

INTELLIGENCE UPDATE

Emerging tech: neuromorphic computing



Max Smolaks

22 Oct 2025

EMERGING TECHNOLOGY SERIES

This report is one of a series on emerging and potentially disruptive technologies that may be deployed in digital infrastructure. Here, Uptime Intelligence considers the use of neuromorphic computing to reduce the power demands of AI workloads radically.

EMERGING TECHNOLOGY: NEUROMORPHIC COMPUTING	
Key innovation	Neuromorphic computing could lead to the creation of smaller, faster, and orders of magnitude more power-efficient AI accelerators.
Potential impact (1 is low, 5 is high)	4 If the technology can be scaled and commercialized, it will change how most AI applications are developed and deployed.
State of maturity	Deployments are mostly limited to research, academic and government organizations; commercial systems are becoming available.
Main drivers to deployment	New types of models could advance state-of-the-art AI capabilities. More efficient hardware architectures could solve the challenges of powering and cooling AI infrastructure.
Main barriers to deployment	Neuromorphic systems are yet to demonstrate a clear advantage over conventional systems. It is challenging to program neuromorphic systems; there are few frameworks, tools and developer environments.
Possible date range for wide deployment	2030-2040
Key innovators/ companies involved	Chip developers: IBM, IMEC, Innatera, Intel, SpiNNcloud, SynSense, Zhejiang Lab. Academic centers: CalTech, Heidelberg University, Technical University of Dresden, Stanford University, Zhejiang University.
Key companies or type of companies most affected (positively or negatively)	Emerging computing architectures will threaten some major chip designers and manufacturers; others will take advantage of them. Power grids will regain stability as data center developers return to lower-density designs.

UPTIME INSTITUTE 2025

uptime
INTELLIGENCE

Context

AI workloads are responsible for a large and rapidly growing share of data center power consumption, increasing the strain on grid operators that cannot build new capacity quickly enough.

At the same time, the high power and cooling requirements of individual rack-scale GPU-based systems are forcing operators to make changes to established data center designs — increasing cost, complexity and risk.

Research into neuromorphic computing could lead to the creation of smaller, faster and orders of magnitude more power-efficient accelerators for AI, while simultaneously offering a solution to the end of traditional silicon scaling techniques.

The future of data centers for AI is expected to be dominated by gigawatt-scale facilities that deliver hundreds of kilowatts to each rack. Neuromorphic computing proposes an alternative — where AI workloads are run in small and efficient facilities filled with systems that consume less power than traditional enterprise IT.

The technology

Neuromorphic computing is a relatively new research discipline, established in the late 1980s to design chips, systems and software that are based on the structure and operating principles of the human brain. These systems use silicon and software representations of artificial neurons and synapses to build spiking neural networks (SNNs) that abstract the characteristics of electrical signals in the brain into mathematical functions.

The human brain is an extremely capable, very efficient biological computer estimated to run on just 20 W of power. The goal of neuromorphic computing is not to fully replicate its functions, but to use what we know to create a practical computing system.

There are two broad approaches to neuromorphic systems:

- **Analogue-digital mixed-signal circuits.** This is where physical neurons are implemented in silicon.
- **Digital circuits.** This is where neurons are simulated in software using specialized chips with many conventional compute cores.

In both types of systems, the computing architecture differs from the Von Neumann design principles that mainstream computers follow — where a compute function is distinct from an operating memory that stores all data, including instructions (software code). In neuromorphic architecture, there is no such distinction. Arithmetic operations and data share the same circuits, and data is operated on in situ, much like how the organic brain works.

The architectural benefit of neuromorphic computing is the removal of data movement to and from the processor's arithmetic units, and complex control logic that directs these flows. A Von Neumann machine, even if fully integrated on a single piece of silicon (die), moves data from memory arrays through control logic to arithmetic units for processing, then stores the results. Inevitably, these movements increase latencies and consume power.

This so-called Von Neumann bottleneck adds inefficiencies and concomitant performance limitations — computing resources often stall during execution as data is missing. The complexities of modern microprocessors and GPUs are rooted in ongoing design efforts put into optimizing data flows to overcome the Von Neumann bottleneck. Much of the emphasis in creating super-dense supercomputing clusters to train the largest generative AI models is on providing the low-latency and extreme bandwidth needed to move data around.

Neuromorphic computing operates to a fundamentally different principle in which data is stored and can be manipulated at the same time. This architecture style lends itself to processing massively parallel data streams for pattern recognition, control and event response. Examples of potential applications are image recognition and analysis, industrial machine vision and robotic controls, autonomous vehicles, natural language processing and anomaly detection.

For appropriate workloads, the efficiency gains promise to be two to three orders of magnitude, while performing at a tiny fraction of the energy of a Von Neumann machine.

Various chip technologies have been proposed to implement neurons in silicon, including transistors, memristors, spintronic memories and threshold switches.

Digital neuromorphic systems bypass many of the complexities of implementing physical neurons but are much larger and do not enjoy the full efficiency benefits of mixed-signal neuromorphic hardware. Notable examples of this approach are the SpiNNaker system at the University of Manchester, powered by one million CPU cores, and its successor, SpiNNaker 2 at the Technical University of Dresden, with five million cores.

Implementations

Notable neuromorphic computing projects, listed in chronological order, are:

- The EU's Human Brain Project, launched in 2013, was a decade-long, €600 million (\$697 million) effort that resulted in the creation of two neuromorphic systems: SpiNNaker (digital) and BrainScaleS (mixed-signal).
- IBM's TrueNorth, launched in 2014, simulated a million neurons and 256 million synapses — roughly equivalent to the brain of a bee. The hardware was used in experiments and applications by more than 40 organizations, including Lawrence Livermore National Laboratory (California, US) and Air Force Research Laboratory (New York State, US).
- The China Brain Project launched in 2016 with a 15-year funding plan and brain-inspired AI technologies among core focus areas.
- Intel's Loihi 2, released in 2021, implemented a million neurons and was accompanied by a software development framework called Lava. In 2024, Intel combined 1,152 chips in a 6U appliance for Sandia National Laboratories (New Mexico, US). It supported 1.15 billion neurons and 128 billion synapses (roughly equivalent to the brain of an owl) while consuming 2.6 kW.
- In 2022, Australian company BrainChip announced that it was taking orders for Akida AI processor PCIe boards, making it the world's first commercially available, off-the-shelf neuromorphic processor.
- In 2023, IBM released NorthPole — a chip that was inspired by the brain and negated the Von Neumann bottleneck, yet was designed to run conventional deep learning workloads, claiming 25 times improvement in energy efficiency of inference.
- In 2025, SpiNNcloud — the company that emerged from the development of SpiNNaker — announced a chip called SpiNNNext. One of the first deployments is at Leipzig University (Germany), where it will be used to simulate 10.5 billion neurons. SpiNNcloud claimed that in some applications its architecture is 78 times more efficient in terms of tokens per watt than contemporary GPUs.
- In 2025, the China Brain Project launched Darwin Monkey, a neuromorphic system powered by the homegrown Darwin 3 chips. It comprised a billion neurons and 100 billion synapses (roughly equivalent to the brain of a macaque) while consuming 2 kW. Demonstrations have included running DeepSeek models.

There are separate efforts to adopt the technology for low-power, on-device applications, including:

- BrainScaleS-2, developed at the Heidelberg University in Germany, features 512 neurons, around 130,000 synapses, and consumes 1 W of power. Chips have been available to researchers using remote access since 2022.
- In 2025, Dutch firm Innatera launched Pulsar, calling it the world's first commercially available neuromorphic microcontroller. The chip can process sensor data and run basic pattern recognition models at a power budget measured in microwatts.

Economics

In 2025, much of the focus of neuromorphic research has shifted from discovering new compute paradigms to developing hardware platforms that can radically improve performance and decrease power consumption of existing deep learning-based AI applications, especially inference workloads.

Hybrid servers equipped with both traditional and neuromorphic processors could make today's AI models faster, cheaper and easier to deploy, before eventually replacing them with native" neuromorphic models. The likely initial use cases include speech and image recognition, natural language processing and brain-machine interfaces.

Separately, neuromorphic chips are being considered for edge processing and robotics applications, bringing advanced AI functionality directly to devices and reducing the need for cloud services.

All neuromorphic chips are currently used for prototyping and demonstration, with commercial applications years, or possibly decades, away. Yet, if some of the early claims are verified, neuromorphic computing could contest a sizeable (>20%) chunk of the compute segment of the semiconductor market, estimated at \$800 billion for 2025.

Commercial activity

Much of the funding for neuromorphic research comes from the public sector and is focused on universities. Some of the programs started in academia have resulted in ambitious commercialization efforts — the most notable being the Technical University of Dresden's SpiNNcloud and Zhejiang University's Zhejiang Lab.

These academic institutions are competing against the research organizations of giants Intel and IBM, but neither appears excited about the prospects of neuromorphic computing. IBM tests show that NorthPole, a prototype chip made using vintage 12 nanometer fabrication processes, delivered much lower latency and higher efficiency in small language model inference than Nvidia's H100. IBM has demonstrated a 2U server that contains four of these, and yet there are no plans to commercialize the system.

The dangers of being too early to the market were illustrated in 2025 by the failure of San Francisco-based Rain AI, a neuromorphic hardware startup backed by Sam Altman. Altman invested \$25 million in a seed round in 2022, and OpenAI committed to buying \$51 million worth of hardware. However, the hardware never materialized, as the startup failed to secure the \$150 million it required to continue development. It is now exploring a sale.

Drivers and barriers

Renewed interest in neuromorphic computing comes at a time when the power consumption of existing approaches to AI is becoming increasingly problematic. A more efficient alternative would be welcomed by broad swaths of the industry, from small-scale data center developers to grid operators, policymakers and the public.

If demonstrated to be superior, the technology poses a risk to investment commitments in conventional AI infrastructure. However, the commercial success of neuromorphic computing appears remote and uncertain.

To take full advantage of a new architecture, researchers will need to reinvent or redevelop existing AI capabilities, for example, teaching spiking neural network (SNN)-based models how to recognize handwriting or speech while not relying on today's artificial neural networks. New programming languages and software will be needed to operate neuromorphic hardware, and a sufficient pool of developers would need to be trained in their use.

The possibility of building systems with more neurons and synapses also depends on advances in nanotechnologies and materials science. Traditional computing architectures keep evolving rapidly to increase compute and data parallelism for AI training and inference workloads while suppressing power use, making neuromorphic advantage" a fast-moving target.

Presently, the neuromorphic hardware and software market is diverse, with no clear leaders that could produce a framework that the rest of the growing industry segment could follow.

At the same time, examples show that neuromorphic systems are scaling very quickly: from millions of neurons a decade ago to billions of neurons today. Connecting these is relatively straightforward, and soon, researchers will be able to build machines that have as many neurons as the human cerebral cortex. It will be the software component that decides whether neuromorphic computing becomes a competitive alternative to the machine learning platforms of today.

The Uptime Intelligence View

Much like the commercial challenge facing quantum computing systems, neuromorphic systems are available but limited to small user bases in universities or industrial research groups. Similarly, there is currently no compelling demonstration of a high-volume application where neuromorphic hardware outperforms the alternative.

To take the next step in development, researchers must demonstrate a neuromorphic advantage in performance and power over conventional GPUs. Such applications appear to be close, and indeed, several vendors already claim performance and efficiency leadership in a narrow and carefully constructed set of tasks.

Funding neuromorphic research carries little risk and could have transformative effects on many sectors. In the most optimistic commercialization scenario, this technology will change how most organizations approach IT — in a transition that will take many years.

ABOUT THE AUTHOR



Max Smolaks

Max is a Research Analyst at Uptime Institute. His expertise spans digital infrastructure management software, power and cooling equipment, and regulations and standards. He has 10 years' experience as a technology journalist, reporting on innovation in IT and data center infrastructure.

msmolaks@uptimeinstitute.com

About Uptime Institute

Uptime Institute is the Global Digital Infrastructure Authority. Its Tier Standard is the IT industry's most trusted and adopted global standard for the proper design, construction, and operation of data centers the backbone of the digital economy. For over 25 years, the company has served as the standard for data center reliability, sustainability, and efficiency, providing customers assurance that their digital infrastructure can perform at a level that is consistent with their business needs across a wide array of operating conditions.

With its data center Tier Standard & Certifications, Management & Operations reviews, broad range of related risk and performance assessments, and accredited educational curriculum completed by over 10,000 data center professionals, Uptime Institute has helped thousands of companies, in over 100 countries to optimize critical IT assets while managing costs, resources, and efficiency.