

 开源软件供应链点亮计划—暑期2020

旷视天元 - MegEngine



主办单位：

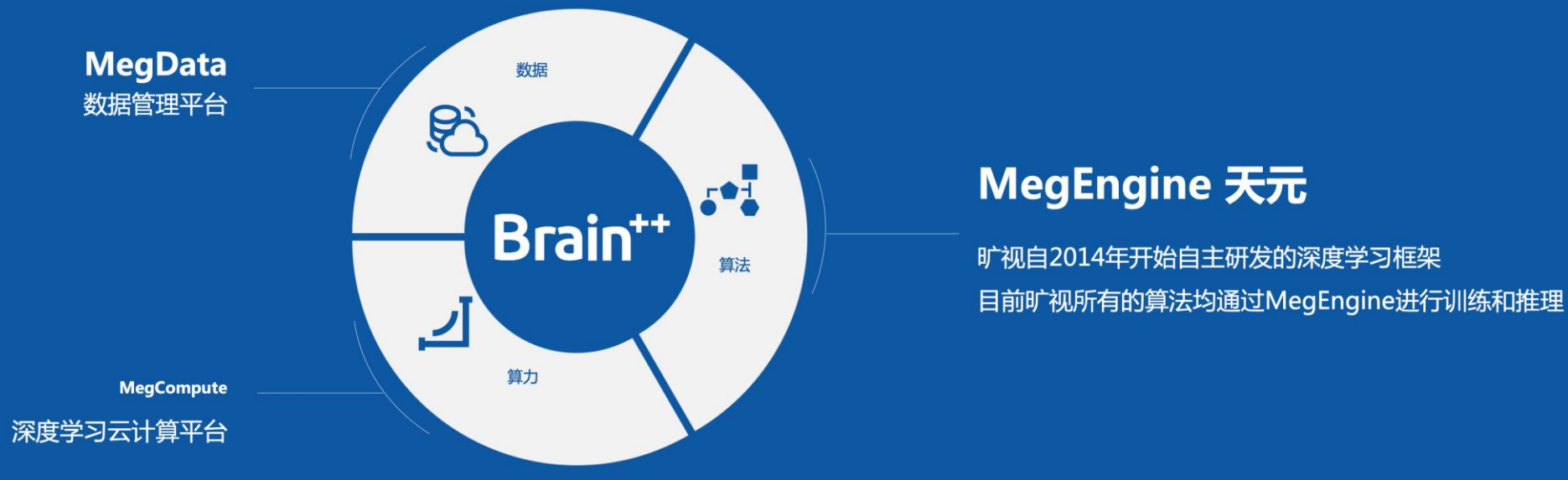


中国科学院软件研究所
Institute of Software Chinese Academy of Sciences



openEuler

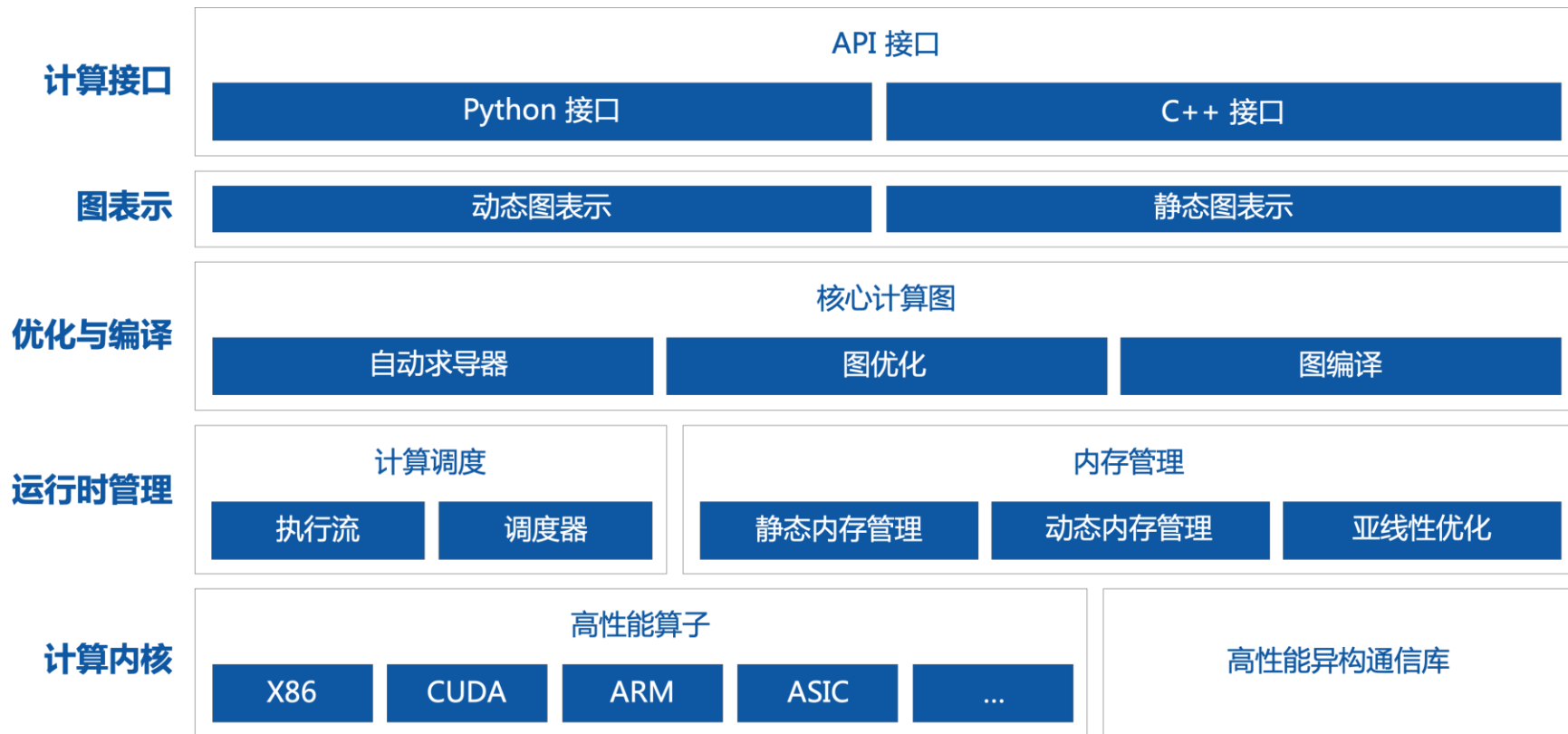
旷视Brain++ “三位一体”



天元深度学习框架

- 研发始于 2014 年
- 35万行源代码
 - C++ (78%)
 - CUDA (17%)
 - Python (5%)
- 采用 Apache 2.0 开源协议
- 旷视内部全员使用

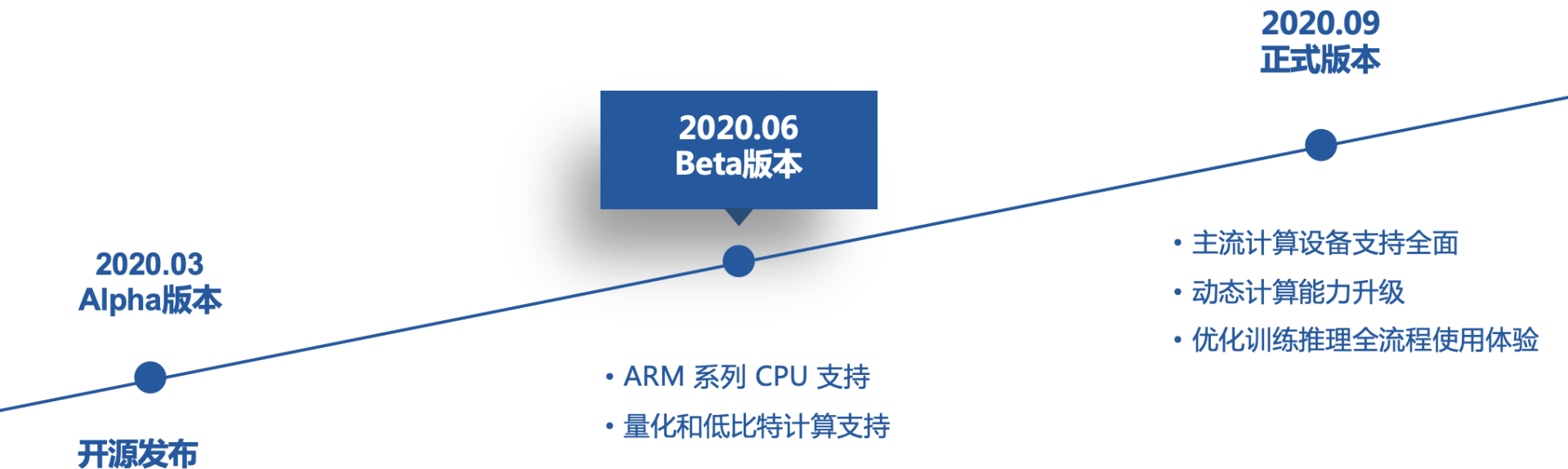
天元深度学习框架



天元深度学习框架 – 特点

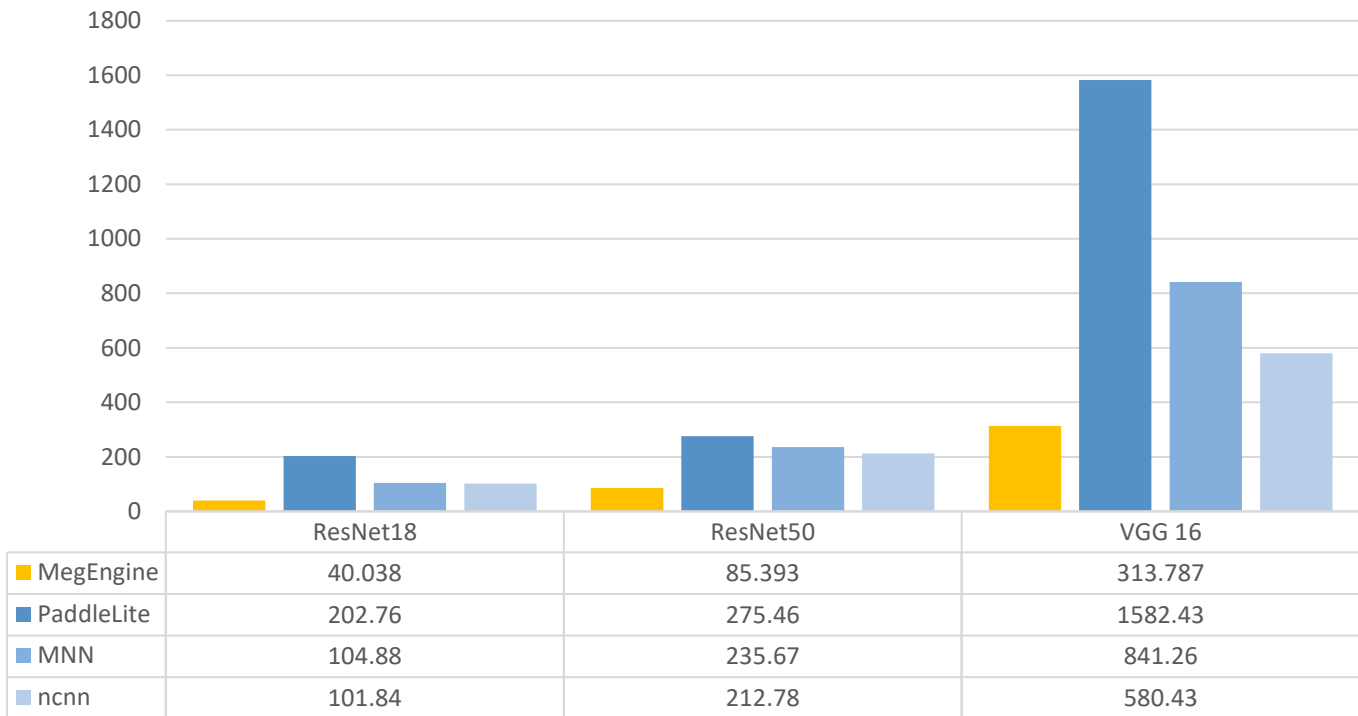
- 训练推理一体
 - 免除转换，保证精度
- 动静合一
- 兼容并包
 - 允许导入 PyTorch Module 作为子图
 - Pythonic 风格 API
- 灵活高效
 - 训练、推理速度有优势
 - 显存占用有领先优势

开发路线图



最新进展(ARM int8)

CPU 推理耗时 (ms/batch)



开源软件供应链点亮计划2020

- Github CI 增加 commit 内容检查
- Models-CI 增加模型检查
- 实现 MegEngine to Caffe 转换器
- 实现 MegEngine to TFLite 转换器
- C++ 文档生成
- MegEngine 网络可视化
- 添加常用 operator
- 实现 assistant 库

Github CI 增强

背景

- 开源项目的 CI 是保证贡献者可以安心贡献代码的重要辅助
- Models Repo 是基于 MegEngine 编写的各类 SOTA 模型
- 目前 Models 项目未检查其中的模型是否可以运行，MegEngine 还处在快速开发的过程中

目的

- 在 Github CI 中增加 commit message checker，规范贡献者 commit 格式
- 添加 Models 模型检查，试运行一个 minibatch 确保可以训练

MegEngine to Caffe / TensorFlow 转换器

背景

- 有大量 NPU / NNIE 等设备需要接受 Caffe / TF Lite 格式的模型
- MegEngine 需要先把自己的模型转换为 Caffe / TF Lite 格式的才便于集成

目的

- 编写转换器脚本，讲 MegEngine 序列化格式转换为 Caffe / TF Lite 格式的模型
- 进行正确性验证

C++ 文档生成

背景

- 目前 MegEngine 的官网使用 Sphinx 生成 Python 文档
- C++ 文档目前基于 doxygen
- 两份文档风格不统一，且无法简单相互 Link

目标

- 希望通过 Breathe 生成 Sphinx 格式的 C++ 文档
- 通过 intersphinx 将 C++ 文档与 Python 文档集成在一起

MegEngine 网络可视化

背景：

- MegEngine 拥有基于 Module 的序列化格式，以及基于静态图的部署用序列化格式
- 需要一种方式可以可视化模型网络结构

目的：

- 基于 Netron 或 TensorBoard 实现模型的网络结构可视化
- 分别针对 Module 序列化格式 以及 基于静态图的 序列化格式编写

MegEngine Assistant

背景介绍:

- 像 `BN.momentum / optimizer.momentum` 这类在业界目前有一定歧义的定义, 未来应该还会遇到很多, 我们应该做一些假设来让帮助用户少犯错, 降低排查成本

目的:

- 希望达到玩游戏的时候的那种辅助提示的效果
- 比如 `momentum<0.5` 的时候, 提示: "The momentum of batch normalization layer rarely uses a value less than 0.5, Please check the document for momentum's definition, which is different from PyTorch."
- 至多提示一次, 如果用户觉得我明白风险, 可以显式用 API 表明"我了解这个风险"

添加常用 Operator

deformable_conv、deformable_ps_roipooling
mask_convolution、matinv
svd、cumsum、batch_renormalization
adaptive_max_pooling
adaptive_avg_pooling
maxout

问答环节

Github: github.com/megengine (欢迎 Star!)

论坛: discuss.megengine.org.cn

联系邮箱: megengine@megvii.com

QQ群: 1029 7417 05

感谢



欢迎关注暑期2020活动官网