The Need for AI at the Edge

A single 4K camera running at 30fps can generate more than 3GB of video data per hour. This data needs to be processed in real-time for many applications in smart cities and smart retail verticals. Traditionally, these applications run on edge servers with integrated GPUs. However, GPUs, which are designed and optimized for data centers, are not optimal for running AI workloads at the edge. We need a different class of compute appliances that provide extremely low latency AI processing at a low cost and low power. Ease of bring up, a mature software ecosystem and maintainability are also important considerations.

Unrivaled Edge AI Efficiency

Everything accelerated! Pre-, post-processing, and inference

Full Visibility through SDK

Detailed power, performance, and accuracy metrics

Automatic Model Optimization

Model mapping without user intervention

KEY BENEFITS

Unrivaled Edge AI Efficiency

> Everything accelerated! Pre-, post-processing, and inference

Full Visibility through SDK

> Detailed power, performance, and accuracy metrics

Automatic Model Optimization

> Model mapping without user intervention

TARGET MARKETS

- Smart Retail
- Smart Cities
- Smart Manufacturing
- Fleet Management
- Drones, Robotics

1, 2: Kinara computations for cost per stream and IPS (inference per second) per dollar are based on comparisons to an Nvidia T4 graphics card on Dec 2022. Product specifications can be found on nvidia.com.

All performance and cost savings claims are provided by Kinara and have not yet been independently verified by AMD. Performance and cost benefits are impacted by a variety of variables. Results herein are specific to Kinara and may not be typical. GD-181.
Support for leading edge AI models

This Edge AI appliance includes a comprehensive SDK with direct support for 100’s of AI models. The SDK’s integrated model zoo is combined with functionally-accurate applications built using GStreamer and C++ APIs that enable prototyping and rapid AI application development. The AMD Kria™ K26 SOM performs the pre- and post-processing functions along with the business analytics, while the Kinara Ara-1 edge AI processors run the inference.

Compute & Storage

- 1x AMD Kria™ K26 SOM
- 4x Kinara ARA-1 Accelerator with 4x 1GB LPPDR4
- 4GB LPDDR4 & 16GB eMMC

INTERFACES

- 3x USB 3.0 Type A
- 1x USB Type B micro for debugging
- 1x DP1.2a and 1x HDMI
- 2x RJ45 Jack to connect external server/ POE Switch
- 1x M.2 connector for SSD and 1x GT Connector
- Status LED and Alarm out

Compact form factor

EAIA comes at a compact form factor of 255mmX205mmX42mm, that makes it suitable for space constrained environments.

FIND OUT MORE

For More Information: www.xilinx.com/kria kinara.ai

For Ordering Information: https://www.vvdntech.com/visio
n/vvdn-edge-ai-appliance

MECHANICAL

- Dimensions: 205mm(w) X 255.2mm (d) X 42mm (h)
- Operating Temperature: 0°C to 45°C

KIT COMPONENTS

- VVDN edge AI appliance unit
- Power adapter 12V at 8.5A

*AMD, the AMD Arrow logo, Kria and combinations thereof are trademarks of Advanced Micro Devices, Inc.