2nd Workshop on Energy for Sustainable Science at Research Infrastructures
Report on parallel session A3

Wayne Salter on behalf of

Dr. Mike Ashworth (STFC)

Talks

- A Comprehensive Approach to Energy Efficiency in Data Centers for High-Performance Computing by Prof. Thomas C. Schulthess (CSCS)
- Exploiting mobile phone technology to build energy efficient supercomputers: the Mont Blanc project by Dr. Simon McIntosh-Smith (University of Bristol)
- Roadmap towards Ultimately-Efficient Datacenters by Dr. Bruno Michel (IBM)
- Energy Savings in CERN's main Data Centre by Wayne Salter (CERN)

Summary - I

- First three talks followed a common theme
  - Computing needs growing rapidly
  - As a result the power needed for computing growing very quickly too
  - Not sustainable
  - Need major change in technology and/or use
- Fourth talk
  - Discussed some concrete measures taken to improve the energy efficiency of an old existing DC

A Comprehensive Approach to Energy Efficiency in Data Centers for High-Performance Computing - I

- Discussed investments in Switzerland for HPC
- One of the goals of HP2C program was to push innovation in algorithm and application software design to take advantage better of capabilities of modern HPC H/W
  - Massive concurrency (multithreading and high node count)
  - Hybrid nodes (CPU+GPU)
- Three pronged approach:
  - Efficient DC (new DC built using cooling from lake already has PUE of 1.2 and this is likely to improve with more loading)
  - Efficient computers (designed specifically for the applications)
  - Efficient applications
- Most energy in computers today is moving data and not compute
- New system (Cray XC30 Piz Daint)
  - Next generation network interconnect developed through DARPA HPCS program
  - Hybrid nodes with Intel CPU and NVIDIA GPU
- Discussed improvements based on COSMO weather predication application

A Comprehensive Approach to Energy Efficiency in Data Centers for High-Performance Computing - II

- Main points:
  - Use of free cooling using lake water has resulted in an efficient data centre at CSCS with a PUE of ~ 1.2 for 2-3 MW
  - The dynamical core of the COSMO weather forecast code used by Meteo Swiss has been adapted to exploit GPU hardware
  - An improvement in efficiency of 10x has been achieved for the COSMO code combining 1.5x from the building, 1.75x from the new system, 1.49x from the new code and 2.64x from use of hybrid nodes
  - Future power efficiency improvements are more likely to come from applications development than from hardware

Exploiting mobile phone technology to build energy efficient supercomputers: the Mont Blanc - I

- HPC systems growing in performance but also power
  - Average power consumption of top 10 in 2008 was 1.5MW and now 6MW (5x increase in 5 years)
  - Future limiting factor not necessary the delivery of power but rather its cost
  - Europe is major HPC player but has no technology of its own
  - However, is strong in embedded computing
- How to build an affordable Exoflop machine?
  - Need revolutionary rather than evolutionary approach
- Mont Blanc project
  - European project led by Barcelona Supercomputing centre
  - Leverage commodity and embedded power-efficient technology from mobile market
  - Proof of concept has been made with a compute card containing Samsung CPU, GPU, DRAM and NAND memory, and NIC.
  - 15 cards/blade, 9 blades per chassis and 4 blade chassis in a rack
  - Delivers 17.2 TFLOPS for 8.2 kW
  - Ambitious development roadmap

Exploiting mobile phone technology to build energy efficient supercomputers: the Mont Blanc - II

- Main points:
  - Following the historical replacement of vector processors by commodity microprocessors – "killer micros", there may be a similar coup by the "killer mobiles"
  - Europe is well placed with strengths in embedded computing for mobile devices
  - The Mont Blanc project aims to produce an Exascale system based on ARM microprocessors with a 200 Pflop/s system consuming 10 MW in 2017 and an Exaflop system at 20 MW around 2020

Roadmap towards Ultimately-Efficient Datacenters - I

- We need three paradigm changes:
  - Moving from cold air cooling to hot water energy re-use
  - Analyse the system in term of efficiency not performance
  - From areal device size scaling to volumetric density scaling – build in 3D not in 2D - vertical integration
- Hot water cooling with waste heat re-use
  - SuperMUC I prototype at ETH Zurich
    - iDataPlex cluster with 3.2PFLOPS (20k CPUs/160k cores)
    - 4MW, PUE 1.15, 90% heat for re-use =>40% less energy consumption
- Analysing systems in terms of compute efficiency and density
  - Shows that we are still 4 orders of magnitude worse than a human brain => use as an example
  - Transistors occupy only 1ppm of system volume
  - Majority of energy used for communication (dependent on wire length and scales quadratically) => need to look at volumetric scaling
  - Some ideas presented on how to move from 2D to high density 3D chip design with interlayer cooling and electrochemical chip powering
  - Aim to develop 1 PFLOPS machine in 10 litres
- Can also learn from allometric scaling in biology

Roadmap towards Ultimately-Efficient Datacenters - II

- Main points:
  - We should move from cold air cooling to hot water energy re-use
  - We must analyse systems in term of efficiency not performance
  - Vertical integration will enable dense architectures which improve efficiency through chip stacking and interlayer cooling – Moore's Law goes 3D
  - Such an ultra-dense 3D system will achieve 1 Pflop/s in 10 litres

Energy Savings in CERN's main Data Centre

- Move from low power density mainframes to rack mounted servers led to cooling issues
  - Solved by cold/hot aisle separation in 2008
- Further improvements made to improve efficiency
  - Modification of air handling to increase substantially the use of free cooling => very low requirement for chillers
  - More aggressive temperature environment
  - Savings of > 6.2 GWh for a 2.6 MW DC
- Achieved with relative simple cost-effective measures
- Further measures foreseen