

Project Report

# How Europe Is Preparing Its Core Solution for Exascale Machines and a Global, Sovereign, Advanced Computing Platform

Mario Kovač <sup>1,\*</sup>, Philippe Notton <sup>2</sup>, Daniel Hofman <sup>1</sup>  and Josip Knezović <sup>1</sup> 

<sup>1</sup> Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, 10000 Zagreb, Croatia; daniel.hofman@fer.hr (D.H.); josip.knezovic@fer.hr (J.K.)

<sup>2</sup> SiPearl, 22, rue Guynemer, 78600 Maisons Laffitte, France; philippe.notton@sipearl.com

\* Correspondence: mario.kovac@fer.hr

Received: 31 May 2020; Accepted: 16 July 2020; Published: 20 July 2020



**Abstract:** In this paper, we present an overview of the European Processor Initiative (EPI), one of the cornerstones of the EuroHPC Joint Undertaking, a new European Union strategic entity focused on pooling the Union's and national resources on HPC to acquire, build and deploy the most powerful supercomputers in the world within Europe. EPI started its activities in December 2018. The first three years drew processor and platform designers, embedded software, middleware, applications and usage experts from 10 EU countries together to co-design Europe's first HPC Systems on Chip and accelerators with its unique Common Platform (CP) technology. One of EPI's core activities also takes place in the automotive sector, providing architectural solutions for a novel embedded high-performance computing (eHPC) platform and ensuring the overall economic viability of the initiative.

**Keywords:** high-performance computing; processor; architecture; accelerator; common platform; exascale computing; automotive computing; low-power computing; artificial intelligence; edge computing

## 1. Introduction

The importance of high-performance computing has been on the rise in recent years, and it is expected that this trend will not only continue at the same pace but rapidly grow. Industry reports show that annual global IP traffic will soon reach several zettabytes, vast amounts of new devices collect and store data, scientists are exploring new computing approaches to solving global challenges and the industry is changing the way products are designed. In the same time, we, as individuals, are constantly expecting more personalized services for the whole spectrum of the aspects of our lives, such as better drugs that we expect to come faster to the market, have fewer side effects and cost less, faster diagnostic tools that enable better medical treatments, autonomous cars that can drive us more safely and more economically and many other upcoming advancements.

The need to collect and efficiently and timely process that vast amount of data comes at a price. The existing approach to HPC system design is no longer sustainable for the exascale era (exascale systems being defined as those capable of executing  $10^{18}$  calculations per second) in terms of both performance and power consumption. Energy efficiency is of enormous importance for the sustainability of future exascale HPC systems.

The importance of sustainable high-performance computing has been recognized by the European Commission and it has strategically initiated efforts to achieve and support activities towards the implementation of European exascale computing systems and related technologies. EU efforts are

synchronized in the establishment of the EuroHPC Joint Undertaking [1], a legal funding entity which will enable the pooling of national and European Union (EU)-wide resources in high-performance computing to acquire, build and deploy the most powerful supercomputers in the world within Europe [2].

The European Processor Initiative (EPI) [2] project is one of the cornerstones of this EU HPC strategic plan. This project is currently implemented under the first stage of the Framework Partnership Agreement signed by the Consortium with the European Commission, whose aim is to design and implement a roadmap for a new family of low-power European processors for extreme-scale computing, high-performance big data and a range of emerging applications. The first phase of EPI started in December 2018 and all the core activities of the project are well underway.

EPI will ensure that the key competence of high-end chip design remains in Europe, a critical point for many application areas. Thanks to such new European technologies, European scientists and the industry will be able to access exceptional levels of energy-efficient computing performance. As recognized by high-level EU officials, EPI will benefit Europe's scientific leadership, industrial competitiveness, engineering skills and know-how—not to mention society as a whole.

The design of a novel HPC processor family cannot be sustainable without thinking about possible additional markets that could support such long-term activities. Thus, EPI will cover other areas, such as the automotive sector, ensuring the overall economic viability of the initiative.

## 2. The Approach

EPI uses a holistic approach to define the system architecture and its components and to develop:

- hardware platform architecture and components;
- system and runtime software (OS, middleware, developers' kits, libraries, etc.);
- end-user applications from various domains not limited to traditional HPC.

To fulfil its objectives of working towards a hybrid exascale system, EPI will develop:

- a novel, exascale HPC-focused low-power processing system unit;
- an accelerator unit to increase energy efficiency for computing-intensive tasks in HPC, artificial intelligence (AI), automotives and many other application domains;
- an automotive demonstration platform to test the relevance of the previous components in this industry sector.

## 3. The Interplay of Activities

EPI activities are organized into three high-level technology domains and two global domains responsible for the integration and coordination of the previous technology domains. All five domains in the project are depicted as streams, as shown in Figure 1.

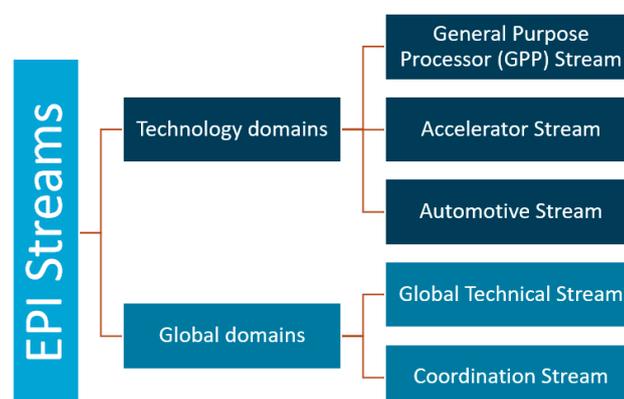


Figure 1. EPI streams organization.

### 3.1. General Purpose Processor (GPP) Stream

The general purpose processor (GPP) stream is focused on the creation of the first implementation of a processor platform targeting the HPC market.

The GPP stream addresses a strong set of common technologies across different application domains. It deals with the selection of cutting-edge process technology, massive parallelism with multi-cores, a memory hierarchy with high bandwidth memories (HBM) integrated using a silicon interposer, a chiplet approach with a high-speed link between silicon dies, a low-power design approach with a low-voltage operating point and fine-grain power management, built-in security to isolate applications and resist against the new cyber threat environment. The software stack will be designed to integrate and take advantage of these features to achieve high-energy efficiency and maximize performance across a wide range of layers from the low-level firmware, all the way up to system software and application run times.

### 3.2. Accelerator Stream

The accelerator stream targets the development and demonstration of fully European processor IPs based on the RISC-V instruction set architecture, providing power-efficient and high-throughput accelerator tiles within the GPP chip. Using RISC-V allows for leveraging open-source resources at the hardware architecture level and software level, as well as ensuring independence from non-European patented computing technologies.

### 3.3. Automotive Stream

The goals of the automotive stream activities are to drive innovation in high performance processing in the automotive domain, including the technology support for autonomous driving (levels 4/5) and the connected car infrastructure. Within this stream, HPC general-purpose processors and accelerators will be integrated into the architectural solutions for a novel embedded high-performance computing (eHPC) platform to demonstrate the approach to be technologically, functionally and economically successful.

New autonomous vehicle network architectures require computing platforms to be able to execute complex vehicle perception algorithms that include sensor/imaging processing, data fusion, environment sensing and modeling, low-latency deep machine learning for object classification and behaviour prediction with seamless, dependable and secure interaction between mobile high-performance embedded computing and stationary server-based high-performance computing.

The expected system requirements for future autonomous vehicles with regards to data produced by the sensors in the vehicles so that the vehicle can properly interact with its environment are shown in Table 1 [3]. The architecture concepts developed in EPI will fill the gap between the processing requirements and the electrical power a vehicle can provide.

**Table 1.** Autonomous driving levels and system requirements.

System Requirements	Level		
	3	4	5
Processing Capacity (TFLOPS)	40	80	120
DRAM (GB)	1	10	20
Non-Volatile Storage (GB)	100	300	500
Data Link Interface (Gbps)	0.1	1	1

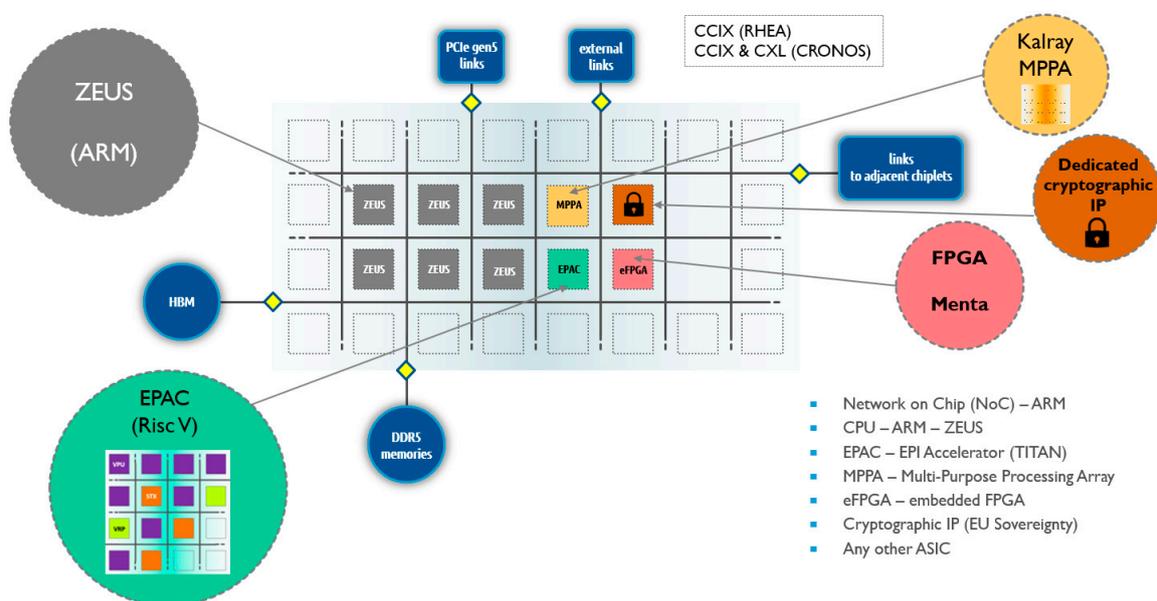
The global technical stream addresses global EPI architecture, co-design methodology, architectural explorations, system and management HW/SW and simulation and modeling tools for benchmarking and validation, while the coordination stream, as its name implies, is focused on the coordination of the whole project.

All activities within the streams are deeply interdependent and thus the interplay of all activities is crucial for the success of the initiative.

#### 4. The Technology

Exascale computing systems need to simultaneously meet challenges related to performance, system cost and energy efficiency. To deliver performance, a vast amount of resources is required, but the wrong choices of components, architecture or implementation might result in a system that is much too expensive and/or too power hungry. To find the right balance, global system level optimization is necessary.

For this purpose, EPI will harmonize the heterogeneous computing environment by defining a common approach: the so-called Common Platform (CP) shown in Figure 2. It will include the global architecture (hardware and software) specification, common design methodology and the global approach for power management and security.



**Figure 2.** EPI Common Platform architecture.

The CP is organized around a 2D mesh network-on-chip (NoC) connecting computing tiles based on general purpose Arm cores with high energy efficiency accelerator tiles, an RISC-V-based EPI accelerator (EPAC), a multi-purpose processing array (MPPA), cryptographic IP and embedded FPGA with different acceleration levels or any other application-specific accelerator.

A common software environment between heterogeneous computing tiles will harmonize the system, as well as act as a common backbone of IP components for IO connection with the external environment, such as memories and interconnected or loosely coupled accelerators. With this CP approach, EPI will provide an environment that seamlessly integrates any computing tile. The right balance of computing resources for application matching will be defined through the ratio of the accelerator and general-purpose tiles.

The HPC general purpose processor (GPP) will be the first implementation of the Common Platform targeting the HPC market. During the project, we will develop the first generation of the GPP with two revisions:

- a real production chip (GPP generation 1–revision 1), tested and validated as suitable to build a pre-exascale HPC system;
- specification and IP ready for the second revision to meet the performance targets (FLOPS/W, FLOPS/socket and bytes per FLOP memory bandwidth).

The EPAC accelerator tile (Figure 3), will be a fully European processor intellectual property (IP) accelerator, based on the RISC-V ISA and aimed at providing very low-power and high-computing throughput accelerator to the general-purpose cores. It will include specialized tiles, such as a vector processing unit (VPU), stencil/tensor accelerator (STX) and variable precision co-processor (VRP).

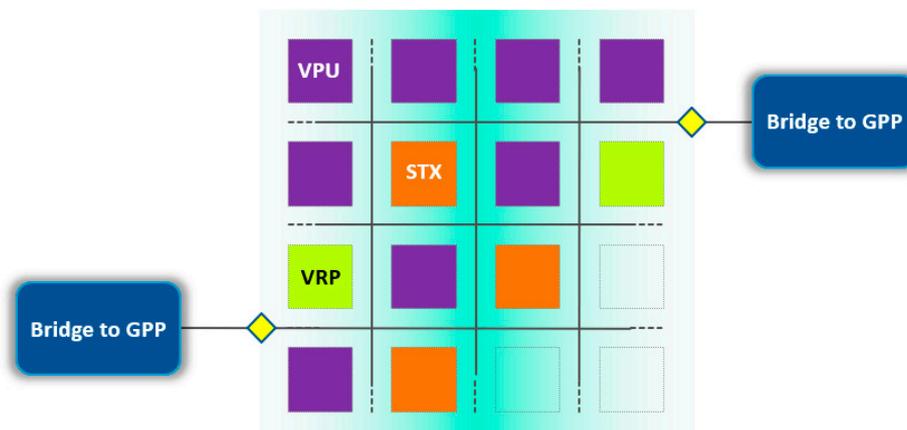


Figure 3. EPAC—the EPI accelerator.

Using RISC-V allows us to leverage many open-source resources at the architecture and system software level and at the same time to innovate and develop new state-of-the-art solutions. The EPAC accelerator will be based on the RISC-V vector instruction set architecture (ISA) and will include specialized blocks for stencil, deep learning and variable precision acceleration. The vector and stencil capabilities will address workloads in HPC centers, while the deep learning block will target AI acceleration.

In the EPI processor, the Kalray MPPA tile targets the acceleration of automated driving perception functions and other high-performance time-critical functions of cyber-physical systems, including machine learning inference. MPPA applies the successful principle of GPU manycore architectures: processing elements are regrouped with a multi-banked local memory and a slice of the memory hierarchy into compute units, which share a global interconnection and access to system memory. The key feature of the MPPA manycore architecture is the integration of fully software-programmable cores for the processing elements, and the provision of an remote direct memory access (RDMA) engine in each compute unit. The cores implement a 64-bit, six-issue VLIW architecture, which is an effective way to design instruction-level parallel cores targeting numerical, signal and image processing applications. Moreover, the implementation of this VLIW core and its caches ensure that the resulting processing element is fully timing compositional, a critical property with regards to computing accurate bounds on the worst-case response time (WCRT).

One of the distinctive core elements in the EPI CP architecture is also Menta embedded FPGA (eFPGA). It is a design adaptive standard cell-based architecture that provides the highest degree of design customization, best-in-class testability and fastest time-to-volume for SoC design, targeting any production node at any foundry. The eFPGA IP cores are 100% standard cell-based, supporting DFT and test coverage up to 99.8%. The embedded FPGA cores will be supplied with a proven EDA tool, Origami Programmer, that supports the design from HDL design to bitstream with synthesis, mapping, place and route.

EPI processors will be designed by the new fabless semiconductor company SiPearl. SiPearl has licensed the Zeus core from Arm and will introduce a chip based on TSMC's 6-nm node in 2021.

#### *Common Platform Vision: A Federation of Accelerators*

One of the distinctive, visionary features of the EPI Common Platform approach is the federation of accelerators (Figures 6 and 7). This unique feature is enabled by the interposer and multi-chiplet concept of the Common Platform (Figure 4). Starting from the introductory SiPearl Rhea family of

processors, chiplets will allow for building a variety of performance/feature options for members of a family. Besides the obvious fact that the number of chiplets will define the processing power provided in a single chip, the CP chiplet approach goes beyond this by also providing the option of HBM configurations and processing IP can even be integrated within a single chiplet as a core in the 2D mesh (as shown in Figure 3).

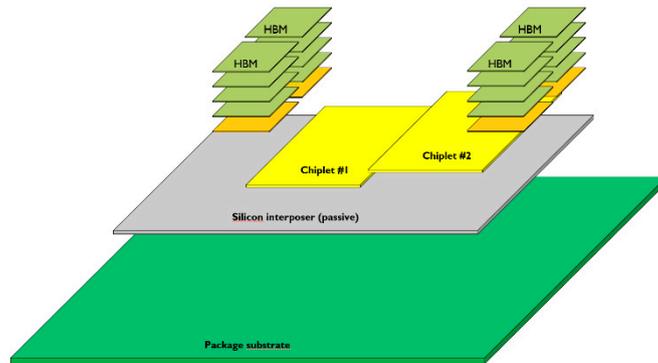


Figure 4. Interposer and multi-chiplet concept.

This design decision has created the opportunity to offer the integration of accelerator IP into a powerful Rhea/Cronos platform to those that could, to date, not be able to consider working their accelerator IP in such a powerful processing platform (Figure 5). The possibilities with this approach are now wide, globally attractive and provide many new business opportunities.

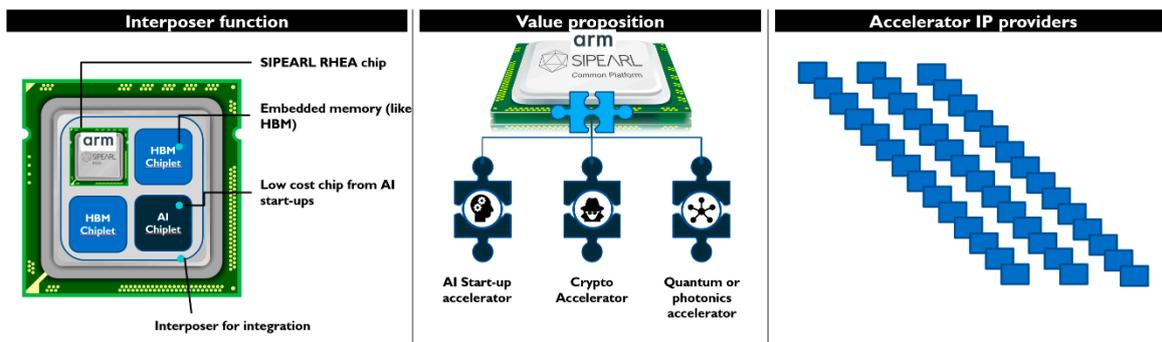


Figure 5. EPI CP federation of accelerators.

From the developers' and users' standpoint, this approach presents new opportunities to map specific applications for certain functions/kernels to dedicated accelerators and still relies on a widely accepted Arm general computing platform for general processing (Figure 6).

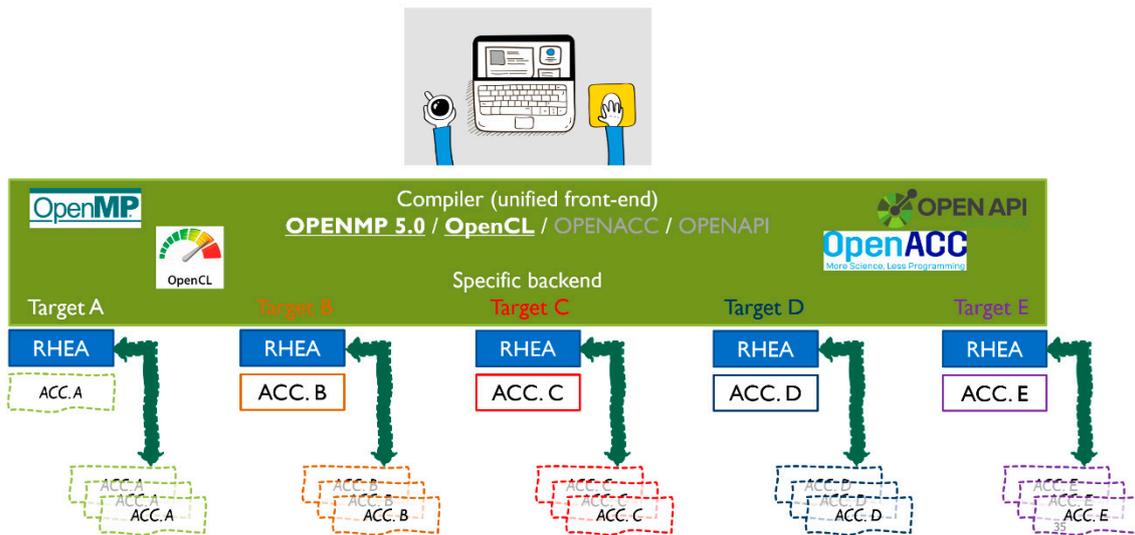


Figure 6. EPI CP federation of accelerators.

The use of specialized accelerators tightly integrated with the powerful Arm GPP cores provides an ideal state-of-the-art platform for today’s and future HPC + AI applications (Figure 7).

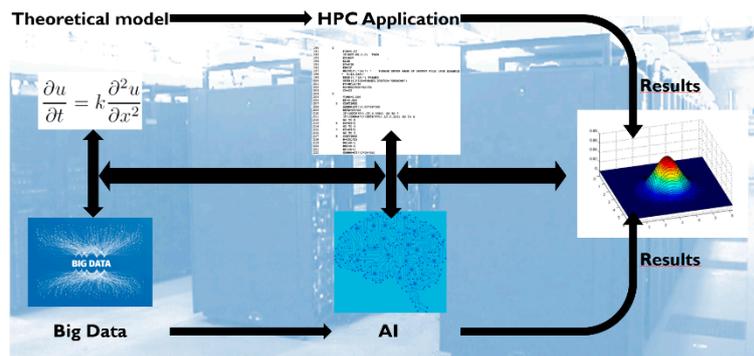
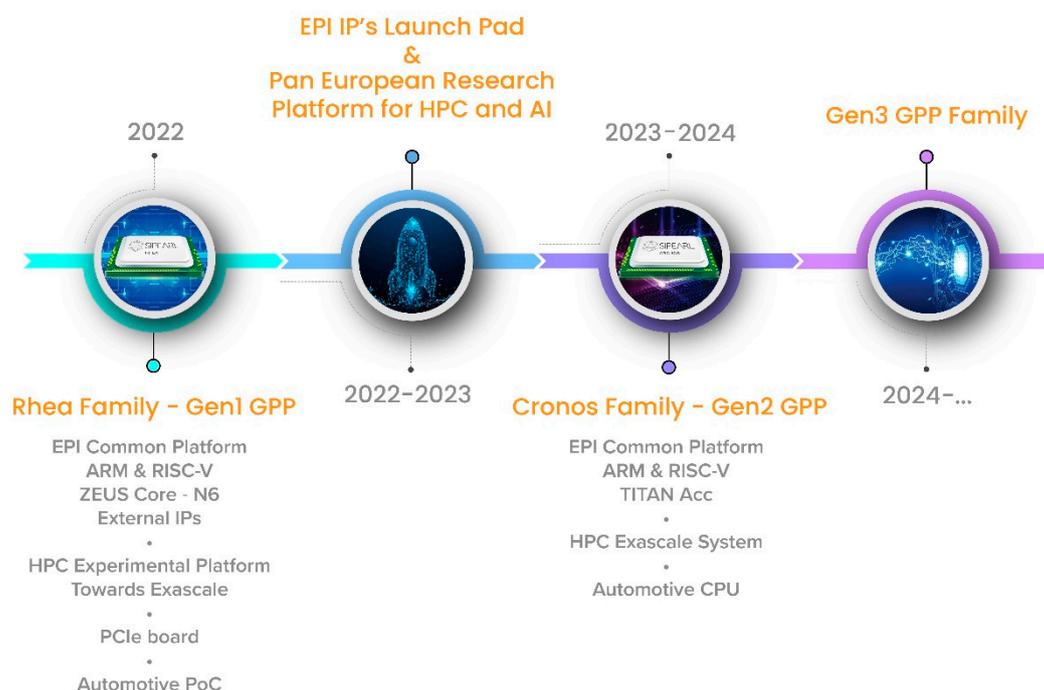


Figure 7. HPC and AI.

### 5. EPI Roadmap

The EPI Roadmap is shown in Figure 8. The first-generation chip family, named Rhea, will include Arm architecture general-purpose cores, EPAC accelerators, MPPA, and eFPGA blocks. The Rhea chips will be integrated into test platforms to validate the hardware units, develop the necessary software interfaces and run applications. Rhea aims to be the European processor for several EU platforms.



**Figure 8.** EPI Roadmap.

## 6. Conclusions

EPI will provide European industry and research with a world-class competitive HPC platform, ecosystem and data processing solutions and capabilities. EPI will harmonize the heterogeneous computing environment by defining a Common Platform that will include the global architecture (hardware and software) specification, common design methodology and global approach for power management and security. The EPI Consortium expects to deliver HPC computing solutions with unprecedented levels of performance, consuming much less power than before by allowing potential users to select the right balance between acceleration and general-purpose processing tiles on our Common Platform.

**Author Contributions:** Conceptualization, M.K., P.N., D.H. and J.K.; methodology, M.K., P.N., D.H. and J.K.; investigation, M.K., P.N., D.H. and J.K.; writing-original draft preparation, M.K., P.N., D.H. and J.K.; writing-review and editing, M.K., P.N., D.H. and J.K.; visualization, M.K., P.N., D.H. and J.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** The European Processor Initiative project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 826647.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. European Commission. The European High-Performance Computing Joint Undertaking—EuroHPC|Digital Single Market. Available online: <https://ec.europa.eu/digital-single-market/en/eurohpc-joint-undertaking> (accessed on 4 May 2020).
2. EPI. European Processor Initiative. Available online: <https://www.european-processor-initiative.eu/> (accessed on 4 May 2020).
3. Gartner. Hype Cycle for Automotive Electronics. Available online: <https://www.gartner.com/en/documents/3883996/hype-cycle-for-automotive-electronics-2018> (accessed on 4 May 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).