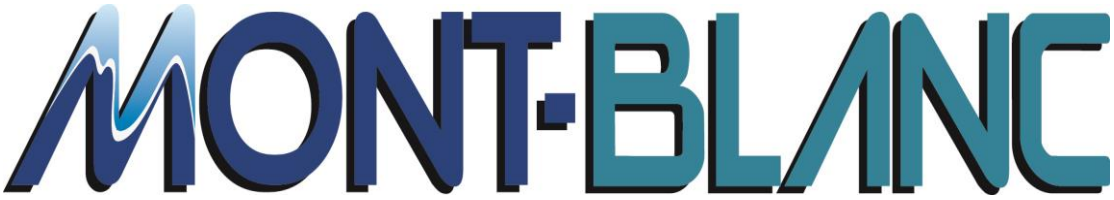

D1.5 - Mont-Blanc 3 Final report

Document Information

Grant Number	671697
Project Website	www.montblanc-project.eu
Period covered	Whole project: M1 – M39 (From 2016, Oct 1st to 2018, Dec. 31th)
Contractual Deadline	PM 39
Dissemination Level	PU
Nature	Report
Author	Etienne WALTER & Pascale BERNIER-BRUNA (Bull/Atos)
Contributors	M.CASAS (BSC), F.MANTOVANI (BSC), G.DOGSA (Arm), R.RUSITORU (Arm), J.WANZA-WELOLI (Bull/Atos) & all partners
Reviewer	
Keywords	Project final report

Notices:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 671697

Change Log

Version	Description of Change
0.1	2018-11-27 - Preliminary draft
0.5	2019-01-31 – Draft version for comments
0.9	2019-02-15 – RFC version

Table of Contents

1. Introduction

Arm low-power processors dominate the mobile world of smartphones, tablets and embedded IoT devices. With data centres consuming ever more power, the idea of using highly energy-efficient Arm chips in servers is enticing, especially for energy-hungry high-performance computing (HPC) configurations. As early as 2011, several pioneering European companies and institutions recognised the tremendous potential offered by embedded processor technology and decided to unite into the Mont-Blanc project to investigate the usage of low-power Arm processors for HPC.

However, making the leap from the mobile market to HPC was not trivial: HPC-optimised libraries, compilers and applications did not exist for Arm platforms. Mont-Blanc partners had to start from scratch building Arm HPC test systems based on 32-bit mobile phone technology and porting and tuning software and tools to create an Arm software ecosystem. In 2015, Mont-Blanc deployed the world's first Arm-based HPC cluster, featuring over 2,000 mobile CPUs. This system helped demonstrate the viability of using Arm technology for HPC.

Eight years on, the landscape has changed dramatically. Arm has introduced its first 64-bit architecture – ArmV8. The Mont-Blanc team put a lot of effort into extending and consolidating the ecosystem developed under the first phase of the project: scientific libraries and runtime systems were ported to ArmV8 and a set of development tools was developed for debugging, performance analysis, performance prediction and automated kernel optimisation. And Arm-based HPC has become a reality, with a choice of commercial products, by Mont-Blanc partner Atos and by several other leading vendors.

Projects Mont-Blanc have a significant share in this success.

[…] so many projects have been funded by the EU to build hardware and while these are all interesting, they are often one-offs that don't emerge to Top 500-level systems. The most notable exception to this statement is the Mont Blanc project, which we have followed diligently over the years. This is ARM focused and built with native European integrator, Atos/Bull, which showed that Europe could indeed roll its own systems with homegrown tech. […]

The Next Platform, 23 January 2019, in "HIPEAC: shifting focal points in European HPC research"

Arm, Bull (Atos group) and the Barcelona Supercomputing Center (BSC) have been the three key partners throughout the various Mont-Blanc projects. However, over the years the projects have benefited from the involvement of a whole host of other European partners, bringing valuable expertise from research labs, academia and industry to contribute to the success of the program.

2. Objectives

The Mont-Blanc projects share the goal of developing a European Exascale approach leveraging commodity power-and cost-efficient embedded technologies.

The third phase of the project, the Mont-Blanc 3 action (also called here Mont-Blanc 3 project or MB3), targets the design of a HPC solution based on an ARM architecture that can deliver cutting edge performance and energy efficiency. The Mont-Blanc and Mont-Blanc 2 projects have shown that increasing the computing efficiency of the cores is an important element and step towards providing competitive solutions. Besides working at the core efficiency level, the project researches and designs an HPC System-on-Chip SoC, to be implemented later in the 2020 timeframe.

Research on a balanced architecture is mandatory for the design of this SoC. The research needs to take into account the evolution of the different technologies used to integrate the compute aspect into the complete solution. The project develops software components that will be essential to enable the ARM based solution to deliver all the expected benefits.

The main objectives of the MB3 project are:

1. **To design a well-balanced architecture** and to design an ARM based SoC or SoP (System on Package) capable of providing pre-exascale performance when implemented in the 2019-2020 timeframe. The performance achieved will be measured using real HPC applications.
2. To **introduce new high-end ARM core and accelerator implementations** to support **HPC applications** efficiently.
3. **To develop the necessary software ecosystem** for the future SoC. This is vital in order to maximize the project impact and to ensure that the ARM architecture will achieve market success.

The project structure is mapped to the above objectives; the project addresses well-balanced architecture (WP3), computing efficiency architecture (WP4) and the development of the needed software ecosystem (WP7), in a holistic approach, fed by a work on targeted applications (WP6) and the creation of a simulation environment enabling the design space exploration (WP5).

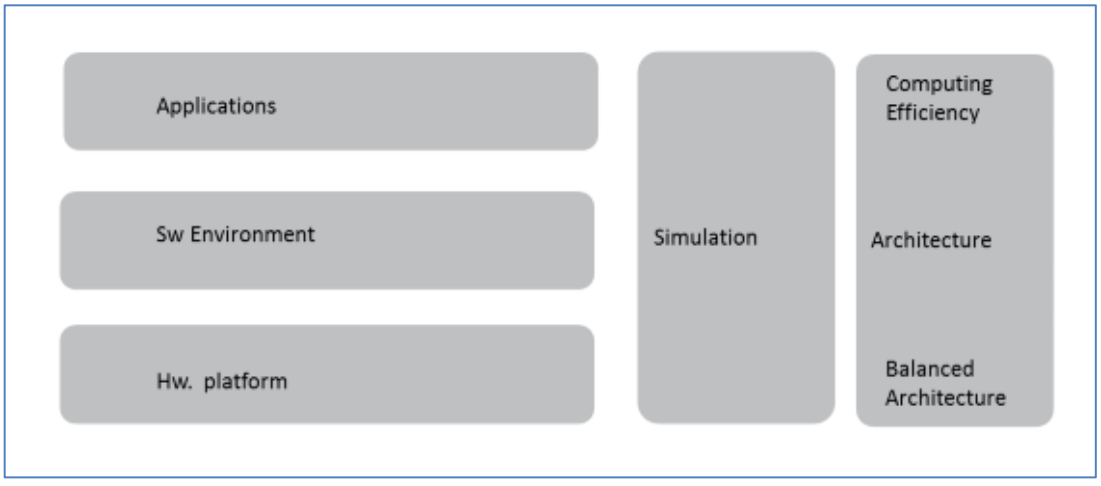

Figure 1: Project structure

3. Main scientific and technological results

The Mont-Blanc 3 project has produced several results on scientific and technological plans. We intend here to highlight a selection of these results

An industrial prototype (the Dibona platform)

An important step for us was the set-up of a medium-sized test platform (as defined in our initial plan), which would make it possible to assed our performance extrapolation with the observation of real applications. Furthermore, this testbed, codenamed Dibona and built by Bull, provides a key support to showcase our holistic work and enhance the impact of our project.

We announced in January 2017 our choice of Cavium's (now Marvell) second-generation ThunderX2 ARM server SoC, and in June 2017 our intent to productize this development. Later in 2017, other providers, such as Cray, HPE, or Gigabyte, also announced their own Thunder X2 based servers. Then, the detailed specifications of the motherboard, the mechanical enclosure and its accessories were implemented, and the manufacturing was completed. The first prototype blades were received in May. Finally, after a bring-up and a test phase, one motherboard was available as of July 2017 for software development.

The integration of the Cavium ARM v8 processor in a BullSequana platform raised some design challenges. The overall density had to be increased to accommodate more memory and larger socket dimensions compared to the version housing Intel CPUs.

A dense design was needed to fit into the BullSequana 1U volume and form factor; with only the necessary components. Management features were implemented by adding FPGA ASIC monitoring board sensors to obtain accurate power measurement at a high sampling rate.

Also, a specific cooling solution had to be designed for this new blade, with new CPU & DDR heat spreaders, and a new cold plate.

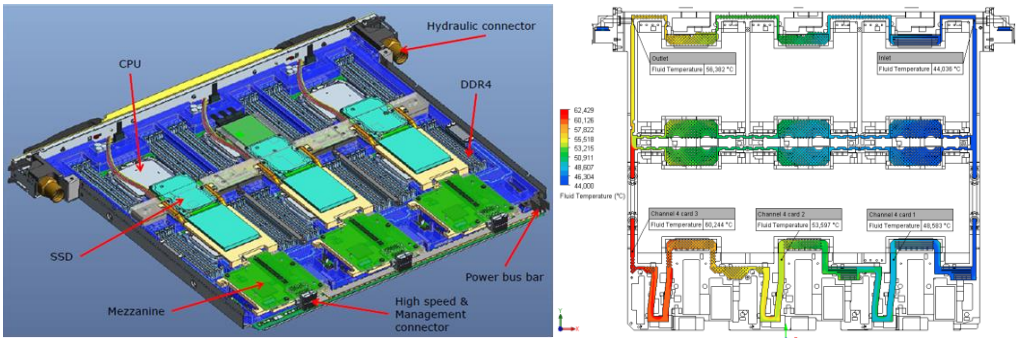

Figure 2 - Blade mechanical and cooling views

Results

As a result, our first motherboard booted an OS (RedHat 7.4) in July 2017.

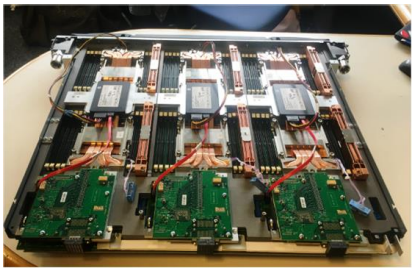

Figure 3 - Dibona blade, the first prototype

Our Dibona BullSequana racks were delivered and installed in S2 2018 as shown in the following pictures:

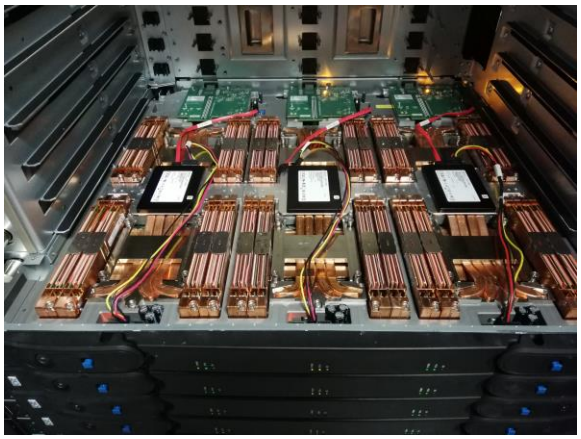

Figure 4 - First blades mounted in the Dibona rack

Figure 5 - The completed Dibona system

Dibona is now fully operational, and accessible to the European HPC community (Prace & other EU project teams for instance).

Software ecosystem

A comprehensive software stack is essential for the successful use and adoption of any system (HPC or otherwise). As Arm-based systems are still, relative to other platforms, a novelty in HPC, we spent a considerable amount of time on developing and enhancing the software ecosystem during Mont-Blanc 3. Our primary goal has been to extend the Arm software ecosystem and fill some of the previously identified gaps. Some of the improvements we made are, however, architecturally-agnostic. Nonetheless, all tools have been tested on Arm-based platforms and most (when relevant) are available for use on Dibona. The work covers both incremental developments (e.g., porting tools to Armv8 systems) and furthers research (e.g., investigate communication and computation overlap within runtimes). Within this work package, we worked on a mix of open-source and commercial solutions. Where possible, open-source contributions have been made.

Prior to the start of the project, we identified gaps across the entire software stack. As shown in Figure 6, we made improvements at all levels, from compilers and scientific libraries, to runtimes and operating systems.

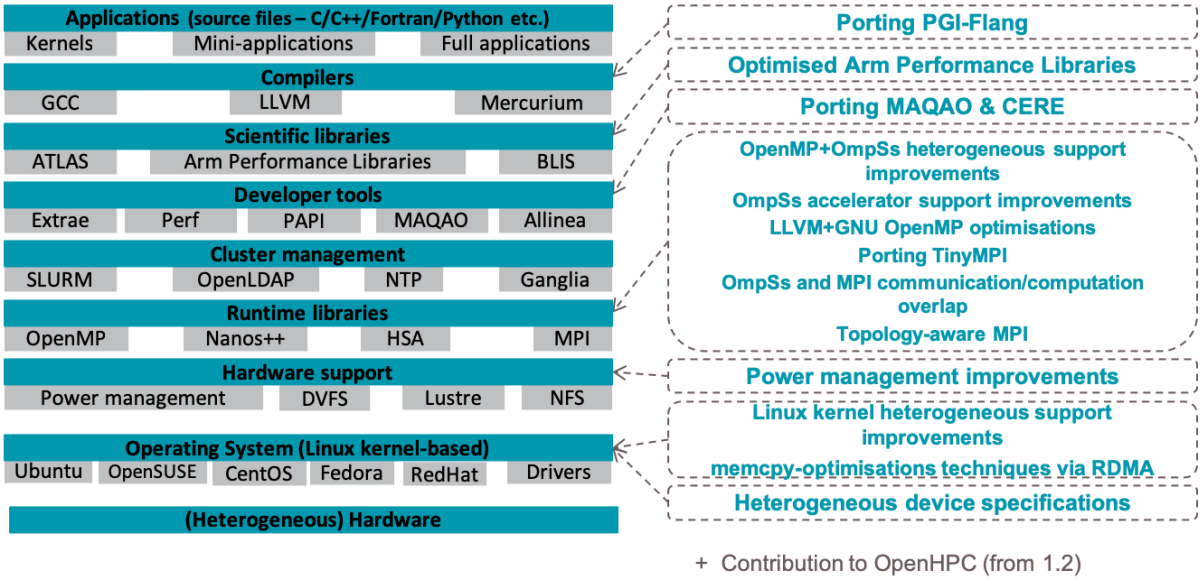

Figure 6 Sample HPC software stack, with a list (right hand side) of all the developments made within the Mont-Blanc 3 project.

Here are the main outcomes in this area:

- PGI-Flang has been ported to Arm and has been incorporated in the Arm Fortran Compiler, and, in parallel with this work, improvements to OpenMP were made;
- The Arm Performance Libraries have been further optimized;
- MAQAO and CERE now have Armv8 support;
- We achieve computation and communication overlap via two methods:
    - MPI-only: by using TinyMPI, a prototype one-sided asynchronous MPI runtime that uses over-subscription to achieve the overlap;
    - MPI+X: by using an MPI+OmpSs library which redefines the synchronous MPI calls using their asynchronous counterparts;
- An Atos Arm Software stack, including improvements to message passing, `memcpy`, power management and topology-aware MPI;
- DAST, a runtime technique that moves the task management operations to an additional thread or to all threads (in the Distributed-DAST version);
- Heterogeneous hardware scheduling policies and Heterogeneous Device Specifications (HDS), runtime and operating system-level scheduling policies and device specifications aimed at improving device information availability and scheduling on heterogeneous platforms, such as those based on Arm big.LITTLE technology.

All improvements, with the exception of HDS and the scheduling policies, which target heterogeneous devices, are available on Dibona.

Further details about some of the presented software tools can be found in the public Work package 7 deliverables.

Table 1 shows a list of which software is or can be open-sourced, has support for multiple architectures and whether it is driven by an industrial or academic entity.

Table 1 - A status list of the main Mont-Blanc 3 software ecosystem contributions in terms of lead developer (academia/industry), open-source status and support for multiple architectures.

Software tool	***Industrial/Academic***	***Open-source***	***Multi-architecture support***
PGI-Flang	Industrial	Yes	Yes
Arm Fortran Compiler	Industrial	Partially (based on some OSS SW)	No
Arm Performance Libraries	Industrial	No	No
MAQAO and CERE	Academic	Yes	Yes
TinyMPI	Academic	Can be OSS	Yes
OmpSs+MPI comm/comp overlap library	Academic	Can be OSS	Yes
Atos Arm Software Stack	Industrial	Partially (based on some OSS SW)	No
DAST	Academic	Can be OSS	Yes
Heterogeneous scheduling and HDS	Industrial	No (can be OSS)	Yes

Application evaluation on Dibona

The Mont-Blanc 3 project evaluated a new production-ready Arm-based machine, Dibona. We structured the evaluation with a bottom-up approach, executing programs with an increasing level of complexity. We evaluated first the simplest micro-benchmarks, exposing the basic architectural features such as the floating-point throughput of Marvell's ThunderX2 CPU, the structure of the memory subsystem and the bandwidth and the latency of the network. We then executed the most relevant high-performance computing benchmarks, LINPACK and HPCG, together with the mini-app Lulesh and commonly used solvers. We finally performed tests with production scientific applications combined with the runtime optimizations.

These studies have been performed at scale, comparing the results with state-of-the-art Tier-0 HPC systems. Also, when possible, we collected figures of energy and performance on different architectures. Depending on the applications, results reflects the architectural features of the Arm-based SoC selected for Dibona: we noticed in fact that the balance between the 30% more bandwidth and the 4x smaller SIMD unit offered by the ThunderX2 compared to Intel Skylake, creates a new balance in the performance of complex applications. The result is a new point in the architectural space that has been evaluated with real production applications, showing results in line with state-of-the-art HPC systems.

This early access to Dibona by the application and benchmarking team has been challenging although an intense teamwork has tremendously accelerated the process towards production readiness of Dibona.

The work performed by the application team within the Mont-Blanc 3 project generated the following impacts:

- **Contribution to the OpenMP5.0 standard** – One of the techniques that we implemented and evaluated within the project are the iterators over dependences that have been added in the new release of OpenMP 5.0 released in November 2018. Using these iterators we are able to define multidependences among tasks (i.e., the number of dependences is decided at runtime, not compile time). We take advantage of the early implementation of multidependences in the OmpSs programming model to evaluate it in a computational fluid and particle dynamic simulation on three different architectures, an Intel-based, an Arm-based, and an IBM-based system, showing benefits in all architectures.
- **Direct impact on existing codes** – The application team developed a shared memory implementation of HPCG. Almost all state-of-the-art HPC architectures have their own closed source HPCG implementation. We started working on an improved version targeting Arm CPUs and we contributed publishing our code into the Arm repository. Further development by Arm, based on the first Mont-Blanc implementation is ongoing.

  Also, the version of Alya including the Dynamic Load Balancing library (DLB) has been improved over the duration of the Mont-Blanc 3 project and it is currently used by the BSC engineering department.

Figure 7 – Example of simulation produced with Alya

- **Indirect impact on best practices** – The Alya+DLB application, developed within Mont-Blanc 3 has been used as demonstrator within the Performance Optimisation and Productivity Centre of Excellence in Computing Applications. As a result, a UK research team working on very high scale biological simulations recently requested a POP2 proof-of-concept service to deploy the techniques in their code.

Performance modeling environment

Simulation is widely used in system design for evaluating different design options. Depending on the abstraction level considered for simulating a given system configuration, there is a trade- off between the obtained precision and speed. Generally, simulating a detailed system model provides accurate evaluation results at the price of potentially high simulation time. On the other hand, less detailed or more abstract system representations usually provide less accurate evaluation results, but in a fast and cost-less manner.

In the Mont Blanc 3 project, we have pursued the idea of ’multi-scale’ simulation where the full simulation workflow consists of various tools to address different abstraction levels of the same simulated system. The results obtained with detailed fine grain simulators (addressing lower abstraction levels) are funnelled into coarser grain tools (addressing higher abstraction levels).

We have explored a rich set of existing simulation tools to select a few of them as core components of our multi-scale simulation infrastructure. We have also implemented new simulation tools and extensions to the existing tools as a result of exploring various options to push the simulation limits both in terms of details/accuracy/speed and size (number of compute nodes) of the simulated HPC systems.

Figure 8 depicts the various components of the Mont Blanc 3 simulation infrastructure and their relationships. The vertical axis represents simulation speed of the tools relative to one another. The "higher" a component is located in the graph the faster is the simulation speed1. The horizontal axis represents the simulation "scope", i.e. what system size and level of details the tool can handle. The "scope" can spread from micro-architecture/instruction level details up to interactions among compute nodes in a full compute cluster. In general, the finer details a tool can capture the lower is the simulation speed. Simulation tools are represented by rectangles (or blobs) labeled with their names. The width reflects the range of simulation scope of the corresponding tool.

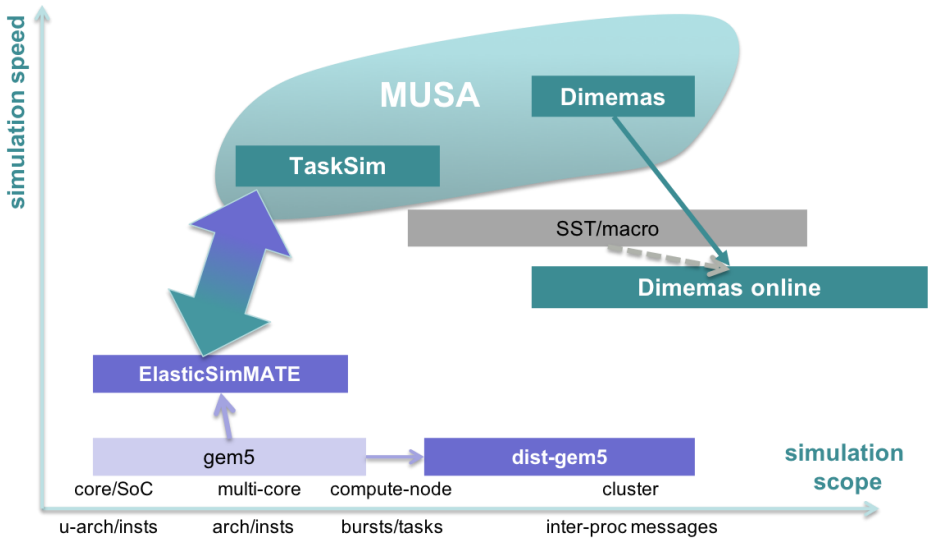

Figure 8 - The Mont Blanc 3 simulation infrastructure

TaskSim, Dimemas, SST/macro and *gem5* existed already at the start of the project. *MUSA, ElasticSimMATE, Dimemas online* and *dist-gem5* were developed as new tools within the project. They are summarised briefly as follows.

- **MUSA:** MUSA [7] is the core of our multi-scale simulation infrastructure. It is an end-to-end MUlti-scale Simulation Approach that integrates the architectural simulator TaskSim [10] with the communication network simulator Dimemas [6]. The result is a versatile tool able to model the communication network, microarchitectural details, and system software interactions. MUSA can model entire HPC systems consisting of several hundreds of compute nodes at a reasonable high speed. MUSA is used extensively for design space exploration within the project.
- **ElasticSimMATE**: ElasticSimMATE [8] is an extension to gem5 to improve simulation speed for cache and memory explorations while preserving all the fine grain details gem5 can provide. It is a trace-based simulation tool. It supports simulation of out-of-order cores and multi-threaded (OpenMP) applications by capturing instruction level dependencies and thread level synchronisation events in the trace, respectively. It also supports task based (OmpSs) applications.

- **Bridging MUSA and ElasticSimMATE**: We implemented a joint simulation workflow with MUSA and ElasticSimMATE (represented by the thick arrow between the tools in Figure 8). The joint workflow enables MUSA to take advantage of more detailed simulation by ElasticSimMATE for some selected regions of high interest. The bridging extends the simulation scope of MUSA at the fine grain end of the spectrum.
- **Dimemas online:** On the other hand, Dimemas online is an attempt to extend the simulation scope to cover even larger HPC systems. It is an extension to the trace based Dimemas network simulator. It replaces the use of application traces with SST/macro style 'skeleton applications'. Dimemas online enables more flexible and larger scale out simulations than Dimemas/MUSA does. Dimemas online successfully completed a Lulesh mini-app simulation consisting of 91,125 MPI processes in less than 24 hours.
- **dist-gem5:** dist-gem5 extends the simulation scope of gem5 towards compute clusters by modelling off-chip interconnect links and devices. As it runs the simulation of many compute nodes in parallel on a cluster of host machines, the simulation speed is roughly the same as that of vanilla gem5. Moreover, the number of simulated compute nodes scales with the size of the host cluster to run the simulation.

We have also explored options to reduce simulation time by finding and simulating only a set of representative sections of a particular application. In particular, we explored how *CERE codelets* [9] and *Barrier points* [11] methodologies can be used to speed up simulations. We have demonstrated that these methods can indeed reduce simulation time significantly achieving 10-100x speed up in many cases.

The publicly available deliverable report "D5.6 Report on Correlated and Fine Tuned Multi-Scale Simulation Infrastructure" provides a more detailed summary about the Mont Blanc 3 multi-scale simulation infrastructure.

Exploration of design space

The Mont-Blanc3 project has performed an exhaustive analysis considering several architectural components relevant for the design of next generation HPC architectures. The project investigated hybrid MPI+OpenMP state of the art benchmarks with native input sets. We extracted several conclusions from this campaign, which can be used to drive the design of next generation HPC systems.

Our results show how **compute node configurations with 512-bit wide FP units yield 20% to 75% performance speed-up and an average power increase of 60%** with respect to 128-bit wide configurations. Consequently, it is appropriate to add 512-bit FP computing units to hardware devices, since these sizes may provide energy reductions for parallel codes that properly exploit both thread level and SIMD level parallelism.

We also observe that in both 32- and 64-core processors, **cache configurations with 1MB shared L3 cache and 512KB private L2 cache per core seem to offer the best trade-off in terms of power and performance**. Moderate Out-of-Order capabilities (in terms of ROB size, Issue width, LD/ST buffers, FP Units) are a good design point in all of our tested applications. Memory bound codes benefit greatly (up to 60%) from enhanced memory bandwidth rates. Doubling the number of DDR channels (and populating them with DIMMs) increases the total node power consumption by only 10%-20%.

4. Impact, general findings, and lessons learned

Impact

MB project, real pioneer in the promotion of Arm based HPC computing, is recognized as one of the pathfinder project that boosted the emergence of this new technological trend - now turning into a mainstream technology.

During MB3 project duration – from end 2015 to end 2018 - this idea was turned into a reality. We started from a situation with no commercial solution available to reach the point where several key IT players – Atos/Bull, but also Cray, HPE and Fujitsu - propose or are about to propose commercial products.

Also, our technological choice, the Cavium/Marvell ThunderX2 chip, has been widely recognized. The Mont-Blanc project was the first to publicly announce, in January 2017, a platform development based upon Cavium (now Marvell) ThunderX2. Atos/Bull was also the first vendor to announce in June , at ISC 2017, that they would productize this design (reported for instance in https://www.top500.org/news/high-performance-computing-in-2017-hits-and-misses/). Cray followed in November, as well as HPE , Gigabyte and other vendors.

The impact of the third phase of Mont-Blanc has been widely recognized by the community, with for instance two prestigious HPCwire awards, and twice the honor of the Top500 “Hits and Misses” yearly report ; at SC 2017 we received the HPCwire Editors’ Choice award attributed to the project as ‘Best HPC Collaboration (Academia/Government/Industry)’, and our technological implementation by Atos received Editors' Choice: ‘Top 5 New Products or Technologies to Watch’.

And, last but not least, we are excited to see our modelization and design work continued within the Mont-Blanc 2020 project, providing a footprint for the European Processor Initiative (see below the *Towards an Exascale level processor* Section)

Products

The Mont-Blanc 3 project has made it possible to turn into products – commercial or Open Source products - a large part of our work. This is extremely important for us, as it now allows users to access the technology we developed and provide feedbacks on the benefits or shortcomings of these developments. We can mention here the above-mentioned Atos’s BullSequana X1310 blade, derived from our Dibona prototype, sold with the Bull SCS software suite. This is also the case for the above-mentioned improvements brought to the Arm HPC software stack.

Industrial collaborations / exploitation

The Mont-Blanc project is committed to enable the emergence of Arm-based HPC systems.

Therefore, beyond the products mentioned in the preceding sub-section), the Mont-Blanc 3 project is pushing at two levels:

a) contributing to the preparation of the future European processor and

b) contributing to the collaborations and alliances that pursue this aim:

- The Mont-Blanc 2020 project, started in December 2017, is exploiting the MB3 design space exploration work (see *Performance modeling environment* section) to

develop core IPs for future chips, with the ultimate aim to have these IPs implemented as part of the EPI (European Processor Initiative).

- We (project per se or project partners) established relationships:
  - with key communities (OpenHPC, Linaro, OpenAlliance) pursuing similar target,
  - with PRACE – invited now to test the Dibona platform.

Dissemination and educational impact

Following on from the significant training and dissemination effort made by Mont-Blanc and Mont-Blanc 2, the Mont-Blanc 3 team dedicated a lot of energy to promoting the project, its vision and results.

The main dissemination highlights for Mont-Blanc3 include:

- the overwhelming success of the press release issued in January 2017 and announcing the processor chosen for the Mont-Blanc 3 prototype [2] ;
- the excellent visibility of the project at ISC 2017, thanks to a) the announcement made by Atos (productization of our prototype in their BullSequana range [3]) and b) Mont-Blanc's presence in the press tour;
- the HPCwire award received at SC17: "Editors' Choice: Best HPC Collaboration (Academia/Government/Industry)" [4];
- Mont-Blanc's visibility at the 2018 Arm Research Summit, where we also organized a dissemination workshop and our Mont-Blanc 3 Final Event [12].
- Participation to two SVE hackathon organized by Arm (Nov. 2018 @ SC18 & Feb. 20198 @ Barcelona [13])

A total of **100 dissemination activities** were held during the lifetime of the project, including 37 scientific, peer-reviewed papers at conferences, 27 miscellaneous invited talks, as well as booths at tradeshows, articles in scientific magazines, workshops, peer-reviewed posters, etc.

Media coverage was outstanding, with 111 articles mentioning Mont-Blanc between October 2015 and December 2018. The phase 3 of the project clearly reaped the benefits of the pioneering investigations into Arm-based HPC started by the previous phases of Mont-Blanc. **As the usage of Arm processors for HPC became a mainstream topic of interest, Mont-Blanc started to be regularly mentioned by journalists and analysts whenever Arm + HPC was discussed.**

Besides the conventional trainings organized by the Mont-Blanc 3 partners, the project continued a key educational activity started by the previous projects: sponsoring and coaching a team in the Student Cluster Competition at ISC, each year. The SCC is a great way to transfer practical HPC knowledge to students, as explained by Mont-Blanc partner BSC during a workshop presentation at SC18 [5].

Towards an Exascale level processor (through Mont-Blanc 2020 & EPI projects)

There are three key hardware challenges to achieve the Exascale performance and power requirements.

First, design an **efficient processing unit** able to deliver large performance in terms of floating point computations

Second, use an **innovative on-die interconnect** able to supply enough bandwidth to the processing units with minimum energy consumption

And, finally, have **a high-bandwidth and low power memory** solution with sufficient capacity and bandwidth for Exascale applications.

The Mont-Blanc3 project explored these key hardware elements by either building an experimental Arm-based HPC platform or by launching large architecture simulation campaigns that investigated these three points as well.

The projects that are continuing the research carried out by Mont-Blanc3, the Mont-Blanc 2020 and the European Processor Initiative (EPI) projects, will deliver hardware designs of these three elements.

In particular, the Mont-Blanc2020 project is focused in delivering, among other things, a new low-power mesh interconnect based on the Coherent Hub Interface (CHI) architecture and an interposer-based high-bandwidth memory (HBM) solution to reduce the energy-per-bit for off-die memory accesses.

EPI will be focused on developing low-power processor technology to be included in a pre-exascale system for Europe in 2020-2021 and exascale in 2022-2023.

The outcomes of Mont-Blanc3 in terms of either real hardware platforms and computer simulations will be used as input elements for both Mont-Blanc2020 and EPI.

Lessons learned / Vision

The Mont-Blanc project contributed – in particular in the present action - to several scientific and technological areas thanks to the collaboration of relevant European representatives in industrial companies and research centers. The present section is dedicated to summarizing the general findings and lesson learned during the project.

Arm based solution have now a serious place in HPC arena

The last two years truly showed the emergence of Arm-based HPC solutions, with Mont-Blanc's Dibona platform and its derived product (Atos' BullSequana X1310 blade), and also with competing solutions proposed by Cray, HPE and soon Fujitsu. Arm was for instance a buzzword at the last two Supercomputing conferences. And – beyond this buzz by suppliers - we can see a vibrant expectation and strong demand from users, with deployments in the US, UK & France for instance, all based on Cavium/Marvell ThunderX2 SoC: i) HPE system deployed in Sandia NL (ranked #205 in top 500), ii) Isambard Cray system deployed in Bristol for GW4, iii) Atos system deployed at CEA).

This trend is expected to continue and get stronger, benefiting from two factors: Arm's Scalable Vector Extension (SVE) and Arm's business model:

- With SVE, Arm is significantly extending the vector processing capabilities associated with its ISA, enabling implementation choices for vector lengths scaling from 128 to 2048 bits,

facilitating code parallelism and reducing software realization, adaptation & deployment effort.
- Arm's business model facilitates the emergence of an eco-system

More precisely, we can observe a two-sided action facilitating the adoption of Arm architecture for HPC.

On one hand, chip providers can leverage existing IP modules choice to facilitate the development of Arm-based SoC with the possibility to select the very best building blocks and integrate them smoothly on a chip.

On the other hand, as Arm Architecture specification guarantees software compatibility across different micro-architectural implementations, the software effort of all providers and users can be mutualized and contribute to the constitution of the software ecosystem – crucial for the adoption.

The work must go on towards Exascale & beyond

All our experience, and our work in the Mont-Blanc project, convince us that exascale level systems are coming soon, and are possible with European technology. However, this major step forward will require – to be truly efficient and beneficial – extensive work on ALL layers:

- SoC (including potential acceleration) level,
- Hardware platform level (Interconnect & I/O's, data),
- Software level (programming models, address space, work orchestration),
- Application level (the applications are clearly part of the scope).

Also some transversal fields require further work; for instance to leverage the convergence with AI/Big Data domains and to improve the overall efficiency (with enhanced power management and resiliency capabilities)

Applications analysis & system modelling are key insights for co-design approach

In view of the application analysis activities performed, we can highlight for instance the following lessons learned:

- We show that having a set of linear algebra libraries optimized for the architecture is beneficial, not only when dealing with HPC codes (e.g., TensorFlow).
- Through our careful benchmarking with complex applications we provided Arm with insights about their positioning with respect to their competitors.
  This application analysis can be combined with modelization tools and techniques A detail analysis of application performance can give hints towards the impact of different architectural choices and potential co-design directions for future architectures. So, gathering numbers on different real systems, as we did on Dibona, combines with our simulation work to give hints on how e.g., SVE vector size or number of memory channels affects a spectrum of applications.

Application evolution: need for an hourglass approach

As part of their long-term evolution, applications must evolve to take advantage of new hardware capacities (i.e. mainly of the very high parallelism we can expect in exascale HPC clusters) while limiting the effects of inter-task synchronizations & latencies. This must be done without creating too many limiting adherences with such or such hardware evolution. Therefore, there is a global need to define new adapted programming paradigms ("taskification", parallelization), with the right abstraction level. We see here OpenMP (with its forerunner OmpSs) and SVE (with its Length Agnosticism - above mentioned) as good examples of such programming models.

Academy – Industry collaboration is good

We had the chance to have, from the early Mont-Blanc project, key EU academic institutes working closely with industrial actors (Arm & Atos among others). Although not always easy, this collaboration, putting together complementary expertise - has proved to be highly beneficial for all partners and for the project. This collaboration however requires – to deliver real benefits – to be considered early (and strongly) enough in the project life.

To conclude we would like to highlight the pleasure we had while pushing our ideas to become a market reality. Doing so, as a kind of milestone, and being able to propose a real product to users is key to assess the benefits – or shortcomings – of these ideas, and prove they are a serious option – or a path - on the road of future supercomputing.

5. Acronyms and Abbreviations

- SVE: Scalable Vector Extension (to Armv8-A architecture)
- Arm: Arm Limited - British semiconductor and software design company.

6. References

1. [DoA] Description of Action of H2020 - 671697 MONT BLANC 3 project.
2. [Mont-Blanc Press Release] 16 January 2017 - Mont-Blanc project selects Cavium's ThunderX2 processor for its new Arm based HPC platform - http://montblanc-project.eu/press-corner/press-releases/mont-blanc-project-selects-caviums-thunderx2-processor-for-its-new-arm-based-hpc-platform-8
3. [Atos Press Release] 19 June 2017 - Atos expands range of supercomputers to include ARM processors with new Bull Sequana X1310 - https://atos.net/en/2017/press-release/general-press-releases_2017_06_19/atos-expands-range-supercomputers-include-arm-processors-new-bull-sequana-x1310
4. [Mont-Blanc Press Release] 13 November 2017 - Mont-Blanc Project Receives 2017 HPCwire Editors' Choice Award for Best HPC Collaboration - http://montblanc-project.eu/press-corner/press-releases/mont-blanc-project-receives-2017-hpcwire-editors-choice-award-for-best-hpc-collaboration-academia-government-industry
5. [Mantovani, F. - Banchelli, F.] SC 18 workshop presentation - Filling the Gap between Education and Industry: Evidence-Based Methods for Introducing Undergraduate Students to HPC - https://upcommons.upc.edu/handle/2117/121732
6. [BSC] Dimenas website http://www.bsc.es/dimemas
7. [Grass, T / Allande, C / Armejach, A / Rico, A / Ayguade, E / Labarta, J / Valero, M / Casas, M / Moreto, M] MUSA: a multi-level simulation approach for next-generation HPC machines https://www.computer.org/csdl/proceedings/sc/2016/8815/00/8815a526.pdf
8. [Nocua Al; Bruguier F; Sassatelli G; Gamatie A] ElasticSimMATE: A fast and accurate gem5 trace-driven simulator for multicore systems - https://ieeexplore.ieee.org/document/8016146
9. [UVSQ] CERE web page - https://benchmark-subsetting.github.io/cere/
10. [A. Rico / A. Duran / F. Cabarcas / Y. Etsion / A. Ramirez / M. Valero] Trace-driven simulation of multithreaded applications https://ieeexplore.ieee.org/document/5762718
11. [T. Carlson, W. Heirman, K. Van Craeynest, and L. Eeckhout] BarrierPoint: Sampled simulation of multi-threaded applications https://ieeexplore.ieee.org/document/6844456
12. [Mont-Blanc @ Arm research summit] https://www.montblanc-project.eu/press-corner/news/mont-blanc-at-arm-research-summit-2018
13. [SVE Hackathon] SVE hackathon (Nov. 2018 @ SC18 & Feb. 2019 @ Barcelona) : https://developer.arm.com/hpc/events/sve-hackathon-sc18 & https://www.montblanc-project.eu/press-corner/events/arm-sve-hackathon-arm-scalable-performance-for-hpc-and-ml