

PyTorch的并行化调研

PB19071501 李平治

I. 关于PyTorch

PyTorch 是一个主要提供Python接口的开源深度学习框架，用于计算机视觉和自然语言处理等应用，由 Facebook 的 AI 研究实验室 (FAIR) 开发。PyTorch以其简单、易用、灵活、高效的内存使用和动态计算图等特性，已经成为目前使用最广泛的深度学习框架。

PyTorch的两个主要计算任务为：1) 通过图形处理单元 (GPU) 进行Tensor计算的加速；2) 基于自动微分系统的深度神经网络计算，包括用反向传播来计算梯度并运行优化算法。

深度学习的迅速发展伴随着不断膨胀的网络模型和数据集，以自然语言处理领域为例：目前最为流行的预训练模型 BERT(-large)具有3亿个参数，在34亿个单词的数据集上训练而得；而OpenAI推出的GPT-3具有1750亿个参数，在大约45TB的数据集上训练而成。训练如此庞大模型，如果没有并行化处理，时间代价显然是难以想象的。

II. PyTorch 并行计算接口

PyTorch主要提供了两个数据并行的模块，分别为DataParallel和DistributedDataParallel，主要区别为：

- DataParallel是单进程、多线程的，只能在单机上工作，而 DistributedDataParallel是多进程的，既可以用于单机训练，也可以用于多机训练。由于跨线程的Python全局锁（GIL）争夺，Python只能运行伪多线程，并且每次迭代的复制模型以及分散输入和收集输出会引入额外开销，即使在单台机器上，DataParallel通常也比DistributedDataParallel慢。
- 如果模型太大而无法放在单个 GPU 上，则必须使用模型并行将其拆分到多个 GPU上。只有DistributedDataParallel与模型并行工作，而DataParallel并不能起作用。当DDP与模型并行相结合时，每个DDP进程将使用模型并行，所有进程共同使用数据并行。

DataParallel严格来说不算在本课程的讨论之内；除此之外，还有RPC模块可用于多机器间并行通信使用。因此，本调研主要关注DistributedDataParallel (DDP, 分布式并行数据) 模块。

III. DDP 的原理与挑战

具体来说，DDP将不同数据分散到每个计算资源 (GPU) 上，并在它们上复制模型、独立生成梯度，然后在每次迭代时同步模型以保持一致性。从v1.5开始，PyTorch 加入了多种加速分布式数据并行的技术，包括分桶梯度、与通信的重叠计算和跳过梯度同步。

对PyTorch这样的框架来说，提供一个通用的分布式数据并行模块面临以下三个挑战：

- 等价性: DDP的意义是为了加速训练，因此训练出的模型必须和单个GPU下的结果完全相同
- 非侵入式和拦截式API: 为了保证模型嵌套和扩展的便利性，API必须是非侵入式的，并且需要内部实现及时获取信号、以进行通信和系统优化
- 高性能: 毋庸置疑，这正是并行计算的最主要目的之一

PyTorch对以上三个问题给出的解答为：PyTorch以nn.Module的形式提供DDP类，这也正是神经网络模块的父类；每块GPU上的模型副本都从模型的相同初始参数开始，并同步梯度以保持参数在训练迭代过程中一致。DDP以本地模型作为构造函数参数，并透明地同步反向传播中的梯度。

PyTorch还提供了多种utils，方便使用者尽可能地提高并行计算性能。例如，DDP支持由多个NCCL或Gloo进程组组成一个Round-Robin进程组，该进程组以RR顺序将集体通信分派到不同的进程组实例，显著降低了通信延迟。

IV. DDP的可扩展性

Facebook AI和UWarsaw的实验表明，在256个GPU以内的情况下（通信带宽瓶颈尚不显著时），DDP呈现出了近乎线性的性能提升表现：

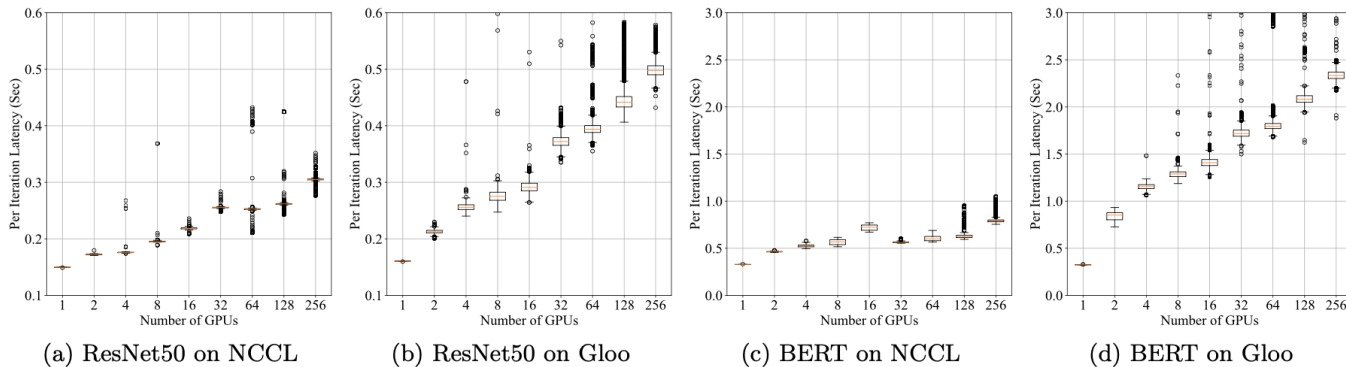


Figure 9: Scalability

在分桶梯度的桶大小选择上，实验表明桶过小或桶过大都会对通信性能有害，并且最佳储存桶大小会随着规模而增加。

V. 总结

PyTorch的分布式数据并行训练设计表明，没有一种单一的配置可以适用于所有案例，因为它很大程度上取决于模型大小、模型结构、网络链路带宽等。

尽管如此，PyTorch作为目前流行的（可能也是最好用的）深度学习脚手架，在各个配置维度上都总结了一些直觉，来帮助开发者快速确定给定用例的较佳解决方案，极大提高了实际训练场景下的计算性能。

附录

1,000,000RMB预算，配置机器:

- 4块NVIDIA Tesla A100 40GB, GPU
 $190000 \times 4 = 760000$ RMB
- 1块Intel Xeon Gold 5320, CPU
95000 RMB
- 3块Samsung 970 EVO NVMe® M.2 SSD 2TB, SSD
 $3800 \times 3 = 11400$ RMB
- 32块Samsung DDR4 64GB, RAM
 $3000 \times 16 = 48000$ RMB
- 主板、风扇、电源等其他
80000 RMB

合计: 994,400 RMB

参考文献:

1. PyTorch Distributed: Experiences on Accelerating Data Parallel Training, Shen et al., VLDB 2020