Архитектура ARM: от мобильных устройств до суперкомпьютеров

V. Nikolskiy, V. Stegailov

International Laboratory for Supercomputer Atomistic Modelling and Multi-scale Analysis

Mobile and embedded market

Architecture evolution

	ARMv6	ARMv7	ARMv8
64-bit	-	-	+
Multicore	-	+	+
Cache hierarchy	-	+	+
SP Scalar Floating-point	Optional	Optional	+
DP Scalar Floating-point	Optional	Optional	+
SP Floating-point SIMD	-	Optional	+
DP Floating-point SIMD	-	-	+

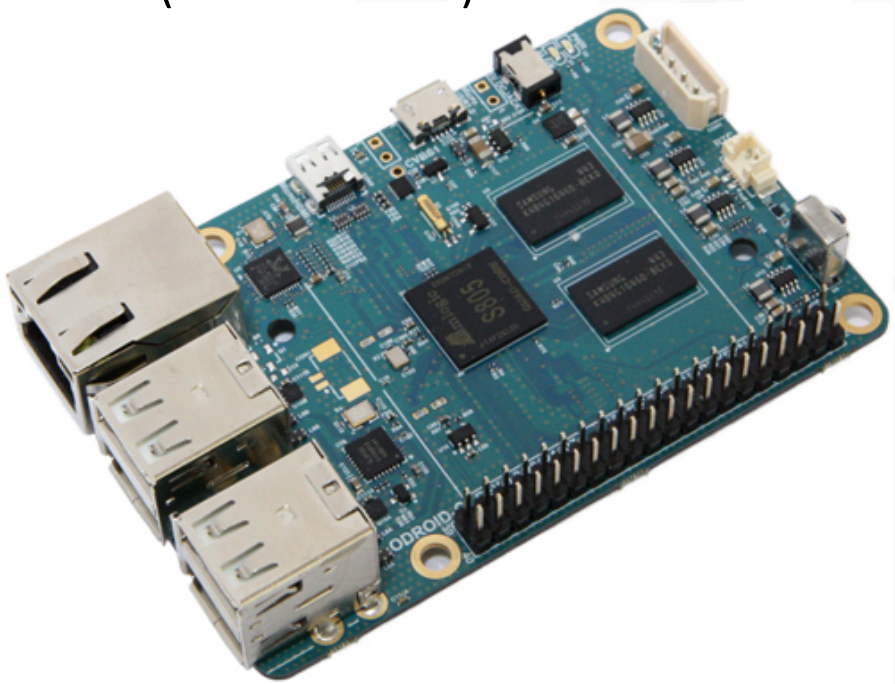
Mont-Blanc 2011-2018:

European approach towards energy efficient high performance computation

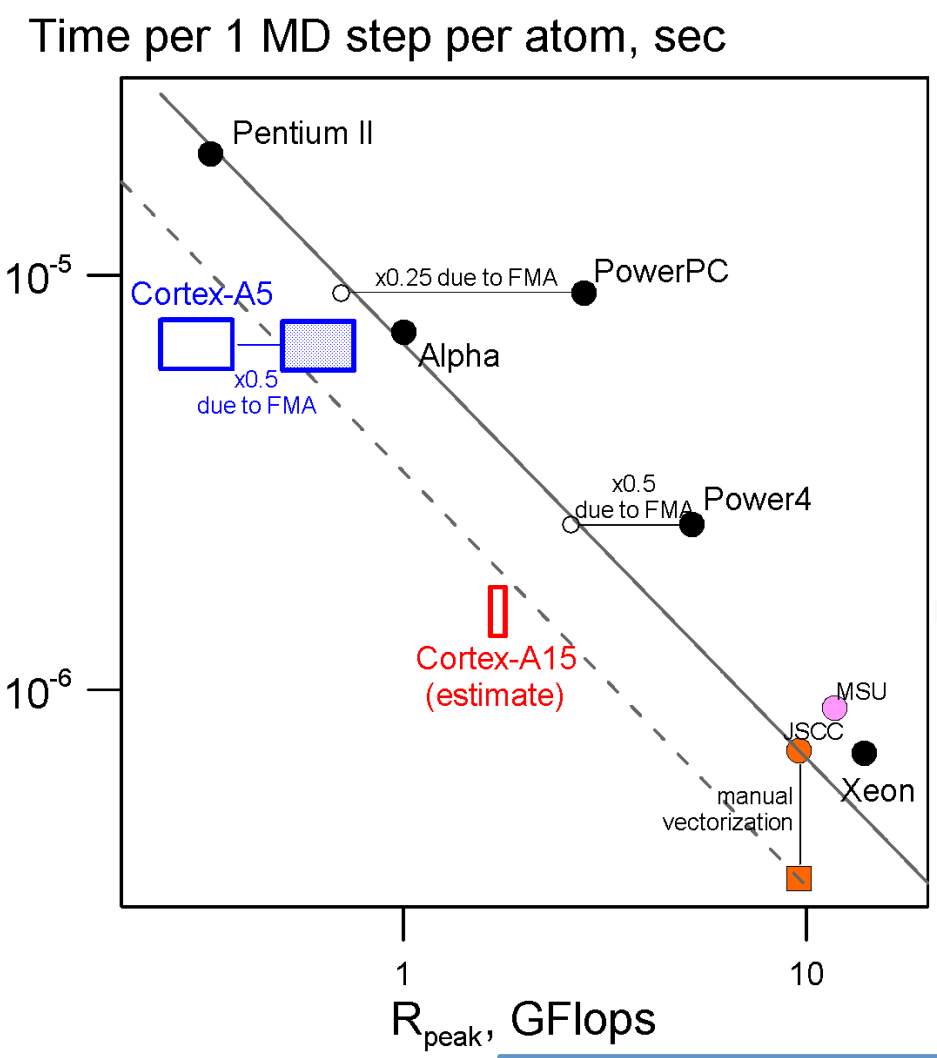
- To produce an architecture that will provide Exascale performance using 15 to 30 times less energy
- To develop the Mont-Blanc software ecosystem, with emphasis on programmer tools and system resiliency
- To design a well-balanced architecture for an ARM based SoC or SoP capable of providing pre-exascale performance when implemented in the 2019-2020

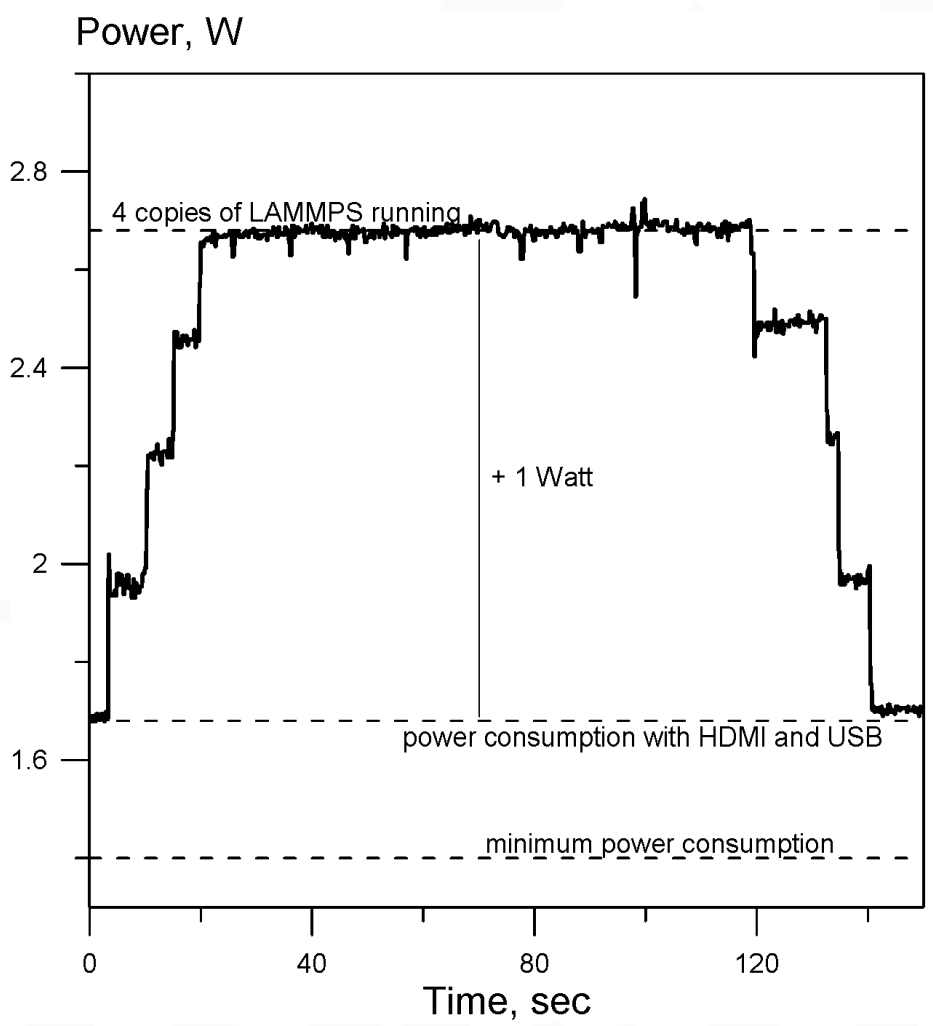
ODROID-C1 (2014-2015)

- Quad core processor Amlogic S805 *Cortex-A5*
  - 1.5GHz
  - VFPv4 modules on each core
  - graphics accelerator Mali-450 MP2 (was not used)
- 1GB DDR3 SDRAM
- OS:
  - Linux Ubuntu 14.05.1
  - or *Android* 4.4.2 OS



LAMMPS performance for different CPUs





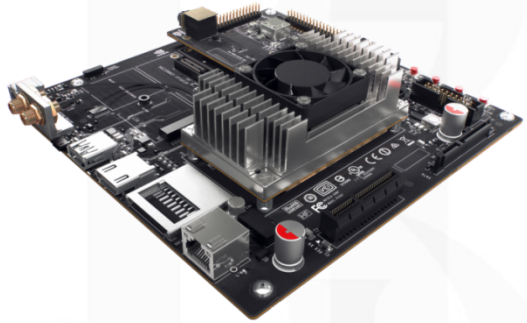
Nikolskiy V., Stegailov V. Floating-point performance of ARM cores and their efficiency in classical molecular dynamics // *Journal of Physics: Conference Series*. 2016. Vol. 681. P. 012049

Nvidia Jetson

Nvidia Jetson TK1 (2014)

- 32-bit Tegra K1 SoC
- 4 GB memory
- 4 Cortex-A15 cores @ 2.3 GHz
- 1 Kepler SM @ 852 МГц, 128 CUDA cores

Nvidia Jetson TX1 (2016)



- 64-bit Tegra X1 SoC
- 4 GB 64-bit memory
- 4 Cortex-A57 cores @ 2.1 GHz
- 2 Maxwell SM @ 998 МГц, 256 CUDA cores

Nvidia Jetson

Nvidia Jetson TX2 (2017)

- 2 Denver2 cores
  and
- 4 Cortex-A57 cores @ 2.0 GHz
- 8 GB 128-bit memory
- Pascal GPU

ARMv8 server-level processors

Cavium ThunderX2
- Up to 54 cores @ 3.0 GHz
- 14 nm FinFET
- 6 DDR4 controllers
- Integrated SATAv3
- Integrated PCIe Gen3

MACOM (AppliedMicro) X-Gene
- 32 cores @ 3.0+ GHz (16 nm FinFET)
- 8 DDR4-2667 memory channels
- 42 PCIe Gen3 lanes with eight controllers

Qualcomm
- 48 cores (10-nm FinFET)
- Multiple DDR4 memory controllers
- PCIe Gen3 x16
- Integrated IO and SATAv3 ports

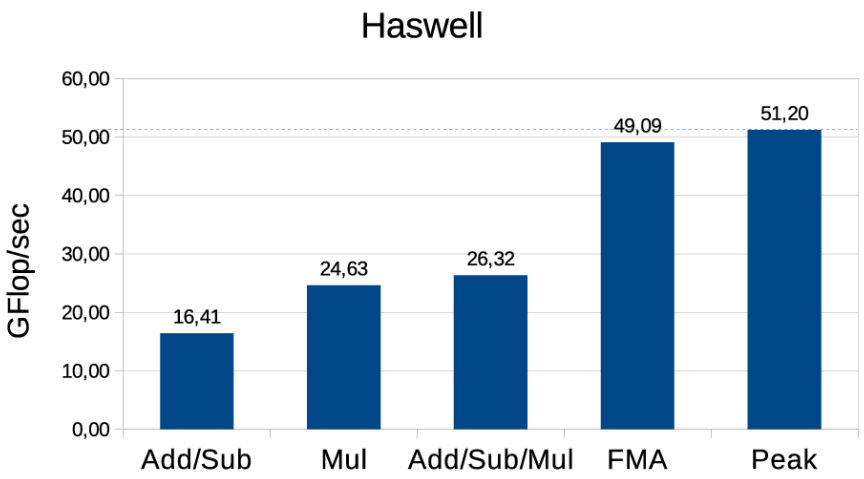
The Cray XC50

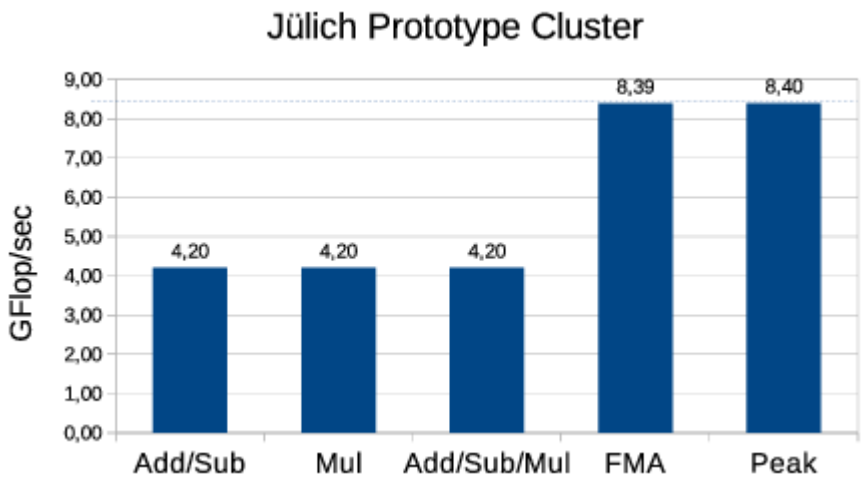
Julich Prototype cluster

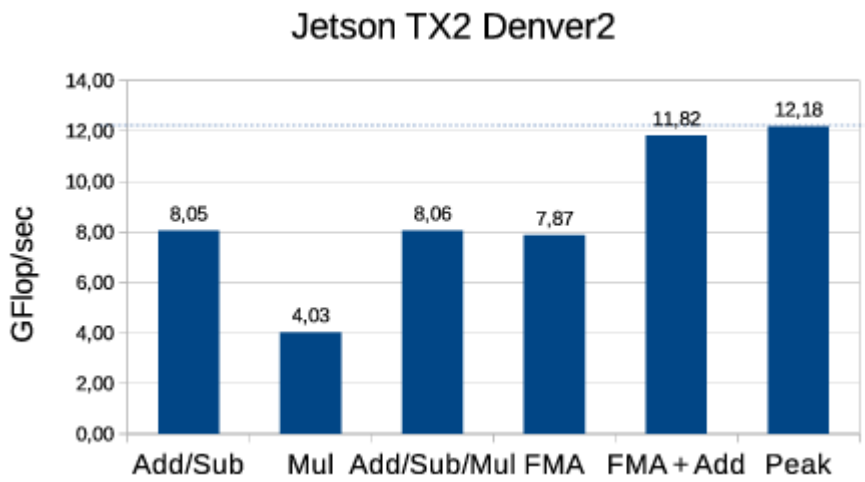
- 32 Cortex-A57 cores @ 2.1 GHz (64-bit)
- 1 MB L2 cache, 32 MB L3 cache
- 128 GB memory

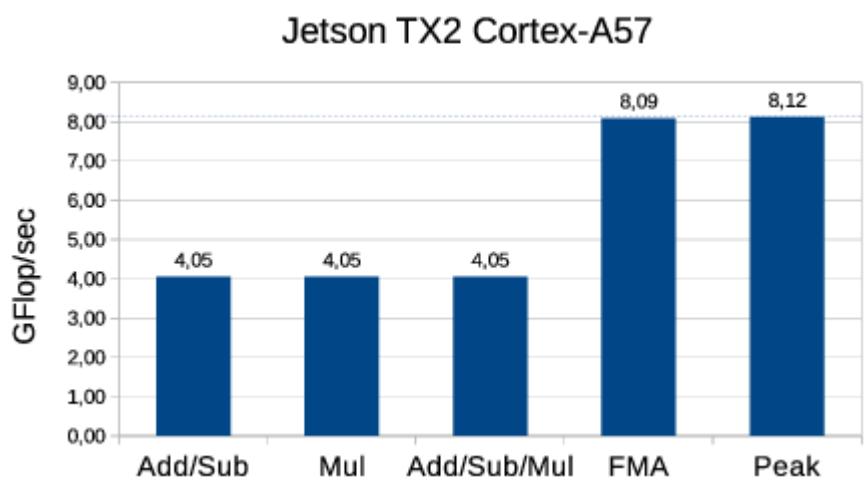
- Ubuntu 16.04 LTS
- gcc v5.3.1

Microbenchmark results

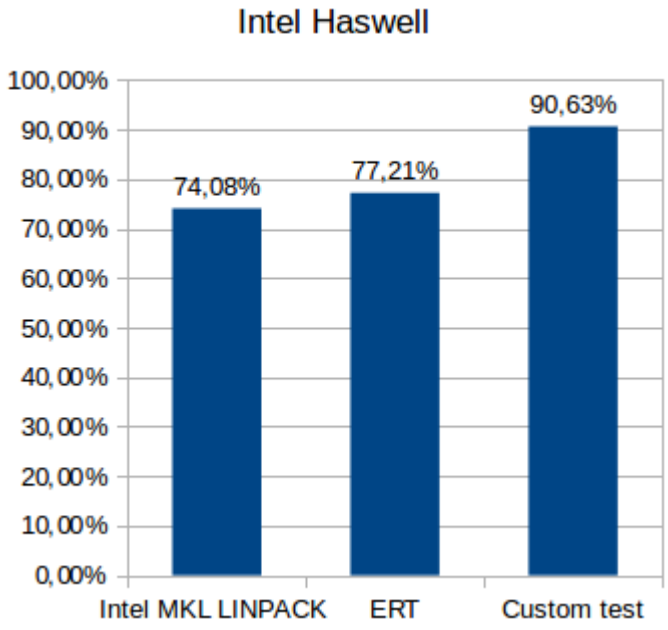


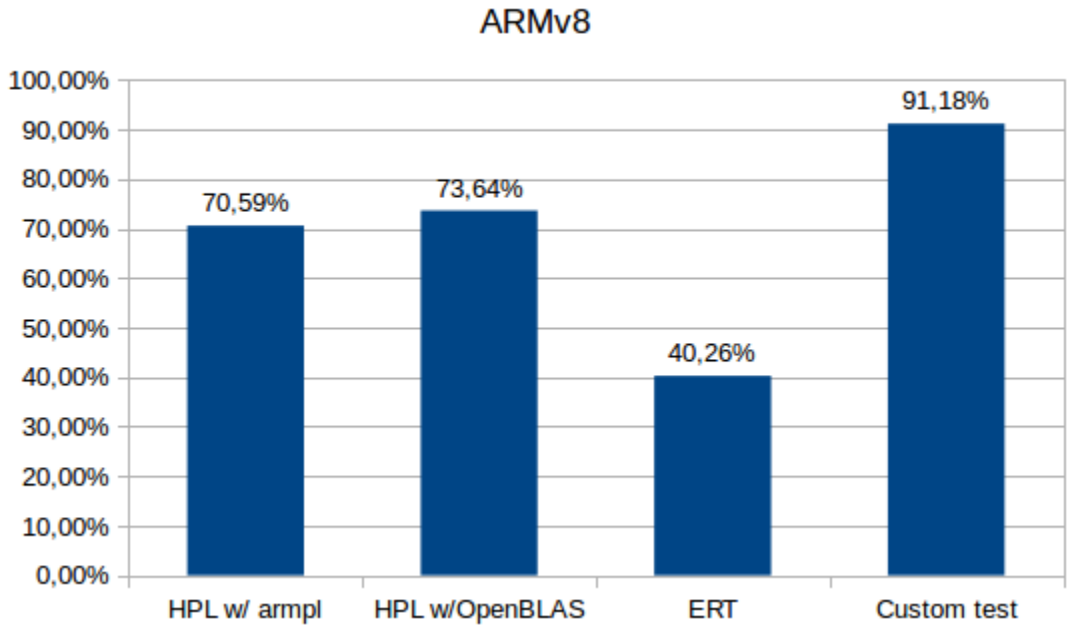






Benchmark comparison



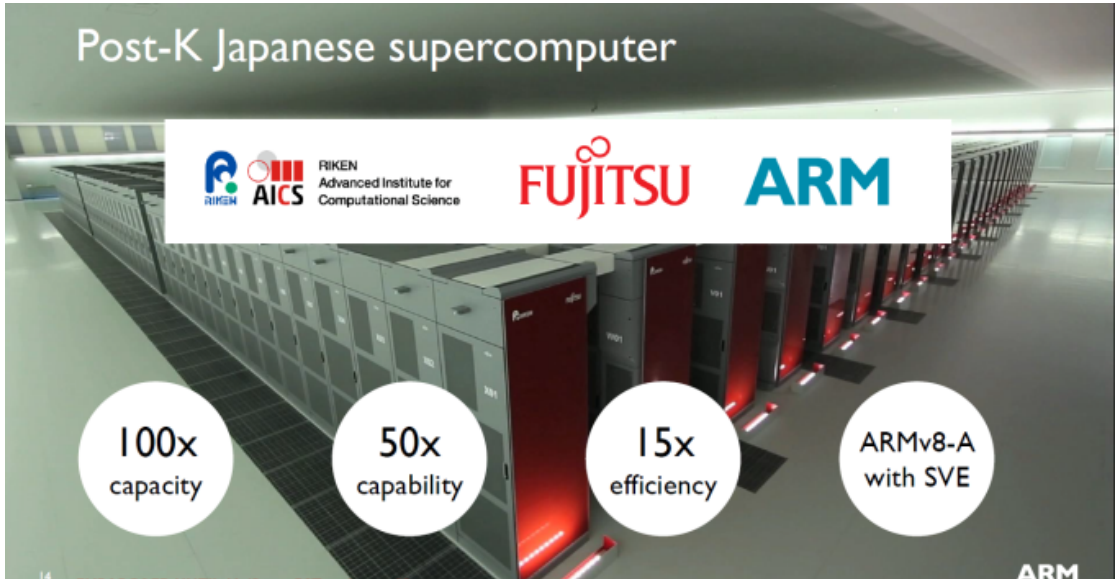


The Scalable Vector Extension (SVE)

Aarch64 extension which expand vector up to 2048 bits

- Expand fine-grain parallelism for HPC scientific workloads
- Reduce software deployment effort
- Vector-length agnostic programming model
- Enables vectorization of complex data structures with non-linear access patterns
- Permits vectorization of uncounted loops with data-dependent exits

Fujitsu Post-K



Be completed by 2022

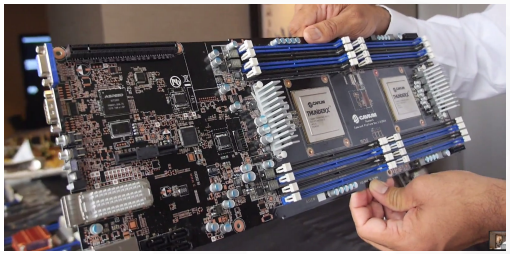
Budget $910m

Power consumption 30-40 MW

- Based on 64-bit ARMv8-A + SVE
- Exaflops supercomputer
- 6D mesh/torus interconnect “Tofu”
- 10nm FinFET

Conclusions

2014 2015 2016 2017



- The efficiency of ARM cores was compared with other CPUs (Intel x86_64 and legacy) for the classical MD algorithm example (LJ liquid in LAMMPS).
- The theoretical peak performance of the processors ARM Cortex and Nvidia Denver2 was calculated
- The highly optimized test code was created, that achieves the highest proportion of the calculated peak performance