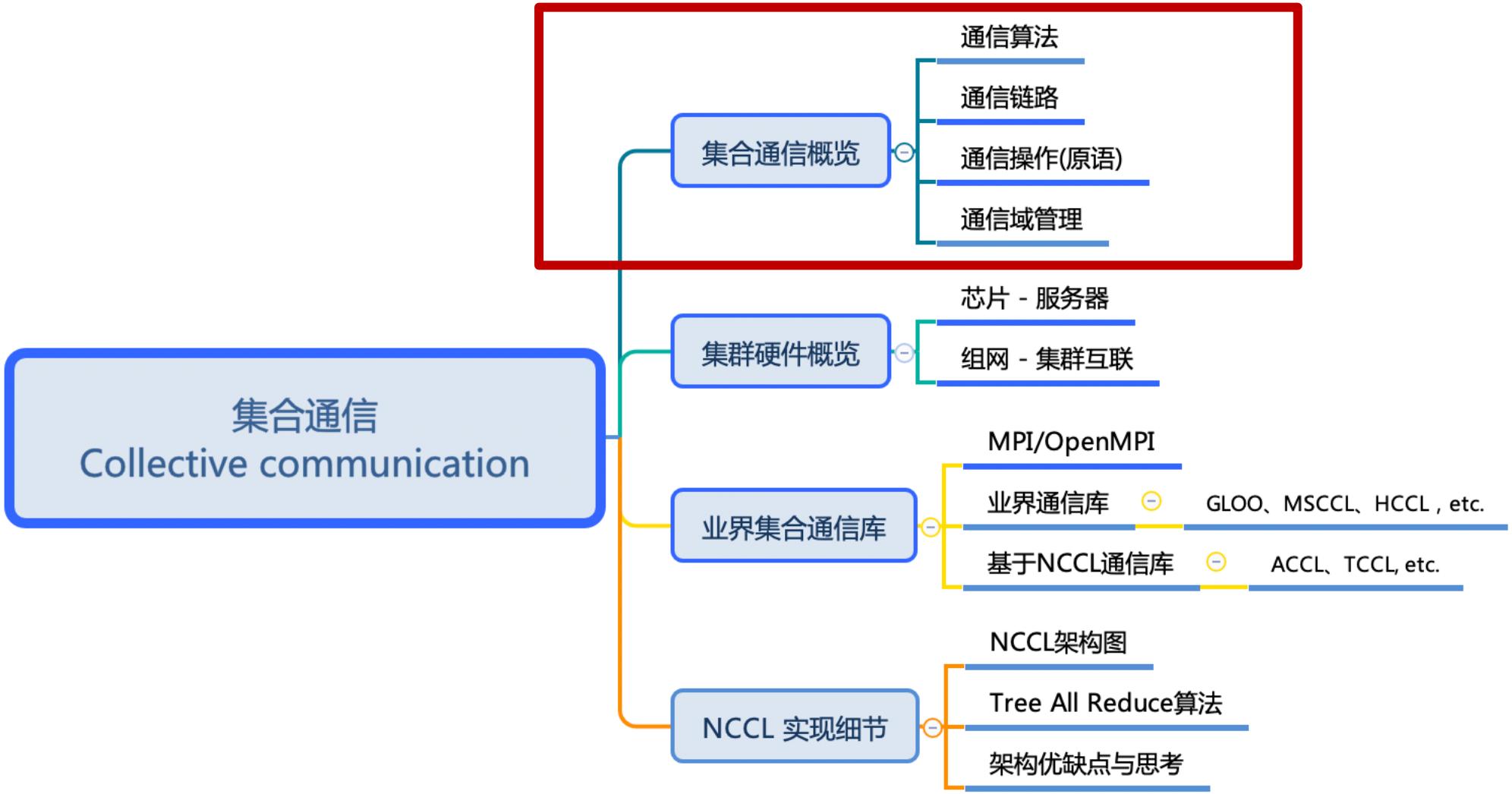
集群算力准备

1. AI 集群建设

- 计算、存储、网络
- AI 集群机房建设
- AI 集群上线与运维



思维导图 XMind


集合通信
Collective communication
集合通信概览
通信算法
通信链路
通信操作(原语)
通信域管理
集群硬件概览
芯片 - 服务器
组网 - 集群互联
业界集合通信库
MPI/OpenMPI
业界通信库
GLOO、MSCCL、HCCL，etc.
基于NCCL通信库
ACCL、TCCL, etc.
NCCL 实现细节
NCCL架构图
Tree All Reduce算法
架构优缺点与思考

集合通信概览

XCCL: XXXX Collective Communication Library

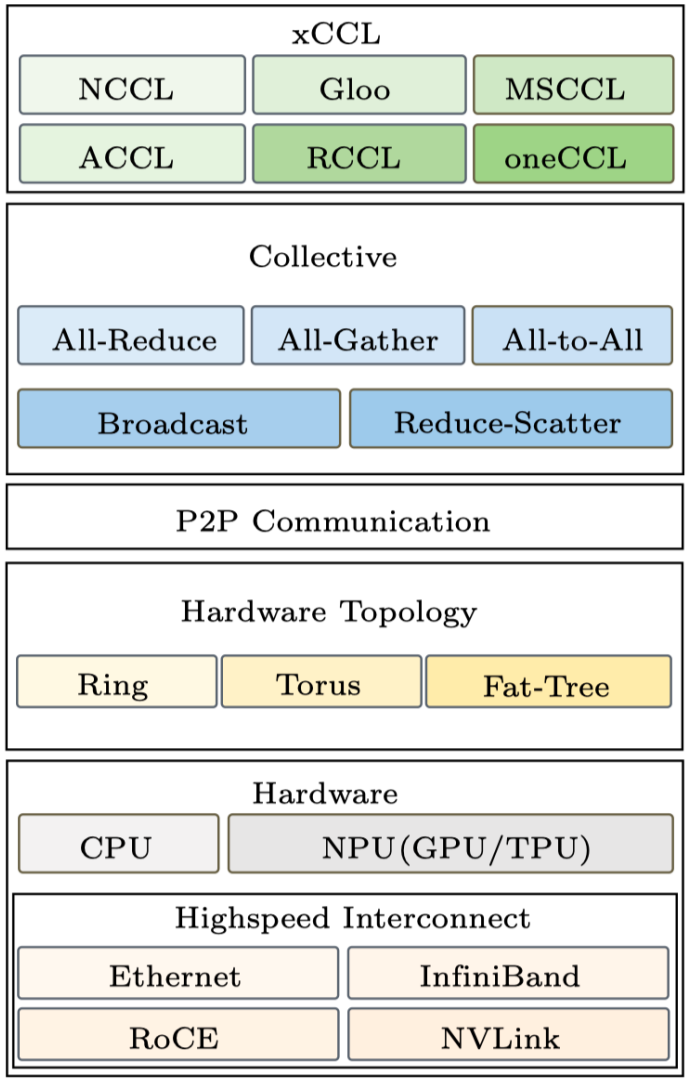
1. AI 与通信关系（AI 基础知识、训练推理、分布式并行）
2. XCCL 基本架构（HPC 通信架构 to XCCL 通信架构）
3. 集合通信原语（All Reduce, etc.）
4. 集合网络拓扑（Hypercube 、Ring 、Torus 、Fat-Tree 、Dragonfly & Dragonfly+）
5. PyTorch 集合通信与计算并行

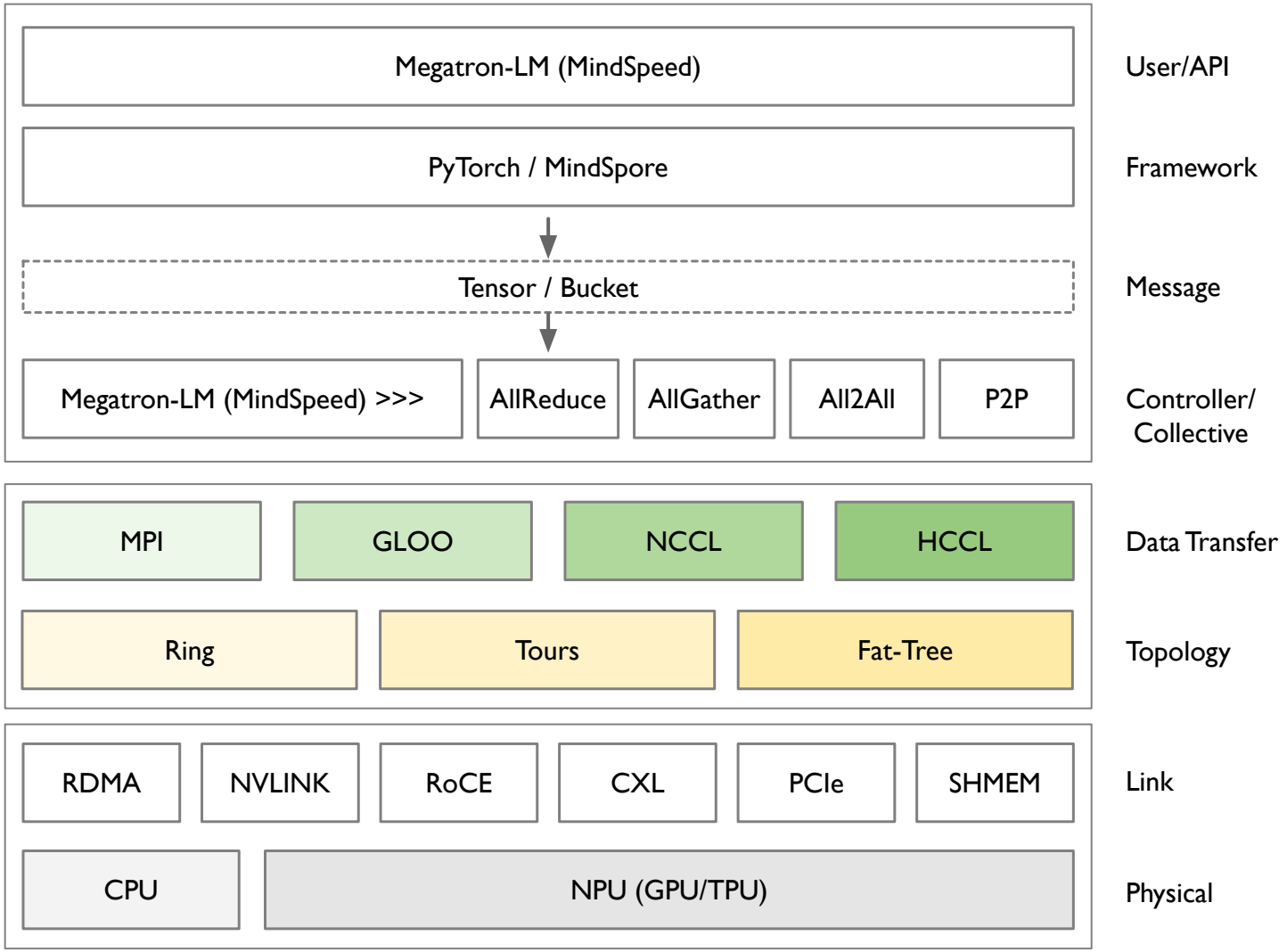
集合通信概览：基本概念

通信特性	HCCL	NCCL
通信算法	ring/mesh + ring/Hav-Doub/Pair-Wise， etc.	ring + Tree ring， etc.
通信链路	NIC / 灵渠总线 / PCIE	NIC / NVLink / NVSwitch / GPU-Direct / PCIE / CMC
通信操作	allreduce、broadcast、reduce、reduce scatter、allgather、all2all、send、recv	allreduce、broadcast、reduce、reduce scatter、allgather、all2all、send、recv
通信域管理	全局通信域、子通信域、基于全局/子通信域配置算法	全局通信域、子通信域、自定义通信域配置算法

01 AI系统中的位置

XCCL 在 AI 系统中的位置





XCCL 对通信操作支持情况

Table 1. Summary of Collective Support Within Libraries

Collective	Discussed in	MPI Function	Implemented in				
			NCCL	MSCCL	Gloo	oneCCL	ACCL
Barrier	Not discussed	`MPI_BARRIER`	No	No†	Yes	Yes	Unknown
Broadcast	Subsection 2.1	`MPI_BCAST`	Yes	No†	Yes	Yes	Yes
Reduce	Subsection 2.5	`MPI_REDUCE`	Yes	No†	Yes	Yes	Unknown
Gather	Not discussed	`MPI_GATHER`	No	No†	Yes‡	No	Unknown
Scatter	Subsection 2.3	`MPI_SCATTER`	No	No†	Yes‡	No	Unknown
All-Gather	Subsection 2.2	`MPI_ALLGATHER`	Yes	No†	Yes‡	Yes	Yes
All-to-All	Subsection 2.4	`MPI_ALLTOALL`	No	Yes	No	Yes	Unknown
All-Reduce	Subsection 2.6	`MPI_ALLREDUCE`	Yes	Yes	Yes	Yes	Yes
Reduce-Scatter	Subsection 2.7	`MPI_REDUCE_SCATTER`	Yes	No†	Yes	Yes	Yes
Scan	Not discussed	`MPI_SCAN`	No	No†	No	No	Unknown

Note: MSCCL is unique because it allows programmers to implement their own collective routines and algorithms. †: algorithm not provided but can be implemented using DSL or called via NCCL API; ‡: algorithm not supported on all accelerator types.

XCCL 对通信操作支持情况

Table 2. Classic and xCCL's Collective Communication Algorithms

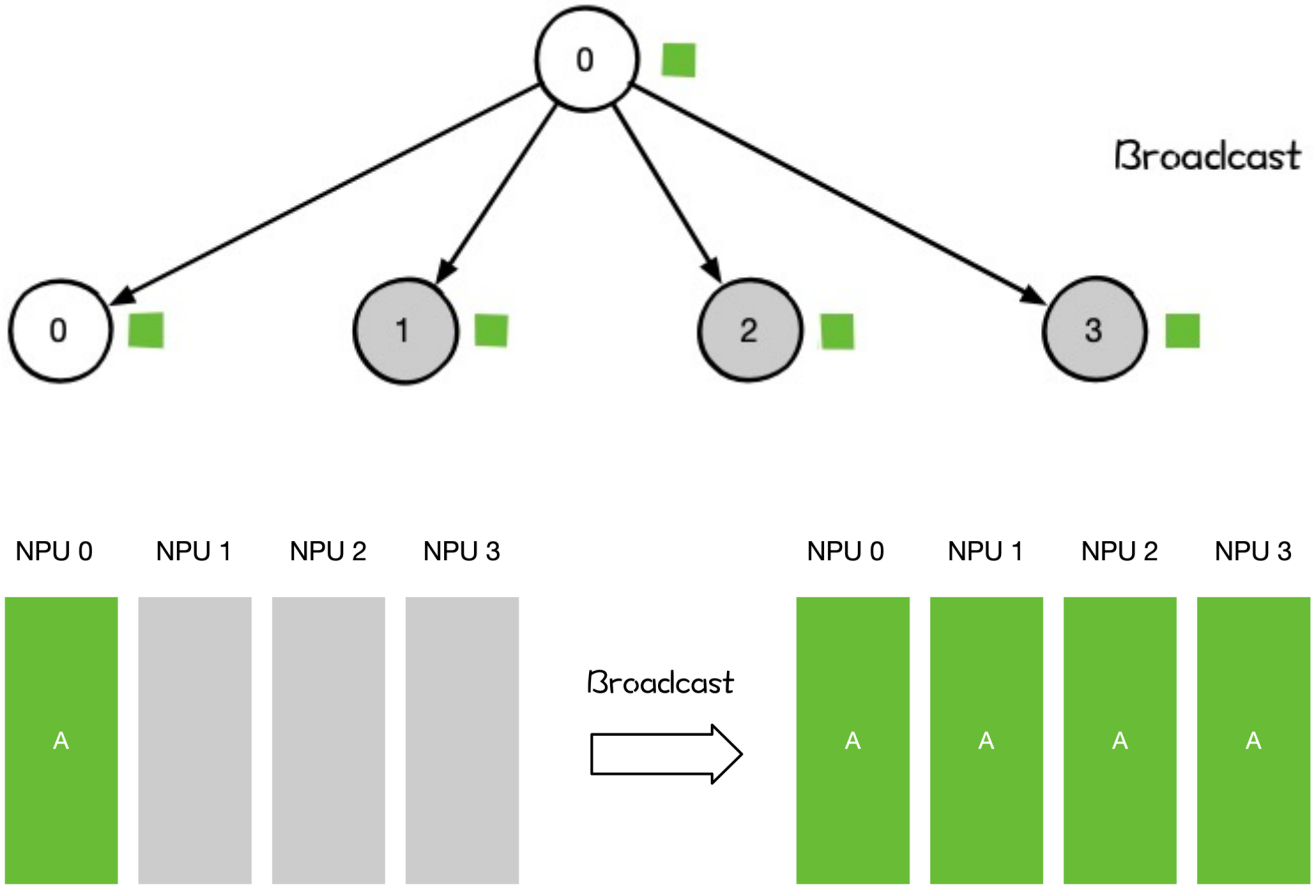
Category	Collective	Algorithm	Description on Suitability (e.g., Message Size, Number of Processes)
Classic	All-to-All	Bruck[42]	Short (e.g., < 32 B)
		Isend-Irecv[43]	Medium (e.g., 32 B to 32 KB)
		Pairwise-Exchange[44]	Long (2^n processes)
	All-Gather	Ring[43]	Long, medium (not 2^n processes)
		Recursive-Doubling[43]	Short, medium (2^n processes)
		Bruck[42]	Short (not 2^n processes)
	Broadcast	Binomial Tree[45]	Short (e.g., < 32 B)
		Van de Geijn[46, 47]	Long (e.g., > 32 KB)
	Reduce-Scatter	Recursive-Halving[44]	Short (commutative reduction)
		Recursive-Doubling[43]	Short (not commutative reduction)
		Pairwise-Exchange[43]	Long (e.g., $\geqslant$ 512 KB for commutative, $\geqslant$ 512 B for noncommutative)
		Binomial Tree and Linear Scatterv[43]	Medium

HCCL 对通信操作支持情况

通信操作/原语

02. 一对多

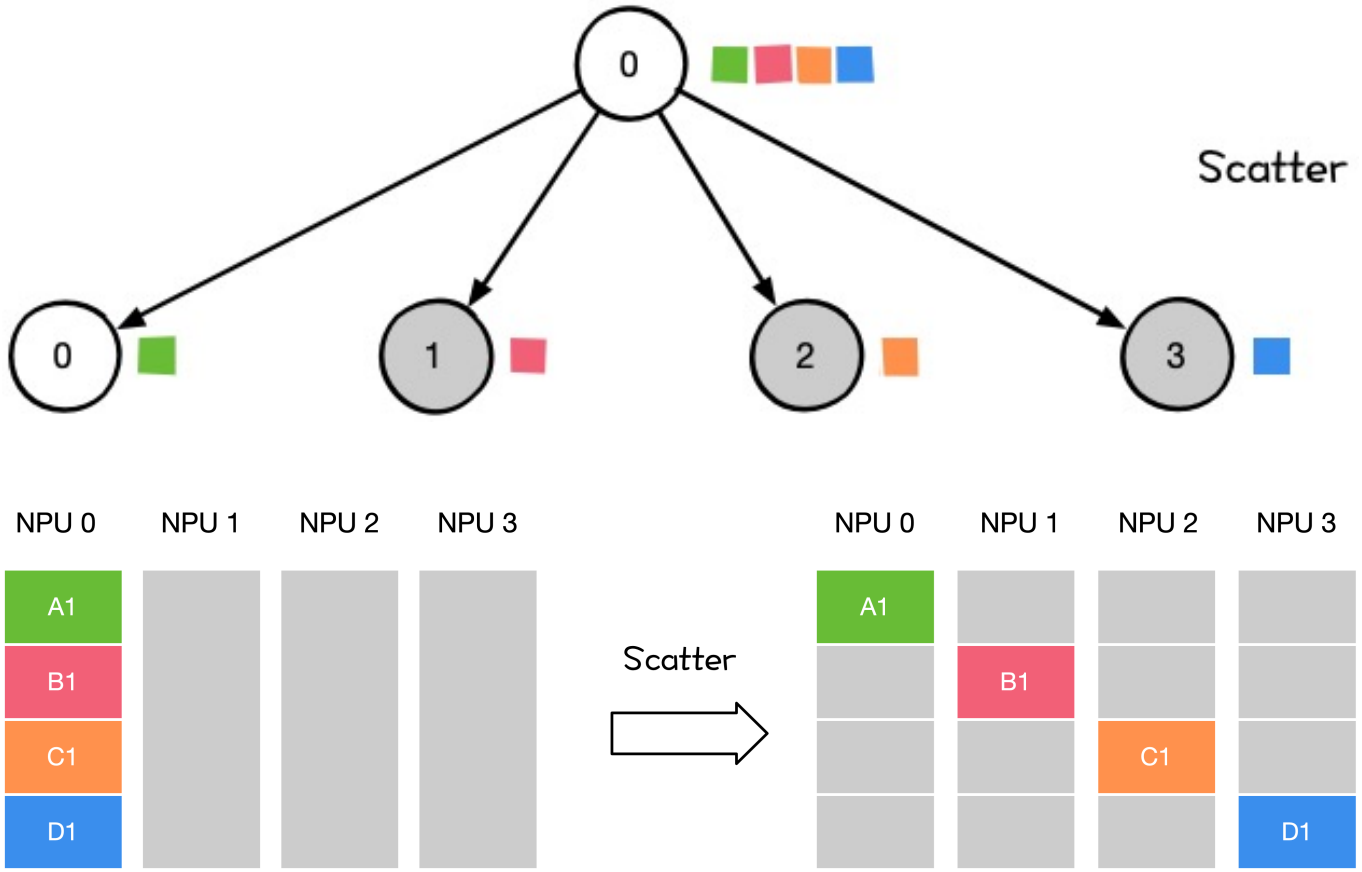
Broadcast



- 单个 rank 把自身的数据发送到集群中的其他 rank。

  - 网络参数 Init Weight 的初始化
  - 数据并行 DP 对数据分发初始化；
  - AllReduce 里的 broadcast + reduce 组合里的 broadcast 操作；

Scatter

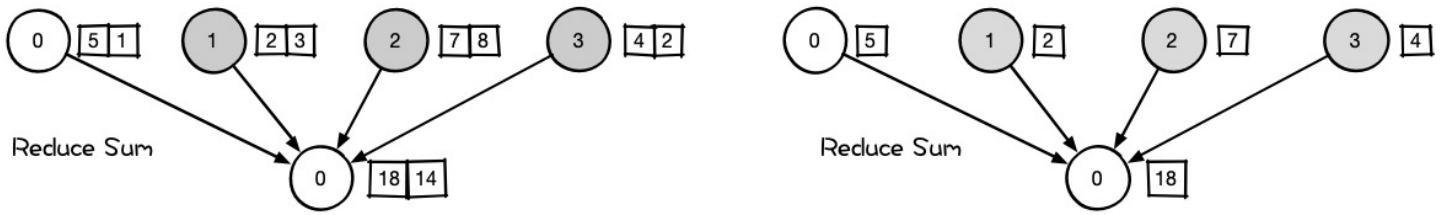


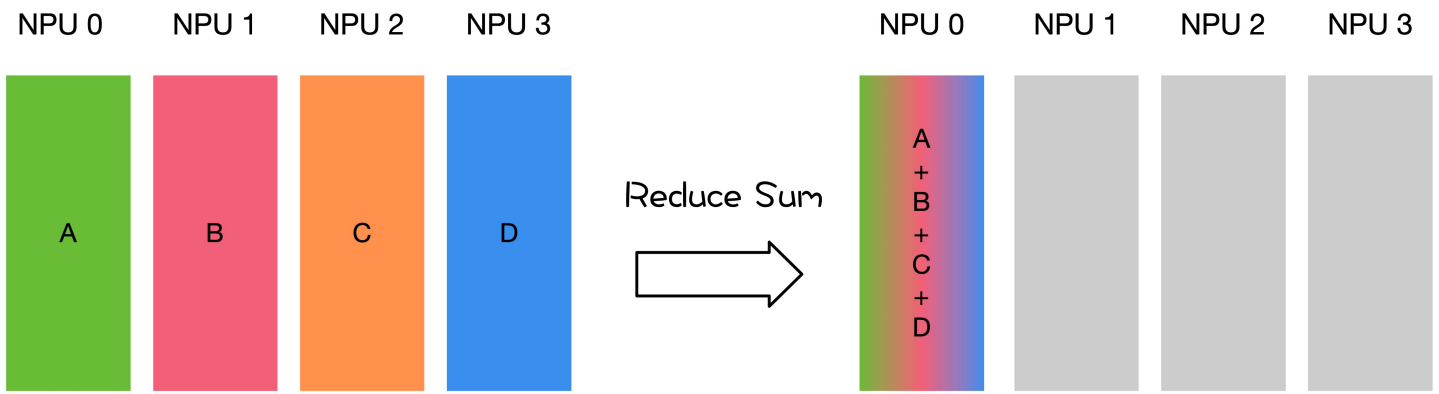
- 将主节点的数据进行划分并散布至其他指定的 Rank。
  - Reduce Scatter 组合里的 Scatter 操作；
  - 流水并行里初始化时将模型 Scatter 到不同 Rank 上；

通信操作/原语

03. 多对一

Reduce

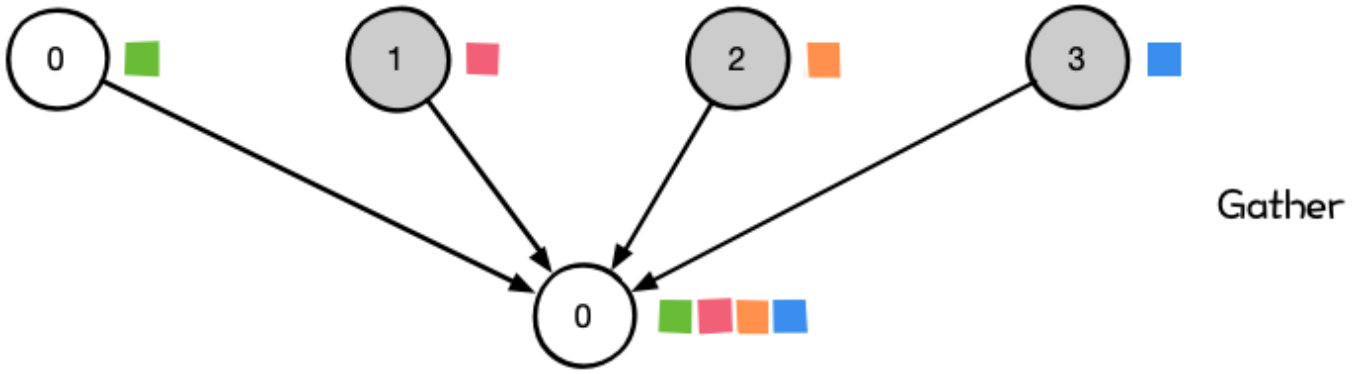


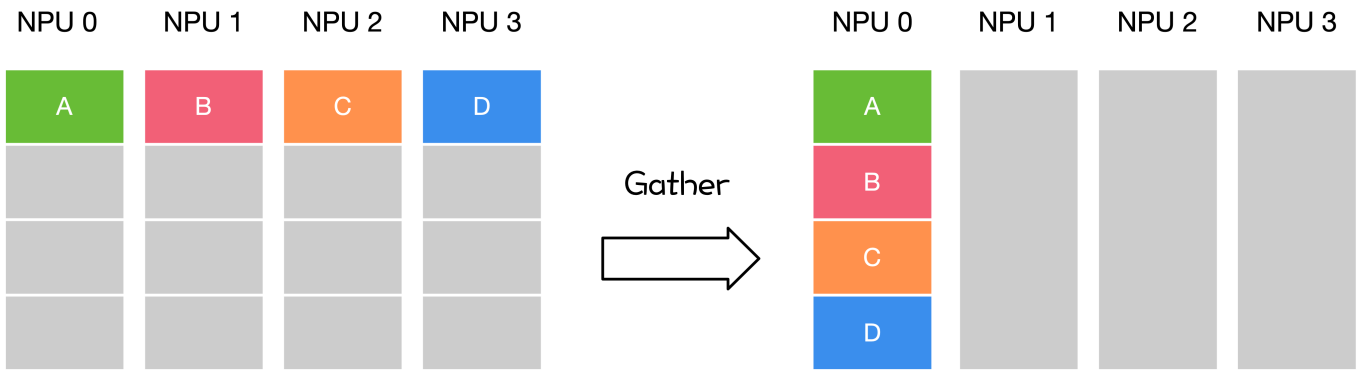


- 把多个 Rank 的数据规约运算到一个 Rank 上。
- 细分可以包括：SUM、MIN、MAX、PROD、LOR 等类型的规约操作。

- AllReduce 里的 Broadcast，Reduce 组合 Reduce 操作；
- Reduce Scatter 组合里的 Reduce操作；
- 大模型训练权重 CKPT 保存；

Gather





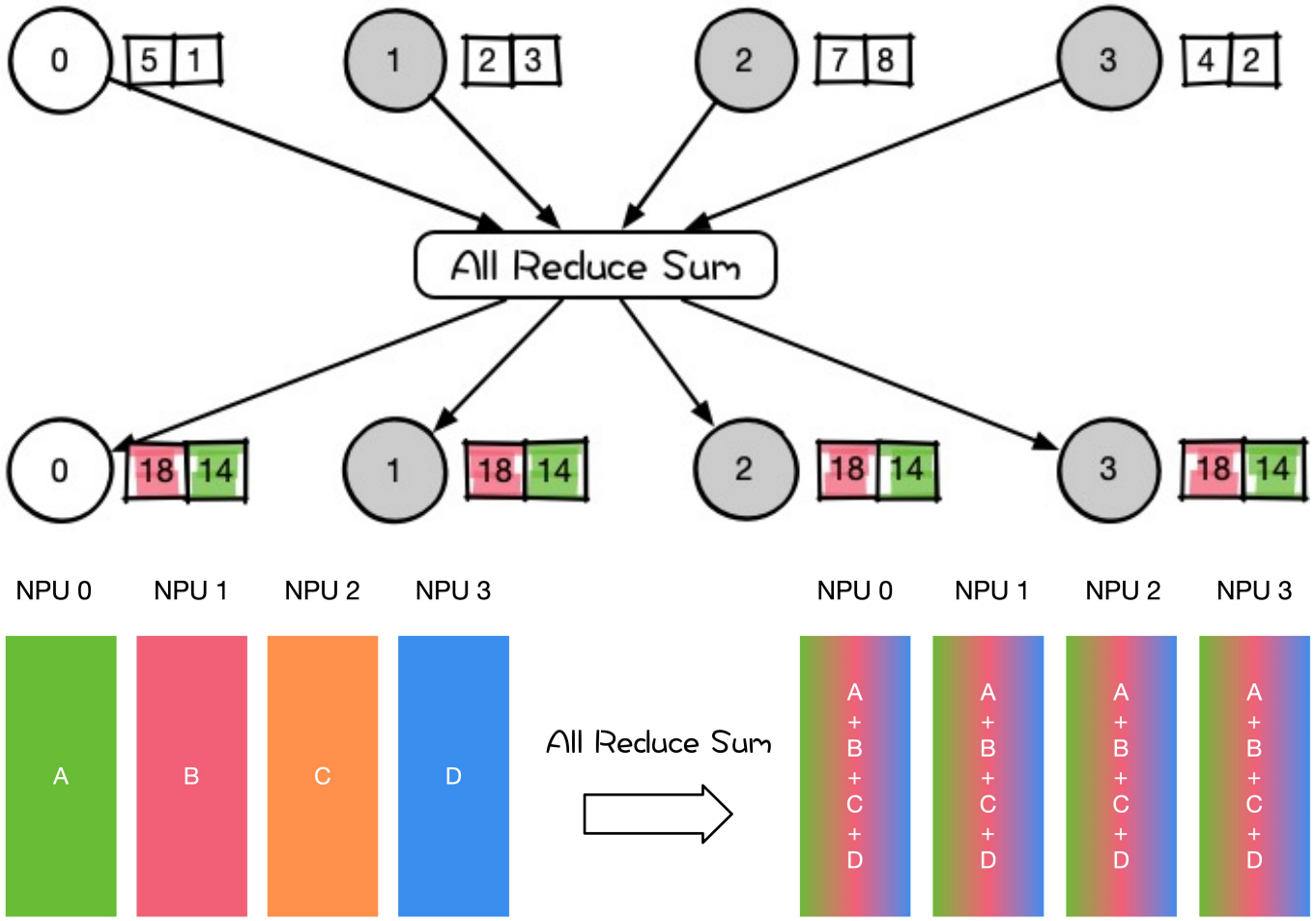
- 将多个 rank 上的数据收集到 rank 上，Gather 可以理解为反向的 Scatter。

  - Gather 相对用得比较少，All Gather 会在张量并行 TP 用得较多；

通信操作/原语

04. 多对多

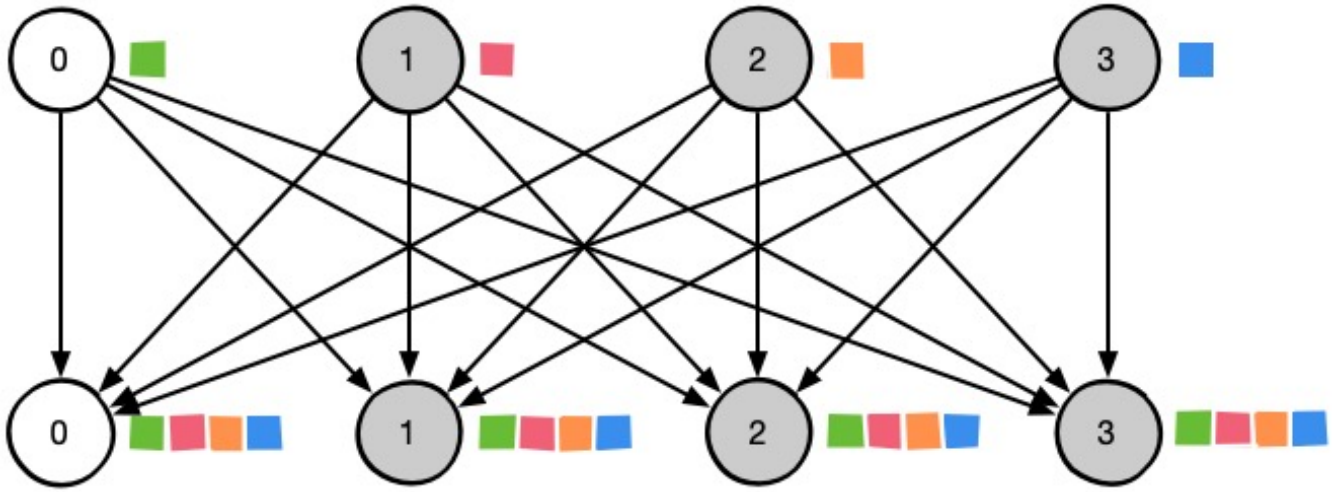
All Reduce

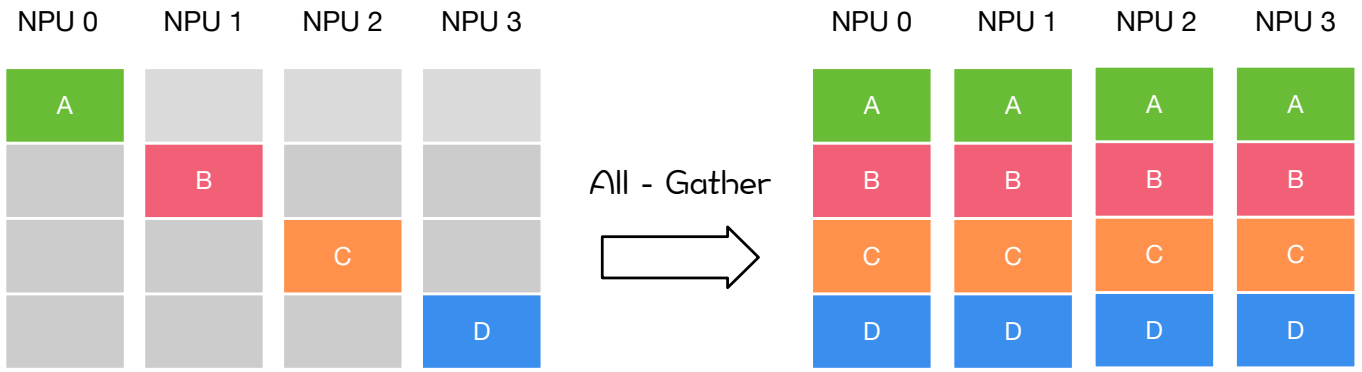


- 在所有 Rank 执行相同 Reduce 操作，将所有 Rank 数据规约运算得到的结果发送到所有 Rank

  - 在专家并行、张量并行、序列并行中大量地使用 All Reduce 对权重和梯度参数进行聚合。
  - 数据并行 DP 各种通信拓扑结构比如Ring AllReduce、Tree AllReduce里的 AllReduce 操作；

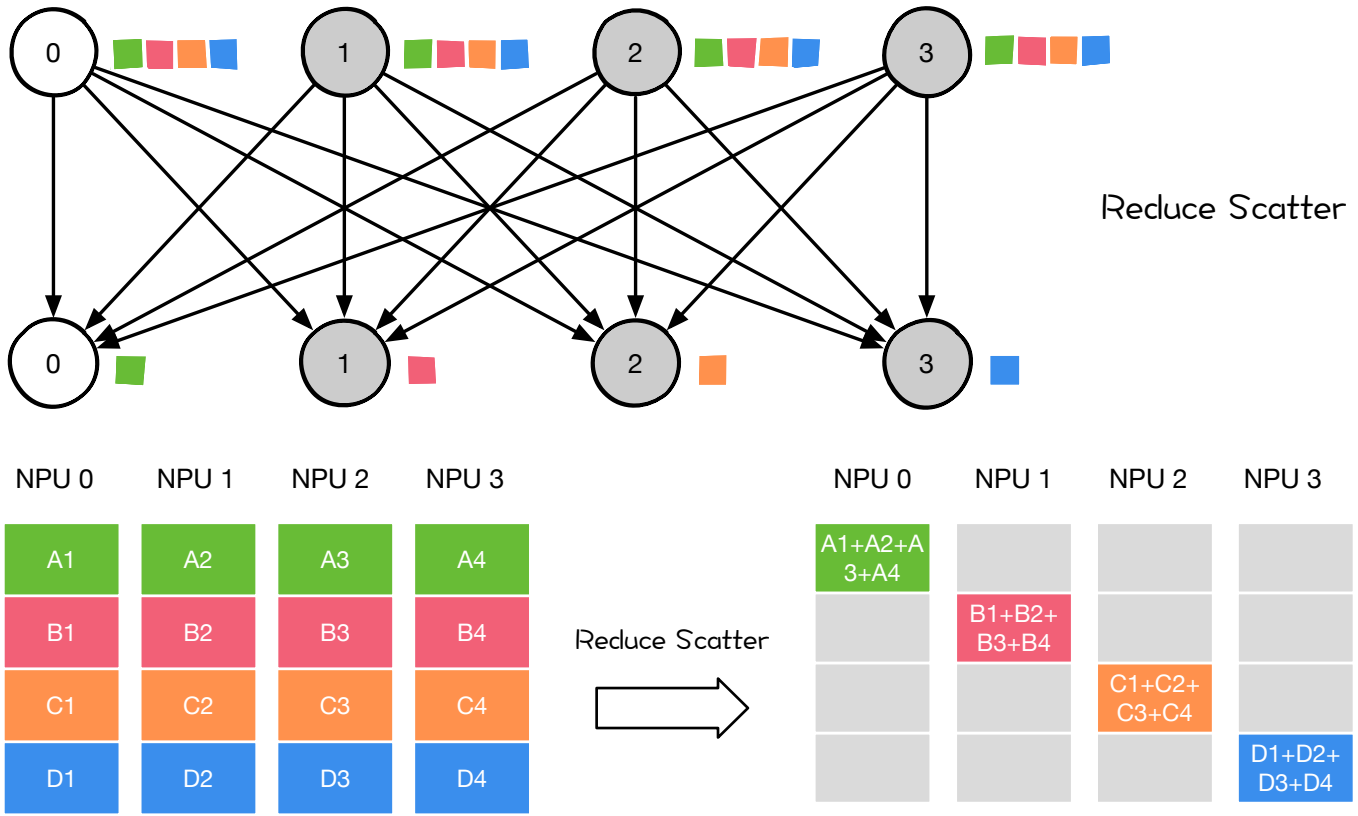
All-Gather





- All Gather 会从所有 rank 收集数据并分发所有 rank 上。 All Gather = Gather + Broadcast。
  - 在专家并行、张量并行、序列并行中大量地使用 All Gather 对权重和梯度参数进行聚合。
  - 模型并行里前向计算里的参数全同步，需要用 All Gather 把模型并行里将切分到不同的 NPU上的参数全同步到一张 NPU 上才能进行前向计算。

Reduce Scatter

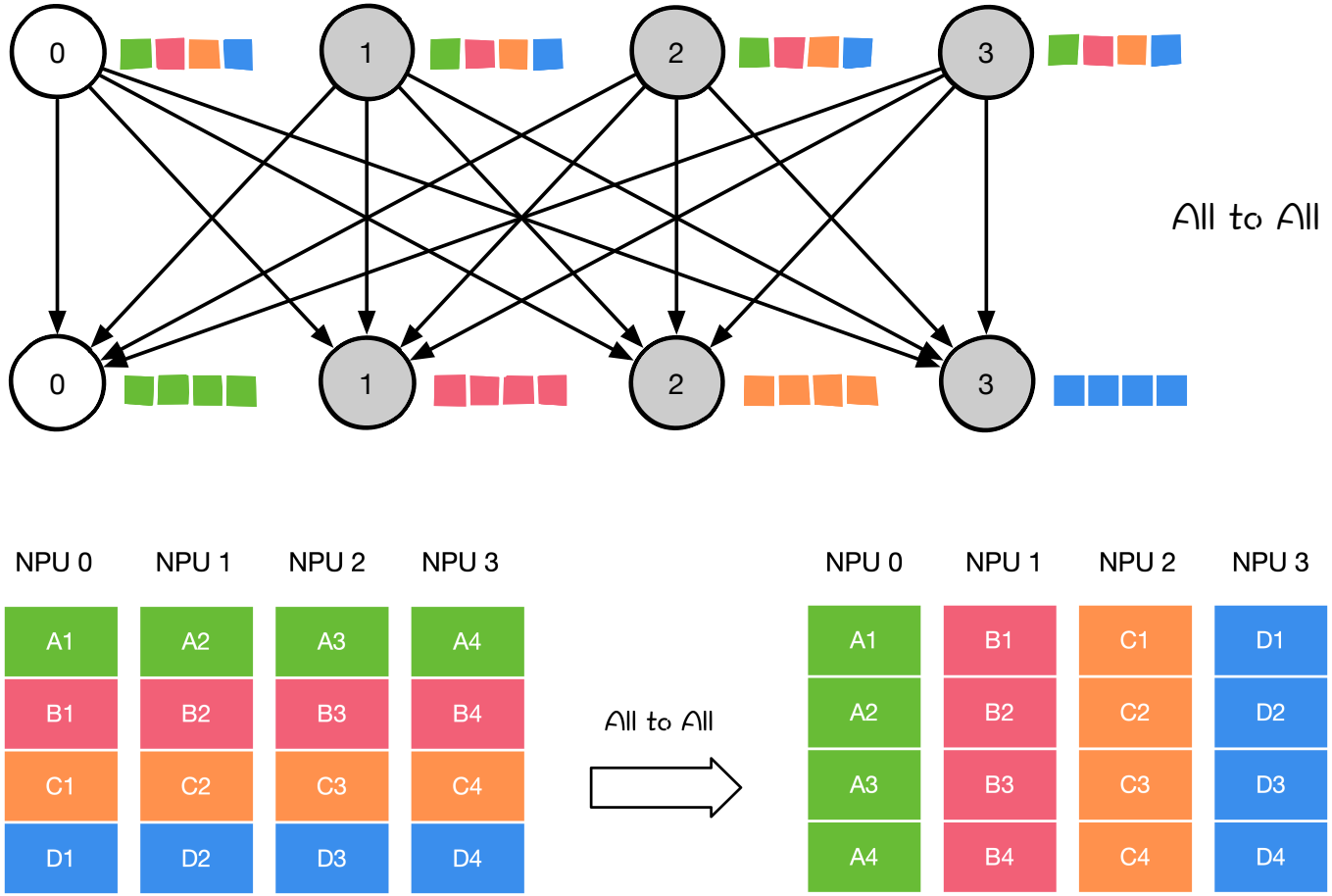


- 所有 Rank 上都按维度执行相同的 Reduce 规约操作，再将结果发散到集群内所有的节点上。

  - 可应用于数据并行 DP 和模型并行 MP：

    1. 数据并行 AllReduce 里的 Reduce Scatter，AllGather 组合里的 Reduce Scatter操作；
    2. 模型并行前向 AllGather 后的反向计算 Reduce Scatter；

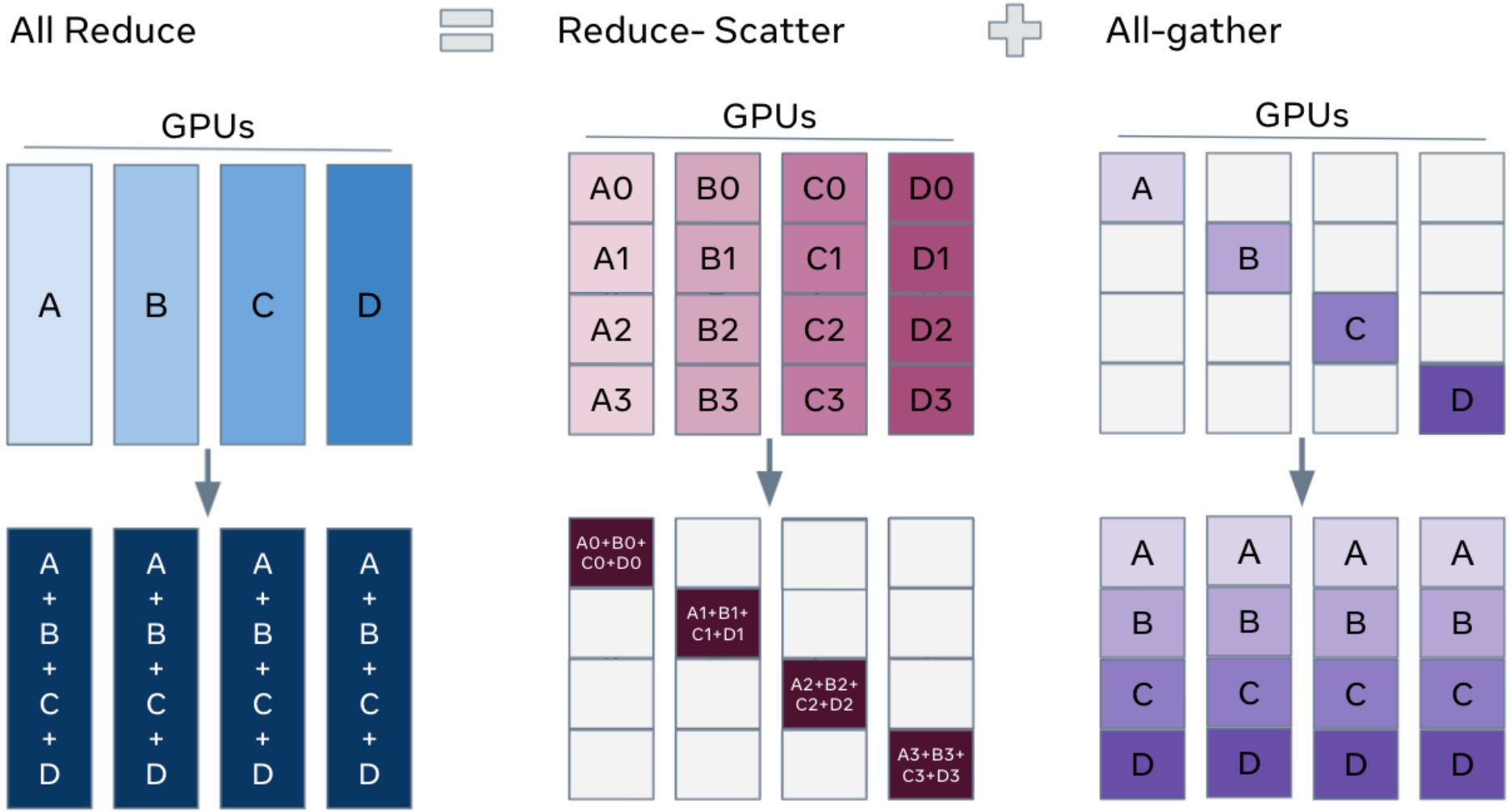
All2All



- 对All-Gather的扩展，但不同的节点向某一节点收集到的数据是不同的。

  - 应用于模型并行中的 TP/SP/EP；
  - 模型并行里的矩阵转置；
  - DP 到模型并行的矩阵转置；

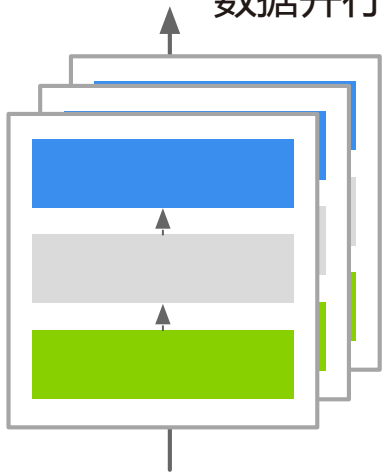
All Reduce 操作分解



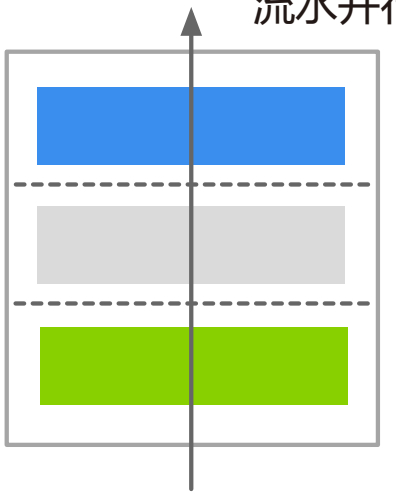
05. 小结与思考

典型通信模型

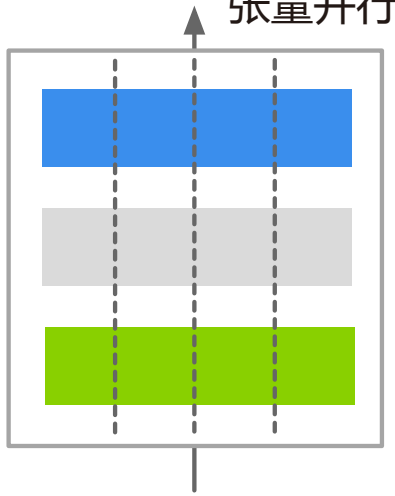
数据并行 DP



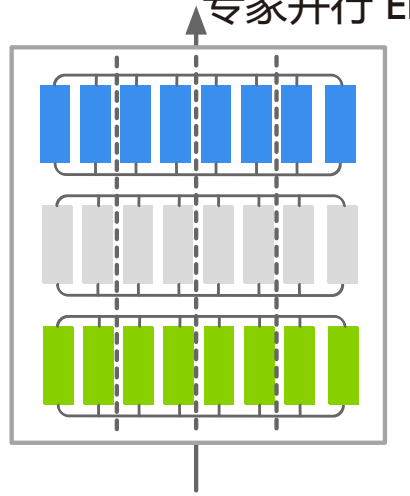
流水并行 PP



张量并行 TP



专家并行 EP



类型	通信操作	节点规模	数据量	备注
数据并行 DP	All Reduce、 Broadcast	~理论无限	~GB	节点规模增长会受线性度影响
张量并行 TP	All Reduce、 AllGather、 Reduce Scatter	2 / 4 / 8 Rank	MB~GB	计算通信可隐藏，节点内进行，不宜跨节点
流水并行 PP	Send、 Recv	2 Rank	~MB	通过多 Micro Batch 实现计算通信可隐藏
序列并行 SP	All Reduce、 AllGather、 Reduce Scatter	~理论无限	MB~GB	计算通信可隐藏
专家并行 EP	All2All	~理论无限	~MB	计算通信串行，不可隐藏

小结与思考

了解完本内容后：

1. AI 神经网络模型学习/训练阶段为什么要通信（AI 基础知识、训练推理、分布式并行）
2. XCCL 在 AI 系统中的位置（HPC 通信架构 to XCCL 通信架构）
3. 集合通信原语（All Reduce, etc.）
    1. 了解集合式通信的3种不同方式
    2. 了解一对多 Scatter/Broadcast，多对一 Gather/Reduce，多对多具体方式
    3. 了解多对多可以由一对多和多对一的方式组合，all-Reduce = Reduce Scatter + All gather
4. 集合网络拓扑（Hypercube 、Ring 、Torus 、Fat-Tree 、Dragonfly & Dragonfly+）
5. PyTorch 集合通信与计算并行

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.

ZOMI

Course chenzomi12.github.io

GitHub github.com/chenzomi12/AISystem