



D6.1 In-kind systems description

Document Properties

Contract Number	101033975
Contractual Deadline	M6 (30/06/2022)
Dissemination Level	Public
Nature	Report
Edited by :	Eric Boyer, Genci
Authors	Eric Boyer, Genci; Florent Neveu, CEA; Maïke Gilliot, CEA; Daniele Gregori E4; Ondrej Vysocky, IT4I; Andrea Bartolini, UNIBO; Marco Aldinucci, UNITO
Reviewers	Daniele Cesarini, CINECA
Date	27/06/2022
Keywords	In-kind resources
Status	Reviewed
Release	Final



EuroHPC
Joint Undertaking

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101033975. The JU receives support from the European Union's Horizon 2020 research and innovation programme and France, Germany, Italy, Greece, United Kingdom, Czech Republic, Croatia.



History of Changes

Release	Date	Author, Organization	Description of Changes
0.1	26/04/2021	P. Bernier-Bruna, Atos	Template Creation
0.2	28/04/2022	Eric Boyer	First Draft
0.3	29/04/2022	Daniele Gregori	Added E4 ARMIDA system description
0.4	10/05/2022	Ondrej Vysocky	Added description of the IT4Innovations systems
0.5	11/05/2022	Andrea Bartolini	Added description of the Monte Cimone systems
0.6	23/05/2022	Marco Aldinucci	Added description of the UNITO system
1.1	25/05/2022	Eric Boyer / Matthieu Hautreux / Maïke Gilliot	Gitlab and Openstack systems in annex and description context addons
1.2	06/06/2022	Eric Boyer / Maïke Gilliot	Version for review
1.3	06/06/2022	Daniele Cesarini	First internal review
1.4	13/06/2022	Eric Boyer	First feed back to Daniele's comments
1.5	21/06/2022	Daniele Cesarini	Review of chapter 5
1.6	26/06/2022	Eric Boyer	Additional adjustments
Final	27/06/2022	Eric Boyer	Version ready for project delivery



Table of Contents

DOCUMENT PROPERTIES	1
HISTORY OF CHANGES	2
TABLE OF CONTENTS	3
1 EXECUTIVE SUMMARY AND INTRODUCTION	5
2 IN-KIND RESOURCES AS FIRST STAGE TOWARDS THE PILOTS FOR EUPEX PARTNERS	6
3 IN-KIND SYSTEMS FOR EUPEX PROJECT	8
3.1 JOLIOT-CURIE-IRENE A64FX	8
3.1.1 Description	8
3.1.2 Available resources	8
3.1.3 In-kind system expected benefit for EUPEX	9
3.1.4 How to access the system	9
3.1.5 Support and documentation	9
3.2 E4 ARMIDA	10
3.2.1 Description	10
3.2.2 Available resources	10
3.2.3 In-kind system expected benefit for EUPEX	11
3.2.4 How to access the system	11
3.2.5 Support and documentation	12
3.3 IT4INNOVATIONS KAROLINA GPU PARTITION	12
3.3.1 Description	12
3.3.2 Available resources	13
3.3.3 In-kind system expected benefit for EUPEX	13
3.3.4 How to access the system	13
3.3.5 Support and documentation	14
3.4 IT4INNOVATIONS BARBORA CPU PARTITION	14
3.4.1 Description	14
3.4.2 Available resources	14
3.4.3 In-kind system expected benefit for EUPEX	14
3.4.4 How to access the system	15
3.4.5 Support and documentation	15
3.5 UNIBO MONTE CIMONE	15
3.5.1 Description	15
3.5.2 Available resources	17
3.5.3 In-kind system expected benefit for EUPEX	17
3.5.4 How to access the system	17
3.5.5 Support and documentation	18
3.6 UNITO EPI-TO	18
3.6.1 Description	18
3.6.2 Available resources	18
3.6.3 In-kind system expected benefit for EUPEX	19
3.6.4 How to access the system	19
3.6.5 Support and documentation	19
4 ANNEX IN-KIND SYSTEMS	20
4.1 TGCC CLOUD	20
4.1.1 Description	20
4.1.2 Available resources	20
4.1.3 In-kind system expected benefit for EUPEX	20



4.1.4	How to access the system	21
4.1.5	Support and documentation	21
4.2	IT4INNOVATIONS GITLAB	21
4.2.1	How to access the system	21
4.2.2	Support and documentation	22
5	OVERVIEW OF SYSTEMS AND EXPECTED BENEFITS FOR EUPEX	23
5.1	JOLIOT-CURIE-IRENE A64FX	23
5.2	E4 ARMIDA	23
5.3	IT4INNOVATION KAROLINA GPU PARTITION	23
5.4	IT4INNOVATION BARBORA CPU PARTITION	23
5.5	UNIBO MONTECIMONE	24
5.6	UNITO EPI-TO	24
6	ACRONYMS AND ABBREVIATIONS	25
Figure 1	- Pilot Stages	7
Figure 2	- Armida front (left) and rear (right) view	10
Figure 3	- Power meter connection diagram	11
Figure 4	- The Monte Cimone cluster	16
Figure 5	- Monte Cimone board and integration	16
Figure 6	- EPI-TO Platform	18
Table 1	- Acronyms and Abbreviations	25



1 Executive Summary and introduction

This document means to provide detailed descriptions of all systems made available in-kind by partners in their supercomputing centers to the EUPEX project.

Overall, 6 partners provide system resources in-kind.

The availability of the in-kind pilot systems as first phase is key to define the requirements for the Exascale pilot architecture and characteristics. Relying on representative application profile analysis and co-design studies performed on the in-kind systems, WP3 will provide valuable inputs to both WP4 and WP5 in order to specify the Pilot hardware and software components.

Applications porting, optimization and scaling tests will be done on these systems made available in-kind, based on state-of-the-art ARM technologies, innovative interconnect, optionally connected to accelerated modules. Some systems although not based upon ARM, offer interesting features, which are exposed.

In chapter 2, "In-kind resources as first stage towards the pilots for EUPEX partners", the key role of in-kind systems in the timeline project is reminded, as "beta" pilot systems design (see figure 1), needs application and software stack requirements input. Phases of In-kind systems presentation and their role in the project are described.

The descriptions given in chapter 3 "In-Kind systems for EUPEX project" cover hardware and software aspects, architecture, software ecosystem, integration and data organisation, services, types and amount of resources available for the EUPEX project, explanation on process to get access (including security and authentication constraints), and resources grant process.

On top of this, two partners also provide OpenStack partitions and Gitlab services as in-kind contribution. These systems will provide annex support for EUPEX activities (cf. Section 4).

- IT4Innovations's will provide GitLab services
- CEA's OpenStack partition will host the OCEAN code repository as well as the EUPEX-internal ticketing system

Users of EUPEX in-kind systems, mainly from WP3 and WP5, will find in this document a hand-book to identify the outcomes targeted for their assessment activities for co-design, and all the information to obtain resources and access.



2 In-kind resources as first stage towards the pilots for EUPEX partners

The EUPEX consortium is designing, building, and validating the first HPC platform integrating European general purpose processor technology (EPI), and European interconnect technology (BXI). EUPEX also covers end-to-end the spectrum of required software, from system software, to development tools, and to the application level.

The EUPEX SW stack will comprise all necessary software components, combining developments of European-funded R&D projects, primarily the DEEP projects, the SEA project series, the Mont-Blanc projects, and others with best-of-breed open source software where needed. This SW stack will enable the efficient exploitation of the resources available in large-scale MSA systems. This way, the co-design applications will be able to demonstrate Exascale readiness on the EUPEX platform.

On the application level, both scientific applications (from various scientific fields) and data-centric applications (AI, HPDA) will be considered. For this, a suite of applications has been selected in order to demonstrate the effectiveness of the EUPEX system for the whole range of workloads of interest.

As of today, some key elements of the future pilot system are not available yet, such as the SiPearl Rhea processor. Nevertheless, the EUPEX partners need software development vehicles. Therefore, in EUPEX we suggest this staged approach in order to allow for a meaningful co-design process and to allow applications and the SW stack to prepare for the future hardware components. The stages, as foreseen in the DoA are as follows (cf. Figure 1 - Pilot Stages):

1. **EUPEX-Alpha:** first phase deployed and available from month six of the project, provided in-kind by the EUPEX partners. One of these systems is an existing Fujitsu A64FX partition, hosted at TGCC-CEA. It is as of today one of the very first alternative to Marvell ThunderX4, which will not be available at the start of EUPEX.
2. **EUPEX-Beta:** Second phase once the first Rhea first sample are available and integrated into the reference board developed in the EPI project. To get a clean version of the board (all problems corrected). A minimal air-cooled packaging compatible with safety standards will be designed for a Pilot1-Beta offering up to 16 Rhea reference boards hosting one Rhea processor connected to one GPU PCIe card and one interconnect PCIe card.
3. **EUPEX-target platform:** the complete HPC system, fully deployed and operational

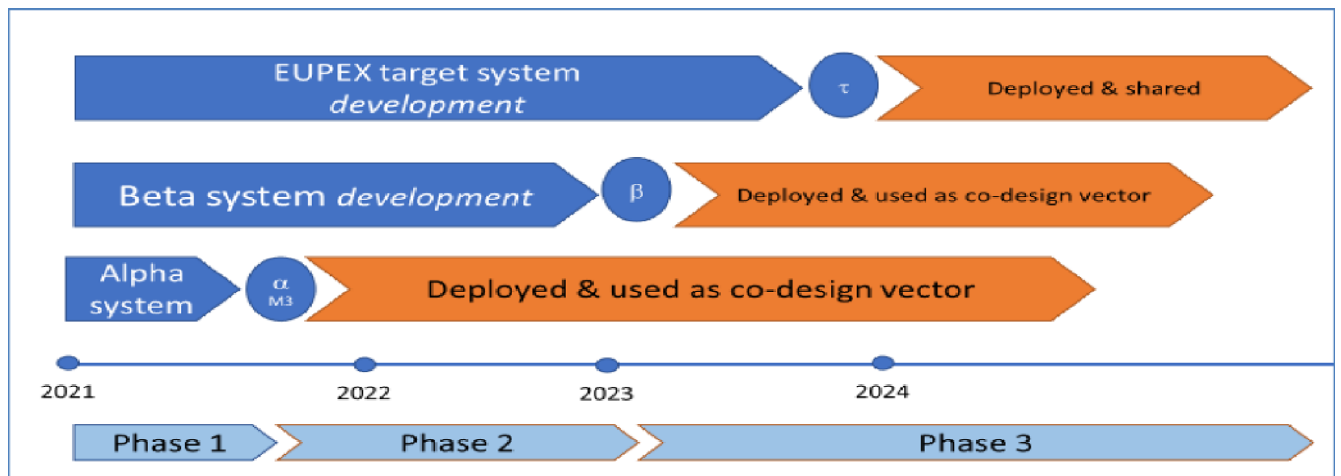


FIGURE 1 - PILOT STAGES

This deliverable focuses on the Alpha systems. It describes in detail the different systems provided in-kind by the partners. In particular section 3 of this deliverable shall serve the EUPEX partners from WP3, and WP5 as user guide and reference for using these systems. Section 5 also summaries the roles of the in-kind system as part of the EUPEX project.

The information on these in-kind systems are also available to all EUPEX partners via the EUPEX wiki^[1]. Moreover, the partners of WP6 have also organised two webinars to present the systems and their particularities to the EUPEX partners.

The webinar on 04/05/2022 focused on the A64FX partition hosted at CEA and on the systems hosted at IT4Innovation (Section 3.1 and 3.4).

The webinar on 17/06/2022 provided more insight on the in-kind systems at E4, the University of Bologna and of the University of Torino (cf. Sections 3.2, 3.5 and 3.6).

Around 30 participants each attended both sessions. The recording and the slides are available on to all partner via the EUPEX wiki.

[1] [https://atos365.sharepoint.com/sites/300000693/wiki/EUPEX In-kind systems.aspx](https://atos365.sharepoint.com/sites/300000693/wiki/EUPEX%20In-kind%20systems.aspx)



3 In-kind systems for EUPEX Project

3.1 Joliot-Curie-Irene a64fx

Based on Fujitsu PRIMEHPC FX700 technology, the "A64FX" partition consists of 80 single-socket DDR-less A64FX compute nodes connected via Mellanox InfiniBand and integrated into GENCI's Joliot-Curie supercomputer.

Joliot-Curie is a machine dedicated to European academic and industrial open research, hosted at the TGCC. Its computer architecture, operation and user support have been entrusted to the teams of the CEA center at Bruyères-le-Châtel.

The A64FX partition give scientists the opportunity to port their applications and prepare for the use of the future European processor by relying on the unique features of the A64FX processor such as its SVE vector instruction set and the use of HBM2 fast access memory (32 GB per node).

3.1.1 Description

Alpha system – Fujitsu FX700:

- 2 A64FX login nodes
- 78 A64FX nodes in a compute partition
- 3744 cores
- 2 IO router nodes for home and Lustre global file systems access

Node specification:

- 1 CPU A64FX (instruction set Armv8.2-A SVE) 48 cores
- 3.072 TFlops @ 2.0 GHz
- 32GB of HBM2 memory @ 1024 GB/s (24 GB for users)
- 1 HDR-100 NIC

3.1.2 Available resources

A computing project will be assigned to each task within a EUPEX WP and the EUPEX task manager will be the TGCC project manager.

For each project the following resources will be allocated:

- 5000 core-hours on the computing partition **a64fx**
- Standard data volumes on the shared file systems (see [TGCC Irene documentation](#) for more technical details):
 - Home (3GB/user)
 - Scratch (100 TB and 2M files / group)
 - Work (5 TB and 500 000 files / group)
 - Store (50 000 files / group)



3.1.3 In-kind system expected benefit for EUPEX

In the preparatory phase of the EUPEX project, this partition was identified as the "alpha-system", i.e. the main system for preparing applications and middleware for the EUPEX project.

Moreover, its positioning in the target infrastructure of the EUPEX pilot within the TGCC allows, as of today, to benefit from an integration with an HPC administration solution: OCEAN. This offers all partners the opportunity to prepare applications, runtimes and tools in the target infrastructure in order to simplify their deployment and integration on the pilot.

3.1.4 How to access the system

Resources allocation

As described above, **a64fx** resources are allocated by project. In the EUPEX context, a project is allocated for each EUPEX task and the TGCC project manager is the EUPEX task leader.

The persons in charge of each EUPEX task can now request a project by contacting ro.tgcc@cea.fr and indicating WP, Task, needs. The TGCC team will initialize the resource allocation, reference the project in: [EUPEX Joliot Curie Irene a64fx accounting.xlsx](#) and contact the project leader to communicate the following steps.

Access

For users who do not have a TGCC account, please fill in this form [Account form TGCC](#) with:

- Type of project: EUPEX
- Project number: See [EUPEX- Joliot Curie Irene a64fx accounting.xlsx](#)
- Project manager: See [EUPEX- Joliot Curie Irene a64fx accounting.xlsx](#)

Sign it and send it to [TGCC Hotline <Hotline.tgcc@cea.fr>](mailto:Hotline.tgcc@cea.fr)

For users with a TGCC account, please ask your EUPEX task manager / TGCC project manager to contact the TGCC Hotline [<Hotline.tgcc@cea.fr>](mailto:Hotline.tgcc@cea.fr) in order to be added to the corresponding project.

3.1.5 Support and documentation

For support on TGCC resources, please contact [TGCC Hotline <Hotline.tgcc@cea.fr>](mailto:Hotline.tgcc@cea.fr)

[TGCC public user documentation](#)

[TGCC Irene documentation](#) (need TGCC account)

3.2 E4 ARMIDA

3.2.1 Description

ARMIDA is an ARM cluster, based on Cavium ThunderX2 CPU. The systems can be reached via the tlnode01 login node. On all nodes the Operating System is Red Hat Enterprise Linux release 8.0 (Ootpa), with Linux kernel 4.18.0-80. The front and rear view are present in the next two figures.

Tcnode0[1-7]

All nodes are dual socket, 32 cores per socket with Cavium ThunderX2 processors, 256 GB RAM, connected with InfiniBand 100 Gb/s network.

Tcnode0[5-6] (nodes with GPU)

These two nodes have the same configuration of the nodes described above but they offer access to extra resources: one NVIDIA V100 GPU per node.



FIGURE 2 - ARMIDA FRONT (LEFT) AND REAR (RIGHT) VIEW

3.2.2 Available resources

Resources are available via slurm resource manager and no quota are set for users. Note : To use Slurm scheduler is mandatory to load the slurm environment module.

The /home and /opt/share directories are shared via NFS protocol towards all nodes.

A peculiarity of the E4 laboratory is the availability of a Zimmer LMG95 power meter that enables to measure the power consumption at the plug of a power supply.

The product is described in the following link:

<https://www.zes.com/en/Products/Discontinued-Products/Energy-and-Power-Meters/LMG95>

This power meter therefore allows to perform measurements of the overall consumption of a server to determine the consumption value in idle and at full load. This value can be compared with the consumption of GPUs alone, the values of which can be obtained with the “nvidia-smi” utility for NVIDIA GPUs.

The meter is available upon request and can be connected to one server at a time.

The connection diagram is shown in the following figure:

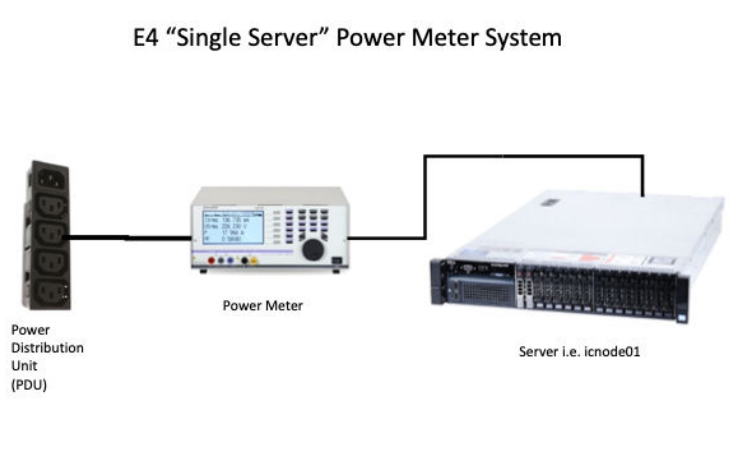


FIGURE 3 – POWER METER CONNECTION DIAGRAM

The LMG95 allows to measure the active power with 0.1 Hz sampling frequency. To allow the simultaneous measurement of more servers, we are evaluating to use a metered PDU, which is able to perform a measurement to the plug of several servers connected to it.

3.2.3 In-kind system expected benefit for EUPEX

The proposed architecture (ARM) is the same as the European Rhea processor. The idea is to be able to develop and test the code in an environment as close as possible to the production one. With the use of energy consumption sensors (power meter, metered PDU, queries to OS from command line), detailed studies on energy efficiency are possible.

3.2.4 How to access the system

E4 System is accessed via VPN, to request VPN the first step is to create an account in the E4 User Portal: <https://service.e4company.com/servicedesk/customer/user/login?destination=portals>

After user creation please log in the portal and chose “E4 R&D Lab Portal” then “Access Request for VPN”.



At this point fill the form with “VPN EUPEX” in the Summary and fill it with all other requested information.

Now to request the cluster access write an Email to marco.cicala@e4company.com and daniele.gregori@e4company.com in copy with subject “ARMIDA Access Request for EUPEX Project”.

3.2.5 Support and documentation

For any issue open a request on the E4 User portal:

<https://service.e4company.com/servicedesk/customer/user/login?destination=portals>

describing the problem and how-to reproduce it.

More documentation on ARMIDA cluster is available in the E4 wiki page:

<http://172.18.16.119:8004/> , E4 VPN is mandatory to access it.

3.3 IT4Innovations Karolina GPU partition

IT4Innovations’ Karolina is a EuroHPC petascale system of overall theoretical peak performance 15.7 PFLOP/s. Its accelerated partition is one of the greenest computing architectures nowadays, using 8 fully connected GPUs per node.

3.3.1 Description

- #71 in top500, #8 in green500, <https://www.top500.org/system/180026/>
- 72 nodes, 2x AMD EPYC 7763 + 8x NVIDIA A100, nvswitch
- Very FAT multi-GPU nodes
- 1TB RAM + 8*40 GB HBM per node
- 4x 200 Gb/s IB port
- 5,017.6 GFLOP/s per compute node
- 9,080.2 TFLOP/s theoretical peak performance

Software:

- Lmod module environment
- PBS Pro scheduler
- Development software stacks
 - GNU, OpenMPI
 - AMD
 - Intel
 - Nvidia
- Profiling tools (not a complete list)
 - Open-source CoE POP tools



- Score-P
- Extra-E
- ARM Forge
- AMD
- Intel
- Nvidia

3.3.2 Available resources

A EuroHPC development project was accepted. This project secures initial 3,000 node hours. Another proposal will be written on demand.

Available storages:

- HOME (25 GB, 500 thousand entries) - 1.93 GB/s write, 3.1 GB/s read
- Short-term parallel Scratch storage shared by the project users (20 TB, 20 Million entries, 90 days) - 730.9 GB/s write, 1198.3 GB/s read
- Mid-term parallel Project storage shared by the project users (20 TB, 5 Million entries, project lifetime) - 39 GB/s

3.3.3 In-kind system expected benefit for EUPEX

Since the future European exascale system will accommodate several accelerators per computing node the Karolina system, which provides 8 fully-connected GPUs per node, give us an opportunity to develop, modify and test the EUPEX applications for efficient usage of such architecture.

Moreover, the system allows to measure energy consumption (AMD RAPL + NVML) and tune CPU and GPU parameters (frequencies, power capping), so it can be used for energy-efficiency analysis of the EUPEX application or development of the energy-efficient/power aware runtime systems.

3.3.4 How to access the system

To create IT4Innovations user account one must send a digitally signed email to support@it4i.cz with the request. The email must contain:

- Full name and affiliation
 - Statement that you have read and accepted the [Acceptable use policy document](#) (AUP)
 - Attach the AUP file
 - Preferred username (4 to 7 characters)
 - ID of the project – DD-22-41
 - Public part of an SSH key
-
- `ssh <login>@karolina.it4i.cz`
 - The cluster uses PBS Pro job scheduler, for a job submission to Karolina GPU partition, one must specify target queue **-q qgpu_eurohpc**.



3.3.5 Support and documentation

General documentation of the IT4Innovations: <https://docs.it4i.cz>

Karolina documentation: <https://docs.it4i.cz/karolina/hardware-overview>

Email to Support: support@it4i.cz

Energy-efficiency support: ondrej.vysocky@vsb.cz

3.4 IT4Innovations Barbora CPU partition

Barbora cluster is a relatively small Intel CPUs-powered cluster. In the context of the EUPEX project this system is interesting for the possibility of hardware parameters tuning and precise power monitoring.

3.4.1 Description

- 192 nodes, 2x Intel Cascade Lake 6240
- 192 GB RAM
- 848 TFLOP/s theoretical peak performance (including accelerated partition)

Software:

- Lmod module environment
- PBS Pro scheduler
- Development software stacks
 - GNU, OpenMPI
 - Intel
- Profiling tools (not a complete list)
 - Open-source CoE POP tools
 - Score-P
 - Extra-E
 - ARM Forge
 - Intel

3.4.2 Available resources

No resources secured at the time of the deliverable submission. Will be handled on request.

3.4.3 In-kind system expected benefit for EUPEX

The non-accelerated Barbora nodes are equipped with out-of-band Atos|Bull High Definition Energy Efficiency Monitoring (HDEEM) system, that allows to precisely measure energy consumption of the applications executed on the node. The system monitors the whole compute node power consumption, and several power subdomains (Voltage Regulators, VRs) such as CPU, DRAM or NIC. The HDEEM stores power samples in its internal memory, so also the power samples can be read during as well as after the



end of an application execution. Compute node is monitored at 1kHz frequency, while the remaining power domains at 100Hz frequency.

Barbora system allows to tune CPU parameters from the userspace (CPU core and uncore frequencies, power capping) and measure energy consumption using HDEEM or Intel RAPL, so it can be used for energy-efficiency analysis of the EUPEX application or development of the energy-efficient/power aware runtime systems.

3.4.4 How to access the system

To create IT4Innovations user account one must send a digitally signed email to support@it4i.cz with the request.

The email must contain:

- Full name and affiliation
 - Statement that you have read and accepted the [Acceptable use policy document](#) (AUP)
 - Attach the AUP file
 - Preferred username (4 to 7 characters)
 - ID of the project – DD-22-41
 - Public part of an SSH key
-
- ssh <login>@barbora.it4i.cz
 - The cluster uses PBS Pro job scheduler, for a job submission to Barbora partition with HDEEM system, one must specify target queue **-q qprod**.
 - Specific PBS parameters
 - Activate HDEEM
 - -l hdeem=2.2.7-1
 - Activate msr-safe to access model specific registers
 - -l msr=1.4.0

3.4.5 Support and documentation

General documentation of the IT4Innovations: <https://docs.it4i.cz>

Barbora documentation: <https://docs.it4i.cz/barbora/hardware-overview/>

Email to Support: support@it4i.cz

Energy-efficiency support: ondrej.vysocky@vsb.cz

3.5 UNIBO Monte Cimone

3.5.1 Description

Monte Cimone is based on the SiFive Freedom U740 RISC-V SoC HiFive Unmatched board integrated in an HPC node form factor.



FIGURE 4 - THE MONTE CIMONE CLUSTER

<http://arxiv.org/abs/2205.03725>

The E4 RV007 blade prototype system, adopted as Monte Cimone building Block, is a dual-board platform server, with a form factor of 4.44 cm (1 RackUnit) high, 42.5 cm width, 40 cm deep. Two 250 W power supplies, one for each board (compute node), are installed inside the case. This allows to turn on every single compute node individually, see Figure 1, and makes the system ready with abundant power headroom for future expansions with hardware accelerators and PCIe Network Card connector.

The board follows the Industry Standard Mini-ITX with a size of 170 mm per 170 mm. It features one SiFive Freedom U740 SoC, 16 GB of 64-bit DDR4 memory operating up to 1866s MT/s and high-speed interconnects with PCIe Gen 3 x16 (but it's limited to x8 lanes), one Gigabit Ethernet, and four USB 3.2 Gen 1, see Figure.

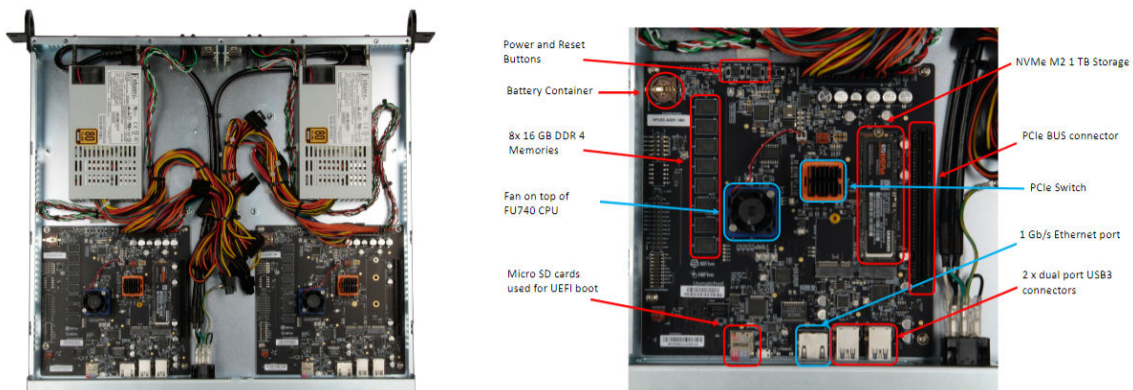


FIGURE 5 – MONTE CIMONE BOARD AND INTEGRATION

In RV007 node the M.2 M-key expansion slot is occupied by a 1 TB NVME 2280 SSD Module storage device used by the Operating System. The Micro SD card is present and used for the UEFI Boot. Two buttons for reset and power up operations are available on top of the board and in front of the case.

The FU740-C000 is a Linux-capable SoC powered by SiFive's U74-MC, the first (to the best of our knowledge) commercially available superscalar heterogeneous multi-core RISC-V Core Complex. The FU740-C000 is compatible with all applicable RISC-V standards.

The U74-MC core complex is composed of four 64-bit U74 RISC-V (Application) cores. Each U74 core has a dual issue in-order execution pipeline, with a peak sustainable



execution rate of two instructions per clock cycle. The U74 core supports Machine, Supervisor, and User privilege modes as well as standard Multiply, Single-Precision Floating Point, Double-Precision Floating Point, Atomic, and Compressed RISC-V extensions.

The SiFive Freedom board features a Microsemi VSC8541 chip to interconnect the SiFive Freedom U740 SoC with a single port gigabit Ethernet copper interface. Moreover, we equipped two of the compute nodes with an Infiniband Host Channel Adapter (HCA) widely used in large-scale HPC systems. We target an Infiniband FDR HCA (56Gbit/s) to leverage RDMA communications among different nodes to improve the network throughput and the communication latency. We used a Mellanox ConnectX-4 FDR HCA interconnect through the PCI-e interface available on the compute node. This HCA support x8 free PCIe Gen 3 lanes, which are currently supported by the vendor. The first experimental results show that the kernel is able to recognise the device driver and mount the kernel module to manage the Mellanox OFED stack. We are not able to use all the RDMA capabilities of the HCA due yet-to-be-pinpointed incompatibilities of the software stack and the kernel driver. Nevertheless, we successfully run an IB ping test between two boards and between a board and an HPC server showing that full Infiniband support could be feasible. This is currently a feature under development.

In addition, the SiFive Freedom U740 SoC features 7 separated power rails including the core complex, IOs, PLLs, DDR subsystem and PCIe one. The HiFive Unmatched board implements separated shunt resistors in series with each of the SiFive U740 power rails as well as for the on-boards memory banks.

3.5.2 Available resources

Resources are available via slurm resource manager and no quota are set for users. Please remember to load the slurm environment module.

The /home and /opt/share directories are shared via NFS protocol towards all nodes.

A peculiarity of the ECS-Lab at UNIBO is the availability of a power meter which can be connected to the blade that enables to measure the power consumption at the plug of a power supply.

3.5.3 In-kind system expected benefit for EUPEX

The proposed architecture (RISC-V) is the first RISC-V cluster worldwide. The idea is to be able to develop and test the code in an environment as close as possible to the production one.

3.5.4 How to access the system

Access will be granted starting Q3 2022. Email to a.bartolini@unibo.it in copy with subject "Monte Cimone Access Request for EUPEX Project".

3.5.5 Support and documentation

Being a research prototype, hosted in a research institution support is given with best-effort by the research group. Additional information can be found <http://arxiv.org/abs/2205.03725>

3.6 UNITO EPI-TO

3.6.1 Description

The EPI-TO platform is a cluster of 4 servers each of them being a GigaByte G242-P32 server equipped with 1 Ampere Altra Q80-30 (80-core Arm Neoverse N1 - https://en.wikichip.org/wiki/arm_holdings/microarchitectures/neoverse_n1) with 512GB RAM, 2 NVidia A100 GPU, 2 NVidia BlueField-2 DPU (200Gb/s DOCA-enabled).

The cluster is connected to an external archive storage system (NFS).

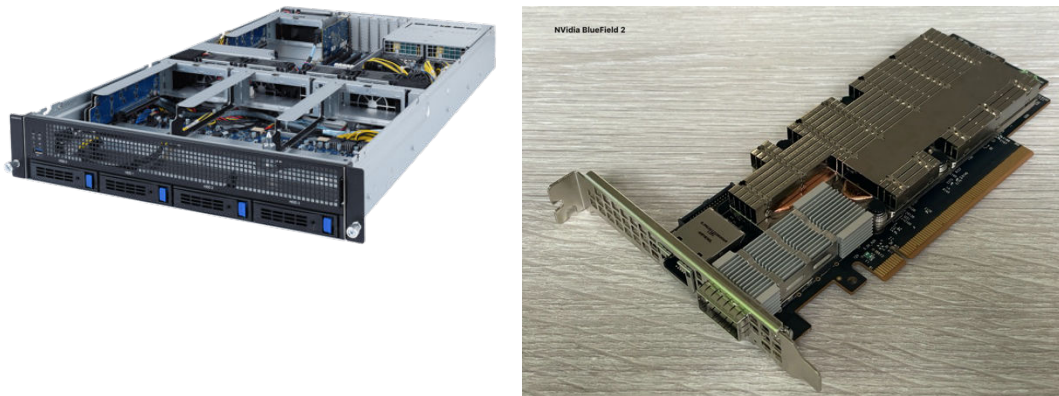


FIGURE 6 – EPI-TO PLATFORM

3.6.2 Available resources

Resources are available via the SLURM-UC, a variant of the SLURM workload manager extend with UNITO's "Urgent Computing" (see Sec. 2.8.4). This project secures initial 5000 core-hours per beneficiary, the initial amount can be extended with a motivated request. The system will not support long production runs as it is a machine devoted to research on system software and development tools.



3.6.3 In-kind system expected benefit for EUPEX

The proposed architecture (ARM) is the same as the European Rhea processor. The EPI-To cluster integrates two high-end GPU per node (NVidia A100), making possible to test accelerated ARM codes, and the 2 NVidia BF2 DPU accelerators per node, which make it possible to test an entirely different accelerator. EPI-TO is part of the NVidia dev-kit which has been distributed to only 100 universities worldwide (UNITO only in Italy).

3.6.4 How to access the system

Access will be granted starting Q3 2022. Email to iacopo.colonelli@unito.it with subject “EPITO Access Request for EUPEX Project”.

EPITO can be accessed via UNITO’s UrgentSLURM system, that makes it possible to book usage slots using a web calendar or submit a standard SLURM job. UrgentSLURM maintains all the SLURM features, and extend them with the possibility to book execution slots in the platform using a web calendar. Booking a number of processing elements guarantee their availability in the booked time slot fully accounting the booked slots to the user, no matter if the processing elements are really used or if the job terminates before the booked period is expired.

3.6.5 Support and documentation

Being a research prototype, hosted in a research institution support is given with best-effort by the research group. Additional information on how access to the EPITO system, installed software, and help desk can be found at:

<https://hpc4ai.unito.it/documentation/>



4 Annex in-kind systems

4.1 TGCC Cloud

In addition to the supercomputers, the TGCC is providing to the European academic and industrial research communities a Cloud platform based on the Open Stack software solution.

4.1.1 Description

Openstack cluster:

- 3 management nodes
- 3 services nodes
- Hosting openstack services
- 4 storage nodes
- Providing 60.8 TB of SSDs (raw)
- Hosting Ceph services
- 20 Hypervisors
- Hosting User allocated/managed VMs

Hypervisor node specification:

- 2 CPUs Intel Cascade Lake G-6240
- 2 * 18 cores @ 2.6 GHz
- 192 GB DDR4
- 10 Gbit/s network

4.1.2 Available resources

A set of resources of this platform will be reserved for the use of EUPEX. It will initially consist of 32 vCPUs, 64 GB of RAM and 160GB of storage and can be adjusted as needed.

4.1.3 In-kind system expected benefit for EUPEX

The EUPEX project partners will thus have the possibility of setting up virtual machines allowing the hosting of services easily accessible from the Internet in order to share and collaborate on their software and/or the use of systems accessible within the framework of the project.

Services accessible in this way will be associated with the DNS domain of the EUPEX project.



4.1.4 How to access the system

Resources allocation

The persons responsible for each EUPEX task can now request a resource allocation by contacting ro.tgcc@cea.fr. Please describe the service to be provided, the requirement (vCPU, disk, RAM), the WP and the corresponding task.

The TGCC team will initialize the resource allocation, reference the project in and contact the project manager to communicate the following steps.

Access

For users who do not have a TGCC account please fill in this form  [Account form TGCC](#), sign it and send it to TGCC Hotline [<hotline.tgcc@cea.fr>](mailto:hotline.tgcc@cea.fr).

For users with a TGCC account please contact the TGCC Hotline [<hotline.tgcc@cea.fr>](mailto:hotline.tgcc@cea.fr) with your task manager as a copy in order to be added to the corresponding project.

4.1.5 Support and documentation

For support on TGCC resources, please contact [TGCC Hotline <hotline.tgcc@cea.fr>](mailto:hotline.tgcc@cea.fr)

Documentation:

- [TGCC public user documentation](#)
- [TGCC Irene documentation](#) (TGCC account needed)

4.2 IT4Innovations Gitlab

Under the EUPEX project the internal git repositories will be maintained at IT4Innovations' gitlab server <https://code.it4i.cz>. This gitlab is possible to use for everyone with IT4I account.

IT4Innovations is holder of the Information Security Management System certificate according to the international ISO/IEC 27001:2013 standard, see <https://www.it4i.cz/en/for-users/information-security>

Every user may create private, internal and public repositories, as well as groups. For the security reasons EUPEX internal git repositories should be created under a specific group of users, to make the repository accessible only to them.

4.2.1 How to access the system

It is not possible to create IT4Innovations user account prior a compute-resources project exists. One must send a digitally signed email to support@it4i.cz to request for IT4I user account. See more details on [email digital signature](#)

The email must contain:

- Full name and affiliation



- Statement that you have read and accepted the [Acceptable use policy document](#) (AUP)
- Attach the AUP file
- Preferred username (4 to 7 characters)

The preferred username must associate with first and last name or be otherwise derived from it. Only alphanumeric sequences and dash signs are allowed. Note that the system will automatically add the it4i- prefix to the username.

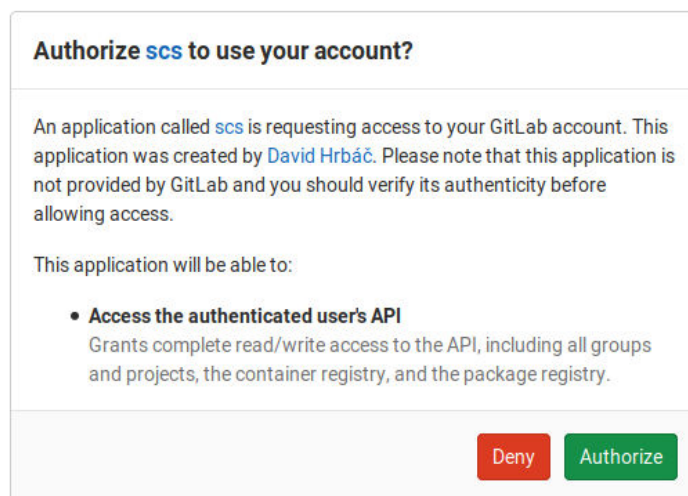
- Public part of an SSH key

If it is not provided in the ticket, it must be [added manually](#) after the account is created.

- In case of choosing an [Alternative way to personal certificate](#), a **scan of photo ID** (personal ID or passport or driver license) is required.
- Self-service portal to manage your HPC resources: scs.it4i.cz

A EuroHPC development project for IT4Innovations' Karolina system was accepted. EUPEX user accounts will be created under this project.

After the very first log in every user must **Authorize** connection of the IT4I and gitlab accounts. The following window should pop up automatically.



4.2.2 Support and documentation

The Gitlab groups documentation: <https://docs.gitlab.com/14.7/ee/user/group/>

To control the permissions, permissions can be specified within the group, and optionally also within each subgroup or repository, as described in <https://docs.gitlab.com/ee/user/permissions.html>

Email to IT4Innovations support: support@it4i.cz



5 Overview of systems and expected benefits for EUPEX

5.1 Joliot-Curie-Irene A64FX

(For precise description refer to 3.1)

This system is identified as the " Software-defined vehicle (SDV) ", i.e., the main system for preparing applications and middleware for the EUPEX project. Its positioning in the target infrastructure and centralized administration (OCEAN) of the future EUPEX pilots within the TGCC, will manage a simplified environment for the partners to deploy and prepare applications, runtimes and on the upcoming pilot.

5.2 E4 ARMIDA

(For precise description refer to 3.2)

ARMIDA hosts ARM processors close to the European Rhea processor. This is of a great interest for application tuning and profile projections. the code in an environment as close as possible to the production one. Moreover, features energy consumption area (power meter, metered PDU, queries to OS from command line), enable energy to solution detailed studies.

5.3 IT4Innovation KAROLINA GPU partition

(For precise description refer to 3.3)

The Karolina GPU partition with 8 GPUs per node, offers a developing testbed for the EUPEX applications and reach efficient usage on accelerated architecture, which will be implemented in the pilot systems. Energy to solution evaluation and optimisation will be allowed by the features available on the system: measure of energy consumption (AMD RAPL + NVML) and tune CPU and GPU parameters (frequencies, power capping).

5.4 IT4Innovation BARBORA CPU partition

(For precise description refer to 3.4)

Barbora nodes are equipped with scalar processors with out-of-band Atos|Bull High Definition Energy Efficiency Monitoring (HDEEM) system. Power consumption of the whole compute blade and range of power subdomains are measured at high frequency (1 kHz blade, 100 Hz on node subdomains). Particularly precise energy profiling can be performed, as well as tuning of the hardware parameters at userspace.



5.5 UNIBO MONTECIMONE

(For precise description refer to 3.5)

The proposed architecture (RISC-V) is the first RISC-V cluster worldwide. This processor architecture, free of licence, is among the best candidates for European technology. MonteCimone opens the opportunity to assess to maturity of RISC V with respect to ARM for HPC usage, where memory performance, cache management, concurrency are critical, as well as energy efficiency. The benefit for EUPEX outcome is to be able to provide feedback to RISC V design and contribute to a wider choice for European technology candidates.

5.6 UNITO EPI-TO

(For precise description refer to 3.6)

The proposed architecture (ARM) is close to the European Rhea processor. The EPI-To cluster integrates two high-end GPU per node (NVidia A100), making possible to test accelerated ARM codes, and the 2 NVidia BF2 DPU accelerators per node, which make it possible to test an entirely different accelerator.



6 Acronyms and Abbreviations

Term	Definition
ARM	Advanced RISC Machines (Processor design)
A64FX	Fujitsu 64-bit ARM architecture microprocessor
CEA	Commissariat aux énergies alternatives
CoE	Center of Excellence
DDR	Double Data Rate (memory type)
GENCI	Grand Equipement National pour le Calcul Intensif
MPI	Message Passing Interface
NFS	Network File System
NVMe	Non volatile Memory Express
OCEAN	Cluster administration & management software stack based on CentOS
PBS	Portable Batch System
RISC-V	Reduced Instruction Set Computer - Vector (Open processor design)
SLURM	Simple Linux Utility for Resource Management
SSD	Solid State Drive
TGCC	Très Grand Centre de Calcul
VPN	Virtual Private Network

TABLE 1 - ACRONYMS AND ABBREVIATIONS