Questions

- ARM – Who are we?
- ARM deployments – what’s already happened and what is announced?
- SVE – How can you make a processor designed for mobile relevant for HPC?
- Software – How do you build an ARM HPC software ecosystem?
High Performance Computing and ARM
An introduction to ARM

ARM is the world's leading semiconductor intellectual property supplier

We license to over 350 partners: present in 95% of smart phones, 80% of digital cameras, 35% of all electronic devices, and a total of 60 billion ARM cores have been shipped since 1990

Our CPU business model:

License technology to partners, who use it to create their own system-on-chip (SoC) products

- We may license an instruction set architecture (ISA) such as “ARMv8-A”
- or a specific implementation, such as “Cortex-A72”
- Partners who license an ISA can create their own implementation, as long as it passes the compliance tests

…and our IP extends beyond the CPU
ARM’s mission

- Deploy energy-efficient ARM-based technology, wherever computing happens…

Leading in wearables and the Internet of Things

~85% share of laptops, tablets, and smartphones

Driving the transformation of the network and data center to an Intelligent Flexible Cloud

Partnering to deliver data center efficiency

Enabling innovation and creativity with embedded intelligence

Taking mobile computing to the next four billion people
ARMv8.0-A

RISC architecture
- Only have 32 bits available for encoding all instructions
- Supports the development of efficient implementations
- Restricts our ability to add lots of specialized instructions

64-bit capable since 2012
- Known as AArch64 (or AArch32 when run in a 32-bit mode)
- Initially 32- and 64-bit floating point (16-bit added later)
- 128-bit vector unit (known as Advanced SIMD / NEON)

Evolving
- Since 2014 we have an “annual tick” of point releases
- New instructions are “auditioned” for potential value in target applications
ARM DYNAMIQ – Multicore redefined

New single cluster design
Greater flexibility with or without big.LITTLE
Redesigned memory sub-system
Advanced compute capabilities
Range of SoCs addressing infrastructure

Highly Accelerated

Massively Multicore

- XILINX
- ZYNQ
- NXP
- QorIQ
- Qualcomm
- Altera
- Stratix 10
- Mellanox
- BlueField
- SoCionext SC2A11
- Applied Micro
- X-Gene 3
- Cavium
- ThunderX2

One size does not fit all
The Road to Exascale

- **GIGASCALE**
  - 1985: 34.6 GFLOPs
  - 1990: 10
  - 1995: 100000
  - 2000: 10000000
  - 2005: 1E+09
  - 2010: 1E+09
  - 2015: 1E+09
  - 2020: 1E+09

- **TERASCALE**
  - 1985: 0.1
  - 1990: 1
  - 1995: 10
  - 2000: 100
  - 2005: 1000
  - 2010: 10000
  - 2015: 100000
  - 2020: 1000000

- **PETASCALE**
  - 1985: 0.1
  - 1990: 1
  - 1995: 10
  - 2000: 100
  - 2005: 1000
  - 2010: 10000
  - 2015: 100000
  - 2020: 1000000

- **EXASCALE**
  - 2020: 93 PFLOPs

*Courtesy of Oak Ridge National Laboratory, U.S. Dept. of Energy*
HPC Leadership: International Exascale Programs

United States
ARM is currently a participant in two Department of Energy funded Exascale projects: Data Movement Dominates and Fast Forward 2

European Union
Through FP7 and Horizon 2020, ARM has been involved in several funded pre-Exascale projects including the Mont Blanc program which deployed one of the first ARM prototype HPC systems

Japan
Fujitsu and RIKEN announced that the Post-K system targeted at Exascale will be based on ARMv8 with new Scalable Vector Extensions

China
James Lin, vice director for the Center of HPC at Shanghai Jiao Tong University claims China will build three pre-Exascale prototypes to select the architecture for their Exascale system. The three prototypes are based on AMD, SunWeiTaihuLight, and ARMv8
Serious ARM HPC deployments starting in 2017

Two big announcements about ARM in HPC in Europe

Bull Atos to Build HPC Prototype for Mont-Blanc Project using Cavium ThunderX2 Processor

Today the Mont-Blanc European project announced it has selected Cavium’s ThunderX2 ARM server processor to power its new HPC prototype.

The new Mont-Blanc prototype will be built by Atos, the coordinator of phase 3 of Mont-Blanc, using its Bull expertise and products. The platform will leverage the infrastructure of the Bull sequana pre-exascale supercomputer range for network, management, cooling, and power. Atos and Cavium signed an agreement to collaborate to develop this new platform, thus making Mont-Blanc an Alpha-site for ThunderX2.

Announcing the GW4 Tier 2 HPC service, 'Isambard': named after Isambard Kingdom Brunel

System specs:
- Cray CS-400 system
- 10,000+ ARMv8 cores
- HPC optimised software stack
- Technology comparison:
  - x86, KNL, Pascal
- To be installed March-Dec 2017
- £4.7m total project cost over 3 years
Japan

Post-K: Fujitsu HPC CPU to Support ARM v8

Post-K fully utilizes Fujitsu proven supercomputer microarchitecture.

Fujitsu, as a lead partner of ARM HPC extension development, is working to realize ARM Powered® supercomputer w/ high application performance.

ARM v8 brings out the real strength of Fujitsu’s microarchitecture.

<table>
<thead>
<tr>
<th>HPC apps acceleration feature</th>
<th>Post-K</th>
<th>FX100</th>
<th>FX10</th>
<th>K computer</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMA: Floating Multiply and Add</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Math. acceleration primitives*</td>
<td>✔ Enhanced</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Inter core barrier</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Sector cache</td>
<td>✔ Enhanced</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Hardware prefetch assist</td>
<td>✔ Enhanced</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Tofu interconnect</td>
<td>✔ Integrated</td>
<td>✔ Integrated</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

* Mathematical acceleration primitives include trigonometric functions, sine & cosines, and exponential...
Introducing the Scalable Vector Extension (SVE)

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

- **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
What’s the vector length?

There is no preferred vector length

- Vector Length (VL) is the CPU implementor’s choice, from 128 to 2048 bits, in increments of 128
- Adopting a Vector Length Agnostic (VLA) code generation style makes code portable across all possible vector lengths
- VLA is made possible by the per-lane predication, predicate-driven loop control, vector partitioning and software-managed speculation features of SVE
- No need to recompile, or to rewrite hand-coded SVE assembler or C intrinsics

https://developer.arm.com/hpc/resources
ARM HPC software ecosystem
Ecosystem for HPC

List of components needed:

▪ Linux OS availability
▪ Compilers
▪ Libraries
▪ Job schedulers
▪ Debuggers
▪ Profilers

Mix of open source and commercial products and applications…

https://developer.arm.com/hpc/hpc-software
OpenHPC is a community effort to provide a common, verified set of open source packages for HPC deployments

**ARM’s participation:**
- Silver member of OpenHPC
- ARM is on the OpenHPC Technical Steering Committee in order to drive ARM build support

**Status:** 1.3 release out now
- All packages built on ARMv8 for CentOS and SUSE
- ARM-based machines are being used for building and also in the OpenHPC build infrastructure

<table>
<thead>
<tr>
<th>Functional Areas</th>
<th>Components include</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base OS</td>
<td>RHEL/CentOS 7.1, SLES 12</td>
</tr>
<tr>
<td>Administrative Tools</td>
<td>Conman, Ganglia, Lmod, LosF, ORCM, Nagios, pdsh, prun</td>
</tr>
<tr>
<td>Provisioning</td>
<td>Warewulf</td>
</tr>
<tr>
<td>Resource Mgmt.</td>
<td>SLURM, Munge, Altair PBS Pro*</td>
</tr>
<tr>
<td>I/O Services</td>
<td>Lustre client (community version)</td>
</tr>
<tr>
<td>Numerical/Scientific Libraries</td>
<td>Boost, GSL, FFTW, Metis, PETSc, Trilinos, Hypre, SuperLU, Mumps</td>
</tr>
<tr>
<td>I/O Libraries</td>
<td>HDF5 (pHDF5), NetCDF (including C++ and Fortran interfaces), Adios</td>
</tr>
<tr>
<td>Compiler Families</td>
<td>GNU (gcc, g++, gfortran)</td>
</tr>
<tr>
<td>MPI Families</td>
<td>OpenMPI, MVAPICH2</td>
</tr>
<tr>
<td>Development Tools</td>
<td>Autotools (autoconf, automake, libtool), Valgrind,R, SciPy/NumPy</td>
</tr>
<tr>
<td>Performance Tools</td>
<td>PAPI, Intel IMB, mpiP, pdtoolkit TAU</td>
</tr>
</tbody>
</table>
ARM HPC software portfolio

**ARM C/C++ Compiler**
COMMERCIALY SUPPORTED FOR HPC APPLICATIONS

**ARM Performance Libraries**
BLAS, LAPACK and FFT MICRO-ARCHITECTURALLY TUNED

**ARM Code Advisor**
ACTIONABLE ADVICE TO OPTIMIZE YOUR CODE

Adding depth with support for next-generation ARM HPC architectures with Scalable Vector Extensions

**ARM SVE C/C++ Compiler**
COMPILER SUPPORT FOR ARM SCALABLE VECTOR EXTENSION

**ARM Instruction Emulator**
DEVELOP SOFTWARE FOR TOMORROW’S HARDWARE TODAY

**Allinea Forge (DDT+MAP)**
PARALLEL DEBUGGING and PROFILING

**Allinea Performance Reports**
FULL APPLICATION PROFILING
Open source in the ARM HPC ecosystem

- Over the past 12 months many more packages and applications have been successfully ported to ARM HPC
Summary

- ARM’s ecosystem is built on **partnership** and **choice**
  - We work with many organisations to drive hardware design and deliver better software
  - This method enables partners to design different products for different markets
- We license IP at all levels of the stack to help customers be successful
- Our 64-bit server platforms are beginning to see large, main-stream deployments
- Building the software ecosystem and tools is an important part of this story
  - We enhance open source software as well as developing commercially supported options
- Architectural enhancements from **SVE** will lead to significant performance increases for many end-user codes
Thank you!

Contact: Florent.Lebeau@arm.com