Accelerated Computing Service in PPD at RAL

Abdeslem DJAOUI

PPD

RAL

Outline

- Motivations and aims
- What we did at RAL in the last 5 years
- Our plan for the near future

Computer performance until ~ 2000

- CPU frequency doubled every 22 months
  - Memory capacity also increased
  - Applications run faster without any code change
  - **Program once, just sit and wait for better computers**
- But power dissipation: $P \sim CV^2f$
  - Power became unsustainable as f increased
    - ~ 4GHz ceiling even with best cooling technologies
- CPU vendors looked for more performance
  - by other means than increasing frequency
  - but without increasing power dissipation

The move to multi-core from 2000

- Rule of thumb based on : $P \sim CV^2f$
  15%🡻*in V results in:* 15%🡻*in f,* 50%🡻*in P,* 10%🡻*in Perf*
  -  *1 core: (area 1, V=1, f=1 ➔ P=1, Perf=1)*
  -  *2 cores: (area 2, V=0.85, f=0.85 ➔ P=1, Perf=1.8)*
- Todays CPUs (up to ~ 20 cores)
- But change to multicore is at price for applications
  - Multithreaded versions needed to exploit added Perf
    - Amdhal's law (even 5% sequential, 20x theoretical speed up)
    - Move to parallel programming tedious compared to sequential
    - Need synchronizations mechanisms (locks) to avoid race conditions
    - Programmer has to avoid thread deadlocks

Science & Technology Facilities Council

Then the advent of co-processor accelerators

- GPGPU: even more cores (1000's)
  - Simpler cores
    - Suitable for data parallel algorithms (small sequential part)
  - Now have heterogeneous architecture: CPU+ Accelerator
    - Run App on CPU, offload data parallel work to accelerator
    - Many dissimilar accelerator architectures exist
- **Harder life for the programmer**
  - **Big new concern in architecture**
    - **Which accelerator to target?**
  - **Big new concerns in the programming**
    - **Multiple programming models needed**
      - **CUDA, OpenCL, …**

Science & Technology Facilities Council

Other co-processors on offer

- Intel Xeon Phi
  - Different architecture model from GPUs
  - Application portability to/from GPUs not always possible
- FPGAs: Another way to get more more performance per Watt
  - FPGA also better performance per Watt than GPUs
  - Not a processor at all unlike CPUs and GPUs
  - A Plate of reconfigurable logic cells with programmable interconnects
  - Can implement any logic function (in hardware, not software)
- **Much harder life for the programmer**
  - **Huge new concern in architecture and programming**
  - **Programmer has to decide what to implement in hardware**
  - **New considerations (hardware replication versus pipelining, …)**

Consequences of these disruptive computing technologies

- Users need early access to gain competitive advantage
  - Access to the new hardware is difficult (cost, availability)
  - No expertise in associated models and tools
- A shared focused service could enable
  - Access for evaluation/prototyping for early adopters
  - Ready to use platforms with reliable support/expertise
  - A development environment following evaluation
  - A learning platform for new commers

Our first GPU in PPD five years ago

- In 2010 ATLAS needed evaluation of GPUs for HLT
  - We acquired a system with a NIVIDIA Tesla C2050
  - Used by ATLAS PPD and Collaborators at Oxford
  - Started with small prototype, which became a full project
    - Eventually moved to the bigger ATLAS testbed for the HLT tracking software
- Close collaboration with users was needed
  - To build the first prototype
  - Less frequent as programmers built expertise

Current resources following additional requirements in PPD

- HEPGPUW1 machine: (First GPU machine)
  - NVIDIA Tesla C2050 'Fermi' (448 GPU cores)
- HEPGPUW4: (OpenCL development on GPU)
  - AMD Firepro V9800 (1600 GPU cores)
- HEPGPW5 machine:
  - Xeon Phi co-processor (56 cores x 4-way hyperthreading)
- HEPACC01 server: (accessible to all PPD)
  - DUAL NVIDIA Tesla K40M GPU (2880 cores)

Use and in-house R&D

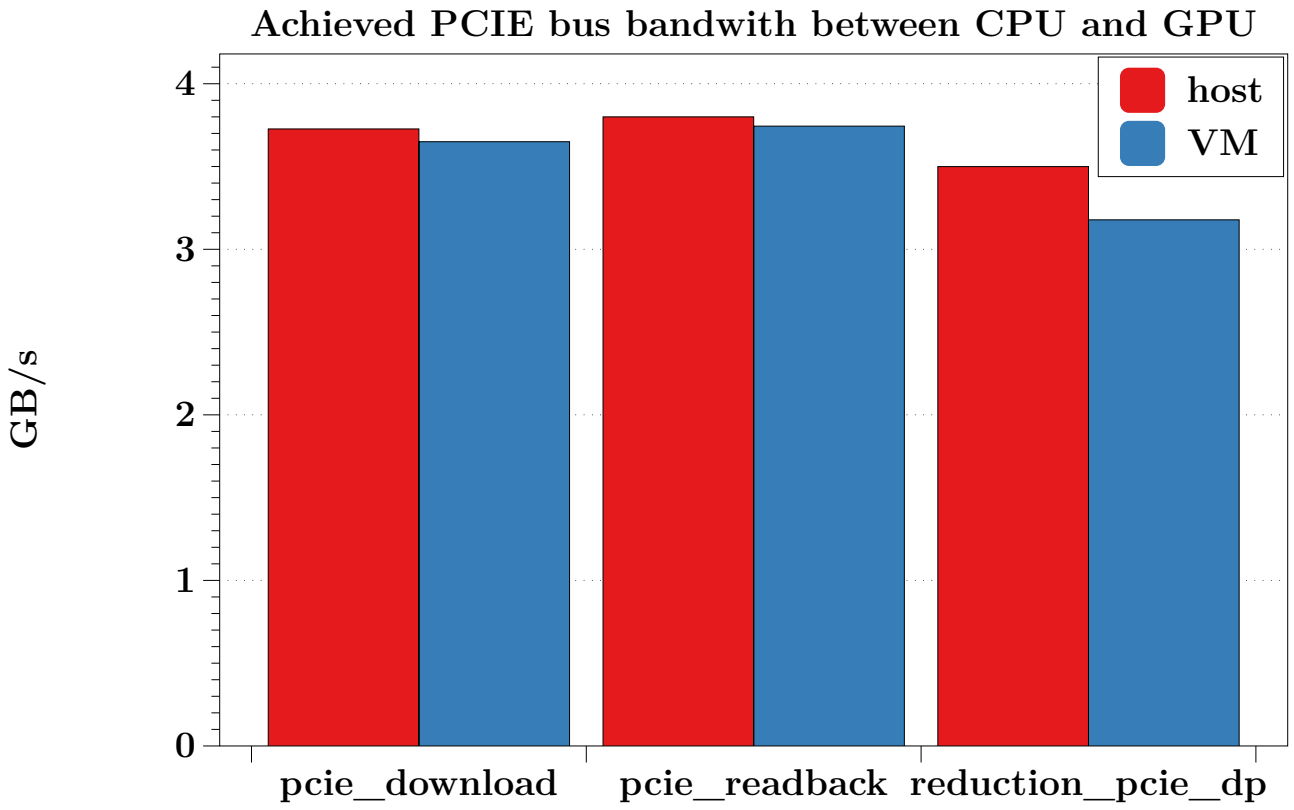
- Used by ATLAS/CMS for (mostly) online computing
  - Track reconstruction for L2 trigger (CUDA, OpenCL)
  - Region-of-interest data preparation (CUDA)
  - Statistical limit setting package (CUDA)
- A lot of summer students projects over the years
- In-house R&D activity
  - Benchmarking and application porting between devices
    - Evaluating OpenCL portability
    - Evaluation of OpenMP (4.0 offload)
  - The use of GPUs in Virtual Machines

GPUs in VM

- Two different uses of GPUs in Virtual environments
  - Desktop Virtualization
    - multiple users share one GPU
  - Our use case: dedicated GPUs per VM (GPU pass-through)
    - Directly connect GPU to VM through the Hypervisor
    - Then GPU not available to host or other VMs
- Virtualization testbed setup:
  - 2 GPU machine configurations
    - One without and One with Vmware ESXi Hypervisor (Type 1)
    - Same OS and tools on both
  - Applications from SHOC benchmark suite

Science & Technology Facilities Council

VM with GPU pass-through effect on PCIe bandwidth

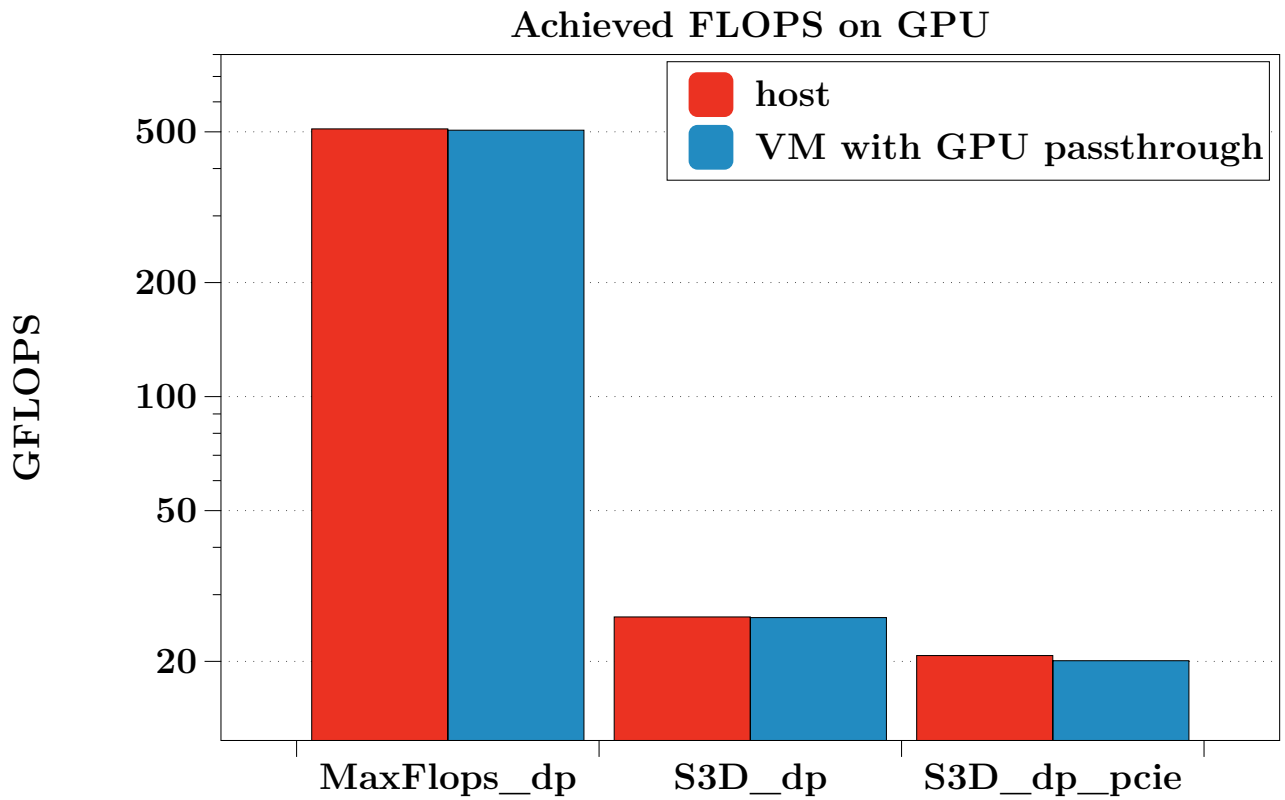


Apps

- Download (H2D) and readback (D2H) bandwidth of PCIe bus
- Sum Reduction includes kernel execution latency as well transfers

Science & Technology Facilities Council

VM with GPU pass-through effect on app performance



Apps

- MaxFlops: Maximum throughput possible
- S3D: Chemical rates application for turbulent combustion
- S3D_dp_pcie: includes PCIe transfers overhead

Science & Technology Facilities Council

In Planning

- Technology refresh needed every 3-4 years
  - In 2010: C2050 (448 cores, 1.15 GHz, 1TFLOP, 238W)
  - In 2014: K40 (2880 cores, 745 MHz, 4.28TFOPS, 235W)
- A lot of interest in OpenCL/FPGA from experiments
  - OpenCL on FPGA already part of ATLAS 10 year plan
  - Expect similar roadmap to GPUs over many years
  - Promising times for reconfigurable computing (FPGA)
    - Intel acquisition of Altera and IBM/Xilinx partnership
    - New devices (CPU and FPGA on same chip)
    - Portable productive programming models
      - » OpenMP on FPGA???

Science & Technology Facilities Council

Accelerated computing service beyond PPD

- In PPD the initiative has proved very effective in
  - Easing the path to productivity on new technologies
  - Enabling a competitive advantage
- Would like the same for the UK-wide HEP community
  - Not in competition with GRIDPP (but could be part of)
  - Not all Uni groups can secure access to early new technologies
  - Currently we could accommodate a couple of external users
  - More would require additional effort and resources
  - Some issues to think about: (authentication, file system, …)
- Eventually could be open STFC-wide

Science & Technology Facilities Council

Summary

- New disruptive technologies need a re-think of Computing services and support
- Early access to innovative for evaluation, prototyping and development should benefit the UK HEP community at large
  - Not just the bigger HEP groups
- RAL as a central Lab can take the lead in enabling this
- We would like ho hear from external users interested in any of the existing platfroms

Science & Technology Facilities Council