http://www.montblanc-project.eu

High Performance Computing Based on Mobile Embedded Technology

Filippo Mantovani

Barcelona Supercomputing Center

Technical Coordinator

This project and the research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement n° 288777.

Mont-Blanc project in a glance

2012 2013 2014 2015 2016

Mont-Blanc

Mont-Blanc 2

SEVENTH FRAMEWORK PROGRAMME

HORIZON 2020

To develop an **European** Exascale **approach**
leveraging **commodity** and **embedded** cost-effective technology

Mont-Blanc objectives

2012	2013	2014	2015	2016

Mont-Blanc

- HPC prototype based on current mobile embedded technology
- Learn from the experience, plan for future architecture
- Port real scientific applications

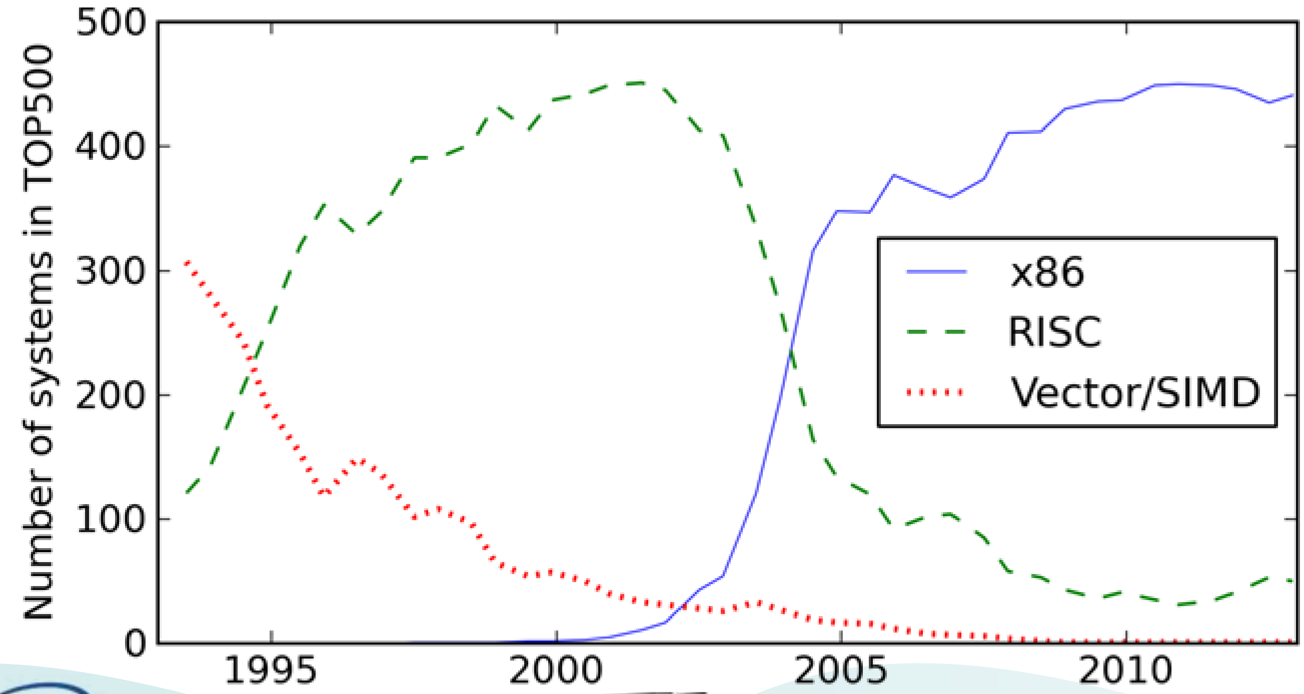
Mont-Blanc 2

Extend
- Set of scientific applications
- OmpSs programming model
- Productivity tools
- Next generation Mont-Blanc architecture

Explore
- ARM 64-bit
- Fault tolerance and resiliency
- Market of ARM-based platforms for mini-clusters

Disseminate: End-User Group

Why are we doing this?

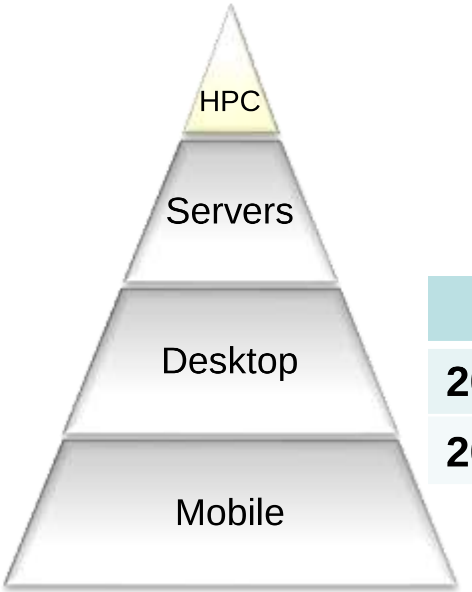


1 teraFLOPS supercomputer
ASCI Red
(Sandia – 1997)
Pentium Pro

1 petaFLOPS supercomputer
Roadrunner
(IBM / Los Alamos NL - 2008)
AMD Opteron + PowerXCell 8i

>10 petaFLOPS supercomputer
Titan
(Cray / Oak Ridge NL - 2012)
AMD Opteron + Nvidia K20

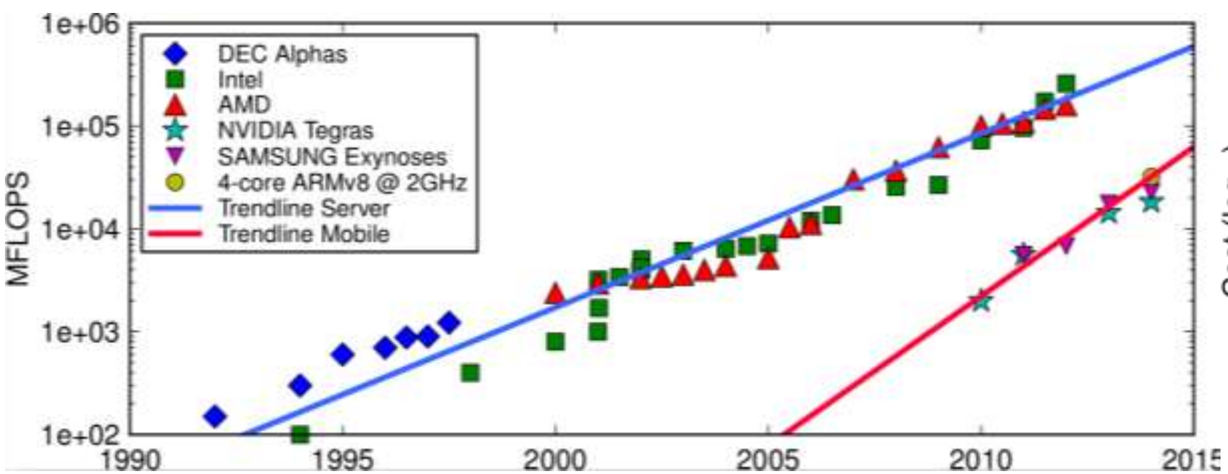
What is commodity nowadays?

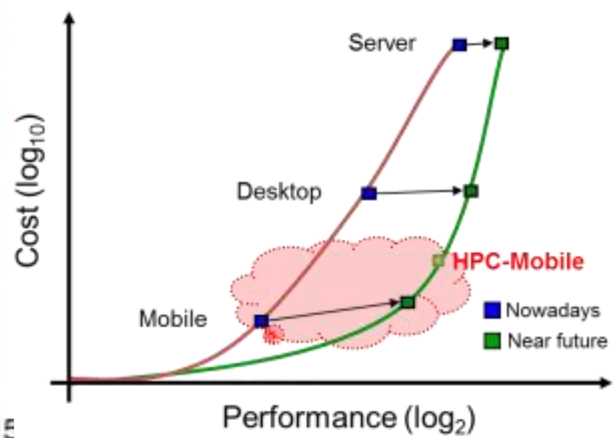


TOP 500 The List. ~23M cores (Nov 2014)

	Servers		PC		Smartphones	
2013	9.0M		316M		1000M	
2014	9.3M	+3%	314M	-1%	1300M	+30%

...and we are still ignoring tablets: >200M





Outline

1. Hardware

- History
- Mont-Blanc prototype
- Limitations

2. Software

- System software
- OmpSs programming model
- Power monitor

3. Applications

- Scalability
- Energy to solution
- End-User Group

4. Conclusions

- Project status
- Mont-Blanc impact
- Remarks

Outline

1. Hardware

- History
- Mont-Blanc prototype
- Limitations

2. Software

- System software
- OmpSs programming model
- Power monitor

3. Applications

- Scalability
- Energy to solution
- End-User Group

4. Conclusions

- Project status
- Mont-Blanc impact
- Remarks

The Mont-Blanc prototype ecosystem

Tibidabo:
ARM multicore

Carma:
ARM +
external
mobile GPU

Pedraforca:
ARM +
HPC GPU

Arndale:
ARM + embedded GPU

Odroid:
ARM bigLITTLE
In-kernel switcher

Odroid Octa:
ARM bigLITTLE
Heterogeneous
multi-processing

NVIDIA Jetson
ARM 4+1 + K1 GPU

Mont-Blanc protoype:

2011 2012 2013 2014

Prototypes are critical to accelerate software development
System software stack + applications

Mont-Blanc Server-on-Module (SoM)

CPU + GPU + Memory + Local Storage + Network

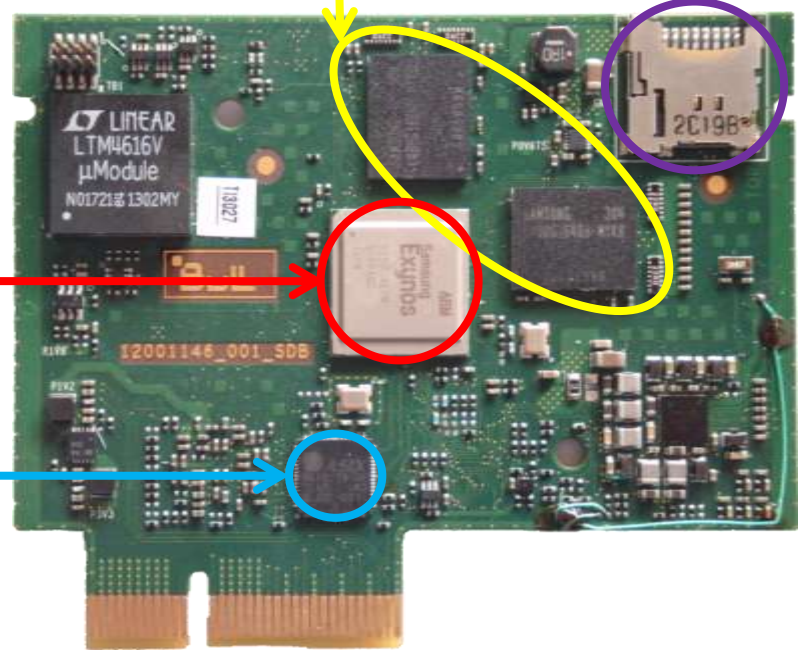
Form factor: 8.5 x 5.6 cm

Memory
4 GB LPDDR3-1600

Local Storage
microSD up to 64 GB

Exynos5 Dual SoC
2x cores ARM Cortex-A15
1x GPU ARM Mali-T604

Network
USB3.0 to 1GbE bridge



The Mont-Blanc prototype

Exynos 5 compute card
2 x Cortex-A15 @ 1.7GHz
1 x Mali T604 GPU
6.8 + 25.5 GFLOPS
15 Watts
2.1 GFLOPS/W

Carrier blade
15 x Compute cards
485 GFLOPS
1 GbE to 10 GbE
300 Watts
1.6 GFLOPS/W

Blade chassis 7U
9 x Carrier blade
135 x Compute cards
4.3 TFLOPS
2.7 kWatts
1.6 GFLOPS/W

Rack
8 BullX chassis
72 Compute blades
1080 Compute cards
2160 CPUs
1080 GPUs
4.3 TB of DRAM
17.2 TB of Flash

35 TFLOPS
24 kWatt

	Mont-Blanc [GFLOPS/W]	Green500 [GFLOPS/W]
Nov 2011	0.15	2.0
Nov 2014	1.5	5.2

Limitations of commodity mobile technology

- 32-bit memory controller
  - Even if ARM Cortex-A15 offers 40-bit address space
- No ECC protection in memory
  - But surprisingly enough this is not affecting badly scalability (so far)
- No standard server I/O interfaces
  - No native Ethernet or PCI Express
  - Provide USB 3.0 and SATA (required for tablets)
- No network protocol off-load engine
  - TCP/IP, OpenMX, USB protocol stacks run on the CPU
- Thermal package not designed for sustained full-power operation

Implementation decisions, not unsolvable problems

The only need is a business case to justify the cost
of including the new features (e.g. the HPC and server markets)

Wait for next SoC producer ⟷ Design a new HPC SoC

Outline

1. Hardware

- History
- Mont-Blanc prototype
- Limitations

2. Software

- System software
- OmpSs programming model
- Power monitor

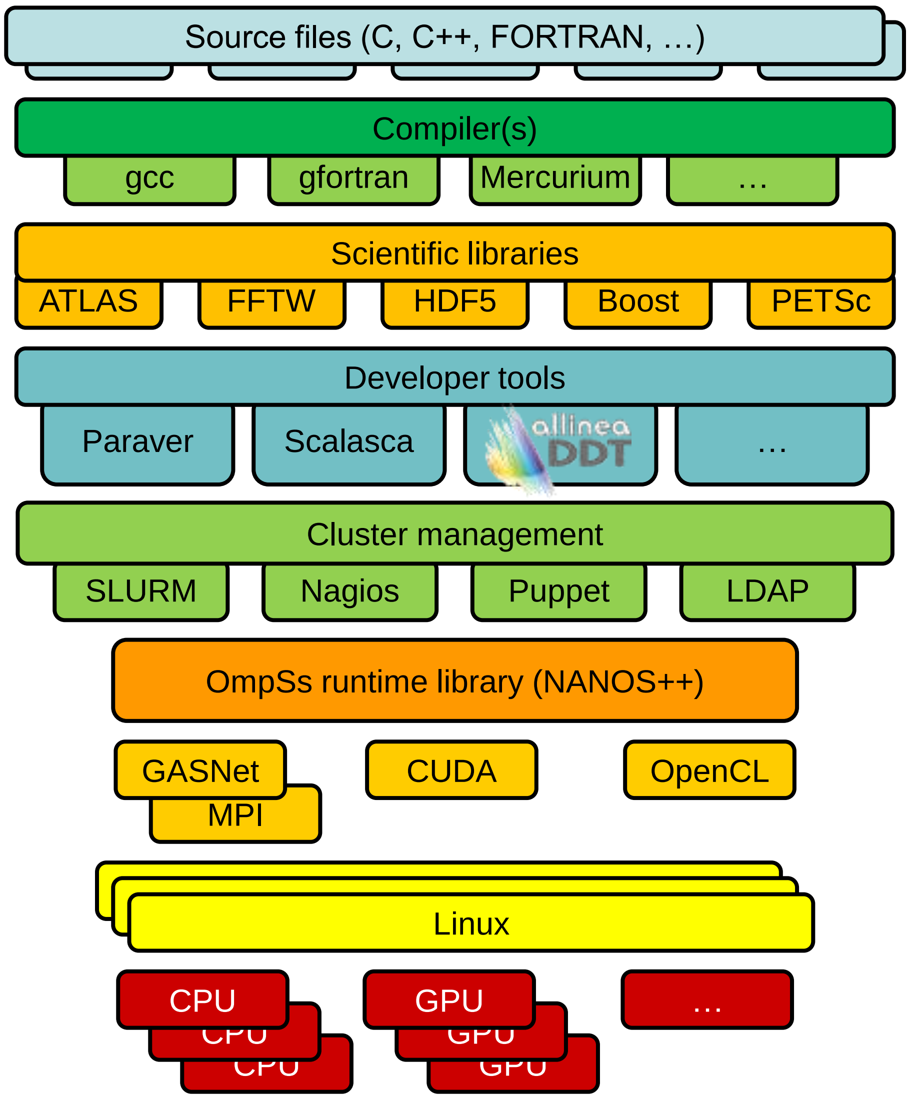
3. Applications

- Scalability
- Energy to solution
- End-User Group

4. Conclusions

- Project status
- Mont-Blanc impact
- Remarks

System software stack

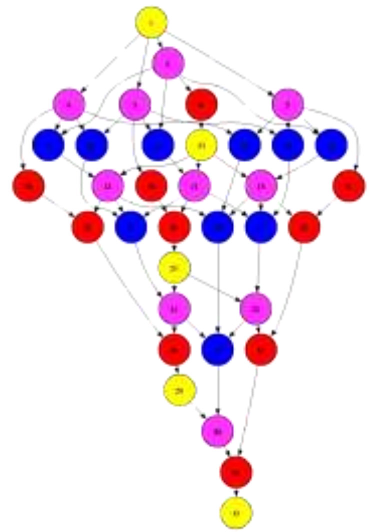


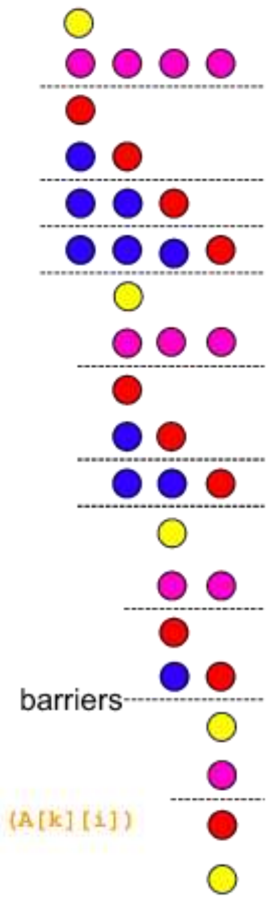
- Open source system software
  - Ubuntu/Debian Linux OS
  - GNU compilers
  - JDK
  - Scientific libraries
  - Cluster management
- Runtime libraries
  - MPICH2, CUDA, OpenCL, OmpSs
- Productivity tools
  - Perf
  - Paraver, Scalasca
  - Allinea DDT debugger

+ Server software stack
(LAMP, Hadoop, OpenStack)

OmpSs programming model

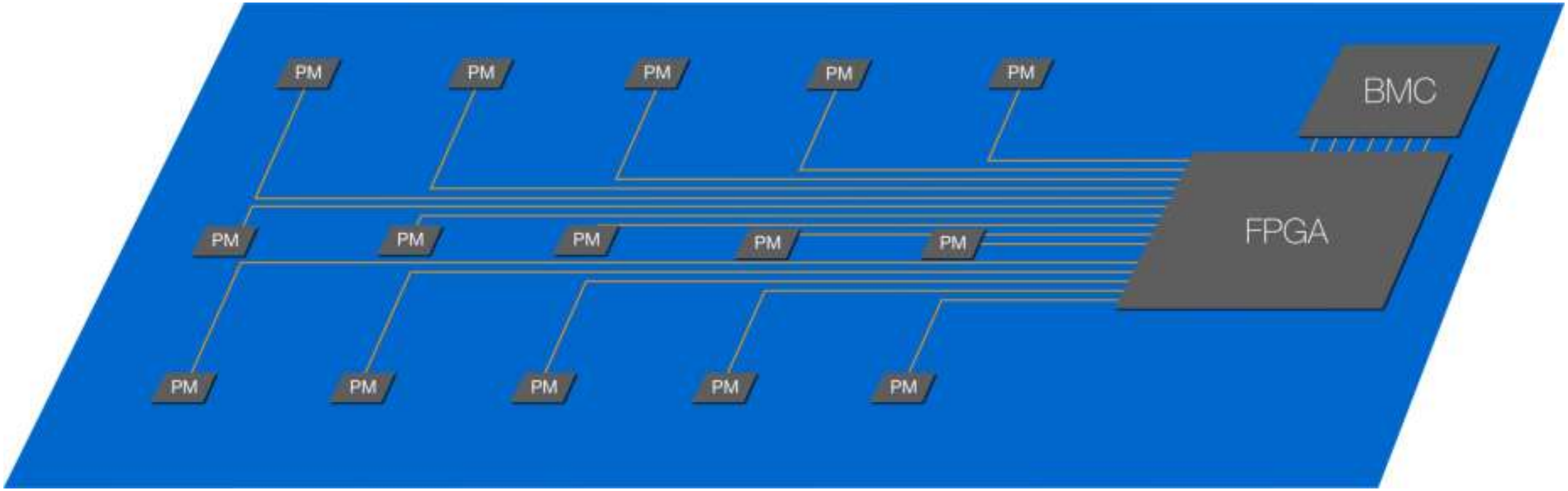
- Programmer exposed to a simple architecture
  - Tasks
  - Data dependencies
  - Target devices (heterogeneity)
- Task graph provides look ahead
  - Exploit knowledge about the future
  - Allows exploration of scheduling policies
- It helps handling limitations of the hardware
  - Heterogeneity
  - Multiple address spaces
  - Low interconnect bandwidth
  - Synchronization





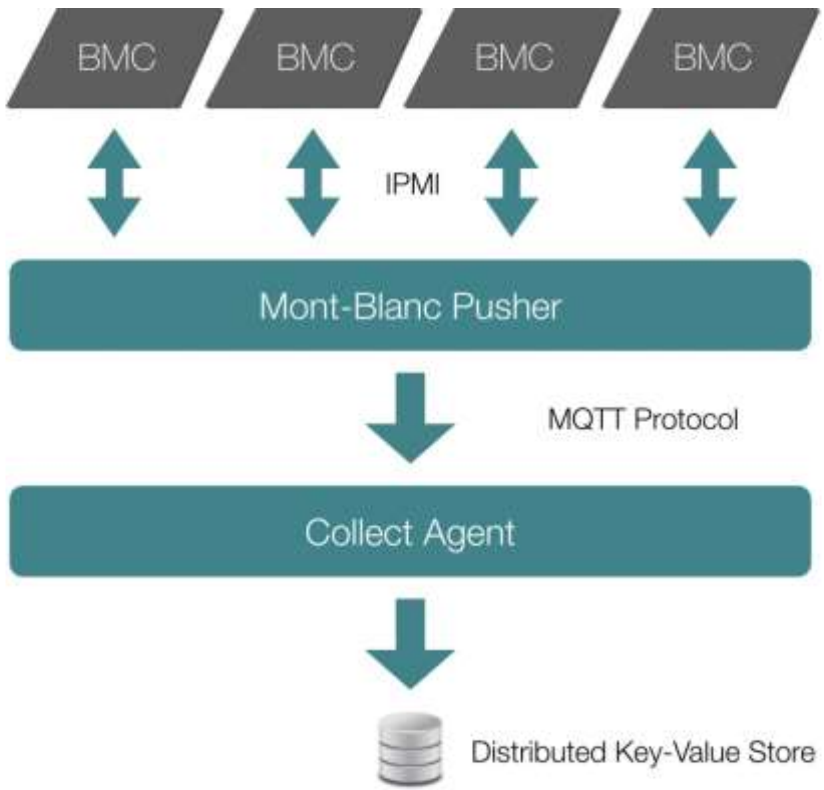

```
void Cholesky( float *A[NT][NT] ) {
int i, j, k;
for (k=0; k<NT; k++) {
   #pragma omp task inout (A[k][k])
   spotrf (A[k][k]) ;
   for (i=k+1; i<NT; i++) {
      #pragma omp task in (A[k][k]) inout (A[k][i])
      strsm (A[k][k], A[k][i]);
   }
   for (i=k+1; i<NT; i++) {
      for (j=k+1; j<i; j++) {
         #pragma omp task in (A[k][i], A[k][j]) inout (A[j][i])
         sgemm( A[k][i], A[k][j], A[j][i]);
      }
      #pragma omp task in (A[k][i]) inout (A[i][i])
      ssyrk (A[k][i], A[i][i]);
   }
}
```

Power monitor – HW infrastructure

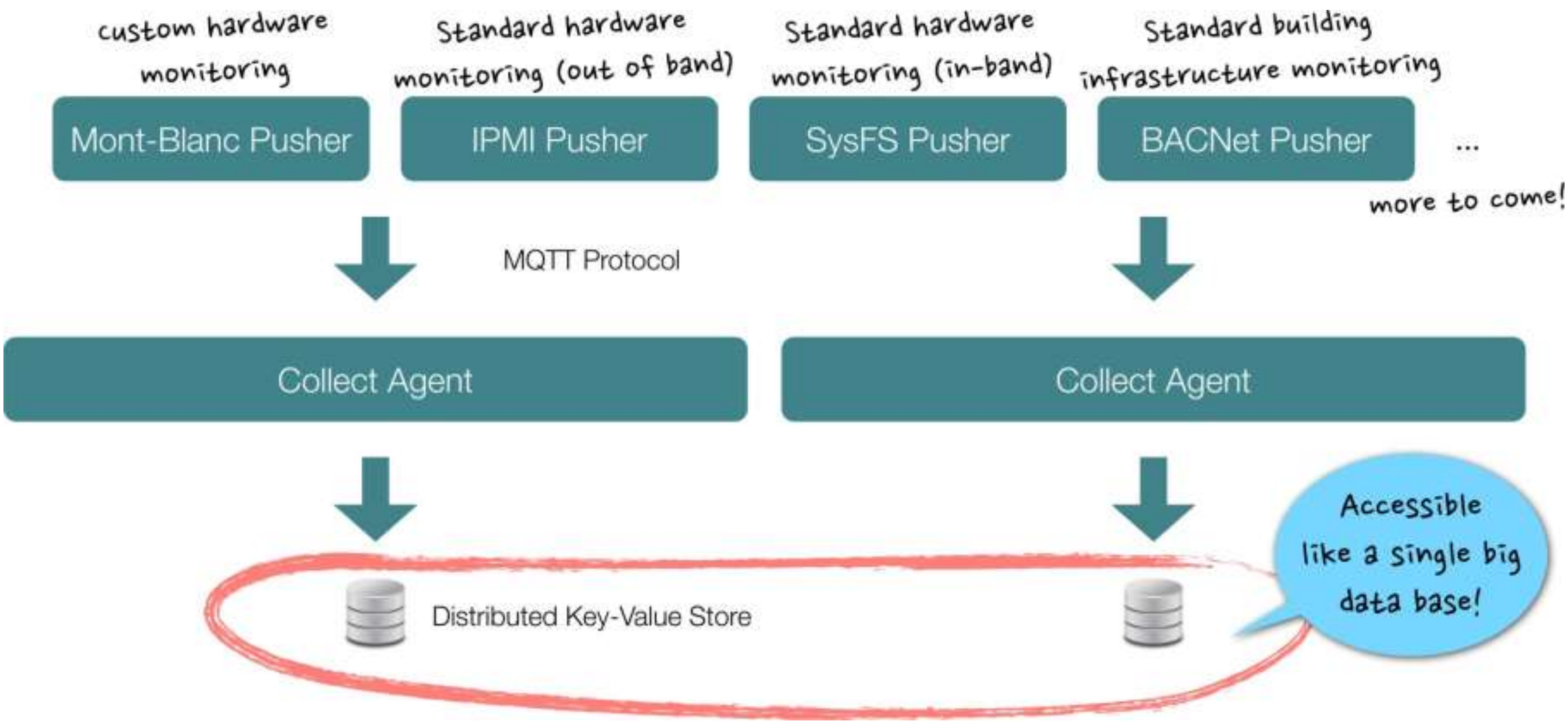


Power monitor – HW / SW interface

- Field Programmable Gate Array (FPGA)
  - Collects power consumption data from all 15 power measurement / sample interval: 70ms
- Board Management Controller (BMC)
  - Collects 1s averaged data from FPGA
  - Stores measurement samples in FIFO
- Mont-Blanc Pusher
  - Collects measurement data from multiple BMCs using custom IPMI commands
  - Forwards data using MQTT protocol through Collect Agent into key-value store



Power monitor – Block diagram



Outline

1. Hardware

- History
- Mont-Blanc prototype
- Limitations

2. Software

- System software
- OmpSs programming model
- Power monitor

3. Applications

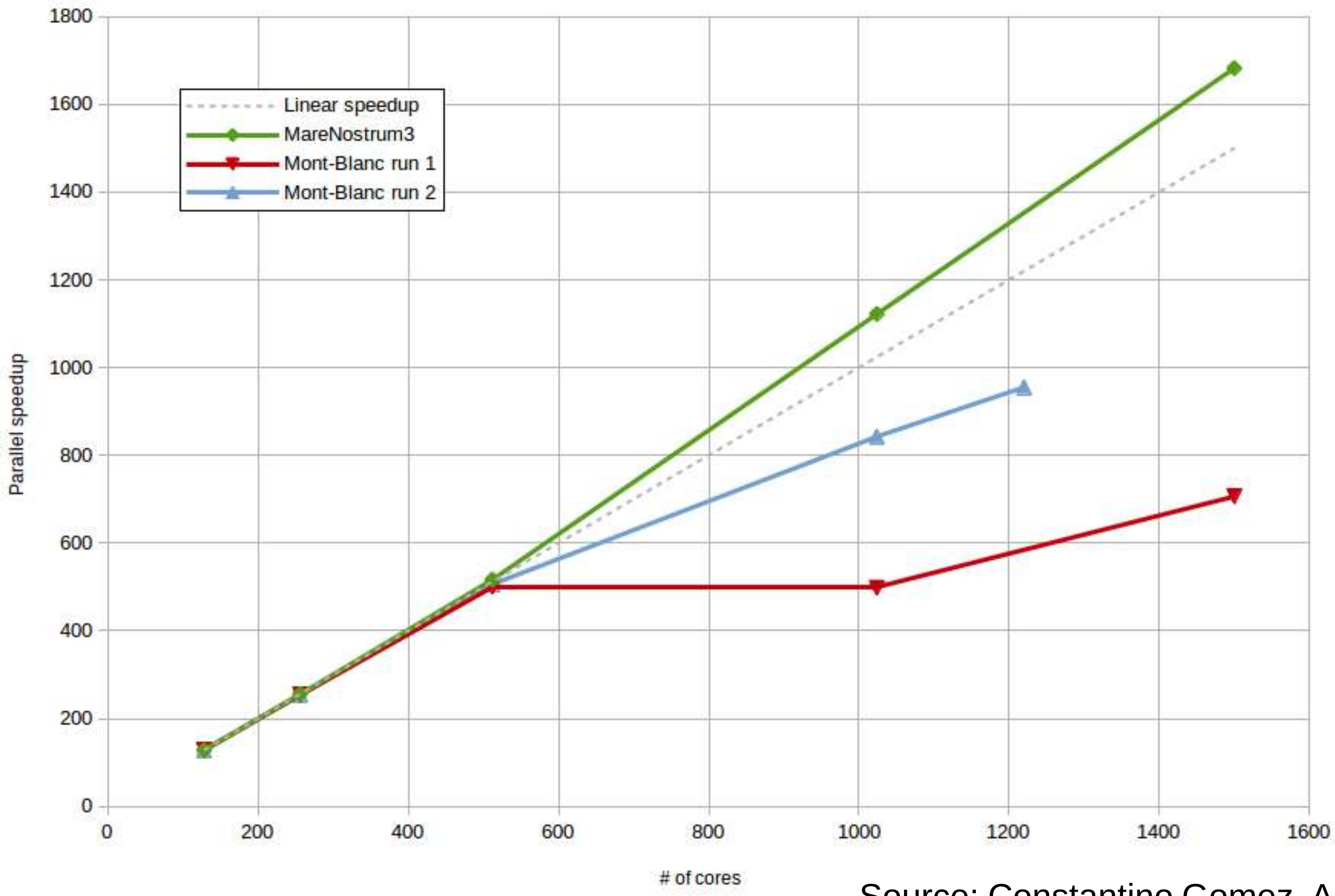
- Scalability
- Energy to solution
- End-User Group

4. Conclusions

- Project status
- Mont-Blanc impact
- Remarks

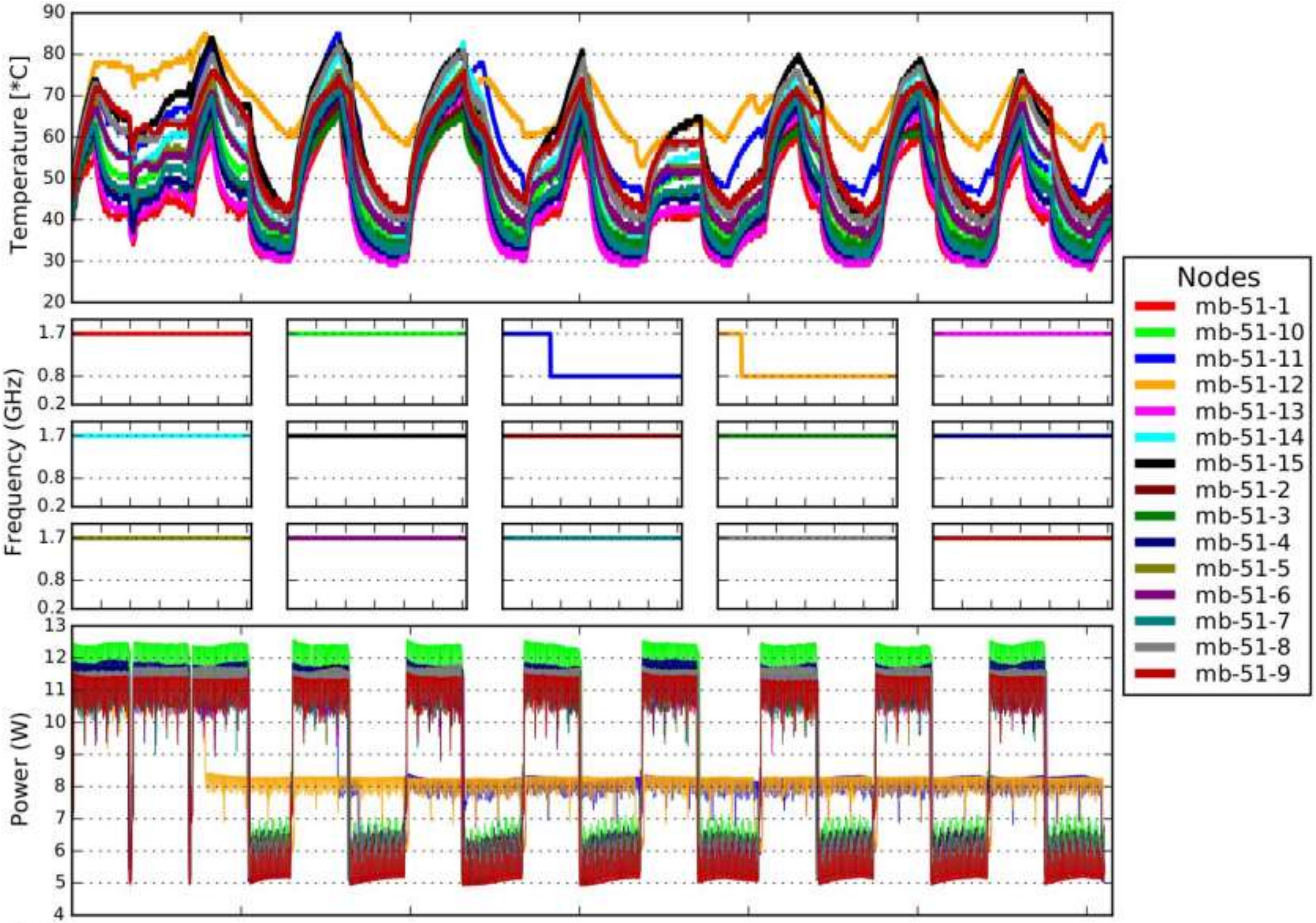
One application running: Alya RED

- Electro mechanics of the heart (of a rabbit)
- 10 steps measuring the parallel region of the code (MPI only)

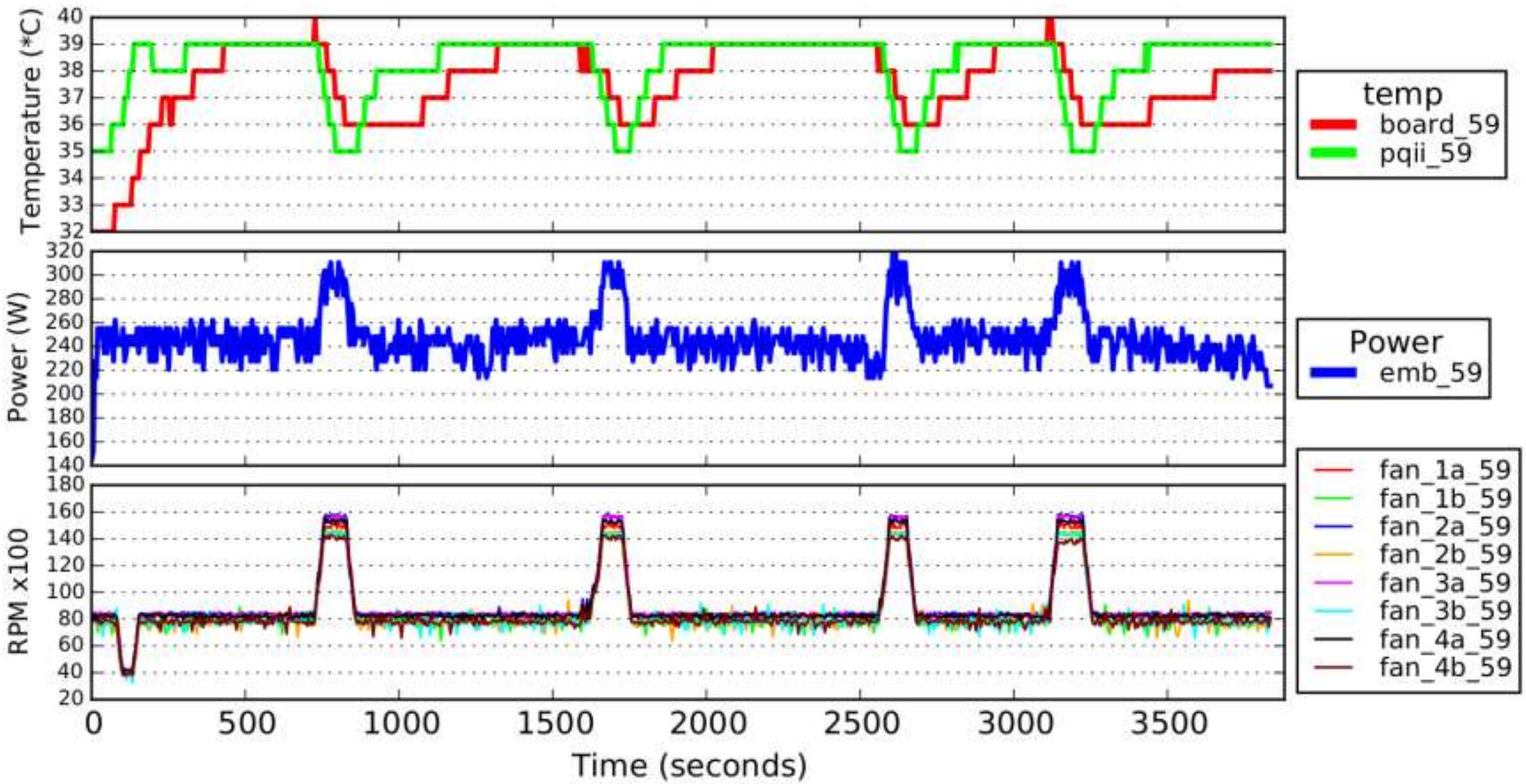


Source: Constantino Gomez, Alejandro Rico

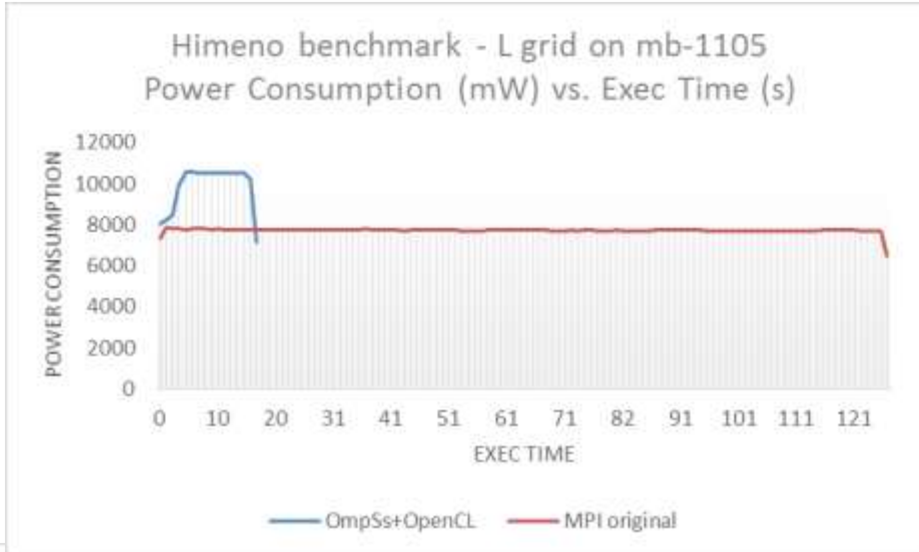
Understanding… at node level

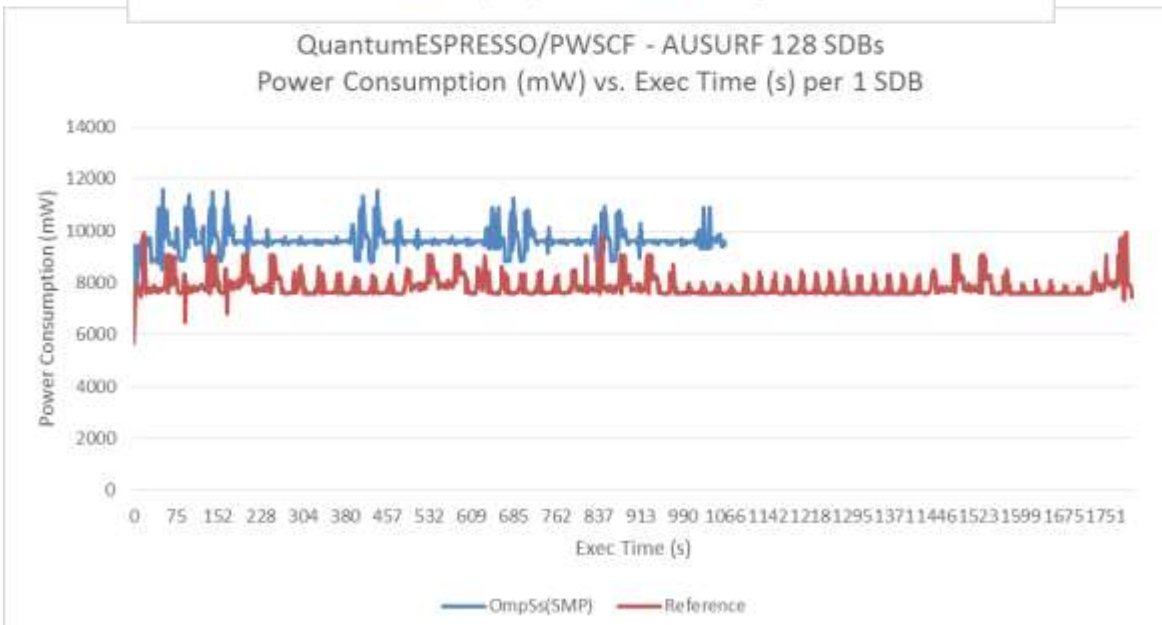


Understanding… at blade level



Energy to solution





End-User Group

- Develops a synergy among industry, research centers and partners of the project
- Validates the novel HPC technologies produced by the project
- Provides feedback to the project

Mont-Blanc provides EUG members with:

- Remote access to Mont-Blanc prototype platforms
- Support in platform evaluation and performance analysis
- Invitation to the Mont-Blanc training program

Outline

1. Hardware
- History
- Mont-Blanc prototype
- Limitations

2. Software
- System software
- OmpSs programming model
- Power monitor

3. Applications
- Scalability
- Energy to solution
- End-User Group

4. Conclusions
- Project status
- Mont-Blanc impact
- Remarks

Mont-Blanc project status

DONE!

- Prototype: design, development, deployment, monitor
- Deployment of HPC software stack on ARM
- Porting of HPC kernels and applications
- Test of non-HPC workload (Hadoop, OpenStack)

ON GOING…

- Next-generation architecture modelling
- ARM 64-bit exploration (mobile and server market)
- Porting of new applications
- Programming model enhancement
- Monitoring prototype for fault tolerant techniques

Mont-Blanc impact

MONT-BLANC

Lower cost

Higher cost

Cheap computing
- Industrial applications
- Small labs
- Small datacenters

New HPC node architecture
- Modular (IP)
- Requires industry involvement
- European architecture

Mobile technology
- Cheap due to volumes
- Fast evolution
- Constrained by mobile requirements
- More aggressive

Automotive
- Sensor fusion
- Autonomous driving
- Hybrid fuel
- …

Server technology
- Expensive
- Slower evolution
- More oriented to HPC (?)
- More conservative

Student Cluster Competition – ISC'15

Visit us @ booth #422

- First team participating with ARM-based cluster
- First Spanish team
- 3 kW power budget
- 5+ applications
  - HPCC
  - LAMMPS
  - PyFR
  - Octopus
  - Some "secret" applications
- 3 awards
  - Highest HPL
  - Fan favorite
  - 1st, 2nd, 3rd overall places

Conclusions

- Mont-Blanc project explores the use of mobile embedded technology for scientific computing
- Hardware contributions
  - Several ARM-based prototypes have been developed
  - Companies and research institutions (EUG) can test MB prototypes
- Software contributions
  - System software + Power monitor
  - Programming model
  - Applications
- Even with “cheap” hardware it is possible to perform “decent” scientific computing

> “The secret is to win going as slowly as possible.”
> Niki Lauda

Mont-Blanc project

montblanc-project.eu

MontBlancEU

@MontBlanc_EU

filippo.mantovani@bsc.es

"The secret is to win going as slowly as possible."
Niki Lauda