Project N^{o}: 610456

D7.9 Review of Industry trends – Competitive analysis

February 28th, 2017

Abstract:

This deliverable provides an update on the competitive analysis performed by the EUROSERVER consortium. As of the time of writing, it is clear that it has not been easy to commercialise ARM-based micro-servers, and many large companies have been unable to bring a viable solution to market. The remaining providers that already have or intend to launch server-ready products are Qualcomm and Cavium. The decision by Fujitsu and RIKEN to base the Post-K supercomputer on the ARM architecture is a positive development.

Document Manager: John Thomson (Editor)	
Authors	AFFILIATION
John Thomson	OnApp Ltd. (ONAPP)
Paul Carpenter	Barcelona Supercomputing Center (BSC)
Manolis Katevenis, Iakovos Mavroidis	FORTH
Denis Dutoit	CEA
Internal Reviewers	
Per Stenstrom	CHALMERS
Isabelle Dor	CEA

Document Id N°:		**Version:**	0.7	**Date:**	28/2/2017

Filename:	Euroserver_D7.9_v0.7.docx

Confidentiality

This document is public and was produced by EUROSERVER contractors. Some of the graphics and information belongs to third-parties and this information is highlighted where appropriate. The commercial use of any information contained in this document may require a license from the proprietor of that information.

The EUROSERVER Consortium consists of the following partners:

Participant no.	Participant organisation names	short name	Country
1	Commissariat à l'énergie atomique et aux énergies alternatives	CEA	France
2	STMicroelectronics Grenoble 2 SAS	STGNB 2 SAS	France
3	STMicroelectronics Crolles 2 SAS	STM CROLLES	France
4	STMicroelectronics S.A	STMICROELECTRONICS	France
5	ARM Limited	ARM	United Kingdom
~~6~~	~~Eurotech SPA~~	~~EUROTECH~~	~~Italy~~
7	Technische Universitaet Dresden	TUD	Germany
8	Barcelona Supercomputing Center	BSC	Spain
9	Foundation for Research and Technology Hellas	FORTH	Greece
10	Chalmers Tekniska Hoegskola AB	CHALMERS	Sweden
11	ONAPP Limited	ONAPP LiMITED	Gibraltar
12	NEAT SRL	NEAT	Italy

Revision history

Version	Author	Notes
0.1	John Thomson	Preparation of the D7.9 document
0.2	Paul Carpenter	Corrections and modifications to text
0.5	John Thomson	Prepared for review
0.6	D. Dutoit, I. Mavroidis, M. Katevenis	European projects
0.7	John Thomson	Implementing review comments

This document is Public, and was produced under the EUROSERVER project (EC contract 610456).

Contents

1. Introduction

This deliverable provides an update on the competitive analysis performed by the EUROSERVER consortium for the DATE'16 paper publication.[1] Since that time, there have been many developments in the market, which have been captured in this document. The EUROSERVER project had originally intended to build a hardware prototype that integrated four chiplets, each with two quad-core Cortex-A53 clusters, which would have resulted in a 32-core, low-power micro-server solution. Given reasons that have been detailed in other project deliverables this has not been possible in the scope of the project. Reviewing other industrial solutions it is clear that in general it has not been easy to commercialise ARM-based micro-servers, and many large companies have been unable to bring a viable solution to market. Qualcomm and Cavium are the main players still active in the ARM-based server market. In addition, Fujitsu will supply ARM-based HPC chips for the Japanese Post-K supercomputer at RIKEN, a development that may "trickle down" in some ways to the server space.

Where ARM-based solutions have not generally entered the market, the need for low-power solutions has resulted in Intel having had time to develop power-efficient x86-based hardware. The other large winner of the delayed entry of power-efficient computing solutions has been the co-accelerator market, where NVidia and Intel have launched CUDA and Xeon Phi solutions, respectively.

[1] Manolis Marazakis , John Goodacre, Didier Fuin, Paul Carpenter, John Thomson, Emil Matus, Antimo Bruno, Per Stenstrom, Jerome Martin, Yves Durand and Isabelle Dor. EUROSERVER: Share-Anything Scale-Out Micro-Server Design. *Design, Automation and Test in Europe (DATE), 2016*.

2. Competitive Analysis

As part of the DATE'16 paper, the EUROSERVER consortium analysed the competitive ARM based solutions. The table with some corrections based on information not available at the time of the original publication is included below in Table 1. At the time the DATE'16 paper was prepared there were at least seven other vendors intending to build ARM-based solutions for the server market, other than the ARM Juno reference platform and the EUROSERVER proposed platform. As is discussed in the rest of this document, many of the vendors have either changed the target market for their devices or pulled out completely. Regarding power efficiency it is also important to note that the silicon fabrication technology has a large bearing on the power utilised by the System on Chip (SoC) solutions. Intel has complete control from design, through to fabrication plants and so have used fabrication technologies beyond the reach of most of these vendors to lower the power envelope

Table 1: ARM-based server solutions

Company	Product	Location	Silicon Tech	# Cores	Core Type	Chip power	Memory support	I/O support	Chip- to-chip support	Status
ARM	JUNO SoC	UK	TSMC 28HPM	6	Dual-core A72 + quad-core A53 + quad-core Mali-T624	N/A	2x channel DDR3L 32-bit 1600	USB 2.0, Other I/O in IOFPGA: 10/100 Eth, PCIe not functional (r1)	Yes, TLX400, total BW 0.7 Gbps	Available
EURO-SERVER	HW Pro-totype	Europe	FDSOI 28nm	32	A53	33W max.	4x HMC links: 4x60GB/s	FPGA I/F, 4x10 GB/s towards FPGA (I/O)	Yes, Unimem, RDMA	Proto Mar 2017
Freescale	LS1043 A	U.S.A	TSMC 28nm HPM	4	A53	8W	32-bit DDR3L/4	6x 10G Eth, PCIe Gen2, USB3, SATA3	No	Sampling Q1 2015
Freescale	LS2085 A	U.S.A	TSMC 28nm HPM	8	A57	30W TDP	2x 64-bit DDR4	8x 10G Eth, PCIe Gen3, USB3, SATA3	No	Q3 2015
Broadcom	Vulcan	U.S.A	TSMC 16nm FinFET	32	ARMv8 custom (Vulcan)		DDR4	PCIe	Yes, coherent	2H 2016 [2]
Cavium	Thunder X CN88xx	U.S.A	GF 28nm HKMG	24 to 48	ARMv8 custom	~125 W TDP	4x 72-bit DDR3/4	10/40/100G Eth, SATA3, PCIe Gen3	CCPI Dual Socket NUMA	Available
Cavium	Thunder X2	U.S.A	14nm FinFET	54	ARMv8 custom	130W ?	6x 72-bit DDR4	10/25/40/50/100 GbE, SATA3, PCIe Gen3	CCPI Dual Socket NUMA	2H 2017
Applied Micro	X-Gene1	U.S.A	TSMC 40nm	8	ARMv8 custom	40W	DDR3	4x 10G Eth, 6x PCIe, SATA	No	Available
Applied Micro	X-Gene2	U.S.A	TSMC 28nm	16	ARMv8 custom	25W?	4x DDR3L/4	PCIe Gen3, 1/10Gbe, SATA3	RoCE	Sampling H1 2016
Applied Micro	X-Gene3	U.S.A	TSMC 16nm	32	ARMv8 custom		8x DDR4	42x PCI G3		Sampling

[2] As discussed later, this has been discontinued

Company	Product	Location	Silicon Tech	# Cores	Core Type	Chip power	Memory support	I/O support	Chip- to-chip support	Status
			FinFET							1Q 2017 [3]
AMD	Opteron A1100	U.S.A	GF 28nm HKMG	8	A57	<45W	DDR3/DDR4	2x 10G Eth, 8x PCIe Gen3, 8x SATA3	CCN- 504?	March 2015[4]
HiSilicon	Phosphor- V660 Hip05	China	TSMC 16nm FinFET	2x 32	A57		DDR3	3D integration with IO chip	No	Engineering samples
EZCHIP (Tilera)	TILE-Mx	Israel (U.S.A)	28nm HPM	100	A53	70W	DDR4	1Gbe to 100GbE, PCIe Gen3	Y(Interlaken) on-chip coherency	Sample H2 2016[5]

Changes in the market

There have been many changes since March 2016 when the EUROSERVER paper was published in DATE 2016 in the spectrum of ARM-based servers and throughout energy-efficient computation. Many mergers and acquisitions have occurred as well as changes in roadmaps. The overall trend has been one of consolidation with few new product lines being developed and most ARM-based solutions failing to make a mark on Intel's dominant position in the data center. AMD, which appeared to be the most interesting challenger to Intel's position, stated in 2013 and 2014 when the EUROSERVER project was starting, that it was the right time for ARM to enter the data center. Given that AMD and many others tried and have since moved back to x86, unless there is a significant change in the market, it is unlikely that there will be any large change to Intel's position.

ARM Holdings – now Softbank

ARM was purchased by Softbank Group in September 2016[6] for 31$Bn (US) / 24.3£bn. Following the acquisition, it is expected that ARM will continue its licensing model and continue to operate in a similar manner as before its acquisition.

Freescale – now NXP – soon Qualcomm

Freescale that produces two ARM-based chips that were listed in the competitive analysis was purchased by NXP in December 2015[7].

In turn, NXP is in the process of being acquired by Qualcomm, as of October 2016 for 47$Bn (USD)[8] NXP has focused more on its PowerPC-based, QorIQ product line for its networking and telecommunications products. The QorIQ LS ARM-based products are more focused on embedded solutions and network controllers and it is hard to find products that include them. The LS2088A, a pre-production, 8-core ARM A72 SoC with up to 8x 10GbE and 8x 1GbE, is mentioned in more recent articles but it is also generally hard to find public information on it[9]

[3] Unlikely to reach the market, see APM section.
[4] AMD have changed strategy, so this model is being discontinued.
[5] Tilera is now part of Mellanox and this solution is being converted for use as a network product
[6] http://www.softbank.jp/en/corp/news/press/sb/2016/20160905_01/
[7] http://investors.nxp.com/phoenix.zhtml?c=209114&p=irol-newsArticle&ID=2120581
[8] https://www.qualcomm.com/news/releases/2016/10/27/qualcomm-acquire-nxp
[9] http://linuxgizmos.com/high-end-boards-run-linux-on-64-bit-arm-qoriq-socs/

Applied Micro (APM) now MACOM

MACOM announced that it would acquire APM for 770M$ (USD) in November 2016[10].

X-Gene 1 was promising and had a public offering that was the first server-ready ARM platform. X-Gene 2 however never really materialised and X-Gene 3 was where the promising upgrades looked to be. MACOM intends to sell the APM Compute business, which includes its X-Gene server products. Of particular interest to EUROSERVER was the X-Tend Interconnect technology, which was a proprietary technology for connecting multiple ARM sockets together.

The X-Gene 3 specifications were never formally announced but there were performance targets that were published by the LinleyGroup[11]. The X-Gene 3 was suggested to have 32 ARMv8 cores, 550 SPECint_rate, 32MB L3 cache. 42 PCI Express 3.0 lanes, 8 DDR4 channels. X-Gene 3 was taped out in November 2016 and was expected to be sampling in March 2017, but the situation will depend on its new owner.[12] The X-Gene 3 never entered full production but it was suggested that the process would be TSMC - 16nm FinFET+. With a clock rate of 3GHz this would have led to an estimated thermal design point (TDP) of between 110W and 125W.

The X-Gene 1 had garnered interest from some large Internet firms as an alternative to Intel processors. X-Gene 1 had been used by PayPal and was used in the HP Moonshot platform[13].

Indeed, the EUROSERVER consortium has used X-Gene 1, see Figure 1, for part of its development and analysis work through the Gigabyte MP30-AR0, which has 8x U/RDIMM DDR3 (Un/Registered, dual inline memory module - RAM) slots. 2x 10GbE (network interface). 2x GbE. 4x SATA III (hard disk drive) and supports up to 128GB RAM[14].

[10] https://www.apm.com/news/macom-announces-definitive-agreement-to-acquire-appliedmicro/
[11] https://www.linleygroup.com/uploads/x-gene-3-white-paper-final.pdf
[12] ARM on the Road to HPC Workshop, January 16—17 2017, Barcelona.
[13] http://www.theregister.co.uk/2015/11/16/appliedmicro_x_gene_3/
[14] https://www.servethehome.com/gigabyte-arm-server-based-on-appliedmicro-x-gene/

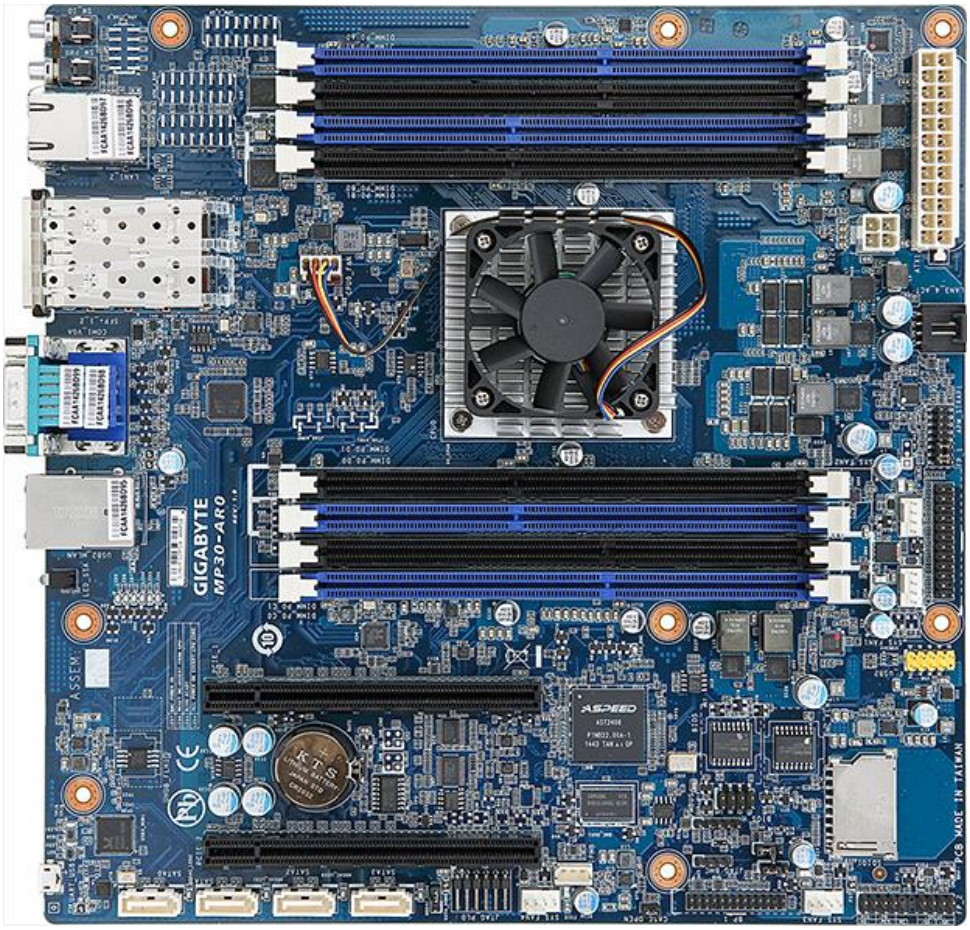


Figure 1: MP30-AR0 X-Gene1 from Gigabyte, copyright Gigabyte, via TechReport

Broadcom now Avago

Broadcom has all but confirmed that it has discontinued its Vulcan project after its acquisition by Avago in December 2016[15][16].

AMD ARM Opteron discontinued

AMD dropped its ARM efforts including the Opteron A1100, see Figure 2, which was delayed by many years in favour of the AMD Zen x86-based platform planned for release in 2H 2017[17].

The Opteron A1100 had been released as a product from SoftIron that had ARM Cortex-A57, 2x RDIMM allowing for 64GB of memory with 2x SATA ports.

[15] https://www.fool.com/investing/2016/12/12/why-broadcom-ltd-abandoned-this-multi-billion-doll.aspx
[16] http://www.theregister.co.uk/2016/12/07/broadcom_arm_processor_vulcan/
[17] http://www.pcworld.com/article/3106852/amd-turns-back-to-x86-for-server-reboot-as-it-downgrades-arm.html

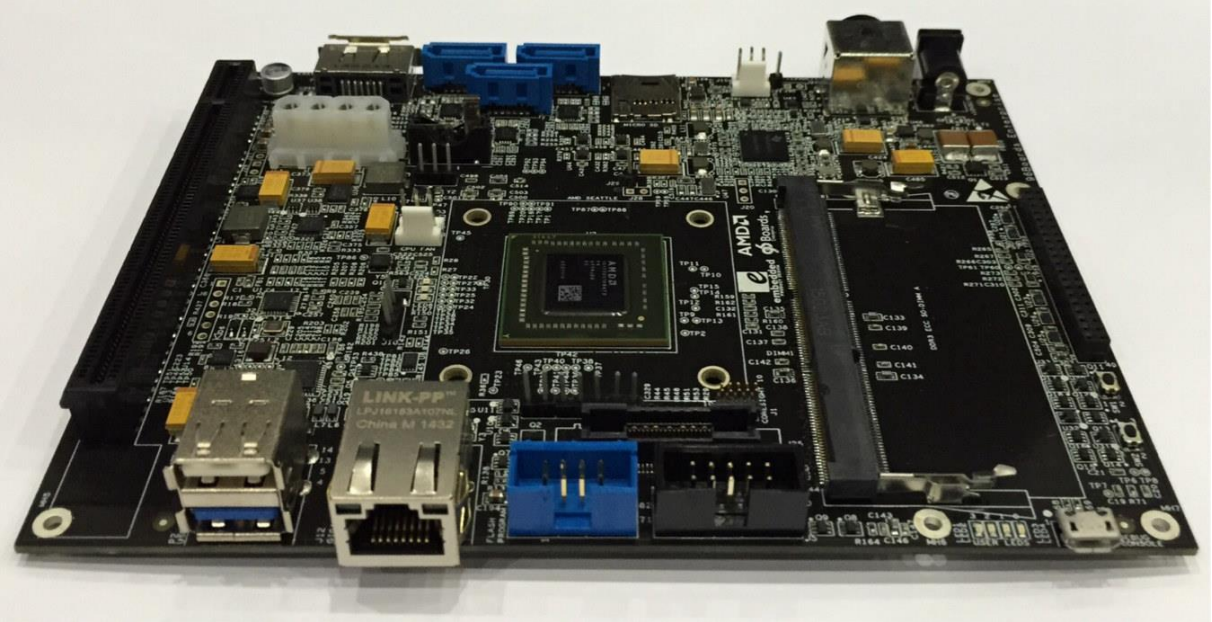


Figure 2: AMD A1100 based HuskyBoard from 96boards, via CNX Software

"AMD believed the power-efficient ARM chips would ultimately replace x86 in servers and have a 20 percent market share by 2017, but that hasn't happened" – source pcworld.com[18].

["ARM servers hit the market for real in 2014, that's the try year," said Andrew Feldman, corporate vice president and general manager of the server business unit at AMD.

…By 2016 or 2017, ARM CPUs will have 20 percent of the server market, Feldman said. Right now the server market is dominated by Intel's x86 chips, but Feldman said large data centers will start porting software from x86 to ARM in 2015.] – source pcworld.idg.com.au[19].

HiSilicon - Huawei

HiSilicon released the D02 board, but it is very hard to get ahold of HiSilicon and its parent company Huawei to understand the progress of the D05.

The D02 board uses the PV660 Hip05 processor, which has a 16-core Cortex-A57, with DDR3 x 2 per socket. The D02[20] is seen in Figure 3. Interestingly, the board supported a 2-socket configuration, leading to a 32-core board, but two-socket solutions have not been found[21].

[18] http://www.pcworld.com/article/3158919/servers/amd-talks-tough-as-it-drums-up-support-for-32-core-zen-server-chip.html
[19] http://www.pcworld.idg.com.au/article/465171/amd_reboots_server_strategy_first_arm_chips/
[20] http://open-estuary.org/d02-2/
[21] http://www.cnx-software.com/2015/02/16/hisilicon-d02-armv8-server-board/#ixzz3RwO4N93v

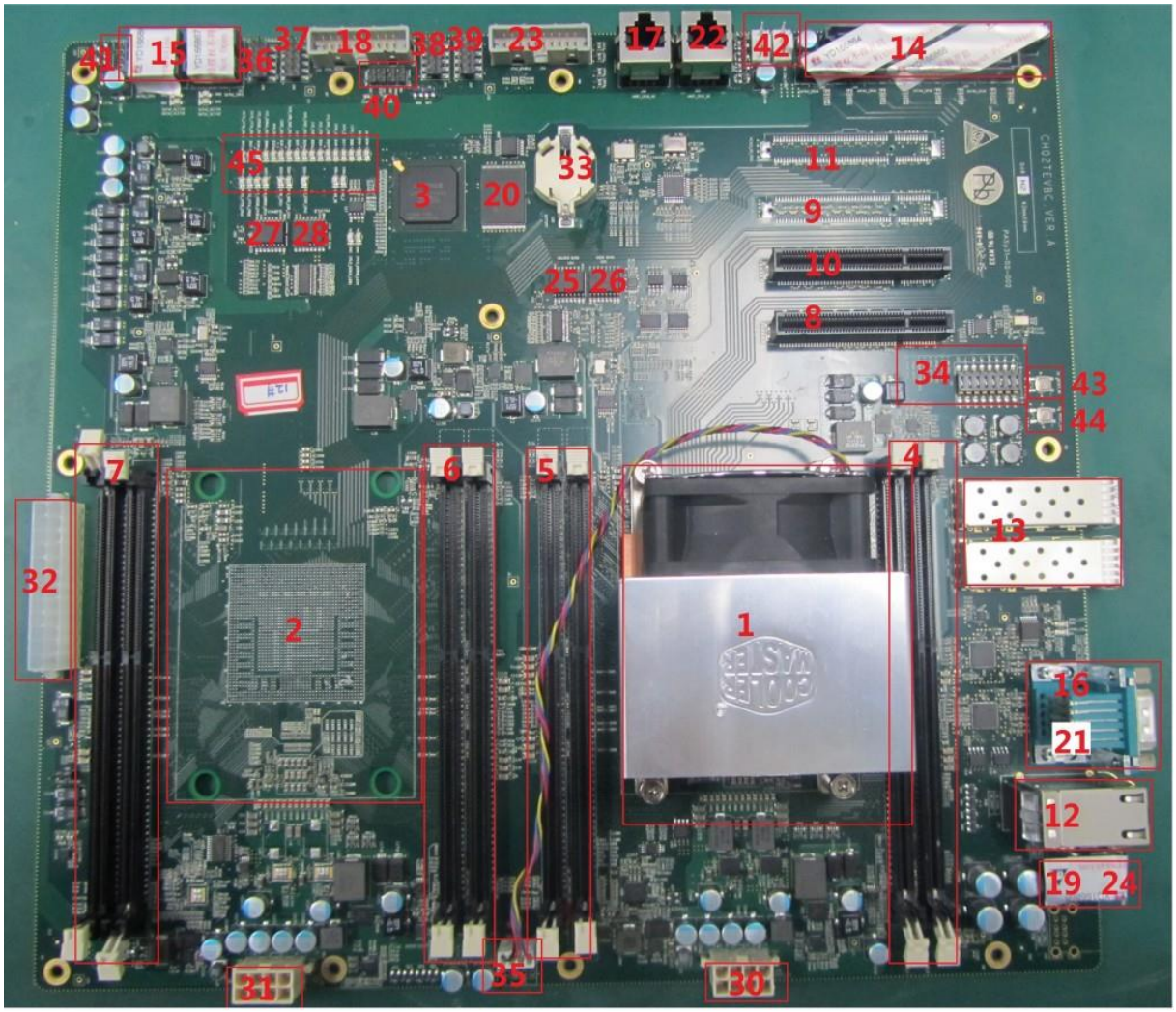


Figure 3: HiSilicon D02 board. Copyright belongs to HiSilicon

Tilera – now Mellanox

Tilera was acquired by Mellanox[22] in Feb 2016 and the Tilera-Mx that was mentioned in Table 1 has since been discontinued.

["As soon as the teams started talking in October last year after the Mellanox deal was announced, it was decided pretty quickly to put the TileMx-100 project on hold because Mellanox has other designs for the technology," Bob Doud, the long-time marketing director at Tilera and then EZchip, tells *The Next Platform*.] – source nextplatform.com[23].

Instead Mellanox is now taking a similar approach to other developers that forayed into ARM designs, which is to work on networking products. The TileMx-100, which was mentioned, has now been merged into the new product line "BlueField" that will have 16 Cortex-A72s combined with high-speed network interconnects to provide high-speed networking devices[24].

[22] http://www.mellanox.com/page/press_release_item?id=1681
[23] https://www.nextplatform.com/2016/06/08/putting-brains-network-frees-compute/
[24] http://www.mellanox.com/related-docs/npu-multicore-processors/PB_Bluefield_SoC.pdf

Altera purchased by Intel

Intel purchased Altera in December 2015[25].

The move to heterogeneous computing as seen through Intel Xeon Phi and Nvidia CUDA has led to resurged interest in FPGAs and ASICs. Users of web-scale workloads seem to be interested in the idea of using programmable co-processors that are energy-efficient for performing specific operations at a reduced power envelope. As part of the strategy to support co-processors, Intel acquired the FPGA manufacturer Altera, which is the second largest FPGA manufacturer behind Xilinx.

Amazon and Google's power efficient compute hardware

The acquisition of Altera by Intel has been echoed by other cloud-suppliers such as Amazon Web Services (AWS)[26] who have released FPGA cloud options to their product offerings. Google announced Tensor Processing Unit, see Figure 4, which it claims is more power efficient than FPGAs and GPUs[27].

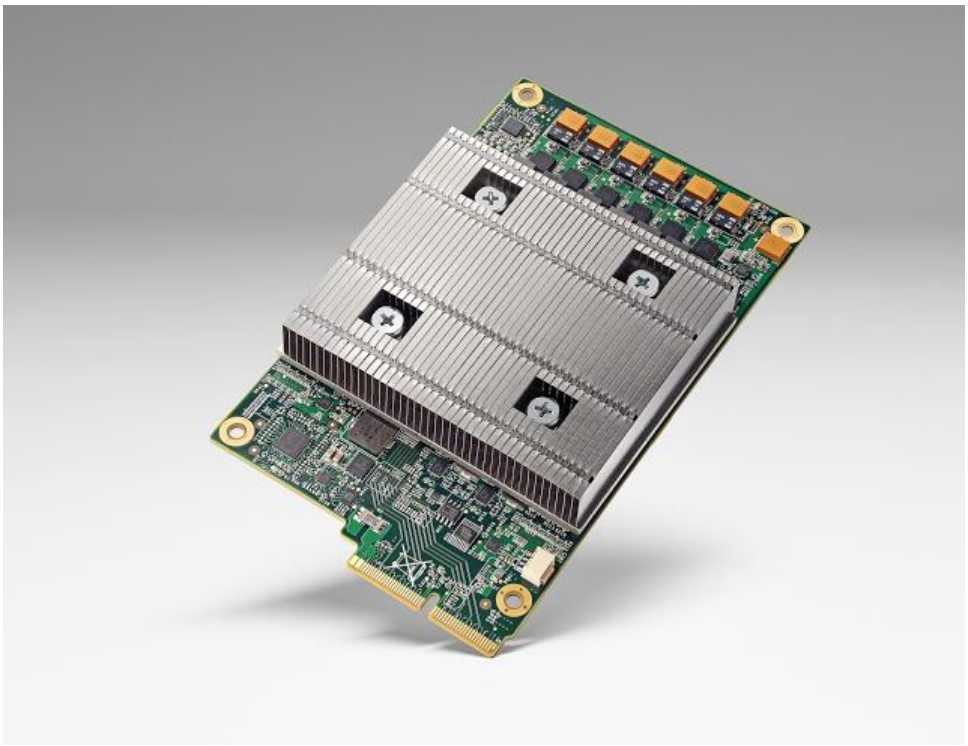


Figure 4: Google Tensor Processing Unit - copyright Google

The remaining challengers

Qualcomm and Cavium are the two big providers that are left in the ARM ecosystem that have or intend to launch server ready products.

[25] http://fortune.com/2015/12/28/intel-completes-altera-acquisition/
[26] https://aws.amazon.com/ec2/instance-types/f1/
[27] http://www.techradar.com/news/computing-components/processors/google-s-tensor-processing-unit-explained-this-is-what-the-future-of-computing-looks-like-1326915

Qualcomm

Qualcomm produced an offshoot of its main business, Qualcomm datacenter technologies that aims to release products by Q3 2017. In particular its focus is to release an ARM based solution[28].

The product range has not formally been announced but it will likely be similar to the Qualcomm Centriq 2400, 48-core solution, see Figure 5, which was announced as the world's first 10nm server solution – source Qualcomm[29].

Figure 5: Qualcomm Centriq 2400, via servethehome

Cavium

Cavium released ThunderX, see Figure 6, which other than APM's X-Gene 1 is the only ARM-based server class hardware that got into production. They offer multiple SKUs that allow some of the 24 to 48 cores in its design to be specialized for networking, security or compute based applications. Although the SoC was designed for 2.5GHz the heat produced at that level of operation was too high and so it is limited to 2GHz and has a TDP of around 130W. The cores are simplified cores with short pipelines that allow for good performance on certain types of workloads, but the performance did not compare well against Intel Xeon platforms for compute intensive workloads.

[28] http://www.theregister.co.uk/2016/11/15/packet_dotnet_arm_cloud/
[29] https://www.qualcomm.com/news/onq/2016/12/07/meet-qualcomm-centriq-2400-worlds-first-10-nanometer-server-processor

Figure 6: Cavium ThunderX via servethehome.com

Cavium has since announced the ThunderX2 solution[30]. The Cavium ThunderX2 is designed to run at 2.4 to 2.8GHz with up to 3GHz turbo. The ThunderX2 is designed to have; 54 cores per socket, 40K I Cache, 64K D Cache, 32MB LLC, 6x 62 bit DDR4 controllers per socket, Dual DIMM per controller and 12 DIMMs per socket allowing for 3TB memory in dual socket configuration. The ThunderX2 is also designed to have fast networking speeds and has multiple ports for 10/25/40/50/100GbE. It can also support multiple SATAv3 and PCI Gen 3 lanes[31].

It is likely that the Cavium ThunderX2 will be produced using the 14nm FinFET process[32].

Cavium also have other ARM-based chip designs for different types of applications including the Octeon-TX range for networking and security based devices[33] with high-speed networking support.

Another of Cavium's offering is the Nitrox V that uses 288 RISC cores for 100Gbps packet processing[34], which looks like it could be superseded with their ARM based solutions.

Given that Cavium has released a many-core ARM solution to the market there have been multiple vendors and solutions providers that have taken the board and have produced an offering. For instance Packet is a developer-focused cloud service provider that allows developers to build against ARM solutions and it offers ARM Cavium ThunderX solutions[35].

[30] http://cavium.com/newsevents-Cavium-Announces-ThunderX2.html
[31] http://www.cnx-software.com/2016/06/01/cavium-introduces-54-cores-64-bit-armv8-thunderx2-soc-for-servers-with-100gbe-sata-3-pcie-gen3-interfaces/
[32] https://www.top500.org/news/cavium-makes-noise-with-new-thunderx2-arm-chip/
[33] http://www.theregister.co.uk/2016/05/02/cavium_octeon_tx/
[34] http://cavium.com/processor_security_NITROX-V.html
[35] http://www.theregister.co.uk/2016/11/15/packet_dotnet_arm_cloud/

Other production ready solutions include the Gigabyte commercial offering – July 19 2016[36].

As with all platforms that have entered the market there were some initial bugs with the Cavium ThunderX but as more development is built around the solution these can be rectified[37].

Other hardware improvements over the course of EUROSERVER

Other than the CPU, there have been various other advances that have contributed towards energy efficiency in server platforms.

Power supply units

PSUs are already highly energy efficient, reaching a minimum of 80% efficiency levels, with many vendors offering 90 and 95%+ energy efficient solutions.

Memory Modules

Hybrid Memory Cubes (HMC) are a stacked (3D), memory type that offer multiple layers of DRAM memory over a logic chip. There are claims that the energy efficiency will be improved by having the energy cost amortised over several DRAM layers.

NVDIMM – non-volatile DIMMs are a solution being offered by vendors such as Micron that allow for storage devices that are almost the speed of standard DRAM modules but have storage backing allowing for persistence such as NAND based flash[38].

Low power DRAM modules such as DDR3L have proved popular in servers, running at lower voltage and hence lower power levels.

One area that EUROSERVER led in this domain has been in the creation of the spin-off, ZeroPoint Technology[39], from Chalmers. An area that was identified as being a problem for the data center market was that memory in the data center is expensive but it is much quicker to access than storage devices. A memory compression system was developed by ZeroPoint that compresses and uncompresses the memory in-line, at the sacrifice of some processing and latency time for 3x higher memory storage. This of course depends on the workload and how much memory can be compressed but for randomised workloads, as found in cloud deployments, they experienced a large compression factor.

Hard disk drives

The general move in the data center to move away from rotating and mechanical drives. The issue with magnetic, rotating disk drives has been that they have high-failure rates, low-tolerance to vibrations, take up a lot of physical space and consume a lot of energy. Mechanical drives are now being phased out in exchange with Solid State Disk (SSD) drives. The main reason for keeping SATA and SAS drives is

[36] http://cavium.com/newsevents-Cavium-and-GIGABYTE-Announce-Official-Release-of-Production-Ready-ThunderX-based-Servers.html
[37] http://www.slideshare.net/Semihalf/freebsd-on-cavium-thunderx-system-on-a-chip
[38] http://www.hotchips.org/wp-content/uploads/hc_archives/hc28/HC28.21-Tutorial-Epub/HC28.21.1-Next-Gen-Memory-Epub/HC28.21.110-Memory-Horizon-Pawlowski-micron-v1-t1-2.pdf
[39] http://zptcorp.com/

that they offer more storage per dollar, but as the technology for SSDs has matured the price of SSD drives has continued to decrease. SSD drives require less power than disk drives. The interconnect used on disk drives is also moving away from the standard that was used for mechanical disk drives, SAS and SATA and instead hardware manufacturers are increasingly using NVMe to connect the rest of the system to the disk drives.

"Given the data center's transition away from hard disks, we expect most customers to rely on NVM Express, which uses PCIe to connect directly to SSDs." – source LinleyGroup.com[40].

Networking devices and infrastructure

Although energy efficiency levels of networking equipment has not been massively improved or marketed, rating systems such as the PUE have put more importance on calculating the underlying energy demands in the infrastructure. Industry alliance groups have also looked into replacing Top-of-rack solutions with more generic hardware, such as carried out in the OpenCompute project, which was spearheaded by Facebook. Power efficiency has not been the number one motivator (cost is), but many solutions are being proposed that look to increase the energy efficiency at the same time as increasing performance. In the short term though, given the emphasis on performance and cost, the solutions that are coming out are likely to be less power efficient until future iterations.

Intel

Where ARM-based solutions have faltered, Intel has taken large steps to producing power-efficient data center, compute solutions. As was previously mentioned, Intel purchased Altera to increase its ability to work on specialized workloads. This complements the approach that they took with what was formerly known as Knights Landing/Crossing, etc., the Xeon Phi. By using low-power Atom-based processors on the Phi, Intel allowed developers to produce code that did not differ as greatly as with CUDA applications, for performing computationally expensive workloads. The Xeon Phi, see Figure 7, is very power efficient, compared to the generic Xeon processors used when programmed correctly.

[40] https://www.linleygroup.com/uploads/x-gene-3-white-paper-final.pdf

Figure 7: Intel Xeon Phi 7290 - copyright Intel

Intel has also moved into the space that the ARM A53 and A57 solutions were looking into by offering power-efficient, SoC solutions as the Intel Xeon-D range. The "XeonD is basically Xeon E3 with integrated [Platform Controller Hub] PCH" – source Anandtech.com[41] "[which] makes it hard to compete in performance / watt." These Broadwell process based chips (14nm) run at low TDP power points and are aimed at Cloud workloads. The PCH device was designed to replace the separate Northbridge and Southbridge chips (collectively known as Intel Hub Architecture) designed by Intel to communicate to peripheral devices on a motherboard. By exposing the PCI Express and other lanes from the integrated controllers, further power-saving can be achieved with respect to having separate chips.

The Intel Xeon-D 1581, see Figure 8, for instance has; 16 cores, 32 threads @1.8GHz, 24MB cache, with a 65W TDP and max 128GB DDR4 RAM[42].

Since the Xeon-D was released in early 2016 there have been subsequent improvements and iterations such as the Intel Xeon-D 1587 that runs at 1.7-2.3GHz and has a price point of 1754$, which is fairly competitive to the ARM solutions, once the cost of training etc. is taken into consideration. The Xeon-D has been used in EUROSERVER as the competitive solution to compare against.

[41] http://www.anandtech.com/show/8357/exploring-the-low-end-and-micro-server-platforms/18
[42] http://ark.intel.com/products/93357/Intel-Xeon-Processor-D-1581-24M-Cache-1_80-GHz

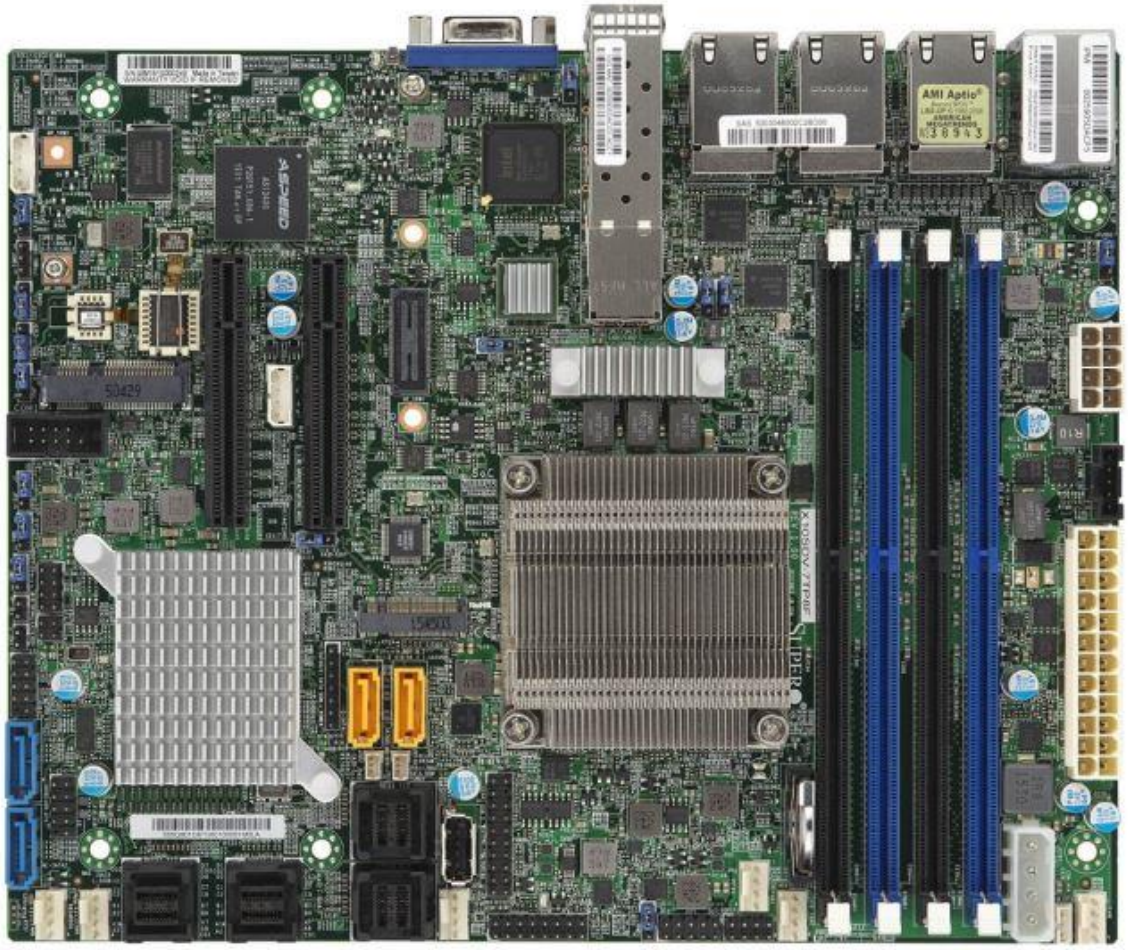


Figure 8: Xeon-D 1581, on a SuperMicro Board, via Anandtech

Intel has also used its low-power expertise into producing other solutions, which will then be incorporated back into future Xeon Phi and other architectures.

The Intel Avoton, which focused on Microservers and storage devices and the Intel Rangeley , which focused on Networks and Comms infrastructure were based on a 22nm design and offered low-power solutions for the entry-level. They both have PCI Express 2.0, DDR3 Memory controllers and 4x GbE.

Intel also has a Pentium N3710 that was released in Q1 2016, which again has a low TDP point but is not aimed at the server market.

Intel has though taken Cavium as a serious contender and has performed multiple benchmarks of their platforms against Cavium ThunderX solutions. “The Intel Xeon E5-2600 V2 and V3 lines can support up to 768GB of RAM per CPU, or about 50% more per chip than the Cavium parts. The older E5-2600 V1 chips were limited to 384GB per CPU. Currently the Xeon D-1500 series is limited to 128GB and single socket configurations, however that is the chip Intel launched to keep most ARM vendors at bay” – servethehome.com[43]. The ThunderX2, which is scheduled for release in the second half of 2017 is

[43] https://www.servethehome.com/exclusive-first-cavium-thunderx-dual-48-core-96-core-total-arm-benchmarks/

expected to be able to address 3TB of memory in a 2-socket configuration. It is mainly the case though that Intel is producing just enough hardware to stop ARM solutions being competitive.

The Xeon Phi with 4 threads per core[44] offers very high power efficiency that competes with the Nvidia P100, K20, K40 and K80 solutions.

Comparison of Intel, ARM and NVidia

Although it is hard to compare the power efficiency of different architectures, we have outlined the leading power-efficient platforms from each of the main solutions providers in the DATE'16 paper. This is now updated to include further hardware platforms and designs that were not available at the time of writing that paper. Please see Table 2 for more details. As a reminder the EUROSERVER platform was designed to have 4 chiplets containing 2x quad core A53s and run at a power envelope of under 33W.

[44] http://insidehpc.com/2017/01/intel-xeon-phi-processor-programming-nutshell/

Table 2: ARM, Intel and Nvidia power efficient compute solutions

Processor	Frequency Cores		Threads Fab.		Cost	TDP	Watts per Thread	Status
A15	1.9GHz	1-4	1-4	28nm	<$15	1.5- 3.7W	0.38- 0.93W	2H 2012
A53	1.6GHz	4-8	4-8	28nm	$17- 25	1.7- 6.2W	0.43- 1.55W	1H 2015
A57	2.1GHz	4-8	4-8	20nm	$30- 60	2.0- 7.1W	0.5- 1.78W	1H 2015
A72	2.3GHz	4-8	4-8	16nm	N/A	1.5- 5.8W	0.4- 1.45W	4Q 2015
X- Gene1	2.4GHz	8	8	40nm	$316	45W TDP	5.63W	2H 2015
X- Gene2	2.8GHz	8	8	28nm	N/A	45W TDP	5.63W	2H 2016
Pentium N3700	1.6GHz	4	4	14nm	$161	6W	1.5W	Q1 2015
Atom C2758	2.4GHz	8	8	22nm	$208	20W	2.5W	Q3 2013
Xeon D-1537	1.7GHz	8	16	14nm	$571	35W	2.19W	Q4 2015
Xeon D-1540	2GHz	8	16	14nm	$581	45W TDP	2.81W	Q1 2015
Xeon D-1559	1.5GHz	12	24	14nm	$883	45W TDP	1.88W	Q2 2016
Xeon D-1571	1.3GHz	16	32	14nm	$1222	45W TDP	1.4W	Q2 2016
Xeon E-1240Lv3	2GHz	4	8	22nm	$278	25W TDP	3.13W	Q2 2014
Xeon E3-1240Lv5	2.1GHz	4	8	14nm	$278	25W TDP	3.13W	Q4 2015
ThunderX	2.5GHz	24- 48	24- 48	28nm	N/A	80W TDP	1.67W	1H 2016
Knights Corner (Phi)	1.1- 1.238 GHz	57- 61	244	22nm	$1695 to $4235	300W TDP	1.23W	Q2 2013
Knights Landing (Phi)	1.5GHz	72	288	14nm	$6401	260W TDP	0.9W	2H 2016
Nvidia Tesla K20X	0.732GHz	2688	2688	28nm	$3799	235W	0.087W	2H 2012
Nvidia Tesla K40	0.745GHz	2880	2880	28nm	$5499	235W	0.082W	2H 2013
Nvidia Tesla K80	0.56GHz	4992	4992	28nm	$5000	300w	0.060W	2H 2014
Nvidia Tesla P100	1.3GHz (shader)	3584	3584	16nm	$9428	250W	0.070W	Q4 2016

The ARM Cortex-A73 announced in June 2016[45] has not been included in the table as it has been primarily designed for mobile phone solutions. It is expected though, given the claimed high-power efficiency, as seen in Figure 9 and Figure 10, that server designs based on the Cortex-A73 should be forthcoming.

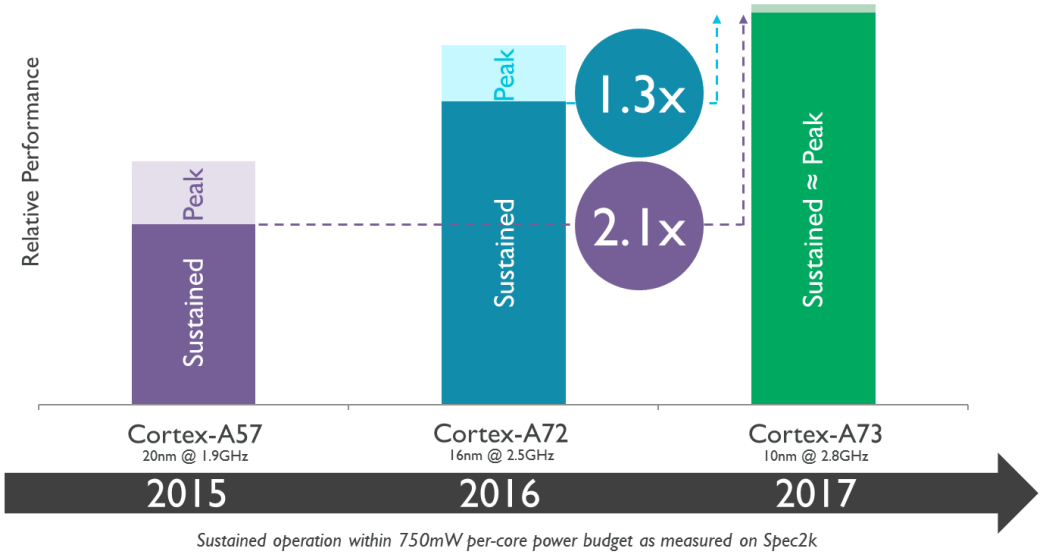


Figure 9: ARM A57, A72 and A73 relative performance comparison

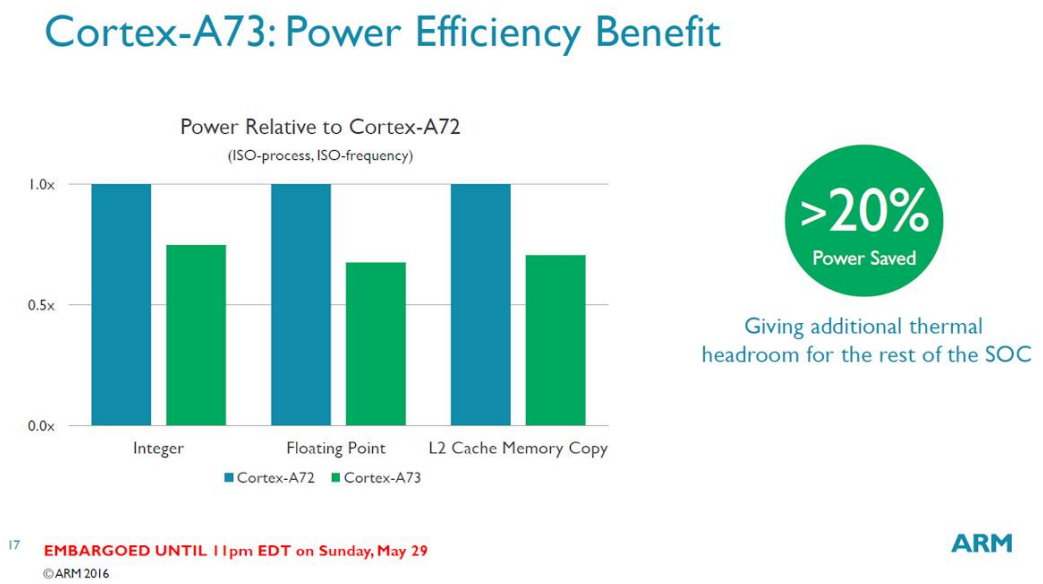


Figure 10: Relative performance of ARM Cortex-A73 vs A72 – source pcper.com[46]

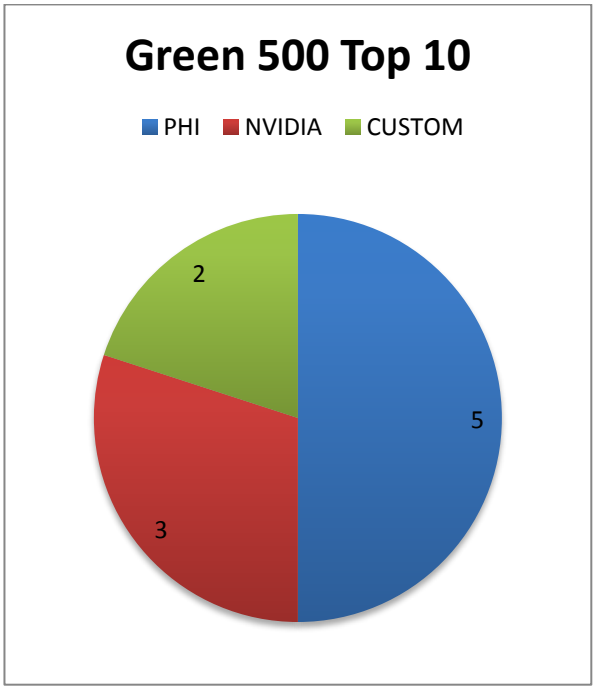
[45] https://www.gizmotimes.com/news/huawei-kirin-960-unveiled-cortex-a73-mali-g71/16766
[46] https://www.pcper.com/reviews/General-Tech/ARM-Tech-Day-2016-Introducing-Cortex-A73-Mali-G71-and-CCI-550

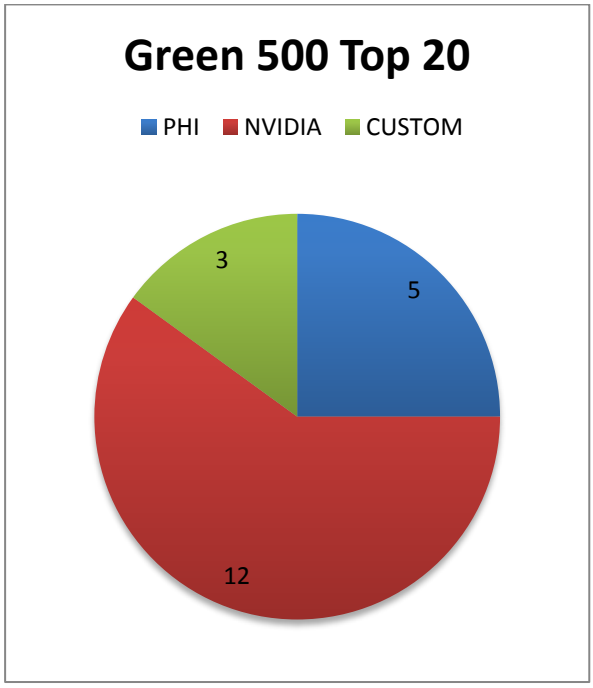
Green500

The Green500 list is a list produced by Top500 that measures the most power-efficient supercomputers. They released their most up-to-date list[47] (at the time of writing) in November 2016.

From this list the top two energy-efficient systems are powered by the Nvidia P100 Tesla. Five of the remaining top 10 most energy efficient supercomputer platforms have Xeon Phi 72XX accelerators. Another uses Nvidia K80 accelerators and the other two (in positions 3 and 4) use custom Many-Core solutions.

As can be seen in the pie-charts below the top 20 and 50 systems predominately have Nvidia-based co-accelerator systems and then the subsequent 50 solutions have a large number of custom interconnect solutions.





[47] https://www.top500.org/green500/list/2016/11/

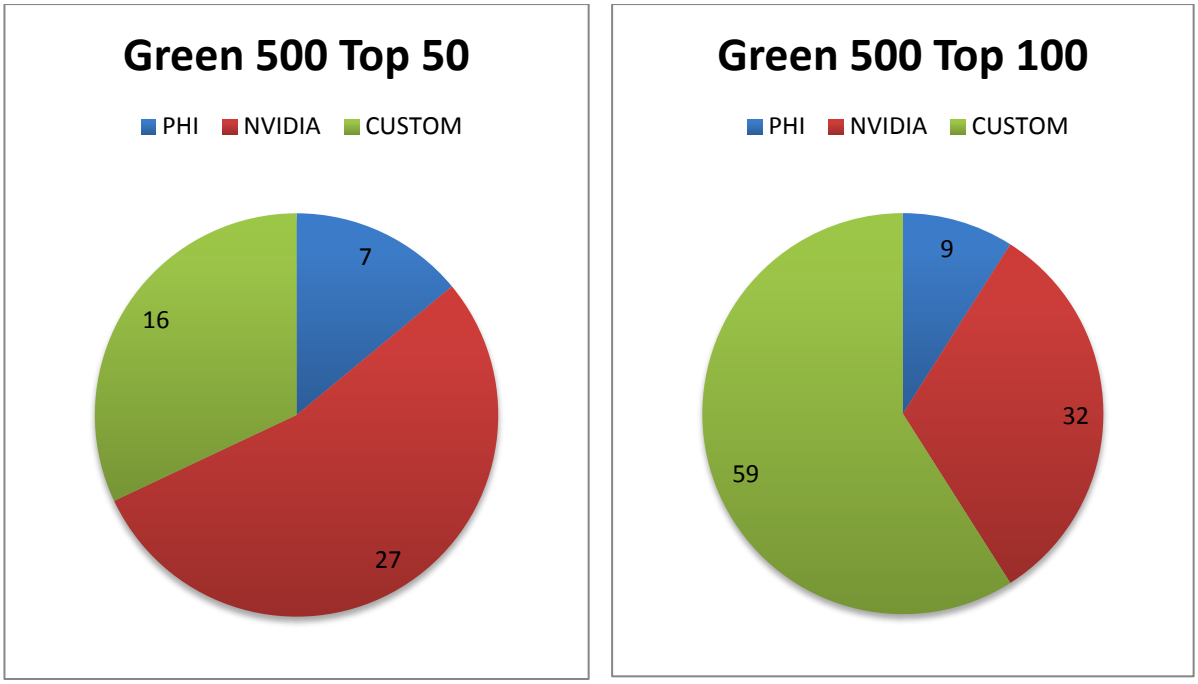


Figure 11: Charts showing the proportion of Intel Phi, NVIDIA and other platforms that make up Top 10, 20, 50 and 100 of the Green 500 list

European Projects investigating micro-servers / energy efficient compute

There are a number of projects that have been carried out under the FP7 and H2020 frameworks investigating low-power and energy-efficient compute platforms. The following is a small subset of projects that the EUROSERVER partners are aware of and/or participating in related to these topics.

ExaNeSt

European Exascale System Interconnect and Storage - http://www.exanest.eu

Duration: 01 Dec. 2015 - 30 Nov. 2018 (36 months).

ExaNeSt develops and prototypes solutions for Interconnection Networks, Storage, and Cooling, as these have to evolve in order for the production of exascale-level supercomputers to become feasible.

ExaNeSt develops its solutions on 64-bit ARMv8 processors, in order to achieve low power consumption, it uses the EUROSERVER-developed UNIMEM Global Address Space and zero-copy send/receive operations, and it tunes its applications for such architectures.

During its first year, ExaNeSt performed all of its development work on the EUROSERVER Juno-based 64-bit Discrete Prototype (DP64) that FORTH made available to the ExaNeSt partners via remote access; this prototype also included firmware and software optimizations, beyond those of EUROSERVER, that were provided by the ExaNoDe project.

ExaNoDe

European Exascale Processor Memory Node Design - http://exanode.eu/

Duration: 01 Oct. 2015 - 30 Sept. 2018 (36 months)

ExaNoDe investigates, develops, integrates and validates the building blocks for a highly efficient, highly integrated, high-performance, heterogeneous compute node aimed towards Exascale computing. The project will deliver a prototype that will enable the deployment of interposer based System-on-Chips and the evaluation, tuning and analysis of HPC critical kernels along with the associated software stack. ExaNoDe is addressing these important challenges through the coordinated application of several innovative solutions deployed in EUROSERVER: ARM-v8 low-power processors for energy efficiency, 3D silicon interposer architecture partitioning for compute density and UNIMEM for low-latency and high-bandwidth memory access.

ExaNoDe already benefits from the EUROSERVER then ExaNeSt multi-board prototype roadmap to develop and optimize the compute node software including UNIMEM scale-out portability with MPI libraries. ExaNoDe prototype will be validated with an ExaNeSt compatible testbed. ExaNoDe will also integrate UNILOGIC from ECOSCALE within its compute node software stack.

EcoScale

Energy-Efficient Heterogeneous COmputing at exaSCALE - http://www.ecoscale.eu/

Duration: 01 Oct. 2015 - 30 Sept. 2018 (36 months)

ECOSCALE will provide a runtime system and a novel heterogeneous architecture, called UNILOGIC, which is an extension of the UNIMEM architecture introduced in EUROSERVER. In the UNILOGIC architecture ARM-based compute nodes will be able to share reconfigurable hardware-based accelerators distributed in the system. In particular UNILOGIC extends the shared partitioned global address space of UNIMEM offering shared partitioned reconfigurable resources. ECOSCALE is already using the firmware for the UNIMEM architecture developed by FORTH in EUROSERVER. Thus ECOSCALE depends on and extends the results of EUROSERVER.

EuroEXA

Project #754337 – Project not yet started

Co-designed Innovation and System for Resilient Exascale Computing in Europe: from Applications to Silicon

Expected to start in Fall 2017; 42-month duration.

EuroEXA will be the continuation of ExaNoDe, ExaNeSt, and EcoScale - which are all three continuations of EUROSERVER.