AMD

AUPERA: SMART WORLD VIDEO AI ANALYTICS

Aupera
Making Video Alive

INTRODUCTION

Many aspects of life are utilizing video analytics to solve real problems and make business and people's lives smarter, whether this is smart City applications that monitor traffic congestion, parking lot availability and detect crime, to monitoring COVID-19 requirements for face masks and social distancing, to a smart building where secure points can be controlled via facial recognition for far better security than a badge. However, these applications are complex and require deterministic low latency, requiring an architecture that can deliver performance while maintaining an advantageous tCo.
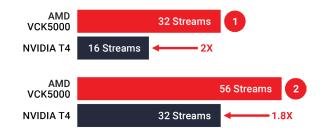
Aupera's smart World solution, built on AMD Alveo™ accelerator cards, Versal™ ACAP, and Zynq™ soCs, delivers high throughput, low latency with unsurpassed performance at a low hardware investment, assuring best-in- class tCo.

Built on the integration of the AMD Video Machine learning streaming server (VMss) framework with Aupera's AupXstream video analytics pipeline processing, these solutions enables hassle free field deployment even with custom models.

SOLUTION OVERVIEW

Aupera smart World video Ai solutions support very high density of video feeds from multiple full HD cameras for object identification and classification. The additional unique proposition of this solution is its ability to run multiple inferencing models concurrently and deliver accurate results at a deterministic low latency, resulting in industry-leading low total Cost of ownership (tCo).

THROUGHPUT ADVANTAGE

AMD VCK5000: 32 Streams **1**
NVIDIA T4: 16 Streams ← **2X**

AMD VCK5000: 56 Streams **2**
NVIDIA T4: 32 Streams ← **1.8X**

TCO ADVANTAGE

Aupera Smart Retail: **30% Lower TCO**
NVIDIA

1. AI MODEL: YOLOV3+: RESNET 18
HW: 1X AMD VCK5000 VS 1 X NVIDIA T4

2. AI MODEL: RESNET 50 & TINYYOLOV3

FEATURES AND BENEFITS

Full FPGA Accelerated Pipeline

> Real-time, low-latency, high throughput
> Concurrent multi-model inference
> Best-in-class accuracy

Best-in-Class TCO

> Fewer servers and lower power
> Fast deploying, turnkey solution

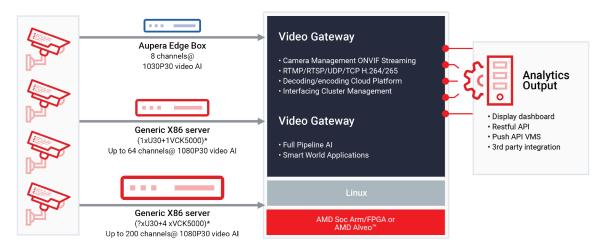### Adaptable & Extensible Framework

> Ready for edge, datacenter, and cloud deployment
> Support for 3rd party plugins
> Standard APi interface

NOTES:
RESULTS FROM POC AT MAJOR GLOBAL RETAILER 32 CAMERAS/STORE
RES:1080P30FPS
AI MODEL: RESNET 50 & TINYYOLOV3
HW: (1 X U30 + 1 X U50LV) VS 3 X NVIDIA T4

SOLUTION DEPLOYMENT

Aupera's video AI solution provides a full software stack that encompasses video gateway (video decoding/encoding/ streaming & processing), ready to deploy AI models for inference acceleration and a standard API for easy integration with 3rd party software platform. The solution can be deployed on the edge or a regional data center.
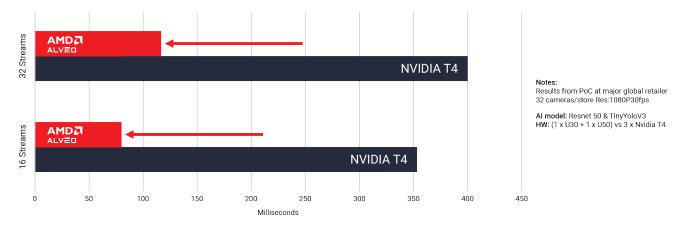

Video Gateway
- Camera Management ONVIF Streaming
- RTMP/RTSP/UDP/TCP H.264/265
- Decoding/encoding Cloud Platform
- Interfacing Cluster Management

Video Gateway
- Full Pipeline AI
- Smart World Applications

Linux

AMD Soc Arm/FPGA or AMD Alveo™

Analytics Output
- Display dashboard
- Restful API
- Push API VMS
- 3rd party integration

Aupera Edge Box
8 channels@ 1030P30 video AI

Generic X86 server
(1xU30+1VCK5000)*
Up to 64 channels@ 1080P30 video AI

Generic X86 server
(?xU30+4 xVCK5000)*
Up to 200 channels@ 1080P30 video AI

* Note: Multiple cards can be flexibly configured. Current validated servers include: Boston StreamAI server, Dell, Lenovo, Gigabyte, Supermicro. (Specific model information upon request)

The Aupera Smart World solution supports high density video feeds from multiple full HD cameras. The Alveo U30 executes the video decode, scaling and preprocessing and the AMD Alveo™ U50 accelerator runs multiple models needed for AI inference. The number of channels (video streams) and concurrent AI models can be increased by simply adding more AMD Alveo™ accelerator cards.

LOWER DETERMINISTIC LATENCY

The VMSS framework delivers end-to-end pipeline acceleration by running pre- and post-processing as well as inferencing on the FPGA fabric, delivering deterministic latency about 75% lower than the GPU-based competition in a typical large retailer 32-camera store deployment.

Notes:
Results from PoC at major global retailer
32 cameras/store Res:1080P30fps

AI model: Resnet 50 & TinyYoloV3
HW: (1 x U30 + 1 x U50) vs 3 x Nvidia T4

TAKE THE NEXT STEP
Visit **www.xilinx.com/smartworld**

AMD READY TO CONNECT? VISIT **AMD SMARTWORLD**