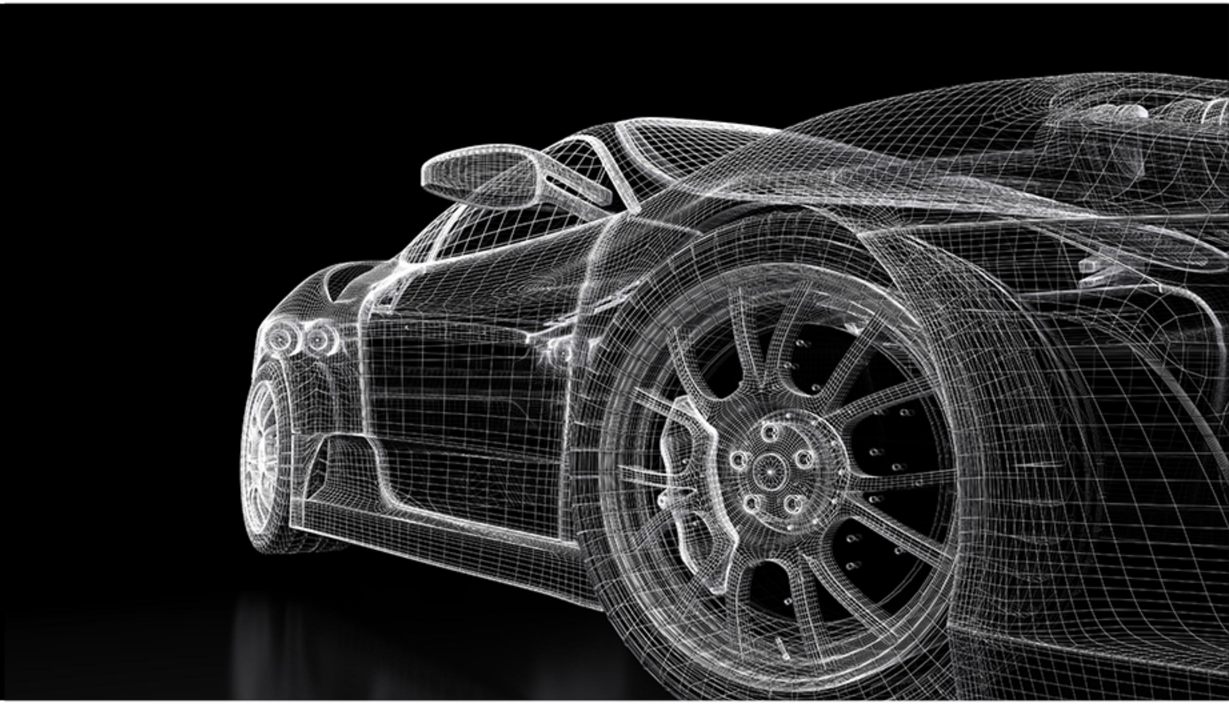
AUTOMOTIVE TECHNOLOGIES
virtual conference
Designing Tomorrow's Vehicles Today

How can efficient, low-latency and high-accuracy inference be performed in ADAS?

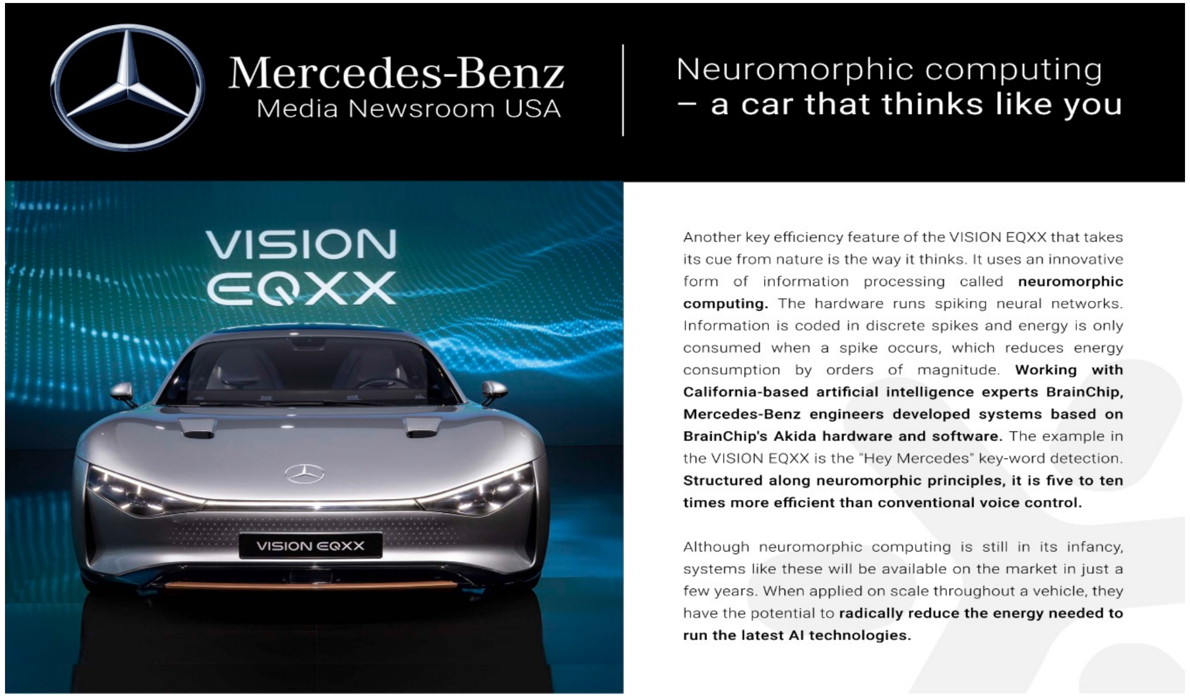


Kristofor Carlson, PhD

Manager of Applied Research

BrainChip Inc.

We Are The Future of Edge AI



Mercedes-Benz
Media Newsroom USA

Neuromorphic computing
– a car that thinks like you

VISION
EQXX
VISION EQXX

Another key efficiency feature of the VISION EQXX that takes its cue from nature is the way it thinks. It uses an innovative form of information processing called **neuromorphic computing.** The hardware runs spiking neural networks. Information is coded in discrete spikes and energy is only consumed when a spike occurs, which reduces energy consumption by orders of magnitude. **Working with California-based artificial intelligence experts BrainChip, Mercedes-Benz engineers developed systems based on BrainChip's Akida hardware and software.** The example in the VISION EQXX is the "Hey Mercedes" key-word detection. **Structured along neuromorphic principles, it is five to ten times more efficient than conventional voice control.**

Although neuromorphic computing is still in its infancy, systems like these will be available on the market in just a few years. When applied on scale throughout a vehicle, they have the potential to **radically reduce the energy needed to run the latest AI technologies.**

brainchip™
Biologically Inspired, Digitally Engineered

In-Cabin AI

- Visual driver authentication
- Keyword spotting
- Voice authentication
- Contextual understanding

The Standard for Edge AI

BrainChip is revolutionizing the future of in-device Artificial Intelligence (AI) and is the worlds first commercial producer of neuromorphic semiconductor chips and IP.

Our technology brings commonsense to the processing of sensor data, freeing machines to do more with less. Accurately. Elegantly. Meaningfully. We call this Essential AI.

Essential is optimizing compute. Maximizing performance. Minimizing power. In the real world. And in real time. We're proving that on chip AI, close to the sensor, has a sensational future, for our customers' products, as well as the planet.

BrainChip. Essential AI.

BrainChip Profile:

- 15 years of AI architecture research.
- World leading team of neuromorphic experts.
- Centers of engineering excellence in Australia, USA, France and India.

Trusted By:

MegaChips Mercedes-Benz NASA RENESAS Valeo

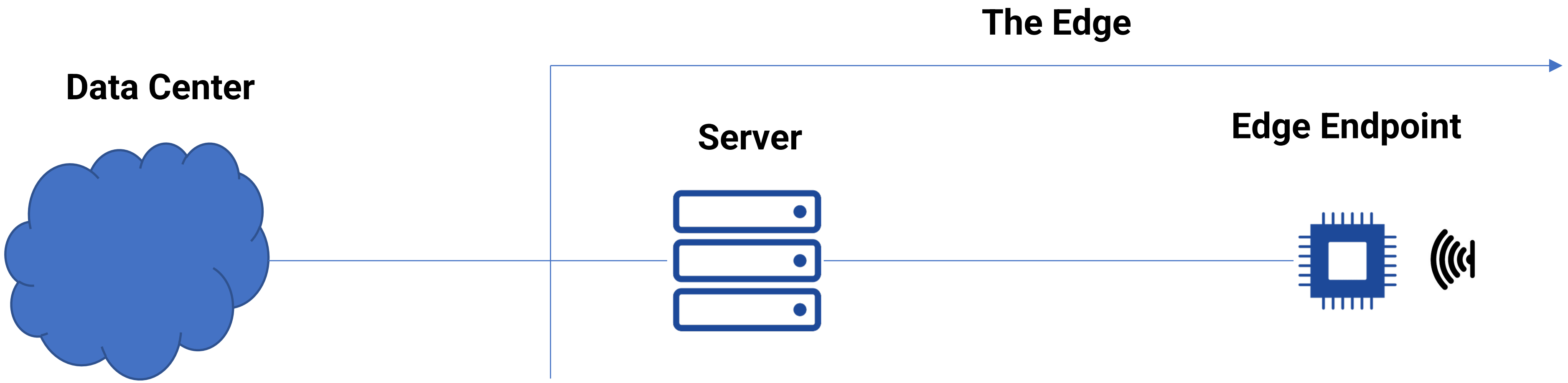
Profiled In:

VentureBeat EETimes Design & Reuse

Traded On:

ASX AUSTRALIAN SECURITIES EXCHANGE
OTCQX

Distributed AI

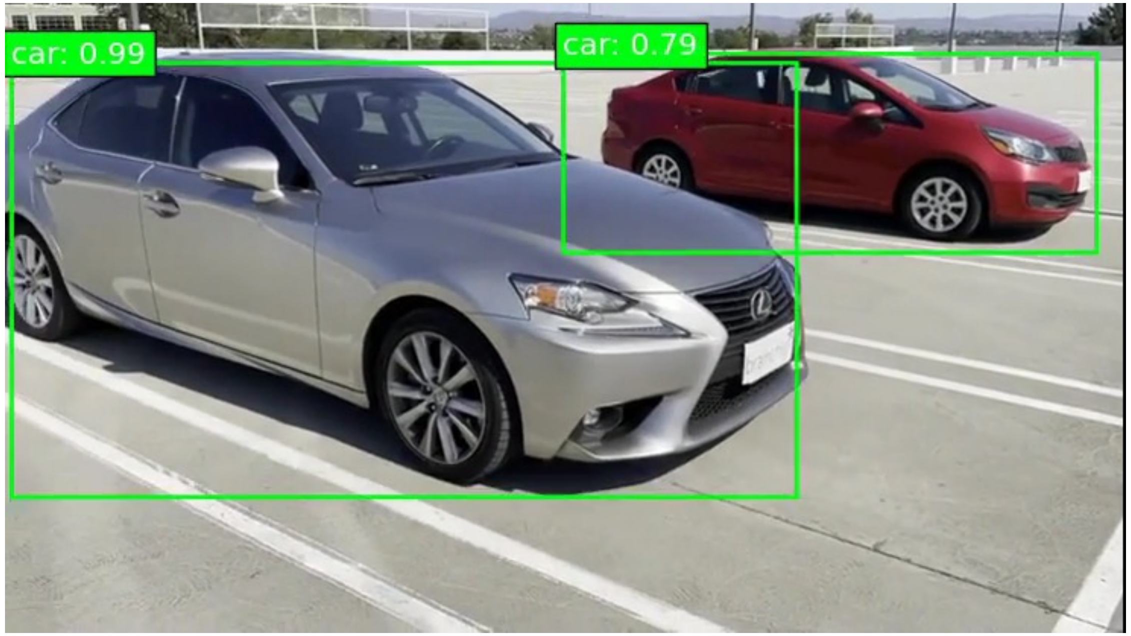


- Power intensive
- High latency
- Huge memory requirement
- Big data inference
- High bandwidth
- Privacy concerns

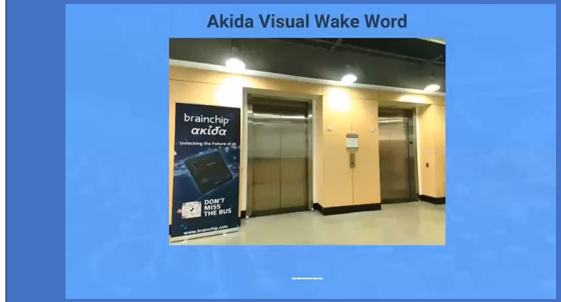
- Power efficient
- Ultra low latency
- Small memory requirement
- Small data inference, one shot learning
- Low bandwidth
- On-chip, in-device
- Privacy enabling

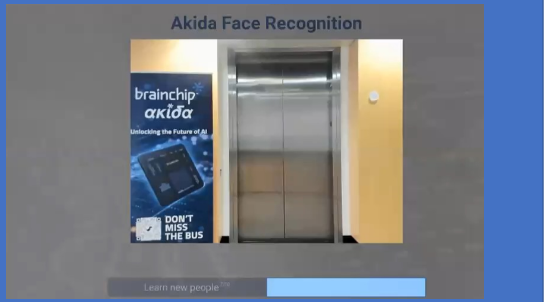
Edge AI In Action

Vehicle and Person Detection



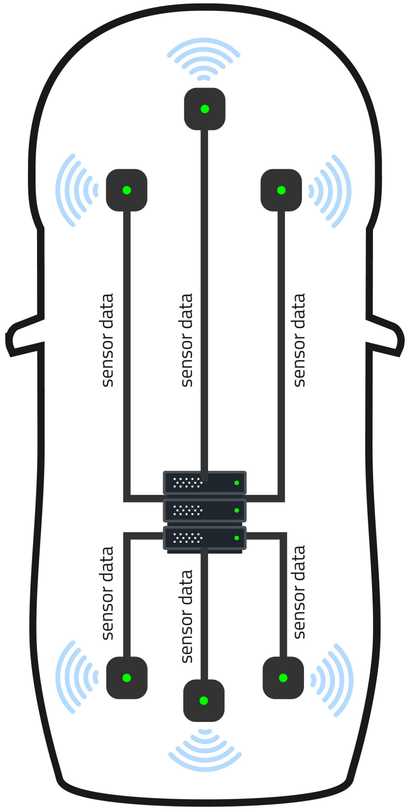
Visual Wake and Facial Recognition

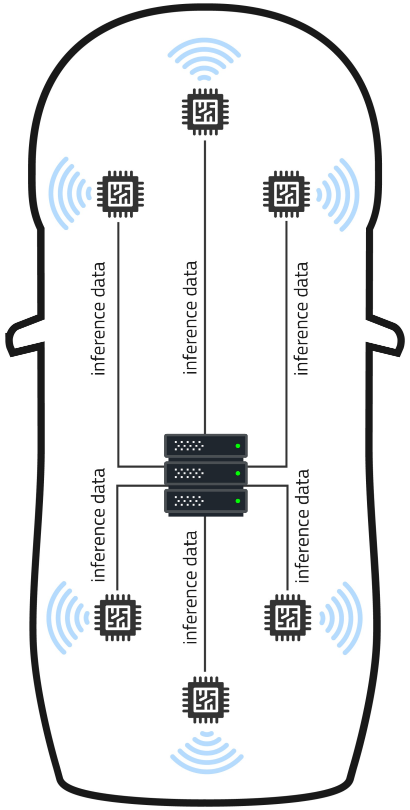




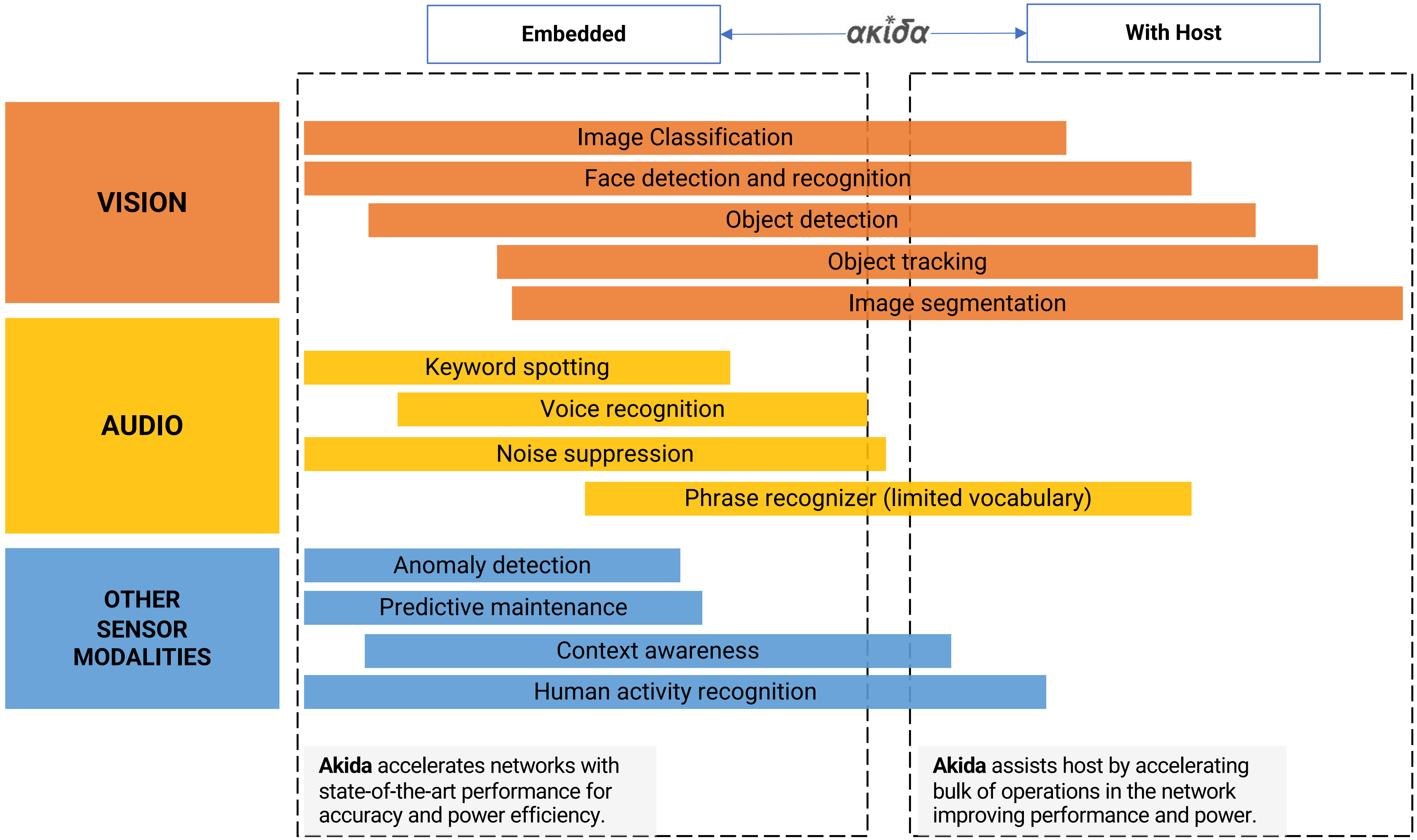
On-Sensor Computing

Conventional Computing vs **On-Sensor** Computing





AI Models to Support Key AI Use Cases



Embedded

ακιδα

With Host

VISION

Image Classification

Face detection and recognition

Object detection

Object tracking

Image segmentation

AUDIO

Keyword spotting

Voice recognition

Noise suppression

Phrase recognizer (limited vocabulary)

OTHER SENSOR MODALITIES

Anomaly detection

Predictive maintenance

Context awareness

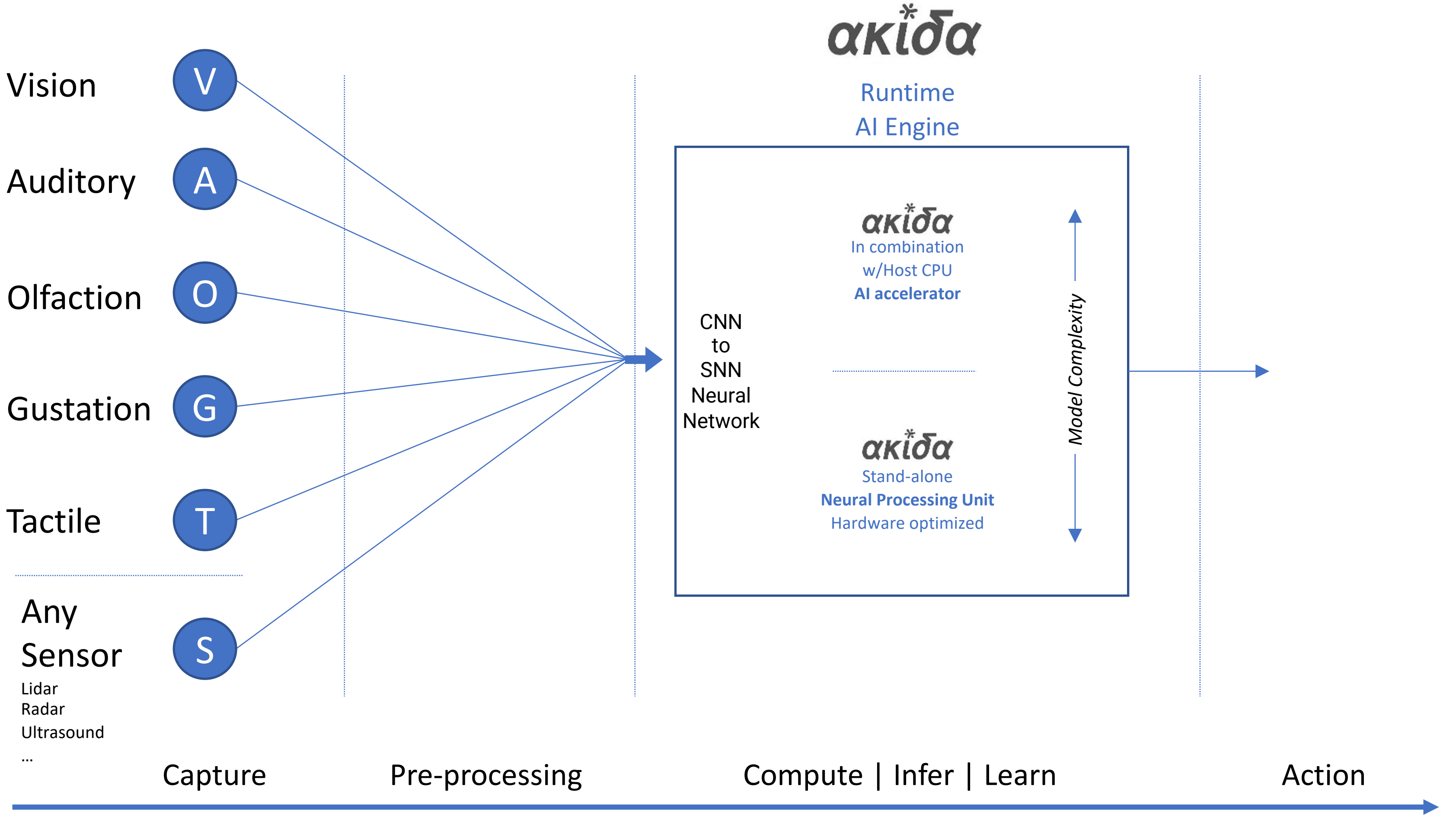
Human activity recognition

Akida accelerates networks with state-of-the-art performance for accuracy and power efficiency.

Akida assists host by accelerating bulk of operations in the network improving performance and power.

From Sensor to Inference

Performant and Efficient



- Akida supports output from any sensor, with its data and network, and applies in-chip neuromorphic AI to efficiently process and infer with radically reduced power consumption.
- Akida can also be deployed as an AI Accelerator to work in combination with a host CPU, still providing power consumption efficiencies.

Optimized compute, radically reduced power…

Features & Benefits

Power Efficient
Microwatts

Ultra Low Latency

Low Memory Footprint

One Shot Learning

Flexible IP & Quick to Deploy

Independent
Of the Cloud

Privacy Enabling
On-Device Processing

Environment
Less Power, Less Carbon

Tech Foundations

Distributed Computation

Each NPU has dedicated compute and memory, reducing data movement.

Event-Based Processing

NPUs perform computation only on events (non-zero values).

Event-Based Communication

Send events over mesh network without host CPU intermediation.

Event-Based Learning

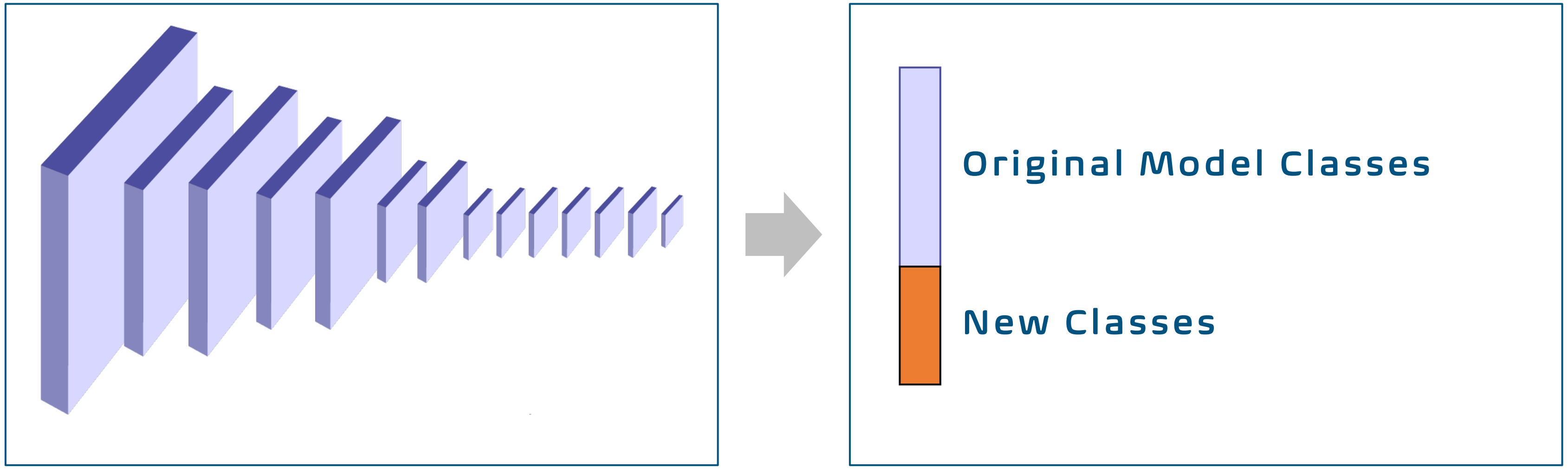
On-chip learning algorithm.

One Shot Learning On-Chip

- Brainchip IP and Akida chips are capable of on-chip learning by leveraging the trained model as a feature extractor then adding new classes to the final layer.
- Demonstrated Edge Learning for:
  - Object detection using MobileNet trained on the ImageNet dataset.
  - Keyword spotting using DS-CNN trained on the Google Speech Commands dataset.
  - Hand gesture classification using small CNN trained on a custom DVS events dataset.

Original Model's Extracted Features

On-Chip Learning Through New Final Layer



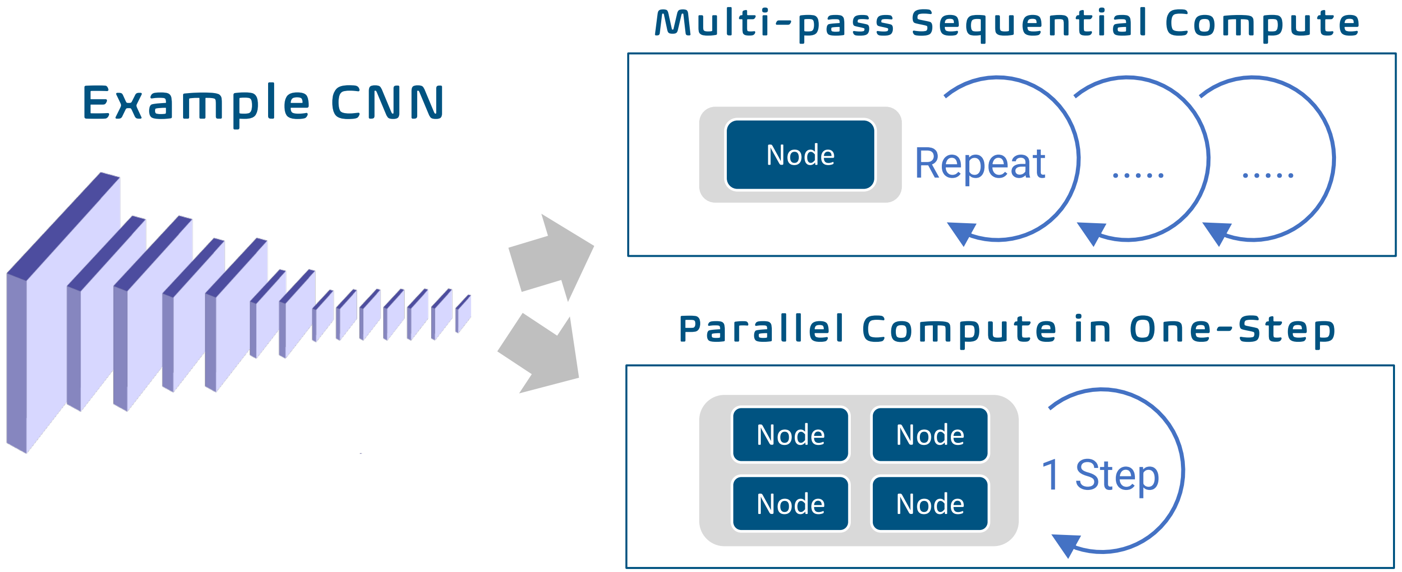
Multi-pass Processing Provides Scalability

- Multi-pass is a component of the Brainchip architecture that reduces the number of neural processing units required for a given compute task by segmenting and processing sequentially.

Benefits of Multi-Pass

- Scalable
- Smaller Memory Requirement *(2X)*
- Power Efficient

How it Works



Faster and More Efficient than Other Solutions

Task	Visual Wake Words	Image Classification	Keyword Spotting	Anomaly Detection
Data	VWW	CIFAR-10	Google Speech	ToyADMOS
Model	MobileNetV1 (0.25x)	Resnet-V1	DS-CNN	FC AutoEncoder
Harvard Arm Cortex-M4 w/FPU*	603ms / 24,320µJ/inf	704ms / 29,207µJ/inf	181ms / 7,373µJ/inf	10ms / 416µJ/inf
PCL RV32IMAC w/FPU*	846ms	1,239ms	325ms	14ms
Syntiant Cortex-M0*	-	-	6ms	-
Google Coral**	-	-	0.5ms / 351µJ/inf	-
NVIDIA Nano**	-	-	2.2ms / 659µJ/inf	-

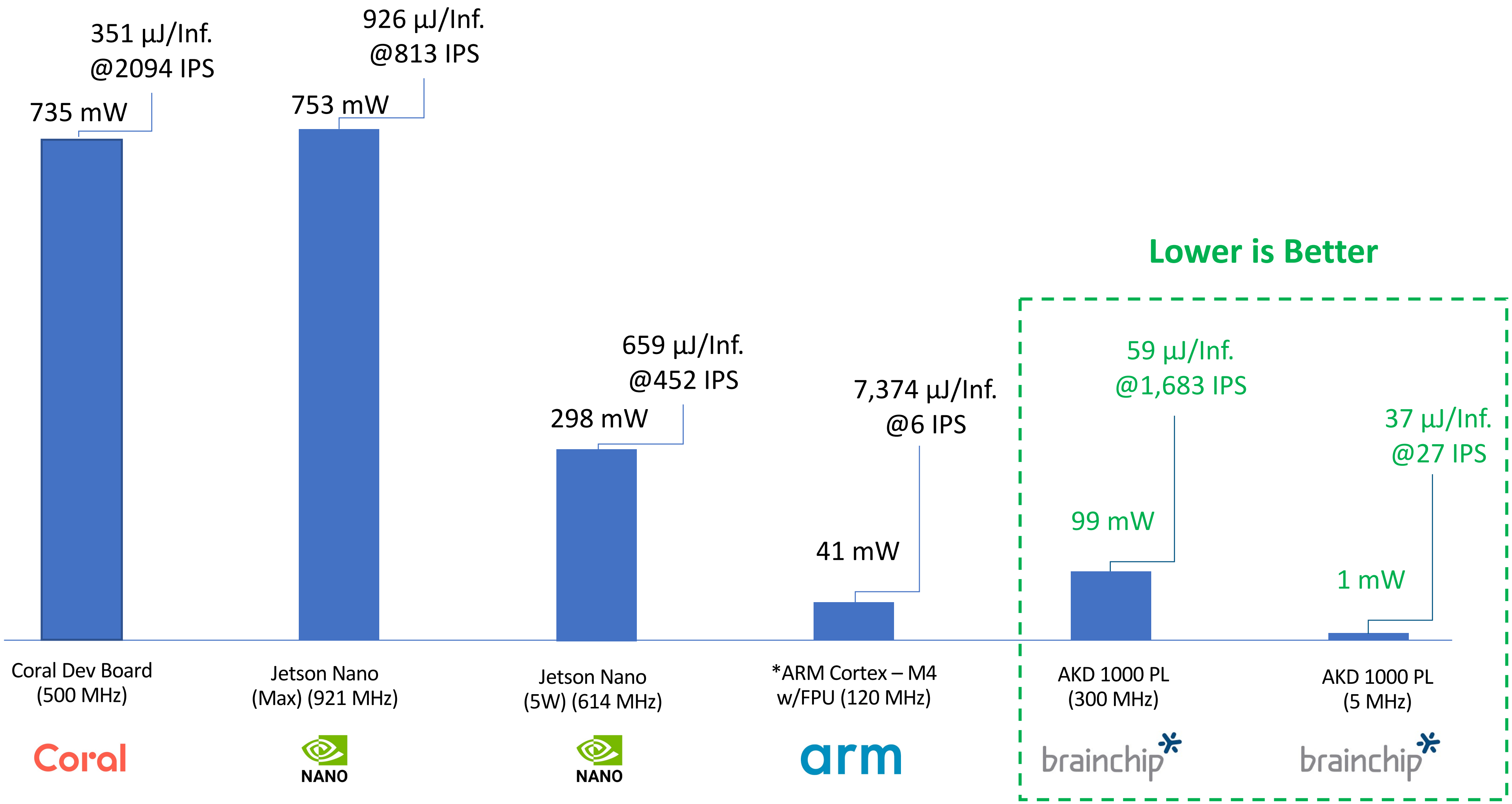
Brainchip Akida1000**	13ms / 259µJ/inf		3ms / 66µJ/inf	- / 4µJ/inf
Brainchip Improvement Over Best Performing Chip* **	**65x Faster 94x Less Power**		**Similar Speed 5-10x Less Power**	**104x Less Power**

*MLPerf™ v0.5 Inference Closed ResNet-v1.5 offline. Retrieved from https://mlcommons.org/en/infererence-tiny-05/ 20 April 2022, entries 0.5-464, 0.5-465, and 0.5-468. The MLPerf name and logo are trademarks of MLCommons Association in the United States and other countries. All rights reserved. Unauthorized use strictly prohibited. See www.mlcommons.org for more information.

**Result not verified by MLCommons.

***Brainchip Improvement denotes the Brainchip Akida1000 performance compared to the next best performing chip in each category given available data.

Power, Efficiency and IPS for Key Word Spotting



*MLPerf™ v0.5 Inference Closed ResNet-v1.5 offline. Retrieved from https://mlcommons.org/en/infererence-tiny-05/ 20 April 2022, entries 0.5-464. The MLPerf name and logo are trademarks of MLCommons Association in the United States and other countries. All rights reserved. Unauthorized use strictly prohibited. See www.mlcommons.org for more information.

** All other results were not verified by MLCommons.

Akida Neuromorphic Processor – Reference Chip

Akida AKD1000 is a reference chip implemented with TSMC at 28nm which proved viability of IP.

Data Input Interfaces

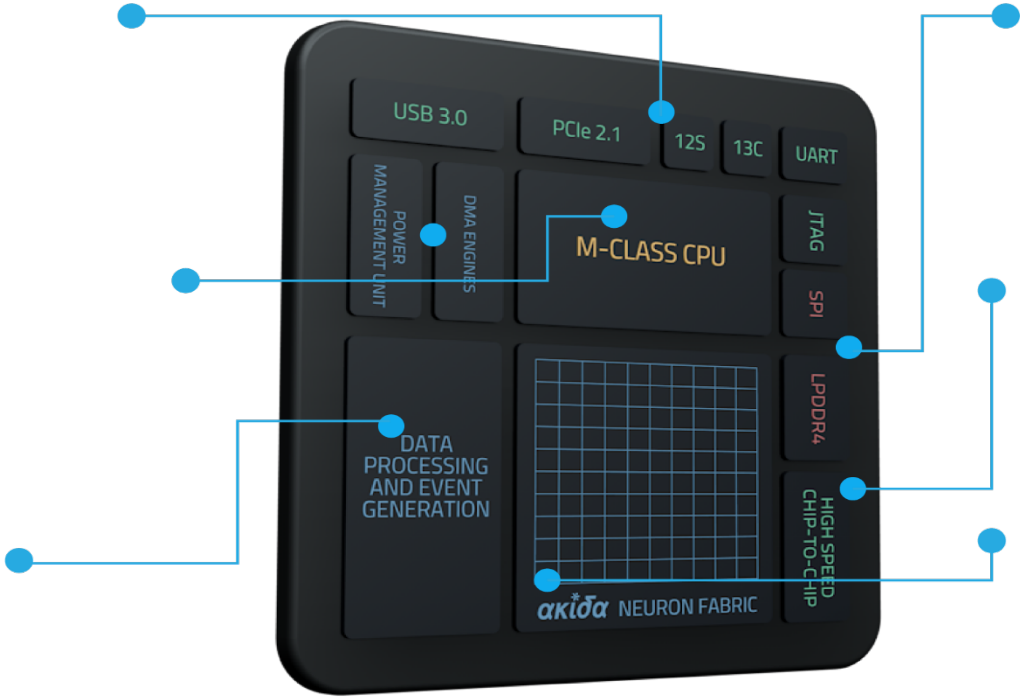
- PCI Express 2.1 x2 Lane Endpoint
- USB 3.0 Endpoint
- I3S, I2C, UART, JTAG

On-Chip Processor

- M-Class CPU with FPU & DSP
- System Management
- Akida Configuration

Data Processing

- Pixel-Event Converter
- SW Data-Event Encoder
- Any multivariable digital data
- Sound, pressure, temp., others



External Memory Interfaces

- SPI FLASH for boot/storage
- LPDDR4 Program/Weights

Multi-Chip Expansion

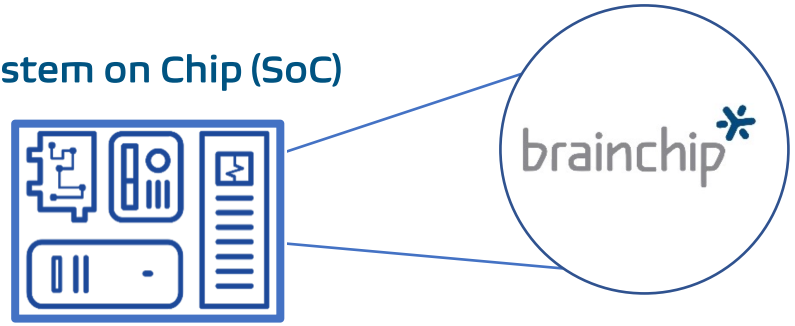
- PCIe 2.1 2 lane root complex
- Connects up to 64 devices

Flexible Akida Neuron Fabric

- Implements 80 NPUs
- All Digital logic with SRAM (8MB)
- Also Available as Licensed IP Core
- First Implementation: TSMC 28nm

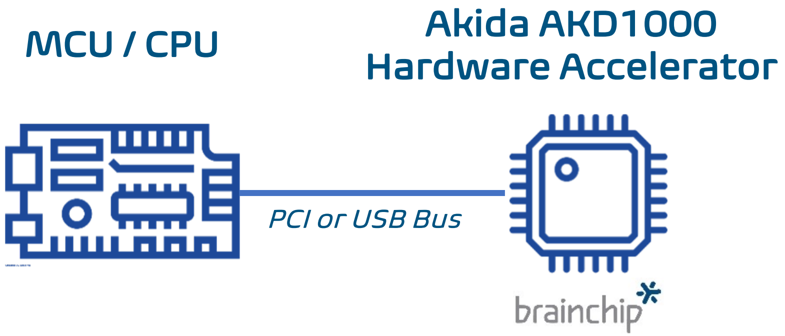
Integrate IP into SOC or Hardware Accelerator

IP Integrated into SoC



- Brainchip IP can be designed into a System on Chip (SoC) to seamlessly accelerate all AI workflows.
- Applications can be ported to SoC and be adapted to market requirements; cost, size, and power.

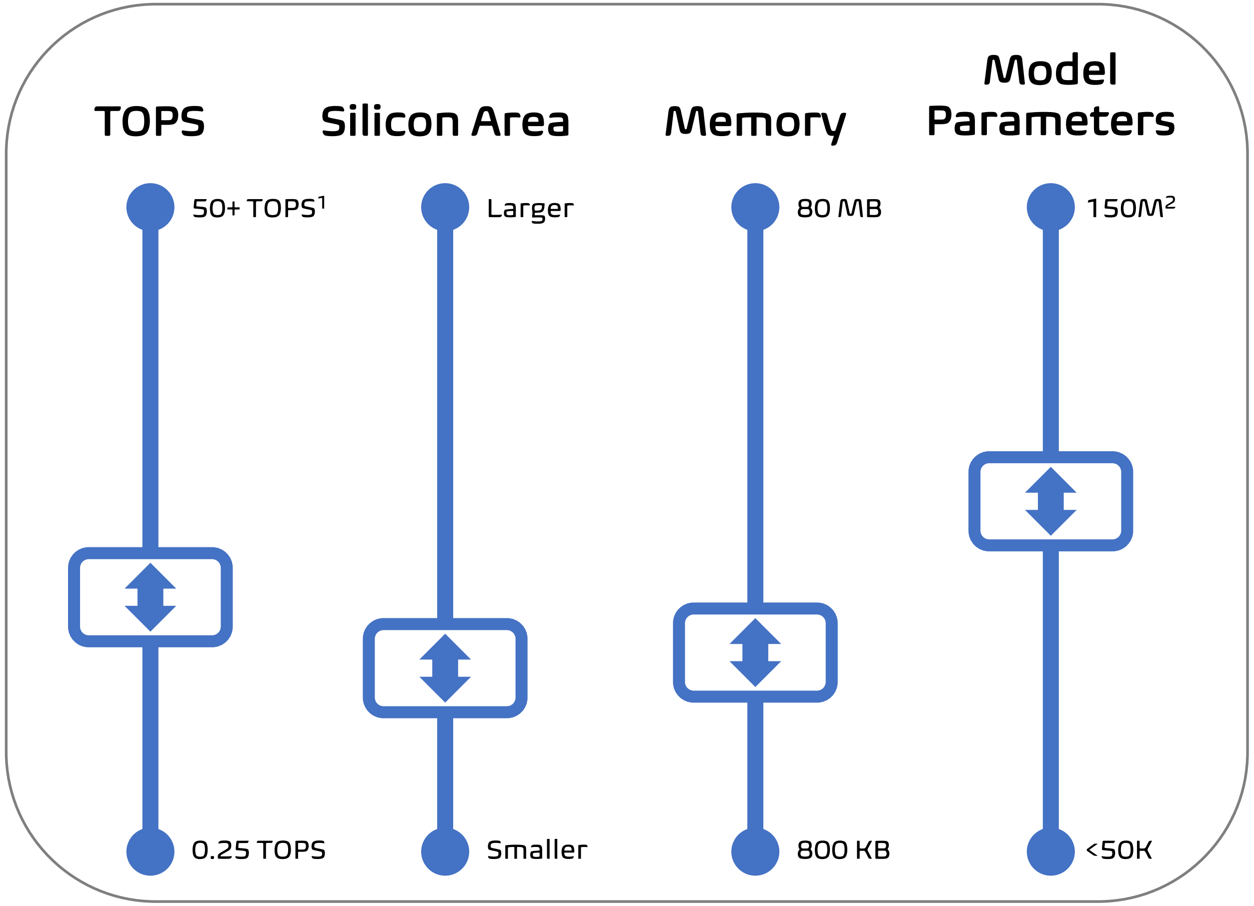
Reference Chip as a Hardware Accelerator



- An Akida AKD1000 with accompanying software development ecosystem can be placed next to an MCU or CPU where it can accelerate AI compute tasks.
- Brainchip's software development ecosystem and runtime is seamlessly compatible with any CPU; hardware and OS agnostic.

Highly Configurable IP Platform

BrainChip IP is flexible and scalable and can be implemented to support multiple edge AI use cases.



TOPS: 50+ TOPS[1] … 0.25 TOPS

Silicon Area: Larger … Smaller

Memory: 80 MB … 800 KB

Model Parameters: 150M[2] … <50K

- Brainchip works with clients to achieve the most cost-effective solution by optimizing the node configuration to the desired level of performance and efficiency.
- Scale down to 2 nodes for ultra low power or scale up to 256 nodes for complex use cases.
- Multi-pass processing provides flexibility to process complex use cases with fewer nodes increasing power efficiency.
- Quantization in MetaTF converts model weights and activations to lower bit format reducing memory requirement.

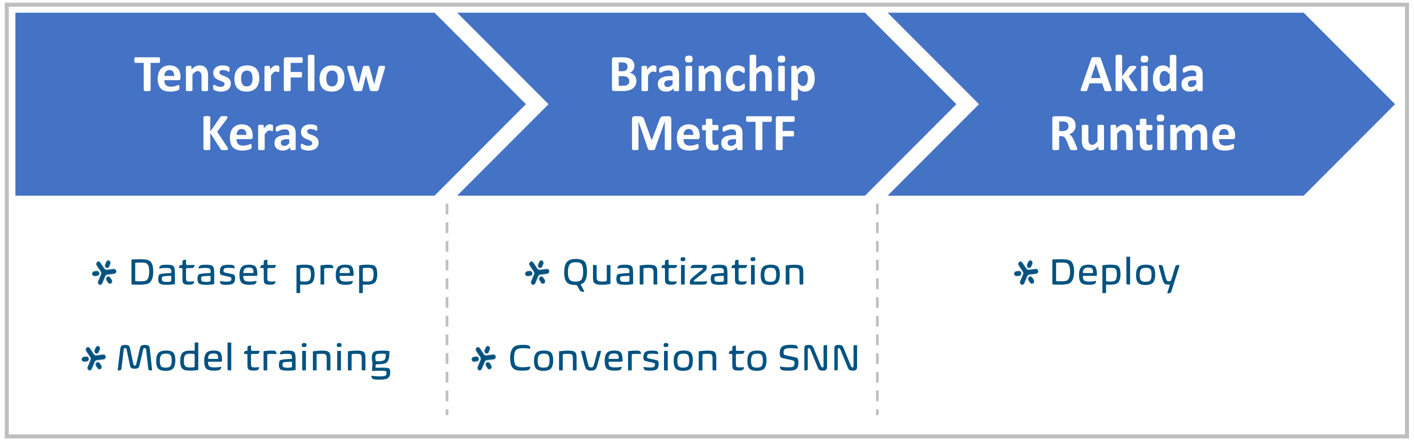
Notes:

1. 50+ TOPS is based on 100 Nodes at 1Ghz.
2. With optional SRAM and host CPU

MetaTF Easily Converts TensorFlow Models

Our open-source software to easily convert models for native Akida runtime.

- Create models in TensorFlow Keras.
- Easily convert TensorFlow Keras models with MetaTF API.
- No need to learn a new ML framework.



TensorFlow Keras

✲ Dataset prep

✲ Model training

Brainchip MetaTF

✲ Quantization

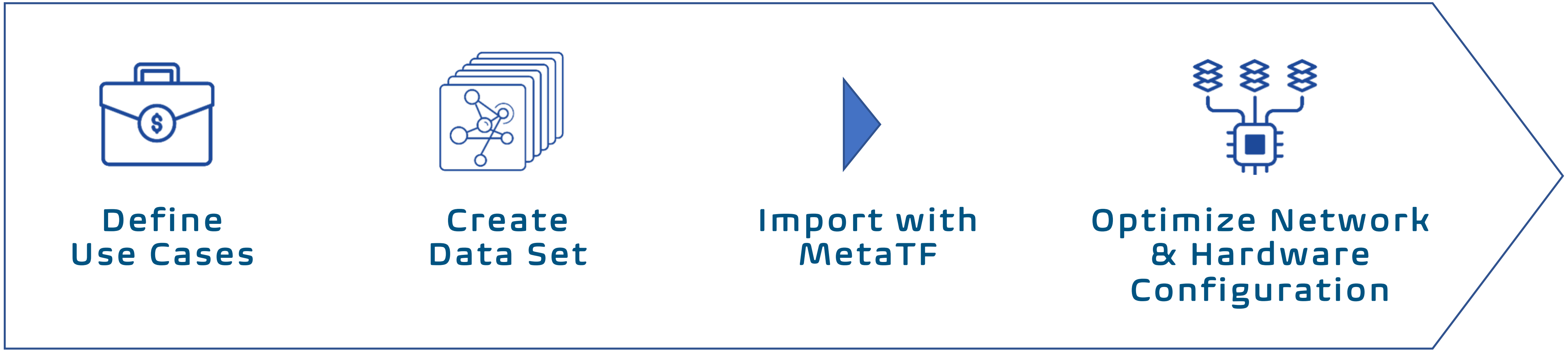
✲ Conversion to SNN

Akida Runtime

✲ Deploy

Consultative Enablement Process

We make AI enablement easy…



Many Cutting-edge Customers

Early Adopters

Mercedes-Benz

NASA

Valeo

VORAGO TECHNOLOGIES
Opening up new possibilities

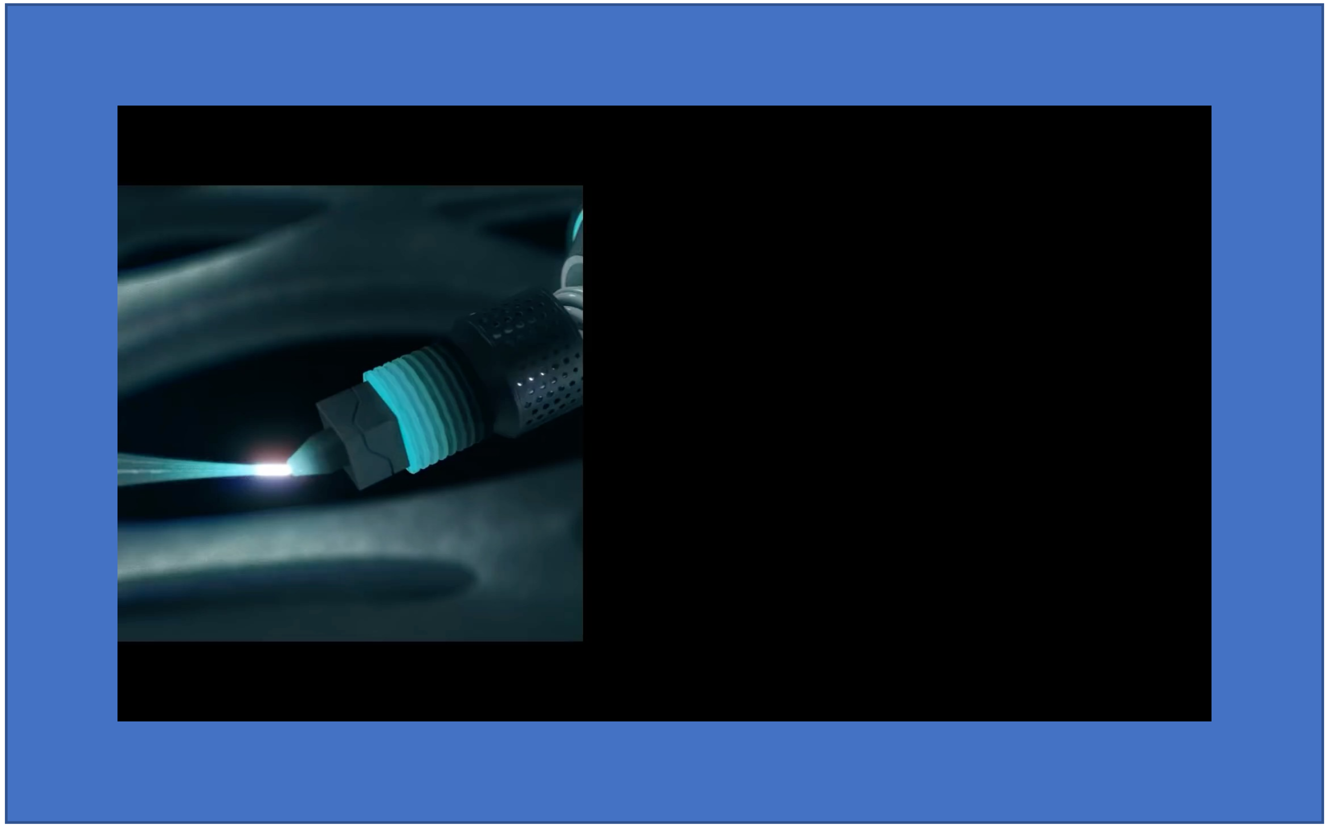
Licensees

MegaChips

RENESAS
BIG IDEAS FOR EVERY SPACE

Most customers cannot be identified due to Non-Disclosure Agreements.

Let’s Create the Future of Edge AI



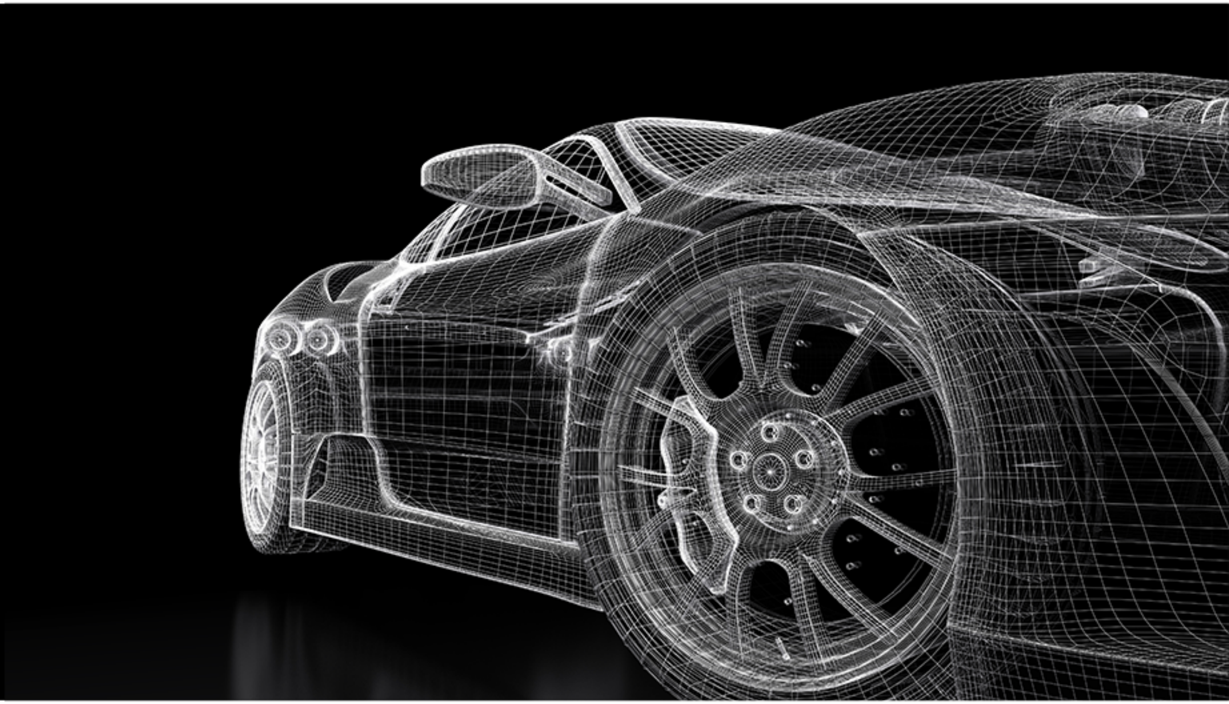
“Hey Mercedes”

AUTOMOTIVE TECHNOLOGIES virtual conference

Designing Tomorrow's Vehicles Today

THANK YOU.

QUESTIONS?



Kristofor Carlson, PhD

Manager of Applied Research

BrainChip Inc