Alternative computing architectures in ATLAS

Gen Kawamura, Arnulf Quadt, **Joshua Wyatt Smith**

11th Annual Helmholtz Alliance Workshop
28.11.2017

Contents

- Why alternative architectures?

- Recent developments and benchmarks

  - Cavium ThunderX2

  - Qualcomm Centric 2400

- ARM in ATLAS

  - Previous studies with ATLAS

  - Where are we now?

- Conclusions

Why alternative architectures?

- Explore new hardware (alternative architectures) that in the future may be more suited to specific task

- Encourages better overall code quality

- Potentially more efficient computing - less energy/computation

- Geopolitics - will all countries use x86?

- Potential for more opportunistic resources

- Business model of ARM is very flexible

  - No silicon produced by ARM

  - IP sold

- Competition, freedom, flexibility

Why alternative architectures?

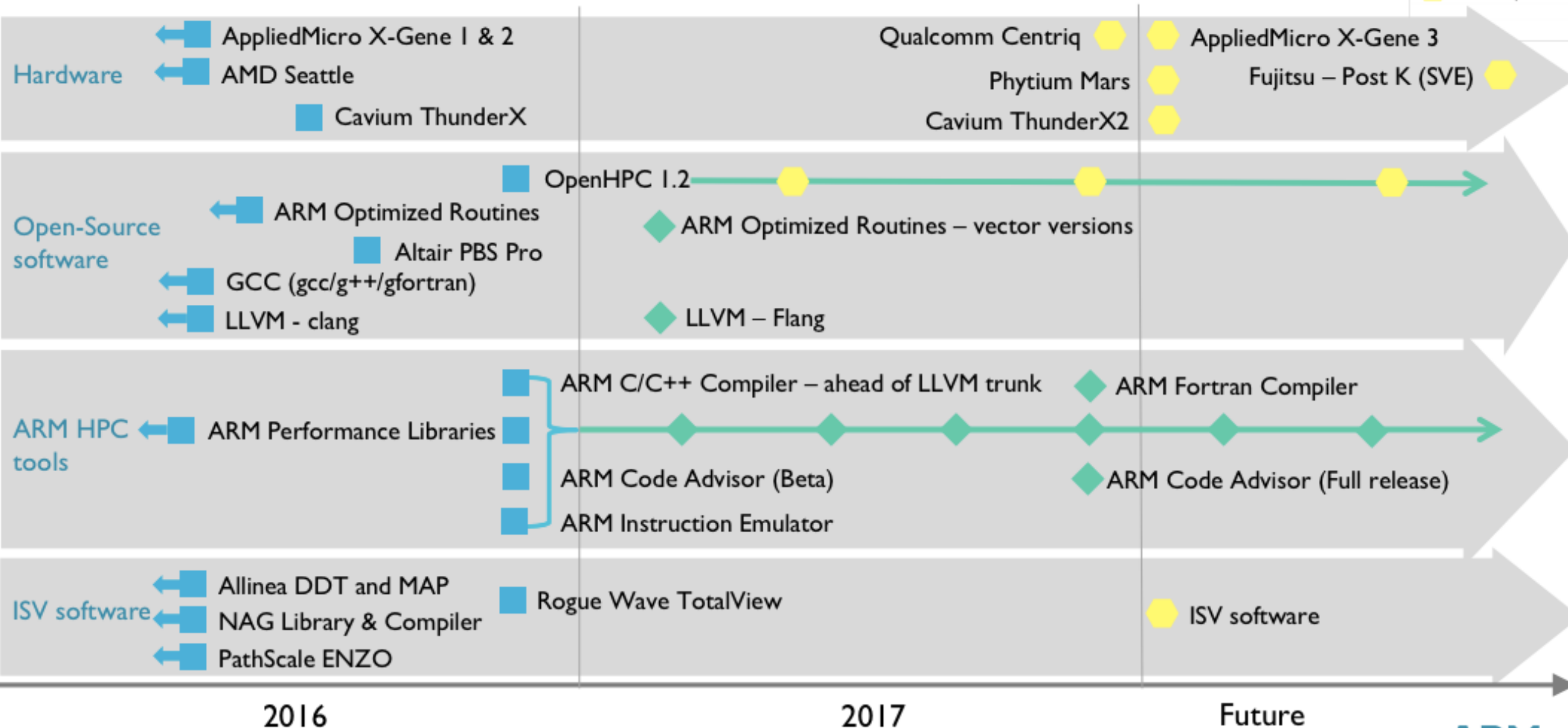**Arm vs x86: Architectural footprint**
Annual shipments (Mu)

■ X86 ■ Arm

100B+
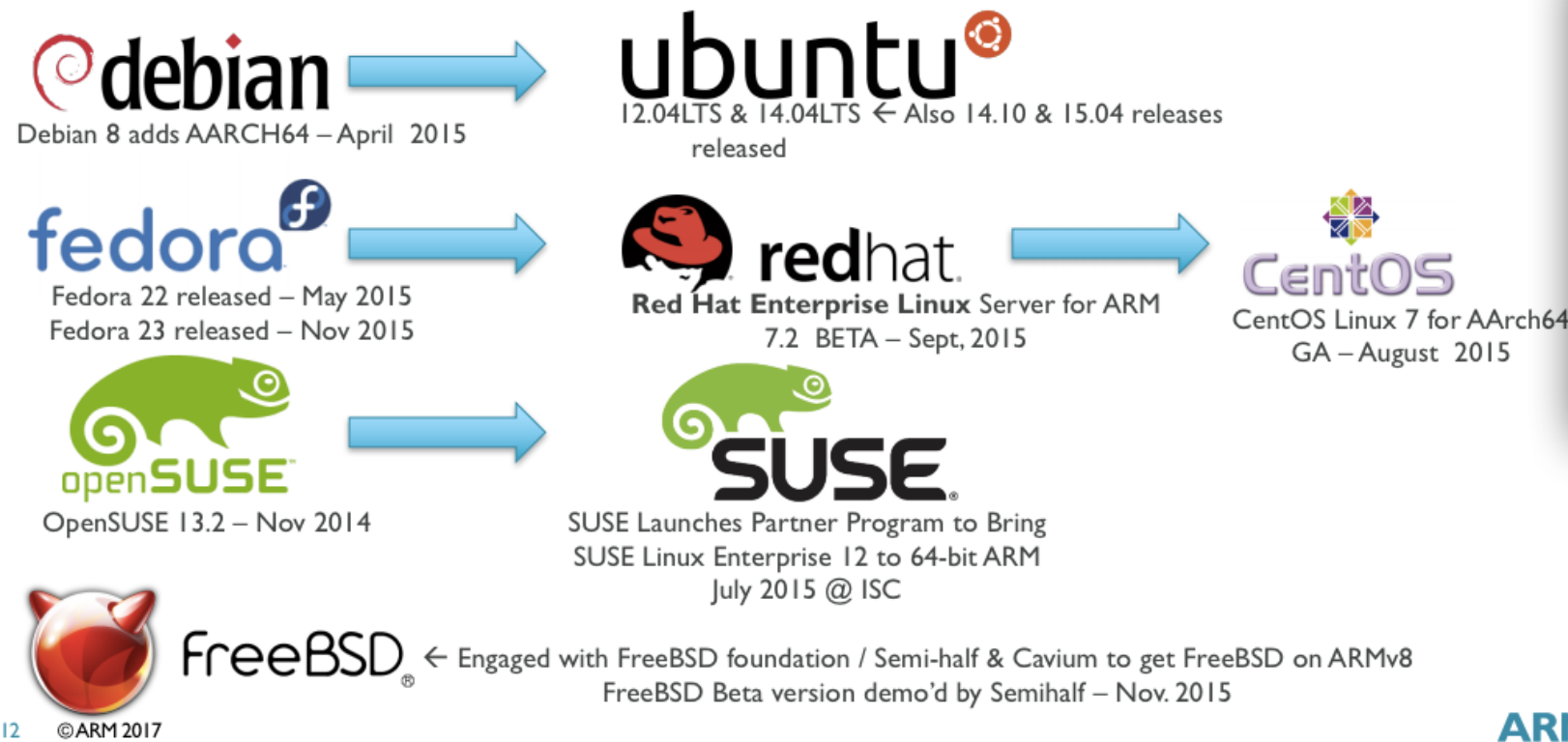Arm-based chips have shipped since 1991
Expect next 100B to ship by 2021
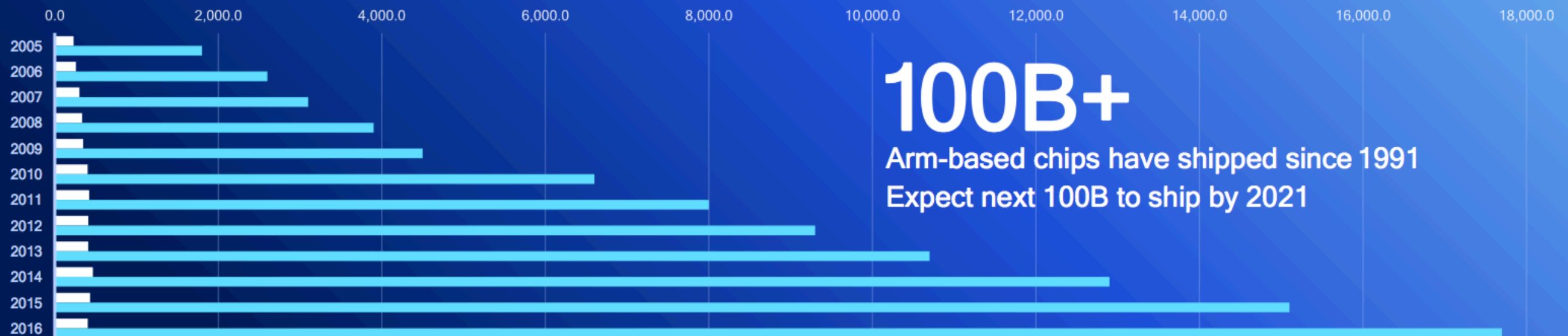
Source: IDC, September 2017.

- Arm-based chip can mean many things!
  - Micro-controllers
  - Embedded devices
  - Servers
  - …

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

OS's

debian
Debian 8 adds AARCH64 – April 2015

ubuntu
12.04LTS & 14.04LTS ← Also 14.10 & 15.04 releases released

fedora
Fedora 22 released – May 2015
Fedora 23 released – Nov 2015

redhat.
Red Hat Enterprise Linux Server for ARM
7.2 BETA – Sept, 2015

CentOS
CentOS Linux 7 for AArch64
GA – August 2015

openSUSE
OpenSUSE 13.2 – Nov 2014

SUSE.
SUSE Launches Partner Program to Bring
SUSE Linux Enterprise 12 to 64-bit ARM
July 2015 @ ISC

FreeBSD ← Engaged with FreeBSD foundation / Semi-half & Cavium to get FreeBSD on ARMv8
FreeBSD Beta version demo'd by Semihalf – Nov. 2015

12 ©ARM 2017

ARM

Filesystems

Software	Status
LUSTRE	Ported
HDFS	Ported
CEPH	Ported
BeeGFS	Ported

Compilers

- **LLVM**
  - C, C++
  - OpenMP 3.1, (4.0 coming soon)
  - Fortran coming Q1 2017

- **NAG**
  - Fortran
  - OpenMP 3.1

- **PathScale**
  - C, C++ Fortran
  - OpenACC
  - OpenMP 4.0

- **ARM C/C++ Compiler**
  - LLVM based
  - Includes SVE

- **GCC**
  - C, C++, Fortran
  - OpenMP 4.0

Joshua Wyatt Smith

6

Recent developments

- Mont-Blanc project using Cavium ThunderX2

- Fujitsu Post-K ARM supercomputer, (more documentation)

- Phytium's Mars processor

- Cray Inc. building 10,000 core cluster with ARM CPUs at University of Bristol - "Isambard"

- Qualcomm unveils Falkor CPU core for the Centriq 2400 SoC
  - Worlds first 10nm server processor!

- …

University of BRISTOL   Isambard   GW4

Isambard system specification (red = new info):

- Cray "Scout" system – XC50 series
  - Aries interconnect
- **10,000+** Armv8 cores
  - Cavium ThunderX2 processors
  - 2x 32core @ >2GHz per node
- Cray software tools
- Technology comparison:
  - x86, Xeon Phi, Pascal GPUs
- Phase 1 installed March 2017
- The Arm part arrives early 2018

I.K.Brunel 1804-1859

Cavium ThunderX2

Setup	# cores	Clock Speed	Memory	Fabrication	TDP
ThunderX2 alpha release	32 or 28	2.5 GHz or 2.0 GHz	2667 MHz DDR4 or 2400 MHz DDR4	14 nm	?
Broadwell Xeon E5-2695 v4	18	2.1 GHz	2400 MHz DDR4	14 nm	120 W
Skylake system Xeon Gold 6152	22	2.1 GHz	2667 MHz DDR4	14 nm	140 W

Software used for ThunderX2:

- UM, NEMO and TeaLeaf: Cray CCE 8.6.4
  - UM and NEMO results produced by the UK's Met Office
- GROMACS, OpenFOAM and STREAM: GCC 7.1
- CloverLeaf: 2D armflang 18.0, 3D armflang 1.4
- SNAP: nang=10 armflang 1.4, nang=136 CCE 8.6.3

Software used for Broadwell:

- UM and NEMO: Cray CCE 8.5.8, produced by the Met Office
- GROMACS: GCC 7.1
- Intel 2017 compiler for everything else

Software used for Skylake:

- Intel 2018 compiler

- GROMACS: Molecular dynamics package
- OpenFoam: Fluid dynamics software
- Nemo: Ocean modelling code
- UM: Climate modelling

http://www.goingarm.com/slides/2017/SC17/GoingArm_SC17_Bristol_Isambard.pdf

Cavium ThunderX2

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

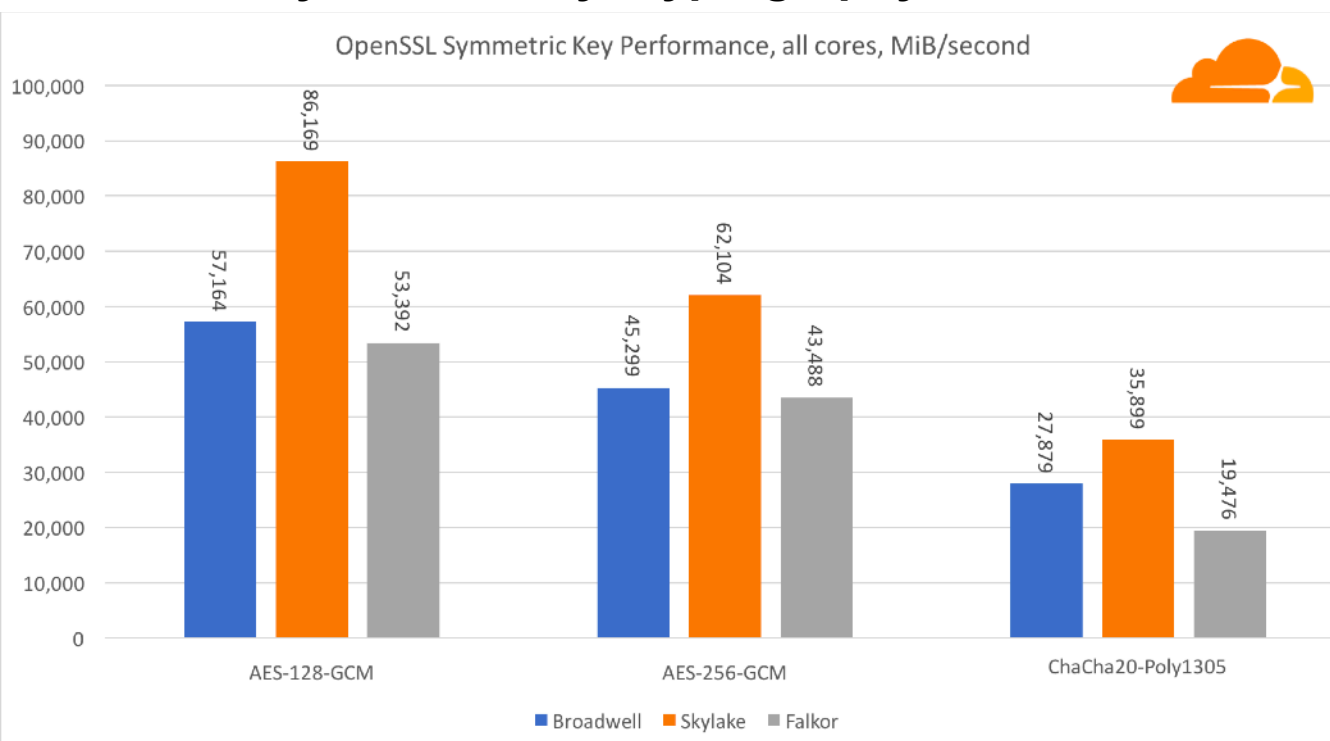mcs   http://gw4.ac.uk/isambard/
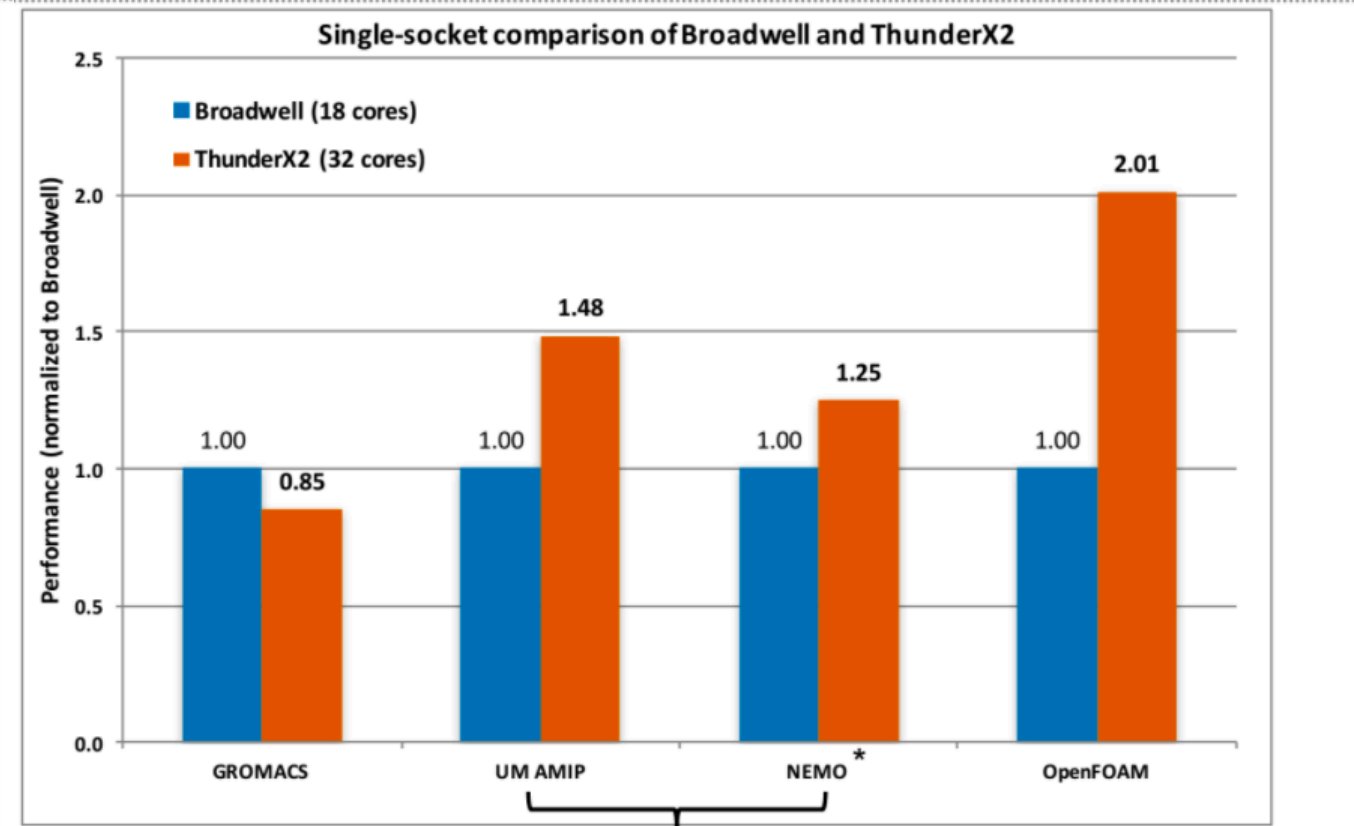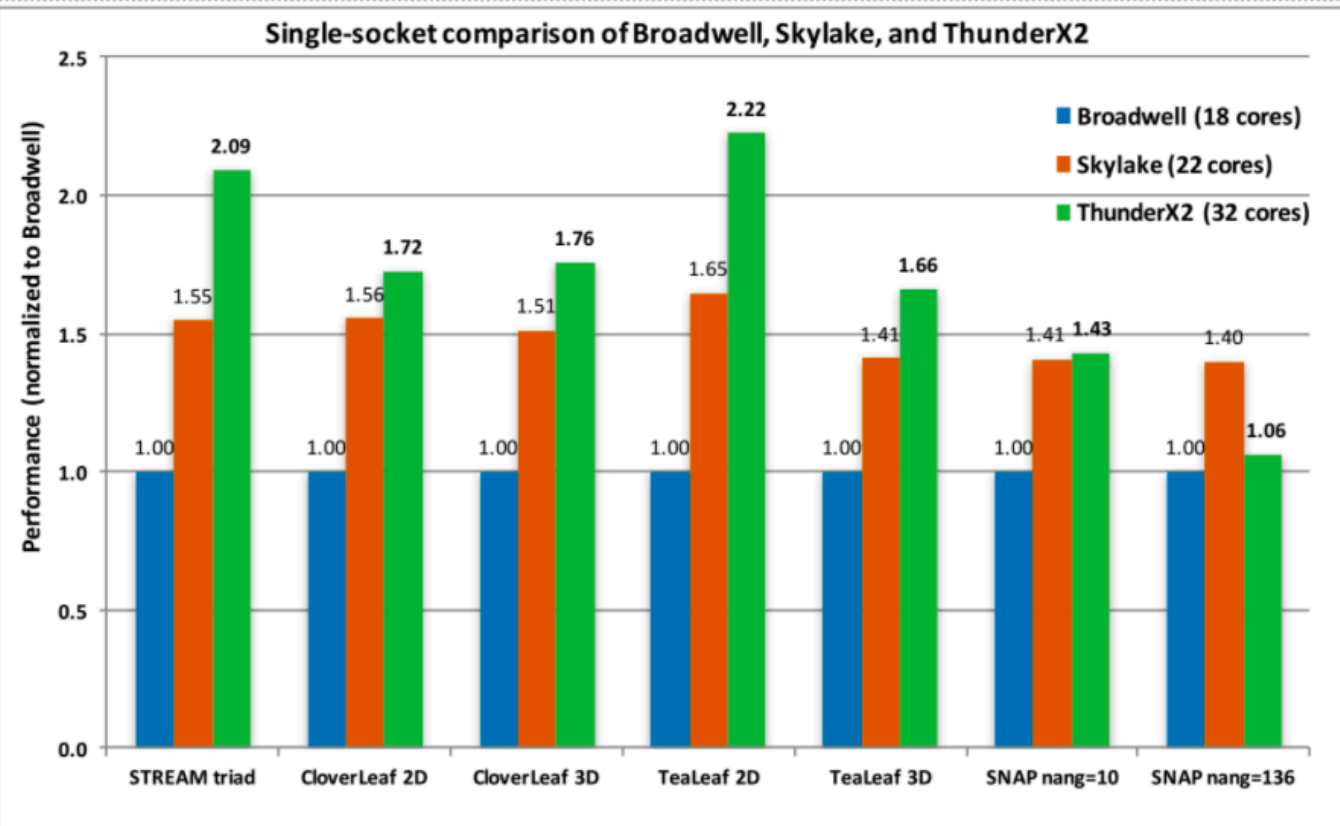
bristol.ac.uk

- In general memory bandwidth dominated benchmarks do better on ThunderX2
- Floating point-heavy benchmarks do better on Skylake and Broadwell
  - Wider vectors

- CPU bound applications closer for different processors due to more cores and higher clock speed for ThunderX2

http://www.goingarm.com/slides/2017/SC17/GoingArm_SC17_Bristol_Isambard.pdf

Qualcomm Amberwing

Setup	# cores	Clock Speed	L3 Cache	Fabrication	TDP
Qualcomm Centriq 2452	46	2.5 GHz	1.25 MB/core, 6 channels	10 nm	120 W
Broadwell Grantley (2016?)	20	2.2 GHz (3.1 GHz turbo)	2.5 MB/core, 4 channels	14 nm	170 W
Skylake Purely (2017)	24	2.1 GHz (3.0 GHz turbo)	1.35 MB/core, 6 channels	14 nm	170 W

Symmetric key cryptography

Gzip compression

https://blog.cloudflare.com/arm-takes-wing/

Joshua Wyatt Smith

10

Qualcomm Amberwing
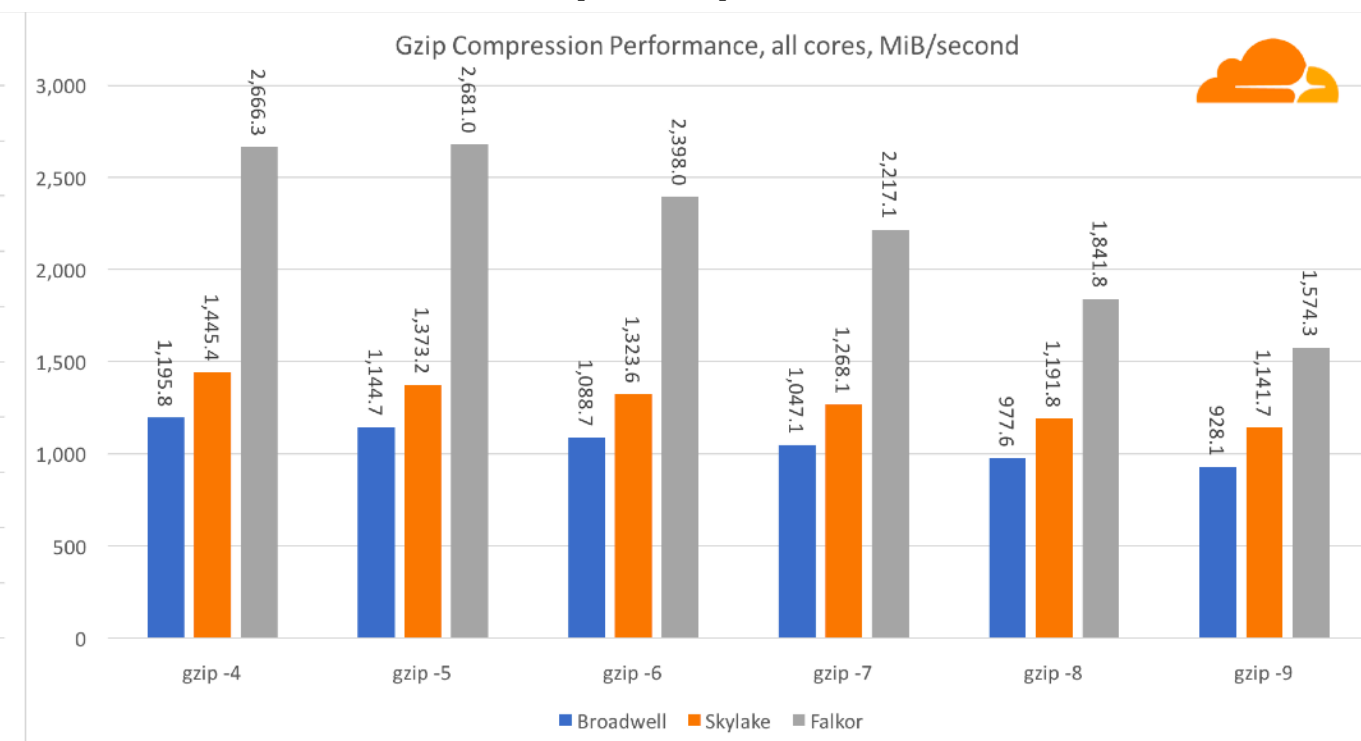
NGINX workload for CloudFlare

Nginx Performance, requests/second and power[W]
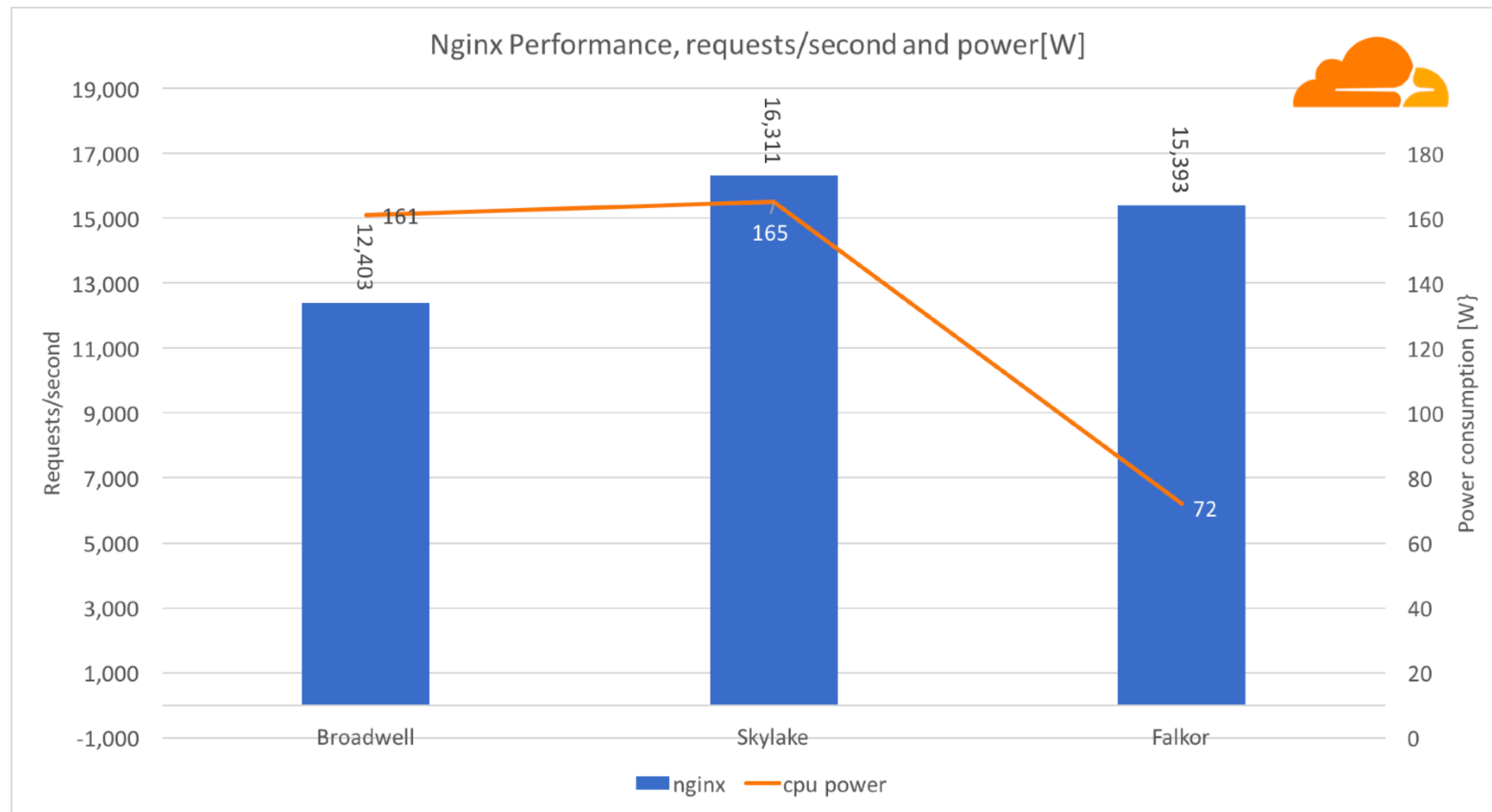
- It really depends on what you benchmark
  - Go libraries perform terribly on the Falkor
- But in general, very impressive results for Falkor

https://blog.cloudflare.com/arm-takes-wing/

But what about in ATLAS or HEP in general?

- We don't really fall into the HPC domain
  - We have different requirements
- Different libraries/OS's need to be optimised, but can gain from HPC work already done
  - I.e. Blas, gcc, clang etc.
- Need RHEL7 (CentOs7)
  - They *do* support ARM64, but not always seamlessly
- Our codebase is also… unique

ARM in ATLAS

- ATLAS's codebase is called *Athena*
- Athena is like the Eierlegende Wollmilchsau… A full release can do everything!

ATLAS ported a subset of Athena to Aarch64
(~400 packages compared to 2500)

~6.5 million lines of code
(written by physicists)

Athena
Gaudi
AtlasExternals
LCG externals

AthSimulation
Gaudi
AtlasExternals
LCG externals

packages like python,
zlib, Root etc.

ARM in ATLAS

- Both ATLAS and CMS have benchmarked many ARM servers over the years
    - 32 bit and 64 bit architectures
    - From generic benchmarks to "the full chain" (event generation, simulation, reconstruction)
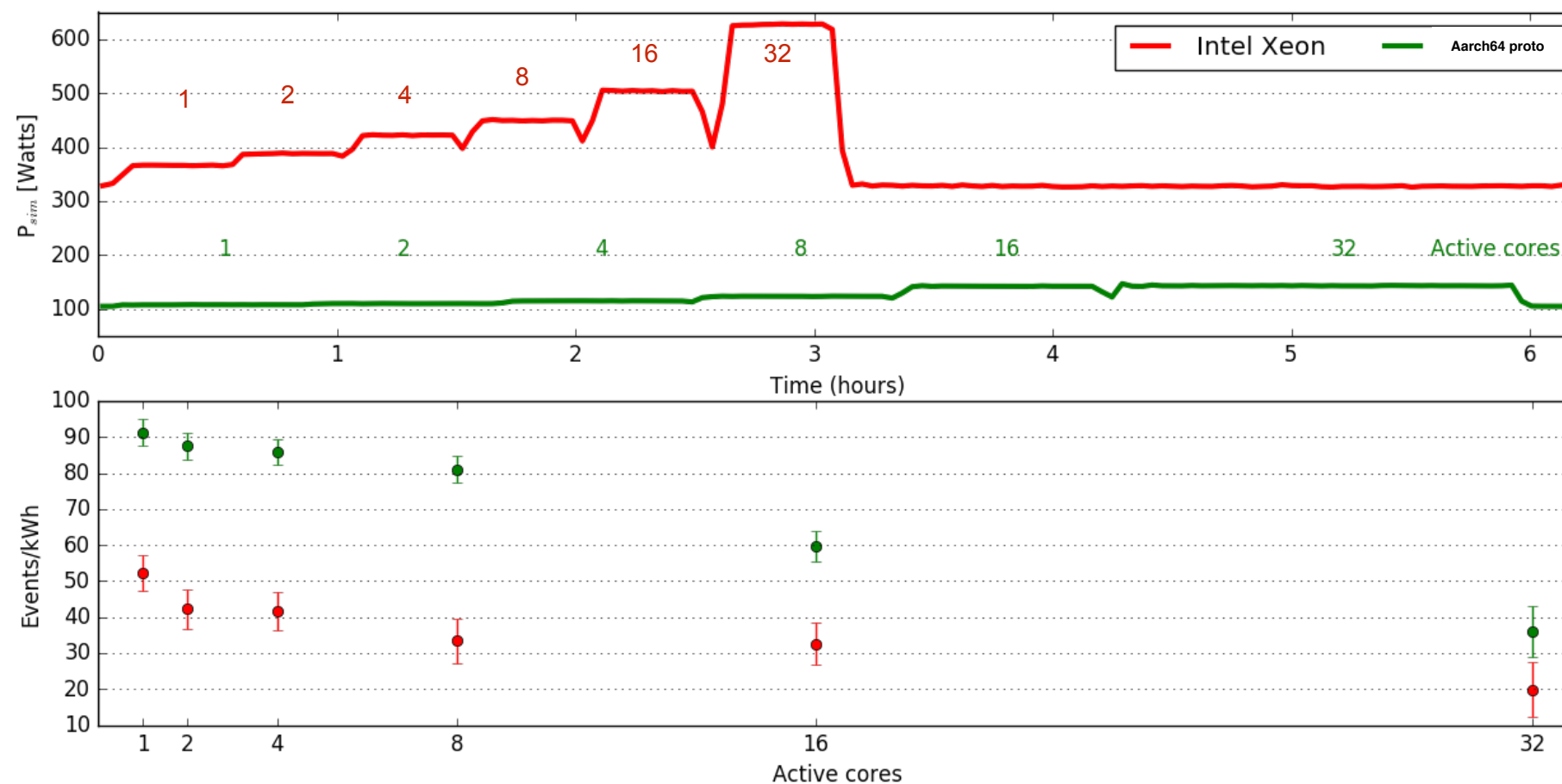
- Both ATLAS and CMS have benchmarked many ARM servers over the years
  - 32 bit and 64 bit architectures
  - From generic benchmarks to "the full chain" (event generation, simulation, reconstruction)

simulated 8 ttbar events while loading an increasing number of cores


Careful: Aarch64 prototype was considerably newer than Intel Xeon E4-4650 (see appendix)

ARM in ATLAS

- Both ATLAS and CMS have benchmarked many ARM servers over the years
  - 32 bit and 64 bit architectures
  - From generic benchmarks to "the full chain" (event generation, simulation, reconstruction)
- Just as importantly what about validation?

Same t$\bar{\text{t}}$ event reconstructed in the ATLAS detector
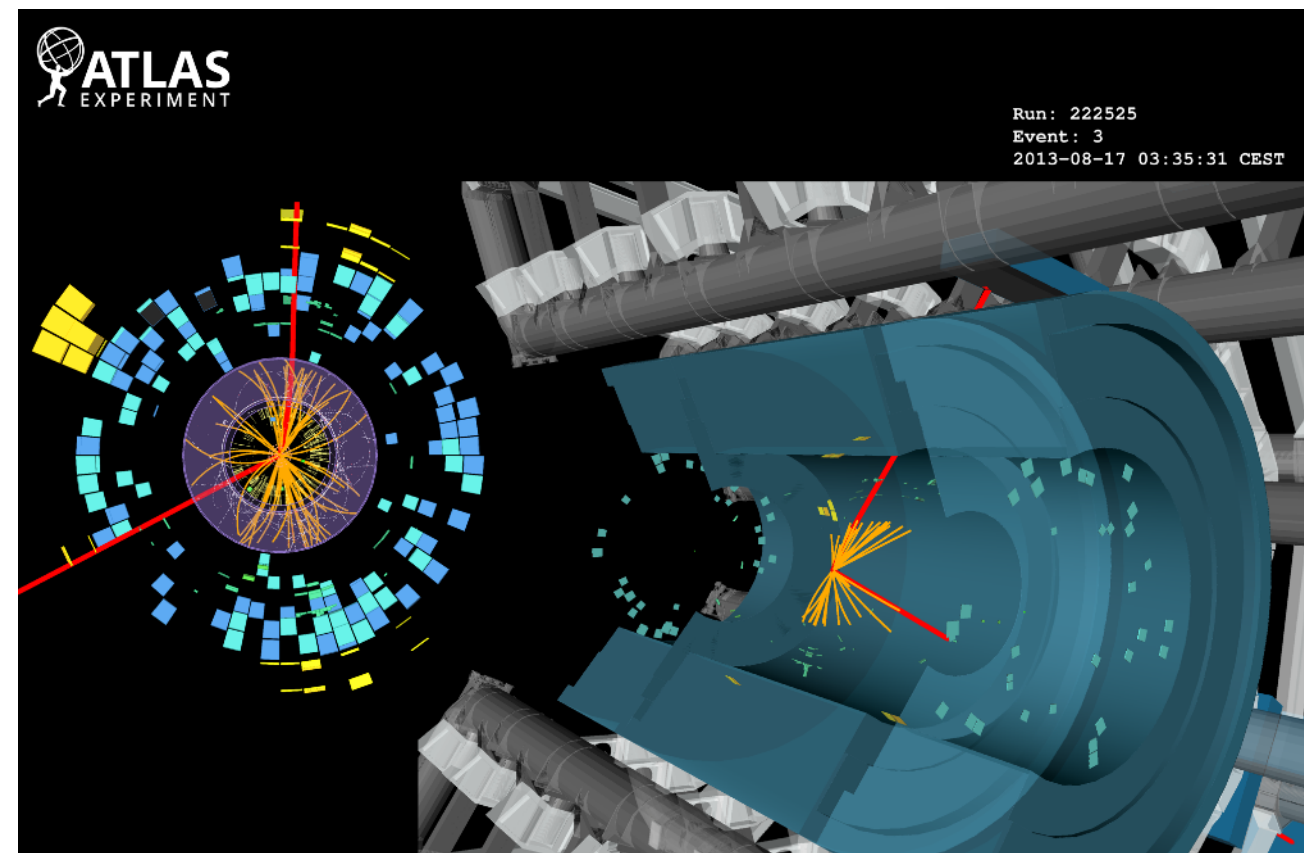
Aarch64 prototype
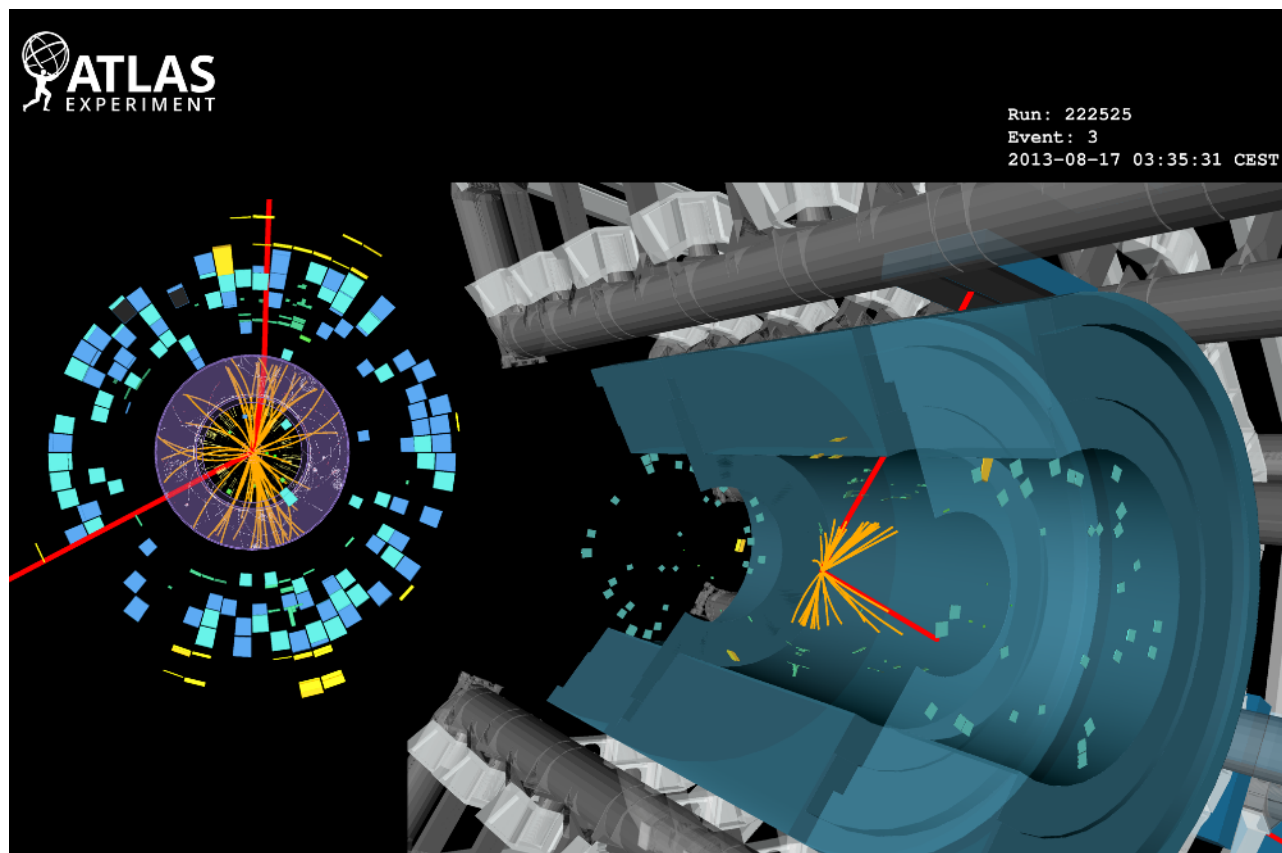
Intel Xeon

ARM in ATLAS

- Both ATLAS and CMS have benchmarked many ARM servers over the years
  - 32 bit and 64 bit architectures
  - From generic benchmarks to "the full chain" (event generation, simulation, reconstruction)
- Just as importantly what about validation?

Hit analysis in the ATLAS Pixel detector

We absolutely 100% expect numerical identity for Intel Atom… except according to Intel we also might not…
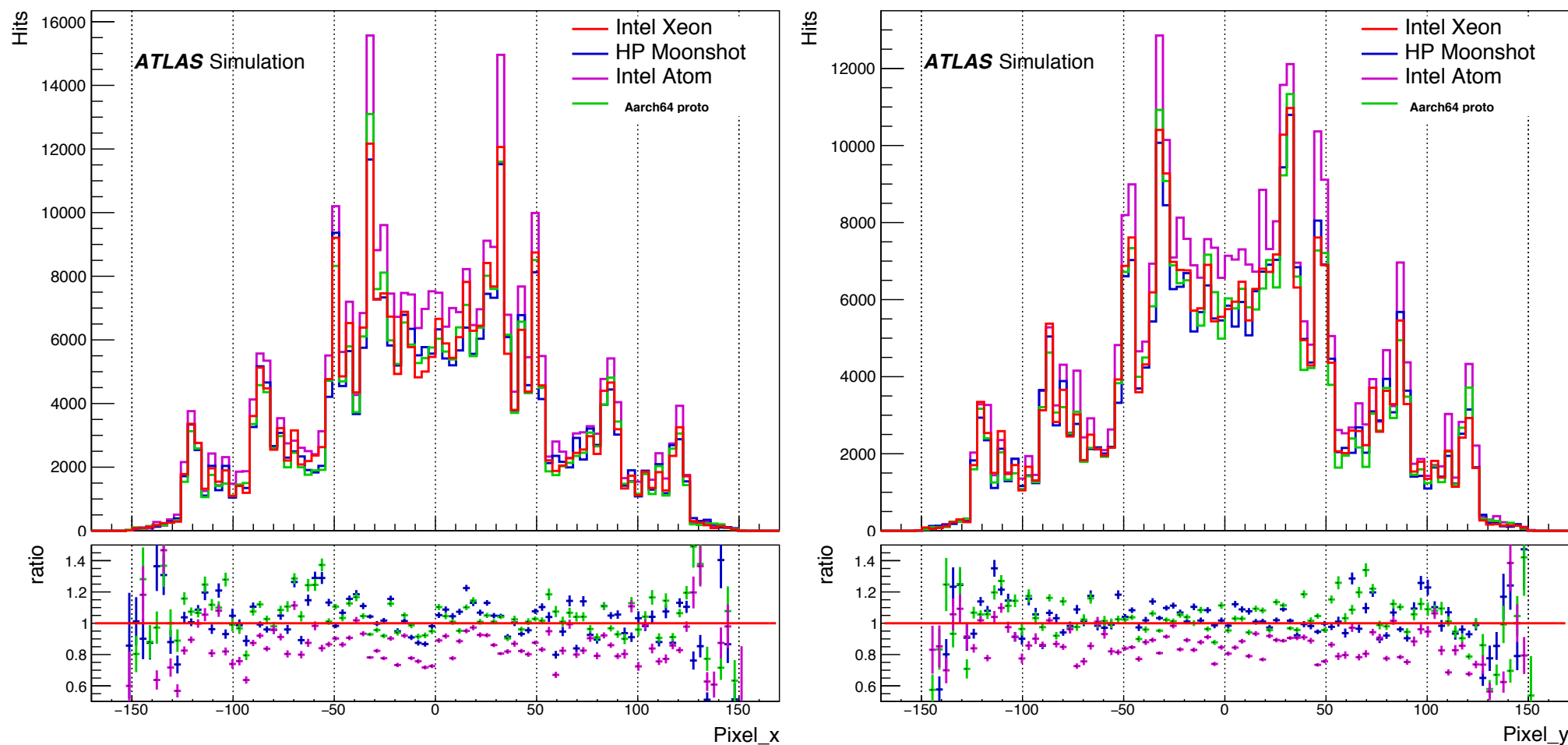Regardless, this got ugly.

ARM in ATLAS

- Both ATLAS and CMS have benchmarked many ARM servers over the years
  - 32 bit and 64 bit architectures
  - From generic benchmarks to "the full chain" (event generation, simulation, reconstruction)
- Just as importantly what about validation?

Hit analysis in the ATLAS Pixel detector

It's also easy to not use the exact run conditions for each full chain test… we needed more stable environment
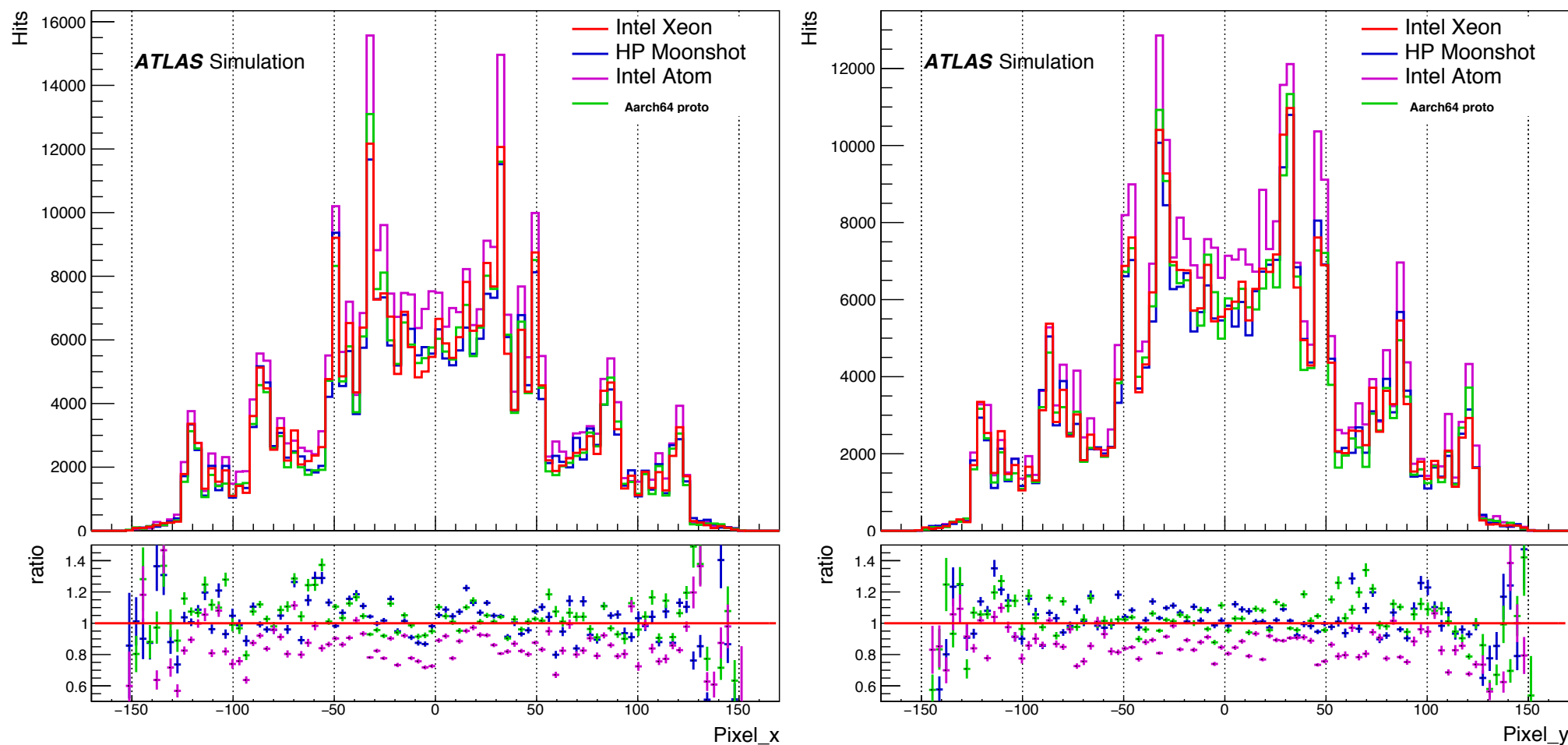
GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

- Both ATLAS and CMS have benchmarked many ARM servers over the years
  - 32 bit and 64 bit architectures
  - From generic benchmarks to "the full chain" (event generation, simulation, reconstruction)
- Just as importantly what about validation?

- Each generation of ARM 64-bit servers got better
  - Till we decided past some point this was largely pointless

- The trend is clear. We want to officially support the Aarch64 architecture
- In the middle of this year ATLAS, CMS and LHCb have officially requested CERN IT to support aarch64 packages
  - Puppet libraries
  - Security libraries
- The goal is to be able to plug in an ARM server into the build farm, and *seamlessly* build and test our respective code bases
- Many details still currently being worked on
  - What does "support" actually mean?
  - Operating system?
  - …

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

- Alternative architectures are becoming more widely used and available
- We like ARM's business model
  - Encourages innovation and competitive prices
- Qualcomm and Cavium are have some very nice ARM 64-bit servers
  - Competitive with latest Intel servers

- ATLAS, CMS, LHCb have officially asked CERN IT to support the ARM 64-bit architecture
 - Details being hashed out
  - Stay tuned…

Thanks for your attention

Backup slides

Name	Processor	Cores	RAM	Cache	Fabrication (Release)	OS
HP Moonshot	X-Gene, 2.4 GHz	8 Armv8	64 GiB DDR3 (1600 MHz)	32 KiB L1/core, 256 KiB L2/core pair, 8 MiB L3	40 nm (2014)	Ubuntu 14.04
Aarch64_Proto	2.1 GHz	32 Cortex-A57	128 GiB DDR3 (1866 MHz)	32 KiB L1, 1 MiB L2	16 nm (-)	Ubuntu 14.04
Intel Atom	Intel Atom Processor C2750, 2.4GHz	8	32 GiB DDR3 (1600 MHz)	24 KiB L1d, 32 KiB L1i, 1 MiB L2	22 nm (2013)	Fedora 21
Intel	Intel Xeon CPU E5-4650, 2.70 GHz	32	512 GiB DDR3 (1600 MHz)	32 KiB L1(d)(i)/core, 256 KiB L2/core, 20 MiB L3	32 nm (2012)	Scientific Linux CERN 6

Joshua Wyatt Smith

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

What's the vector length?

- There is **no** preferred vector length
  - Vector Length (VL) is hardware choice, from 128 to 2048 bits, in increments of 128
  - Does not need to be a power-of-2
  - *Vector Length Agnostic* programming adjusts dynamically to the available VL
  - No need to recompile, or to rewrite hand-coded SVE assembler or C intrinsics
 - Has extensive implications for loop optimizations
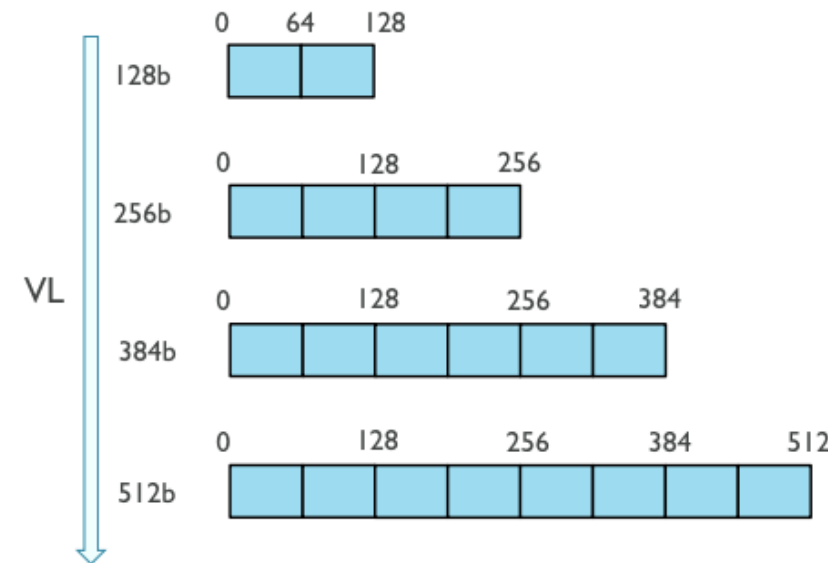
Introducing the Scalable Vector Extension (SVE)

A vector extension to the ARMv8-A architecture; its major new features:

- Gather-load and scatter-store
- Per-lane predication
- Predicate-driven loop control and management
- Vector partitioning and SW managed speculation
- Extended integer and floating-point horizontal reductions

SVE is **not** an extension of Advanced SIMD

- A separate architectural extension with a new set of A64 instruction encodings
- Focus is HPC scientific workloads, not media/image processing

Joshua Wyatt Sm

www.qualcomm.com
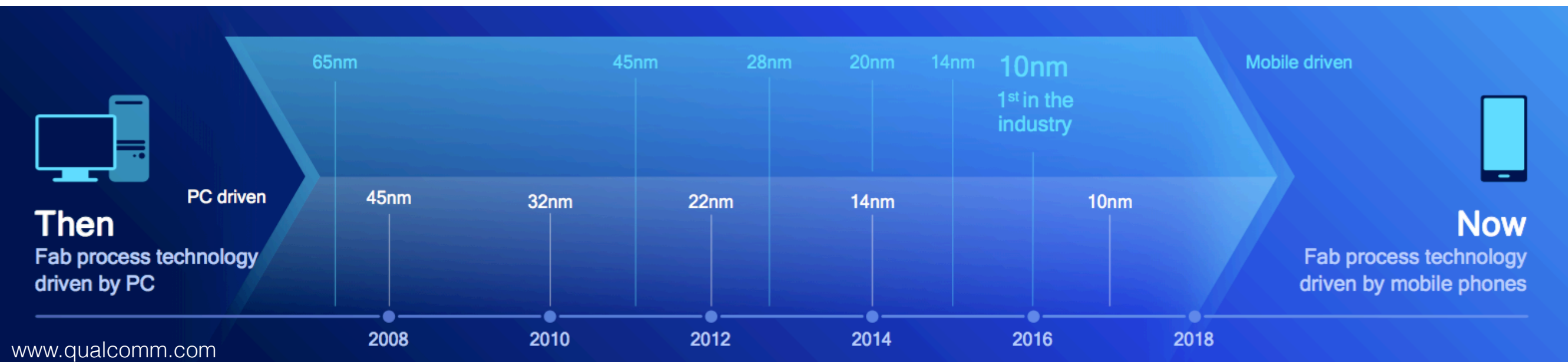
Comparative SKU lineup

Intel Xeon – Top bin Platinum, Gold, Silver SKUs*

SKU	Core Count	L3 Cache	Frequency (base and turbo** freq)	Power (TDP)
Intel Xeon Platinum 8180	28	38.5 MB	2.5 / 3.8 GHz	205W
Intel Xeon Gold 6152	22	30.25 MB	2.1 / 3.7 GHz	140W
Intel Xeon Silver 4116	12	16.5 MB	2.1 / 3.0 GHz	85W

Qualcomm Centriq Processor SKUs

SKU	Core Count	L3 Cache	Frequency (base and peak** freq)	Power (TDP)
Qualcomm Centriq 2460	48	60 MB	2.2 / 2.6 GHz	120W
Qualcomm Centriq 2452	46	57.5 MB	2.2 / 2.6 GHz	120W
Qualcomm Centriq 2434	40	50 MB	2.3 / 2.5 GHz	110W

*Source: https//ark.intel.com. ** Intel Xeon Turbo frequency is lower when all cores are running; Qualcomm Centriq 2400 processor peak frequency is for all cores running.

www.qualcomm.com 25

Qualcomm Centriq performance leadership

Performance per Watt leadership vs. top end Intel Xeon Platinum, Gold, and Silver

*SPECint®_rate2006 estimate extrapolated from published icc numbers using icc to gcc -O2 scale factor derived from internal measurements on Intel Xeon Platinum 8176, Intel Xeon Platinum 8160 and Intel Xeon Silver 4110. Power based on TDP rating; more details are in end notes.

www.qualcomm.com 27