

Trends in High Performance Computing

Simon McIntosh-Smith

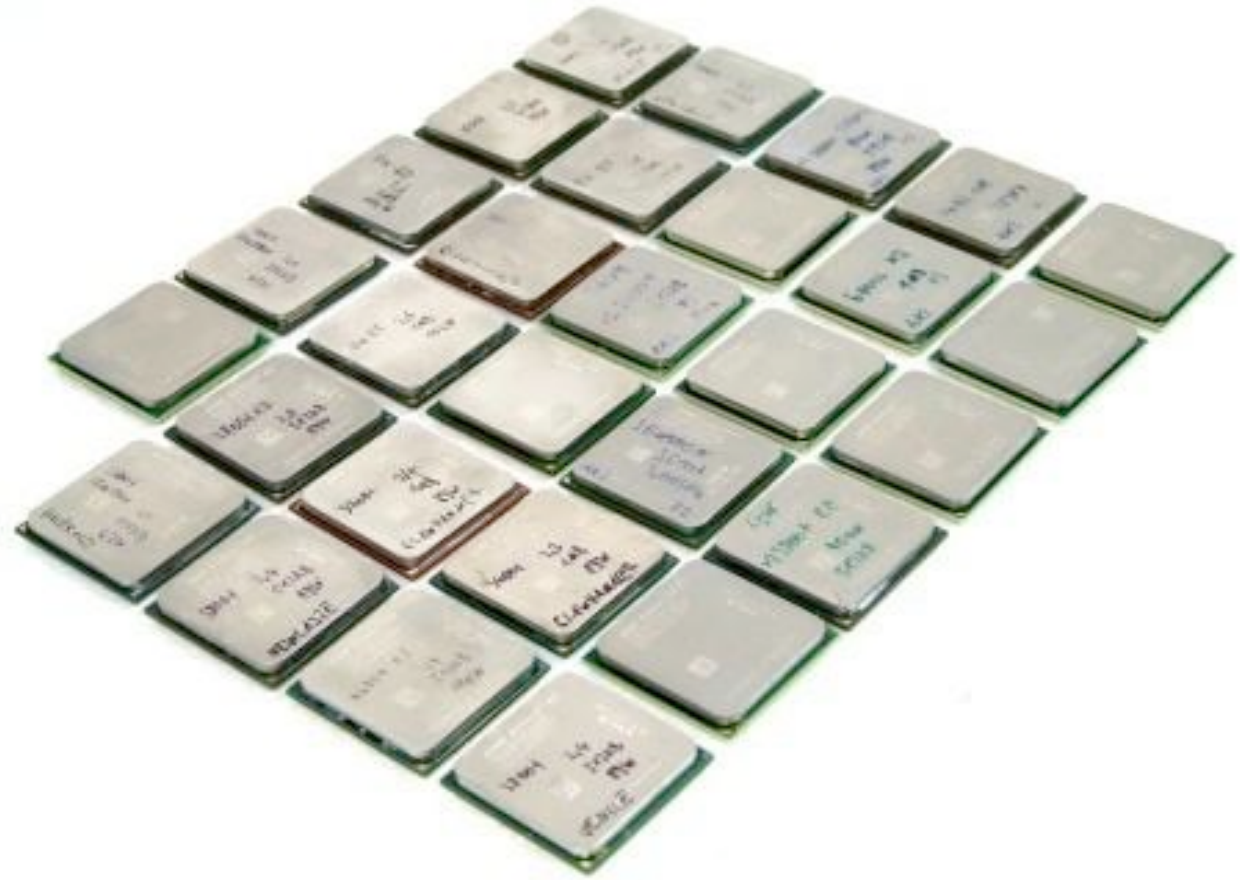
HPC Symposium

21st October 2009

University of BRISTOL

Agenda

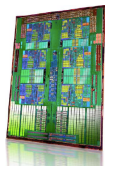
- The latest High Performance Computing trends in
  - Processors
  - Interconnect
  - Storage
- Exascale (10^{18}) computing
- Wrap-up

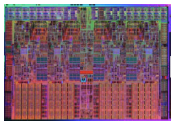


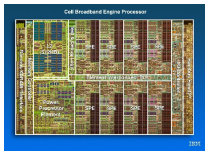
Trends in processors

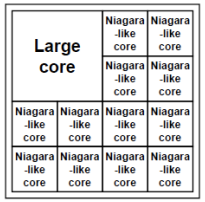
Trends in Processors

- Consolidation
  - Sun, SiCortex, ClearSpeed, …
- Power consumption
  - Keeps going up!
    - HP predicting racks up to 50kW
  - Will require water cooling at multiple levels
- Quad core CPUs already the norm, six and eight cores now appearing
- Multiple heterogeneous systems now in Top500
  - Intel, AMD, IBM all heading in this direction





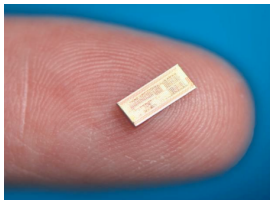




University of BRISTOL

Trends in Processors

- Trend towards 2U servers to reduce power consumption, cost and weight
- First signs of consumer computing as an alternate path to commodity computing
  - CMU's "FAWN: Fast Array of Wimpy Nodes"
    "Links together a large number of tiny nodes built using embedded processors and small amounts (2-16GB) of flash memory into an ensemble capable of handling 1300 queries per second per node, while consuming fewer than 4 watts of power per node."

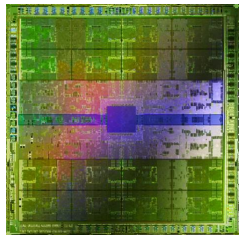


- Hardware faults will become more prevalent as systems scale up in core counts

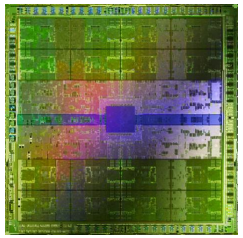
GPUs become HPC monsters

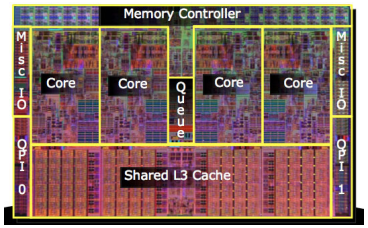
- GPUs have been becoming more HPC-friendly
  - OpenCL, 64-bit capable, passive cooling, IEEE754
  - Systems vendors providing GPU-capable servers

- Nvidia's "Fermi" architecture, announced Sep '09 kicks this up several gears:
  - Adds full IEEE 754-2008 double precision
  - Adds hardware reliability features (ECC on all memories)
  - Adds a single, flat address space, enabling C++
  - Many features specifically for HPC



Comparing Fermi and Nehalem

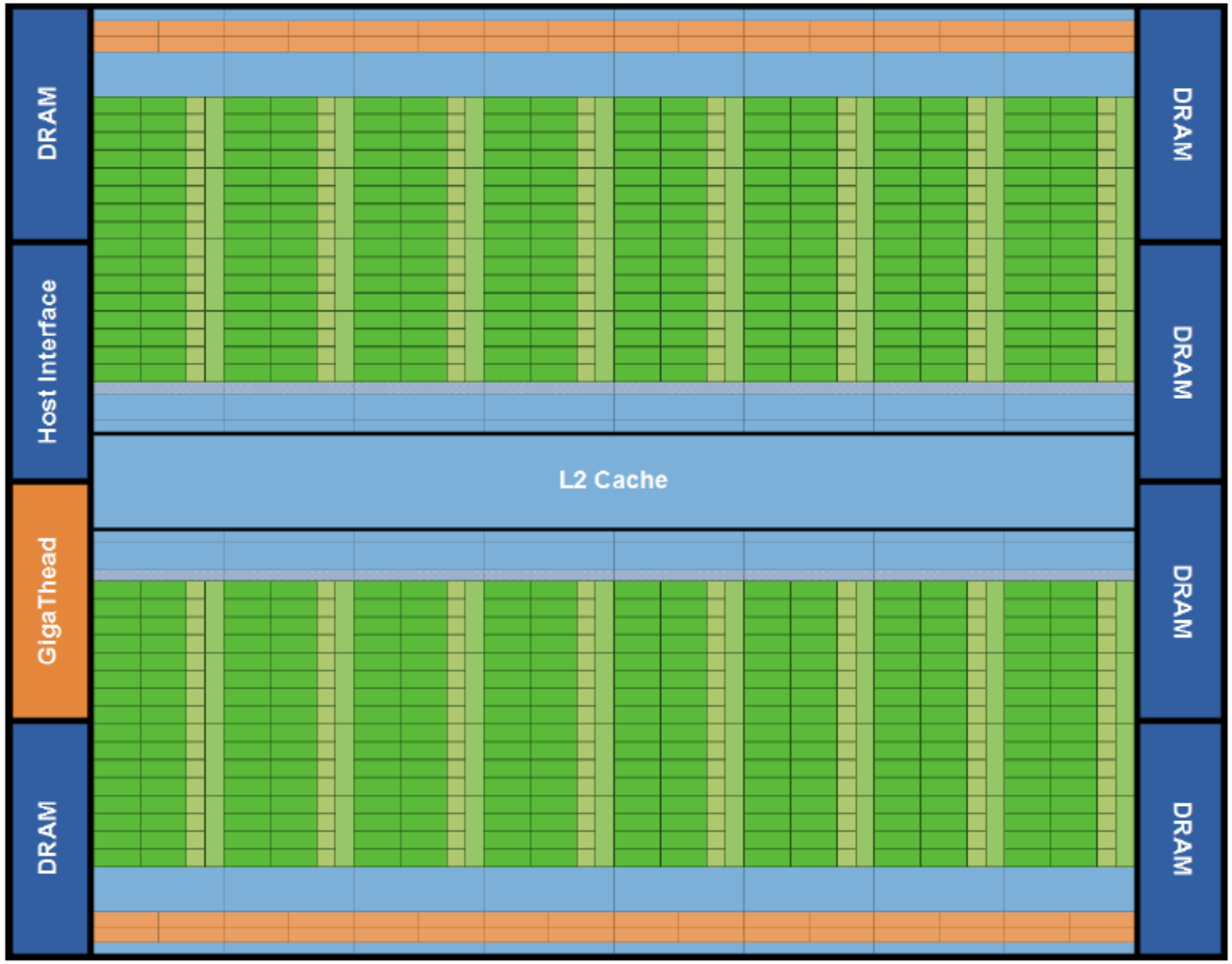




- 512 simple cores
- ~3 billion transistors
- ~1.5GHz
- 1,500 GFLOPS S.P.
- 750 GFLOPS D.P.
- ~190 GBytes/s
- IEEE 754-2008 support
- ECC on all memories

- 4 complex cores
- 731 million transistors
- ~3GHz
- 96 GFLOPS S.P.
- 48 GFLOPS D.P.
- ~30 GBytes/s
- IEEE 754-1985 support
- ECC on all memories

Fermi's major elements



Source: NVIDIA

16 Streaming Multiprocessors (SMs)

Each SM contains 32 CUDA cores (512 in all)

GigaThread scheduler for parallel kernels

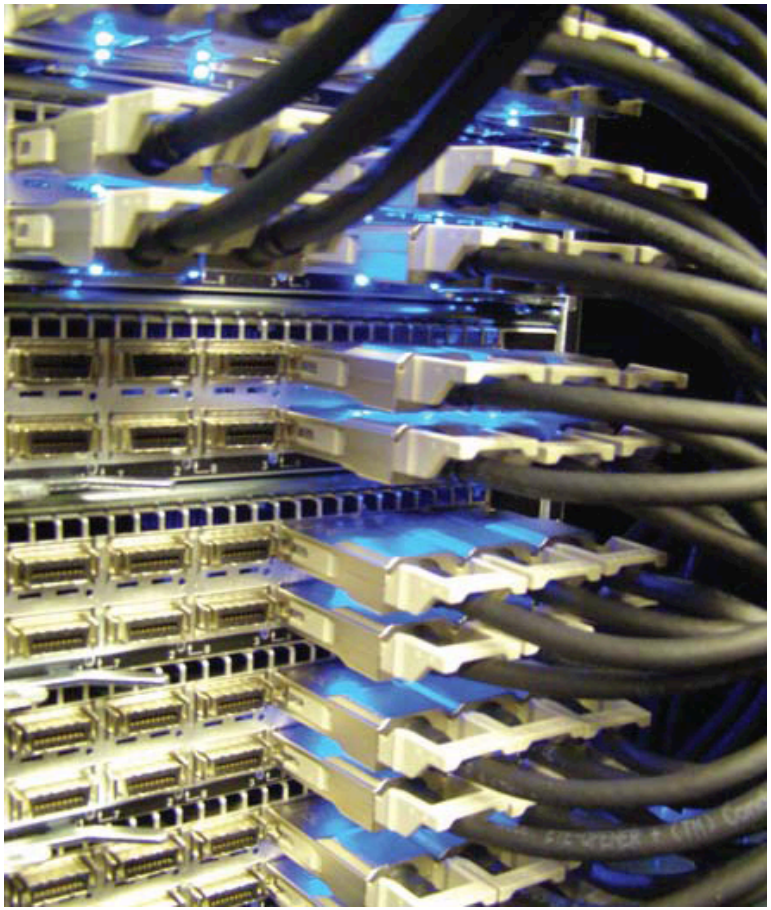
6 GDDR5 memory controllers: ~190 GB/s

New L1 / L2 caches for easier programming

University of BRISTOL

GPU rate of improvements

	2006	2008	2010
GPU	**G80**	**GT200**	**Fermi**
Transistors	681 million	1.4 billion	3.0 billion
CUDA Cores	128	240	512
Double Precision Floating Point Capability	None	30 FMA ops / clock	256 FMA ops /clock
Single Precision Floating Point Capability	128 MAD ops/clock	240 MAD ops / clock	512 FMA ops /clock
Warp schedulers (per SM)	1	1	2
Special Function Units (SFUs) / SM	2	2	4
Shared Memory (per SM)	16 KB	16 KB	Configurable 48 KB or 16 KB
L1 Cache (per SM)	None	None	Configurable 16 KB or 48 KB
L2 Cache (per SM)	None	None	768 KB
ECC Memory Support	No	No	Yes
Concurrent Kernels	No	No	Up to 16
Load/Store Address Width	32-bit	32-bit	64-bit



Trends in interconnect

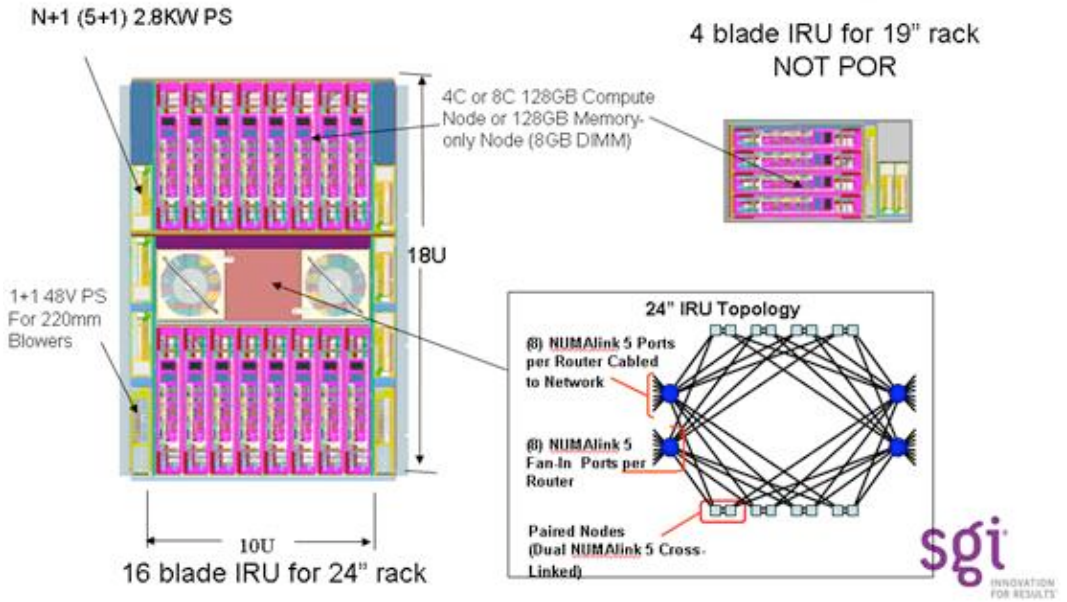
Trends in Interconnect

- Consolidation
  - Quadrics folding, proprietary interconnects disappearing
  - Top 100 mostly InfiniBand (40Gbps being deployed)
  - Bottom 400 mostly Gbit Ethernet
- Disruption?
  - Mass market is >> HPC and dominated by Ethernet
  - InfiniBand currently delivers better latency and bandwidth than Ethernet but at higher cost
  - Opportunity for disruption: better than IB latency and bandwidth via Ethernet?

www.gnodal.com

Trends in Interconnect

- Acceleration in the interconnect
  - E.g. sgi's UltraViolet technology to accelerate MPI

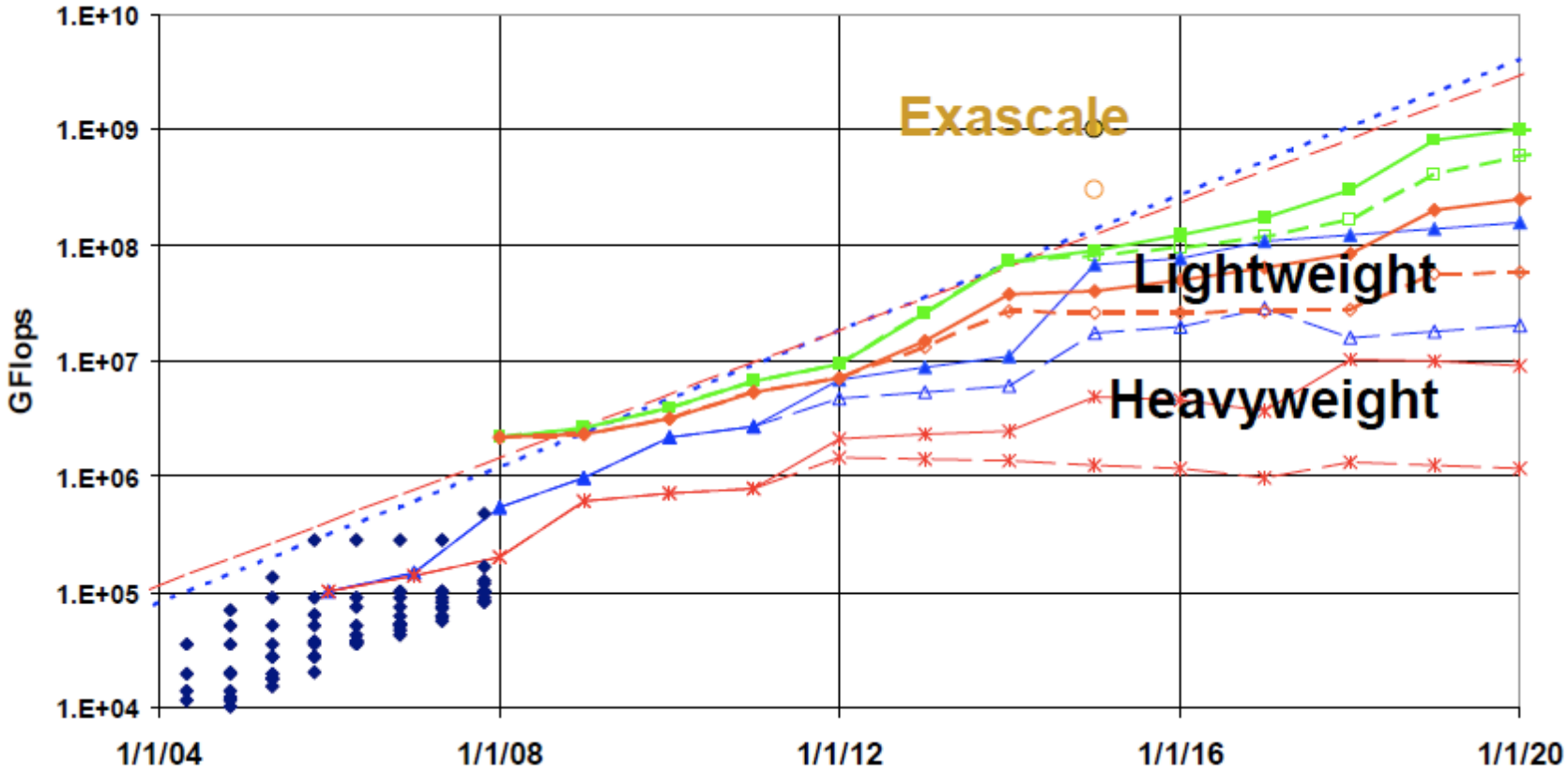


- But (current) MPI won't scale forever

Trends in storage

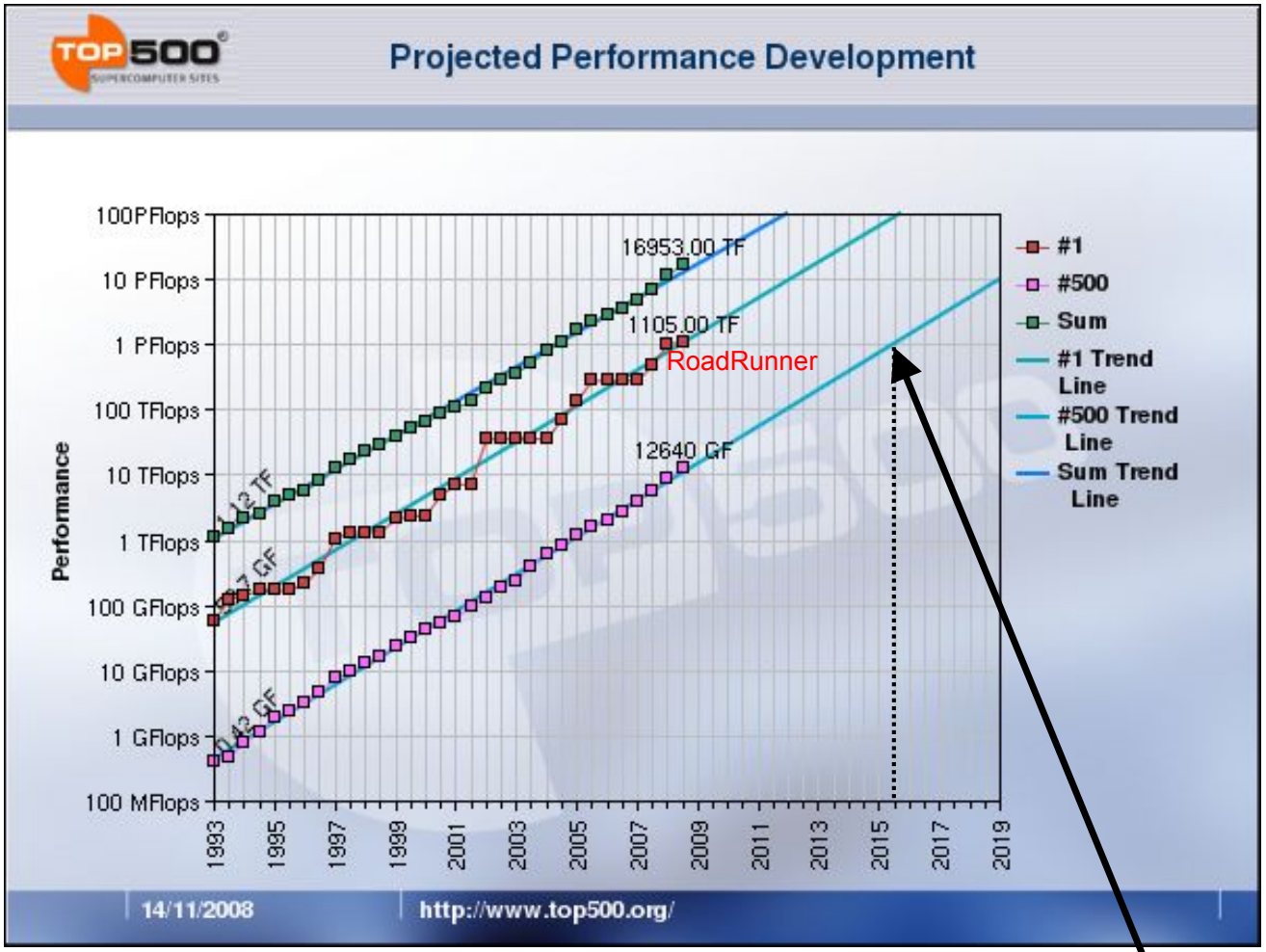
Trends in Storage

- Solid State Disks (SSDs) becoming ubiquitous
- Currently ~10X less capacity but ~3X lower power consumption and 4X smaller form factor
- Random read performance already excellent – future SSDs may be more like DRAM than disks
- Will supplement traditional hard drives in the short term
  - Metadata servers, IOP monsters
- Their greater reliability and lower power consumption will be essential for future, large scale storage solutions
- ASPs on flash memory reducing by nearly 50% per year per GigaByte



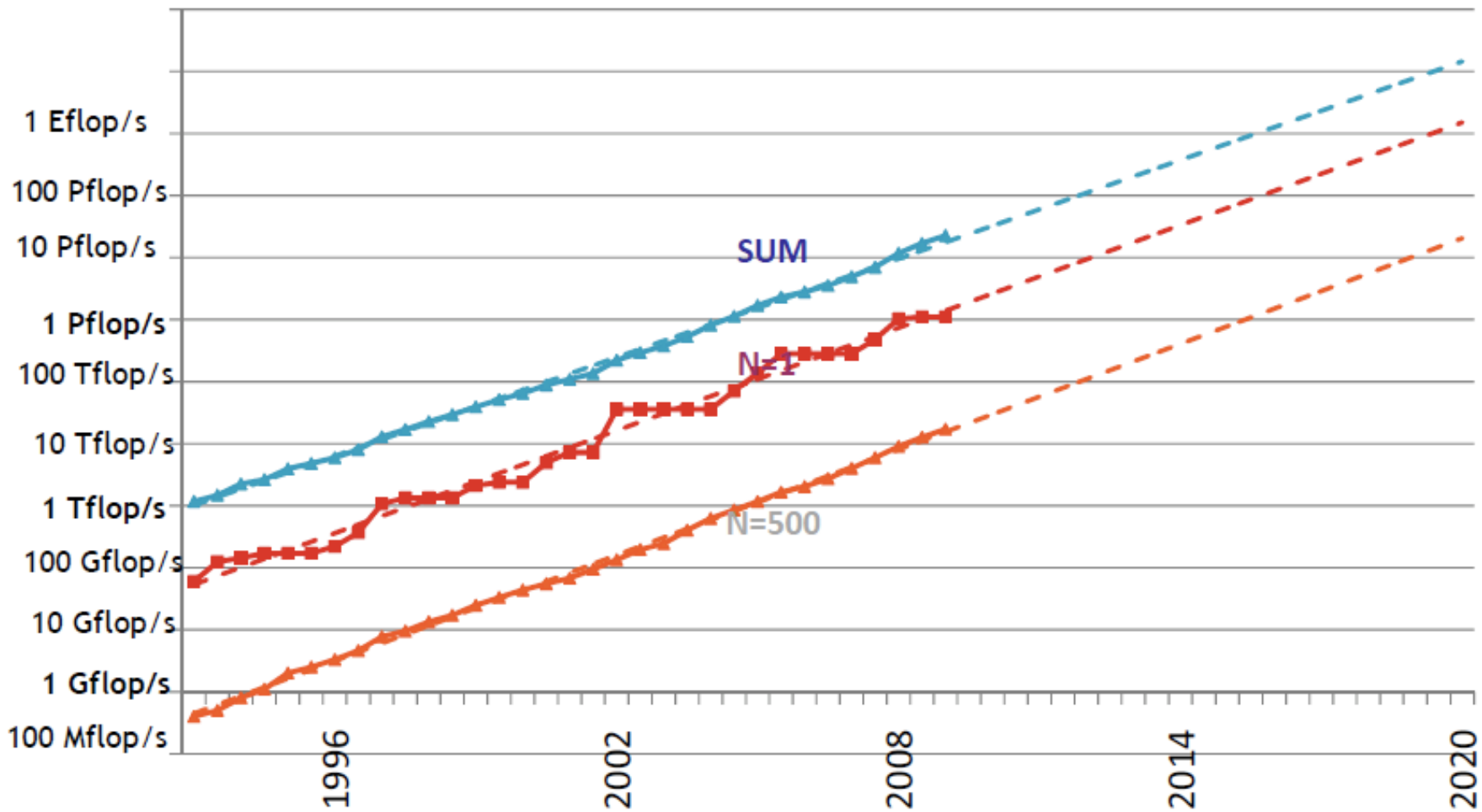
The path to Exaflops

On the way to Exaflops



- Within 6 years *everything* will be Petascale!

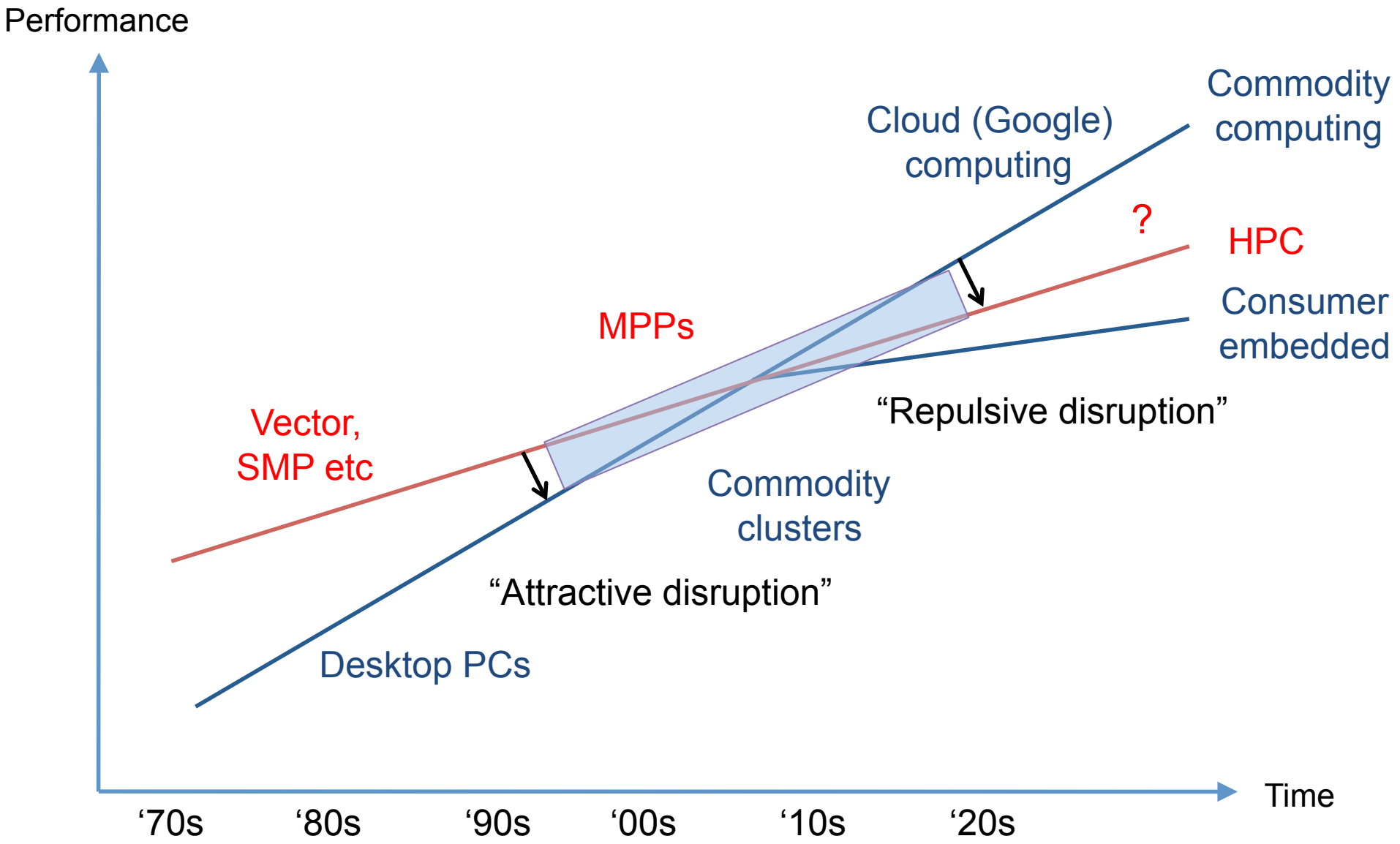
Performance Development



TOP500 SUPERCOMPUTER SITES

University of BRISTOL

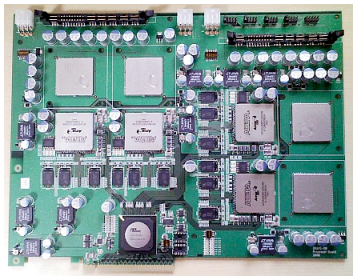
A parting of the ways?



Onwards and upwards

The path to Exascale:

- General purpose to give way to specialist machines (again)?
  - Grape-DR
  - D.E. Shaw's Anton
  - Cray, sgi, NEC, …



- The predicted continued growth in data centre size and power consumption must surely hit a wall?
  - In the June 2009 Green Top500, average machine efficiency increased by 10% while average power consumption increased by 15%
  - Lowest power consumption prediction for first Exaflop machine is *50 MW!*

Summary

- Processors: heterogeneous many cores, FAWN
- Interconnect: IB & Ethernet, time for disruption?
- Storage: SSDs, but where do they fit in the memory hiearchy?
- Systems: millions of cores, heterogeneous, unreliable

The future is predictably massively parallel

HPC@Bristol

University of BRISTOL