D 4.1 Report on the assessment and validation of innovative solutions

Work package:	WP4	Evaluation and Dissemination	
Author(s):	Javier Bartolome		BSC
	Mirko Cestari		CINECA
	François Robin		CEA
	Dirk Pleiter		JUELICH
	Philipp Thörnig		JUELICH
	Nikoleta Kiapidou		BSC
Authorized by:	Dirk Pleiter		JUELICH
Reviewer #1	Dirk Pleiter		JUELICH
Reviewer #2	Philippe Segers		GENCI
Dissemination Level	Public		

Date	Author	Comments	Version	Status
15.07.21	All authors	First draft	V0.1	Draft
10.08.21	All authors	Comments, edits, proofreading	V0.2	Draft
30.08.21	All authors	First complete draft	V0.3	Draft
15.08.21	Internal reviewers	Internal review	V0.4	Internal review
28.09.21	All authors	Final version	V1.0	Final version

The PPI4HPC project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 754271.

Executive summary

Within the PPI4HPC project, a group of leading European supercomputing centres established the formation of a buyers group to run a joint Public Procurement of Innovative Solutions (PPI) in the area of high-performance computing (HPC). The involved HPC centres, namely BSC, CEA/GENCI, CINECA and JUELICH, have a strong track record in providing supercomputing resources at the European level and are founding and active members of the Partnership for Advanced Computing in Europe (PRACE[1]), that already performed a successful Pre-Commercial Procurement (PCP) before this PPI. The joint procurement[2], led by GENCI as lead procurer, was organised in four independent lots leading to the awarding of contracts for systems deployed at four different European research sites. These systems include innovative high-performance supercomputers and storage system for science and engineering applications. PPI4HPC main goals were to significantly enhance the planned pre-exascale HPC infrastructure and set an example for joint European investments in the future:

- More innovative supercomputing resources for science and engineering applications in Europe
- Strengthened European research and development on HPC technologies
- Innovation in the design of the solutions needed by scientists and engineers in Europe

The D4.1 “Report on the assessment and validation of innovative solutions” is related to the activities that fell under Task 4.1 “Evaluation of the PPI process”. The document explains the process of the PPI4HPC coordinated procurement procedure, provides the lessons learnt from this process from a technical point of view, and offers details on the validation and assessment of the purchased technologies.

Overall, PPI4HPC addressed major scientific, industrial and societal challenges. It became the first joint European procurement of innovative HPC systems and contributed to the upgrade of a pan-European HPC infrastructure to serve research and engineering. It created an impact on the market by fulfilling existing market requirements and needs, and set the example for future joint procurement in HPC in Europe. It strengthened R&D on HPC architectures and technologies and newly designed solutions according to the needs of scientists and researchers in Europe and set up a coordinated roadmap and approach for providing optimised HPC resources that can be used for future joint procurements. Furthermore, it emphasised the importance of energy efficiency requirements in the tender procedure in order to achieve display of carbon footprint data by vendors and supercomputing sites.

[1] https://www.ppi4hpc.eu/call-for-tender
[2] https://prace-ri.eu/infrastructure-support/pre-commercial-procurement/

Contents

1. Introduction

The D4.1 report contains the details related to the activities carried out during the evaluation of the PPI process done within Task 4.1 “Evaluation of the PPI process”. This task defined the evaluation process that needed to be done by each of the lots and their innovative technology deployed during the test phase, in order to study the quality and efficiency of the improvements that each site has provided. These evaluation tasks were performed by each of the procurers once each lot was installed.

The results generated out of these evaluations along with the best practices and recommendations that the procurers have gathered from the joint procurement process have been collected and analysed in this deliverable. At the same time, the content of this report was designed to accommodate the European Commission’s (EC) review recommendations received in October 2020 after the review conducted in May 2020. The recommendations generated from the lessons learned in terms of legal issues have been described in a separate Whitepaper titled “Lessons Learned on Legal Aspects", which was published by the consortium in October 2020 and can be downloaded from the PPI4HPC website[3].

The structure of this document is the following: The procurement procedure is described in Section 3. This section also includes a procurement market analysis and an assessment of lessons learnt which was derived from the technical angle of the coordinated procurement procedure. Section 4 explains the methodology used in the evaluation of innovative technologies for each site. Section 5 describes the evaluation results which includes details on the most important innovative technologies per site. Section 6 offers an analysis on the PPI4HPC impact on the various stakeholders (suppliers, users, and HPC centres) as well as the lessons learnt in terms of the quality, benefits, and challenges of the coordinated procurement. Section 7 provides an overview of each PPI site regarding green policies and actions to reduce carbon footprint. Finally, the conclusion sums up the main recommendations from the previous sections.

2. Procurement procedure

A number of activities were carried out by the public procurers in order to establish a common framework - in terms of objectives, requirements, common terminology, evaluation criteria - to coordinate the writing of the technical specifications. This effort has helped to homogenise and improve the procurement tenders of the lots while decreasing the risks associated with the procedures.

After the joint selection of the qualified candidates was completed in August 2018, the public procurers started the procurement process for each local lot. In order to coordinate these activities, regular meetings were conducted to share relevant details for the success of the project or best practice adopted in the procurement process. In particular, these meetings allowed partners to share experiences and provide advice to other partners, and provided the opportunity to ask questions and receive feedback between partners at different stages of a lots’ procurement process. The outcome was a significant sharing of best practice examples between the involved partners and, as a consequence, a shared process with regular checkpoints in order to evaluate the project progress status and likely or unlikely risks. Among the shared examples, the following best practices can be cited: i) organisation of a market consultation event; ii) shared experience and criteria on refining the tendering documents,

[3] https://ppi4hpc.eu/news/ppi4hpc-whitepaper-%E2%80%9Clessons-learned-legal-aspects%E2%80%9D-just-published

during the negotiation phase; iii) management of the bidder's question during the single lot procurement process; iv) Non-disclosure topics were addressed in order to find a balance between sharing information while avoiding breaching any confidentiality agreement with providers; v) organisation of notification and publication of the final award by GENCI in the context of a coordinated procurement.

The procurement procedure went as planned and all the lot contracts were awarded due to successful tendering procedures.

2.1 PPI4HPC procurement market analysis

To better frame the technical specifications and project the possible offered solutions for each local lot, a joint market analysis[4] was conducted by the group of public procurers. The analysis was conducted as one-to-one meetings[5] between the technology provider (vendor) and the group of public procurers, and was open to participation by all interested technology providers in the HPC field.

The market analysis can be considered successful in informing the market on the objectives of the procured lots. Considering the vendors' system share of the top500 list accounted for over 90% of the systems in the list as of November 2018. These vendors were:

- Atos
- Cray (later acquired by HPE)
- Dell EMC
- Fujitsu
- HPE
- Huawei Technologies Co., Ltd.
- IBM
- Inspur
- Lenovo
- Sugon

The procurements of the local lots have seen the participation of five vendors included in the above list (IBM, Atos, Lenovo, HPE, Cray) plus one vendor that failed to submit a proposal for participation in time. Furthermore, two SMEs (Megware and E4) were involved directly as qualified candidates in the procurement process or as service providers in the best and final bids for the lots. Therefore, the market information and analysis conducted within the project can be deemed to be successful in attracting key players in the HPC field, and placing them in a position that allowed them to participate effectively in the procurement procedures. Moreover, the objectives of the joint procurement have benefited from the market analysis, leading to technical requirements more aligned with the evolution roadmap of the main

[4] https://www.ppi4hpc.eu/call-for-tender/questions-and-answers
[5] https://ppi4hpc.eu/events/one-one-technical-meetings

technology providers. This is also confirmed, to some extent, by the response to the common technical needs reported in the table of section 5.2.

2.2 PPI4HPC procurement procedure

The highly innovative offers resulting from the procurement procedures have clearly benefited from the vendors' competition. This is a result of two key aspects:

- the selected procurement procedure
- the coordination process between the public procurers in drafting the technical specifications and sharing best practice

The selected procurement procedure was the competitive dialogue that provided sufficient flexibility during the procurement of particularly complex projects. These include highly innovative solutions that have still not entered the market or have just recently entered the market with a small share.

A number of activities were carried out by the public procurers in order to establish a common framework - in terms of objectives, requirements, common terminology, and evaluation criteria - to coordinate the writing of the technical specifications. This effort has helped to homogenise and improve the procurement tenders of the lots, and decrease the risks associated with the procedures.

A summary of the purchased systems per lot is collated in the tables below.

Site	Manufacturer	Cores	Memory	Processor	Interconnected	Linpack Performance (Rmax)	Theoretical Peak (Rpeak)	Power
CEA/GENCI (main compute partition)	ATOS	293,376	394 TB	AMD Rome 7H12 64C 2.6GHz	Mellanox InfiniBand High Dynamic Range (HDR)100	6,988.04 TFlop/s (partial)	12,039.4 TFlop/s	1,436.00 kW (Submitted)
CINECA	IBM	347,776	252 TB	IBM POWER9 16C 3GHz	Dual-rail Mellanox EDR Infiniband	21,640 TFlop/s	29,354* TFlop/s	1,476.00 kW (Submitted)
JUELICH	ATOS	449,280	628 TB	AMD EPYC Rome 7,402 24C 2.8GHz	Mellanox HDR InfiniBand/ ParTec ParaStation ClusterSuite	44,120 TFlop/s	70,980 TFlop/s	1,764.22 kW (Submitted)

Table 1: HPC systems and features

Site	Manufacturer	Data storage	Metadata storage	Tape Storage	Network Interconnect	Service Servers

BSC	IBM	4.54 PB, ESS GL5S (HDD Based)	120 TB, ESS GS2S (Flash Based)	32 PB, 2,640 LTO8, 3 Tape Drives	Mellanox SN3800 64-port 100GbE	2x Archive servers + 6x Data mover servers 4x Cloud servers 1x HPDA[6] server

Table 2: HPC storage and features

3. Evaluation of innovations: Methodology

This section explains the common methodology used in the evaluation of the innovative technologies along with the evaluation results for each PPI4HPC site.

When specifying the procurement technical requirements, a mandatory requirement for each provider was to establish the definition of an innovative technology to include in their solutions. This definition helped to clearly identify innovation components provided by each solution/vendor. In order for the sites to evaluate each of those innovative components, each center had to collect the following information:

- Innovation: Brief technical description of the innovative technology
- Criteria to assess the innovation: Description of the criteria and characteristics that serve to assess the innovation technology
- Assessment tasks for validating the innovation: List of concrete tasks that are performed to validate that the component delivers innovation and improvement compared to previous/current technologies or solutions
- Impact on the market: Description of possible impact of the innovative technology on the HPC market/solutions once the new technologies are put in production and deployed

The table below provides a summary of the innovations identified and the information that was needed to perform the evaluation (criteria to assess the innovation, assessment tasks for validating the innovation and impact to the market).

BSC			
Innovation	**Criteria to assess the innovation**	**Assessment tasks for validating the innovation**	**Impact on the market**
Software: Spectrum Archive: • Integration of the data flow cycle inside a unique filesystem vision	• Adaptation of data user workloads independently of storage tier used • Number of parallel operations of copy and retrieve from tape storage	• Compare with other HSM[8] software capabilities or/and previous archive storage at BSC • Increase researchers storage capacity without affecting time-	Use of standard tape format that could permit migration of data between centers without vendor-locking

[6] High Performance Data Analytics
[8] Hierarchical Storage Management

• Usage of an open format for tape storage (LTFS[7]) • Parallelism and scalability in terms of Archive servers distribution and failover	• Reliability and failover during production in front of an archive server	to-solution of their workloads	
Software: ESS[9] Distributed RAID (SSD[10] or HDD[11] based): • Reduction of recovery time after hard drives failures • Instead of only 2 drives participating in the recovery, all disks do it	Mean recovery time in front of a HDD failure	Compare recovery vs other storages based on classic RAID6 (time recovery per TB[12] to restore)	Most vendors are introducing distributed RAID[13] technologies in their portfolio
Hardware/Software: Use of RDMA[14] over Converged Ethernet (RoCE[15]) in the: • RDMA permits movement of data between applications memory (low latency) • RoCE implements that protocol on Ethernet networks	• Raw performance (iperf / nsdperf) improve in the use of RoCE • Real workload filesystem performance (ior or similar) with the use of RoCE	Comparison with performance in same network without RoCE	Improve performance of classic Ethernet networks for data bulk transfers
CEA/GENCI			
Innovation	**Criteria to assess the innovation**	**Assessment tasks for validating the innovation**	**Impact on the market**
Hardware: AMD Rome with 64 cores @2.6GHz	Proper and performant functioning of the processors on real workloads in production	Common HPC benchmarks results, issues running usual workloads on these systems	Diversify the x86_64 processor offer to help users/communities design applications/libraries for HPC oblivious to particular vendors First system based on AMD Rome processors

[7] Linear Tape File System
[9] Elastic Storage System: https://www.ibm.com/products/elastic-storage-system
[10] Solid State Drive
[11] Hard Disk Drive
[12] TeraByte
[13] Redundant Array of Inexpensive Disks
[14] Remote Direct Memory Access
[15] RDMA over Converged Ethernet

			available to European scientists (nov 2019)
Hardware: DragonFly+ interconnect topology	Proper and performant functioning of the fabric using this new type of topology	Common HPC benchmarks results, validation of the positive effect of the adaptive routing to avoid performance reduction in such topology	Vendors working on the support of new topologies requiring performant dynamic routing thus enabling to reduce network cost while ensuring optimal efficiency of the available hardware (links/switches)
Hardware: Support of hot processors (280W) into dense nodes	Proper functioning of the nodes in production without power consumption throttle	Proper functioning of nodes equipped with AMD Rome 7H12 (280W) in production	Improve cooling capacity/efficiency of vendors HPC rack solutions
Hardware: Efficient water-cooled power supplies	Proper and performant functioning of the racks using this technology	Proper functioning and evaluation of the gain compared to traditional air-cooled power supplies	Optimised Power Usage Effectiveness (PUE[16]) by reducing energy waste at all the levels when possible.
Software: Advanced energy management (BEO[17], BDPO[18] tools)	Effective usage on a production cluster	Successful integration and power constraints on real production systems	Introduce software defined power consumption control of datacenter to guarantee consistency with power/cooling production capability over time
Hardware: New generation storage with SBF (Smart Bunch of Flash)	Efficient provisioning of temporary per-user file systems	Applications adapted to benefit from these new mechanisms	Introduction of a cheaper but efficient first storage level for transient needs in the HPC architecture
Hardware: New generation of ARM processors	Proper and performant functioning of the processors on real workloads in production	Applications adapted to run on this new architecture and showing state-of-the-art performances.	Diversify the processor architectures offer to help users design applications/standards for HPC oblivious to particular technologies/vendors.
Hardware: BXIV2[19] interconnect	Proper and performant functioning of the fabric using this new interconnect topology	Common HPC benchmarks results, applications running on this interconnect and showing no performance degradation compared to similar interconnect technologies (100Gb/s).	Diversify the interconnection network offer to help users/developers design applications/standards for HPC oblivious to particular technologies/vendors.
CINECA			
Innovation	**Criteria to assess the innovation**	**Assessment tasks for validating the innovation**	**Impact on the market**

[16] Power Usage Effectiveness
[17] Bull Energy Optimser.
[18] Bull Dynamic Power Optimizer
[19] Bull eXascale Interconnect v2

Hardware: Highly dense system providing significant performance/watt ratio	Proper and performant functioning of all hardware components that allow to reach the performance target(GPUs, Dragonfly+ Interconnect, CPU-GPU interconnect) for the given power envelope	Common HPC benchmarks run on the system and validation of the benchmark's commitments for the cost performance analysis	With the PPI4HPC, leading European HPC centers have for the first time procured hybrid (CPU/GPU) Tier-0 class systems, fostering their adoption in Europe. Hybrid systems are providing among the highest flops/watt ratio according to the most recent top500 lists. At least two precursors of exascale systems procured by EuroHPC are hybrid systems based on GPUs. Main technology providers (AMD, NVIDIA, Intel among others) all include in their roadmap evolution or new GPU devices.
Hardware/Software: Innovative storage NVMe[20] over fabric for Burst Buffer capability	Create a high-performance filesystem at user level to benefit IO intensive user workloads	IOR and other IO benchmark tools performed on Beeond filesystems	IO intensive workloads are typical in AI and deep learning analysis. Introduction of a cheaper but efficient first storage level for transient needs in the HPC architecture.
Software: Online performance and energy monitoring with Examon framework	Provide an online energy and performance monitoring to identify hot spots (such during the installation phase) or malfunctioning nodes.	Continuous data collection over the lifetime of the HPC system	Framework to track energy and performance of applications running on HPC systems are available on the market (such as BEO from ATOS)
JUELICH			
Innovation	**Criteria to assess the innovation**	**Assessment tasks for validating the innovation**	**Impact on the market**
Hardware: Next-gen. processors improving energy-to-solution for applications & increasing performance-per floor space • AMD EPYC Rome processors & NVIDIA A100 GPUs	Performance and energy-to-solution comparison to previous systems	Validation of benchmark commitments in acceptance (i.e., verification of application performance improvements w.r.t. to successor system)	Promote diversity of different manufacturers and keep user applications portable
Software: Energy management with BEO and BDPO: • Power capping capabilities were	Power monitoring capabilities. Ability to	Power monitoring of benchmarks. Comparison of benchmark performance	Strategies to meet a set power budget without noticeably reducing the performance for users. Improved power monitoring capabilities.

[20] NonVolatile Memory Express

very important feature for JUELICH ● BEO provides advanced power management capabilities for CPUs & GPUs	maintain a fixed upper limit power consumption with low/acceptable application performance degradation.	with and without power caps.	
Hardware/Software: Next-gen. NVM storage based on SBB and SBF[21] technology ● 400 TB NVM storage with >250 GB/s accumulated ● SBB & SBF advantaged storage management software	Support for data-intensive applications though node-local fast random access storage. Convenient storage management via the workload manager	Comparison of benchmark using NVM vs. parallel file system	A Slurm integrated storage tier to support a demanding data I/O during application runtime
Hardware/Software: Mellanox DragonFly+ topology	Effective bandwidth vs. costs	Performance analytics using benchmarks	More efficient and high-performance software algorithms that lead to more cost-effective and more flexible hardware solutions

Table 3: Summary of innovations per site

The innovations that are in bold letters are the ones that each site identified as the most important and were, therefore, chosen to be analysed along with their assessment criteria and validation tasks in detail. An overview of evaluation results, including the features of each innovation, the specific technical criteria and the methods employed by each partner to assess each innovation is provided in detail in the individual sub-sections of the chapter that follows. Those detailed evaluation results are provided in section 4.

4. Evaluation of innovations: Main results

4.1 BSC

4.1.1 Spectrum archive

Innovation

In order to assess the innovative aspects of the Spectrum Archive software, we will first describe previous archive solutions deployed at BSC and their flaws.

In 2006 a tiering storage solution installed at BSC was based on Sun SAM-QFS with a capacity of 3PB, which was decommissioned in 2012 due to technology constraints that were not fulfilling user needs. The main limitations are listed below:

[21] SBB – Smart Burst Buffer / SBF – Smart Bunch of Flash

- Could not be integrated seamlessly with HPC filesystems, implied a copy of data between HPC and a specific filesystem with very limited capacity on disk for archiving. Only the NFS protocol could be used for accessing this data.

- Tiering policies were the same for the whole archiving system, and limited control over how data could be migrated to tape.

- Archive system offered limited scalability as only 2 servers could be devoted to archiving tasks, which could easily be overloaded from accessing the filesystem while data movement operations were taking place between the tape and storage tier systems.

After a market analysis in 2013, BSC decided to acquire a new archive storage solution built exclusively out of hard drives, eliminating the need for any tiering solutions. This solution was based on a parallel file system that could be integrated completely with BSC HPC clusters.

In 2019, after 6 years of production, archive capacity was becoming one limitation factor. However, this was not the only issue, as the archiving technology had a high maintenance and operational cost with a high failure rate on some specific components, namely metadata disks based on FiberChannel configured as RAID1.

The PPI4HPC project offered the opportunity to BSC to acquire an innovative storage solution that was sustainable in terms of cost per PB and maintenance, and with the expectation to solve all technological flaws seen in previous archive solutions deployed on site.

A tiered solution based on Spectrum Archive software was the solution selected in the PPI4HPC competitive dialogue to cover these needs. The following section discusses these technological innovations in detail.

Assessment criteria and tasks to validate the innovation

Spectrum Archive provides the ability to define a multitude of unique data migration policies for how data is transferred to tape, which can then be applied to different research group's directory/fileset. Each policy is defined using an SQL-like statement that provides a rich and flexible way to define the policy.

All this has fostered a better adaptation to the different research groups and their data management plan needs, making the tiering solution more transparent than any other previous archive solution. It has also enabled BSC to implement diverse policies to data originating from a range of sources such as HPC workloads, Earth-Dust simulation data[22] and EGA genomic data[23].

[22] See: https://ppi4hpc.eu/news/ppi4hpc-storage-system-bsc-used-explore-how-dust-storms-affect-societies-major-climate-project
[23] See: https://ppi4hpc.eu/news/ppi4hpc-system-bsc-employed-massive-european-genome-phenome-archive-0

The new archive solution proposed to use the innovative LTFS open standard format[24] to store data on tapes. This standard was defined in 2010, and spectrum archive is the first HSM tool by IBM to support this type of format.

Previous tape formats, including old BSC archive solutions based on tape, used a proprietary format to store information on tape with the index indicating the information stored in each tape is saved externally. When tapes with data needed to be exchanged between institutions, they were either forced to have the same archive solution on both ends, or extract information from tapes to disk and then manually store it again in a general format (tar).

BSC had to follow this cumbersome procedure several times during the previous tape archive solution which contributed to a large waste of time and manpower.

The LTFS open standard format includes index information in each tape, being able to read the data that is contained on the tape without any other extra procedures for conversion. Many of the companies have implemented the format in their solutions, favoring interchange of data between sites regardless of their tape library vendor. This includes the following solutions:

- IBM Linear Tape File System
- Oracle's StorageTek Linear Tape File System
- HP Linear Tape File System
- Quantum Linear Tape File System

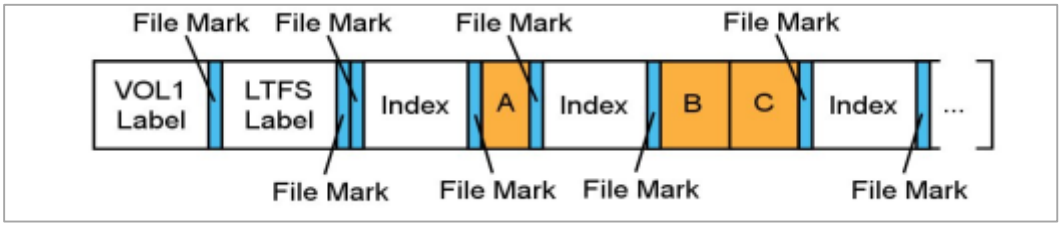

Figure 1: LTFS open standard format

Finally, an obvious improvement came from the operational cost reduction. The new archive solution provided to BSC, compared to previous archive solutions, finally provided an obvious improvement from operational cost reductions due to things such as the elimination of tasks such manual data conversions between archives across multiple sites and vendors.

Another innovation is the combination of Flash + HDD and tape, which also reduced operating costs (maintenance and also power consumption). In the table below, we summarise a comparison of the different archive solutions deployed at BSC and the PPI4HPC system:

[24] See: https://www.snia.org/tech_activities/standards/curr_standards/ltfs

	Technology	PB	kW / PB	K€ maintenance year / PB
Archive 2006	HDD + Tape	3	4	20
Archive 2013	HDD	6	16	25
PPI4HPC 2019	Flash + HDD + Tape	120	0.6	1.42

Table 4: Comparison of the archive solutions at BSC

4.1.2 ESS Distributed RAID

Innovation

To provide additional reliability in terms of hardware failures, IBM Spectrum Scale Native RAID is used. This is a software-based RAID solution, where information is distributed in a way that recovery time in case of hardware issues (Hard drive or SSD failures) is significantly reduced.

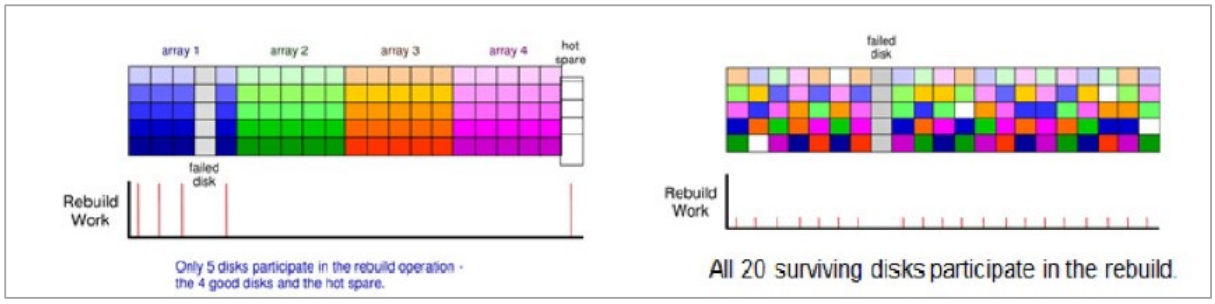

Figure 2: RAID software recovery function

As can be seen in the figures above, with this technology all disks participate in the rebuild when a failure occurs. For the PPI4HPC system, this applies to HDD and flash devices that store data and metadata respectively.

Assessment criteria and tasks to validate the innovation

A comparison was made in terms of hardware failures comparing PPI4HPC archive solution with the previous BSC archive solution, which was not based on distributed RAID during the first year of operation.

The previous archive infrastructure reported 22 hard drive failures out of a total of 1440 failures that occurred during the first year of production in 2013.

Distributed RAID provides the possibility to diagnose and predict possible failures of components. During the first year of production of the PPI4HPC system, 140 events of a component that entered in a diagnosed state, were registered. From these, 37 events were concluded that needed to be drained of their data to other drives, but only 14 events

concluded in a drive finished as failed. Also during this year, none of the SSD/Flash drives needed any replacement.

All these intermediate states of components were not found in previous BSC archive solutions based on classic RAID controllers, in which a drive could be healthy or in failure state that triggered a recovery of the RAID to a hot-spare device.

Upon a hardware failure, recovery time is lower in a distributed RAID solution compared to a classical RAID controller. This has also been compared between recovery failure times in the PPI4HPC system and the previous BSC archive solution.

In the previous archive solution, a recovery of a failing drive had a mean duration of 8 hours, which also impacted on the performance of the RAID controller. With the new distributed RAID technology, the time for recovery of the failing drive depends on the amount of data stored. At the same time, the mean recovery times observed are approximately 10 minutes, considering that most of the time data is drained from suspicious disks previous to its real failure.

Finally, as all hard drives participate in this recovery there is no measurable impact on the performance of the whole system that has been detected during the rebuilding process.

4.2 CEA/GENCI

4.2.1 Support of hot-processors (280W) into dense nodes

Innovation

Supercomputers' increased compute power comes with an increase in the associated power consumption. required to make the compute nodes components work as expected. Additional power is consumed from the energy required to extract the generated heat to keep them at their proper functioning temperature.

Advanced and highly efficient processors, network cards and accelerators tend to have a increased power consumption with each new generation. This pushes vendors to innovate in the way they are designing and manufacturing their solutions to keep up with this trend.

The GENCI/CEA PPI4HPC machine was the first large production machine to employ the AMD EPYC Rome 7H12 (280 Watts) processors in a dense cabinet. The innovation proposed by Atos, to integrate those processors, was the design of a direct liquid cooled rack consisting of up to 32 blades, each capable of integrating 3 dual-processor nodes at full thermal capacity: the XH2000 rack series associated with the X2410 AMD blades.

Assessment criteria and tasks to validate the innovation

The High Performance Linpack (HPL) is a traditional benchmark that pushes the node's power consumption to high values that exceed values commonly monitored in day-to-day production. The following diagrams illustrate the proper behavior of a single XH2000 cabinet with 96 AMD EPYC Rome dual-processors nodes running the HPL benchmark. During the 80 minutes of the HPL execution, the nodes are performing their computation at full speed. The power consumption of each node is monitored and aggregated to evaluate the whole rack power usage in kW. The higher the power consumption, the larger the hydraulic valves of the rack cooling system are opened to cool down the internal coolant (secondary loop) and keep the incoming water at the desired temperature into the blades.


Power consumption
80 kW
70 kW
60 kW
50 kW
40 kW
30 kW
20 kW
10 kW
10:50
11:00
11:10
11:20
11:30
11:40
11:50
12:00
12:10
12:20
min
max
Total power consumption
20.4 kW
69.6 kW
Cooling fluid temperature (Rack internal)
30 °C
25 °C
20 °C
15 °C
10 °C
5 °C
0 °C
10:50
11:00
11:10
11:20
11:30
11:40
11:50
12:00
12:10
12:20
min
max
Cooling inlet temperature of instance
24.00 °C
28.00 °C
Cooling outlet temperature of instance
20.00 °C
22.00 °C

Figure 3: Behaviour of a single XH2000 cabinet with 96 AMD EPYC Rome dual-processors nodes running the HPL benchmark

The criteria used to assess this innovation was to ensure the proper execution of code on the nodes without having to throttle down the power consumption by lowering the computing power of the nodes. throttling due to cooling limitations were not observed during the HPL benchmarks, the initial tests nor code executions in production, thus validating this innovation.

4.2.2 Efficient water-cooled power supplies

Innovation

Providing electricity from the grid down to the nodes involves different steps of electricity conversions. These conversions can have different power efficiencies and impact the global power usage effectiveness of the whole system.

The Atos XH2000 series integrates power supply units that convert the 220V AC input from the datacenter to the 54V DC used as the power input on the compute blades and the interconnection network switches of the rack.

In order to improve the efficiency of the PSUs, Atos used a new generation of PSU with a high conversion efficiency as well as an innovative cooling design consisting of a recirculated hydraulic cooling loop in the rack, the one used for direct liquid cooling of the compute blades and switches, to also cool down these electrical converters. The PSUs can be used in both cold and warm water cooling scenarios.

Assessment criteria and tasks to validate the innovation

Using DLC[25] to cool down the PSUs has multiple benefits. First, it enables them to increase their efficiency by reducing the amount of electricity once wasted on spinning fans that were traditionally used for cooling. Second, it reduces the quantity of air cooling needed and therefore both improves PUE and reduces the investments needed in terms of air cooling equipment. Less hot air going out of PSUs means less hot air to be cooled down by the traditional air cooling system. Third, by injecting more energy into the cold or warm cooling loop of the datacenter, it enables the reuse of this waste energy to heat up buildings when necessary, further reducing the PUE.

To assess the interest of this innovation, studies were conducted to evaluate the gain of using such a new technology in that area. In the studies, comparisons of thermal power were performed on two prototype racks, one using traditional air cooled PSUs, the other using the new approach. The following diagram illustrates the observed efficiencies of the two types of PSU for different power load. It should be noted that the better efficiency of water-cooled PSU comes mainly from the gain of efficiency due to water-cooling (removal of fans, better cooling of components) but also from a never design of PSU.

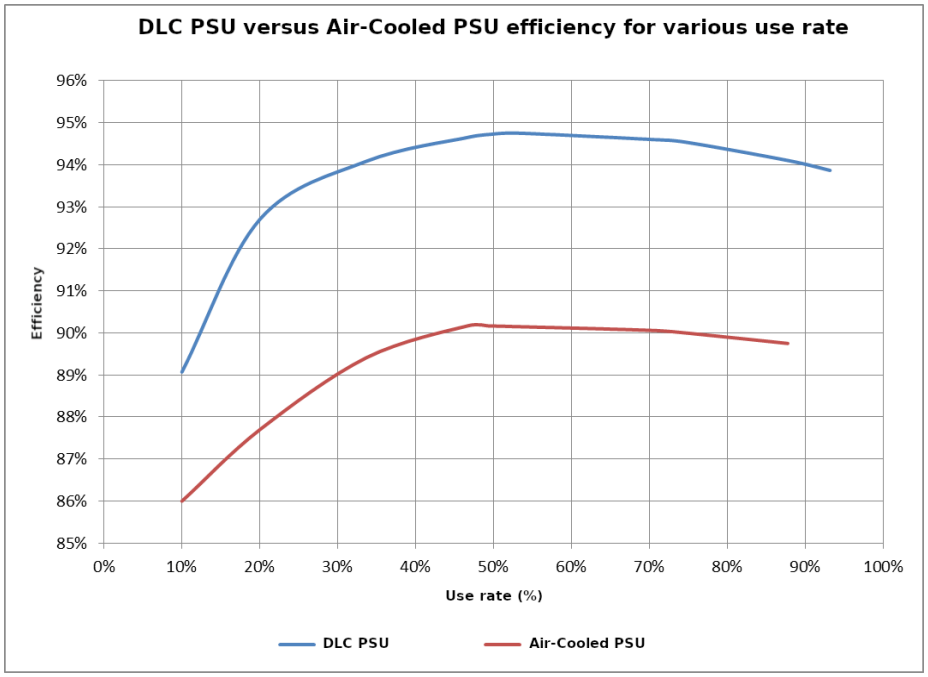

Figure 4: Efficiencies of the two types of PSU for different power load

[25] Direct Liquid Cooling

We can see a 4 to 5% improvement of the efficiency of the new generation of PSU compared to the previous one.

Comparing the measures of the DLC PSU with the 80 PLUS grades, we can see that the following type of PSU is placed between the Platinium and the Titanium grades in this certification system:

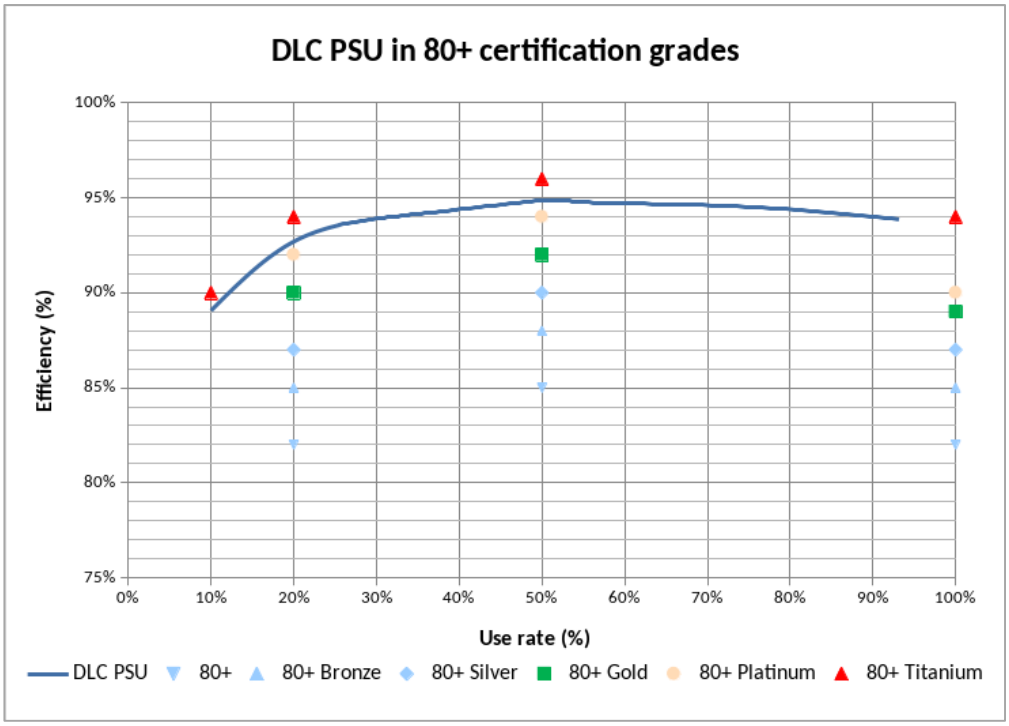

Figure 5: DLC PSU in the 80 PLUS grades

Overall, the efficiency improvement of these new PSUs, combined with the dissipated energy injected in the primary cooling loop of the datacenter as well as the reduction of the amount of dissipated energy in the ambient air, leads to a total gain of about 7% of the global power consumption of a warm water cooled compute center, validating the interest of this innovation.

Finally, it is important to note that the R&D of this water-cooled PSU had been funded by a previous EC PCP (Pre- Commercial Procurement), performed by PRACE, involving GENCI, CEA and other PRACE partners (JUELICH in Germany, CINECA in Italy, EPCC in UK and CSC in Finland) towards Energy Efficient HPC systems and components[26]. It is the first time in the HPC domain in Europe that a full cycle with a PCP and then a PPI is used.

4.3 CINECA

4.3.1 Innovative storage solution

Innovation

[26] https://prace-ri.eu/infrastructure-support/pre-commercial-procurement/

This part of the document aims to describe a technology solution based on high performance storage NVMe[27] in order to provide a parallel filesystem (PFS) with burst buffer capabilities.

Supercomputers' computing power has been increasing exponentially for decades, however, the I/O performance has only increased linearly over the same time. This has generated an I/O bottleneck that can cause large delays when large amounts of data need to be written in the parallel filesystem. To alleviate this issue, storage layers have been added between the main dynamic random-access memory (DRAM) and the PFS. Such layers are often based on Non-Volatile Memories (NVM).

NVM based storage achieves performance (bandwidth and IOPS) that are superior to a common hard disk (HD) by an order of magnitude. In the last few years, the scientific community started to leverage these devices as "Burst Buffers" (BBs) to speed up data processing and analysis in the HPC environment.

BBs can act as a fast storage system to support various use cases for HPC applications:

1. hidden cache for PFS access
2. checkpoint/restart (C/R)
3. Staging data in BBs can support operations such as in-transit data processing (interactive computing), which may include data visualisation or analysis

Currently each NVMe on Marconi100 has been formatted with the XFS file system with a block size of 4k and accessible via path '/scratch_local' for each computing node. This small block size allows to rapidly access tiny files thus utilising all the available IOPS of the NVMe. For example, high IOPS value is critical for deep learning programs, e.g., usually the "Image Recognition" algorithms need to process simultaneously many images < 1 MB, a high IOPS value allows to load and unload rapidly these tiny files to DRAM (of CPUs or GPUs).

In this initial configuration, each NVMe provides storage that is local to the compute nodes (/scrath_local) and therefore each of them are independent. However, as we will see in the next section, it is possible to use them in conjunction with BeeOND.

Assessment criteria and tasks to validate the innovation

BeeOND, which stands for BeeGFS ON Demand, allows for the aggregation of multiple NVMes in order to combine their performances and capacity in a single parallel filesystem. BeeOND enables easy on-the-fly creation of Multiple BeeGFS instances in user space.

In general, it can be used to aggregate storage capacity coming from:

- RAMs
- NVMes
- SSDs
- Hard Disks (based on BeeOND config)

[27] NVMe is a protocol (e stands for express in NVMe) that works on NVM types of physical storage.

A BeeOND instances can be built statically with admin privileges, or dynamically via workload manager (Slurm) at a user level thus aggregating resources from the nodes for the duration of a job. In Marconi100, BeeOND was adopted to build a BeeGFS PFS on top of the NVMes filesystem (aggregating several /scratch_local across multiple nodes). When a BeeOND instance is created, it is typically required to tune some additional variables:

- Numtargets: amount of storage devices to aggregate
- Chunksize: size of chunks in which a file is split and shared among the targets
- TuneFileCacheType: defines if the client server has to create a buffered file cache
- TuneFileCacheBufSize: define the size of the buffered file cache on the client server

These variables have an impact on the communication between client and storage servers in several ways. The steps involved in a write process are:

1. Client queries the metadata servers
2. Client locates the storage target with a data chunksize to modify
3. Client sends a modification message to the storage target

The maximum size of such messages is determined by chunksize. If the amount of data written to the file is larger than chunksize, more modification messages will be required to be sent to the servers. If the number of modification messages increases too much, it may cause performance loss. Increasing the chunksize has the effect of reducing the quantity of modification messages and this can have a positive performance impact, even in a system with a single storage target.

In any case, there is no universal value for the chunksize, but it depends on the kind of use cases the hardware should support. For example, a chunksize of 4k (same blocksize of the underlying xfs) would be appropriate for dealing with many tiny files. In this case, the number of requested IOPS would also increase given the higher number of modifications messages (such effect can be seen in the plot above "IOPS vs Blocksize [kiloBytes]"). This tiny chunksize would not be convenient for I/O operations on large files (hundreds of GB) for which a larger chunksize would reduce the IOPS requested. Currently in Marconi100, BeeOND is set with a chunksize of 1 MB.

In summary:

1. Small chunk size (~kB):
    1. High IOPS
    2. Many system calls
    3. Good for I/O bandwidth on many tiny files
    4. Bad for I/O bandwidth on few large files possibly due to the overhead from system calls
2. Big chunk size (~MB):
    1. Low IOPS

2. Few system calls

3. Bad for I/O bandwidth on many tiny files, inefficient modification messages

4. Good for I/O bandwidth on few large files, no system calls overhead

RDMA buffer size is an important constraint. It is important to make sure that a data chunk fits into the RDMA buffer available on the client in order to prevent the modification messages from being split before they are transmitted over RDMA. If the client server is using the buffered cache (tuneFileCacheType = buffered), it accumulates in a file cache buffer. This data is sent to the storage servers only when data from outside the boundaries of that buffer is needed by the client. The larger this buffer, the less communication traffic that will be needed between client and servers. In order to reduce unnecessary network traffic, it is important to set this buffer size to a multiple of the data chunk size: e.g. adding tuneFileCacheBufSize = 2097152 to the client configuration file will raise the file cache buffer size to 2 MB and it can accommodate 2x multiple of a chunksize of 1 MB.

After a BeeOND instance is created (mounted on /mnt/BeeOND), standard tools such “cp” or “rsync” can be used to transfer data from/to this high-performance storage. It is worth noting that when BeeOND is used inside Slurm, the temporary data created during the job run are removed automatically when the job ends, so the user is required to move non-temporary data (stage-out) before the job terminates.

The FIO benchmark tool was used to collect performance figures of an NVMe mounted as “/scratch_local” in a single node of Marconi100. FIO features three main execution variables:

1. Blocksize: It it is the size of the pieces the I/O is issued

2. Numjobs: number of clones of the job. Each clone gets an independent thread. Used in combination with the flag “--group_reporting” shows the statistics of all the clones together

3. Iodepth: number of I/O units of each job clones

In the following section, tests performed for a BeeOND PFS on 4 computing nodes are reported. For testing purposes, these computing nodes were separated from the rest of the production partition. Moreover, to test the workload manager interaction with BeeOND, a container with a custom version of Slurm was specifically prepared. In this way, modifications on Slurm were made possible without impacting the production environment. A Slurm plugin was set up and tested for the purpose of the tests described in this section.

The following tests have been executed using a BeeOND chunksize of 1 MB.

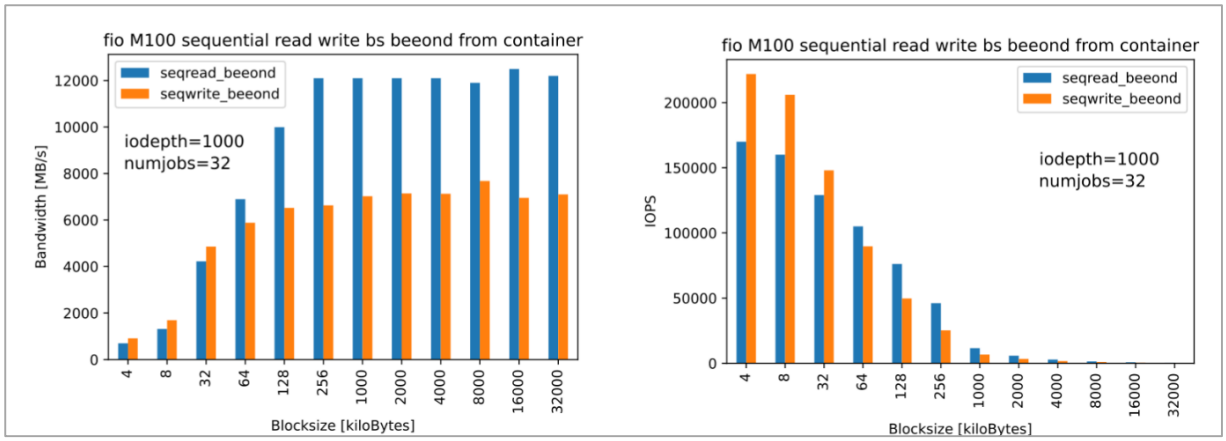

Figure 6: Tests performed for a BeeOND PFS

In these plots, we can see the bandwidth for different blocksize. The plot on the left proves that BeeOND is able to aggregate the performances of the 4 NVMes together reaching a maximum bandwidth of respectively 12 GB/s in read and 7 GB/s in write. Interestingly, as opposed to the single node with XFS, with BeeOND the write IOPS overtake the read IOPS.

The top write IOPS (160k) is lower with respect to the benchmarks in /scratch_local (600k). Although it is likely that the IOPS depletion is due to the current chunksize of BeeOND, further analysis is necessary to obtain final conclusions.

We performed additional analysis in order to find the optimal chunksize to maximise the IOPS. The striping values that we are planning to sample are {4kB, 8kB, 32kB, 64kB, 128kB, 256kB, 1MB, 2MB}. These additional tests require the permissions to create and destroy customs BeeOND instances. Typically, these changes are done only by root and not in user space because the normal user will not have the rights to modify all the parameters of BeeOND, but only the overall number of aggregated nodes.

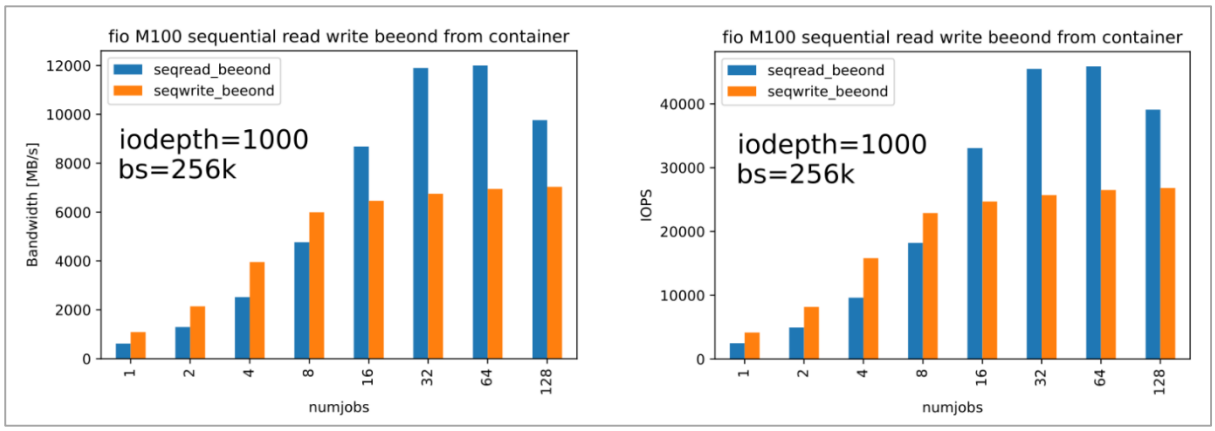

Figure 7: Additional tests performed for a BeeOND PFS

In the plots shown above, we see the effect of a different number of jobs on the bandwidth and IOPS of BeeOND. In comparison with the test on a single /scratch_local (not shown here) we need 32 numjobs instead of 8 to reach the maximum bandwidth. If we use 128 jobs, we can notice a depletion of both bandwidth and IOPS, probably this means that with BeeOND too many queued jobs slow down the overall I/O performance.

4.3.2 Online monitoring

Innovation

Like for other recently deployed systems at CINECA, ExaMon (Exascale Monitoring) which is an integrated tool for monitoring, was installed on Marconi100. ExaMon is a highly scalable framework for the performance and energy monitoring of HPC servers. It is designed to collect a broad spectrum of the heterogeneous data that is generated in a high-performance computing center and to offer a common data interface to the end users.

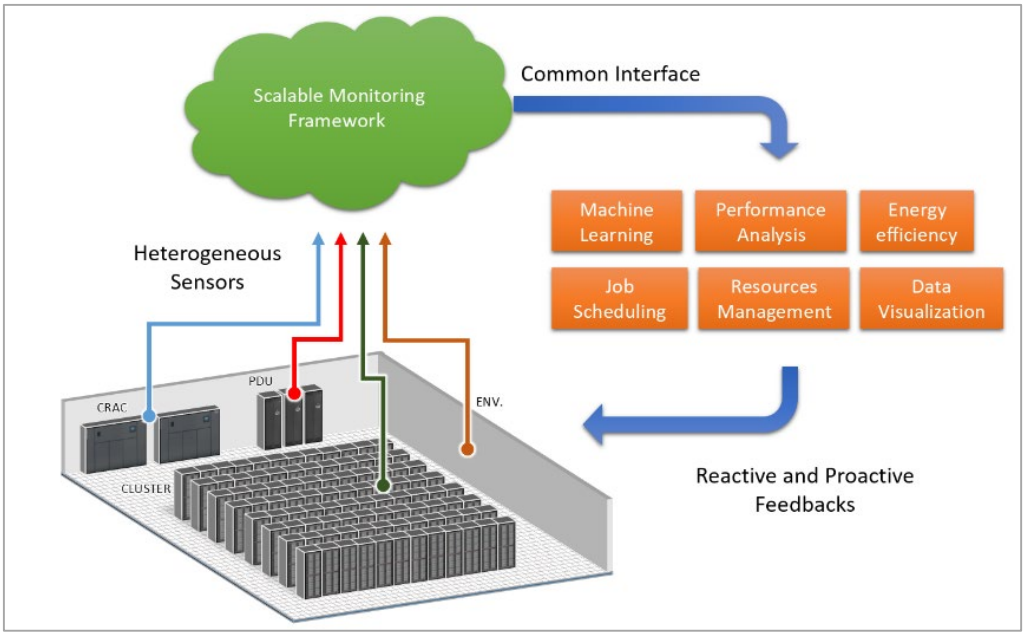

Figure 8: ExaMon framework

In a complex scenario such as that defined by a data center room, data can be: very large in size (Volume), heterogeneous, structured and unstructured (Variety), inaccurate (Veracity) and real-time (Velocity). For these reasons, ExaMon is implemented as a big data framework.

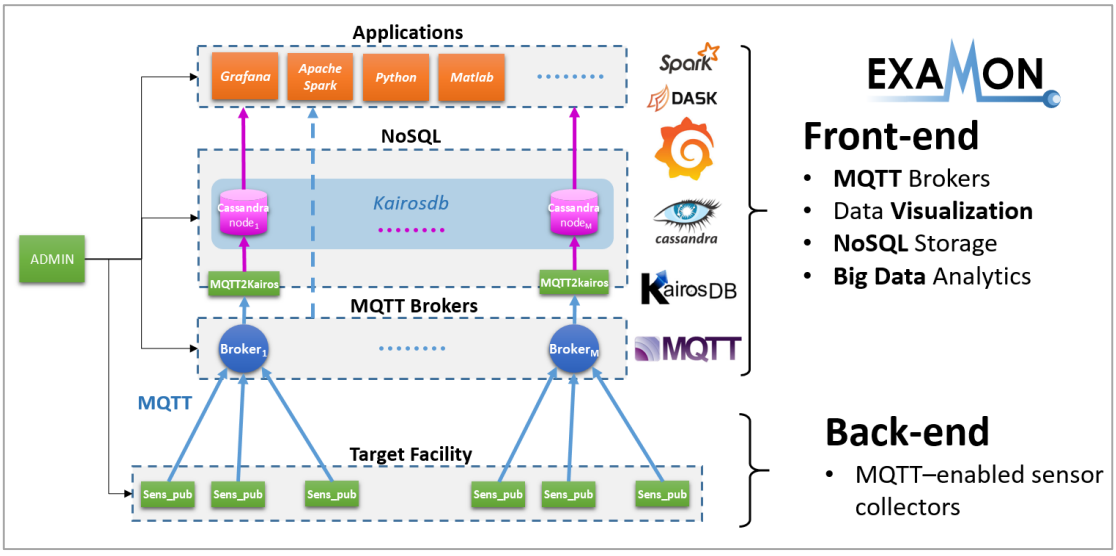

Figure 9: ExaMon components

Briefly, ExaMon consists of several components that include:

- A communication layer based on IoT protocols;
- A set of data collection agents (plugins) that periodically measure and deliver physical and micro-architectural quantities;
- A storage layer implementing a distributed and scalable time-series database using NoSQL technologies;
- An application layer that exposes the collected data to different uses, such as visualisation, data analysis, machine learning and artificial intelligence

Assessment criteria and tasks to validate the innovation

ExaMon was initially developed to monitor the Galileo cluster (538 nodes) at CINECA in 2015. In the following years it was extended to monitor Marconi (3200 nodes) in 2016, D.A.V.I.D.E (45 nodes) in 2017, and an updated version of Galileo (1022 nodes) and Marconi100 (980 nodes) in 2020. Currently, ExaMon is managing an input data rate of 12 GB/day (compressed) and more than 1 Million unique sensors.

ExaMon was installed during the initial cluster setup stages and some plugins (data collectors) were specifically developed for Marconi100.

GPU plugin

A node of Marconi100 is equipped with 4 Nvidia Volta GPUs resulting in a maximum power consumption of 1200 Watts per node when fully utilised. To ensure system stability under all conditions, an optimal cooling system setup is required. In this stage of the project, a new ExaMon plugin (nvidia_pub) was developed to collect all the needed GPU sensors measurements (power consumptions, frequencies, temperatures and performance) of the 3920 GPUs running on Marconi100.

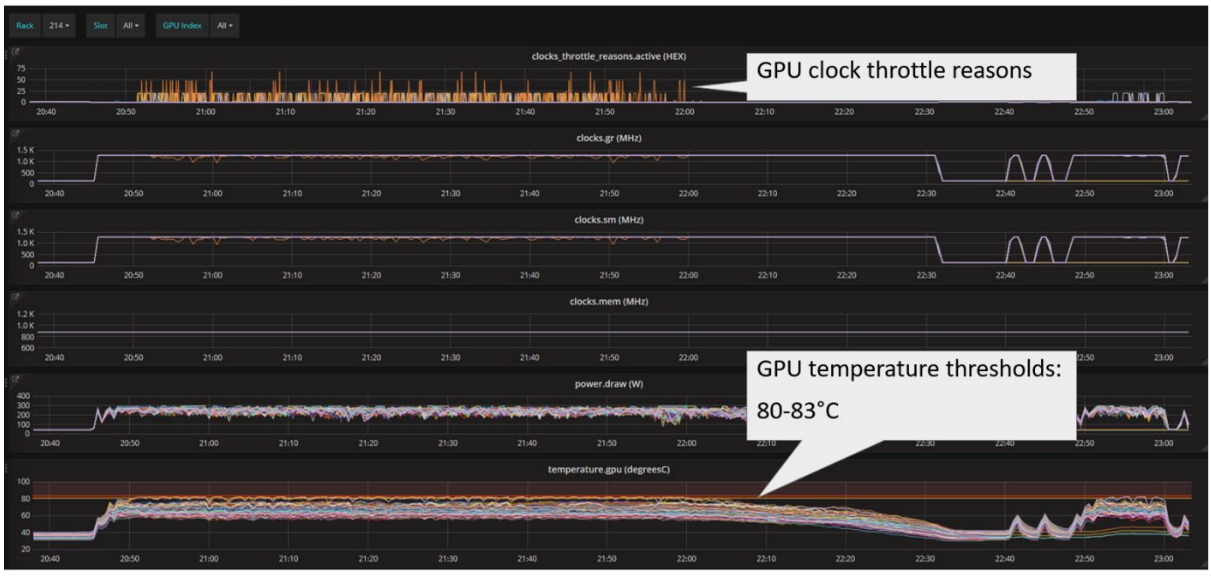

Figure 10: GPU sensor measurements by the new ExaMon plugin (nvidia_pub)

The data collected by this plugin is displayed in real time on “drill down” interactive dashboards that help in finding hot spots while running verification tests. As an example, for the GPUs that

showed anomalous high temperatures and frequency throttling in the dashboards, a quick fix was found by both looking at reliability sensors (clock_throttle_reasons) and physical inspection of the GPU cooling system.

IPMI Plugin

The IPMI plugin was installed to collect the sensor data available on each node using the out-of-band interface available through the board's BMC. The type of the available data mainly concerns temperatures and power consumption of some components present on the Marconi100 nodes.

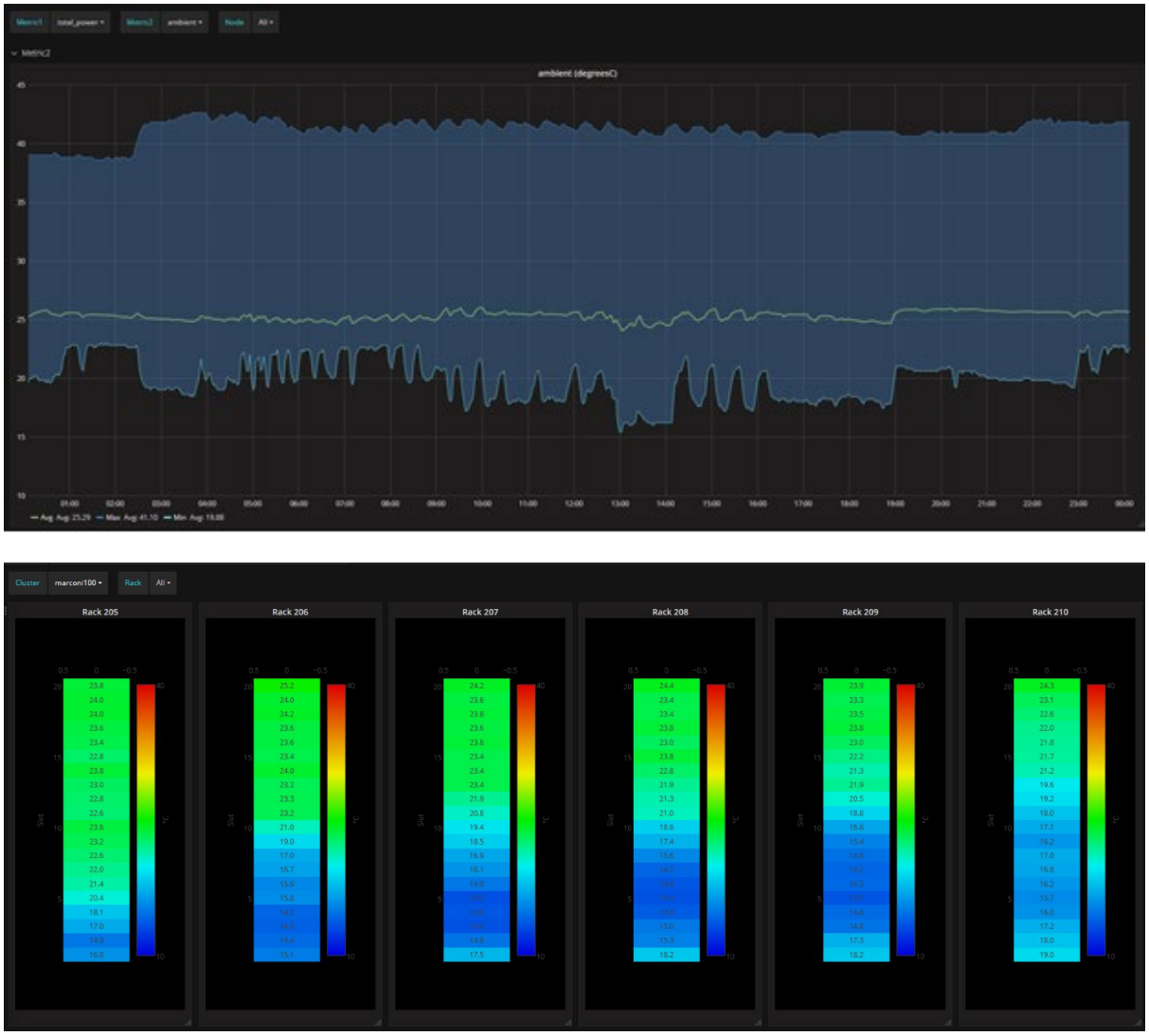

Figure 11: Sensor data by the IPMI plugin

During the installation of the system, the value distribution of the "Ambient" sensor was used in order to measure the air temperature at the inlet of each node. This helped to find the optimal set point of the output temperature of the room cooling system.

OCC Plugin

This plugin was developed specifically for Marconi100 to enable fine grain power profiling. The OCC (On Chip Controller) is an embedded unit found in every Power9 CPU that has the role of measuring and controlling system power and chip temperatures for efficient fan control,

power capping, power saving and performance profiling applications. This unit communicates with the BMC chip using a dedicated bus. Normally, monitoring applications can access the data provided by the OCC sensors both from the CPU cores (in band) and from remote hosts, leveraging the BMC/IPMI interface (out of band). The first solution features low latencies but, in general, it may affect users' applications. The second solution does not cause interference by using the same resources as the user applications but introduces high latencies in the sensors sampling process. The ExaMon plugin developed for Marconi100 instead was designed to execute the data collection routine entirely from within the BMC. This approach benefits from the advantages of both out of band and in band solutions. In particular, the lower impact on latencies of this solution enables non-intrusive fine grain power profiling applications.

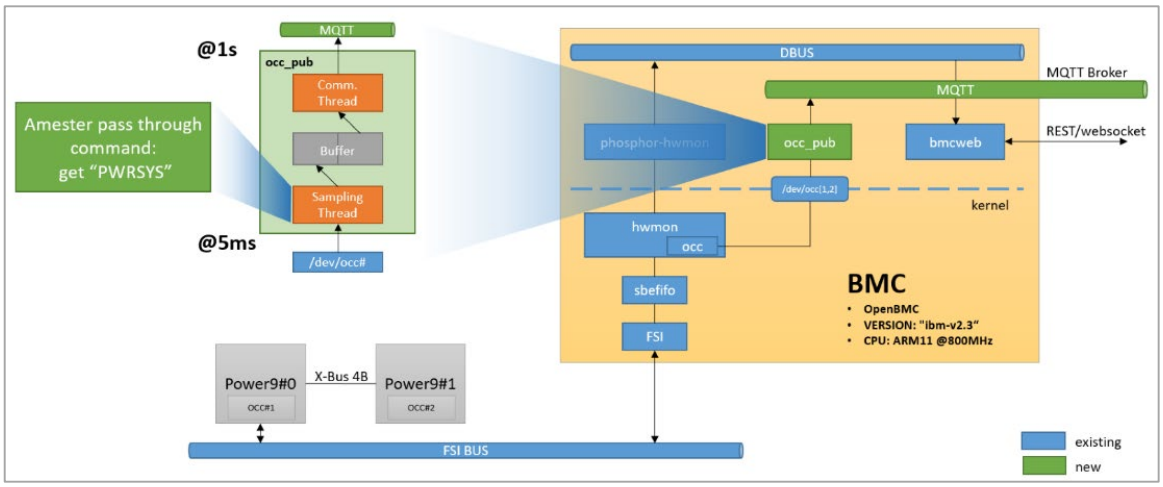

Figure 12: Data collection process by OCC Plugin

Using this approach, a sampling ratio of up to 200Hz of the system's power sensor (PWRSYS) can be achieved, which is a 50X improvement over a standard out-of-band solution (IPMI/REST).

Figure 13: Time series databases by OCC Plugin

Like any other ExaMon plugin, collected data is then stored in the time series database and can be used in AI/ML applications, anomaly detection algorithms, or can simply be displayed in real time dashboards.

4.4 JUELICH

4.4.1 Next-gen. processors improving energy-to-solution for applications & increasing performance-per floor space

Innovation

AMD EPYC Rome support – up to 64 cores and 2.5 GHz at BullSequana Ultra dense and innovative technologies

The JURECA-DC Module is based on an ultra-dense BullSequana XH2000 architecture equipped with AMD x86_64 64-core EPYC Rome CPUs, code-named “Rome”. These CPUs have been selected as they feature good performance for highly parallel MPI applications. This CPU offers 64 cores and memory has been sized according to JUELICH requirements. Due to the large number of cores available per node (128), the solution also allows users to work in depopulated mode, allowing memory hungry jobs to adapt their configuration in terms of bandwidth, capacity per core and frequency.

AMD EPYC Rome is innovative in various respect such as memory bandwidth, PCIe Gen 4 support and of course core count. These advances are enabled by new modular chiplet-based design and a new 7nm process. To take advantage of these characteristics Atos developed a new fully water-cooled blade, supporting three dual-CPU nodes.

The BullSequana XH2000 Platform is Bull’s new iteration of the best-of-breed DLC Sequana architecture. Still offering a fully water-cooled system, enabling one of the best PUEs on the market through hot water (up to 40°C) the BullSequana XH2000 platform further improves density and flexibility over its predecessor. As of today, the BullSequanaXH2000 Platform is the best-of-breed energy-efficient cooling solution, fully meeting the most demanding market requirements in terms of price, performance and environment friendliness. Atos software suite BEO and BDPO is part of our offer and will complete the solution to achieve the best “Energy to Solution” for user needs and applications at JUELICH.

Assessment criteria and tasks to validate the innovation

A Linpack performance of 9.33 PFlop/s was measured with 186 DC GPU-equipped compute nodes. This result placed the JURECA-DC module on spot 43 in the June 2021 Top500 list (Top500, 2021). On the High-Performance Conjugate Gradients (HPCG) benchmark, JURECA-DC achieved 273.784 TFlop/s in 2021 corresponding to place 29 in the June 2021 HPCG list (HPCG, 2021).

JURECA-DCs Footprint is 15,62 m²

In the tender for the JUELICH lot, for a set of application benchmarks reference performance numbers have been provide for the old system, i.e. JURECA-Cluster. During the acceptance tests, the performance numbers have been measured on the system procured in the context of PPI4HPC, i.e. JURECA-DC. The performance is expressed in terms of execution time Δ t^(b) and number of nodes n^(b) and results can be found in Table 5 below. As can be seen from this table, the benchmarks could be executed in a similar or less time on a significantly smaller number of nodes. This means that it was possible to significantly reduce the number of nodes by using the AMD EPYC processors that feature a significantly larger number of cores.

	JURECA-Cluster		JURECA-DC	
Application Benchmark	n^(b)	Δt^(b)	n^(b)	Δt^(b)
GROMACS	58	19.69	3	19.85
ZFS	160	425,239	19	509.22
CP2K	128	819,326	48	882.5
ParFlow	64	81	16	98.23
ICON	100	24,326	9	38.16

Table 5: JUELICH Application benchmarks - JURECA-Cluster reference values / JURECA-DC measured acceptance values

4.4.2 Next-gen. NVM storage based on SBB and SBF technology

Innovation

New Generation storage with SBF (Smart Bunch of Flash) and SBB (Smart Burst Buffer)

Atos SBF (Smart Bunch of Flash) provides dynamically allocated NVMe over Fabric temporary storage. This proximity caching layer can be used to check pointing activities and provide fast access Storage to fulfill GPU's high data bandwidth demands. SBF is connected to the same Level1 switches where the compute nodes are also connected.

Compute and data have long evolved in separate worlds. Moving from teraflopic storage to petaflopic storage involved switching from shared to parallel repositories. But to answer the next generation of HPC systems as well as the coming exascale generation, a new leap is needed: bring data to compute. To address this challenge Atos has developed a whole strategy based on the most recent technologies:

- Abstract data locality and unify RAM, NVRAM, flash, disks, etc. (project Dino)
- Object storage to abstract access from hardware (project Raptor)
- NVMe over Fabric to unify local and remote access
- Etc.

Putting together all these technologies is achieved through FastIO libraries, which embed and abstract the smart placement of data. Two components first issued from this global strategy are SBF and SBD. While SBF provides remote storage that is seen as local from the compute nodes, SBB uses the same technology to offer remote, shared burst buffer to compute nodes.

Assessment criteria and tasks to validate the innovation

The Graph500 benchmark, when using external NVM storage, is used to show the attainable performance by specific data-intense analytics workloads that leverage out-of-core computations.

We present here the results of the Breadth First Search (BFS) implementation for

Graph500 benchmark[28]

The benchmark was run on 1 node with 1 MPI process and 64 threads per process. Two different placements were used: the first placement is limited to 1 socket, with 1 core per thread ('compact'). The second placement is scattered over 2 sockets, with the threads spread over the 8 NUMA domains of a node ('scattered'). The scattered run has twice the available memory bandwidth than the compact run. The table below shows the harmonic mean GigaTEPS.

	compact	**scattered**
NVM	3.94	4.06
/dev/shm (in-memory)	3.98	4.09

Table 6: Harmonic mean GigaTEPS

The use of /dev/shm can be seen as an upper limit of the performance, where the internal memory is used for hosting a filesystem that stores the input files. It can be seen that the external NVM adds almost no overhead.

The usage of the external NVM storage needs to be requested as part of the SLURM commands in the job script.

4.4.3 Mellanox DragonFly+ topology

Innovation

HDR was selected as the interconnect technology. Based on the well-known, highly proofed Infiniband technology, it provided particular high RDMA-based communication performance between nodes at the time of installation. To cope with a dense system while keeping infrastructure costs under control (switches, fiber cables), the topology selected for the cluster is Dragonfly+. This solution features a high bisection bandwidth while requiring a relatively small number of switches.

DragonFly+ can be traced in papers as long as 10 years back, its applications to production HPC clusters are just beginning. This is due to the new "dynamic routing" algorithms. Compared to Fat Tree (with or without pruning) or "classical" Dragonfly is the only topology to retain as many advantages, combining the advantages of Dragonfly and Fat Tree:

- Scalability
- Cost, with bigger groups of nodes
- Bi-section bandwidth
- Best worst-case performance
- Less costly routing loop prevention

[28] https://github.com/htsst/netalx

Dragonfly+ is the default topology with the new BullSequana XH2000 infrastructure. Increased radix (40-port switches) further reduce the cost while participating fully in the DLC environment.

Assessment criteria and tasks to validate the innovation

JURECA-DC Infiniband Components:

- NVIDIA Mellanox InfiniBand HDR (HDR100/HDR) DragonFly+ network
- NVIDIA Mellanox ConnectX-5 single and dual-port host channel adapters in nodes
- 102 times 40-port Mellanox HDR switches
- 7 times Fat tree non-blocking interconnection networks inside each compute cell (each cell consists of two racks)
    - 8 times HDR leaf switches (L1)
    - 6 times HDR spine switches (L2)
- Service/Storage Island (attached via Up/Down chain routing)
    - 8 times Mellanox Skyway Storage (GPFS) Gateways
    - 4 times 40-port Mellanox HDR switches

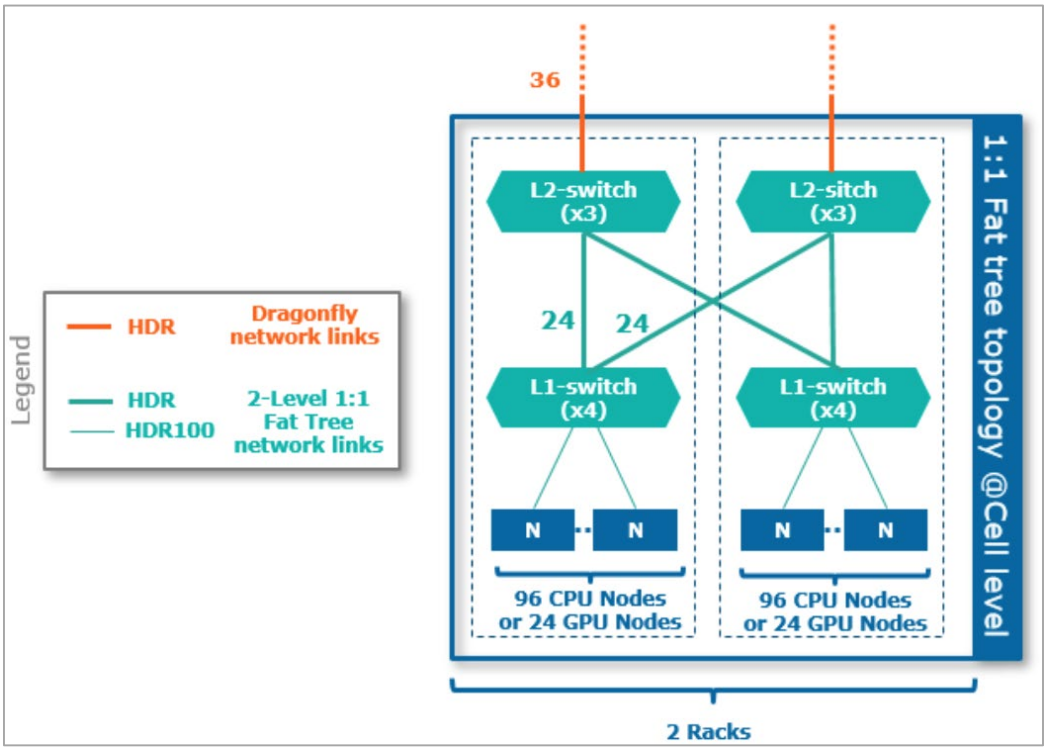

Figure 14: Structure of one JURECA-DC compute Sequana Cell

A dedicated communication analysis took place to represent benchmarks on three different infiniband fabric levels:

Level-1: Compute node Switch Level (involving L1 connections)

L1 → 24 (all/max-amount) Computes behind one lowest Layer L1 Switch:

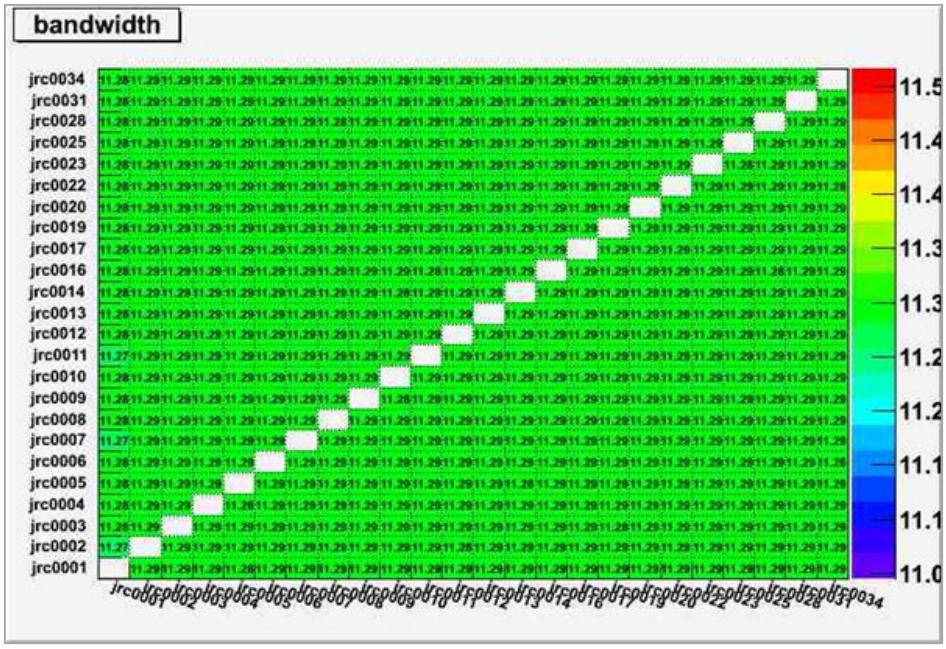

Figure 15: Bandwidth matrix of CPU only compute nodes distributed at one L1 Switch layer

Level-2: Cell-Level (involving L1 to L2 connections)

L2 → 24 (same amount to bedder compare) computes inside one cell spread over all available 8 L1 Switches which then brings the L1 to L2 Connection into the game (inside a Cell we have Fat tree non-blocking):

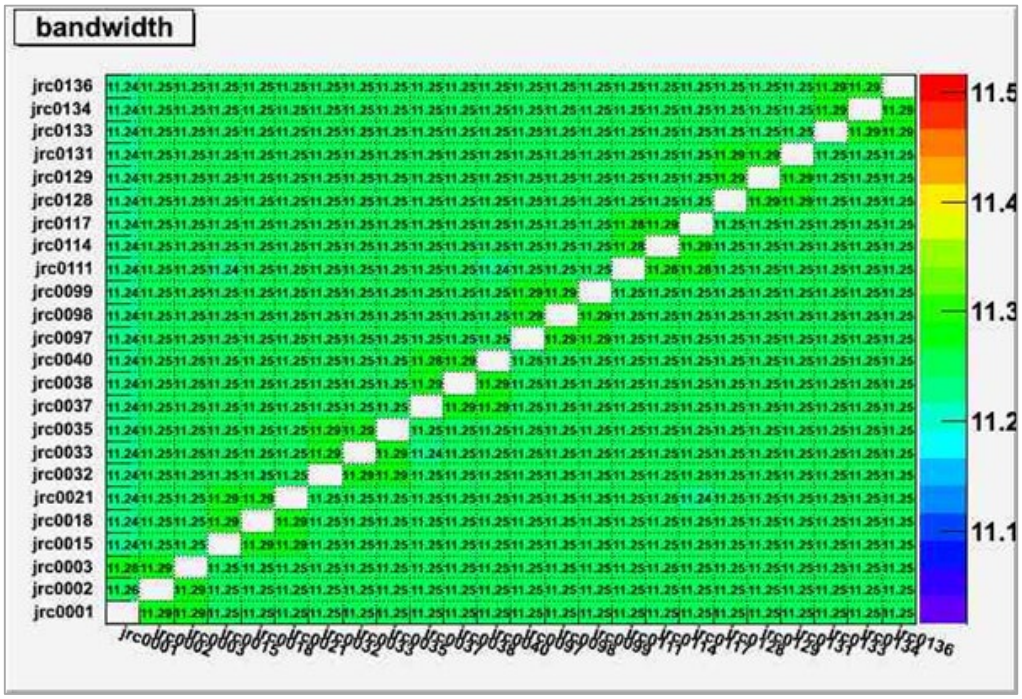

Figure 16: Bandwidth matrix of CPU only compute nodes distributed at Cell - Full Fat Tree layer

Level-3: Cluster Level (involving L2 to L2 connections in addition)

L3 → 24 nodes spread over all CPU compute cells (3) and their 24 Switches inside which brings the L2 to L2 connections into the game (to compare with the DragonFly+ layer):

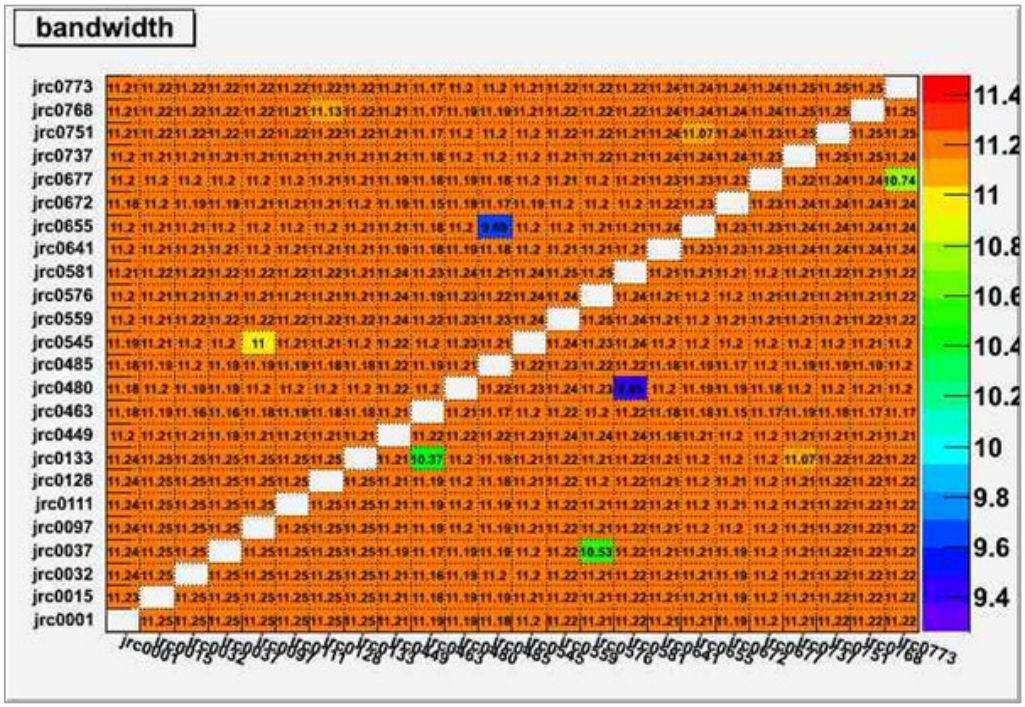

Figure 17: Bandwidth matrix of CPU only compute nodes distributed at DragonFly+ layer

To also analyse the DragonFly+ performance from application benchmarks, we ran a benchmark of the Chroma QCD code[29].

This is supported by a domain-specific software stack called the USQCD software stack. The benchmark is the Hybrid Monte-Carlo part of an LQCD simulation ---- used to generate the background gluonic fields upon which we calculate physical quantities. It is a Markov chain process update that is proposed then accepted or rejected in a Metropolis test (updated which increase the system energy more are less likely to be accepted.) Proposing updates involves solving a very large, sparse linear system (dim ~10^8 in this case).

The problem is formulated on a 4-dimensional lattice, which is split between the nodes. In the solver, each node works on its portion of the lattice but must exchange a halo of data with neighboring nodes to include nearest neighbor interactions of points at the edge of each node's domain. Often it makes sense to divide the problem across fewer than 4 dimensions, as that reduces the size of the halo (and amount of communication), speeding up the solver.

We ran a realistic lattice size of Nx*Ny*Nz*Nt = 64*64*64*96. The files read and written are about 15GB to stress the infiniband fabric as much as possible:

[29] https://jeffershttps://jeffersonlab.github.io/chroma/onlab.github.io/chroma/

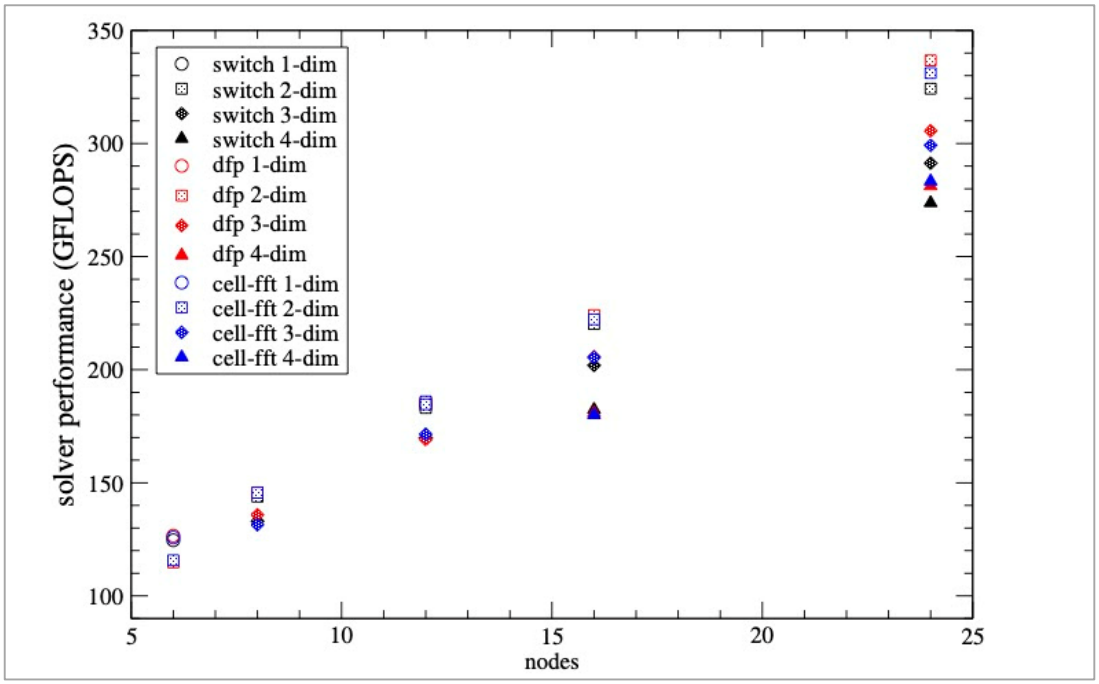

Figure 18: Switch[1-4]-dim=Level-1, dfp [1-4]-dim=Level-3, cell-fft [1-4]-dimm=Level-2

In addition, a JUELICH-Linktest[30] (all to all node communication test) was successfully performed to demonstrate the full fabric performance:

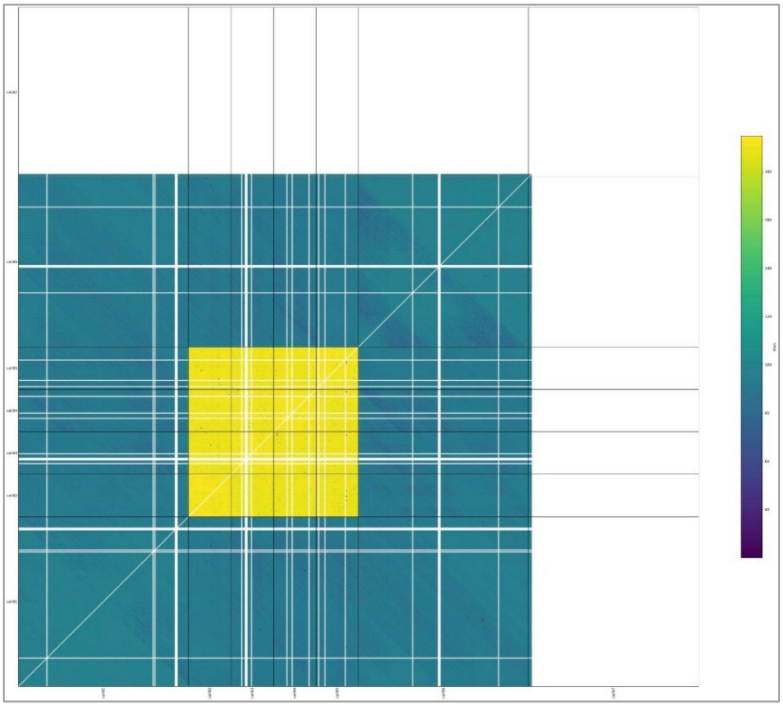

Figure 19: Bandwidth matrix: values are in Gb/s at right colour legend. Each coloured area where the X-axis hits the Y-axis is representing one (parallel) node to node communication bandwidth. The inner more performant area represents the GPU equipped compute nodes with dual port HCAs (2 times HDR100). The CPU only compute nodes are equipped with single port HDR100 HCAs

[30] See: https://www.fz-juelich.de/ias/JUELICH/EN/Expertise/Support/Software/LinkTest/_node.html

To verify that the DragonFly+ algorithm is functioning as expected we enhanced our monitoring to be able to identify the congestion level inside the hole whole fabric over time:

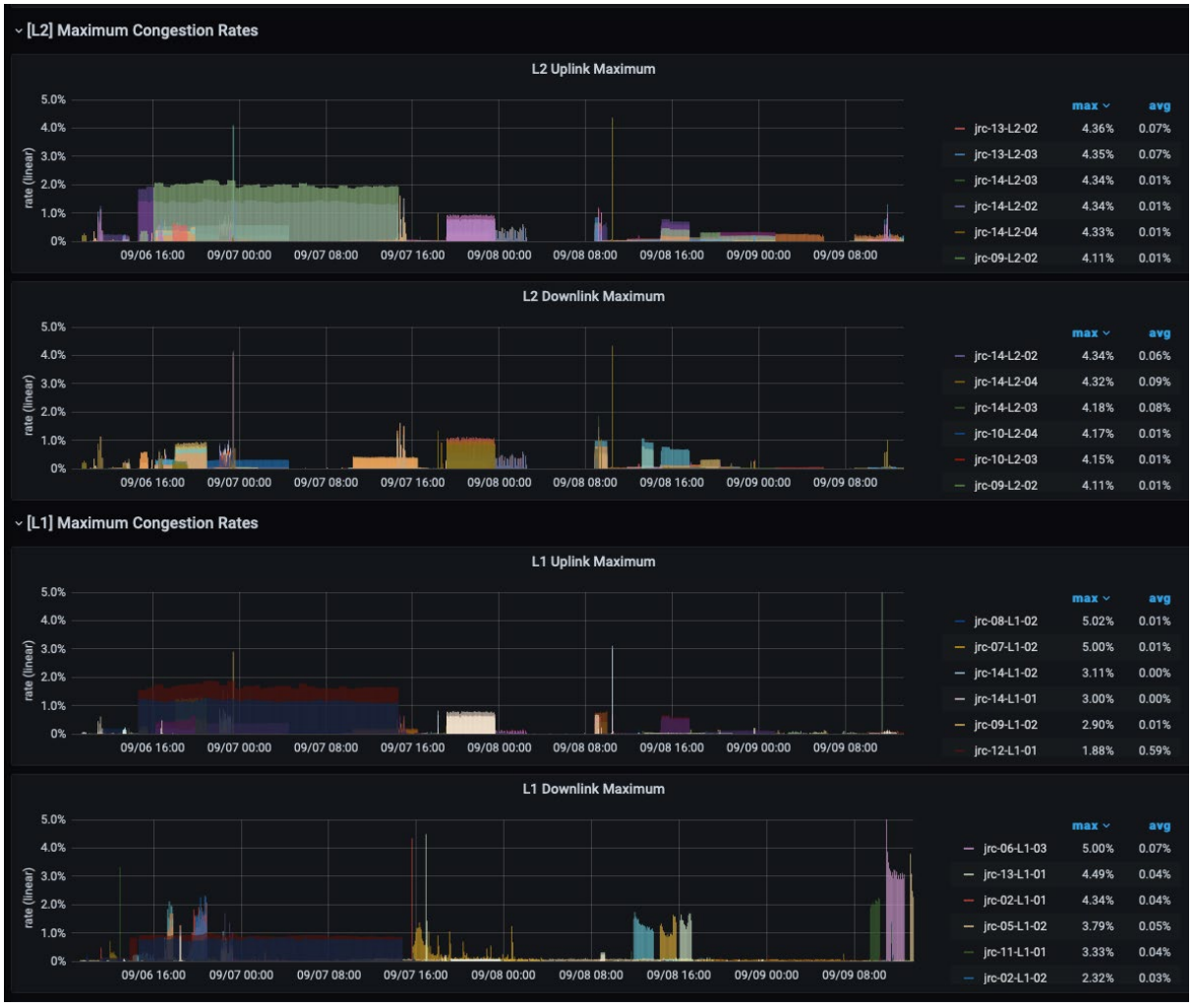

Figure 20: JURECA-DC IB Switch congestion between 2021-09-06 08:31 and 2021-09-09 13:22 during production

5. PPI4HPC Impact and Lessons learnt

In this section we first analyse the impact, i.e. the changes that PPI4HPC expects to have caused in the relevant context. At this point of time it is difficult to provide empirical evidence for the claimed impact and the analysis therefore has to be considered tentative. We consider more specifically the impact on the market in terms of the offered solutions, which we can measure in terms of responses to our common technical needs. Finally, we summarise the lessons, which we learned as project partners. In this deliverable we focus on the technical aspects of PPI4HPC, while for the legal aspects we refer to the white paper “Lessons learned on legal aspects”[31].

5.1 PPI4HPC general impact

The most important impact is the increased willingness of multiple publicly funded entities to work together on some form of joint and coordinated procurements in the area of HPC after PPI4HPC has successfully demonstrated its feasibility. Some of the partners involved in

[31] See: https://ppi4hpc.eu/sites/default/files/public/file-uploads/PPI4HPC_white_paper_2_0.pdf

PPI4HPC engaged in coordinated procurement efforts involving also other stakeholders in the context of the ICEI project[32]. We, furthermore, believe that the Joint Undertaking EuroHPC did benefit from this demonstration when running procurements involving the Joint Undertaking as well as different national stakeholders.

A few changes to the way of procuring HPC solutions, which have been introduced and/or further improved in the PPI4HPC procurement, is likely to impact future procurements.

Firstly, the innovation criteria, which was introduced as a requirement of the PPI, allowed to strengthen the focus on innovations as it forced suppliers to clearly document the innovative features of the offered solution. In the case of the CINECA this made it even attractive for an industrial supplier to integrate an innovative solution from a European SME.

Secondly, the methodologies for awarding a score based on a total cost of ownership (TCO) approach, which were developed at different sites, could be refined and consolidated. TCO does, in particular, also take electricity costs into account and therefore favours energy-efficient solutions. We expect this approach based on the methodologies used within the PPI4HPC procurement to become common practice.

Thirdly, the joint effort on formulating the technical part of the tender documents resulted in a much clearer structure and homogenisation of the vocabulary. This resulted in a better and more precise formulation of the needs and requirements of the public procurers and, therefore, improved transparency from the suppliers' perspective. Another improvement worth mentioning are the improved tracking of risks through a RAC-I model as well as the advanced acceptance test procedures introduced at all sites. Also here, we expect that the PPI4HPC tender documents have set a new quality level, which did already have an impact on the EuroHPC procurements.

5.2 Impact through common technical needs

While the PPI4HPC procurement had been split in different lots, the buyers group nevertheless aimed on maximising the impact on the market through the joint procurement. The strategy was to formulate a set of common technical needs, which were endorsed by all procurers even if they did not necessarily play a role in each of the lots. The table below documents the result of a self-assessment of how strongly the market responded to the common technical requirements. A "strong" or "very strong" indicates that the offers received by the suppliers during the dialogue did address the common technical needs well or even very well, while "very weak" indicates that the supplier did not address this common need in the offer. In case ranges are given, it indicates that different suppliers responded at different levels. Note that the reasons for not addressing a common need have not been investigated. Suppliers may have considered addressing the need technically feasible but did not have a technical solution available as a product. Possibly suppliers do not expect a sufficiently large market for such solutions to be available despite a buyers group comprising of important customers ask for such solutions.

We observe that in a number of cases the responses have been considered rather weak. This was, in particular, the case when no solutions were available on the market to address the identified need like the improved application performance monitoring capabilities. This indicates that the impact on the market and the availability of suitable solutions through the formulation of common needs had been limited.

[32] ICEI is a project funded by the European Commission through the H2020 program under the Framework Partnership Agreement of the Human Brain Project (http://www.fenix-ri.eu). The ICEI project is delivering e-infrastructure services federated as the Fenix Infrastructure.

Common Need	Market response
BSC	
The offered solution must not contain any single-point of failure	Strong
The build time (recovery from fault) in storage infrastructure must be improved	Strong
The monitoring capabilities must allow to foresee future failures and to take preventive mitigation measures	Weak
Availability of tools to get a global awareness of the usage of the data infrastructure	Borderline
Need of optimised and faster access to storage from computing processes	Weak
CEA	
Support for node power and energy measurements	Very Strong
Operation of the system within a low guaranteed power envelope(power capping capability)	Very Strong
Integration of job energy accounting with the Workload Manager (WLM)	Strong
Improved application performance monitoring capabilities	Strong
Support of complex software environments	Strong
Security	Strong
CINECA	
Support for node power and energy measurements	Strong
Operation of the system within a low guaranteed power envelope(power capping capability)	Strong - very strong
Integration of job energy accounting with the Workload Manager (WLM)	Very weak - strong
The build time (recovery from fault) in storage infrastructure must be improved	Strong - very strong
Improved application performance monitoring capabilities	Very weak - borderline
Support of complex software environments	Borderline - strong
JUELICH	

Significant reduction of the Energy-to-Solution	Strong
Operation of the system within a low guaranteed power envelope(power capping capability)	Borderline - strong
Significantly higher performance-per-floor space density	Strong - very strong
Support of complex software environments	Borderline - strong
Tight integration of NVM into the system	Strong - very strong
Improved application performance monitoring capabilities	Weak - borderline

Table 7: Common technical needs and market response per site

5.3 Lessons learnt

Based on the experiences made during the PPI4HPC project, the following lessons had been learned from a technical perspective:

1. Using a competitive dialogue allowed for in-depth analysis of the technical options during the tendering phase. This was, in particular, helpful to address technical challenges related to data centre integration like power control and capping or the usability of innovative solutions like integration of storage devices based on non-volatile memory.

2. Within a joint procurement it is possible to operate with a common set of benchmarks, which helps to reduce the efforts of suppliers compared to the case of separate and uncoordinated procurements with disjoint sets of benchmarks.

3. The idea of establishing a collaboration during contract execution with the provider, which was not common practice before the PPI4HPC procurement, turned-out to be fruitful for all sites as it allowed to continue supporting the development of innovations or improvement of the quality and/or usability of the innovations integrated in the procured solutions.

4. The impact on the market caused by a joint procurement of this size in terms of new solutions, which meet the common needs of the procurers, is limited. One reason is likely the short amount of time between publication of the tender documents and delivery of the solutions, which is short compared to typical development cycles. Anyhow, discussions derived from PPI4HPC procurers and providers could help in defining future products to cover HPC common requirements/goals needs.

6. PPI4HPC Carbon footprint

6.1 Introduction

Global warming trend, observed since the mid-20th century and subject to a strong acceleration during the last decades, is widely recognised as a major threat for the future. Scientists have identified that the main cause of this trend is the release of greenhouse gases in the atmosphere due to human activity among which CO_2 plays an important role.

In response to this critical situation and more generally to the challenges of a sustainable future, the European Commission issued in 2019 a communication[33] to "sets out a European Green Deal for the European Union (EU) and its citizens" with the goal, among others, to reach a situation in which "there are no net emissions of greenhouse gases in 2050".

In a more recent document[34], the European Commission, in line with the Green Deal, states that "Data centres and telecommunications will need to become more energy efficient, reuse waste energy, and use more renewable energy sources".

This, indeed, is an important goal, as, in the same document, it is mentioned that "the environmental footprint of the [ICT] sector is significant, estimated at 5-9% of the world's total electricity use". It should however be noted that the usage of electricity by data centres tends to be stable over time and the usage by conventional data centres (category to which HPC centres belong) tends to decrease[35].

The goal of this chapter is to show how the PPI4HPC project and sites involved in the project work towards the realisation of this statement:

- At site level, by, as much as possible, selecting low carbon energy sources and reusing waste energy.
- At system level, by selecting systems with high performance to power consumption ratio and reducing the facility overhead.

In addition, this chapter contains suggestions for future projects and procurements in order to further reduce the carbon footprint of HPC data centres.

6.2 PPI4HPC site level activities

The four sites involved in the PPI4HPC projects are large sites hosting European Tier-0 (PRACE) systems for a number of years. As such, all four sites, depending on the local conditions, have a strategy of reduction of carbon footprint at the level of energy sources and waste energy reuse as well as a policy of development and dissemination of best practices in terms of hosting of large systems.

6.2.1 Energy sources

Regarding the energy sources, the main concern is the use of low-carbon electricity, possibly with specific contracts with suppliers for the provision of renewable energy.

The figure below shows the energy mix of the electricity source of the four sites involved in the PPI4HPC project in terms of fraction of low carbon (including renewable and nuclear), fraction of renewable and average of gCO_2/kWh. It should be noted that this average is lower than the national value for BSC and JUELICH, as both sites are getting electricity with a specific contract making it possible to get electricity with a lower carbon footprint than at national level (respectively of 160 and 320 gCO_2/kWh).

[33] Communication from the Commission to the European Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of regions - The European Green Deal - Brussels, 11.12.2019.
[34] Shaping Europe Digital Future, European Commission, February 2020.
[35] Recalibrating global data centre energy-use estimates, Science Vol. 367, Issue 6481, 28 February 2020.

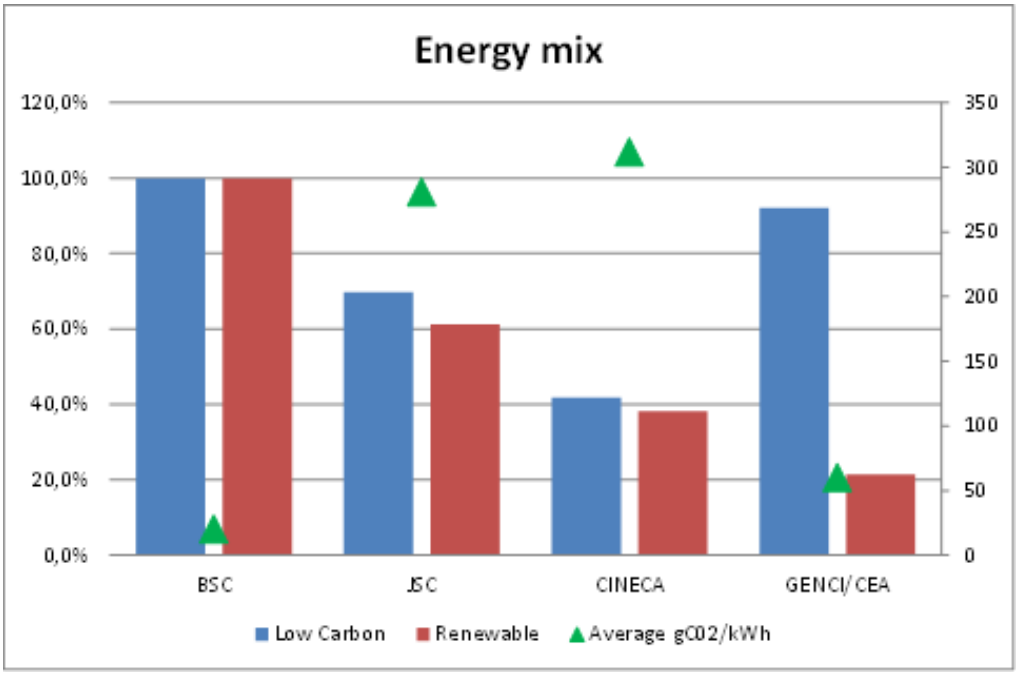

Figure 21: Energy mix of the electricity source per site

6.2.2 Reuse of waste energy

The heat dissipated by the HPC systems can be reused in two different ways: for heating or for cooling. In the first case, heat pumps are usually needed in order to increase the temperature of the warm water available for heat reuse, in the second case, adsorption chillers are needed to produce cold from a hot source.

At this time, only CEA implements heat reuse for one of the two data centres located in the site hosting the PPI4HPC system. The heat reuse system makes it possible to heat entirely the offices of the site (close to 200 employees) except during a few very cold days when the assistance of a gas fired heating system is needed.

JUELICH and CINECA are planning to implement a heat reuse system in the near future for future systems.

6.2.3 Development and dissemination of best practices

Several sites involved in the PPI4HPC projects, BSC, CINECA and CEA are members of the programme committee of the Annual European Workshop on HPC infrastructures, which is supported by PRACE[36]. This workshop is widely recognised as an event where best practices in terms of hosting supercomputers are developed and exchanged between specialists.

In addition, CEA is a participant to the EU Code of Conduct for Data Centre[37].

6.3 PPI4HPC system level activities

At system level, two elements contribute to the reduction of carbon footprint: the PUE (ratio between total power usage including facility usage and IT equipment power usage) on the one

[36] See: https://prace-ri.eu/infrastructure-support/european-workshops-on-hpc-infrastructures/
[37] See: https://ec.europa.eu/jrc/en/energy-efficiency/code-conduct/datacentres

hand, the energy intensity on the other hand while the increase of capacity of the systems has an inverse effect.

While the PUE is meaningful for all four sites involved in the PPI4HPC, the energy intensity (expressed in MW/Pflops for the HPL benchmark) and the capacity (expressed in Pflops for the HPL benchmark) makes sense only for the three sites that have installed supercomputers in the context of the PPI4HPC projects[38].

It should be noted that the large increase of energy intensity is the consequence of the inclusion in the technical specifications of requirements regarding the energy efficiency while the improvement of the PUE is the consequence of improvements of the site facility and of the move from air cooling to direct liquid cooling or from warm to warmer direct liquid cooling.

The figure below, inspired in its form by an article published in Science magazine[39], presents the situation of the different sites involved in the project and the average.

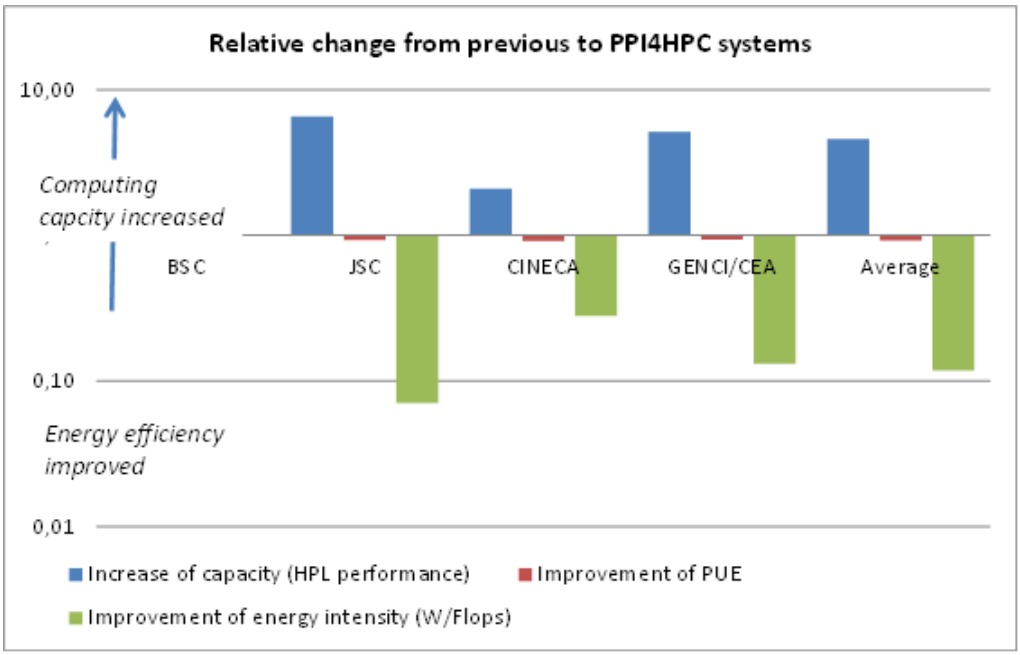

Figure 22: Relative change of technical features per site

It turns out that in all cases for these three sites, the total power consumption is similar or lower to the one of the systems that were replaced while the computing capacity increased by a factor of close to 5 on average (see table below). It should be noted that the large increase of performance related to the previous system for JUELICH can be explained by the fact that a system based on CPUs is replaced by a system using GPUs and for GENCI/CEA that the system replaced is older than the others.

	JUELICH	**CINECA**	**GENCI/CEA**
Previous system	Jureca (2015)	Marconi KNL (2017)	Curie TN(2012)

[38] PPI4HPC system at BSC is a storage system only so metrics related to computing are not applicable to that system and are not shown in the graphics below.
[39] Recalibrating global data center energy-use estimates, Science Vol. 367, Issue 6481, 28 February 2020.

Relative change of power consumption (including PUE)	0.43	0.9	0.62
Relative change of performance	6.55	2.08	5.14

Table 86: Relative change of power consumption and performance at JUELICH, CINECA, and GENCI/CEA

6.4 PPI4HPC carbon footprint (for operation)

Based on the previous information, it is possible to compute the carbon footprint for 5 years of operation for the three sites that acquired supercomputers in the context of the PPI4HPC project considering that the average power usage is 80% of the power consumed during the HPL test and that the system is running 100% of the time. The result is shown in the table below:

	JUELICH	**CINECA**	**GENCI/CEA**
Power consumption in production mode (including PUE) (MW)	0.34	1.40	1.58
Average gC0$_2$/kWh	281	312	60
Ton of CO_2 released during 5 years of operation	4160	19180	4170

Table 9: Carbon footprint of the JUELICH, CINECA, and GENCI/CEA systems for 5 years of operation

For reference, one ton of CO_2 corresponds to one return flight from Paris to New York.

6.5 Suggestions for future projects or procurements

The analysis presented above focuses on the carbon footprint related to the operation of the systems procured in the context of the PPI4HPC project, the carbon footprint of the construction of the systems (including transport and dismantling) is not considered.

In order to investigate how important this carbon footprint was, compared to the one of five-years operation of the systems, the sites involved in the project asked their suppliers (ATOS and IBM) for information about this carbon footprint.

Since no answers were received, the only usable information found for making such analysis was information regarding Dell server (dual intel processors) respectively model r640 and r740[40][41]. The analysis performed by Dell concludes that construction of the system accounts for around 15% of the total carbon footprint of the server for 4 years of operation for a model r640 and around half for a model r740. The analysis of Dell highlights the fact that the SSD present in the model r740 compared to the r640 explains most of the difference.

It turns out that the carbon footprint of the construction of a system is significant compared to the carbon footprint of the operation. This means that it may be useful to require information about the carbon footprint of the construction of a system in the context of a procurement even

[40] See: https://i.dell.com/sites/csdocuments/CorpComm_Docs/en/carbon-footprint-poweredge-r640.pdf
[41] See: https://corporate.delltechnologies.com/content/dam/digitalassets/active/en/unauth/data-sheets/products/servers/lca_poweredge_r740.pdf

if how this information could be used and even verified is unclear. At least vendors should be encouraged to publish information as Dell is doing with at least two of its servers.

7. Conclusions

This report offered a detailed analysis on the assessment and validation of the PPI4HPC innovative solutions. It explained the process of the PPI4HPC coordinated procurement procedure and early market analysis, the common methodology and assessment results of the purchased technologies from each site, the overall impact and specific lessons learnt from this process from the technical side of things, and, finally, the carbon footprint of the project.

PPI4HPC made new and innovative technologies and resources available to European researchers after succeeding in attracting offers from the main players in the HPC field through a joint tendering process. The PPI4HPC project was the first common procurement of four of the main European HPC sites. This is a coordinated effort that has never been made before in the HPC sector. The four sites worked together to achieve a joint tendering process from a legal and technical point of view.

The joint procurement offered several benefits, including:

- A common vocabulary and benchmarks when it comes to coordinating the joint tender procedure, identifying the needs in resources, and assessing the new technologies
- An established collaboration with the key HPC industrial stakeholders and vendors to kickstart new innovative solutions for future products or improve available products
- A coordinated roadmap and approach for providing optimised HPC resources that can be used for future joint procurements. PPI4HPC documents are already being used for EuroHPC pre-exascale tenders making the case of the project's contribution in the field of joint HPC procurements
- Inclusion of a RAC-I model as a requirement from vendors for improved tracking of risks
- More and innovative supercomputing resources that are already being efficiently exploited for science and engineering applications by users and projects in Europe
- Strengthened R&D on HPC architectures and technologies and newly designed solutions according to the needs of scientists and researchers in Europe
- Emphasising the importance of energy efficiency requirements in the tender procedure in order to achieve display of carbon footprint data by vendors and supercomputing sites. This helps infrastructures to operate with the highest effectiveness of power usage possible, and makes for an overall stronger case for social responsibility by both the manufacturers and supercomputing sites

By achieving these goals, PPI4HPC addressed major scientific, industrial and societal challenges. PPI4HPC, the first joint European procurement of innovative HPC systems, contributed to the upgrade of a pan-European HPC infrastructure to serve research and engineering, created a strong impact on the market by fulfilling existing market requirements and needs, and set the example for future joint procurement in HPC in Europe.