Prof. Simon McIntosh-Smith
PI Isambard Tier-2 HPC service
Head of the HPC research group
University of Bristol

Looking at new processors: The Isambard 2 Arm-based supercomputer

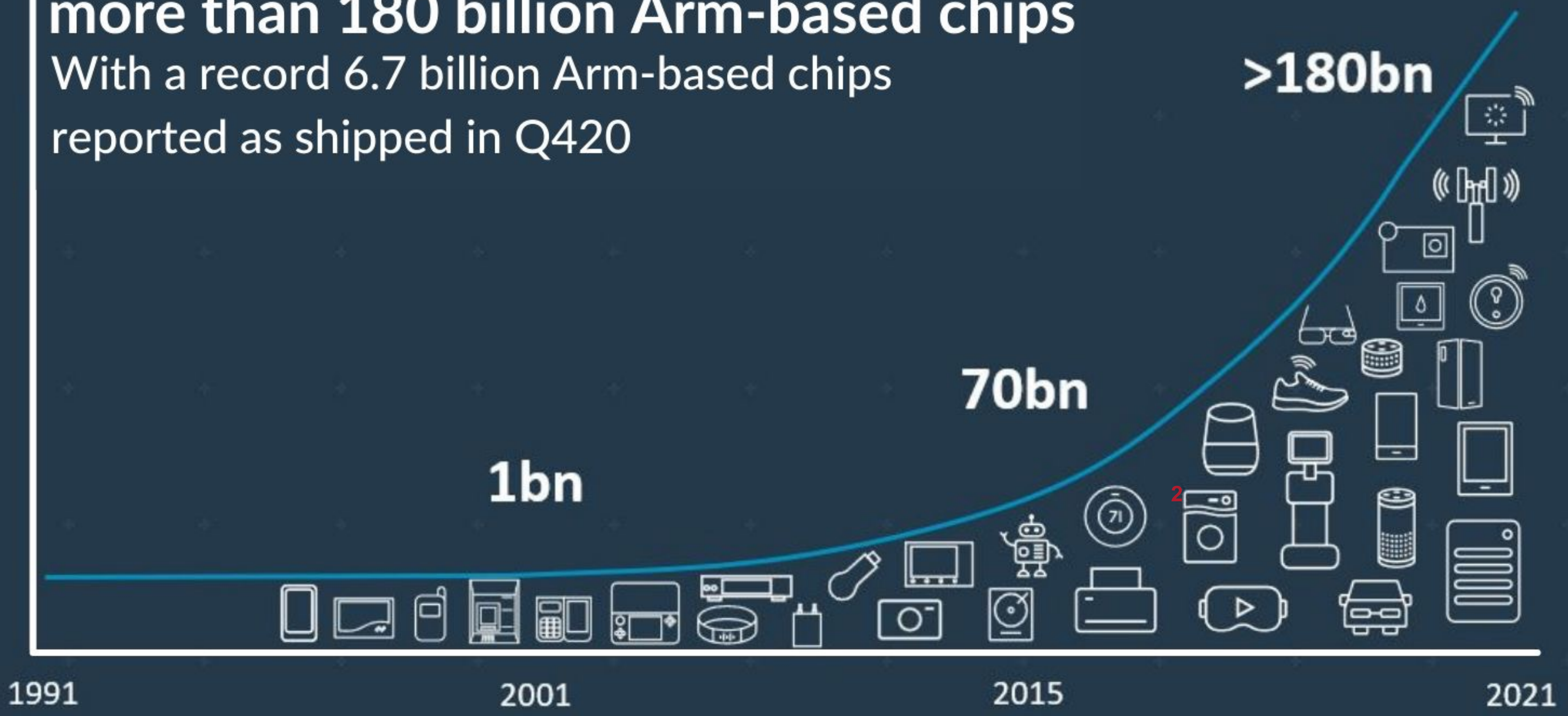
University of BRISTOL

http://uob-hpc.github.io

GW4

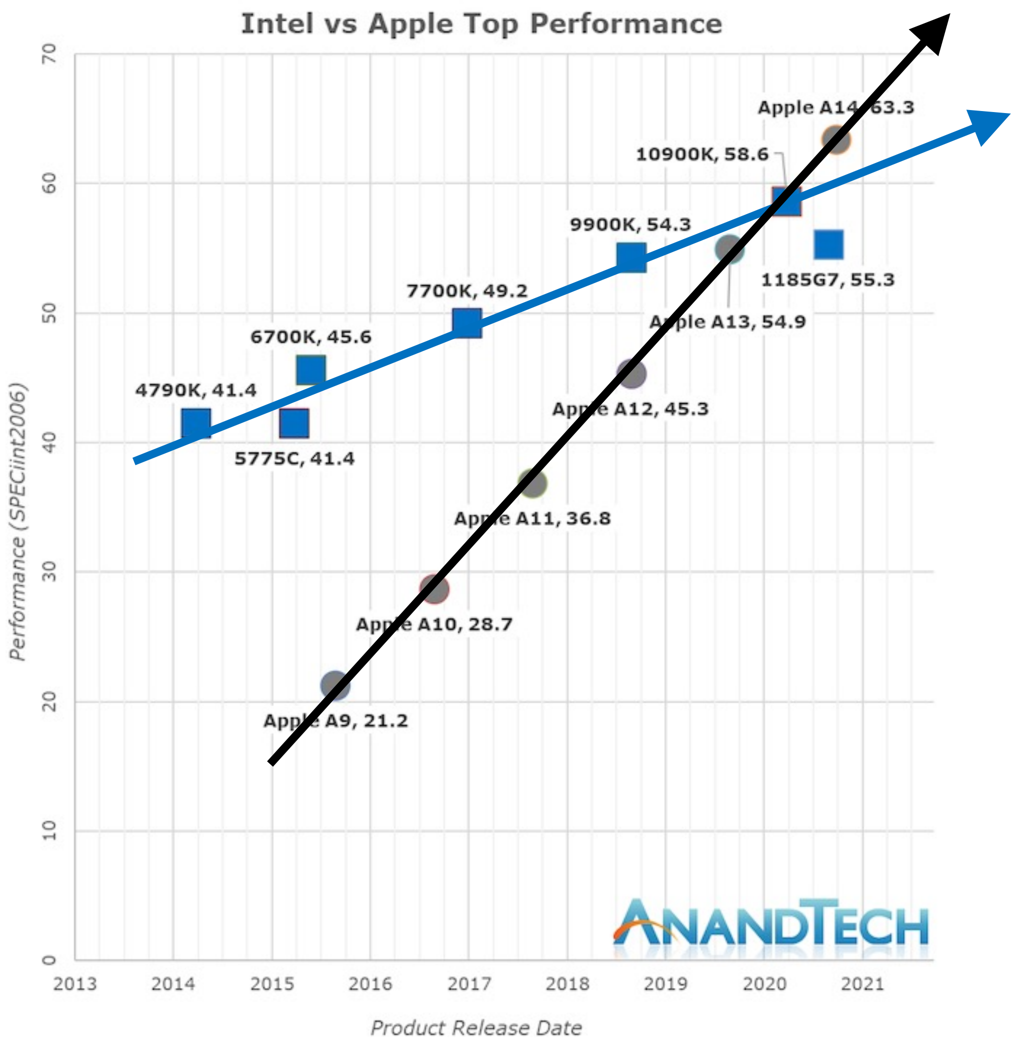
To-date, Arm partners have shipped more than 180 billion Arm-based chips

With a record 6.7 billion Arm-based chips reported as shipped in Q420



Source: https://www.arm.com/company/news/2021/02/arm-ecosystem-ships-record-6-billion-arm-based-chips-in-a-single-quarter

Mobile processors driving market disruption



Classic "Christensen" disruption of the laptop/desktop (and server?) space by the faster-growing and much larger mobile space

https://hbr.org/2015/12/what-is-disruptive-innovation

https://www.anandtech.com/show/16226/apple-silicon-m1-a14-deep-dive/4
https://www.anandtech.com/show/16252/mac-mini-apple-m1-tested

University of BRISTOL

GW4

The UK’s HPC service ecosystem is intentionally diverse

Tier-1
National

Tier-2
Regional

University of BRISTOL

https://epsrc.ukri.org/research/facilities/hpc/tier-2-hpc-centres/

GW4

Introduction to the GW4 Isambard 2 supercomputer

- **Isambard 2** is a £4.1M EPSRC project, run by a consortium of the GW4 Alliance, the Met Office, HPE/Cray, Fujitsu and Arm, to deliver a Tier-2 HPC service to researchers across the UK and around the world
- Funded in late 2019, Isambard 2 builds on Isambard 1’s achievements as the world’s first Arm64-based production supercomputer (£3M system, service began summer 2018)
- **Isambard 1** has been a huge success, proving for the first time that Arm works for supercomputing in production environments

University of BRISTOL

UKRI Engineering and Physical Sciences Research Council

GW4

Isambard 2 production system

- **21,504** ARMv8 cores (336n x 2s x 32c)
  - Marvell ThunderX2 32 core @ 2.5 GHz
- Cray XC50 ‘Scout’ with Aries interconnect
- Cray HPC optimised software stack
  - Compilers, math libraries, CrayPAT, …
  - Also comes with all the open source software toolchains: GNU, Clang/LLVM etc.
- Multi-Architecture Comparison System
- Hosted for the Consortium by the Met Office in Exeter

University of BRISTOL

UKRI Engineering and Physical Sciences Research Council

GW4

Isambard 2's A64FX system

- HPE Apollo 80 system with A64FX CPUs from Fujitsu
  - 72 nodes connected with 100 Gbps IB
    - 3,456 cores, 72 TB/s memory bandwidth, 202 TFLOP/s 64-bit f.p.
  - Comes with a Cray software stack
    - CCE, GNU, Arm, Fujitsu Compilers

University of BRISTOL

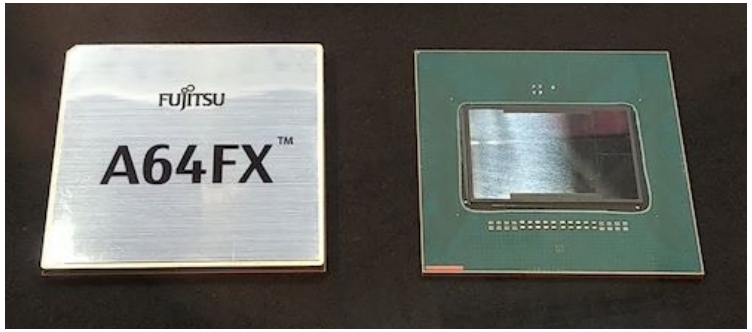
UKRI Engineering and Physical Sciences Research Council

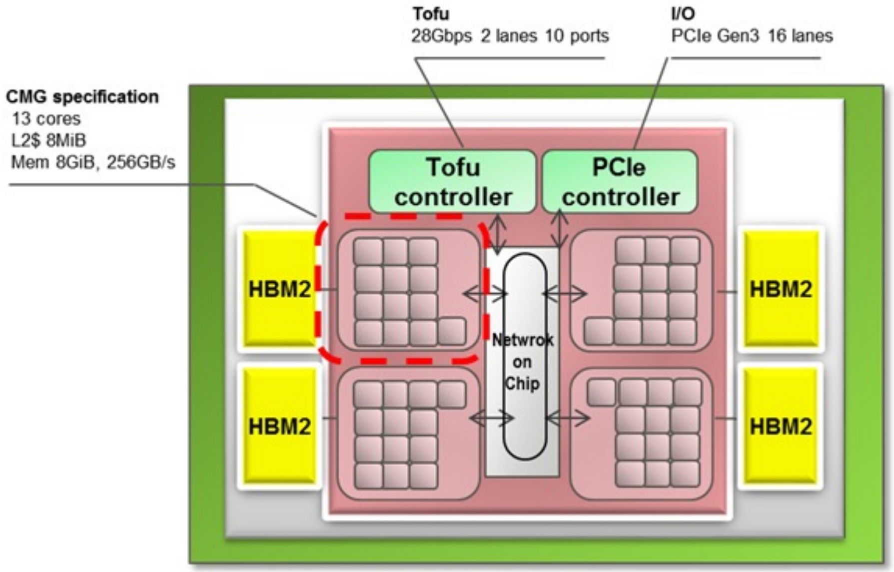
Fujitsu's A64FX

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

https://github.com/fujitsu/A64FX

- 48 cores, 1.8 – 2.2GHz
  - 4 NUMA regions (CMGs)
- >2.7 TFLOP/s double precision
- 2x 512-bit vector pipelines per core
  - ARMv8.3-A + SVE
- 1 TByte/s main memory bandwidth
  - 4 stacks of HBM2
- ~170 Watts
- High speed interconnect
- 8.7B transistors, 7 nm





University of BRISTOL

GW4

A new generation of Arm-based processors

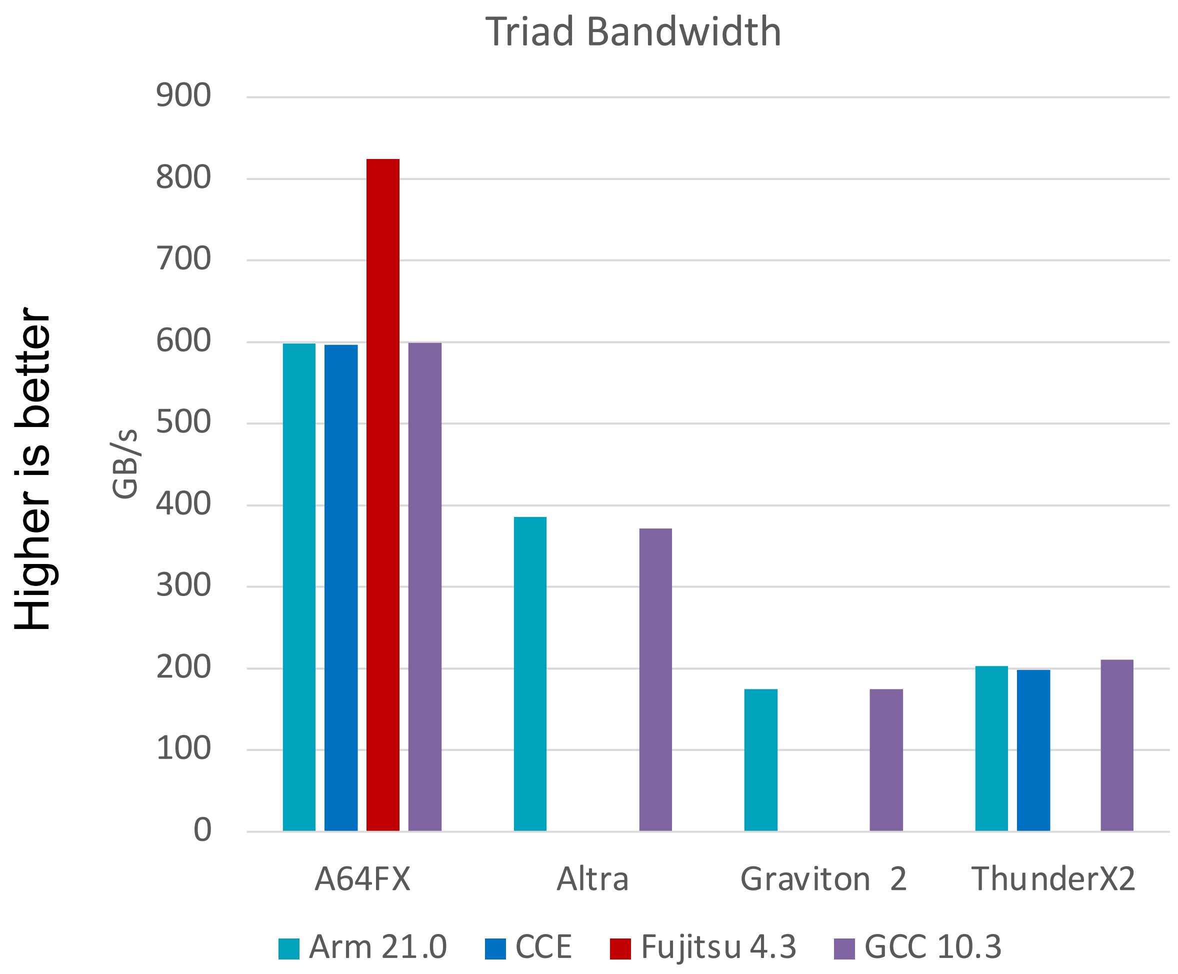
Platform	Cores	Clock Speed	Peak FLOP/s (d.p.)	Peak memory BW
Fujitsu A64FX	1 x 48	1.8 GHz	2.8 TFLOP/s	1,024 GB/s
AWS Graviton 2	1 x 64	2.5 GHz	1.3 TFLOP/s	205 GB/s
Ampere Altra	2 x 80	3.0 GHz	3.8 TFLOP/s	410 GB/s
Marvell ThunderX2	2 x 32	2.5 GHz	1.3 TFLOP/s	320 GB/s

No special configuration required on any platform

- Arm and GNU compilers support all
- Cray supports A64FX† and TX2
- Fujitsu supports A64FX
  - Traditional (Trad) and Clang mode

† SVE support in CCE is still beta

BabelStream – memory bandwidth benchmark



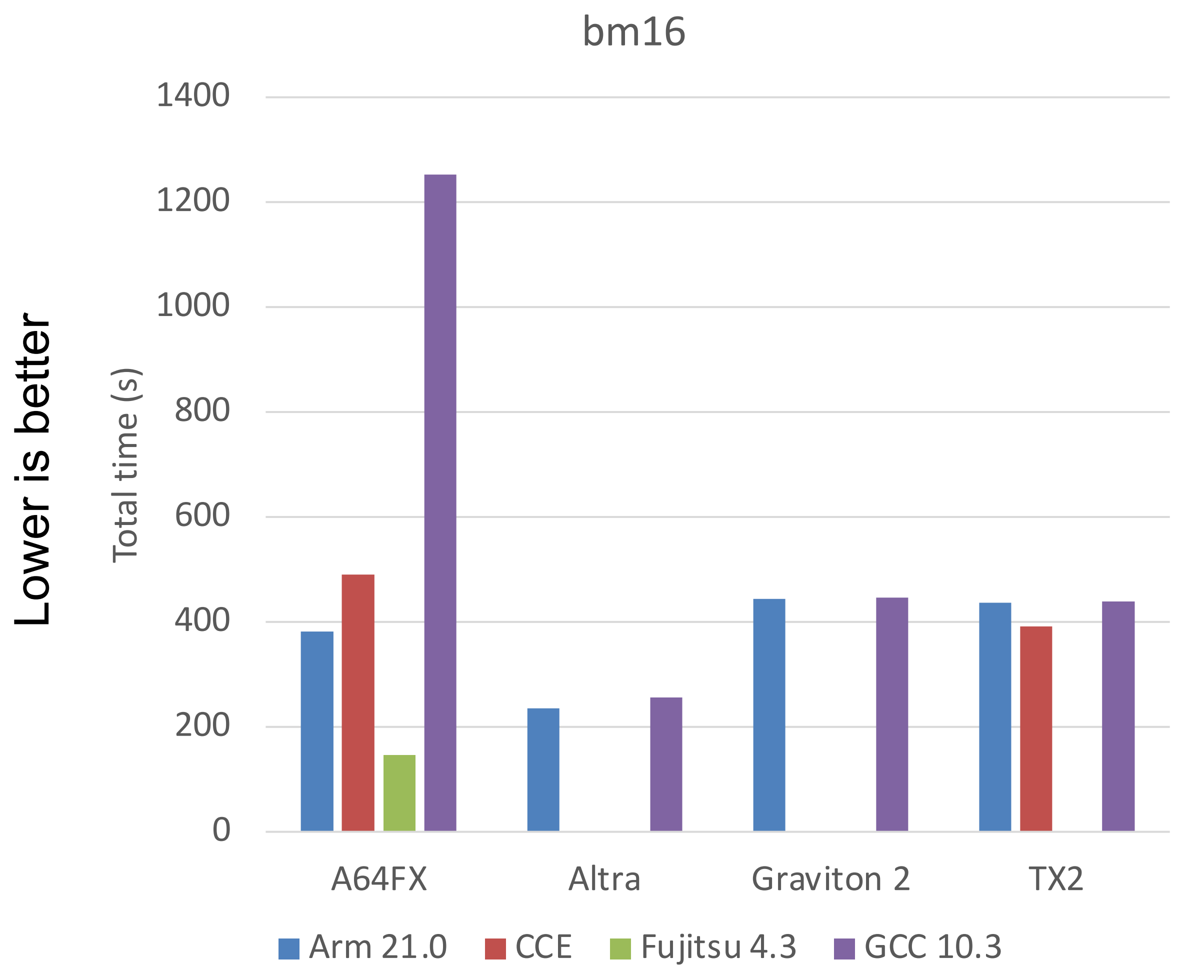
- A high fraction of A64FX's HBM2 peak bandwidth is achievable in practice
- In order to avoid reads into cache before streaming writes, FCC uses zero-fill instructions
  - Compiler flags can be used to tune prefetching if desired

University of BRISTOL

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

GW4

CloverLeaf



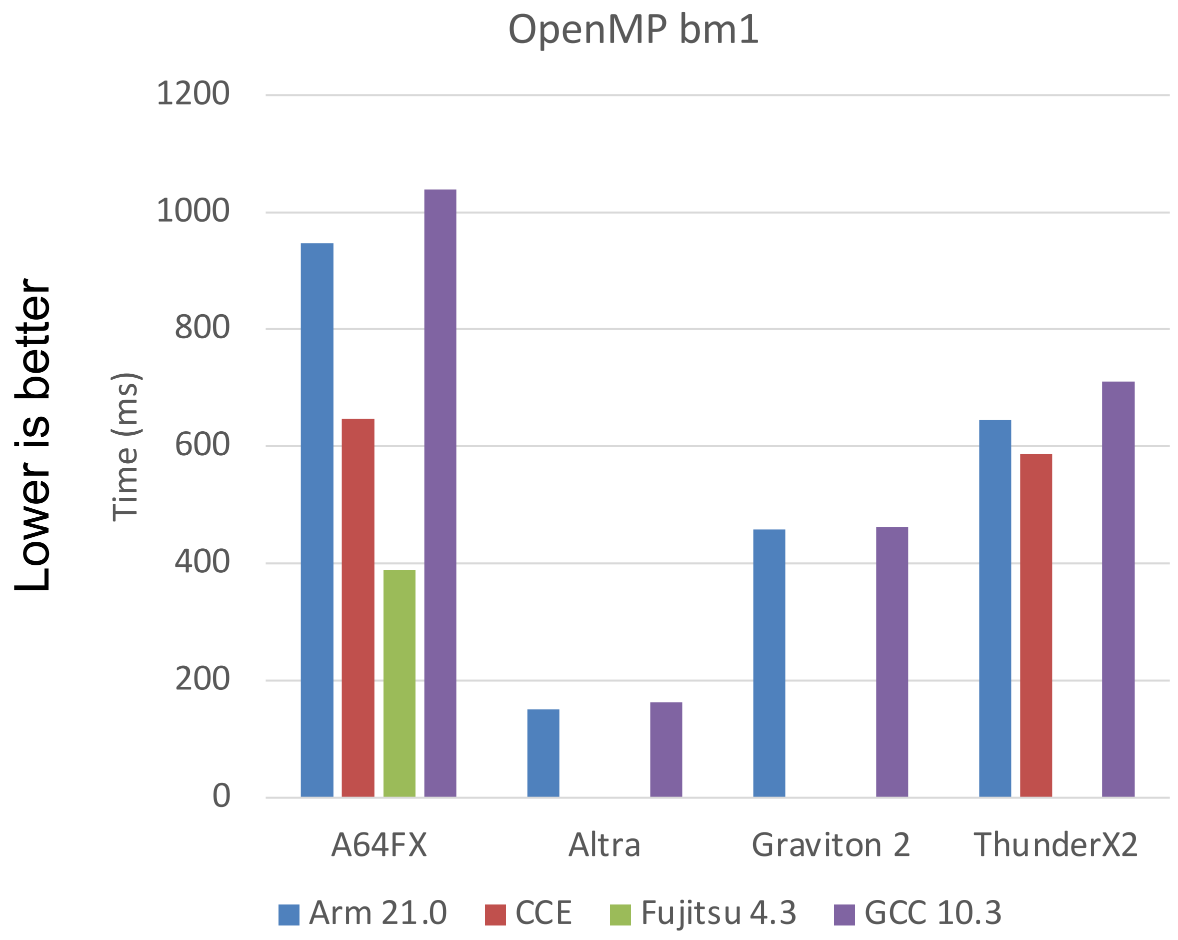
- CloverLeaf is generally bandwidth-bound, but there are also divisions and trigonometry
  - With the Fujitsu Compiler, the benefit of the HBM2 is clear
  - Arm compiler optimises division (`-fiterative-reciprocal`)
  - GCC produces bad SVE vector code
- CloverLeaf is hybrid MPI+OMP
  - Results here are the best in each case

University of BRISTOL

https://github.com/UK-MAC/CloverLeaf_ref

GW4

miniBUDE: Molecular Docking



- miniBUDE is heavily compute-bound [1]
  - Achieves ~60% of peak FLOP/s on CXL
- On A64FX it benefits greatly from compiler unrolling, interleaving, and software pipelining
- On Graviton 2 and TX2, short vectors are the main limitation
- Altra performance is competitive with Cascade Lake and Rome

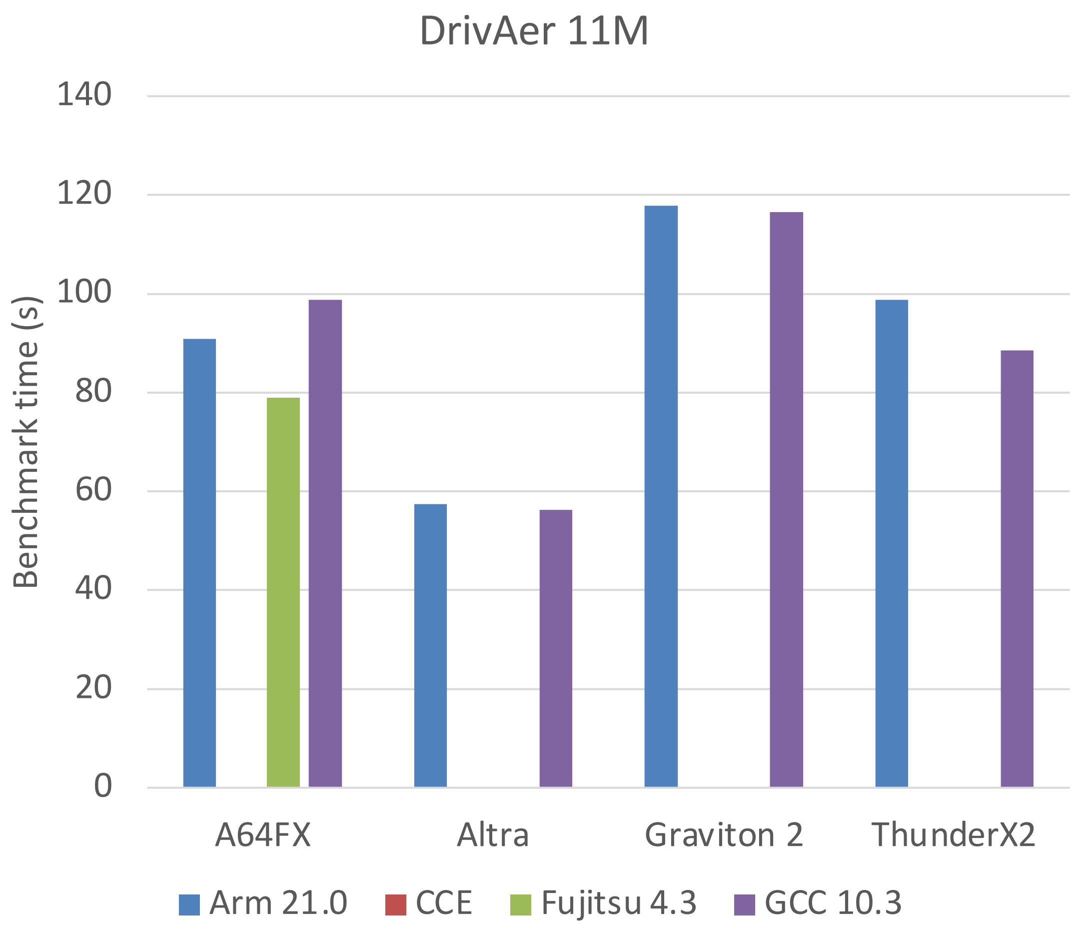
University of BRISTOL

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

GW4

OpenFOAM

Lower is better



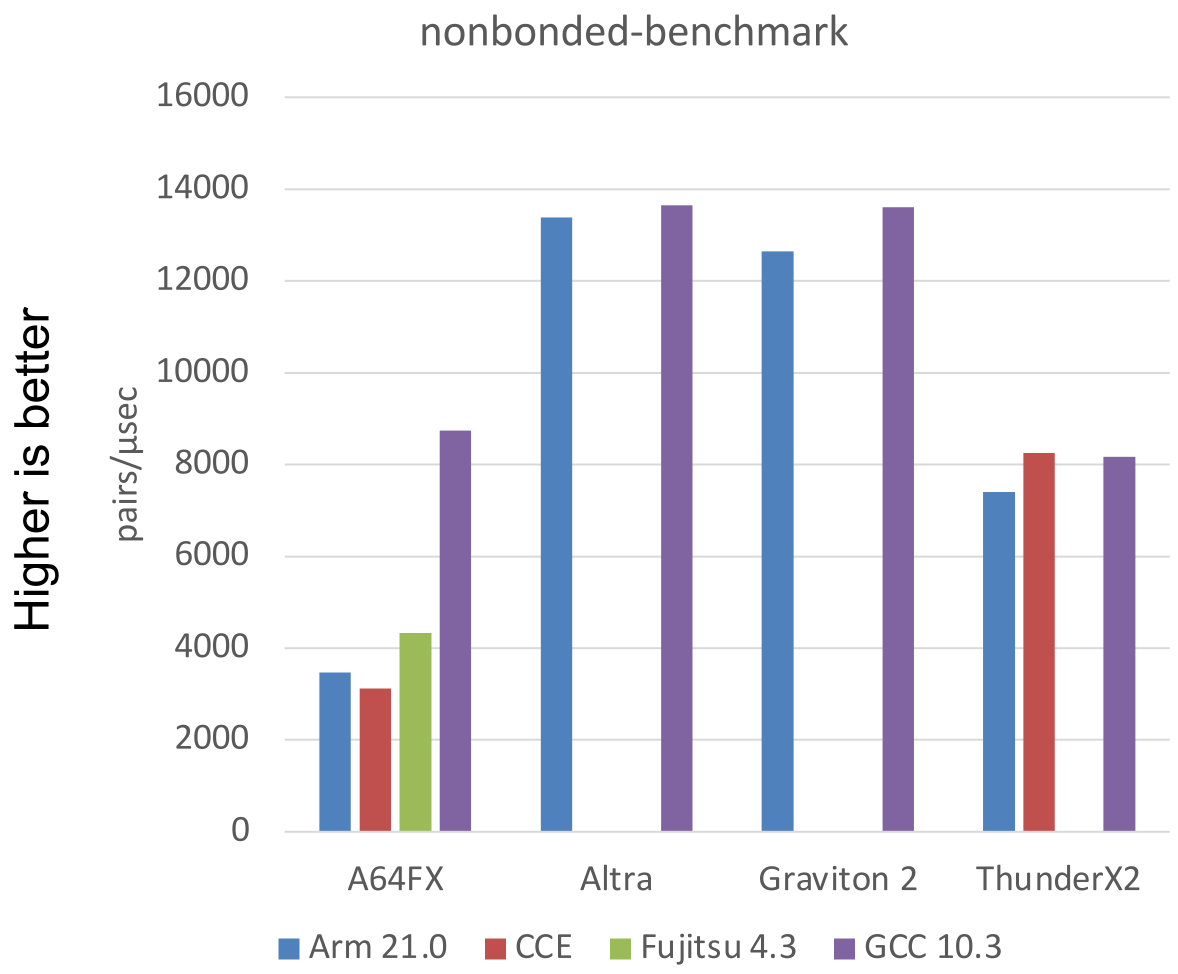
- OpenFOAM v2006, DrivAer test case, `simpleFoam` solver
- Benchmark time: last – first
- CCE: compilation errors
- In general very memory-bandwidth-bound, but compute performance does matter
  - Vectorisation is low
- A64FX result faster than TX2, but shows external memory bandwidth is not every thing
- Rome and Altra are the fastest
  - Both have large caches…

University of BRISTOL

https://www.openfoam.com

GW4

GROMACS



- 2021.1, integrated PME benchmark
  - Limited to 64 threads
- Only GNU supported for SVE
  - With the other compilers, need to use NEON version on A64FX
- A 40 core Cascade Lakes platform achieves around 28,000 pairs/usec, so A64FX should be able to achieve around 39,000 pairs/usec

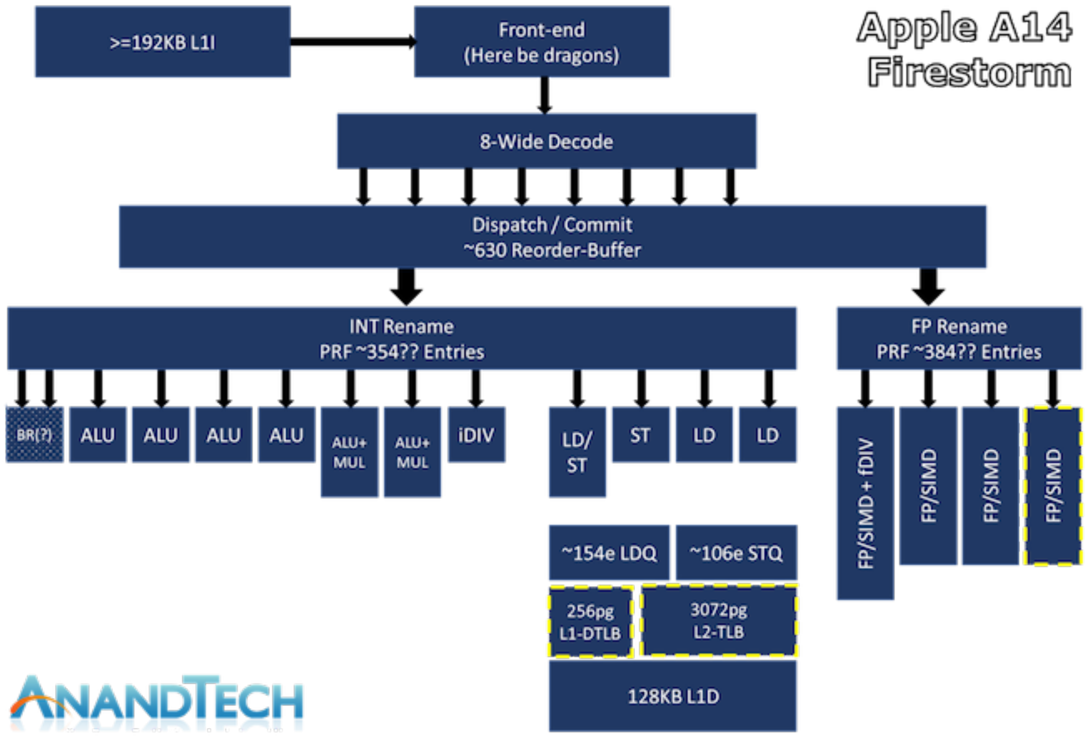
Experience with workloads on Isambard 2's A64FX

- Everything works out-of-the-box!
  - The same experience we had with ThunderX2 in Isambard 1
  - No specific programming model or language needed
- Compilers and libraries are already available
  - Cray, Arm, Fujitsu support A64FX
  - GCC support in 10.3 and 11 (released late April '21)
  - All already stable but room to improve performance
- Optimised libraries and higher-level frameworks are continuously being improved

University of BRISTOL

GW4

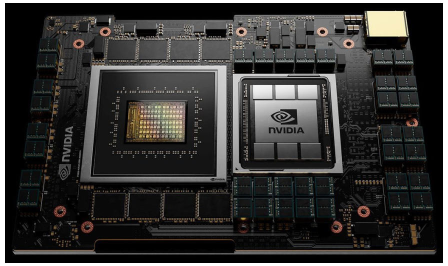
Arm performance looking strong – e.g. Apple's latest core in M1 / A14



- 8-way superscalar, 2.9GHz, 5nm
- Units: 7 integer, 4 load/store, 2 branch, 4 FP/vector
- 192KB L1 I-cache (huge!)
- 128KB L1 D-cache (also big)
- 8MB L2 cache (shared by cores)
- 16MB LLC (system level)
- Class-leading single thread performance, faster than most of the latest CPUs from Intel and AMD

https://www.anandtech.com/show/16226/apple-silicon-m1-a14-deep-dive/2

Recent announcement: NVIDIA's 'Grace' Arm-based CPUs



- Lots of fast Arm Neoverse™ cores
- 450GB/s x2 for direct, cache coherent GPU links
- ~500GB/s per socket to energy-efficient LPDDR5 memory
- Performance "*over 300 on the SPECrate2017_int_base test, which is on par with a pair of 28-core "Skylake" and "Cascade Lake" Xeon SP processor from Intel.*" – The Next Platform, April 12th 2021
- Shipping in 2023

https://www.nextplatform.com/2021/04/12/nvidia-enters-the-arms-race-with-homegrown-grace-cpus/

University of BRISTOL

GW4

Summary

- **Arm processors have proven competitive in HPC since 2018**
- Today A64FX looks very promising, beating cutting-edge dual-socket nodes in many tests
  - Easy to use – in most cases running unmodified flat MPI, or hybrid MPI+OpenMP
  - Performance in same ballpark as GPUs, but with a significantly lower barrier to entry in ease of use
- **Arm processors have exciting roadmaps** in both HPC **and** cloud
- Isambard 2 makes most of the major technologies available in one place, enabling rigorous comparative benchmarking

For more information

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

Isambard: https://gw4-isambard.github.io/

Email: **S.McIntosh-Smith at bristol.ac.uk**

Twitter: **@simonmcs**

[1] Andrei Poenaru, Wei-Chen Lin and Simon McIntosh-Smith. 'A Performance Analysis of Modern Parallel Programming Models Using a Compute-Bound Application'. In: 36th International Conference, ISC High Performance 2021. Frankfurt, Germany, 2021. In press.