REVIEW PAPER

The international race towards Exascale in Europe

Fabrizio Gagliardi[1] · **Miquel Moreto**[1] · **Mauro Olivieri**[2] · **Mateo Valero**[1]

Abstract

In this article, we describe the context in which an international race towards Exascale computing has started. We cover the political and economic context and make a review of the recent history in high performance computing (HPC) architectures, with special emphasis on the recently announced European initiatives to reach Exascale computing in Europe. We conclude by describing current challenges and trends.

1 Introduction

Over the last several years HPC computing has become essential not only for big science applications and mission critical defence, but for almost any aspect of modern life and economy. This is particularly true after the advent of Machine Learning applied to Big Data. Obviously, the HPC systems need to be completely trustworthy. For this to be true the entire HPC ecosystem from the processors, communication and all the way up through the various software layers needs to be based on technology that is readily available, not subject to foreign export licencing, and completely trustworthy. For all these reasons as early as the beginning of the decade, all major developed regions have started to invest heavily in HPC technology. Particularly remarkable is the case of China, which started from scratch and, in a few years, has managed to become a major player in HPC and operates some of the top systems in the world. More traditional HPC players, such as Japan and the US, have also continued to invest and launch programmes such as the Post-K in Japan, and the Exascale Computing Project (ECP) from the Department of Energy (DoE) in the US. Finally, the European Commission recently announced the EuroHPC initiative to reach Exascale performance with European technology in the upcoming years.

✉ Mateo Valero
mateo.valero@bsc.es
http://www.bsc.es/cv-mateo/

1 Polytechnic University of Catalonia, Barcelona Supercomputing Center, Barcelona, Spain

2 Sapienza University of Rome, Rome, Italy

2 The application breakthrough of Exascale computing

Besides the traditional mission critical applications of HPC, in the defence and security sectors, Exascale computing has enabled yet another major step for scientific applications. The unprecedented power of the latest generation of HPC systems allows the simuation of physical phenomena with high accuracy. Moving to Exascale, science will be capable of implementing simulation-based predictability. Computers will become in practical terms time machines capable to describe with almost no error what the future will be. From climate and weather prediction, to personalized medicine, to entire complex ecosystems. This evolution has convinced policy makers and governments worldwide that those countries which will own the most powerful HPC systems will dominate the international economy and rule the world. The potential impact on the society of scaling the performance of the most powerful computers of a factor at least tenfold in the next 3–5 years is considerable and strategic. European economy is already sustaining considerable losses from the important climate change of these last decades. The EU has established the Copernicus programme over 20 years ago. Today it is in full operation supported by a latest funding of 5.8 billions of € (B€). The system produces several 10's of Terabyte per day of satellite information. This information needs to be processed in real-time and compared to the data produced by very sophisticated models. Similarly, in personalized medicine the need for supercomputing performance is great with obvious impact on the citizen wellbeing and the cost of health care, which is becoming a major issue in the increasingly ageing European society.

3 From domain-specific to application-specific architectures and co-development

The history of HPC systems, spanning more than 50 years, shows the progression of recurring design concepts adapted to evolving scenarios:

- Fabrication technology advancements and varying limitations, mostly influenced by the trend of Moore's Law, characterized by short-medium term periods of evolution;
- Application landscape of HPC, generally characterized by a growing range but relatively slow in its evolution.

Subsequent to HPC architecture design evolution, the programming paradigms have generally evolved to meet the requirements for approaching peak performance of the architecture. Figure 1 depicts the main drivers of HPC development in the last decades.

Here we analyse part of the HPC evolution from the above perspective, through a set of representative cases, in the view of establishing the expected way for European Exascale computing system development.

3.1 The first rise of hardware acceleration: vector computers (1974–1993)

At the end of the mainframe age, which had gone through more than two decades after World War II, many foundational ideas of modern computer science such as compilers, operating systems, floating point arithmetic, virtual memory and memory hierarchy, had already been conceived and experimented. Mainframes were general purpose systems, although the range of computer applications was limited with respect to the present idea of general purpose computer, and they had struggled for years with the limited memory capacity.
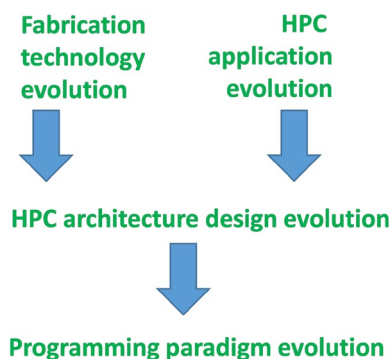

Fig. 1 Drivers of HPC system progression

3.1.1 Technology and application drivers

On the technology side, bipolar transistors and emitter-coupled logic (ECL) families had been already adopted to boost the speed at the expense of power efficiency. With the advancement of memory technology and memory sub-system design, the memory address space size ceased to be the limiting factor for application development. On the application side, the presence of heavy matrix algebra processing in scientific and military applications opened the way for acceleration of arithmetic operations on vectors of real numbers.

3.1.2 Architecture design advancement

Instructions operating on vector operands, rather than the more conventional scalar operands, were introduced along with hardware support within the CPU for executing vector operations. Vector processors exploit data level parallelism (DLP), where a Single Instruction operates over Multiple Data streams (SIMD) (Asanovic 1998; Espasa et al. 1998). This constitutes the first representative example of hardware acceleration of computational kernels, especially in the form of dedicated and parallel (SIMD-organized) functional units and of dedicated vector register file. For this reason vector computers can be considered the first form of domain-specific machine. Vector machines appeared in the early 1970s and dominated supercomputer designs for two decades.

There are two main classes of vector processors, depending on the location of their vector operands. Vector memory–memory architectures locate all vector operands in memory, while vector register architectures provide vector instructions operating on registers, while separate vector loads and store move data between memory and the register file. Some relevant vector memory–memory machines that appeared in such a period of time are the ILLIAC IV supercomputer (Barnes et al. 1968), the TI advanced scientific computing (ASC) supercomputer (Watson 1972), and the CDC STAR 100 (Hintz and Tate 1972) and successors (Lincoln 1982). In contrast, representative vector register architectures include the Cray series (Russell 1978; Cray Research 1984). These designs exploited DLP with long vectors of thousands of bits.

Many applications can potentially benefit from vector execution for better performance, higher energy efficiency and greater resource utilization. Ultimately, the effectiveness of a vector architecture depends on its ability to vectorize large quantities of code. However, the code vectorization process incurs in several obstacles, such as horizontal operations, data structure conversion or divergence control. As a result, a significant effort was done in improving automatic vectorization of scientific codes (Callahan et al. 1988). However, autovectorizing large scientific codes requires the

programmer to perform some degree of code annotation, modification or even complete rewrite.

3.2 The rise of massive homogeneous parallelism (1994–2007)

3.2.1 Technology and application drivers

The progresses of CMOS technology allowed the growth of general purpose single-chip microprocessor performance, pushed by increasing clock frequencies and by microarchitecture advances, targeting the high-volume personal computer (PC) market during the 80s. Microprocessor speed grew faster than memory speed, which represented the first appearance of a *memory wall* limiting the performance. With the appearance of on-chip caches, made possible by the increasing scale of integration of CMOS, the initial memory wall was overtaken and the performance of general purpose complex CPUs, organized in multi-processor parallel architecture, became comparable with vector computer systems. Vector computers relied on specialized hardware boards composed of multiple integrated circuits. Because of their relatively limited market the development of powerful CMOS-based single-chip vector microprocessors was not justified. In order to maintain the pace with the speed of general purpose multi-processor, vector computers remained bonded to bipolar technology that again was not in the mainstream of semiconductor industry. More recently, a further significant advance on memory technology was the advent of eDRAM, allowing on chip L3 caches for example, thus pushing back the memory wall in commodity CPU based HPC systems [eDRAM].

On the application side, the market favoured the availability of general purpose parallel architectures that were not intrinsically devoted to a special class of algorithms like vector CPUs.

3.2.2 Architecture design advancements

The rise of massively parallel architectures based on off-the-shelf processors soon opened the way to shared memory symmetric multiprocessor and subsequently to the advent of distributed architectures to overcome the limit of memory bandwidth with respect to the traffic generated by multiple CPUs (the new memory wall). Cluster-based architectures, with multi-processor shared-memory nodes connected by different topologies and technologies of interconnection networks have represented the dominant paradigm of HPC systems for the last 25 years.

On the software side, programming massively parallel architectures has produced a number of approaches and APIs (application programming interface) broadly divided into shared-memory paradigms and message-passing paradigms, with the support of compiler technology.

3.3 The renaissance of acceleration units (2008– 2018)

3.3.1 Technology and application drivers

The arrival of the *power wall*, due to the increasing power consumption density, has definitely limited clock frequencies in favour of the increase of the number of cores integrated in the same silicon die. This phenomenon has been characterized by technology node progress through geometry scaling, accompanied by voltage scaling (a.k.a. Dennard scaling), while maintaining practically the same clock frequency and increasing performance by augmenting the number of cores on the die. Yet, due to the need for acceptable noise margins in the logic gates, Dennard scaling has demonstrated to be unfeasible, thus slowing supply voltage scaling with respect to geometry scaling. This effect has led to the impossibility of further increasing the number of active cores on the die, again due to excessive power density, a situation known as the *dark silicon* necessity (Esmaeilzadeh et al. 2012). Dark silicon refers to the need of maintaining part of the silicon die inactive or active at lower frequency than the CPU cores. One main way of facing this design complication is the adoption of specialised hardware acceleration to dramatically gain in power efficiency. To clarify the real foundations of the above trends, Table 1 evidences the actual gain in efficiency of hardware specialization on a simple computational kernel.

On the memory side, a decisive step to push back the memory wall has been the introduction of 3D-stacked DRAM technology created for graphic applications, basically High Bandwith Memory and Hybrid Memory Cube, which permit up to 512 GB/s data rate (HBM3).

3.3.2 Architecture design advancements

While the trend towards increasing parallelism continues, the designers of HPC systems have rediscovered hardware acceleration with the need for power efficiency. In fact, the consequences of this technology scenario, while limiting the proliferation of cores in large Systems-on-Chip, like those employed in HPC systems, move in two directions:

- Limiting the number of cores and the die area, leveraging acceleration of the computation by means of specialized external units hosted on the same board or even on daughterboards connected by high speed links like PCIe (Peripheral Component Interconnect Express);

Fig. 2 Parallelism progression
in HPC system

Parallelism

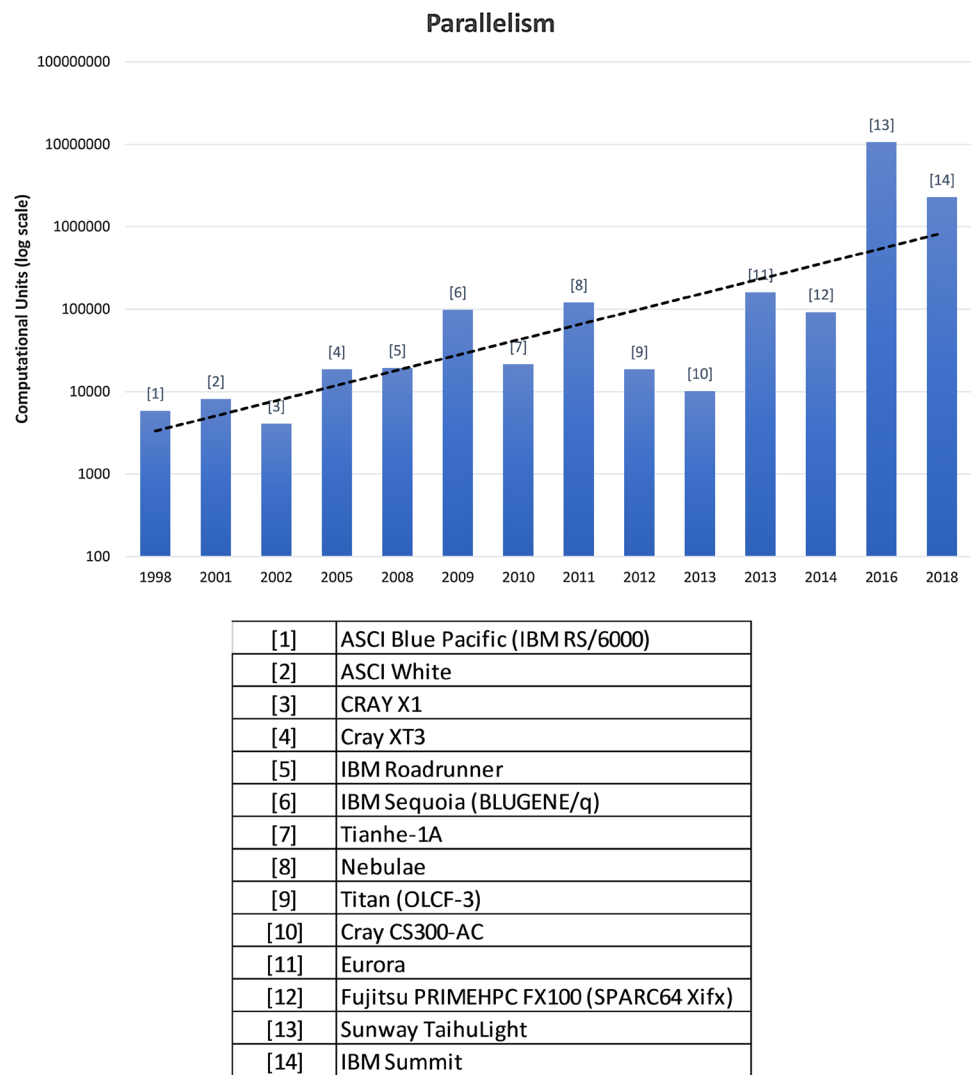|	
[1]	ASCI Blue Pacific (IBM RS/6000)
[2]	ASCI White
[3]	CRAY X1
[4]	Cray XT3
[5]	IBM Roadrunner
[6]	IBM Sequoia (BLUGENE/q)
[7]	Tianhe-1A
[8]	Nebulae
[9]	Titan (OLCF-3)
[10]	Cray CS300-AC
[11]	Eurora
[12]	Fujitsu PRIMEHPC FX100 (SPARC64 Xifx)
[13]	Sunway TaihuLight
[14]	IBM Summit

Table 1 Performance and energy efficiency of hardware acceleration on a numerical application design case (Adapted from Jason Cong, Keynote presentation at IEEE/ACM ISLPED'14)

128 bit AES encryption unit	Throughput	Efficiency Gb/s per W
Integrated circuit (IC)	3.84 GB/s	11
FPGA	1.32 Gb/s	2.7
Arm (assembly coded)	31 Mb/s	0.13
Pentium III (assembly coded)	648 Mb/s	0.015

- Leveraging on-chip hardware acceleration units that allow increased throughput at relatively low clock frequency, thus with considerably higher energy efficiency.

The first direction has been followed since 2008, with the development of the first supercomputer equipped with GPU (graphic processing unit) acceleration boards allowing a tremendous increase of parallelism while relieving the CPU from the highest computational load (Fig. 2). Notably, the advent of GPUs in HPC systems has been possible thanks to the availability of off-the-shelf top-level performance silicon chips fabricated for existing high-volume markets (high-end personal computers). This trend followed the same trend with the appearance of off-the-shelf high-speed CPUs in the HPC market 15 years before.

The second direction of development of acceleration units, i.e. on-chip accelerators, while promising a new boost in performance, thanks to the less severe limitation imposed by data transfers, necessarily requires a similar support from a high-volume market in order to sustain the cost of dedicated chip design and fabrication, to avoid the decline experienced by vector architectures in the 90s. Yet, this opportunity will be provided by the emergence of new application domains, not traditionally related to HPC, which may benefit from the usage of such dedicated high-speed processing chips. Such applications are already widely

documented and consist of Artificial Intelligence (Deep Learning), statistical analysis (Big Data), biology, etc.. Notably, in addition to extend the applicability of the HPC solutions to a larger market, the new applications also present particular requirements that demand attention by HPC system designers. Examples are reduced floating precision, integer processing, and bit-level processing. In this scenario, it is significant that big market drivers—like Google—are designing their own hardware solutions to solve their needs for high performance computing.

On the CPU side, the need for higher power efficiency along with the availability of high-performance and established software eco-system in ARM architectures, have opened the way to ARM-based supercomputers in the search for less power consuming CPUs. In the last decade, the Barcelona Supercomputing Center (BSC) has pioneered in the adoption of ARM-based systems in HPC. The Mont-Blanc projects (Online: http://montblanc-project.eu/), together with other European projects, allowed the development of an HPC system software stack on ARM. The first integrated ARM-based HPC prototype was deployed in 2015 with 1080 Samsung Exynos 5 Dual processors (2160 cores in total) (Rajovic et al. 2016). The clear success of these projects has influenced the international roadmaps for Exascale: the Post-K processor, designed by Fujitsu, will make use of the Arm64 ISA (Yoshida 2016), while Cray and HPE are developing supercomputers together with ARM in the US.

A special case of hardware acceleration units is represented by FPGA-enhanced systems. Some relevant examples of the adoption of this technology are already at a mature level of development. Most of such systems are using FPGA acceleration to cut down the communication latency in HPC networks—such as the case of Novo-G# architecture—leveraging the long experience of FPGA based routers. More generally, FPGA can be used at node-level processing acceleration, by reconfiguring data-paths within the FPGA connected to each CPU. This approach relies on the availability of the reconfigurable on-chip connections and on-chip memory structures within the FPGA, which in principle allows the exploitation of a higher degree of parallelism at the expense of one-order-of-magnitude decrease in clock frequency. An example of systems going in this direction is the Catapult V2 from Microsoft, which employs FPGA for local acceleration, network/storage latency acceleration, and remote computing acceleration. Yet the most interesting impact appears to be in network latency reduction. A rather exhaustive list of systems that have experimented FPGA accelerators in HPC systems and other related systems can be found in http://www.bu.edu/caadlab/HPEC16c.pdf.

On the programming paradigm side, the advent of GPU accelerators as well as on-chip accelerators and possibly FPGA accelerators, has surely complicated the scenario. In general, this trend pushes towards a more and more

collaborative development between the hardware and the software designers. In this scenario, a holistic strategy matches applications requirements with the technical implementation of the final design. In the last years, we have observed a rise in the popularity of novel programming models targeting HPC systems, especially focused on managing parallelism at node level. Parallelism between nodes at large scale still relies on the standardized Message Passing Interface (MPI).

Traditional programming models and computing paradigms have been successful on achieving significant throughput from current HPC infrastructures. However, more asynchronous and flexible programming models and runtime systems are going to be used to support huge amounts of parallelism with the hardware. With that goal, several task-based programming models have been developed in the past years for multiprocessors and multicores, such as Cilk, Intel Threading Building Blocks (TBB), NVIDIA's CUDA, and OpenCL, while task support was also introduced in OpenMP 3.0. These programming models allow the programmer to split the code into several sequential pieces, denoted *tasks*, by adding annotations to identify potentially parallel phases of the application.

More recently, emerging data-flow task-based programming models allow the programmer to explicitly specify data dependencies between the different tasks of the application. In such programming models, the programmer (or the compiler) identifies tasks that have the potential to run in parallel and specify the required input and output parameters. Then, the runtime (or the programmer) builds a task-dependency graph to handle data dependencies and expose the parallel workload to the underlying hardware transparently. Therefore, the application code does not contain information on how to handle the workload besides specifying data dependencies. Representative examples of such programming models are Charm++, Codelets, Habanero, OmpSs), and the task support in OpenMP 4.0.

4 Architecture and technology trend for European Exascale path

Presently installed supercomputing facilities in Europe are examples of the above described evolution. According to the Top500 list published in June 2018, there are six systems installed in continental Europe in the top 25 positions of the list (The TOP500 List 2018). The **Piz Daint supercomputer** is the fastest European supercomputer with 25.3PFLOPS peak compute power with the Linpack benchmark. Piz Daint, resembling the accelerator trend, is a Cray XC50 system featuring nodes composed of a Xeon host and a Tesla P100 NVIDIA GPU accelerator, totalling 361,760 cores. The **HPC4 supercomputer** follows a similar

approach, combining Intel Xeon hosts with NVIDIA Tesla P100 GPU accelerators.

Other European supercomputers follow the general purpose trend, such as the **MareNostrum supercomputer** hosted by BSC, the **JUWELS supercomputer** hosted by Juelich, the **Marconi-A1 supercomputer** hosted by CINECA, or the **TERA-1000-2 supercomputer** hosted by CEA. The MareNostrum supercomputer is a general purpose supercomputer consisting of 3456 compute nodes; every node has two Intel Xeon Platinum 8160 processors, each with 24 cores and 96 GB of DDR4-2667 main memory. MareNostrum has a total of 165,888 cores and 390 TB of main memory. The JUWELS supercomputer, the Marconi-A1 supercomputer are similar in their design as well as in performance; based on Lenovo NeXtScale nodes featuring Intel Xeon Phi processors, totalling 312,936 cores (The TOP500 List 2018). The TERA-1000-2 supercomputer follows a very similar design with Xeon Phi processors totalling 561,408 cores.

Installed supercomputers in Europe are still far from reaching Exascale performance, as this would require deploying a system 40 times more powerful than the Piz Daint. The path for building an Exascale system must take into account existing technology drivers and constraints. The latter come from the power wall, which means having the possibility of improving power efficiency. If we look at the trend in power efficiency evolution of representative systems, there is a clear exponential growth as shown in Fig. 3. This figure represents the energy efficiency in GFLOPS per Watt of the fastest supercomputer since 1996 according to the top 500 list. In the mid 90s, the largest supercomputers had a power efficiency of few MFLOPS/W. The fastest supercomputer in June 2018, the IBM Summit hosted at Oak Ridge National Laboratory already has a power efficiency over 10GFLOPS/W. If we take into account the trend in Fig. 3, we can expect future systems reaching between 33 GFLOPS and 50 GFLOPS per W, which would mean building a 20 MW to 30 MW system capable of one ExaFLOPS performance.

There is evidence that further VLSI technology scaling is the way for achieving higher power efficiency (and not higher clock frequency) by allowing low voltage operation at similar propagation delay obtained with presently available technology nodes. The promising technologies for achieving this shift in power efficiency are 7 nm or 5 nm FinFET processes, getting close to 5 ps gate delay at 0.5 V supply voltage. In this perspective, it is not realistic for Europe to renounce leveraging non-European top-performance silicon technology suppliers, which in the present scenario appear to be Samsung and TSMC. In addition to processing devices, memory technology plays a primary role in reducing the Memory Wall impact, by the use of recently developed 3D-stacked DRAM solutions like High Bandwidth Memory (HBM). Such memory technologies provide a significant increase in memory bandwidth available to the processor, which is one of the main bottlenecks for future supercomputers. Notably, the opportunity of stacking the memory directly on the CPU is subject to further analyses when we consider 7 nm or 5 nm processes, due to the power density consequent to geometry scaling.

On the architecture side, the European Union's encouragement for a large range of applications, involving linear algebra computing, favours re-thinking dedicated solutions for vector processing. The present short-term effort spent by part of the EPI project (European Processor Initiative 2018) consortium in developing a vector accelerator tightly coupled to a RISC-V scalar core (see next section) goes in this direction, targeting pre-Exascale performance.
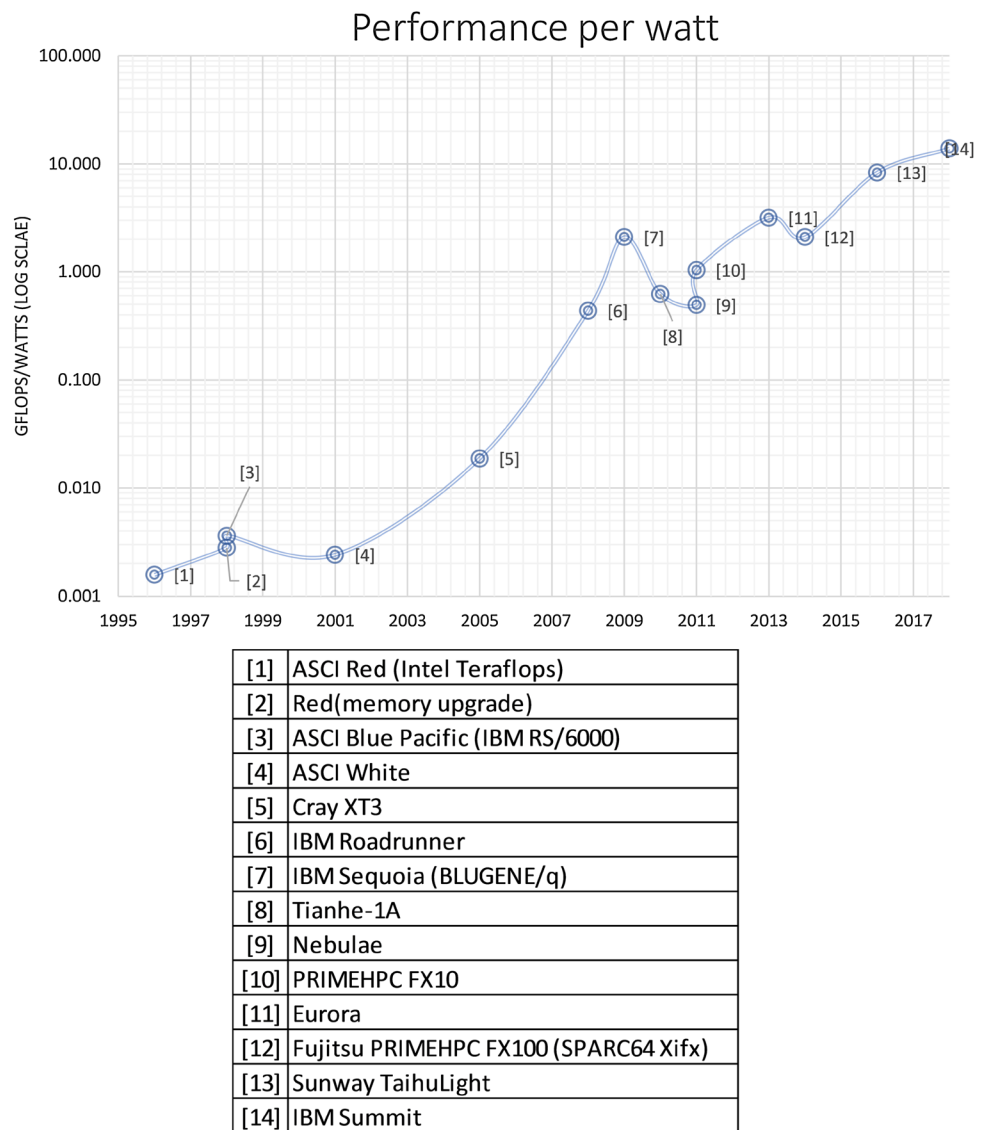
In a medium-term perspective, novel solutions for hardware and compiler support for reducing the register access overhead of traditional vector processing are a credible way for reaching the 33 to 50 GFLOPS/W target, in 7–5 nm finFET technology node.

5 The rise of an open instruction set architecture

The history of HPC shows that in order for supercomputer microprocessor to survive on the market, it is essential that they can share part of the high volume market of general purpose microprocessor. At the same time, the advent of Artificial Intelligence applications in the embedded system market has raised the demand for high performance computing capabilities in embedded processors. This demand is presently coming from cloud based computing services for embedded applications and will involve more and more edge computing devices featuring high computing power close to the final user or even to Internet-of-Things nodes. This reality opens a particularly favorable scenario for sharing horizontally the same computing platforms (instruction set architectures) in different contexts, ranging from the embedded market to the HPC market.

In this view, the appearance of the RISC-V instruction set (Patterson 2018; Waterman et al. 2016) on the embedded system scene is of particular interest as it allows processor designers to join the high volumes of embedded solutions with the advantages of an open instruction set and software technology.

RISC-V originated in 2010 from a research project at the University of California, Berkeley and is now supported by the RISC-V Foundation counting over 100 partners, among which numerous major industrial actors in the ICT market (http://riscv.org). It is composed of a base instruction set—divided into user and privilege set—that has been finalized and will never change, extended in a modular fashion by a

Fig. 3 Energy efficiency progression in HPC system

[1]	ASCI Red (Intel Teraflops)
[2]	Red(memory upgrade)
[3]	ASCI Blue Pacific (IBM RS/6000)
[4]	ASCI White
[5]	Cray XT3
[6]	IBM Roadrunner
[7]	IBM Sequoia (BLUGENE/q)
[8]	Tianhe-1A
[9]	Nebulae
[10]	PRIMEHPC FX10
[11]	Eurora
[12]	Fujitsu PRIMEHPC FX100 (SPARC64 Xifx)
[13]	Sunway TaihuLight
[14]	IBM Summit

number of dedicated instruction sets targeting higher performance or specialized application domains. The present status of the instruction set extensions is summarized in Table 2. Interestingly, there is a vector processing extension being defined, already at advanced stage of maturity.

The RISC-V initiative inherits the long history of RISC processors started in the 80s from the research initiatives led by John Hennessy and David Patterson. In 2018, they received the ACM Turing award "for pioneering a systematic, quantitative approach to the design and evaluation of computer architectures with enduring impact on the microprocessor industry".

On the political side, the openness of the platform is of special interest in Europe as it may guarantee a total independence from US-based or Far East based intellectual property. In fact, while the RISC-V foundation is established in the US, there are already plenty of open source

as well as commercial products based on RISC-V, making sure that RISC-V instruction set will remain always open in the future.

At present, while several RISC-V commercial and open products are ready for the embedded market (Cheikh et al. 2019; Patsidis et al. 2018; Flamand et al. 2018; Schiavone et al. 2017; Keller et al. 2017; Gautschi et al. 2017; Olivieri et al. 2017), there is limited availability of RISC-V HPC processors, either implementing vector computation (Lee et al. 2014, 2016; Zimmer et al. 2016) or targeting general purpose HPC applications. The 45 nm vector-accelerated Hwacha chip (Lee et al. 2014) achieves 16.7 Double-Precision GFLOPS/W, while the 28 nm FDSOI Hwacha chip (Zimmer et al. 2016) employs aggressive voltage biasing techniques to achieve power efficiency in variable load conditions. The U540 Quad-Core Processor [ROS18] addresses general purpose performance and exhibits a 6 GFLOPS theoretical peak

Table 2 The RISC-V user instruction set extensions as of July 2018

Name	Description	Version	Status
RV32I	Base Integer Instructions, 32 bit	2.0	Final
RV32E	Base Integer Instructions, 32 bit, embedded	1.9	Open
RV64I	Base Integer Instructions, 64 bit	2.0	Final
RV128I	Base Integer Instructions, 128 bit	1.7	Open
Q	Standard Extension Quad-precision Floating Point	2.0	Final
L	Standard Extension Decimal Floating Point	0.0	Open
C	Standard Extension Compressed Instructions	2.0	Final
B	Standard Extension Bit Manipulation	0.36	Open
M	Standard Extension Integer Multiply and Divide	2.0	Final
A	Standard Extension Atomic Instructions	2.0	Final
F	Standard Extension Single-precision Floating Point	2.0	Final
D	Standard Extension Double-precision Floating Point	2.0	Final
J	Standard Extension Dynamically Translated Languages	0.0	Open
T	Standard Extension Transactional Memory	0.0	Open
P	Standard Extension Packed SIMD Operations	0.1	Open
V	Standard Extension Vector Operations	0.2	Open
N	Standard Extension User Level Interrupts	1.1	Open

at a 4.63 W board power consumption, showing the need for accelerator support to target breakthrough power efficiency.

The EPI (European Processor Initiative) project (European Processor Initiative 2018), where the Barcelona Supercomputing Center is a main partner along with other 22, is the first significant effort in Europe explicitly addressing the goal of an HPC RISC-V microprocessor. One RISC-V tile of the EPI System-on-Chip, to be manufactured in 7 nm, targets a peak performance of at least 192 GFLOPS and a power efficiency over 30 GFLOPS/W, featuring a multi-lane vector acceleration engine addressing very long vector size.

Looking further at a successive chip generation, we estimate that with a tile architecture based on superscalar cores and 1024-lane EPI-like RISC-V vector accelerators running at 2.3 GHz, employing SRAM near memory and HBM3 far memory, a system featuring 64 tiles per node and 4096 nodes may reach Exascale performance with an effective power (processing and communication) below 16 MW and total system power (including power supply chain and cooling) below 24 MW.

Notably, the adoption of RISC-V is a key factor for having the maximum freedom in the development of the necessary acceleration technology, without limitations coming from hardware and software commercial eco-systems. Yet, the potential difficulty of adopting RISC-V in HPC relies in its intrinsic novelty. Due to the need of changing an entire software and hardware ecosystem, presently based on $\times 86$ and GPU accelerators, it is unlikely that a RISC-V based Exascale machine will be ready in the short term. However, the possibility of reaching such goal in the mid term is realistic. In fact, BSC already experienced all these difficulties when starting the Mont-Blanc project in 2011. When the

project began, no HPC software stack existed for ARM-based systems. After 7 years of significant efforts made by both industry and academia, we can state that the ARM HPC ecosystem is by now mature. The success of adopting RISC-V in HPC will depend on convincing more and more partners to contribute to this open source HPC initiative.

6 EuroHPC, EPI and EOSC

In the last decades, the European Union has invested a significant amount of money on developing cutting edge HPC technology. The Centres of Excellence for HPC, the strategic research agendas developed by ETP4HPC think tank, the access to world class supercomputers in PRACE (Partnership for Advanced Computing in Europe), together with multiple European projects led by European industries, research centres and universities has positioned Europe as a world leader in HPC. However, there are no local European vendors capable of supplying EU customers with HPC hardware based on domestic supercomputing technology. As a result, Europe is currently purchasing all its HPC processors from non-European companies.

At the major European funding agency event in 2015 in Lisbon, ICT2015, the EU Commission Director General, Roberto Viola, conveyed three top European scientists to discuss the impact of science on society. There, the director of BSC, Mateo Valero proposed to start a European plan for the Exascale race. DG Viola enthusiastically supported this idea and in the following weeks, the leadership of the Commission all the way to the President Jean Claude Junker confirmed the firm intention of Europe to

enter in the competition with the rest of the world to reach Exascale performances based on European technology by 2023.

At the Rome celebration of the 60th anniversary of the Rome Treaty, which established the framework for what later became the European Union, on March 23rd 2017 the **EuroHPC initiative** was lunched with the signature of the first 7 participating member states [The European declaration on high-performance computing (EuroHPC 2018). Up to date other 18 European countries have joined for a total of 25 signatories. In an unusually fast pace the Commission has managed to establish a specific instrument (Joint Undertaking or JU) for EuroHPC which has been officially approved by the European Council of Ministers on September 28th 2017. This JU will be funded with a first investment of 1 billion euros (B€) till 2020, to be followed by an additional investment of 2.7 B€ in the next funding phase of 2021–2027. This level of funding is in line with what the other developed regions of the world have announced.

As a first action a specific funding grant of 120 million euros (M€) has been allocated to a consortium of 23 partners to develop in the next 3–4 years a European processor. The initiative has been named **European Processor Initiative (EPI)** (2018) and accepted for funding by the Commission on March 23rd. Three lines of development have been so far proposed: (i) a general purpose processor and common platform that will leverage existing technology to deliver high performance systems in the short term; (ii) an accelerator based on RISC-V ISA that will develop a complete European design in the mid term; (iii) an automotive processor that will provide a high performance system suitable for future autonomous cars satisfying real-time constraints. The Barcelona Supercomputing Center will contribute to the three designs aiming at developing European high performance systems for different markets, including HPC, deep learning and automotive. BSC will leverage the experience from the Mont-Blanc and other European projects to deliver a high performance general purpose processor. Moreover, BSC is leading the development of the RISC-V accelerator together with other academic and industrial partners.

Another important aspect of the European HPC strategy is the **European Open Science Cloud (EOSC)** as a vehicle to provide access to the entire European Scientific Community especially to the less affluent normally referred to as the *long tail of research.* A considerable amount of funding, in the order of 300 M € for the reminder of the H2020 programme (till 2020), has been allocated to develop a strategy and implement a roadmap for scientific data sharing. The official launch of the European Open Science Cloud took place in Vienna on November 23rd, 2018 (European Open Science Cloud, EOSC 2018).

7 A look at the future and conclusions

Europe has a long way to go to obtain complete independence in the HPC market, traditionally dominated by US and Japan. The tremendous progress by Chinese scientists and industry in the last decade clearly demonstrates that with adequate resources and well-focussed planning it is possible to fill the gap and count in the top players. Clearly, Europe lacks a large domestic industry in HPC, but a significant part of the HPC ecosystem already exists and on the research side several centres have the necessary competences to develop processors and the other components necessary to commission large Exascale systems.

The economic landscape is complicated by the presence of international global vendors that have traditionally dominated the European HPC market. Differently from the US, Chinese and Japanese markets, essentially closed to foreign vendors, the European one is completely open. It will not be easy for the Commission to privilege an emerging domestic HPC industry in the European procurements. The fragmentation of the EU in at least 27 member states, assuming the exit of the UK, is another fundamental problem. Any major initiative needs the approval of the majority of the member states making the investments often not very effective. In any case, for the first time in many years the strategic importance of Europe gaining a leading place in the worldwide Exascale race has been recognized. It is now up to all the stakeholders from the public and private sectors to make this wish a reality. The Barcelona Supercomputing Center will make its best for this vision to become a reality.

References

Asanovic, K. (1998) Vector microprocessors. Ph.D. thesis, (1998)

Barnes, G.H., Brown, R.M., Kato, M., Kuck, D.J., Slotnick, D.L., Stokes, R.A.: The ILLIAC IV Computer. IEEE Trans. Comput. **C-17**(8), 746–757 (1968)

Callahan D, Dongarra J, Levine D (1988) Vectorizing compilers: a test suite and results. Proceedings of the 1988 ACM/IEEE Conference on Supercomputing, pp. 98–105, November 12–17, 1988, Orlando, Florida, United States.SC (1988)

Cheikh, A., Cerutti, G., Mastrandrea, A., Menichelli, F., Olivieri, M.: The microarchitecture of a multi-threaded RISC-V compliant processing core family for IoT end-nodes. Lecture Notes Elect. Eng. **512**, 89–97 (2019)

Cray Research: Cray X-MP Series Model 48 Mainframe Reference Manual. (1984)

Esmaeilzadeh, H., Blem, E., Amant, R.S., Amant, R.S., Sankaralingam, K., Burger, D.: Dark silicon and the end of multicore scaling. IEEE Micro **32**(3), 122–134 (2012)

Espasa, R., Valero, M., Smith, J. E.: Vector architectures: past, present and future. In Proceedings of the 12th International Conference on Supercomputing (ICS), pp. 425–432, (1998)

European Open Science Cloud (EOSC): https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud. (2018)

European Processor Initiative (2018): Consortium to develop Europe's microprocessors for future supercomputers. (https://ec.europa.eu/digital-single-market/en/news/european-processor-initiative-consortium-develop-europes-microprocessors-future-supercomputers). (2018)

Flamand, E., Rossi, D., Conti, F., Loi, I., Pullini, A., Rotenberg, F., Benini, L.: GAP-8: a RISC-V SoC for AI at the edge of the IoT. In: Proceedings of the International Conference on Application-Specific Systems, Architectures and Processors. IEEE, Piscataway (2018)

Gautschi, M., Schiavone, P.D., Traber, A., Loi, I., Pullini, A., Rossi, D., Flamand, E., Gürkaynak, F.K., Benini, L.: Near-threshold RISC-V core with DSP extensions for scalable IoT endpoint devices. IEEE Trans Very Large Scale Integr (VLSI) Syst **25**(10), 2700–2713 (2017)

Hintz, R.G., Tate, D.P.: Control Data STAR-100 processor design. In Proc. COMPCON, pp. 1–4. IEEE, Piscataway (1972)

Keller, B., Cochet, M., Zimmer, B., Kwak, J., Puggelli, A., Lee, Y., Blagojević, M., Bailey, S., Chiu, P.-F., Dabbelt, P., Schmidt, C., Alon, E., Asanović, K., Nikolić, B.: A RISC-V processor SoC with integrated power management at submicrosecond timescales in 28 nm FD-SOI. IEEE J. Solid-State Circuits **52**(7), 1863–1875 (2017)

Lee, Y. et al.: A 45 nm 1.3 GHz 16.7 double-precision GFLOPS/W RISC-V processor with vector accelerators. In: ESSCIRC 2014—40th European Solid State Circuits Conference (ESSCIRC), Venice Lido, pp. 199–202, (2014)

Lee, Y., Zimmer, B., Waterman, A., Puggelli, A., Kwak, J., Jevtic, R., Keller, B., Bailey, S., Blagojevic, M., Chiu, P.-F., Cook, H., Avizienis, R., Richards, B., Alon, E., Nikolic, B., Asanovic, K.: Raven: A 28 nm RISC-V vector processor with integrated switched-capacitor DC–DC converters and adaptive clocking, (2016). In: 2015 IEEE Hot Chips 27 Symposium, HCS (2015)

Lincoln, N.R.: Technology and design trade offs in the creation of a modern supercomputer. IEEE Trans. on Comput. **C-31**(5), 349–362 (1982)

Olivieri, M., Cheikh, A., Cerutti, G., Mastrandrea, A., Menichelli, F (2017) Investigation on the optimal pipeline organization in RISC-V multi-threaded soft processor cores. In: Proceedings—2017 1st New generation of CAS, NGCAS 2017, pp. 45–48 (2017)

Patsidis, K., Konstantinou, D., Nicopoulos, C., Dimitrakopoulos, G.: A low-cost synthesizable RISC-V dual-issue processor core leveraging the compressed Instruction Set Extension. Microprocess. Microsyst. **61**, 1–10 (2018)

Patterson, D.: 50 years of computer architecture: from the mainframe CPU to the domain-specific tpu and the open RISC-V instruction set. In: Digest of Technical Papers—IEEE International Solid-State Circuits Conference, 61, pp. 27–31 (2018)

Rajovic, N., Rico, A., Mantovani, F., Ruiz, D., Vilarrubi, J.O., Gomez, C., Backes, L., Nieto, D., Servat, H., Martorell, X., Labarta, J., Ayguadé, E., Adeniyi-Jones, C., Derradji S., Gloaguen, H., Lanucara, P., Sanna, N., Méhaut, J.-F., Pouget, K., Videau, B., Boyer, E., Allalen, M., Auweter, A., Brayford, D., Tafani, D., Weinberg, V., Brömmel, D., Halver, R., Meinke, J.H., Beivide, R., Benito, M., Vallejo, E., Valero, M., Ramírez, A.: The Mont-Blanc prototype: an alternative approach for HPC systems. In: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, pp. 444–455, IEEE, Piscataway (2016)

RISC-V Foundation home page. Online: http://riscv.org

Russell, R.M.: The CRAY-1 computer system. Commun. ACM **21**(1), 63–72 (1978)

Schiavone, P.D., Conti, F., Rossi, D., Gautschi, M., Pullini, A., Flamand, E., Benini, L.: Slow and steady wins the race? A comparison of ultra-low-power RISC-V cores for internet-of-things applications. In: 2017 27th International Symposium on Power and Timing Modeling, Optimization and Simulation, PATMOS 2017, pp. 1–8 (2017)

The European declaration on High-Performance Computing (EuroHPC): https://ec.europa.eu/digital-single-market/en/news/european-declaration-high-performance-computing). (2018)

The TOP500 List. Online: https://www.top500.org/, June 2018

Watson, W.J.: The TI ASC: A highly modular and flexible super computer architecture. In Proceedings of the December 5–7, 1972, Fall Joint Computer Conference, Part I (AFIPS), pp. 221–228, (1972)

Waterman, A., Lee, Y., Avizienis, R., Cook, H., Patterson, D., Asanovic, K.: The RISC-V instruction set. In: 2013 IEEE Hot Chips 25 Symposium, HCS (2016)

Yoshida, T: Introduction of Fujitsu's HPC processor for the Post-K computer. In: Hot Chips. (2016)

Zimmer, B., Lee, Y., Puggelli, A., Kwak, J., Jevtić, R., Keller, B., Bailey, S., Blagojević, M., Chiu, P.-F., Le, H.-P., Chen, P.-H., Sutardja, N., Avizienis, R., Waterman, A., Richards, B., Flatresse, P., Alon, E., Asanović, K., Nikolić, B.: A RISC-V vector processor with simultaneous-switching switched-CAPACITOR DC–DC converters in 28 nm FDSOI. IEEE J. Solid-State Circuits **51**(4), 930–942 (2016)

Dr. Fabrizio Gagliardi spent the first 30 years of his career (1975–2005) at the European Centre for Particle Physics, in Geneva, Switzerland, and 8 years (2005–2013) at Microsoft and Microsoft Research in the US, Switzerland and in the UK. He is now Distinguished Research Director at the Polytechnic University of Catalonia, Spain. In this position he advises the Barcelona Supercomputing Centre leadership on their long term international strategy. He also leads BSC participation in a major EU supported projects (RDA-EU 4) and plays a major role in the EuroHPCinitiative to develop a European high performance supercomputer to list in the top 3 supercomputers in the world by 2023. He also leads the organisation of the 2019 ACM HPC summer school in Barcelona, as part of his role in the ACM Europe Policy board. Formerly he was visiting professor, at the Gran Sasso Scientific Institute, L'Aquila, Italy and Independent computing consultant, in Geneva, Switzerland. In his long career, Fabrizio held senior scientific and managerial positions in prestigious scientific research organisations (CERN in CH, Stanford Accelerator in the US) and consulted with scientific organisations such as Fermilab NationalLaboratory and NSF in the US, INFN in Italy, STFC in the UK). His research interests focus on exploring new advanced computing technologies which could better support science: AI, real-time data acquisition systems, data management system design, distributed systems, grid and cloud technology, High Performance Computing and computer architecture. He was

awarded a number of prizes, most notably the ACM President Award in 2013 and again in 2018.

Miquel Moreto is a Ramon y Cajar Researcher at UPC and an associate researcher at BSC. Prior to joining UPC, he spent 15 months as a post-doctoral fellow at the International Computer Science Institute (ICSI), Berkeley, USA. He received the B.Sc., M.Sc., and Ph.D. degrees from UPC. His research interests include high performance computer architectures and hardware-software co-design for future massively parallel systems.

Miquel Moreto (IEEE M'98 SM'16) received the Master (Laurea) degree in electronics engineering "cum laude" in 1991 and the Doctorate degree in electronics and computer engineering in 1994 from the University of Genoa, Italy, where he was assistant professor from 1995 to 1998. In 1998 he joined Sapienza University of Rome as associate professor, teaching digital electronics and VLSI system architectures. He is a visiting researcher at the Barcelona Supercomputing Center within the European Processor Initiative. His research interests are digital system-onchip design, microprocessor core design and digital nano-scale CMOS circuits. He authored over 100 papers and a textbook in three volumes. He has been a member of the TPC of IEEE DATE and was General Co-Chair of IEEE/ACM ISLPED'15. He is an evaluator for the European Commission in the ECSEL Joint Undertaking.

Mateo Valero Director of the Barcelona Supercomputing Center. His research focuses on high performance architectures. Prof. Valero has been honoured with several awards, among them the 3 most relevant awards in Computer Architecture field : the Eckert-Mauchly Award 2007 by the IEEE and ACM; Seymour Cray Award 2015 by IEEE; Charles Babbage 2017 by IEEE; Harry Goode Award 2009 by IEEE:; the Spanish National awards "Julio Rey Pastor" and "Leonardo Torres Quevedo". "Hall of the Fame" member of the ICT European Program (selected as one of the 25 most influents European researchers in IT during the period 1983-2008. Lyon,November 2008); Honoured with "Condecoración de la Orden Mexicana del Águila Azteca" 2018, highest recognition granted by the Mexican Government. He is Honorary Doctorate by 9 Universities. He is member of 9 academies. He is a fellow of IEEE and ACM and he is also Intel Distinguished Research Fellow.