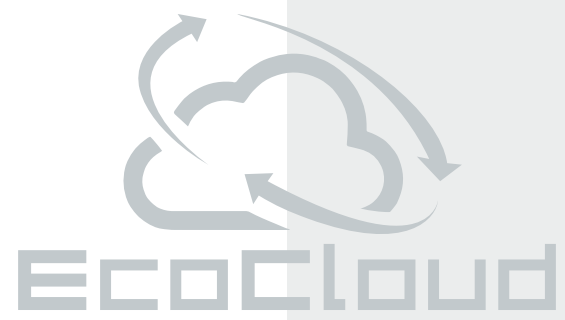


EPFL



2026





CONTENTS

Welcome	4
Research activities	5
Faculty members	7
Industry affiliates	10
Current EcoCloud affiliates	13
On power for computing and computing for power	14
The EcoCloud experimental facility	16
GemaTEG: Taking thermal management to the next level	18
Do we really need big data centers for AI?	20
Midgard: Mission accomplished, but the story goes on	22
SwissChips: an ongoing initiative	24
Chips from a higher dimension: the next generation of generative AI computing	26
Flexible data centers could support Europe's net-zero transition	27
SwissChips Annual Meeting: A place to share advances in silicon design	28
Visiting Speakers of 2026	29
Contact	30



WELCOME TO ECOCLOUD



Welcome to EcoCloud, an academic center unique of its kind, promoting research and development of hardware and software computing technologies as key enablers for a sustainable future.

Our mission revolves around design and management of cloud and edge computing to achieve effective services with reduced environmental impact. Our projects include, but are not limited to, digital twins for smart cities, sustainable artificial intelligence technologies, and energy-aware scientific computing. EcoCloud oversees the research at EPFL on the SwissChips federal project. The centre is also working towards the definition and creation of a new circular economy of large IT infrastructures and datacentres to minimize carbon footprint and long-term environmental damage.

Our strategy is to create and grow a community around EPFL and its key industrial players to support environmentally sustainable information technology. Our ambitions and success rely on synergy, collaboration and support by private industry. With our strong emphasis on industry teamwork and technology transfer, the Industry Affiliates Program (IAP see page 10) aims to build long-term partnerships founded on research collaborations, large-scale University-Industry partnered research grants and PhD student dissemination activities with industry, internships, and fellows.

Overall, the EcoCloud centre aims to be a landmark in digital transformation for a sustainable society.



RESEARCH ACTIVITIES OF THE CENTER

Prof. Giovanni De Micheli has been Director of EcoCloud since 2024.

"EcoCloud is a research centre open to academics in all schools of EPFL, including professors, assistants and students. The centre cooperates with other EPFL's centres, such as the Artificial Intelligence (AI) centre. Indeed, "the benefits that AI will bring to science, the economy and the society will be mitigated if not combined with new hardware and software designs that are energy efficient," says Prof. Giovanni De Micheli, a pioneer in low-energy consumption for electronic circuits.

For this reason, EcoCloud is the focal point of EPFL's research on the federal SwissChips project in a bid to boost the Swiss chip manufacturing industry (see page 26) — in a collaboration with EPFL, ETHZ and CSEM.

EcoCloud is also a partner in the Heating Bits project, a multi-disciplinary activity that aims at increasing the efficiency of computation by harvesting energy from the heat generated by silicon chips. We are also deeply involved with the UrbanTwin project, searching for planning and operation solutions for sustainable smart cities and transportation systems.

Moreover, we are involved in several other research projects, ranging from low-energy consumption computational astronomy to the design and realization of accelerators for computation.

EcoCloud interacts with scientists, technologies and managers of several companies, including partners of the Industry Affiliates Program (IAP): Eaton, GemaTEG, Hewlett Packard, Huawei, IBM, IMO Ventures, Intel and Oracle.

A research facility for sustainable computing

We have assembled an experimental facility dedicated to multi-disciplinary research on sustainable computing at EPFL (see page 16). Over the last year, we have expanded the monitoring capabilities of our systems to include accurate measurements of performance, power, energy and temperature, among others. In particular, EcoCloud has collaborated with the SCITAS scientific computing platform in the efficiency characterization of their GPU cluster, named Kuma. The fruit of our labors: in 2024 Kuma broke into the Green500 list of supercomputers reaching, with an energy efficiency of 55 GFLOP/J, the 23rd position in the world.

Best practices for IT infrastructure

As the engine of our digital world, IT must embrace its responsibility to support economic progress and societal development without compromising the future of our planet. In that line, EcoCloud has collaborated with the EPFL Sustainability Team and SCITAS to define a methodology for the characterization of the CO₂ impact of IT services. This methodology will allow the different actors in an organization - IT providers and final service users - to assess the CO₂ emissions associated to the use of an IT service over its complete lifespan.

The dissemination of best practices for sustainable IT infrastructure is another of our core missions. EcoCloud has collaborated with the LEARN center in the design of "Utop'IA," a graphic novel to understand the environmental impact of AI. We have also participated in numerous

RESEARCH ACTIVITIES OF THE CENTER

communication activities such as conferences, talks and scientific publications.

Continuing collaboration on research projects

EcoCloud has continued promoting and collaborating in large scale projects. UrbanTwin is embodying an actionable digital model of the city of Lausanne, with the goal of producing a tool that will allow policy makers to evaluate the impact of environmental policies before actually undertaking them, and under different climate scenarios.

Heating Bits strives to reduce the CO₂ impact of EPFL data centers and is currently preparing to test the prototypes of the first new-generation cold plates. Work is progressing on the organic Rankine cycle (ORC), that will convert part of the energy recovered as heat in the cooling of servers directly back into electricity. Other important works on efficient distribution of direct current over the data center, workload management, and centralized planning and control of resources are also approaching the prototyping phase.

EcoCloud is the coordinator of the SwissChips project in EPFL. SwissChips, funded by SERI, has the goal of ensuring that new generations of chip designers are educated in Switzerland to ensure its digital sovereignty.

Finally, SEAMS, funded by SNSF and the French ANR, is making progress towards new ways of conducting large scale scientific computing with much larger energy efficiency - which has particular societal impact since some large scale experiments take place in areas where the need for new scientific discoveries must be balanced with the energy needs of the local population.

The key pillars of EcoCloud activity

<< IT infrastructure as enabler for a sustainable society

- Energy-constrained and sustainable deep learning
- Sustainable smart cities and transportation systems
- Computational and data storage sustainability
- Energy-constrained trustworthy systems

<< Sustainable IT infrastructure

- Minimal-energy computing and storage cloud platforms
- Sustainable use of renewable energy in IT infrastructures

<< Dissemination of best practices for IT infrastructure in a sustainable society

<< Preparation of courses and focused programs on sustainable computing, for IT professionals

<< Annual EcoCloud event on sustainable computing trends and forward-looking research

EcoCloud Staff



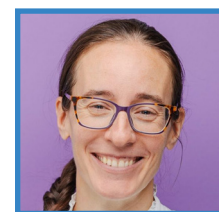
Valérie Locca
Administrator



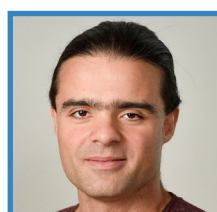
John Maxwell
Media Specialist



Dr. Xavier Ouvrard
Systems Specialist



Julia Paolini
Research Engineer



Dr. Miguel Peón-Quirós
Research Scientist



Alex Widderski
Graphic Designer

FACULTY MEMBERS

AND LABS - IN ALPHABETICAL ORDER



Anastasia Ailamaki

*Data-Intensive Applications
and Systems Laboratory*

Enabling discoveries in scientific domains through automating physical database design, revolutionizing exploration algorithms in very large data repositories



Alexandre Alahi

*Visual Intelligence
for Transportation*

Socially aware AI applying computer vision, deep learning and human-robot interaction to transportation applications



David Atienza

*Embedded
Systems Laboratory*

Efficient machine-learning based resource management in servers and data centers. Low-power design of edge AI and heterogeneous server architectures



Antoine Bosselut

*Natural Language
Processing Lab*

Natural language processing, machine learning, artificial intelligence



Thomas Bourgeat

*Verification and Computer
Architecture Laboratory*

We leverage the power of formal methods and high-level hardware programming languages to ensure the correctness and the security of tomorrow's computers



Maria Brbic

*Machine Learning for
Biomedicine Lab*

Developing machine learning methods that solve real-world data challenges, paving the way for new biomedical discoveries



Andreas Burg

*Telecommunications
Circuits Laboratory*

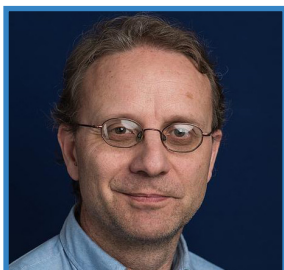
Design of technology systems, prototypes and demonstrators for the development of robust, reliable and energy efficient systems



Volkan Cevher

*Laboratory for Information and
Inference Systems*

Robust machine learning and optimization, reinforcement learning, game theory, and deep learning



Edoardo Charbon

*Advanced Quantum
Architecture Lab*

3-D vision, LiDAR, FLIM, FCS, NIROT, super-resolution microscopy, time-resolved Raman spectroscopy, and cryo-CMOS circuits and systems for quantum computing



Drazen Dujic

*Power Electronics
Laboratory*

Ensuring reliable, compact, and efficient power electronics-based power supplies for data centers. From the power grid to the chip



Babak Falsafi

*Parallel Systems
Architecture Laboratory*

Computer architecture, vertically integrated data center systems, post-Moore server design



Olga Fink

*Intelligent Maintenance and
Operations Systems Laboratory*

Development of intelligent algorithms for complex infrastructures and industrial systems. Deep learning and hybrid algorithms for intelligent maintenance systems

FACULTY MEMBERS AND LABS



Pascal Frossard

*Signal Processing
Laboratory*

Computer vision, medical imaging,
network machine learning, robust
machine learning, deep learning



Rachid Guerraoui

*Distributed
Computing Lab*

Distributed machine learning with
Byzantine resilience and privacy.
Distributed algorithms for new
technologies: RDMA and NVRAM



Josie Hughes

*Computational Robot Design
and Fabrication*

New fabrication and
computational design tools for the
development of robots with new
and exciting capabilities, and to
better understand the physical
world



Paolo Ienne

*Processor
Architecture Laboratory*

Computer and processor
architecture, FPGAs and
reconfigurable computing,
electronic design automation,
computer arithmetic



Martin Jaggi

*Machine Learning
and Optimization Laboratory*

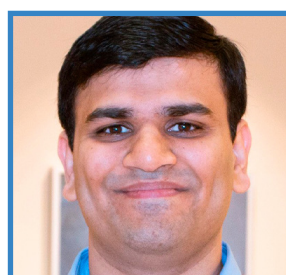
Machine learning, optimization
algorithms and text understanding,
as well as several application
domains



Colin Jones

*Automatic
Control Laboratory*

Theory and practice of
optimization-based, or model
predictive control with a particular
emphasis on problems arising from
renewable energy challenges



Sanidhya Kashyap

*Robust Scalable
Systems Software Lab*

Robust and high-performance
software for heterogeneous
hardware: concurrency, scheduling,
networks, analytics and fuzzing



Anne-Marie Kermarrec

*Scalable Computing
Systems Laboratory*

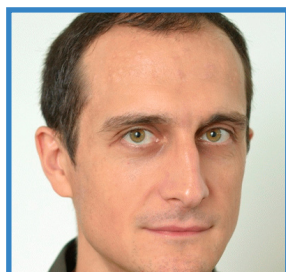
Large-scale distributed systems,
failure resilience, performance
and privacy-preservation, frugal
distributed learning systems



Jean-Paul Kneib

*Laboratory
of Astrophysics*

Reliable transport and precise
integration of a flow of 707
Petabytes per year of data from
large arrays of radiotelescopes



Viktor Kuncak

*Lab for Automated
Reasoning and Analysis*

Precise automated reasoning
techniques: tools, algorithms and
languages, for the construction of
reliable computer systems



Zhengmao Lu

*Energy Transport
Advances Laboratory*

Towards a deeper understanding
of phase change phenomena,
creating sustainable energy and
water technologies by optimizing
interfacial transport



Gabriele Manoli

*Laboratory of Urban and
Environmental Systems*

Analysis and conceptualization of
complex urban and environmental
dynamics, to guide the design
of greener and more sustainable
territories



François Maréchal

*Industrial Process and Energy
Systems Engineering*

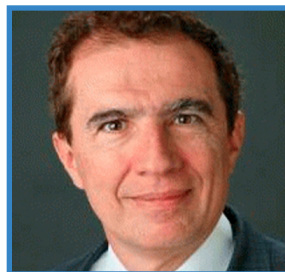
Process and energy system
engineering for efficient use
and reuse of energy, efficient
energy conversion, integration of
renewables and complex system
integration

FACULTY MEMBERS AND LABS



Elisa Matioli
POWERlab

Microchannel liquid cooling of data center components, ultra-efficient purpose-built cooling solutions



Giovanni De Micheli
Laboratory of Integrated Systems

Modelling of hardware with dedicated languages, co-design of software and hardware, system-level optimization with efficient performance, energy consumption and yield



Christophe Moser
Laboratory of Applied Photonics Devices

Non-linear transformation in fiber optics to simplify machine learning tasks



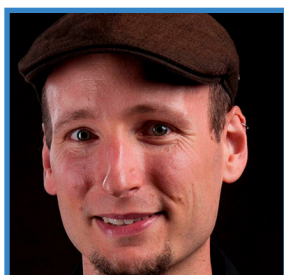
Martin Odersky
Programming Methods Laboratory

The design and implementation of Scala, to achieve a fusion of object-oriented and functional programming, compatible with platforms such as Java and .NET



Mario Paolone
Distributed Electrical Systems Laboratory

Developing smart grid concept solutions to efficiently deliver sustainable, economic and secure electricity supply



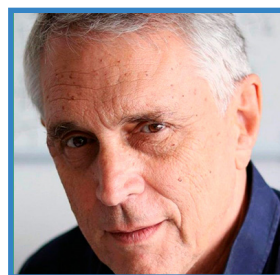
Mathias Payer
HexHive Laboratory

Software testing to discover security bugs. Sanitization for memory, type, and API violations. Fuzzing of complex code to trigger bugs



Clément Pit-Claudel
Systems and Formalisms Lab

Can we leverage compilers, languages, and proofs to build more robust, more efficient, and more trustworthy systems?



Demetri Psaltis
Optics Laboratory

Optical systems such as spatiotemporal nonlinearities in multimode optical fibers, used as neuromorphic neural networks



Jürg Schiffmann
Laboratory for Applied Mechanical Design

Design and experimental investigation of small scale turbomachinery for decentralized energy conversion



Mirjana Stojilovic
Parallel Systems Architecture Laboratory

Electronic design automation, reconfigurable computing, electromagnetic-compatibility and signal-integrity issues, hardware security



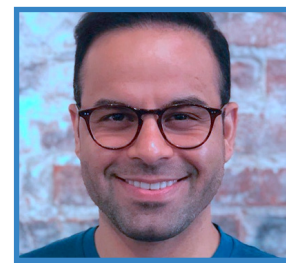
Giulia Tagliabue
Laboratory of Nanoscience for Energy Technologies

Exploring innovative mechanisms and concepts for light harvesting, conversion and storage by exploiting non-equilibrium (ballistic) transport processes and by leveraging strong light/solid/liquid coupling at the nanoscale



Anastasios Vassilopoulos
Group of Composite Mechanics

Using new composite materials and structural concepts to bring innovation to structural engineering. Renewable energy, wind energy conservation and rethinking manufacturing processes



Amir Zamir
Visual Intelligence and Learning Lab

The development of computer vision models that can function as part of larger intelligent systems

INDUSTRY AFFILIATES

MISSION

The EcoCloud Industry Affiliate Program (IAP) offers companies a unique opportunity to collaborate with EPFL faculty, students and researchers. Affiliates are given unparalleled access to new technologies and ideas as they move from laboratory to marketplace. The ideal platform for communication and discovery between the research and corporate communities, the program catalyzes collaborative research, customizes educational programs and facilitates graduate recruiting.

The EcoCloud IAP was created to enable connections and strengthen collaborations between EcoCloud and industry. While there are numerous benefits to joining the IAP and becoming an affiliate, the advantages boil down to three key reasons.

<< To gain early awareness of the latest research

Through meetings, visits and online resources, EcoCloud Industry Affiliates get to preview the latest research findings from across our labs before they are published.

<< To explore potential research collaborations and sponsorships

Companies can get much more by sponsoring EcoCloud research directly. Becoming an affiliate allows you to see how we work and what we do, giving you the insights you need to identify research partnerships.

<< To recruit EcoCloud students

Our students are one of our most valuable assets. They can add substantial value to your company as interns or employees. Our Annual Event is an excellent chance to meet our students, many of whom present their work in the posters exhibition. EcoCloud can also host talks geared towards students, and distribute job and internship announcements.



INDUSTRY AFFILIATES

MEMBER BENEFITS

The EcoCloud IAP is a corporate membership program whereby companies pay an annual membership fee in return for facilitated access to the research programs, the researchers and the graduate students, offering the ability to capitalize on the unique, dynamic trans-disciplinary innovation culture at EPFL. At the heart of EcoCloud's IAP activities are entrepreneurial startups emerging from our research, executive education which leads to industry collaboration and technology transfer.

Specific benefits of becoming an affiliate include:

<< Annual Event

An annual conference exclusively designed for our existing and prospective Industry Affiliates and research colleagues to showcase the activities of the center. The event's program includes activities such as keynotes by prestigious guests; presentations by EcoCloud faculty and researchers on the latest research results organized around chosen research themes; student poster sessions; lab tours; demonstrations; discussions on grand challenges, applications and technology roadmaps; and opportunities to meet and network with EcoCloud researchers, students and colleagues.

<< Research monitoring

Throughout the year, EcoCloud will enable affiliates to remain engaged by means of newsletters, seminars, talks and virtual meetings. When feasible, these events will be broadcast live for our IAP member companies, so participants can join in remotely to hear from different thought leaders at EcoCloud and keep informed about the latest research. Affiliates also have access to the comprehensive information about research outcomes, EcoCloud events, video and publication archives, and other research outputs.

<< Graduate student recruiting

EcoCloud will organize events (including the Annual Event), during which affiliates have access to soon-to-be graduating students to facilitate recruiting. We work with our affiliates to facilitate recruiting activities throughout the year, including advertising job and internship announcements, hosting talks and seminars, and other student-targeted networking events.

<< Occasional visits

Affiliates may arrange visits to EcoCloud's affiliated laboratories and the experimental facility created in the new data center of EPFL. Visits enable previews of EcoCloud's research programs and results, as well as demonstrations of emerging technologies. We work with our Affiliates to identify appropriate EcoCloud researchers, ongoing projects and potential opportunities for collaboration.

<< Joint research projects

Member companies have opportunities to engage with EcoCloud in research projects and collaborations into deployable technology. These include, but are not limited to: the opportunity to contribute and formally participate in EcoCloud research projects, customization of educational programs and the opportunity to develop and sponsor structured research programs.

<< Advertising

Our affiliates have the opportunity to promote their company's brand on EcoCloud's website and reports.

INDUSTRY AFFILIATES

MEMBER BENEFITS

<< Outreach and Executive/Continuing Education

EcoCloud develops and hosts outreach programs and workshops. Past events included a summer workshop on cooling technologies, a winter school on data-centric systems, and co-hosting the 3D silicon integration conference. Moreover, EcoCloud is eager to develop special executive courses, as well as continuing education courses, to address the specific needs of our affiliates.

<< Technical Advisory Board

EcoCloud Affiliates designate a technical staff to the EcoCloud Technical Advisory Board which meets once a year to discuss grand challenges, research and industrial trends, and EcoCloud research direction.

<< Visiting Scholars and Fellows Program

The EcoCloud Visiting Scholar and Fellows Program stimulates and supports our research by engaging promising scholars and practitioners in order to foster exchange. Each year, a number of distinguished academics (Visiting Scholars) and junior faculty and students (Fellows) will be selected on the basis of their qualifications, the quality of their research plans, and the relevance to both EcoCloud's mission and targeted research objectives. EcoCloud's Visiting Scholars and Fellows will work on projects that offer joint collaborative opportunities.

CURRENT ECOCLOUD AFFILIATES

IN ALPHABETICAL ORDER



IBM **Research** Europe



ORACLE

As observer:





ON POWER FOR COMPUTING AND COMPUTING FOR POWER

Prof. Giovanni De Micheli, EcoCloud Director

"It's a UNIX system, I know this" screamed Lex in the epic finale of Spielberg's Jurassic Park. Then, with few command-line clicks she saved her family, friends and possibly the whole of humanity from a miscarriage of technology aimed to recreate dinosaurs for profit. The movie described a questionable achievement of computing and biology, ruined by a power and communication breakdown triggered by dishonest behavior. The UNIX system, with a transparent interface between the human and the computer system, saved the day!

Over the last fifty years, progress has been fueled by the synergy of computing and communications. Most advances in engineering and business are enabled by powerful interconnected computing units housed in data centers. The unprecedented growth of computing power of processors and their interconnections is the physical enabler of Artificial Intelligence (AI) systems, that have come to everyone's attention because of their strong ability in problem solving. As a scientist, I enthusiastically embrace the development of this technology, but I raise concerns about its boundaries for the best of the planet and humanity.

The increasingly larger power of computational systems has evolved through a constant technological progress in device manufacturing and in architectural design. Processors have evolved into multi-processors connected by Networks on Chips (NoCs). In turn, computing racks and data centers leverage advanced optical and electrical data communication schemes. The Internet enables seamless communication among users, local edge systems and data centers, thus providing the backbone for distributed operation. Social networks have not just enabled communication among individuals and groups but have fundamentally changed the nature of human social behavior. As a consequence, computing and communication are the keys to power in its broad sense, including political, military and economic power.

From a very high-level of abstraction, we can see interconnected computing as a machine. This machine provides us with an organization of information of all types, rendering us a useful service. But, as all machines, it requires energy to operate. The entropic cost (i.e., the energy to be paid for organizing information) is the minimum energy contribution required by physics. But the overall energy cost is much larger because of our current inability to design machines optimal in terms of energy efficiency.

Decades of research efforts (for integrated circuits and software) have tackled the energy consumption problem, and despite progress we are still far from an optimum, because it is an inherently hard problem in theory and in practice. Whereas the energy cost may be small for a portable system, like a cell phone, it is significant for a large data center, where electric power demand is going up to 7GW, now requiring a dedicated power generation plant. Thus, the scaling up of AI systems and data centers has to take into account the large cost that they entail in terms of emissions and heat disposal.

Comprehensive plans should include useful disposal and reusing of waste heat for both data centers and power generation plants. From a broader perspective, political and economic power used to be rooted around energy sources (e.g., oil) fifty years ago. Today power is centered around knowledge, as said by the metaphor: "Data is the oil of the 21st century". But in turn, data acquisition is tied to availability of energy resources, thus making knowledge and energy the key pillars for economic growth and political stability.

Correctness is quintessential in the operation of systems, whether they leverage AI or not. As systems are applied to domains affecting human integrity (e.g., vehicular control, medical systems, defense), correct design and operation are necessary. Verifying correctness is known to be a computationally hard problem and realistically only some properties can and must be verified. Thus, it is imperative that systems (including the hardware and software

ON POWER FOR COMPUTING AND COMPUTING FOR POWER

components) have models supporting the formulation of verifiable properties. This is not the case of most AI systems, whose surprising excellent performances are not matched by a full understanding of their operations and thus limiting our ability to verify correctness.

Much has been said and written about the responsibility of physicists who conceived atomic weapons. Much attention is given to potential nuclear offensive capability. But little is said about the responsibility of engineers and computer scientists who design and operate systems where errors are possible and not quantified. Liability lays also in the hand of decision makers who replace human operation by agents with no guarantee of correctness. With any replacement of human activities by agents, there must be individuals who are responsible in front of the law. This issue is exacerbated by the application of AI to mass surveillance, questioning freedom of individuals and expanding the exercise of power beyond ethical limits. Furthermore, lethal autonomous weapon systems (LAWS) are still underregulated and becoming increasingly more powerful. Thus, they represent a large, possibly existential danger to humanity.

As the world evolves toward the broader use of AI systems, it is important to raise global awareness about

the importance of free will. Free will is the major asset in most societies. Relinquishing free will to decisions made by systems that may be smarter than us – even for small decisions – may represent a terrible loss for humanity. While this can seem apparently obvious, freedom from choice is a situation desired by many, because of laziness, lack of focus and search for convenience. Thus, education about freedom of decision should go along with technical education. Next, legislation should set clear boundaries on risks, transparency and ethical AI practices. The EU AI act is a first step in this direction, but we will need new regulatory steps as technology advances.

Whereas the previous arguments hint diffidence towards recent advances in AI, many positive goals have shown to be reachable or within reach. Technological progress is unstoppable, but it can be steered. The balance between Techno-optimism and Techno-pessimism was well presented recently by M. Vardi. The crucial factor is the reward function. The type of reward, i.e., the overall goal of the cyber world, has clearly a socio-political color. Even more for this reason, education on the advantages and risks of technology should be prioritized and be positioned in the present historical context of the world.





THE ECOCLOUD EXPERIMENTAL FACILITY

The capabilities of EcoCloud's experimental space have steadily expanded since its opening two years ago, with multiple experiments currently running within. The facility is ideally located, in the CCT building, alongside the EPFL data center and above the campus central heating plant. The experimental space has a set of eight racks identical to those in the production side of the data center, which are equipped with rear-door liquid-air exchangers. Additionally, two of the racks were recently equipped with a tertiary cooling loop, served by a coolant distribution unit (CDU): this upgrades their cooling capacity from 35 kW to 135 kW. These racks host experiments with direct liquid cooling (DLC) and enable a full comparison with air cooling (AC) on state of the art servers.

Two additional mobile racks, one of them over a large retention bucket, provide a safe "sand box" for experiments with liquid cooling. The experimental space has 150kW of cooling allocation for the row of eight racks, and 2 x 50kW for the rest of the experimental space.

EcoCloud has acquired this year an immersion tank to plunge servers in dielectric oil, with a cooling capacity of 40kW. This new experiment allows for the benchmarking of AC vs. DLC vs. immersion cooling.

The set of computing resources available to experiment on the energy efficiency of algorithms and IT systems was extended last year with donations of servers from the SCITAS computing platform, the Faculty of Engineering IT team and the former Blue Brain project; with three new Cisco servers funded by EcoCloud and a switch donated by Cisco; and with the donation of a large Alveo FPGA by AMD Xilinx.

During 2026, the largest project using the experimental space has been Heating Bits, financed by EPFL solutions 4 Sustainability and involving six laboratories: DESL, ESL, IPESE, LAMD, PEL and POWERlab. The project has entered an exciting new phase with the prototyping of the newly developed cold plates by Corintis, through a

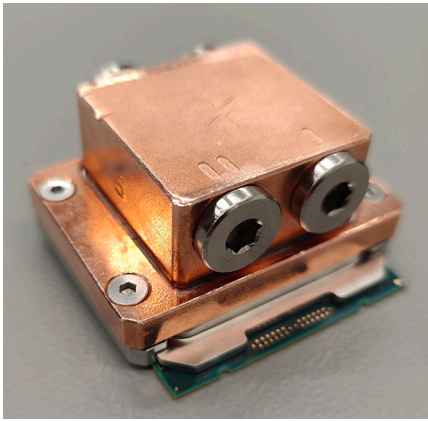
demonstrator for the end of Phase 1. These new microfluidic cold plates enable cooling of the server processors using high-temperature coolant - which opens the door to more efficient energy recovery. The Phase 1 demonstrator includes 16 CPUs fully retrofitted with microfluidic cold plates. Full benchmarks are being run at the moment to assess the retrofitting, compared to the current trend of pushing hot air within servers. Micro-channel DLC shows exemplary versatility to push the boundaries of cooling, while increasing the performance of servers.

During the second phase of this project, the heat retrieved from the servers will be delivered to an organic Rankine cycle (ORC) to recover part of the energy directly as electricity. Alternatively, the recovered heat will be injected into the campus heating system in the lower floor. Meanwhile, the project is developing an intelligent controller that will modulate workloads to ensure the maximum utilization of electricity from renewable sources, e.g., by monitoring a set of batteries and the carbon intensity of the electric grid at each moment while acting upon the server power allocation and performance.

Among the projects run within the facility, an optical switch developed by the LAPD, and currently under incubation in Zigzag, allows for fast switching of videos between screens, and perfect broadcasting.

New projects are foreseen for the coming year, including new techniques with cold plates, peak heat absorption, and automating tasks with robots.

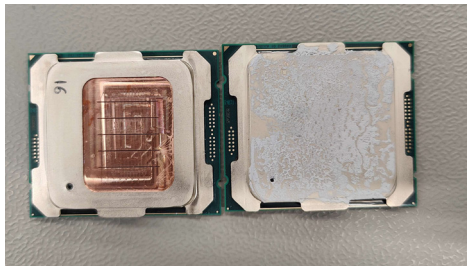
The EcoCloud experimental facilities welcome research projects that explore new avenues for IT sustainability, and the use of IT for sustainability. Any idea, proposal or project that fits our mission is welcome.



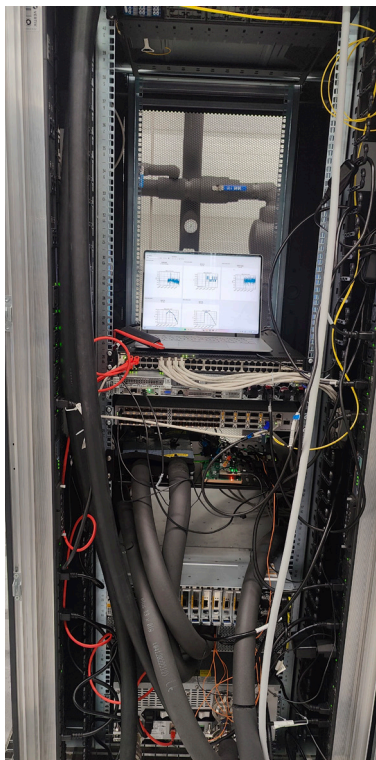
Direct Liquid Cooling (DLC) assembled coldplate



Optical switch plate



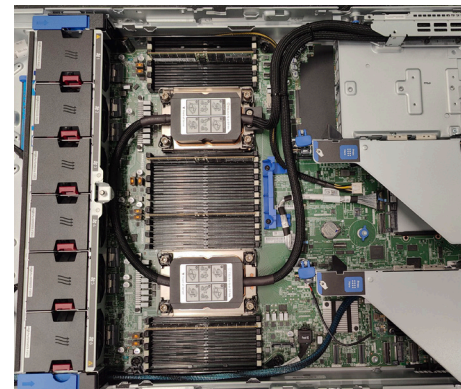
DLC microchannels inlid vs. normal CPU



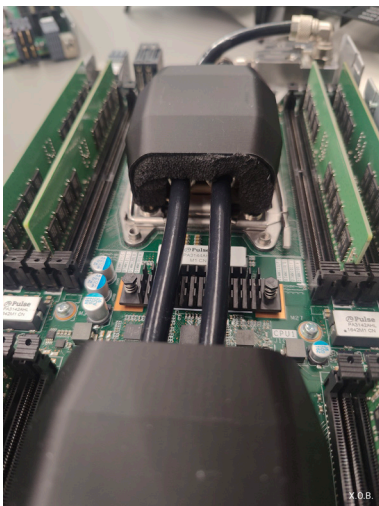
Heating Bits demo rack rear view



Cisco server ready for immersion



DLC server with coldplates



Coldplate blade side



Cisco server in full immersion



DLC versatile CDU provisioning



GEMATEG: TAKING THERMAL MANAGEMENT TO THE NEXT LEVEL

We spoke to Maurizio Miozza, CCO and co-founder of GemaTEG, about their company and the challenges it was created to address.

"This is an exciting time," says Miozza. "Artificial intelligence is here to stay, but its extraordinary capabilities come with an equally extraordinary demand for computing power.

"As workloads grow, CPUs and GPUs require more electrical power, most of which ultimately becomes heat. At the same time, increasingly complex chiplet architectures create localized hotspots and significant temperature variations across the package.

"Data centers therefore face a dual challenge: removing more heat while limiting the energy and water required to do it. Cooling is essential, but if we define the problem simply as cooling, we also limit the solutions available to us."

Not cooling, but thermal management

"We prefer to think in terms of thermal management. The distinction matters because cooling removes heat; thermal management gives you control over where, when, and how that heat is handled."

GemaTEG developed DaTEG to provide active, localized temperature control across advanced chiplet architectures. The system regulates individual thermal zones in real time, responding to the different demands of compute cores and high-bandwidth memory.

"It is almost like having a thermostat directly on the chiplet," explains Miozza. "Using advanced control methods, including machine-learning algorithms, we can react dynamically to operating conditions and manage temperature with much greater precision."

That level of control requires more than one cooling technique.

"Our mantra is integration. Next-generation computing cannot be supported by a single thermal approach. We combine solid-state micro heat pumps, conventional single-phase liquid cooling, advanced heat-transfer methods, and machine-learning-based control algorithms into a single thermal-management platform."

Our mantra is integration

Building such a platform also requires expertise across multiple disciplines.

"This is fundamentally an integrated product. Thermal science, electronics, materials, controls, software, and system engineering all have to come together to make it work."

GemaTEG established its research center in Perugia, Italy, because of the strong expertise available in systems engineering and integration.

"One of the reasons we chose Europe for our R&D activities is the ability to bring highly specialized expertise together around complex engineering problems.

"Our collaboration with academia has been equally important. It gives us access to deep knowledge in specialized fields and allows us to bring that knowledge directly into product development."

Building the technology from the ground up has also given GemaTEG a detailed understanding of the entire thermal path.

"We now look at the problem from the chip and package all the way to the cooling loop, control architecture, and surrounding infrastructure. That perspective is becoming increasingly important as thermal constraints emerge as one of the key limitations of AI infrastructure."

Miozza also sees value in bridging the different cultures of academia and industry.

"Academia brings scientific depth and exploration; industry brings execution, reliability, and measurable outcomes. The real opportunity is combining the strengths of both and turning advanced research into solutions that can be engineered, validated, manufactured, and deployed."

An opportunity to create an ecosystem

Miozza is cautious about claims of simple breakthroughs in a field governed by unforgiving physics.

"There is no shortage of ambitious claims in cooling. You often see a new cold plate or cooling solution presented as a breakthrough, only to find that the underlying thermal problem has simply moved somewhere else."

"That is why we look at the entire heat path. The laws of thermodynamics have not changed. Whatever innovation you introduce, heat still has to move from one point to another, and the physics has to work across the complete system."

For GemaTEG, that complexity also creates an opportunity for collaboration.

"We work with our research center in Perugia, leading Italian R&D institutions, and organizations such as EcoCloud. No single organization needs to solve every part of the problem. The opportunity is to bring complementary expertise together around a common objective."

Miozza believes that combination of scientific rigor, engineering discipline, and collaboration is essential as the demands placed on AI infrastructure continue to grow. Artificial intelligence is already raising concerns about energy consumption and the expansion of data-center infrastructure. Miozza nevertheless remains strongly optimistic about its potential.

"This is an extraordinary time. If developed and used responsibly, AI can expand human knowledge and capabilities in ways that were difficult to imagine only a few years ago."

"I strongly believe it can improve people's lives and create enormous value for society. The challenge is to ensure that the infrastructure supporting it evolves just as intelligently."

Taking it to the next level

"We are already moving into the next phase of the AI revolution. Generative AI created enormous momentum, but rapid technological change inevitably brings disruption, resistance, and new challenges."

"As a physicist, I tend to think in terms of equilibrium. Periods of rapid change eventually lead to a new balance, and I believe we are beginning to see happening in computing infrastructure."

One example is the changing relationship between cloud and on-premises computing.

"Hyperscale infrastructure will remain essential, but it will increasingly coexist with on-premises computing and storage. The two models serve different needs and will develop alongside one another rather than one replacing the other."

"That evolution creates an important opportunity for GemaTEG. More distributed computing will require infrastructure that is efficient, manageable, and capable of operating within increasingly demanding thermal constraints."

"If we solve that challenge, we can make advanced computing more broadly accessible while taking energy efficiency and thermal performance to an entirely new level."



GemaTEG founders Maurizio Miozza, CCO, and Manfred Markevitch, CEO.

Maurizio Miozza will return to give a talk at the EcoCloud Annual Event 2026: "Beyond the Thermal Wall: Intelligent Cooling Architecture for the AI Era."

A portrait of Prof. Rachid Guerraoui, a man with dark hair and a beard, wearing a dark suit and a light-colored shirt. He is smiling and has his hands clasped in front of him. The background is a blurred office setting with a large screen and some architectural elements.

DO WE REALLY NEED BIG DATA CENTERS FOR AI?

With Prof. Rachid Guerraoui

EPFL researchers have developed new software – now spun-off into a start-up – that eliminates the need for data to be sent to third-party cloud services when AI is used to complete a task. This could challenge the business model of Big Tech.

The use of artificial intelligence for everyday tasks has grown rapidly in the past three years. AI models are increasingly useful to process sensitive data such as patient records, customer requests or any work-related confidential documents.

Each time AI is asked to perform a task, the query starts locally on a user's computer and is then sent into the cloud, where the AI generates an answer using powerful hardware, a process known as inference. The answer is finally sent back to the user's local computer.

Huge data processing capabilities are nowadays used to do this, requiring enormous data centers, of the same type as those which are needed to train AI models, such as ChatGPT, Gemini and Claude. This means that both inference and training are currently controlled almost exclusively by Big Tech.

Privacy, sovereignty and sustainability

When a user's data is sent to the cloud, there are crucial questions around security and privacy, particularly whether that data is used to further train or improve AI models. As well, the reliance on large, global cloud providers for AI services raises questions around AI sovereignty because it transfers control over critical national assets—data, algorithms, and infrastructure—from a domestic entity to transnational corporations.

Moreover, the immense computing power needed to answer AI queries in the cloud – it is estimated that inference accounts for 80 to 90% of AI-related computing power – is driving the rapid expansion of the huge data centers for AI that consume enormous amounts of energy and water.

Plug and Play: distributed AI made simple

Now, EPFL researchers Gauthier Voron, Geovani Rizk and Rachid Guerraoui, from the Distributed Computing Laboratory (DCL) in the School of Computer and Communication Sciences, have released new software that allows users to download open-source AI models and use them locally, with no need for the cloud to answer questions or complete tasks.

The new software, called Anyway Systems, coordinates and combines distributed machines on a local network into an on-premise cluster. It uses robust self-stabilization techniques to optimize the usage of underlying local hardware, contradicting the common belief that huge data centers are needed to deploy AI models.

It can be installed in just half an hour on a network of local machines with no data leaving the network, guaranteeing privacy and sovereignty. A very big AI model like GPT-120B, the latest and largest open model from OpenAI, can be downloaded and deployed on Anyway Systems in a few minutes, requiring no more than 4 machines with 1 commodity GPU each (costing around 2300 CHF each), instead of an expensive specialized rack enclosure (costing around 100,000 CHF), which until now has been assumed necessary to run an AI model.

"For years people have believed that it's not possible to have large language models and AI tools without huge resources, and that data privacy, sovereignty and sustainability were just victims of this, but this is not the case and smarter, frugal approaches are possible," said Professor Rachid Guerraoui, head of the DCL.

"Anyway Systems shines on inference but it could also help reduce the resources needed for training," explained Guerraoui. Pilot testing has shown that when a model is downloaded and run on scattered local machines instead of a huge cloud, we may lose a bit of latency – that is, time

DO WE REALLY NEED BIG DATA CENTERS FOR AI?

to respond to a prompt – but not any accuracy.”

From blockchain to AI?

“We say that Anyway Systems is simple, scalable and safe,” continues Guerraoui. Earlier variants of the Anyway algorithm were developed by the DCL many years ago where researchers have long focused on distributed computing, fault tolerance, optimization and privacy.

The DCL's earlier algorithms were solutions for other challenges for technologies such as blockchain and cryptocurrency. Three years ago, Guerraoui and his colleagues had the idea to apply self-stabilization techniques to AI, finding an almost perfect fit.

“As a lab we might be unique in working on robust distributed computing and machine learning together from both a theoretical and a practical perspective and we turned our attention to using self-stabilization techniques for AI. They worked! We thought, let's optimize and optimize and they worked even better, the result is almost too good to be true,” Guerraoui said.

Looking ahead – your own AI at home

Anyway Systems was recently chosen as one of six inaugural grantees of the Startup Launchpad AI Track – powered by UBS, Switzerland's first grant program dedicated to AI. Selected from over 50 proposals, these projects are receiving funding and tailored support to accelerate their journey from prototype to market readiness.

The software has gone beyond the prototype phase and is now being tested in companies and administrations across Switzerland, including at EPFL. Early users are currently evaluating any trade-offs in terms of speed, accuracy and quality.

“Anyway Systems represents an interesting and appealing technology that optimizes resource usage while ensuring data security and sovereignty and could be an AI gamechanger,” said Professor David Atienza, associate vice-president of research centers and technology platforms at EPFL. “Its sustainable approach aligns perfectly with the needs of EPFL's advanced computing platforms and will play a pivotal role in shaping the trajectory of future AI development at EPFL to consume fewer resources with the new deployment of LLM models such as Apertus.”

For now, Anyway Systems won't work on a single desktop or laptop at home but the history of computing shows that optimization often occurs quickly.

“Your phone contains crazy amounts of information that would have been unimaginable a few years ago and now you do everything on it. It can beat all the best 100 chess champions at the same time, whereas the computer that was needed to beat Kasparov was enormous. History tells us this is the way things go. What we're saying is that we will be able to do everything locally in terms of AI. We could download our open-source AI of choice, contextualize it to our needs, and we, not big-tech, could be the master of all the pieces,” concluded Guerraoui.

Question and answer session

What is the difference between Anyway Systems and Google AI Edge?

Google AI Edge is meant to be run on mobile phones for very specific and small Google made models with each user running a model constrained by the phone's capacity. There is no distributed computing to enable the deployment of the same large and powerful AI models that are shared by many users of the same organization in a scalable and fault-tolerant manner. The Anyway System can handle hundreds of billion parameters with just a few GPUs.

What is the difference between Anyway Systems and other solutions that allow people to run local LLMs such as Llama or msty.ai?

Most of these approaches help deploy a model on a single machine, which is a single source of failures. To deploy the most powerful models they need to buy and invest in very costly machines, basically the same kind you would find in a data center. Hence, if you have a single commodity machine, the cited solutions are useful for deploying small models. If you have several commodity machines, you can't combine them efficiently with the cited solutions to deploy a large model, and even if you could, it would require a team to manage and maintain the system. Anyway Systems does this transparently, robustly and automatically. The fact that a machine fails, or leaves or joins the network is transparent in Anyway Systems, saves a small change in latency (the time to respond to a request).

AI models are constantly being improved and fed, how are these improvements reflected locally?

As Anyway Systems allows any open-source AI model to be deployed locally, feeding it with local and sensitive data becomes completely safe and acceptable, handing control back to the user.



MIDGARD: MISSION ACCOMPLISHED BUT THE STORY GOES ON

**We spoke to Prof. Babak Falsafi,
founder of EcoCloud and Midgard**

Midgard, the international collaborative research initiative, has completed its run of five years. The technology pioneered in this collaboration represents a paradigm shift, questioning the fundamental virtual memory system that has endured since the 1960s. At the same time it has demonstrated a unique form of research funding, coming to light at a time when Swiss universities had lost the right (recently regained) to lead most European research projects.

All modern computers run a trusted piece of software called the operating system, that enables the safe and secure sharing of hardware. Computer memory, which accounts for half of the cost of a server's hardware today, is where a program code and its data are kept. Modern operating systems are based on principles of "virtual memory" that were formed in the 1960s to build operating systems for safe and secure use of memory.

What happened? Firstly, CPUs grew exponentially faster and memory became exponentially bigger. Secondly, there are now tens of thousands of services (programs) running on the CPU that need to be directed to their own space. Thirdly, memory has become more expensive: in the last 12 months, memory prices have jumped 500%.

The process of ensuring and managing protected access to memory forms a bottleneck.

Researchers at EPFL saw the possibility to innovate a new virtual memory design that, with CPU support, can accelerate managing and securing access to memory. This new design would be available to application developers on mobiles or servers, allowing full compatibility with today's software running on Linux, Android, iOS/macOS and HarmonyOS.

A paradigm shift like this required a high level of collaboration between many institutes. As mentioned above, when

Midgard started, Switzerland had recently lost the right to lead projects in almost all European Union research funding organisations. For its time, Midgard was therefore an outlier in being able to federate a global team of experts, for a simple reason: the funding came from industry.

Transformative research needs transformative financing

"For a project in transformative architecture we needed a big funding instrument, in terms of the amount of funding, and in terms of the wide range of institutions we wanted to get involved," explains Prof. Babak Falsafi. "EcoCloud had secured funding from Microsoft and Facebook in the past, but this time it was Intel who came forward. So we could make a collaboration between Yale, Edinburgh University, Korea University, HES-SO and EPFL, with the backing of the world's biggest CPU manufacturer.

"That was the whole purpose of creating EcoCloud: to be able to work with companies, while at the same time creating funds that would allow us to work with other Academic institutions."

In a recent paper, presented at the 52nd Annual International Symposium on Computer Architecture in Tokyo, it was shown how Midgard technology could be applied to FaaS (Function as a Service), a popular software development paradigm offered by cloud service providers such as Amazon, Microsoft and Google.

FaaS was originally conceived so developers from various organizations (and potentially untrusted parties) could write programs (in the form of just a software function) that worked together seamlessly with the operating system managing secure access to memory. "Unfortunately, today's FaaS systems suffer from millisecond-level performance bottlenecks that arise from isolating functions using conventional virtual memory," explains Prof. Falsafi.

"We strove for efficient in-process memory isolation - a

MIDGARD: MISSION ACCOMPLISHED BUT THE STORY GOES ON

new hardware mechanism that exceeds the capability of traditional virtual memory systems and forgoes syscalls, page table manipulation, and TLB shutdown: processes that take tens or even thousands of microseconds to complete.

"We demonstrated that our implementation could meet microsecond-level Service Level Objectives for microservice workloads while performing within 16% of an idealized but insecure baseline, delivering over two times higher throughput compared to enhanced state-of-the-art academic operating system proposals."

The future lies in efficient rack technology

Marking the end of Midgard's five year run, as the largest computer systems project in the history of EPFL, are a raft of publications, an in-house demonstrator showing order of magnitude gains, a showcase presentation at Intel and - of course - talk of a sequel.

"We are looking at rack technology," explains Prof. Falsafi. "Apple make purpose-built processors for their smartphones and laptops. Most data centers, however, have been using desktop hardware and operating system technology. It's now time to create cloud-native hardware and software technologies. The next wave of innovation is in "rack-scale" integration. Today's cloud datacenters bring together several dozens of nearly identical servers into a "rack". Because the services have diverse use for hardware, much of this hardware ends up being fragmented and not used across dozens of servers. This is not only a waste of silicon but also of idle power.

Rack-scale integration with the help of Midgard can help improve utilization of the hardware. As far as software is concerned, Midgard is compatible with modern operating systems including Linux, Apple iOS/macOS and HarmonyOS.

"There's always left-over memory that data centers don't



use," concludes Falsafi. "Microsoft reported last year, in a paper, that 50% of GPUs are underutilised in AI. We're trying to create technologies to help operators use their hardware more efficiently. That is the goal that drives us on."

So, will there be Midgard 2? No name as yet. "Midgard" comes from the ancient Norse, meaning Middle Earth. Whether the next one is Asgard, or even Valhalla, we can be sure that the paradigm will shift again.





SWISSCHIPS AN ONGOING INITIATIVE

In conversation with Dr. Alexandre Levisse-Willommet,
co-lead of IC Design Infrastructure at SwissChips

The SwissChips initiative is coming to a close, but not to a full stop.

At a time when Switzerland's scientists were being deprived of many sources of European research funding, Swiss national initiatives needed to be identified, and financed, at short notice. It was also clear that many countries had started to bolster their chip design and manufacturing industries, and to seek more autonomy, as demand continued to soar.

In order to tackle these two challenges of getting funding to scientists, and building up Switzerland's chip industry, the SwissChips initiative was launched in 2024.

"This is an initiative that was started by Swiss universities," explains Dr. Alexandre Levisse-Willommet, "and the goal was to support academic research in Switzerland. So the

institutions involved are mainly EPFL, ETHZ, CSEM plus our collaborators from universities of applied sciences. There are work packages dedicated to fundamental research, while at the same time creating a collaborative network between institutions."

Collaboration is a simple concept, in theory. "Chip design is an extraordinarily complicated field in general, but it becomes even more complex when it comes to working together between several institutions or companies," explains Levisse-Willommet. "For this reason, it is vital that there is a solid design framework. You need to ensure that everyone is working with similar design flows, software, technologies, in similar design environments, with compatible infrastructure.

"Then there are the legal aspects. We needed to setup the legal framework that would allow us to share data, share



SWISSCHIPS

AN ONGOING INITIATIVE



Prof. Alberto Sangiovanni-Vincentelli of UC Berkely with EcoCloud Director Giovanni De Micheli.

Prof. Sangiovanni-Vincentelli was keynote speaker at the SwissChips Event 2026. He was also the supervisor of Prof. De Micheli's PhD thesis.

The third annual SwissChips event will take place in Lugano on 4th-5th May, 2027.

designers, returns of experience, or fabricated chips.

"At the same time, we invested in open source hardware and open source software, so we could mitigate some of these constraints.

"These are just some of the ways we have learned to work together in SwissChips."

It is clear from talking with Dr. Levisse-Willommet that he sees a lot of potential in the Swiss chip design ecosystem, in which the SwissChips Initiative has been playing a leading role.

"From an educational aspect we did something fantastic this year: we hosted high school students for a week. They came up with great ideas, and in a week designed their first chip, that is currently being fabricated. This is not something you see often.

"At EPFL the number of students choosing to study electrical engineering has been consistently increasing, so there is general momentum in this direction, mostly pushed by the hype and innovation in fields such as AI, robotics or autonomous driving.

"Our students at EPFL get the opportunity to learn through projects. They could carry these projects over from Bachelor to PhD. With these projects we are able to bring them to a level where they make real innovations, and even to incubate startups. It is a self-perpetuating cycle."

The SwissChips Initiative has also organized summer schools and winter schools, training and networking

events. "Young scientists can participate in cutting edge science, present posters, participate in challenges set by experts in the field.

"The SwissChips Initiative annual event did attract hundreds of people in 2026, with many poster presentations. This event can compete with the best conferences in the world in our field, thanks to the enthusiasm of everyone involved in this excellent initiative."

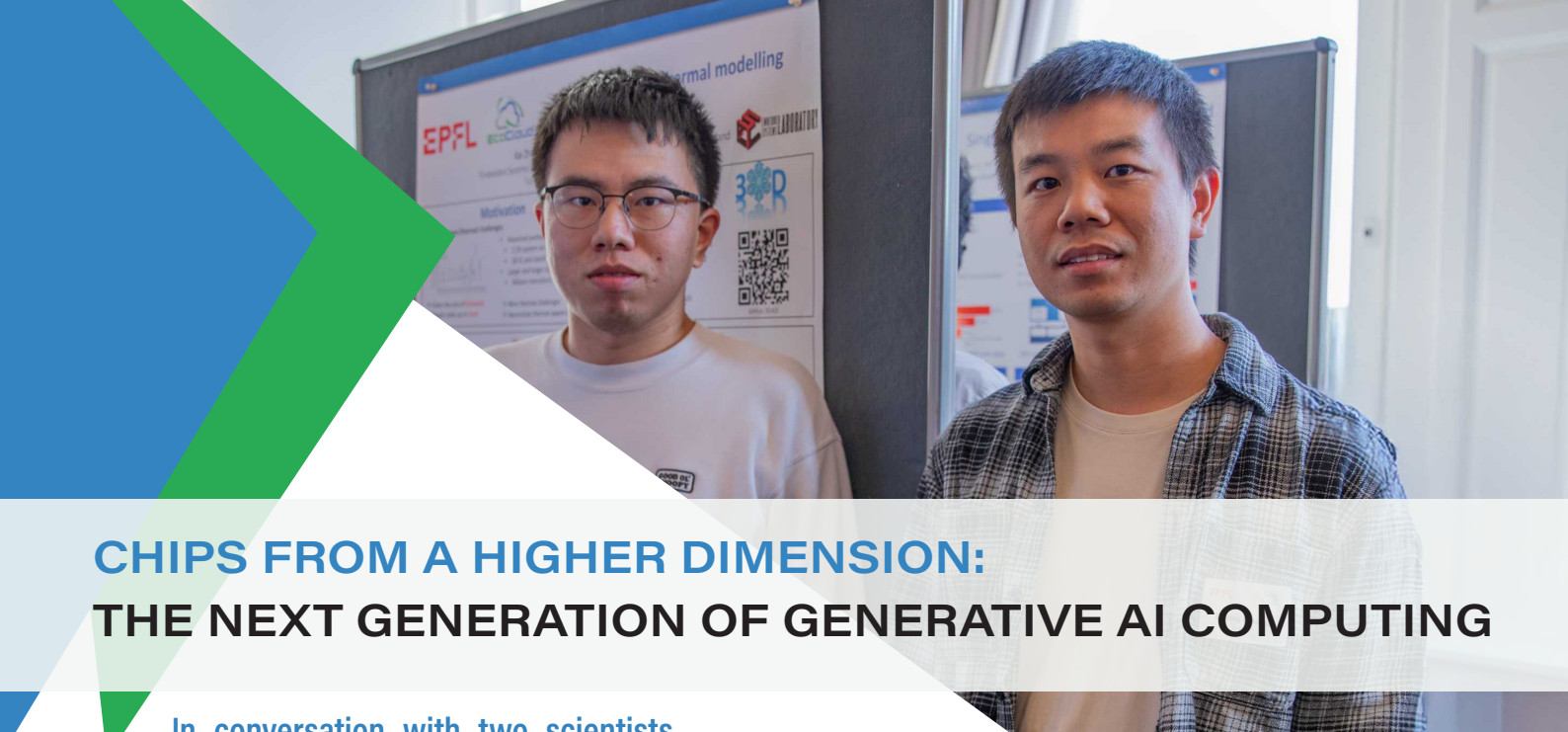
So what next? SwissChips itself has a limited run - ending in 2027, but future scientists will benefit greatly from its legacy.

"We have now established a framework of collaboration, across institutes in Switzerland, and we know how to work together, sharing data and information and, most importantly, we have designed together state-of-the-art chips that get published in top-notch venues! We have set up a strong baseline, and this will continue," explains Levisse-Willommet.

"Partly because of the nature of microelectronics design, teams in Zurich, Neuchâtel and Lausanne used to work in silos. We now have a great framework for collaboration.

"The SwissChips Initiative highlighted the importance of chip design, from a strategic and sovereignty standpoint, in Switzerland. At EPFL, capitalizing on SwissChips, we have created a new platform that supports chip design inside and outside EPFL: the CHIPS platform," explains Dr. Levisse-Willommet, who is Operational Director.

SwissChips has therefore followed the scientific method: theories were put forward, experiments were carried out, knowledge was gained, lasting progress made.



CHIPS FROM A HIGHER DIMENSION: THE NEXT GENERATION OF GENERATIVE AI COMPUTING

In conversation with two scientists
from Prof. David Atienza's
Embedded Systems Lab

For many years, Prof. John Thome of the Institute of Mechanical Engineering pioneered the use of heat sensing technology, and collaborated with Prof. David Atienza of the Institute of Electrical Engineering to model ways in which 3D chips could be cooled efficiently, using techniques like micro-channels of liquids. This work has been further developed as actual products by Prof. Elison Matioli, and EPFL spin-off Corintis.

Prof. David Atienza, head of the Embedded Systems Lab (ESL), is continuing to explore this important technology. He has just completed a twenty-year survey of power and thermal modelling and management with the help of his ESL team members, which was published in IEEE Design & Test this month. ESL is developing a new thermal modelling emulator with more capabilities than have been seen thus far.

The survey was performed in collaboration with Guest Prof. Luis Costero, Dr. Darong Huang and PhD candidate Kai Zhu. As well as providing a detailed history of the technology, which has seen a lot of changes in the last twenty years, the authors were able to make predictions about future trends and challenges.

Emulators are vital to future chip designs in the generative AI era of computing. In the same way that Formula 1 teams use virtual models (or digital twins) of their cars — a less expensive and time-consuming process than testing an actual car over and over — chip developers need to be able to test a digital version of their new AI chips before sending them for manufacture.

The modelling of power, thermal, and cooling techniques remains key in the development of new strategies, according to the new survey. It is also vital to be able to model the electrical, mechanical and thermal effects which 3D chips will endure, and to include complex chips such as 3D MPSoCs (Multi-Processor Systems on Chips).

Back in 2012, Atienza's laboratory unveiled the 3D-ICE

thermal emulator - the first of its kind for micro-channels of liquid cooling. 3D-ICE is still going strong, and forms the basis for the latest version, 3D-ICE 4.0, which Atienza, Huang and Zhu presented at last year's EcoCloud Annual Event. It promises to be a game changer for next-generation 2.5D, 3D, and wafer-scale technologies.

"We include analysis of a very careful material distribution, over different layers," explains Kai Zhu. "We also examine resistance and heat gradients over different layers with differing materials." This is a vital new development, since many other researchers (such as Prof. Andras Kis and Prof. Adrian Ionescu) are pioneering the use of other materials beyond simple silicon.

"We also have a feature for multi-thread acceleration," adds Dr. Huang, "with more efficient and precise simulation."

As chips themselves get more complicated, so too must their emulators!

"Being able to rapidly (and very accurately) model thermal spreading and propagation in complex 3D stacks is incredibly important work in terms of finding ways to lower overall energy consumption in complex chipsets, for both cooling and computing," explains Prof. Atienza. "Manufacturers need solutions to control the heat surge in 3D many-core computing systems for generative AI workloads, but only researchers who fully understand computing engineering co-design (hardware and software co-optimization) and what is possible on the cooling side (single- and two-phase cooling, in-package many-fold capillary cooling, etc.) can forge new paths in the right direction."

3D chip technology is alive and well. It is evolving quickly, due to the need for generative AI computing, and EPFL researchers are at the cutting edge, as they have been since its inception.

**Congratulations to Dr. Darong Huang,
who is now Assistant Professor at the
University of Hong Kong**



FLEXIBLE DATA CENTERS COULD SUPPORT EUROPE'S NET-ZERO TRANSITION

Although electricity-hungry data centers are often seen as a growing energy burden, an EPFL School of Engineering study suggests that they could also help Europe cut carbon emissions by integrating renewable power and reducing energy waste.

According to the study, led by the group of Industrial Process and Energy Systems Engineering (IPESE) and published in *Cell Reports Sustainability*, data centers could account for nearly 8% of Europe's electricity demand by 2050. These facilities, which support cloud computing, streaming services, and artificial intelligence, are projected to consume about 667 TWh of electricity annually, roughly equivalent to the electricity use of a large European country.

Yet data centers could also help support the renewable energy transition. Using continent-scale energy-system modeling, the IPESE team examined data centers as more than just large electricity loads. They accounted for waste heat generated by cooling systems, and for the temporal and geographic flexibility of deferrable computing workloads, which can be rescheduled or migrated across data centers without materially degrading service quality.

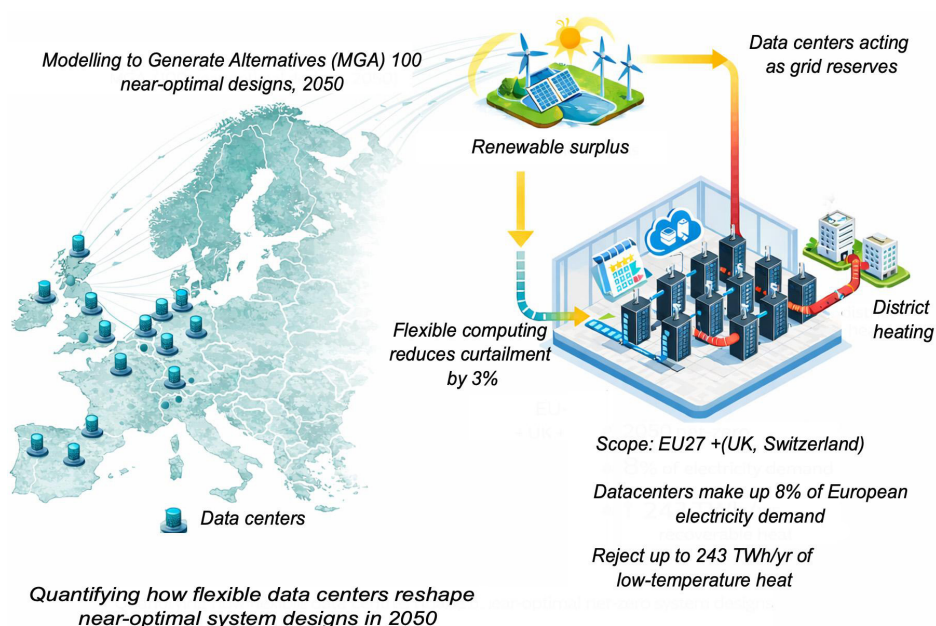
They found that data centers could generate up to 243 TWh of recoverable heat each year, much of which could be used in cities with district-heating networks rather than being released into the environment. And by scheduling certain data center activities – such as data analysis, backups, and AI training tasks – during periods of abundant wind and solar generation, operators could help absorb surplus renewable electricity that might otherwise go unused.

“Shifting flexible computing workloads to renewable-rich periods could help reduce Europe's curtailment rate, or

the share of renewable power that goes unused, from just over 9% to about 6.5%,” says lead author and IPESE PhD student Sai Sudharshan Ravi. “This could help countries make better use of growing amounts of wind and solar power, and avoid energy waste.”

The researchers say their results highlight the need to integrate information and communications technology infrastructure into long-term decarbonization planning and policy design in Europe. Currently, most energy system models tend to omit data-center demand, or treat it as a relatively minor addition to overall electricity use.

“Policies that encourage workload flexibility, grid interaction, and waste-heat recovery could help transform data centers from passive electricity consumers into active contributors to a cleaner and more efficient energy system,” Ravi summarizes.





SWISSCHIPS ANNUAL MEETING

A PLACE TO SHARE ADVANCES IN SILICON DESIGN

After the success of the first annual meeting organized by ETHZ, in 2026 EPFL took the baton to organize the second SwissChips annual meeting in Lausanne. Thanks to its large volume - this year the event gathered over 300 participants - and openness, the annual meeting has become a forum for interaction and exchange of ideas, needs and points of view.

The SwissChips annual meeting does not only attract participants from the involved universities: it is also open to industrial visitors, who can interact directly with the young researchers.

The meeting featured an invited talk from Prof. Alberto Sangiovanni-Vincentelli, from the University of California, Berkeley, and co-founder of Cadence Design Systems and Synopsys, two of the main EDA companies. Based upon

decades of experience, his talk served as inspiration to all attendees, lightening in particular the role of AI agents in the future of EDA tools.

In addition to featured talks, the leads of the work packages in SwissChips presented the advances made over the last year in their respective working groups. However, perhaps the most important part of the event were the slots reserved specifically for discussion. During them, the students and researchers of the different groups had the chance to engage in lively discussion and present their work on posters.

As much as semiconductors are based on interconnections between functional units, semiconductor research and industrial communities articulate through people connections. SwissChips is becoming the central catalyzer for that communication.



ECOCLOUD VISITING SPEAKERS



In November we welcomed Dr. Andrea Bonetti of Sony, who gave a talk on AI for Chip Design



Prof. Saron Hu of the University of Notre Dame gave a talk on Cross-Layer Co-Design in June



Prof. Luca Carloni of Columbia University (New York) came in April to give a talk on Chip Design with Open Source Platforms



In April, Prof. Prashant Nair of the University of British Columbia gave a talk on Scaling the Memory Wall



In June, Dr. Peter de Bock of Eaton gave a seminar on Engineering the AI Stack



Dr. Javier Picorel of Huawei, Zurich gave a talk on Scaling Interconnects in September

CONTACT

ECOCLOUD



Exploring new paradigms in data center hardware technology, pioneering strategies of cloud data management and innovative security techniques, in collaboration with industrial partners

EcoCloud

EPFL ECOCLOUD
INJ 234 (INJ Building)
Station 14
CH-1015 Lausanne
Tel: +41 21 693 13 24
contact.ecocloud@epfl.ch
<https://ecocloud.epfl.ch>

Prof. Giovanni De Micheli

Chief Scientific Officer
EPFL ECOCLOUD
INF 341 (INF Building)
Station 14
CH-1015 Lausanne
Tel: +41 21 693 09 11
giovanni.demicheli@epfl.ch

Valérie Locca

Administrative Assistant and IAP Manager
EPFL ECOCLOUD
INJ 234 (INJ Building)
Station 14
CH-1015 Lausanne
Tel: +41 21 693 13 24
support.ecocloud@epfl.ch



myclimate.org/01-24-942321

CREDITS AND RIGHTS



Texts copyright EcoCloud 2026, written by

4	Giovanni De Micheli
5-6	Giovanni De Micheli, Xavier Ouvrard, Miguel Peón-Quirós
10-12	Babak Falsafi and David Atienza
14-15	Giovanni De Micheli
16-17	Xavier Ouvrard, Miguel Peón-Quirós
18-19	John Maxwell
20-21	Tanya Petersen
22-26	John Maxwell
27	EPFL School of Engineering
28	Miguel Peón-Quirós

Image Credits - All Rights Reserved

1-2	Theerapong Khaotong, Vecteezy: 68464666
4	Alex Widerski
5	Eduardo Garlant, Adobe Stock: 317397107
6-9	EPFL
10	Alex Widerski
11	Eduardo Garlant, Adobe Stock: 317397107
12	Issaranow, Adobe Stock: 280047831
14-16	Alex Widerski
17	Xavier Ouvrard, Alex Widerski
18-19	GemaTEG
20	Alain Herzog
22-25	Alex Widerski
26	John Maxwell
27	iStock, EPFL School of Engineering
28-29	Alex Widerski
30	Ipopba, Adobe Stock: 480306328

