

arm

# Arm HPC Ecosystem

Hardware, Software and tools

Srinath Vadlamani, Field Application Engineer  
SEA, April 8, 2019

# Arm Technology Already Connects the World



## Arm is ubiquitous

21 billion chips sold by partners in 2017 alone

Mobile/Embedded/IoT/  
Automotive/Server/GPUs

## Partnership is key

We design IP, not manufacture chips

Partners build products for their target markets

## Choice is good

One size is not always the best fit for all

HPC is a great fit for co-design and collaboration

**arm**

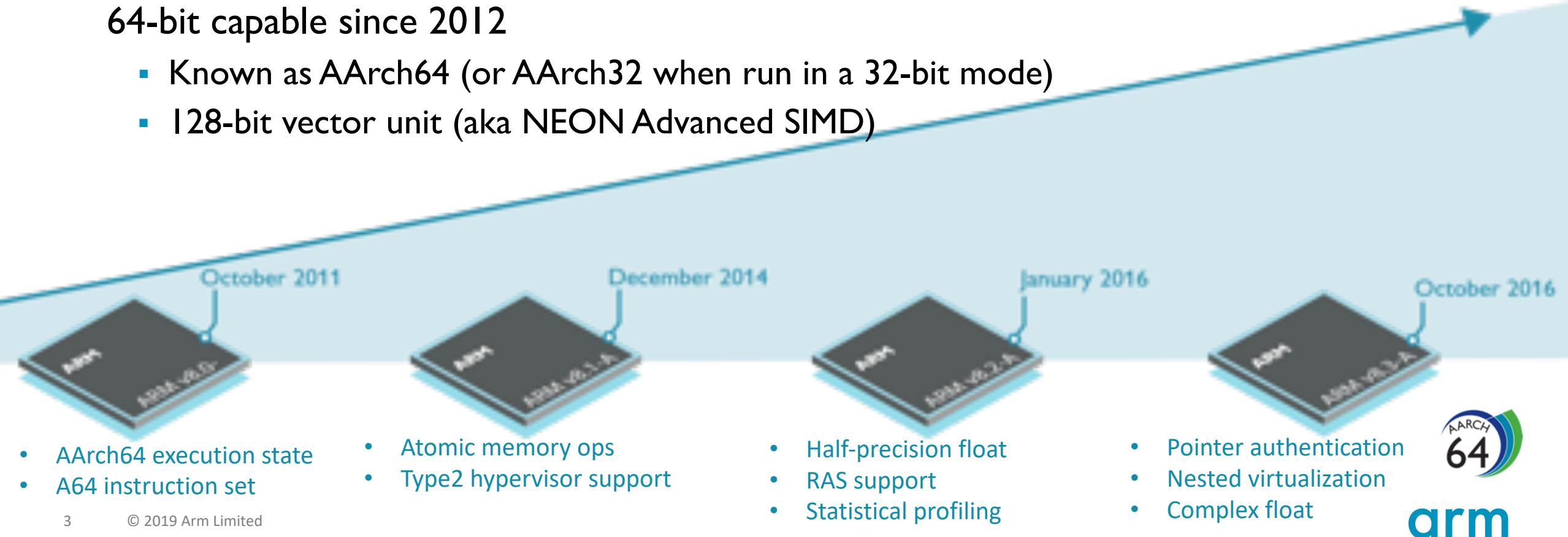
# Armv8-A Architecture Evolution

## RISC architecture

- Only have 32 bits available for encoding all instructions
- Supports the development of efficient implementations

64-bit capable since 2012

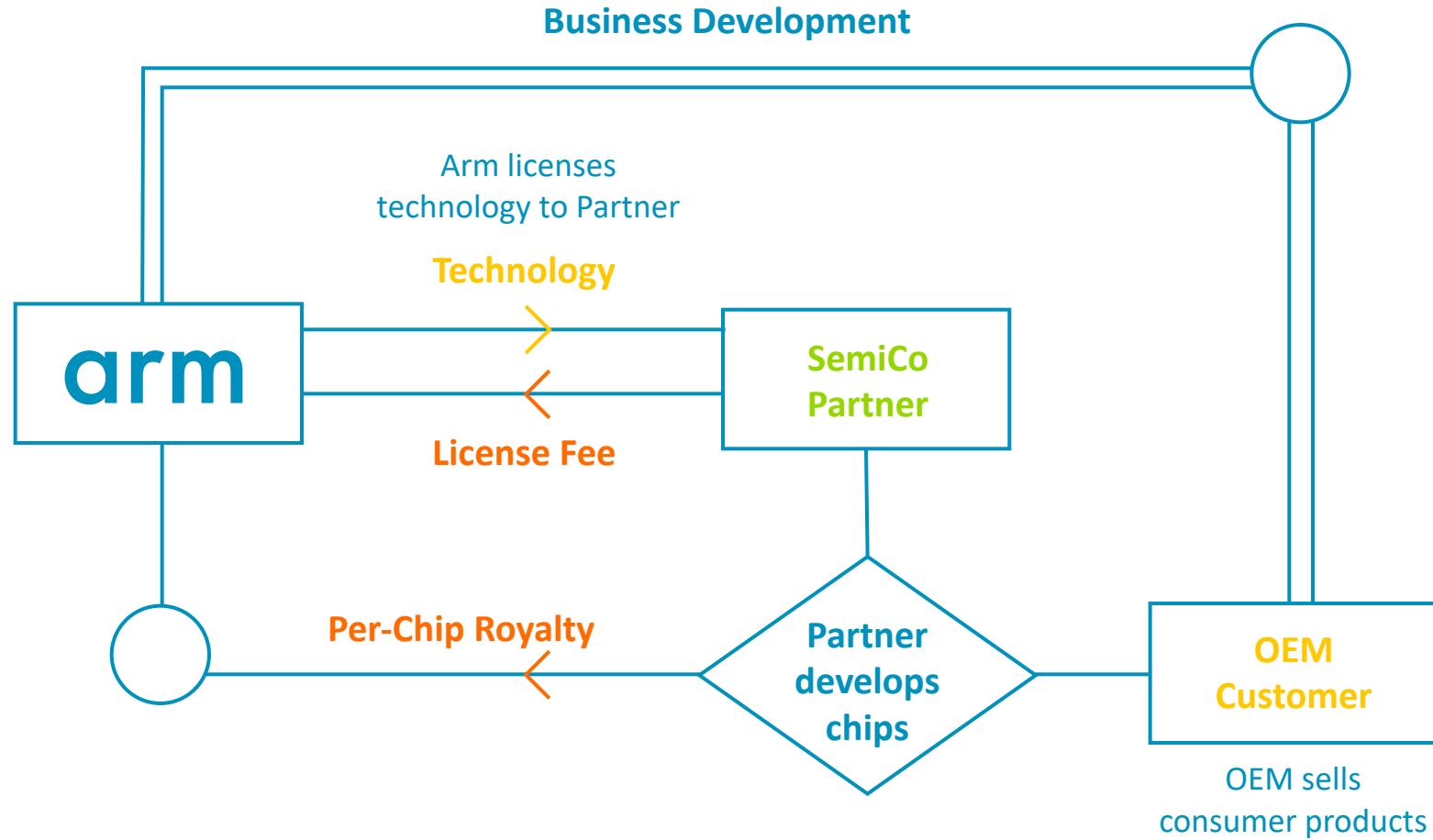
- Known as AArch64 (or AArch32 when run in a 32-bit mode)
- 128-bit vector unit (aka NEON Advanced SIMD)



# Arm business model

Arm develops technology that is licensed to semiconductor companies.

Arm receives an upfront license fee and a royalty on every chip that contains its technology.



# CPU Engagement Models With Arm

## Core License

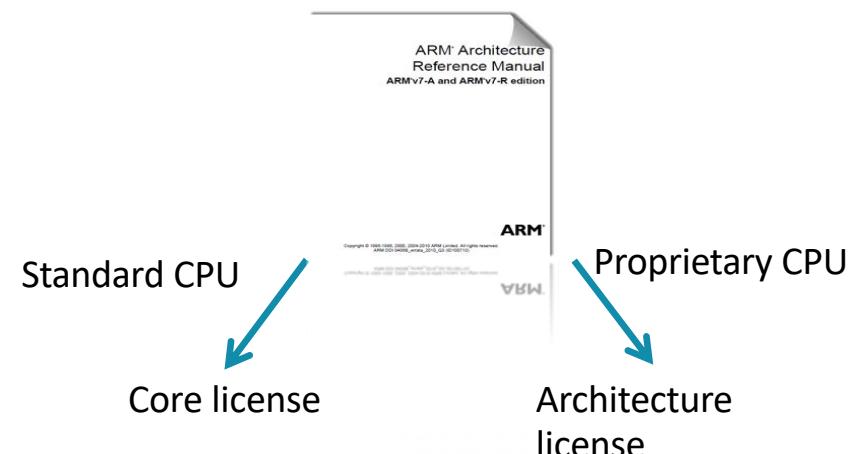
Partner licenses complete microarchitecture design

- Wide choices available
- Many different A, R & M products

CPU differentiation through:

- Flexible configuration options
- Wide implementation envelope with different process technologies

Range of licensing & engagement models possible



## Architecture License

Partner designs complete CPU microarchitecture from scratch

- Clean room – no reference to Arm core designs

Freedom to develop any design

- Must conform to the rules & programmers model of a given architecture variant
- Must pass Arm architecture validation to preserve software compatibility

Long term strategic investment

# HPC on Arm – What's new in 2018/19

## Powerful hardware for now and future

- Marvell ThunderX2 now GA
- Fujitsu announced details of A64FX (with SVE) for Post-K
- Arm announces Neoverse brand for infrastructure and core IP roadmap (Ares, Zeus, Poseidon) with each generation delivering 30% perf boost. N1 platform details announced.

## Mature toolchains and ISV Software

- Three mature toolchains available –Arm Commercial, GNU and Cray CE
- ISVs start porting to Arm – Altair RADIOSS, ANSYS Fluent and LS-DYNA

## Deployments

- New deployments across the EU and USA
- USA - Sandia Astra (Top 500), Comanche Clusters
- EU – Catalyst and Isambard in UK, GENCI and Dibona (MontBlanc 3) in France



arm

Arm Hardware for  
Infrastructure  
(including HPC)

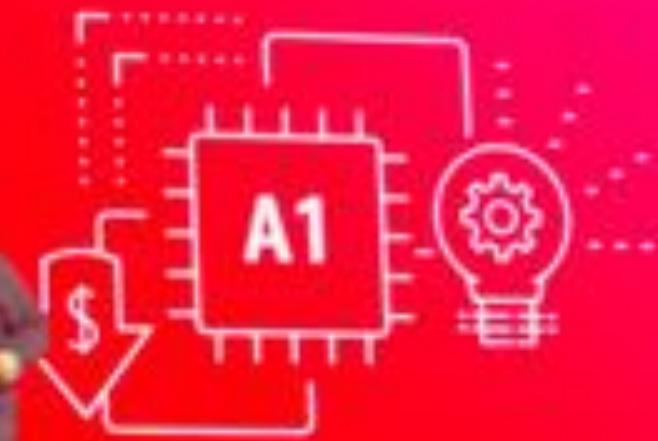
# AWS Graviton by Amazon

NEW!

## Amazon EC2 A1 Instances Powered by the AWS Graviton Processor

Up to 45% lower cost for scale-out workloads

AVAILABLE IN US EAST (N. VIRGINIA), US EAST (OHIO),  
US WEST (OREGON), AND EUROPE (IRELAND) REGIONS



# AWS Graviton by Amazon



## Introducing Amazon EC2 A1 Instances Powered By New Arm-based AWS Graviton Processors

Posted On: Nov 26, 2018

Amazon EC2 A1 instances deliver significant cost savings and are ideally suited for scale-out and Arm-based workloads that are supported by the extensive Arm ecosystem. A1 instances are the first EC2 instances powered by AWS Graviton Processors that feature 64-bit Arm Neoverse cores and custom silicon designed by AWS.

# Huawei unveils KunPeng 920 CPU and TaiShan Servers



TaiShan 2280



TaiShan 5280/5290



TaiShan X6000

*“Use ARM-based CPU in areas like cloud and servers where they are better.”*  
– William XU, Chief Strategy Marketing Officer, Huawei

# arm NEOVERSE

The Cloud to Edge Infrastructure Foundation  
for a World of 1T Intelligent Devices

# Broad SoC system design options within Arm Ecosystem

## Arm IP

High performance CPUs  
Data plane CPUs  
CMN Fabric  
Other IP

## Arm Architectural design

Custom Arm High performance CPU  
Custom Fabric & IP

## Accelerators

ML, on-die FPGA  
Networking, security, encryption  
Video, Custom

## Memory

DDR, HBM, Flash, Storage Class memory

## IO

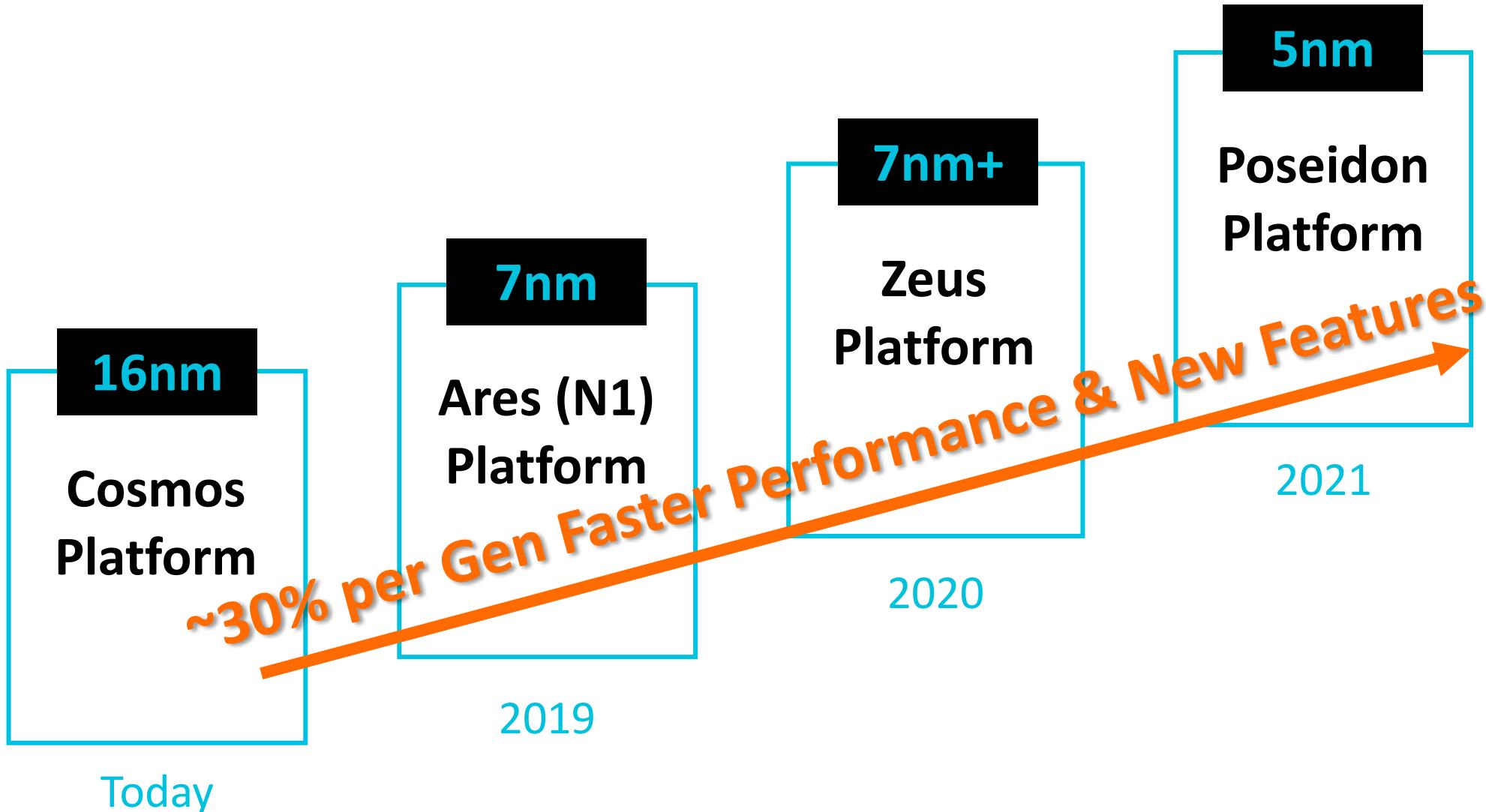
PCIe, CCIX, 100G+ ethernet

## Foundry

TSMC 7FF, Samsung 7LPP, UMC

Common Software Platform and Ecosystem  
Arm Architecture v8.x-A

# Arm IP : Commitment to Infrastructure segment



# Neoverse N1 platform

Accelerating the transformation to a scalable cloud to edge infrastructure



**Revolutionary compute performance**



**Platform features specific to infrastructure**



**Extreme range of scale and diversity of compute**



# Neoverse N1 platform: Revolutionary compute performance



*Improved cloud to edge TCO through revolutionary workload performance*

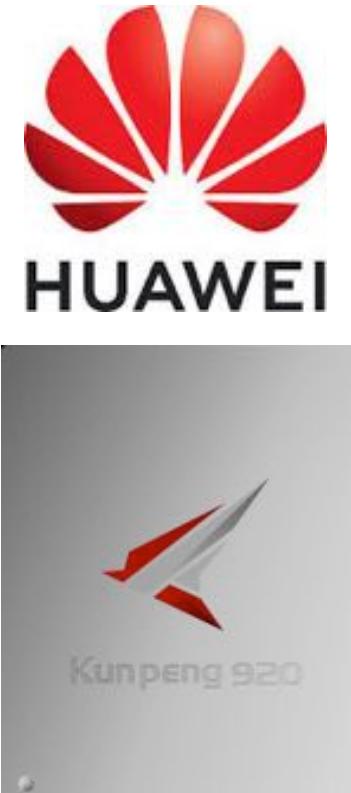
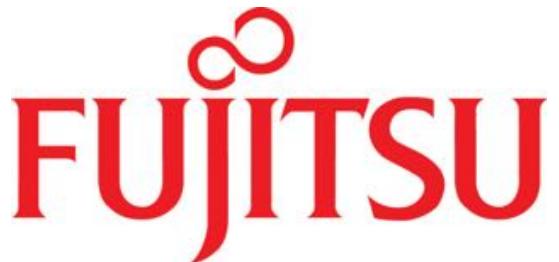
The background of the image is a dark, star-filled night sky. In the lower right foreground, a person is sitting on the edge of a rocky cliff, looking up at the stars. The sky is filled with numerous stars of varying brightness, and there are some faint, glowing nebulae or galaxy clusters visible.

arm

Arm Hardware for  
HPC

# Arm Architecture Partner SoC for HPC

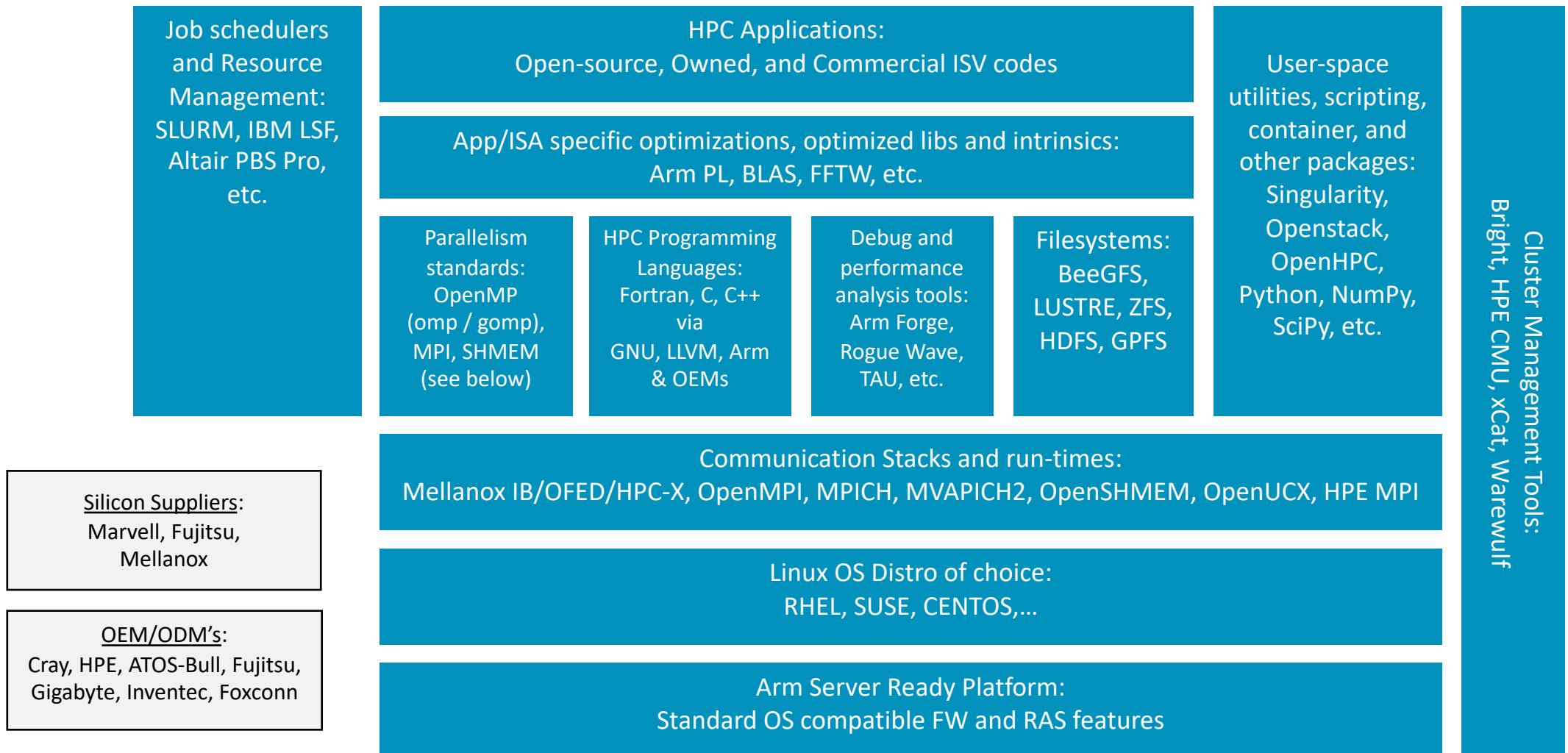
Available or Announced in 2018-19



arm

HPC Software  
Ecosystem

# Arm HPC Ecosystem – Overview



# Common HPC applications now available

GROMACS	LAMMPS	CESM2	MrBayes	Bowtie
NAMD	AMBER	Paraview	SIESTA	UM
WRF	Quantum ESPRESSO	VASP	MILC	GEANT4
OpenFOAM	GAMESS	VisIT	DL-Poly	NEMO
BLAST	NWCHEM	Abinit	BWA	QMCPACK

Build recipes online at <https://gitlab.com/arm-hpc/packages/wikis/home>

# ISVs codes on Arm

Porting underway



Available

cādence



# : Typical HPC packages available for Arm

OpenHPC is a community effort to provide a common, verified set of open source packages for HPC deployments

Arm and partners actively involved:

- Arm is a silver member of OpenHPC
- Linaro is on Technical Steering Committee
- Arm-based machines in the OpenHPC build infrastructure

Status: 1.3.6 release out now

- Packages built on Armv8-A for CentOS and SUSE

Functional Areas	Components include
Base OS	CentOS 7.5, SLES 12 SP3
Administrative Tools	Conman, Ganglia, Lmod, LosF, Nagios, pdsh, pdsh-mod-slurm, prun, EasyBuild, ClusterShell, mrsh, Genders, Shine, test-suite
Provisioning	Warewulf
Resource Mgmt.	SLURM, Munge
I/O Services	Lustre client (community version)
Numerical/Scientific Libraries	Boost, GSL, FFTW, Metis, PETSc, Trilinos, Hypre, SuperLU, SuperLU_Dist, Mumps, OpenBLAS, Scalapack, SLEPc, PLASMA, ptScotch
I/O Libraries	HDF5 (pHDF5), NetCDF (including C++ and Fortran interfaces), Adios
Compiler Families	GNU (gcc, g++, gfortran), LLVM
MPI Families	OpenMPI, MPICH
Development Tools	Autotools (autoconf, automake, libtool), Cmake, Valgrind, R, SciPy/NumPy, hwloc
Performance Tools	PAPI, IMB, pdtoolkit, TAU, Scalasca, Score-P, SIONLib

# Arm HPC Ecosystem website: [www.arm.com/hpc](https://developer.arm.com/hpc)

Starting point for developers and end-users of Arm for HPC

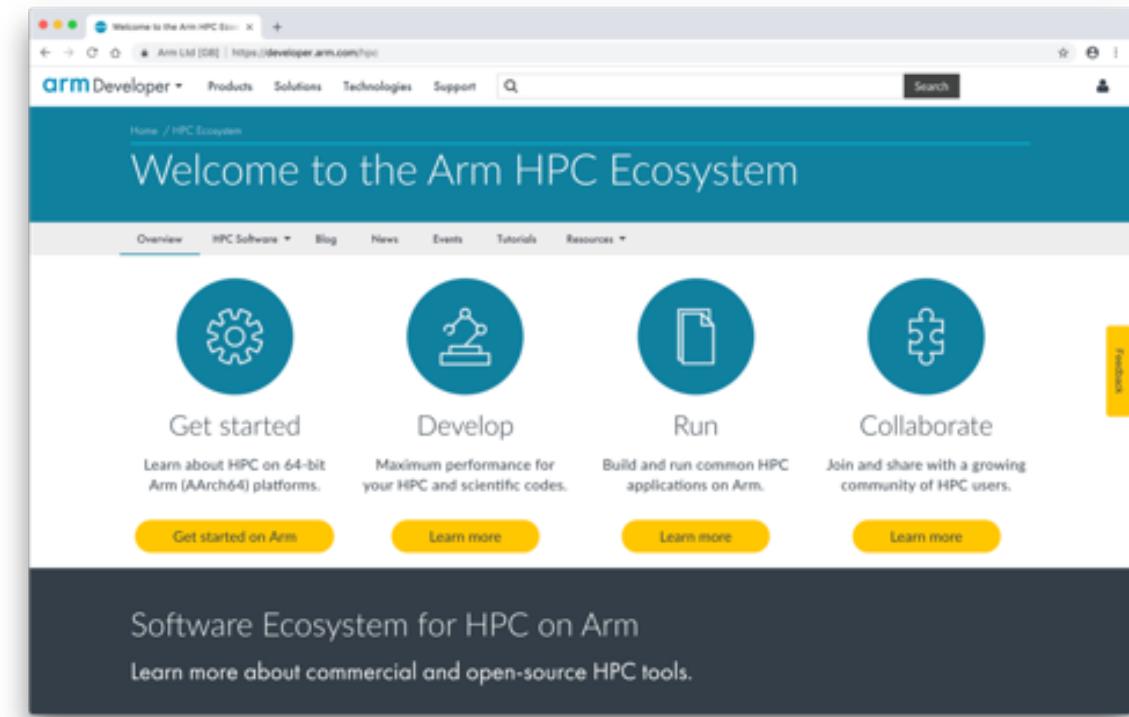
Latest events, news, blogs, and collateral including whitepapers, webinars, and presentations

Links to HPC open-source & commercial SW packages  
Guides for porting HPC applications

Quick-start guides to Arm tools

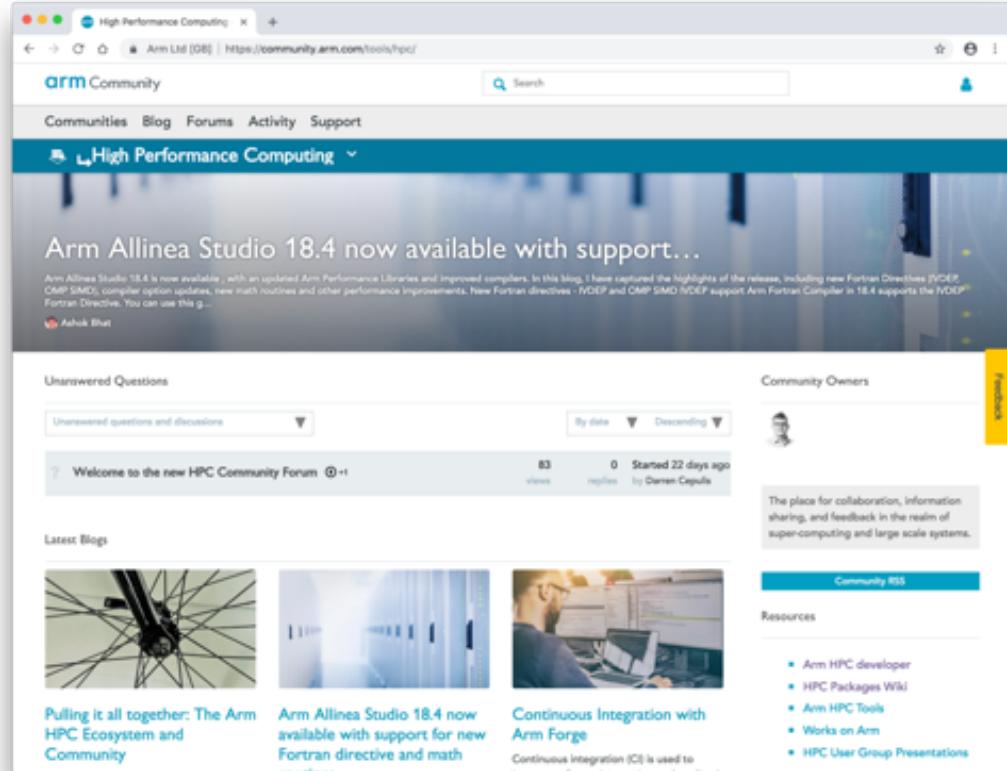
Links to community collaboration sites

Curated and moderated by Arm



# Arm HPC Community: [community.arm.com/tools/hpc/](https://community.arm.com/tools/hpc/)

## HPC Community-driven Content



**Blogs** by Arm and our HPC community

**Calendar** of upcoming events such as workshops and webinars

**HPC Forum** with questions & posts curated and moderated by Arm HPC technical specialists

**Ask, answer, share progress and expertise**

# Arm HPC Packages wiki

[www.gitlab.com/arm-hpc/packages/wikis](https://www.gitlab.com/arm-hpc/packages/wikis)

- Dynamic list of common HPC packages
- Status and porting recipes
- **Community** driven
- **Anyone can join** and contribute
- Provides **focus for porting** progress
- Allows developers to **share** and **learn**

The screenshots illustrate the structure and content of the Arm HPC Packages wiki. The top screenshot shows the main 'Home' page, which includes a sidebar for navigating projects, repositories, merge requests, CI/CD, and the wiki. The main content area displays a summary Excel spreadsheet and a list of categories such as Allpackages, Apex, Application, Benchmark, Closed source, Compiler, Coral2, and Debugger. The bottom screenshot shows a detailed table titled 'All Packages' listing various software packages along with their last modified date, build maturity status, and compatibility with the ARM Compiler and Open Compiler (OCC). The sidebar on the right provides a comprehensive list of categories.

package	last modified	BuildMaturity	CompilesARMCompiler	CompilesOCC
abinit	2018-04-04	NeedsPatch	Yes	Yes
ados	2017-12-04	-	Yes	Yes
adventure	2017-12-04	-	-	-
allinea-dbt	2017-12-06	Supported	N.R.	N.R.
allinea-map	2017-12-06	Supported	N.R.	N.R.
alya	2017-07-17	-	-	Yes
arpack	2017-07-17	-	Yes	-
awranch-blis	2017-07-10	-	-	-
atlas	2017-07-10	-	-	Yes
augustus	2017-12-07	-	-	-
autoconf	2017-12-04	-	Yes	Yes
automake	2017-12-04	-	Yes	Yes
AWP-OCC	2018-07-25	NeedsPatch	Yes	Yes

# Open source libraries for helping increase performance

## Arm Optimized Routines

<https://github.com/ARM-software/optimized-routines>

These routines provide high performing versions of many math.h functions

- Algorithmically better performance than standard library calls
- No loss of accuracy

## SLEEF library

<https://github.com/shibatch/sleef/>

Vectorized math.h functions

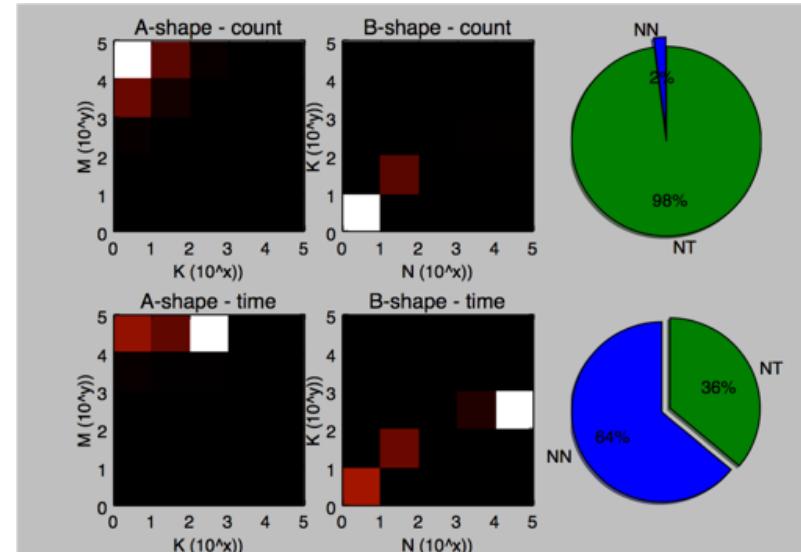
- Provided as an option for use in **Arm Compiler**

## Perf-lbs-tools

<https://github.com/ARM-software/perf-lbs-tools>

Understanding an application's needs for BLAS, LAPACK and FFT calls

- Used in conjunction with **Arm Performance Libraries** can generate logging info to help profile applications for specific case breakdowns



Example  
visualization:  
DGEMM  
cases called

arm



arm

Arm HPC  
deployments

# Deployments



Sandia  
National  
Laboratories



GW4



THE UNIVERSITY  
*of* EDINBURGH



# Arm Supercomputer Makes Top500 List!

**SCIENTIFIC COMPUTING WORLD**

Search Site  
For scientist who

News Analysis & Opinion Features Issues Events Resources

**SOFTIRON** 1PB for 17.5¢/GB LIMITED TIME OFFER

---

NEWS  
Tags: HPC

Sandia Labs supercomputer the fastest Arm-based TOP500 system  
16 November 2018

[Tweet](#) [Share](#)

Astra, the world's fastest Arm-based supercomputer according to the TOP500 list, has achieved a speed of 1.529 petaflops, placing it 203rd on a ranking of top computers announced at SC18 conference in Dallas.

The supercomputer is also ranked 36th on the High-Performance Conjugate Gradients benchmark, co-developed by Sandia and the University of Tennessee Knoxville, with a performance of 66.942 teraflops. (One thousand teraflops equals 1 petaflop.)

The conjugate benchmark uses computational and data access patterns that more closely match the simulation codes used by the National Nuclear Security Administration.

*“Astra, the world’s fastest Arm-based supercomputer according to the TOP500 list, has achieved a speed of 1.529 petaflops, placing it 203rd on a ranking of top computers …”*

# Vanguard Astra at Sandia

MOST POWERFUL ARM SUPERCOMPUTER, IN TOP 500 (#203 in HPL and #36 in HPCG)

- 2,592 HPE Apollo 70 compute nodes
  - 5,184 CPUs, 145,152 cores, 2.3 PFLOPs (peak)
- Marvell ThunderX2 Arm SoC, 28 core, 2.0 GHz
- Memory per node: 128 GB (16 x 8 GB DR DIMMs)
  - Aggregate capacity: 332 TB, 885 TB/s (peak)
- Mellanox IB EDR, ConnectX-5
  - 112 36-port edges, 3 648-port spine switches
- Red Hat RHEL for Arm
- HPE Apollo 4520 All-flash Lustre storage
  - Storage Capacity: 403 TB (usable)
  - Storage Bandwidth: 244 GB/s



# Isambard in Production at Bristol/GW4

Largest EU Arm HPC cluster to date

- Cray XC50 system w/ 168 nodes with Marvell ThunderX2 (32C)
- 10,752 total cores
- High-speed ARIES interconnect
- Cray HPC SW Stack including CCE, CrayPAT, Cray MPI, libs, ...
- Production deployment reached @ SC18



# Deployments: Catalyst UK



- **HPE**, in conjunction with **Arm** and **SUSE**, announced in April the “**Catalyst UK**” program: deployments to accelerate the growth of the Arm **HPC** ecosystem into three universities
- Each machine will have:
- 64 HPE Apollo 70 systems, each with two 32-core Cavium ThunderX2 processors (i.e. 4096 cores per system), 128GB of memory and Mellanox InfiniBand interconnects
- SUSE Linux Enterprise Server for HPC



THE UNIVERSITY  
of EDINBURGH



UNIVERSITY OF  
LEICESTER

**Bristol:** VASP, CASTEP, Gromacs, CP2K, Unified Model, Hydra, NAMD, Oasis, NEMO, OpenIFS, CASINO, LAMMPS

**EPCC:** WRF, OpenFOAM, Rolls Royce Hydra opt, 2 PhD candidates

**Leicester:** Data-intensive apps, genomics, MOAB Torque, DiRAC collab

# Deployment: Mont Blanc

## The Mont-Blanc prototype ecosystem

Prototypes are critical to accelerate software development  
System software stack + applications

MONT-BLANC

- Mini-clusters
  - Arndale
  - Odroid XU
  - Odroid XU-3
  - NVIDIA Jetson

- PRACE prototypes
  - Tibidabo
  - Carma
  - Pedraforca

2011      2012      2013

- Mont-Blanc prototype
- 1080 compute cards
    - Dual Cortex-A15
    - GPU Mali-T604
    - 4 GB LPDDR3
    - Up to 64 GB local storage
    - USB-to-Eth network
    - Fine grained power monitoring system
    - Installed between Jan and May 2015
    - Operational since May 2015 @ BSC



2014      2015      2017

Bull  
Asia Technologies

- Mont-Blanc 3**
- Bull Sequana™
  - Cavium ThunderX2™



## BSC KEEPS ITS HPC OPTIONS OPEN WITH MARENOSTRUM 4

December 1, 2016    Timothy Prickett Morgan



When it comes to supercomputing, you don't only have to strike while the iron is hot, you have to spend while the money is available. And that fact is what often determines the technologies that HPC centers deploy as they expand the processing and storage capacity of their systems.

A good case in point is the MareNostrum 4 hybrid cluster that the Barcelona Supercomputing Center, one of the flagship research and computing institutions in Europe, has just commissioned IBM to build with the help of partners Lenovo and Fujitsu. The system balances the pressing need for more general purpose computing while at the same time allowing for researchers to explore how applications might be sped up on hybrid CPU-GPU machines that mix processors from IBM and accelerators from Nvidia and alternatively on nodes based on Intel's "Knights" family of Xeon Phi manycore processors. There is even a slice of the system that will be based on a baby version of the "Post-K" supercomputer that Fujitsu is building for the Japanese government and that is based on its own homegrown ARM processors with vector extensions it is developing with ARM Holdings.

# Deployments: HPE's Comanche Collaboration

Early access to Cavium ThunderX2 systems that became Apollo 70



Engagements in HPE Comanche program have accelerated adoption

- We have been able to assess the state of fundamental software stacks, such as MPI and NUMA capabilities
  - Collaborative work here especially great with all partners focusing on interoperability issues
  - Examples include fixing bugs with kernels, MPI drivers and OpenMP thread placement
  - Optimization of packages, environment and execution configurations



Over 1,000 processors delivered | LLNL TOSS stack ported and demoed | InfiniBand optimized

arm

Performance  
results on Arm



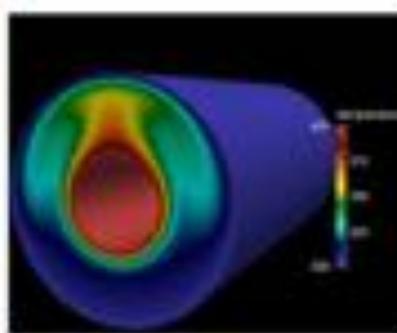
## Early Results from Astra

System has been online for around two weeks , incredible team working round the clock, already running full application ports and many of our key frameworks

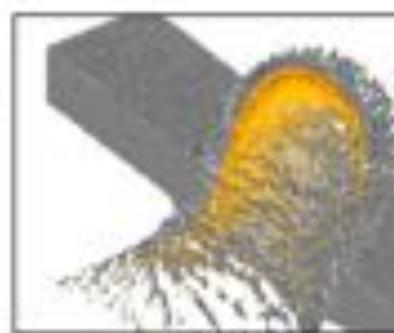
Baseline: Trinity ASC Platform (Current Production (LANL/SNL)), dual-socket Haswell



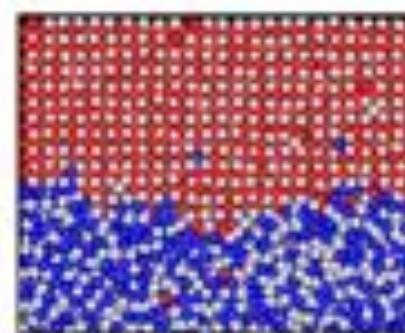
Monte Carlo



CFD Models



Hydrodynamics



Molecular Dynamics



Linear Solvers

1.60X

1.45X

1.30X

1.42X

1.87X

# Single node results from GENCI - France



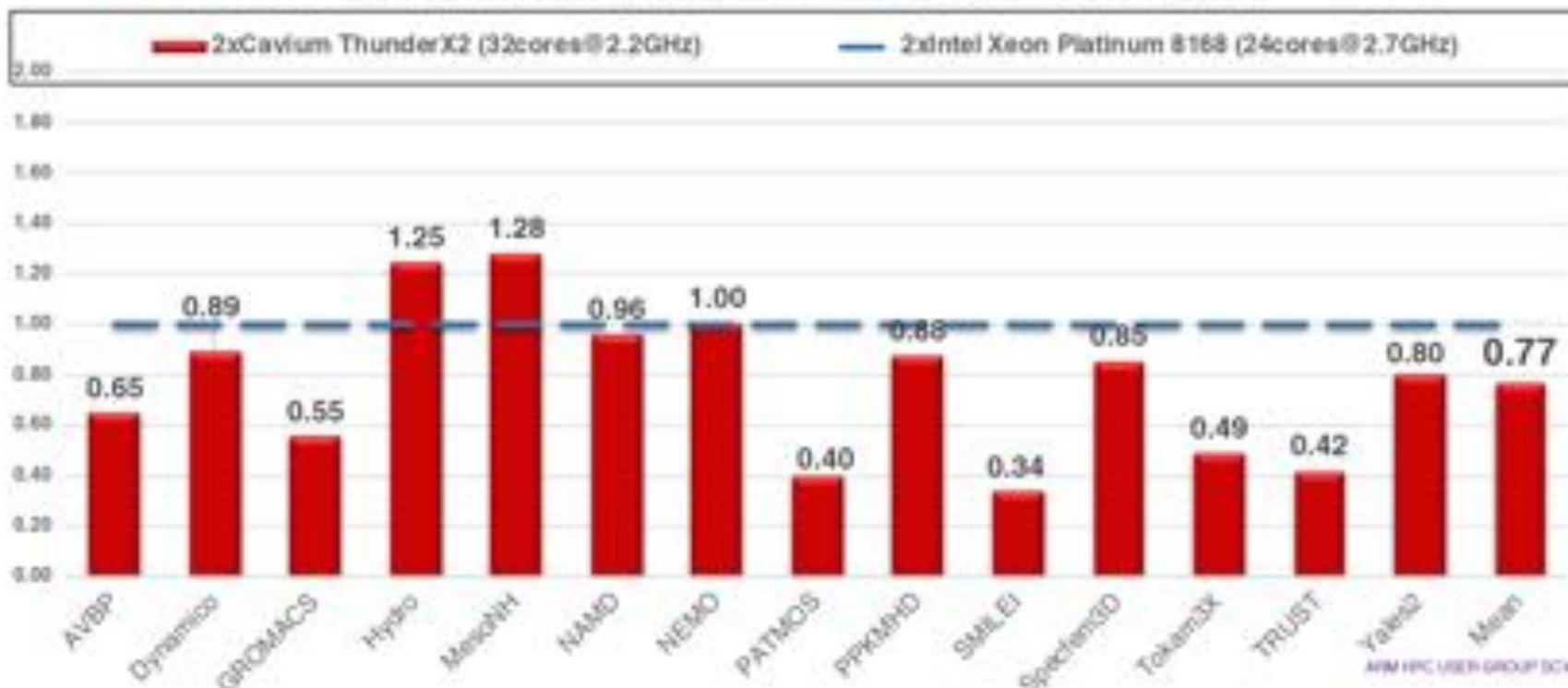
## PERFORMANCE SUMMARY

### Preliminary results

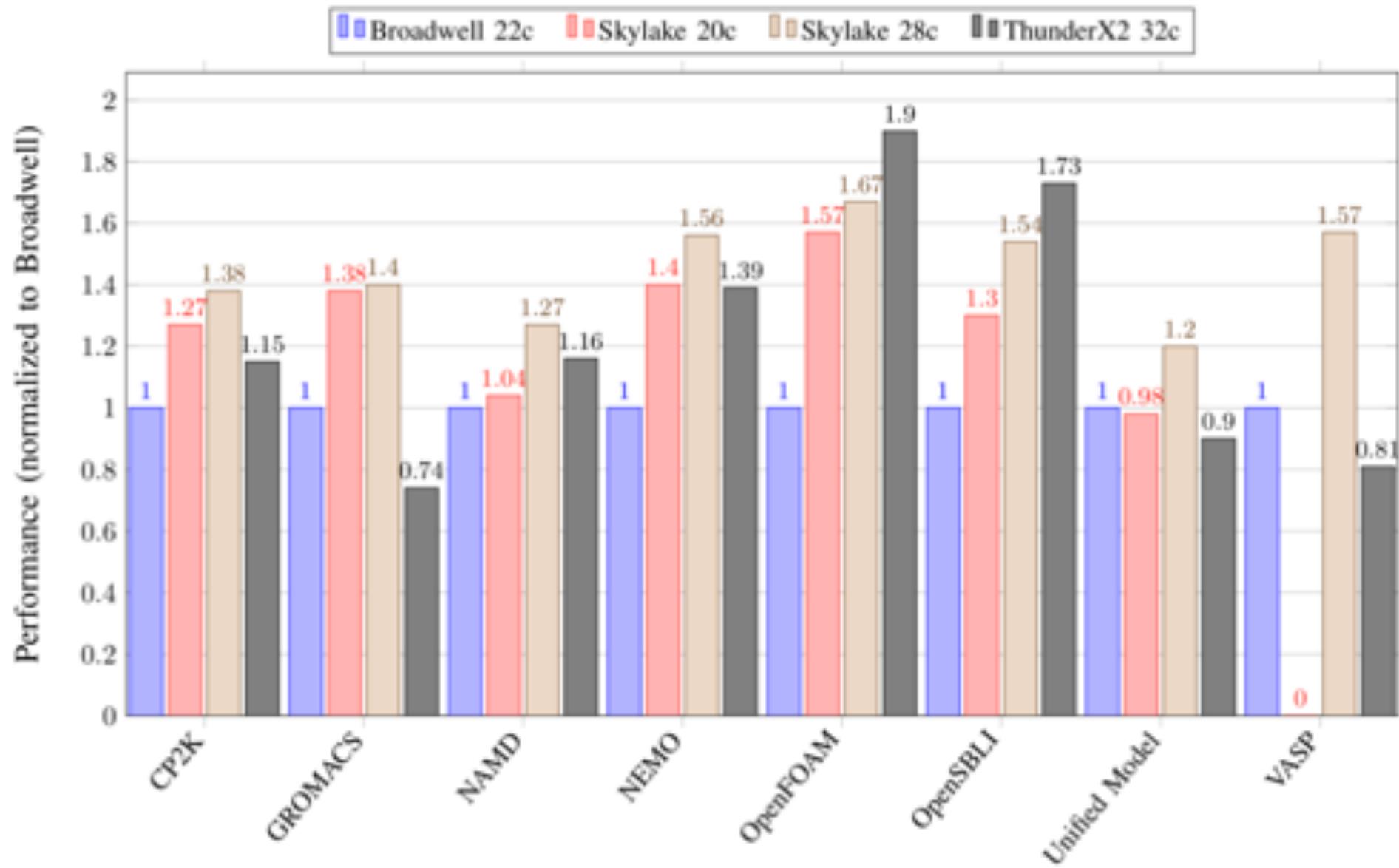
□ The overall performance for those applications at the moment:

- Comparison to Tier-0 machine Irene Joliot-Curie @ CEA/TGCC, Bruyères-le-Châtel (France)
- ARM compiler v18.4.2 vs Intel compiler 18.x

NODE TO NODE SPEED-UP THUNDER-X2 VS SKYLAKE

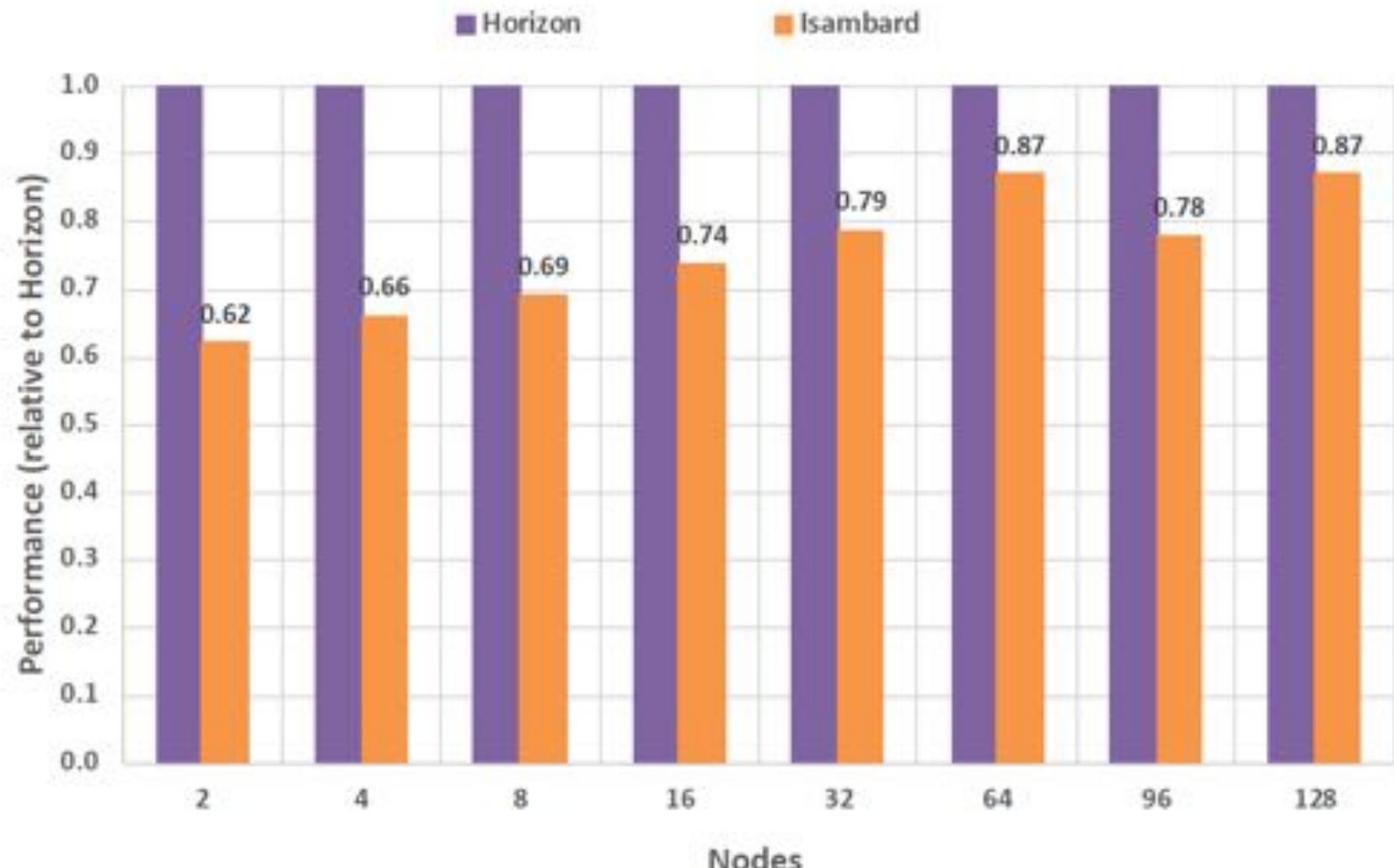


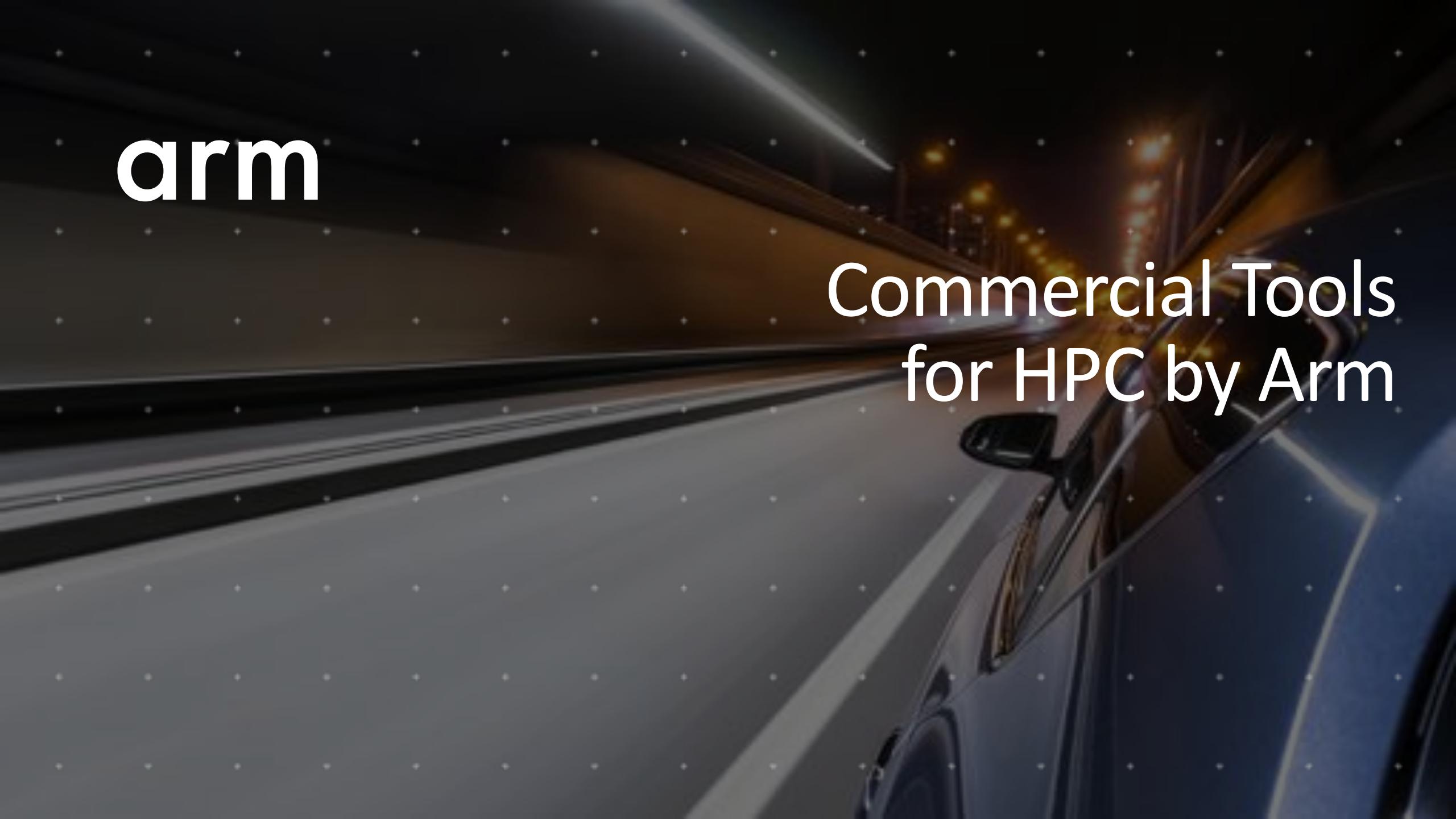
# Isambard, UK – Single node results



# Isambard, UK – Multi-node results

Gromacs (42M atoms) on Horizon (Intel Skylake, 20C) vs Isambard (Marvell ThunderX2, 32C)



The background of the slide is a blurred photograph of a car driving on a road at night. The motion blur creates streaks of light along the road's edge and the car's body. The scene is set against a dark sky with small white stars.

arm

Commercial Tools  
for HPC by Arm

# Our Solution for *any* Architecture, at *any* Scale

Commercial tools for AArch64, x86\_64, ppc64 and accelerators

## Arm Cross-Platforms Tools

### Arm DDT

Slash your time to debug on  
any hardware, at any scale.

### Arm MAP

Speed-up applications with a  
lightweight scalable profiler

Debug, optimise and analyse any platform

### Arm Forge Professional

Arm DDT and MAP in  
One Single Package

### Arm Performance Reports

Find the most efficient  
settings for your workloads.

## arm ALLINEA STUDIO

All-inclusive development toolkit for Arm hardware

### Arm HPC Compiler

Linux user space compiler  
for HPC applications

### Arm Performance Libraries

BLAS, LAPACK and FFT

### Arm Forge Professional

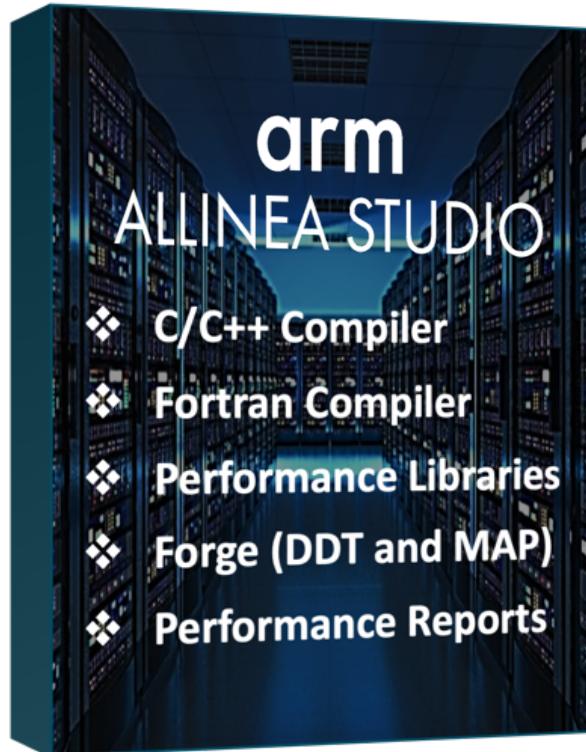
Multi-node interoperable  
profiler and debugger

### Arm Performance Reports

Interoperable application  
performance insight

# Arm Allinea Studio

Built for developers to achieve best performance on Arm with minimal effort



**Comprehensive and integrated tool suite** for Scientific computing, HPC and Enterprise developers

**Seamless end-to-end workflow** from getting started to advanced optimization of your workloads

**Commercially supported** by Arm engineers

**Frequent releases** with continuous performance improvements

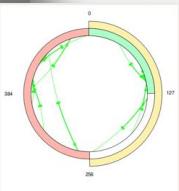
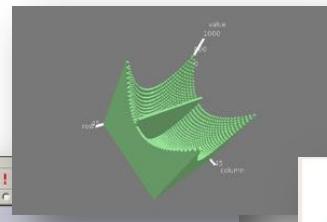
**Ready for current and future generations** of server-class Arm-based platforms

Available for a wide-variety of Arm-based server-class platforms

# arm ALLINEA STUDIO

Meets the requirements of HPC developers on Arm

**Arm Performance Libraries**  
BLAS, LAPACK, FFT  
Scalar and vector math functions



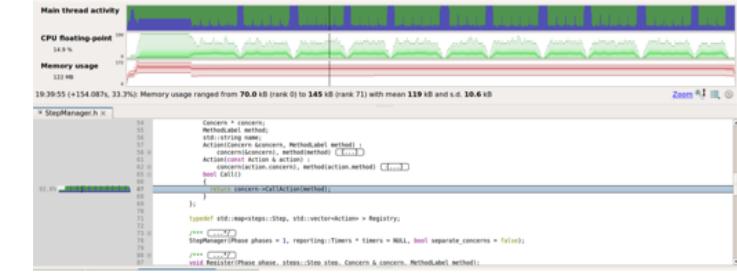
**Arm DDT**  
Cross-platform parallel debugger

Optimize

Profile

Debug

Develop  
and build

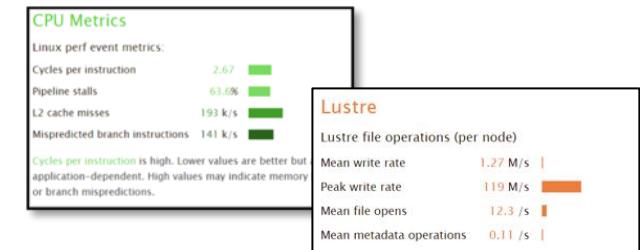


## Arm MAP

Cross-platform lightweight profiler

## Arm Performance Reports

Maximize System Efficiency

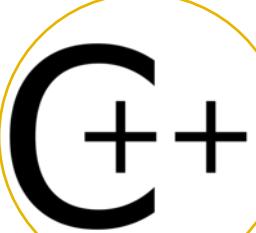


## Arm Linux Compiler

For C, C++ and Fortran codes

# arm Allinea Studio

## A quick glance at what is in Arm Allinea Studio



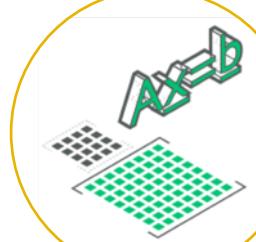
### C/C++ Compiler

- C++ 14 support
- OpenMP 4.5 without offloading
- SVE ready

### Fortran

### Fortran Compiler

- Fortran 2003 support
- Partial Fortran 2008 support
- OpenMP 3.1
- SVE ready



### Performance Libraries

- Optimized math libraries
- BLAS, LAPACK and FFT
- Threaded parallelism with OpenMP
- Scalar math routines



### Forge (DDT and MAP)

- Profile, Tune and Debug
- Scalable debugging with DDT
- Parallel Profiling with MAP



### Performance Reports

- Analyze your application
- Memory, MPI, Threads, I/O, CPU metrics

Tuned by Arm for a wide-range of server-class Arm-based platforms

# Progress in the last year

A fully integrated tools suite for deployment on Arm systems



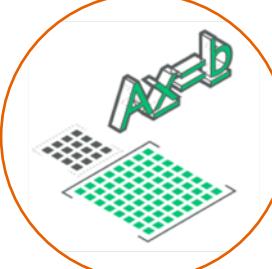
## Arm C/C++ Compiler

- Porting and tuning guides for common applications
- Optimizations and bug fixes

## Fortran

## Arm Fortran Compiler

- New Fortran Directives
- Improved Fortran 2008 support
- Support for vectorization of loops with math calls



## Arm Perf Libraries

- BLAS, FFT and LAPACK Improvements
- Sparse routine SPMV support
- Scalar math routines



## Forge and Perf Reports

- General cross-platform improvements
- Python profiling
- Better interop with Arm Compiler and Libraries



## GNU8 toolchain

- GCC and Gfortran
- 2<sup>nd</sup> toolchain in the studio
- Better suited for certain applications
- Beta support for HPC users

Support and tuning for Arm server-class platforms

# arm COMPILER

Commercial C/C++/Fortran compiler with best-in-class performance



Compilers tuned for Scientific Computing and HPC



Latest features and performance optimizations



Commercially supported by Arm

## Tuned for Scientific Computing, HPC and Enterprise workloads

- Processor-specific optimizations for various server-class Arm-based platforms
- Optimal shared-memory parallelism using latest Arm-optimized OpenMP runtime

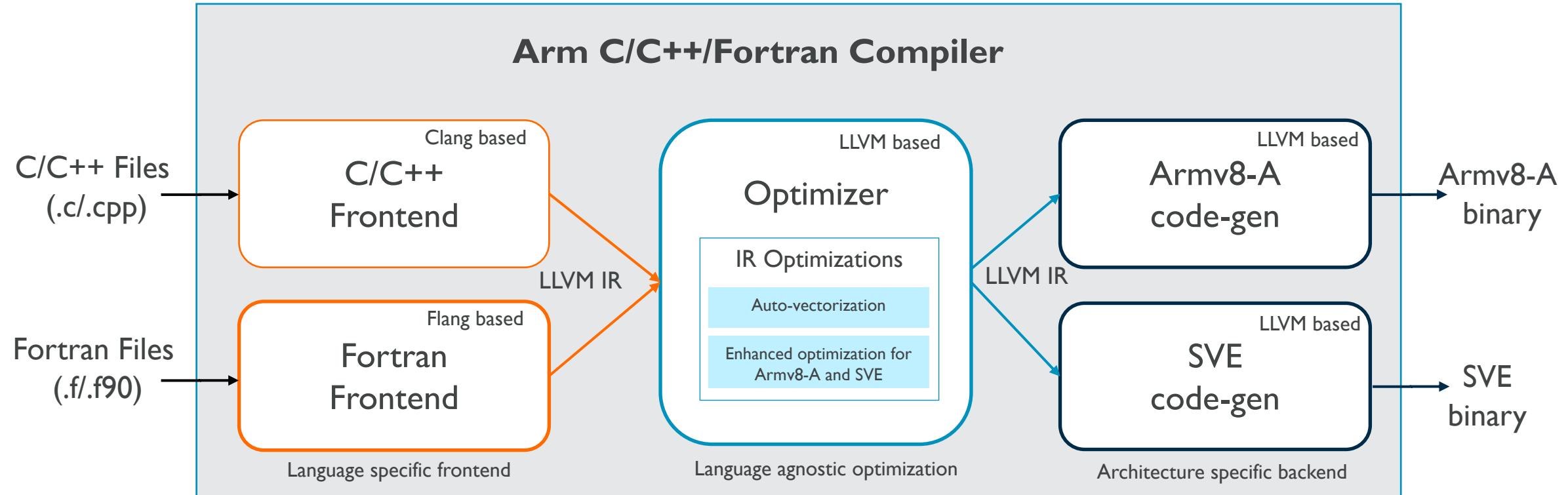
## Linux user-space compiler with latest features

- C++ 14 and Fortran 2003 language support with OpenMP 4.5
- Support for Armv8-A and SVE architecture extension
- Based on LLVM and Flang, leading open-source compiler projects

## Commercially supported by Arm

- Available for a wide range of Arm-based platforms running leading Linux distributions – RedHat, SUSE and Ubuntu

# Arm Compiler – Building on LLVM, Clang and Flang projects



# Arm Linux Compiler – What's new in 2018/19?

## Overall - Better code generation

- For Arm platforms for current generation (Marvell ThunderX2) and future (SVE based)
- Base compiler technology upgrade (Clang/LLVM 7, GNU8, Latest Flang)
- Vectorization of loops with math function calls

## Fortran – Increase in maturity

- Enable key Fortran applications (open source, in house and commercial)
- Improved auto vectorization
- Fortran vectorization directives like IVDEP

# arm PERFORMANCE LIBRARIES

## Optimized BLAS, LAPACK and FFT



Commercially supported  
by Arm



Best in class performance



Validated with  
NAG test suite

### Commercial 64-bit Armv8-A math libraries

- Commonly used low-level math routines - BLAS, LAPACK and FFT
- Provides FFTW compatible interface for FFT routines
- Batched BLAS support

### Best-in-class serial and parallel performance

- Generic Armv8-A optimizations by Arm
- Tuning for specific platforms like Cavium ThunderX2 in collaboration with silicon vendors

### Validated and supported by Arm

- Available for a wide range of server-class Arm-based platforms
- Validated with NAG's test suite, a de-facto standard

# Arm Performance Libraries progress

Progress and additions since SC17

## Key improvements in since 18.0

- Massive improvements in FFT performance
  - All basic, advanced and guru interface FFTW calls now supported
- Many functions have had extra serial and parallel performance improvements targeting ThunderX2
- Addition of libamath
  - High performing implementations of certain key math.h functions

## New features in 19.0

- Sparse linear algebra for higher performing SpMV calls
- FFTW MPI interface for FFT calls added
- Parallelisation of many FFTW plans
- Parallel scaling improvements, especially for ThunderX2
  - Particular focus on GEMMs and POTRF, GETRF and GETQR

# Compiler and Libraries - Future roadmap

Focus on current and next generation hardware

Libraries : Vector Math routines and more scalar math routines

Fortran Compiler : Directives & new Fortran 2008/OpenMP features

All compilers : Vectorization and optimization report improvements

More features in compilers and libraries

- Application specific tuning and optimization
- For Marvell ThunderX2 and other server-class Arm-based platforms

More optimizations for current hardware

- SVE enabled Performance Libraries
- Application specific tuning and optimization in Compilers and Libraries for SVE

Getting ready for SVE-based future hardware

arm

Toolchain  
performance  
results

# Arm Compiler and Libraries – 19.1 release

Progress and additions since SC18 (19.0 release)

## Arm C/C++/Fortran Compilers

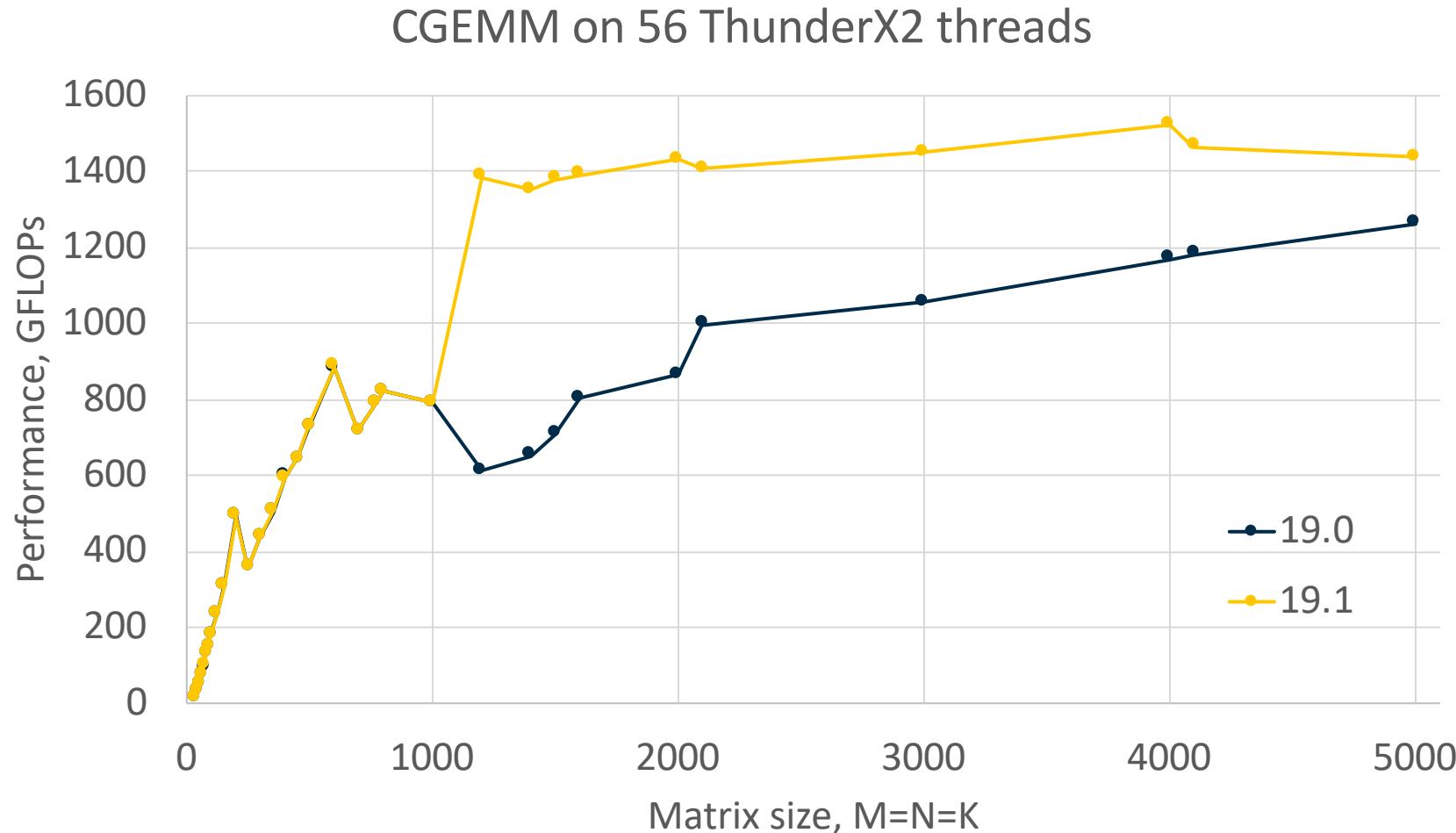
- Fortran: TRAILZ intrinsic, a Fortran 2008 feature, now supported
- Fortran: Runtime I/O performance improvement when handling formatted text data
- Fortran: New UNROLL directive to provide unrolling hints to the compiler
- Bug fixes

## Arm Perf Libraries

- BLAS - Improved GEMV and GEMM (SCZ variants)
- FTW Fortran MPI interface now supported
- FFT MPI parallel scaling has been improved.
- SpMV - Support for CSC and COO formats; Improved single-precision performance; Fortran Interface now supported.
- Math routines (in libamath) – Vector routines support with optimized logf and expf; Arm Compiler uses libamath by default; A GNU compatible version provided

# BLAS improvements to many GEMM routines in 19.1

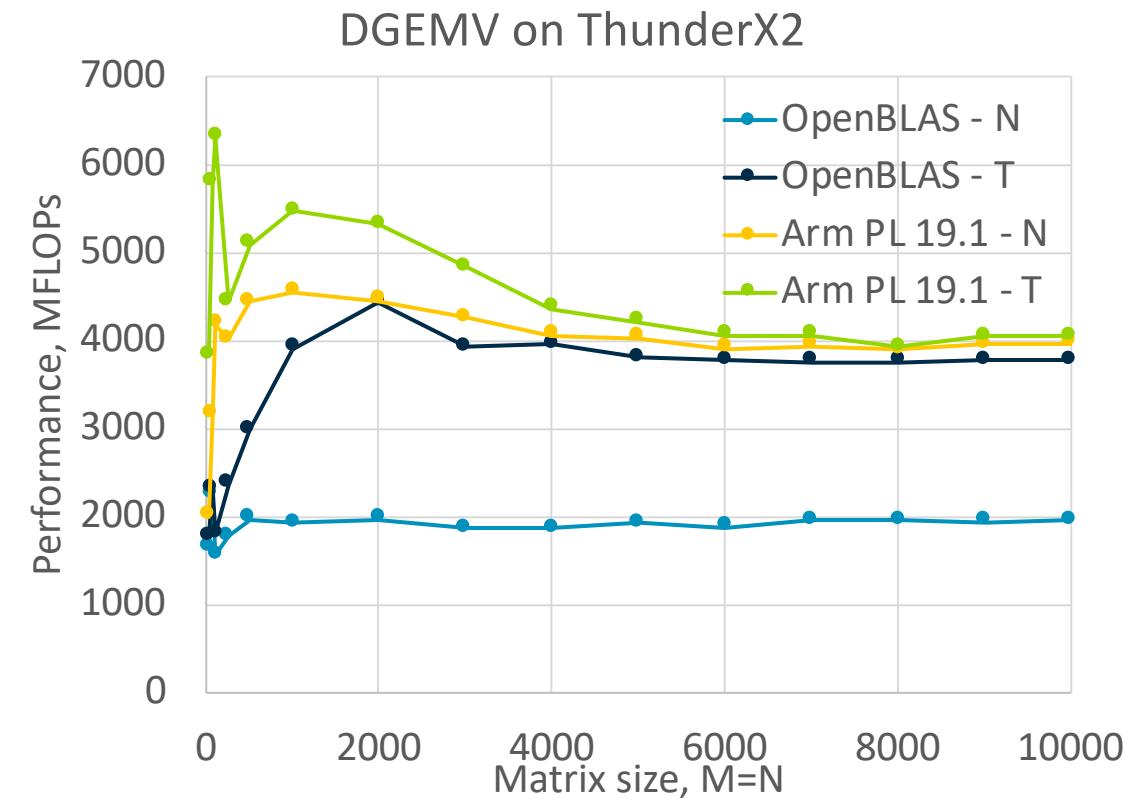
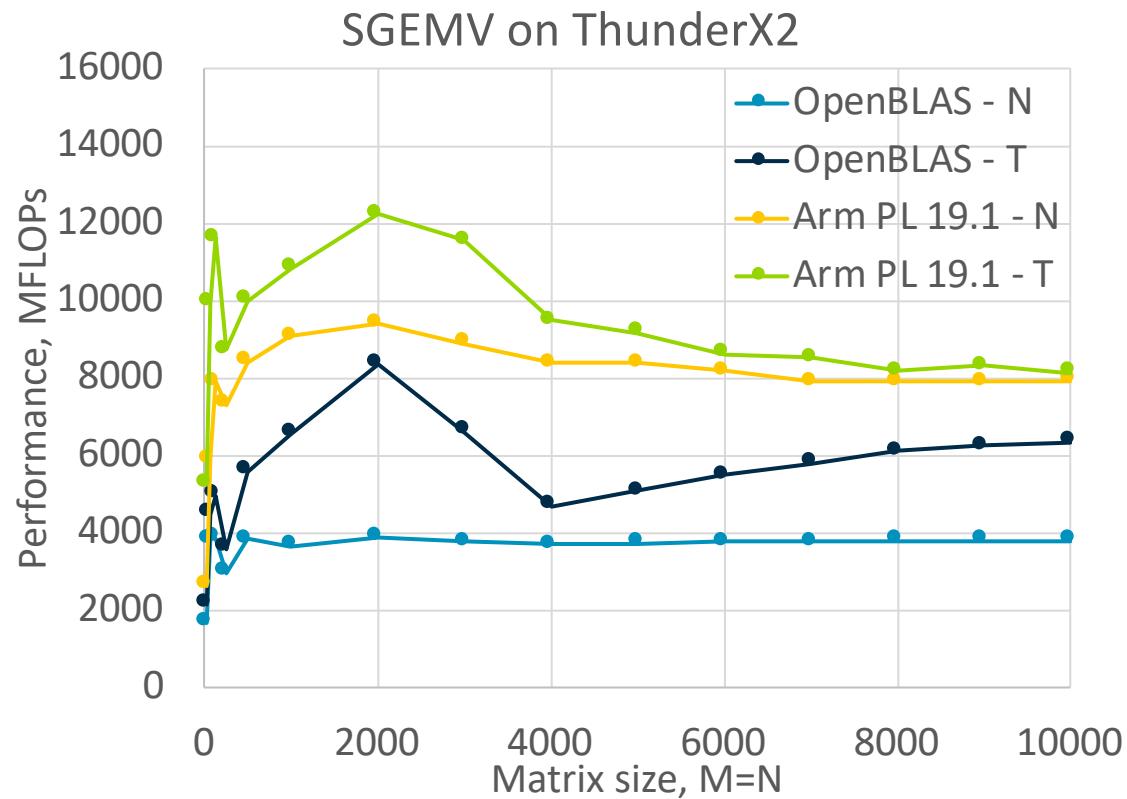
Shown below: CGEMM on Marvell ThunderX2 run using 56 threads



# BLAS improvements to GEMV routines in 19.1

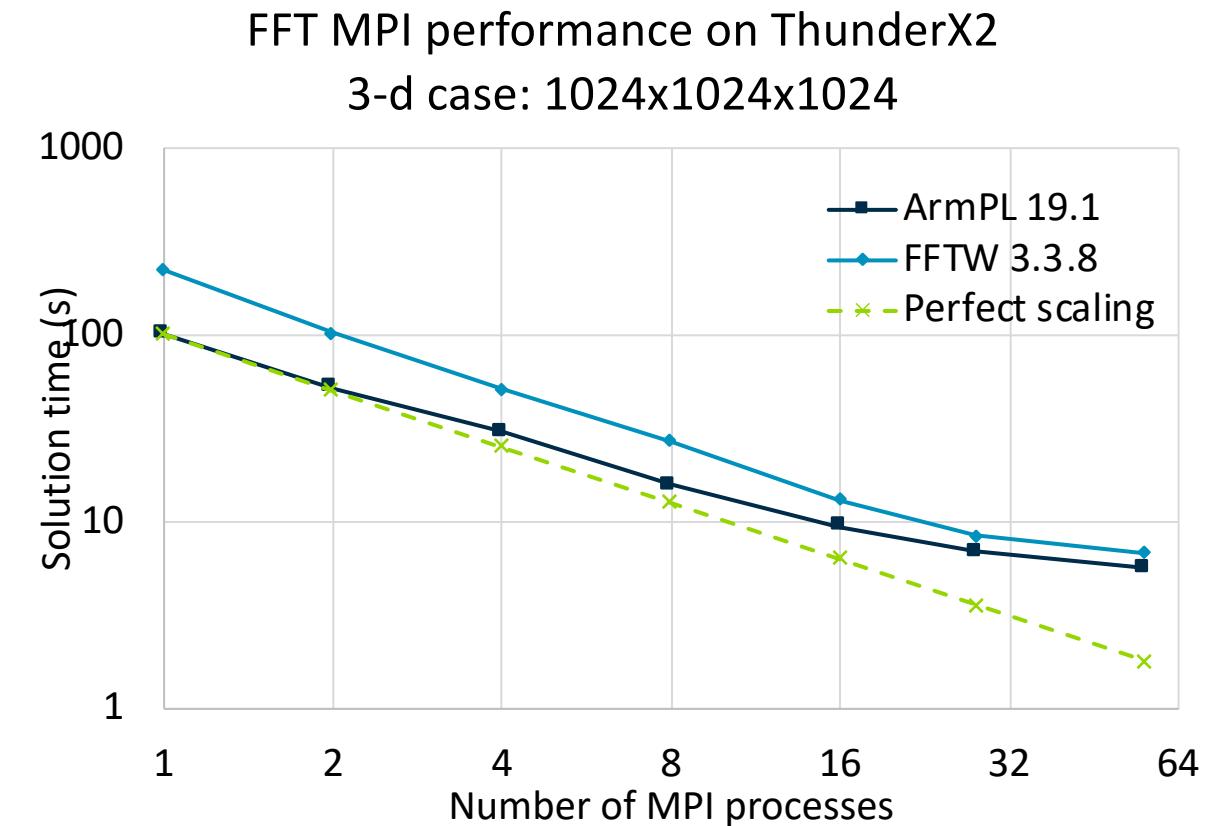
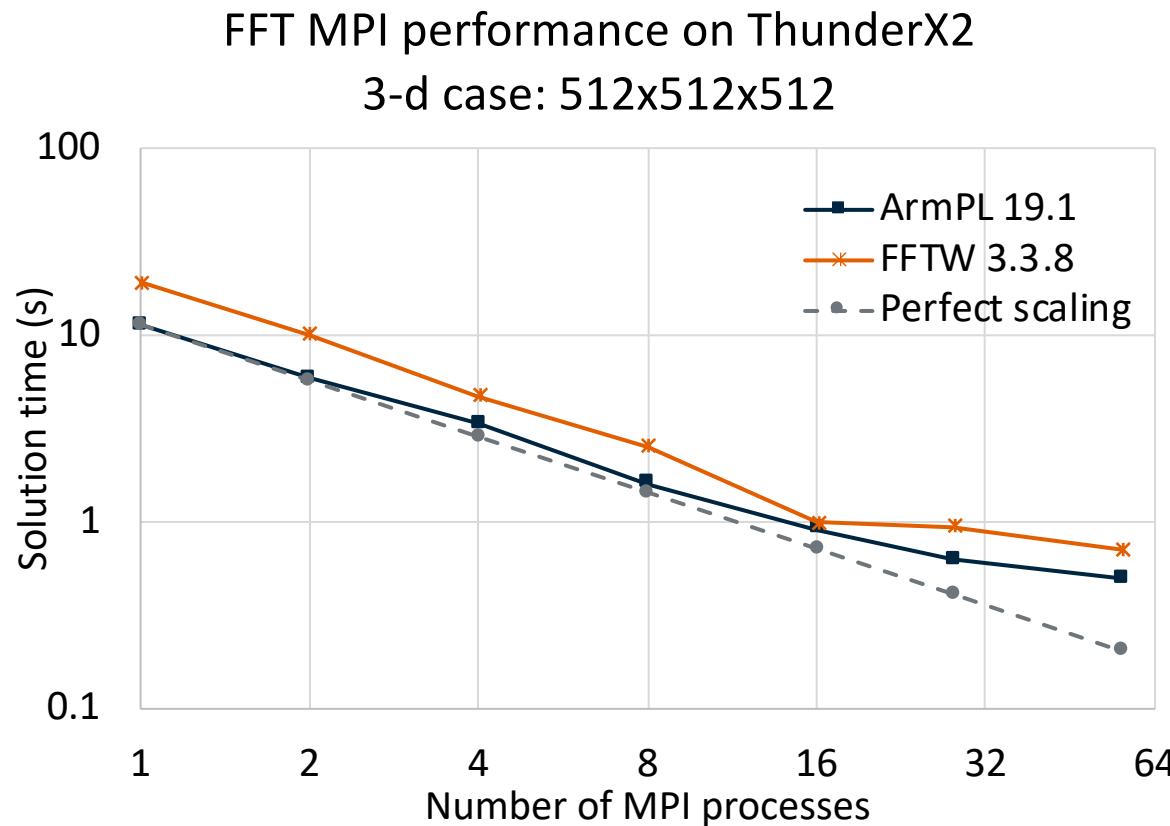
All cases improved for both serial and parallel.

Comparison shown on ThunderX2 for serial SGEMV and DGEMV against OpenBLAS



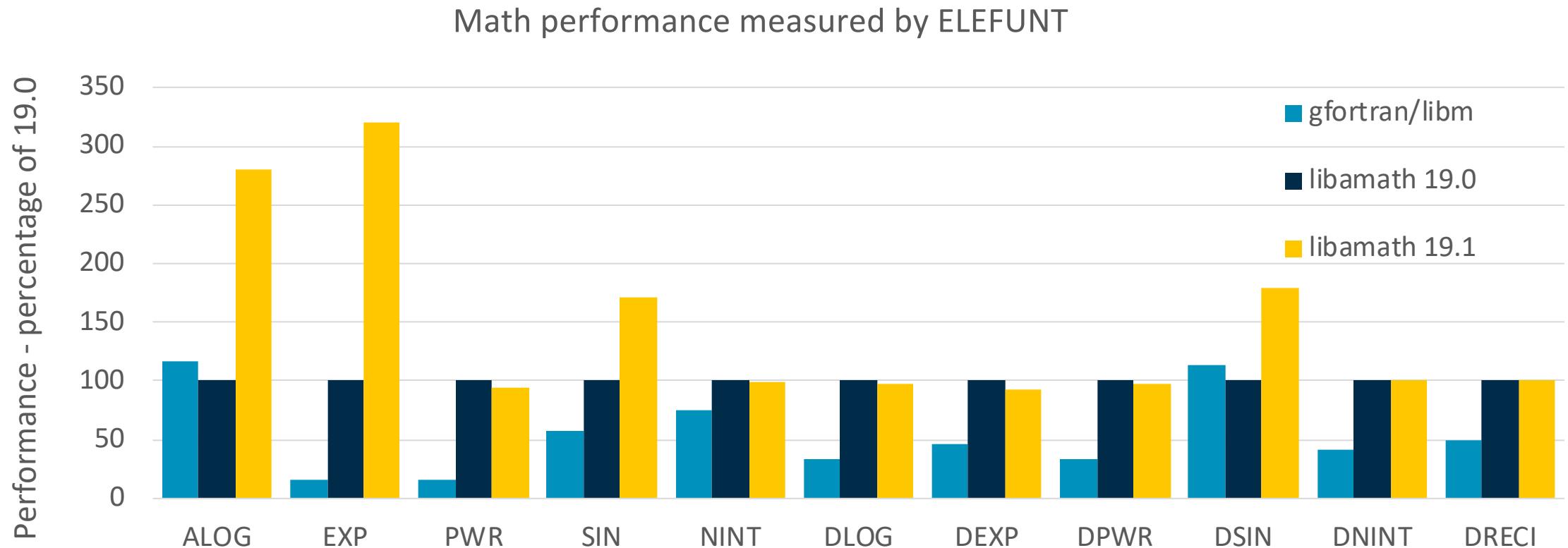
# FFT MPI performance in 19.1

Scaling using FFTW MPI interface improved; now similar scaling to FFTW



# Libamath – increased performance for math.h functions

ELEFUNT run on ThunderX2: cases no libamath, Arm Compiler with libamath 19.0 and 19.1



arm

# Cross-Platform tools

Arm Forge and Arm Performance Reports

# By Choosing Arm, You Choose a State-of-the-art Solution

- Interoperable
  - Available on the vast majority of HPC platforms, including AMD, IBM, Intel, Nvidia... and of course Arm!
- Performant
  - Fast, lightweight and transparent tools that help focus on the real issues that count
- Comprehensive
  - Packed with the best features to slash the development overhead spent on debugging and optimising issues

# Arm Forge Professional

A cross-platform toolkit for debugging and profiling



Commercially supported  
by Arm



Fully Scalable



Very user-friendly

The de-facto standard for HPC development

- Available on the vast majority of the Top500 machines in the world
- Fully supported by Arm on x86, IBM Power, Nvidia GPUs, etc.

State-of-the art debugging and profiling capabilities

- Powerful and in-depth error detection mechanisms (including memory debugging)
- Sampling-based profiler to identify and understand bottlenecks
- Available at any scale (from serial to petaflopic applications)

Easy to use by everyone

- Unique capabilities to simplify remote interactive sessions
- Innovative approach to present quintessential information to users

**arm**

# Arm Performance Reports

Characterize and understand the performance of HPC application runs



Commercially supported  
by Arm



Accurate and astute  
insight



Relevant advice  
to avoid pitfalls

Gathers a rich set of data

- Analyses metrics around CPU, memory, IO, hardware counters, etc.
- Possibility for users to add their own metrics

Build a culture of application performance & efficiency awareness

- Analyses data and reports the information that matters to users
- Provides simple guidance to help improve workloads' efficiency

Adds value to typical users' workflows

- Define application behaviour and performance expectations
- Integrate outputs to various systems for validation (e.g. continuous integration)
- Can be automated completely (no user intervention)

# Key highlights in Forge & Performance Reports

Latest 19.0 version released in Dec 2018

	Forge		Performance Reports
	DDT	MAP	
Packaging	<b>Creation of Allinea Studio</b> <i>A new solution for aarch64 platforms that includes the Arm Compiler, Arm Performance Libraries, and the former Allinea tools!</i>		
Platforms	<b>Full support for IBM systems</b> Arm v8 support CUDA 9 support	<b>Full support for IBM systems</b> Arm v8 support CUDA 9 support	<b>Full support for IBM systems</b> Arm v8 support CUDA 9 support
Improvements	<b>Usability Improvements</b> Memory debugging optimizations	Optimizations for many-core systems	Optimizations for many-core systems
New Features	Combined C/C++/Fortran and Python Debugging	<b>Python profiling</b> Backfill Custom Metrics On-kernel GPU profiling Ability to profile selected ranks	<b>Python performance analysis</b> Ability to profile selected ranks

# Forge and Performance Reports – Future roadmap

Why do our tools matter and what will we focus on this year?

## Reduce migration costs and increase portability

Finding and using the right hardware is hard, even more so because of porting and migration costs.

We will keep providing **cross-platform tools** to enable **choice** and innovation in HPC.

## Slash down code validation costs and time

For every run in production, codes are run 3 to 5 times to validate they meet standards.

We will assist the community reduce their testing costs by promoting **best practices** and tightening the link between tools **agile continuous delivery**.

## Provide capabilities on demand

Too often, users are stopped in their work by licence sizes limitations.

We will work on providing **capabilities** to users **on demand** at any time.

# Forge/Performance Reports Roadmap 2018-2019

Key highlights for Forge/PR 19.1 and 19.2

Continuous work	19.1	19.2/20.0
<ul style="list-style-type: none"><li>• Support for latest software environments (MPI, compilers, etc.)</li><li>• Support for popular HPC systems (Intel, Arm, Power, GPUs...)</li><li>• Developing exclusive features in collaboration with vendors (e.g. HPE, etc.)</li></ul>	<ul style="list-style-type: none"><li>• Arithmetic evaluations of CPU metrics</li><li>• Assembly views to Forge</li><li>• Integration with DynamoRIO for low-level instrumentation of operations</li></ul>	<ul style="list-style-type: none"><li>• Addition of a “burst mode” in the tools</li><li>• Simplify the integration of tools within scripts</li><li>• Add the json, xml, csv outputs of the “offline” tools features</li></ul>

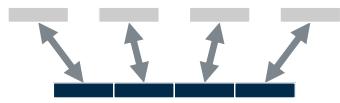


arm

# SVE - Introduction, tools and workflow

# Scalable Vector Extension (SVE)

A vector extension to the ARMv8-A architecture with some major new features



	1	2	3	4
	5	5	5	5
pred	1	0	1	0
=	6	2	8	4

```
for (i = 0; i < n; ++i)
INDEX i    n-2 | n-1 | n | n+1
CMPLT n    1   | 1   | 0 | 0
```

	1	2		
	1	2	0	0
pred	1	1	0	0

$$\begin{array}{cccc} 1 & + & 2 & + \\ \hline 1 & + & 2 & + \\ \hline 3 & & & \end{array} = \begin{array}{cc} 3 & + \\ \hline 7 & \end{array} =$$

## Gather-load and scatter-store

Loads a single register from several non-contiguous memory locations.

## Per-lane predication

Operations work on individual lanes under control of a predicate register.

## Predicate-driven loop control and management

Eliminate scalar loop heads and tails by processing partial vectors.

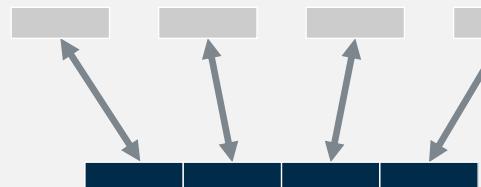
## Vector partitioning and software-managed speculation

First Faulting Load instructions allow memory accesses to cross into invalid pages.

## Extended floating-point horizontal reductions

In-order and tree-based reductions trade-off performance and repeatability.

# SVE is Arm's next generation SIMD ISA



	1	2	3	4
<i>pred</i>	5	5	5	5
=	1	0	1	0

	6	2	8	4
--	---	---	---	---

Per-lane predication

```
for (i = 0; i < n; ++i)
```

INDEX <i>i</i>	n-2	n-1	n	n+1
CMPLT <i>n</i>	1	1	0	0

Predicate-driven loop  
control and management

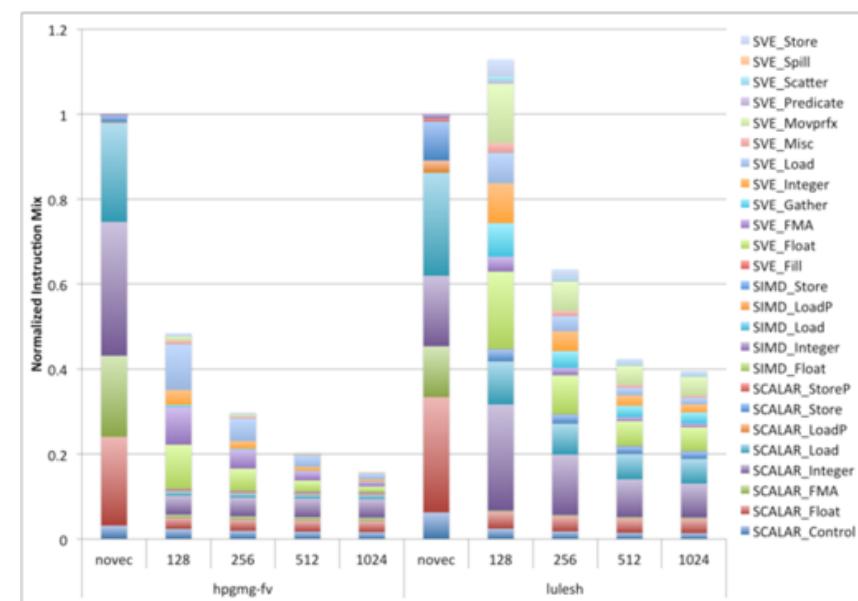
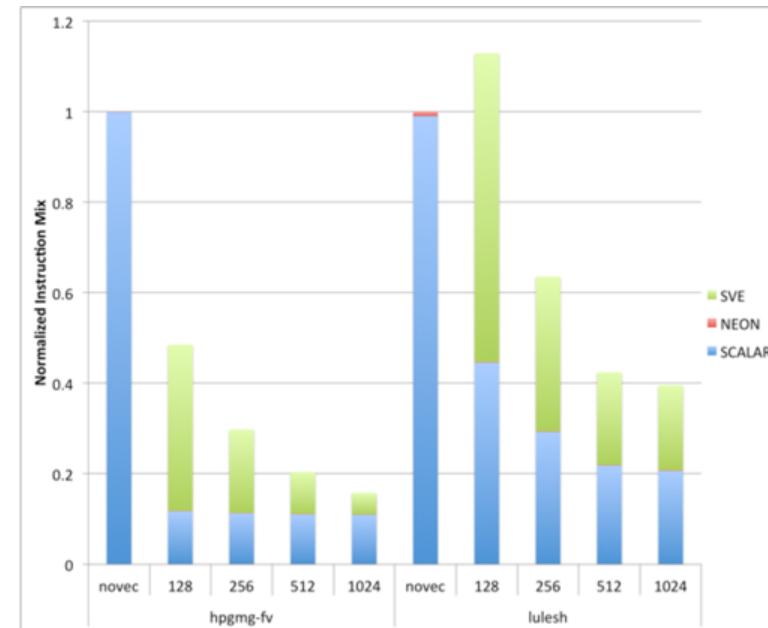
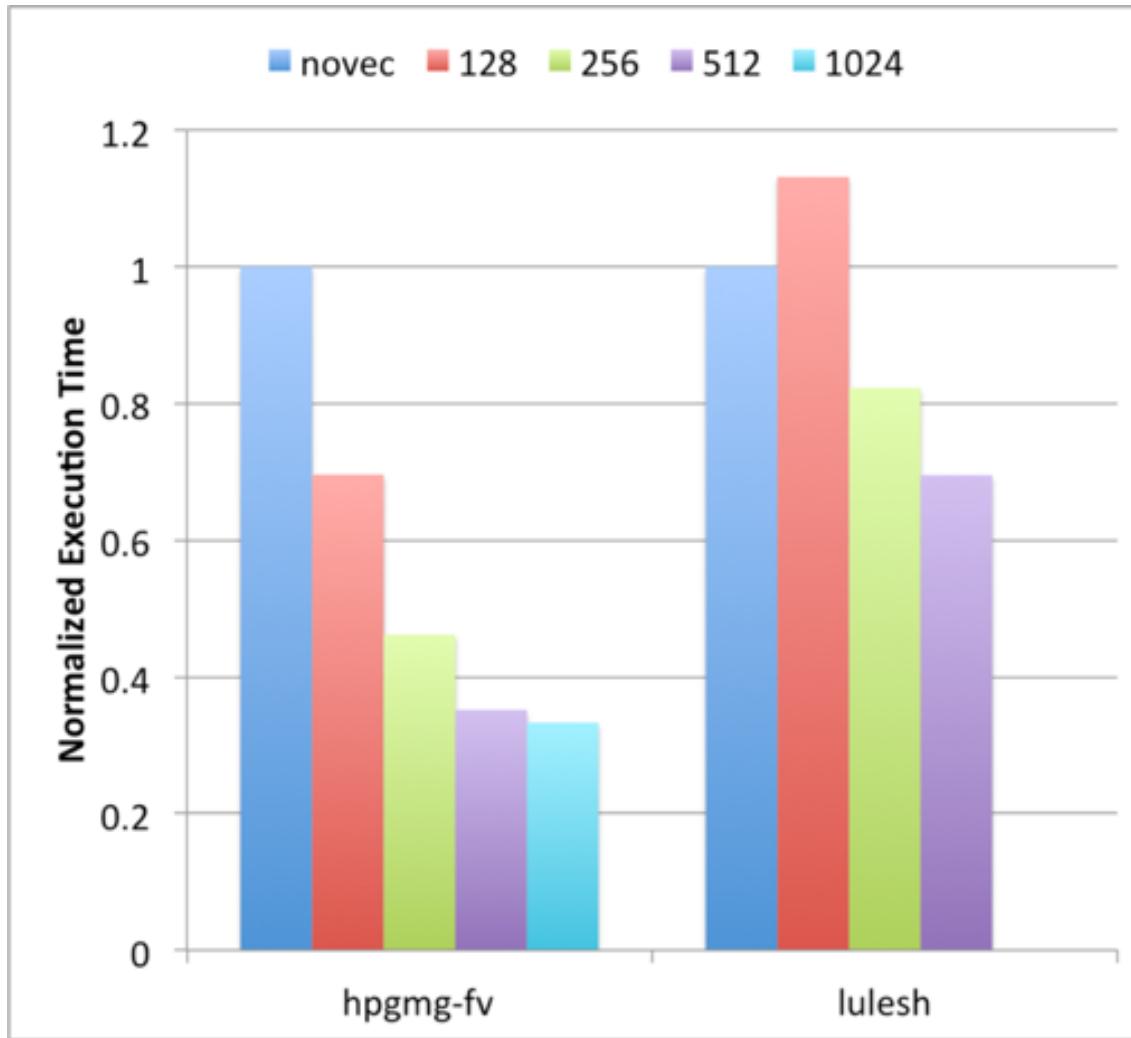
  
$$\begin{matrix} & \text{1} & \text{2} \\ + & \text{1} & \text{2} \\ \text{pred} & \text{1} & \text{1} \end{matrix} = \begin{matrix} \text{0} & \text{0} \end{matrix}$$

Vector partitioning and  
software-managed speculation

$$\begin{matrix} \text{1} & + & \text{2} & + & \text{3} & + & \text{4} \\ \text{1} & + & \text{2} & + & \text{3} & + & \text{4} \\ = & & = & & = & & = \end{matrix} = \begin{matrix} \text{3} & + & \text{7} \end{matrix} = \begin{matrix} \text{10} \end{matrix}$$

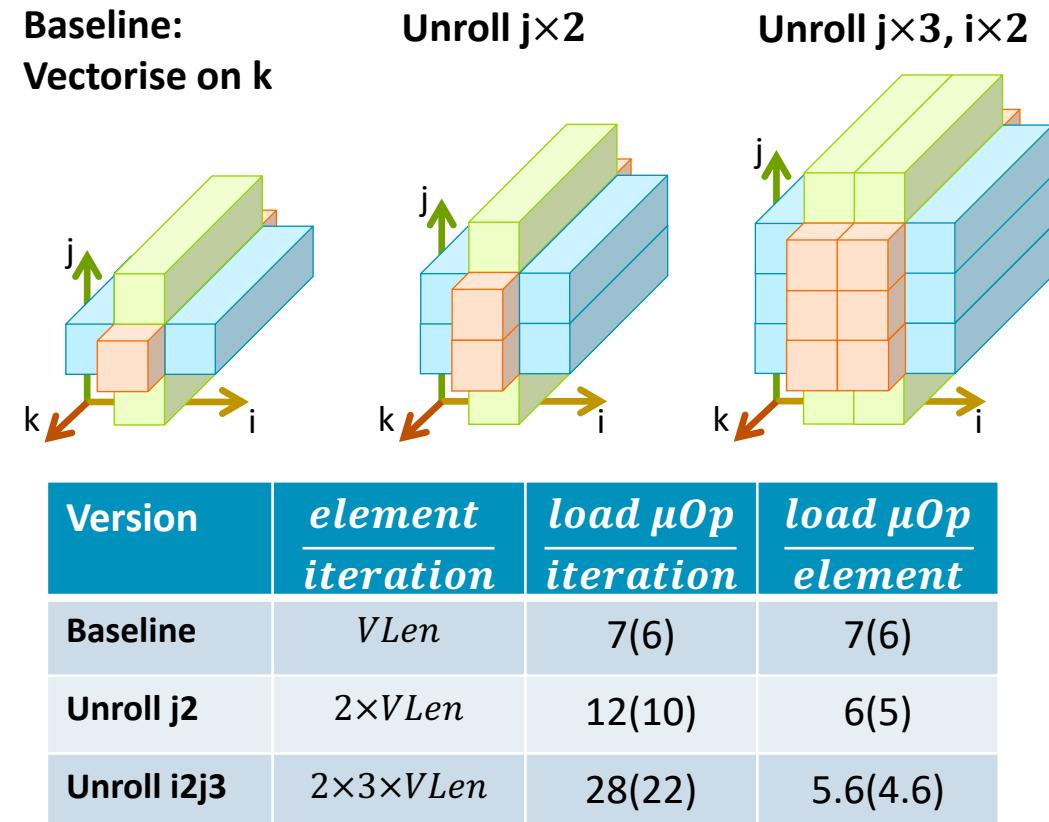
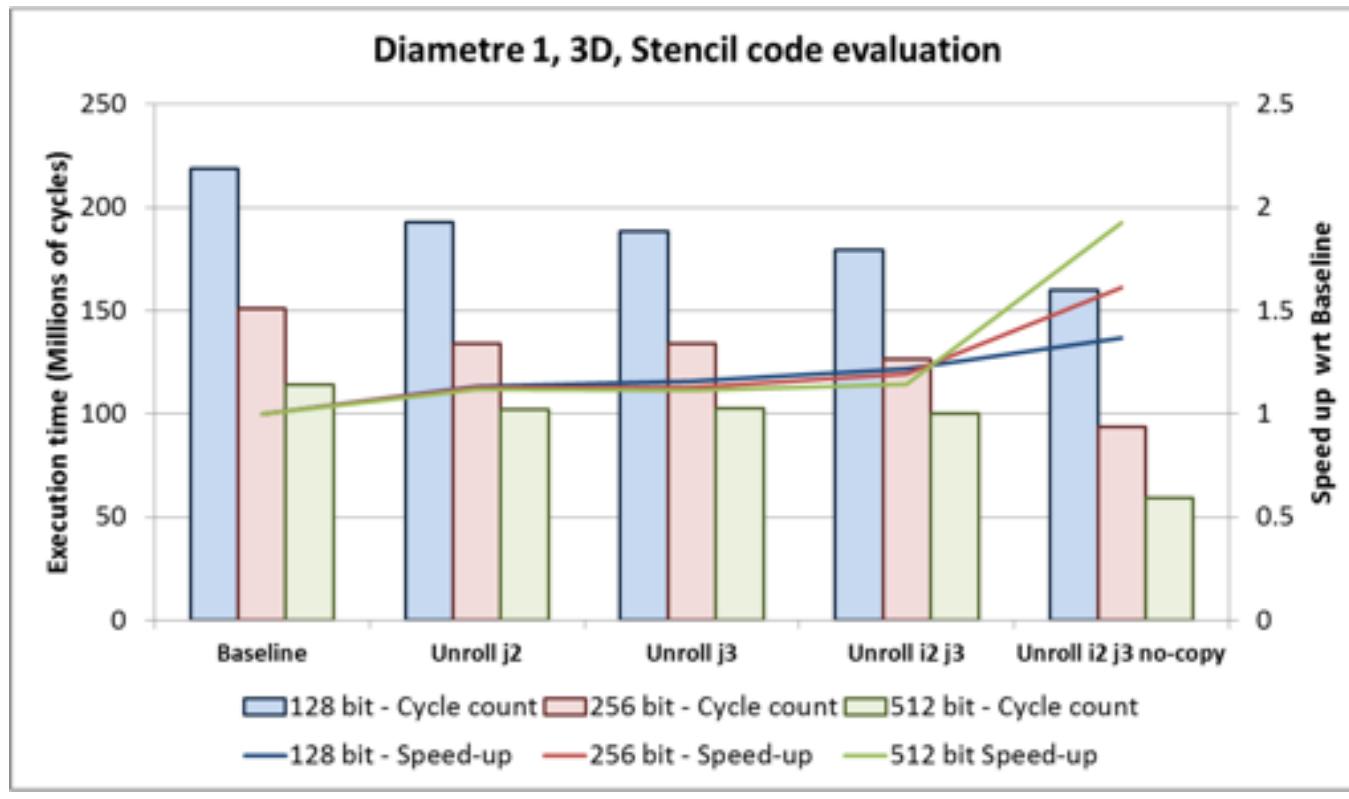
Extended floating-point  
horizontal reductions

# SVE: HPGMG & Lulesh



# SVE: Optimizing Stencil

- What are the effects of *Vector Length Agnosticism*?
- How well suited is the ISA to express the semantics of stencil codes?



# Open source support

- **Arm actively posting SVE open source patches upstream**
  - Beginning with first public announcement of SVE at HotChips 2016.
- **Available upstream**
  - GNU Binutils-2.28: released Feb 2017, includes SVE assembler & disassembler.
  - GCC 8: Full assembly, disassembly and basic auto-vectorization
  - GDB 8.2 SVE support
  - LLVM 7: Full assembly, disassembly
  - Linux kernel: since Mar 2017
  - QEMU 3.1: SVE support (user-space and system mode)
- **Under upstream review**
  - LLVM: since Nov 2016, as presented at LLVM conference.

# Compiler support

Feature	Upstream GCC	Upstream LLVM	Arm Compiler 6 (For bare metal)	Arm Linux Compiler (for Linux user-space)
SVE asm and disasm	Yes	Yes	Yes	Yes
SVE code generation	Yes	No Planned for 2019-20	Yes	Yes
SVE ACLE	No Planned for GCC10 (2020)	No Planned for 2019-20	Yes	Yes
Auto-vectorization	Basic More improvements planned for GCC9	None Planned for 2019-20	Advanced	Advanced

# Getting ready for SVE



## Port to Arm

- Port to current Arm hardware – Single node and multi-node
- Tune it for current Arm hardware



## Get ready for SVE

- Port to SVE using QEMU and/or ArmIE on current Arm hardware



## Tune for SVE

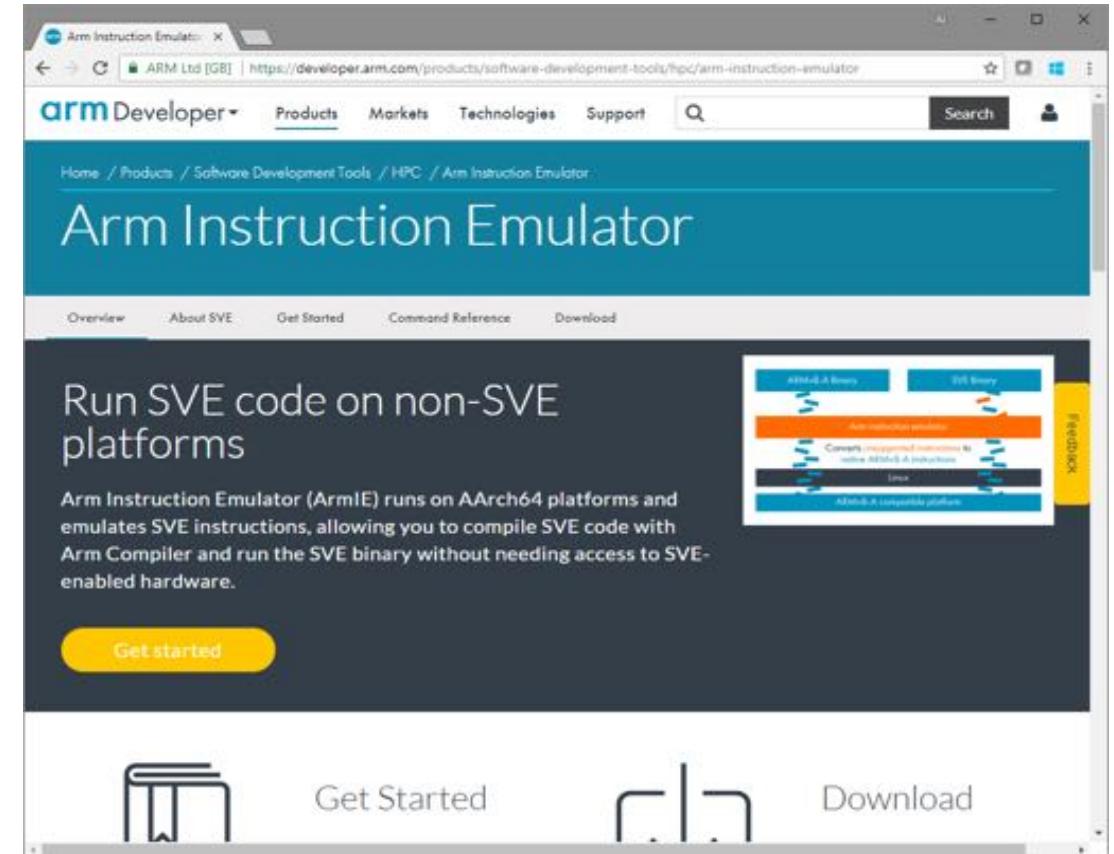
- On real SVE hardware

Co-work with Arm tools and professional services team

# Arm Instruction Emulator for SVE

*Develop tomorrow's software on today's hardware*

- Simple “black box” tool aimed at userspace software developers
  - \$ **armclang hello.c --march=sve**
  - \$ **./a.out**
  - Illegal instruction**
  - \$ **armie -msve-vector-bits=256 -- ./a.out**
  - Hello**
- Runs userspace application binaries at close to native speed
  - runs multithreaded applications
  - transparent to system calls
- Intercepts and emulates use of ARM instructions newer than hardware



# Arm Instruction Emulator

Develop your user-space applications for future hardware today



Develop software for  
tomorrow's hardware today



Runs at close to  
native speed



Commercially Supported  
by Arm

Start porting and tuning for future architectures early

- Reduce time to market, Save development and debug time with Arm support

Run 64-bit user-space Linux code that uses new hardware features on current Arm hardware

- SVE support available now.
- Tested with Arm Architecture Verification Suite (AVS)

Near native speed with commercial support

- Emulates only unsupported instructions
- Maintained and supported by Arm for a wide range of Arm-based SoCs

# DynamoRIO

Dynamic Binary Instrumentation

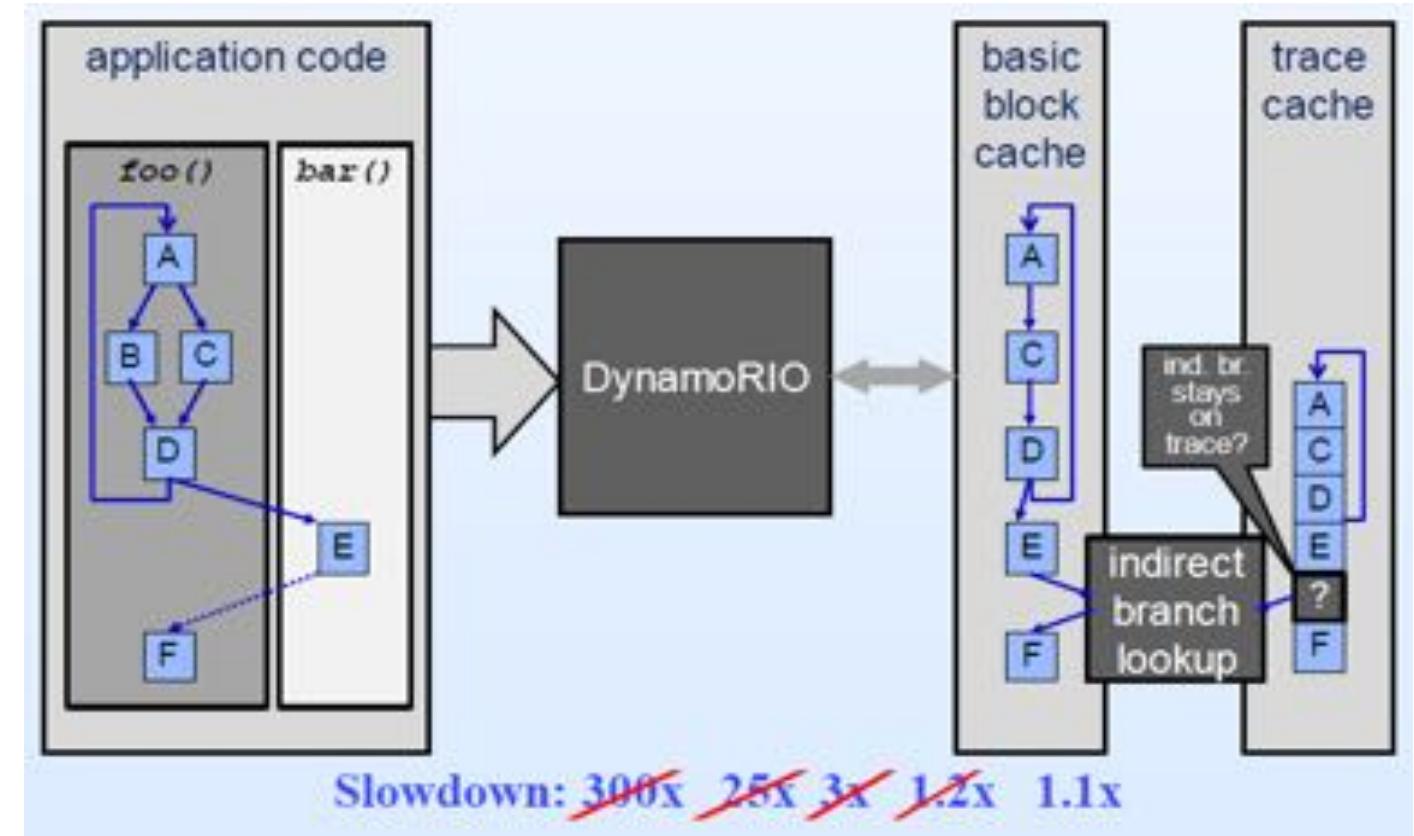
Fast code translation in userspace

Originally developed in MIT

Now managed by Google

Used for

- profiling
- valgrind-like checking
- architecture emulation



# Key points of contact

Visit [www.arm.com/hpc-tools](http://www.arm.com/hpc-tools) for further information

## Product team

David Lecomber

Sr Director, Infrastructure tools

Ashok Bhat

Sr Product manager – Compiler and Libraries

Patrick Wohlschlegel

Sr Product manager – Forge and Perf Reports

## Sales team

Rob Rick and Andrew Westergren – Americas

Marcin Krzysztofik – EMEA, India and China

Toshinori Kujiraoka – Japan