

郑东阳

个人总结

- 微软实习经历以及开源项目 (Google Summer of Code、中科院开源之夏) 开发、维护经验；完成**百度飞桨黑客松**两个项目并成为 **PFCC(Paddle Framework Contributor Club)** 成员。
- 热爱技术，代码风格规范，平时有自己的**博客记录**；对深度学习框架开发、算子开发很感兴趣。

专业背景

- 了解深度学习框架开发、算子开发流程；
- 掌握 C++ 语法、STL 容器、算法库的使用及 Modern C++ 特性；了解基于 TCP / IP 的网络编程、多线程编程的相关知识；知道 CMake / Bazel，了解常用 CMake / Bazel 指令以及构建项目的流程。

实习经验

微软中国 (苏州) —— Windows 365 (Cloud PC)

2022.05 - 2022.08

- **目标**：优化数据处理模块的资源消耗与计算延迟
- **过程**：参与了 Windows 365 Data Team 中 CPC Data Platform 项目优化，将不同模块中公共的查询逻辑抽取出来复用以减少计算资源消耗；利用 Spark 处理大量流式数据与批式数据，在每秒收到 1000 条数据的负载下确保数据不丢不错，将数据处理延迟从 3-4 分钟减少至 1 分钟以内。

项目经历

百度深度学习框架飞桨 PaddlePaddle 算子开发

2023.03

技术栈: 深度学习框架开发, C++, 推理优化, GPU

1. 为 Paddle-TensorRT 添加 bitwise 算子 (代码 PR 见此处)

- **意义**：TensorRT 是英伟达的深度学习推理加速库，能最大化旗下显卡的推理性能。将 Paddle 框架中的算子映射到 TensorRT 的实现，在使用 Paddle 进行模型推理时最大化利用 NVIDIA GPU 的计算能力实现硬件加速。
- **过程**：TensorRT 中没有 `bitwise_not` 操作的算子，无法直接映射。此项目中利用 TensorRT 已有算子组合出了 `bitwise_not`。完成 C++ 端逻辑代码以及 Python 测试代码，算子测试行为与 numpy 接口对齐。
 - 对于 bool 类型，使用 `UnaryOperation::kNOT`
 - 对于 int 类型， $\sim n = -n - 1$ 成立，使用 `ElementWiseOperation::kPROD` 和 `kSUM` 组合而成。

2. 为神经网络编译器 Paddle CINN 增加 bitcast_convert 算子 (代码 PR 见此处)

- **意义**：CINN 在对接上层框架 (Paddle、Pytorch、TensorFlow) 时，会将上层的框架复杂算子进一步拆分为若干基础算子便于编译器执行算子融合，同时减少开发成本，该项目实现了基础算子 `bitcast_convert`。
- **过程**：理解 `bitcast_convert` 操作，在不改变底层存储的情况下，强制转换输入张量数据类型。若转换前后数据类型的字节大小不相同，则改变其形状。完成 C++ 端逻辑代码以及 Python 测试代码，算子测试行为与 numpy 接口对齐。

GSoC (Google Summer of Code): Introduce WebSocket to rTorrent

2022.04 - 2022.09

- **技术栈**：Modern C++、多线程、网络编程、unix domain socket
- **目标**：优化下载工具 rTorrent 的网络性能
- **过程**：
 - 用 websocket 替换原项目中的轮询式的网络模型，提高网络连接利用率。
 - 在 C++ websocket 库中添加监听在 unix domain socket 上的 feature。
 - 优化加锁逻辑，将全局互斥锁改为共享锁，使 rTorrent 可以同时执行多个只读命令，提高并发度。
- 其他：[官方公布中选结果](#)、[开发过程\(Young-Flash\)](#)