THE **AIOT** 'BIG BANG'

With billions of devices, using zetabytes of data, coming to market a new class of processor looks to meet the demands of the growing AIoT. By **Neil Tyler**

With devices getting smarter and intelligence now becoming an essential element in homes, vehicles, and workspaces, what's referred to as the Artificial Intelligence of Things (AIoT) is fast becoming a fact. According to market research, by 2025 there is expected to be 65 billion connected devices generating 180 zetabytes of data, all of which will require complex and diverse processing capabilities.

That brings with it a number of challenges for design engineers who will have to address problems such as speed, reliability and security. The growth in the AIoT will also see more AI-enabled decisions taking place on-device as opposed to in the cloud which is energy hungry and expensive.

Too many of today's smart products are reliant on processing in the cloud and the growing adoption of natural voice interfaces, imaging and presence detection, for example, not only raise performance issues but will create further challenges in the form of reliability, privacy and cost.

In fact, as voice control becomes mainstream issues around latency will be exacerbated as the number of connected devices grows.

"Device proliferation and the growing diversity of demands means that there has been a need for a new type of processor," suggests Mark Lippett, CEO of XMOS, the Bristol-based fabless semiconductor company behind voice solutions, audio products, and multicore microcontrollers.

"There is a huge market opportunity for a device that is able to address the needs of a range of applications delivering both performance and functionality while, at the same time, offering ease of use, low power and real-time operation.

"This is a space that's dominated by performance and price and there has, and will remain, a necessary process/performance trade-off that will need to be addressed. What's required is the ability to match the features of a processor to the requirements of the end product and then ensure that it's affordable," he explains.

Lippett makes the point that there's no point in coming up with a solution that costs too much to be deployed. "First and foremost, you have to keep a very careful eye on cost."

According to Lippett too many devices are not delivering the levels of performance needed.

"Devices increasingly need to be on all the time, so power management is critical. But there's also a third comparator," according to Lippett, "and that's the growing diversity of customer demands. That is making it increasingly problematic for companies who are looking to differentiate their products in what are becoming commoditised markets, especially those that lack the necessary flexibility to address dynamic markets that require products getting to market quickly."

The company's xcore.ai is its response to this fast developing AIoT market and it has been designed to deliver high-performance AI, DSP, control and IO in, critically, a single device.

"To date these types of capabilities have tended to be deployed either using a powerful (and costly) applications processor or a microcontroller with additional components to accelerate key capabilities," according to Lippett.

"What we've been able to do with the xcore.ai is to provide a crossover processor that can deliver real-time inferencing and decision making at the edge, as well as signal processing, control and communications.

"This will enable manufacturers to integrate high-performance processing and intelligence economically into their products."

The device employs deep neural networks that use binary values for activations and weights rather than full precision values, which has helped to dramatically reduce execution time.

"By using binary neural networks, we're able to deliver 2.6x to 4x more efficiency than is the case with traditional 8-bit counterparts," says Lippett.

A new generation of embedded platform

According to Lippett, the xcore.ai, "heralds an entirely new generation of embedded platform. We've designed it to be versatile, scalable, cost-effective and easy-to-use."

Fast processing and neural network capabilities means that the xcore.ai can process data locally, within nanoseconds.

"In the evolving AIoT ecosystem, that capability means that manufacturers can build smarter
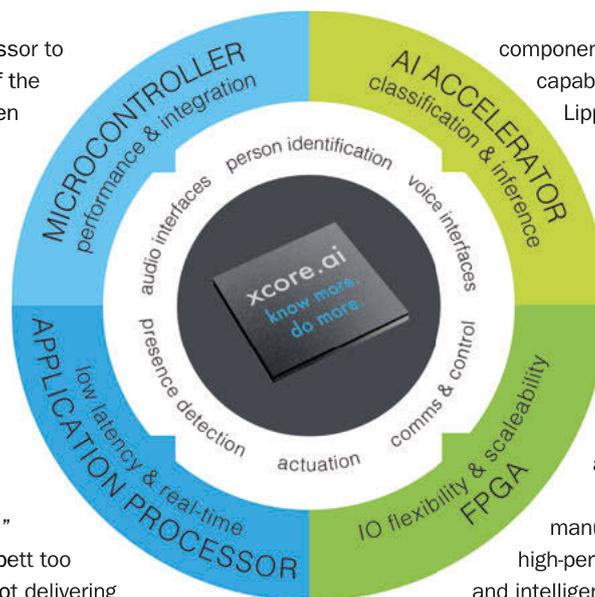
Above: The processor has been designed to deliver performance, flexibility, scalability and is AI enabled

> **"Device proliferation and the growing diversity of demands means that there has been a need for a new type of processor."**
> Mark Lippett

sensing technologies that will be able to fit seamlessly into smart devices," Lippett explains. "Not only that, the xcore.ai is delivering, what we believe, is record processing power at the dollar price-point. That means electronics manufacturers, whatever their size, will be able to embed multi-modal processing into all sort sorts of smart devices."

The xcore.ai looks to deliver solutions that will be able to address challenges associated with the AIoT, such as latency, connectivity, privacy and energy consumption, while at the same time keeping costs low and, crucially, keeping design potential high, according to Lippett.

"At XMOS we aspire to be the only digital piece of silicon in the box."

The device is flexible so that it can be used across a wide range of markets from asset tracking to personal health and well-being, as well as a host of smart appliances in the lighting, security and audio-visual space.

But what does this all mean in reality? Can the xcore.ai help to change the user experience for the better?

Lippett highlights its versatility with the example of a smoke detector.

"The smoke detector was designed to ensure our safety. They were developed over a hundred years ago but while they have come a long way, are they able to do enough?

"While they may alert us to danger and we instinctively know to exit the premises when we hear one, how can we truly optimise the smoke detector to dramatically reduce risk and support rescue and recovery?"

According to Lippett, by using the xcore.ai it will be possible for a smoke detector to deploy radar and imaging to identify whether there are people in an affected building and, if so, determine how many there are and where they are located.

"Using voice interfaces, the detector could communicate with those inside, while vital sign detection could identify whether they are breathing. Put together, this builds an intelligent picture of the environment that can be fed straight to the emergency services, enabling an informed rescue operation, improving accuracy and speed of response," he explains.

Another possible example comes with the need to develop affordable, unobtrusive and easy-to-use healthcare solutions for a fast ageing population.

"Think of a smart personal health

Above: The xcore.ai processor

companion that can constantly monitor vital signals giving the earliest possible indication of the need for specialist care," explains Lippett.

"Personal data is held securely on device, without being sent to the cloud, and can be shared with trusted medical or care staff through appropriate permissions. The system can even detect medical emergencies such as a fall, and take appropriate action without delay."

Scalable and multi-core, the xcore.ai processor is able to deliver improved levels of performance and enables embedded software engineers to deploy every different class of processing workload on a single multicore crossover processor that's able to interpret data without having to communicate with the cloud.

"It's fully programmable in 'C' and there are specific features such as DSP and machine learning that are accessible through optimised c-libraries," explains Lippett.

"It has 1Mb of embedded SRAM on chip, which is complemented by an LPDDR interface providing simple memory extension where required.

"It also supports the FreeRTOS real-time operating system, enabling developers to use a much broader range of familiar open-source library components. It is also compatible with TensorFlow Lite which allows easy prototyping and deployment of neural network models."

In terms of connectivity xcore.ai has up to 128 pins of flexible IO giving designers access to a much wider variety of interfaces and peripherals, which can then be tailored to the precise needs of the application.

"We've also integrated hardware, such as USB 2.0 PHY and an MIPI interface, so it's possible to collect and process data from a wide range of sensors."

"What we've looked to deliver with the xcore.ai is an economic solution that combines higher performance and flexibility with energy efficiency " says Lippett.

Left: A block diagram of XMOS's xcore.ai processor