Unlock a New World of AI Applications into Edge and Endpoint Devices

Build low-cost, highly efficient AI solutions in a wide range of embedded devices with Arm’s Ethos-U microNPUs which enable systems based on Arm Cortex and Arm Neoverse. Ethos-U provides a scalable range of performance and memory interfaces and integrates low-power Cortex-M SoCs, as well as SoCs based on high-performance Arm Cortex-A, Cortex-R, and Arm Neoverse. Develop, deploy, and debug AI applications with the Arm endpoint AI solution using a common toolchain across Arm Cortex and Ethos-U processors.

Highlights

- **Energy Efficient**
  Provides up to 90% energy reduction for ML workloads, such as ASR, compared to previous Cortex-M generations.

- **Network Support**
  Flexible design supports a variety of popular neural networks, including CNNs and RNNs, for audio processing, speech recognition, image classification, and object detection.

- **Future-Proof Operator Coverage**
  Heavy compute operators run directly on the micro NPU, such as convolution, LSTM, RNN, pooling, activation functions, and primitive element-wise functions. Other kernels run automatically on the tightly coupled Cortex-M using CMSIS-NN.

- **Reduce Memory Footprint**
  Up to 70% reduction in model size with compression allows for the execution of larger networks and speeds up the execution of networks.

- **Offline Optimization**
  Offline compilation and optimization of neural networks, performing operator, and layer fusion, as well as layer reordering to increase performance and reduce system memory requirements by up to 90%. Delivers increased performance and lower power compared to non-optimized ordering.

- **Develop Multiple Products**
  Target multiple markets with a single processor IP architecture that provides the system flexibility to configure the performance and desired memory system and OS.

- **Quickly Create Applications**
  Accelerate time to market by leveraging the Arm AI ecosystem with partners developing optimized algorithms ahead of hardware availability.

**Ethos-U65 can target numerous different applications with use in high-performance Cortex-A and Neoverse systems or low power Cortex-M based embedded devices.**
KEY USE CASES FOR THE ETHOS-U PROCESSOR SERIES

- Object classification
- Object detection
- Face detection/identification
- Human pose estimation
- Image segmentation
- Image beautification
- Super resolution
- Speech recognition
- Sound recognition
- Noise cancellation
- Speech synthesis
- Language translation
- Natural Language Processing

MARKET SEGMENTS

- Mobile
- Smart Camera
- Smart Home
- STB/DTV
- Consumer
- AR/VR
- Medical
- Robotics
- Drones
- Rich IoT
- Automotive (IVI/ADAS)
- Infrastructure

Specifications

<table>
<thead>
<tr>
<th>Key Features</th>
<th>Ethos-U55</th>
<th>Ethos-U65</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance (At 1 GHz)</td>
<td>64 to 512 GOP/s</td>
<td>512 GOP/s to 1 TOP/s</td>
</tr>
<tr>
<td>MACs (8x8)</td>
<td>32, 64, 128, 256</td>
<td>256, 512</td>
</tr>
<tr>
<td>Internal SRAM</td>
<td>18 to 50 KB</td>
<td>55 to 104 KB</td>
</tr>
<tr>
<td>System Interfaces</td>
<td>Two 64-bit AXI</td>
<td>Two 128-bit AXI</td>
</tr>
<tr>
<td>External Memory</td>
<td>SRAM and Flash</td>
<td>SRAM, DRAM, and/or FLASH</td>
</tr>
<tr>
<td>Recommended Host CPU</td>
<td>Cortex-M55, Cortex-M7, Cortex-M4, Cortex-M33</td>
<td>Cortex-M55, Cortex-M7</td>
</tr>
<tr>
<td>Operating Systems</td>
<td>Bare-metal or RTOS</td>
<td>Bare-metal, RTOS, or Linux</td>
</tr>
</tbody>
</table>

To find out more about the Ethos-U processor series, visit developer.arm.com/ethos-u