montblanc-project.eu | @MontBlanc_EU

Paving the way to a scalable, modular and power-efficient European HPC processor

Arm HPC User's Group @ISC'19

Said Derradji (Atos)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement n° 671697 & 779877

Mont-Blanc roadmap

Vision: leverage the fast growing market of mobile technology for scientific computation

2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020

European Processor Initiative

- **Focus on SoC design**
- **Develop critical IPs**

Prepare an industrial solution
Test market acceptance

Mont-Blanc 2020

Extend the concept and explore new possibilities

Mont-Blanc 3

Proof of concept : HPC computing based on mobile embedded technology

Mont-Blanc 2

Mont-Blanc

MONT-BLANC

Mont-Blanc 2020 partners

Atos

arm

BSC

DE LA RECHERCHE À L'INDUSTRIE
cea

KALRAY

JÜLICH
FORSCHUNGSZENTRUM

semidynamics
silicon design and verification services

MONT-BLANC

MB2020 mission

- **Trigger the development of the next generation of industrial processor for Big Data and High Performance Computing**

- **Mont-Blanc DNA**
  - economic sustainability
  - power efficiency
  - heterogeneity
  - strong codesign approach for real applications
  - building a prototype

MONT-BLANC

Mont-Blanc prototypes

2015 | 2016 | 2017 | 2018 | 2019 | 2020

Mont-Blanc prototype

1080 Dual Cortex-A15 (+ Mali)

- **2160 Armv7 cores**

Dibona

48 nodes bi-socket THX2

- **3072 Armv8 cores**

MB2020 Demonstrator

Only one MB2020 SoC

- **32 SVE cores**

MONT-BLANC

Mont-Blanc Dibona open to more users

- **The Mont-Blanc Dibona prototype:**
  - 48 nodes bi-socket THX2 @2GHz (3072 cores)
  - 12.3 TB RAM
  - featuring Mellanox EDR 100Gb/s
  - Theoretical peak performance of > 49 Tflops

- **Usage:**
  - +70 users
  - Allocations with SLURM (including bash): 15k
  - Application executions with SLURM: 12.5k
  - Multi-node jobs with SLURM: ~5.7k

- **Events:**
  - Hackathon by Arm @SC18 +40 users
  - Hackathon by BSC Feb. 21st +30 students

MONT-BLANC

Dibona SW stack today

- Atos ARM Lab still updating the SW stack providing more OpenMPI flavors
- ARM provides new Compiler, Performance Libraries
- HW support :Atos fine grain power monitoring, DVFS
- RedHat: More recent and optimized kernel
- Mellanox: Improved drivers

- Preparing the porting/developing of Atos BXI drivers on Arm

MONT-BLANC

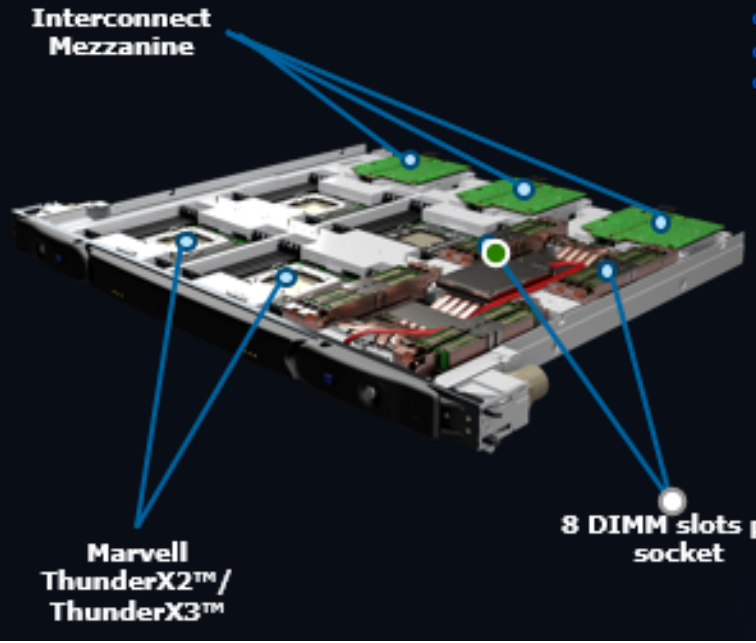
Introducing BullSequana XH2000

MONT-BLANC

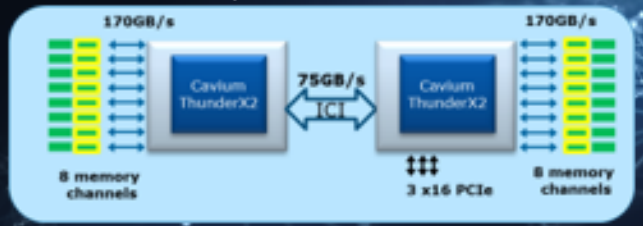
The Arm blade

BullSequana XH2000 Blades

BullSequana X1310: Marvell ThunderX2™/ThunderX3™ (ARMv8) processor



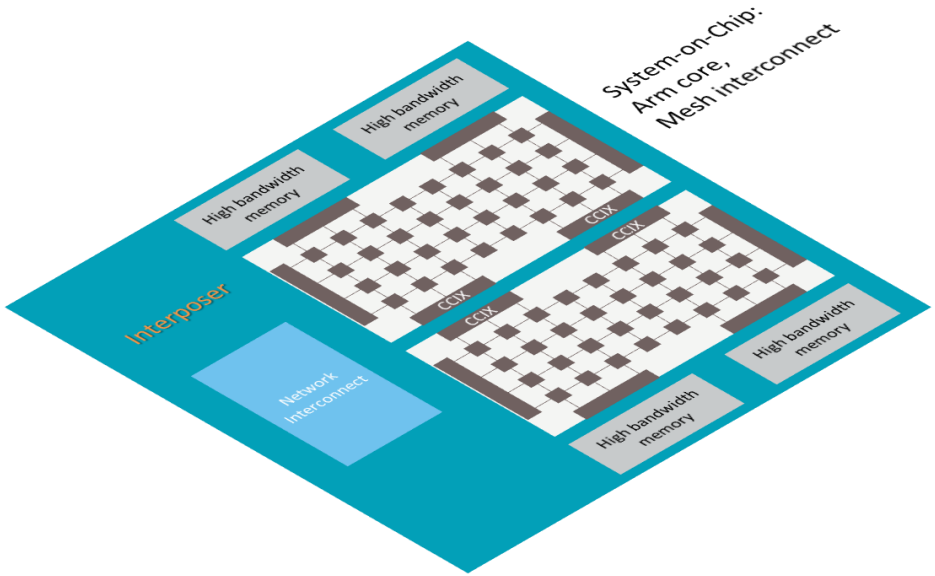
- 1U form factor
- Direct liquid cooling
- 3 compute nodes per blade with:
  - 2 Marvell ThunderX2/ThunderX3 processors with
    - Up to 32 custom ARMv8 cores per CPU, 4 threads per core, up to 2.5GHz
  - 16 DDR4 @2666MT/s DIMM slots
  - 1 Interconnect mezzanine board (High-speed Ethernet, InfiniBand EDR, HDR or HDR100)
  - 1 optional PCIe switch mezzanine
  - Disks
    - 1 x 2.5" 7 mm SATA SSD
    - Or 1x 2.5" PCIe flash SSD, connected to an embedded PCIe switch



Project Objectives

Mont-Blanc 2020 will provide new IPs, such as a new low-power "Network-on-Chip" based on the Coherent Hub Interface (CHI) architecture:

- **supporting up to 128 SVE cores;**
- **each core embedding 2 vector units from 128 to 1024 bits wide;**
- **interfacing HBM2/3 memory controllers to sustain high memory bandwidth.**



3 Key Challenges

To achieve the desired performance with the targeted power consumption, the project will need to:

understand the trade-offs between vector length, NoC bandwidth and memory bandwidth to maximize processing unit efficiency;

come up with an innovative on-die interconnect that can deliver enough bandwidth with minimum energy consumption;

opt for a high-bandwidth / low power memory solution with enough capacity and bandwidth for Exascale applications.

MONT-BLANC

MB2020 main choices

- **To use the ARM ISA**

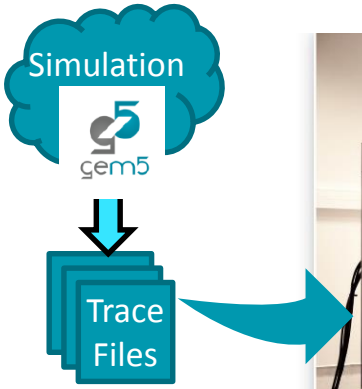
- **To design, implement and leverage new technologies to improve the performance, power efficiency, reliability and security of the processor :**
  - large SVE compute units
  - specific NoC
  - specific High Bandwidth Memory controller
  - specific SoC Power Management controller
  - innovative packaging technologies (Interposer)
  - state of the art silicon process 7nm FF

- **To improve on the economic sustainability of processor development through a modular design that allows to retarget our SoC for different markets adding specific accelerators.**

MONT-BLANC

From design to HW emulation

- **Objectives: verification of MB2020 design (interconnection network and memory hierarchy) and performance extractions**
- **Methodology: emulation of the MB2020 design with applicative workflow based on simulation traces injected into the design**
  - Trace injector are playing the role of ARM cores in the design



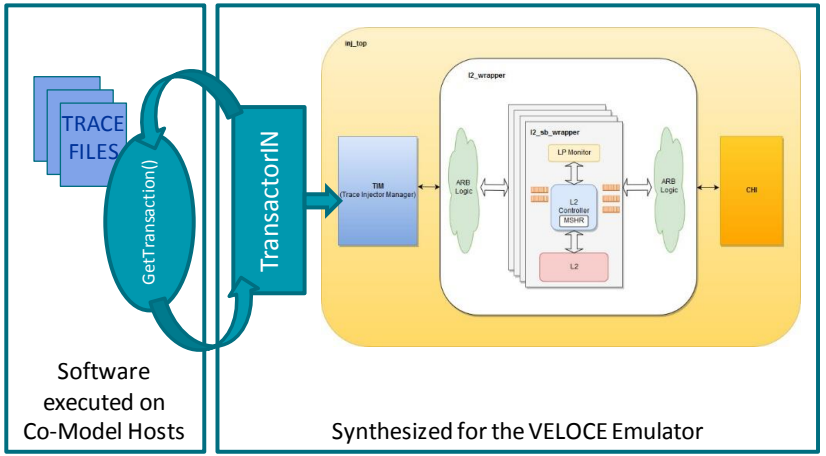
Mentor Graphics Veloce2: Powerful emulator environment hosted by CEA

- Up to 2MHz emulation speed
- 200 Mgates max capacity
- 32 Gb memory capacity

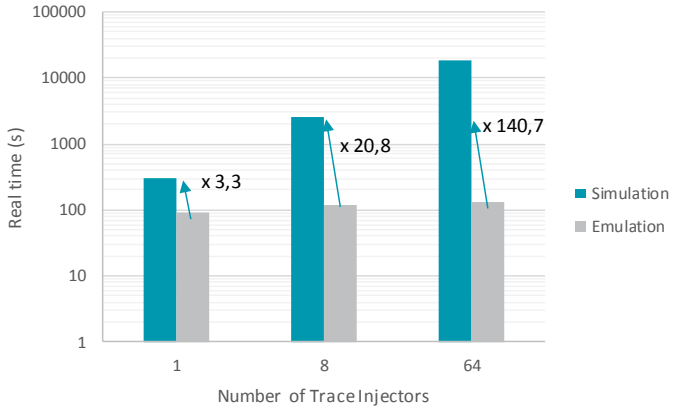
MONT-BLANC

HW emulator accelerated simulation

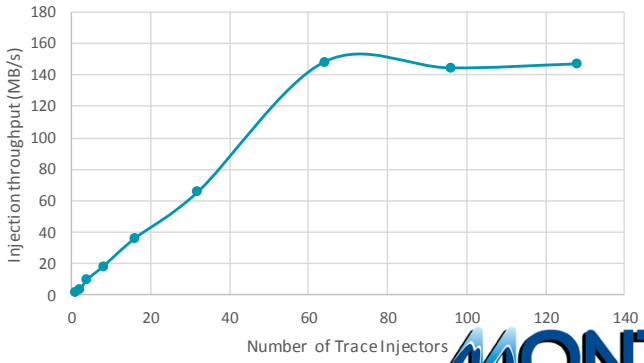
- **Validation of the trace injector design on the emulator**
  - HW/SW interface for trace file injection in the emulator
  - RTL code fully synthesizable for the HW part
  - Configurable number of Trace Injectors
- **Measurement of emulation performance compared to simulation**
- **Measurement of aggregated injection throughput during the emulation**
- **Emulation environment ready to integrate the MB2020 design**



RTL simulation time compared to emulation time for the same design



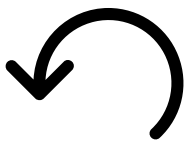
Aggregated injection throughput during the emulation

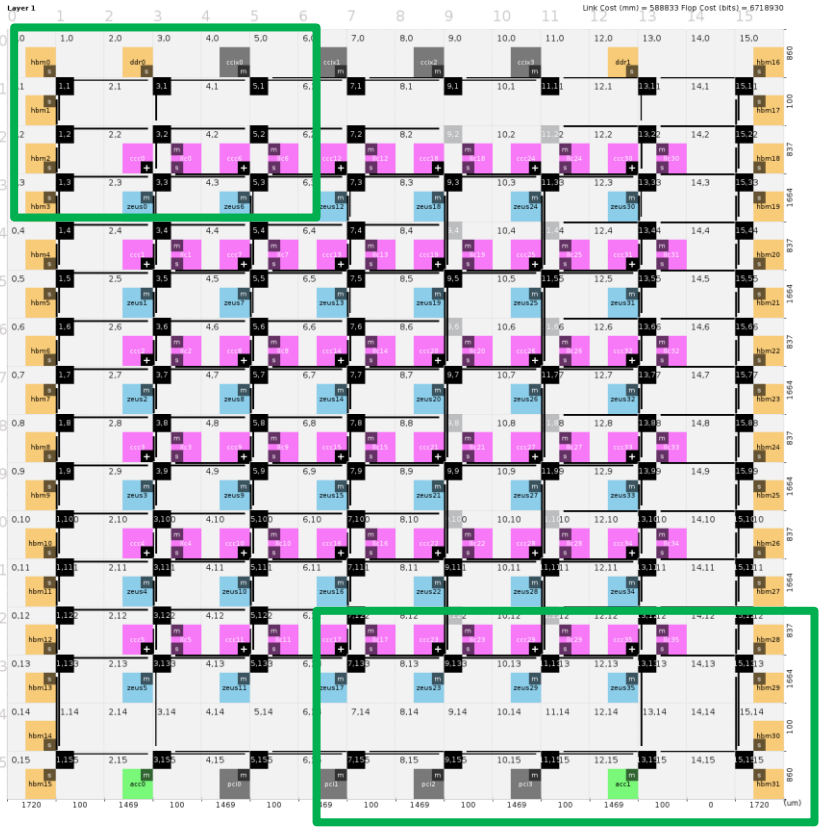


MONT-BLANC

Codesign from Apps to architecture

setting NoC parameters
evaluate performance using traces
identify HW bottleneck
then try new parameters or new features





Message class/BW: the CHI coherence protocol is mapped over several physical layers:

- Data, Response, Request, Snoop
- Each message class can be mapped to several lanes to sustain the BW required per channel

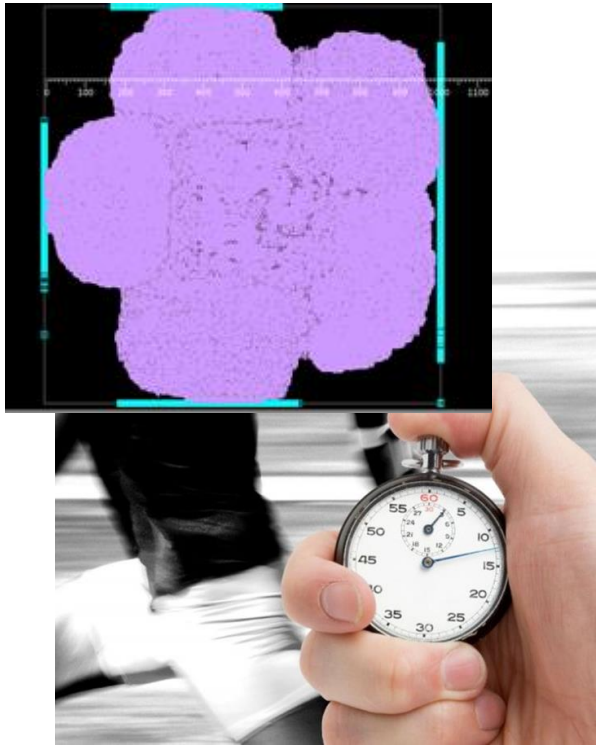
Memory mapping/affinity :

- the home agent/snoop filter must be distributed
- the interleaving function (per group per cache line) must assure well-balanced traffic spreading to memory (home agents and controllers).

MONT-BLANC

Demonstrating the performance

- Power, Performance and Area (PPA) metrics in 7nm
- Using IP RTL, back-end design team will run topographical synthesis (=taking physical info into account) using Design Compiler NXT from Synopsys
- Providing estimations for all IPs, including SVE cores



MONT-BLANC

Beyond Mont-Blanc 2020

- **Mont-Blanc 2020 is at the heart of the European exascale supercomputer effort**
- **Most of the IP developed within Mont-Blanc 2020 will be reused and productized in the European Processor Initiative (EPI)**

MONT-BLANC


MONT-BLANC 2020
More information:
montblanc-project.eu
@MontBlanc_EU