Prof Simon McIntosh-Smith
Isambard PI
University of Bristol /
GW4 Alliance

How Arm's entry into the HPC market might affect meteorological codes

University of Bristol

http://gw4.ac.uk/isambard/

GW4

Recent processor trends in HPC

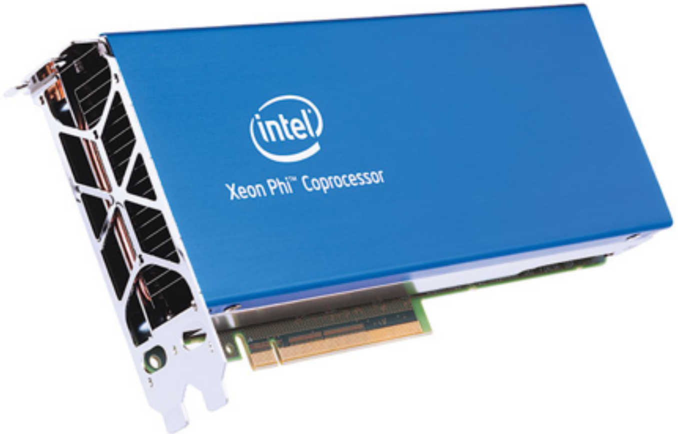
- Most of the world's supercomputers are large collections of servers based on commodity processors, typically Intel's x86 CPUs

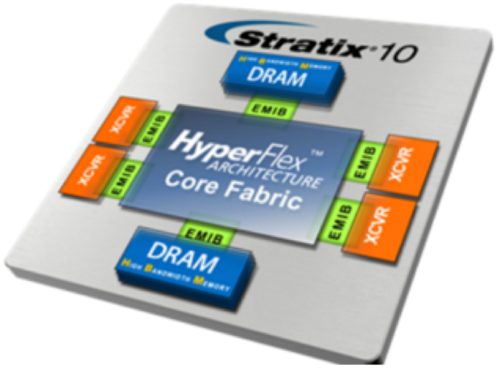
- **New computer architectures** have emerged in the last few years, exploring diverse ways to provide the next jump in performance

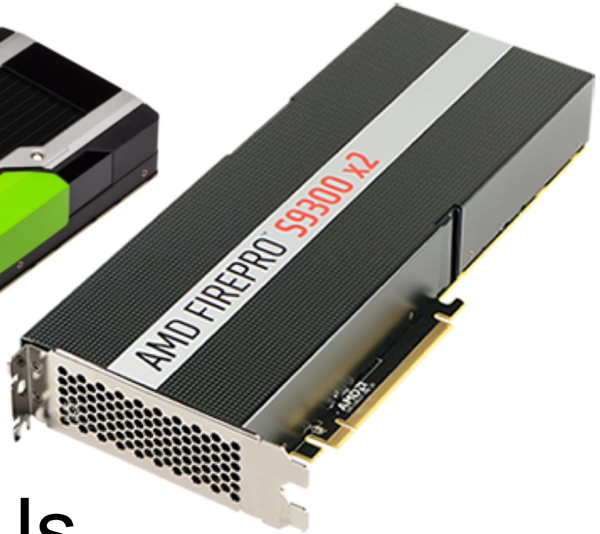
Emerging architectures

FPGAs

Many-core CPUs

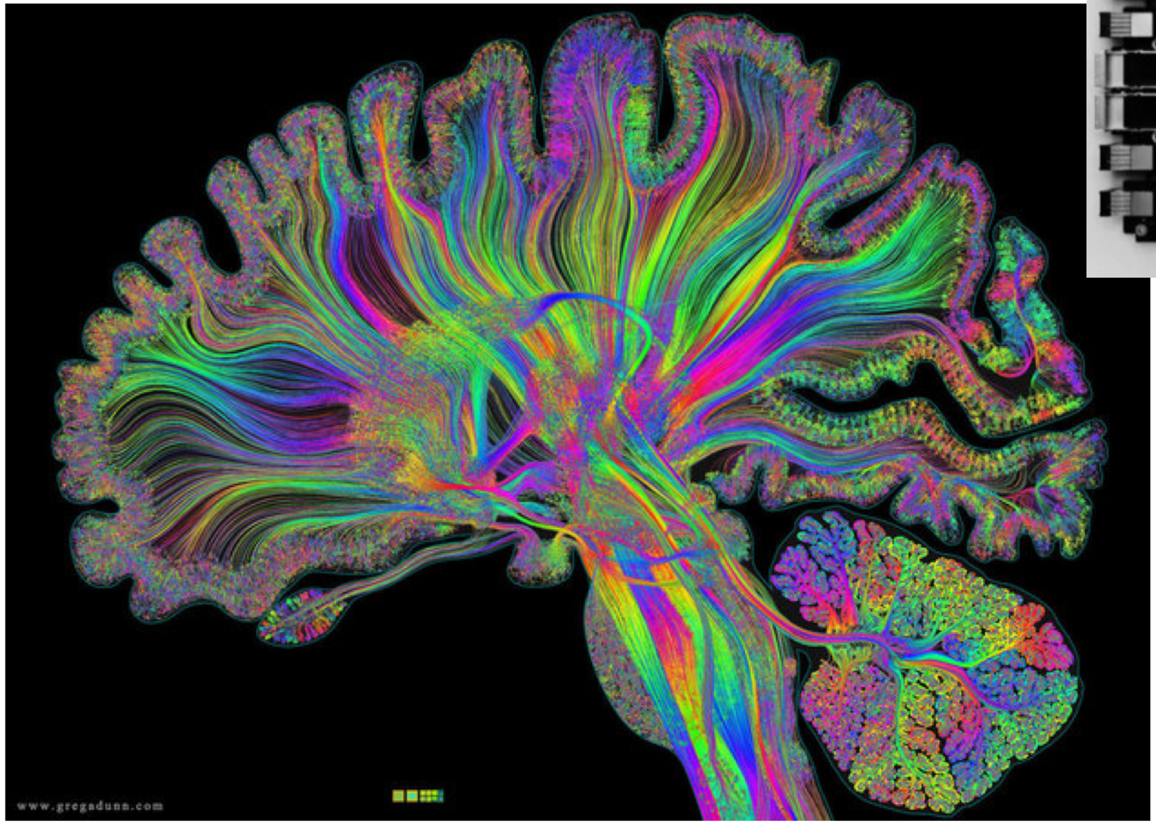


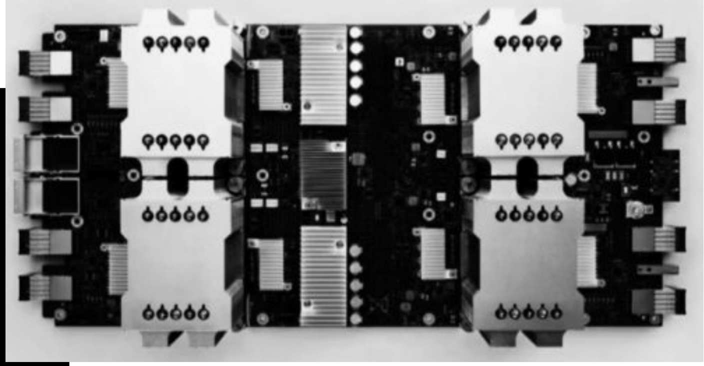


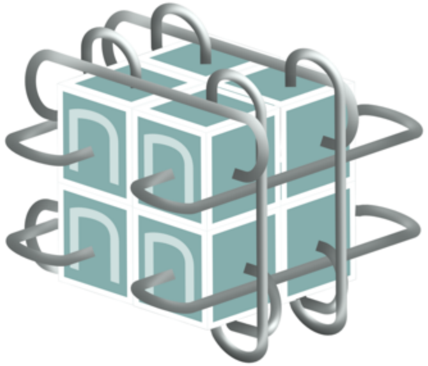


GPUs

Emerging architectures





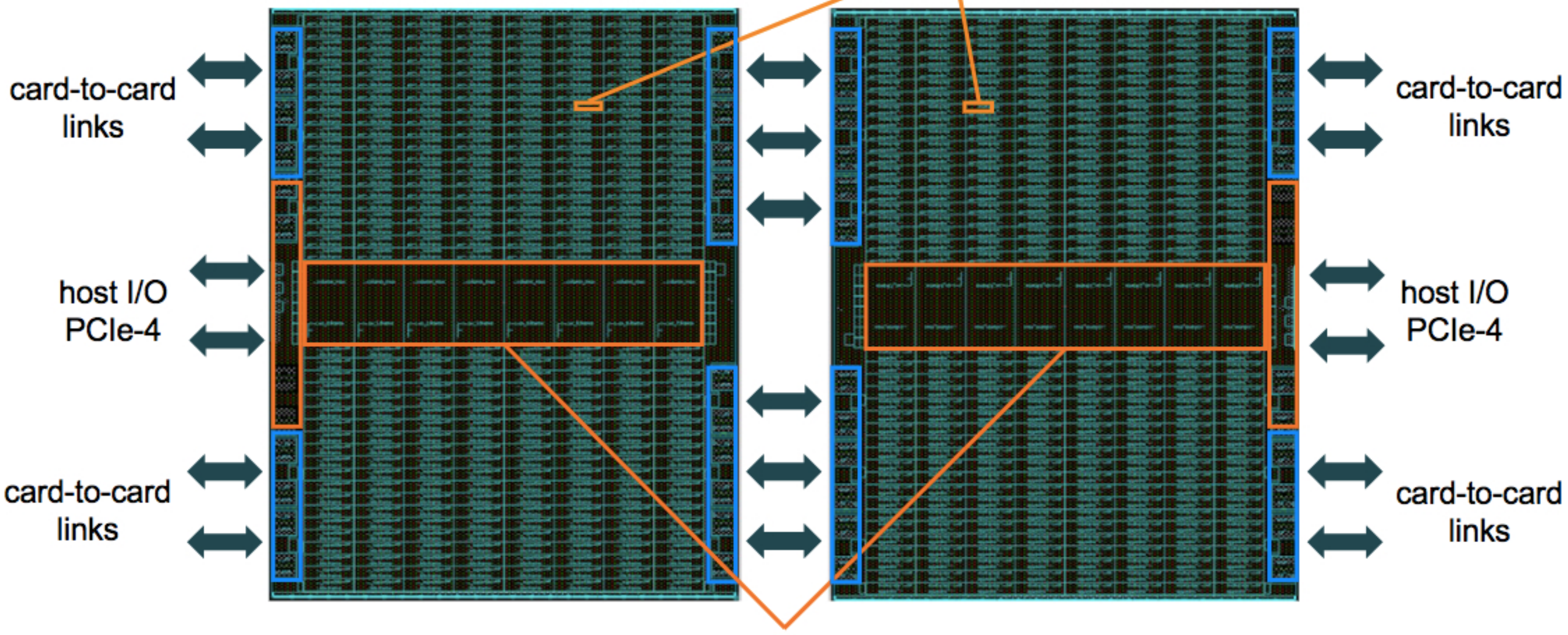


Google’s Tensorflow Processing Unit (TPU), GraphCore, Intel’s Nervana

GRAPHCORE IPU pair – 600MB @ 90TB/s

"Colossus" IPU pair
(300W PCIe card)

2432 processor tiles >200Tflop$_{16.32}$ ~600MB



all-to-all exchange spines each ~8TBps
link + host bandwidth 384GBps/chip

Recent CPU trends

- CPUs have evolved to include **lots of cores** and **wide vector units**
- The latest Intel Skylakes have up to 28 cores each
  - 56 cores, 256 GB/s, >3.7 TFLOP/s (dual socket node)
    - Intel Xeon Platinum 8176 (Skylake), 2.1GHz
  - **~$18,000 list price** just for two CPUs!
- Rate of improvement in CPU performance is at a ***historical low***
  - At least, for today's mainstream CPU vendors…

So why explore Arm-based CPUs?

- The architecture development is driven by the ***fast-growing mobile space***
- Multiple vendors of Arm-based CPUs:
  - Greater **competition**
  - More **choice**
  - Exciting **innovations**, e.g. in vector instruction set

- Current vendors include Cavium, Fujitsu, Ampere, Huawei
- At least three of the first Exascale machines will use Arm

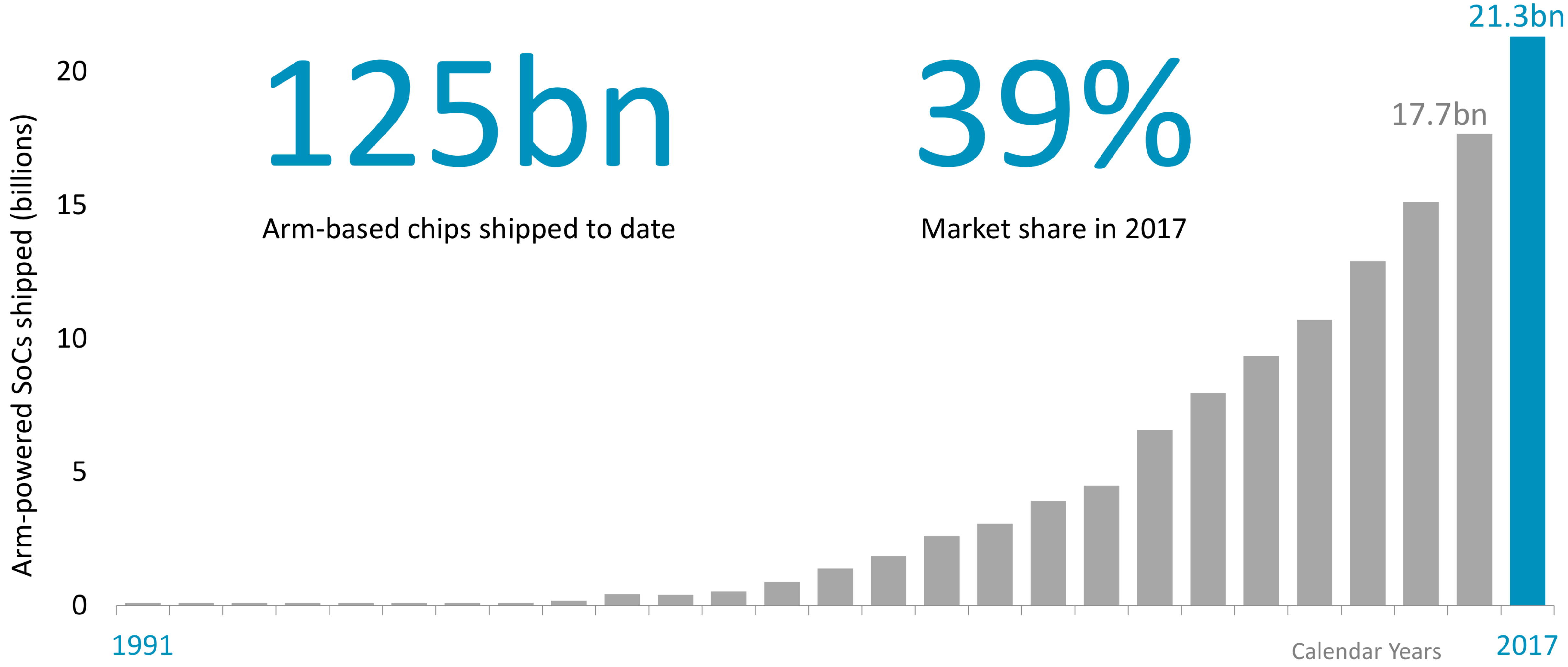
Arm-based chip shipments

125bn

Arm-based chips shipped to date

39%

Market share in 2017

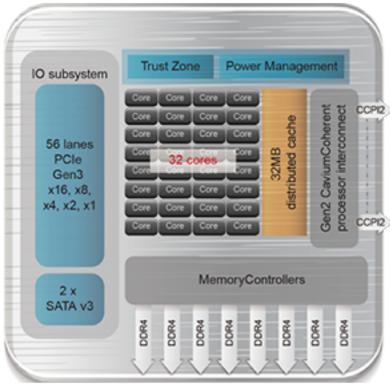


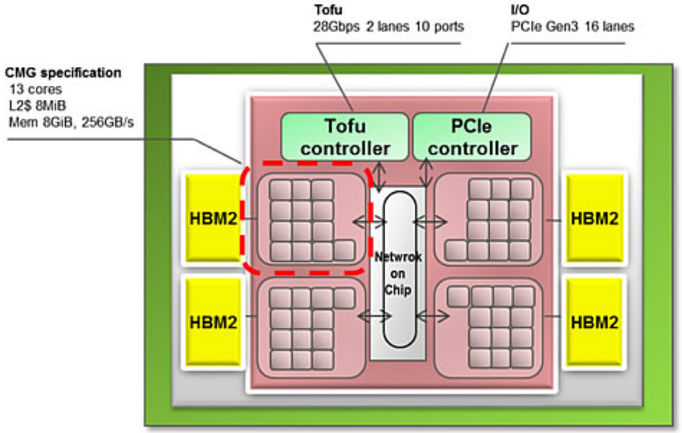
 © Arm 2018

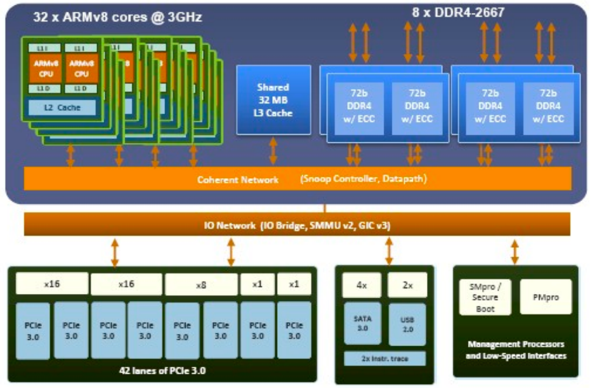
https://www.arm.com/-/media/global/company/investors/PDFs/Arm_SB_Q1_2018_Roadshow_Slides_Final.pdf

arm

Current Arm server CPU vendors

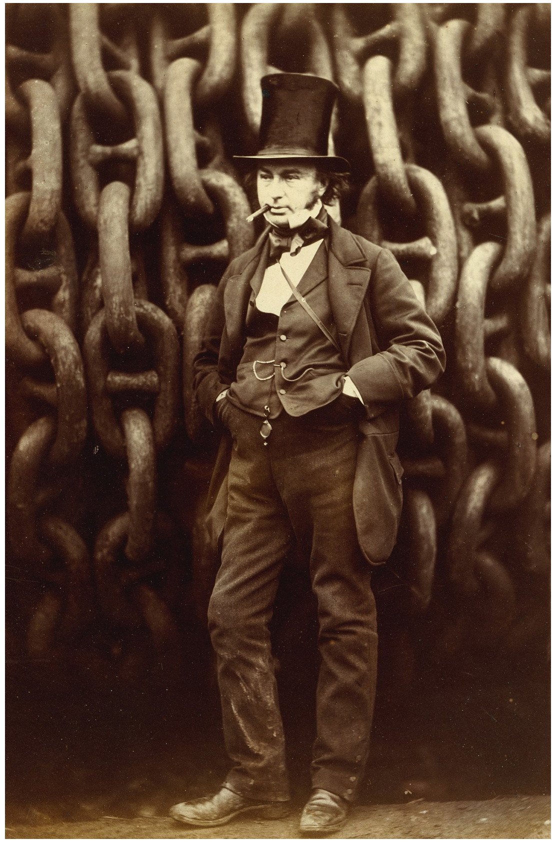






http://gw4.ac.uk/isambard/

'Isambard' is a new UK Tier 2 HPC service from GW4



Isambard Kingdom Brunel
1804-1859

http://gw4.ac.uk/isambard/

The Great Western railway was one of the first high-speed information networks

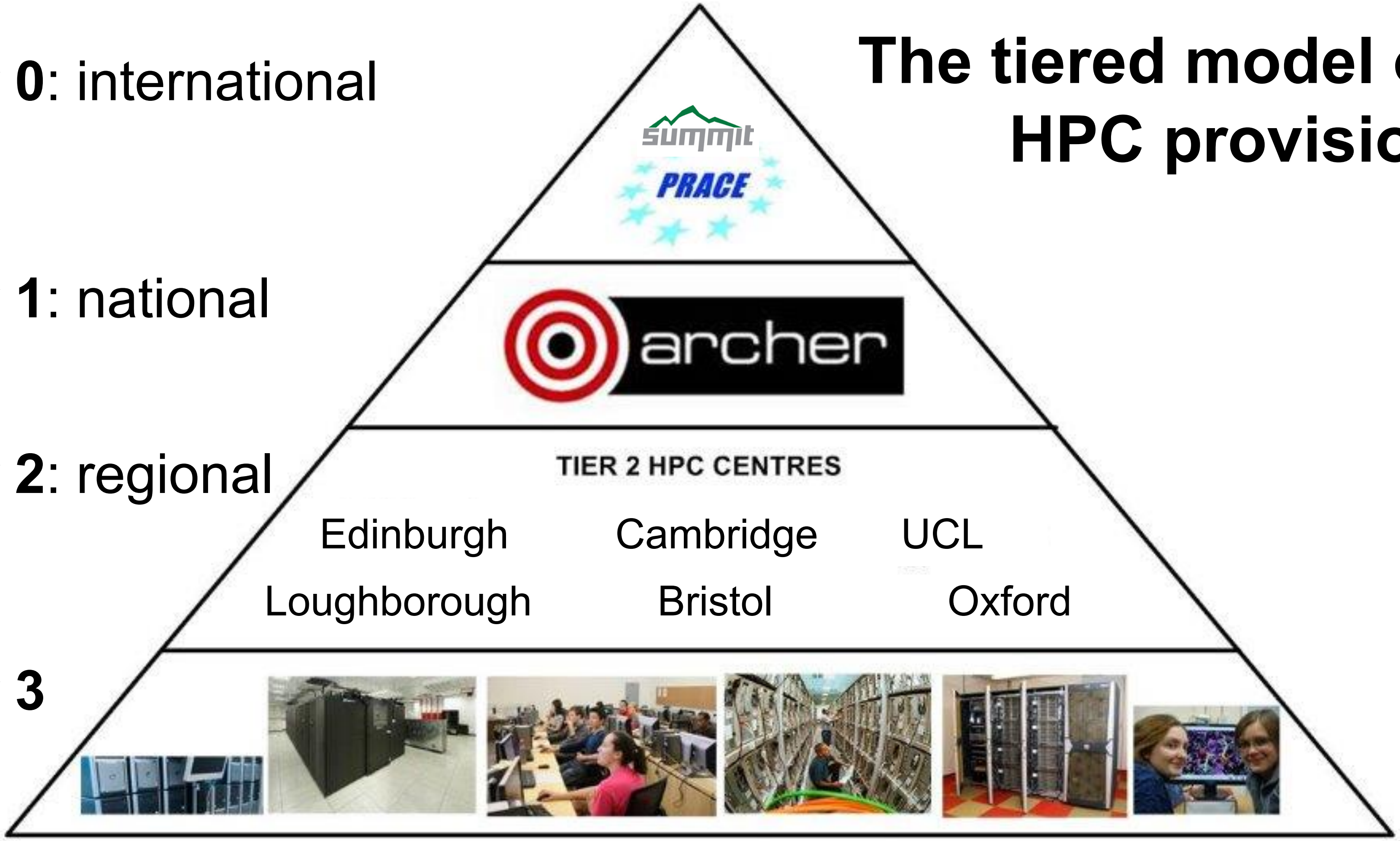
The tiered model of HPC provision

Tier 0: international

Tier 1: national

Tier 2: regional

Tier 3



Isambard system specification

- **10,752** Armv8 cores (168 x 2 x 32)
  - **Cavium ThunderX2 32core 2.1GHz**
- Cray XC50 Scout form factor
- High-speed **Aries** interconnect
- Cray HPC optimised software stack
  - CCE, CrayPAT, Cray MPI, math libraries, …
- **Technology comparison:**
  - **x86, Xeon Phi, Pascal GPUs**
- Phase 1 installed March 2017
- Phase 2 (the Arm part) ships Oct 2018
- £4.7m total project cost over 3 years

University of BRISTOL

http://gw4.ac.uk/isambard/

Isambard system specification

- **10,752** Armv8 cores (168 x 2 x 32)
  - **Cavium ThunderX2 32core 2.1GHz**
- Cray XC50 Scout form factor
- High-speed **Aries** interconnect
- Cray HPC optimised software stack
  - CCE, CrayPAT, Cray MPI, math libraries, …
- **Technology comparison:**
  - **x86, Xeon Phi, Pascal GPUs**
- Phase 1 installed March 2017
- Phase 2 (the Arm part) ships Oct 2018
- £4.7m total project cost over 3 years

University of BRISTOL

http://gw4.ac.uk/isambard/

Isambard's core mission: evaluating Arm for production HPC

Starting with some of the most heavily used codes on Archer

- **VASP, CASTEP, GROMACS, CP2K, UM, NAMD, Oasis, SBLI, NEMO**
- Note: many of these codes are written in FORTRAN

Additional important codes for project partners:

- **OpenFOAM, OpenIFS**, WRF, CASINO, LAMMPS, …


RAISING
STEAM
1ST ISAMBARD HACKATHON - BRISTOL
NOVEMBER 2ND & 3RD 2017


Isambard
STOKING
THE FIRE
2ND ISAMBARD HACKATHON-BRISTOL
MARCH 19TH & 20TH 2018

OpenCFD®

University of York

University of Southampton

KTH Vetenskap och Konst

ETH Zürich

University of Vienna

University of Bristol

http://gw4.ac.uk/isambard/

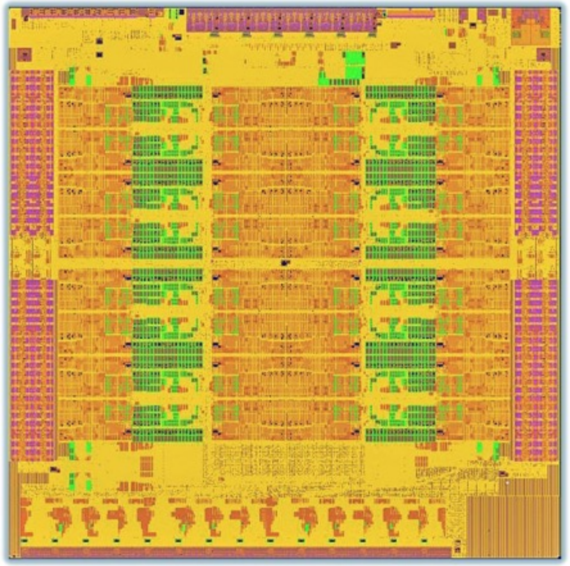
GW4

Benchmarking platforms

Processor	Cores	Clock speed GHz	TDP Watts	FP64 TFLOP/s	Bandwidth GB/s
Broadwell	2×22	2.2	145	1.55	154
Skylake Gold	2×20	2.4	150	3.07	256
Skylake Platinum	2×28	2.1	165	3.76	256
ThunderX2	2×32	2.2	175	1.13	320

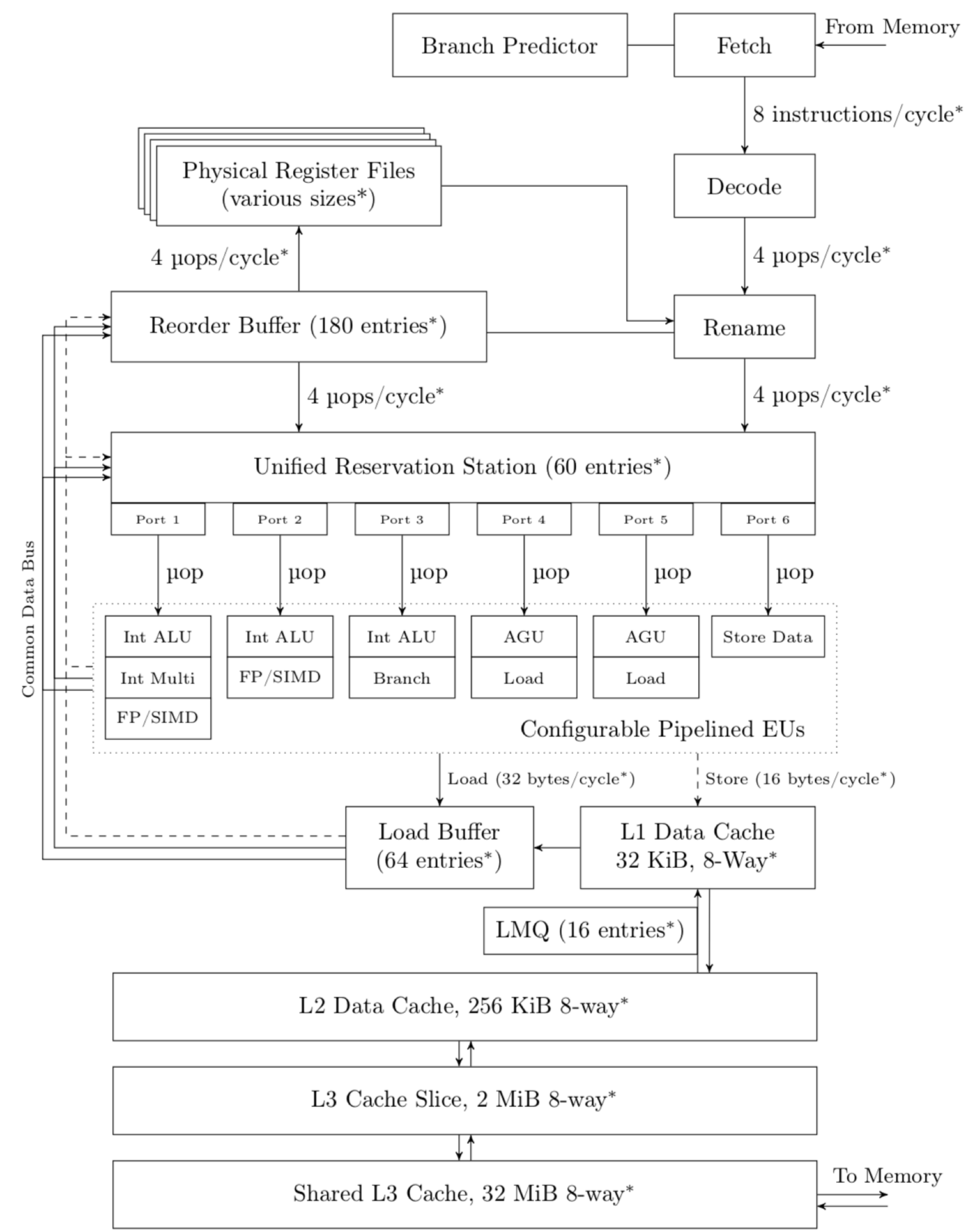
BDW 22c Intel Broadwell E5-2699 v4, **$4,115** each (near top-bin)
SKL 20c Intel Skylake Gold 6148, **$3,078** each
SKL 28c Intel Skylake Platinum 8176, **$8,719** each (near top-bin)
TX2 32c Cavium ThunderX2, **$1,795 each** (near top-bin)

Cavium ThunderX2, a seriously beefy CPU

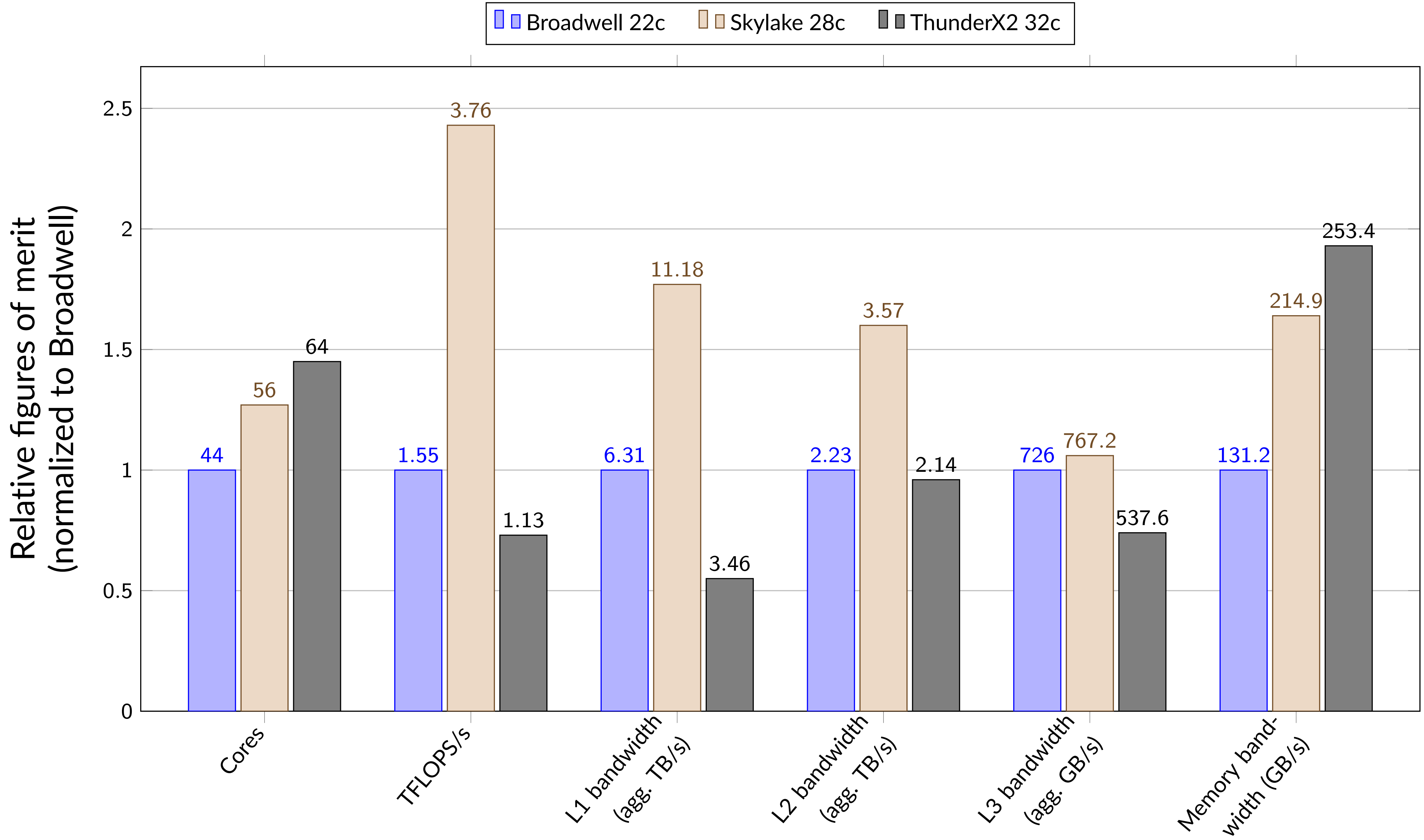


- 32 cores at up to 2.5GHz
- Each core is 4-way superscalar, Out-of-Order
- 32KB L1, 256KB L2 per core
- Shared 32MB L3
- Dual 128-bit wide NEON vectors
  - Compared to Skylake's 512-bit vectors, and Broadwell's 256-bit vectors
- 8 channels of 2666MHz DDR4
  - Compared to 6 channels on Skylake, 4 channels on Broadwell
  - AMD's EPYC also has 8 channels

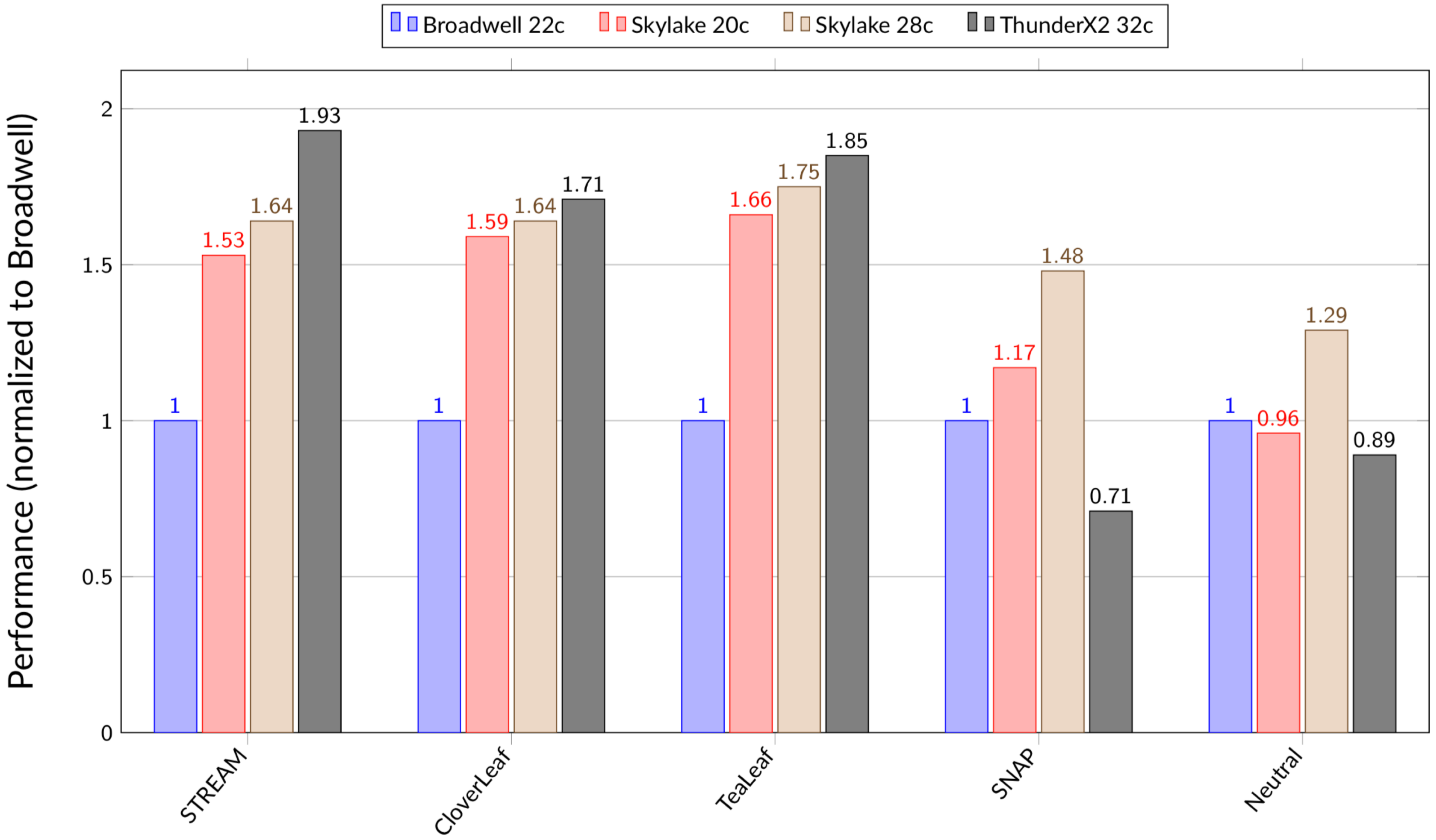
ThunderX2 architecture



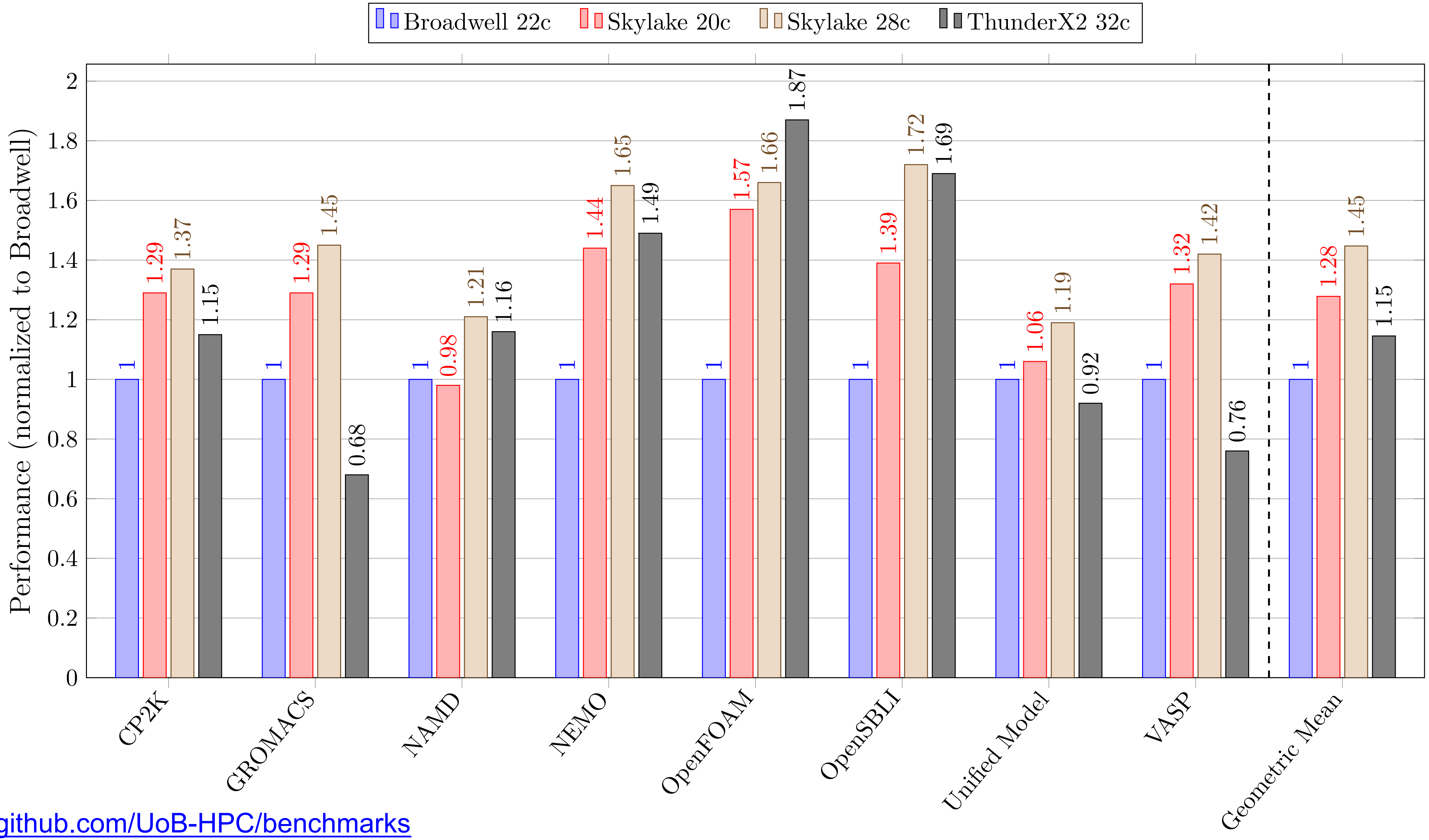
Key architectural comparisons (node-level, dual socket)



Performance on mini-apps (node level comparisons)



Performance on heavily used applications from Archer



https://github.com/UoB-HPC/benchmarks

Performance summary

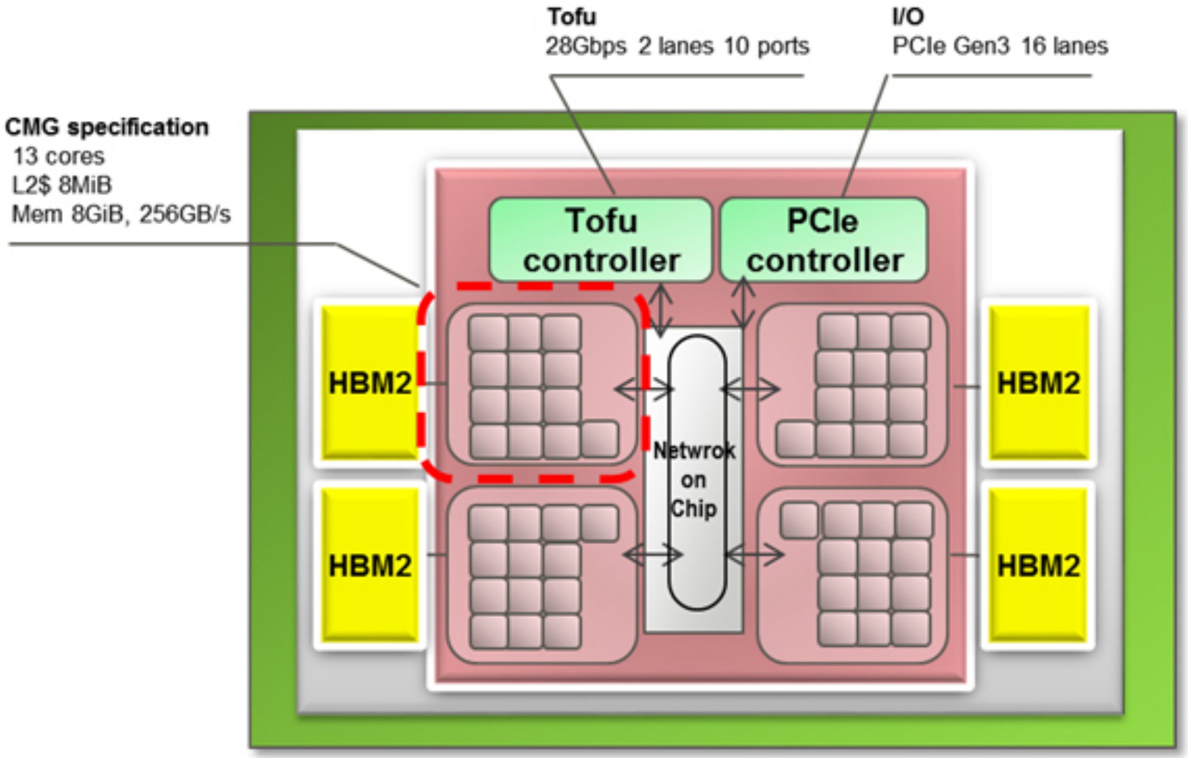
- ThunderX2 is competitive with contemporary x86 processors
  - ThunderX2 is **faster** when external memory bandwidth is critical
  - Skylake is **faster** when FLOP/s and L1 cache bandwidth dominate
  - **Performance per dollar is very compelling for ThunderX2**
- Next-gen Arm CPUs will increase FLOP/s and cache bandwidth
  - Introduction of SVE will allow vector width of up to 2048-bits
  - E.g. Fujitsu A64FX chip unveiled recently with 512-bit SVE
  - Expecting 512-bits to be a common choice for server chips

Future opportunities

- Important to note that Arm is the main driver of the System-on-Chip ecosystem than underpins most mobile computing
- Benefits:
  - Fast-growing → **rapid innovation, investment, competition**, …
  - Focus on customization → **enables real co-design of future processors**
- Future innovations:
  - Scalable Vector Extensions (SVE), e.g. Fujitsu A64fx CPU
  - Application-optimized accelerators/co-processors
  - Advanced memory systems, e.g. HBM

An example forthcoming Arm-based CPU: Fujitsu's A64fx

- 48 cores
- 2.7 TFLOP/s double precision (vs. SKL/s 1.9 TFLOP/s)
- 1 TeraByte/s main memory bandwidth (vs. SKL's 128 GB/s)
- ~170 Watts
- High speed interconnect
- 512-bit wide vectors
- First silicon now
- 8.7B transistors, 7nm



University of BRISTOL

GW4

http://www.fujitsu.com/global/about/resources/news/press-releases/2018/0822-02.html

Arm software ecosystem

- Three mature compiler suites:
  - GNU (gcc, g++, gfortran)
  - Arm HPC Compilers based on LLVM (armclang, armclang++, armflang)
  - Cray Compiling Environment (CCE)
- Three mature sets of math libraries:
  - OpenBLAS + FFTW
  - Arm Performance Libraries (BLAS, LAPACK, FFT)
  - Cray LibSci + Cray FFTW
- Multiple performance analysis and debugging tools:
  - Arm Forge (MAP + DDT, formerly Allinea)
  - CrayPAT / perftools, CCDB, gdb4hpc, etc

University of BRISTOL

GW4

Which compiler was fastest on each code?

Benchmark	ThunderX2	Broadwell	Skylake
STREAM	Arm 18.3	Intel 18	CCE 8.7
CloverLeaf	CCE 8.7	Intel 18	Intel 18
TeaLeaf	CCE 8.7	GCC 7	Intel 18
SNAP	CCE 8.6	Intel 18	Intel 18
Neutral	GCC 8	Intel 18	GCC 7
CP2K	GCC 8	GCC 7	GCC 7
GROMACS	GCC 8	GCC 7	GCC 7
NAMD	Arm 18.2	GCC 7	GCC 7
NEMO	CCE 8.7	CCE 8.7	CCE 8.7
OpenFOAM	GCC 7	GCC 7	GCC 7
OpenSBLI	CCE 8.7	Intel 18	CCE 8.7
UM	CCE 8.6	CCE 8.5	CCE 8.7
VASP	GCC 7.2	Intel 18	Intel 18

https://github.com/UoB-HPC/benchmarks

Exact same issues on x86

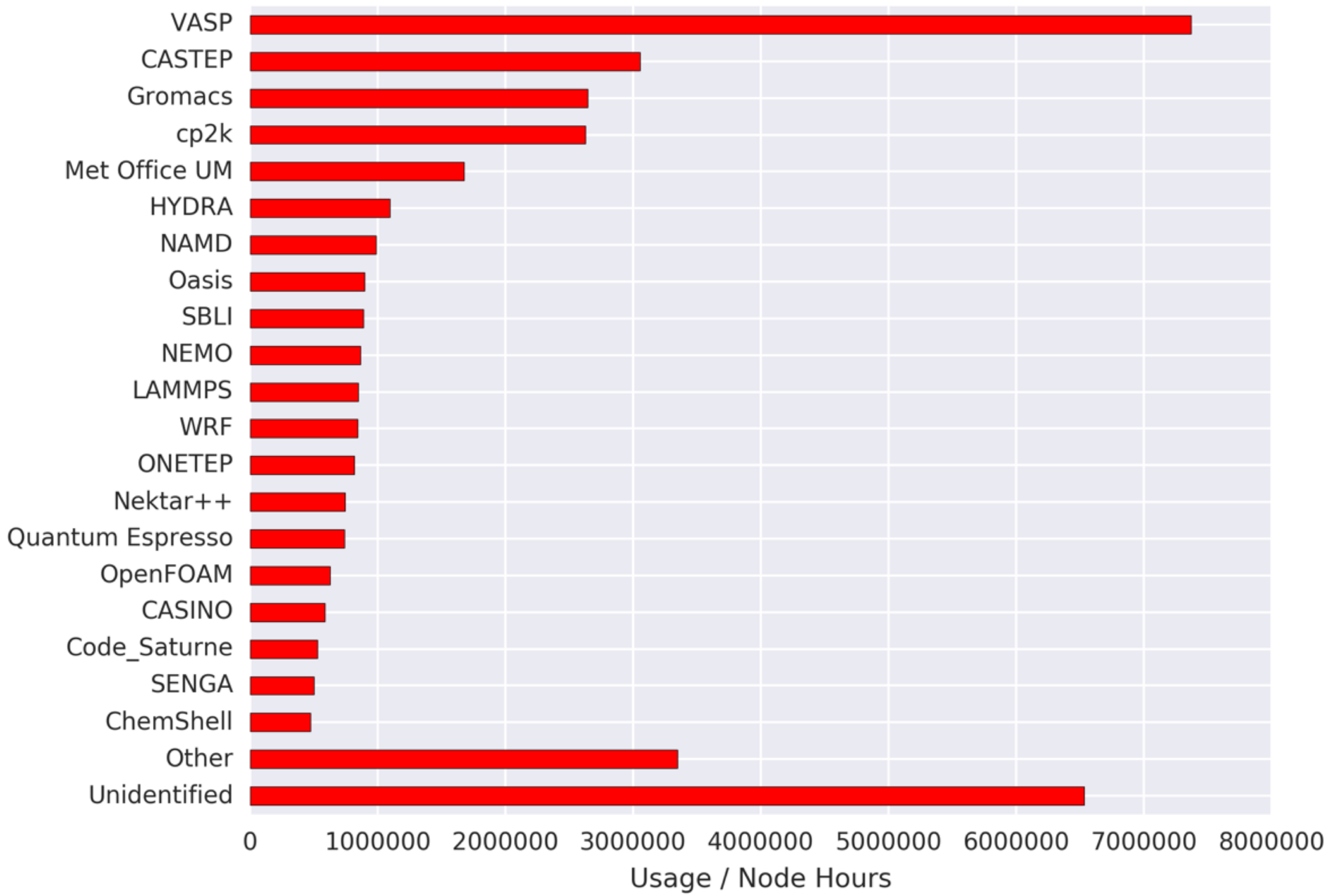
Comparison of compilers on Arm

	GCC	Arm	CCE
STREAM	97%	100%	99%
CloverLeaf	92%	95%	100%
TeaLeaf	99%	95%	100%
SNAP	74%	87%	100%
Neutral	100%	94%	85%
CP2K	100%	BUILD	CRASH
GROMACS	100%	91%	CRASH
NAMD	83%	100%	BUILD
NEMO	-	-	100%
OpenFOAM	100%	97%	BUILD
OpenSBLI	-	-	100%
Unified Model	84%	72%	100%

University of BRISTOL

https://github.com/UoB-HPC/benchmarks

Future opportunities: HBM, how much would we need?



Archer usage from a 12 month study.

Archer has 24 IVB cores and 64 GiB per node (2.67GiB/core).

A. Turner and S. McIntosh-Smith. "A survey of application memory usage on a national supercomputer: an analysis of memory requirements on ARCHER." In PMBS, IEEE/ACM SuperComputing 2017.

University of BRISTOL

GW4

Future opportunities: HBM, how much would we need?

Fujitsu's "Post-K" A64fx CPU has 32GB HBM2 for 48 cores, 0.67GB/core

Max. memory use (GiB/node) vs. Job size (nodes)

Max. memory use (GiB/node)	[1,1]	[2,2]	(2,4]	(4,8]	(8,16]	(16,32]	(32,64]	(64,128]	(128,256]	(256,512]	(512,1024]	(1024,2048]	(2048,4096]	(4096,4290]
[0,12)	2.0%	1.4%	5.6%	10.4%	10.1%	9.4%	8.2%	5.4%	4.0%	2.3%	1.4%	0.5%	0.0%	0.0%
[12,24)	0.1%	0.6%	0.8%	1.9%	4.9%	3.3%	2.4%	2.0%	1.4%	0.6%	0.3%	0.2%	0.0%	-
[24,48)	0.2%	0.2%	0.6%	0.9%	1.3%	2.1%	1.5%	1.2%	1.5%	1.3%	0.2%	0.4%	0.0%	-
[48,96)	0.1%	0.2%	0.1%	0.3%	0.6%	1.1%	1.0%	2.1%	0.9%	0.3%	0.0%	0.2%	0.0%	0.0%
[96,128]	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	0.1%	1.0%	0.2%	0.1%	-	0.0%	0.0%

GiB/core	Σ
[0-0.5)	65.4%
[0.5-1.0)	86.8%
[1.0-2.0)	96.2%
[2.0-4.0)	99.9%
[4.0-8.0]	100.0%

University of BRISTOL

A. Turner and S. McIntosh-Smith. "A survey of application memory usage on a national supercomputer: an analysis of memory requirements on ARCHER." In PMBS, IEEE/ACM SuperComputing 2017.

GW4

Implications for meteorological codes

- **More choice and diversity in architectures**
  - Significant improvements in performance and cost are possible
- Arm-based CPUs with **GPU-like levels of performance** are coming
- Make sure codes remain (performance) portable
- Ensure that memory requirements can be kept at 0.5-1.0 GB/core
  - Will enable the use of ~1TByte/s high bandwidth memories
- Include at least one Arm-based hardware platform in your plans
  - And make sure all your software builds and runs well with Arm's port of Clang/Flang/LLVM, as well as GNU

Conclusions

- Results show **ThunderX2 performance is competitive with current high-end server CPUs,** while **performance per dollar is compelling**
- **The software tools ecosystem is already in good shape**
- The full Isambard XC50 Arm system is coming up now, we're aiming to have early results to share at SC18
- The signs are that **Arm-based systems are now real alternatives for HPC,** reintroducing much needed competition to the market
- Added benefits include **real opportunity for co-design**

University of BRISTOL

GW4

For more information

Comparative Benchmarking of the First Generation of HPC-Optimised Arm Processors on Isambard

S. McIntosh-Smith, J. Price, T. Deakin and A. Poenaru, CUG 2018, Stockholm

http://uob-hpc.github.io/2018/05/23/CUG18.html

Bristol HPC group: https://uob-hpc.github.io/

Isambard: http://gw4.ac.uk/isambard/

Build and run scripts: https://github.com/UoB-HPC/benchmarks

University of BRISTOL

GW4