MONT-BLANC

http://www.montblanc-project.eu

High Performance Computing based on embedded processors

Filippo Mantovani

Barcelona Supercomputing Center

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 288777.

Mont-Blanc project goals

- To develop an **European** Exascale approach
- Leverage **commodity** and **embedded** power-efficient technology

Supported by EU FP7 with 16M€ under two projects:

- Mont-Blanc: October 2011 – September 2014
  14.5 M€ budget (8.1 M€ EC contribution), 1095 Person-Month
- Mont-Blanc 2: October 2013 – September 2016
  11.3 M€ budget (8.0 M€ EC contribution), 892 Person-Month

Mont-Blanc: Project objectives

- To deploy a prototype HPC system based on **currently available** energy-efficient embedded technology
  - Scalable to 50 PFLOPS on 7MWatt
  - Competitive with Green500 leaders in 2014
  - Deploy a full HPC system software stack

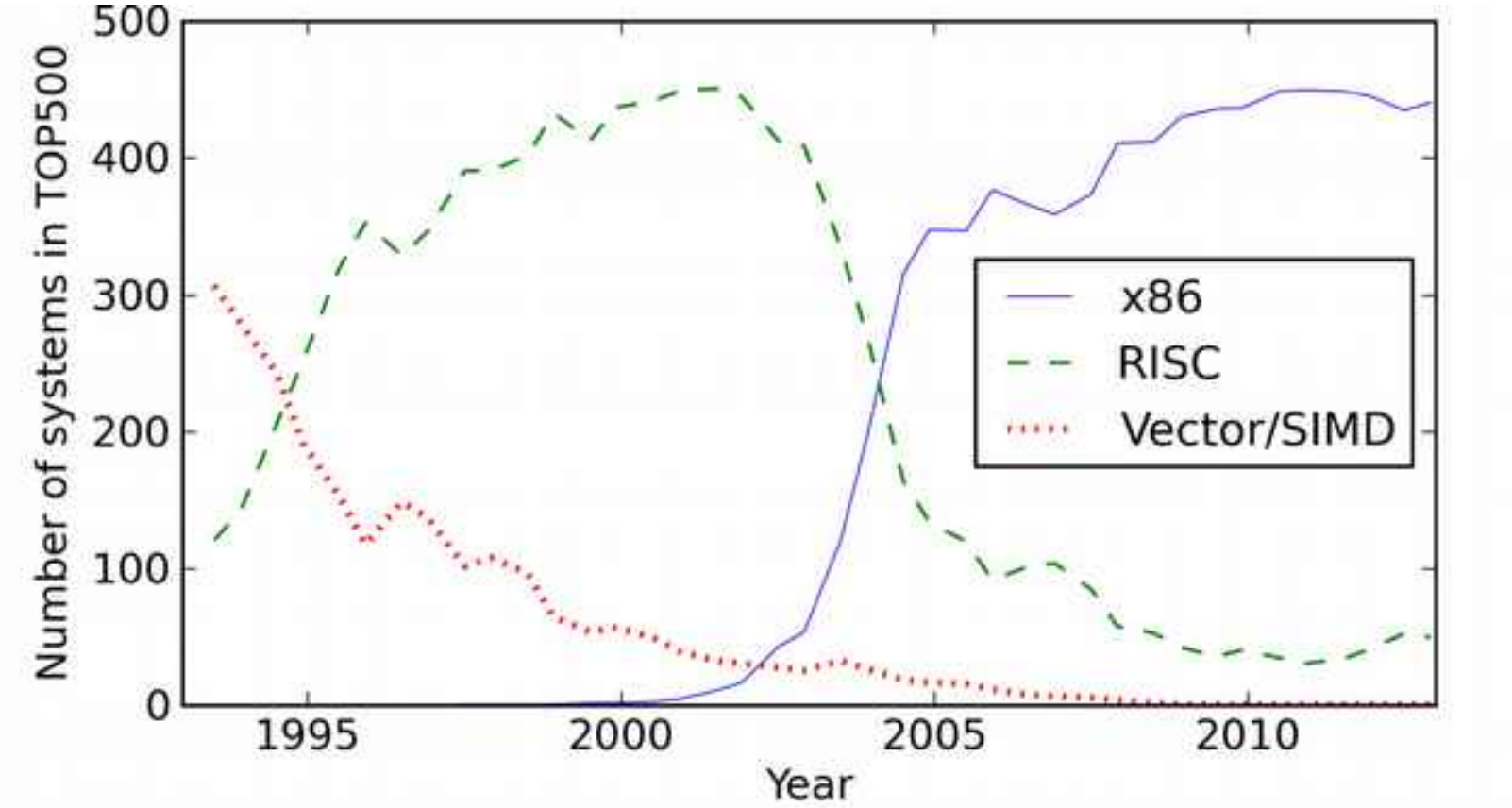
- To design a next-generation HPC system and new embedded technologies targeting HPC systems that would **overcome most of the limitations** encountered in the prototype system
  - Scalable to 200 PFLOPS on 10MWatt
  - Competitive with Top500 leaders in 2017

- To port and optimize a small number of **representative Exascale applications** capable of exploiting this new generation of HPC systems
  - Up to 11 full-scale applications

Mont-Blanc 2: Project objectives

- Continue **support for** the **Mont-Blanc** consortium
  - Mont-Blanc prototype(s) operation
  - OmpSs developer support
  - Increased dissemination effort (End-User Group)
- Complement the effort on the Mont-Blanc **system software stack**
  - Development tools: debugger, performance analysis
  - Resiliency
  - ARMv8 ISA
- Continue tracking and **evaluation of ARM**-based products
  - Deployment and evaluation of small developer kit clusters
  - Evaluation of their suitability for HPC
- Initial definition of future Mont-Blanc **Exascale architectures**
  - Performance & power models for design space exploration

Why are you doing this?



- RISC processors replaced vectors
- x86 (commodity) processors replaced RISC
  - Vector processors survive as (widening) SIMD extensions

Why are you doing this?

Commodity

HPC

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

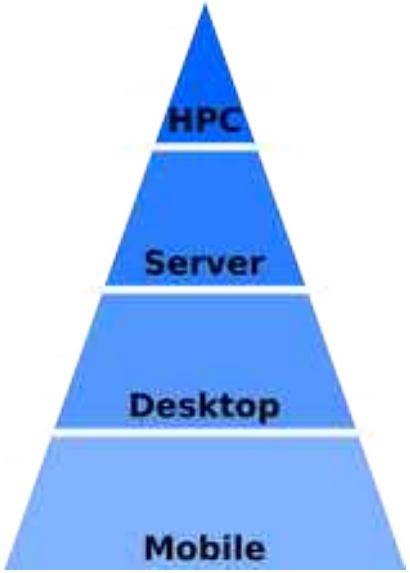
First petaFLOPS supercomputer
Roadrunner (IBM / Los Alamos NL - 2008)
AMD Opteron + PowerXCell 8i

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

What's commodity nowadays?

~22M cores (June '14)

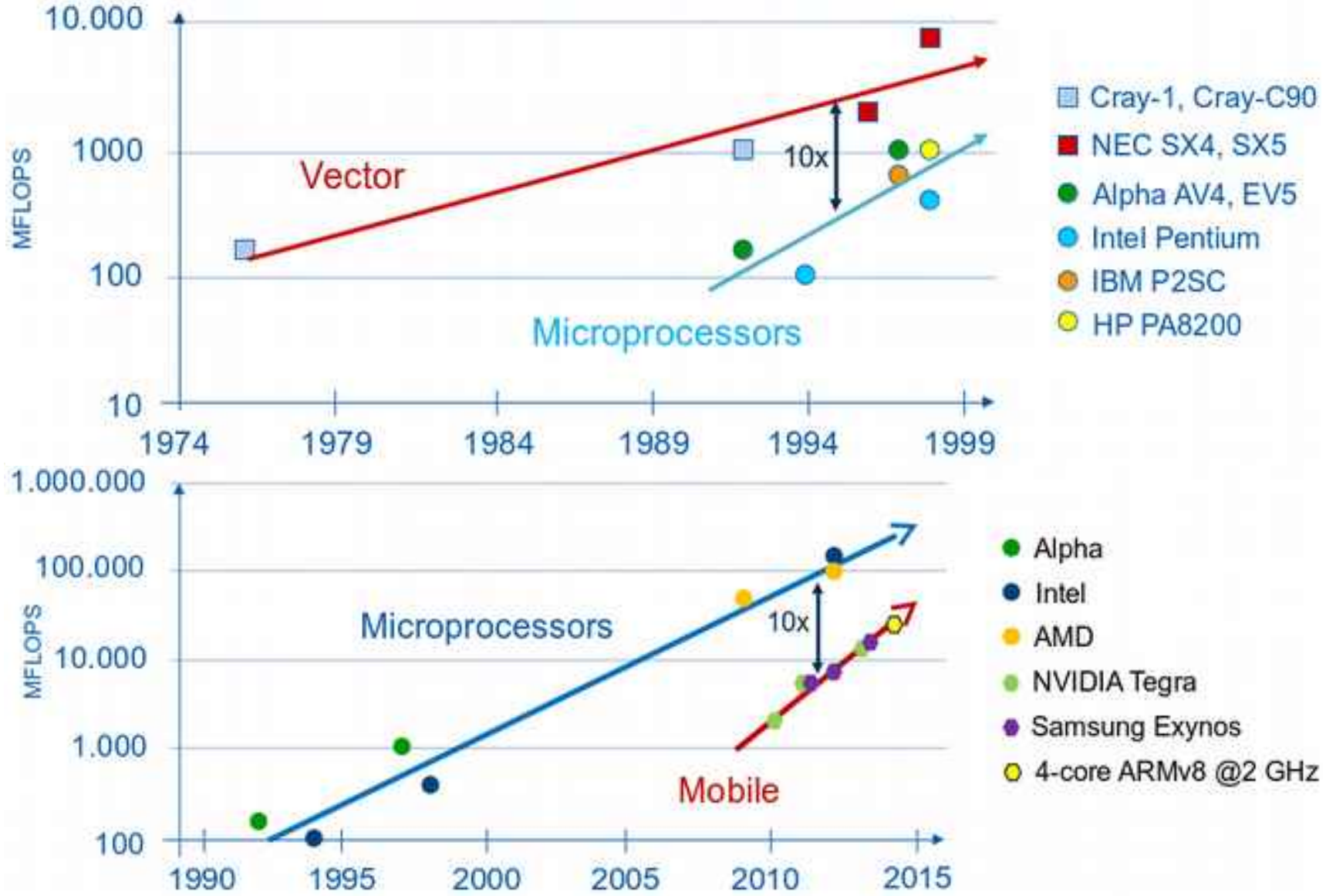
	Servers		PC		Smartphones	
2012	8.7M		350M		725M	
2013	9.0M	+3%	315M	-9.8%	1000M	+38%



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

Source: International Data Corporation

In case of “invasion”, we want to be ready...



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

Arndale Octa:
ARM bigLITTLE
Heterogeneous
multi-processing

Nvidia Jetson
ARM 4+1 + K1 GPU

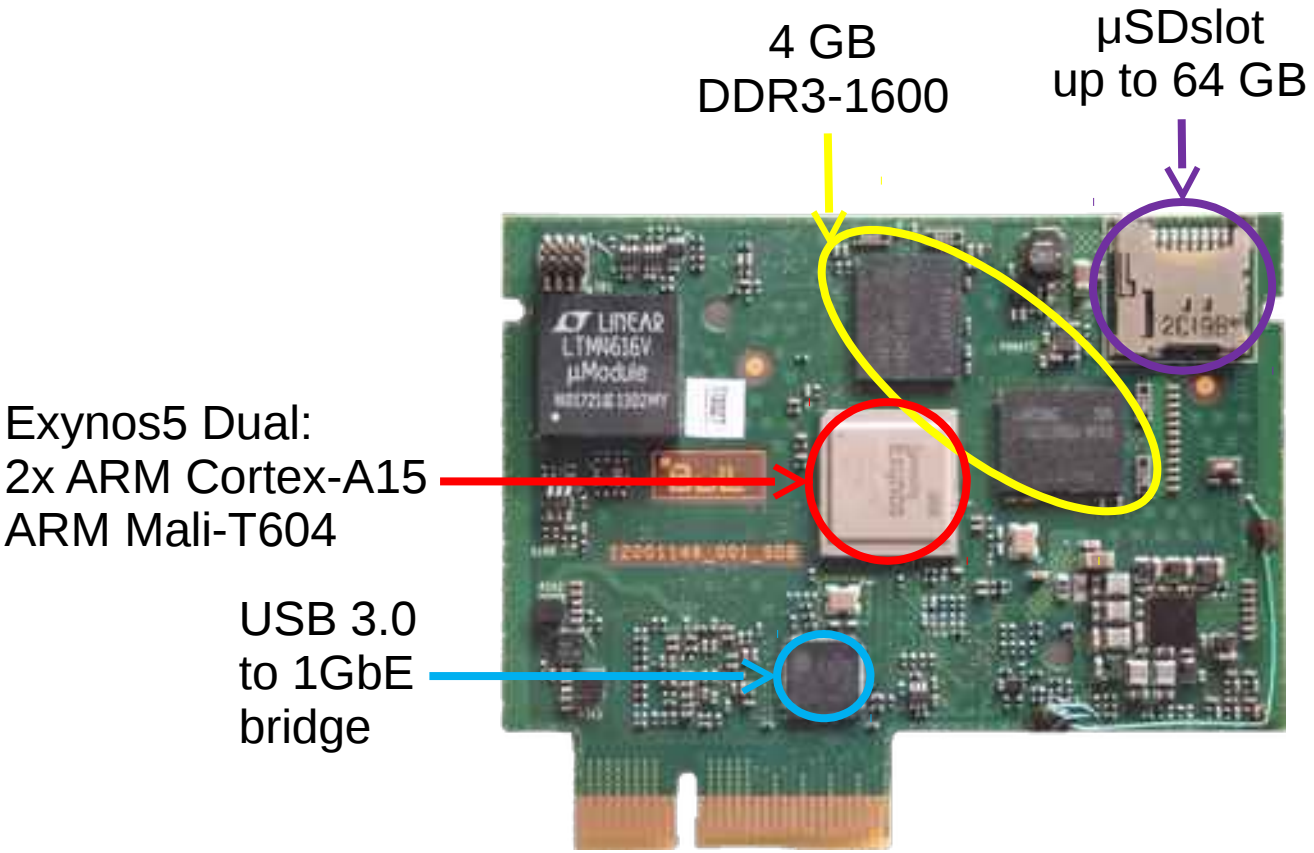
Mont-Blanc protoype:

2011 2012 2013 2014

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

Mont-Blanc Server-on-Module (SoM)

CPU + GPU + DRAM + storage + network
all in a compute card just 8.5 x 5.6 cm



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

GPU ~ 2/3 peak
CPU ~ 1/3 peak

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

6 BullX chassis

54 Compute blades

810 Compute cards

1620 CPU

810 GPU

3.2 TB of DRAM

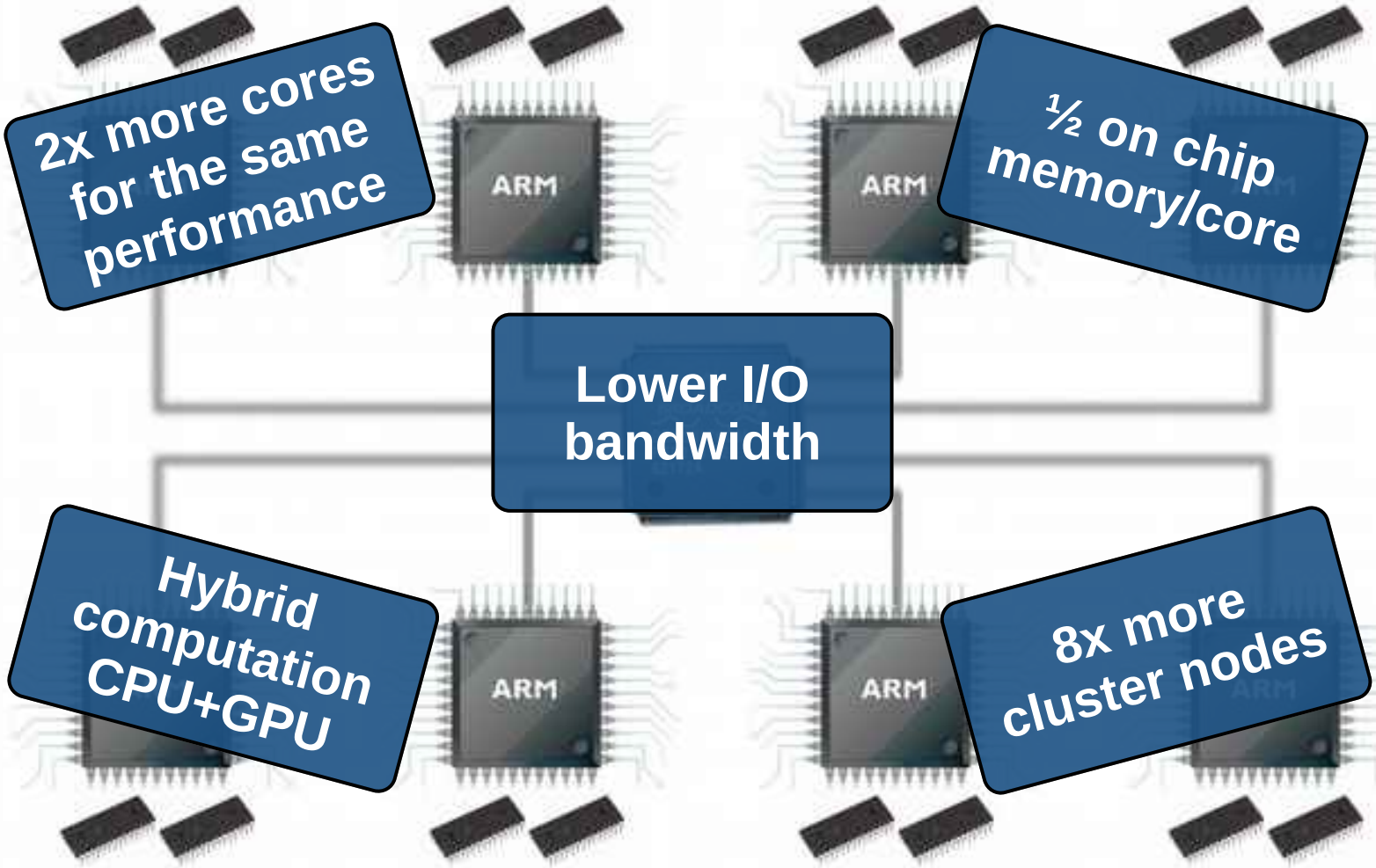
52 TB of Flash

26 TFLOPS

18 kWatt

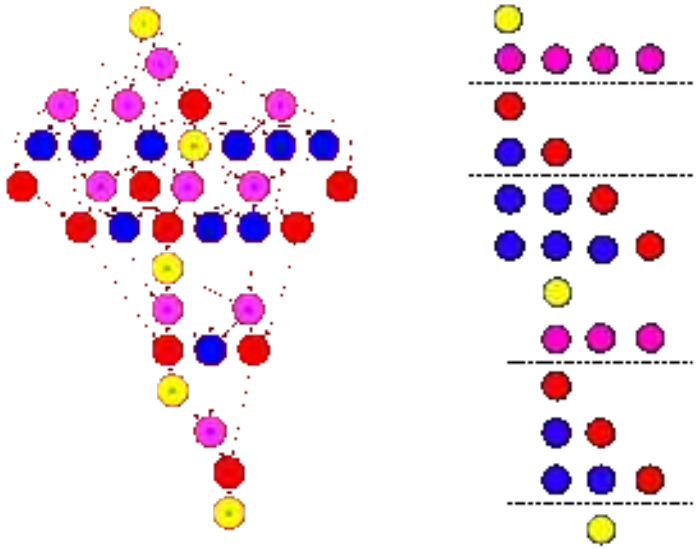
	Mont-Blanc [GFLOPS/W]	Green500 [GFLOPS/W]
Nov 2011	0.15	2.0
Jun 2014	1.5	4.4

Everything perfect, then?



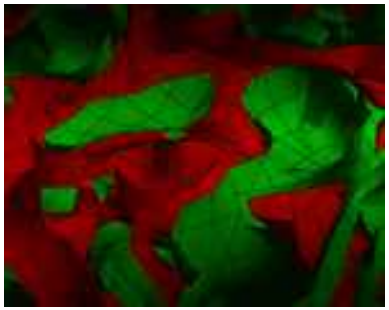
OmpSs runtime manages architecture complexity

- Programmer exposed to a simple architecture
  - Tasks
  - Data dependencies
  - Target (heterogeneity)
- Task graph provides lookahead
  - Exploit knowledge about the future
- Automatically handle many of the architecture challenges
  - Strong scalability
  - Multiple address spaces
  - Low cache size
  - Low interconnect bandwidth




```
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]);
  }
}
```

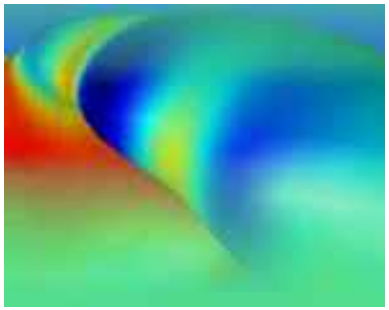
Porting applications to Mont-Blanc



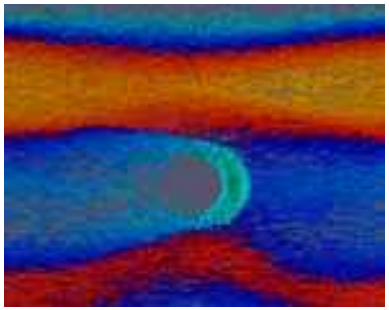
BQCD
Particle physics

BigDFT
Elect. Structure

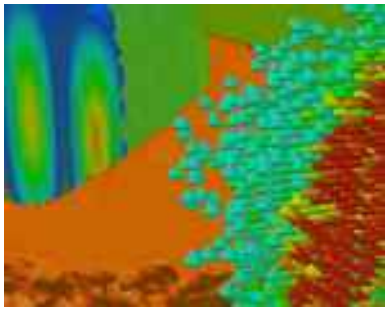
COSMO
Weather forecast



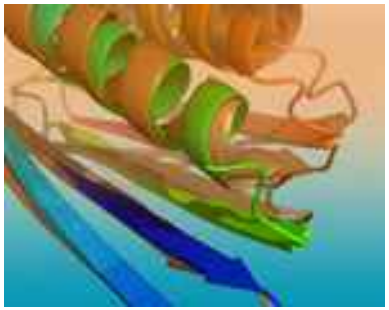
EUTERPE
Fusion



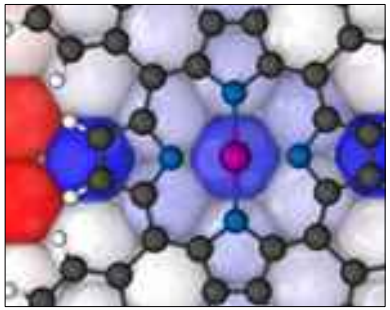
MP2C
Multi-particle collisions



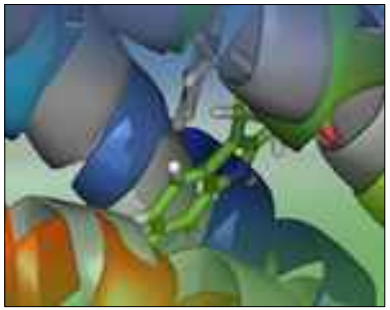
PEPC
Coulomb + Grav. Forces



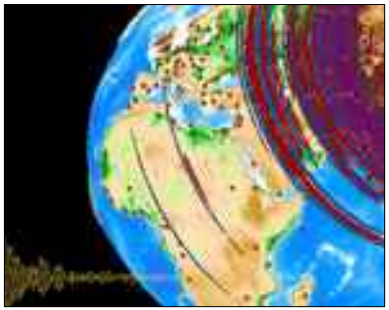
ProFASI
Protein folding



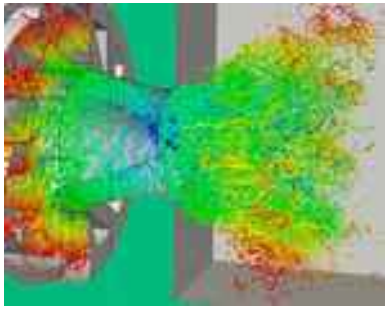
Quantum ESPRESSO
Elect. Structure



SMMP
Protein folding



SPECFEM3D
Wave propagation



YALES2
Combustion

GPU capable
(CUDA or OpenCL)
OmpSs capable

+ 4 applications of
Mont-Blanc2

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

Micro-benchmarks results (multicore + GPU):

Architecture summary:

2011

2 x ARM Cortex-A9 @ 1GHz
1 x 32-bit DDR2-333 channel
32KB L1 + 1MB L2

2012

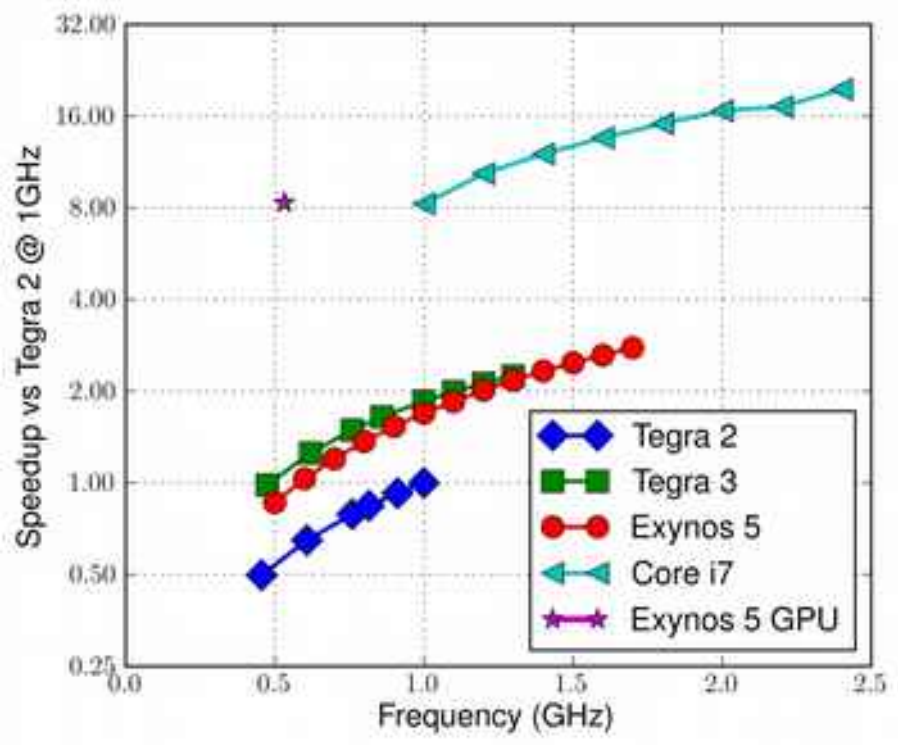
4 x ARM Cortex-A9 @ 1.3GHz
1 x 32-bit DDR3-750 channel
32KB L1 + 1MB L2

2012

2 x ARM Cortex-A15 @ 1.7GHz
2 x 32-bit DDR3-800 channels
32KB L1 + 1MB L2
GPU Mali T604 (4 shader)

2012

4 x Intel SandyBridge @ 2.4GHz
2 x 64-bit DDR3-800 channels
32KB L1 + 1MB L2 + 6MB L3



Under evaluation:

Samsung Exynos 5 Octa big.LITTLE

- IKS (in-kernel switcher)
- MP (multi-processing)

NVIDIA Tegra K1

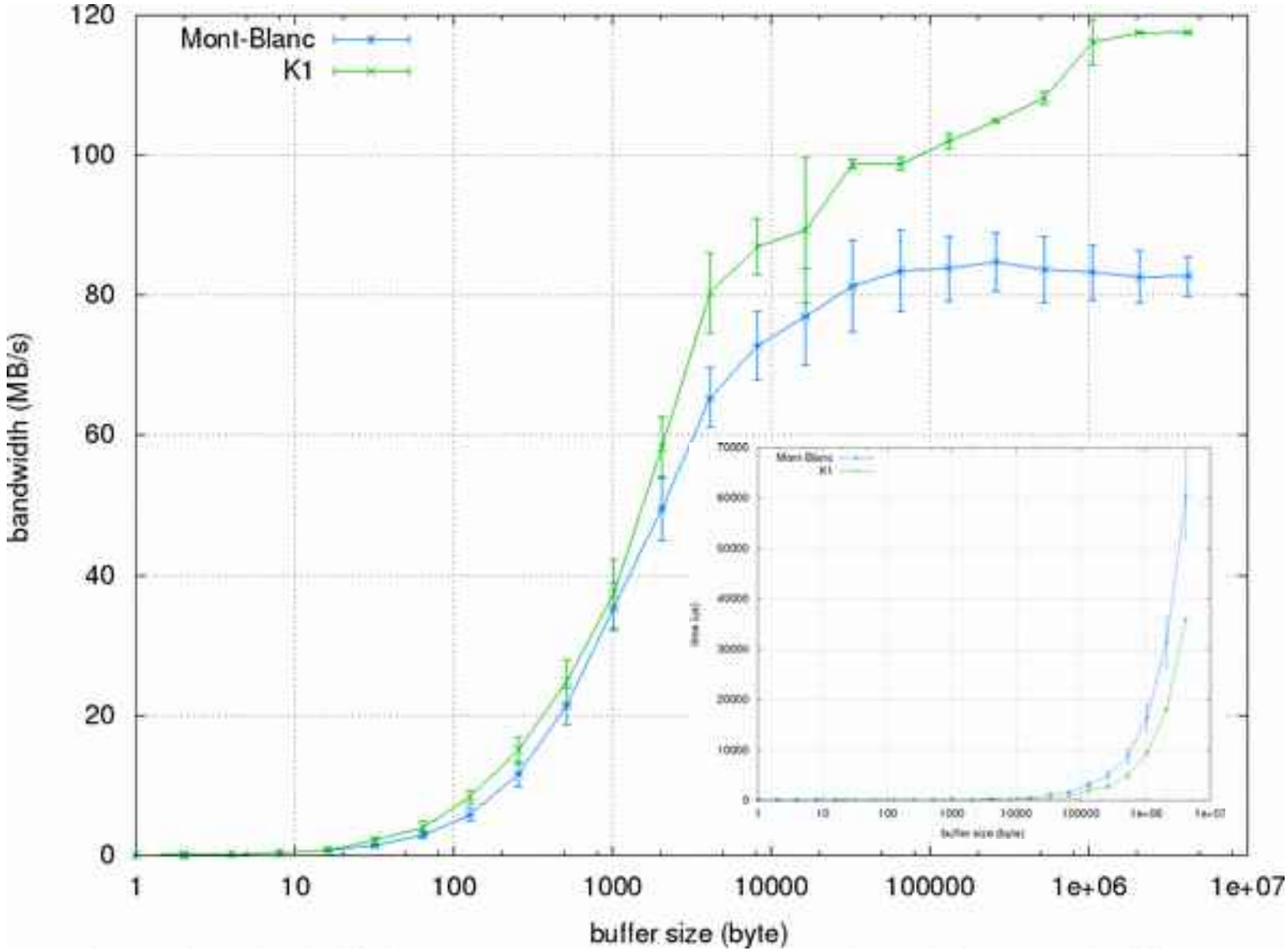
What do we learn from this?

What we learned so far:

- 32-bit memory controller
  - Even if ARM Cortex-A15 offers 40-bit address space
- No ECC protection in memory
  - Limited scalability, errors will appear beyond a certain number of nodes
- No standard server I/O interfaces
  - Do NOT provide 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

All these are **implementation decisions, not unsolvable problems**.
Only need a business case to justify the cost of including the new features
(e.g. the HPC and server markets)

e.g. in the meantime one issue is “solved”...



Thanks to D. Ruiz, J. O. Vilarrubi

Conclusions:

- Need sustainable EFLOPS technology
  - min(power + space + cost + …)
- Europe has a strong position in embedded computing
  - Energy efficiency
  - Commodity market
- BSC has a strong position in parallel programming models
  - OmpSs tasking model extends OpenMP 4.1
- Mont-Blanc offers a nice bag of Lego bricks to play with
  - Mont-Blanc prototype(s)
  - HPC software stack for embedded devices
- Leverage on all this to build
  a new class of sustainable computer
  faster, cheaper, more efficient

montblanc-project.eu

MontBlancEU

@MontBlanc_EU