

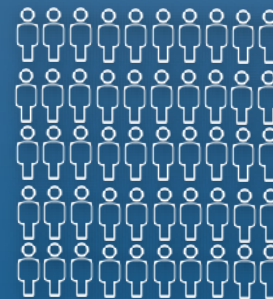
雪湖科技

深度学习神经网络加速器专家



关于雪湖科技：专注于深度学习神经网络加速器的开发

Snowlake-tech Confidential



美图公司投资

公司成立于2017年
已获得包括美图在内的两轮
风险投资

被认可的技术实力

曾为NASA提供FPGA硬件
加速的技术团队

专注神经网络加速器开发

专注于各类神经网络IP开发
，已成功为公有云 私有云
自动驾驶等提供神经网络加
速IP

国内最大

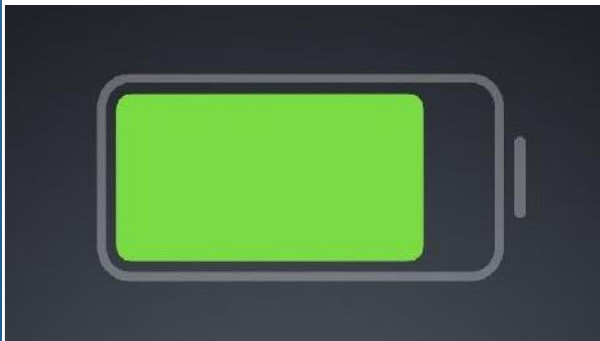
雪湖科技拥有国内最大的独
立FPGA逻辑开发团队

媒体关注

由人工智能行业著名
媒体《机器之心》评
选

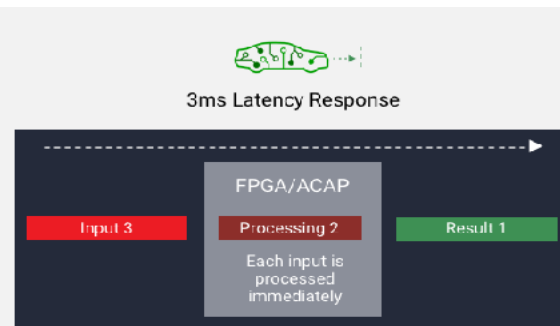
雪湖科技专注于基于FPGA的加速器开发

低功耗



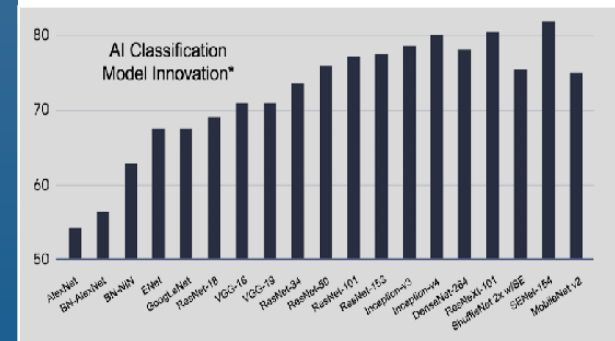
- 用于神经网络运算的GPU功耗极高，达数百瓦
- FPGA加速器可以在GPU 1/100的功耗下完成同等推断运算加速
- 除了数据中心，也适合各类边缘计算的场景

低延时高吞吐



- 神经网络运算对吞吐量要求高
- GPU的缺点是在运算前必须等待所有的输入就位，从而导致吞吐量越高时延 (Latency) 严重。
- FPGA在每个输入就位时就立即完成运算，即使在高吞吐量的情况下，时延远小于GPU和CPU

灵活迭代



- 平均每3.8个月就有1种新的算法模型问世
- FPGA+算法，最快只需2周就能完成AI硬件平台打造，快速适配新模型，无需打造新芯片
- AI芯片 (ASIC) 周期长达18个月，无法适配算法迭代。

主流FPGA神经网络加速硬件



芯片&嵌入式



FPGA加速卡



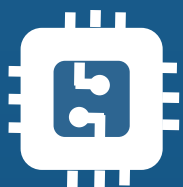
FPGA云

深度学习神经网络 (算法)



AlexNet, VGG, ResNet, Inception,
MobileNet, Yolo, FastRCNN,
Deep&Wide, etc

FPGA 芯片/加速卡/云



赛灵思, Intel Altera, Enfinix, 安路, 京微齐力

雪湖科技

神经网络加速器



Yolo, ResNet, Wide&Deep
FasterRCNN, LSTM
等各类神经网络模型

支持各种主流神经网络模型
支持Tensorflow Caffe PyTorch 等主流框架
支持各品牌FPGA: Xilinx Altera 等

应用于人工智能场景 AI推断加速

公有云



FPGA云 (FaaS)
FPGA加速卡

边缘计算



智能安防, 智慧零售
工业控制, 机器人 无人机

私有云



FPGA加速卡
FPGA FARM服务器

智能驾驶



L2-L4自动驾驶
ADAS

标准神经网络加速器



无需开发，即插即用

- 支持主流框架 Caffe TensorFlow, Pytorch等
- 已支持各类算法模型 Yolo, ResNet, FastRCNN, MobileNet等
- 标准神经网络模型

定制神经网络加速器



根据需求定制加速IP

- 支持非标准神经网络
- 支持参数量化
- 支持低精度，混合精度
- 支持各主流框架 Caffe TensorFlow, Pytorch等
- 支持赛灵思 Altera等FPGA

科学计算加速器



各类科学计算加速IP开发

- 金融交易，航空航天
- 生物医药，气象科学
- 地质勘探，国防科学等
- 开发速度快，加速性能优

产品：标准神经网络加速器（神经网络&FPGA硬件平台）

Snowlake-tech Confidential

神经网络模型

FPGA
芯片型号

GoogleNet 检 VU9P, GX2800, GX1150, GX1150	YoloV2 检 类 ZYNQ7100	YoloV3 检 类 ZU5, ZU7, ZYNQ7100	FasterRCNN 检 类 ZU3, ZU7	MobileNet 移 ZYNQ7020, ZU5
RestNet50/101/512 检 VU9P, ZYNQ7100, GX2800, GX1150	Inception_V3/V4 检 VU9P, GX2800, GX1150	FaceNet 脸 ZU5, ZYNQ7100	SENet 割 检 计划中	
BiSeNet 割 计划中	Mask RCNN 类 实 检 计划中	RefineNet 类 实 检 计划中	Deeplab_V3 割 计划中	CornerNet 检 计划中

- 检 目标检测
- 割 语义分割
- 移 移动端
- 类 图像分类
- 脸 人脸识别
- 实 实例检测

标准加速器案例：Yolo_V3 Tiny加速器

Snowlake-tech Confidential

Yolo_V3 Tiny Accelerator

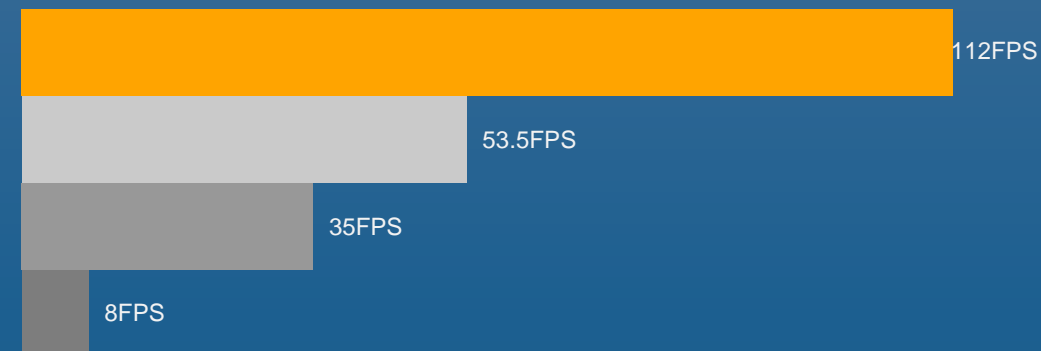
Xilinx Zynq® UltraScale+™ ZU7

112FPS@608*608

8.9ms Latency

Plug and Play

support



■ Snowlake Accelerator ■ Tesla V100 ■ Tesla P100 ■ Jetson-TX2



自动驾驶



工业



机器视觉



机器人



金融交易



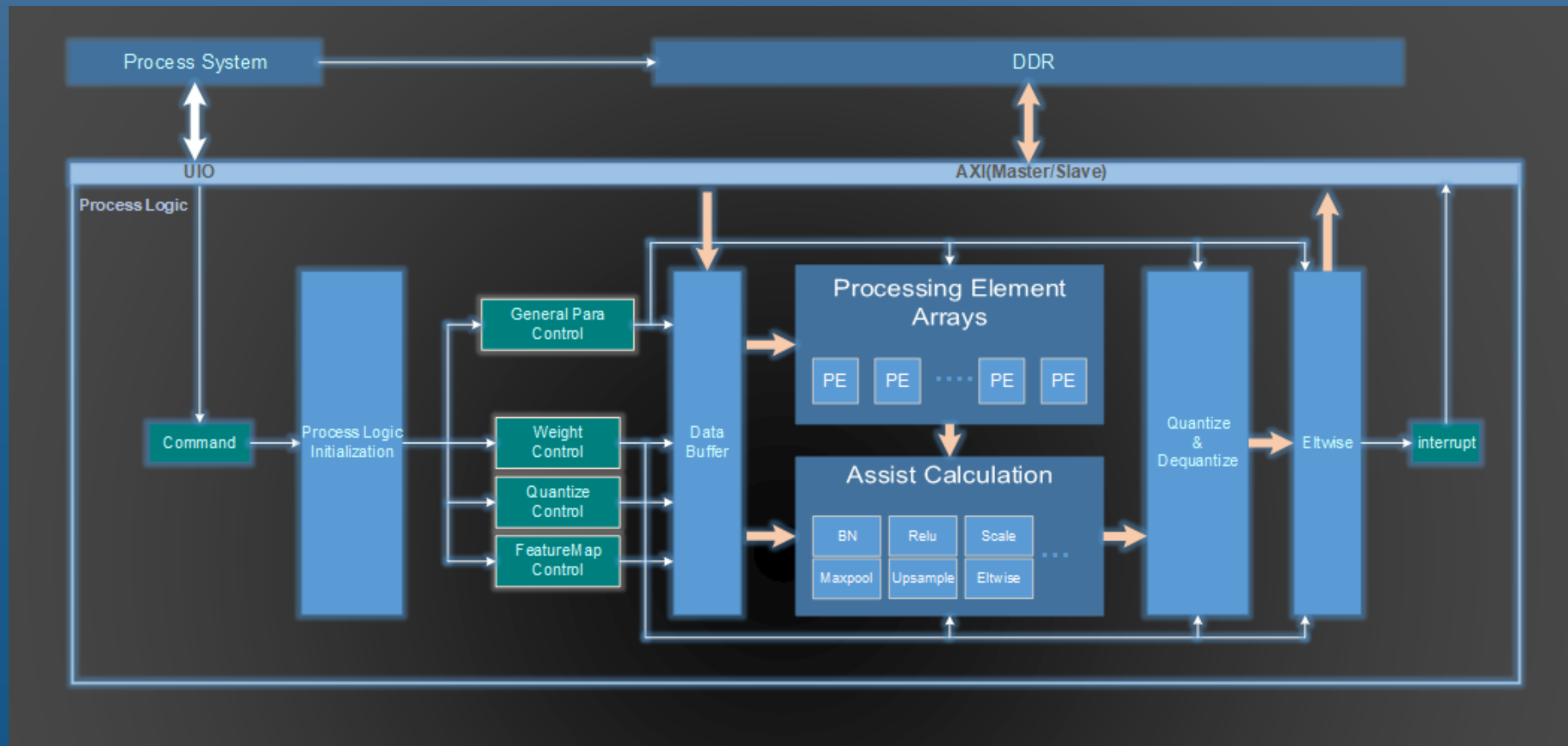
无人机



云计算

Yolo_V3 Tiny加速器架构

Snowlake-tech Confidential



产品：定制神经网络加速器

Snowlake-tech Confidential

深度学习神经网络
(算法)

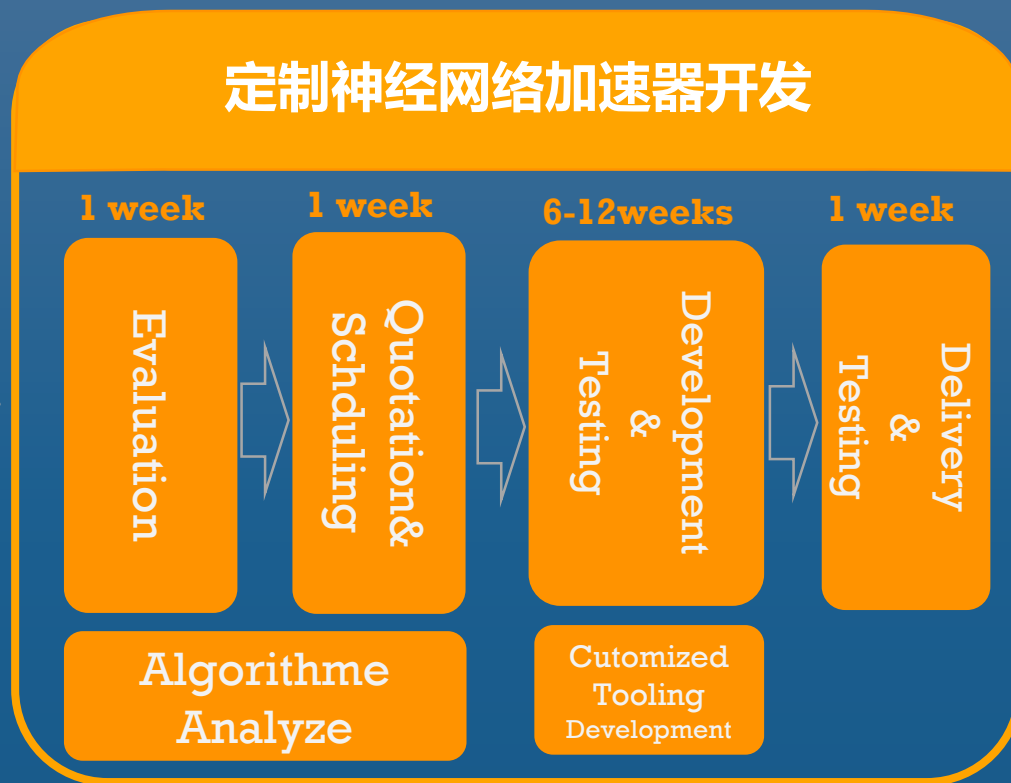


AlexNet, VGG, ResNet, Inception,
MobileNet, Yolo, FastRCNN,
Deep&Wide, etc

FPGA芯片选择

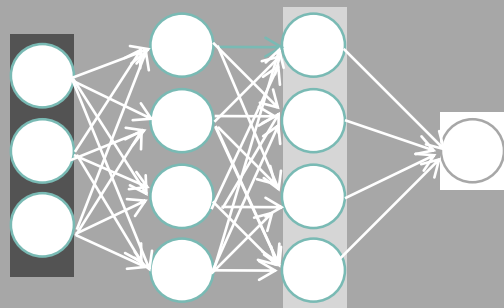


Support: Xilinx Intel Altra,
Enfinix, Anlogic, hercules-
micro



语音 机器视觉 机器人 金融交易 工业 自动驾驶 无人机

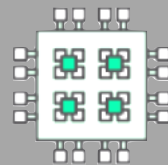
传统开发方式



算法



应用开发



FPGA芯片

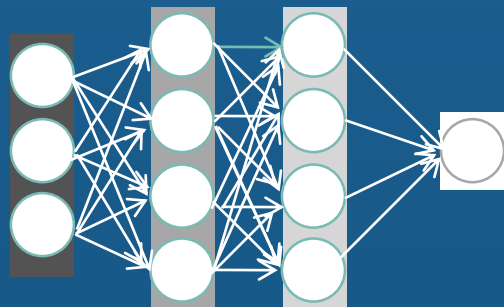
开发周期长



资源利用率低



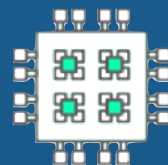
雪湖开发方式



算法



应用开发



FPGA芯片

开发周期缩短

50%



芯片资源使用率

98%





《科普：揭秘美图T9背后的FPGA计算平台处理速度提升30倍！》
<https://www.ithome.com/html/it/369425.htm>



《高出GPU服务器30倍，雪湖科技要做FPGA性能的颠覆者》
<http://dy.163.com/v2/article/detail/E6AAVHC50518UKHH.html>



《上交会今天开幕 国内首台FPGA人工智能服务器亮相》
<https://m.yicai.com/image/100165379.html>



《AI 算法在 FPGA 芯片上还有这种操作？》
<https://baijiahao.baidu.com/s?id=1631307534423387196&wfr=spider&for=pc>



《上海：浦软创业营毕业项目路演 优秀创业者遭“爆灯”》
http://www.stdaily.com/02/difangyaowen/2018-09/14/content_708721.shtml

联系方式

Snowlake-tech Confidential



扫一扫联系我们

地址：上海市长宁区天山路600弄思创大厦4号楼7层

邮件：wangyun@snowlake-tech.com，电话/微信：18616552812

THANK YOU

愿算力与你同在

