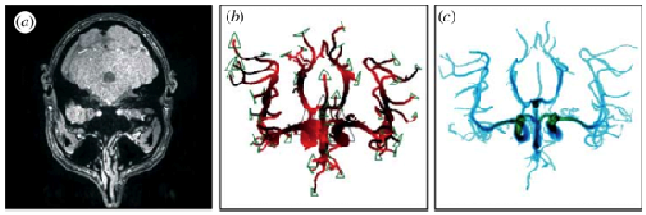
SUPERCOMPUTING

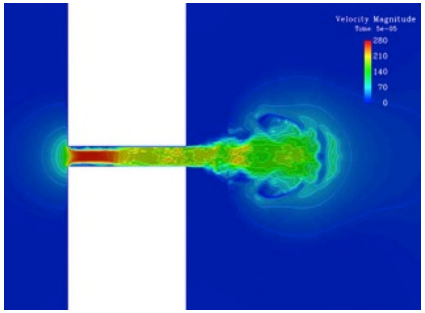
Adrian Jackson

adrianj@epcc.ed.ac.uk

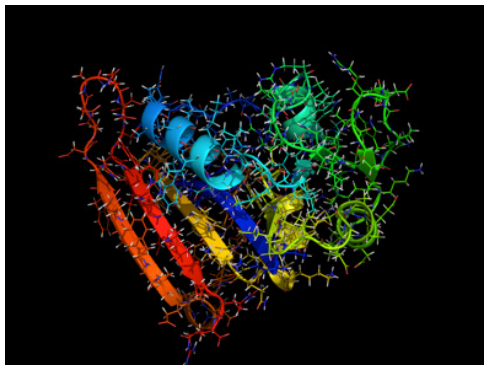
epcc

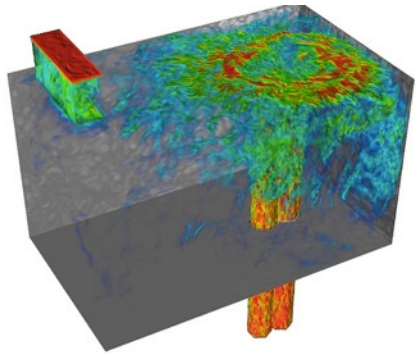
EPCC in miniature



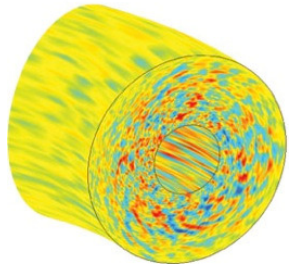


- The University of Edinburgh
- 25 years old – 80 staff
- Fully self-sustaining
- UK National HPC Service provider
- Wide range of work from HPC to Data Analytics and Cloud
- Work with academia and industry





- Facilities:
  - ARCHER: Cray XC30
    - 2.6 PFlop/s
  - DiRAC: IBM BG/Q
    - 6 frames, 1.2 PFlop/s
  - Indy:
    - 1600 core machine for industry simulations
    - Linux and windows dual deploy
  - SGI: Large shared memory system
  - EDIM1: Data intensive research machine
    - 240 cores (dual core ATOM), 480GB RAM, 30TB SSD, 720TB HD



epcc

Intel's IPCC program

- Collaboration between Intel and leading Universities around the world
- "Intel® Parallel Computing Centers are universities, institutions, and labs that are leaders in their field, focusing on modernizing applications to increase parallelism and scalability through optimizations that leverage cores, caches, threads, and vector capabilities of microprocessors and coprocessors."

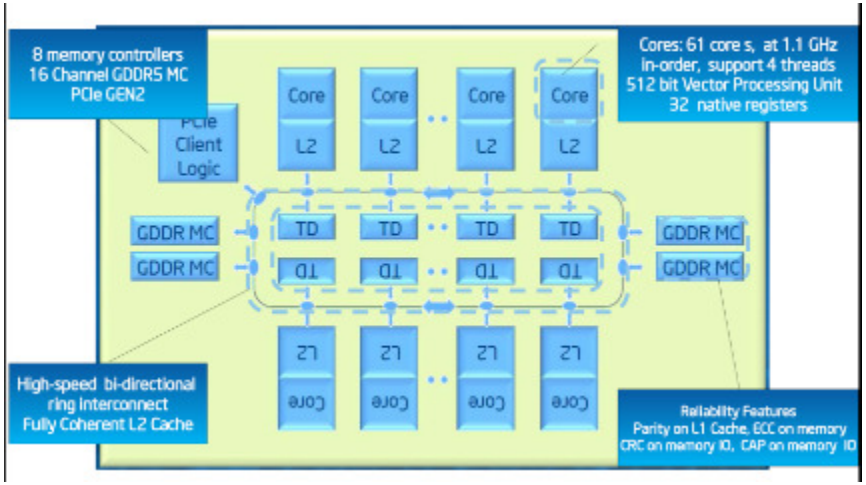
epcc

ARCHER

- UK national HPC for Engineering, Physical Sciences, and Natural Sciences
- Cray XC30 Hardware
  - Nodes based on 2×Intel Ivy Bridge 12-core processors
  - 64GB (or 128GB) memory per node
  - 4920 nodes in total (118,080 cores)
  - Linked by Cray Aries interconnect (dragonfly topology)

epcc

Xeon Phi Processor

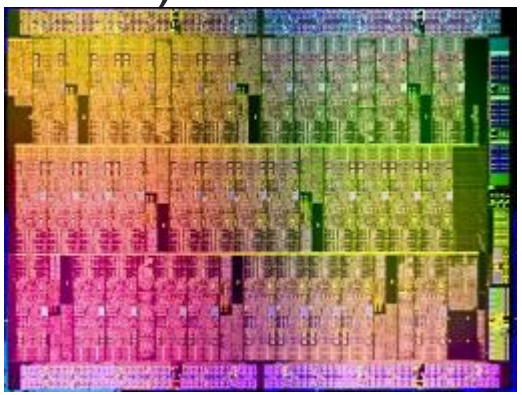


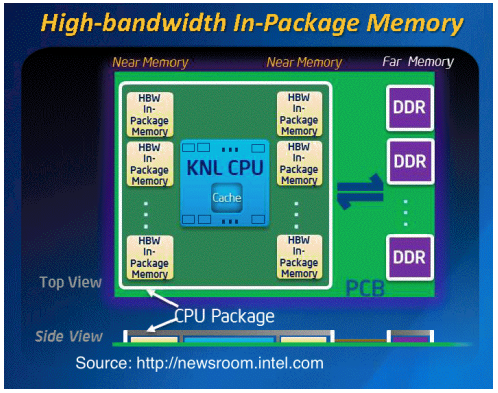
	3100 series	5100 series	7100 series
cores	57	60	61
Clock frequency	1.100 GHz	1.053 GHz	1.238 GHz
DP Performance	1 Tflops	1.01 TFlops	1.2 TFlops
Memory Bandwidth	240 GB/s	320 GB/s	352 GB/s
Memory	6 GB	8 GB	16 GB

epcc

KNL – Knights Landing

- Successor to Xeon Phi (Knights Corner)
  - Due to be released later this year
  - ~3 TFLOP/s double precision
  - 72 Airmont cores
  - 2 vector units per core
  - Two cores share 1MB L2 cache form a tile
  - A mesh fabric routes between the tiles
  - On package-DDR4 memory controller support for up to 384 GB main memory
  - Up to 16GB of on-package stacked RAM



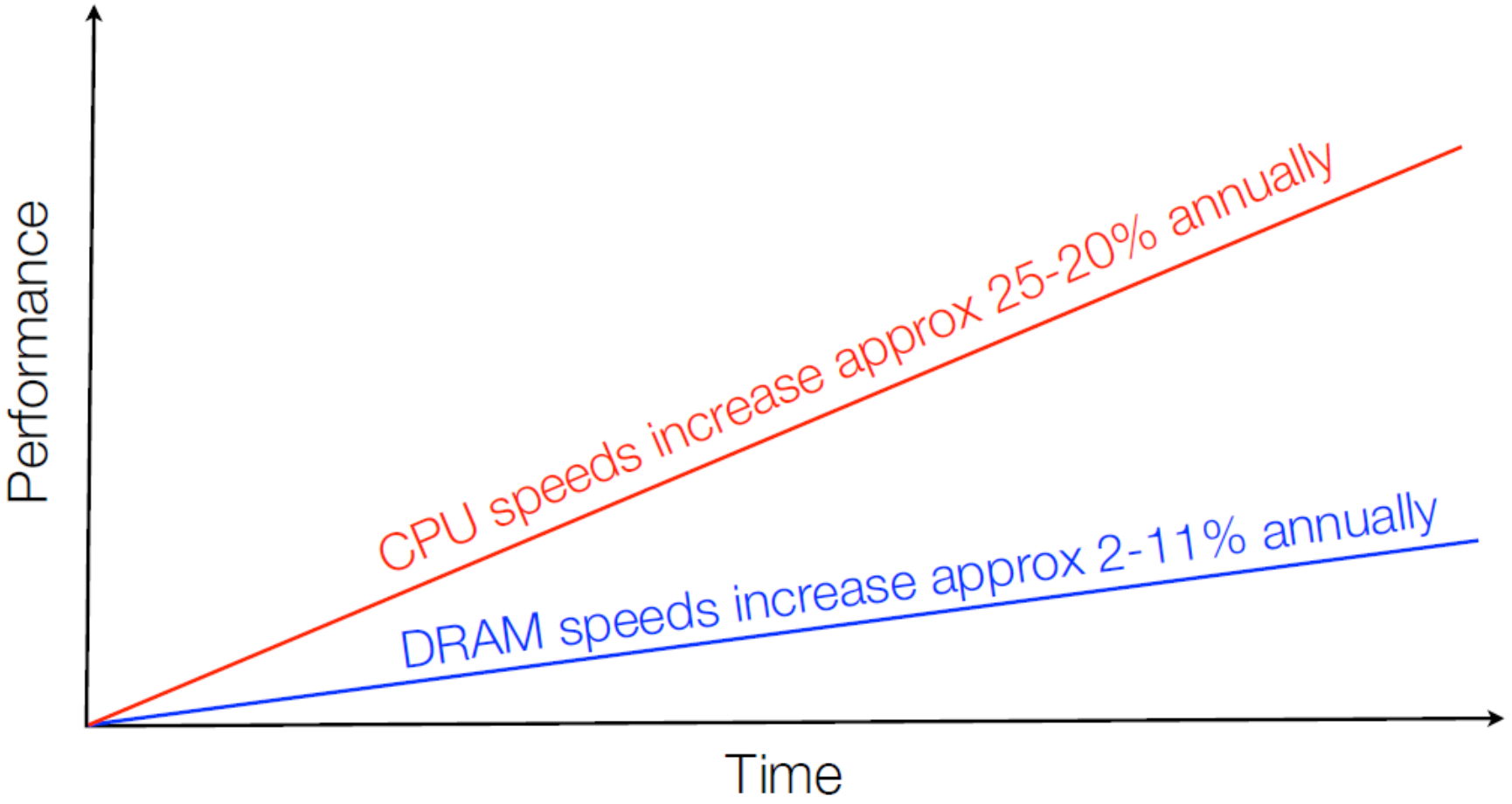


epcc

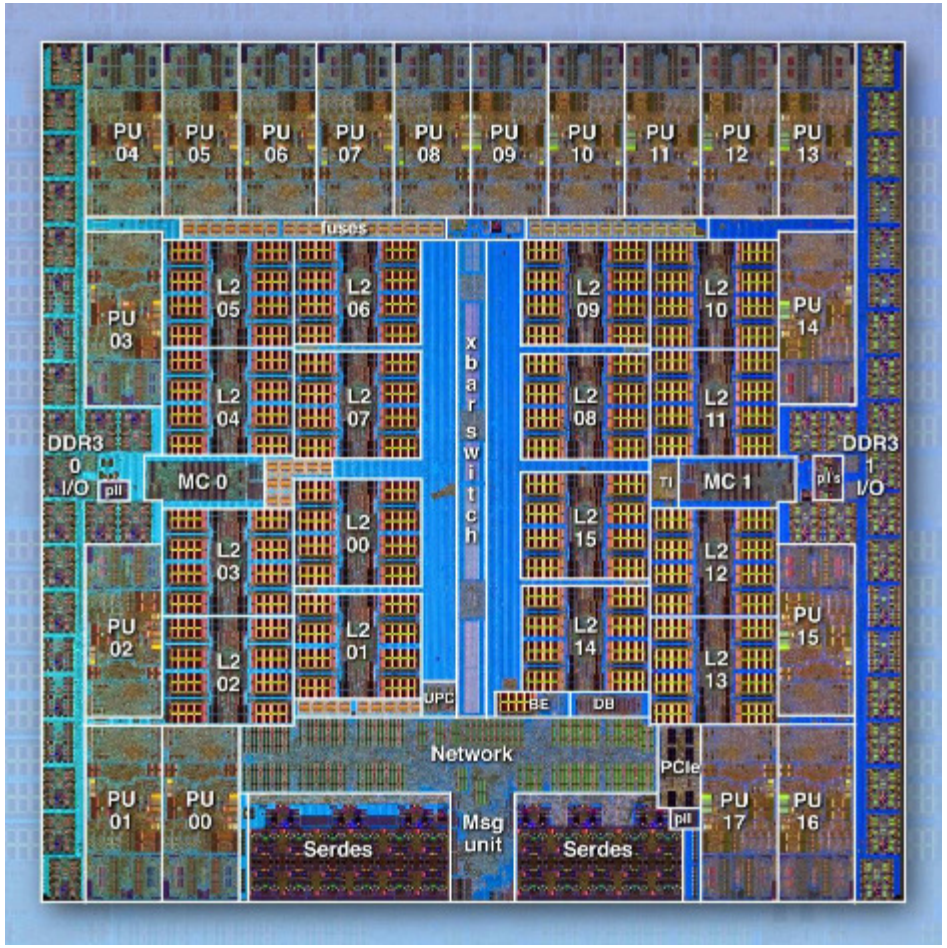
Processors

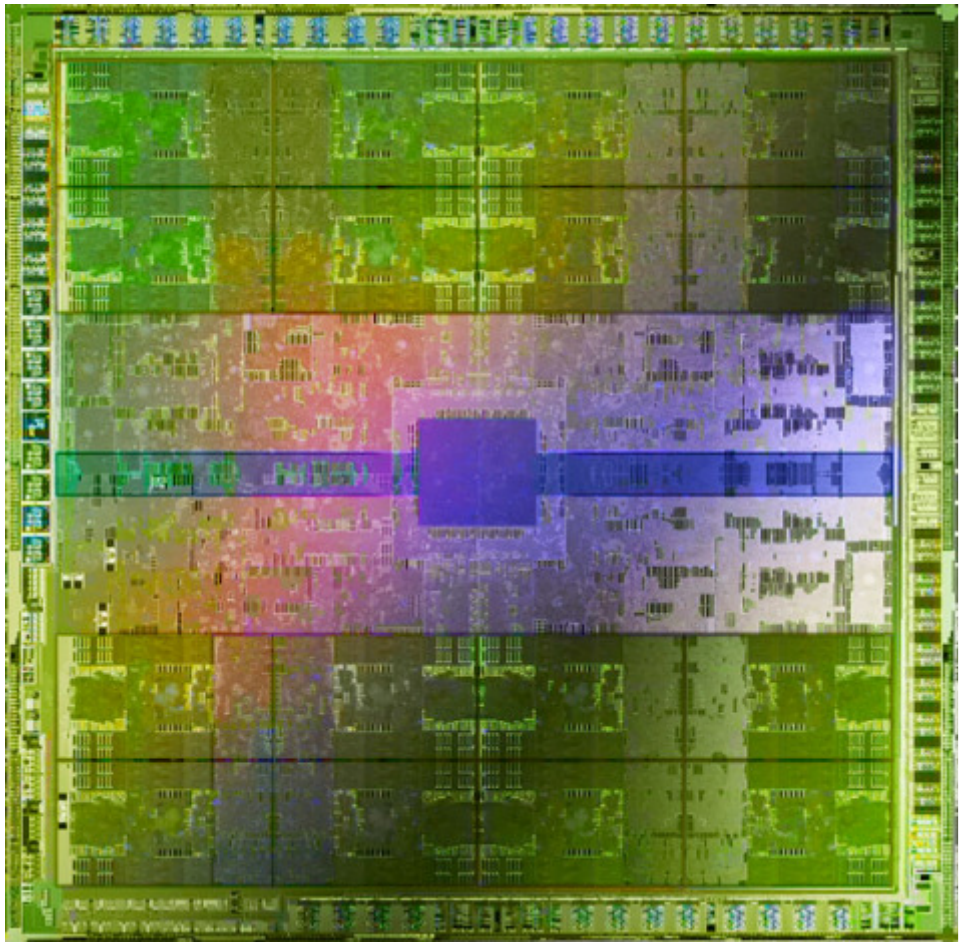
- The power used by a CPU core is proportional to Clock Frequency x Voltage^2
- In the past, computers got faster by increasing the frequency
  - Voltage was decreased to keep power reasonable.
- Now, voltage cannot be decreased any further
  - 1s and 0s in a system are represented by different voltages
  - Reducing overall voltage further would reduce this difference to a point where 0s and 1s cannot be properly distinguished
- Other performance issues too…
  - Capacitance increases with complexity
  - Speed of light, size of atoms, dissipation of heat
- And practical issues
  - Developing new chips is incredibly expensive
- Must make maximum use of existing technology
- Now parallelism explicit in chip design
  - Beyond implicit parallelism of pipelines, multi-issue and vector units

Performance Trend



Multicore processors





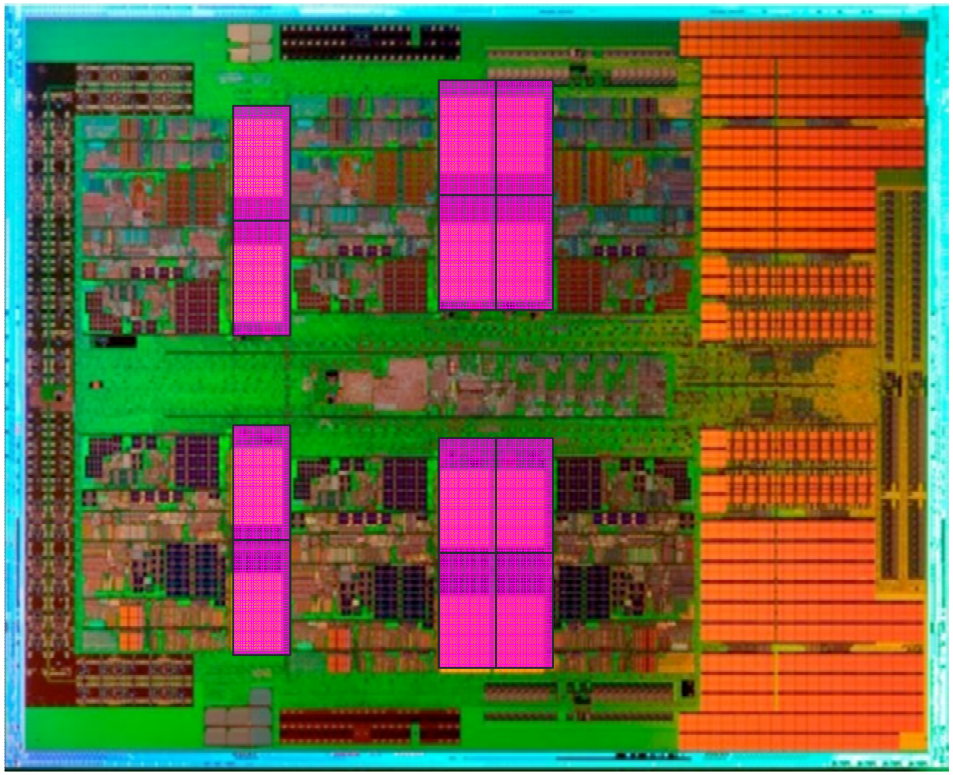
epcc

Accelerators

- Need a chip which can perform many parallel operations every clock cycle
  - Many cores and/or many operations per core
  - Floating Point operations (FLOPS) what is generally crucial for computational simulation
- Want to keep power/core as low as possible
- Much of the power expended by CPU cores is on functionality not generally that useful for scientific simulation
  - Branch prediction, out-of-order execution etc

AMD 12-core CPU

- Not much space on CPU is dedicated to compute



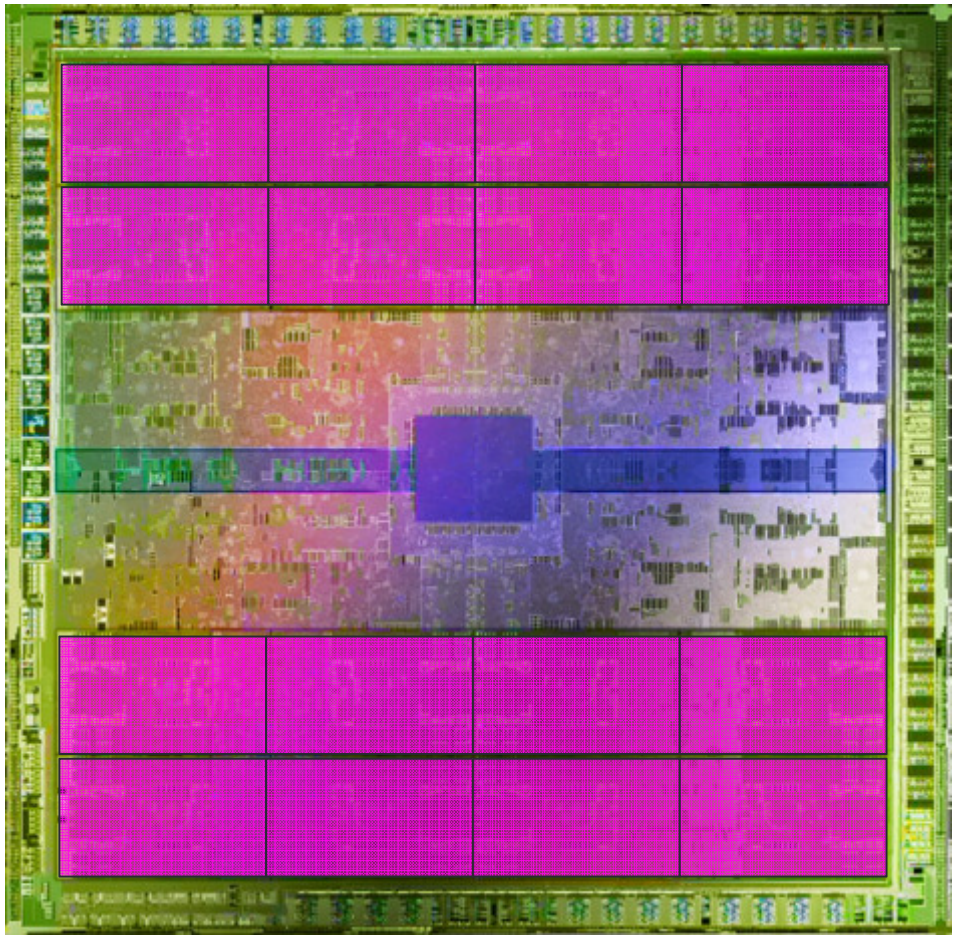
= compute unit
(= core)

epcc

Accelerators

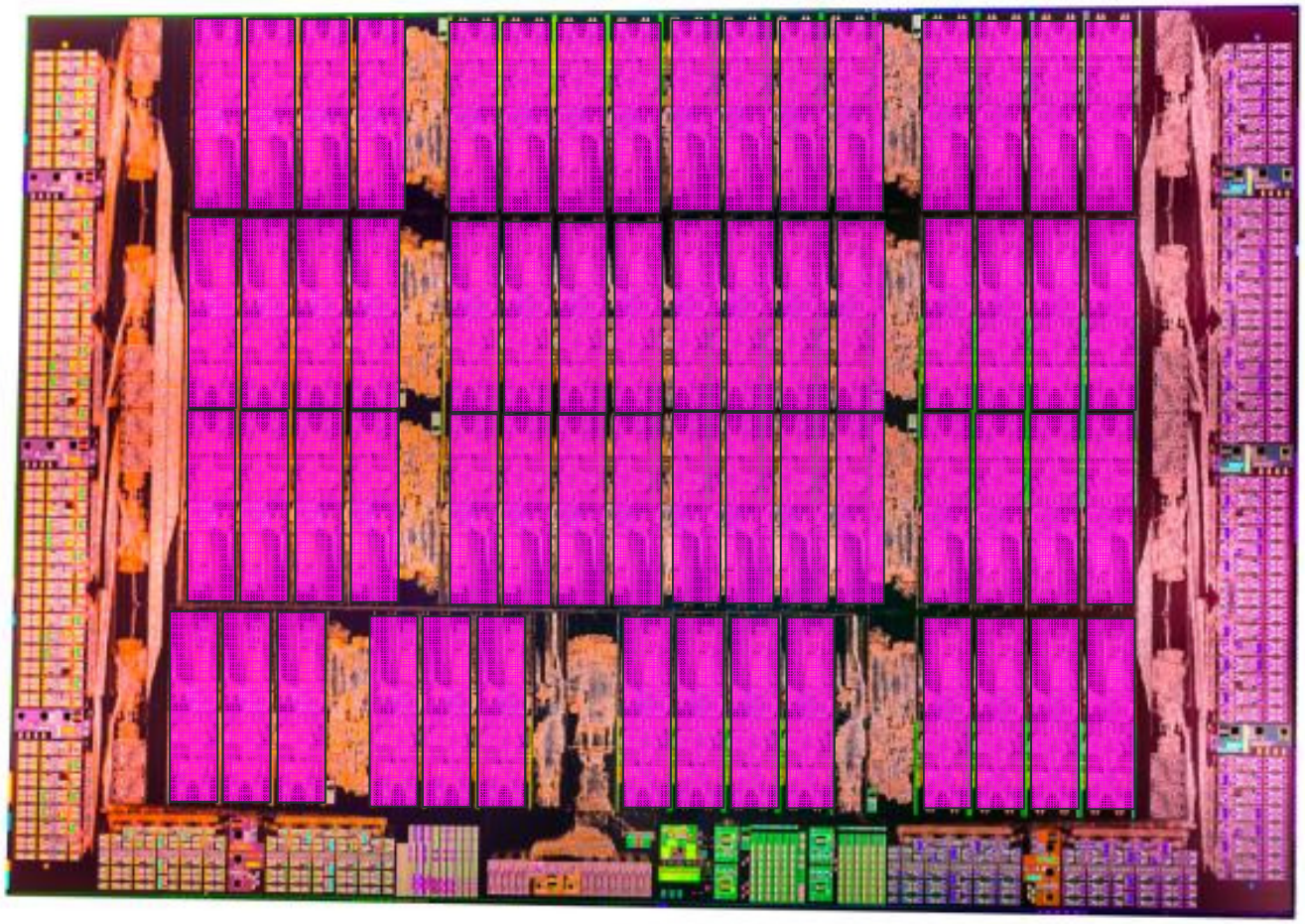
- So, for HPC, we want chips with simple, low power, number-crunching cores
- But we need our machine to do other things as well as the number crunching
  - Run an operating system, perform I/O, set up calculation etc
- Solution: “Hybrid” system containing both CPU and “accelerator” chips

NVIDIA Fermi GPU

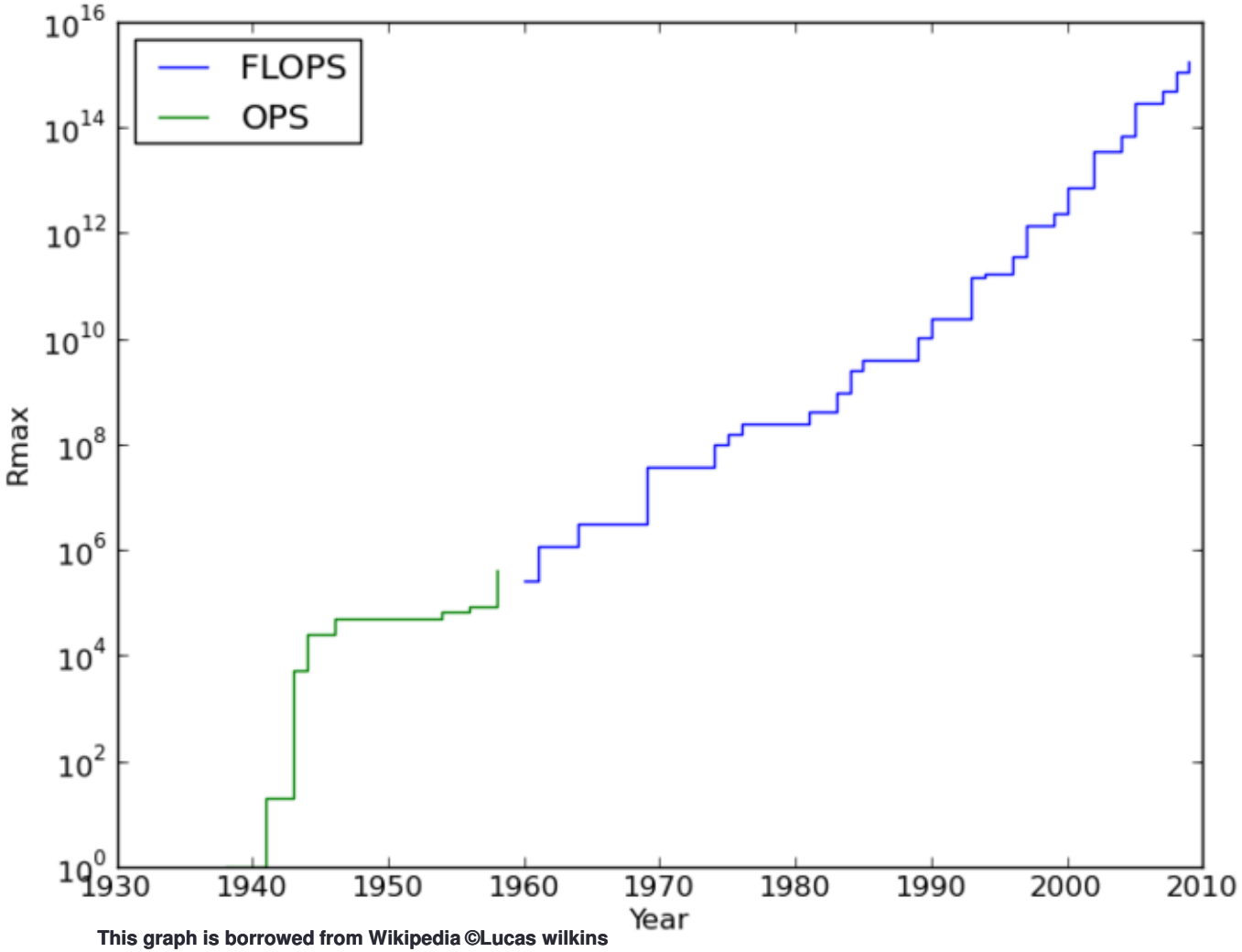


= compute unit
(= SM
= 32 CUDA cores)

Intel Xeon Phi



= compute unit (= core)



This graph is borrowed from Wikipedia ©Lucas wilkins

FLOPS

- Yotta: 10^{24}
- Zetta: 10^{21}
- Exa: 10^{18}
- Peta: 10^{15}
- Tera: 10^{12}
- Giga: 10^{9}
- Mega: 10^{6}
- Kilo: 10^{3}

epcc

Cray T3D

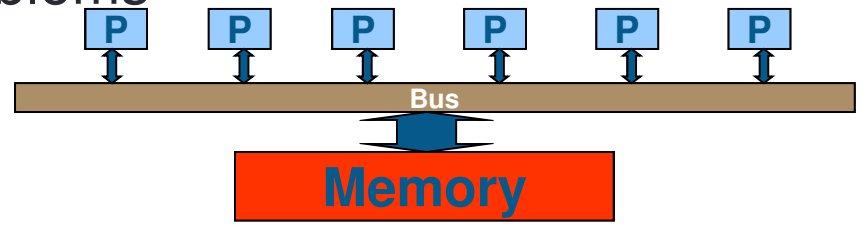
- First national parallel computing service
- EPCC: 1994-1999
  - 512 x 150 MHz EV5 Alphas processors
  - 76 GFlop/s peak
  - 32 GB memory
- Distributed memory system
- Supplemented with Cray T3E
  - 344 x 450 MHz EV56 Alpha processors
  - 310 GFlop/s peak, 40 GB memory

- **C**omputer **S**ervices for **A**cademic **R**esearch
- Manchester: 1998-2006
- Shared memory systems
  - Number of concurrent systems
- SGI Altix 3700
  - 512 Itanium 2 processors
  - 1Terabyte of memory
  - 2.7 TFlop/s peak
- SGI Origin 3800
  - 512 MIPS R12000 processors
  - 400 GFlop/s peak, 512 GB of memory.

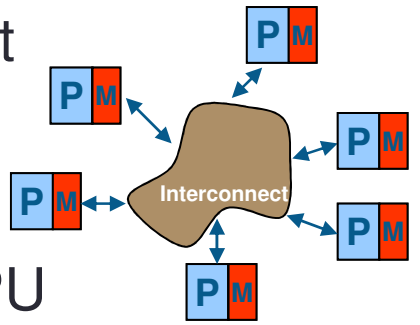
epcc

Types of HPC systems

- Shared-memory: OpenMP
  - Multiple processors share a single memory space
  - Simple to program for many problems
  - Scaling is problematic



- Distributed memory: MPI
  - Each processing unit has its own memory space
  - Excellent scaling properties
  - Can be more complex to program due to explicit communications



- Accelerators (GPU, Intel MIC)
  - Specialist processing units attached to main CPU
  - Can be difficult to extract good performance
  - (Conceptually similar to old vector architectures.)

epcc

HPCx
CAPABILITY COMPUTING

- The University of Edinburgh (EPCC), STFC (Daresbury), and IBM: 2002-2010

- Phase 1:
  - 80 nodes, 1280 processors
  - 16 x 1.3 GHz Power4
  - 6.7 TFlop/s peak
  - 1280 GB memory
- Phase 2:
  - 50 nodes, 1600 processors
  - 32 x 1.7 GHz Power4+
  - 10.8 TFlop/s peak
  - 1.6 TB memory
- Phase 3:
  - 160 nodes, 2560 cores
  - 8 x Dual core 1.6 GHz Power5
  - 15.36 TFlop/s peak, 5.12 TB memory

epcc

Distributed Shared Memory (clusters)

- Dominant architecture is a hybrid of these two approaches: *Distributed Shared Memory.*
  - Due to most HPC systems being built from commodity hardware – trend to multicore processors.
  - Each Shared memory block is known as a *node*.
  - Usually 16-64 processors per node.
  - Nodes can also contain accelerators.
- Majority of users try to exploit in the same way as for a purely distributed machine
  - As the number of cores per node increases this can become increasingly inefficient…
  - …and programming for these machines can become increasingly complex

HECToR

epcc

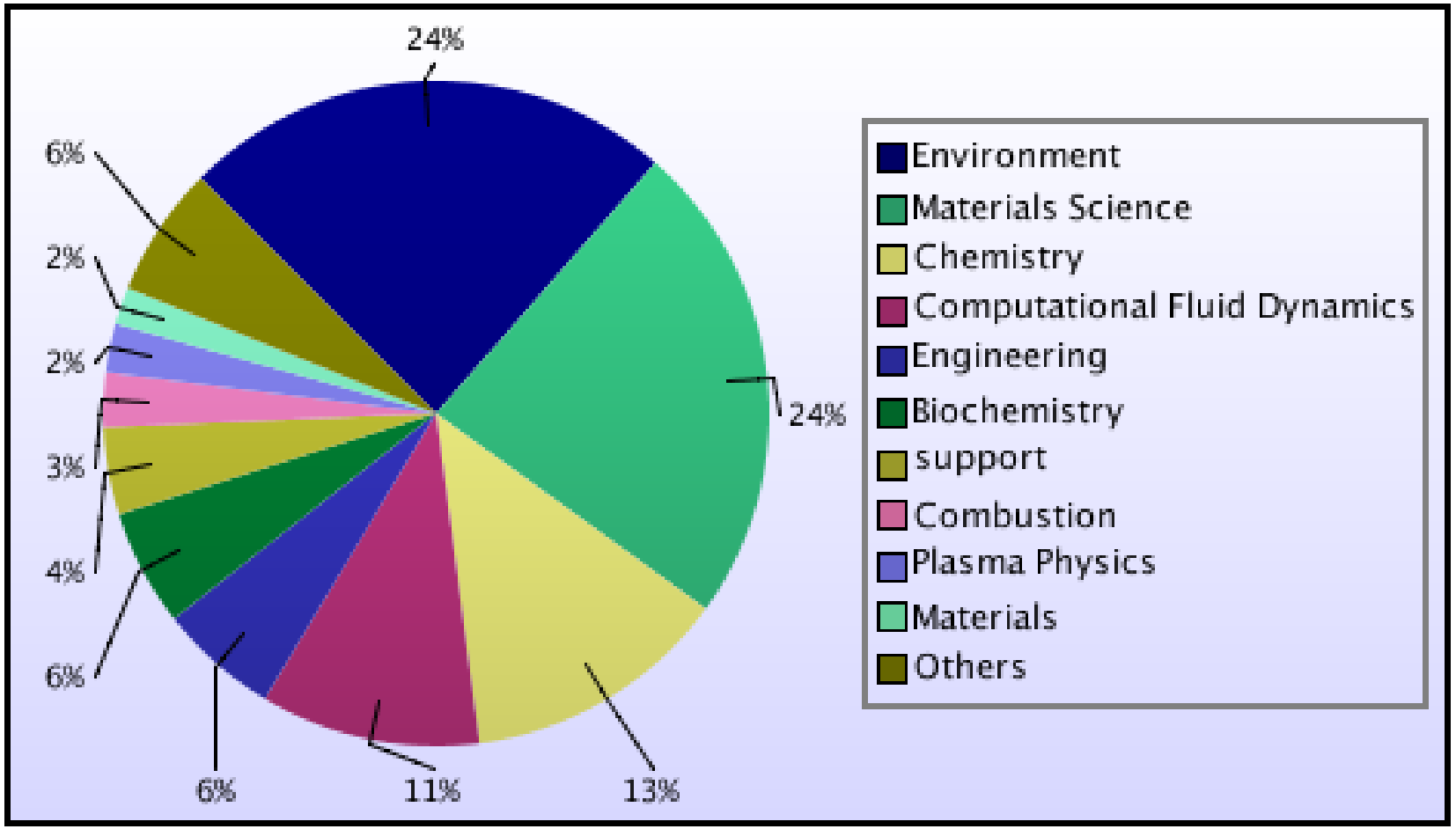
HECToR

- Current national service: 2007 – 2014
- Phase 1:
  - Cray XT4
  - 1416 nodes, 11,328 cores
  - 4 x 2.8 GHz dual-core Opterons
  - 63.4 TFlop/s peak
  - 34 TB memory
- Phase 2a:
  - 1416 nodes, 22,656 cores
  - 4 x 2.3 GHz quad-core Opterons
  - 208 TFlop/s peak
  - 45.3 TB memory
- Phase 2b:
  - Cray XT6
  - 1856 nodes, 44,544 cores
  - 2 x 2.1 GHz 12-core Opterons
  - 373 TFlop/s peak
  - 58 TB memory
- Phase 3:
  - Cray XE6
  - 2816 nodes, 90,122 cores
  - 2 x 2.3 GHz 16-core Opterons
  - > 800 TFlop/s peak
  - 90 TB memory

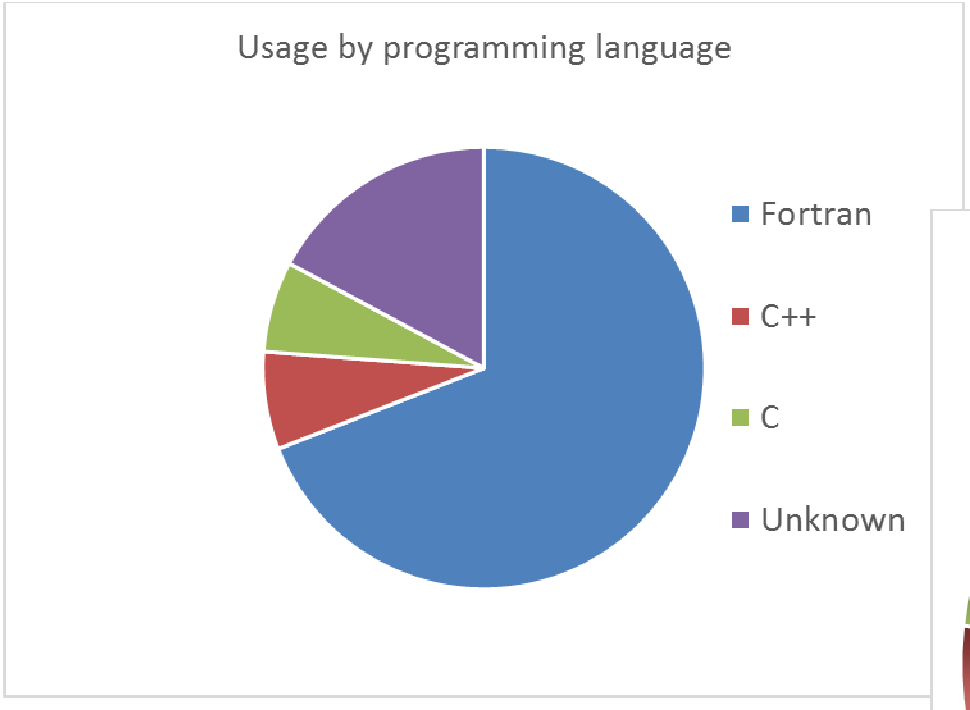
ARCHER

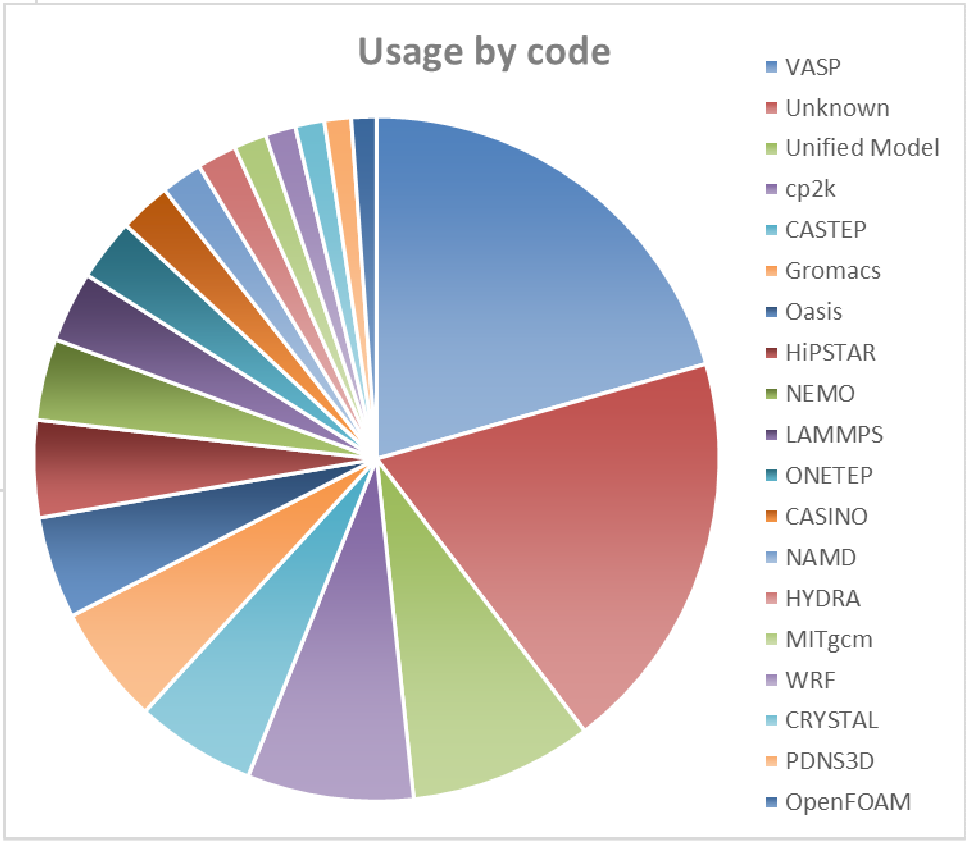
- UK national HPC for Engineering, Physical Sciences, and Natural Sciences
- Cray XC30 Hardware
  - Nodes based on 2×Intel Ivy Bridge 12-core processors
  - 64GB (or 128GB) memory per node
  - 4920 nodes in total (118,080 cores)
  - Linked by Cray Aries interconnect (dragonfly topology)
  - 2.6 Pflop/s

What is it used for?



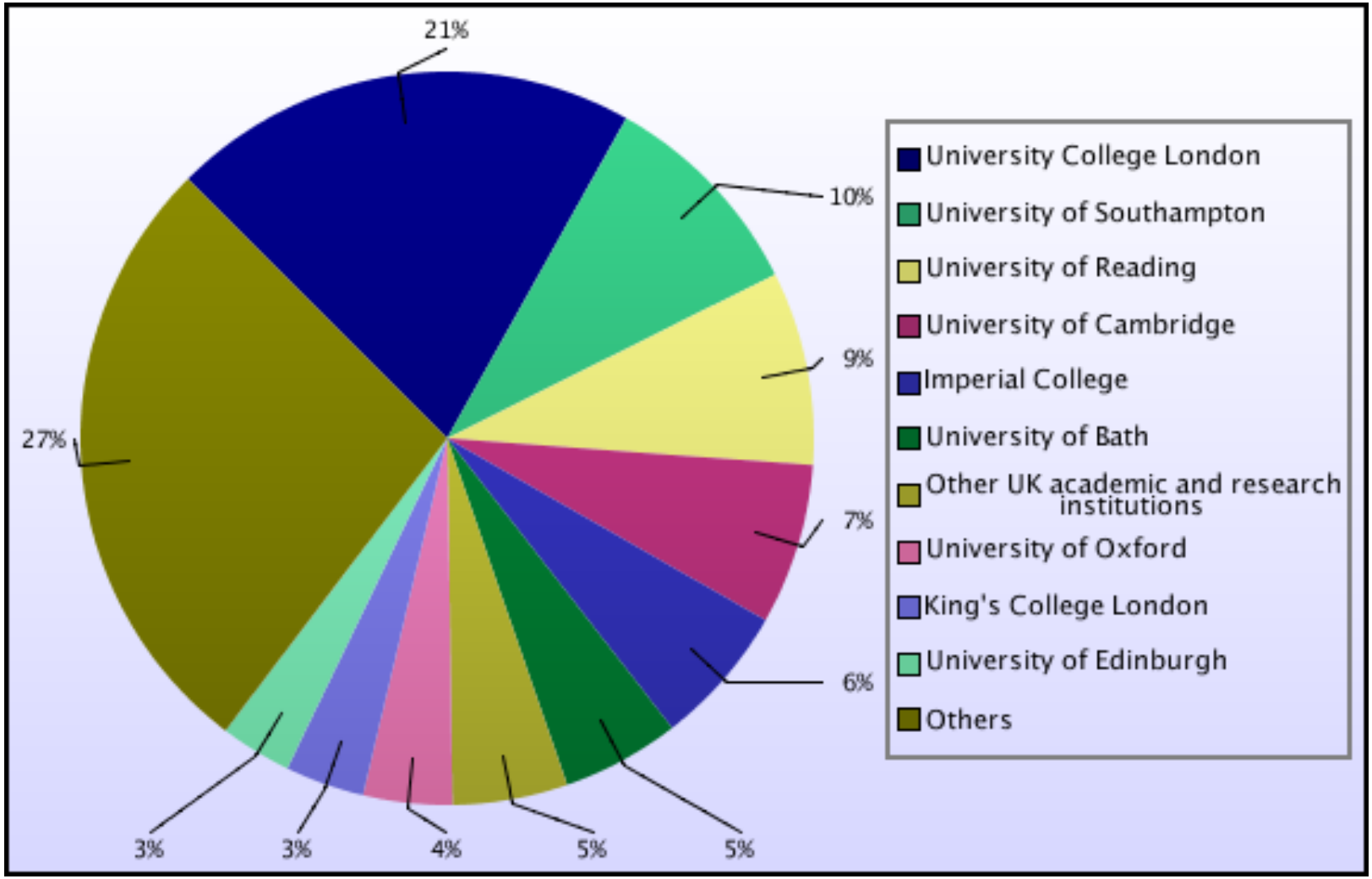
Usage by codes





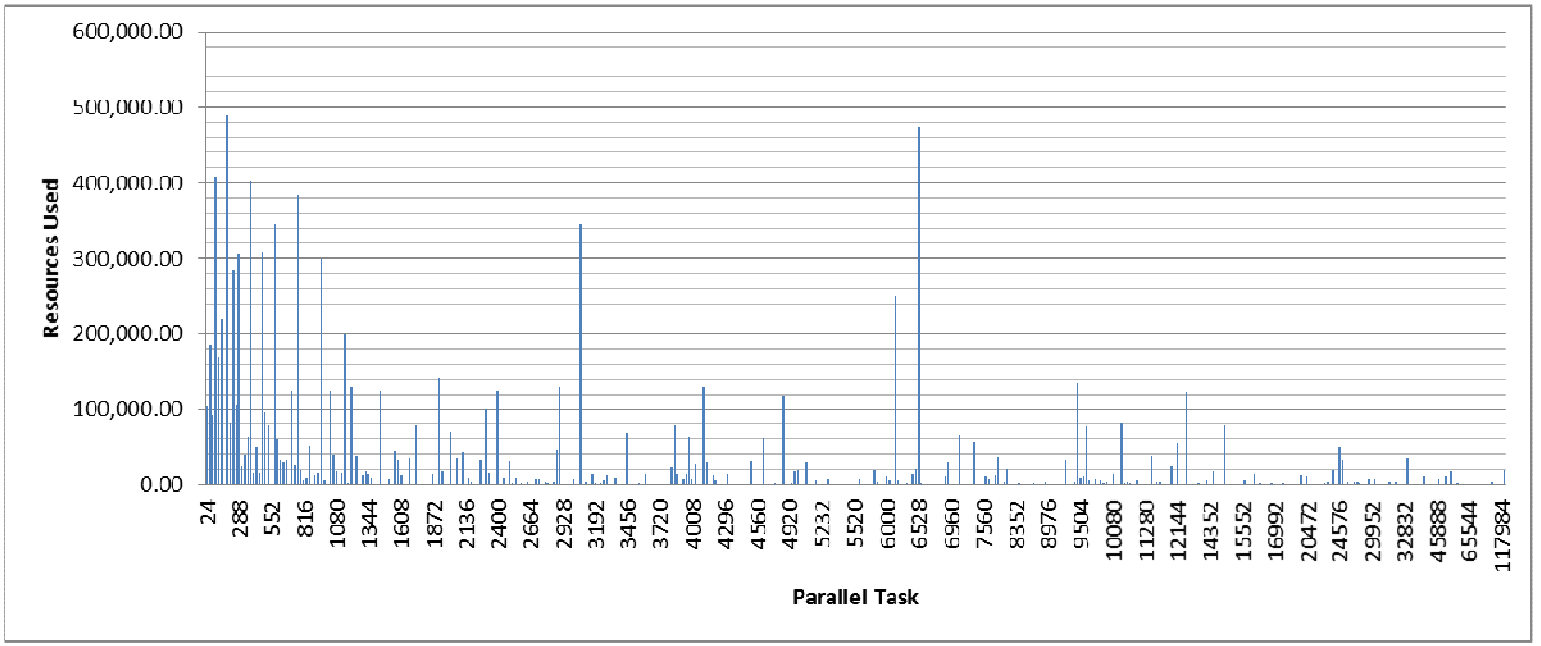
epcc

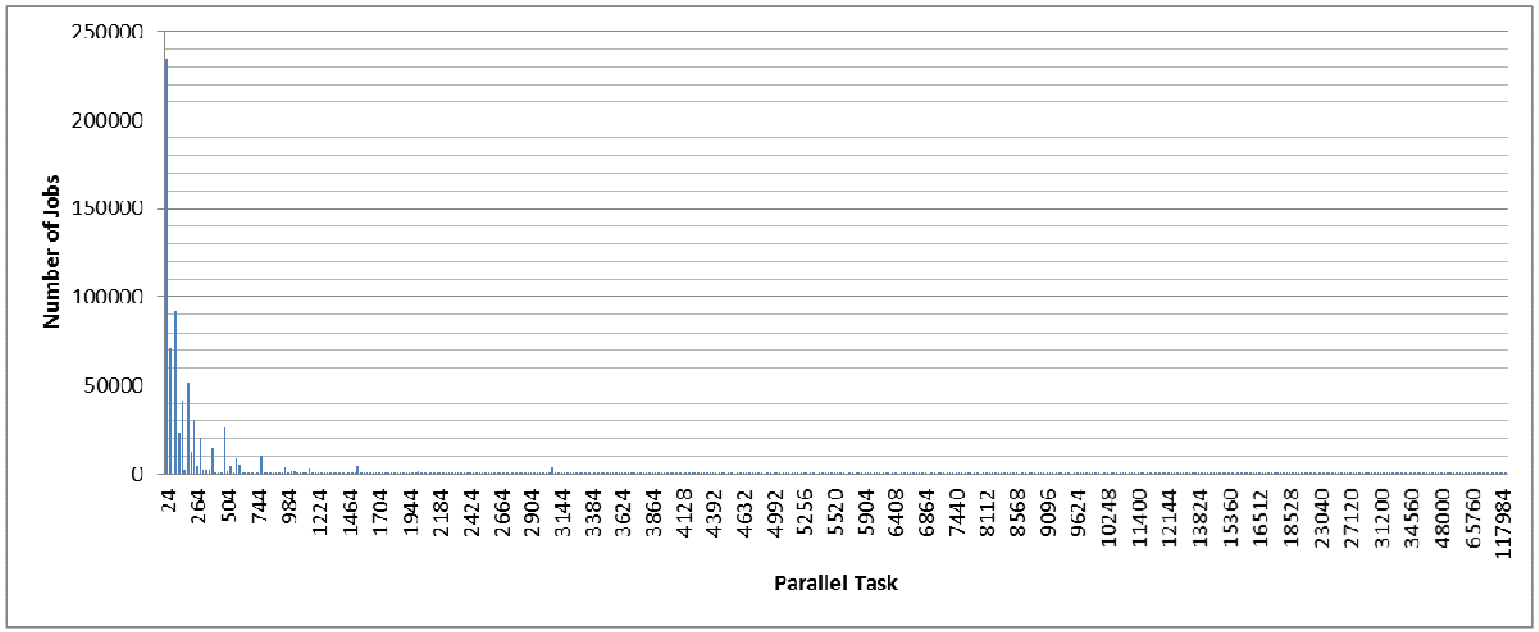
Institutional use



epcc

Machine Usage

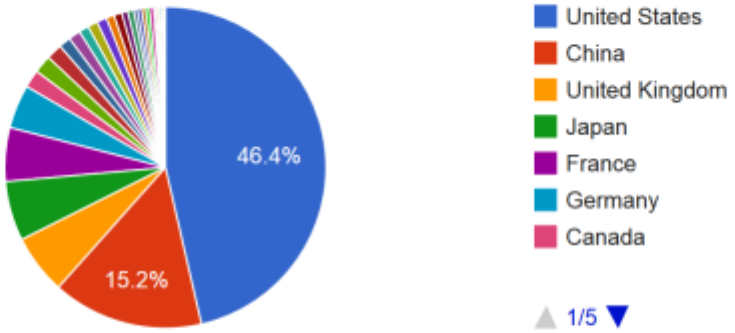




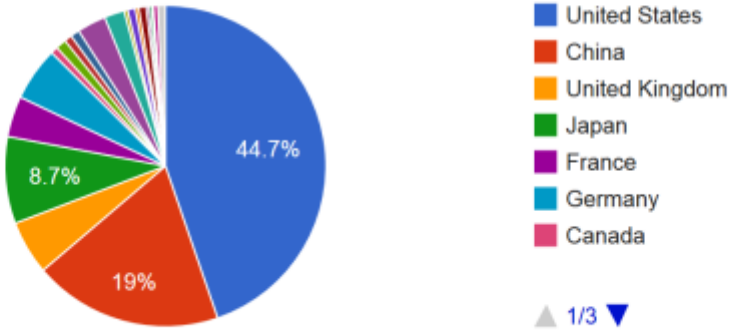
epcc

International Aspect

Countries System Share



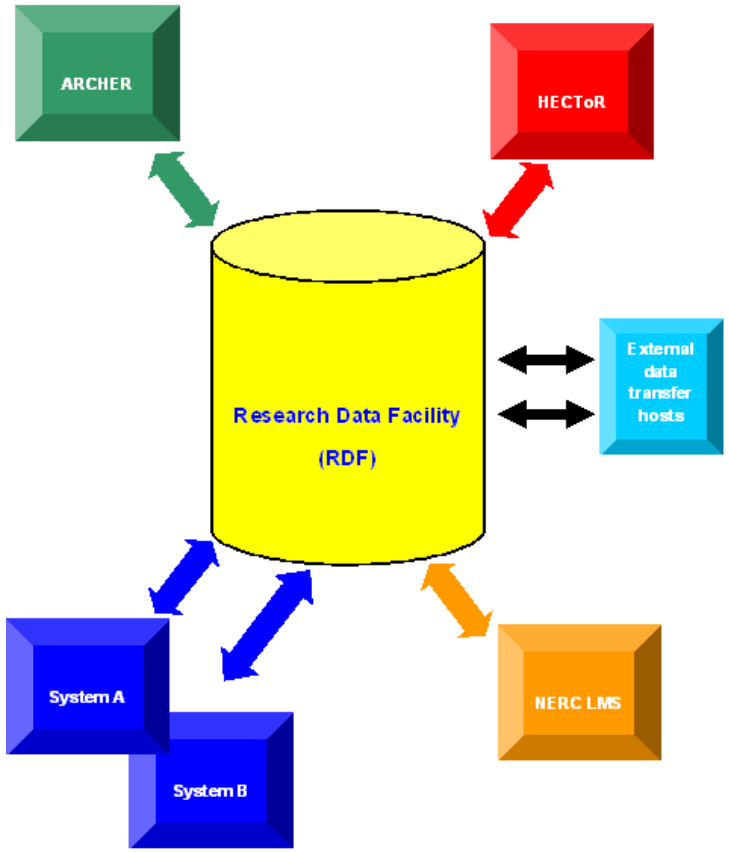
Countries Performance Share



epcc

Research Data Facility (RDF)

- RDF is designed for long term data storage
  - Connected to ARCHER
- RDF consists of
  - 13.8PB disk
  - 19.5 PB backup tape
  - Provide a high capacity robust file store;
  - Persistent infrastructure - will last beyond any one national service



epcc

Differences from Cloud computing

- Performance
  - Clouds usually use virtual machines which add an extra layer of software.
  - In cloud you often share hardware resource with other users – HPC access is usually exclusive.
- Tight-coupling
  - HPC parallel programming usually assumes that the separate processes are tightly coupled
  - Requires a low-latency, high-bandwidth communication system between tasks
  - Cloud usually does not usually have this
- Programming models
  - HPC use high-level compiled languages with extensive optimisation.
  - Cloud often based on interpreted/JIT.

Getting access to ARCHER

- Standard research grant
  - Request Technical Assessment using form on ARCHER website
  - Submit completed TA with notional cost in J-eS
  - Apply for time for maximum of 2 years
- ARCHER Resource Allocation Panel (RAP)
  - Request Technical Assessment using form on ARCHER website
  - Submit completed TA with RAP form
  - Application for computer time only
- Instant Access – Pump-Priming Time
  - Request Technical Assessment using form on ARCHER website
  - Submit completed TA with 2 page description of work
  - Office decision on application

Consortia

Consortium	URL	PI
Materials Chemistry Consortium	http://www.ucl.ac.uk/klmc/mcc/	Richard Catlow, Scott Woodley, UCL
UK Car-Parrinello Consortium	http://www.cse.scitech.ac.uk/cmg/NETWORKS/UKCP/	Matt Probert, York
UK Plasma Physics Consortium	http://www.ccpp.ac.uk/hec/index.html	Tony Arber, Warwick
Consortium for Biomolecular Simulation		Adrian Mulholland, Bristol
UK Turbulence Consortium	http://www.turbulence.ac.uk	Richard Sandberg, Southampton
UK Consortium on Mesoscale Engineering Sciences		Kai Luo, UCL
National Oceanography Centre	http://noc.ac.uk/	Andrew Coward, Southampton
National Centre for Atmospheric Science	http://www.ncas.ac.uk/index.php/en/	Grenville Lister, Reading
Computational Mineral Physics Consortium		John Brodholt, UCL

What now?

- You can attempt the ARCHER driving test
  - www.archer.ac.uk/training/course-material/online/driving_test.php
- On successful completion, eligible users can apply for
  - account on ARCHER
  - 1,200 kAUs of time (80,000 core-hours) over 12 months
- Further information
  - Helpdesk: support@archer.ac.uk
  - Training: http://www.archer.ac.uk/training/.
- IPCC
  - https://www.epcc.ed.ac.uk/research/computing/intel-parallel-computing-centre

epcc

Women in HPC

- http://www.womeninhpc.org.uk

Software sustainability institute

- http://www.software.ac.uk/