HPC future trends from a science perspective

Simon McIntosh-Smith
University of Bristol
HPC Research Group
simonm@cs.bris.ac.uk

University of BRISTOL

Business as usual?

- We've all got used to new machines being relatively simple evolutions of our previous machines

- … This will **no longer be true** from 2016 onwards
- … With huge implications for the scientific software community

What's changing?

- Mainstream multi-core CPUs will continue to evolve, but more slowly…
  - IVB → Haswell → Broadwell CPUs
  - 12 core → 18 core → 22+ cores …

- To retain the levels of performance increase we have historically enjoyed, we will have **no choice** but to adopt radically different architectures

What are the options?

Many-core CPUs:

- Intel Xeon Phi
  - Knights Landing (KNL) launching 2016
  - Large KNL machines going into US national labs
  - O(70) cores, 2x 512-bit vectors per cycle per core, on-package high bandwidth memory

- Other many-core CPUs emerging
  - Mostly based on ARM64 architecture
  - Multiple vendors: AMD, Broadcom, Cavium, AMCC, …

What are the options?

GPUs:

- Nvidia Pascal / Volta
  - Tightly couples with IBM Power CPUs

- AMD
  - Often have more memory bandwidth and FLOPs than Nvidia
  - Interesting focus on tight CPU/GPU integration ("fat APUs")

What other big changes are coming?

- **Deeper memory hierarchies**
  - Could have up to ***5*** levels of memory in-node
    1. Registers
    2. Caches (L1 to L3 or even L4), O(10MB)
    3. On-package high bandwidth memory, O(10GB)
    4. Regular DRAM, O(100GB)
    5. NVRam, O(1-10TB)

- *How does software use this efficiently?*

What other big changes are coming?

- **Integrated interconnects**, e.g.
  - Intel Omni-Path fabric (on-chip in KNL)
  - Nvidia Nvlink (direct CPU-GPU connection)
  - Cray and Mellanox must be working on this too

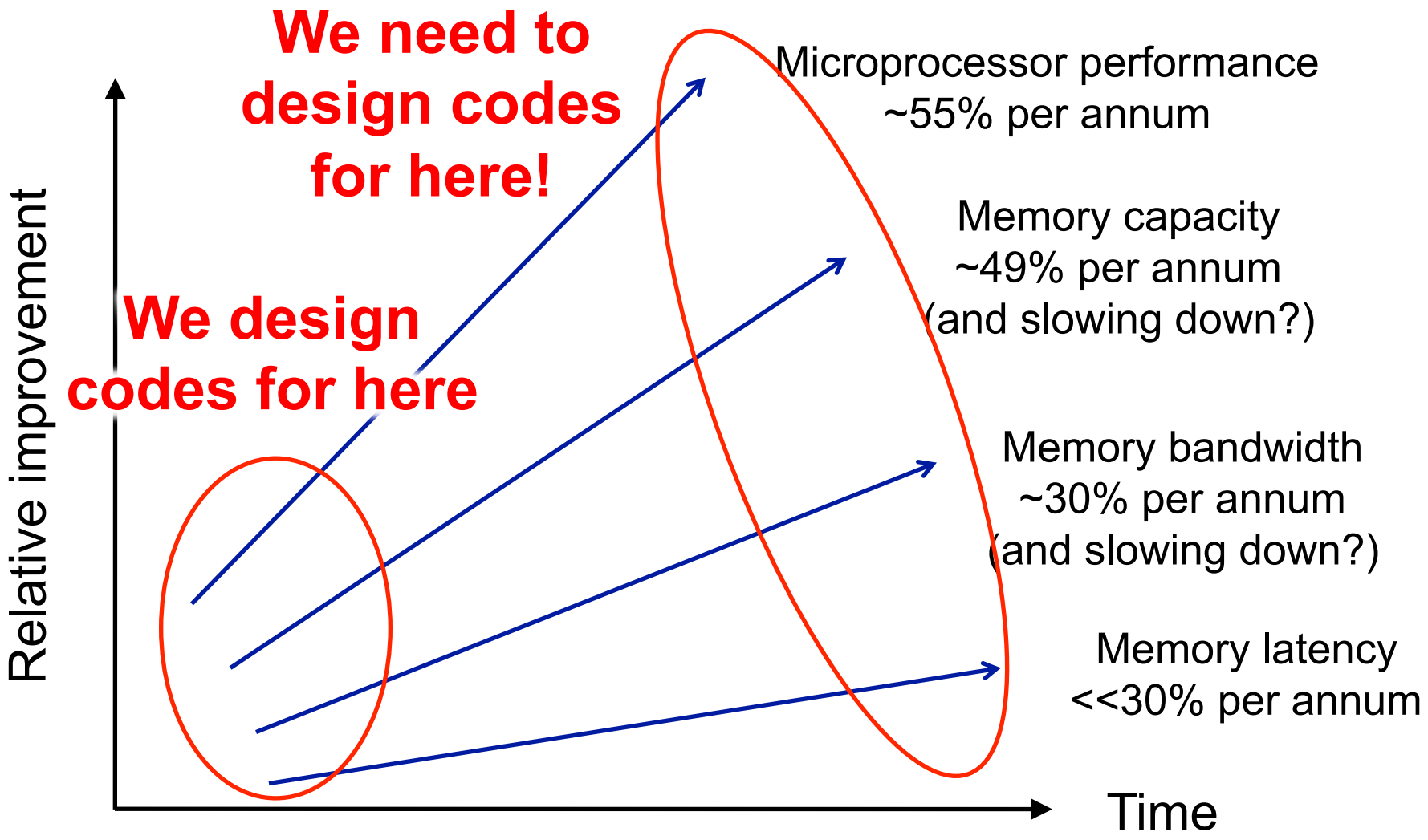
- Once the interconnect is integrated on-chip, the next step would be to integrate **inside** the virtual memory system
  - Dramatically reduces overhead (no more drivers)
  - Should enable efficient **PGAS** implementations
  - *Likely this will be a feature of Exascale machines*

What other big changes are coming?

- **FPGAs**
  - Now supporting OpenCL
  - Intel recently acquired Altera
  - Hybrid CPU/FPGA systems are possible
  - Microsoft using FPGAs for Bing searches

- Could cast some key algorithms into optimised hardware in FPGA form
- *I'm still not sure how big a role FPGAs will play in HPC's future*

Long-term fundamental trends



University of BRISTOL

"Will this affect my code?"

What are we going to have to do?

- **Expose maximum parallelism**:
  - Task, data, vector, …
- Explicitly manage the memory hierarchy
  - Can't leave it all to the caches anymore
- Be ready for major changes to interconnects in the next few years
  - PGAS revival?

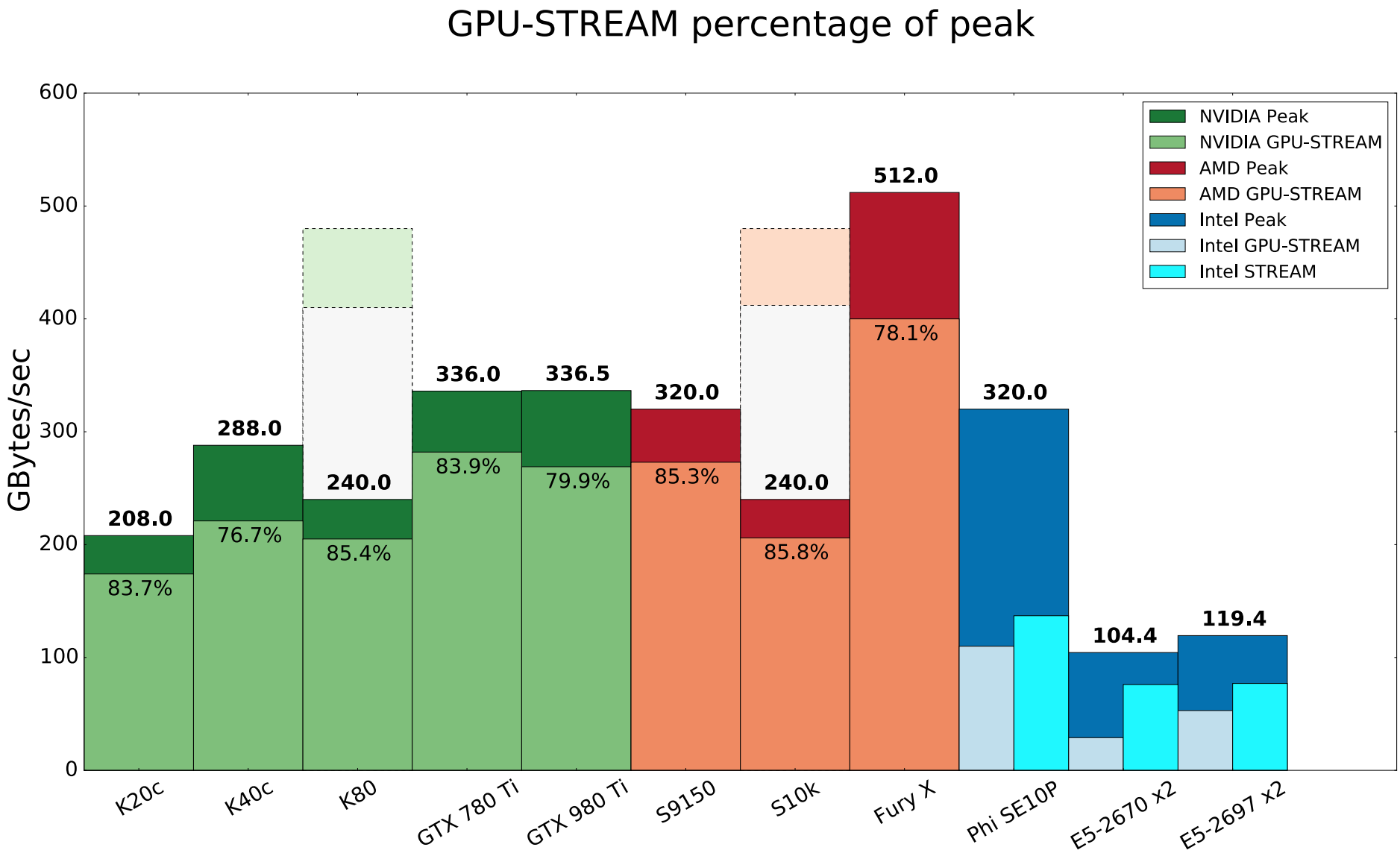
HPC nirvana

- (Re)write your code once, run efficiently everywhere
- Realistic?
- We don't even really have this today on CPU-based systems
  - E.g. between x86 clusters and Blue Gene
- Expect to see a lot more (i) **Domain Specific Languages (DSLs)**, (ii) **code generation** and (iii) **autotuning** to help make this a reality

Performance Portability - How close can we get?

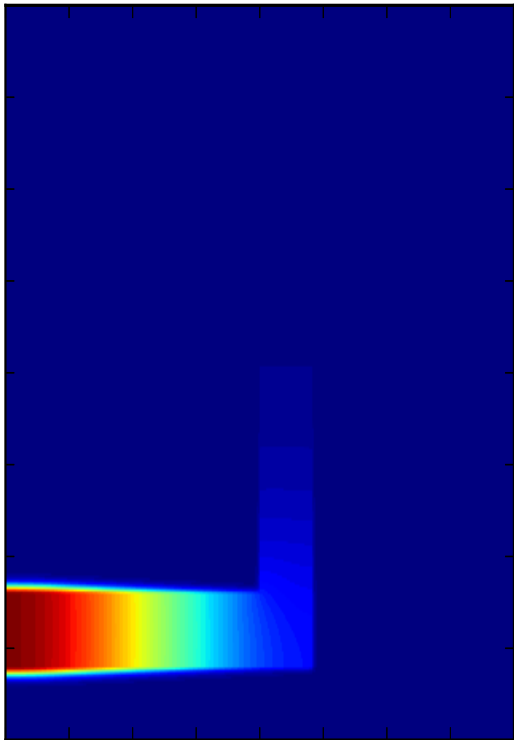
- My group has been looking at this problem for 6 years
- Only one truly cross platform parallel programming API we could use in the past
  - OpenCL 
- Finally seeing **OpenMP 4.x** being adopted for GPU programming, a big step forwards
- Also Kokkos, Raja, OmpSs, …

See: https://github.com/UoB-HPC/GPU-STREAM/wiki



TeaLeaf – Heat Conduction

- Mini-app from Mantevo suite of benchmarks

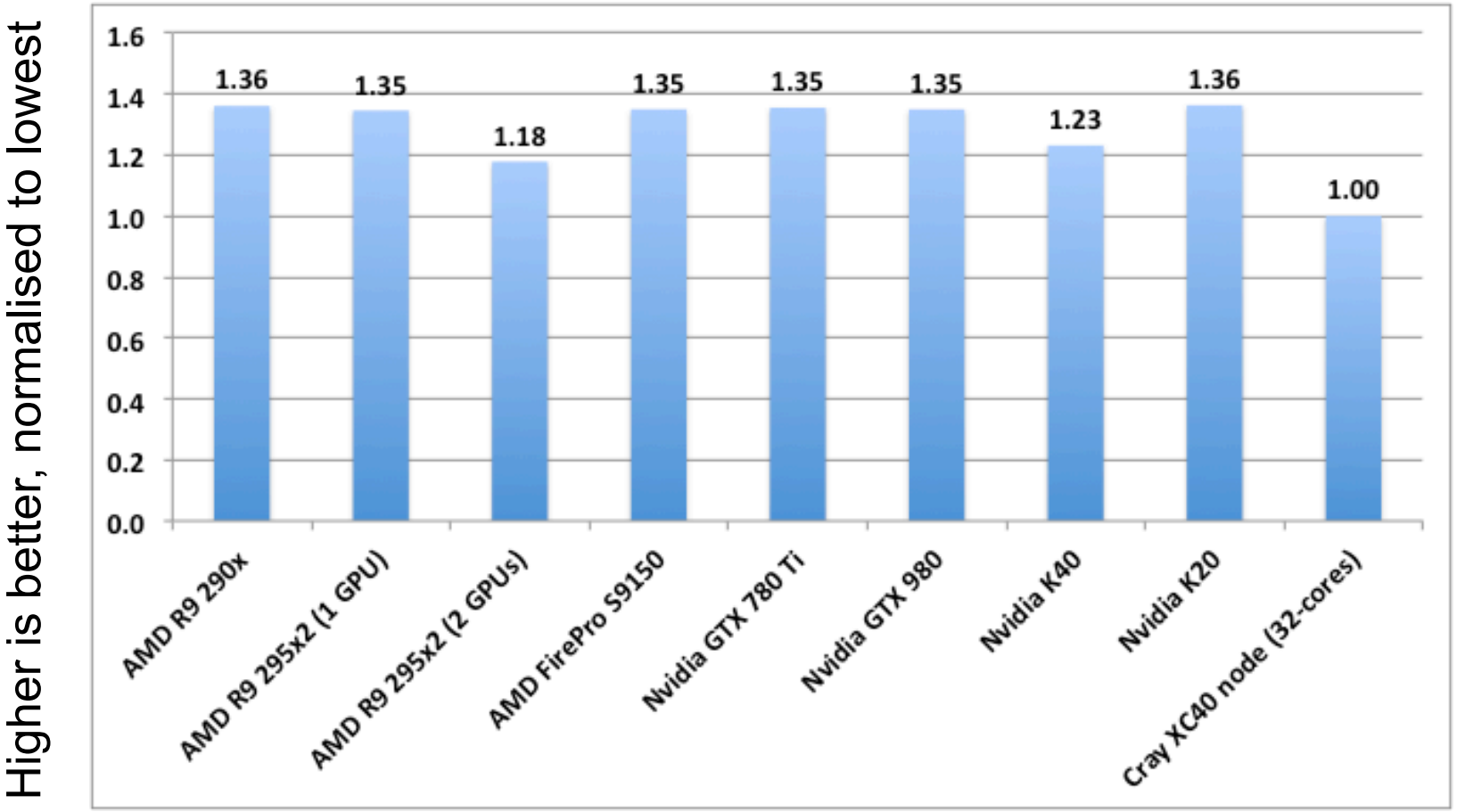


- Implicit, sparse, matrix-free solvers on structured grid
  - Conjugate Gradient (CG)
  - Chebyshev
  - Preconditioned Polynomial CG (PPCG)
- Memory bandwidth bound
- Good strong and weak scaling on Titan & Piz Daint

University of BRISTOL

Download **TeaLeaf** from:
http://uk-mac.github.io/TeaLeaf/

TeaLeaf CG performance



Most GPUs sustaining ~50% of peak memory bandwidth.
Achieved **excellent performance portability** across *diverse* architectures

University of BRISTOL

"*An Evaluation of Emerging Many-Core Parallel Programming Models*", *Martineau, M., McIntosh-Smith, S., Boulton, M., Gaudin, W., PMAM, Barcelona, March 2016.*

Evidence?

- 5 of top 10 leadership class machines already rely on Xeon Phi or GPUs
- Holdouts are mostly based on IBM's BlueGene/Q, but this is end-of-lifeing…
- From 2016 onwards, expect almost all of the Top 10 machines world-wide to rely on these radically different architectures
- **Performance chasm** will then open up between "haves" and "have nots"

Summary

- A "business as usual" approach to scientific software development will result in being left in the ***slow lane***
- Developers are faced with the challenging issue of developing performance portable code on increasingly complex and diverse architectures
- This will require the most significant software re-engineering effort in a generation…
- **BEWARE** vendor specific languages! OpenACC, CUDA, ... Invest in ***open standards*** instead: OpenMP, OpenCL, ...

References

Evaluating OpenMP 4.0's Effectiveness as a Heterogeneous Parallel Programming Model
Martineau, M., McIntosh-Smith, S. & Gaudin, W.
To appear at HIPS Workshop (IPDPS), Chicago, May 2016

On the performance portability of structured grid codes on many-core computer architectures
S.N. McIntosh-Smith, M. Boulton, D. Curran and J.R. Price. ISC, Leipzig, June 2014.

Assessing the Performance Portability of Modern Parallel Programming Models using TeaLeaf
Martineau, M., McIntosh-Smith, S. & Gaudin, W.
Concurrency and Computation: Practice and Experience (April 2016), to appear

University of BRISTOL