The Mont-Blanc Project

Daniele Tafani
Leibniz Supercomputing Centre
Outline

• A bit of history...
  • Microprocessors killed vector supercomputers
  • Next step in commodity chain: killer mobile processors?

• The Mont-Blanc Project
  • General overview and project objectives
  • System architecture
  • Power aspects
  • Cooling aspects

• Conclusions, Q/A
In the beginning there were only supercomputers...

- **Built to order**
  - Very few of them
- **Special Purpose Hardware**
  - Very expensive!
- **Control Data, Convex,...**
- **Cray-1**
  - 1975, 160 MFlops, 80 units, approx. 5-8M $
- **Cray X-MP**
  - 1982, 800 MFlops
- **Cray-2**
  - 1985, 1.9 GFlops
- **Cray Y-MP**
  - 1988, 2.6 GFlops
- **Fortran + vectorizing compilers**
The killer mobile processors™

- Microprocessors killed the Vector supercomputers
  - They were not faster ...
  - ... but they were significantly cheaper and greener

- History may be about to repeat itself ...
  - Mobile processor are not faster ...
  - ... but they are significantly cheaper
ARM Processor Improvements in DP Flops

- IBM BG/Q and Intel AVX implement DP in 256-bit SIMD
  - 8 DP ops / cycle
- ARM quickly moved from optional floating-point to state-of-the-art
  - ARMv8 ISA introduces DP in the NEON instruction set (128-bit SIMD)
**ARM Processor Efficiency vs Intel / IBM / Nvidia**

- **Cortex-A15 @ 2 GHz**: Not an ARM commitment.
- **Cortex-A9 @ 1 GHz**: 
- **ARM11 @ 482 MHz**: 
- **BG/Q @ 1.6 GHz**: 

*Based on ARM Cortex-A9 @ 2GHz power consumption on 45nm, not an ARM commitment.*
The Mont-Blanc Project Goals

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

Fundied under FP7 Objective ICT-2011.9.13 Exascale computing, software and simulation
  • 3-year IP Project (October 2011 - September 2014)
  • Total budget: 14.5 M€ (8.1 M€ EC contribution)
Hardware: Samsung Exynos 5 Dual

- 32nm HKMG
- Dual-core ARM Cortex-A15 @ 1.7 GHz
- Quad-core ARM Mali T604
  - OpenCL 1.1
- Dual-channel DDR3
- USB 3.0 to 1 GbE bridge

All in a low-power mobile socket!
Hardware: Insignal Arndale development board

- Exynos 5 Dual SoC, full profile OpenCL
  - 2x ARM Cortex-A15, ARM Mali-T604, 2GB DDR3
- 100 Mbit Ethernet, NFC, GPS, HDMI, SATA 3, 9-axis sensor, ...
- uSD, USB 3.0
- Available today, priced at $249
What about performance?

Sandy Bridge + Nvidia K20

Samsung Exynos 5 Dual

1 Gb/s

10-40 Gb/s
There is no free lunch...

Sandy Bridge + Nvidia K20

- 2x more cores for the same performance!
- 8x address space!
- 1/2 on-chip memory/core!
- 1 GbE inter-chip communication!

Samsung Exynos 5 Dual

1 Gb/s

10-40 Gb/s
“We’re only in it for the money”...and energy!

Sandy Bridge + Nvidia K20

- $ > 3000
- W > 400

Samsung Exynos 5 Dual

- $ < 200
- W < 100
BullX Carrier Blade

- Each blade is a cluster on its own
  - 15 compute nodes + integrated GbE switch
Prototype architecture

**Exynos 5 Compute card**
1x Samsung Exynos 5 Dual
2 x Cortex-A15 @ 1.7GHz
1 x Mali T604 GPU
6.8 + 25.5 GFLOPS (peak)
≈10 Watts
3.2 GFLOPS / W (peak)

**Carrier blade**
15 x Compute cards
485 GFLOPS
1 GbE to 10 GbE
200 Watts (?)
2.4 GFLOPS / W

**7U blade chassis**
9 x Carrier blade
135 x Compute cards
4.3 TFLOPS
2 KWatt
2.2 GFLOPS / W

**1 Rack**
4 x blade cabinets
36 blades
540 compute cards
2x 36-port 10GbE switch
8-port 40GbE uplink
17.2 TFLOPS (peak)
8.2 KWatt
2.1 GFLOPS / W (peak)

- Mont-Blanc prototype limited by SoC timing + availability
  - Exynos 5 Dual is the 1st ARM Cortex-A15 SoC
- Better mobile SoCs keep appearing in the market ...
  - Exynos 5 Octa, Tegra 4, Snapdragon 800 ...

26th June 2013
Power Aspects

- Power gating, clock gating
- Voltage and Frequency Scaling (VFS)
  - Allows considerable energy savings by reducing the frequency at which the CPU is clocked
  - Preliminary test performed running the Hydro Benchmark on the Arndale Board
Power Aspects

Arndale Board Hydro Benchmark Energy-to-Solution

Energy-to-Solution (J)

CPU Frequency (MHz)

SWEET SPOT
Cooling Aspects

- **Air cooling**
  - Remove waste heat by blowing air into the rack and redirecting it outdoors.
  - Can be further improved with the adoption of heat exchangers

- **Liquid cooling**
  - Use a liquid coolant for removing the waste heat.
  - Different solutions: direct liquid cooling (coldplate, pipeline, etc.), indirect liquid cooling, immersion cooling

Bull Newsca compute unit (Coldplate)  
LRZ SuperMUC compute unit (cooling pipeline)
Cooling Aspects

Liquid Cooling vs Air Cooling...

- Thermal conductivity water = 21.5x Air!
- Thermal capacity water = 4.12x Air
- Maximize computing package density
- Better opportunities for free cooling

Liquid Cooling wins 4-0...

...however...
...Air Cooling is still a viable option because of different reasons...

- Heat dissipation profile
  - The prototype will have different heat dissipation profile than standard x86 systems.

- Daughterboard system packaging
  - The prototype will reuse Bull system architecture

- Air-cooled components
  - Power supplies, network switches,...

- Maintainance costs...

...and we still have rear-door heat exchangers...
HPC System software stack on ARM

- Open source system software stack
  - Ubuntu Linux OS
  - GNU compilers
    - gcc, g++, gfortran
  - Scientific libraries
    - ATLAS, FFTW, HDF5,...
  - Slurm cluster management
- Runtime libraries
  - MPICH2, OpenMP
  - OmpSs toolchain
- Performance analysis tools
  - Paraver, Scalasca
- Allinea DDT 3.1 debugger
  - Ported to ARM
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

* Already GPU capable (CUDA or OpenCL)
Conclusions

- **Objective 1:** to deploy a prototype HPC system based on currently available energy-efficient embedded technology.

- **Objective 2:** to design a next-generation HPC system together with a range of embedded technologies in order to overcome the limitations identified in the prototype system.

- **Objective 3:** to develop a portfolio of Exascale applications to be run on this new generation of HPC systems.

Stay tuned!

www.montblanc-project.eu

MontBlancEU

@MontBlanc_EU
Thank you for your attention!

...Questions?