

大模型系列 – 集合通信库

XCCL 通信库



ZOMI

时事热点

智能体

RAG

具身智能

智能驾驶

工业大模型

...

大模型训推

6. 大模型数据&算法

数据&模型评估

Prompt 工程, 模型评估算法和测评体系

大模型算法

Scaling Law、Transform 结构, LLM/MLM 模型

7. 大模型训练

分布式训练

TP/DP/PP/SP/EP 并行, Megatron、DeepSpeed 分布式并行库介绍

微调

全参微调、底参微调(LoRA/QLoRA 等)、指令微调

8. 大模型推理

推理框架

VLLM、推理框架的架构, 推理框架线程池等构架

推理优化

大模型推理加速(XXXAttention)、长序列推理优化算法

编译计算架构

4. 计算架构

传统编译器

传统编译器 GCC与LLVM

AI 编译器

AI编译器发展与架构定义, 未来挑战与思考

前端优化

前端优化(算子融合、内存优化等)

后端优化

后端优化(Kernel优化、Auto Tuning)

多面体

复杂的循环依赖关系映射到高维几何空间

5. 通信架构

集合通信

通信原语、通信原理、集合通信算法

NCCL/HCCl

集合通信库、网络拓扑、通信方式、通信算法, NCCL 架构

3. AI 集群

集群管理运维

K8s集群运维、K8s容器、集群监控等工具

集群性能指标

稳定性、吞吐、线性度等

集群训推一体化

训练、推理大模型执行, 训练推理显存分析

机房建设

风火水电、夜冷、柜板等知识

硬件体系结构

1. AI 芯片

AI 计算体系

AI 计算模式与计算体系架构

AI 芯片基础

CPU、GPU、NPU等芯片体基础原理

英伟达GPU

英伟达GPU TensorCore、NVLink剖析

国外AI芯片

谷歌、特斯拉等专用AI处理器核心原理

国内AI芯片

寒武纪、燧原科技等专用AI处理器原理

2. 通信与存储

通信

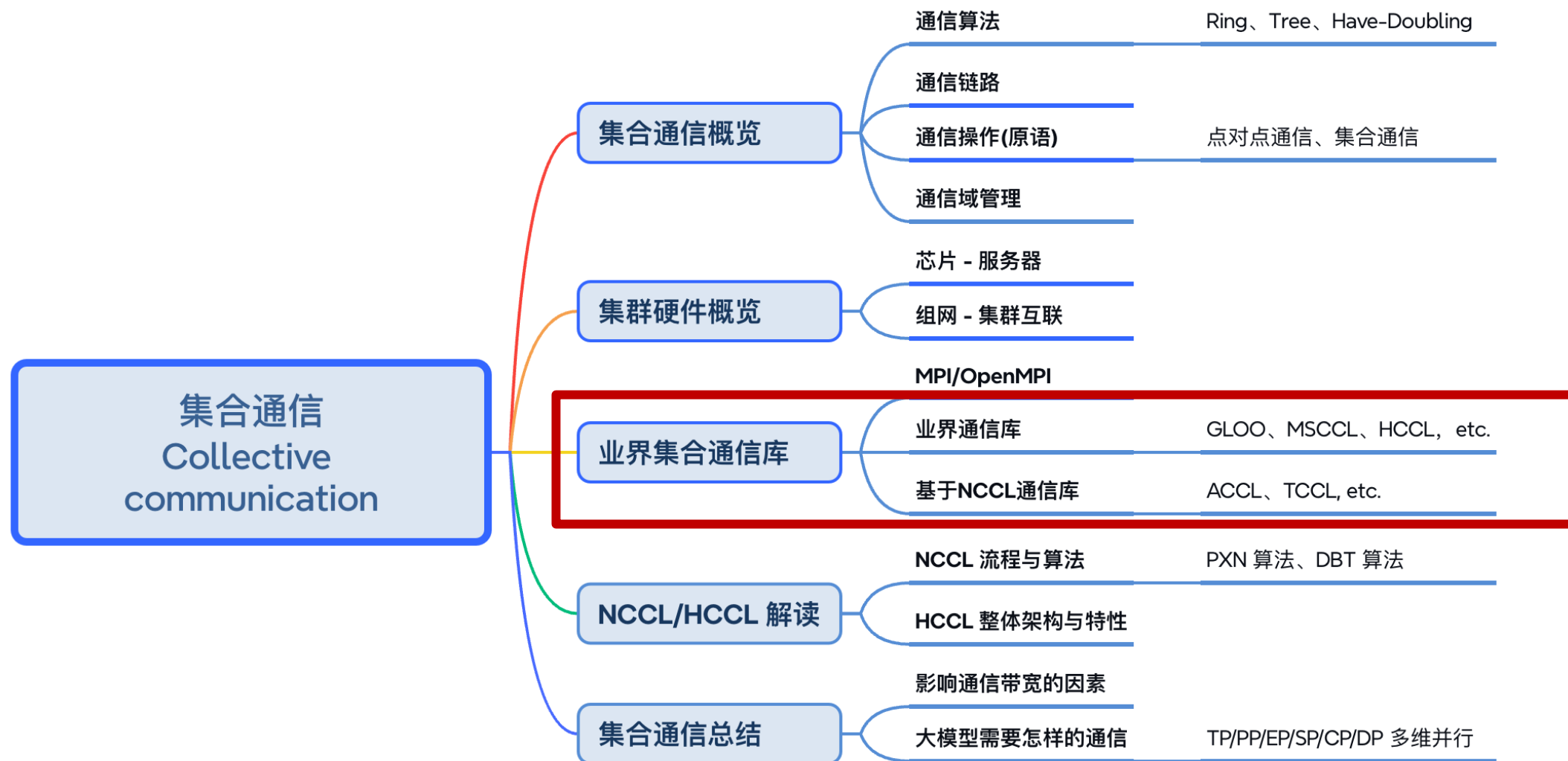
路由器、交换机基本原理和网络拓扑

存储

DRAM、SRAM、存储 POD 到大模型存储 CKPT 算法



思维导图 XMind



本节内容

1. XCCL 基本概念
2. NVIDIA NCCL
3. 阿里 ACCL
4. 腾讯 TCCL
5. AMD RCCL
6. Intel OneCCL
7. Meta Gloo
8. 总结与思考

04. 腾讯 TCCL

- <https://cloud.tencent.com/document/product/1646/93319>
- 基于NCCL (Nvidia Collective Communication Library) 开发的集合通信库

TCCL, Tencent 腾讯集合通信库

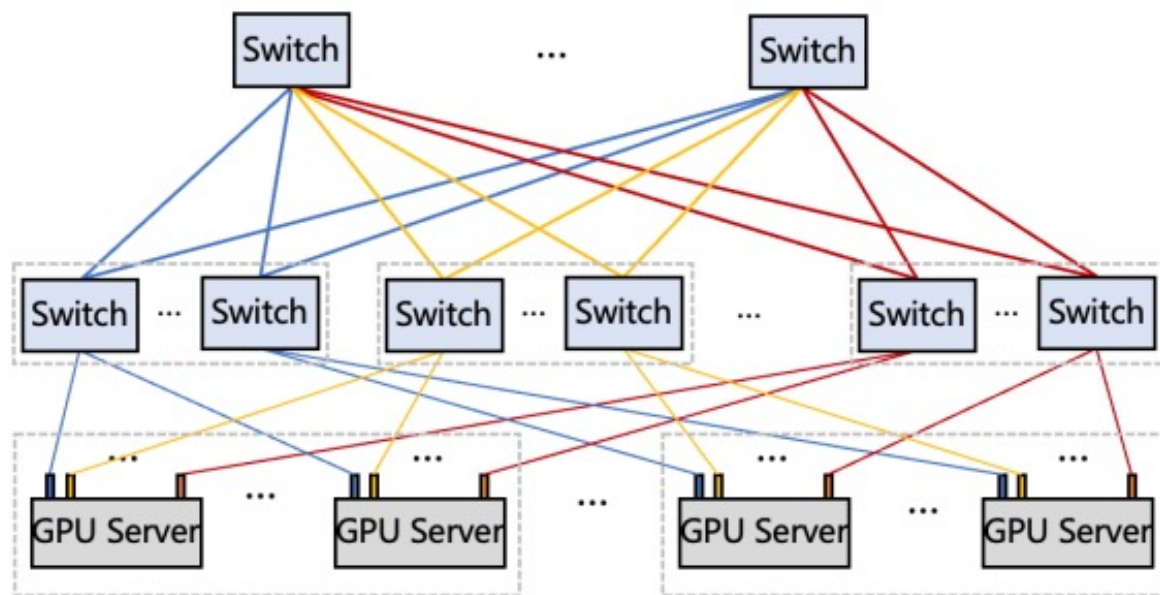
- TCCL (Tencent Collective Communication Library) 针对腾讯云星脉网络架构的通信库。依托星脉网络硬件架构，为 AI 大模型训练提供更高效率的网络通信性能，同时具备网络故障快速感知与自愈的智能运维能力。

1.6T超带宽接入

多轨道聚合网络架构

异构网络自适应通信

定制加速通信库



业务级性能监控

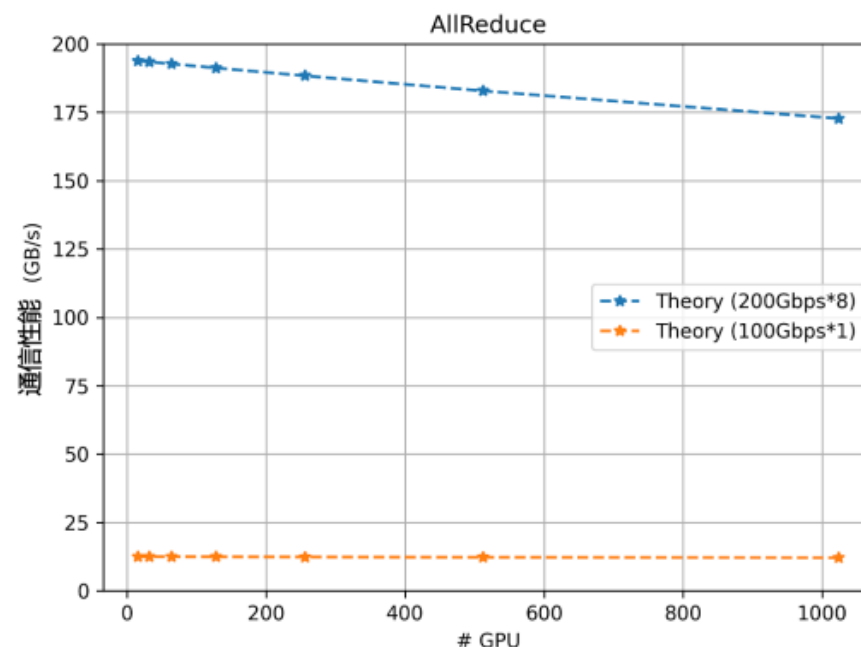
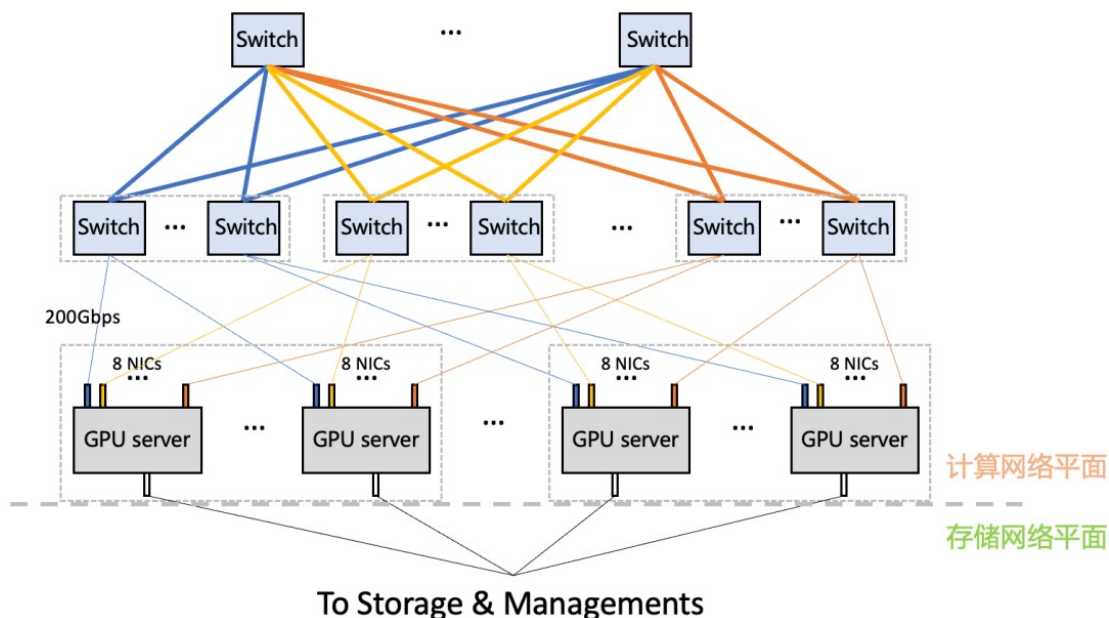
业务无感网络自愈

全栈巡检一键故障定位

端网一体化部署

TCCL, Tencent 腾讯集合通信库：主要特性包括

1. 双网口动态聚合优化，端网协同自研协议栈，发挥 GPU 性能极限。
2. 全局 Hash 路由（Global Hash Routing），负载均衡，避免拥塞。
3. 拓扑亲和 fat-tree 组网，流量调度监控，最小化流量绕行。
4. 自研交换机，可编程RDMA拥塞控制算法。



Question

I. 可以打开讲讲腾讯云的星脉网络架构吗?

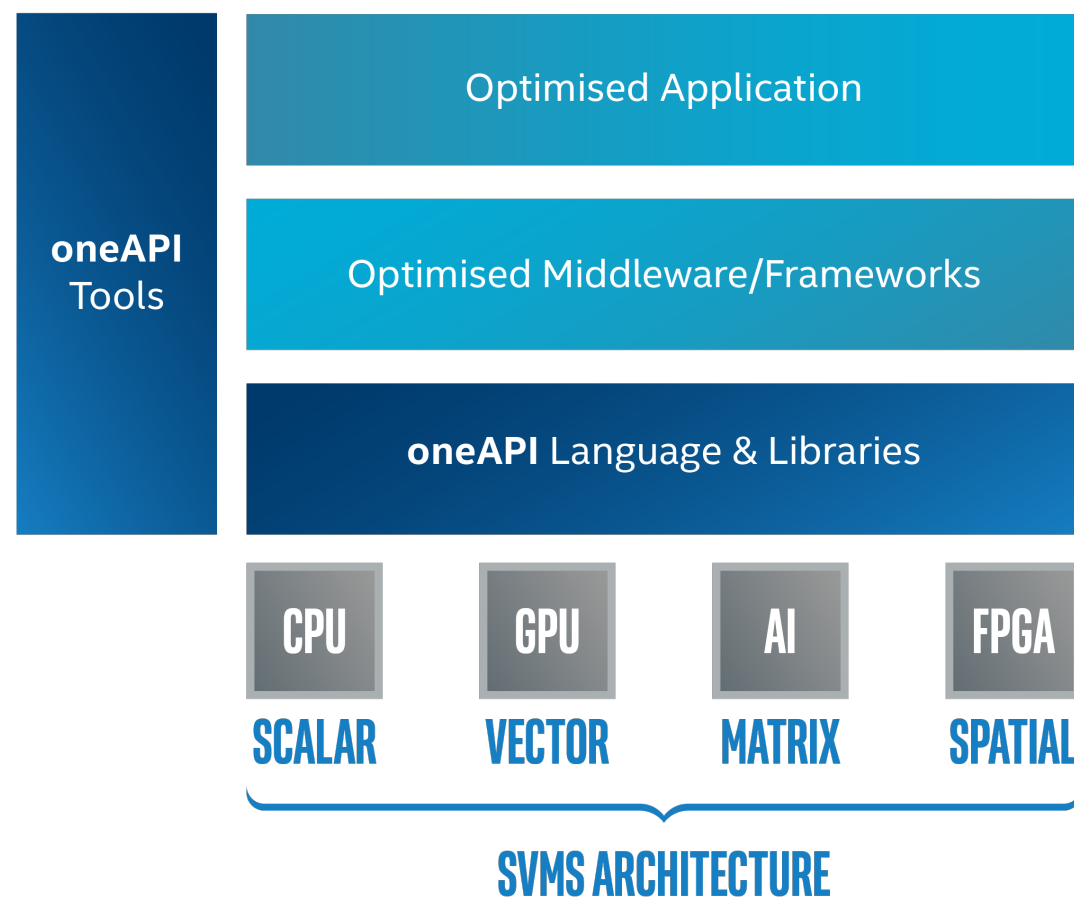


05. Intel oneCCL

- oneAPI Collective Communications Library (oneCCL)
 - <https://github.com/intel/torch-ccl>
- <https://oneapi-src.github.io/oneCCL/introduction/sample.html>

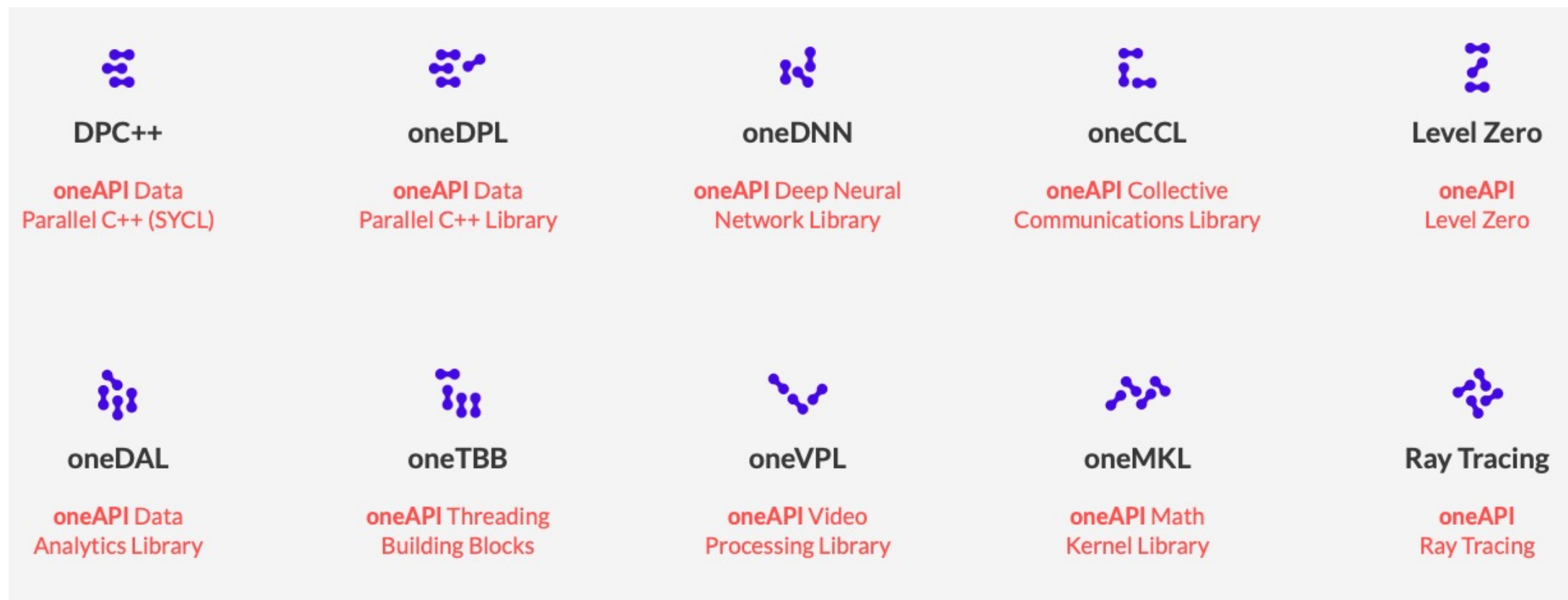
Intel oneCCL

- oneCCL 是 Intel 开发的一个高效通信库，作为 oneAPI 生态系统的一部分。



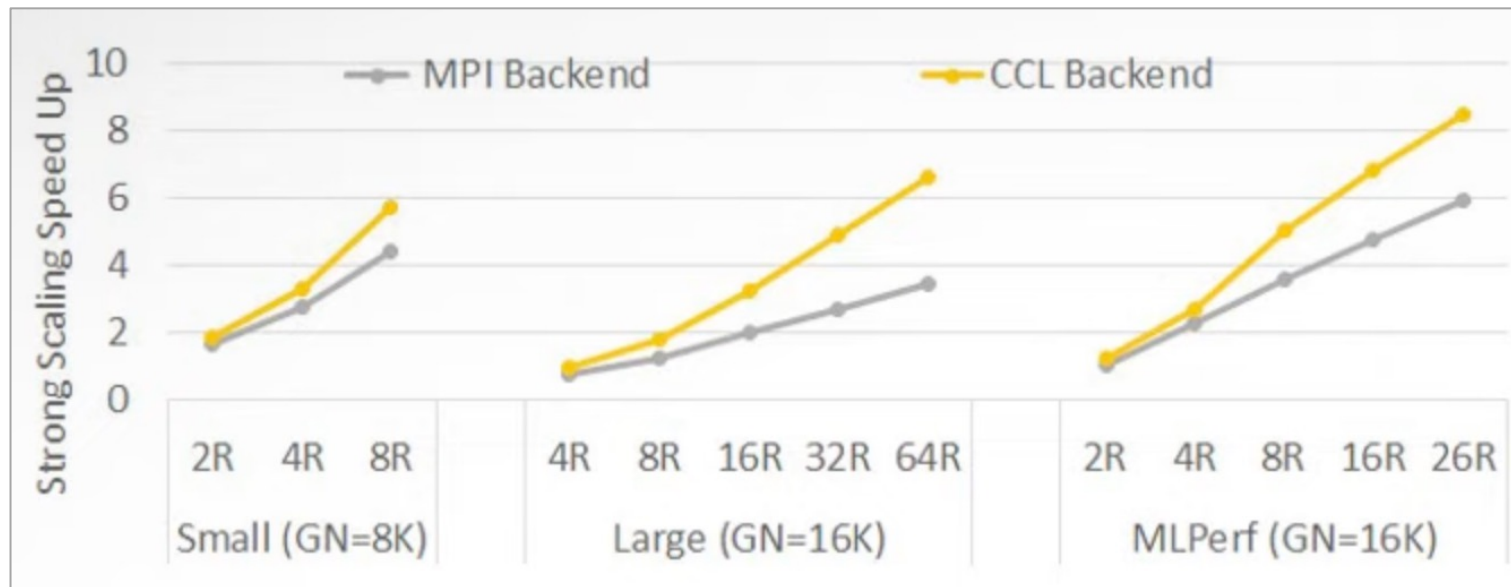
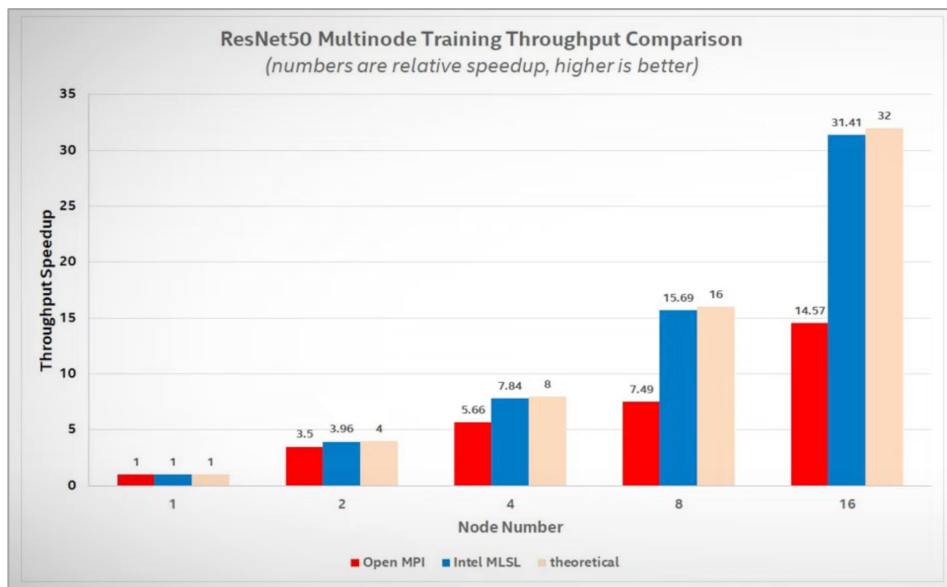
Intel oneCCL

- oneCCL 是 Intel 开发的一个高效通信库，作为 oneAPI 生态系统的一部分。



Intel oneCCL

- oneCCL 支持多种通信模式，包括点对点、广播、规约等：
 1. oneCCL 支持 SYCL，可与 Level Zero 驱动层配合，充分利用 Intel 的 NPU 和 GPU 硬件资源。
 2. 提供自动线程亲和性设置功能，可以优化多核处理器上的并行性能。
 3. 对于低精度数据类型，e.g. BF16 提供专门的优化支持。



06. AMD RCCL

- AMD ROCm Communication Collectives Library(RCCL)
 - <https://github.com/ROCm/rccl>
 - <https://rocm.docs.amd.com/projects/rccl/en/latest/>



AI HARDWARE + SOFTWARE STACK OPTIONS



LLM Training



Infiniband / RoCE

Networking Across Nodes

Infiniband / RoCE

NVLink / NVSwitch

Networking Within Nodes

Infinity Fabric

NCCL

Cross-GPU Operations

RCCL

CUDA / CuDNN

GPU Kernels for ML Workloads

ROCm

A100, H100

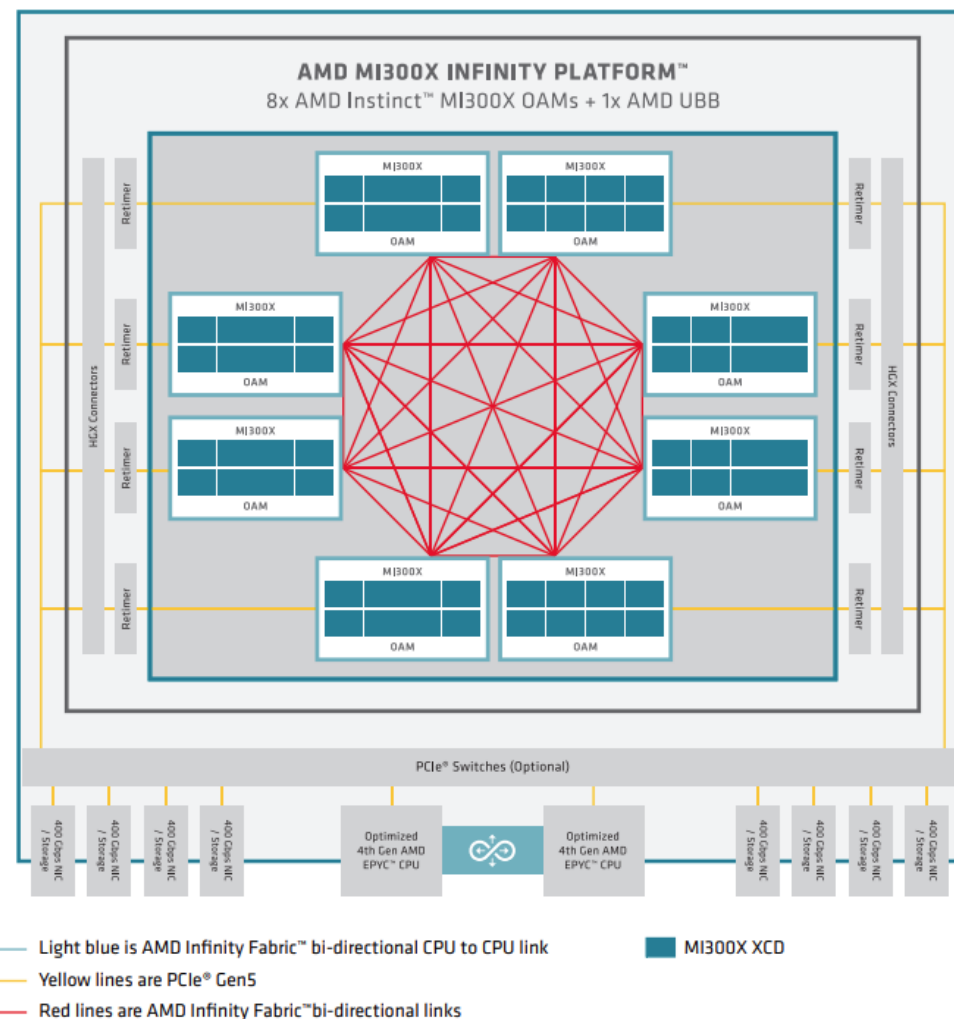
Accelerator

MI250, MI300



AMD RCCL

- RCCL (pronounced “Rickel”) , AMD ROCm 生态的一部分。
- RCCL替代NCCL, Infinity Fabric替代节点内的NVSwitch。



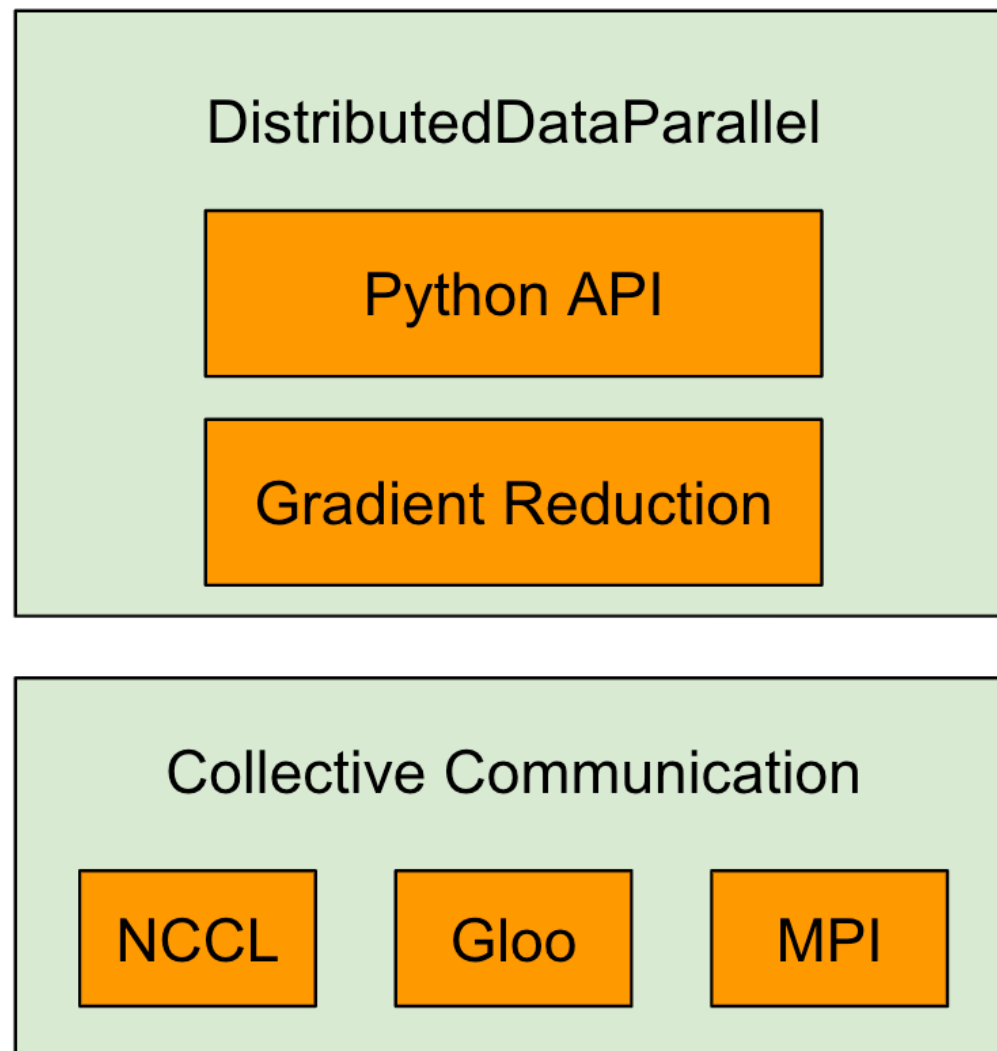
AMD RCCL 主要特性

- GPU 到 GPU (P2P) 直接通信操作
- 支持通过 PCIe 和 xGMI 互连在本地传输数据
- 支持通过 InfiniBand Verbs 和 TCP/IP 套接字在网络上传输数据。
- 对标 CUDA 使用NCCL相同 API，所以RCCL还支持通信器和拓扑等特性
- RCCL 中流特性与 NCCL 中流特性不同，它使用 HIP 流而不是 CUDA 流

07. Meta Gloo

- <https://github.com/facebookincubator/gloo>
- Gloo is a collective communications library

PyTorch 通信后端



Facebook GLOO

1. Gloo 是 facebook出品的一个类似MPI的集合通信库
2. Gloo 提供系列基于 CUDA-aware 的 All reduce 实现;
3. Mate 开源机器学习集合通信库, 对 all reduce 提供多种实现方案:
 - allreduce_ring / allreduce_ring_chunked / allreduce_halving_doubling / allreducube_bcube
4. 集成 GPU-Direct, 实现多 GPU 内存缓冲区间数据传输, 减少 Host – Device 内存复制;
5. Gloo支持通过 PCIe 和 NVLink 等互连进行多 GPU 通信
6. Gloo支持不同的数据传输方式进行节点间和节点内的数据通信

什么时候用 GLoo 什么时候用 NCCL

Which backend to use? <https://pytorch.org/docs/stable/distributed.html>

In the past, we were often asked: “which backend should I use?”.

- Rule of thumb
 - Use the NCCL backend for distributed **GPU** training
 - Use the Gloo backend for distributed **CPU** training.
- GPU hosts with InfiniBand interconnect
 - Use NCCL, since it's the only backend that currently supports InfiniBand and GPUDirect.
- GPU hosts with Ethernet interconnect
 - Use NCCL, since it currently provides the best distributed GPU training performance, especially for multiprocess single-node or multi-node distributed training. If you encounter any problem with NCCL, use Gloo as the fallback option. (Note that Gloo currently runs slower than NCCL for GPUs.)
- CPU hosts with InfiniBand interconnect
 - If your InfiniBand has enabled IP over IB, use Gloo, otherwise, use MPI instead. We are planning on adding InfiniBand support for Gloo in the upcoming releases.
- CPU hosts with Ethernet interconnect
 - Use Gloo, unless you have specific reasons to use MPI.

08. 总结与思考

各种 XCCL 库优缺点对比

	厂商	是否基于 NCCL	是否开源	是否支持异构通信	拓扑感知强化	流量监控强化	集群容错	核心设计目标
NCCL	NVIDIA	业界标杆						
oneCCL	Intel	否	是	是	是	是	是	为 Intel 的异构芯片进行加速
RCCL	AMD	否	是	否	是	是	是	为 AMD 的 NPU 硬件进行通信加速
Gloo	Mate	否	是	是	是	是	是	为 CPU-GPU 之间通信进行加速
BCCL	百度	是	否	是	是	是	是	异构通信故障诊断容错性能优化
TCCL	腾讯	是	否	是	是	是	是	为腾讯星脉网络定制通信库极致优化性能
ACCL	阿里	是	否	是	是	是	是	面向阿里云灵骏架构设计优化性能
HCCL	华为	否	是	否	是	是	是	基于昇腾硬件 NPU 之间通信进行加速

基于 NCCL 进行创新

1. 对基础/集合通信算法创新:

- 对开源网络拓扑编排算法进行优化
- 对基础通信算子&集合通信算法进行优化

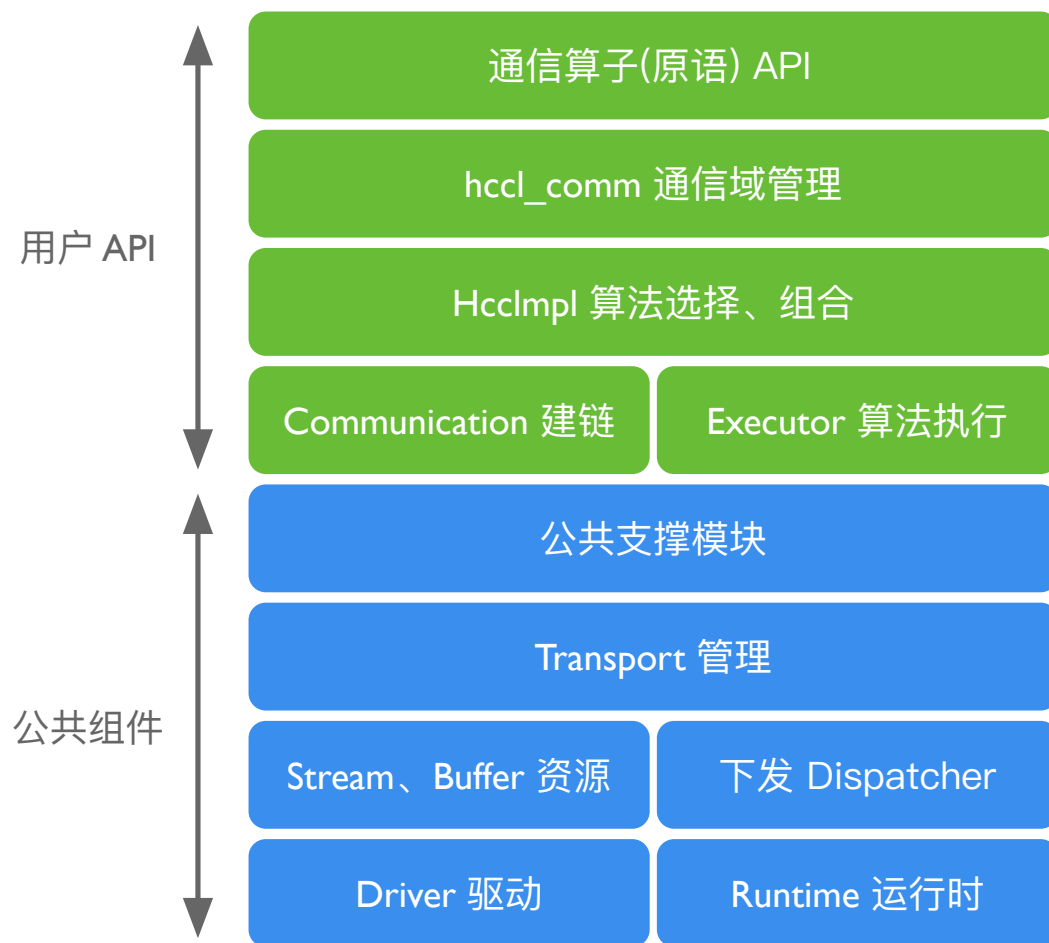
2. 网络传输优化:

- 对拥塞控制、路由控制进行优化
- 对传输协议层、路由协议层等进行细化操作

3. 端网协同专用优化:

- 端网协同自研协议栈, 匹配自研交换机
- 适配自研的网络架构进行优化

HCCL 架构图





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/AIFoundation