ActualTech Media

brainchip™
Train. Infer. Learn.

How BrainChip Is Changing the AI Industry

Brien M. Posey

CONTENTS

IN THIS PAPER

BrainChip is a global technology company that has developed the revolutionary advanced Akida neural networking processor that brings artificial intelligence (AI) to the edge in a way that existing technologies are not capable of. This paper looks at the high-performance, small, ultra-low power solution that enables a wide array of edge capabilities that include on-device learning and inference. It also discusses how BrainChip is leading the industry to bring true intelligence to "artificial intelligence."

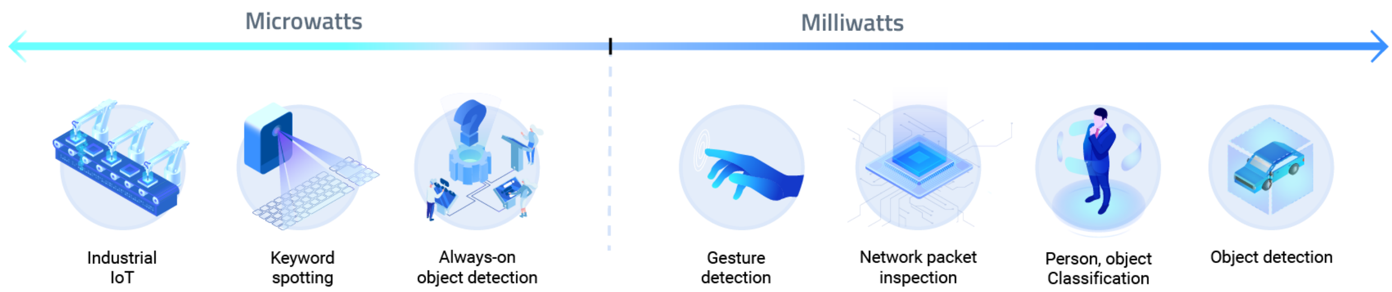

Figure 1: The Akida Neural Processor supports a wide range of use cases spanning from 100 micro-watts to a few hundred milli-watts

Over the last few years, manufacturers have been increasingly adding artificial intelligence (AI) and machine learning capabilities to a huge variety of products. Today, machine learning is used in everything from antivirus software to driverless cars. Even so, conventional techniques such as deep learning neural networks (DNNs) have their limits. DNNs depend on vast computing resources and require extensive training prior to use, which limits their use cases.

It has become increasingly clear in the evolution of AI that Neuromorphic Computing, inspired by the human brain, will drive the next generation of AI, especially the newly emerging area of edge AI. Based upon the fundamentals of Neuromorphic Computing, BrainChip has created the Akida event-based processor. The Akida processor's revolutionary architecture makes it possible to bring machine learning capabilities to edge devices while consuming ultra-low power, in the range of microwatts to low milliwatts—orders of magnitude more efficient than traditional approaches (**Figure 1**).

The Akida processor performs inference and incremental learning, requires no external process support or memory, and eliminates the need for connectivity to back-end resources. The combination of these features is essential for enabling mass adoption of edge AI where power constraints are limited to coin-sized batteries and energy harvesting, where personalization is required on each individual device through incremental learning, and where physical space is at a premium. Applications enabled by the Akida neural processor span a wide range of critical use cases.

The Akida Event-Based Neural Processor

The Akida event-based neural processor is a fundamentally different approach that breaks the linear relationship between high power consumption and performance seen in traditional accelerators. The Akida processor is 10x to 30x more energy-efficient than its nearest competitor for inferencing on industry-standard benchmarks such as MobileNet and Google Keyword Spotting DNNs, and is easy to use. Trained on MobileNet's Imagenet 1000 data set, the Akida neural processor can classify all 1.2 million images, and 1,000 classes, at 30 frames per second within a power budget of just 156 milliwatts in 28nm, compared to several watts for a Google Edge TPU. Audio keyword recognition using the Google keyword database runs at an extremely low power of 150 microwatts.

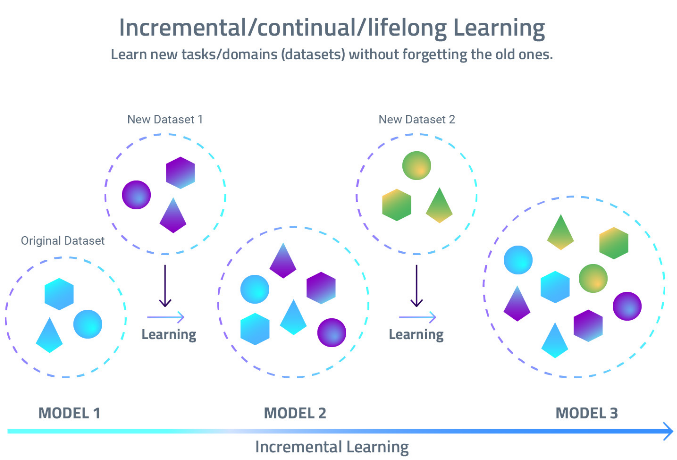

Figure 2: The Akida event domain neural processor supports multiple types of on-chip learning including incremental learning

The Akida event domain neural processor is capable of on-chip incremental, continuous learning so devices learn new data without the need for retraining in the data center—a breakthrough that fundamentally changes today's paradigm of deep learning neural network-based AI (**Figure 2**).

> The Akida event-based neural processor is a fundamentally different approach that breaks the linear relationship between high power consumption and performance seen in traditional accelerators.

BrainChip Team

BrainChip's ongoing success can be directly attributed to its staff. Members of BrainChip's Board of Directors and executive management team have more than 200 years of combined experience in the highly-competitive technology sector. BrainChip also employs a scientific advisory board consisting of neuroscientists, cognitive scientists, and other industry experts.

LOCATIONS

BrainChip is a multinational corporation headquartered in New South Wales (Sydney), Australia. BrainChip has two North American offices with expert chip development teams in Aliso Viejo and San Francisco, California. BrainChip's equally renowned software development teams are based in Toulouse, France and Hyderabad, India. BrainChip's Research and Innovation Center is in Perth, Australia. BrainChip's executive team and Board of Directors reside in Northern California, Southern California, Eastern Australia, and Western Australia. (See **Figure 3**.)

Figure 3: BrainChip corporate locations

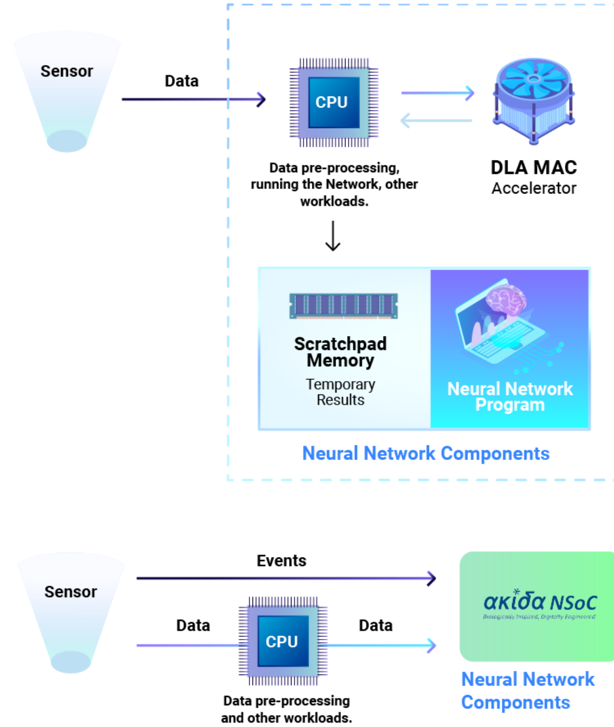

Figure 4: The Akida neural processor is fully integrated, requiring no host CPU to run the neural network algorithm, nor external memory to hold weights and intermediate values

BrainChip's Financial Strength

BrainChip was founded in 2013 and has been a publicly traded company on the Australian Stock Exchange (ASX) since 2015. It is traded under the ticker name BRN.

To date, BrainChip has received seven rounds of post-IPO equity funding totaling $39 million USD. These proceeds have been used to support specific achievements and milestones, the development of the Akida neural processor, and to fund the 2016 acquisition of SpikeNet, a French company focusing on computer vision systems.

Akida Product Offerings

BrainChip's flagship products in its Akida series are the Akida neural processor IP and the Akida neural processing System on a Chip (SoC). Both are engineered to mimic biological neuronal processes, delivering far better efficiency than conventional DNN solutions can provide (**Figure 4**).

AKIDA NEURAL PROCESSING IP

Inspired by the biological function of neurons, this event-based neural processor IP consumes far less power than traditional deep learning neural network accelerator IP that requires a host processor and external memory. When used with the BrainChip CNN2SNN conversion flow, the event-based neural processor enables CNNs to be implemented with very low power consumption and high throughput. Because the IP is based upon the fundamentals of neuromorphic computing, it enables unsupervised learning, allowing for autonomous edge applications and personalization at the edge (**Figure 5**).

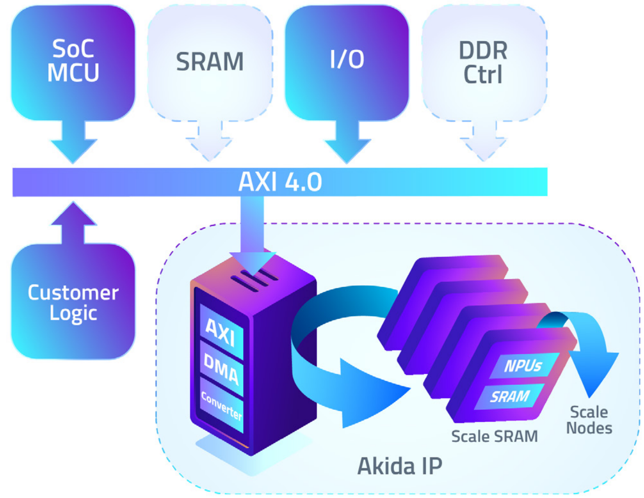

Figure 5: Typical SoC integrating the Akida neural processor. The solution can scale both the number of nodes as well as the embedded memory in each node.

AKIDA NEURAL PROCESSING SYSTEM ON A CHIP

The Akida NSoC represents a revolutionary new breed of neural processing computing for edge AI devices and systems. Each Akida NSoC has effectively 1.2 million neurons and 10 billion synapses, representing massively better efficiency than other neural processing devices on the market (**Figure 6**). Comparisons to leading DNN accelerator devices show better images per second per watt when running industry-standard benchmarks, while maintaining excellent accuracy.

Each Akida NSoC has effectively 1.2 million neurons and 10 billion synapses, representing 100 times better efficiency than neuromorphic test chips from other suppliers

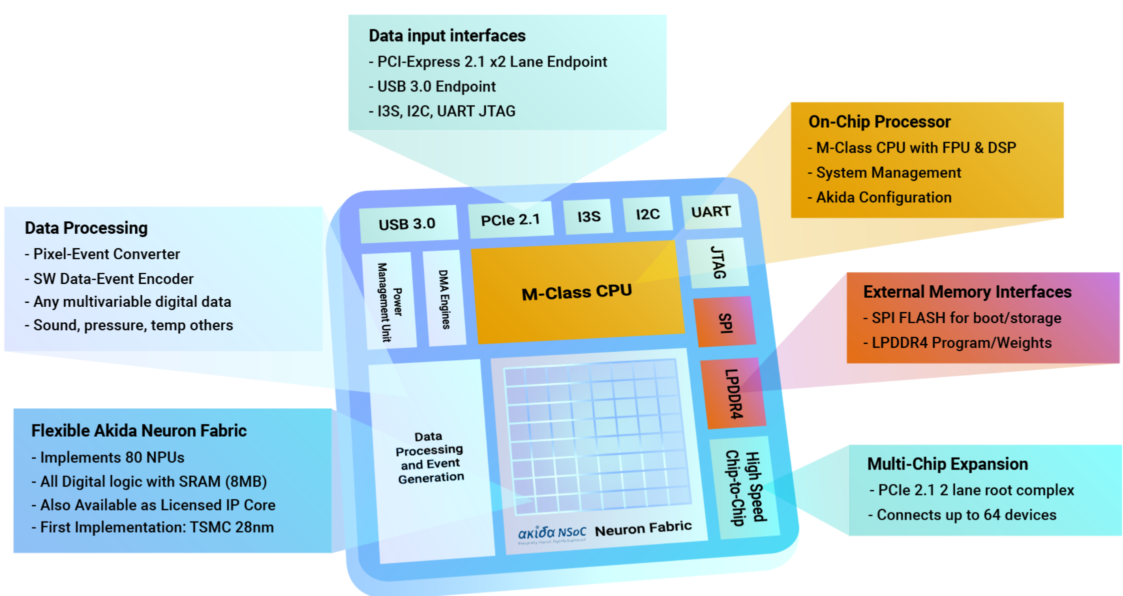

Figure 6: The Akida neural processing SoC

such as Intel and IBM. Comparisons to leading conventional neural network accelerator devices show similar performance gains of an order of magnitude better when running industry standard benchmarks such as CIFAR-10 and MobileNet, with comparable accuracy.

> Inspired by the biological function of neurons, this event-based neural processor IP consumes far less power than traditional deep learning neural network accelerator IP that requires a host processor and external memory.

The Akida NSoC is designed for use as a standalone embedded accelerator or as a co-processor. It includes sensor interfaces for traditional pixel-based imaging, dynamic vision sensors (DVS), Lidar, Radar, audio, and analog signals. It also has high-speed data interfaces such as PCI-Express and USB. Embedded in the NSoC are data-to-spike converters designed to optimally convert popular data formats into spikes to train and be processed by the Akida Neuron Fabric (**Figure 7**).

Product Roadmap

BrainChip's roadmap for the future includes a 2.0 release of its Akida processor in 2020, with a 3.0 release currently slated for a 2022 release. BrainChip is also planning a

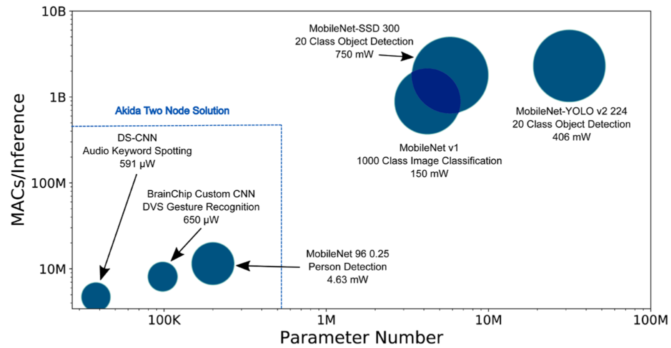

Figure 7: The Akida Neural Processing IP and NSoC is highly scalable and flexible. This chart shows a variety of solutions and the relationship between number of parameters, MACs/inference, and power for 30 FPS throughput.

2020 release of its NSoC AI Silicon product, with a 2.0 version planned for 2022.

2020 should also see the release of the 2.0 versions of Akida Scalable IP and the Akida Development Environment, with subsequent versions being released in 2021 and 2022. BrainChip is planning to release an application-specific version of its Akida processor for automotive applications in 2021. BrainChip's full product roadmap is illustrated in **Figure 8**.

Next-Gen Architecture

Although the Akida processor mimics certain biological processes, it does not fully align with the way that the human brain works. Current neural networks, both CNN and SNN, consist of successive "flat" layers of neurons. This is not how the brain is organized.

Product	2018	2019	2020	2021	2022
AI Architecture	**Akida 1.0** CNN and SNN Feed forward layers On-chip training & learning		**Akida 2.0** CNN, RNN, Transformer and SNN All layers Enhanced On-chip training & learning		**Akida 3.0** Learning, Experience, Organization
AI IP Solutions		**Akida 1.0 Scalable IP** 10uW-200mW Acoustic, IoT	**Akida 2.0 Scalable IP** 7uW-150mW Plus: Vision, System Health Monitoring	**Akida 3.0 Scalable IP** 5uW-100mW Plus: Behavior analysis	**Akida 4.0 Scalable IP** 2uW-50mW Plus: Behavior prediction
AI Development Tools	**Akida Development Environment (ADE) Version 1.0**	**ADE Version 1.5** Plus: Model Zoo	**ADE Version 2.0**	**ADE Version 2.5** Plus: Model Zoo	**ADE Version 3.0**
AI Test Chip			**Akida 2.0 NSoC** 28nm TSMC		**Akida 3.0 NSoC** 16 / 12nm

Figure 8: The Akida product roadmap

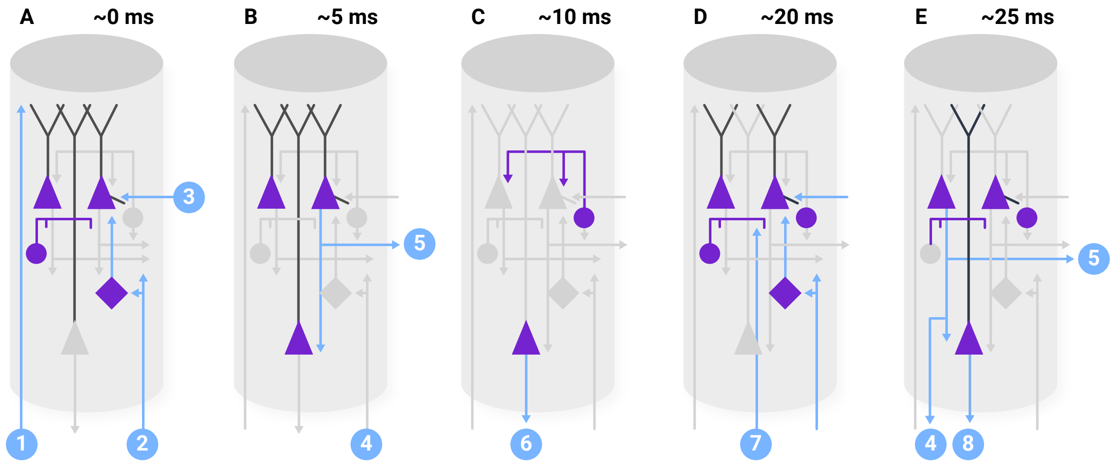

Figure 9: Neurons in the brain are organized in columns that trigger each other like dominoes, resulting in sequential memory and prediction of the next sensory perception a few milliseconds before it arrives. Prediction is an important facilitator of learning in the brain that has been overlooked for more than 30 years.

The neocortex accounts for about 76 percent of the brain's overall volume. The neocortex is a component of the cerebral cortex, and controls functions such as conscious thought, spatial and sensory perception, and motor commands.

Unlike current generation CNNs and SNNs, the neocortex is equipped with both excitatory and inhibitory neurons, and is divided into six separate horizontal layers. Each of these layers are tied together through neuronal connections.

The neocortex is organized into vertical columns called cortical columns. Cortical columns consist of dense vertical connections between neurons within a single column. Cortical columns also feature lateral neuronal connections to neighboring columns (**Figure 9**).

The neocortex column structure, combined with the fact that 80 percent of the neurons are excitatory and 20 percent are inhibitory, allows the neurons to perform predictive functions. Neuronal stimulation within the neocortex has the effect of partially depolarizing a neuron in an adjacent column. Neuronal depolarization refers to the point at which a neuron reaches its action potential, thereby releasing neurotransmitters across the synaptic cleft. This is the point at which the neuron "fires."

When a neuron partially depolarizes a neuron in an adjacent column, it puts that neuron into a predictive state. When sensory stimuli arrive a few milliseconds later, the prediction is either confirmed or denied. If the prediction is false, connections are weakened through an inhibitory response. If on the other hand the prediction is confirmed, the connection between columns are strengthened by increasing weight values. This process, known as Spike Time Dependent Plasticity (STDP), continues to form linked lists of cortical columns, forming sequence memory. BrainChip believes that this process exists everywhere in the cortex, including the visual cortex, the sensory-somatic cortex, and the prefrontal cortex.

> "We're going through a transition right now in the world of machine intelligence that's similar to the transition from analog to digital computing back in the 1940s."
>
> — *Jeff Hawkins, inventor of the Palm Pilot*

The company eventually expects machine learning to evolve into a new artificial neural network architecture, a spiking cortical network, that is superior to current

CNN and SNN architectures and appropriate for industrial application in visual object learning and classification, big data analysis, feature extraction and classification, industrial automation, and robotics. A spiking cortical neural network is capable of real-time learning, and will require significantly less training time than a CNN.

Much work needs to be done to reach this goal. After all, each neuron in the human brain has 10,000 synaptic connections to other neurons. Despite the complexity, however, some progress has been made toward creating a spiking cortical network.

> The solution is high-performance, small, ultra-low power and enables a wide array of edge capabilities that include on device learning and inference.

BrainChip has created and tested a simple multi-compartment neuron that is related to the current Integrate-and-Fire (I&F) neural model used in the Akida processor. This neural model makes it possible to partially depolarize the model without generating a spike. A realistic connectome between Akida II neurons within cortical mini-columns and hyper-columns needs to be developed, as well as the lateral connections between columns. So far, 58 journals and publications on columnar organization have been reviewed to form a processing theory. It is expected that this work, including building models of the processing method, will continue for another nine months before a tangible result is attained.

The Industry Leader

BrainChip is a global technology company that has developed the revolutionary advanced Akida neural networking processor that brings AI to the edge in a way that existing technologies are not capable of. The solution is high-performance, small, ultra-low power and enables a wide array of edge capabilities that include on device learning and inference. BrainChip is leading the industry to bring true intelligence to "Artificial Intelligence."