



Aupera's Video and AI Acceleration Platform Disrupts Data Center Streaming Market

Purpose-built, Distributed Architecture Delivers 33X Performance Improvement at 1/10th the Energy Cost and Rack Space of x86 Systems

AT A GLANCE:

Aupera Technologies is an emerging player in data center video processing systems. The Aup2600 is a purpose-built, distributed video processing system that contains 48 Xilinx Zynq® UltraScale+™ MPSoCs. It also features a complete video+AI software framework based on the Xilinx Vivado environment and Deep Learning Processor Unit (DPU) engine for neural network processing.

Customer: Aupera Technologies

Industry: Data Center

www.auperatech.com

CHALLENGE:

Displace x86 servers in the data center for video processing, transcoding and AI analytics applications

SOLUTION:

Zynq MPSoC-based distributed processing architecture and software framework enabled by Xilinx tools, video and machine learning IP.

RESULTS:

A single Aup2600 replaces 30 x86 E5 servers with 33X better performance and only 10 percent of the power and space requirements. (Figure 1).



Figure 1: The Aup2600 features 48 Xilinx Zynq MPSoC devices

CHALLENGE:

Change the Game for Video Processing in the Data Center

Aupera Technologies, founded in 2014, is an emerging player in data center video processing systems with the mission to “make video alive” for streaming applications. The company is focused on implementing heterogeneous computing architecture at the system level to build a highly efficient video processing platform. The traditional x86 systems are becoming the compute bottleneck as Internet video grows to more than 80 percent (source: Cisco) of total network traffic by 2021. “Our objective is to change the video processing landscape in the data center with a completely new architecture and software framework that addresses the concurrency pain point of live streaming,” said Roy Liao, CEO.

Video processing functions, such as decoding and encoding, are very compute intensive. Use of generic CPUs, which perform all processing in software, has reached the breaking point with the popularity of streaming increasing exponentially. Even the stacking of CPUs has proven inefficient, especially for real-time streaming video applications. To remove the CPU data center bottleneck, Aupera worked closely with YY, Inc, the largest live streaming video company in China with more than 100M active users, to design its Aup2600 system for large scale, real-time video transcoding and content analytics.

The Aup2600 is a purpose-built, distributed video processing system that contains 48 Xilinx Zynq[®] UltraScale+™ MPSoCs and handles 380 high definition 1080p concurrent video streams (H.264/265 compatible) transcoding simultaneously. In addition to its unique architecture, the Aup2600 features a complete video+AI software framework based on the Xilinx Vivado environment and Deep Learning Processor Unit (DPU) engine for neural network processing. The software framework includes customized boot loader and accelerator for the xfOpenCV computer vision library, video codec, and deep learning algorithms for object detection and feature extraction. Aupera also successfully ported the FFmpeg streaming media platform onto the Arm application processing core, and built a complete video transcoding application with region of interest (ROI) optimization.

SOLUTION:

Rapid Deployment Enabled by MPSoC-Based Architecture with Software Framework Built on Vivado and AI Environment

The Aup2600 project officially started in April of 2018. After briefly considering ASIC alternatives, which proved problematic for video processing flexibility, Aupera’s engineering team chose the Zynq UltraScale+ MPSoC. The engineers possessed strong expertise in FPGA development based on a long history of work conducted for MILCOM Telecom and NASA. In just six months, the Aup2600 project moved through initial lab testing, system integration development and testing, field testing, and commercial testing, resulting in the first product order and deployment with YY.

Liao stated, "FPGAs provide both hardware compute speed and software flexibility. In particular, the ZU7EV MPSoC is a complex, heterogeneous device, but a very state-of-the-art design—4 ARM processor cores, video codec unit, with plenty of FPGA logic resource. We compared various devices and found this MPSoC was the best fit for designing our innovative system optimized for video processing." (Figure 2).

During its early engagement with YY, Aupera needed to address rapid deployment of low latency, high efficiency video transcoding and future seamless upgrade of AI functions running on the FPGA based system. Furthermore, to support YY's live streaming and broadcasting, the company required monitoring and filtering for inappropriate content. They also needed to know what standard content to push to meet customer interests. For the ROI application that will be deployed in the coming quarter, this includes the detection of the human face where the FPGA handles the optimization, video encoding, and then addition of special effects that make the live broadcast more attractive to the end user.

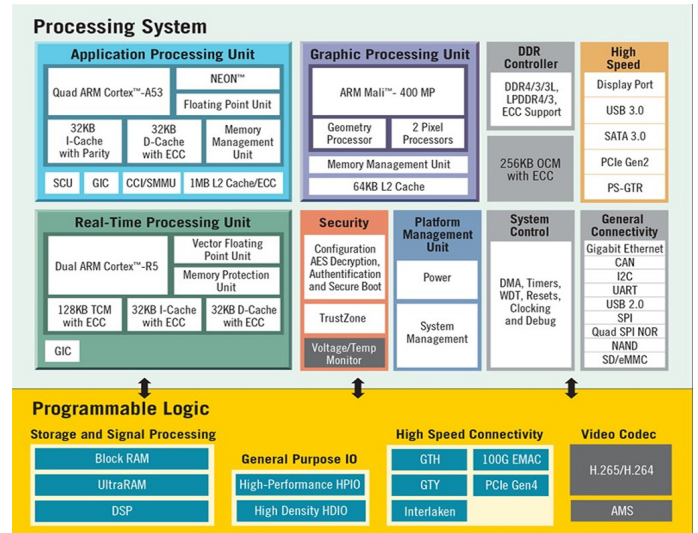


Figure 2: Block diagram of the MPSoC

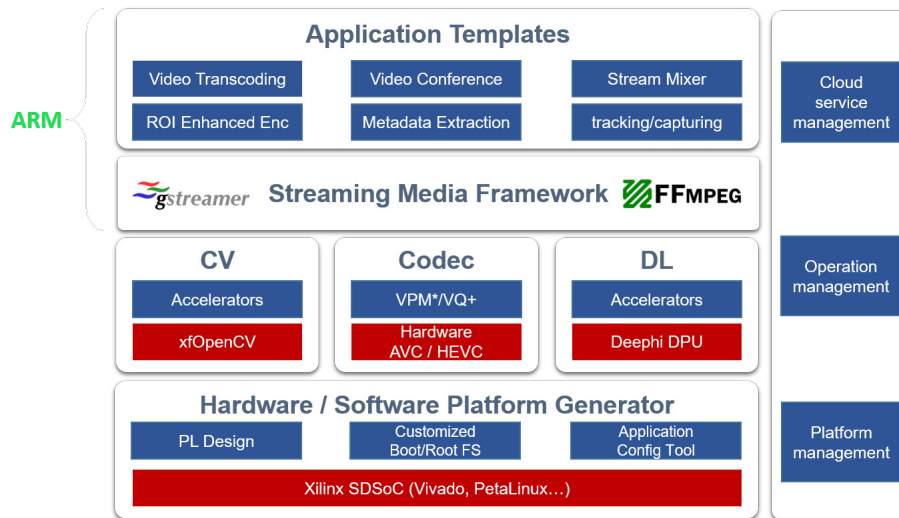


Figure 3: Aupera's video+AI software framework

Aupera leveraged the work Xilinx has done to provide a comprehensive AI environment for common model frameworks like Caffe. "With our video+AI software strategy, we are making the customer's life easier to enable the fastest deployment," said Liao. "On top of this application, we also provide templates so customers can continue to develop and design their own application based on these templates. This makes it easier for customers to adopt our new architecture, because it is very efficient. We call this 'video genius' compute architecture and think it's the future for the data center," he added. (Figure 3).

“For live streaming applications like YY’s with millions of users, controlling the latency is also very important.”

RESULTS:

Higher Performance and Lower Cost, Power, and Footprint than x86 Systems

Aupera achieved significant improvements in all the most critical metrics with the Aup2600. Performance increased 33X compared to x86-based transcoding systems and the

Aup2600 only requires 1/10th of the space and power of traditional server-based approaches. For YY, this translated into very high-quality video service at much lower cost per channel. With a single Aup2600 running a unified video+AI capability, YY could eliminate not only the traditional servers dedicated to video transcoding, but also part of the servers used for video content analytics. At the same time, accelerating the object and feature detection neural network algorithms on the FPGA resulted in real-time video analytics. (Figure 4).

Liao commented, “Although Xilinx’ video codec unit (VCU) is hard coded IP, it provides sufficient flexibility to a growing number of video workloads. Our collective solution is capable of addressing applications that not only require high density, but also low latency. For live streaming applications like YY’s with millions of users, controlling the latency is also very important.”

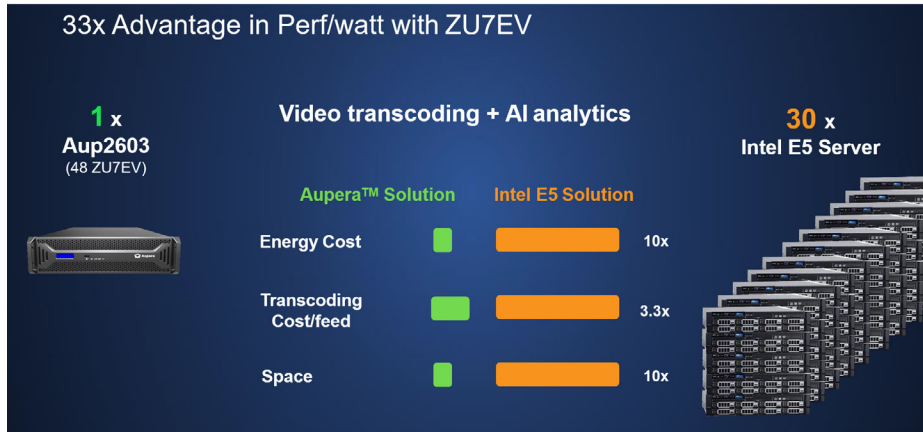


Figure 4: The Aup2600 displaces 30 x86-based servers for video processing in the data center.

CONCLUSION:

Overall, Aupera is very satisfied with its partnership with Xilinx. Liao concluded, “Xilinx is very open to its partners. Their ecosystem is extremely helpful and putting the data center first is a great strategy. We’re facing an enormous deluge of video data with 5G and IoT emerging markets. There will be a huge amount of processing required in the data center. FPGA-based video processing is going to be the most important computing capability in the future.”

Corporate Headquarters

Xilinx, Inc.
2100 Logic Drive
San Jose, CA 95124
USA
Tel: 408-559-7778
www.xilinx.com

Xilinx Europe

One Logic Drive
Citywest Business Campus
Saggart, County Dublin
Ireland
Tel: +353-1-464-0311
www.xilinx.com

Japan

Xilinx K.K.
Art Village Osaki Central Tower 4F
1-2-2 Osaki, Shinagawa-ku
Tokyo 141-0032 Japan
Tel: +81-3-6744-7777
japan.xilinx.com

Asia Pacific Pte. Ltd.

Xilinx, Asia Pacific
5 Changi Business Park
Singapore 486040
Tel: +65-6407-3000
www.xilinx.com

India

Meenakshi Tech Park
Block A, B, C, 8th & 13th floors,
Meenakshi Tech Park, Survey No. 39
Gachibowli(V), Seri Lingampally (M),
Hyderabad -500 084
Tel: +91-40-6721-4747
www.xilinx.com