

Endpoint AI is Critical to Our AI-enabled Future

Carlos Morales, Ambiq

AI Day

September 21st, 2023



What is Endpoint AI? (aka Why Cloud AI Round-trips Suck)



CLOUD-ASSISTED AI

Voice processed in the cloud

2-5s response time

0.5j BLE xmit/rcv



ENDPOINT AI

Voice processed locally

100ms response time

<.0005j compute

TAANSTAFL: to avoid the round-trip tax, you need efficient system and enough local memory

Endpoint AI Will Be LLM's Senses



TODAY

AI-assisted
"Microphone"



Audio
(~0.05 joules)

Good!

Speech
Recognition



Text
Prompt

ChatGPT/LLM

COMING SOON

AI-assisted
"Microphone"

+

Speech
Recognition
Front End



Tokens
(~0.005 joules)

Better!

Speech
Recognition
Back End



Text
Prompt

ChatGPT/LLM

Endpoint AI is Already Here

AI is the **easiest** way to perform difficult tasks

Endpoint AI is benefiting from **major tailwinds**

The “AI” bit is **only part** of Endpoint AI solutions

Endpoint AI is **practical today**



Endpoint AI Requires a Complete System in a Tiny Power Budget

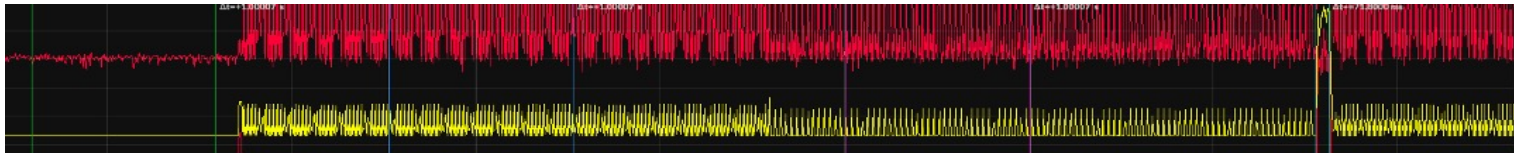


TYPICAL SYSTEM ELEMENTS MUST FIT WITHIN 1mW FOR 1 YEAR LIFE ON 2 AA LITHIUM BATTERIES



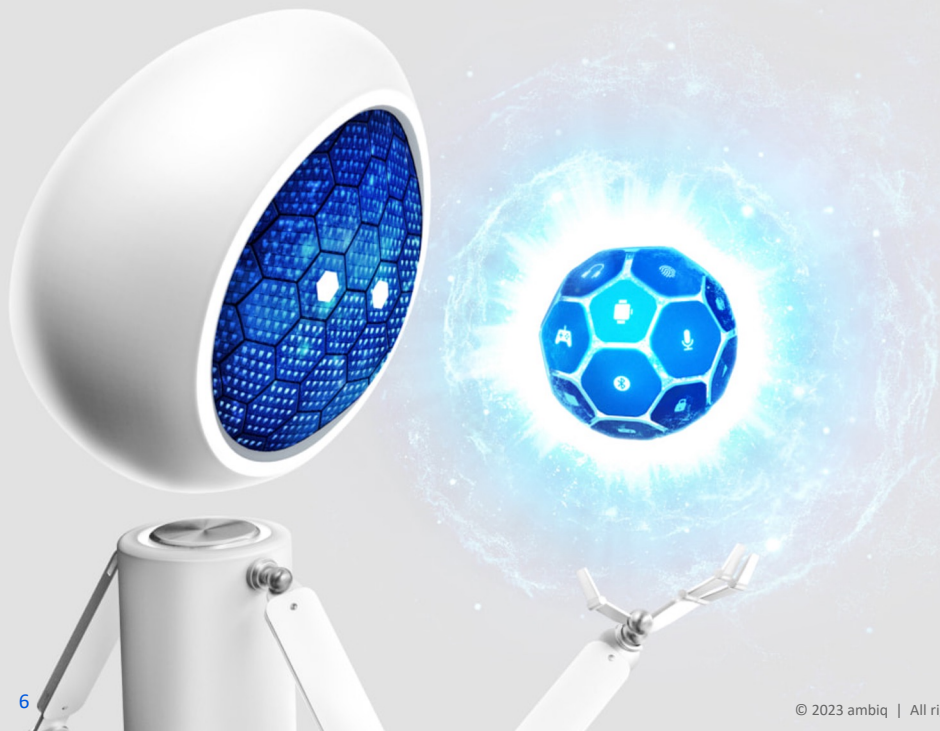
Endpoint AI is about meeting ALL system needs in a mW power budget

Capturing 6 second audio utterance



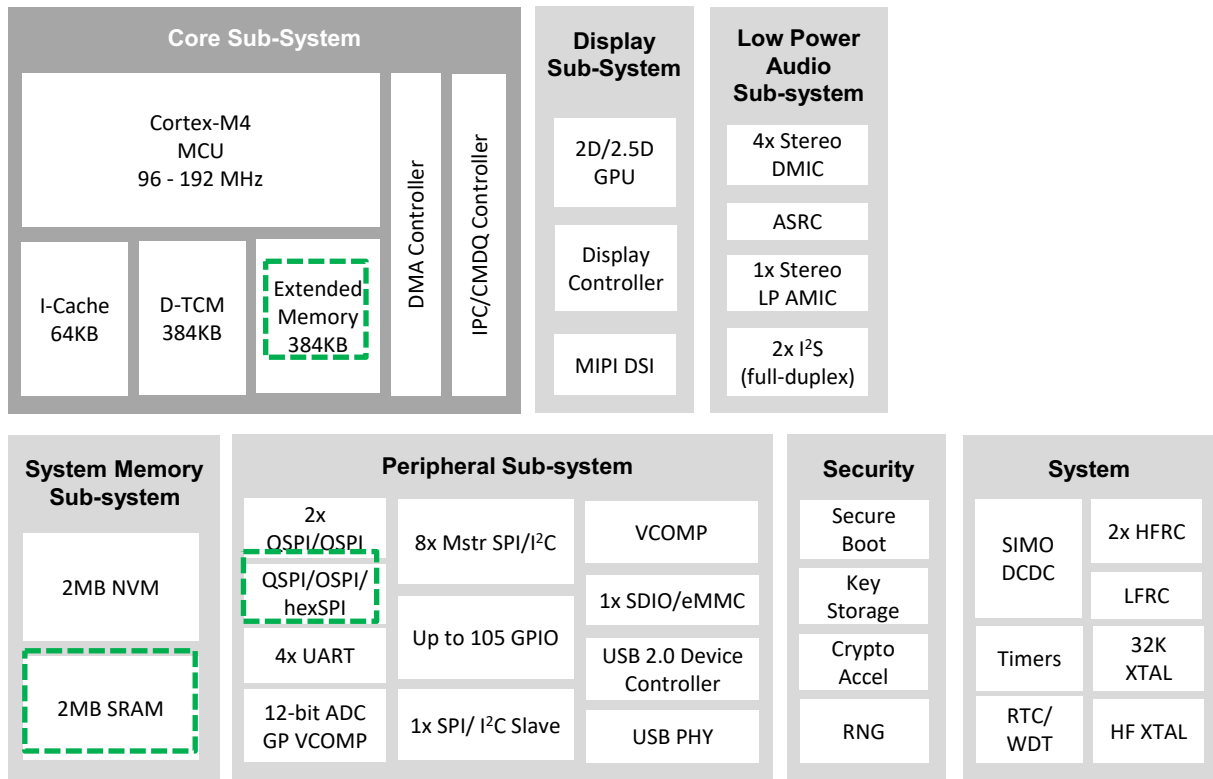
76ms Inference

The World's Most **Energy Efficient** Solutions



- ◆ A pioneer of ultra-low-power semiconductor solutions built on our **patented** and **proprietary** Subthreshold Power Optimized Technology (SPOT®) platform
- ◆ SPOT provides a game-changing, **multi-fold improvement** in energy efficiency
- ◆ Our integrated SoCs are designed to act as the **brains** for our end-customers' offering
- ◆ SPOT fuels further growth by defining and enabling **endpoint AI**

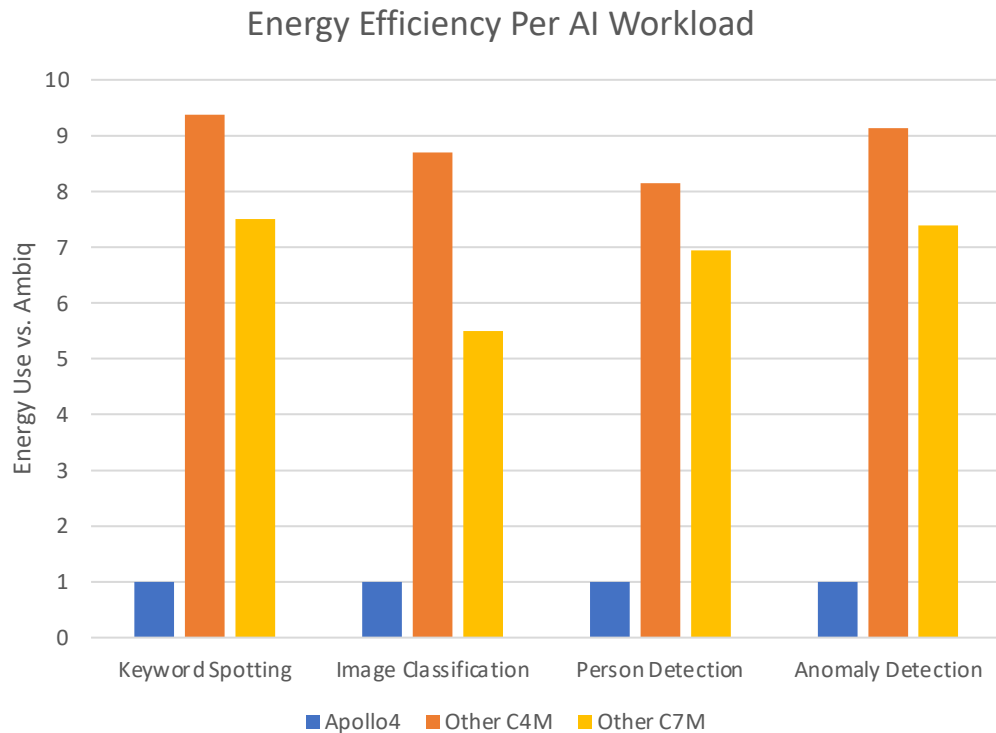
Apollo4 Plus



Feature Highlights

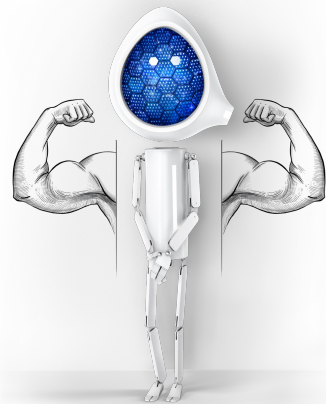
- 4μA/MHz executing from MRAM
- Up to 192 MHz clock frequency with TurboSPOT
- Improved data caching; up to 32 line buffers with programmable read/write data persistence
- 2MB MRAM, 2.75MB SRAM
- 2D/2.5D Graphics Accelerator with anti-aliasing and dithering
- MSPI enhancement with HexSPI
- Display Controller supporting MIPI DSI 1.2 with up to two lanes at 500Mbps
- 4-Layer alpha blending, up to 500x500 resolution with 60fps
- Ultra-low power analog microphone for truly always on voice processing
- 4 PDM stereo channels, 2 full duplex I2S channels with ASRC, and an ultra-low power ADC for analog mics

Ultra-low Power, High Performance Inference



Apollo4 Plus

Up to **9x**
better AI efficiency



Ambiq's Endpoint AI Models for Speech

VOICE ACTIVITY DETECTION

- Detects presence of speech
- Lightweight and power efficient
- Used to activate other, heavier models

KEYWORD SPOTTING

- Detect pre-programmed phrase (e.g. "ok computer")

SPEECH-TO-INTENT

- Detects intent in spoken phrases
- An intent is an action performed on an object
- "Turn on lights in kitchen"
- Phrase can be spoken in many forms

SPEECH ENHANCEMENT

- Eliminates background noise such as leaf-blowers, music, and restaurant noise from speech
- Can be used to increase intelligibility and improve speech recognition accuracy

SPEAKER ID

- Identify speech as belong to a specific individual
- Supports multiple individual
- Can be trained on new individuals on-device

Open Source

Real-time

OS-agnostic

Heavily
Documented

Ambiq's Endpoint AI Model for Health/Fitness

HEARTKIT®

- Detects heart statistics and conditions based on ECG sensor readings
- Heart Rate, HRV, Afib detection
- More added regularly

PHYSIOKIT®

- Coming Soon
- Optical-sensor based health statistics
- SpO2, Heart Rate, Afib, cuff-less blood pressure monitoring

SLEEPKIT®

- Coming Soon
- Used PhysioKit and IMU reading to detect sleep state

HUMAN ACTIVITY RECOGNITION

- Detects physical activity based on IMU readings
- Sit, Walk, Run, etc.

Open Source

Real-time

OS-agnostic

Heavily
Documented

Endpoint AI is Awesome



Optimal User Experience
with of Battery Life



Expanded Social Media
Video/Image Experience



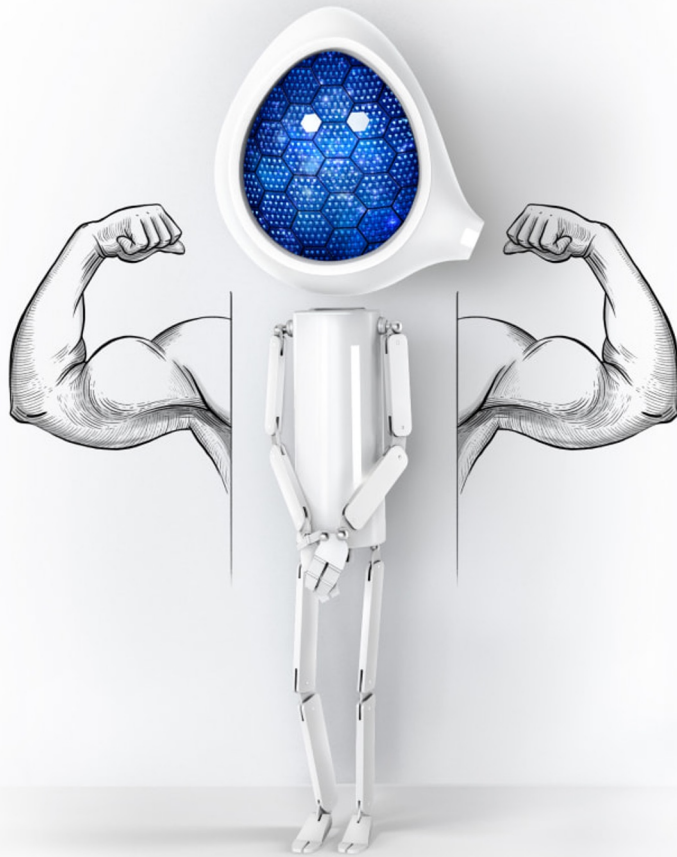
Natural Language UI



Intuitive **On-Device** Health and
Fitness Tracking and Coaching



Dazzling UI with
3D Graphics and 120fps





Thank You!