Building supercomputers from embedded technologies

Alex Ramirez
Barcelona Supercomputing Center
Technical Coordinator
The problem … and the opportunity

- Europe represents ~35% of the HPC market
  - Yet, it does not have HPC technology of its own

- Nobody really knows how to build a sustainable EFLOPS supercomputer

- Consensus about it being revolutionary, not evolutionary
  - Everyone starts from square zero, hence equal opportunities

- Europe is very strong in embedded computing
  - The most energy- and cost-efficient computing technology today
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
    - 11.3 M€ budget (8.0 M€ EC contribution), 892 Person-Month
Mont-Blanc 1: Project objectives

• Objective 1: To deploy a prototype HPC system based on currently available energy-efficient embedded technology
  • Scalable to 50 PFLOPS on 7 MWatt
    • Competitive with Green500 leaders in 2014
  • Deploy a full HPC system software stack

• Objective 2: 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 10 MWatt
    • Competitive with Top500 leaders in 2017

• Objective 3: To port and optimise 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

- **Objective 1:** Complement the effort on the Mont-Blanc **system software stack**
  - Development tools: debugger, performance analysis
  - Resiliency
  - ARMv8 ISA
- **Objective 2:** Initial definition of the Mont-Blanc **Exascale architecture**
  - Performance & power models for DSE
- **Objective 3:** Continued tracking and **evaluation of ARM-based products**
  - Deployment and evaluation of small developer kit clusters
  - Evaluation of their suitability for HPC
- **Objective 4:** Continued **support** for the Mont-Blanc consortium
  - Mont-Blanc prototype(s) operation
  - OmpSs developer support
  - Increased dissemination effort
Commodity components drive HPC

- Microprocessors replaced Vector/SIMD supercomputers
  - They were not faster
  - They were cheaper
Samsung Exynos 5 Dual Superphone SoC

- 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
Mont-Blanc SoM

- CPU + GPU + DRAM + storage + network ... all in a compute card just 8.5x5.6 cm
Mont-Blanc server blade

- 15 node-cluster in a standard Bull B505 enclosure

15 compute cards
30 x ARM Cortex-A15
15 x ARM Mali-T604 GPU
120 GB DDR3

Cluster management

1 GbE crossbar switch

2 x 10 GbE links
Mont-Blanc server chassis

• 9 blades in a standard 7U BullX chassis
  • Shared cooling, PSU, chassis management

9 x Compute blades:
135 x Compute cards + 36 x 10 GbE links
270 x ARM Cortex-A15
135 x ARM Mali-T604 GPU
540 GB DDR3-1600
The Mont-Blanc prototype

- 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
There is no free lunch

- 2X more cores for the same performance
- 1/2 on-chip memory / core
- 8X more cluster nodes
- Hybrid CPU + GPU
- Lower I/O bandwidth
OmpSs runtime layer manages architecture complexity

- Programmer exposed a simple architecture
- Task graph provides lookahead
  - Exploit knowledge about the future
- Automatically handle all of the architecture challenges
  - Strong scalability
  - Multiple address spaces
  - Low cache size
  - Low interconnect bandwidth
- Enjoy the positive aspects
  - Energy efficiency
  - Low cost
Advances over State of the Art

• Developed an entire HPC software ecosystem on ARM
  • Including OpenCL driver for GPU accelerator
  • Performance monitoring + tracing + analysis
  • Advanced parallel programming models
    • MPI + OmpSs @ OpenCL + FORTRAN

• Fine-grain per-node power monitoring
  • Open the door to future per-node power management
  • Combined with SoC power management features

• Microserver architecture for HPC
  • Built on commodity SoC + commodity network (Ethernet)
Limitations of current mobile processors for HPC

• 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 … such as the HPC and server markets
Get ready for the change, before it happens …

- Embedded processors have qualities that make them interesting for HPC
  - FP64 capability
  - Performance increasing rapidly + energy efficient
  - Large market, many providers, competition, low cost

- Lots of other embedded IP blocks to build the perfect HPC SoC
  - Embedded GPU accelerators, DSP, FPGA, …
  - Network interfaces + protocol offloads

- Current limitations are due to target market conditions
  - Not real technical challenges

- A whole set of ARM server chips is coming
  - Solving most of the limitations identified
Conclusions

• The convergence of Embedded and HPC technologies has happened already
  • We have enabled the software already

• Leverage on Europe’s leadership position in parallel programming models & tools

• Leverage on all of the embedded systems technology to build a new class of HPC system
  • Automated SoC design
  • Automatic core customization
  • SoC power management
  • Decouple IP provider from semiconductor provider

montblanc-project.eu
MontBlancEU
@MontBlanc_EU