

DIVIDEND

Project presentation

Distributed Heterogeneous Vertically Integrated
Energy Efficient Data centres

Madrid, 20th March 2015



Project Overview



Distributed Heterogeneous Vertically Integrated Energy Efficient Data centres

- Context:
 - New digital technologies & data usages are quickly growing worldwide
 - The energy footprint of Data Centers is getting critically high in a context of global warming
 - Hardware providers propose specialized hardware, energy efficient but more difficult to use !

Challenge: Address the energy efficiency bottleneck and the programmability of Heterogeneous data centers

- Project Timeline:
 - Expected starting date: 1st May 2015
 - Project Duration: 24 months

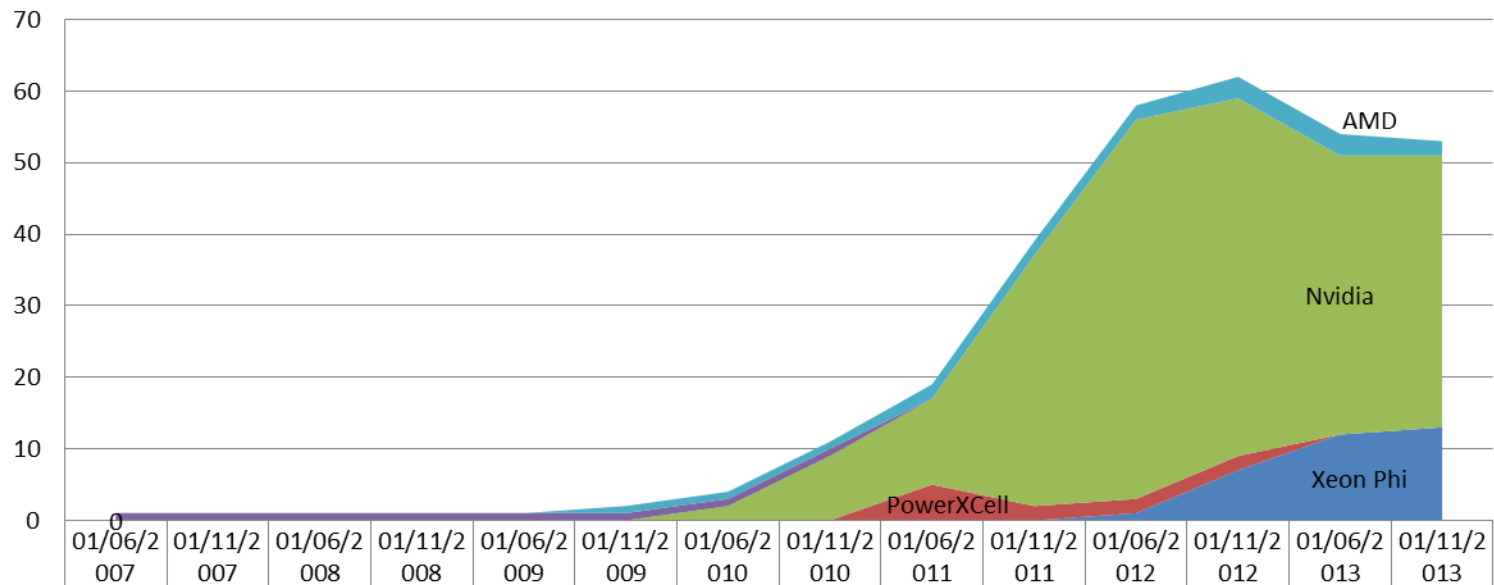
More and more accelerators available...



(LMorin)

Number of Accelerator by Family in TOP 500

source: top500.org 2014

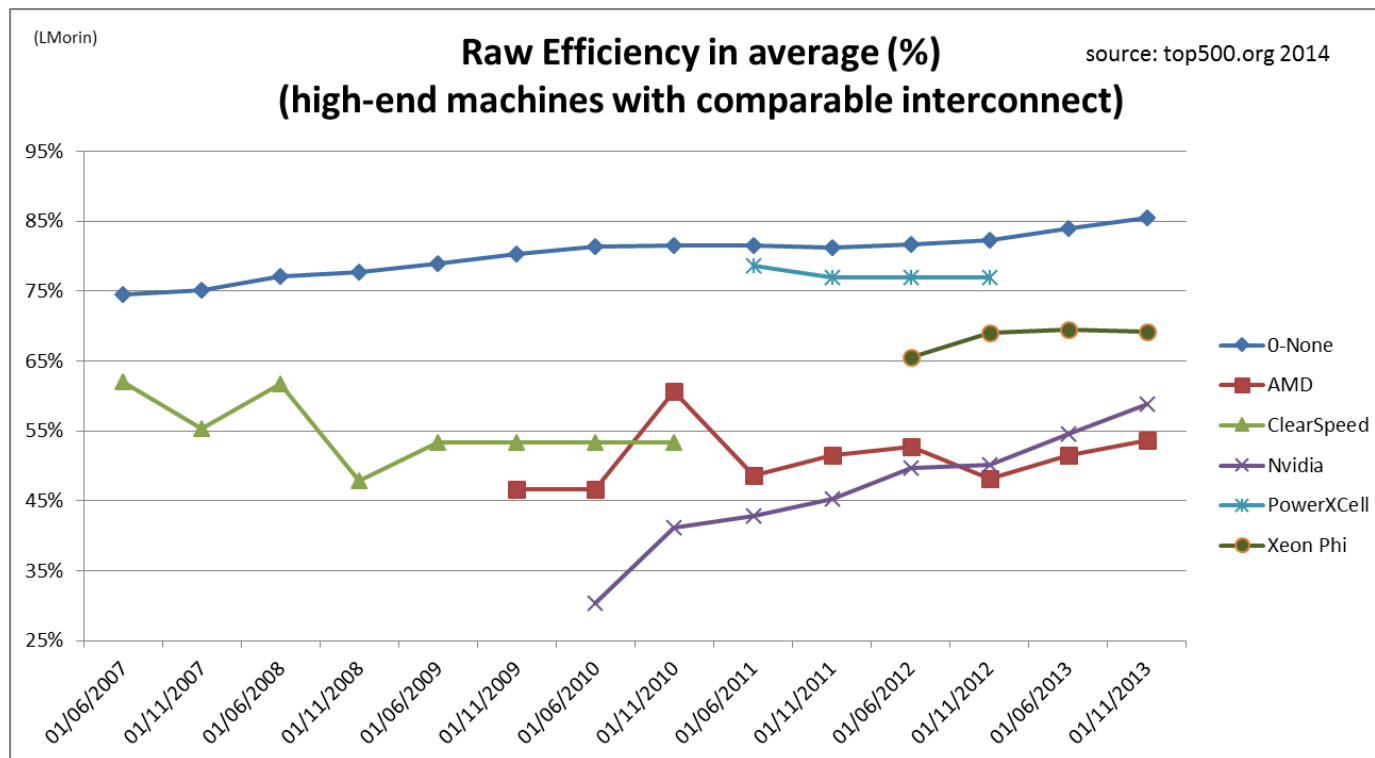


	01/06/2007	01/11/2007	01/06/2008	01/11/2008	01/06/2009	01/11/2009	01/06/2010	01/11/2010	01/06/2011	01/11/2011	01/06/2012	01/11/2012	01/06/2013	01/11/2013
AMD						1	1	1	2	2	2	3	3	2
ClearSpeed	1	1	1	1	1	1	1	1						
Nvidia							2	9	12	35	53	50	39	38
PowerXCell									5	2	2	2		
Xeon Phi											1	7	12	13

... with less raw efficiency despite a better energy efficiency!



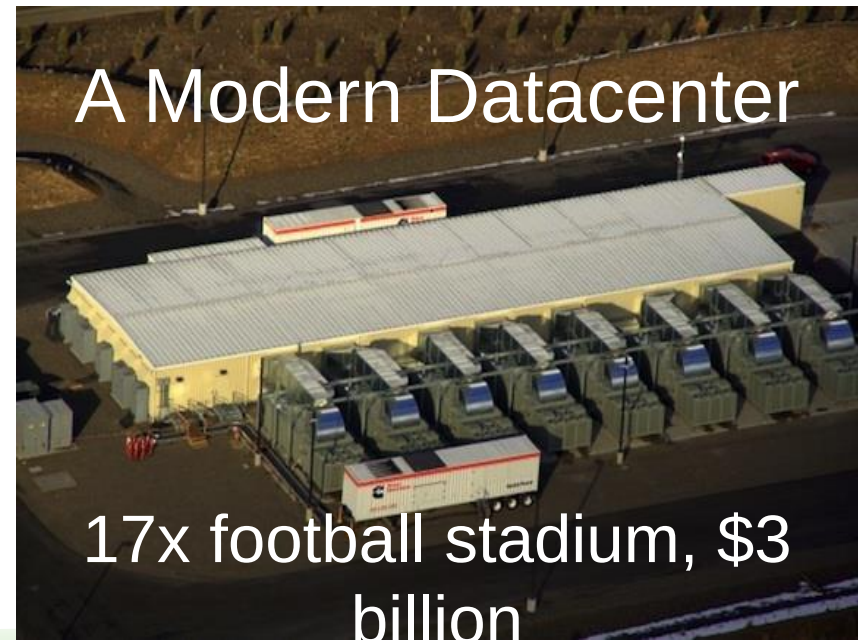
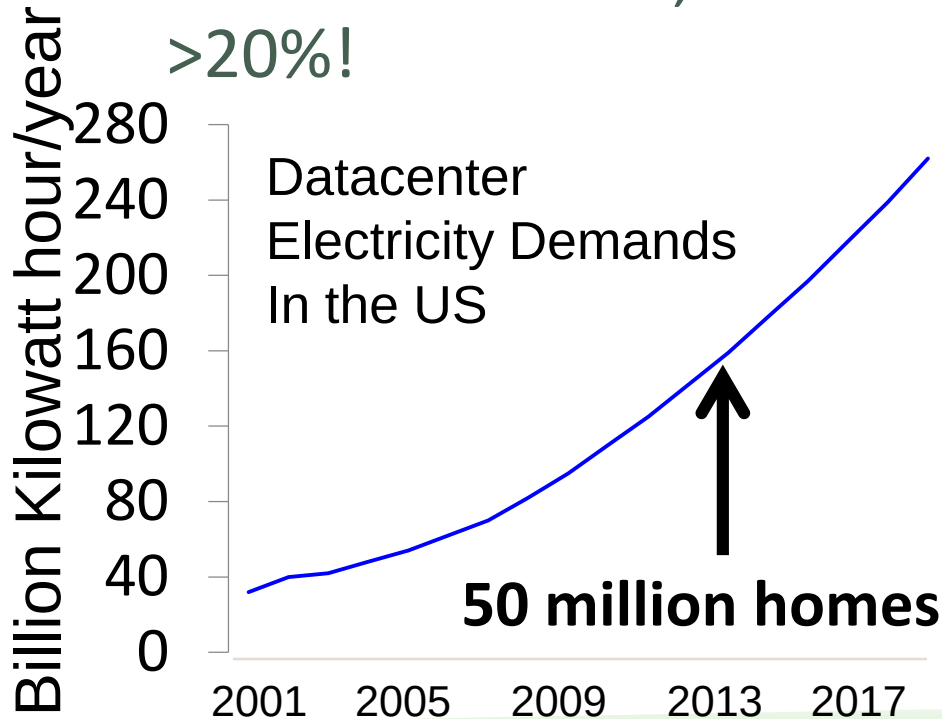
- Modern HPC accelerators unusable for Data Centers



Higher Demand + Lower Efficiency: Datacenter Energy Not Sustainable!



- Modern datacenters → 20 MW!
- In modern world, 6% of all electricity, growing at >20%!



Dividend Objectives



- Expand the energy performance of large-scale analytic applications using vertical integration & cross-layer optimizations

Three key issues addressed...

- Energy consumption: at all scales of the computations
- Heterogeneity: CPUs, GPUs, and other accelerators
- Execution Distribution: via data partitioning & movement control

... with a global framework supporting:

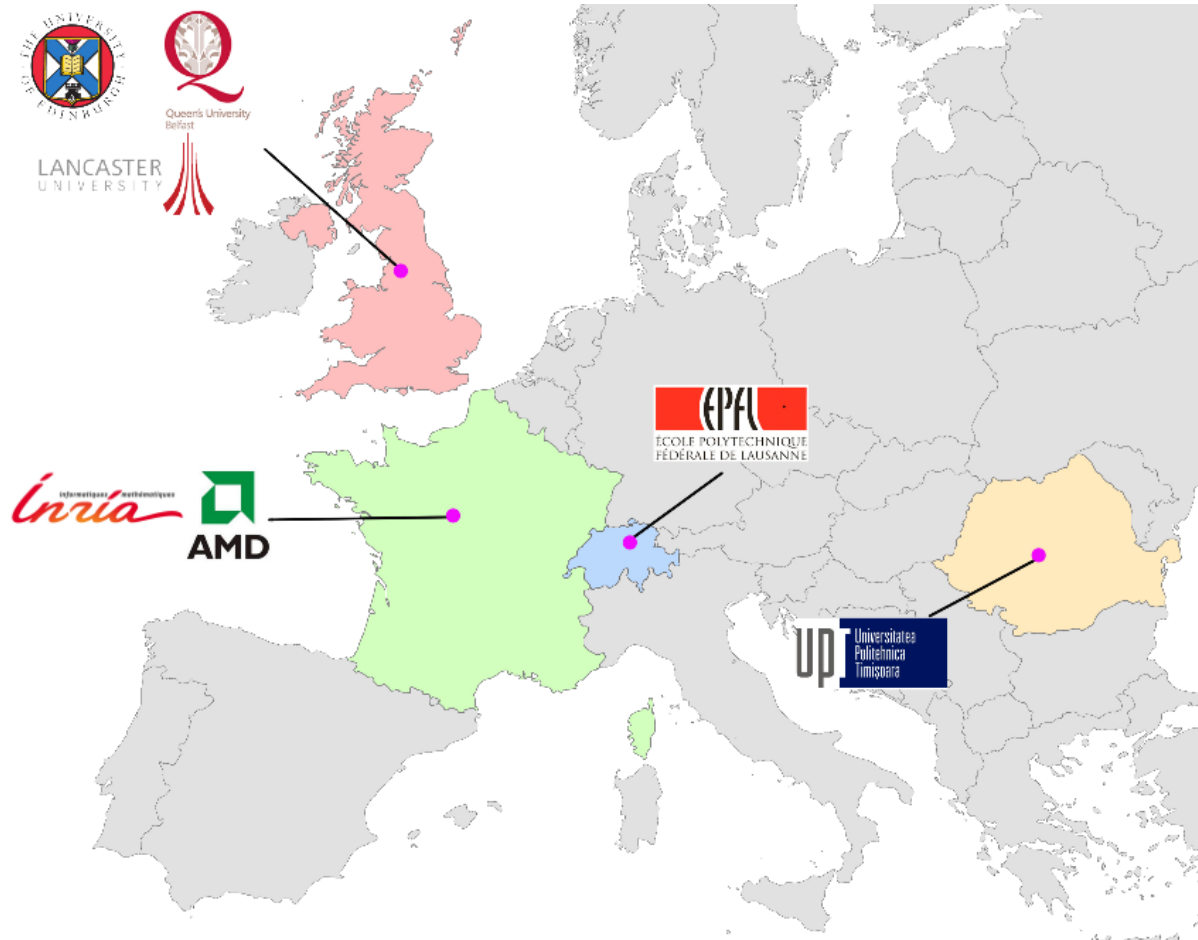
- Vertical and distributed energy accounting
- Data center scale scale-out NUMA interconnect
- Vertical integration of the optimization process
- Energy aware compilation and runtime optimization tool chain

Project Partners (1)



- University of Edinburgh – UoE
 - Prof. Michael O’Boyle
- Queens University of Belfast – QUB
 - Prof. Dimitrios S. Nikolopoulos
- University of Lancaster – LANC
 - Dr Zheng Wang
- Ecole polytechnique Fédérale de Lausanne – EPFL
 - Prof. Babak Falsafi
- Advanced Micro Devices – AMD
 - Dr Mauricio Breternitz
- Universitatea Politehnica Timisoara – UPT
 - Dr. Alexandru Amaricai
- Institut National de Recherche en Informatique et en Automatique – INRIA
 - Dr. Albert Cohen

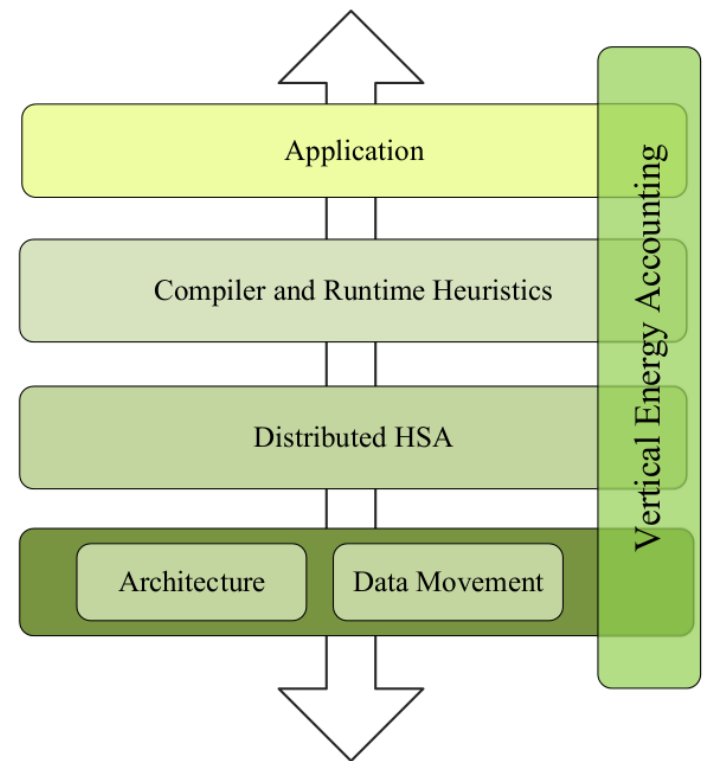
Project Partners (2)



Architecture & Technologies



- HSA: Heterogeneous System Architecture
- Vertical Energy Accounting
- Application Specification and Implementation
- Compilation and Runtime Heuristics
- Distributed Heterogeneous Architectures
- Data center scale Distributed Memory System
- Heterogeneous Accelerators



HSA Technology (1)



Heterogeneous System Architecture – HSA

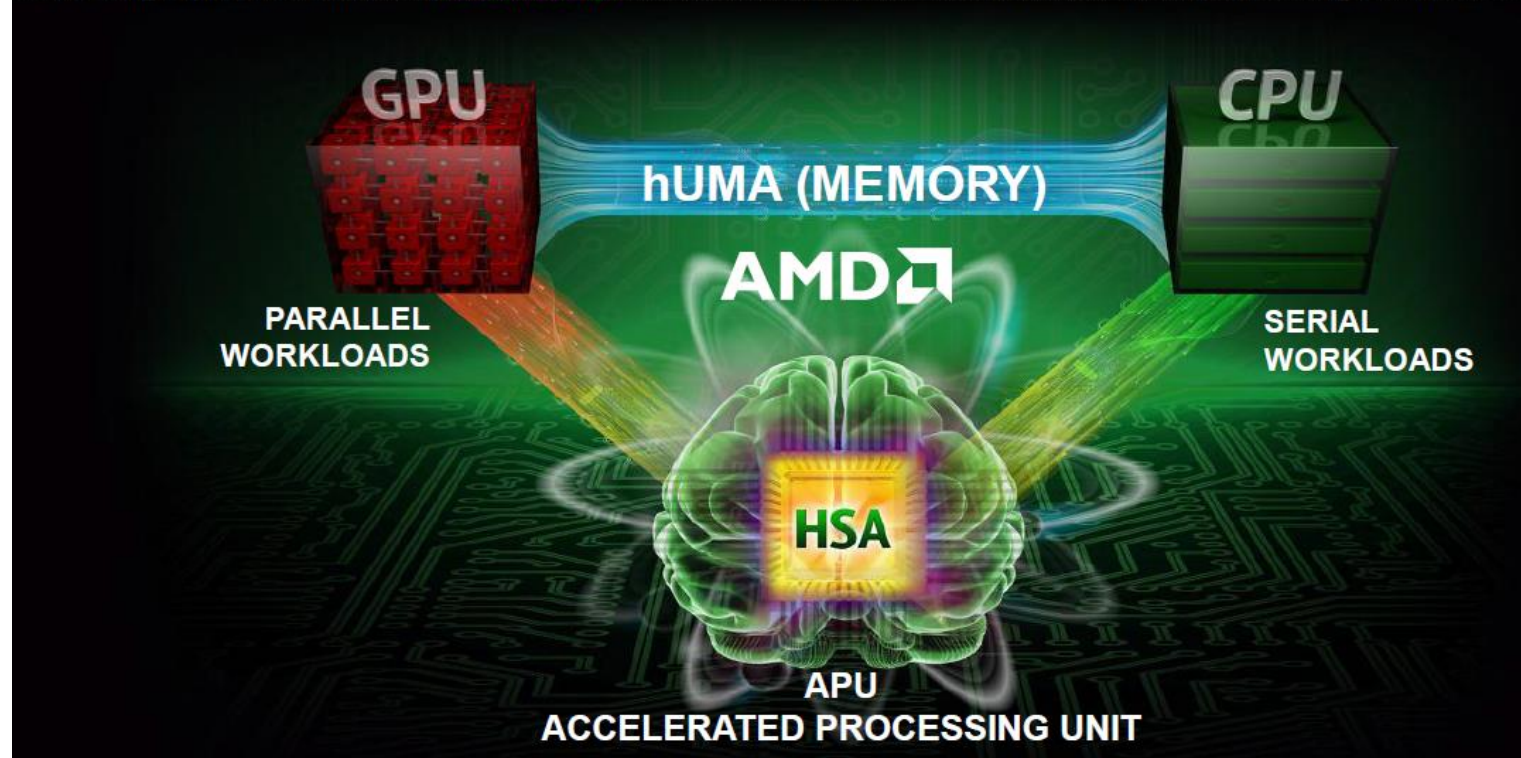
- Family of computer processor architectures combining CPUs and GPUs
 - All sharing memory and tasks with a virtual address space and an “Heterogeneous uniform memory access” (hUMA)
- The standard proposes:
 - An intermediate language (HSAIL)
 - A Memory model specification
 - A Runtime specification
- Supported by the HSA Foundation (2012)
 - A not-for-profit engineering organization composed of both industry and academia associates
 - Founded by AMD and ARM Holdings,
 - But also Imagination Technologies, MediaTek, Qualcomm, Samsung, and Texas Instruments

HSA Technology (2)



WHAT IS HSA?

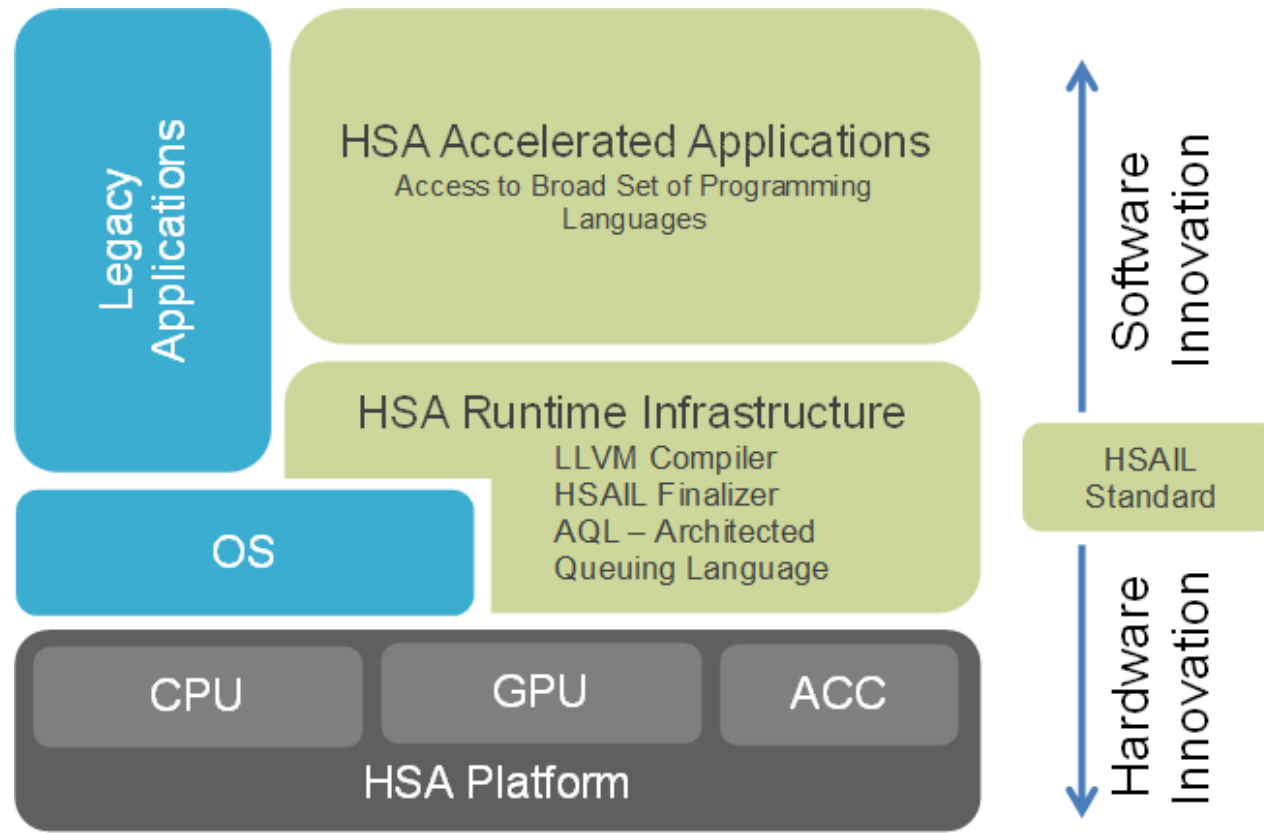
An *intelligent computing architecture* that enables CPU, GPU and other processors to work in *harmony* on a single piece of silicon by *seamlessly* moving the right tasks to the best suited processing element



HSA Technology (2)



HSAIL: HSA Intermediate Language



Vertical Energy Accounting

(WP2)



- Vertical EPA: Energy Performance Accounting
 - Develop energy modeling & measures for data centers
 - New energy models of shared resources for virtual machines
 - New data transfers energy consumption measurement approaches
 - New energy measurement approaches based on activity counters
 - Virtualization of Energy as a resource
 - Provide abstraction level energy accounting (MapReduce)
 - Develop a runtime energy monitoring tool

Key Technology:

- ALEA: Programmable Energy-Efficiency for Many-Core Systems
 - Statistical profiling framework targeting energy and time consumption for Intel and ARM architectures

Application Specification and Implementation (WP4)



- Fundamental building blocks of data-centric computation in data centres
 - Key-value stores and Big Data analytics
- Deployment and exposure of optimisation opportunities
 - Implement workloads and deploy them on heterogeneous architectures
 - Identify bottlenecks in performance and energy consumption
- Mapping to low-level elements of the architectures
 - Optimisation through adaptive system reconfiguration

Key Technology:

- Hadoop Distributed Processing Framework



Compilation and Runtime Heuristics (WP3)



- Portable & energy aware compiler and runtime heuristics
 - Develop energy-first static and dynamic optimization heuristics (code & runtime)
 - Develop an automatic adaptation of compiler & runtime heuristics for distributed systems based on predictive models
 - Introduce machine learning technics to optimize task mapping and data movements
 - Main target: heterogeneous data centers

Key Technologies:

- LLVM Compilation infrastructure (HSA backend)
- State of the art Machine learning technics

Distributed Heterogeneous Architectures (WP6)



- Accounting for heterogeneity at the system level
 - HSA prototype drivers for energy estimation & accounting, and integration with the vertical accounting
- Enriching system with distributed processing primitives
 - Runtime and API support for language and system-level hooks for high-level application and system configuration
- Accelerating distributed heterogeneous systems
 - Enrichment of the distributed architecture platform and interconnect features to propagate energy information between layers (from the application to the hardware)

Key Technology:

- HSA compilation infrastructure

Data center scale Distributed Memory System (WP5)



- Provide architectural and micro-architectural support for the Scale-Out NUMA model
 - Design a rack-scale SO-NUMA
 - Bridge rack-scale systems via an inter-rack “global” interconnect layer (data center scale)
 - Develop architectural and micro-architectural support for Distributed HSA

Key Technology:

- SO-NUMA: an architecture, programming model, and communication protocol for low-latency, distributed in-memory processing

Heterogeneous Accelerators (WP7)

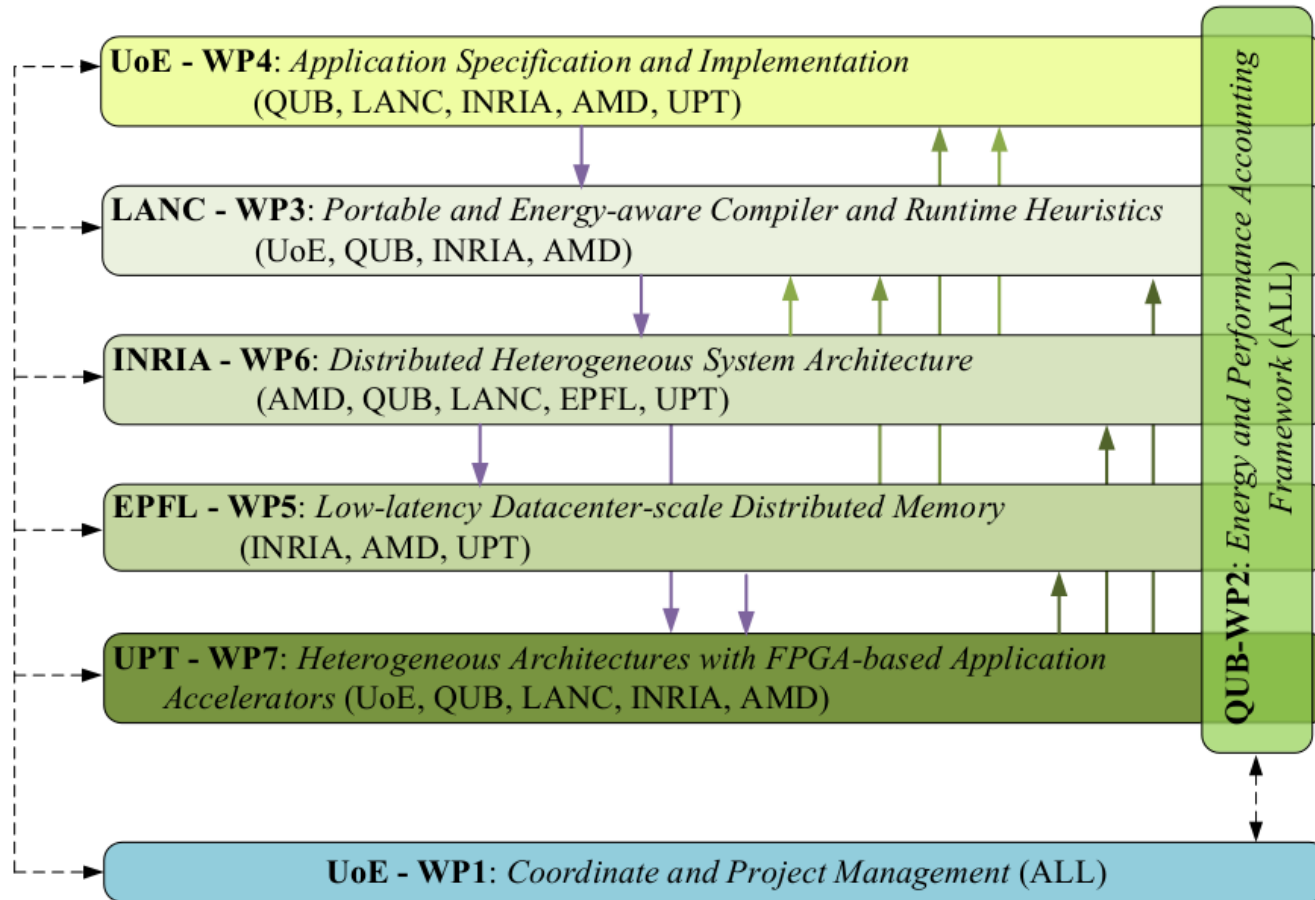


- Heterogeneous Architectures with FPGA-based Application Accelerators
 - Design of FPGA Accelerators
 - Targeting a reconfigurable acceleration of selected data center applications
 - Develop Energy and performance monitoring infrastructure for FPGA based accelerators
 - Monitor hardware activity counters,
 - Feed the vertical energy accounting information system
 - Develop FPGA layer support for HSA
 - Integrate FPGA into the HSA shared address space and queues

Key Technologies:

- Xilinx FPGA and HSA Architecture model

Project Organization (1)



Project Organization (2)



- Total effort per WP and partner (Person-months)

Project partner	WP1	WP2	WP3	WP4	WP5	WP6	WP7	Total
1 (UoE)	3	3	8	14	0	0	2	30
2 (QUB)	1	18	8	2	0	3	2	34
3 (LANC)	1	2	18	2	0	3	4	30
4 (EPFL)	1	2	0	0	40	8	0	51
5 (AMD)	0	0	0	0	0	0	0	0
6 (UPT)	1	4	0	2	6	12	15	40
7 (IN-RIA)	1	4	32	12	6	30	23	108
Total	8	33	66	32	52	56	46	293

Expected Outcomes & collaborations

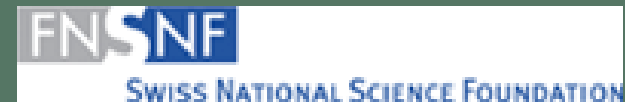


- Majors Scientific outcomes:
 - Develop methods and tools for energy-aware optimization throughout the hardware and software stack
 - Involve and help the programmer in energy management and energy efficient parallel programming
 - Develop the industrial HSA standard and technology to support energy accounting and power consumption optimizations
- Strong Collaborations should emerge from:
 - the development of the vertical energy accounting system across all work packages
 - the use of the HSA technology with the support of the industry

Dissemination & Exploitation



- Dissemination, a major objective of the project:
 - Dissemination of results in international conferences, workshops, and journals:
 - TECS, TPDS, TACO, IEEE TC, DATE, DAC, CASES, ISCA, PPOPP, PACT, CGO, ...
 - Web-portal development and maintenance
 - Demonstrator systems
- Exploitation via education...
 - Between one and three PhD Students planned per partner
- ... and partnerships with associations & industrial partners
 - Direct exploitation with the AMD HSA product line,
 - CriticalBlue and Vector Fabrics in compiler technologies (LANC),
 - EcoCloud Research center at EPFL,
 - Xilinx (UPT)



DIVIDEND – A CHIST-ERA Project