The Future HPC will be Open

Designing and Building Supercomputers @ BSC

Russian Supercomputing Days

John D. Davis, PhD,
Prof. Mateo Valero

September 22th, 2020
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

BSC-CNS objectives

- Supercomputing services to Spanish and EU researchers
- R&D in Computer, Life, Earth and Engineering Sciences
- PhD programme, technology transfer, public engagement

BSC-CNS is a consortium that includes:

- Spanish Government 60%
- Catalan Government 30%
- Univ. Politècnica de Catalunya (UPC) 10%
People evolution

BSC Staff evolution 2005 - 2020

August 31, 2020
<table>
<thead>
<tr>
<th>MareNostrum</th>
<th>Year</th>
<th>Peak Performance</th>
<th>Rank</th>
<th>World Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>MareNostrum 1</td>
<td>2004</td>
<td>42.3 Tflops</td>
<td>1st</td>
<td>4th</td>
</tr>
<tr>
<td>MareNostrum 2</td>
<td>2006</td>
<td>94.2 Tflops</td>
<td>1st</td>
<td>5th</td>
</tr>
<tr>
<td>MareNostrum 3</td>
<td>2012</td>
<td>1.1 Pflops</td>
<td>12th</td>
<td>36th</td>
</tr>
<tr>
<td>MareNostrum 4</td>
<td>2017</td>
<td>11.1 Pflops</td>
<td>2nd</td>
<td>13th</td>
</tr>
</tbody>
</table>
Today’s technology trends

Massive penetration of Open Source Software
• IoT (Arduino),
• Mobile (Android),
• Enterprise (Linux),
• HPC (Linux, OpenMP, etc.)

Moore’s Law + Power = Specialization (HW/SW Co-Design)
• More cost effective
• More performant
• Less Power

New Open Source Hardware
Momentum from IoT and the Edge to HPC
• RISC-V
• OpenPOWER
• MIPS
Linux History

- 1991: Started development, release
- 1992: X Windows released
- 1998: Adopted by many major companies
- 2004: BSC Supercomputer OS
- 2009: Basis for many new business systems and the cloud
- 2013: Android in 75% of the world smart phones
- 2015: De facto OS for IoT, mobile, cloud, and supercomputers
HPC Today

- Europe has led the way in defining a common open HPC software ecosystem
- Linux is the de facto standard OS despite proprietary alternatives
- Software landscape from Cloud to IoT already enjoys the benefit of open source
- Open source provides:
  - A common platform, specification and interface
  - Accelerates building new functionality by leveraging existing components
  - Lowers the entry barrier for others to contribute new components
  - Crowd-sources solutions for small and larger problems
- What about Hardware and in particular, the CPU?
Mont-Blanc HPC Stack for ARM

Industrial applications

Applications

System software

Hardware
The Exascale Race – The Japanese example

Co-design from Apps to Architecture

- Architectural Parameters to be determined
  - #SIMD, SIMD length, #core, NUMA node, O3 resources, specialized hardware
  - cache (size and bandwidth), memory technologies
  - Chip die-size, power consumption
  - Interconnect
- We have selected a set of target applications
- Performance estimation tool
  - Performance projection using Fujitsu FX100 execution profile to a set of arch. parameters.
- Co-design Methodology (at early design phase)
  1. Setting set of system parameters
  2. Tuning target applications under the system parameters
  3. Evaluating execution time using prediction tools
  4. Identifying hardware bottlenecks and changing the set of system parameters

Target applications representatives of almost all our applications in terms of computational methods and communication patterns in order to design architectural features.

<table>
<thead>
<tr>
<th>Program</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENESIS</td>
<td>M84+ problems</td>
</tr>
<tr>
<td>Genome</td>
<td>Genomic processing (Genome alignment)</td>
</tr>
<tr>
<td>GAMESA</td>
<td>Earthquake simulator (RFMin understate &amp; 8x8-cracked grid)</td>
</tr>
<tr>
<td>NUCAM-LSTK</td>
<td>Weather prediction system using Big-data (unstructured grid &amp; ensemble ensemble climate)</td>
</tr>
<tr>
<td>NCTChem</td>
<td>Molecular electronics (structured 3d-calculator)</td>
</tr>
<tr>
<td>FFB</td>
<td>Large body simulation (unstructured grid)</td>
</tr>
<tr>
<td>HSDFT</td>
<td>an ab-initio program (almost functional theory)</td>
</tr>
<tr>
<td>Advent</td>
<td>Computational Materials System by Large-scale Analysis and Design (unstructured grid)</td>
</tr>
<tr>
<td>CCS-QCD</td>
<td>Lattice QCD simulation (structured grid)</td>
</tr>
</tbody>
</table>
The Exascale Race – The Japanese example

“Post-K” Arm64fx Processor is...

- an Many-Core ARM CPU...
  - 48 compute cores + 2 or 4 assistant (OS) cores
  - Brand new core design by Fujitsu
  - Near Xeon-Class Integer performance core
  - ARM V8.2 --- 64bit ARM ecosystem

- but also a GPU-like processor
  - SVE 512 bit vector extensions (ARM & Fujitsu)
    - Integer (1, 2, 4, 8 bytes) + Float (16, 32, 64 bytes)
    - Cache + access localization (sector cache) -- similar to scratchpad
  - HBM2 OPM – Massive Mem BW (1TByte/s, Bytes/DPF ~0.4 same as K)
    - Streaming memory access, strided access, scatter/gather etc.
  - Intra-chip barrier synch. and other memory enhancing features
  - 40GByte/s Tofu-.D interconnect + PCIe 3

- GPU-like High performance in HPC, AI/Big Data, Auto Driving…
The European Open System Stack

Applications
Federation/Cloud Services

Tools - debugging, performance tuning

Mathematical, Data Analytics and AI Libraries

Compilers, Programming Environments, Communication Middleware

Operating System, Schedulers, Management Software, Cyber Security

System-Level Composability / Modularity

Chiplet, Board and System Integration / Cooling

Memory
Storage
Interconnects

CPUs
Accelerators
Neuromorphic
Quantum

TRL 8-9
TRL 5-7
TRL 3-4
TRL 1-2
Why Europe needs its own processor

- Processors now control almost every aspect of our lives

- **Security** (back doors, etc.)

- Possible **future restrictions on exports to EU** due to increasing protectionism

- A competitive **EU supply chain** for HPC technologies will create jobs and growth in Europe
RISC-V History

• 2010: Started development and initial proposal
• 2015: RISC-V Foundation formed
• 2019: Adopted by many major companies
  • Starting in the embedded market with already over 1 Billion CPUs

• 2020
  • RISC-V Foundation moves to Switzerland

The time is now to embrace and support RISC-V from IoT to HPC
RISC-V is democratizing chip-design

More and more global IT actors are adopting RISC-V architectures to be vendor independent

- Google
- Amazon
- Western Digital
- Alibaba

And of course the entire IoT ecosystem for lower performance, lower energy applications.

Major opportunity for ICT industry also in Spain
Europe can lead the way to a completely open SW/HW stack for the world

RISC-V provides the open source hardware alternative to dominating proprietary non-EU solutions

Europe can achieve complete technology independence with these foundational building blocks

Currently at the same early stage in HW as we were with SW when Linux was adopted many years ago

RISC-V can unify, focus, and build a new microelectronics industry in Europe.
“Open Source has become mainstream across all sectors of the software industry during the past 10 years. To a large extent, open software re-use has proven economically efficient. The level of maturity of Open Source Hardware (OSH) remains far lower than that of Open Source Software (OSS). However, business ecosystems for OSH are developing fast so that OSH could constitute a cornerstone of the future Internet of Things (IoT) and the future of computing.”

- DG Connect & DG IT Workshop, Brussels, Nov. 14-15, 2019
From IoT, Edge Computing, Clouds to Supercomputers
Current BSC RISC-V European Accelerator & CPU Landscape

RISC-V Accelerators

RISC-V CPUs

MEEP

EPI - SGA1

European Open Source Ecosystem

Open Source SW

Open Source HW

EPI - SGA2

The European PILOT (RISC-V)

Note: Collaboration between all RISC-V projects possible

Exploitation

eProcessor

Arm GPP EuroHPC-2020-01 subtopic A

2020 2021 2022 2023 2024 2025 2026
Rebuilding the European CPU Industry

Closed + Open
Research

MEEP
IP/Infra

eProcessor
Design
IP/Infra

EPI RISC-V PILOT
IP/Infra

Exascale
Production
The European Processor Initiative

• In the same way BSC led the development of ARM processors for HPC in the various MontBlanc projects, now it leads the RISC-V HPC accelerator development in EPI

• EPI is a 100% funded EuroHPC project (120 M€) to develop European processor technology by 2022

• BSC was the original initiator of EPI and most active proponent in the scientific and technical community

• EPI is led by Atos/Bull with 28 partners from leading HPC industrial and academic centres
EPI Partners
EPAC architecture

- **Objective:** Develop & demonstrate fully European processor IPs based on the RISC-V ISA. Build on existing EU IP, leverage EU background and vision

- Provide a very **low power** and high computing throughput **accelerator for HPC & Emerging -> Automotive**

**Main Contributors:**

- **RISC-V in the Tile:** 8 Vector cores, 8 STX, 1 VRP

- **Vector Core:**
  - Spain + Italy + Croatia
  - C: RISC-V Scalar core
  - V: RISC-V Vector core: 8 lanes

- **STX**
  - Switzerland + Germany
  - RISC-V NTX (AI/ML/DL) + Stencil

- **VRP**
  - France
  - RISCV + Extended precision FPU
Rebuilding the European CPU Industry

Closed + Open

Research

Open

Production

Exascale

MEEP

EPI

IP/Infra

Design

eProcessor

EPI RISC-V PILOT

IP/Infra

IP/Infra

IP/Infra

IP/Infra

BSC

Supercomputing Center
Centro Nacional de Supercomputación
200 Petaflops peak performance (200 x 10^{15})

Experimental platform to create supercomputing technologies “made in Europe”

217 M€ of investment

Hosting Consortium:
- Spain
- Portugal
- Turkey
- Croatia

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC Joint Undertaking, through the European Union’s Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States Spain, Portugal, Croatia, and Turkey.
MEEP: MareNostrum Experiment Exascale Platform

MEEP is a flexible FPGA-based emulation platform that will explore hardware/software co-designs for Exascale Supercomputers and other hardware targets, based on European-developed IP. The project provides two functions:

- An evaluation platform of pre-silicon IP and ideas, at speed and scale.

- A software development and experimentation platform to enable software readiness for new hardware.

MEEP will enable software development, accelerating software maturity, compared to the limitations of software simulation. IP can be tested and validated before moving to silicon, saving time and money.
Accelerator Compute and Memory Engine (ACME) Architecture
Mapping ACME Architecture to an FPGA

IoT to HPC, AI/ML/DL
Rebuilding the European CPU Industry

Closed + Open
Research

Open
Production

Exascale

EPI RISC-V PILOT

IP/Infra

IPA

eProcessor

MEEP

IP/Infra

EPI

IP/Infra

BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación
eProcessor Stack

HPC
Applications
Middleware

AI
Applications
Middleware

Bioinformatics
Applications
Middleware

Tools (compiler, performance monitoring, debugging)
Runtime (OpenMP, Tensorflow, Apache Spark, etc.)
Linux

64,32,16-bit mixed precision

8, 4, 3, 2, 1-bit mixed precision

2-way OOO Multicore + Adaptive caches and scratch pad + Bioinformatics Accelerator + Low Power

Fault Tolerance
eProcessor System diagram

HPC/HPDA SW

Sys SW

interconnect

LEGaTO PCB

CPU0
Core
Core
Accel

CPU1
Core
Core
Accel

Accel
Accel
Accel

Coherent off-chip interface
Rebuilding the European CPU Industry

Closed + Open
Research

Open
Design

Exascale
IP/Infra

MEEP
IP/Infra

EPI
IP/Infra

eProcessor
IP/Infra

EPI RISC-V PILOT
IP/Infra

Production
The European PILOT

Stream 2: PILOT SW Stack

- HPC Vector
- ML/AI/Stencil

Applications:
- GROMACS
- EC-EARTH
- MD + AI
- Video
- BERT

SW Libraries:
- FFT
- BLIS
- oneDNN
- Neural Net (MLS-DNN)

AI Framework:
- Tarantella
- Tensorflow
- Dace

Runtime & Schedulers:
- MPI
- OpenMP
- DLB
- TAMPI

System SW:
- BeeGFS
- SLURM
- DROM
- Boot, Drivers, Linux
- BBQUE

Toolchain:
- Inference Engine
- LLVM

Stream 3: European RISC-V Accelerators

Accelerator Chassis

Accelerator PCB Cards

Accelerator Chips

Stream 4: European PILOT

Arm (Rhea, etc.) and/or x86 host

Immersive Cooling
The European PILOT Accelerator Architecture

- **Accelerator Card**
  - 3 C2C interfaces per chiplet, 8 lanes each
  - 1 PCIe 6.0 w/ CXL 3.0 interface per chiplet, 4 lanes
  - 4x16 LPDDR controllers per chiplet
  - Homogeneous or Heterogeneous PCB
  - Optional FPGA/PCIe Switch

---

**Accelerator Chiplet**

<table>
<thead>
<tr>
<th>Compute Tile</th>
<th>Compute Tile</th>
<th>Compute Tile</th>
<th>Compute Tile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
</tr>
<tr>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
</tr>
<tr>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
</tr>
<tr>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
<td>Compute Tile</td>
</tr>
</tbody>
</table>

**Common Uncore**

- AXI
- IOTile
- PCIe w/CXL
- Serdes Phy 32Gbps
- C2C Bridge
- Serdes Phy 32Gbps
- C2C Bridge
- Serdes Phy 32Gbps
- C2C Bridge
- Serdes Phy 32Gbps
- LPDDR Ctrl
- X16 Phy
- LPDDR Ctrl
- X16 Phy
- LPDDR Ctrl
- X16 Phy
- LPDDR Ctrl
- X16 Phy

**PCB**

- FPGA or PCIe Switch
- LPDDR4 x64
- 33GB/s
- LPDDR4 x64
- 33GB/s
- C0
- C2
- 33GB/s
- C1
- C3
- 33GB/s
- LPDDR4 x64
- 33GB/s

---

**Barcelona Supercomputing Center**
Centro Nacional de Supercomputación
Rebuilding the European CPU Industry

Closed + Open
Research

Open

Exascale

EPI

MEEP

eProcessor

EPI RISC-V PILOT

IP/Infra

IP/Infra

IP/Infra

IP/Infra
The Future: Flagship RISC-V Exascale Accelerator

RISC-V Accelerators

RISC-V CPUs

MEEP
EPI - SGA1

Note: Collaboration between all RISC-V projects possible

Exploitation

European Open Source Ecosystem

Open Source SW

Open Source HW

EPI - SGA2

The European PILOT (RISC-V)

Flagship Accelerator Program (RISC-V)

Flagship CPU Program (RISC-V)

eProcessor

Arm GPP EuroHPC-2020-01 subtopic A
The Future is Wide Open!

➡️ There is an urgent need, from mobile phones to supercomputers: more compute at lower power

➡️ The RISC-V ecosystem is in the nascent period where it can become the de facto open hardware platform of the future

➡️ An opportunity for Europe to lead the charge to creating a full stack solution for everything, from supercomputers down to IoT devices

➡️ Our main goal: Create European chips that meet the needs of future European and global markets across HPC, cloud, automotive, mobile to IoT

➡️ This is the framework for the Exascale Supercomputing Initiative at BSC
BSC full stack

- HPC Applications
- Specialization using HW/SW Co-Design
- HPC Hardware
LOCA Goals

• Mechanism to extend open source ecosystem to include H/W
  • Add H/W expertise to BSC and European partners, leverage existing S/W expertise
  • Provide proven/usable Open Source H/W
  • Intersection of academia and industry
  • Open European IP repository → rapid implementation
  • Catalyst to reinvigorate European ICT industry
  • Global collaboration and training center
  • Incubator for European IP
Traditional chip design is done in a Master/Apprentice environment

LOCA recreates this environment by bringing in Masters from industry to collaborate with a variety of people, pushing beyond RTL

Professors, students, and industry veterans all together

Ideal sandbox for creative and innovative work

Research and Design to chip fabrication
LOCA is Bigger than Barcelona!

We must stand on the shoulders of giants to build great things.

We are assembling many giants and hope you can join us.
Thank you

mateo.valero@bsc.es