

WHEN PERFORMANCE MATTERS

# E4 Experience with RISC-V in HPC

Daniele Gregori Ph.D.

ISC 2024

# E4 Experience with RISC-V in HPC

## INDEX:

- Company profile
- Monte Cimone: The first RISC-V HPC cluster
- Monte Cimone Update 2024
- Ongoing RISC-V projects

# Company Profile

# E4 IN A NUTSHELL



2002 - 2022



Strategic Members  
<https://riscv.org/>

## WHO WE ARE

E4 Computer Engineering is an **Italian** Company, designs and manufactures highly technological solutions for HPC Clusters, Cloud, Data Analytics, Artificial Intelligence and Hyper-Converged infrastructure for the Academic and Industrial markets. We have been collaborating for years with the main research centers at national and international level (Cineca, CERN, ECMWF, LEONARDO) and we are involved in national and European projects in the HPC and AI fields (EuroHPC JU EPI, EUPEX, Horizon Europe)

## VISION

We explore future scenarios to find solutions for highly performing computational needs in application areas that are unimaginable today.

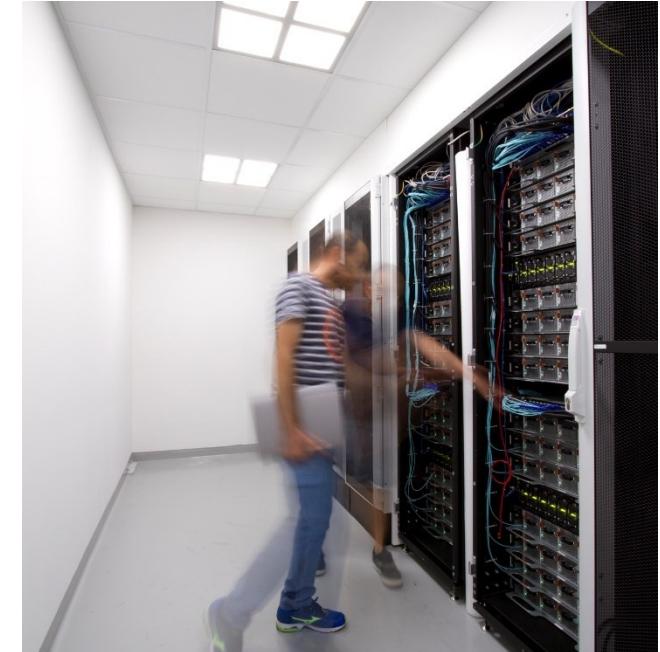
## MISSION

We anticipate the ever-accelerating disruptive transformation of our era, providing mature solutions in sophisticated technological contexts with a dizzyingly innovative approach

## APPROACH

Each E4 solution is **UNIQUE**, like each of our customers; **TESTED** in every single component; **VALIDATED** to verify the actual performance of each system and **SERVED** by technicians who provide assistance in the most extensive and complex Italian and European computing infrastructures.

# E4 TECH FACTORY



- Integration Facility where our technicians build servers or storage systems
- Burn In Room to improve E4 systems reliability with at least 72 hours of test that involves all components
- R&D Lab, with 6 standard racks with heterogeneous systems, 100kW, remote access available on demand to perform benchmarking, co-design, prototyping

# Monte Cimone:

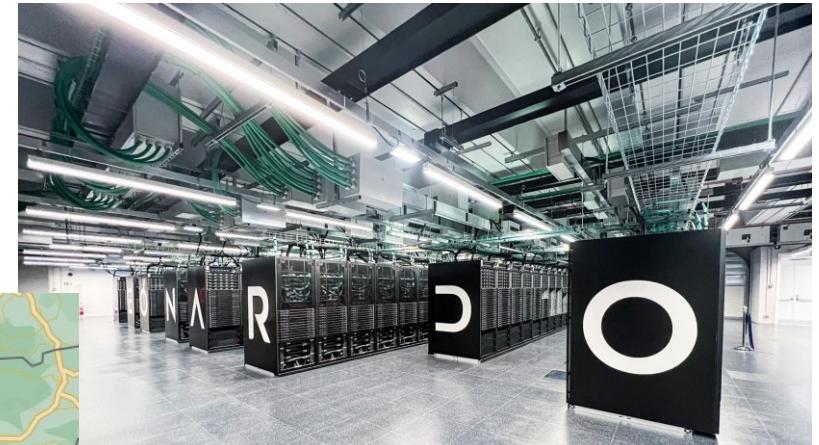
## The first RISC-V HPC cluster

### (2021)

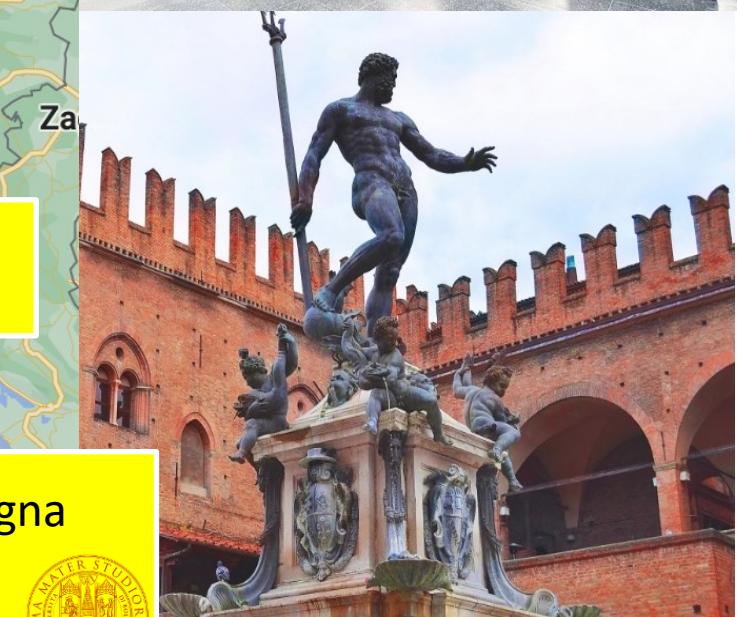
# E4 – UNIBO – CINECA and the Data Valley: A strong cooperation

Bologna New Technopole - 60MWatt datacentre

- CINECA Leonardo – The Italian Pre-exascale
  - 240 Pflops, 150PBytes, 4th Top500@Jun. 2023
- ECMWF HPCE – The new ECMWF supercomputer
  - 40+ Pflops



University of Bologna  
Monte Cimone  
1st RISC-V Cluster



# Monte Cimone Project

The **first physical prototype** and test-bed of a **complete RISC-V (RV64) compute cluster** integrating **compute, interconnect, a complete software stack for HPC and a full-featured system monitoring infrastructure.**

1. Ported and assessed the maturity of a HPC software stack composed of:
  - SLURM job scheduler, NAS filesystem, Spack package manager
  - compilers toolchains, scientific and communication libraries,
  - a set of HPC benchmarks and applications,
  - ExaMon datacenter automation and monitoring framework.
2. Characterized the HPL and STREAM benchmarks w. the toolchain and libraries installed by SPACK.
3. Extended the ExaMon monitoring framework to monitor the Monte Cimone cluster. Power consumption characterization of Monte Cimone.
4. «In Production» since May 2021.
  1. Access to external user (>40 users).
  2. Used in University Master courses and in two PhD summer school (> 100 students/year).
5. Now extended with SG2042 computing systems and accelerator cards

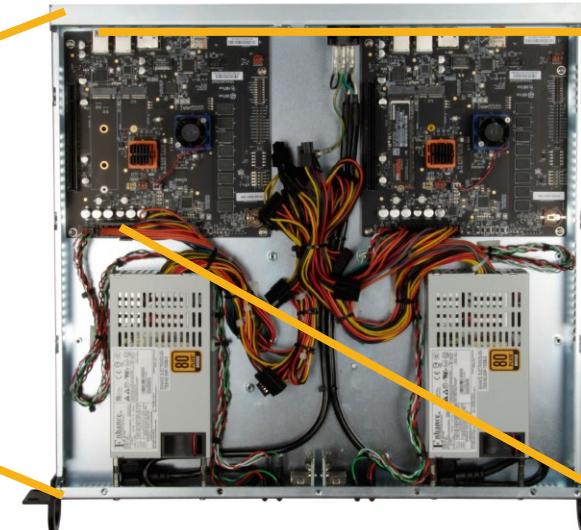


A. Bartolini *et al.*, "Monte Cimone: Paving the Road for the First Generation of RISC-V High-Performance Computers," *IEEE SOCC'22*,

F.Ficarelli et al. «Meet Monte Cimone: exploring RISC-V high performance compute clusters,» *ACM CF'22*

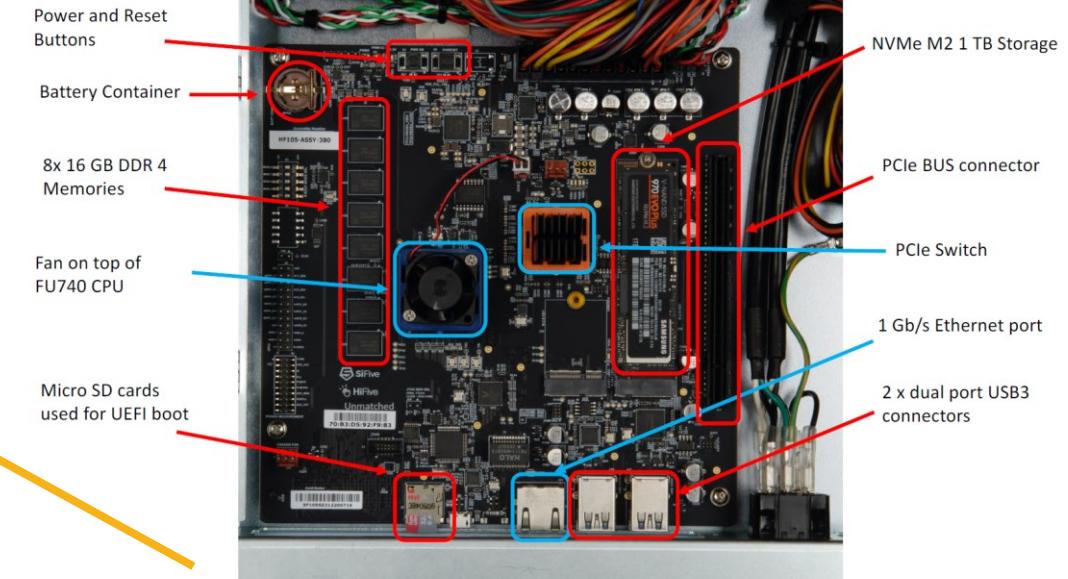
G. Mittone et al. «Experimenting with Emerging RISC-V Systems for Decentralised Machine Learning» *CF'23*

# Monte Cimone v1 Hardware



E4 RV007 blade prototypes

SiFive HiFive Unmatched board



## 4x E4 RV007 1U Custom Server Blades:

- 2x SiFive U740 SoC with 4x U74 RV64GCB cores
- 16GB of DDR4
- 1TB node-local NVME storage
- PCIe expansion card w/InfiniBand HCAs
- Ethernet + IB parallel networks

## SiFive U740 SoC w. 7 separated power rails:

- Core complex, IOs, PLLs, DDR subsystem and PCIe one.
- Board implements distinct shunt resistors

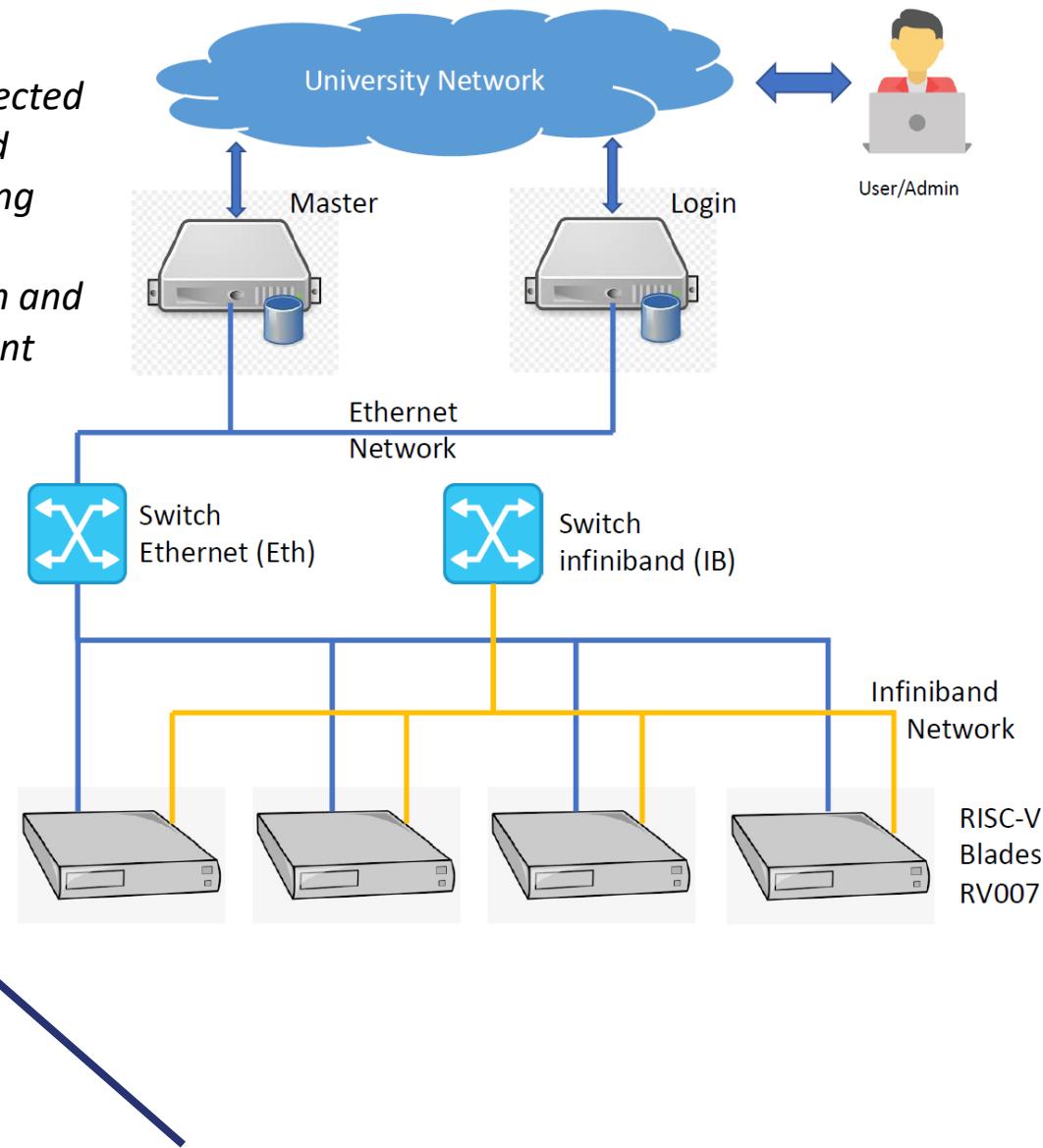


# Monte Cimone Software Stack:

## Production-level HPC software stack

- SLURM job scheduler, NFS filesystem, Nagios
- User-space deployed via **Spack** package manager
- Upstream and custom **toolchains**
- **Scientific libraries**
- Industry-standard **HPC benchmarks and applications** (e.g.: **quantumESPRESSO** suite)
- The **ExaMon** datacenter automation and monitoring framework

*The cluster is connected to a login node and master node running the job scheduler, network file system and system management software.*



## Monte Cimone: User-facing software stack

Package	Version
gcc	10.3.0
openmpi	4.1.1
openblas	0.3.18
fftw	3.3.10
netlib-lapack	3.9.1
netlib-scalapack	2.1.0
hpl	2.3
stream	5.10
quantumESPRESSO	6.8

- All software stack installed w. SPACK with the already present linux-sifive-u74mc
- Ubuntu 20.04 Linux O.S. installed with riscv64 image

# Monte Cimone Update 2024

## NEW CHIP SG 2042

&gt; 1 TFLOPS(FP64)

- 64 Cores
- 2 GHz
- 120 W TDP
- 3200 MHz (Max DIMM Frequency)
- 1 Gbit Ethernet
- 1 LPC



- Up to 256 GB RAM
- 4 MB L1 Cache
- 16 MB L2 Cache
- 64 MB L3 Cache
- 2 SPI Flash Interface
- 2 General SPI Controller

# NEW SERVER



### SG2042 BMC

The screenshot shows the BMC interface for an SG2042 server. The left sidebar includes sections for Overview, Logs, Hardware status, Operations, Settings, Security and access, and Resource management. The main area displays System information, Network information, and Status information. Under System information, there are sections for Server information (Model: --, Manufacturer: --) and Firmware information (Running: 2.13.0-dev-6420-g640f3622b, Backup: --). Under Network information, Hostname is ast2600-sophgo and Link status is LinkDown. Under Power information, Power consumption is Not available and Power cap is Disabled. In the Status information section, Event logs show 0 Critical and 0 Warning events. Inventory and LEDs indicate the System identify LED is Off.

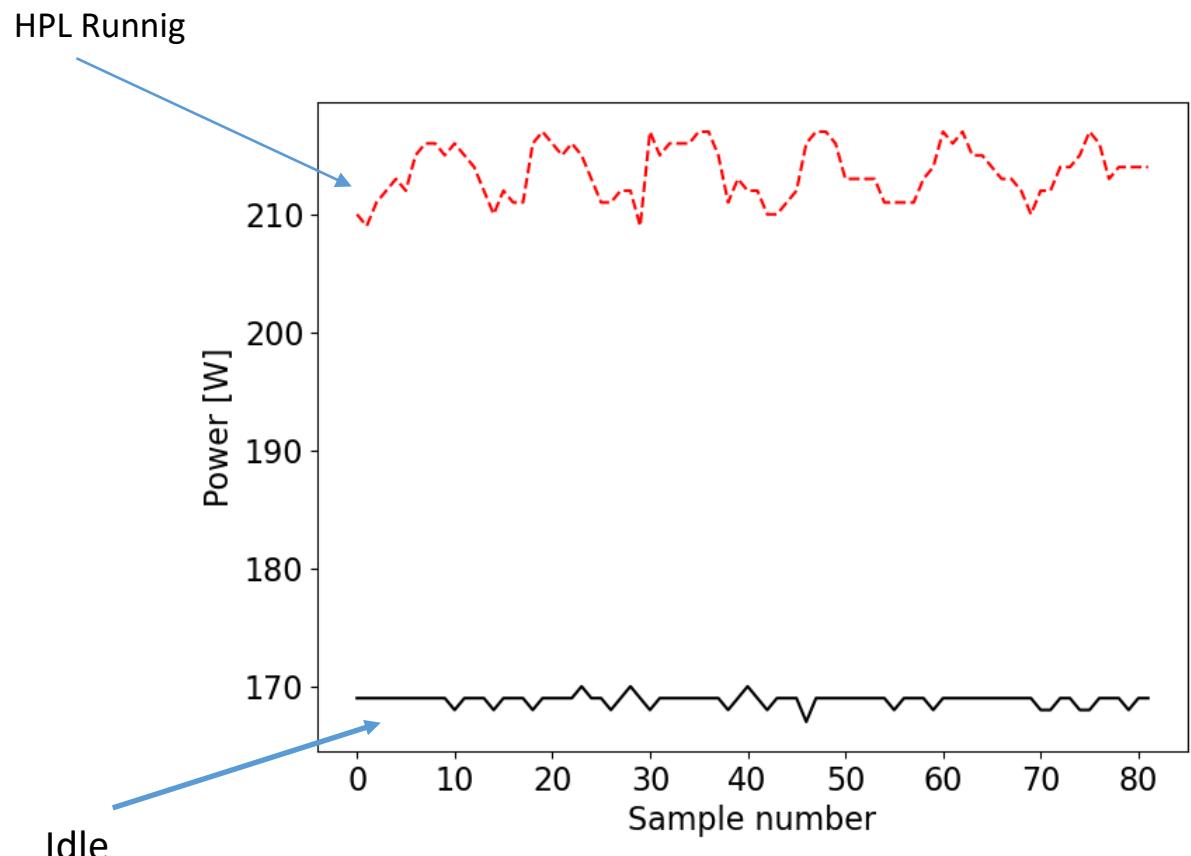
### SG2042 BMC

The screenshot shows the Sensors page for the SG2042 BMC. The left sidebar lists Overview, Logs, Hardware status, Inventory and LEDs, Sensors (which is selected), Operations, Settings, Security and access, and Resource management. The main area displays a table of 48 sensor items. The columns include Name, Status, Lower critical, Lower warning, Current value, Upper warning, and Upper critical. Examples of sensors listed include PSU1 Input Current (OK, 0.386 A), PSU1 Output Current (OK, 5.335 A), PSU1 Input Power (OK, 75.125 W), PSU1 Output Power (OK, 64.375 W), PSU2 Input Voltage (OK, 227.75 V), PSU2 Output Voltage (OK, 12.097 V), NVMe 1 Temp (OK, 47 °C), PSU2 Fan Speed 1 (OK, 4424 RPM), Pwm PSU2 Fan 1 (OK, 30 Percent), PSU2 Temperature (OK, 27.343 °C), PSU1 Input Voltage (OK, 227.25 V), PSU1 Output Voltage (OK, 12.007 V), PSU1 Fan Speed 1 (OK, 4392 RPM), Pwm PSU1 Fan 1 (OK, 30 Percent), and PSU11 Temperature (OK, 50 °C).

Name	Status	Lower critical	Lower warning	Current value	Upper warning	Upper critical
PSU1 Input Current	OK	-- A	-- A	0.386 A	-- A	-- A
PSU1 Output Current	OK	-- A	-- A	5.335 A	46 A	52 A
PSU1 Input Power	OK	-- W	-- W	75.125 W	1000 W	1200 W
PSU1 Output Power	OK	-- W	-- W	64.375 W	-- W	-- W
PSU2 Input Voltage	OK	-- V	-- V	227.75 V	-- V	-- V
PSU2 Output Voltage	OK	-- V	-- V	12.097 V	-- V	-- V
NVMe 1 Temp	OK	0 °C	5 °C	47 °C	110 °C	115 °C
PSU2 Fan Speed 1	OK	-- RPM	-- RPM	4424 RPM	-- RPM	-- RPM
Pwm PSU2 Fan 1	OK	-- Percent	-- Percent	30 Percent	-- Percent	-- Percent
PSU2 Temperature	OK	-- °C	-- °C	27.343 °C	50 °C	55 °C
PSU1 Input Voltage	OK	-- V	-- V	227.25 V	-- V	-- V
PSU1 Output Voltage	OK	-- V	-- V	12.007 V	-- V	-- V
PSU1 Fan Speed 1	OK	-- RPM	-- RPM	4392 RPM	-- RPM	-- RPM
Pwm PSU1 Fan 1	OK	-- Percent	-- Percent	30 Percent	-- Percent	-- Percent
PSU11 Temperature	OK	-- °C	-- °C	50 °C	55 °C	55 °C

```
[root@rcnode01 ~]# lscpu
Architecture:          riscv64
Byte Order:            Little Endian
CPU(s):                128
On-line CPU(s) list:  0-127
NUMA:
  NUMA node(s):        8
  NUMA node0 CPU(s):   0-7,16-23
  NUMA node1 CPU(s):   8-15,24-31
  NUMA node2 CPU(s):   32-39,48-55
  NUMA node3 CPU(s):   40-47,56-63
  NUMA node4 CPU(s):   64-71,80-87
  NUMA node5 CPU(s):   72-79,88-95
  NUMA node6 CPU(s):   96-103,112-119
  NUMA node7 CPU(s):   104-111,120-127
[root@rcnode01 ~]# free -m
              total        used        free      shared  buff/cache   available
Mem:       256890         2068     227769          1517        27052     251699
Swap:        8191           48        8143
[root@rcnode01 ~]# uname -a
Linux rcnode01.e4red 6.1.31 #1 SMP Sun Oct 22 00:58:22 CST 2023 riscv64 GNU/Linux
[root@rcnode01 ~]# cat /etc/redhat-release
Fedora release 38 (Thirty Eight)
[root@rcnode01 ~]#
```

# HPL PRELIMINARY TEST



Under NDA - Private & Confidential

The following parameter values will be used:

```
N      : 165504
NB     : 64
PMAP   : Row-major process mapping
P      : 8
Q      : 16
PFACT  : Right
NBMIN  : 4
NDIV   : 2
RFACT  : Crout
BCAST  : 1ringM
DEPTH   : 1
SWAP   : Mix (threshold = 64)
L1     : transposed form
U      : transposed form
EQUIL  : yes
ALIGN  : 8 double precision words
```



- The matrix A is randomly generated for each test.
- The following scaled residual check will be computed:  

$$\|Ax-b\|_{\infty} / (\text{eps} * (\|x\|_{\infty} * \|A\|_{\infty} + \|b\|_{\infty}) * N)$$
- The relative machine precision (eps) is taken to be 1.110223e-16
- Computational tests pass if scaled residuals are less than 16.0

---

T/V	N	NB	P	Q	Time	Gflops
WR11C2R4	165504	64	8	16	28272.89	1.0690e+02

---

HPL\_pdgesv() start time Wed May 15 20:47:51 2024

HPL\_pdgesv() end time Thu May 16 04:39:04 2024

---

$\|Ax-b\|_{\infty}/(\text{eps} * (\|A\|_{\infty} * \|x\|_{\infty} + \|b\|_{\infty}) * N) = 1.72116637e-03 \dots \text{PASSED}$

---

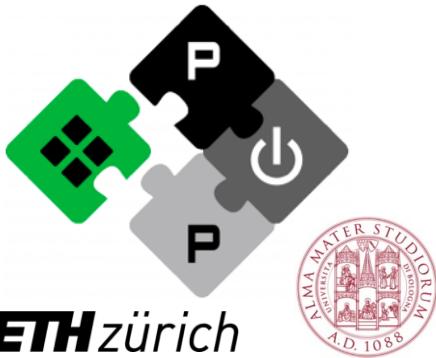
Finished 1 tests with the following results:

1 tests completed and passed residual checks

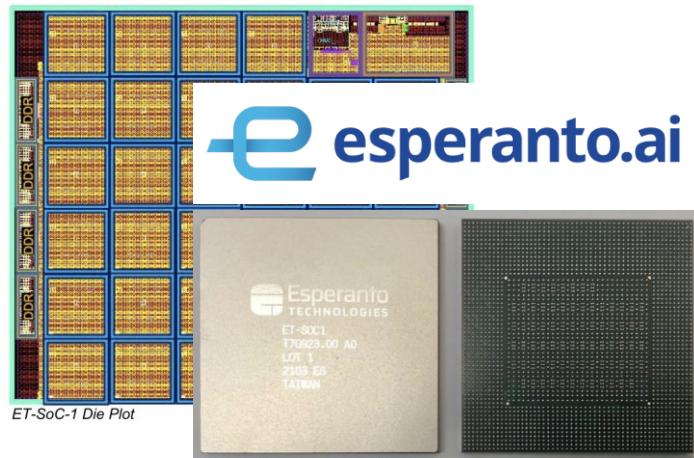
# Looking Forward - Systems

Build a **Petascale** class RISC-V Supercomputer  
Explore **RISC-V accelerated HPC platforms in production.**

**Goal:** Currently working with **several RV accelerator provider**. Among those:



**PULP Platform energy efficient accelerators<sup>[1]</sup>**  
[STX, Occamy, ...]



**Esperanto Technologies ET-SoC-1<sup>[2]</sup>**



**Axelera Metis AIPU**

[1] <https://pulp-platform.org>

[2] Accelerating ML Recommendation With Over 1,000 RISC-V/Tensor Processors on Esperanto's ET-SoC-1 Chip, David R. Ditzel, the Esperanto team, DOI: 10.1109/MM.2022.3140674

[3] Co-Design of the Kalray Manycore Accelerator for Edge Computing, Benoît Dupont de Dinechin, HiPEAC CSW Autumn 2021

# Ongoing RISC-V Projects

## NEXT STEPS IN THE RISC-V WORLD

E4 is member of the **TRISTAN & ISOLDE** (2023-2025) consortia to develop a European RISC-V Framework for the Space and Industrial Use Case

<https://tristan-project.eu/>

<https://www.isolde-project.eu/>



E4 has won a grant from the Italian state to develop the **FUTURE** project. The project aims to build a cluster of 9 RISC-V (State of the art) servers with accelerators also based on RISC-V.

# E4 Monte Cimone Team

Cosimo Gianfreda,  
Elisabetta Boella,  
Francesco Beneventi,  
Mattia Paladino,  
Marco Cicala,  
Daniele Gregori

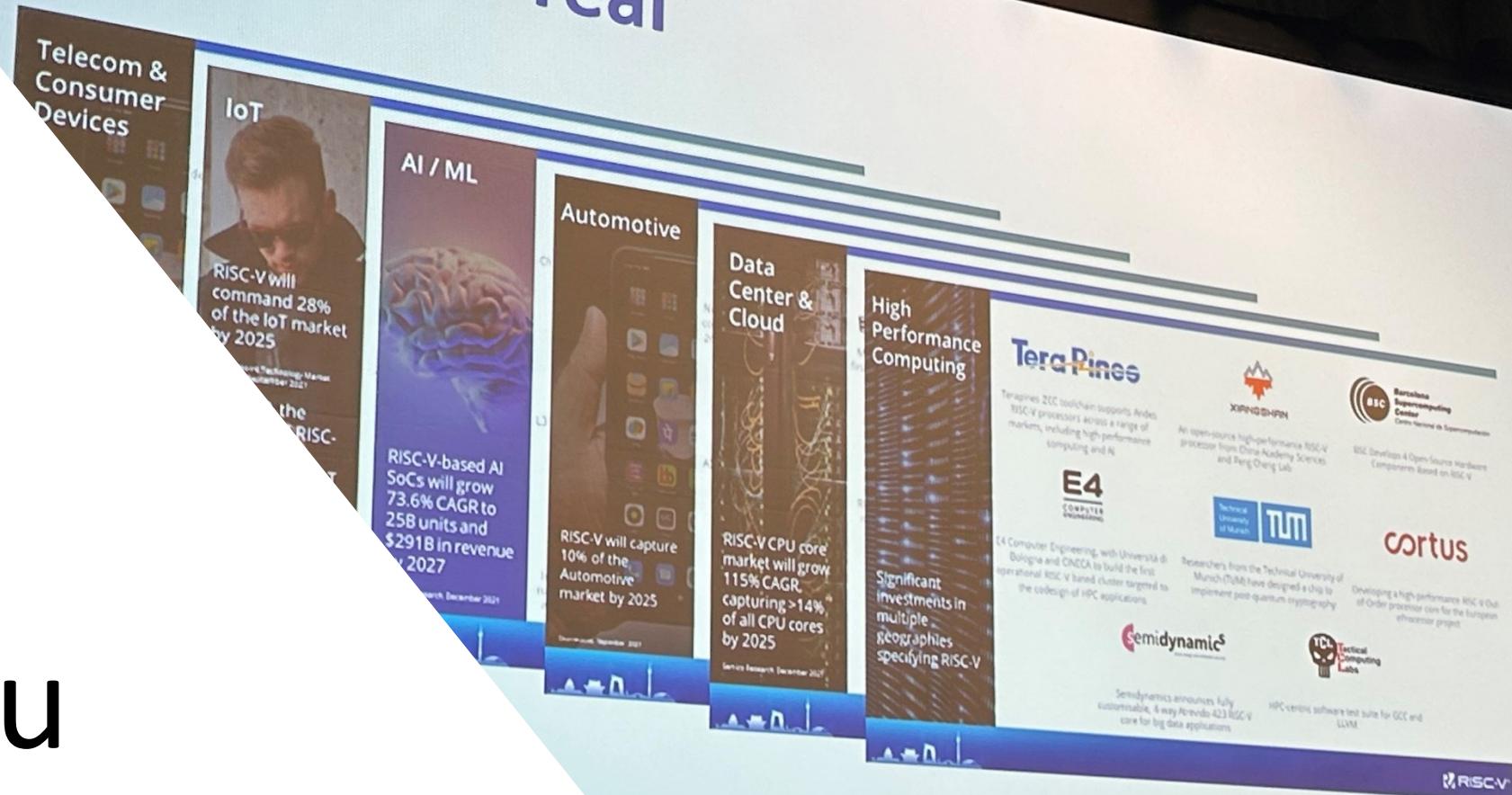


# E4

COMPUTER  
ENGINEERING

# Thank you

## The magic is real



Picture from RISC-V Summit 2023 Santa Clara CA

## CONTACTS

Email contacts

[info@e4company.com](mailto:info@e4company.com)

[support@e4company.com](mailto:support@e4company.com)

[sales@e4company.com](mailto:sales@e4company.com)

E4 Computer Engineering SpA

Via Martiri della Libertà, 66 . 42019 Scandiano (RE) - Italy

Tel. +39 0522 991811

