HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

THE NEXT COMPUTING PARADIGM

Funded by the European Union

Building the next computing paradigm

The ‘next computing paradigm’ is the convergence of technologies including the web, cyber-physical systems (CPS), cloud computing, the internet of things (IoT), digital twins and artificial intelligence (AI) into a coherent, federated ecosystem.

European academic and industry leaders need to **act fast** to establish made-in-Europe technologies in this rapidly changing landscape. Technological offerings should **meet the needs of European markets**, while ensuring that European technology is synonymous with **quality and trustworthiness** in the minds of consumers across the globe.

The HiPEAC Vision for the European computing ecosystem is characterized by the following factors, which play to European strengths and establish a ‘European’ flavour of computing:

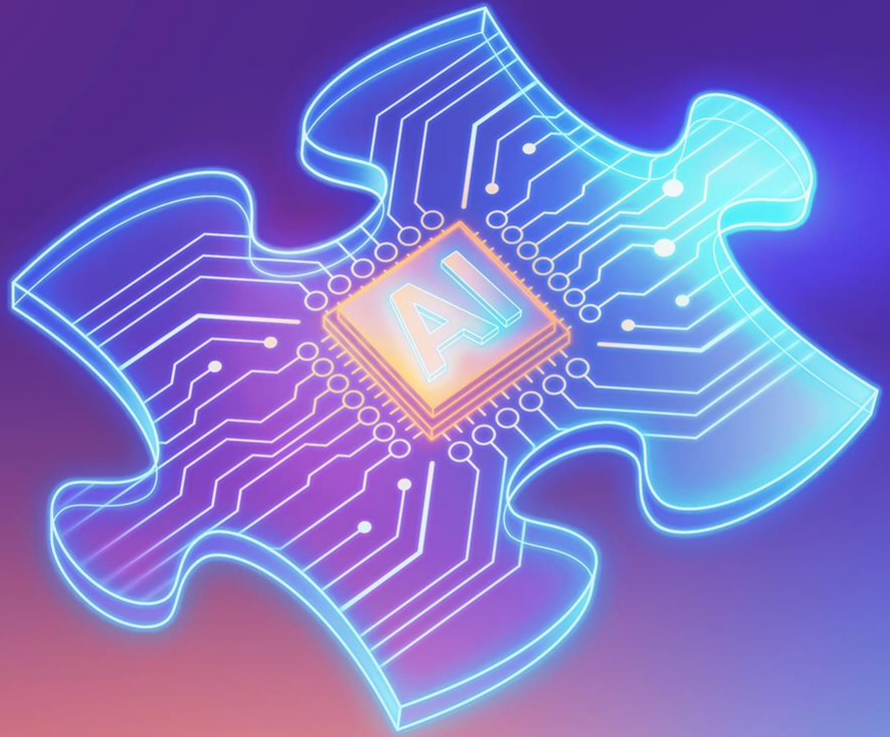
- Collaborative
- Federated
- Distributed
- Interoperable
- Open source
- Trustworthy (i.e. explainable, reliable, secure, safe and privacy-preserving)
- Sustainable

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

INDUSTRY: AI CHIPS

Funded by the European Union



HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

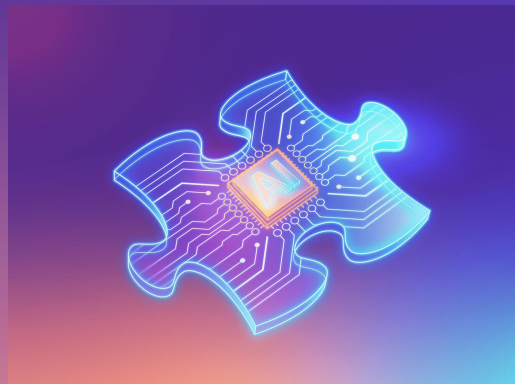
RECOMMENDATIONS INDUSTRY: AI CHIPS

Funded by the European Union

1. Execute specialized action models

Aim: Ensure that AI hardware can efficiently execute specialized action models (SAMs), support fine-tuning of SAMs and enable the deployment of SAMs in federated networks.

Actions: Ensure chips support on-device learning, lightweight model updates, and distributed execution of AI agents on a low power budget, among other functions.



HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

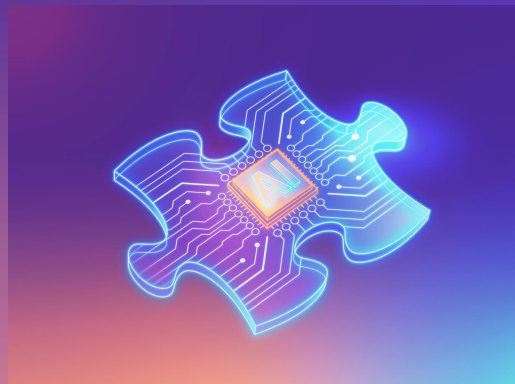
RECOMMENDATIONS INDUSTRY: AI CHIPS

Funded by the European Union

2. Optimize for the edge

Aim: Optimize hardware for edge deployment, using architectures which reduce energy use and unnecessary data transfers.

Actions: Focus on highly efficient custom accelerators which reduce overhead through non-volatile memory, near- or in-memory computing, or direct-from-storage execution, for example.



HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI CHIPS

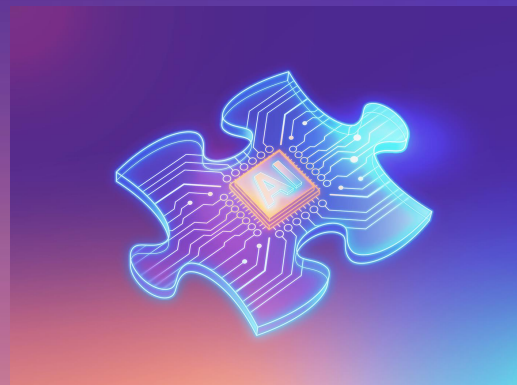
Funded by the European Union

3. Support AI orchestration technologies

Aim: Ensure that hardware can enable orchestrators to break down user goals into actionable service flows, as well as running SAMs and supporting federated genAI.

Actions:
Embed accelerators optimized for low-latency inferencing and orchestration logic at the edge, co-designing them with software for optimal efficiency.

Enable orchestrators to dynamically combine SAMs to execute personalized applications.



HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

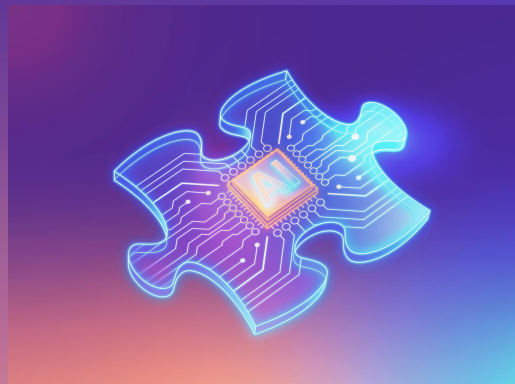
RECOMMENDATIONS INDUSTRY: AI CHIPS

Funded by the European Union

4. Make it interoperable

Aim: Support collaborative, sandboxed execution of services and enable the live migration of compute. Clear, contract-based APIs – defining performance guarantees as well as functionality – should be used for services and devices to communicate.

Actions: Ensure hardware natively supports virtualization, runtime adaptability, and containerized AI execution, so that it is plug-and-play with evolving digital service ecosystems.



HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI CHIPS

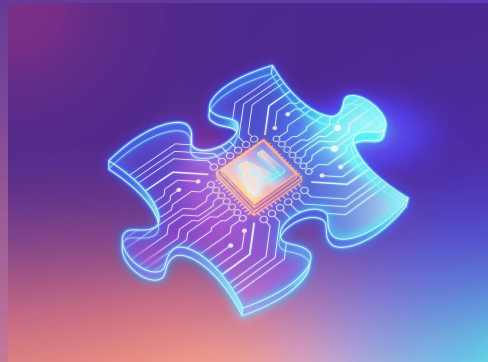
Funded by the European Union

5. Reboot security, from supply chain to design

Aim: Integrate cybersecurity into hardware design and management from the outset, adapting to the distributed setting of the next computing paradigm.

Actions:
Harden firmware / software against supply-chain attacks. Use zero-trust architectures that support encrypted boot, secure enclaves, and trusted execution.

Partner with cybersecurity researchers to implement proactive, self-healing hardware security measures and ensure the secure orchestration of AI workloads.



HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI CHIPS

Funded by the European Union

6. Prioritize sustainability by design

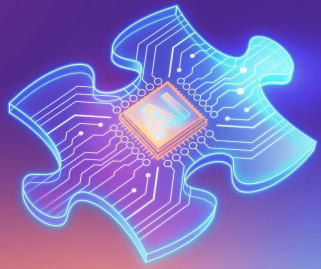
Aim: Adapt production and processes to be sustainability first and make sustainability a selling point of your hardware.

Actions:
Develop life-cycle sustainability models for hardware products, taking into account both production (including water use, raw materials, chemicals used in production, end-of-life processing) and operation (including energy use).

Create digital product passports for hardware products based on sustainability data to allow customers to make informed choices.

Make sustainability a first-class citizen in hardware design choices.

Develop new business models based on life-cycle services including maintenance, repair and disposal.



HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

INDUSTRY: AI MODELS AND SOFTWARE

Funded by the European Union

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI MODELS AND SOFTWARE

Funded by the European Union

1. Develop agentic AI (specialized action models, or SAMs)

Aim: Develop SAMs acting as a service, which should operate in a distributed infrastructure.

Actions:
Optimize general foundation models to create lightweight SAMs capable of interacting with their environment.

Position optimized SAMs in a marketplace so they can be dynamically discovered by orchestrators.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI MODELS AND SOFTWARE

Funded by the European Union

2. Develop orchestrating technologies for distributed agentic AI

Aim: Develop orchestration technologies which should analyse application requirements, then select and activate SAMs.

Actions:
Develop edge-native orchestration technologies that break down user requirements into actionable tasks and that dynamically combine SAMs for personalized applications.

Ensure that orchestrators can collaborate and share resources.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI MODELS AND SOFTWARE

Funded by the European Union

3. Develop open protocols for distributed agentic AI systems

Aim: Establish open protocols to facilitate seamless interaction among distributed AIs.

Actions:
Incorporate non-functional requirements and provide the orchestrator with sufficient information to select appropriate services.

Include information about response time, accuracy, environmental impact, localization, privacy, etc.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI MODELS AND SOFTWARE

Funded by the European Union

4. Use models for cybersecurity and harden against attacks

Aim: Support cybersecurity in distributed systems with tailored models, while strengthening defences against prompt injection and meeting AI-focused cybersecurity standards.

Actions:
Develop made-in-Europe federated AI models for threat detection and mitigation.

Defend models against prompt injection using techniques such as input validation, output encoding and defensive prompting.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: AI MODELS AND SOFTWARE

Funded by the European Union

5. Differentiate AI models with sustainability

Aim: Place sustainability at the heart of model development and execution.

Actions:
Create life-cycle models for AI models with comprehensive data on environmental impact, including both embodied and operational emissions.

Use life-cycle models to optimize AI model designs to lower their environmental impact.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

INDUSTRY: SYSTEM-LEVEL INTEGRATION

Funded by the European Union

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: SYSTEM-LEVEL INTEGRATION

Funded by the European Union

1. Create digital envelopes based on XaaS

Aim: Create digital envelopes to seamlessly integrate users and devices into a cohesive ecosystem, allowing access to services in XaaS mode.

Actions:
Enable live migration of compute components and runtime-evolving infrastructure to support deployment across the computing continuum.

Ensure interoperability across diverse systems and platforms.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: SYSTEM-LEVEL INTEGRATION

Funded by the European Union

2. Expand web-level protocols

Aim: Expand web-level protocols and associated standards to be both spatially aware and time-sensitive.

Actions:
Enhance existing communication stacks with WebRTC-like capabilities for time-sensitive data.

Adopt spatial web standards such as OpenUSD and IEEE P2874 for 3D context awareness.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: SYSTEM-LEVEL INTEGRATION

Funded by the European Union

3. Implement contract-based API specifications

Aim: Design and implement interoperable contract-based API specifications with clear expectations for functional and non-functional performance.

Actions: Develop contracts similar to SLAs detailing the conditions under which services will perform optimally, including non-functional properties such as latency and cost.

Work with standardization bodies to promote and consolidate standards, ensuring an interoperable ecosystem of services and orchestrators.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: SYSTEM-LEVEL INTEGRATION

Funded by the European Union

4. Ensure robust, end-to-end cybersecurity

Aim: Tackle cybersecurity proactively across federated, distributed systems, from production to supply-chain to source code.

Actions: Identify and address vulnerabilities at the component, interconnection, agent and orchestration levels.

Use dedicated cybersecurity AI tools where appropriate to triage, analyse, and mitigate threats autonomously.

HiPEAC Vision 2025

HIGH PERFORMANCE, EDGE AND CLOUD COMPUTING

RECOMMENDATIONS INDUSTRY: SYSTEM-LEVEL INTEGRATION

Funded by the European Union

5. Address sustainability holistically

Aim: Take a system-wide approach to sustainability, ensuring global optimization.

Actions: Use validated life-cycle models to evaluate the environmental impact of the whole system, including both embodied and operational emissions.

Use orchestrators to make sustainability-focused operational choices (e.g. selecting the most energy-efficient paths or devices).

Shift to service-based business models offering maintenance and repair to prolong device lifetimes and ensure system longevity.