

傲睿智存：智能零售视频 AI 分析

引言

零售业“萎缩”或偷窃造成损失，如今这一问题比以往任何时期都严重。2020年，零售萎缩导致产业损失617亿美元，占零售商利润约2%。AI视频分析应用能提供强大的零售萎缩应对工具，但此类应用非常复杂，需要确定性的低时延，且其架构需要在维持较低的总拥有成本(TCO)的同时确保高性能。

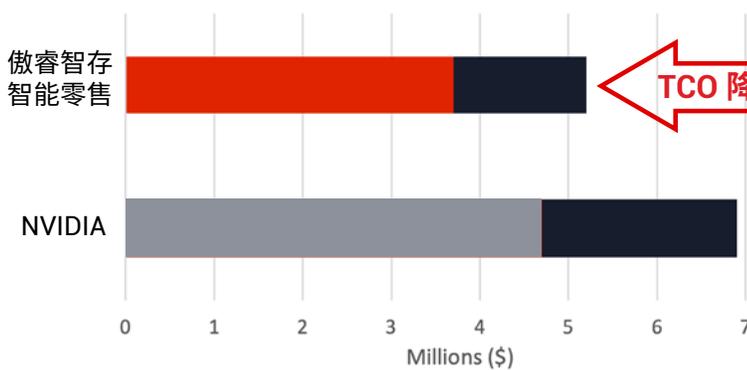
傲睿智存(Aupera)智能零售解决方案基于赛灵思Alveo加速器卡和Zynq SoC，能以较低硬件投资实现低时延和无与伦比的高性能，确保同类产品中最优的总拥有成本。

该解决方案整合了赛灵思视频机器学习流服务器(VMSS)框架和傲睿智存的AupXStream视频分析流水线处理技术，并通过Boston Stream AI和傲睿智存设备实现。

解决方案概览

傲睿智存智能零售解决方案支持多个全高清摄像头的视频输入，用于防损应用中零售目标的识别和分类。该解决方案的独到之处在于能同时运行多个推断模型，并以确定性低时延提供准确的结果，从而实现业界领先的总拥有成本(TCO)。

总拥有成本优势



注：
全球大型零售商 PoC 结果
每家店 32 个摄像头 分辨率：1080P30fps

AI 模型：Resnet 50 与 TinyYoloV3
硬件：(1 x U30 + 1 x U50) 与 4 x Nvidia T4



特性和优势

全 FPGA 加速流水线

- ▶ 实时、低时延、高吞吐量
- ▶ 并发型多模型推断
- ▶ 同类最佳的准确性

同类最佳的总拥有成本

- ▶ 减少服务器数量，降低功耗
- ▶ 快速部署，交钥匙解决方案

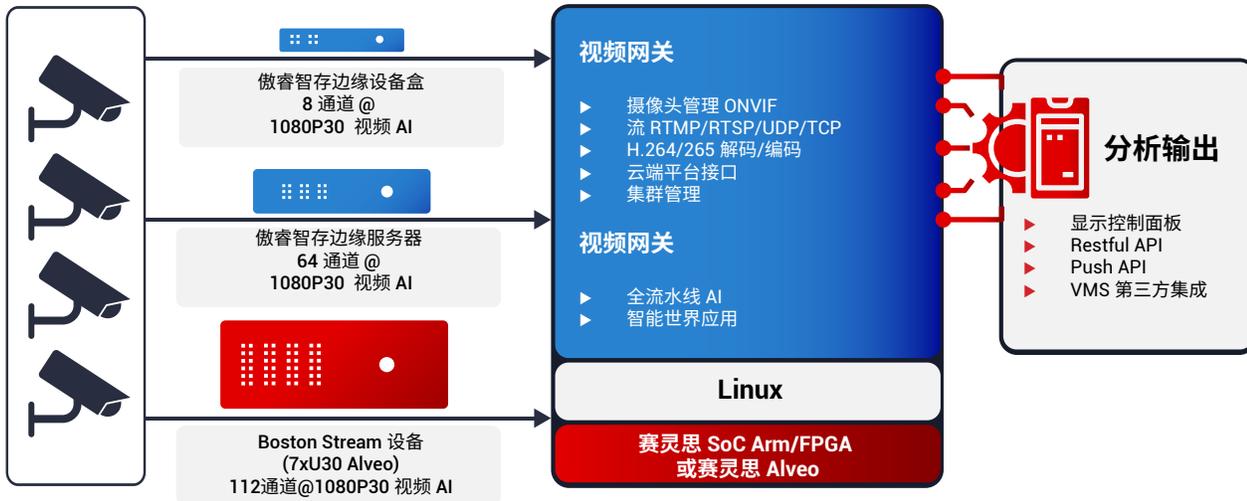
自适应和可扩展框架

- ▶ 支持边缘、数据中心和云端部署
- ▶ 支持第三方插件
- ▶ 标准 API 接口

- ▶ 32 个摄像头的配置需要两个赛灵思 Alveo 卡或 4 个 Nvidia T4 支持
- ▶ 该解决方案每处部署地点需要一个服务器，而 NVIDIA 解决方案需要两个
- ▶ 低维护、低功耗和低散热成本

解决方案部署

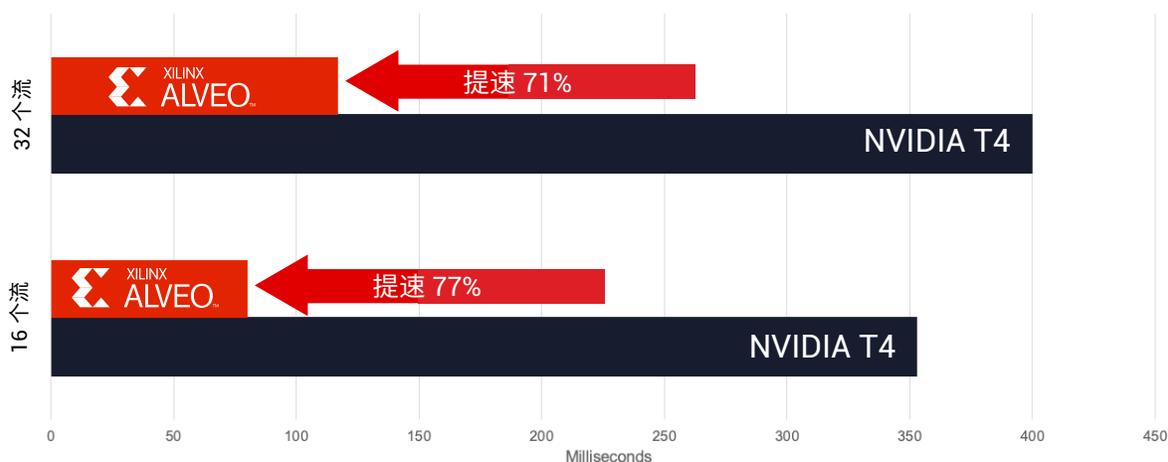
傲睿智存视频 AI 解决方案提供全软件堆栈，包括视频网关（视频解码/编码/流和处理），能部署 AI 模型用于推断加速，并通过标准化 API 支持整合于第三方软件平台。该解决方案可部署在边缘或地区性数据中心，从而支持多个零售店。



傲睿智存智能零售解决方案在 Boston Stream AI 设备上支持最多 32 个通道的 1080p、1080p30 视频，并采用一个 Alveo U30 和一个 Alveo U50 加速器卡。Alveo U30 用于执行视频解码、缩放和预处理功能，而 Alveo U50 加速器运行 Resnet 50 与 TinyYoloV3 模型，用于 AI 推断。通道（视频流）的数量和同时支持的 AI 模型在单个 Stream AI 设备中可通过添加更多的 Alveo 加速器卡来增加（最多 7 个）。

确定性低时延

VMSS 框架提供端到端的流水线加速功能，可运行预处理与后处理并通过 FPGA 架构进行推断，就典型的大型零售商 32 个摄像头商店部署场景而言，其确定性时延比基于 GPU 的竞争性解决方案低大约 75%。



注：
全球大型零售商 PoC 结果
每家店 32 个摄像头分辨率：1080P30fps
AI 模型：Resnet 50 与 TinyYoloV3
硬件：(1 x U30 + 1 x U50) 与 4 x Nvidia T4

建议后续步骤 >

访问 china.xilinx.com/smartworld



© Copyright 2021 年赛灵思公司版权所有。Xilinx、赛灵思标识、Alveo、Artix、Kintex、Spartan、Versal、Virtex、Vivado、Zynq 及本文提到的其它指定品牌均为赛灵思在美国及其它国家的商标。所有其它商标均是各自所有者的财产。在美国印刷。EW 021421

灵活应变.万物智能