



腾讯云携手NVIDIA Mellanox挑战分布式AI训练新高度

背景

近日,腾讯云宣布,为了提升云上弹性分布式AI训练能力,采用了NVIDIA Mellanox网络事业部的 ConnectX-5 25G高性能网卡,通过RDMA技术赋能容器场景的分布式AI训练算力,大幅提升腾讯云上AI分布式训练性能,同时实现了AI训练集群的大规模扩展。

弹性容器服务 EKS 产品

EKS 是无服务器的 Kubernetes 服务,支持多种资源的高效弹性,包括基于 RDMA 网络的GPU 容器,是周期性训练任务运行的最佳容器托管平台。

介绍:<https://cloud.tencent.com/product/eks>

智能钛TI-ONE机器学习平台

TI-ONE 是为 AI 工程师打造的一站式机器学习服务平台,为用户提供从数据预处理、模型构建、模型训练、模型评估到模型服务的全流程开发支持。支持多种的 CPU/GPU 资源以及RDMA 网络,符合用户对差异化算力的场景需求。

介绍:<https://cloud.tencent.com/product/ti>

RDMA技术是一种可以在不同服务器的内存之间直接进行数据访问的通信方式,通过服务器和服务器之间或服务器和存储之间的直接内存访问,可以Bypass操作系统内核,在数据传输过程中不需要CPU的参与,在提升性能的同时也消除了集群扩展时由于CPU性能不足而导致的扩展性问题。GPU Direct RDMA则实现了不同服务器之间GPU和GPU之间的内存之间通信,不需要将数据拷贝到CPU再传给网络,大大降低了GPU之间通信的延时和提升了通信的带宽,目前已经成为了提升AI分布式训练性能的关键技术。

作为腾讯智能钛机器学习平台的关键技术之一,就是如何实现AI的训练和推理性能达到在云上和在本地操作同样的性能,整个集群的线性可扩展性和系统资源调度的灵活性成为首要解决的问题。腾讯云研发团队面对挑战,从架构设计到产品开发都做了大量的技术调研和细致的设计规划,分析和讨论了市场上各种主流的计算和通信技术,选定了GPU Direct RDMA技术作为解决云上分布式训练性能的主攻方向。在架构设计上,通过与网卡的Switchdev模式的整合,实现了在一张物理网卡上RDMA流量与VPC流量的并行传输,并且各自的带宽可以按需设定,实现了对于GPU资源的灵活应用、训练和推理的资源一体化。在具体开发问题上,由于RDMA/GPU Direct RDMA与传统的TCP和应用的接口API不一样,要想充分利用RDMA的优势,需要对应用做相应的修改。通过EKS团队的技术攻关,不仅实现了将网卡RDMA设备和PCIE拓扑透传到Kata Container,还通过对网卡相关功能的深度优化,进一步提升了应用在Kata Container中的性能。

从测试结果来看,腾讯云研发团队通过将AI训练框架(包括通用的TensorFlow和腾讯自己的训练框架)的通信库从TCP移植到GPU DirectRDMA,解决了腾讯云上AI集群训练的扩展性问题,同时将应用性能提升了56%,在成本、功耗不变的情况下,大大提升了腾讯云智能钛弹性AI集群的性价比,为腾讯自研业务上云提供了高性能的基础架构保障。

目前,除了腾讯云弹性容器服务和智能钛机器学习平台上用到了GPU DirectRDMA技术,腾讯其它产品线也逐步将RDMA技术和GPU Direct RDMA技术用在其它的应用中来帮助提升性能。

腾讯云总监裴超谈到:

“打造高性能分布式机器学习平台是腾讯云的重要发展方向之一, TI-ONE和EKS作为承担这个任务的核心产品,进行了大量的技术调研和软件开发工作,通过将传统的云数据中心内集群通信基于TCP的架构改到了采用RDMA和GPU Direct RDMA来通信,极大的提升了应用的性能,并将我们的AI分布式训练系统扩展到了上千台GPU服务器的规模,这在采用TCP通信的训练集群上是根本无法达到的。非常感谢产品团队的优异表现,让我们实现了在腾讯云上弹性分布训练和推理任务,优化了整体计算资源、提升了计算力,大大降低了总体运营成本。”

NVIDIA Mellanox网络销售与市场开拓副总裁Vadim Balahovski先生说到:

“腾讯是我们非常重要的合作伙伴, RDMA和GPU Direct RDMA技术在分布式训练中至关重要,非常高兴能和腾讯一起合作,将NVIDIAMellanox网络事业部的网络和RDMA/GPU Direct RDMA技术成功应用在腾讯云的分布式训练集群中,助力于腾讯自研业务上云。NVIDIAMellanox网络事业部一直致力于为服务器、存储和超融合架构提供高性能的智能端到端InfiniBand和以太网互连解决方案,通过我们的网络计算技术,数据中心从以CPU为计算核心的传统格局走向了以数据为核心,实现了CPU计算、网络计算和GPU加速等的多维一体新型数据中心架构。我们期待与腾讯云的进一步合作,将更多的先进网络技术应用到腾讯云上。”



北京市朝阳区望京东园七区保利国际广场T1 15层
电话: 010-5789 2000
www.mellanox.com

