

OCTOBER 2026

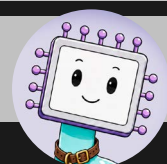
**Register for  
HiPEAC 2027  
Glasgow**



**Adventures in quantum, from qubits to software**

**How Roofline is bringing mixture-of-experts to the edge**

**Saving Swiss chips from extinction, with Chipsaur**

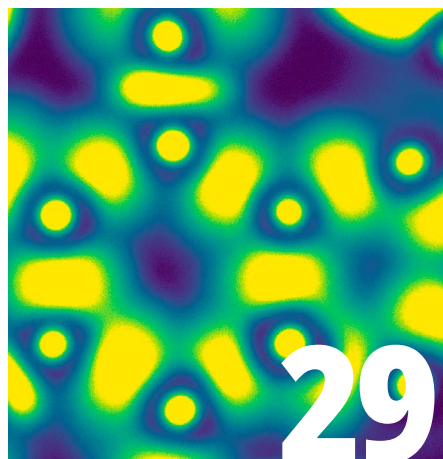




**Carmen G. Almudéver on scaling up a quantum computer**



**Enabling quantum, from hardware to software to access**



**Identifying quantum-appropriate workloads**

|    |                                                                                                                                                                                                                                                                                        |    |                                                                                                                                                                                                                                                                                                    |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3  | <b>Welcome</b><br><i>Koen De Bosschere</i>                                                                                                                                                                                                                                             | 34 | <b>Special feature: Quantum</b><br><b>'Just as in the classical stack, we need to co-design the system'</b><br><i>Swamit Tannu and Wisse Hettinga</i>                                                                                                                                              |
| 4  | <b>HiPEAC news</b><br><b>DISCOVER-US workshop, PUMPS+AI, ACACES 2026</b>                                                                                                                                                                                                               | 36 | <b>SME snapshot</b><br><b>Qside: Building verifiable trust infrastructure for modern cryptography</b><br><i>Marta Alcaide</i>                                                                                                                                                                      |
| 10 | <b>Ecosystem evolution</b><br><b>SiPearl, Quobly, CEA, TechWolf, Openchip, EUCLYD, Infininode, Nearby Computing, Multiverse, Racyics, Cloudberry, SambaNova, Chipmind, Ximplic, Zeropoint, NOFire AI, TenX Semi, ZuriQ, QuTech, Applied Quantum</b>                                    | 37 | <b>Technology opinion</b><br><b>Unblocking global quantum communication networks from geopolitical limitations</b><br><i>Alexandra Pinto and Myrna James</i>                                                                                                                                       |
| 13 | <b>Community news</b>                                                                                                                                                                                                                                                                  | 38 | <b>Peac performance</b><br><b>Bringing mixture-of-experts to the edge: Native MoE compiler support in IREE</b><br><i>Jan Moritz Joseph</i>                                                                                                                                                         |
| 20 | <b>Special feature: Quantum</b><br><b>'We would love to demonstrate that our modular approach can be used to scale up quantum computers'</b><br><i>Carmen G. Almudéver</i>                                                                                                             | 40 | <b>Computing in practice</b><br><b>vision 4rescue: The technologies behind a new era of emergency response</b><br><i>Thomas Regis</i>                                                                                                                                                              |
| 23 | <b>Special feature: Quantum</b><br><b>Laying the foundations for quantum, from hardware to software to access</b><br><i>Rui Pereira, Katrin Tina Möbius, Nikolay Tcholtchev, Sebastian Bock, Jan Swiatkowski, Jakub Konvička, Ladislav Foltyn, Jan Martinovič and Branislav Janský</i> | 42 | <b>Innovation Europe</b><br><b>Greenshift</b><br><i>The Greenshift consortium</i>                                                                                                                                                                                                                  |
| 29 | <b>Special feature: Quantum</b><br><b>Preparing today for quantum tomorrow: Why knowing your workloads will be key to exploiting quantum acceleration</b><br><i>Max Erné</i>                                                                                                           | 43 | <b>HiPEAC futures</b><br><b>Chipsaur asks: How do you build a chip-design community in Switzerland?</b><br><b>From HiPEAC Student Challenge to dream job</b><br><b>Three-minute thesis: Design automation of approximate machine learning circuits for ultra resource-constrained environments</b> |
| 31 | <b>Special feature: Quantum</b><br><b>You can't patch a safety case: Future-proofing safety-critical systems against Q-Day</b><br><i>Adam Mackay</i>                                                                                                                                   |    |                                                                                                                                                                                                                                                                                                    |



**Look beyond the qubit: Wisse Hettinga interviews Swamit Tannu**



**Roofline's native MoE compiler support in IREE**



**How Chipsaur is resurrecting Swiss chips**

Spanning the compute continuum from edge to cloud, HiPEAC (High Performance, Edge And Cloud computing) is a network of over 2,500 world-class computing systems researchers, industry representatives and students. First established in 2004, the project is now in its eighth edition. HiPEAC8 focuses on networking, roadmapping and innovation activities: bringing the computing community together in Europe, exchanging ideas, inspiring entrepreneurs and exploring the long-term vision for computing systems.

 [hipeac.net](https://hipeac.net)

 [hipeac.net/linkedin](https://hipeac.net/linkedin)

 [@hipeac](https://twitter.com/hipeac) / [@hipeacjobs](https://twitter.com/hipeacjobs)

 [hipeac.net/tv](https://hipeac.net/tv)



**Funded by  
the European Union**

The HiPEAC project has received funding from the European Union's Horizon Europe research and innovation funding programme under grant agreement number 101069836. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

**Cover image:** Bartek Wróblewski | [stock.adobe.com](https://stock.adobe.com)

**Design:** [www.magelaan.be](https://www.magelaan.be)

**Editor:** Madeleine Gray



The theme of this magazine is quantum computing – a technology with significant disruptive potential, but whose path forward has not always been clear. In her interview for this issue, Carmen G. Almudéver reminds us of Nobel winner Michel Devoret's quote that 'technology is less predictable than the weather'. The first association that popped up in my brain was 'VUCA', which stands for 'volatility, uncertainty, complexity, and ambiguity'. It was first used in US leadership literature in 1985; I learned about it in the aftermath of the 2008 financial crisis, which was called a 'VUCA shock'.

However, compared to the situation today, the world in 2010 was pretty stable and predictable: the UK was still part of the EU, there was no war with Russia, the US and the EU were still strong allies, energy prices were predictable, globalization was driving growth, climate change was considered a future problem, humans were still smarter than artificial intelligence (AI). During the last decade a lot has changed: Trumpism and geopolitical changes, COVID-19, a hot war at the border of Europe, acceleration of climate change, the rise of radical political parties, and the relentless race for AI supremacy. Compared to 2010, this is at least a (VUCA)<sup>2</sup> world.

Even the HiPEAC Vision has been impacted by VUCA. Until 2023, HiPEAC produced a vision every two years. Since 2024, we've produced a yearly vision. In the past we used a 10-year horizon for the vision. Today, it is difficult even to forecast the next five years with high confidence due to acceleration in different domains (technological, climate change, geopolitical, military, ...). The impact of the acceleration is so fundamental that the solutions of the past to manage the resulting changes are no longer effective today. Global warming is accelerating and reducing emissions no longer suffices to avoid the immediate effects of climate change. The United Nations has become powerless in many conflicts. Even education is impacted by the acceleration. In 2040, a PhD might be considered a relic of a bygone era because all research is done by superintelligent AIs...

The only confident prediction I dare to make is that this acceleration will continue, driven by progress in AI – which does not (yet) show any sign of slowing down. This acceleration will impact every human on the planet – some for the better, some unfortunately for the worse. We will have to adapt in the next decade, and faster than we are used to. Adaptation will become the new normal.

Koen De Bosschere, HiPEAC coordinator

## DISCOVER-US UAV workshop explores AI and decentralized intelligence for drone sector



On 3-5 June, the European-funded DISCOVER-US project hosted a workshop featuring high-level participants from the European Commission along with representatives from industry and academia working on next-generation drone technologies. The workshop explored recent innovations in AI-powered drones, with a focus on key sectors such as agriculture, energy, logistics, water and underwater logistics, transportation and maintenance.

Participants at the workshop, which took place in Brussels and included a visit to the DronePort test centre, considered how to balance autonomy, coordination, and robustness in real-world, resource-constrained environments. Among the main topics that were discussed were:

- the design of lightweight, energy-efficient AI
- strategies for resilient communication and information sharing
- coordination in swarm behaviour
- learning paradigms in decentralized environments

The impact of recent technology shifts such as software-defined everything (SDX) and AI were considered. Key challenges for the sector were identified, such as balancing flight-critical functions and data processing functions, establishing a modular architecture for unmanned aerial vehicles (UAVs) and developing abstraction levels.

Real-world scenarios, such as search-and-rescue operations, were presented during the workshop, illustrating how local intelligence, decentralized cooperation, and distributed planning come together in practice.

 [discover-us.eu](https://discover-us.eu)



## PUMPS+AI 2026 offers up-to-the-minute training on programming for HPC and AI

Taking place every year in Barcelona, the PUMPS+AI ACM Europe Summer School allows participants to enrich their skills in developing applications for many-core processors with massively parallel computing resources such as graphics processing unit (GPU) accelerators, with a particular focus on artificial intelligence (AI).

Organized by Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS), the Universitat Politècnica de Catalunya-Barcelona Tech, the University of Illinois at Urbana-Champaign and HiPEAC, the 16th edition of the summer school took place in Barcelona on 13-17 July. Long-standing lecturer Wen-mei Hwu (University of Illinois at Urbana-Champaign / NVIDIA) was joined by Juan Gómez Luna and Carlos Villavieja (both NVIDIA) and local lecturers Antonio J. Peña, Marc Jordà, Leonidas Kosmidis, Xavier Martorell and Xavier Teruel. The summer school is co-directed by BSC Director Mateo Valero, HiPEAC co-founder and first coordinator.

'This year's edition had a significant proportion of completely new content, including more applied GPU techniques for AI,' said Antonio J. Peña, one of the PUMPS+AI organizers and a HiPEAC member. 'We welcomed a new professor, Carlos Villavieja from NVIDIA, although we missed Professor Izzat el Hajj, who was unable to join us due to the armed conflict in Lebanon. Stephen Jones from NVIDIA also gave a great keynote on the future of CUDA features.'

Three poster awards were given during the summer school; the winners received a copy of the latest edition of Wen-mei Hwu's book *Programming Massively Parallel Processors*.

 [pumps.bsc.es](https://pumps.bsc.es)

## Preparing the next generation of computing researchers: ACACES 2026

Returning to Fiuggi, Italy, in July, ACACES 2026 once again offered participants a rich, engaging programme of courses by experts from around the world. Unsurprisingly, artificial intelligence (AI) was a major topic: Ruchir Puri (IBM) kicked off the event with a high-octane keynote setting out the latest developments in the field and pointing towards his vision for future artificial useful intelligence. Uday Bondhugula (Indian Institute of Science / Polymage Labs) guided students through compiler design for AI chips and programming frameworks. Nathaniel Pinckney (NVIDIA) and Stijn De Pauw (TechWolf) led courses on large language models (LLMs) for hardware and software, respectively. Meanwhile, Morley Mao (University of Michigan) considered the interplay between AI, cybersecurity, and cyber-physical systems, particularly in the automotive domain, and Vlad Constantin Coroamă (Roegen Centre for Sustainability) explored the relationship between AI and sustainability.

Novel forms of acceleration were also a key feature. Angelo Garafolo (University of Bologna) focused on RISC-V architectures for the edge, Lesley Shannon (Simon Fraser University) investigated the principles and engineering frameworks of composable heterogeneous computing, Sagar Karandikar (UC Berkeley) considered domain-specific hardware in scale-out systems, and Swamit Tannu (University of Wisconsin, Madison) introduced students to quantum computer architecture (see p.34 for an interview with Swamit).

Rapid changes in the field notwithstanding, the summer school was testament to the enduring importance of the fundamentals of computer architecture, as evidenced by the popularity of Boris Grot's (University of Edinburgh) course on central processing units (CPUs), which emphasized efficiency, and Daniel Sorin's (Duke University) course on memory consistency and cache coherence.

In addition to the technical curriculum, ACACES aims to encourage and support would-be entrepreneurs. Drawing on his significant experience, Omar Mohout's (Deloitte) course on entrepreneurship offered practical advice and guidance, while Rolf Segger (SEGGER Microcontroller) gave an invited talk outlining how his unique journey as a founder has enabled him to live life on his own terms.

As usual, students had ample opportunities to develop, whether by presenting posters to peers and senior attendees, discussing European digital autonomy in a paper discussion session focusing on the HiPEAC Vision, or simply discussing research over relaxed meals.

Applications for ACACES 2027 will open in the spring – check the HiPEAC website and LinkedIn page for further information.

🔗 [hipeac.net/acaces](https://hipeac.net/acaces)



# SiPearl delivers first Rhea1 CPUs to Bull for integration into Jülich's JUPITER supercomputer



In September, SiPearl announced that the first Rhea1 central processing units (CPUs) had been delivered to Bull, the French high-performance computing (HPC) company. Rhea1 is the first European-designed CPU to be deployed in a European-built platform.

Bull has begun integration of Rhea1 onto its BullSequana XH3000 platform for JUPITER, Europe's first exascale supercomputer. Procured by the EuroHPC Joint Undertaking, JUPITER is hosted at Jülich Supercomputing Centre (JSC). According to SiPearl's announcement, this milestone will contribute to the development of JUPITER's CPU blade, which complements the GPU partition, paving the way for full modular integration of the machine.

The most complex server CPU ever designed in Europe, Rhea1 includes 80 Arm® Neoverse V1 cores (each with 2 x 256-bit Scalable Vector Extension), four stacks of high-bandwidth memory (HBM), and four DDR5 memory interfaces.

'The delivery of the first Rhea1 CPUs marks an important milestone for Europe's supercomputing sovereignty. Through the European Processor Initiative, EuroHPC JU has supported the development of a European processor that is now moving from innovation to deployment in JUPITER, Europe's first exascale supercomputer. This achievement demonstrates what Europe can accomplish by bringing together research, industry and world-class infrastructure to build a stronger European HPC ecosystem,' said Anders Dam Jensen, executive director of EuroHPC JU.

SiPearl was founded as part of the European Processor Initiative, an EU-funded project to develop designed-in-Europe chips for HPC. SiPearl Chief Executive Philippe Notton will be giving a keynote at HiPEAC 2027 in Glasgow in January.

[sipearl.com](https://sipearl.com)

# CEA-Leti and Quobly strengthen collaboration through access to the FAMES Pilot Line

Quobly, a French company developing silicon-based quantum computers, and the research institute CEA-Leti, which specializes in microelectronics and information technologies, recently announced that they are strengthening their collaboration through access to the FAMES Pilot Line, which complements Quobly's industrial collaboration with STMicroelectronics and supports its technology roadmap.

Thanks to FAMES, an initiative which supports the emergence of higher-performance, more energy-efficient and more sustainable chip designs, Quobly can access advanced semiconductor equipment and process modules to develop and evaluate new process flows and patterning approaches for later processor generations.

Founded in 2022, Quobly's technology is based on 15 years of research at CEA-Leti and the French national research centre CNRS. In 2016, CEA-Leti and its research partners demonstrated a silicon spin qubit fabricated using an industrial complementary metal-oxide semiconductor (CMOS) process based on fully depleted silicon-on-insulator

(FD-SOI) technology on a 300 mm wafer. Since Quobly's founding, the company and CEA-Leti have continued to collaborate on silicon spin-qubit technologies and the technology building blocks underlying Quobly's QSOI® architecture. QSOI® combines silicon spin qubits with an FD-SOI semiconductor platform and is being industrialized with STMicroelectronics using their 300 mm manufacturing infrastructure.

Quobly plans to commercialize its first quantum computer, Alloy Pioneer, by the end of 2026, with a roadmap extending toward large-scale, fault-tolerant architectures targeting one million qubits by 2032. In September, the company announced new results showing qubit readout, single-qubit gates and two-qubit gates on a single quantum chip fabricated on its proprietary QSOI® technology. According to the company, this demonstrates that key quantum operations can be performed on devices manufactured through an industrial semiconductor production process.

[bit.ly/Quobly\\_FAMES\\_2026](https://bit.ly/Quobly_FAMES_2026)

## SAP to acquire Belgian company TechWolf

Belgian startup TechWolf and German enterprise software company SAP have entered into a definitive agreement for SAP to acquire TechWolf. The Belgian startup is the provider of an artificial intelligence (AI) work intelligence platform, based on a proprietary data model called context graph for work.

According to the announcement, the acquisition will bring this context graph, along with TechWolf's AI models and applied AI research team, into SAP, establishing work intelligence as a strategic SAP asset.

The deal is expected to close in the last quarter of 2026, subject to customary closing conditions. Terms of the deal were not disclosed.

The announcement also stated that TechWolf is expected to remain as an independent entity at its Ghent headquarters, with offices in London, New York and San Francisco. TechWolf's Stijn De Pauw is a HiPEAC member and was a tutor at HiPEAC's ACACES 2026 summer school.

🔗 [techwolf.ai/press/sap-to-acquire-techwolf](https://techwolf.ai/press/sap-to-acquire-techwolf)

## Openchip launches BER10 chip

In July, Openchip, a European semiconductor company headquartered in Barcelona, announced its chip BER10. Built on RISC-V, BER10 is a Linux-capable 64-bit processor with four out-of-order central-processing unit (CPU) cores and a vector processing unit, fabricated with leading-edge sub-2nm gate-all-around (GAA) process technology. According to the Openchip announcement, BER10 paves the way for Openchip's next generation of RISC-V compute accelerators, designed for high-performance computing (HPC) and datacentre-scale AI workloads.

'When we started Openchip, the goal was never just to build a chip, but to prove that Europe can design and deliver world-class silicon on open standards, on its own terms,' said Cesc Guim, the chief executive of Openchip. 'BER10 is that proof, made real. It's a foundation for the digital sovereignty Europe needs, built on a more open and resilient value chain that ultimately benefits the global technology industry.'

'BER10 confirms that 2026 is the year of highly powerful RISC-V application processor silicon and that RISC-V has already moved beyond experimentation into high-performing industrial-grade, real-world computing,' said Andrea Gallo, CEO of RISC-V International. 'Seeing a



European team deliver Linux-capable silicon of this calibre reflects the momentum we're witnessing across the ecosystem worldwide, and it's exactly the kind of innovation an open standard is meant to enable.'

Previously, in June, it was announced that the Spanish Society for Technological Transformation (SETT), an entity attached to the Ministry for Digital Transformation and Public Administration, had approved a €115.77 million public investment in Openchip.

🔗 [openchip.com](https://openchip.com)

## EUCLYD raises over €200 million to accelerate AI inference chip development

Dutch startup EUCLYD, which is headquartered in Eindhoven, has secured Series A financing of more than €200 million, in a round co-led by Somerset Capital Partners, Scaleup Europe Fund, which is managed by EQT, and Innovation Industries. Peter Wennink, former president and chief executive of ASML, has also joined EUCLYD as chairman of the board.

EUCLYD's craftwerk chip is based on a dataflow architecture. It aims to optimize token production for artificial intelligence (AI) agents. EUCLYD claims that craftwerk offers radical improvements in energy

use and cost efficiency compared to competitors such as NVIDIA graphics processing units (GPUs).

The financing will expand EUCLYD's engineering organization, accelerate its silicon and systems roadmap, strengthen ecosystem partnerships, and prepare the company for commercial deployment across enterprise, sovereign, and hyperscale AI markets.

🔗 [www.euclid.ai](https://www.euclid.ai)

# HiPEAC Tech Transfer Award winner InfiniNode secures €2 million in financing

In July, InfiniNode, a semiconductor intellectual property (IP) company spun off from Chalmers University of Technology which develops advanced interconnect and memory-subsystem technologies for RISC-V, multicore, and chiplet-based designs, announced that the company had secured €2 million in funding and grants.

Co-founded by HiPEAC founding member Per Stenström and HiPEAC member Ioannis Sourdis, InfiniNode's semiconductor IP enables efficient communication across modern compute architectures, helping improve system scalability, performance and energy efficiency while reducing system-on-chip (SoC) design complexity. The technology targets data-intensive computing applications, including artificial intelligence (AI) and high-performance computing (HPC).

InfiniNode was born out of an observation by researchers at Chalmers that existing interconnect and caching solutions targeting RISC-V did not scale effectively for manycore processors. To address this issue, the team embarked on developing a set of IP blocks as part of several national and EU-funded projects, including the European Processor Initiative. In 2024, the InfiniNode team won a HiPEAC Technology Transfer Award for successfully commercializing this technology.

This funding round was financed by Nordic deep-tech investors including Turbine Capital, PSV Hafnium, Norrskan Evolve, Chalmers



Ventures and Almi Invest, together with public innovation funding. The investment will enable InfiniNode to advance its semiconductor IP portfolio, deepen customer engagements and scale its commercial operations.

'The performance and energy efficiency of future compute systems will increasingly be determined by how efficiently data moves between processors, accelerators and memory. Our mission is to enable efficient data movement in future compute systems,' said Patrik Millsjö, chief executive officer of InfiniNode. 'This financing allows us to accelerate our technology roadmap, strengthen customer collaborations and expand our semiconductor IP portfolio.'

[infininode.com](https://infininode.com)

# Nearby Computing secures €1.4 million funding round and launches AI Mesh

Nearby Computing, a spinoff of the Universitat Politècnica de Catalunya-Barcelona Tech (UPC) and Barcelona Supercomputing Center (BSC), has closed a €1.4 million funding round, including €680,000 from the Spanish Society for Technological Transformation (SETT).



Announced on 17 September, the funding is structured through the Next Tech fund, financed with Next Generation EU funds, and will be used to support the company's growth and the development of its NearbyOne platform.

In October, the company launched AI Mesh, a NearbyOne module for deploying, connecting and governing AI workloads across the cloud to edge continuum.

Nearby Computing was founded in 2018 to commercialize internet of things (IoT) and 5G technology developed at UPC and BSC. The company provides a cloud-to-edge orchestration platform that runs infrastructure, networks, applications and artificial intelligence (AI) from a single control plane.

[nearbycomputing.com](https://nearbycomputing.com)

## Racyics and Cloudberry team up to offer design enablement and finance



Racyics, a German company providing semiconductor-design services, and Cloudberry VC, a European venture-capital fund dedicated to semiconductors, have entered into a collaboration to connect their two platforms.

Racyics is a Design Enablement Team (DET) within the European Chips Design Platform (EuroCDP), part of the European Chips Act. Through its makeChip platform, Racyics provides startups with proven silicon design flows, electronic design automation (EDA) tooling, process design kits (PDKs), intellectual property (IP), and full supply-chain access. As a DET, Racyics helps startups prepare technically sound EuroCDP applications and delivers hands-on design services to bring a concept to silicon.

Cloudberry VC is Europe's first venture capital fund dedicated to semiconductors, photonics, and advanced materials. Alongside capital,

Cloudberry is assembling a platform of industrial and technical partners spanning EDA and simulation, foundry access, design enablement and EuroCDP access.

EuroCDP offers startups support, tools, training and financing. Grants for EuroCDP require a private co-funding component which Cloudberry, as a pre-seed and seed investor, can supply.

According to the Racyics announcement, fabless teams working with Racyics on a EuroCDP application can be introduced to Cloudberry VC for the private co-funding component, while teams in Cloudberry's pipeline can be introduced to Racyics for design enablement. In addition, the two organizations also plan to collaborate on events and community initiatives across the European semiconductor ecosystem.

[bit.ly/Racyics\\_Cloudberry\\_2026](https://bit.ly/Racyics_Cloudberry_2026)

## Multiverse targets €500 million Series C funding

In July, Multiverse Computing, a company specializing in artificial intelligence (AI) model compression using a methodology inspired by quantum physics, announced a €500 million Series C funding round led by Forgepoint Capital International, BNPP SIVF, and Bullhound Capital at a €1.5 billion pre-money valuation. The company, which is based in Donostia, Spain, claims that its CompactifAI technology shrinks a high-fidelity model by 95%, from 258 MB to 8 MB, while preserving accuracy, further opening up the possibility of integrating these models at the edge.

'The AI industry has accepted a false constraint for years — that powerful models require expensive infrastructure,' said Enrique Lizaso, co-founder and chief executive of Multiverse Computing. 'That constraint is gone. We have proven that AI can run at full performance on a smartphone, inside a sovereign data centre, on a factory floor with no cloud connection. This round is the moment we scale that proof across every industry that needs it.'

[bit.ly/Multiverse\\_Series\\_C](https://bit.ly/Multiverse_Series_C)

## SambaNova closes \$1 billion funding round

In July, SambaNova announced that it had completed the first close of \$1 billion in strategic financing as part of a Series F round led by General Atlantic, valuing the company at \$11 billion post-money. Co-founded by HiPEAC associate member Kunle Olukotun, who is also the company's chief technologist, SambaNova provides a full-stack platform – including chips, systems, and cloud services – for AI inference.

SN50, SambaNova's fifth-generation reconfigurable dataflow unit (RDU) was unveiled in February 2026 and is described as 'purpose-

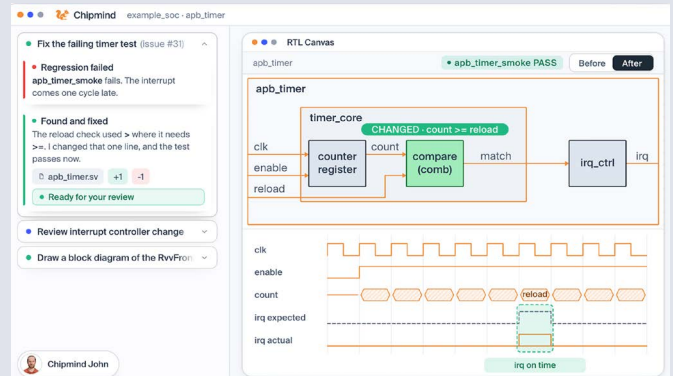
built hardware for the memory-bound phase of inference'. Speaking to EE Times' Sally Ward-Foxton, SambaNova's chief executive Rodrigo Liang said: 'Big models need memory capacity. Even the open-source models are multi-trillion parameters now, and as far as I can see, most people are not willing to give up quality for quantized versions of the model... that's what we're solving for.'

[bit.ly/SambaNova\\_EETimes\\_2026](https://bit.ly/SambaNova_EETimes_2026)

## Chipmind launches RTL Canvas and secures CHF 1.37 million in funding

Chipmind has released RTL Canvas, a generative user interface which turns register transfer level (RTL) code into an interactive graphic interface and allows engineers to prototype, plan, and simulate chip-design variations without making immediate changes to the underlying codebase. This allows engineers to explore architectural alternatives before committing these changes in code. According to Chipmind, the canvas generates the optimal visual format on-the-fly, tailored to the engineering task at hand; for example, displaying relevant parts of the architecture, showing finite-state-machine or waveform diagrams.

Chipmind was founded by two doctorates from ETH Zurich, Harald Kröll and Sandro Belfanti; the company's scientific advisor is HiPEAC member Luca Benini. Based in Zürich, the company is building artificial intelligence (AI) agents for chip-design teams. In October, Chipmind announced that it had been awarded CHF 1.37 million in funding



through the Start-up Innovation Projects programme of Innosuisse, the Swiss Innovation Agency.

[chipmind.ai](https://chipmind.ai)

## Launch of Ximplic Evaluations



Based in Groningen, Ximplic is a startup licensing artificial intelligence (AI) inference intellectual property (IP) based on compute-in-memory (CiM) architectures. Ximplic was founded to reduce data movement in semiconductor architectures and thereby improve efficiency, reduce latency and improve overall system performance.

The company has now launched Ximplic Evaluations, which allows customers to evaluate the Ximplic's CiM AI inference technology before committing to hardware. According to Ximplic, the evaluation flow

allows designers to run AI models on a virtual representation of the Ximplic architecture, estimate power and efficiency early in the design process, and explore how CiM can reduce data movement.

Ximplic was co-founded by HiPEAC member Farhad Merchant with Simranjeet Singh. The company won a HiPEAC Technology Transfer Award in 2025 for commercializing technology developed by the two researchers.

[ximplic.com](https://ximplic.com)

## ZeroPoint Technologies launches ZeroStream



ZeroPoint Technologies AB, which develops semiconductor technology that integrates memory compression and memory management, has released ZeroStream™, a hardware intellectual property (IP) block that increases effective memory bandwidth across a variety of use cases, including delivering more tokens per second from artificial intelligence (AI) accelerators.

ZeroStream uses hardware-accelerated compression to move more data across existing memory and streaming interfaces, improving performance for bandwidth-bound applications. According to the company's announcement, it offers lossless compression with no quantization or accuracy loss, up to 1.5x compression on large language model (LLM) weights, and 20-35% effective bandwidth improvement (up to 50 percent for some workloads).

'What sets an AI system apart today is how much performance and context you can pull from the memory you already have,' said Brett Cline, chief executive of ZeroPoint. 'ZeroStream lets chip designers do more with that same memory, increasing tokens per second, enabling more context held in memory, and fitting larger foundation models in a smaller footprint.'

ZeroPoint was founded in 2016 by HiPEAC founding member Per Stenström and HiPEAC member Angelos Arelakis, building on years of research at Chalmers University of Technology. The commercialization of the underlying technology was recognized by a HiPEAC Technology Transfer Award in 2016.

[zeropoint-tech.com](https://zeropoint-tech.com)

## NOFire releases Brig, an open-source microVM sandbox for AI coding agents



**NOFire.ai** NOFire AI has released Brig, an open-source tool that runs artificial intelligence (AI) coding agents inside a microVM, a lightweight virtual machine (VM), on the developer's own computer. Brig supports Apple silicon Macs and Linux, and is available under the Apache 2.0 licence.

AI coding agents install packages, run commands, use the network and act with the developer's credentials. To work faster, many engineers also turn on auto-approval, so the agent's actions run without review. A mistake, or a malicious instruction hidden in a repository or a web page, then runs with access to the developer's files, keys and accounts.

Brig keeps the agent inside a VM instead. The agent runs on a kernel of its own, isolated by the processor's hardware virtualization support (Hypervisor.framework on macOS, KVM on Linux). It sees only the project

directory it is given and the credentials its profile names. It does not see the rest of the disk, the keychain or the SSH agent. A network policy, enforced on the host outside the VM, can limit which hosts the agent reaches. When the work is done, the developer deletes the sandbox.

'I lead the development of the urunc CNCF (Cloud Native Compute Foundation) Sandbox project, a container runtime for unikernels and microVMs. The virtual machine technologies I've been working on since 2007 are now more relevant as ever, due to the rise of AI agents and agentic workloads, which require the security and trust boundary to move lower without paying a significant performance tax,' said Anastassios Nanos, co-founder and chief scientist of NOFire AI and a HiPEAC member.

[brig.sh](https://brig.sh)

## TenX Semi officially launches after seed round

TenX Semi Inc, co-founded by HiPEAC associate member Subhasish Mitra (Stanford University), has emerged from stealth following a seed round co-led by South Park Commons and Khosla Ventures. The company develops autonomous chip-design tools and artificial intelligence (AI) agents which leverage 'self-verify-and-fix' technology, which the company says can dramatically shorten design timelines while simultaneously improving design quality.



In a post on the TenX Semi website, founders Subhasish Mitra, Shwetabh Verma and SukHwan Lim said: 'If AI creates accelerator designs 100x faster but we can't verify them 100x faster, then we have simply amplified the verification bottleneck. ... For the first time in decades, some of the fundamental questions in chip design and verification are back on the table. For example: How does "intent" become silicon? What should humans actually specify? How do we establish correctness when machines can generate designs no human could fully inspect? What becomes possible if we finally break the verification bottleneck?'

Further information is available at [tenxsemi.com](https://tenxsemi.com)

View Subhasish Mitra's keynote talk from HiPEAC 2023 on HiPEAC TV  
[youtu.be/Pd8NNSv6Guo](https://youtu.be/Pd8NNSv6Guo)

## ZuriQ closes \$25.5 million seed round



ZuriQ, a Swiss quantum computing company spun out of ETH Zürich, announced in July that it had raised \$25.5 million in seed funding to scale its trapped-ion quantum processors. Led by Quantonation, the round builds on the company's \$4.2 million pre-seed round in 2025.

ZuriQ is developing a native two-dimensional trapped-ion quantum processor architecture, tested on a demonstrator comprising a 3x3 array of nine individually controlled ions. The chips behind it were

fabricated with ZuriQ's manufacturing partner, Infineon. According to ZuriQ, the native two-dimensional approach is inherently easier to scale.

According to the announcement, ZuriQ will use the new capital to expand its team alongside research and development efforts as it works to significantly increase the number of qubits its architecture can support.

[zuriq.com](https://zuriq.com)

# QuTech releases Tuna-17 quantum processor and launches Quantum Systems Integration Unit

Research institute QuTech has released Tuna-17, an open-architecture superconducting quantum processor, available through the Quantum Inspire cloud platform. Developed by the DiCarlo Lab at QuTech, Tuna-17 features 17 transmon qubits in a 'ninja-star' configuration, connected by 24 flux-tuneable couplers. It is the third system released in a year, following Tuna-5 and Tuna-9.

Recently, QuTech established a Quantum Systems Integration Unit to develop full-stack superconducting quantum computers. According to QuTech's announcement, the new unit will build open-architecture quantum computers using a European quantum value chain. It is seed-funded by QuTech Business Development, with Quantum Inspire providing the environment for deploying integrated systems and making them accessible to users.

Delft has attracted significant amounts of public funding, both from the Netherlands and the European Union (EU), including the Dutch QDNL programme and the EU's Quantum Flagship. To date, five generations of superconducting quantum systems have been made publicly accessible through Quantum Inspire. Bringing QuTech's knowledge of integrating full-stack quantum systems together, the team plans to create a repeatable system-integration blueprint, build an intellectual property (IP) base around system architecture, integration methods, tooling and interfaces, and work with potential customers and partners to understand where this capability is most valuable.



Image credit: QuTech

Rebecca Mayrhofer, head of business development at QuTech, said: 'Delft has built up years of experience in developing and operating complete quantum systems. We now want to turn that experience into a repeatable integration capability, build the IP around it and understand where customers see the greatest need. That creates a strong basis for future venture creation.'

Marco Matters, QDNL CAT-1 lead, said: 'Europe has excellent companies and research groups across many parts of the quantum computing stack. A strong system integrator can connect those technologies into complete practical and usable machines. Expanding that capability in Delft strengthens both the Dutch ecosystem and Europe's ability to develop its own quantum computing infrastructure.'

[quantuminspire.com](https://quantuminspire.com)

## Applied Quantum relocates to Europe

In August, Applied Quantum Chief Executive Marin Ivezic announced that the quantum-systems integration company was relocating from Singapore to Europe. In a blogpost on his personal website, Postquantum, he explained that, despite bureaucratic and funding hurdles, Europe was chosen for a variety of reasons, including its hosting of multiple quantum-hardware providers, market demand spurred by EU sovereignty policy and regulation – including an open-architecture model, a less geopolitically contentious jurisdiction, highly qualified engineers, and the growing maturity of the EU quantum ecosystem.

In the blogpost, Ivezic describes the move as 'a bet that for a quantum systems integration and advisory company, Europe's advantages now outweigh its well-documented disadvantages'.

To read the full blogpost, visit

[bit.ly/Postquantum\\_Applied\\_Quantum\\_EU](https://bit.ly/Postquantum_Applied_Quantum_EU)

## BZL announces second Cinco Ranch test chip

In June, Barcelona Zettascale Lab, an initiative of Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS), announced the arrival of the second Cinco Ranch test chip (TC2), allowing the BZL team to move onto the bring-up phase.

According to the announcement, TC2 incorporates a dual-core architecture with three levels of on-chip cache memory and a vector processing unit compatible with the RISC-V Vector 1.0 (RVV 1.0) extension, featuring eight vector lanes equipped with four floating-point units (FPUs) each. The design has been optimized for Intel 18A technology and reaches a target frequency of 1.5 GHz.

During this new phase, the BZL team will carry out the first experimental tests of the chip using the validation boards prepared for the TC2. These tests will make it possible to verify the correct functioning of the processor's various blocks, evaluate its performance, and gather key information for future generations of the Cinco Ranch family.

HiPEAC member Miquel Moretó, the head of hardware at BZL and director of the HPC Architectures Research Area at BSC-CNS, said: 'The arrival of TC2 turns months of design and validation into a real chip that we can now put to the test. It's a fundamental step toward our vision of next-generation European RISC-V processors.'

### FURTHER INFORMATION:

🔗 [bit.ly/BZL\\_Cinco\\_Ranch\\_TC2](https://bit.ly/BZL_Cinco_Ranch_TC2)

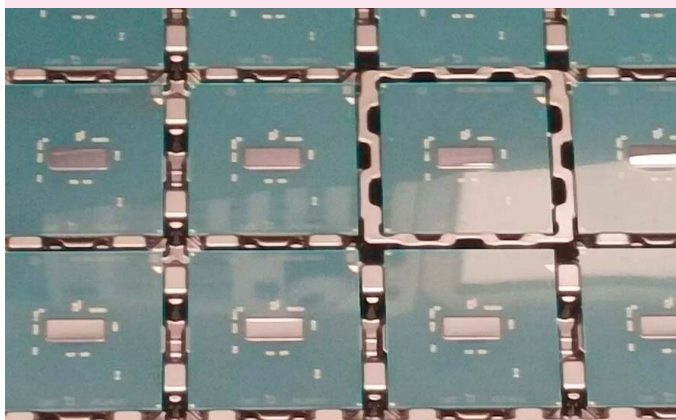


Image credit: BSC

## TRU-STACK to bring BSC's uHyper hypervisor to market



Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS) has been awarded €2.4 million from the European Innovation Council (EIC) Transition fund to coordinate TRU-STACK. The project will validate and bring to market uHyper, a novel user-space hypervisor that lets multiple parallel applications run simultaneously and efficiently on shared computing nodes. TRU-STACK was one of 40 projects – three of which are from Spain – selected from 611 submissions across Europe for funding in the 2025 EIC Transition call.

uHyper was developed to address the issue of idle capacity in parallel-computing environments such as high-performance computing (HPC) centres and data centres. Currently, according to the TRU-STACK announcement, workload-distribution inefficiencies cost hundreds of millions of euros, as well as wasting energy.

Developed at BSC as part of European research projects including EUPILLOT, uHyper is a lightweight software layer that sits between the operating system and parallel runtimes such as MPI, OpenMP, Python, and artificial intelligence (AI) frameworks. It dynamically coordinates concurrent workloads on the same node, preventing interference and filling idle resource gaps automatically, with no changes required to existing applications.

According to the announcement, early results show over 30% shorter execution times compared to best-practice baselines and at least 25% lower energy consumption per unit of work on real HPC and AI workloads.

HiPEAC member Vicenç Beltran, principal investigator of TRU-STACK and leader of the System Tools and Advanced Runtimes (STAR) Group at BSC, said: 'This funding validates years of research and lets us test uHyper at scale with real operators. Our goal now is to move it from the lab into production use.'

TRU-STACK is seeking enterprises, supercomputing centres and HPC users interested in exploring pilot deployments to make better use of on-premise infrastructure, cloud spending or awarded computing time. Potential partners can learn more about the pilot programme at [🔗 trustack.eu/pilots](https://trustack.eu/pilots)

## EPCC establishes pilot UK AI Factory Antenna

Jano van Hemert, EPCC, University of Edinburgh

Launched in February 2026, the UK AI Factory Antenna (Pilot-UKAIFA) at EPCC is a €10million initiative funded by the European Union and UK government created to support the growing demand for AI adoption. It supports industry and the public sector to build, test, and scale AI solutions using cutting-edge infrastructure and expertise.

Rooted in UK capability and capacity, Pilot-UKAIFA connects users to the broader European AI Factory network, opening up further opportunities for collaboration and access to complementary resources. It focuses on the development and implementation of artificial intelligence applications, particularly in the fields of manufacturing, engineering, and health, by start-ups, small/medium enterprises (SMEs), larger industrial companies and the public sector.

The AI Factory Antenna is part of the EuroHPC Joint Undertaking AI Factories initiative, which fosters innovation, collaboration and research in the field of AI by bringing together computing, data and talent. AI Factory Antennas allow countries outside the EU to partner with an AI Factory to bring the benefits to their countries. The UK's AI Factory Antenna at EPCC is collaborating with the HammerHAI AI Factory, based at the High-Performance Computing Center Stuttgart (HLRS), in partnership with a consortium of leading German academic research centres.

'EPCC and HLRS have led the use of national supercomputing services in Europe by industry over the past 30 years. The emergence of AI as a key application of supercomputing, and this funding from EuroHPC and the UK Government, is now enabling our organizations to embark on this new collaboration,' said Professor Mark Parsons, EPCC director, UKAIFA lead coordinator and HiPEAC member.

 [ukaifa.ai](https://ukaifa.ai)

## ICE and AUTH joint seminar: 'From virtual prototypes to intelligent chips'

Georgios Keramidas (AUTH, Thessaloniki) and Rainer Leupers (RWTH Aachen)

In May 2026, researchers from the Institute for Communication Technologies and Embedded Systems (ICE) at RWTH Aachen University visited the THESSIS research group (Kostas Siozios, Vasilis Pavlidis, Georgios Keramidas, and Thomas Noulis) at Aristotle University of Thessaloniki (AUTH) for a joint scientific seminar on the latest developments in designing the next generation of artificial intelligence (AI) chips. The event featured technical talks covering accelerator-centric AI chip design, design automation tools, advanced packaging (2.5D/3D), and secure and trustworthy chips. The seminar provided valuable opportunities for collaboration and knowledge exchange between the two HiPEAC research groups.

Opening presentations introduced the research focus areas of both THESSIS and ICE, with welcoming remarks by Georgios Keramidas (AUTH) and Rainer Leupers (RWTH Aachen). These were followed by a series of in-depth technical sessions. ICE researchers presented innovations in virtual prototyping, software security, and compute-in-memory (CIM), and emerging architectures. THESSIS showcased advances in deep neural network / convolutional neural network (DNN/CNN) compression, system-on-chip (SoC)-level simulation of accelerator-centric systems based on RISC-V, the design of reusable interposers, chip design tools, edge-to-cloud continuum methodologies, and tensor contraction optimizations. An additional presentation by Panagiotis Katsaros (AUTH) focused on correct-by-design embedded software, while Vasilios Kelefouras (University of Plymouth) presented compiler-level techniques for optimizing DNN kernels on edge central processing units (CPUs).

More than 15 PhD students participated in the event, benefiting from direct interaction and networking opportunities that will support their scientific development and future collaborations.



## TASKING-led team wins top award in AWS agentic AI hackathon for SDVs

A team representing TASKING, Infineon and the German Aerospace Center (DLR) won the top award at the Amazon Web Services (AWS) 'Accelerating the V-Cycle with Agentic AI' hackathon, which took place in Munich in June. The award was for the team's demonstration of a governed AI-assisted engineering workflow for software-defined vehicle (SDV) development.

'To meet the increasingly complex requirements demanded by software-defined vehicles, we created a prototype that shows how AI can orchestrate engineering tasks without sacrificing the rigour or essential human judgement required for safety-critical development,' said Travis Bone, principal solutions architect at TASKING, who created the prototype together with Felix Martin, research engineer at TASKING, DLR graduate student Pablo Jahnen and Simon Achatz, principal engineer at Infineon. 'By combining TASKING's compliance and verification technologies with trusted semiconductor products and knowledge from Infineon, we showed how engineering teams can reduce manual handoffs, accelerate verification and generate the evidence needed throughout the V-cycle while keeping engineers firmly in control.'

### Technical innovations

According to TASKING's announcement, the prototype developed for the hackathon used AWS Kiro AI agents to orchestrate engineering activities across the software development lifecycle, connecting Infineon device specifications, product updates and errata with implementation, TASKING coding-standard analysis, automated unit-test generation, virtual electronic control unit (ECU) execution, structural coverage analysis and evidence reporting. Using the TASKING AI Framework, agents iteratively evaluated results and refined code or tests until coding-standard findings, requirements coverage and structural coverage were resolved or appropriately documented. The engineers maintained oversight and approval of every critical decision.

Designed as a proof of concept for deployment through AWS Marketplace, TASKING states that the architecture illustrates how cloud-delivered compliance and verification capabilities can make advanced safety-critical engineering workflows more accessible to distributed development teams.

'AI can generate code, but in safety-critical development the real challenge is ensuring that implementation, verification and compliance remain connected to trusted engineering



knowledge,' said Simon Achatz. 'Our contribution was to bring Infineon's device specifications, product updates and errata directly into the workflow so that the AI-assisted decisions are grounded in the actual hardware context. Combined with deterministic verification and human oversight, this provides a practical path toward scaling software-defined vehicle development.'

'As judges, we look for solutions that pair technical innovation with real-world impact,' said Stefano Marzani of AWS. 'The winning project stood out because it paired a clear understanding of the challenges facing software-defined vehicles with the ability to bring the right technologies together into a single, executable workflow that can actually solve them.'

### TASKING's integrated toolchain is now available on SDVerse

In September, TASKING announced that it had joined the SDVerse business-to-business (B2B) marketplace for automotive software. Through its participation in SDVerse, TASKING will make its automotive software development capabilities more accessible to original equipment manufacturers (OEMs) and Tier 1 suppliers developing safe, secure and increasingly complex software-defined vehicles.

🔗 [sdverse.auto](https://sdverse.auto)

### TASKING and WITTENSTEIN partner on OS-aware timing analysis

In September TASKING announced a tighter integration between its unified toolchain and WITTENSTEIN high integrity systems' pre-certified safety real-time operating system SAFERTOS®. According to the announcement, using TASKING's embedded software trace tool (SWAT) developers can perform comprehensive OS-aware timing analysis without requiring dedicated on-chip hardware trace support.

🔗 [tasking.com/compile-debug-test](https://tasking.com/compile-debug-test)

## A manifesto for embedded systems and smart manufacturing

Roberto Giorgi, University of Siena

The CINI National Laboratory on Embedded Systems and Smart Manufacturing has launched the *Manifesto on Embedded Systems and Smart Manufacturing*, an Italian community initiative aimed at highlighting a vision for a strategic role of embedded systems in Europe's digital and industrial future. CINI is the Italian National Interuniversity Consortium for Informatics, bringing together academic competences across the country; its national laboratories promote collaboration between universities, research groups, industry and institutions. The manifesto is promoted by the CINI National Laboratory on Embedded Systems and Smart Manufacturing and is already supported by an initial group of more than 100 scientists from Italy and across Europe.

The manifesto starts from a simple but strategic observation: digital and technological sovereignty cannot be reduced to semiconductor production, cloud infrastructures or artificial intelligence alone. These are essential ingredients, but they become industrial value only when integrated into complete embedded and cyber-physical systems able to operate reliably in the physical world.

Embedded systems are the layer where hardware, software, data, artificial intelligence (AI) models, communication, timing constraints, safety, security, energy efficiency, certification, and maintainability converge into real operational behaviour. This makes them central to the next generation of intelligent distributed systems, from edge-cloud infrastructures to autonomous machines and industrial platforms.

The vision is particularly relevant for smart manufacturing, where innovation is not simply a matter of adding sensors, dashboards or isolated AI tools. Instead, factories must evolve into trustworthy cyber-physical ecosystems that connect machines, operators, digital twins, industrial data spaces, and production processes. Key use cases include advanced manufacturing, mobility, healthcare technologies, robotics, energy systems, environmental monitoring and agri-food applications.

By promoting openness, interoperability, security, sustainability, and advanced training, the manifesto aims to strengthen collaboration among academia, industry, and institutions, helping Italy contribute actively to Europe's technological capabilities.

To read the manifesto, visit [manifesto-embed-smart.it](https://manifesto-embed-smart.it)

## Torsten Hoefler's DaCeGPT ERC to explore LLMs for performance optimization with Daisytuner



ETH Zurich Professor Torsten Hoefler has been awarded an ERC Proof of Concept grant for the project DaCeGPT (DGPT). This project further develops work from his ERC Consolidator grant, which developed the data-centric (DaCe) parallel programming paradigm, separating scientific program logic from performance optimization. The project developed new programming models for large-scale computing and introduced a spatial perspective on program structure.

Building on this foundation, DaCeGPT applies large language models (LLMs) to performance optimization in scientific computing, with the goal of making advanced optimization accessible to researchers without expertise in high-performance computing. It targets applications such as weather, climate and quantum simulations, which must run efficiently on increasingly complex hardware, including graphics processing units (GPUs) and specialized accelerators.

DGPT combines artificial intelligence (AI) with the DaCe framework, using language models as an explainable interface rather than an unreliable direct code generator. Scientific code written in languages such as Python or Fortran is translated into DaCe's intermediate representation, where performance aspects such as data movement and parallelism can be analysed and optimized. The system then tests and explains optimization strategies using benchmarking feedback.

The project is carried out in collaboration with Daisytuner, a start-up co-founded by Lukas Trümper, a former member of Torsten Hoefler's research group and winner of a HiPEAC Technology Transfer Award in 2024.

### FURTHER INFORMATION:

[bit.ly/Torsten\\_Hoefler\\_DaCe\\_GPT](https://bit.ly/Torsten_Hoefler_DaCe_GPT)



## ETH Zurich team 'RACKlette' wins ISC 2026 Student Cluster

In June, the ETH Zurich team 'RACKlette' won first place in the 15th Annual ISC Student Cluster Competition, held at ISC 2026 in Hamburg. Ten teams reached the finals from universities around the world; RACKlette finished ahead of Tsinghua University (China) and Nanyang Technological University (Singapore). This is the second time that RACKlette has won the ISC Student Cluster Competition.

The team prepared for the competition over several months with guidance from HiPEAC member Torsten Hoefer (ETH Zurich) and Hussein Harake, a senior HPC system engineer at CSCS (Swiss National Supercomputing Centre).

'Congratulations to the ETH Zurich students, their mentors and everyone who contributed to this outstanding achievement. The victory highlights ETH Zurich's strength in educating the next generation of HPC experts, supported by CSCS's hands-on expertise in designing and operating advanced computing systems', said CSCS Associate Director Maria Grazia Giuffreda.

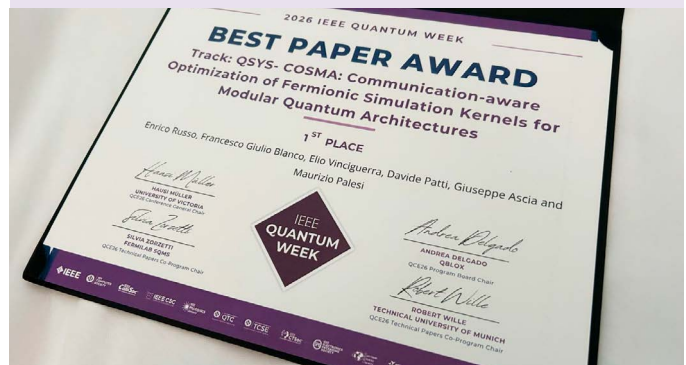
Further information [bit.ly/Racklette\\_ISC2026](https://bit.ly/Racklette_ISC2026)

## COSMA wins best paper award at IEEE Quantum Week

The paper 'COSMA: Communication-aware Optimization of Fermionic Simulation Kernels for Modular Quantum Architectures' by Enrico Russo, Francesco G. Blanco, Elio Vinciguerra, Davide Patti, Giuseppe Ascia, and Maurizio Palesi won best paper in the Quantum System Software track at IEEE Quantum Week in September. In a post announcing the win, Enrico said: 'As quantum computers scale toward modular multi-QPU architectures, communication between quantum cores becomes an increasingly important bottleneck. With COSMA, we explore a communication-aware compilation approach that jointly optimizes fermion-to-qubit mapping, Pauli scheduling, and qubit allocation for fermionic simulation workloads, and we hope it can contribute to ongoing work toward more practical large-scale quantum simulation.'

[arxiv.org/abs/2607.09381](https://arxiv.org/abs/2607.09381)

This research was undertaken as part of the QUADRATURE project. To learn more about QUADRATURE, turn to our interview with Carmen G. Almudéver on p.20.



## PULP Platform team win best paper at EuroMicro DSD

A team of PULP Platform researchers from the University of Bologna, STMicroelectronics, ETH Zurich and the Fondazione Chips-IT were awarded best paper at Euromicro Conference Series on Digital System Design (DSD), which took place in Kraków on 2-4 September. The paper, 'A Fine-Grained Multithreaded RISC-V Core for Efficient and Predictable Real-Time Microcontrollers', was authored by Arbi Gunbardi, Riccardo Tedeschi, Filippo Grillotti, Fabio De Ambroggi, Elio Guidetti, Luca Benini, and Davide Rossi.

The slides are available at [pulp-platform.org/docs/dsd2026/FGMT\\_RISCV\\_CORE\\_Arbi\\_Gunbardi.pdf](https://pulp-platform.org/docs/dsd2026/FGMT_RISCV_CORE_Arbi_Gunbardi.pdf)



## Antonio J. Peña, Roberto Giorgi and Francesco Flammini named IEEE Distinguished Contributors



HiPEAC members Antonio J. Peña (Barcelona Supercomputing Center–Centro Nacional de Supercomputación), Roberto Giorgi (University of Siena) and Francesco Flammini (University of Florence / Dalle Molle Institute for Artificial Intelligence–IDSIA USI–SUPSI) have been recognized as distinguished contributors by the IEEE Computer Society. HiPEAC associate members Rajiv Gupta (University of California, Riverside) and Timothy M. Pinkston (University of Southern California) were also included in the class of 2026.

The title recognizes IEEE members who have made technical contributions to the computing profession, with a significant portion of those technical contributions made through association with the IEEE Computer Society. Eligible contributions include keynotes and presentations, publications, papers presented, standards, patents, working groups, and training events.

The number of members recognized as distinguished contributors in a calendar year is limited to 0.5% of the total IEEE Computer Society membership.

[bit.ly/IEEE\\_DC\\_2026](https://bit.ly/IEEE_DC_2026)

## Alastair Donaldson elected to Royal Academy of Engineering



HiPEAC member Alastair Donaldson (Imperial College London) has been elected as a fellow of the UK's Royal Academy of Engineering. According to the Royal Academy of Engineering's announcement, the 2026 cohort of fellows are making an 'impact across engineering and technology with innovations in clean energy, AI, medicine and public health'.

'Right now, artificial intelligence is radically transforming what software engineering is, what it enables us to build, and the engineering challenges that follow,' said Alastair in a post announcing the election. 'One major challenge is whether we can build software systems that are secure against the increased cyber-threat posed by AI. At the same time, AI itself is a major enabler of formal methods – a suite of techniques that allow us to build systems with the requisite security guarantees. I am grateful to the academy for recognizing software engineering and fundamental computer science through this year's appointment of new fellows – and in particular appointing several of us who specialize in formal methods.'

[raeng.org.uk/news/new-fellows-2026](https://raeng.org.uk/news/new-fellows-2026)

## Aida Todri-Sanial and Hai (Helen) Li win IEEE CEDA Outstanding Service Recognition

In April, HiPEAC member Aida Todri-Sanial (Eindhoven University of Technology) was presented the IEEE Council on Electronic Design Automation (CEDA) Outstanding Service to the EDA Community, in recognition of her role as the general chair of the Design Automation and Test in Europe (DATE) conference in 2025.

At DAC 2026 in May, HiPEAC associate member Hai (Helen) Li (Duke University) received the same award for her role as general chair of DAC 2025.

[bit.ly/IEEE\\_CEDA\\_OS\\_2026](https://bit.ly/IEEE_CEDA_OS_2026)



## Tushar Krishna wins 2026 Maurice Wilkes Award

HiPEAC associate member Tushar Krishna (Georgia Tech) has won the 2026 Maurice Wilkes Award, which is given by the Association for Computing Machinery's (ACM) Special Interest Group on Computer Architecture (SIGARCH) for an outstanding contribution to computer architecture by an early- or mid-career researcher. The award, which was presented during the International Symposium on Computer Architecture (ISCA) 2026, was made in recognition of Tushar's 'outstanding contributions to architectures and modelling tools for large-scale AI systems'.

[bit.ly/Maurice\\_Wilkes\\_2026](https://bit.ly/Maurice_Wilkes_2026)



## Sarita Adve receives 2026 Alan D. Berenbaum Distinguished Service Award



In June, HiPEAC associate member Sarita Adve (University of Illinois Urbana-Champaign) was named the recipient of the 2026 ACM SIGARCH Alan D. Berenbaum Distinguished Service Award. The award, which was presented at ISCA 2026, was given to

Sarita for her 'sustained and transformative contributions to ACM SIGARCH, the broader architecture community and, via CARES, the ACM SIG ecosystem'.

'Research awards reward discoveries; this one specifically honors the work of building and sustaining a community – the mentoring, organizing, governance, and culture-building that often happen quietly behind the scenes,' said Sarita, on receiving the award.

[bit.ly/Illinois\\_Sarita-Adve-ADB\\_2026](https://bit.ly/Illinois_Sarita-Adve-ADB_2026)

## Dates for your diary

### HiPEAC 2027: High Performance, Edge And Cloud computing

18-20 January 2027, Glasgow, UK

[hipec.net/conference](https://hipec.net/conference)

Sponsorship opportunities available

[sponsorship@hipec.net](mailto:sponsorship@hipec.net)



Photo: © Glasgow Life

### HiPEAC webinars

Check the HiPEAC website to keep up to date on forthcoming dates

[hipec.net/webinars](https://hipec.net/webinars)

### REACH 2026: Reach Emerging Architectures in Computing Horizons conference

9-10 November 2026, London, UK

[reach.theiet.org](https://reach.theiet.org)

### EF ECS 2026

2-3 December 2026, Helsinki, Finland

[efecs.eu](https://efecs.eu)

### ARC 2027: 23rd International Symposium on Applied Reconfigurable Computing

31 March to 27 April 2027, Berlin, Germany

[arc2027.org](https://arc2027.org)

### EuroSys 2027: European Conference on Computer Systems

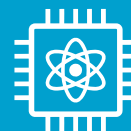
19-23 April 2027, Rabat, Morocco

[2027.eurosys.org](https://2027.eurosys.org)

### ISC High Performance

7-11 June 2027, Hamburg, Germany

[isc-hpc.com](https://isc-hpc.com)



In 2023, we spoke to HiPEAC member Carmen G. Almudéver (Universitat Politècnica de València) about QUADRATURE, the four-year proof-of-concept (Pathfinder) project funded by the European Innovation Council (EIC) to explore the possibilities of scaling a quantum computer from a full-stack perspective. We caught up with Carmen to learn about what's changed in the field since then, whether quantum is reaching a tipping point, and what QUADRATURE has achieved as it enters its final phase.

# 'We would love to demonstrate that our modular approach can be used to scale up quantum computers'

*Since we last spoke [see HiPEACinfo 70], what have been the main developments in the field of quantum computing?*

One of the most significant developments has been in quantum error correction. Current quantum processors typically have physical error rates in the range of  $10^{-2}$  to  $10^{-4}$  per operation (i.e. between one in 100 and one in 10,000 quantum operations), depending on the technology and type of operation. While this represents substantial progress, these error rates still need to be reduced and, more importantly, combined with effective quantum error-correction schemes to achieve reliable, large-scale computation.

There have been significant achievements in encoding logical qubits and in performing fault-tolerant logical operations. A major milestone was Google's demonstration with its Willow processor in 2024, which showed that the logical error rate can decrease exponentially as the number of physical superconducting qubits contributing to a logical qubit is increased, a key requirement for scalable quantum error correction. In addition, important results have been reported across several quantum hardware platforms. A team of researchers at Harvard, MIT and QuEra demonstrated error correction across 96 logical qubits on a reconfigurable neutral-atom architecture, while Quantinuum has demonstrated computations on trapped-ion platforms using 48 to 94 encoded logical qubits.

### A note about qubits: Physical and logical

Microsoft defines a qubit as follows: 'A qubit, short for "quantum bit", is the basic unit of information in quantum computing. Unlike a classical bit, which can represent either a 0 or a 1, a qubit can exist in a combination of both states at the same time through a property known as superposition'. A physical qubit uses quantum properties to do computations. By bundling several physical qubits together into a virtual 'logical qubit', it is possible to increase fault tolerance.

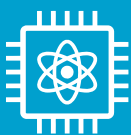
Another major direction we have seen over the last few years is the integration of quantum computers into high-performance computing (HPC) infrastructures. As we have discussed previously, quantum computers are unlikely to operate as standalone systems. Instead, they are increasingly being conceived as accelerators within heterogeneous computing environments, combining central processing units (CPUs), graphics processing units (GPUs) and quantum processing units (QPUs). IBM refers to this vision as 'quantum-centric supercomputing' while NVIDIA is developing infrastructure such as NVQLink to tightly connect GPUs and QPUs. This integration will become particularly important for quantum error correction, which requires substantial classical computing resources for tasks such as real-time decoding, calibration and control. In this sense, the quantum computer of the future is likely to be not an isolated machine, but one component of a tightly integrated classical-quantum HPC system.

In Europe, we've seen several examples of quantum computers being integrated into HPC facilities – for example at Jülich and Barcelona Supercomputing Center. These initiatives are still largely at an experimental stage; the quantum systems are relatively small, and the integration of QPUs with classical



The QUADRATURE project team

Photo courtesy of Universitat Politècnica de València, 2025



HPC resources and workflows is an active area of development. The real benefits of this integration will become more apparent as quantum processors scale up and can be incorporated into hybrid quantum-classical workloads involving computationally demanding classical and quantum tasks. Importantly, these facilities are also developing access models similar to those used for HPC resources, allowing researchers to access quantum processors as shared computational resources alongside conventional supercomputing infrastructure.

***Has one qubit implementation emerged as a clear frontrunner, or are there still various technologies in play?***

Over the last year, the landscape of qubit implementations has remained broadly unchanged. Spin, superconducting, trapped-ion, neutral-atom and photonic qubits each offer distinct advantages and face different challenges. There is no clear winner across all dimensions; rather, different platforms may be better suited to different applications and requirements. Some researchers envision hybrid quantum systems that combine different qubit technologies, while others expect a particular implementation to dominate a given application, whether in quantum computing or quantum communication.

***There have been multiple funding rounds for quantum computing companies in the last couple of years. Are we reaching a tipping point in terms of commercial viability?***

There has been a significant amount of investment in quantum computing companies in recent years, alongside increasing knowledge and technology transfer from universities to industry. However, the main customers today are still national labs or large research facilities, such as HPC centres, which acquire quantum computers primarily for research and experimentation. Also, companies in pharmaceuticals, chemicals, finance, automotive, aerospace, energy and telecommunications are experimenting with quantum computing, generally through partnerships or cloud access rather than buying a complete quantum computer.

The real test of commercial viability will come when quantum computers can solve useful problems. It is still difficult to predict when this point will be reached. As Nobel laureate physicist Michel Devoret said, ‘technology is less predictable than the weather’.

***Has the rapid evolution in artificial intelligence (AI) had an effect on the field?***

I would say there is a two-way relationship between quantum computing and AI. AI is already being used to improve quantum systems, particularly in the laboratory. For example, machine-learning techniques can help with qubit calibration, error detection, system stabilization and the simulation and optimi-

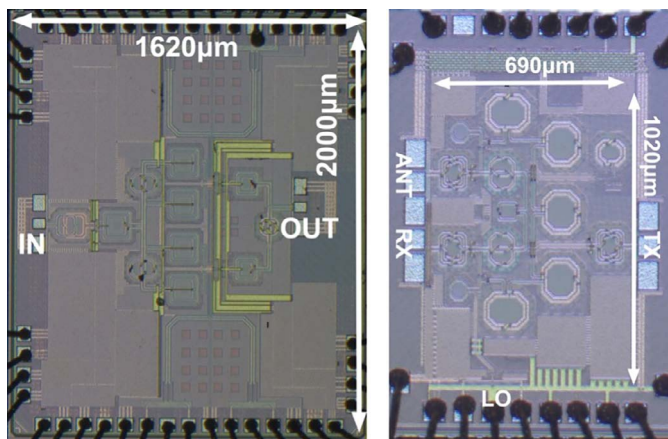
zation of quantum circuits. As a result, AI is having an impact on both quantum hardware and quantum software.

The use of quantum computing to improve AI is a more emerging field, but it could become very promising. One area of interest is using quantum computers for optimization problems that arise in AI and machine learning. There is also growing research into hybrid quantum-classical approaches, in which classical machine-learning models, such as neural networks, are combined with quantum circuits. The goal is to determine whether quantum resources can provide a practical advantage for specific machine-learning tasks.

***What has QUADRATURE achieved in the time since we last spoke?***

The goal of QUADRATURE [as discussed in *HiPEACinfo* 70] is to explore the feasibility of a modular quantum computing architecture in which multiple moderate-sized quantum chips, or ‘Qcores’, are interconnected to scale quantum computers, rather than integrating a very large number of qubits into a single QPU. This approach relies on a communications network with two complementary components: a quantum link based on a microwave cavity, which enables quantum communication between QPUs, and a classical wireless network that orchestrates and synchronizes the quantum transfers. Both networks are designed to operate at cryogenic temperatures.

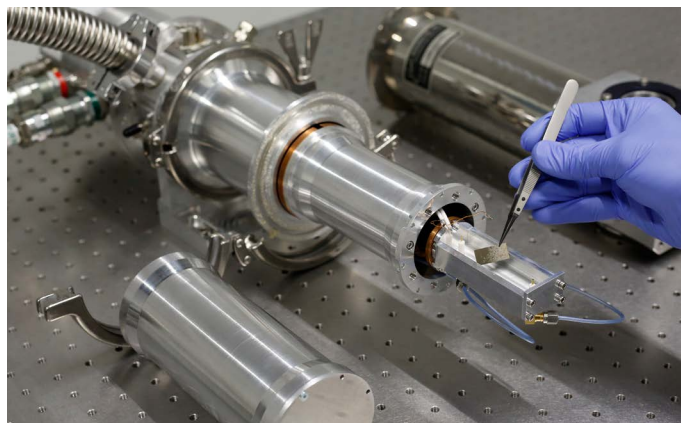
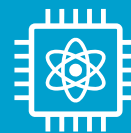
The project encompasses an experimental part and a simulation part. The experimental part involves fabricating the qubit device and the quantum link, which features microwave cavity channels. The crucial part is the integration of both to determine whether it’s possible to transmit quantum data between the Qcores. With regard to the classical network, most of the components (including cryo-antennas and wireless transceivers) have now been designed and fabricated. We are now in the phase of integrating the different elements. ▶



*Images of the QUADRATURE transmitter chip (left) and circulator chip (right)*

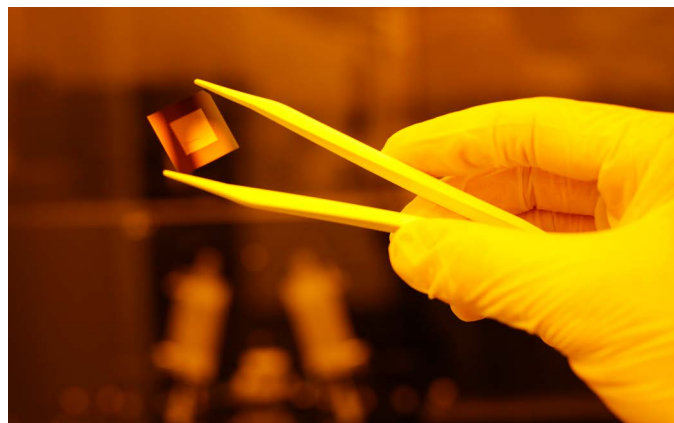
*Images courtesy of Delft University of Technology, 2026*

*Acknowledgements: Fabio Sebastiano, Masoud Babaie, Heqi Deng & Zenghui Li*



University of Siegen laboratory

Photos courtesy of University of Siegen, 2026. Acknowledgements: Peter Haring-Bolivar & Emrah Dursun



On the simulation side, we are exploring and optimizing the proposed architecture and determining under which conditions a modular architecture can outperform a single-core design. To study large-scale quantum systems, we have developed a scalable simulation framework specifically for modular quantum computing architectures. The framework will allow us to derive design guidelines and identify configurations that offer a performance advantage. It integrates different models, including cavity models, and is being used to evaluate approaches to distributed quantum computing. In parallel, we have been developing tensor-network techniques that are aligned with the project's modular architecture and help us investigate the behaviour of larger quantum systems.

*Have there been changes to the original project plan? What are the main technical challenges you've faced?*

One of the main changes has been including quantum error correction. We were already aware of its importance when we wrote the project proposal, as it is one of the main motivations for scaling up the number of qubits, but its relevance has grown significantly in recent years. The project now includes an exploratory task focused on how error correction could be integrated into the modular architecture.

***"Our goal is to demonstrate that the modular approach and the use of quantum links to transmit quantum information between QPUs together with a classical network can provide a viable path towards scaling quantum computers"***

In terms of technical challenges, we have experienced some delays in the fabrication process, which are not unusual in hardware projects. Another challenge has been scaling the simulation of this novel quantum architecture to larger systems.

***What are the next steps for the project?***

The next steps can be summarized as integration on the experimental side and scaling up the simulations. QUADRATURE is a proof-of-concept project: our goal is to demonstrate that the modular approach and the use of quantum links to transmit quantum information between QPUs together with a classical network can provide a viable path towards scaling quantum computers.

The nice thing about this project is the level of engagement from the full consortium. All the partners participate in the discussions, and this has really helped all the pieces of the puzzle to come together.

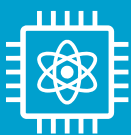
## QUADRATURE

*Scalable Multi-Chip Quantum Architectures Enabled by Cryogenic Wireless/Quantum-Coherent Network-in-Package*

### FURTHER INFORMATION:

QUADRATURE project website [quadrature-project.eu](https://quadrature-project.eu)

QUADRATURE has received funding from the European Union's Horizon Pathfinder programme, under grant agreement no.101099697. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.



# Laying the foundations for quantum, from hardware to software to access

As the quantum computing roadmap progresses, there is work to be done across the computing stack. In this article, members of the HiPEAC community share recent advances across three different areas: hardware manufacturing, software programming, and access to quantum computers.

## The future of quantum computing will be determined by manufacturing

Rui Pereira and Katrin Tina Möbius, Fraunhofer EMFT



The path to future high-performance quantum computers runs through semiconductor manufacturing. Moving from

individual high-performance qubits to powerful quantum processors demands reproducible, scalable, and industrially viable manufacturing processes. At Fraunhofer EMFT, researchers are developing CMOS-compatible processes that combine cutting-edge component performance with industrial scalability.

A critical challenge remains for the industrial application of quantum computers: manufacturing thousands of nearly identical qubits with consistent quality. Only when manufacturing processes achieve the same level of reliability and control as modern semiconductor technologies will economically viable quantum computers become possible. On its 200-mm CMOS pilot line, Fraunhofer EMFT is developing manufacturing processes for superconducting qubits that bridge

quantum physics requirements with established semiconductor manufacturing standards.

Among the most critical circuit elements in superconducting qubits are coplanar microwave resonators and Josephson junctions. Both significantly impact overall system performance. Using a fully aluminium-based process, Fraunhofer EMFT has demonstrated energy relaxation times ( $T_1$ ) of up to 200 microseconds – an important milestone for high-performance superconducting quantum devices.

However, the real challenge lies in achieving reproducible performance at the wafer level. Therefore, further development of the fabrication process focuses on several key areas: optimizing material quality, wafer-to-wafer reproducibility, characterizing and optimizing the Josephson junction oxide barriers, and continuously improving the quality factors of coplanar resonators across various superconducting materials. Simultaneously, the institute is working on the CMOS-compatible integration of alternative superconductors as well as on the selective incorporation of non-CMOS-compatible materials when technological advantages warrant it.

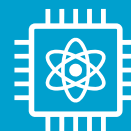


200 mm wafer with Al Josephson junction test structures for superconducting qubits © Fraunhofer EMFT/ Bernd Müller



Wafer processing in the Fraunhofer EMFT cleanroom

© MQV / Jan Greune



## Special feature: Quantum computing

A recent Fraunhofer EMFT publication (see ‘Further reading’) on CMOS-compatible titanium nitride (TiN) resonators fabricated on 200-mm wafers describes just how critical individual process steps are to component performance. The investigations showed that it is not the superconducting material itself that has the greatest influence on losses, but rather the fabrication steps during patterning. Parameters such as etching processes, photoresist removal, and oxide removal strategies influence the resulting superconducting properties. By specifically optimizing these parameters, the research team developed a scalable, low-loss process achieving internal quality factors of nearly one million.

In line with its materials and process research, Fraunhofer EMFT is working on developing qubit chip-scaling strategies for the next generation of superconducting quantum processors. This includes concepts for 3D integration, chip-to-chip interconnects, and through-silicon vias. This approach is complemented by pre-characterization and measurement techniques at room temperature, enabling faster process evaluation and early prediction of device performance.

Fraunhofer EMFT’s process toolkit covers the entire manufacturing workflow – from the blank silicon wafer to the fully 3D-integrated quantum chip. Supported by statistical process control and functional testing at the wafer level, the modular process modules enable flexible integration into customer-specific or partner processes within defined design and contamination limits. This manufacturing expertise is already being applied in collaborations with industry partners such as Oxford Quantum Circuits (OQC).

At the same time, as a partner in Munich Quantum Valley, Bavaria’s quantum computing ecosystem, Fraunhofer EMFT

contributes its expertise in semiconductor manufacturing to the joint effort of establishing a competitive quantum technology landscape in Germany and Europe.

Quantum computing is entering a new phase in which manufacturing capability is becoming a strategic asset. As the field moves beyond laboratory demonstrations toward industrial-scale systems, scalable fabrication processes will be essential for unlocking the technology’s full potential. With its expertise in semiconductor manufacturing, Fraunhofer EMFT contributes the process know-how needed to support this transition and to strengthen Europe’s role in the emerging global quantum ecosystem.

Dr Rui Pereira is head of the Quantum Technologies Program at Fraunhofer EMFT. Katrin Tina Möbius is manager for marketing and strategic communication at Fraunhofer EMFT.

### FURTHER READING:

‘Wafer-based Superconducting Qubit Manufacturing Processes’, Fraunhofer EMFT [bit.ly/FraunhoferEMFT\\_qubit\\_manufacturing](https://bit.ly/FraunhoferEMFT_qubit_manufacturing)

S.J.K. Lang, et al. ‘Optimizing CMOS-compatible, superconducting titanium nitride resonators: Deposition conditions and structuring processes’. Appl. Phys. Rev. 1 September 2026; 13 (3): 031403  
[doi.org/10.1063/5.0331943](https://doi.org/10.1063/5.0331943)

‘Oxford Quantum Circuits (OQC) and Fraunhofer EMFT Partner to Strengthen Quantum Fabrication Capabilities’ (press release, November 2025)  
[bit.ly/FraunhoferEMFT\\_OQC](https://bit.ly/FraunhoferEMFT_OQC)

Munich Quantum Valley [munich-quantum-valley.de/](https://munich-quantum-valley.de/)

## Programming quantum computers made easy

How Eclipse Qrisp offers a European open-source solution for technological sovereignty in quantum computing



Nikolay Tcholtchev,  
Hochschule RheinMain  
– University of Applied  
Sciences and Arts and  
Fraunhofer Institute for  
Open Communication

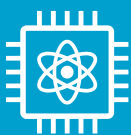
Systems (FOKUS), and Sebastian Bock, Fraunhofer FOKUS

### Introduction

Eclipse Qrisp is a high-level (embedded) programming language designed to provide low-threshold access to quantum computer programming for a broad range of developers and end users

– including those beyond the quantum community currently dominated by physicists and mathematicians. Initiated at Fraunhofer FOKUS in Berlin and now an open-source project within the Eclipse Foundation, Eclipse Qrisp abstracts complex quantum operations on qubits and gates as user-friendly constructs (branches, loops, variables, data types, algorithmic primitives, etc.) in the sense of high-level programming.

Eclipse Qrisp is implemented as an extension of the Python programming language in the form of an embedded domain-specific language (eDSL) that enables the development of hybrid quantum algorithms, facilitating the interplay of different types



of computing units – such as quantum processing units (QPUs), graphics processing units (GPUs) and central processing units (CPUs) – and thus the efficient, demand-oriented solution of complex problems. This is also made possible by the fact that, through the Python embedding, the potential of the entire Python ecosystem in terms of libraries and available developer resources can be leveraged.

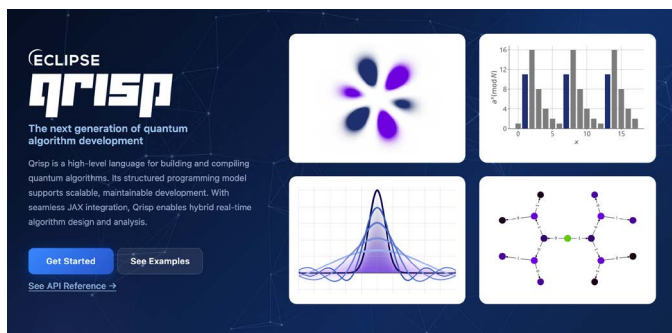
We are firmly convinced that the eDSL approach for high-level quantum programming will prevail over the idea of defining specific languages – with their own grammar and ecosystem – because it can be more easily integrated into existing information technology (IT) products and services.

### The Eclipse Qrisp programming model

The Eclipse Qrisp programming model builds on several core concepts aimed at enabling easy access to quantum computer programming for experts beyond the physicist and mathematician communities. The following main concepts are significant:

- Use of abstracted constructs from higher-level programming languages, such as loops, branches, variables, various (hybrid) computing environments, and data structures.
- Handling variables and functions instead of assembler-like qubits and gates.
- Abstract data types such as QuantumFloat, QuantumChar, QuantumString, as well as data structures built on them, such as QuantumArray and QuantumDictionary.
- Abstract operators for quantum data types (e.g. assignment operator, multiplication operator...).
- Provisioning of algorithmic primitives such as Oracle and Diffuser (for the Grover algorithm), QFT (Quantum Fourier Transform), and others that are relevant for rapid programming of hybrid algorithms and workflows based on current developments in the domain.
- Automatic cleanup and release of qubits – similar to garbage collection in classical algorithms – through specialized algorithms for uncomputation. The term ‘uncomputation’ refers to the mathematical construction of algorithmic circuits such that certain qubits can be reversed to an initial state (i.e. qubits are freed).
- Use of special ‘QuantumSessions’ to manage the state of exchange between the classical code portion and various quantum computing backends.

Based on these constructs, it is possible to quickly and efficiently program various algorithms from current research. Currently, the Eclipse Qrisp distribution offers a broad range of available algorithms and algorithmic frameworks, such as BlockEncoding, Shor, Grover, VQE, LCU, QAOA, QUBO, MaxCut, Quantum Backtracking, Monte Carlo Simulation, Numerical Integration, Stim Simulation and many more.



*The VLQ quantum computer in the IT4Innovations data room*

### Qrisp in action

To demonstrate the capabilities of the Qrisp programming model, we present a short code snippet that showcases the elegance and ease of use.

In this example, we see the multiplication of two floating-point numbers on a quantum computer – in this specific case, the simulator included with Qrisp is used. Two three-qubit floating-point numbers are declared, and the values 0.5 and 3 are assigned to them. Subsequently, the two numbers are multiplied together, and the result is output, with a measurement on the quantum backend being initiated directly as well. The final result is 1.5 with a probability of 1.0.

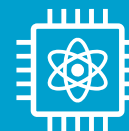
Although this is a very simple example, one can see the power of Eclipse Qrisp in direct comparison to other frameworks that still rely heavily on assembler-like direct circuit programming based on qubits and gates. Multiplying two numbers in Qrisp is significantly shorter than interconnecting gates and modelling floating-point numbers on the available qubits, which would be necessary in the case of a low-level programming framework. By utilizing the additional constructs of the Qrisp programming model mentioned above, a powerful toolkit emerges for the efficient and exploratory development of hybrid quantum algorithms.

### Technological sovereignty

Eclipse Qrisp provides strong support for strategic technological sovereignty in the field of quantum computing. By formulating algorithms and hybrid workflows at a higher abstraction level, the associated program code is platform-independent. This means that the algorithm description contains no specifics (e.g. qubit topology, native gates...) of the target platform for execution. If this target platform is no longer available in the future, a new QPU can be deployed.

The prerequisite for this is special adapters that ensure the translation between the higher-level Eclipse Qrisp constructs and the interfaces of the specific quantum backends. These are





## Special feature: Quantum computing

easy to implement due to the modular Qrisp architecture. In fact, Eclipse Qrisp already provides tested adapters for IQM, AQT and IBM (Qiskit), as well as further adapters that are currently in the process of implementation, such as access to established platforms like KiPU Quantum Hub and NVIDIA CUDA-Q for GPU simulation.

Ultimately, we envision the creation of an abstract operating system-like platform for various QPU and simulation architectures following the open-source paradigm.

To advance this work, Qrisp is the foundation of the Q-VERSE project, which will develop a European quantum software stack comprising components and tools for scalable, technologically sovereign and reliable software solutions that harness the computing power of quantum computers.

### Next steps

As part of the further development and dissemination of Qrisp, teaching at colleges and universities will play an increasingly important role. By using Qrisp in exercises, students can become familiar with modern concepts and tools for quantum computer software development at an early stage.

In parallel, a viable governance structure has been established that enables coordinated and collaborative further development of Qrisp by a growing developer community. A central concern in this context is the continuous strengthening of quality assurance processes as well as the consistent elaboration and

implementation of architecture specifications. These measures serve the long-term maintainability, extensibility, and security of the software.

We are also working intensively on standardization activities, particularly through active participation in international committees such as the working group on 'Quantum Computing and Simulation' in CEN/CENELEC JTC22 on Quantum Technologies. As of now, a specific agile standard activity on a meta-model for a high-level programming language for quantum computing has been established and is actively being pursued.

The goal is to sustainably integrate Qrisp into existing and future technological frameworks and to help shape their further development, thereby establishing Eclipse Qrisp as the open-source programming language and software stack for programming and operating quantum computer devices. In short, we envision Eclipse Qrisp becoming the 'Linux for quantum computing'.

### FURTHER READING:

Qrisp website [↗ qrisp.eu](https://qrisp.eu)

Q-VERSE (Quantum Virtual Ecosystem for a unified Runtime for agnostic Software Engineering) on the Quantum Flagship website [↗ qt.eu/projects/computing/Q-VERSE](https://qt.eu/projects/computing/Q-VERSE)

## The quantum access trail: How QHEAppE empowers users across Europe to leverage QPUs

Jan Swiatkowski, Jakub Konvička, Ladislav Foltyn, Jan Martinovič, Branislav Jansík, IT4Innovations

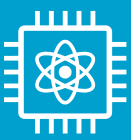
Today's quantum computers (QC) are mostly built as single-user instruments suited to serving small research groups. Scaling this model to thousands of users across many institutions in the EuroHPC Joint Undertaking (EuroHPC JU) landscape, who are executing hybrid quantum-classical workloads, raises two distinct challenges.

The first is a multi-tenancy and resource-governance problem: users must be organized into projects, projects must be allocated quantum processing unit (QPU) time, and policies governing consumption, duration, and priority must be enforced, all while accurately accounting for resource usage per project. Although the QPU is inherently a serial device, processing one

job at a time, users submit jobs concurrently from heterogeneous endpoints such as high-performance computing (HPC) clusters, workstations and PCs. The system must guarantee job scheduling and strict isolation, ensuring no user can access data, jobs, or results belonging to another.

The second challenge concerns co-scheduling: jointly scheduling workloads across HPC and QC resources to achieve optimal overall utilization, taking into account the fundamentally different execution models of classical and quantum processors.

In this article, we present QHEAppE, a quantum-as-a-service architecture that directly addresses multi-tenant access, project-based resource allocation, policy enforcement, and usage accounting to address the first challenge, while estab-



lishing the architectural foundations necessary to adopt a co-scheduling solution to the second challenge.

### Enabling quantum-as-a-service with QHEAppE

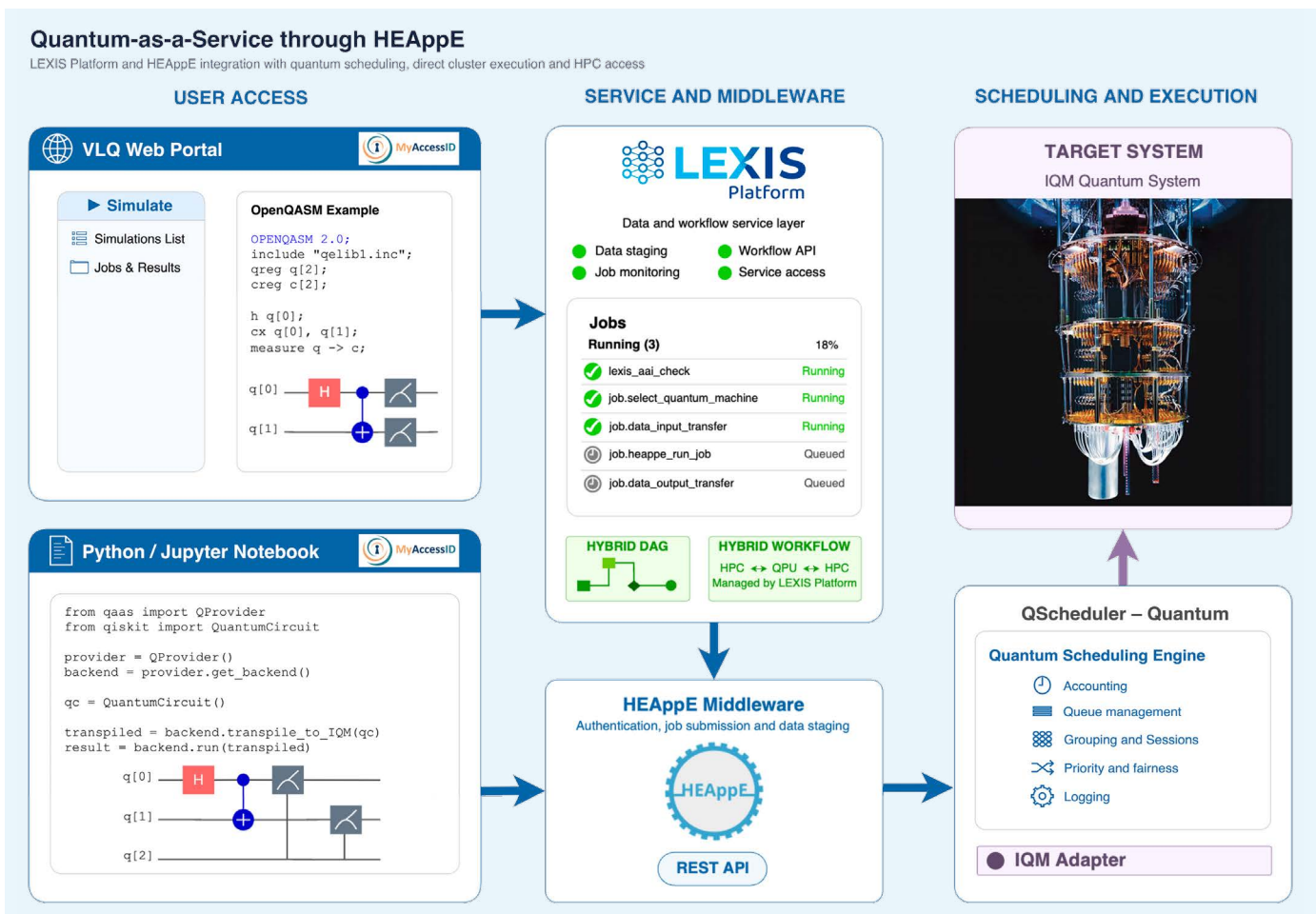
QHEAppE integrates quantum computing resources with HEAppE Middleware to meet the computational demands of both national and EuroHPC JU users. Four core components, including a client-side Python package, called qaas, as a part of QHEAppE, HEAppE Middleware, QScheduler and the LEXIS Platform (see 'Further information'), are integrated to realize the architecture outlined in the figure below. These components work together to enable an end-to-end workflow covering user authentication, quantum workload submission and scheduling, QPU execution, and resource accounting.

From the user's perspective, the workflow starts with the QHEAppE client, provided through the qaas Python library. Users can submit quantum circuits described using Qiskit or OpenQASM 3.0. Pulse-level control is also supported via Pulla and, prospectively, IQM Pulse. The client exposes a Qiskit-compatible interface and abstracts the underlying networking, authentication, and job-submission mechanisms. To access a quantum resource, the user is authenticated and authorized

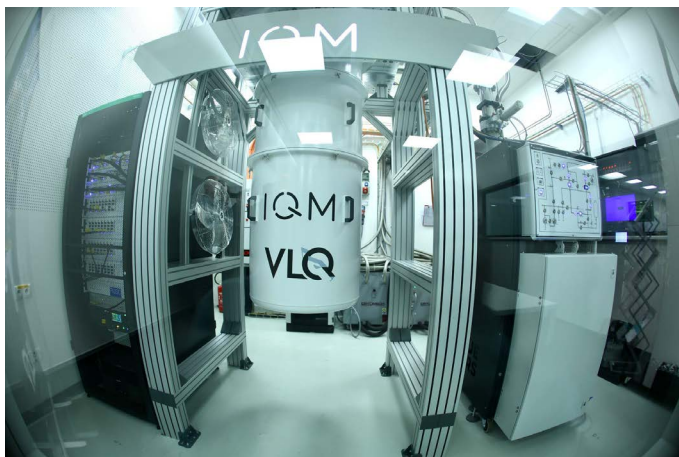
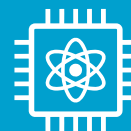
via the LEXIS Platform using MyAccessID, a federated identity provider proxy. Subsequent operations are authorized based on a JSON Web Token (see 'Further information'), used within an interactive OAuth2/OIDC authentication flow.

Once authenticated, the client submits the quantum workload through the HEAppE REST application programming interface (API). HEAppE Middleware acts as an HPC-as-a-service abstraction layer extended to quantum systems and manages the asynchronous job lifecycle, including job definition, payload validation, command staging, submission, state synchronization, and output retrieval.

After authorization, HEAppE forwards the job to QScheduler, a low-latency Rust-based scheduler. QScheduler forms the final scheduling layer between HEAppE and the QPU, isolating the QPU control system from direct external access and brokering execution over a secured internal network. It enforces quantum-specific scheduling and allocation policies, including circuit grouping, job prioritization, resource reservations, runtime limits, maximum shot counts, and execution timeouts. In this way, the QHEAppE system provides a unified



The architecture of QaaS through HEAppE Middleware



*The VLQ quantum computer in the IT4Innovations data room*

workflow, from user authentication and workload submission to controlled QPU execution, result retrieval, and resource accounting.

The QHEAppE architecture is successfully deployed at IT4Innovations to provide access to IQM's Star 24 architecture QPU installed on the VLQ quantum computer, both for Czech users and for LUMI-Q and EuroHPC JU international users.



**HEAppE Middleware** (High-End Application Execution) is an HPC-as-a-service middleware developed by IT4Innovations that provides unified RESTful API access to super-computing infrastructure. By abstracting complex cluster schedulers such as Slurm and PBS, it handles authentication, job submission, monitoring, and data transfer. Recently extended to support quantum processing unit (QPU) tasks, HEAppE enables external applications to securely manage both classical and quantum workflows without manual cluster management.



The **LEXIS Platform** is an advanced engineering framework that integrates HPC, cloud resources, and big-data management into a unified environment. By combining modern web portals, APIs, distributed orchestration, and iRODS-based data storage, it enables users to seamlessly design, manage, and execute complex scientific and industrial workflows across heterogeneous European research infrastructures.

### Proven in a production environment

The QHEAppE quantum-as-a-service architecture addresses the challenges of multi-tenant access, project-based resource allocation, policy enforcement, and usage accounting for shared quantum computing resources. We have developed and implemented this solution in production to serve the VLQ quantum machine operated at the IT4Innovations National Supercomputing Center on behalf of the LUMI-Q consortium and EuroHPC JU, demonstrating its practical viability at scale.

Lightweight and non-intrusive, the QHEAppE architecture integrates with existing vendor quantum stacks without requiring modification, and is easy to deploy across heterogeneous institutional environments. It supports multiple authentication authorities, including pan-European federated identity through MyAccessID, enabling seamless onboarding of users from many institutions while preserving strict data and result isolation.

By establishing this operational foundation, the architecture is a natural fit for hybrid quantum-classical computations and provides the readiness required to adopt future co-scheduling solutions that jointly optimize HPC and QC resource utilization, positioning it as a robust and extensible basis for the wider deployment of shared quantum computing infrastructure across the European HPC ecosystem.

### FURTHER READING:

V. Svatoň, J. Martinovič, J. Křenek, T. Esch and P. Tomancak. 'HPC-as-a-Service via HEAppE Platform', *Advances in Intelligent Systems and Computing*, vol. 993, p. 280, 2019 [bit.ly/HEAppE\\_2019](https://bit.ly/HEAppE_2019)

M. Golasowski, J. Martinovič, J. Křenek, K. Slaninová, M. Levrier, P. Harsh, M. Derquennes, F. Donnat and O. Terzo. 'The LEXIS Platform for Distributed Workflow Execution and Data Management', in *HPC, Big Data, AI Convergence Toward Exascale: Challenge and Vision*, Boca Raton, CRC Press, 2022, p. 17 [bit.ly/LEXIS\\_2022](https://bit.ly/LEXIS_2022)

M. Jones, J. Bradley and N. Sakimura, "JSON Web Token (RFC 7519)," 2015 [rfc-editor.org/info/rfc7519](https://rfc-editor.org/info/rfc7519)

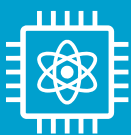
QScheduler on GitHub [github.com/It4innovations/qscheduler](https://github.com/It4innovations/qscheduler)

IBM Qiskit [ibm.com/quantum/qiskit](https://ibm.com/quantum/qiskit)

OpenQASM Live Specification [openqasm.com](https://openqasm.com)

IQM Resonance [iqm.tech/products/iqm-resonance](https://iqm.tech/products/iqm-resonance)

VLQ quantum computer [it4i.cz/en/infrastructure/vlq-quantum-computer](https://it4i.cz/en/infrastructure/vlq-quantum-computer)



Although mainstream quantum computing is still some way off, Max Ern  (StreamHPC) argues that understanding which workloads can be optimized using quantum accelerators is an important exercise to hit the ground running when these accelerators are eventually put to use.

# Preparing today for quantum tomorrow

## Why knowing your workloads will be key to exploiting quantum acceleration

Quantum processors are emerging as specialized experimental accelerators, although the path to practical advantage remains uncertain and workload dependent. How should a company with no experience in quantum computing prepare for this?

The answer will not be ‘port today’s work to specific hardware’. Instead, it is about understanding your workloads well enough to recognize when quantum computing might become relevant and building enough expertise to test that quickly. In particular, accelerator companies are in the perfect position to improve this testing cycle.

### Quantum is different

The evolution from single-threaded, to multi-threaded, to massively parallel environments has come with many changes to algorithm design. But the underlying rules have stayed comparable. You have bits, you can share data, you can store to disk, etc.

With the advent of quantum computers, we don’t have those guarantees anymore. Qubits are fickle creatures, and the rules are convoluted linear algebra. If you so much as look at qubits the wrong way, they get messed up. You cannot attach a debugger to a quantum program to look at its state. At scale, it’s even worse. You can’t copy generic quantum data. And want to send a qubit over the network? That opens an entirely different can of worms about quantum channels.

But where some see problems, we should see opportunities! It is such a different world that it gives completely new ways to attack certain problems.

***“Qubits are fickle creatures, and the rules are convoluted linear algebra. If you so much as look at qubits the wrong way, they get messed up”***

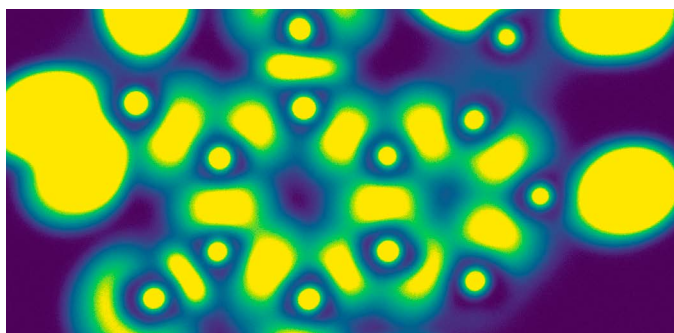


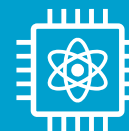
*Illustration of an electron localization function. Quantum chemistry tries to predict molecular properties better by taking quantum effects into account*

### Quantum’s superpower

You may have heard quotes like ‘quantum computers use superposition to calculate all inputs at once’. If this were the case, we could easily port any algorithm that runs on graphics processing units (GPUs) to also run on a quantum processing unit (QPU). Unfortunately, this is too much of an oversimplification. A very important caveat that is often left out is that once you look at a qubit, you only get one output sampled from its internal state. Making proper use of this requires reformulation of the problem so that quantum mechanics can actually help.

There is a lot of literature about which workloads can be optimized by quantum computing already, and at least as much ongoing research. We can divide problems roughly into:

1. Problems with strong theoretical evidence for quantum advantage. Take for instance quantum’s poster child, Shor’s factoring algorithm.
2. Problems with promising algorithms, where practical advantage remains uncertain, like some problems in quantum chemistry.
3. Problems where the quantum formulation is interesting, but with no advantage established yet, such as quantum machine learning.
4. A much larger category where we don’t have a reason to believe quantum helps. This includes most business needs.



## Special feature: Quantum computing

A lot of this depends on specifics. A problem that involves probabilistic methods like Monte-Carlo or random walks might be anywhere from the first to the last category, with similar stories for optimization.

Importantly, just like GPUs were not a replacement for central processing units (CPUs), you cannot, in general, take a GPU workload and expect it to work in a quantum environment. Knowing which category your problem falls into is a matter of experience, which is crucial for knowing when you might need a QPU.

### Quantum today

Currently, qubits are too noisy to really satisfy the mathematical model of quantum computing. Depending on the hardware, it may take many actual qubits or significant computational overhead to form one logical qubit you can work with. Error propagation makes long-running, reliable circuits challenging. As a result, production-scale quantum computing remains out of reach.

But it's not just quantum hardware that's in its infancy – the entire stack is. There's a lot of active research going into optimizing quantum algorithms and compiler development. Higher-level abstractions to improve the circuit creation process are being developed, but many useful analyses (performance, resources, reliability) remain hard.

With young or hard-to-access hardware, simulation is very attractive. Unfortunately, in general every qubit added doubles the internal state. If you restrict your simulation, the runtime gets better, but as soon as the exponential factor disappears, you might not be simulating anything interesting anymore.

On the positive side, large parts of the stack can benefit from today's acceleration already. Of course, being able to simulate larger states better helps. But compilation, optimization, graph-based formulations, and benchmarking are all also classical workloads that can be improved. So even if you do not directly benefit from a QPU, you can still contribute significantly.

---

***“This is the goal of quantum readiness: to understand both quantum computing and your own needs well enough to experiment as the technology develops, and benefit the moment the quantum version overtakes, or even complements, the classical version”***

---

### Quantum tomorrow

So should we just wait until quantum accelerators are available before we do anything with them?

No!

There are some concrete steps we can take to be ready for tomorrow. First, familiarize yourself with the field. Build expertise. And importantly, keep that expertise up to date. The field is still evolving rapidly, and today's 'impossible' might be possible tomorrow.

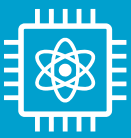
Next, if you encounter a problem that quantum might be good at... make a classical baseline first. Get a thorough understanding of the theoretical and practical boundaries of your problem. Quantum is hard enough as it is; you do not want to have to deal with unrelated problems at the same time.

Only then try making a quantum version. A small-scale proof-of-concept, to see if there is reason to believe you can beat the classical baseline, and whether your problem is actually suitable for quantum acceleration. It does not have to run on a real quantum computer, just a simulation on a small dataset and some informed assumptions are enough. Only then does it make sense to start on the full version.

But be prepared that the full version should be flexible. Hardware is in flux, not much is standardized, theoretical improvements keep compounding, so it is likely you will still have to make changes once quantum truly arrives.

Still, making these changes will be a much quicker process than starting from scratch and playing catch-up. This is the goal of quantum readiness: to understand both quantum computing and your own needs well enough to experiment as the technology develops, and benefit the moment the quantum version overtakes, or even complements, the classical version.

To find out more about StreamHPC's performance optimization services, including quantum-prepared software, visit [streamhpc.com](https://streamhpc.com)



With mainstream quantum computers approaching, the race is on to develop cryptographic algorithms which would be secure against an attack carried out by such machines. However, when it comes to safety-critical systems, implementing quantum-proof algorithms – which may have higher overheads and no fixed worst-case execution time – is not trivial. In this article, Adam Mackay (Head of AI, QA-Systems) speaks to experts working at the intersection of safety and security to consider the way forward.

# You can't patch a safety case

## Future-proofing safety-critical systems against Q-Day

The advent of a capable quantum machine is drawing closer, and with it the ability of quantum computers to break standard public-key cryptography – an event known as Q-Day. The European Union has set an end-of-2030 target to move all critical infrastructure onto a quantum-safe footing.

Post-quantum cryptography (PQC) replaces public-key algorithms that are vulnerable to quantum computers with a new set of algorithms. In 2024, the US National Institute of Standards and Technology (NIST) standardized three PQC algorithms, and the UK's National Cyber Security Centre recommends PQC as the primary response to Q-Day; see 'Further reading'.

For readers who can recall the fuss over the Y2K 'bug', it is tempting to treat this as a similar abstract future risk that can be solved should it become an issue later. However, given the typically long lifespan of the technology deployed in fields such as power management, automotive and avionics, it is prudent to act now to ensure the future security of these systems. Bad actors can (and do) collect encrypted data today in the hope that the encryption can be cracked tomorrow. Quantum systems

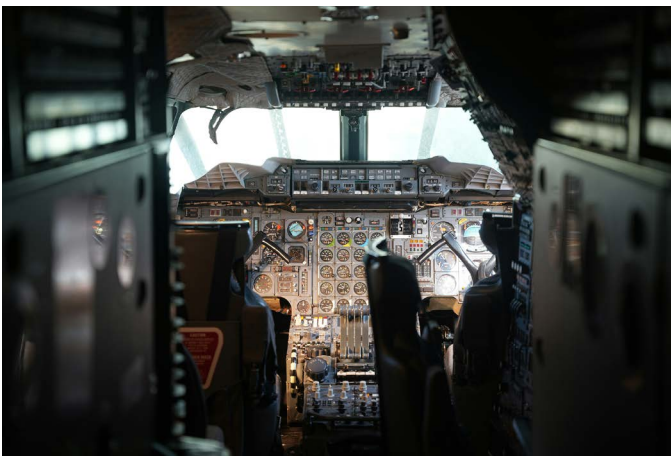
engineer Kieran Bjergstrom expects the first capable machines to be scarce enough that attackers will have to choose which ciphertext is worth their time: 'You don't want to get the prime minister's shopping list.' Long-lived critical infrastructure is exactly the kind of target worth waiting for. Consider an aircraft entering service before the PQC deadline: it is likely still to be flying decades later.

### Building around the algorithm

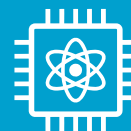
Engineers of safety-critical software deliver both code and a safety argument backed up by evidence about timing, interfaces, independence and failure behaviour. If any of those things change, then safety cases need to be reworked and the system recertified, which is not a simple task. Five specialists from related fields were approached to discuss PQC through this lens.

According to Wanja Zaeske, an avionics researcher at DLR, Germany's national aerospace research centre, cryptography 'can become defunct without wrongdoing, purely through an external observation' – that is, someone, somewhere, discovering a new weakness in the algorithm. This runs counter to the intuitive safety perspective where 'proven in use or relevant service history is usually seen as a good thing'. Cryptographic assurance can weaken over time without anything failing within the envelope of the aircraft.

July 2026 brought an early glimpse of what this type of failure mode might look like for a PQC system when an AI-assisted attack found a vulnerability in one of the candidate algorithms (HAWK). The weakness discovered virtually halved the key strength of HAWK, a post-quantum signature candidate that had survived two years of expert review. Fortunately, no deployed system was affected; the post-quantum schemes standardized by NIST were also unaffected (see 'Further reading'). However, should similar attacks materialize once PQC is deployed, system designers need to retain the ability to patch and upgrade systems with minimal downtime.



*Concorde cockpit at the Museum of Flight. Long-lived aircraft make cryptographic change an assurance problem as much as a software one*



### Missing deadlines

A major consideration for resource-constrained embedded systems is that post-quantum cryptography can take longer than the classical algorithms it replaces. Signing with NIST's Module-Lattice-Based Digital Signature Algorithm (ML-DSA) reveals an even trickier problem: it lacks a fixed worst-case execution time (WCET), which is a significant part of a real-time safety case.

ML-DSA co-designer Peter Schwabe of the Max Planck Institute for Security and Privacy describes the process: an ML-DSA signer creates a candidate signature, which is checked and rejected if its value could reveal information about the secret key. This process is repeated until a leak-free signature emerges. Four or five attempts are typical, depending on the parameter set, but, as Schwabe put it, 'there is no strict upper bound'.

For critical systems, deadlines for every sub-process are a vital part of the certification evidence. Without a deadline, safety certification becomes problematic. However, the Federal Information Processing Standard (FIPS) 204 says implementations 'should not' bound these loops. At first glance, it seems engineers are left with either a system that can be safety-certified or quantum-safe, but not both.

Fortunately, FIPS does make a small allowance here. If the signing loop needs an upper bound, FIPS's published minimum is 814 iterations (NIST's 31 July 2026 list of potential updates would raise that to 821). The bound was chosen so that the chance of a signature running out of attempts is vanishingly small: roughly  $2^{-256}$ , or  $10^{-77}$ , per signature – about the same odds as guessing a 256-bit secret key at the first try.

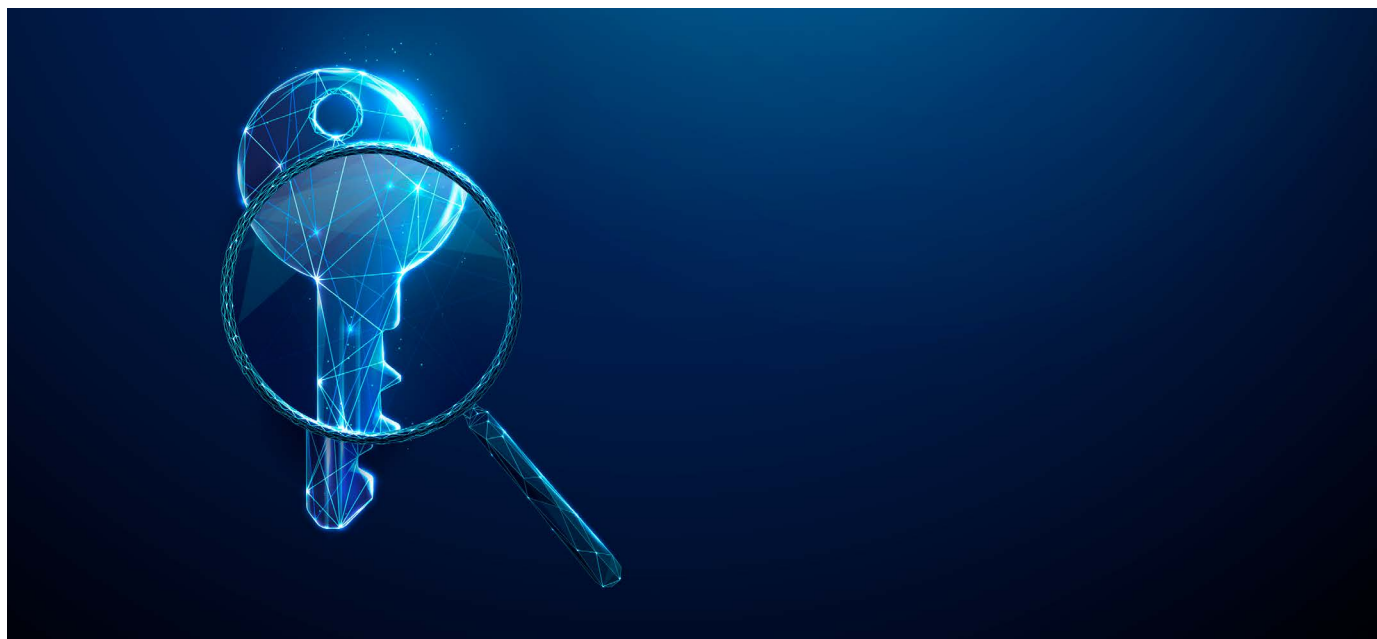
For perspective, aviation calls a catastrophic failure condition 'extremely improbable' at  $10^{-9}$  per flight hour. The quantities are not directly comparable, but the loop-exhaustion probability is about 68 orders of magnitude smaller. The pragmatic engineer would much prefer to bound iterations at a much lower number, quantify the small probability of failure and arrange for the system to enter a safe degraded mode. Schwabe agrees that this makes sense at the application level but would almost certainly preclude FIPS certification.

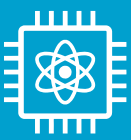
Mark Everitt, founder of Loughborough University's Quantum Systems Engineering Group, describes the larger tension: 'If you're cryptographically agile, you are certificationally challenged to some extent... there's always a trade space.' It appears that ML-DSA signing makes this trade literal.

### Move the boundary, not the risk

Matthias Kannwischer of zeroRISC, a maintainer of the open-source mldsa-native implementation, describes a compromise devised by fellow maintainer Hanno Becker of AWS: make signing restartable. Instead of lowering the overall limit, each call performs a bounded amount of work. If no signature is ready, it pauses and resumes at the same attempt, using the same randomness. 'This way, you produce exactly the same signatures as an unbounded signing loop and at least get a bounded runtime per attempt,' Kannwischer explained. FIPS 204 explicitly permits mathematically equivalent procedures that return the correct output for every input.

In this deployment scenario, the unbounded tail has not disappeared: the total time to obtain a signature remains indeterminate. What has shifted is where the uncertainty is handled.





Rather than a single blocking call, restartability provides a series of bounded schedulable calls. If the application is designed appropriately, it can protect downstream deadlines through a defined degraded mode.

The feature was merged on 14 July 2026 (see GitHub link in 'Further reading'). As of 7 August, Kannwischer was not aware of any public deployment. Whether the solution is enough to satisfy the regulators of safety-certified systems remains to be determined.

Similar tensions between safety certification and quantum safety recur one layer up. Zaeske and colleagues placed a post-quantum-secure cryptographic service inside an ARINC 653 partition, which provides time and space isolation between avionics applications. Applications access the service through a defined interface, allowing the cryptography to be swapped out without changing those application partitions. The caveat is that this only works if the partition has enough timing and memory headroom. Such a modular approach must, of course, be in place before the algorithm changes.

Modular solutions help control risk, facilitate reuse and make parallel development more efficient. Zaeske stresses that pursuing modularization and incremental certification while co-hosting independent applications on one computer is 'where big money is likely to be sunk, and hence it only makes sense to do so if big savings are achievable at all over the entire lifecycle (which requires a certain scope of the overall aircraft program)'.

Bjergstrom raises a similar point when discussing trust migration in such modular systems: 'You need to implement a hybrid key approach whilst the original keys remain valid and remain secure'. The original key must authorize its successor: if you wait until it can no longer be trusted, the migration path closes.

Everitt frames it as a question: 'could the more important requirement be assurance agility?' – an architecture and safety case modular enough that changing the algorithm invalidates only a defined part of the argument. Crypto-agility allows an engineer to replace an algorithm, whereas 'assurance agility' ensures that the surrounding certification evidence is still usable when they do. Post-quantum migration in safety-critical systems is possible, but the work to enable this needs to happen years earlier: in the interfaces, the timing budget, the recovery behaviour and the trust path. You can patch a safety case, but only if you bought into that option at design time, when there may have been no business case.

### FURTHER READING:

'EU reinforces its cybersecurity with post-quantum cryptography', European Commission, 2025  
[bit.ly/EU\\_PQC\\_PR\\_2025](https://bit.ly/EU_PQC_PR_2025)

'NIST Releases First 3 Finalized Post-Quantum Encryption Standards', NIST, 2024  
[bit.ly/NIST\\_3\\_PQC\\_standards](https://bit.ly/NIST_3_PQC_standards)

Quantum networking technologies, UK National Cyber Security Centre, 2025  
[bit.ly/NCSC\\_QNT\\_2025](https://bit.ly/NCSC_QNT_2025)

'What Is Post-Quantum Cryptography?', NIST, updated 2026  
[bit.ly/NIST\\_what-is-PQC](https://bit.ly/NIST_what-is-PQC)

Aircraft end-of-life and average retirement age in 'Sustainability in the end-of-life phase of aircraft', European Union Aviation Safety Agency, 2024  
[bit.ly/EASA\\_aircraft\\_EoL](https://bit.ly/EASA_aircraft_EoL)

'Discovering cryptographic weaknesses with Claude', Anthropic, 2026  
[bit.ly/Claude\\_cryptographic\\_weakness](https://bit.ly/Claude_cryptographic_weakness)

Status Report on the Second Round of the Additional Digital Signature Schemes for the NIST Post-Quantum Cryptography Standardization Process, NIST, 2026  
[nvlpubs.nist.gov/nistpubs/ir/2026/NIST.IR.8610.pdf](https://nvlpubs.nist.gov/nistpubs/ir/2026/NIST.IR.8610.pdf)

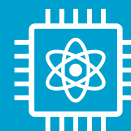
AMC 20-193 Use of multi-core processors, EASA, 2023  
[bit.ly/EASA\\_AMC20-193](https://bit.ly/EASA_AMC20-193)

FIPS 204: Module-Lattice-Based Digital Signature Standard, NIST  
[csrc.nist.gov/pubs/fips/204/final](https://csrc.nist.gov/pubs/fips/204/final)

CS-25 AMC 25.1309 safety objectives, EASA, 2023  
[bit.ly/EASA\\_CS-25\\_AMC25\\_1309](https://bit.ly/EASA_CS-25_AMC25_1309)

mldsa-native on GitHub, including example and applications  
[github.com/pq-code-package/mldsa-native](https://github.com/pq-code-package/mldsa-native)  
mldsa-native signing hooks for observing signature attempts and implementing restartable signing  
[bit.ly/MLDSA-native\\_signing\\_hooks](https://bit.ly/MLDSA-native_signing_hooks)

Varner, K., Zaeske, W., Friedrich, S. et al. Agile, post-quantum secure cryptography in avionics. CEAS Aeronaut J 17, 133–163 (2026)  
[doi.org/10.1007/s13272-025-00806-5](https://doi.org/10.1007/s13272-025-00806-5)



Intrigued by ACACES 2026 teacher Swamit Tannu's exhortation to 'look beyond the qubit', The Quantum Factor's Wisse Hettinga interviewed Swamit, an assistant professor at the University of Wisconsin-Madison to learn about the implications of building an efficient quantum computing stack.

# 'Just as in the classical stack, we need to co-design the system'

This interview has been edited for length and clarity; for the full version visit [youtu.be/SF3Ljpr-89I](https://youtu.be/SF3Ljpr-89I)

*When you say, 'look beyond qubits and consider the full system stack', what are you trying to communicate?*

In classical computing, we have transistors at the bottom of the stack, we build circuits on top of that using transistor properties, and we use these circuits to build blocks, then we have a microarchitecture and a compiler compiles down to the instruction set architecture (ISA). A very similar stack is required for quantum: in fact, it's even more powerful than a classical setting because hardware is continuously changing, so you need a stable system stack that you can use to address all these different quantum hardware modalities.

So even now, in the early days of quantum computing, I think the system architecture and software stack can play an enormous role in terms of scaling computation, making it more programmable, and reducing its costs. Quantum computers promise to solve really hard problems but right now they are extremely costly, so we need a lot of system engineering that can reduce this cost and help democratize quantum computing, so that we can apply this technology to a broad sector of problems.

*Going back to qubits – a lot of research is focusing on this area, and the conclusion seems to be 'the more qubits, the better'. How many qubits do you need for the calculation part and the error-correction part?*

That's a great question, and the answer is changing every day, partly because people are inventing new error-correction



codes that are more and more efficient. For a really long time there was this notion that we have quantum error-correction codes that can build fault-tolerant systems, but that are really expensive in terms of how many physical qubits you need to create a fault-tolerant logical qubit that can tolerate errors and you can perform a large number of computations – which is what we need to run quantum applications. Lately, people have been developing these LDPC [low-density parity check] codes, where you don't need as many physical qubits to create a fault-tolerant logical qubit.

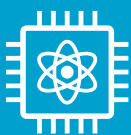
Also, the answer depends on how good your physical qubit is – so if your physical qubit is inherently more reliable, then you don't need a large number of physical qubits to create a logical qubit. So we need both quality and quantity of qubits.

For a one in a thousand error rate, you probably need from 100,000 to one million qubits to run interesting applications.

*Building up from qubits, what does the quantum computing stack look like?*

Today's quantum computers use a range of technologies. IBM, Google, and Rigetti build devices based on superconducting circuits, while companies such as Infleqtion, IonQ, and Quantinuum use atoms or ions as their building blocks. But no matter which technology you use, you need a classical computer which orchestrates computations using these devices.

***"The system architecture and software stack can play an enormous role in terms of scaling computation, making it more programmable, and reducing its costs"***



---

***“All the abstractions that we’ve developed for classical computing could become very important for quantum computing”***

---

Physically, to perform quantum operations, if you are using atomic qubits you use a laser pulse – you shine a laser and that changes the state of the atom which we are using as a qubit – or, in the superconducting setting, you use a microwave signal, a telecom frequency signal to change the state of your qubit. This process of performing physical gates is typically done by a field-programmable gate array (FPGA). So you can think of a quantum computer as having two parts: quantum devices, based on either atoms or superconducting circuits, connected to an FPGA via a laser system or via a microwave-signal generation system.

Today, the programmer writes a very low-level quantum circuit which is equivalent to a quantum program. The compiler takes this quantum-circuit description and compiles it down to laser pulses or microwave pulses. Those pulses are sent to the FPGA, which plays those pulses to the qubits and you see the computation happening.

How do we measure qubit states? If you have atomic qubits, you typically take a picture of the qubit and if it’s emitting a photon, that means it’s in a certain state. In a superconducting setup, you have something similar – some kind of sensor to detect its state.

***Do things that exist in classical computing – such as the memory – have a place in tomorrow’s quantum world?***

That’s another open question, but what we’re seeing on quantum computers is that there is still the separation of memory and compute. They are becoming more like von Neumann architectures, and the reason is that memory, even on classical computers, is a massively dense system: you can store a lot of information, but you don’t require active control. That’s basically what’s happening with quantum computers now: people are realising that if we separate memory and compute, we can pack more physical qubits into a smaller area, and we don’t need as much control overhead.

For example, modern dynamic random-access memories (DRAMs) have trillions of transistors on a chip, but we still use 512 wires to control all these transistors. So the buses are quite small and we have this nice multiplexing scheme for our memories. Today that’s not possible with quantum computers: the qubits we use as both computer and memory elements need active control, and every qubit requires at least one wire, so it becomes very complicated. But with atomic qubits – where many qubits could be addressed with one laser pulse – we have this nice setup where input / output (I/O) complexity could be reduced. And there are lots of reasons why people want to separate memory and compute, one of them being that it would allow qubit scaling without blowing the FPGA budget (considering that one FPGA only controls five qubits). All the abstractions that we’ve developed for classical computing could become very important for quantum computing.

***What do you think is most challenging about building the quantum computing stack?***

I would say there are three main pain points in scaling quantum computers. We need systems that can run fast error correction, which involves decoding errors in real time. We need compilers that can find parallelism – currently, almost all of a quantum program is run sequentially. And we need the right data structures to reduce the number of quantum operations, which currently run into the billions. So just as in the classical stack, compilation, algorithms and programming languages need to come together and we need to co-design the system.

At Wisconsin, we’re really lucky to have physics labs which are actively building quantum computers. So we talk to them, understand their problems, and try to solve some of these problems by using architectural techniques.

Follow The Quantum Factor on LinkedIn:

 [linkedin.com/company/the-quantum-factor](https://www.linkedin.com/company/the-quantum-factor)



A spinoff of the photonics institute ICFO, Barcelona-based Quside exploits quantum phenomena to provide a veritable physical foundation for cryptography.

## Quside: Building verifiable trust infrastructure for modern cryptography

**COMPANY:** Quside

**MAIN BUSINESS:** Trust infrastructure for modern cryptography

**LOCATION:** Barcelona, Spain

**WEBSITE:** [quside.com](https://quside.com)

Modern digital systems depend on cryptography. It protects communications, identities, transactions, devices and critical infrastructure. Yet many of the foundations underneath that cryptography remain invisible and are largely assumed to work, rather than being verified.

Quside is building the infrastructure to make those foundations stronger, measurable and verifiable. Founded as a spin-off from ICFO, the Institute of Photonic Sciences, Quside began by solving one of cryptography's most fundamental dependencies: randomness. The company developed quantum technologies capable not only of generating high-quality entropy – which is used to quantify the amount of randomness or uncertainty in the state of a system – but also of measuring and monitoring the physical process behind it. This introduced a simple principle that now guides Quside's broader technology strategy: trust should be verified, not assumed.

Today, Quside is expanding that principle across the cryptographic stack. Its technology brings together quantum entropy generation, runtime assurance, hardware roots of trust, cryptographic execution and acceleration. The goal is to provide a verifiable physical foundation for cryptography across devices, networks, data centres and critical infrastructure.

Hardware remains central to this approach. Cryptography ultimately executes somewhere physical, keys must be generated and protected, algorithms must run, and systems need trusted foundations from which secure operations can begin. Quside develops this infrastructure across silicon, embedded modules, Peripheral Component Interconnect Express (PCIe) and FPGA Mezzanine Card (FMC) cards, cryptographic intellectual property (IP) and integrated hardware platforms, allowing its technology to operate from the semiconductor level through to large-scale infrastructure.

Quantum random number generation remains one of Quside's core technologies and an important part of this platform. Its QRNG technology generates entropy from measurable quantum phenomena and is designed for industrial deployment at scale. Quside's approach goes beyond generating random bits by providing visibility into the physical entropy source itself, including runtime monitoring and verification. This turns a traditionally hidden cryptographic dependency into something that can be observed and managed.

Quside also integrates its technology into existing cryptographic and computing environments, rather than requiring customers to build new systems around it. Its solutions support established software interfaces and ecosystems and are designed to work alongside hardware security modules (HSMs), public key infrastructure (PKI), security platforms, telecommunications infrastructure, cloud and data-centre systems, embedded devices and accelerated computing environments.

This industrial focus is backed by certification and deployment experience. Quside has developed entropy sources certified under NIST SP 800-90B, while products including Garnet™ and Nellite™ have achieved qualification under Spain's National Security Framework. Its technology is used and evaluated across demanding environments including telecommunications, defence, space, cybersecurity and advanced computing.

The transition to post-quantum cryptography makes this physical layer increasingly important. New algorithms will change how cryptography works, but they will still depend on secure entropy, trusted hardware, reliable execution and the ability to adapt as cryptographic requirements evolve. Quside is building infrastructure for that transition, combining its foundations in quantum entropy with a broader platform for cryptographic assurance and execution.

The company's evolution therefore goes beyond bringing quantum randomness from the laboratory into real-world systems. Quside is building verifiable trust infrastructure underneath modern cryptography, rooted in physics, implemented in hardware and designed to operate at scale.

Image courtesy of Quside. Photo: Raquel Puras



For quantum networks to be viable, the challenge of synchronization must be solved. In this opinion piece, Alexandra Pinto and Myrna James (Xairos Systems, Inc.) set out how Xairos, with teams in the US, the UK and the Netherlands, is addressing this challenge to enable quantum networks.

# Unblocking global quantum networks from geopolitical limitations

How entanglement and precise synchronized timing are foundational infrastructure for quantum networks

Over the last few decades, many companies and governments have transitioned from radio frequency (RF) to optical links for data transfer, as optical networks provide higher bandwidth communication links. They also provide protection from the jamming and spoofing to which RF signals are prone to and therefore photonics has become essential for upgrading security. Utilizing quantum technologies via quantum key distribution (QKD) has enabled a further level of encryption for data transfer. The next step is full-scale quantum networks, which use quantum phenomena such as entangled photons to distribute quantum information and connect quantum devices, enabling both security and accuracy at scale.

QKD provides one layer of security. It is superior to public key encryption (PKE) because QKD can reveal eavesdropping attempts, and if the transmitted single photon links are interrupted for any reason, both parties are aware of the intrusion through real-time efficient post-processing software.

Approaches to QKD can be divided into two categories: prepare-and-measure protocols such as BB84 (1984), and entanglement-based protocols such as E91 (1991). QKD is an early application of quantum networking capabilities but the no-cloning theorem, and thus the need of intermediate trusted nodes, has limited wide-scale adoption.

For future global communications and networks, additional security is required to be able to trust different devices, systems, and governments. Commercially, a missing piece for global trust were protocols for untrusted nodes which allowed for the use of uncharacterized or untrusted devices. These protocols, called Device-Independent QKD (DIQKD) or Measurement-Device-Independent QKD (MD-IQKD), were proven experimentally in 2013.

What is the most important next step on this roadmap toward global quantum networks? The critical element is timing synchronization and to make this timing synchronization secure, we use entanglement distribution.

Global timing synchronization is the foundation of all digital infrastructure. It directly affects telecommunications, high-performance computing, financial infrastructure, distributed sensing and defence systems. Current sensors and future quantum sensors, as well as current integrated circuits and future photonic integrated circuits are all part of this digital infrastructure. Satellites, precision network protocols and atomic clocks are used to guarantee global timing synchronization.

Quantum Time Transfer (QTT) is a novel way of using entangled photons for secure and accurate timing, invented and commercialized by Xairos. The Xairos protocol is the first to solve the primary engineering challenge of quantum networks: secure and accurate synchronization. It aligns classical control operations with the timing requirements of quantum processes, and coordinates preparation, transmission and measurement of quantum states across interconnected nodes.

High-accuracy timing is essential in quantum networks because quantum states are extremely sensitive to temporal misalignment. Also, picosecond synchronization is often required to ensure control pulses, photon arrival times, and to ensure that state-transfer operations preserve fidelity, interference visibility, and reliable entanglement distribution rates across the network.

QTT is a foundational enabling technology for any distributed system where large volumes of geographically distributed data require precise and highly accurate timing, security and resiliency. The added security layers create trust among sovereign nations, removing geopolitical limitations. Entanglement is the required gateway to trusted global quantum networks.

For further information, contact  myrna@xairos.com and  alexandra@xairos.com.





Based in Cologne, Germany, and spun out of RWTH Aachen University, Roofline builds scalable edge artificial intelligence (AI) deployment infrastructure, including a compiler, runtime, and benchmarking tools. In this article, Roofline Co-Founder and Chief Executive Jan Moritz Joseph explains how the company has provided native compiler support for mixture-of-experts (MoE) models.

# Bringing mixture-of-experts to the edge: Native MoE compiler support in IREE

Roofline's compiler is the core technology behind its broader deployment infrastructure. It's built on IREE (Intermediate Representation Execution Environment), an open-source MLIR-based framework that targets a wide range of hardware from a shared codebase. Our work extends that foundation with the model and hardware capabilities needed to support new AI architectures and hardware targets including central processing units (CPUs), graphics processing units (GPUs), and neural processing units (NPUs).

One of our most recent additions is native compiler support for mixture-of-experts (MoE) models, an architecture increasingly common in newly released LLMs. MoE introduces a compilation challenge that dense-model paths were not designed for: expert selection happens at runtime, so the amount of work assigned to each expert depends on the input.

## Faster inference, same memory footprint

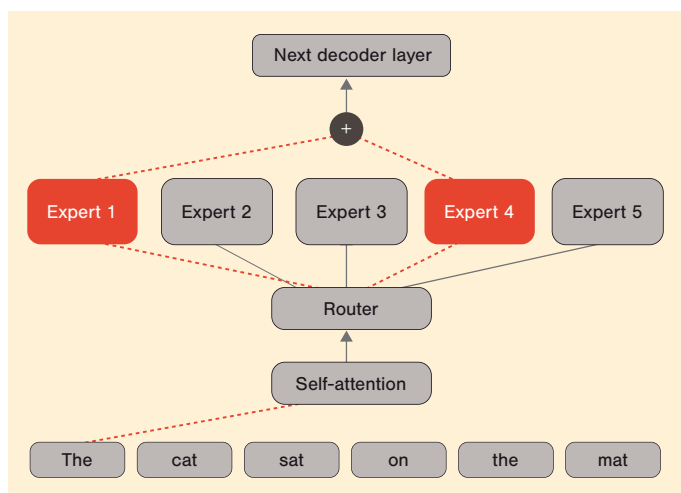
MoE changes what larger models cost to run. Instead of a single feed-forward network per transformer block, an MoE model uses many parallel expert networks, with a router selecting a

subset of experts for each token. Activating only a subset of the model per token can accelerate inference by an order of magnitude compared to a dense model of similar size. A model might have, for example, 30B total parameters but only 3B active for a given token; only those active experts need to be read from memory, so inference speed tracks active parameter count rather than total model size.

The full set of expert weights still has to fit in memory, even though only part of it is read per token. That has kept MoE mostly in the data centre. As edge devices gain more memory while bandwidth remains the binding constraint, MoE becomes relevant to automotive, robotics, and other embedded applications that need more capable models without a proportional rise in memory traffic.

## Compiling for a workload that changes at runtime

A dense model's feed-forward computation is the same for every token, so a compiler can lay it out and schedule it ahead of time. The same is not true for MoE models, where the workload is data-dependent: each expert can receive a different number of tokens depending on what the router decides.



*Token-level expert routing in a mixture-of-experts (MoE) layer*

Supporting MoE's runtime expert routing starts with grouped matrix multiplication, a matmul that takes a different number of tokens per expert rather than a fixed shape known ahead of time. That's what allows the compiler to represent MoE's sparse execution pattern, and it's what turns selective activation into an actual speedup rather than just a routing decision. It also works around a real limitation in frameworks like PyTorch, which don't handle tensor shapes that depend on the data rather than being fixed in advance.

Getting tokens to the right expert also depends on a set of supporting operators for selecting, counting, and re-ordering tokens by expert. Making all of these work correctly, not just the headline operator, is what makes end-to-end MoE execution possible.

***“The pace of AI model development means we cannot rebuild the deployment stack for every new architecture. Shared compiler infrastructure lets us solve these problems once and reuse that work across the edge AI ecosystem”***

Native compiler support gives MoE models a common execution path that can be reused across different models and hardware targets without per-model workarounds. A key part of that support in IREE is grouped matrix multiplication, which the compiler translates into lower-level executable operations and runs through the existing runtime without MoE-specific tuning at execution time.

## One compiler path, any MoE model

We validated this approach end to end on three open MoE models spanning a wide range of sizes:

- Granite 3.1 1B-A400M (1 billion total parameters, 400million active, 32 experts with top-8 routing)
- LFM2-8B-A1B from Liquid AI (8.3 billion total parameters, around 1.5 billion active, 32 experts with top-4 routing)
- Qwen3-REAP-15B-A3B (15 billion total parameters, 3 billion active, produced through REAP (Router-weighted Expert Activation Pruning), a method that removes low-impact experts without changing the active path)

Testing across three different model types demonstrates that the approach is generic: the same compiler support for MoE’s grouped-computation pattern handled all three without any extra work needed per model. This generality separates the approach from runtimes that rely on hand-written, per-architecture implementations, where adding a new model can take months of specialized work.

Because this support is implemented at the compiler level in IREE, it is reusable across MoE architectures and hardware targets. Our goal is to upstream these capabilities, so they become shared infrastructure available to the wider IREE community.

## From an x86 proof of concept to real edge hardware

Our initial milestone was functional: we compiled and ran Granite 3.1 1B-A400M, LFM2-8B-A1B, and Qwen3-REAP-15B-A3B end to end on an x86 CPU from AMD, with MoE’s dynamic execution pattern handled through compiler-level support rather than MoE-specific runtime changes.

We deliberately left performance optimization for later. Getting the models to run correctly came before making them fast. We focused on showing that the same compiler path could support multiple MoE architectures end to end on an AMD x86 CPU.

## Next: performance, then upstreaming to IREE

With functional MoE support demonstrated on an AMD x86 CPU, our next steps are to improve performance, extend the work to edge targets, and upstream the work to IREE.

- Optimize the current MoE compiler path for Arm Scalable Matrix Extension and Scalable Vector extension (SME/SVE) and RISC-V Vector Extension (RVV) CPUs, as well as Vulkan/SPIR-V GPUs. We have shown that the models can run; the next step is to make MoE execution fast on these targets.
- Contribute our MoE compiler support to IREE, extending the open compiler ecosystem that other IREE-based projects build on.
- Extend MoE support to NXP Neutron and Qualcomm Hexagon NPUs. These targets will still require backend-specific work, built on top of the foundation we have now.

The pace of AI model development means we cannot rebuild the deployment stack for every new architecture. MoE is one example of how quickly model architectures are changing, and why edge AI deployment needs shared infrastructure that can evolve with them, letting us solve problems once and reuse that work across the edge AI ecosystem. Building that support at the compiler level and contributing it to IREE makes each new model and hardware target easier to support than the last.

Further information is available at [roofig.ai](https://roofig.ai).



In this article, Thomas Regis (Program Director, Software République) introduces vision 4rescue, a platform which collects and analyses data to provide emergency response teams with actionable insights in real time. Developed by Software République, an open innovation ecosystem comprised by Atos, Dassault Systèmes, JCDecaux, Renault Group, STMicroelectronics and Thales, vision4rescue aims to transform emergency response and help save more lives.

# vision 4rescue: The technologies behind a new era of emergency response

When a wildfire spreads, a flood cuts off a neighbourhood, or an industrial accident unfolds, emergency services face the same challenge: making fast critical decisions with incomplete information. The problem is rarely the absence of data. On the contrary, responders are often confronted with an overwhelming volume of information coming from different systems, teams and communication channels.

This is exactly what inspired vision 4rescue, developed by Software République, an open innovation ecosystem dedicated to smart, secure and sustainable mobility. It brings together six leaders: Atos, Dassault Systèmes, JCDecaux, Renault Group, STMicroelectronics and Thales.

The ambition behind vision 4rescue is to turn this complexity into a coherent, real-time operational picture. This is made possible by the complementary expertise of each consortium partner, bringing together industrial know-how and technological capabilities to tackle complex and ambitious challenges collectively.

At the centre of the project lies a simple idea: create a digital ecosystem capable of collecting information from the field, analysing it instantly and delivering actionable insights to those coordinating the response. To achieve this, vision4rescue combines connected vehicles, drones, artificial intelligence, advanced sensors and resilient communication networks into a single platform designed to support first responders throughout a crisis with real-time analysis.

The process begins in the field with the combination of a specialized vehicle equipped with drones and a mobile command centre, all connected to a digital hub that can analyse information in real time. Data is collected by sensors embedded in the vehicle and drones, as well as by fixed urban infrastructure such as cameras and other connected systems.

Drones can be rapidly deployed from the vehicle to provide an aerial view of the surrounding area, helping emergency teams detect and assess specific threats; from the early signs of a forest fire to situations in areas that are difficult or dangerous to access.

At the same time, mobile and / or fixed cameras, LiDAR systems and environmental sensors collect information ranging from infrastructure conditions to potential hazards. But unlike traditional emergency management systems, the objective is not simply to gather data. The challenge is to transform that information into operational intelligence that enables actionable decisions while events are still unfolding.

This is where artificial intelligence and edge computing become critical. Rather than sending all data to distant servers for processing, part of the analysis occurs directly on connected devices and operational equipment. Images captured by drones can be processed in near real time from the deployed vehicle or in the backend of the emergency system to identify damaged infrastructure, detect anomalies or locate points of interest.



Photo credit: Agence CÉTADI Prod



Photo credit: Agence CÉTADI Prod

By analysing information closer to where it is generated, the system reduces delays and provides responders with faster, more relevant insights.

The true value of the platform, however, emerges when these different streams of information are combined. A drone image alone provides limited context. A geolocation feed or an environmental sensor reading tells only part of the story. By fusing multiple sources of data, vision 4rescue creates a shared operational picture that can be accessed by emergency teams and decision-makers simultaneously. Artificial intelligence helps prioritize the most critical information and support decision-making, allowing responders to focus on taking actions rather than managing data.

Importantly, this flow of information would be impossible without robust communications. During major emergencies, conventional networks can quickly become saturated or even unavailable. vision 4rescue therefore relies on a combination of communication technologies, including 5G, private networks, Wi-Fi, satellite links and vehicle-to-everything connectivity. Temporary communication bubbles and mesh-networking capabilities can maintain connectivity between responders, vehicles, drones and command centres even when existing infrastructure is compromised.

These technologies come together in a vehicle that functions as a mobile command centre capable of deploying drones, processing information and maintaining communications directly from the field. It acts as a bridge between data collection, AI analysis and operational coordination, bringing computing power as close as possible to where decisions are being made.

---

***“vision 4rescue offers a glimpse of how emergency response could evolve from a collection of disconnected tools to a fully connected, data-driven ecosystem”***

---

Another important dimension of vision 4rescue is its ability to look beyond the immediate emergency. By combining real-time information with virtual twins and simulation technologies, the platform can model the evolution of a crisis and evaluate different response scenarios. Emergency managers can therefore anticipate developments, assess risks and adapt their strategy before conditions deteriorate.

Ultimately, vision 4rescue is not defined by a particular drone, communication network or AI model. Its innovation lies in the way these technologies work together. Data collected on the ground is transmitted through resilient networks, processed by artificial intelligence, enriched by predictive simulations and transformed into actionable decisions. In doing so, vision 4rescue offers a glimpse of how emergency response could evolve in the coming years: from a collection of disconnected tools to a fully connected, data-driven ecosystem designed to help responders act faster, coordinate better and save more lives.

### About Software République

Software République is an open innovation ecosystem dedicated to smart, secure and sustainable mobility. It is made up of six members: Atos, Dassault Systèmes, JCDecaux, Renault Group, STMicroelectronics and Thales.

Software République develops joint projects focused on the mobility of the future through its unique model of horizontal collaboration. The ecosystem stands out for its innovative approach, combining established companies and start-ups from diverse backgrounds to bring to market products and services that address the new challenges of connected vehicles, smart cities and energy. These projects rely on the complementary expertise of its partners in semiconductors, data analytics, artificial intelligence, cybersecurity, connectivity and digital twins, as well as on the ambition to invent a new model of innovation that keeps people and the environment at the heart of its motivation.

For further information, visit [softwarerepublique.eu](https://softwarerepublique.eu)

In this article, the GreenShift consortium tells us how GreenShift offers a flexible master's and micro-credential pathway that builds advanced artificial intelligence (AI) and high-performance computing (HPC) skills for green mobility, embedded in real use cases across Europe.

# How GreenShift is upskilling the transport sector with advanced digital skills for the green transition

## GreenShift

Smarter Journeys. Cleaner Europe.

GreenShift is establishing the first joint European master's programme in artificial intelligence (AI) and high-performance computing (HPC) dedicated to the green transition of the transport and mobility sector. The programme supports startups, small and medium enterprises (SMEs), companies and public authorities in designing and deploying smarter, lower-carbon mobility solutions that contribute to Europe's net zero targets by 2050. The transport sector remains one of Europe's largest sources of greenhouse gas emissions, and GreenShift addresses a growing skills gap by equipping professionals and graduates with advanced digital capabilities. It aims to upskill more than 1,500 professionals, students and public-sector staff over the project's lifetime while also engaging larger corporations driving sector decarbonization, and building lasting capacity in the universities and industry partners that deliver the programme.

Offered as a nine-month full-time or 12-month part-time postgraduate degree (60 European Credit Transfer and Accumulation System (ECTS) credits), the programme is complemented by flexible upskilling pathways – micro-credentials, standalone modules and short courses – and hands-on formats such as bootcamps, hackathons and innovation challenges designed to give practitioners immediately applicable skills. The curriculum also includes a dedicated quantum computing module, reflecting the technology's growing relevance to transport optimization and simulation. To widen access, the programme includes dedicated scholarships for under-represented groups and is working towards EU-recognized joint academic accreditation alongside industry certification.



Members of the GreenShift consortium

The consortium brings together higher education institutions (HEIs), HPC centres, SMEs, industry players and training providers, reflecting the cross-disciplinary collaboration needed to align computational science with real-world decarbonization challenges. GreenShift has been awarded the European Commission's STEP Seal in recognition of its strategic importance to European technological sovereignty. The project is funded under the Digital Europe Programme (call DIGITAL-2024-ADVANCED-DIGITAL-07), with the European Health and Digital Executive Agency (HaDEA) as granting authority.



**CONTACT:** Horacio González-Vélez, GreenShift project coordinator, professor of computer systems and founding head of the Cloud Competency Centre, National College of Ireland

**PROJECT NAME:** GreenShift – EU MSc in AI & HPC for Green Digital Innovation in Transport

**START / END DATE:** 1 October 2025 – 30 September 2029 (48 months)

**KEY THEMES:** artificial intelligence, high-performance computing, green digital skills, decarbonization of transport and mobility

**PARTNERS AND ASSOCIATED PARTNERS:** **Ireland:** National College of Ireland (**coordinator**), Digital Technology Skills, FIT (Fasttrack into Information Technology), Matrix Internet; **Belgium:** EBN (European Business and Innovation Centre Network); **Croatia:** Hrvatski Telekom d.d. (Croatian Telecom Inc.); **Germany:** HLRS High Performance Computing Center Stuttgart; **Italy:** Università di Pisa, CINECA, Adecco Italia; **Netherlands:** ScoutinScience BV, SURF; **Poland:** Transition Technologies PSC S.A.; **Portugal:** NOVA Information Management School, Unicorn Factory Lisboa; **Romania:** University Politehnica of Bucharest

**BUDGET:** € 4.8M

[thegreenshift.eu](https://thegreenshift.eu)



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.



For this issue of *HiPEACinfo*, we have a special guest: Chipsaur. With a little help from his herder, Sophie Zweifel (ETH Zurich), Chipsaur helps us get to the bottom of how Swiss chip design – which some had thought in danger of extinction – has been given a new lease of life thanks to the SwissChips initiative.

## Chipsaur asks: How do you build a chip-design community in Switzerland?

These days, almost everyone has heard of chips. Some of us can even confidently state that the latest iPhone twelve-billion Max Pro ++ Ultra Fold MK2 contains an M-series chip or an A-series chip, before quietly praying nobody asks a follow-up question.

Chips sit inside our phones, cars, and computers. They also power medical equipment, communication networks, satellites, industrial systems, and energy grids. In short, modern life would grind to a rather immediate halt without them.

**Chipsaur: I would survive.**

You almost cried when your Wi-Fi stopped working for eleven minutes.

**Chipsaur: I don't appreciate you bringing up my trauma like that.**

Your head is literally a chip.

**Chipsaur: That is offensive.**

What are you doing here anyway?

**Chipsaur: I heard you were talking about SwissChips.**

I am the mascot. It seemed only right that I attend this little shindig.

Are you aware there are no snacks?

**Chipsaur: Dinosaurs have rights too, you know.**

Despite the importance of chips, most of us rarely ask who designs them or how somebody learns to design one in the first place. These questions lie at the heart of SwissChips, a national initiative co-led by the Swiss Federal Institute of Technology Zurich (ETH Zurich), Swiss Federal Institute of Technology in Lausanne (EPFL), and Swiss Center for Electronics and Microtechnology (CSEM) that is strengthening integrated-circuit design, research, education, and collaboration across Switzerland.

### Why does SwissChips matter?

Switzerland did not wake up one morning to discover that somebody had misplaced all its chip designers. The expertise was already there, spread across universities, universities of applied sciences, and research centres around the country. What was missing was a way to connect it all, to bring together the people, tools, and infrastructure needed to turn strong individual groups into a national chip-design community.

Working at the forefront of chip design also requires more than just a clever idea; researchers need specialized design software, manufacturing access, packaging and testing solutions, and most importantly: people who know how to use them.

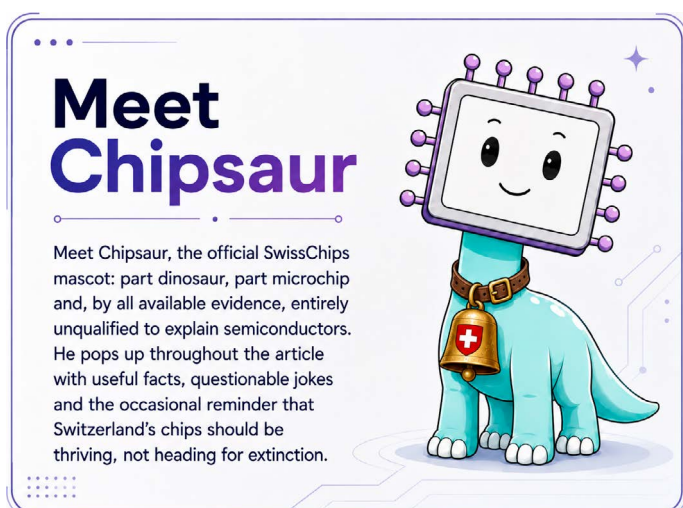
SwissChips was launched in 2024 as a three-year initiative funded by the Swiss State Secretariat for Education, Research, and Innovation. Its purpose was not simply to finance separate projects, but to connect the people, infrastructure, and knowledge, protect national capabilities, and open advanced chip design to more researchers and students.

### So how do you build a chip-design community?

Apparently, you begin with a great many meetings.

**Chipsaur: Have you never heard of email?**

The meetings quickly became a place where the different teams could swap hard-won lessons. One group might have found a workable packaging solution, another a better way to test a new design, and another had already wrestled with the particular



problem now causing trouble elsewhere. In chip design, that kind of shared experience can save a great deal of time, money, and frustration.

Design support teams are now working together to address shared challenges in infrastructure, design flows and the engineering processes involved in integrated circuit (IC) design. Open-source design flows have also been introduced into research projects and teaching, lowering some of the barriers to a field traditionally associated with costly commercial software.

### **Chipsaur: Like SimCity?**

Not quite. Though for the price of some design tools, you could build a sizeable REAL metropolis.

## From the classroom to real silicon

Tools alone do not create chip designers. At some point, people have to sit down and design something.

At ETH Zurich, two very large-scale integration (VLSI) courses are giving students hands-on experience across different stages of integrated-circuit design:

- In the 2025 edition of the VLSI 5 course, 19 students designed 218 cells for an open standard-cell library based on a 130nm open-source semiconductor process – the cells will be released to the public very soon.
- In the 2026 edition of the VLSI 2 course, this standard-cell library was then used by 75 students, who developed 37 integrated-circuit designs. Eight of these designs were selected for tape-out, taking the students' work from the computer screen to fabrication in real silicon.

CSEM also organized specialist training in layout and chip finishing using a 22nm technology, with more than 50 participants at the ETH Zurich edition. The 2026 EFCL Winter School then brought 100 participants together for hands-on work in open-source digital IC design and field-programmable gate array (FPGA)-based systems, while over 200 attended the keynote programme. Finally, a Tiny Tapeout workshop gave participants the chance to create small digital designs using open-source tools.



*EPFL TechXperience 2025 workshop as part of the Semaines d'études (Study Weeks) programme run by Science et Jeunesse*

### **Chipsaur: Field-programmable what now?**

It's hardware you can reconfigure on the fly. Sort of like software-defined hardware, if you will.

### **Chipsaur: Also, is tape-out some kind of chewing gum?**

It's the point at which a completed design is sent for manufacturing.

### **Chipsaur: So the peaceful end of the process?**

Almost precisely the opposite. Once a design is sent for fabrication, the work is far from over. The finished chip still has to be brought to life, measured, tested, and verified – and any unexpected behaviours need to be carefully traced back through the design.

At EPFL's TechXperience programme, six high-school students also worked with doctoral researchers to design circuits that were then prepared for fabrication. Experiences like these make chip design tangible, and harder to dismiss as mysterious work carried out by wizards in windowless laboratories.

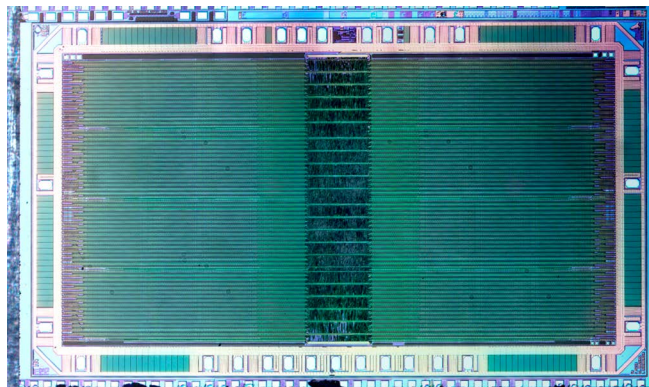
### **Chipsaur: You're telling me chip designers are not wizards?**

Let's not rule anything out.

## Helping designs reach the foundry

Designing a chip is only half the challenge. Manufacturing even a small batch can cost tens or hundreds of thousands of Swiss francs, putting fabrication beyond the reach of many university groups and student projects.

SwissChips therefore created a national support programme for universities and universities of applied sciences. In its first two years, 16 external application-specific integrated circuit (ASIC) designs were approved, and more than CHF 500,000 was allocated towards fabrication. A fast-track procedure also helps smaller and student-led projects move forward more quickly.



*Zervreilahorn is a SwissChips-funded ASIC developed at OST Rapperswil for energy-efficient deep neural network processing. Designed by Gian-Luca Brazzol and Flavio Peter, with support from the IMES ASIC Design Group and supervision by Paul Zbinden and Lars Kamm*

### Chipsaur: Application-specific what now?

ASICs are chips that are tailored to a specific task. They're more efficient than general-purpose processors, but their application is also more limited. Sort of like a dedicated factory machine as opposed to a Swiss Army knife.

The supported designs range from neuronal-culture interfaces and biomedical sensors to optical communications, quantum technologies and brain-inspired computing. The chip Zervreilahorn, for example, was designed by two master's students at OST (Eastern Switzerland University of Applied Sciences) and performs analogue matrix multiplication.

In SwissChips' second year alone, the three institutions taped out a combined 51 ASICs, with a further 11 coming from other Swiss universities and universities of applied sciences with support from the initiative. More than half had already been successfully tested at the time of reporting.

### Chipsaur: Sixty-two chips? That will never fit on a sticker.

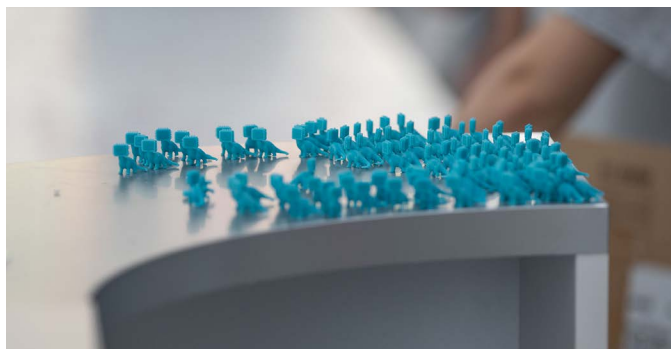
## Bringing the community together

A community also needs places where people can share their work, meet collaborators and discover that somebody three train stops away has already solved the problem keeping them awake at night.

The first SwissChips Annual Event brought more than 200 participants to Zurich's Landesmuseum. The second welcomed over 270 people to the SwissTech Convention Center at EPFL. Across the two events, almost 100 posters gave students and researchers a national platform to connect with colleagues from across academia, universities of applied sciences, and industry.



*The SwissChips Event 2026 was held at the SwissTech Convention Center (STCC) at Lausanne with over 270 participants*



*Spreading the joy of chip design across Switzerland one tiny chipsaur at a time*

SwissChips has also supported winter schools, technical workshops, researcher spotlights, and video interviews. Then, of course, there is Chipsaur Asks.

### Chipsaur: Finally, the important bit.

The series explains chips, semiconductors, transistors and wafers in language that does not require three engineering degrees and a cold compress afterwards. Together with the SwissChips website, LinkedIn channel, and outreach events, it has taken the initiative to a much wider audience. The LinkedIn community grew to more than 3,600 followers in its first year and is now approaching 5,000.

## More than the sum of its chips

Building a chip-design community is not simply about producing more chips. It means giving high-school students their first glimpse of the field, helping university students fabricate their first designs, enabling researchers to share expertise, and allowing institutions across Switzerland to work together rather than in isolation.

SwissChips may have begun as a three-year initiative, but the infrastructure, experience, and relationships it has created are intended to last far longer.

### Chipsaur: So Swiss chips are no longer at risk of extinction?

Not if we have anything to say about it.

**Chipsaur: Excellent. I did not survive the Cretaceous for Switzerland to run out of chips.**

Want more from Chipsaur? Check out the fantastic 'Chipsaur Asks' series on the SwissChips website: [bit.ly/SwissChips\\_Chipsaur\\_Ask](https://bit.ly/SwissChips_Chipsaur_Ask) And don't forget to follow the SwissChips LinkedIn page for updates:

[in linkedin.com/company/swisschips](https://www.linkedin.com/company/swisschips)



Participating in the HiPEAC Student Challenge introduced Maria Zerva to Barcelona Supercomputing Center–Centro Nacional de Supercomputación (BSC) – so when an opportunity came up at BSC, she jumped at it. We caught up with Maria at ACACES 2026 to find out more.

# From HiPEAC Student Challenge to dream job

Growing up, Maria was always interested in STEM subjects. ‘I had a big love for maths – my father was a mathematician, and I participated in competitions when I was at school,’ she explains. At 18, she was accepted into the School of Electrical and Computer Engineering at the National Technical University of Athens. ‘That was a bit of a shock,’ she laughs, ‘because suddenly I was just one of a large number of students with high grades.’

Over the course of the five-year integrated master’s degree, she learned a huge amount – including programming for the first time – and made close friends. During their last semester, they found out about the HiPEAC Student Challenge, due to take place at HiPEAC 2025 in Barcelona. ‘We saw it as a great opportunity to keep working together after our course finished,’ she smiles.

The hardest part was finding a topic. ‘We had an algorithms guy, an AI guy, people working on hardware, software, embedded systems,’ says Maria. ‘In the end we came up with something where everyone could contribute: an indoor navigation system. The hardware specialists programmed beacons, trying to make them as energy efficient as possible. Algorithms calculated the position of the device based on the beacons’ signals. AI models

were used to analyse the whole building and identify different features – doors, windows, walls, etc. – and to find the shortest path. Lastly, the software people created an interface for the app developers, but also for the clients and for the mobile application.’

If this sounds like a lot of work, it was, says Maria – but she adds that it was highly rewarding. ‘Everyone had an interest in working on this project. It was an adventure. We spent a whole day contacting embedded systems companies just to get hardware – and eventually we got lucky with Nordic Semiconductors, who kindly sent us some beacons and microcontrollers.’

The hard work paid off when Maria’s team were selected to present their project at HiPEAC in January 2025. ‘HiPEAC was my first conference, and, looking back, it was a nice place to be as a student for the first time: with no obligations to present apart from the challenge, we could just enjoy the presentations and the networking.’

As well as presenting their project at the student challenge session, the team took part in the STEM Student Day, including an interactive tour of the industry exhibition. ‘We won the STEM Day competition, and one of the prizes was a guided tour of the supercomputer at BSC,’ says Maria. ‘It was the first time I’d seen a supercomputer, and I found it inspiring.’

After finishing her thesis, Maria decided to explore job opportunities for the next year while she planned applications for her PhD. ‘I saw an open position for a research engineer post working on register-transfer level (RTL) for the vector processing unit (VPU) being developed at BSC. I was delighted when I was offered the post; I moved to Barcelona and found myself surrounded by a critical mass of researchers working on computer architecture.’ Now several months into her job at BSC, she is very happy: ‘I think there was a lot of defining moments involved in how I got here – and I’m really glad for the opportunity.’



*Maria presenting during the Student Challenge at HiPEAC 2025 and showing her poster at ACACES 2026*



In this edition of our regular series showcasing early career researchers' work, we learn how Georgios Armeniakos developed novel, cross-layer design methodologies to make machine learning feasible on ultra-resource-constrained devices.

## Three-minute thesis

**NAME:** Georgios Armeniakos

**RESEARCH CENTRES:** National Technical University of Athens (NTUA)

**SUPERVISORS:** Dimitrios Soudris (NTUA)

**THESIS TITLE:** Design automation of approximate machine learning circuits for ultra resource-constrained environments

Machine learning (ML) is increasingly moving away from large cloud servers and towards the devices that collect data. This shift enables faster decisions, greater privacy, and lower communication costs. Yet at the extreme edge, inside disposable sensors, smart packaging, wearable patches, or battery-free devices, the usual silicon solution can become too expensive, too rigid, or physically unsuitable.

Printed and flexible electronics (PFEs) offer a different path: circuits that can be manufactured on polymers, paper or other flexible substrates at very low cost. The problem is that printed transistors are large, slow, power hungry, and difficult to integrate densely, making even simple machine-learning computations disproportionately expensive.

This PhD started from a basic question: can meaningful machine learning be brought to hardware that is orders of magnitude more constrained than conventional electronics? Simply shrinking existing processors is not enough. The problem is not only size; the entire design philosophy must change. Indeed, even basic ML building blocks can be dramatically larger, slower and less energy-efficient in printed technologies than in silicon.

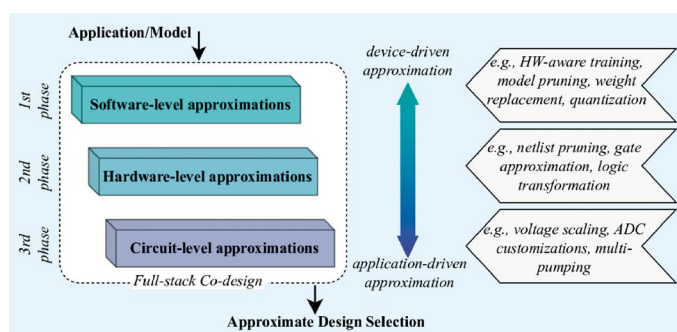
This research therefore treats accuracy, hardware and technology as parts of the same optimization problem. Instead of first

designing an accurate ML model and then asking whether it fits the hardware, we developed automated cross-layer methodologies in which the algorithm, architecture and circuit are optimized together. Approximation becomes a design tool: coefficients can be replaced with hardware-friendly values, unnecessary logic can be pruned, arithmetic can be simplified, and voltage can be reduced, accepting tiny, controlled inaccuracies where they provide large physical savings.

The originality of this work lies in systematically bringing approximate computing into bespoke printed ML circuits across multiple abstraction levels. Because these circuits are customized for a specific model, their parameters can effectively become part of the hardware itself – more like a tailor-made suit than a general-purpose processor. The dissertation extends this idea further through printing-aware neural-network training and on-sensor co-design, where even the analogue-to-digital converters are customized together with the classifier. The resulting approaches reduce resource demands by more than an order of magnitude while keeping accuracy loss below 1%.

The practical vision is intelligence embedded directly into everyday objects: a food package that recognizes spoilage, a flexible health patch that interprets sensor signals locally without continuously transmitting private data, or an environmental tag that classifies conditions using harvested energy by ultra-low power requirements. These are applications where electronics may need to be cheap enough to discard, flexible enough to conform to the body, and efficient enough to operate almost autonomously.

Finally, the ideas are not limited to printed electronics. The same approximate, cross-layer philosophy is also extended to a programmable mixed-precision RISC-V accelerator, demonstrating that it can improve efficiency while retaining programmability. Overall, this research provides a practical design methodology for bringing machine learning to hardware platforms that were previously considered too constrained to support meaningful intelligence.



*Approximate computing stack hierarchy applied*



# HiPEAC conference

18-20 January 2027

Glasgow

compilation in the LLM era · cognitive fog · AI meets safety-critical systems ·  
sustainable computing systems · next-generation interconnects  
for AI and HPC · open hardware · post-quantum cryptography ·  
self-aware neuromorphic architectures · wafer-scale computing ·  
Student Challenge · STEM Day · industry exhibition ... and much more

[hipeac.net/conference](https://hipeac.net/conference)

#HiPEAC27



Funded by  
the European Union