arm

# Energy-efficient HPC
## A full system story

Roxana Rusitoru | Senior Research Engineer

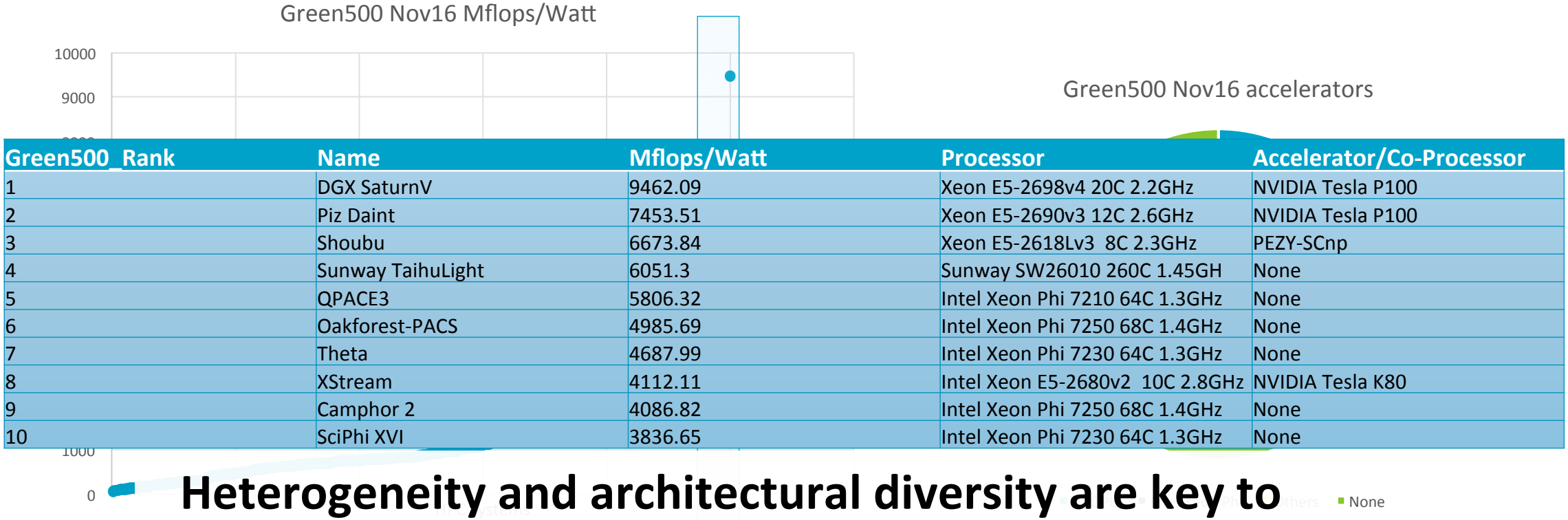
EEHPCWG@SC17

11/12/2017

Copyright © 2017 Arm Limited

# Towards a more energy-efficient future

1 ExaFLOP / 20 MW = 50 GFLOPS/W




| Green500_Rank | Name | Mflops/Watt | Processor | Accelerator/Co-Processor |
|---|---|---|---|---|
| 1 | DGX SaturnV | 9462.09 | Xeon E5-2698v4 20C 2.2GHz | NVIDIA Tesla P100 |
| 2 | Piz Daint | 7453.51 | Xeon E5-2690v3 12C 2.6GHz | NVIDIA Tesla P100 |
| 3 | Shoubu | 6673.84 | Xeon E5-2618Lv3 8C 2.3GHz | PEZY-SCnp |
| 4 | Sunway TaihuLight | 6051.3 | Sunway SW26010 260C 1.45GH | None |
| 5 | QPACE3 | 5806.32 | Intel Xeon Phi 7210 64C 1.3GHz | None |
| 6 | Oakforest-PACS | 4985.69 | Intel Xeon Phi 7250 68C 1.4GHz | None |
| 7 | Theta | 4687.99 | Intel Xeon Phi 7230 64C 1.3GHz | None |
| 8 | XStream | 4112.11 | Intel Xeon E5-2680v2 10C 2.8GHz | NVIDIA Tesla K80 |
| 9 | Camphor 2 | 4086.82 | Intel Xeon Phi 7250 68C 1.4GHz | None |
| 10 | SciPhi XVI | 3836.65 | Intel Xeon Phi 7230 64C 1.3GHz | None |

**Heterogeneity and architectural diversity are key to energy efficiency**


© 2017 Arm Limited
arm Research

# Why ARM in HPC?

**Energy efficiency**

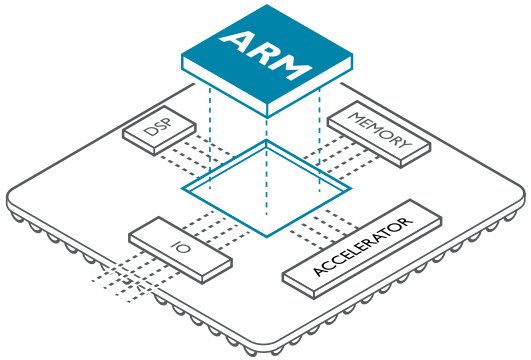
- Energy has always been a first-class design constraint at ARM, the goal in ARM's uArchitecture IP is to maintain that power efficiency advantage whilst increasing performance. The CPU core is only one element of the equation, and ARM are also working to explore efficiency in the memory and on-chip interconnect space.

**Choice**

- Independent silicon providers within a shared software ecosystem.
- Wide range of price/performance/power/size options available.

**Customisation**

- ARM's partners can build SoCs very quickly.





© 2017 Arm Limited

arm Research

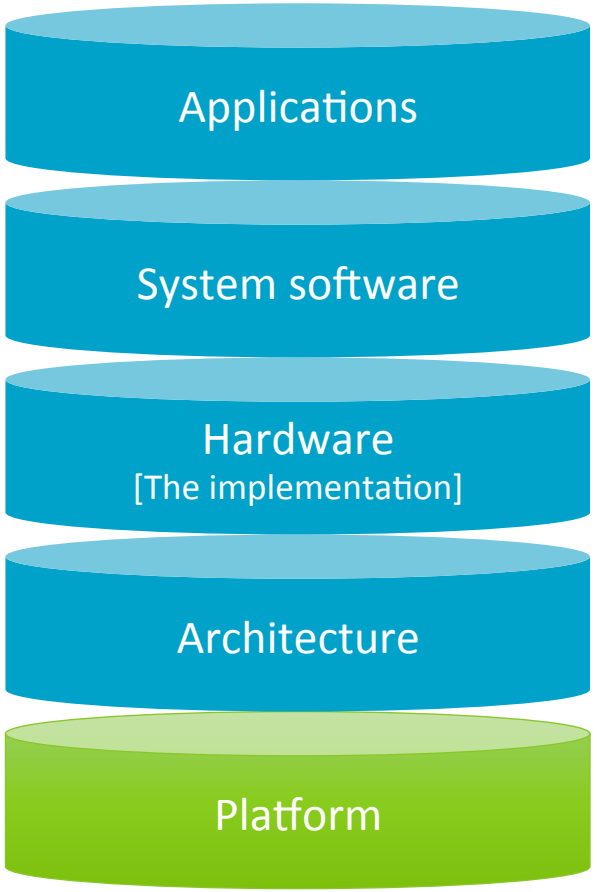
# What we're building


© 2017 Arm Limited

# The platform

## Bull Sequana

- MontBlanc-3 test platform basis
- Ultra-energy efficient, using Direct Liquid Cooling
- Optimized in density for 2-socket systems in 1U (4.45cm)
  - Cavium ThunderX2 – 256 hw threads





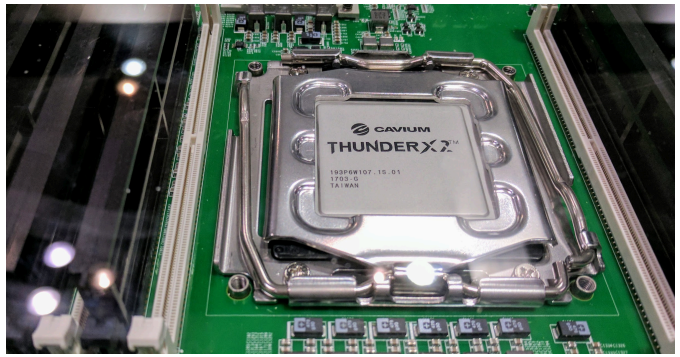
© 2017 Arm Limited

arm Research

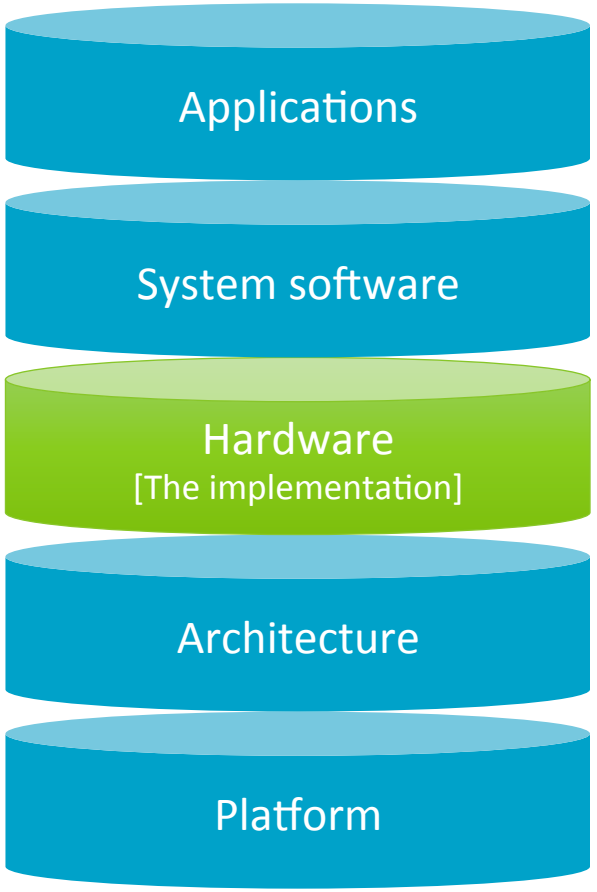
# The hardware

**Cavium ThunderX2**

- Multi-threaded, fully out-of-order high performance ARMv8 cores
- High memory bandwidth & capacity
- Extensive power management
- Comprehensive hardware and software ecosystem




Courtesy of https://goo.gl/1VJWGz

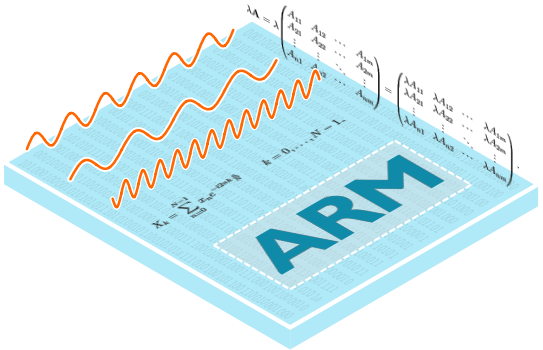




© 2017 Arm Limited

arm Research

# The system software




## Arm Performance Libraries

Enable the wide variety of ARM cores available today without adding complexity to the software ecosystem.

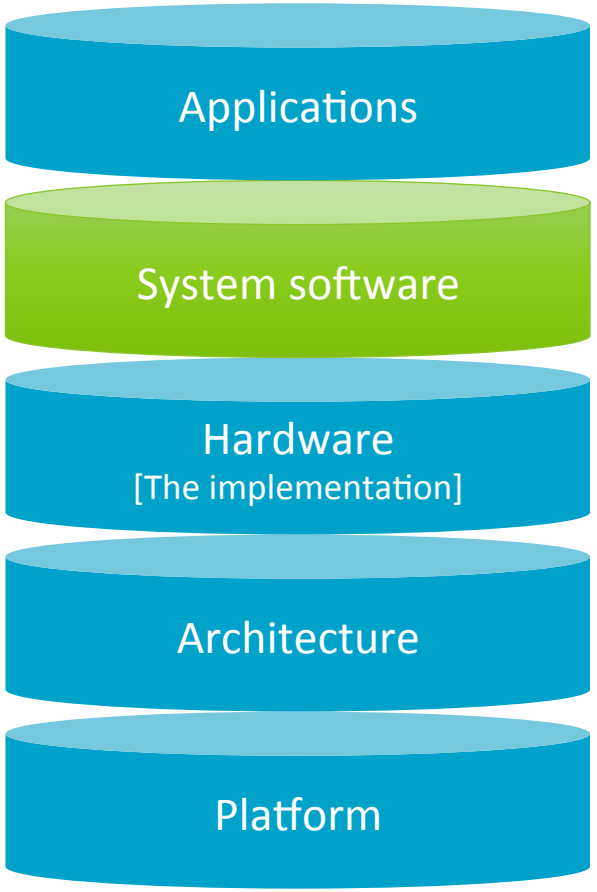
- Commercially supported 64-bit ARMv8 vendor math libraries for scientific computing.
- Arm silicon partners provide us with tuned kernels.

**Capabilities:**

➔ **BLAS**
➔ **LAPACK**
➔ **FFT**

**Tuned for:**

- Cortex-A53, A57, A72
- Applied Micro X-Gene®
- Cavium® ThunderX





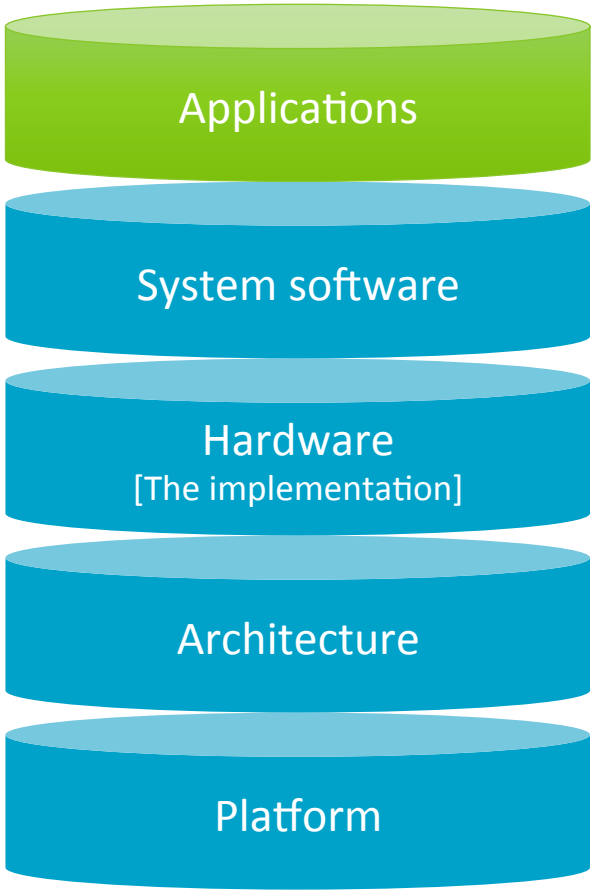
© 2017 Arm Limited

arm Research

# The applications

## Alya

- Barcelona Supercomputing Centre simulation code for multi-physics problems. It solves time-dependent Partial Differential Equations (PDEs).
- **Optimizations**
  - Parallelized Finite Element codes
    - Helps increase IPC
  - Taskified
    - Move towards throughput computing
    - Dynamic load balancing





© 2017 Arm Limited
arm Research

# Application issues

- Data layout
- Unused hardware resources
- Vectorization
- Memory bandwidth
- Dynamic load balance
- & more…

Courtesy of https://goo.gl/H56VpH


© 2017 Arm Limited

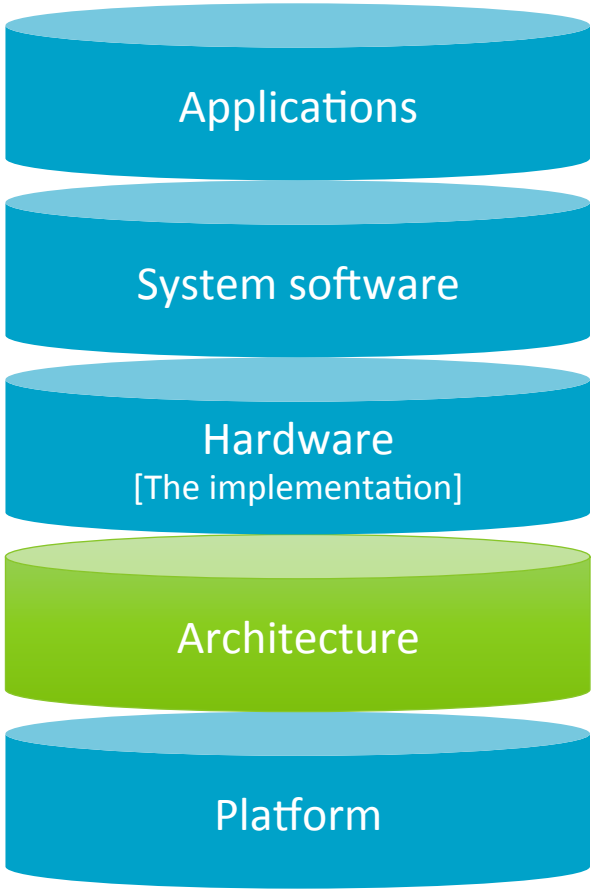
# Research


© 2017 Arm Limited

# The architecture

**Scalable Vector Extension (SVE)**

- Vector length is a hardware choice, from 128 to 2048 bits, in increments of 128 bits.
- Vector length agnostic programming adjusts dynamically to the available vector length.
- No need to recompile, or to re-write hand-coded SVE assembler or C intrinsics.
- SVE begins to address some of the traditional barriers to auto-vectorization (e.g. control flow).




 © 2017 Arm Limited

arm Research

# SVE – Key Architectural Features

- Scalable vector length (up to a current maximum of 2048 bits)
- Vector length Agnosticism
- Per-lane predication
- Gather-load and scatter-store
- Fault-tolerant speculative vectorization
- Vector partitioning
- Horizontal and serialized vector operations

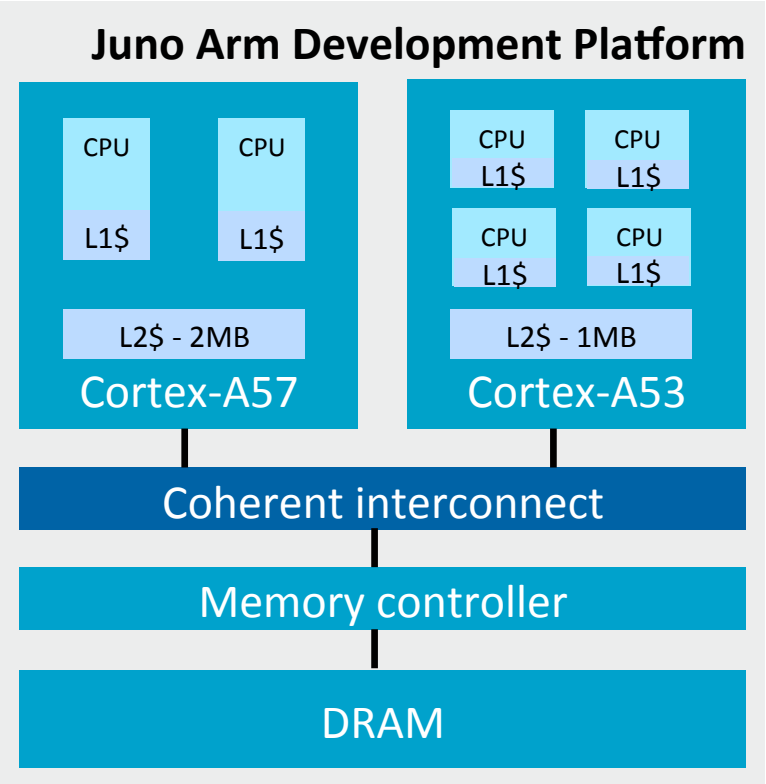
**But, I won't be talking about SVE any more today. Lots of talks about this later in the week.**

 © 2017 Arm Limited

arm Research

# Arm big.LITTLE

- High performance and high efficiency cache coherent CPU clusters
- Use the right processor at the right time
- Potential for energy savings
- Flexible and transparent to the applications
  - Enabled by the Global Task Scheduling (GTS) or Energy Aware Scheduling (EAS)
- Idea: maintain performance, save power, save area.
  - Mileage may vary depending on workload.
- Example platform: Juno Arm Development Platform

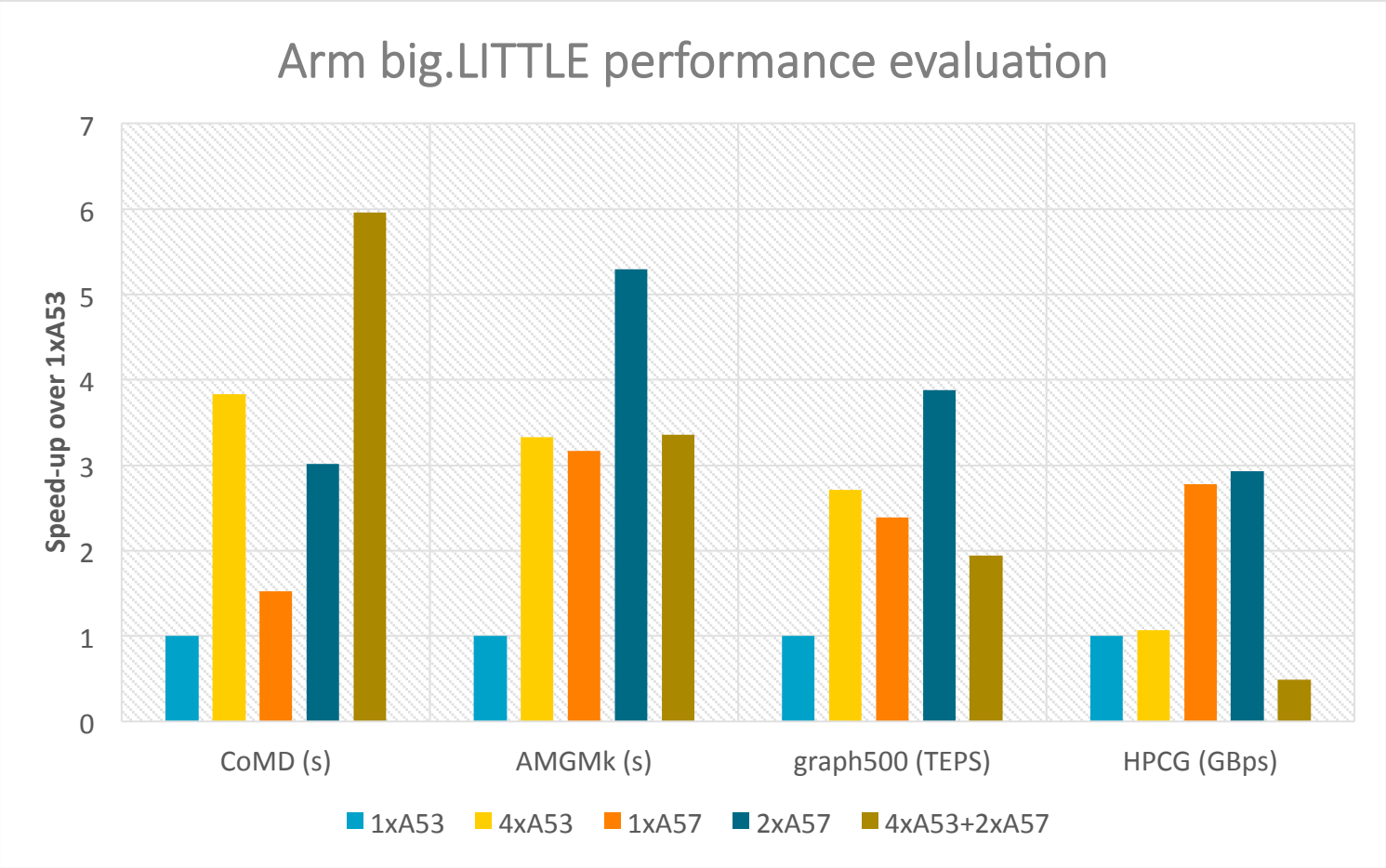




© 2017 Arm Limited

arm Research

# A few examples

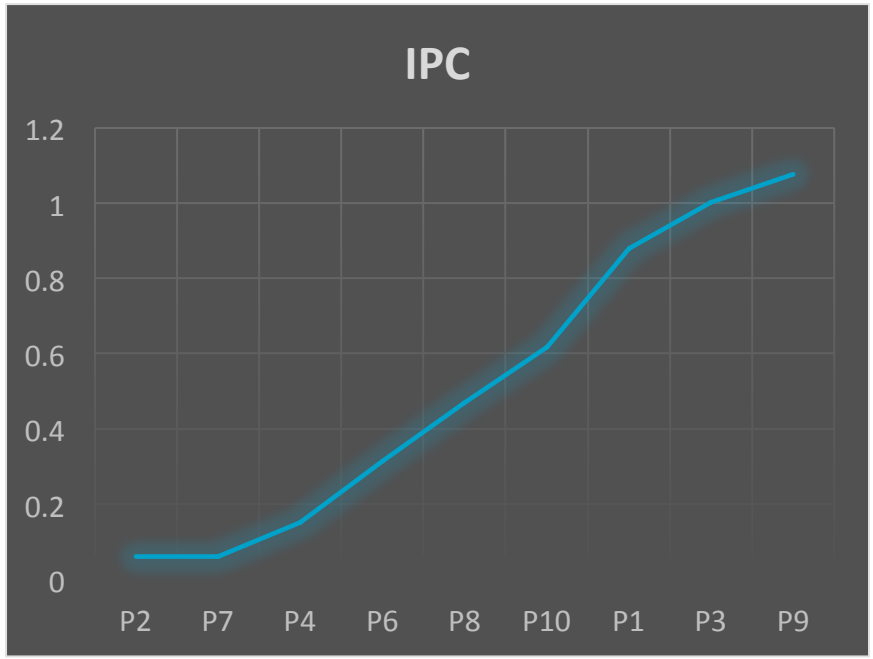



 © 2017 Arm Limited

arm Research

# Meabo

| | | |
|---|---|---|
| **Compute kernels** | 10 phases covering a range of Instructions-per-Cycle (IPC) points | |
| **Instrumentation** | Performance counter readings (via PAPI) | Power & Energy readings (currently proprietary) |
| **Infrastructure** | Multi-threaded support via OpenMP | Thread pinning |
| **Validation** | Validation for serial consistency | |



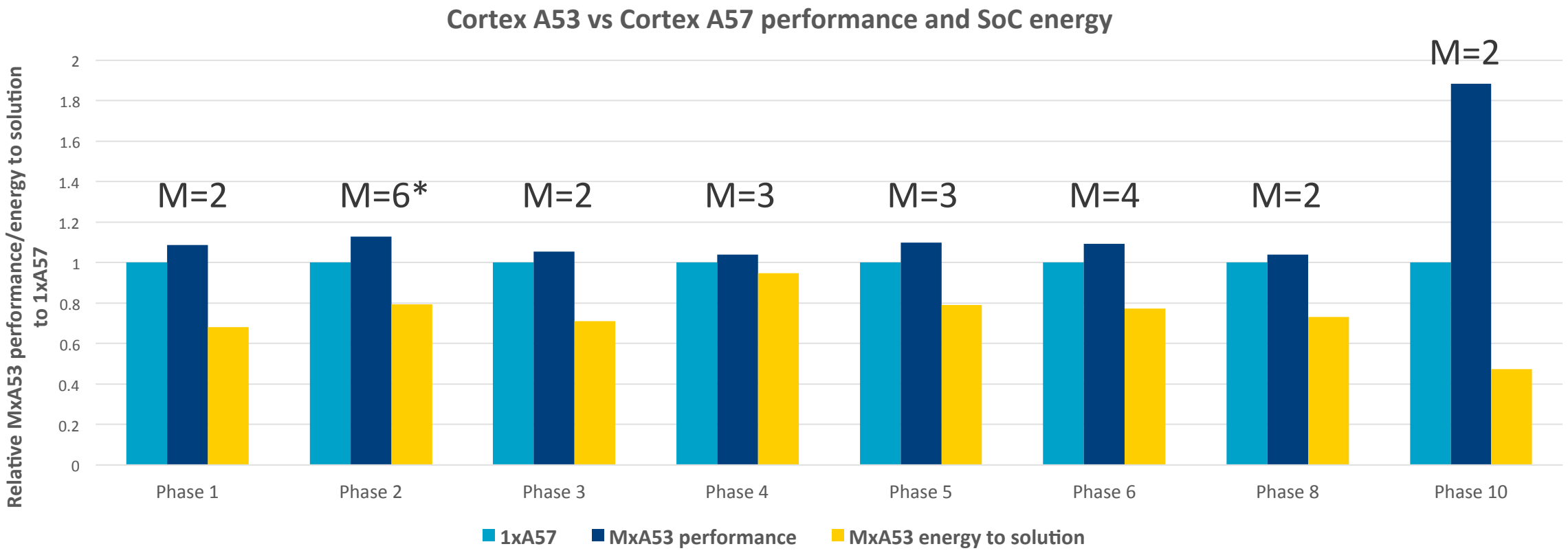

Available today on github: https://github.com/ARM-software/meabo


© 2017 Arm Limited

arm Research

# A57 vs A53 – A Meabo perspective of performance

## … and energy




* The 6-core Cortex-A53 numbers duration numbers are estimated based on trends observed from running on 1 to 4 Cortex-A53s.

All presented results were obtained on a Juno Arm Development Platform.


© 2017 Arm Limited

arm Research

# Arm big.LITTLE Observations

- There's more energy effciency to be gained by further exploring tightly-coupled heterogeneous systems, without having to sacrifice performance.
  - Mileage may vary depending on the workload.
- We need application parallelism.

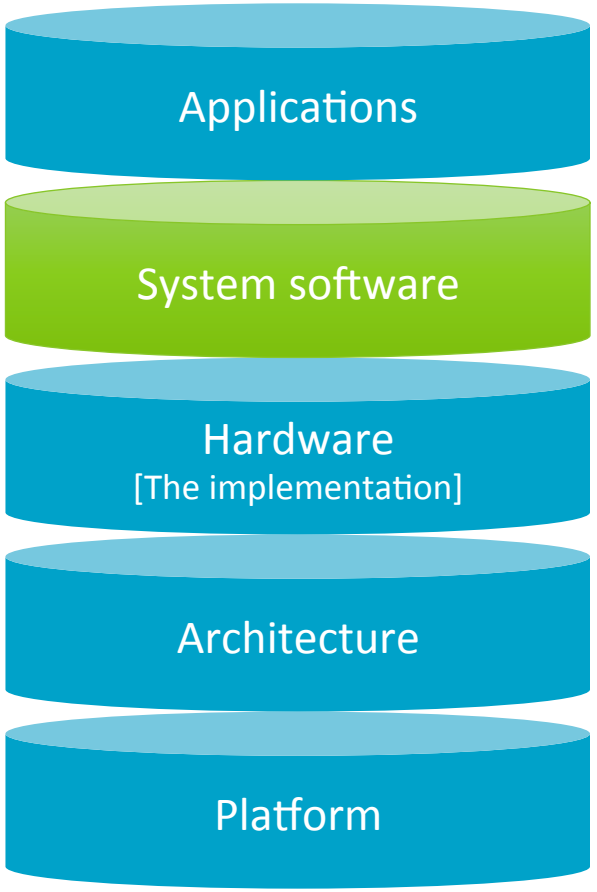
- Whilst maybe not the most popular, energy-efficient cores have the potential of being competitive in HPC.
  - For most Meabo phases, 2-3x A53s are on par with 1xA57.
- This does not exclude big cores, necessary for their higher single-thread performance for serial operations.


© 2017 Arm Limited

# The system software

## DAST

- Implemented on top of OmpSs (a research runtime similar to OpenMP)
  - Unit of compute: task
- Task management operations (e.g. allocation, submission) are managed by a separate thread.
- Use big or LITTLE cores to execute DAST
- Platform: Odriod-XU3, ARMv7 (4xArm Cortex-A7, 4xArm Cortex-A15, 2GB RAM)




 © 2017 Arm Limited

arm Research

# Conclusions

- We have a long way to go (~5x in system-wide energy efficiency)
- Energy efficiency is a full system story
- Heterogeneity and architectural diversity are key
- Improvements need to be made across all levels:
  - Platform
  - Architectural
  - Hardware
  - System software
  - Applications


© 2017 Arm Limited

The Arm trademarks featured in this presentation are registered trademarks or trademarks of Arm Limited (or its subsidiaries) in the US and/or elsewhere. All rights reserved. All other marks featured may be trademarks of their respective owners.

www.arm.com/company/policies/trademarks

arm Research

 © 2017 Arm Limited

Thank You!
Danke!
Merci!
谢谢!
ありがとう!
Gracias!
Kiitos!

arm Research


© 2017 Arm Limited