



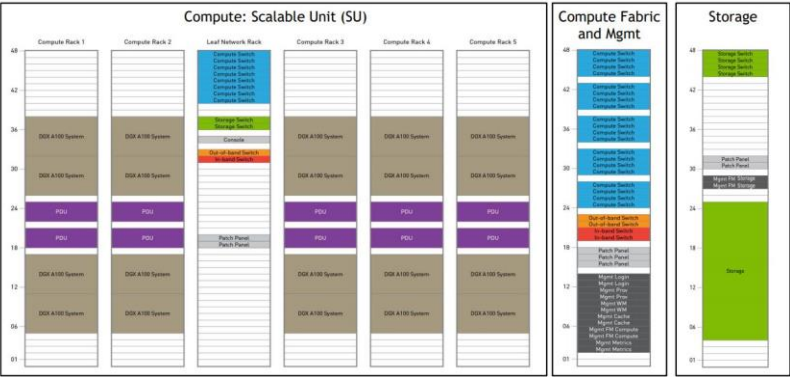
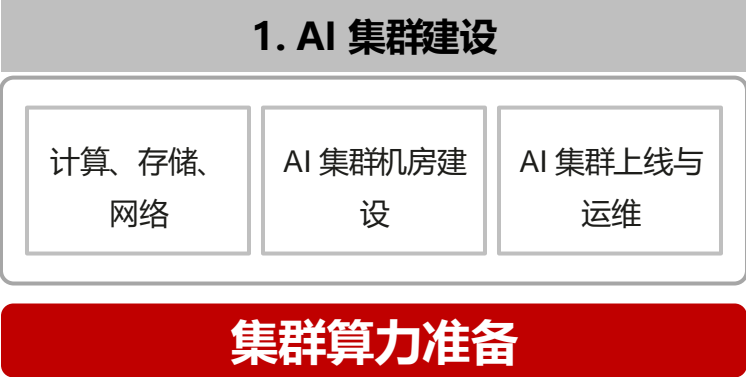
ZOMI

大模型：分布式训练

DeepSpeed & Megatron

通信后端

# 大模型业务全流程



# 01. DeepSpeed

## 后端通信

## 使用 DeepSpeed 分布式后端

- 如果 torch 分布式通信后端已经初始化，DeepSpeed 直接使用；
- 如果 torch 分布式通信后端未初始化，利用 MPI 获取和设置环境变量后初始化 torch NCCL 分布式通信后端，供 DeepSpeed 使用。

# DeepSpeed 核心步骤

1. Global cdb; 初始化 Communication Data Backend 用于跟踪和管理分布式通信后端状态;
2. Check cdb status; 如果为空或者未被初始化, 设置初始化通信后端 `dist_init_required` is true;
3. If cdb none, 设置通信后端 NCCL/GLOO/MPI 等;
4. if cdb none & torch 分布式后端已初始化, 设置 DS 通信后端为 torch 通信后端; 完成初始化, 函数退出!
5. 如果不需要初始化通信后端 `dist_init_required` is False, 函数退出!
6. 如果需要初始化通信后端:
  - a) 如果开启MPI自动发现且必须环境变量没有, 根据运行环境设置变量, 支持 PyTorch 分布式训练时使用 NCCL 后端。
  - b) 如果检查通信后端已被初始化, 通信后端初始化完成, 函数完成退出!
  - c) 如果检查通信后端未被初始化, 初始化torch通信后端, 并赋值给cdb, 函数完成退出!

# DeepSpeed 核心步骤

- <https://github.com/microsoft/DeepSpeed/blob/master/deepspeed/comm/comm.py>

# 对 PyTorch 通信接口封装

- Torch Backend 对 torch 通信接口封装，将 `torch.distributed.all_reduce()` 封装为 deepspeed 自己的通信接口 `all_reduce()`

# 对 PyTorch 通信接口封装

- <https://github.com/microsoft/DeepSpeed/blob/master/deepspeed/comm/torch.py>



# 02. Megatron-LM

## 后端通信

# Megatron-LM 后端通信

- Megatron 也对 torch 通信后端做了封装，但属于功能性封装，如封装了 `torch.distributed.broadcast()`，不是把它封装为 Megatron 自己 broadcast 集合通信接口，而是实现特定功能。

# Megatron-LM 后端通信

- [https://github.com/NVIDIA/Megatron-LM/blob/main/megatron/core/distributed/distributed\\_data\\_parallel.py](https://github.com/NVIDIA/Megatron-LM/blob/main/megatron/core/distributed/distributed_data_parallel.py)

# 03. PyTorch

## Ddistributed 后端





# 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](https://github.com/chenzomi12)

GitHub [github.com/chenzomi12/DeepLearningSystem](https://github.com/chenzomi12/DeepLearningSystem)