Towards EXtreme scale technologies and accelerators for euROhpc hw/Sw supercomputing applications for exascale: The TEXTAROSSA approach☆

Giovanni Agosta [a], Marco Aldinucci [f], Carlos Alvarez [h], Roberto Ammendola [n], Yasir Arfat [f], Olivier Beaumont [g], Massimo Bernaschi [c], Andrea Biagioni [j], Tommaso Boccali [l], Berenger Bramas [g], Carlo Brandolese [a], Barbara Cantalupo [f], Mauro Carrozzo [c], Daniele Cattaneo [a], Alessandro Celestini [c], Massimo Celino [b], Iacopo Colonnelli [f], Paolo Cretaro [j], Pasqua D'Ambra [d], Marco Danelutto [e], Roberto Esposito [f], Lionel Eyraud-Dubois [g], Antonio Filgueras [h], William Fornaciari [a], Ottorino Frezza [j], Andrea Galimberti [a], Francesco Giacomini [k], Brice Goglin [g], Daniele Gregori [o], Abdou Guermouche [g], Francesco Iannone [b], Michal Kulczewski [i], Francesca Lo Cicero [j], Alessandro Lonardo [j], Alberto R. Martinelli [f], Michele Martinelli [j], Xavier Martorell [h], Giuseppe Massari [a], Simone Montangero [m], Gianluca Mittone [f], Raymond Namyst [g], Ariel Oleksiak [i], Paolo Palazzari [b], Pier Stanislao Paolucci [j], Federico Reghenzani [a,*], Cristian Rossi [j], Sergio Saponara [e], Francesco Simula [j], Federico Terraneo [a], Samuel Thibault [g], Massimo Torquati [e], Matteo Turisini [j], Piero Vicini [j], Miquel Vidal [h], Davide Zoni [a], Giuseppe Zummo [b]

a *Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Italy*
b *ENEA, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile, Italy*
c *Istituto per le Applicazioni del Calcolo (IAC), Consiglio Nazionale delle Ricerche, Rome, Italy*
d *Istituto per le Applicazioni del Calcolo (IAC), Consiglio Nazionale delle Ricerche, Naples, Italy*
e *University of Pisa, Italy*
f *University of Torino, Italy*
g *Institut national de recherche en informatique et en automatique (Inria), France*
h *Barcelona Supercomputing Center, Spain*
i *Poznan Supercomputing and Networking Center, Poland*
j *Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Italy*
k *Istituto Nazionale di Fisica Nucleare, CNAF, Italy*
l *Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Italy*
m *University of Padova and Istituto Nazionale di Fisica Nucleare Sezione di Padova, Italy*
n *Istituto Nazionale di Fisica Nucleare, Sezione di Roma Tor Vergata, Italy*
o *E4 Computer Engineering Spa, Italy*

☆ This manuscript is an extended version of Agosta et al. (2021) [1].
* Corresponding author.
E-mail addresses: giovanni.agosta@polimi.it (G. Agosta), marco.aldinucci@unito.it (M. Aldinucci), carlos.alvarez@bsc.es (C. Alvarez), roberto.ammendola@roma2.infn.it (R. Ammendola), yasir.arfat@unito.it (Y. Arfat), olivier.beaumont@inria.fr (O. Beaumont), massimo.bernaschi@cnr.it (M. Bernaschi), andrea.biagioni@roma1.infn.it (A. Biagioni), tommaso.boccali@pi.infn.it (T. Boccali), berenger.bramas@inria.fr (B. Bramas), carlo.brandolese@polimi.it (C. Brandolese), barbara.cantalupo@unito.it (B. Cantalupo), mauro.carrozzo@cnr.it (M. Carrozzo), daniele.cattaneo@polimi.it (D. Cattaneo), alessandro.celestini@cnr.it (A. Celestini), massimo.celino@enea.it (M. Celino), iacopo.colonnelli@unito.it (I. Colonnelli), paolo.cretaro@roma1.infn.it (P. Cretaro), pasqua.dambra@cnr.it (P. D'Ambra), marco.danelutto@unipi.it (M. Danelutto), roberto.esposito@unito.it (R. Esposito), lionel.eyraud-dubois@inria.fr (L. Eyraud-Dubois), antonio.filgueras@bsc.es (A. Filgueras), william.fornaciari@polimi.it (W. Fornaciari), ottorino.frezza@roma1.infn.it (O. Frezza), andrea.galimberti@polimi.it (A. Galimberti), francesco.giacomini@cnaf.infn.it (F. Giacomini), brice.goglin@inria.fr (B. Goglin), daniele.gregori@e4company.com (D. Gregori), abdou.guermouche@inria.fr (A. Guermouche), francesco.iannone@enea.it (F. Iannone), kulka@man.poznan.pl (M. Kulczewski), francesca.locicero@roma1.infn.it (F. Lo Cicero), alessandro.lonardo@roma1.infn.it (A. Lonardo), albertor.martinelli@unito.it (A.R. Martinelli), michele.martinelli@roma1.infn.it (M. Martinelli), xavier.martorell@bsc.es (X. Martorell), giuseppe.massari@polimi.it (G. Massari), simone.montangero@pd.infn.it (S. Montangero), gianluca.mittone@unito.it (G. Mittone), raymond.namyst@inria.fr (R. Namyst), ariel@man.poznan.pl (A. Oleksiak), paolo.palazzari@enea.it (P. Palazzari), pierstanislao.paolucci@roma1.infn.it (P.S. Paolucci), federico.reghenzani@polimi.it (F. Reghenzani), cristian.rossi@roma1.infn.it (C. Rossi), sergio.saponara@unipi.it (S. Saponara), francesco.simula@roma1.infn.it (F. Simula), federico.terraneo@polimi.it (F. Terraneo), samuel.thibault@inria.fr (S. Thibault), massimo.torquati@unipi.it (M. Torquati), matteo.turisini@roma1.infn.it (M. Turisini), piero.vicini@roma1.infn.it (P. Vicini), miquel.vidal@bsc.es (M. Vidal), davide.zoni@polimi.it (D. Zoni), giuseppe.zummo@enea.it (G. Zummo).

ABSTRACT

In the near future, Exascale systems will need to bridge three technology gaps to achieve high performance while remaining under tight power constraints: energy efficiency and thermal control; extreme computation efficiency via HW acceleration and new arithmetic; methods and tools for seamless integration of reconfigurable accelerators in heterogeneous HPC multi-node platforms. TEXTAROSSA addresses these gaps through a co-design approach to heterogeneous HPC solutions, supported by the integration and extension of HW and SW IPs, programming models, and tools derived from European research.

1. Introduction

High Performance Computing (HPC) is a strategic asset for countries and large companies alike. Such infrastructures are of key importance to support a variety of applications in domains such as oil & gas, finance, and weather forecasting. Recently, emerging domains have been gaining traction, such as bioinformatics, medicine, security and surveillance. These newer applications tend to fall in the classes of High Performance Data Analytics (HPDA) and High Performance Computing for Artificial Intelligence (HPC-AI). The trend in the design of such infrastructures is more and more exploiting heterogeneous hardware architectures to cope with the request of peak performance and to meet the need of achieving a "Green HPC". These paths have prompted Europe to align its research priorities in HPC along a Strategic Research Agenda (SRA[1]) resulting from wide consultations within the European Technology Platform for HPC (ETP4HPC), the PRACE initiative,[2] and the PlanetHPC[3] initiative.

The need to achieve high efficiency, while remaining within reasonable power and energy bounds, is extensively discussed in the SRA, focusing also on the main technology challenges posed by these objectives. Such challenging goal can only be addressed with an holistic approach that takes into account multiple factors across the HPC hardware/software stack, including the use of extremely efficient application-specific hardware accelerators, software management of resources, applications, and cooling systems. Together, these components can provide the desired computational power while keeping under control the power consumption of the supercomputer.

The TEXTAROSSA project [1] originates from such a vision and aims at providing key technological advances on all three aspects and to validate them on new development platforms representative of future HPC systems, using a wide range of applications, selected considering traditional HPC application fields as well as coming from emerging domains. Fig. 1 highlights the strategic goals of TEXTAROSSA, set against the EuroHPC initiative and the broader framework of European research in High Performance Computing. More in detail, TEXTAROSSA will implement the above-mentioned approach by pursuing the following goals:

- **Technical goals**

  1. *Energy efficiency and thermal control*. A two-phase cooling technology deployed at both node and rack levels will be fully integrated with a multi-level runtime resource manager. The on-board sensor data feeds the power, energy, and thermal models.
  2. *Sustained application performance*. The applications and underlying ecosystem (libraries, programming models, …) are optimized through an efficient exploitation of parallel accelerators (GPUs and FPGAs) and by exploiting data/stream locality, efficient algorithms, and innovative IPs.

  3. *Seamless integration of reconfigurable accelerators*. The tools necessary for the design and implementation phases, such as Vitis and OmpSs@FPGA, will be extended to support new IPs and methodologies, including power/energy monitoring and control, in a mixed-precision computing context.
  4. *Development of new IPs*. Mixed-precision computing is applied to several fields, including artificial intelligence, data compression, security, power monitoring, and workload scheduling.
  5. *Integrated Development Platforms*. Two heterogeneous Integrated Development Vehicles (IDVs) are in development, one as a dedicated testbed for two-phase cooling technology, and one to support a wider range of technical goals.

- **Strategic goals**

  1. *Alignment with the European Processor Initiative (EPI)* by testing, extending and boosting key technologies applicable to future EPI evolution.
  2. *Supporting the objectives of EuroHPC* as reported in ETP4HPC's Strategic Research Agenda (SRA) for open HW and SW architecture.
  3. *Building over European expertise* gained through past research projects as well as through the Centers of Excellence in HPC.
  4. *Opening of new usage domains*, including High Performance Data Analytics (HPDA) and High Performance Artificial Intelligence (HPC-AI) applications, alongside support for traditional HPC domains.

1.1. The TEXTAROSSA consortium

TEXTAROSSA is a three-year project co-funded by the European High Performance Computing (EuroHPC) Joint Undertaking. The project is led by ENEA (Italy) and aggregates 17 institutions and companies, including the linked third parties, located in 5 European countries: CINI, an Italian consortium grouping together three leading universities, Politecnico di Milano, Università degli studi di Torino, and Università di Pisa, Fraunhofer (Germany), INRIA (France), ATOS (France), E4 Computer Engineering (Italy), BSC (Spain), PSNC (Poland), INFN (Italy), CNR (Italy), In Quattro (Italy), Université de Bordeaux (France), CINECA (Italy) and Universitat Politècnica de Catalunya (UPC). The three Italian universities are part of the lab of CINI,[4] created in 2021, that is grouping together the main academic and research entities working in the field of high-performance and Exascale computing in Italy. CINI is providing also the technical

---

1 https://www.etp4hpc.eu/sra.html (last accessed March 2022)
2 https://prace-ri.eu (last accessed March 2022)
3 https://cordis.europa.eu/project/id/248749 (last accessed March 2022)

4 https://www.consorzio-cini.it/index.php/it/laboratori-nazionali/hpc-key-technologies-and-tools (last accessed March 2022)
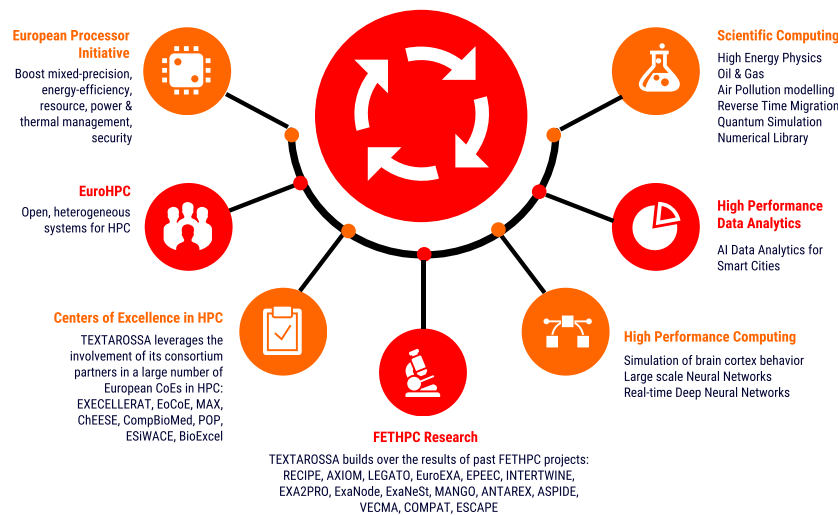
Fig. 1. Impact strategy of TEXTAROSSA and positioning within the landscape of the European initiatives on HPC.

leadership of the project. More information on the activities carried out during the execution of TEXTAROSSA can be found in the project website.[5]

1.2. Organization of the paper

The paper is structured as follows: Section 2 introduces the TEXTAROSSA co-design approach, and the key technological innovations are described in Section 3. Section 4 provides an overview of the industrial application and use cases. Section 5 concludes the article by highlighting future research directions.

2. TEXTAROSSA co-design approach

From a methodological point of view, TEXTAROSSA adopts a co-design process as key strategy for Fast Forward and Exascale computing, considering the entire system stack from underlying technologies to applications. The co-design process includes five layers covering the whole HPC stack:

1. *User Application*: representing a wide range of scenarios, from mathematical libraries, to mini applications and flagship codes for numerical modeling with massive parallelism in HPC/HPDA/AI applications. The performance of HPC applications depends on the level of optimization of them and also of the underlying tools. Research usually focuses on both facets separately: optimizing the applications on one side and improving hardware/middleware layers on the other side. This is not the case with our co-design approach: We aim at connecting both aspects by evaluating the impact of novel technologies on a set of target HPC applications, and, at the same time, studying which improvements at lower levels could be beneficial to these applications.

2. *Runtime Services*: ensuring that application requirements are dynamically satisfied and mapped onto system resources, including execution models with workload handling, fault tolerance, and data management. These services are well established in traditional HPC data centers, and the users, mainly belonging to academic and big enterprises research, are aware enough to use the computing resources. Because there has been a growth of edge applications in the last years aimed at effectively analyzing big data in a timely manner, a cloud edge continuum enabling

HPC/HPDA is becoming a business case. As the levels and fidelity of instrumentation increase and the types and volumes of available data grow, new classes of applications are being explored that seamlessly combine real-time data with complex programming models and data analytics to monitor and manage systems of interest.

3. *Programming Models*: underlying the applications, they define the toolchains and SW development tools able to implement applications in parallel architectures. The HPC community provided several programming models that have demonstrated their potential to develop efficient applications. However, parallelization models are used separately, and most of them target CPUs or combinations of CPUs/GPUs, but not FPGAs. However, FPGAs open new possibilities that can be highly beneficial to almost any HPC application, providing toolchains to handle heterogeneous architectures.

4. *System Architecture*: including the processor core's micro-architecture, the arrangement of cores within a chip, memory hierarchy, system interconnect, and storage subsystems. Future HPC platforms increasingly depend on heterogeneous node architectures to meet power and performance requirements. In the HPC landscape, two main approaches have appeared as viable solutions for a possible system architecture bridging the current gaps in terms of power and performance that are required in Exascale computing: the first approach relies on multi-core processors whose high performance is boosted by the use of GPU-based accelerators; the second approach aims at integrating FPGA-based accelerator within the host architecture. Additionally, interconnection networks featuring very low latency are going to be essential to support the high performance of the computation nodes.

5. *Platforms*: concerning the HW platform at node and rack levels, they need to be able to achieve performance requirements in terms of computing power and energy consumption. High-density computing power at node/rack level requires new technologies of direct cooling on the chip able to remove heat with high efficiency in order to reduce energy consumption. Direct cooling provides a more efficient method to transfer the heat from these hot components to the building chilled water loop and then outside with very little additional energy, compared to transferring the heat first to air and then to the building chilled water system. In addition, in a direct cooling system, the water temperature returning after cooling the IT equipment is much higher than typically found in data centers, and provides more opportunity for heat reuse or the ability to reject this heat
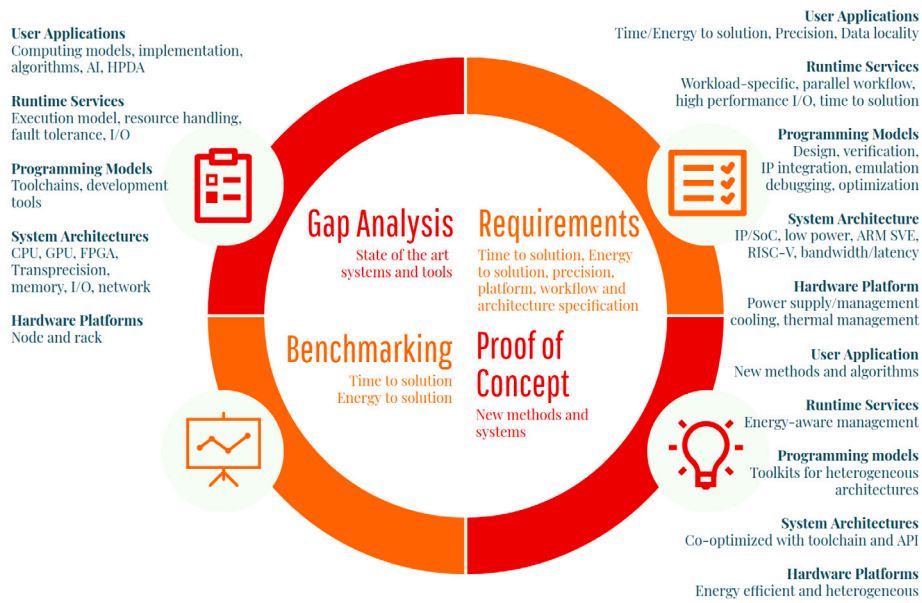

Fig. 2. The TEXTAROSSA Co-Design Approach.

to the atmosphere with a dry cooler, thereby eliminating the requirement of a cooling tower or chiller plant in most climates.

Fig. 2 provides an overview of the co-design approach adopted in TEXTAROSSA, showing how the five layers of the HPC stack are addressed in each of the four main stages of the co-design process:

1. *Gap Analysis*: to compare the current state-of-art of the technological assets with the objectives of the project in order to identify the gap to be filled by developments or update in co-design process.
2. *Requirements*: to define specifications and requirements of the technological solutions for designing and developing.
3. *Proof of Concept*: to develop HW/SW prototype solutions able to achieve the KPIs (Key Performance Indicators) of the project objectives.
4. *Benchmarking*: to provide performance results of the technological solutions by means of benchmark tools.

3. TEXTAROSSA technologies

TEXTAROSSA develops, starting from the results of previous European research activities mentioned in Fig. 1, a set of technologies to deal with each of the four layers of the HPC stack, as well as applications, which will be covered in full in Section 4. Fig. 3 provides an overview of the primary technology bricks adopted by TEXTAROSSA. In the rest of this section, we provide insights on specific technology bricks developed within the project

3.1. Programming models & toolchains

Vitis based HLS flow. Vitis [2] is the software development platform developed by Xilinx which translates an application written as a C/C++ program into a bitstream able to run on an FPGA card. Vitis also supports the communication between the FPGA and programs running on the host node. In order to facilitate the development process, Vitis allows the developers to emulate the whole system's behavior (host program and FPGA code) by running and debugging the C/C++ code in an user-friendly IDE. After the verification of the functional correctness, i.e., the program produces the expected results, the HLS engine translates the C/C++ code into an equivalent and optimized hardware implementation for the FPGA. The host node, memory banks,
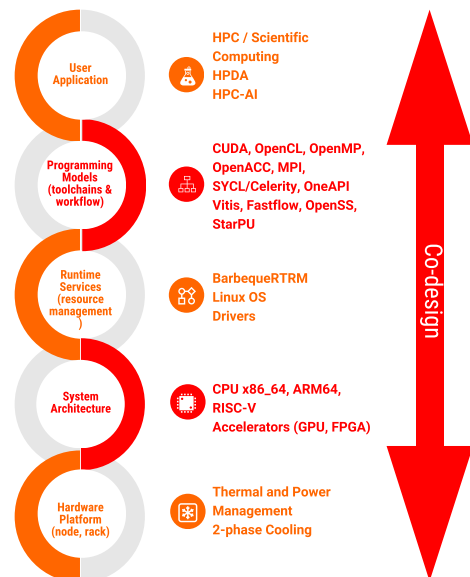

Fig. 3. The TEXTAROSSA key technologies.

and other external world components are accessed from the FGPA via a presynthesized layer containing the interested interfaces, such as PCIe, DDR, HBM, etc. Vice versa, the FPGA resources are accessed by the host node through the runtime and APIs provided by Xilinx. Vitis is very customizable, because based on the open-source LLVM-based front-end[6] and many open-source libraries (e.g., math, video processing, AI, signal processing).[7] In the TEXTAROSSA project, Vitis will be extended by adding multi-precision arithmetic. In this way, we enable the usage of new data types based on the Posit format, and we provide a library used for inter-FPGA direct communications. The inter-node communication layer, namely MPI, will also be updated to transfer data without requiring memory copies between FPGAs and hosts. A new set of APIs, named TEXTAROSSA APIs, will be defined over the previously

---

[6] https://github.com/Xilinx/HLS (last accessed August 2022)
[7] https://xilinx.github.io/Vitis_Libraries (last accessed August 2022)

mentioned APIs, in order to create an homogeneous interface for the accelerators. These extensions will require the definition of new IPs to be integrated into the flow of the Vitis program. Consequently, we will develop the hardware modules, their C++ implementation and we will integrate all the necessary information needed by the HLS tool to properly manage (i.e., schedule and connect) the new functionalities.

StarPU, OmpSs. Task-based programming models have demonstrated their potential to address the programming challenges of heterogeneous computing nodes. Indeed, in current computing nodes, due to the automated management of energy and therefore of the speed of the computing resources and the difficulty to precisely predict the duration of the communications, it is very difficult to build a reliable model that would allow to design an optimal placement and scheduling of computing tasks. On the contrary, task-based systems make scheduling and placement decisions at runtime, relying on low-cost strategies but on a precise vision of the platform state. These dynamic strategies have shown their efficiency, especially in the context of regular applications consisting of a small number of task types, such as linear algebra [3–5]. Moreover, the task-based approaches allow the developers to achieve high productivity by providing higher-level abstractions and still generating high-performance applications. In Textarossa we will improve and study two task-based runtime sytems: OmpSs[8] [6] and StarPU[9] [7]().

Task-based programming models can simplify FPGA programming [8]. This set of approaches has the main advantage of making the memory allocation and data copies automatic and transparent to the users. Additionally, it makes the automation of the code generation of the CPU and FPGA binaries possible. Finally, it can transparently run open or vendor tools when advantageous. Using high-level tools like tracing directly from inside the FPGA [8] can subsequently improve performance. In addition, kernel invocations can be replaced by tasks described by data dependencies [9]. This is why, we plan to move part of the runtime system work inside the hardware to create a fast task scheduler [9,10] with the objective of enhancing the performance. It has been shown [10] that this type of mechanism can significantly improve the performance for many applications when using FPGAs directly from high-level language programs. In addition, we plan to broaden the range of programs that prove these optimal results with the tools used by further adapting the environment to TEXTAROSSA platforms. New features obtained from these platforms will then be ported back to the frameworks in order to benefit the whole range of systems targeted by our tools.

Training Deep Neural Network (DNN) is a memory-intensive operation. Indeed, the training algorithms of most DNNs require to store both the model weights and the forward activations in order to perform back-propagation [11]. In practice, training is performed automatically and transparently to the user through autograd tools for back-propagation available in frameworks such as PyTorch or TensorFlow. Unfortunately, the memory limitation of current nodes often prevents data scientists from considering larger models, higher resolution input date or larger batch sizes, especially in recent NLP models [12]. Our goal is to extend StarPU to enable inference and training, taking advantage of both the heterogeneity of the architecture to select where to place layers and the memory architecture in order to minimize transfers. In particular, we will rely on the RoToR framework[10] [13] which allows to control the memory consumption and to minimize the energy consumed by data exchanges.

Several problems should be solved to use FPGAs more efficiently and control energy consumption. In the TEXTAROSSA project, we will extend these two task-based runtime systems, OmpSs and StarPU, which use different approaches to support FPGA. This will allow the integrated development vehicle to benefit from their existing features, but also to study their complementarity while validating the robustness of the new HW against different runtime systems.

FastFlow. Stream processing is gaining increasing industrial attention for real-time data analytics and data-driven applications [14]. FastFlow [15] is a C++ programming library targeting multi/many-cores. It offers both a set of high-level ready-to-use parallel pattern implementations and a set of mechanisms and components (called building blocks) to support low-latency and high-throughput data-flow streaming networks. FastFlow simplifies the development of parallel applications modeled as a structured, directed graph of processing nodes. The graph of concurrent nodes is constructed by the assembly of sequential and parallel building blocks and higher-level components (i.e., parallel patterns) modeling recurrent schemas of parallel computations (e.g., pipeline, task-farm, parallel-for, etc.). FastFlow efficiency stems from the optimized implementation of the base communication and synchronization mechanisms and its layered software design. Besides, stream processing is the natural paradigm for event-driven distributed applications that need to communicate with each other via message passing. Finally, some data streaming paradigms are naturally suited for implementation on reconfigurable platforms [16], e.g. the dataflow/actor paradigm. In TEXTAROSSA, we aim at exploiting reconfigurable platforms to accelerate HPDA tasks leveraging the FastFlow framework [17].

StreamFlow. The StreamFlow framework [18,19] is a container-native Workflow Management System (WMS) written in Python 3 and based on the Common Workflow Language (CWL) Standard [20]. StreamFlow has been designed around two main principles: 1. Allowing the execution of tasks in multi-container environments, in order to support concurrent execution of multiple communicating tasks in a multi-agent ecosystem; 2. Relaxing the requirement of a single shared data space, in order to allow for hybrid workflow executions on top of multi-cloud or hybrid cloud/HPC infrastructures. StreamFlow source code is available on GitHub under the LGPLv3 license. A Python package is downloadable from PyPI and Docker containers can be found on Docker Hub. More details about the tool and its applications can be found in the StreamFlow website.[11]

Compiler technology for mixed-precision support. Many applications of HPC are error-tolerant, a characteristic that can be exploited – either in hardware or in software – to achieve important savings in system costs or power efficiency improvements [21]. Hardware-based proposals take advantage of inherent sensor limitations, redundant data, and reduced precision input, or introduce additional uncertainty by adopting design features that produce approximate results. On the other hand, software-based approximate-computing approaches such as floating-point optimization or loop perforation allow to trade-off algorithm exactness for a more efficient implementation [22]. To this end, high complexity of modern applications makes approximate compilers increasingly important. We aim to focus on mixed-precision compilers [23], which are a subset of approximate compilers. Recently, new data types specifically geared for mixed-precision are emerging, such as Posit [24] and BFloat16 [25]. Posit is a new compressed floating-point data format for which University of Pisa has developed a software library called CppPosit [26,27]. From the first results of applying the CppPosit library to AI/DNN problems, Posit can lead to the same processing accuracy of single-precision floating-point but with a data compression from a factor 2 to 4 [26,27]. This means that applying Posit to the application cases (HPC, HPDA and AI/CNN) has the potential to reduce data storage issues and allows for fast data movement. BFloat16 (Brain Floating Point) is used in upcoming Intel AI processors (NERVANA), XEON processors, Google Cloud TPU and ARMv8.6-A, as well as in RISC-V extensions [28]. In contrast to other standardized 16-bit floating point formats, BFloat16 offers a greater dynamic range and higher compatibility with the conventional single-precision floating-point format defined by the IEEE-754 standard. In TEXTAROSSA, we

---

aim at exploiting and extending the tools for precision tuning developed as part of the H2020 FETHPC ANTAREX project [29] collected in the TAFFO framework [30,31]. The TAFFO framework is implemented as a set of plugins for the LLVM compiler. Based on programmer hints expressed as attributes, TAFFO performs value range analysis, data type and code conversion, and static estimation of the performance impact, automatically producing a mixed-precision application with statically-guaranteed error bounds. TAFFO is language-independent, supports data types ranging from fixed-point to standard floating-point formats, and allows the user to finely tune the performance-precision trade-off to their needs [32]. The extensions to TAFFO will allow it to cover a wider range of target platforms, such as FPGAs through integration with the TEXTAROSSA High Level Synthesis (HLS) toolchain. Additionally, we aim to include support for emerging data types such as Posit and BFloat16, expanding the use of the tools to heterogeneous systems with reconfigurable components, and to improve the performance estimation by exploiting recent analysis techniques [33] as well as a deeper understanding of the target processor pipeline.

3.2. Runtime services: Energy/power management

The linear increment of power consumption, due to the end of the Dennard's scaling, is not sustainable for the next generation of Exascale computing systems. This, in fact, would have a huge impact. The energy consumption of the HPC centres would increase dramatically, other than posing severe issues in terms of thermal and power management. Overall, the deriving increase of costs and energy requirements is far from being an attractive perspective for an HPC center. Moreover, this increase of power requirements would represent an additional contribution to growth of the global carbon footprint, due to the information technologies [34]. Therefore, the design and deployment of energy-efficient HPC infrastructures is a primary concern, at all the possible levels. In this regard, the scientific community converged on the idea that providing computing nodes equipped with heterogeneous processing is an effective solution. This means that also the HPC infrastructures must be characterized by the presence of different processing units, such that, the workload allocation could follow an energy efficiency driven schema. This means to deploy a resource manager or job dispatcher capable of selecting the most efficient unit with respect to the specific job or application to serve [35]. However, energy efficiency can then be maximized if we can actually rely on a suitable resource management strategy at different levels. Although the state-of-the-art already includes some solutions, recent projects, like MANGO [36] and RECIPE [37], have shown that we need to take into account the platform-specific characteristics, other than develop suitable knobs to profile and monitor the execution of the workload, at run-time [38]. This enables more accurate resource management policies [39,40]. In addition, by integrating the programming model with the resource management layers we could dynamically tune the numerical accuracy (precision) of the algorithms' implementation, with respect to the actual application requirements and power/energy constraints [41]. Similarly, this applies to the problem of guarantee real-time requirements to time-critical applications [42,43]. In general, given the reference hardware platform and the application use-cases, the TEXTAROSSA project would represent an extremely interesting test-bench for exploring novel power and energy management strategies, by operating at both software and hardware level. At the hardware level, for example, we can provide a major contribution, by instrumenting the computing architecture with ad-hoc power monitors [44] and controllers [45,46], in order to minimize the monitoring latencies, while increasing the effectiveness of the management policies. On the software side, we partially start from an already existing resource management framework [47], and aim at revising its implementation, by introducing the support for the new target hardware, integrating the precision tuning mechanisms and the developing new platform-specific resource management policies. Overall, this would allow us to explore all the possibilities offered by the platform, and the application-side integration, to increase the FLOPS-per-Watt ratio, with respect to state-of-the-art HPC solutions.

3.3. Posit hardware accelerators

Within EPI and the European programs, accelerators based on RISC-V cores enable energy-efficiency for applications using stencil patterns or neuromorphic algorithms. Such accelerators implement floating point units able to work on classic fp32 float or new BFloat16 formats. To increase performance, one could try to reduce the number of bits in the floating point representation. But when reducing the precision of the used arithmetic, iterating over many timesteps, the derivations may become unstable and hence affect the final result of the simulations. The novel Posit binary arithmetic format can offer higher precision while using less bits than standard IEEE floating-point numbers. Recent literature shows [26,27] that 16 bit Posit can leverage comparable results like fp32 and Posit with 8 bit precision outperform in terms of accuracy fp16 (for CNN 8 bit Posit can leverage comparable results like fp32). Calculations can be done even with simple bit manipulations on the Posit format without extraction, further decreasing the complexity of the operations. It is thus possible to enhance memory bandwidth, and lower level cache utilization, power footprint, and throughput of the arithmetic units.

Within TEXTAROSSA, a RISC-V unit will be extended with support for alternative data representations, including fine-grained reduced precision floating point [28] and Posit arithmetic. To complement the hardware IP developments, the LLVM compiler, also adopted in EPI, will be extended for Posit and data compression support and real-world HPC applications and CNN kernels will be ported to leverage the IP. Fast software co-design will be enabled through software implementation of Posit in the CppPosit library.[12] Two approaches will be followed: the first called light PPU (Posit Processing Unit) will add Posit to/from float and Posit to/from integer converter to a RISC-V 64b 6 stages core with ALU and FPU. This way, with minimal circuit overhead, estimated in less than 1%, and exploiting the instruction customization of RISC-V, the support of Posit8 and Posit16 is also added. Since with the light PPU, Posit processing replies on ALU or FPU back-end the advantage is mainly in the compression properties of Posits that allows a compact storage of large CNNs, while the inference part will be done using the FPU with floats. In the second approach a full PPU will be added to the RISC-V 64b core, replacing the FPU. This version will be suited to use Posits for both storage and inference. Beside the implementation of accelerators for new data formats, other analysis will be carried out to provide basic blocks to implement novel cryptosystems onto FPGAs, like multipliers for large binary polynomials [48] and decoders for post-quantum cryptography [49].

The generated IPs will be considered for the sake of the global energy and power management and optimization, thanks to the exploitation of a methodology capable to automatically identify the power model of a generic hardware module and to augment the original design with its RTL implementation [50]. Note that such power monitors are properly designed in order to prevent the creation of an entry point for side-channel attacks.

The novel IPs will be ported to FPGAs for benchmarking. Both techniques have not been implemented on top of a completely co-designed accelerator so far and will provide a huge benefit for the European IP portfolio.

3.4. Hardware platform optimization

HPC systems have historically always been limited by thermal considerations and computing architectures need both optimized heat dissipation solutions and runtime thermal control policies to operate reliably and efficiently. The TEXTAROSSA project will provide contributions in both areas.

For what concerns heat dissipation improvements, the company InQuattro has developed an innovative thermal management solution

---

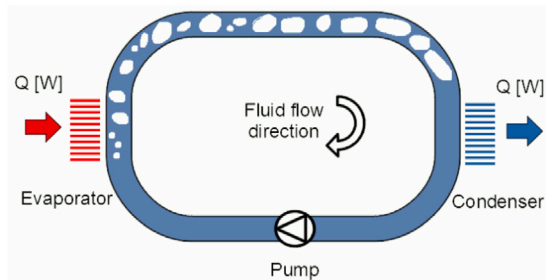[12] https://github.com/eruffaldi/cppPosit (last accessed March 2022)

Fig. 4. Schematic of the two-phase cooling system (left) and its embedding in a server (right).

(patent pending) based on two-phase mechanically pumped loops, which uses a flow boiling heat transfer for cooling electronics in a more efficient way. This allows the use of the latent heat of vaporization so that flow rates are significantly reduced, temperature gradients are kept small, and heat transfer coefficients are increased compared to not only air cooling but also liquid cooling systems. Two-phase cooling systems using evaporation and condensation are known to be the best way to meet demanding cooling requirements in terms of compactness, weight, and energy-consumption. A possible configuration of the two-phase cooling system for a single node is shown in Fig. 4.

In TEXTAROSSA, an optimized two-phase cooling solution for HPC is being developed and customized to fit the requirements of node and rack levels for Exascale applications. The development cycle consists of an initial cooling solution design phase complemented by the design of a dynamic model to be used for simulation, as well as for the design of thermal control policies. The following step consists of the actual production of a prototype cooling solution, to be fully characterized resulting in refinements of its model through experimental data. The final step will be the testing on HPC node(s) provided by E4 and ATOS and developed with the objective to serve an entire rack. Such a testing will be performed jointly with the thermal control policies. It is foreseen to develop two solutions of the two-phase cooling system that could be patented during the project on the principal components of the cooling system (evaporator, condenser). This innovative technology is expected to improve the cooling efficiency up to 70% compared to traditional air cooling, and up to 30% compared to existing liquid cooling and will be tested on both ATOS and E4 infrastructures.

It should be stressed that both the design of cooling solutions and thermal control policies critically relies on thermal models. Thermal simulators for CPUs/MPSoCs have been proposed, but most of them can only represent a limited range of heat dissipation solutions, and are not easy to extend towards two-phase liquid cooling solutions or to encompass the simulation of an entire rack. For this reason, in TEXTAROSSA, we are taking advantage of a recently introduced paradigm for the modeling of MPSoC cooling systems [51]: the co-simulation of highly optimized IC thermal model with heat dissipation models expressed by means of equation-based object-oriented modeling languages. This approach allows at the same time to leverage the performance of traditional IC thermal model, and to be able to simulate the complex, nonlinear thermal phenomena that arise when bringing evaporative cooling into play.

A key aspect of thermal modeling is the need for experimental data. As some of the thermal phenomena to be modeled rely on empirical correlations, experiments are needed to provide correct parameter values in order to accurately model heat dissipation phenomena. Moreover, the availability of high quality experimental data allows to perform validation of thermal models. However, the direct use of MPSoCs such as processors or GPUs to perform thermal experiments is made extremely difficult by the uncertainty in the power spatial distribution across the silicon die during computational workloads in modern processors, as well as due to the insufficient number of temperature sensors to fully reconstruct the temperature spatial distribution. In TEXTAROSSA, we are overcoming these limitations by relying on a thermal test chip

platform [52] developed at Politecnico di Milano. Thermal test chips are integrated circuits dedicated to thermal testing, providing an array of on-silicon heating elements and temperature sensors, allowing to capture accurate silicon thermal maps connected to the proposed heat dissipation solution, considering both spatial (e.g. hot spots) and temporal temperature variations. The integration of the two-phase cooling system with the TTC is currently underway (Fig. 5). The collected experimental data will be used to improve and validate the thermal models as well as for the design of the control policies.

For what concerns thermal control policies, in TEXTAROSSA the integration of two-phase cooling will be made by means of a multilevel thermal control strategy, aiming to overcome the complexity of controlling an HPC platform from node to system level. To reduce overhead, we will exploit and extend the use of event-based control policies (patent pending) developed at Politecnico di Milano to provide effective thermal control with minimal overhead. As the fastest temperature gradients occur at the silicon active layer, we will use fast event-based control loops [53] acting on DVFS to limit the maximum operating temperature of compute elements. These inner control loop will in turn interact with higher level control loops operating the two phase cooling infrastructure of the node, which is comparatively slower and has higher overheads but has the capability to increase the heat transfer coefficient on-demand, thus allowing to relieve the need to reduce frequency using DVFS, in turn improving performance. A further supervisory control layer will allow to set the desired temperatures at the rack level based on reliability metrics. Multilevel control allows thus to partition the system level control problem into multiple interacting control loops, each optimized for the specific thermal dynamics to control.

3.5. Low-latency communication between FPGAs

The usage of FPGAs as accelerators is getting so widespread that even big cloud providers are now installing reconfigurable devices in their instances (e.g. on Microsoft Azure and Amazon EC2). Interaction of hundreds to thousands of FPGAs requires a scalable approach to hold them together, in order to guarantee a low latency connection among them. A definitive approach has to be found, to provide users with the best flexibility, meanwhile easing the usage for software developers. As an example, the latest version of the Microsoft Catapult fabric, puts a Stratix 10 device between each NIC on the x86 servers and the ToR switch, enabling a fast path for accelerators to communicate among themselves with a few microseconds latency. The Brainwave project [54] leverages this architecture to provide a deep learning platform for real-time AI inference on the cloud. While this framework offers a very friendly interface for users to deploy their models on top of this architecture, it loses the flexibility of delivering the cores as black boxes, and providing an implementation of only a few pre-trained models. Our approach, on the other hand, is more flexible: users can define processing tasks freely using C++, the configuration tools developed in the project will generate automatically all the glue logic needed for the FPGA bitstream generation. This will allow developers
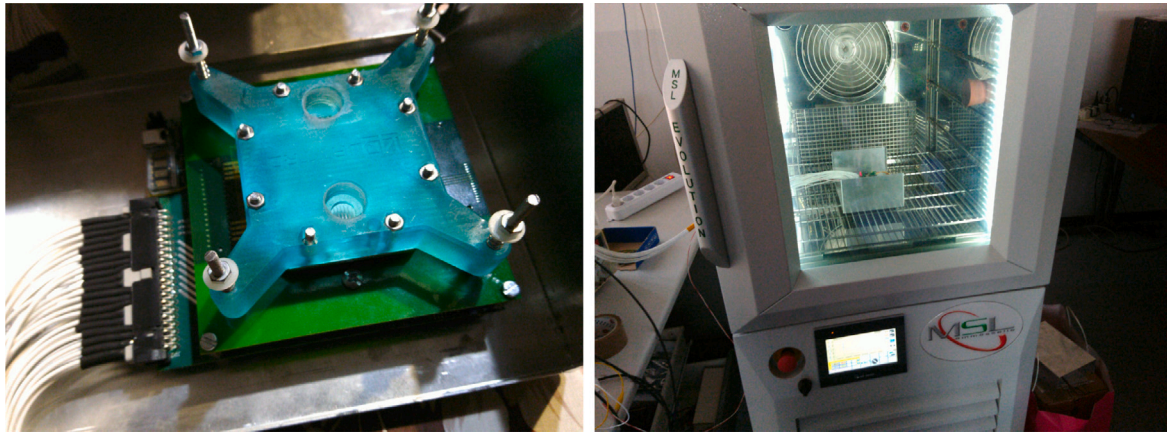
Fig. 5. Two-phase evaporator prototype installed on the Thermal Test Chip (left) and thermal chamber calibration of the TTC on-die temperature sensors (right).
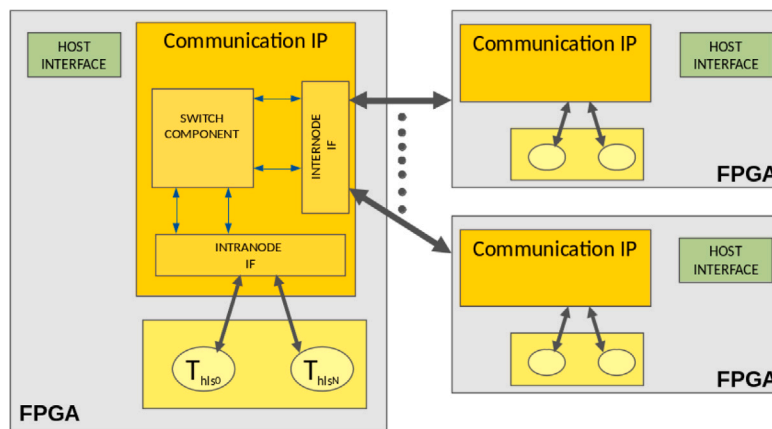

Fig. 6. Architectural partition of Communication IP.

to define a scalable application using a streaming programming model (Kahn Process Network [55]) that can be efficiently deployed on a multi-FPGAs system. TEXTAROSSA is developing a communication IP and its software stack, providing the implementation of a direct network that allows low-latency communication between processing tasks deployed on FPGAs, even if hosted in different computing nodes. The communication IP, based on the ExaNet IPs (switch, router, high-speed channels [56]) developed in ExaNeSt H2020 project [57] and EuroEXA H2020 [58], can be split into two Interface Blocks – InterNode IF and IntraNode IF – which manage data flow, and a switch_component, in charge of dynamically interconnect all ports of IFs (Fig. 6). The Communication IP meets the interface requirements to be used as an RTL kernel within the Vitis IDE.

To better take advantage of the framework and to easily allow the interconnection of computational kernels, we are developing a tool to simplify the inclusion and the configuration of the interconnection IP without the need to add an ad-hoc custom project. The user can specify the number of the ports as parameter, as well as how to connect the different kernels to the available ports in a custom configuration file. This configuration file is then used to generate the project files. This way the interconnection logic is built automatically by the configuration tool based on the application needs, allowing the end user to focus on the processing/computing kernels. The direct communication between tasks deployed on FPGAs avoids the involvement of the CPUs and system bus resources in the data transfers, improving the platform's energy efficiency and reducing communication latency.

4. TEXTAROSSA applications

In this section, we briefly provide an overview of the use case applications adopted in TEXTAROSSA. To address the variety of application domains of future Exascale systems, TEXTAROSSA applications include basic mathematical libraries (*MathLib*), traditional HPC applications (*UrbanAir*, *TNM*, *HEP*, and *RTM*), and applications from emerging domains (*RAIDER*, *DPSNN*, *Danger Detection*).

4.1. MathLib

Mathematical libraries are key components of the SW stack for HPC applications. They are building blocks implementing methods and algorithms for solving main mathematical problems that domain scientists reuse in form of basic components and then are critical to the accuracy, performance and energy efficiency of HPC applications. Any new generation of computer architecture requires new generation of mathematical SW libraries and the race to exascale, featured by computing elements of increasing complexity and heterogeneity, is no exception. Within the TEXTAROSSA project, extensions and improvements of high-performance algorithms and SW libraries for kernels in numerical linear algebra [3,59] and graph computation, such as iterative [60–63] and direct linear solvers, edge weighted graph matching, and fast multipole methods [5] will be deployed. We will provide new high-performance algorithms and SW modules for some of the so-called Colella's dwarves, that classified numerical methods crucial for science and engineering. In particular, we will focus on algorithms and SW for sparse linear algebra, where data sets include many zero-values and are usually stored in compressed data structures to reduce storage and

memory bandwidth requirements. Those data structures are generally accessed with indexed loads and stores, and main computational kernels are communication bound. We intend to develop and benchmark a library of kernels optimized for hybrid nodes embedding NVDIA GPUs in order to pursue node-level performance further than multi-node scalability. Scalability will be improved by leveraging overlap between communication and computation. We will also investigate the chance of using mixed-precision floating-point arithmetic, which offers many advantages in terms of memory footprint and computational efficiency, by applying strategies for preserving robustness and correctness, such as defect-correction approaches to meet application's accuracy and reproducibility needs. The library kernels will be of immediate use in a wide range of applications, ranging from classical scientific simulation to AI techniques, including automatic pattern recognition in complex systems, and will be tested on well established benchmarks.

4.2. HPDA and HPC-AI applications

Real-time AI-based data analytics on hetEROgeneous distributed systems (RAIDER). A Proof-of-Concept shall be provided as a typical High Energy Physics (HEP) real-time AI-based data analytics on heterogeneous distributed system, a processing paradigm that INFN APE Lab has already successfully adopted in the context of the NA62 experiment at CERN where we have developed the GPU-RICH system [64,65]. GPU-RICH integrates one CPU, one FPGA and one GPU in a single computing node and implements a real-time track reconstruction pipeline in which the FPGA-based NIC receives eight data streams from the RICH Cherenkov detector, processes them and injects the output stream directly into the GPU memory through the PCIe bus (GPUDirect RDMA) for the track reconstruction parallel processing.

Future experiments, which will handle high volume of data and high costs, push the need for new techniques in Trigger and Data Acquisition (TDAQ) systems to improve particle identification and further suppress background events in trigger systems, or to perform an efficient online data reduction for trigger-less ones. Architecturally, data streams from different channels/sources/detectors can be recombined through some processing layers using a low-latency, modular and scalable network infrastructure (configurable in number of channels, topology and size). Each processing layer performs feature extraction through machine learning leveraging Deep Neural Networks implemented on heterogeneous devices. This implementation must take into account the limited memory and floating point resources of some of the devices (usually those located at the edge) deploying resource-demanding Neural Network layers (e.g. CNN) in subsequent processing layers and studying reduced precision and/or compression techniques for input data.

These concepts define the RAIDER software application and its reference hardware platform, whose combination yields a distinguishing feature that is the capability of performing this distributed inference scheme with real-time constraints. FPGA devices are the key architectural elements enabling the implementation of the general RAIDER architecture in application scenarios characterized by very low-latency classification requirements. In fact, these devices allow the implementation of data transport and processing stages characterized by a highly predictable and low latency. Wrapping up overview and studies concerning available low-latency communication IPs and frameworks for Neural Networks (NNs) deployment on FPGA, it was straightforward to actualize a preliminary but functional testbed for distributed data-analytics. Henceforth having the possibility to evaluate pros and cons of a programming model focused on the realization of a real-time AI-based data analytics heterogeneous distributed system. For this purpose, a distributed HLS development framework is required, hence we will work toward the extension of one of those frameworks, namely Xilinx Vitis, in order to be able to deploy in a straightforward manner multi-FPGA distributed low-latency applications, such as RAIDER. It shall be done integrating INFN set of Interconnection IPs (switch, low-latency data channels) in the Xilinx Vitis framework also through the definition

and implementation of the full software stack supporting those IPs for the very low-latency data transfer between processing tasks deployed on the same FPGA (intra-node communication) and on different FPGAs (inter-node communication), in order to offer hardware support for the execution of applications developed according to the streaming programming models on a system made of multiple interconnected FPGAs. The careful design and implementation of the HW/SW interface, through the Vitis HLS flow to define the communication protocol (HLS communication primitives) and to map the I/O ports of the Vitis HLS flow on those provided by the communication IPs will be of great importance to fully exploit the potential of the hardware.

Brain simulation (DPSNN). INFN started tackling neural simulations with its own engine, the Distributed and Plastic Spiking Neural Network (DPSNN), which is a scalable C++/MPI code for HPC platforms at extreme scales simulating the spiking dynamics of a brain cortex modeled as a grid of cortical columns populated with neurons and their interconnecting synapses. It has been used to first model brain cortex behavior – with a special focus on sleep-like states [66] – and to gauge compute and power efficiency on different architectures [67]. INFN has now transitioned to another, more versatile tool, the NEST Simulator [68]. This is a C++/MPI/OpenMP code by the NEST Initiative[13] that empowers an user with a domain-specific language to design a virtual neurophysiology experiment, from the equations driving the dynamics of the components of interest in the cortex (with a rich library of many types of either neurons and synapses ready to be used) to the topology of their interconnections – the so-called *connectome* – all the way to the necessary supporting tools, like probing or stimulating electrodes that read or inject electrical currents into the simulated cortex. NEST offers to the experimenter an intuitive Python interface to easily setup a detailed and complex protocol of interaction between a simulated cortex and a set of external stimuli. Coupled with the huge set of tools for analysis, visualization and data transformation available to the Python user, NEST allows for a compact yet expressive way to perform even very involved neural simulations. INFN has used NEST to implement a biologically-inspired thalamo-cortical model which is able to be trained in classification of handwritten digits from the MNIST data set and then mimic the wake-sleep cycle, in order to test the enhancing effects of sleep on the quality of learning and recognition [69], even in noisy environments [70].

Smart cities video surveillance. Implementing distributed video surveillance systems in the context of smart cities provides a huge amount of data for processing, using also AI techniques for detection and classification. Possible applications are related to safety versus Covid-19, e.g., by detecting in real-time people, their body temperature, the respect of social distancing, or if they are wearing protective facial masks. A design activity will be carried out for a real-time danger alarm system composed of a network of smart cameras where a pre-processing stage plus an AI algorithm (e.g. a Yolo based CNN) is implemented on an EDGE server.

4.3. Traditional HPC applications

Air pollution (UrbanAir). The UrbanAir concerns the modeling and forecasting of the concentration and dispersion of pollutants. It is a 3D multiscale model that combines a numerical weather prediction (NWP) model, running at larger scale (e.g. mesoscale), with a city-scale geophysical flow solver (EULAG) for accurate prediction of contaminant (e.g. NO2, PM2.5, PM10) transportation through the street corridors, over buildings and obstacles. A design activity shall be carried out to use mixed-precision computing and energy-efficient accelerators for faster response while preserving results accuracy. The EULAG is a Fortran application which exploits message passing parallelization in

---

[13] https://nest-initiative.org (last accessed March 2022)

two (2P) or three (3P) dimensions. In each version, the computational domain is balanced between CPU workers — horizontally for the former version, horizontally and vertically for the latter. Some of the kernels of 3P version have been already adapted to GPU to reduce processing time. The computational domain can be divided between GPU accelerators only or between CPUs and GPUs. In the latter case the subdomain sizes are carefully chosen to avoid unbalanced computations. Moreover, the 3P version has been adapted with the BarbequeRTRM runtime resource manager to manage computing resources allocation, to dynamically switch between CPUs and GPUs to increase energy efficiency, and to increase reliability with checkpoint/restore mechanism [71].

Quantum simulation (TNM). Tensor Network Methods (TNM) are a class of powerful numerical methods developed to study the equilibrium and out-of-equilibrium properties of strongly correlated many-body quantum systems [72]. TNM are complementary to Monte Carlo methods as they do not suffer from the sign problem. A design activity shall be performed to overcome current TNM limitations studying how to push the simulation boundaries towards high-dimensional systems with the support of HPC infrastructure. As important steps in this direction, the first tensor network simulations of lattice gauge theories in (2+1) and (3+1) dimensions have been performed [73,74]. In particular, lattice Quantum Electrodynamics (QED) in the Hamiltonian formulation including dynamical matter has been considered and, by using the sign-problem-free TNM, it has been possible to compute the ground states of the model at zero and finite charge densities, and to address fundamental questions such as the characterization of collective phases of the model, the presence of a confining phase at large gauge coupling, and the study of charge-screening effects. These simulations have been performed on computer clusters by taking advantage only of OpenMP parallelization on single multi-core nodes. Further developments shall be carried out to simulate larger system sizes, by exploiting large-scale parallelization of TNM by means of MPI or GPU acceleration.

High Energy Physics (HEP). The HEP community has the transition to heterogeneous computing on its roadmap for the next decade, in order to profit from the large investments in HPC systems, and in general to access resources with a better cost/performance ratio [75]. A number of solutions have appeared in the last 4–5 years in the market; they, with their own peculiarities and strong and weak points, promise a seamless utilization of a variety of platforms, with a clear separation between a frontend part (seen by the user/programmer) and a backend part (taken care by the framework via a toolset of libraries and compilers). The HEP community, and in particular the LHC experiments, have started an experimentation with a few products, like Alpaka [76], Kokkos [77], Intel OneApi and SYCL. In TEXTAROSSA, we plan to select a representative set of applications, mission critical in their scientific domains, and evaluate the porting to the frameworks. Whenever successful, a thorough benchmarking will follow on the main target platform, and auxiliary ones. A first target for migration is certainly the collection of high-level software libraries used by the LHC experiments for simulation and data analysis. They include Geant4 [78] and Fluka [79] for particle–matter simulation, the use of high-level analysis tools like those in ROOT [80], and simulation packages of high-energy collisions. A design activity will be focused to optimize these software frameworks for the realization of code bases able to execute on multiple architectures, including the next generations of pre-Exascale and Exascale European HPC systems.

Biomedical application (HPC-Drugs). Nowadays, Ligand binding affinity predictions carried out with Molecular Dynamics simulation is one of the main research focus in computational chemistry, due to its potential impact in industrial drug discovery. A design activity shall be carried out for HPC-backed pharmaceutical applications based on n-body kernel functions running in specialized cores of GPU

```
_kernel void kforce(global double i-th part.);
for each particle j-th do
 │  evaluate the force on j-th part. by i-th part.;
end
end _kernel kforce;
```

Algorithm 1: n-body kernel function algorithm.

and FPGA, relying on recently discovered non-equilibrium thermodynamics theorems and capable of delivering absolute binding free energies of drug-size molecules in a predictable wall-clock time with a credible confidence interval, hence bypassing the limitations of the traditional equilibrium-based Free Energy Perturbation (FEP) alchemical approaches [81].

Heterogeneous architecture based on GPU/FPGA is already used to reduce the huge processing time in MD simulations. Its programming model provides a top-level abstraction for low-level hardware routines as well as consistent memory and execution models for dealing with massively-parallel code execution. A standard programming model is OpenCL, which is composed of one CPU-based "Host" controlling multiple "Compute Devices" such as GPUs and FPGAs. Each of these coarse-grained compute devices consists of multiple "Compute Units", and within these are multiple "Processing Elements". At the lowest level, these processing elements all execute OpenCL "Kernel Functions". These kernel functions compute the forces by involving in an n-body interaction a particle system on which all coordinates are fixed during the force computation that can be parallelized for the different values of each particle. In terms of streams and kernels, this can be expressed as the kernel function shown in the pseudo-code of Algorithm 1.

Reverse Time Migration (RTM). The Reverse Time Migration application and mini-kernels are used within EPI to co-design the STX Accelerator and have been ported to FPGAs within the EuroEXA project. The respective kernels will be analyzed to which extent they can leverage the new, energy-efficient, capabilities like Posit arithmetic and lossy compression to enhance performance and energy efficiency. The RTM kernels are stencil-based kernels. Hence, they provide conclusions on many stencil based applications. Reverse Time Migration by FHG for HPC applications to Oil & Gas and Geo-Services.

5. Conclusions

The EuroHPC TEXTAROSSA project addresses technology gaps targeting pre-Exascale and Exascale scenarios. In particular, the project aims at developing new IPs, algorithms, methods, and software components for HPC, HPC-AI, and HPDA applications. A co-design methodology is employed to reach the technical goals: from the optimization of the user application to the development of runtime services and programming models able to exploit the parallelism of the architecture. Then the CPU architecture and HW platforms, including at the rack-level, are optimized to achieve performance targets not only in terms of computing power, but also of power and energy consumption. The majority of the TEXTAROSSA tools will be open-source and able to be adopted as standalone building blocks or to interoperate with other Exascale-ready components developed within the EuroHPC initiative. The use-case applications adopted in TEXTAROSSA include mathematical libraries, traditional HPC software, and applications from emerging domains.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Authors report financial support was provided by Horizon 2020.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Acknowledgment

This work is supported by the TEXTAROSSA project (G.A. n. 956831), as part of the EuroHPC initiative.

References

[1] Giovanni Agosta, Daniele Cattaneo, William Fornaciari, Andrea Galimberti, Giuseppe Massari, Federico Reghenzani, Federico Terraneo, Davide Zoni, Carlo Brandolese, Massimo Celino, et al., TEXTAROSSA: Towards EXtreme scale Technologies and Accelerators for euROhpc hw/Sw Supercomputing Applications for exascale, in: 2021 24th Euromicro Conference on Digital System Design, DSD, IEEE, 2021, pp. 286–294.

[2] Vinod Kathail, Xilinx vitis unified software platform, in: Proceedings of the 2020 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, 2020, pp. 173–174.

[3] Emmanuel Agullo, et al., Achieving high performance on supercomputers with a sequential task-based programming model, IEEE TPDS (2017).

[4] Olivier Beaumont, Philippe Duchon, Lionel Eyraud-Dubois, Julien Langou, Mathieu Verite, Symmetric block-cyclic distribution: Fewer communications leads to faster dense cholesky factorization, in: SC'22: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, IEEE and ACM, 2022, (selected as best paper candidate).

[5] Emmanuel Agullo, et al., Task-based FMM for heterogeneous architectures, Concurr. Comput.: Pract. Exper. 28 (9) (2016) 2608–2629.

[6] Alejandro Duran, et al., Ompss: a proposal for programming heterogeneous multi-core architectures, Parallel Process. Lett. 21 (02) (2011) 173–193.

[7] Cédric Augonnet, et al., StarPU: a unified platform for task scheduling on heterogeneous multicore architectures, Concurr. Comput.: Pract. Exper. 23 (2) (2011) 187–198.

[8] Jaume Bosch, et al., Application acceleration on FPGAs with OmpSs@FPGA, in: FPT 2018, Naha, Okinawa, Japan, December 10-14, 2018, 2018, pp. 70–77.

[9] Xubin Tan, et al., A hardware runtime for task-based programming models, IEEE Trans. Par. Distributed Syst. 30 (9) (2019) 1932–1946.

[10] Juan Miguel De Haro, et al., OmpSs@FPGA framework for high performance FPGA computing, IEEE Trans. Comput. 70 (12) (2021) 2029–2042.

[11] Julia Gusak, Daria Cherniuk, Alena Shilova, Alexander Katrutsa, Daniel Bershatsky, Xunyi Zhao, Lionel Eyraud-Dubois, Oleg Shlyazhko, Denis Dimitrov, Ivan Oseledets, et al., Survey on large scale neural network training, in: Proceedings of IJCAI-ECAI 2022, the 31st International Joint Conference on Artificial Intelligence and the 25th European Conference on Artificial Intelligence, IJCAI, 2022.

[12] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 4171–4186.

[13] Olivier Beaumont, et al., Optimal Checkpointing for Heterogeneous Chains: How to Train Deep Neural Networks with Limited Memory, Report RR-9302, Inria Bordeaux Sud-Ouest, 2019.

[14] Marco Aldinucci, Salvatore Ruggieri, Massimo Torquati, Porting decision tree algorithms to multicore using FastFlow, in: Conference in Machine Learning and Knowledge Discovery in Databases, ECML PKDD, in: LNCS, vol. 6321, 2010, pp. 7–23.

[15] Marco Aldinucci, et al., FastFlow: high-level and efficient streaming on multi-core, in: Programming Multi-Core and Many-Core Computing Systems, in: Parallel and Distributed Computing, 2017.

[16] Stephen Neuendorffer, Kees Vissers, Streaming systems in FPGAs, in: Embedded Computer Systems: Architectures, Modeling, and Simulation, 2008, pp. 147–156.

[17] Marco Aldinucci, et al., Design patterns percolating to parallel programming framework implementation, Int. J. Parallel Program. 42 (6) (2014) 1012–1031.

[18] Iacopo Colonnelli, et al., StreamFlow: cross-breeding cloud with HPC, IEEE Trans. Emerg. Topics Comput. (2020).

[19] Iacopo Colonnelli, et al., HPC application cloudification: The StreamFlow toolkit, in: PARMA-DITAM HiPEAC Workshop, 2021), OASIcs, vol. 88, Dagstuhl, Germany, 2021, pp. 5:1–5:13.

[20] Michael R. Crusoe, et al., Methods included: Standardizing computational reuse and portability with the common workflow language, Commun. ACM 65 (6) (2022) 54–63.

[21] Swagath Venkataramani, et al., Approximate computing and the quest for computing efficiency, in: 2015 52nd ACM/EDAC/IEEE Design Automation Conference, DAC, IEEE, 2015, pp. 1–6.

[22] Phillip Stanley-Marbell, et al., Exploiting errors for efficiency: A survey from circuits to applications, ACM Comp. Surv. 53 (3) (2020).

[23] Stefano Cherubin, Giovanni Agosta, Tools for reduced precision computation: a survey, ACM Comput. Surv. 53 (2) (2020).

[24] John L. Gustafson, Isaac T. Yonemoto, Beating floating point at its own game: Posit arithmetic, Supercomput. Front. Innovat. 4 (2) (2017) 71–86.

[25] Neil Burgess, et al., Bfloat16 processing for neural networks, in: 2019 IEEE 26th Symposium on Computer Arithmetic, ARITH, IEEE, 2019, pp. 88–91.

[26] Marco Cococcioni, et al., A fast approximation of the hyperbolic tangent when using posit numbers and its application to deep neural networks, in: International Conference on Applications in Electronics Pervading Industry, Environment and Society, 2019, pp. 213–221.

[27] Marco Cococcioni, et al., Vectorizing posit operations on RISC-V for faster deep neural networks: experiments and comparison with ARM SVE, Neural Comput. Appl. (2021) 1–11.

[28] Davide Zoni, Andrea Galimberti, William Fornaciari, An FPU design template to optimize the accuracy-efficiency-area trade-off, Sustainable Comput.: Inform. Syst. 29 (2021) 100450.

[29] Cristina Silvano, et al., The ANTAREX domain specific language for high performance computing, Microprocess. Microsyst. 68 (2019) 58–73.

[30] Stefano Cherubin, Daniele Cattaneo, Michele Chiari, Antonio Di Bello, Giovanni Agosta, TAFFO: Tuning assistant for floating to fixed point optimization, IEEE Embedded Syst. Lett. (2019).

[31] Stefano Cherubin, et al., Dynamic precision autotuning with TAFFO, ACM Trans. Archit. Code Optim. 17 (2) (2020).

[32] Daniele Cattaneo, Michele Chiari, Nicola Fossati, Stefano Cherubin, Giovanni Agosta, Architecture-aware precision tuning with multiple number representation systems, in: 2021 58th ACM/IEEE Design Automation Conference, DAC, 2021, pp. 673–678.

[33] Eva Darulova, Viktor Kuncak, Towards a compiler for reals, ACM Trans. Program. Lang. Syst. 39 (2) (2017) 8:1–8:28.

[34] Michele Zanella, Giuseppe Massari, Andrea Galimberti, William Fornaciari, Back to the future: Resource management in post-cloud solutions, in: Proceedings of the Workshop on INTelligent Embedded Systems Architectures and Applications, INTESA '18, 2018, pp. 33–38.

[35] Michele Zanella, Filippo Sciamanna, William Fornaciari, BarMan: A run-time management framework in the resource continuum, Sustain. Comput.: Inform. Syst. 35 (2022) 100663.

[36] J. Flich, et al., Exploring manycore architectures for next-generation HPC systems through the MANGO approach, Microprocess. Microsyst. 61 (2018) 154–170.

[37] William Fornaciari, Giovanni Agosta, David Atienza, Carlo Brandolese, Leila Cammoun, Luca Cremona, Alessandro Cilardo, Albert Farres, José Flich, Carles Hernandez, Michal Kulchewski, Simone Libutti, José Maria Martínez, Giuseppe Massari, Ariel Oleksiak, Anna Pupykina, Federico Reghenzani, Rafael Tornero, Michele Zanella, Marina Zapater, Davide Zoni, Reliable power and time-constraints-aware predictive management of heterogeneous exascale systems, in: Proceedings of the 18th International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation, SAMOS '18, Association for Computing Machinery, New York, NY, USA, 2018, pp. 187–194.

[38] Davide Zoni, Luca Cremona, William Fornaciari, All-digital energy-constrained controller for general-purpose accelerators and CPUs, IEEE Embedded Syst. Lett. 12 (1) (2020) 17–20.

[39] Giovanni Agosta, William Fornaciari, Giuseppe Massari, Anna Pupykina, Federico Reghenzani, Michele Zanella, Managing heterogeneous resources in HPC systems, in: Proceedings of the 9th Workshop and 7th Workshop on Parallel Programming and RunTime Management Techniques for Manycore Architectures and Design Tools and Architectures for Multicore Embedded Computing Platforms, PARMA-DITAM '18, Association for Computing Machinery, New York, NY, USA, 2018, pp. 7–12.

[40] Carlo Brandolese, Simone Corbetta, William Fornaciari, Software energy estimation based on statistical characterization of intermediate compilation code, in: IEEE/ACM International Symposium on Low Power Electronics and Design, 2011, pp. 333–338.

[41] Federico Reghenzani, Giuseppe Massari, William Fornaciari, A probabilistic approach to energy-constrained mixed-criticality systems, in: 2019 IEEE/ACM International Symposium on Low Power Electronics and Design, ISLPED, 2019, pp. 1–6.

[42] Federico Reghenzani, Giuseppe Massari, William Fornaciari, The misconception of exponential tail upper-bounding in probabilistic real time, IEEE Embedded Syst. Lett. 11 (3) (2019) 77–80.

[43] Federico Reghenzani, Giuseppe Massari, William Fornaciari, Timing predictability in high-performance computing with probabilistic real-time, IEEE Access 8 (2020) 208566–208582.

[44] Davide Zoni, et al., PowerTap: All-digital power meter modeling for run-time power monitoring, Microprocess. Microsyst. 63 (2018) 128–139.

[45] Davide Zoni, Luca Cremona, William Fornaciari, All-digital control-theoretic scheme to optimize energy budget and allocation in multi-cores, IEEE Trans. Comput. 69 (5) (2020) 706–721.

[46] Federico Reghenzani, Simone Formentin, Giuseppe Massari, William Fornaciari, A constrained extremum-seeking control for CPU thermal management, in: Proceedings of the 15th ACM International Conference on Computing Frontiers, CF '18, Association for Computing Machinery, New York, NY, USA, 2018, pp. 320–325.

[47] Patrick Bellasi, Giuseppe Massari, William Fornaciari, Effective runtime resource management using linux control groups with the barbequertrm framework, ACM Trans. Embed. Comput. Syst. 14 (2) (2015) 39:1–39:17.

[48] Davide Zoni, Andrea Galimberti, William Fornaciari, Flexible and scalable FPGA-oriented design of multipliers for large binary polynomials, IEEE Access 8 (2020) 75809–75821.

[49] Davide Zoni, Andrea Galimberti, William Fornaciari, Efficient and scalable FPGA-oriented design of QC-LDPC bit-flipping decoders for post-quantum cryptography, IEEE Access 8 (2020) 163419–163433.

[50] Davide Zoni, Luca Cremona, William Fornaciari, Design of side-channel resistant power monitors, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. (2021) 1.

[51] Federico Terraneo, et al., 3D-ICE 3.0: efficient nonlinear mpsoc thermal simulation with pluggable heat sink models, IEEE Trans. Comput.-Aided Des. Integrated Circ. Syst. (2021) 1.

[52] Federico Terraneo, et al., An open-hardware platform for mpsoc thermal modeling, in: SAMOS'19, 2019, pp. 184–196.

[53] Alberto Leva, et al., Event-based power/performance-aware thermal management for high-density microprocessors, IEEE Trans. Control Syst. Technol. 26 (2) (2018) 535–550.

[54] Eric Chung, Alii, Serving DNNs in real time at datacenter scale with project brainwave, IEEE Micro 38 (2) (2018) 8–20.

[55] Gilles Kahn, The semantics of a simple language for parallel programming, in: Information Processing, 6th IFIP Congress 1974, 1974, pp. 471–475.

[56] Roberto Ammendola, et al., Large scale low power computing system: Status of network design in ExaNeSt and EuroExa projects, Adv. Parallel Comput. 32 (2018) 750–759.

[57] Manolis Katevenis, Roberto Ammendola, et al., Next generation of exascale-class systems: ExaNeSt project and the status of its interconnect and storage development, Microprocess. Microsyst. 61 (2018) 58–71.

[58] Andrea Biagioni, Paolo Cretaro, Ottorino Frezza, Francesca Lo Cicero, Alessandro Lonardo, Pier Stanislao Paolucci, Luca Pontisso, Francesco Simula, Piero Vicini, EuroEXA custom switch: an innovative FPGA-based system for extreme scale computing in Europe, EPJ Web Conf. 245 (2020) 09004.

[59] Terry Cojean, et al., Resource aggregation for task-based Cholesky Factorization on top of modern architectures, Parallel Comput. 83 (2019) 73–92.

[60] Massimo Bernaschi, et al., A factored sparse approximate inverse preconditioned conjugate gradient solver on graphics processing units, SIAM J. Sci. Comput. 38 (1) (2016) C53–C72.

[61] Massimo Bernaschi, Pasqua D'Ambra, Dario Pasquini, AMG based on compatible weighted matching for GPUs, Parallel Comput. 92 (2020) 102599.

[62] Massimo Bernaschi, Pasqua D'Ambra, Dario Pasquini, BootCMatchG: An adaptive algebraic MultiGrid linear solver for GPUs, Softw. Impact. 6 (2020) 100041.

[63] Pasqua D'Ambra, Salvatore Filippone, A parallel generalized relaxation method for high-performance image segmentation on GPUs, J. Comput. Appl. Math. 293 (2016) 35–44.

[64] Roberto Ammendola, et al., NaNet: a flexible and configurable low-latency NIC for real-time trigger systems based on GPUs, J. Instrum. 9 (02) (2014) C02023.

[65] R Ammendola, M Barbanera, A Biagioni, P Cretaro, O Frezza, G Lamanna, F Lo Cicero, A Lonardo, M Martinelli, E Pastorelli, P S Paolucci, R Piandani, L Pontisso, D Rossetti, F Simula, M Sozzi, P Valente, P Vicini, Real-time heterogeneous stream processing with NaNet in the NA62 experiment, J. Phys. Conf. Ser. 1085 (2018) 032022.

[66] Elena Pastorelli, et al., Gaussian and exponential lateral connectivity on distributed spiking neural network simulation, in: 26th Euromicro PDP, IEEE, 2018, pp. 658–665.

[67] Roberto Ammendola, et al., The brain on low power architectures-efficient simulation of cortical slow waves and asynchronous states, Adv. Parallel Comput. 32 (2018) 760=769.

[68] Robin de Schepper, Jochen Martin Eppler, Anno Kurth, Pooja Nagendra Babu, Rajalekshmi Deepu, Sebastian Spreizer, Guido Trensch, Jari Pronold, Stine Brekke Vennemo, Steffen Graber, Aitor Morales-Gregorio, Charl Linssen, Mohamed Ayssar Benelhedi, Håkon Mrk, Abigail Morrison, Dennis Terhorst, Jessica Mitchell, Sandra Diaz, Itaru Kitayama, Mahdi Enan, Nilton Liuji Kamiji, Hans Ekkehard Plesser, NEST 3.2, 2022.

[69] Cristiano Capone, et al., Sleep-like slow oscillations improve visual classification through synaptic homeostasis and memory association in a thalamo-cortical model, Sci. Rep. 9 (1) (2019) 1–11.

[70] Bruno Golosio, Chiara De Luca, Cristiano Capone, Elena Pastorelli, Giovanni Stegel, Gianmarco Tiddia, Giulia De Bonis, Pier Stanislao Paolucci, Thalamo-cortical spiking model of incremental learning combining perception, context and NREM-sleep, in: Maxim Bazhenov (Ed.), PLoS Comput. Biol. 17 (6) (2021) e1009045.

[71] Giovanni Agosta, William Fornaciari, David Atienza, Ramon Canal, Alessandro Cilardo, José Flich Cardo, Carles Hernandez Luz, Michal Kulczewski, Giuseppe Massari, Rafael Tornero Gavilá, Marina Zapater, The RECIPE approach to challenges in deeply heterogeneous high performance systems, Microprocess. Microsyst. 77 (2020) 103185.

[72] Ahmed Omran, et al., Generation and manipulation of Schrödinger cat states in Rydberg atom arrays, Science 365 (6453) (2019) 570–574.

[73] Timo Felser, Pietro Silvi, Mario Collura, Simone Montangero, Two-dimensional quantum-link lattice quantum electrodynamics at finite density, Phys. Rev. X 10 (4) (2020).

[74] Giuseppe Magnifico, Timo Felser, Pietro Silvi, Simone Montangero, Lattice quantum electrodynamics in (3+1)-dimensions at finite density with tensor networks, Nature Commun. 12 (1) (2021).

[75] Tommaso Boccali, Computing models in high energy physics, Rev. Phys. 4 (2019) 100034.

[76] Erik Zenker, Benjamin Worpitz, René Widera, Axel Huebl, Guido Juckeland, Andreas Knüpfer, Wolfgang E. Nagel, Michael Bussmann, Alpaka - An abstraction library for parallel kernel acceleration, CoRR (2016) http://arxiv.org/abs/1602.08477.

[77] Sivasankaran Rajamanickam, Seher Acer, Luc Berger-Vergiat, Vinh Q. Dang, Nathan D. Ellingwood, Evan Harvey, Brian Kelley, Christian R. Trott, Jeremiah J. Wilke, Ichitaro Yamazaki, Kokkos kernels: Performance portable sparse/dense linear algebra and graph kernels, CoRR (2021) https://arxiv.org/abs/2103.11991.

[78] S. Agostinelli, et al., GEANT4 Collaboration Collaboration, GEANT4: A simulation toolkit, Nucl. Instrum. Meth. A506 (2003) 250–303.

[79] A. Ferrari, Paola R. Sala, A. Fassò, Johannes Ranft, FLUKA: A Multi-Particle Transport Code (Program Version 2005), in: CERN Yellow Reports: Monographs, CERN, Geneva, 2005.

[80] Rene Brun, Fons Rademakers, ROOT - An object oriented data analysis framework, in: Nucl. Inst. & Meth. in Phys. Res. A, AIHENP'96 Workshop, Lausane, vol. 389, 1996, pp. 81–86.

[81] Marina Macchiagodena, et al., Virtual double-system single-box: A nonequilibrium alchemical technique for absolute binding free energy calculations: Application to ligands of the SARS-CoV-2 main protease, J. Chem. Theory Comput. 16 (11) (2020).

Giovanni Agosta (Ph.D. 2004) is currently Associate Professor at Politecnico di Milano, Italy. His research interests focus on compiler technologies for assessing and enforcing extra-functional properties of programs in both HPC and embedded systems. He has published over 100 papers in international journals and conferences. He is a member of the CINI consortium, where he is a member of the managing board of the CFC Lab on training and competence certification. He is associate editor of SoftwareX (Elsevier).

Marco Aldinucci is a full professor and P.I. of the Parallel Computing research group at the University of Torino. He has participated in over 15 EU-funded research projects on parallel and cloud and high-performance computing attracting over 10M€ of research funds to the University of Torino. He has led the design of the HPC4AI laboratory. He is the founding director of the CINI HPC Key Technologies and Tools national laboratory, gathering researchers from 38 Italian Universities. He is co-leading FutureHPC, the technological spoke of the Italian National Centre on HPC (ICSC). He co-designed the FastFlow programming framework and several other libraries for parallel computing.

Carlos Alvarez received his M.S. and Ph.D. degrees in Computer Science from Universitat Politecnica de Catalunya (UPC), Spain, in 1998 and 2007, respectively. He is currently a Tenured Assistant Professor in the department of Computer Architecture at UPC and an associated researcher at the BSC. His research interests include parallel architectures, runtime systems, and reconfigurable solutions for high performance multiprocessor systems. He has participated on several EU projects related to the use of FPGAs for HPC: AXIOM, EuroEXA, LEGaTO. He is currently participating in the Textarossa and MEEP projects. He has coauthored more than 80 publications in international journals and conferences. He has co-advised four Ph.D. theses and he is currently co-advising three Ph.D. students.

Roberto Ammendola received his Master's Degree in Physics and Ph.D. in Electronic Engineering from "Tor Vergata" University in Rome, Italy. He is with INFN since 2003, but worked also for Centro Ricerche Enrico Fermi and "Tor Vergata" University. His main research interests are reconfigurable devices and high speed network interconnects, HPC systems and infrastructure design and deployment, high-speed parallel trigger and readout acquisition systems, integer and floating point arithmetics in-network acceleration. He collaborates with CERN and ASI.

Yasir Arfat received his Bachelor's degree in Software Engineering from the University of Azad Jammu & Kashmir, Muzaffarabad, Pakistan in 2011 and a Master's degree in computer science from King Abdulaziz University, Jeddah, Saudi Arabia in 2017. He is pursuing his Ph.D. in modeling and data science from the University of Turin (Unito). He also has experience in working with different companies as a software engineer. His main areas of interest are data analytics, federated Learning, distributed and parallel computing, machine learning, data science, and software engineering.
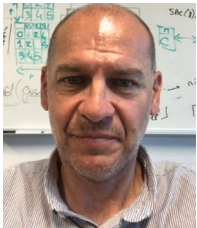
Olivier Beaumont holds a Senior Scientist position at Inria in the Hiepacs team. He defended his Ph.D. thesis in Rennes in 1999 and his "Habilitation à diriger des recherches" in Bordeaux in 2004. He is the author of 25 journal papers and 76 papers in International conferences. Currently, he acts as Associate Editor in Chief of JPDC and he acted as track chair of the main HPC conferences (SC, IPDPS, ICPP, HIPC). His main research interests are combinatorial optimization problems arising in HPC (load balancing, scheduling, data distribution), with applications in Numerical Linear Algebra and Training of Deep Neural Networks.

Massimo Bernaschi graduated in Physics in 1987 working with the first prototype of the APE computer developed by INFN. He worked ten years at the IBM European Center for Scientific and Engineering Computing, where he developed the PVMe product and received two Outstanding Technical Achievement Awards. In 1998 he joined the National Research Council of Italy as Chief Technology Officer of the Institute for Applied Computing. In 2012 he has been named CUDA Fellow for his contributions to multi-GPU programming. He is also an adjunct Professor of Computer Science at LUISS University in Rome.

Andrea Biagioni received his Master's degree in Physics in 2006 and his Ph.D. in Information and Communications Technologies in 2018 from University Sapienza in Rome, Italy. Since 2007 he has worked in the APE Group at INFN Roma as hardware developer and interconnection network designer, collaborating at several international (SHAPES, EURETILE, NA62, KM3NeT, ExaNeSt, HBP, EuroEXA, TEX-TAROSSA, RED-SEA) and national (APEnet, NaNet, COSA, GAP, APEIRON) projects. His main research interests are High Performance Computers, massively parallel computers, Network Interface Card design, communication protocols, FPGA, fault tolerance and routing algorithms. He co-authored about 100 international peer-reviewed journal and conference papers (total citations: 2729; h-index: 22)
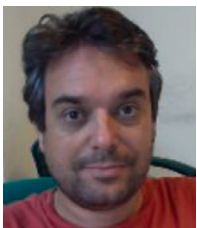
Tommaso Boccali I am a Senior Researcher at INFN (Italian National Institute for Nuclear Physics), currently working in Pisa. I have a Ph.D. in Particle Physics from Scuola Normale Superiore, and I spent the last 20 years working mostly on High Energy Physics experiments at CERN, initially at LEP (Large Electron Positron Collider) and now at LHC (Large Hadron Collider). With time, I moved my interests from pure physics research, to the design, construction and operations of large computing systems. I have been CMS (one of the two LHC experiments discovering the Higgs Boson in 2012) Offline and Computing coordinator, responsible for all the software and computing related issues. I was also in charge, as Deputy Upgrade Coordinator, of the design of the

computing system for the upgraded CERN Collider, expected in operations by 2030.Currently I am supervisor of the CNAF-reloaded project, which aims to move Italy's biggest data intensive Computing Center to the new Tecnopolo in Bologna. I am part of the INFN Computing management, and involved in projects with PRACE and the other realities in computing at national scale. I am INFN PI for the H2020 Project ESCAPE, after having participated to INDIGO-DataCloud and EOSC-Hub. In INFN, I work on the Infrastructure Working Group which analyzes and proposes the deployment and the governing of distributed computing centers in Italy. I am part of the Science Europe Working Group on "Data Sharing and Supporting Infrastructures".

Bérenger Bramas holds a researcher position (Chargé de Recherche) at Inria Nancy since October 2018. He is also a member of the ICube laboratory. He defended his Ph.D. thesis in 2016 on the parallelization and optimization of the time-domain boundary element method for the wave equation. Subsequently, he worked as an HPC Expert at the Max Planck supercomputing center (MPCDF). His research interests focus on scientific computing, runtime systems, scheduling, software engineering for HPC, and automatic optimization/parallelization.

Carlo Brandolese is a researcher at Politecnico di Milano since 2004 and his research interests are focused on ultra low power design methodologies for embedded systems. In the last few years he started studying IoT and distributed heterogeneous embedded systems and their interaction in the context of high performance computing applications.

Barbara Cantalupo is a research engineer at the Computer Science Department of the University of Torino. She received her master's degree in Computer Science from the University of Pisa in 1994, and she has been a researcher in the parallel computing field at University and at the Italian National Research Council (CNR). Afterwards, she worked for 15 years in several private companies in the areas of supercomputing, mobile networks and space, acquiring knowledge in different application fields. She also gained experience in the entire cycle of software development, project and team management, and research project development. A few years ago, she returned to University to investigate new evolution in the parallel computing environment.

Mauro Carrozzo is a software developer at the Institute for Applied Computing, National Research Council of Italy. He holds a Bachelor degree in Physics from the University of Rome "La Sapienza".

Daniele Cattaneo earned his M.Sc graduation cum laude in Computer Science and Engineering in December 2018, with a thesis work based on the architecture and implementation of a mixed precision compiler which exploits the fixed point numerical representation. Currently he is a Ph.D. student at Politecnico di Milano. His research interests involve embedded systems, compiler technology and software-hardware codesign.

Alessandro Celestini is a researcher at the Institute for Applied Computing "M. Picone" of the National Research Council of Italy. He holds both a Bachelor and Master's degree in Computer Science, from the University of Rome "La Sapienza", and a Ph.D. in Computer Science and Engineering, from the School for Advanced Studies IMT Lucca. His research interests include complex network analysis, graph algorithms and high performance computing.

Massimo Celino (ENEA). He obtained his Ph.D. in Condensed Matter Physics at the University "L. Pasteur" of Strasbourg (France). Since 1996 staff scientist at ENEA, his main interests are in the field of high performance computing and its applications. He promotes and manages several national and European projects in the field of HPC and data science. Principal Investigator of the EuroHPC project TEXTAROSSA, he is currently involved in the Energy oriented Centre of Excellence in computing applications (www.eocoe.eu), the EERAdata project (www.eeradata.eu) and the Italian Mission Innovation initiative. He is the author of more than 80 scientific papers in international journals. He delivered talks and communications in a significant number of national and international conferences.

Iacopo Colonnelli is a Computer Science research fellow. He received his Ph.D. with honors in Modeling and Data Science at the University of Torino with a thesis on novel workflow models for heterogeneous distributed systems. His research focuses on both statistical and computational aspects of data analysis at a large scale and workflow modeling and management in heterogeneous distributed architectures.

Paolo Cretaro received his master's degree in Physics in 2015 at Sapienza University of Rome (Italy), with a thesis about the implementation of a custom data acquisition system for a fast astronomical photometer (SiFAP). In 2020 he obtained his Ph.D. in Computer Science on reconfigurable computing and heterogeneous architectures. Since 2016 he is working for INFN as hardware/software developer. His main research interest is the exploitation of reconfigurable devices as accelerators for High Performance Computing.

Pasqua D'Ambra is senior researcher at the Naples branch of the Institute for Applied Computing "M. Picone" of the National Research Council of Italy. She earned her Ph.D. in Applied Mathematics and Computer Science from University of Naples Federico II in 1995. Her research interests are in algorithms and software for parallel numerical linear algebra with applications to Computational and Data Analysis. Her activities are carried on in the context of national and international projects, also with PI roles. She is also adjunct Professor of Numerical Computing at the University of Campania since 2003 and member of the governing board of HPC-Key Technology and Tools laboratory at the Italian Interuniversity Consortium for Informatics (CINI).

Marco Danelutto is a full professor at the Univ. of Pisa, Dept. of Computer Science. His main research interests are in the field of structured parallel programming models and include design of parallel programming frameworks, tools to support parallel program development, autonomic management of non functional features, software components, parallel design patterns and algorithmic skeletons. He is the author of about than 200 papers in international refereed journals and conferences.

Roberto Esposito received the M.Sc. and Ph.D. degrees in computer science from the University of Turin, Italy, in 1999 and 2003, respectively. He is presently an Assistant Professor working within the Machine Learning group at the Computer Science Department of the University of Turin. He has more than 20 years of research experience in the field of Machine Learning and published more than 50 research products, including articles in prestigious journals (JMLR, Information Systems, Neurocomputing, Nature Methods, IEEE Access, …) and conferences (ICML, IJCAI, …). Today his research interests mainly focus on non-functional aspects of Machine Learning (e.g., fairness and explainability) as well as on theory of neural networks and federated learning. He participated in numerous European Projects and served in the program committee of numerous prestigious journals and conferences (Machine Learning Journal, IJCAI, ECML-PKDD, ECAI, SAC, …).

Lionel Eyraud-Dubois is a Researcher at the Inria Center of the University of Bordeaux since 2008, as a member of the TOPAL team. He graduated his Ph.D. at the University of Grenoble in 2007. He has a broad research interest in approximation algorithms for combinatorial optimization problems, with a focus on resource management in large scale computing platforms. He has worked on HPC, Grid and Cloud platforms, and more recently on optimization for Deep Learning applications.

Antonio Filgueras earned a degree in Computer Science from Universitat Politècnica de Catalunya (UPC) in 2012. Currently, he is a Ph.D. student in the department of Computer Architecture at UPC and working in the OmpSs@FPGA team within the Programming Models group at BSC, focused primarily on programming models for reconfigurable systems in High Performance Computing. He has been a researcher in the AXIOM, Legato, EuroEXA and MEEP European projects among others.

William Fornaciari is professor at Politecnico di Milano, Milano, Italy. He won in 2016 the HiPEAC technology transfer innovation award and the Switch to Product contest in 2019, with the result of founding two startups. He published six books and more than 300 papers, collecting 6 best paper awards, one certification of appreciation from IEEE. He filed five international patents on power/thermal management. He has been involved in around 20 EU-funded international projects, in FP7 he coordinated the HARPA project and in H2020 the FET-HPC RECIPE project. Since 2021 he is the Project Technical Manager of the Euro-HPC TEXTAROSSA project. He is in the board of directors of the CINI (consortium of informatic and computing engineering in Italy). His main research interests cover multi/many core architectures, High-Performance-Computing, Design of low power hardware and software, runtime resource management, thermal and power management, and security of IoT/Embedded systems.

Ottorino Frezza received his Master's Degree in Physics in 2005 from University Sapienza in Rome, Italy with a thesis regarding VHDL FPGA-based design of a Read out board for a high energy physic experiment (NEMO). He worked from 2007 to 2009 for Eurotech s.p.a. as technical employee. Since 2009 he has worked for INFN as hardware developer. His main research interest is the development of network architecture on FPGA for massively parallel processing systems, focusing on high speed interfaces.

Andrea Galimberti, M.Sc., is a Ph.D. student at Politecnico di Milano, Italy. He received his M.Sc. degree in 2019 in Computer Science and Engineering at Politecnico di Milano. His research interests include computer architectures, hardware-level countermeasures to side-channel attacks and design of hardware accelerators.

Francesco Giacomini is a senior technologist at INFN-CNAF, in Bologna, where he leads the Software Development group. His activity mainly concerns the design and implementation of software components in the domains of distributed computing systems, data acquisition systems and applications for physics experiments.

Brice Goglin is a senior research scientist at Inria Bordeaux – Sud-Ouest. He earned his Ph.D. at École normale supérieure de Lyon (France) in 2005. He then worked for Myricom, inc. (Oak Ridge, TN) as a software architect for low latency networks. His research interests at Inria now include the modeling of HPC platforms, the management of data locality as well as high performance I/Os. He is the main developer of hwloc, the de-facto standard library for managing hardware topology and locality information in parallel applications.

Daniele Gregori, Ph.D. in Physics, is a parallel file system, automation tool and data tiering expert. During his career, he collaborated with the LHCb experiment at CERN, operationally managed the INFN-CNAF Tier1 computing center and taught General Physics at the University of Bologna. Daniele started working at E4 Computer Engineering in 2015 and is currently involved in the design of HPC computing infrastructures and in the management of European projects.
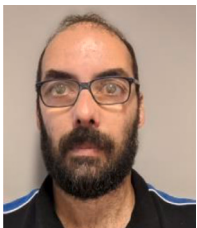
Abdou Guermouche is an associate professor at university of Bordeaux since 2005 and a member of the TOPAL team. He graduated his Ph.D. at ENS Lyon in 2004 and defended his Habilitation à diriger des recherches in Bordeaux in 2017 in Bordeaux. His main research interests are in parallel computing and more precisely high-performance solvers, runtime systems and scheduling strategies. He is one the developers of software packages like MUMPS and qr_mumps.

Francesco Iannone, ENEA senior researcher physicist, graduated cum laude in physics. He is currently permanent senior researcher of ENEA and for 20 years more, he worked in ENEA Nuclear Fusion Department as specialist in controls and data acquisition systems of the main Italian experiment in magnetic confinement nuclear fusion (Frascati Tokamak Upgrade). He designed and developed Instrumentation and Controls in hard real-time operating systems as well as geographical distributed data handling in large physics experiments. Since 2016 he is member in Operation and Project Committees of HPC facilities of EUROfusion. Since 2017 he is head of HPC laboratory responsible of ENEA HPC systems operations.

Michal Kulczewski (Male) works in the Applications Departmentat Poznan Supercomputing and Networking Center since 2005. He received his M.Sc. degree in Computer Science, Computer Network and Distributed Systems, from the Poznan University of Technology, Poland in 2005. His research interests include mainly numerical weather predictions, modeling air quality in urban environments, code optimization, programming multi- and many-core computing systems, advanced in-situ data analysis and visualization. He has actively contributed to many EU and national projects such as InteliGrid, ACGT, BREIN, OGF-Europe, POWIEW, PL-Grid+, ESCAPE, RECIPE, Change2Twin, TEXTAROSSA.
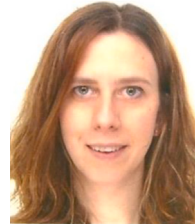
Francesca Lo Cicero received her Master's Degree in Electronics Engineering in 2005 from University Tor Vergata in Rome, Italy with a thesis regarding VHDL FPGA-based design of a Calibration System for a Digital Beam Forming Network. Since 2006 she has worked for INFN as hardware developer. Her main research interest is the development of network architecture on FPGA for massively parallel processing systems, focusing on hardware accelerations and optimizations.

Alessandro Lonardo, M.Sc. in Physics and Ph.D. in Information and Communications Technologies from University of Rome "La Sapienza". He is in Italy with INFN Roma APE Lab as senior technology researcher. His current research interests include reconfigurable computing architectures for High Performance Computing and Data Analytics, Artificial Intelligence and Cognitive Networks. He co-authored more than 190 international peer-reviewed journal and conference papers (total citations 2068, h-index 23).

Alberto R. Martinelli received his Master's Degree in 2020 with a master thesis on CAPIO (Cross-Application Programmable I/O), a novel method for in-transit HPC computing. While he was a M.Sc. student he worked as a junior research engineer in the parallel computing research group of the university. He is currently Ph.D. student at the Computer Science Department of the University of Turin working on storage for HPC, parallel computing, distributed computing and heterogeneous computing.

Michele Martinelli obtained his Master's Degree and Ph.D. in Computer Science from University of Rome "La Sapienza", Italy. He is currently working in the APE Group at Istituto Nazionale di Fisica Nucleare (INFN). His research activities are in the field of hardware/software co-design, device driver and software optimization, while his interests focus on HPC, operating systems, GPGPUs and networks.

Xavier Martorell received the M.S. and Ph.D. degrees in Computer Science from the Universitat Politecnica de Catalunya (UPC) in 1991 and 1999, respectively. Since 1992 he has lectured on operating systems, parallel runtime systems, OS administration, and systems for data science. He has been an associate professor in the Computer Architecture Department at UPC since 2001. His research interests cover the areas of operating systems, runtime systems, compilers and applications for high-performance multiprocessor systems. Since 2005 he is the Manager of the Parallel Programming Models team at the Barcelona Supercomputing Center. He has participated on several EU projects related to the use of FPGAs for HPC: AXIOM, EuroEXA, LEGaTO. and he is currently participating in the Textarossa and MEEP projects. He has coauthored more than 80 publications in international journals and conferences. He has co-advised eight Ph.D. theses and he is currently co-advising three Ph.D. students.

Giuseppe Massari received his Ph.D. in Information Engineering in 2015 at the Politecnico di Milano. He has worked on several EU-funded research projects, FP7 (2PARMA, CONTREX, HARPA) and H2020 (MANGO, RECIPE). He has been the lead developer of frameworks and libraries released with the projects' outcomes. He is currently a researcher at Politecnico di Milano, working on techniques for reducing energy/power consumption and improving the reliability of computing systems.

Simone Montangero leads the line of research on quantum science at the Department of physics and astronomy of the University of Padova, where he is full professor since 2017. Prof. Montangero has been a Heisenberg Fellow of the German Science Foundation, a Humboldt Fellow and an IQOQI visiting fellow of the Institute for Quantum Optics and Quantum Information of the Austrian Academy of Science. He has authored more than 150 scientific publications at the interface of quantum science, many-body physics, and condensed matter. He has pioneered the development of tensor network methods for quantum technologies and the application of optimal control theory to many-body quantum systems. He sits in the Science and Engineering Board of the EU Quantum Flagship and has coordinated and participated in several national and European research projects. He co-leads the Quantum Computing spoke in the Italian National Center for HPC, Big Data and Quantum computing funded but the EU Recovery plan.

Gianluca Mittone received his Bachelor's Degree in Computer Science in 2017 with a thesis on the handling of exceptions in Description Logics, proposing the implementation of an algorithm for the automatic revision of ontologies exploiting a Typicality operator. He also received the Master's Degree in Computer Science in 2019 with a master thesis on a novel distributed approach for deep learning, named NNT (Nearest Neighbors Training), which takes advantage of a locally synchronous approach to achieve a better trade-off between computational time and learning results. After eight months as a research engineer at the Computer Science Department of the University of Turin, he is now a Ph.D. student in Modeling and Data Science at the same university, working on different projects involving HPC and Machine Learning techniques.

Raymond Namyst is Professor of computer science at the University of Bordeaux. He is member of the "Storm" Inria research team, devoted to the design of high performance runtime systems for parallel architectures. His main research interests are in parallel computing, task scheduling over heterogeneous architectures and load balancing strategies. He has played a major role in the development of the StarPU runtime system. He authored numerous papers about the design of efficient runtime systems.

Ariel Oleksiak, Ph.D., D.Sc. (Male) works in the Applications Division at PSNC since 2002. He received his M.Sc. degree in Computer Science, Intelligent Decision Support Systems, from the Poznan University of Technology, Poland in 2001 and Ph.D. in Computer Science from the same institution in 2009. He leads an energy-efficiency technologies department. His research interests include mainly efficiency of data centres and clouds, distributed systems, intelligent methods of resource management and data processing, and simulations. He has actively contributed to many research projects such as FiPS. Dolphin, CATALYST, ASPIDE and RECIPE. He was a project coordinator of the EU-funded CoolEmAll project dealing with energy-efficiency of data centres. another EU-funded M2DC project developing a modular, highly efficient, cost-optimized server architecture composed of heterogeneous microservers. He has been also a member of the management committee of the European COST Actions IC0804 and recently IC1305 on ultrascale computing.

Paolo Palazzari (Laurea Degree'89, Ph.D. '94) is a researcher at ENEA, co-founder of Ylichron Srl (an ENEA spin-off company), and has been for eight years (2010–18) responsible for the HLS activities in Accelize. Formerly, he was the CTO for the SW activities of Ylichron Srl, where he developed HCE, the Ylichron C to HDL compiler. He has been one of the main designers of QuickCompiler, the HLS engine of the QuickPlay framework, developed by PLDA/Accelize. His current research interests are related to HLS, parallelism extraction along with its management, and efficient implementation of algorithms onto FPGA through HLS methods.

Pier Stanislao Paolucci currently serves as deputy leader of the "networks underlying cognition and consciousness" WP in the Human Brain Project (HBP) and coordinates the participation of the INFN APE Parallel/Distributed Computing lab to the HBP and to the EBRAINS-Italy projects. The current focus of its research is on the simulation of the cognitive functions of the different brain states in biological neural networks, on the reproduction of the spatio-temporal features of cortical waves, on the analysis of experimental brain activity and on the study of the brain architecture as source of inspiration for the development of advanced parallel computing solutions and bio-inspired artificial intelligence. In the past, the focus of his activity has been on the co-design of massive parallel hardware, VLIW and MPSOC architectures, system software and numerical algorithms.

Federico Reghenzani received his Ph.D. degree from Politecnico di Milano in January 2021. He is currently a Postdoctoral Researcher working on real-time systems and embedded computing. His research interests are on hard real-time scheduling, probabilistic timing analysis, mixed-criticality systems, and resource management on heterogeneous platforms.
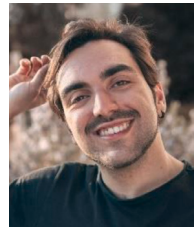
Cristian Rossi, M.Sc. in Physics (Particle and Astroparticle curriculum) from University of Rome "La Sapienza", Italy. He is currently working with INFN Roma APE Lab as research fellow. His current research interests include Real-time data analytics on heterogeneous distributed system, Artificial Intelligence, and the possible application of these in High Energy Physics experiments.

Sergio Saponara received the master's (cum laude) and Ph.D. degrees from the University of Pisa. He is currently a Full Professor of electronics at the University of Pisa. He is Director of: I-CAS Laboratory, Summer School Enabling Technologies for IoT, President of B.Sc. and M.Sc. programmes in Electronics Engineering. He coauthored more than 300 scientific publications and 20 patents. He is the leader of many funded projects by EU and by companies like Intel, Marelli, Magna, Ericsson. In 2012, he was a Marie Curie Research Fellow of IMEC. He is an IEEE Distinguished Lecturer and co-founder of special interest group on the IoT of both IEEE CAS and SP societies. He is an associate editor of several IEEE and Springer journals.

Francesco Simula received his Master's Degree in Theoretical Physics in 2006 from Sapienza University in Rome, Italy with a thesis regarding spin glass simulations performed on the massively parallel computers designed by the APE Group, the supercomputing initiative internal to Istituto Nazionale di Fisica Nucleare (INFN). Since 2006, he has been working at the Department of Physics of Sapienza University as temporary research fellow and then for INFN – as permanent staff as of October 2018 – as developer of high performance numerical simulations of scientifically interesting codes, ranging from Lattice QCD

to brain simulations with spiking neuron dynamics. Parallel computing is his main research interest, focusing on high performance networking and GPU acceleration on both HPC and embedded systems. His very latest effort concerns the optimized, power-efficient porting of brain simulation code onto FPGA-accelerated hybrid platforms.

Federico Terraneo has obtained the Master of Science Degree in Engineering of Computing Systems at the Politecnico di Milano with the thesis "Control based design of OS components", and the Research doctorate in Information Technology with the thesis "Thermal and energy management techniques for multi-core and many-core systems". His current research line focuses on thermal management for microprocessor-based computing systems and improving time determinism of distributed embedded systems. He has multidisciplinary research interests, including embedded systems and operating systems, with a lively interest in designing hardware and software components applying a control-theoretical methodology. Since 2008 he is the original designer, main developer and maintainer of Miosix (http://miosix.org), an operating system kernel targeting microcontroller-based embedded systems.

Samuel Thibault is Professor at the University of Bordeaux, and part of the Inria STORM team. His researches revolve around task and data transfer scheduling in parallel and distributed runtime systems. He is currently focused on the design of the StarPU runtime, and more particularly its scheduling heuristics for heterogeneous architectures and for distributed systems.

Massimo Torquati is an assistant professor in computer science at the University of Pisa. He has published more than 100 peer-reviewed papers in conference proceedings and international journals, mostly in the fields of parallel and distributed programming and run-time systems for parallel computing platforms. He has been involved in several Italian, EU, and industry-supported research projects. He is the maintainer and the main developer of the FastFlow parallel programming framework.

Matteo Turisini is a postDoc technologist at INFN in Rome (Ph.D. in 2017). His research interest include innovative data acquisition and data analysis systems for scientific research. He has been working for a decade in the field of particle detectors R&D, both for fundamental physics experiments at international accelerator facilities and for medical and environmental applications. Since 2020 he joined the APE team focusing his activity on neural network algorithms implementation on FPGA devices using HLS techniques.

Piero Vicini, M.Sc. in Physics, "La Sapienza" Roma University, joined Istituto Nazionale di Fisica Nucleare (INFN) in 1993. He is senior researcher and head of the APE Lab (INFN Supercomputing initiative). Since 2006, he coordinated the INFN participation to the SHAPES, EURETILE, H2020 FET-HPC ExaNeSt and EuroExa, and EuroHPC RED_SEA projects. Current research interests include: development of parallel processing systems optimized for scientific simulations, dedicated network architecture on FPGA and computational accelerators. He co-authored more than 250 international peer-reviewed journal and conference papers (total citations 4170, h-index 34).

Miquel Vidal received a B.Sc. degree in Computer Engineering in 2015 and an M.Sc. in High-Performance Computing in 2017 from the Universitat Politècnica de Catalunya (UPC). Since 2015 has been part of the Programming Models group at Barcelona Supercomputing Center (BSC), first as a resident student and currently as a research engineer. Current research is focused on programming models for heterogeneous and reconfigurable solutions for high-performance computing.

Davide Zoni is an Assistant Professor at Politecnico di Milano, Italy. He published more than 50 papers in journals and conference proceedings. His research interests include RTL design and verification of single- and multi-cores at the edge with emphasis on low power, hardware security, and deep learning. He filed two patents on cyber-security and is co-founder at Blue Signals, a spin-off of Politecnico di Milano.

Giuseppe Zummo, Mechanical Engineer, Ph.D., University of Rome "La Sapienza", Italy, has been researcher of ENEA (Italy) and founded the technological start-up In Quattro. He is Contract Professor of Thermofluid-Dynamics at Guglielmo Marconi University. His main field of interest is heat transfer and two-phase flow in microgravity and microgeometries for advanced thermal management of power electronics.