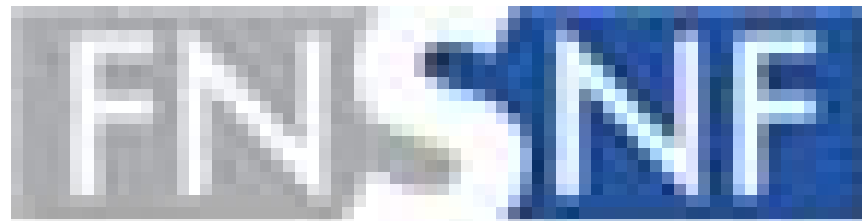
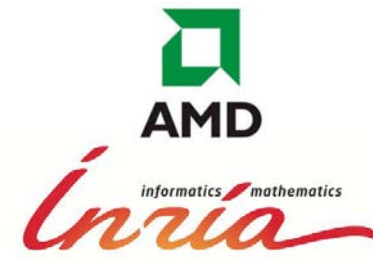


DIVIDEND

Distributed Heterogeneous
Vertically Integrated Energy
Efficient Data centres



Data is Shaping Future of IT



- Data growth (by 2015) = 100x in ten years [IDC 2012]
 - Population growth = 10% in ten years



Datacentres Growing Fast

Source: James Hamilton, 2012

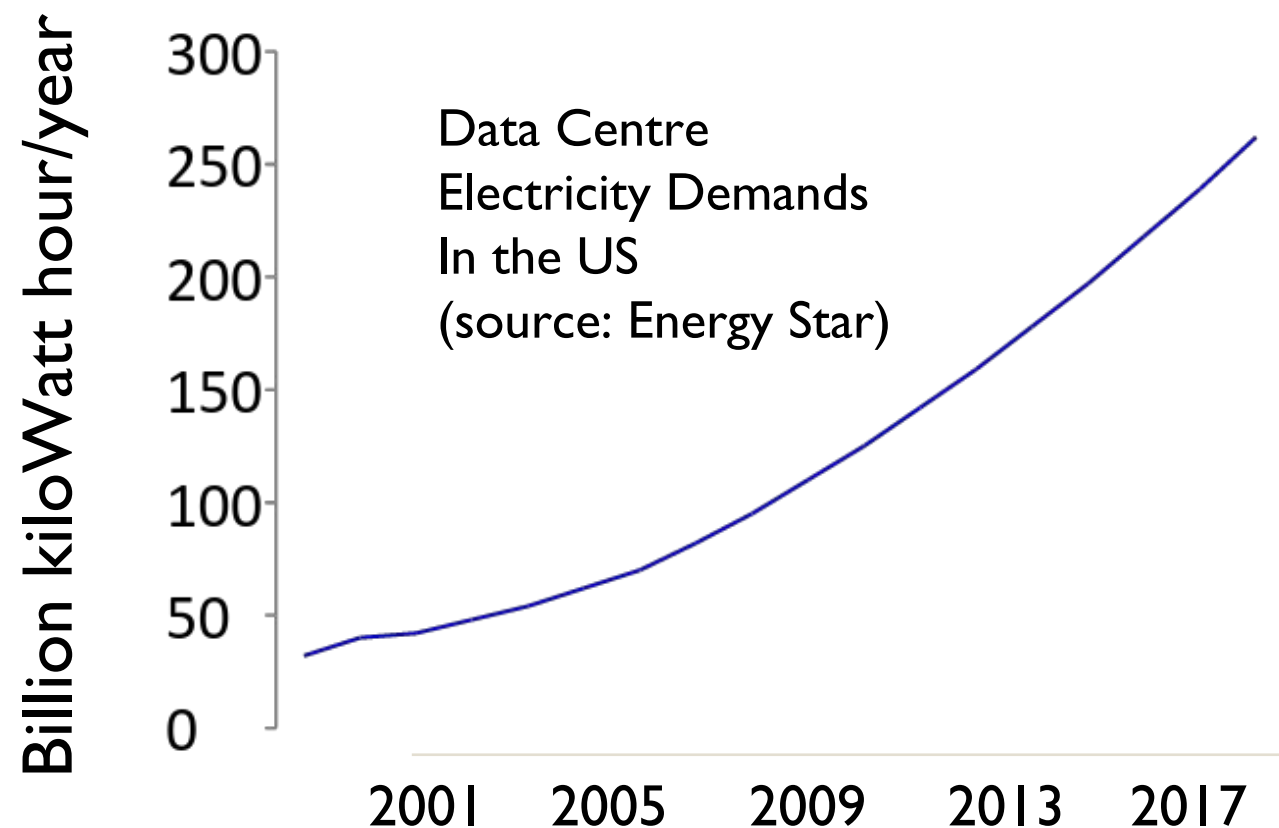


Each day Amazon Web Services adds enough new capacity to support all of Amazon.com's global infrastructure through the company's first 5 years, when it was a \$2.76B annual revenue enterprise

Daily IT growth in 2012 = IT first five years of business!

Higher Demand + Lower Efficiency: Data Centre Energy Not Sustainable!

~20MW each, 1.5% globally,
growing exponentially



Energy efficient heterogeneous DC is hard

- CPUs, GPUs, FPGAs - complex interactions
- Massive distribution - complex network
- Must optimise at multiple scales



What we need

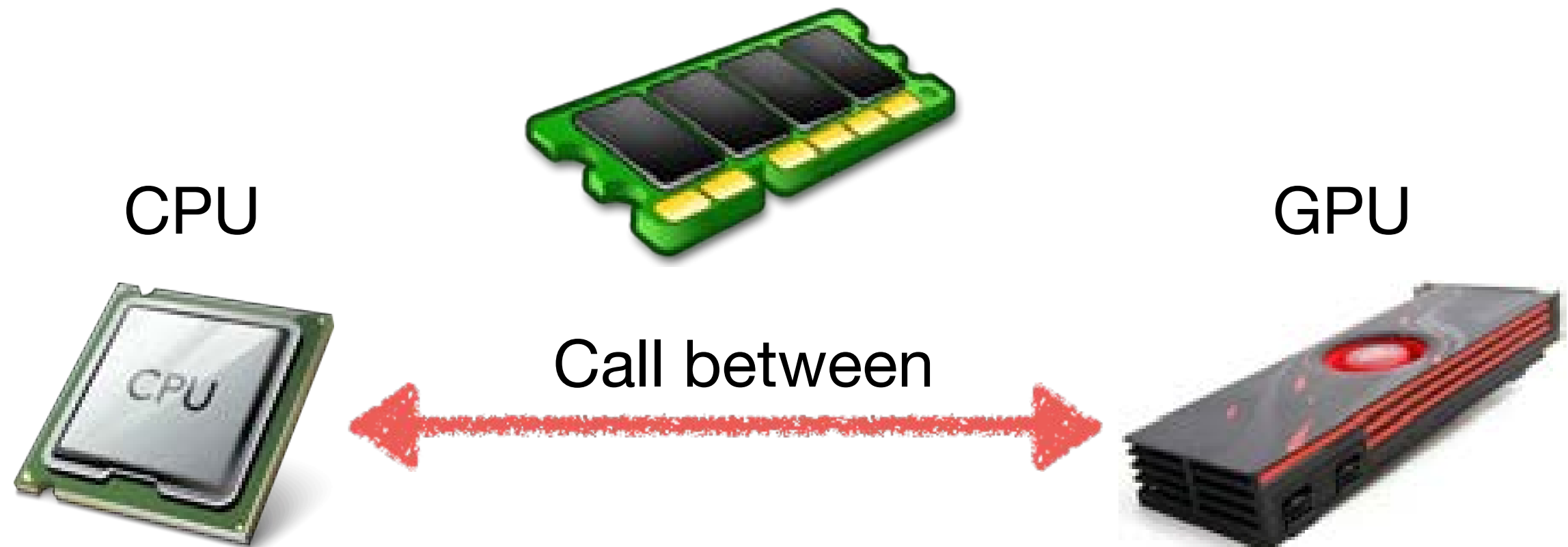
- Single paradigm to program heterogeneous systems
- Efficient communication
- Multi-scale/crosscutting energy profiling
- Multi-scale/crosscutting energy aware optimisation



Heterogeneous Systems Architecture

HSA

Single memory address space

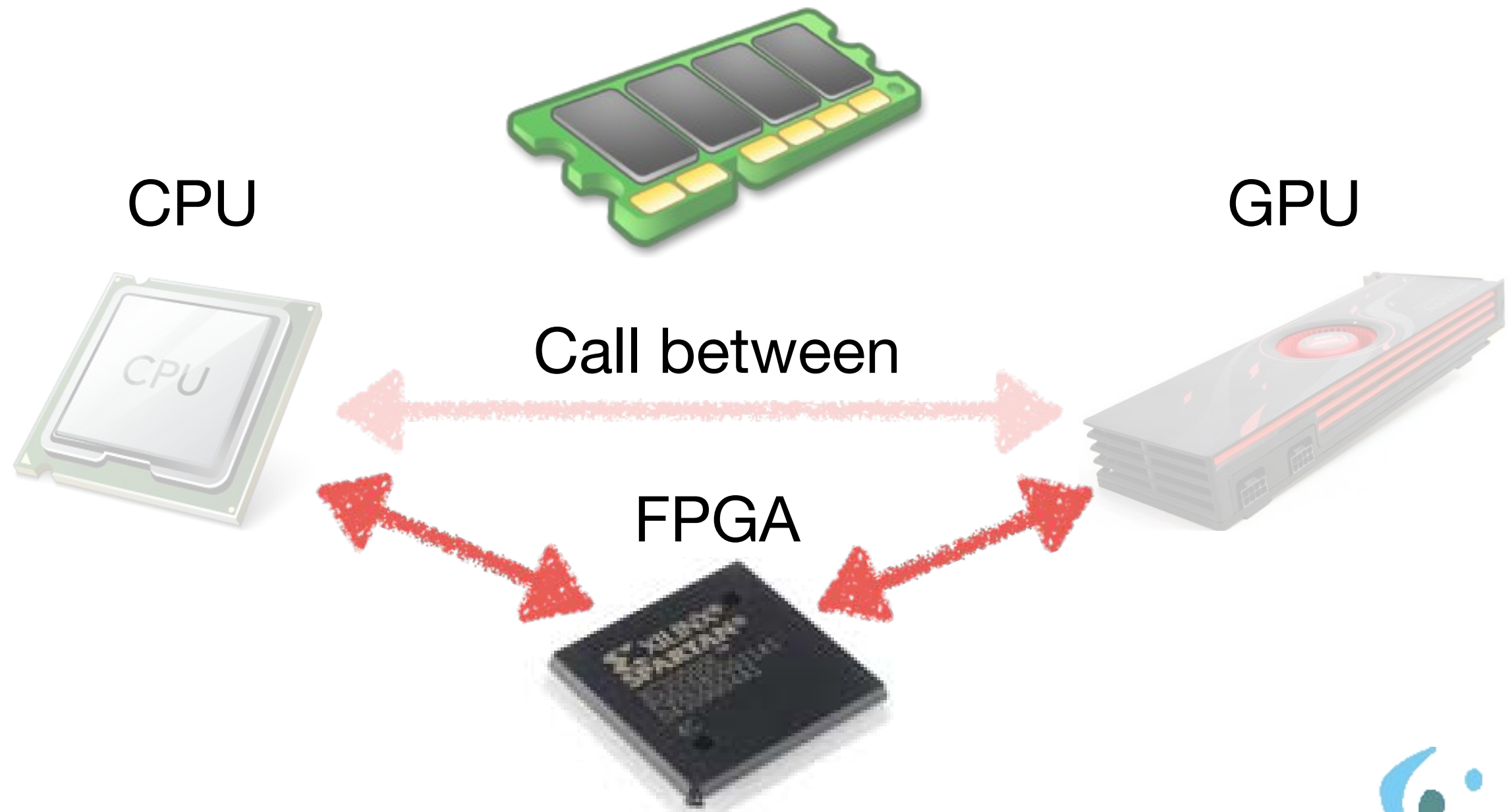


Single programming paradigm



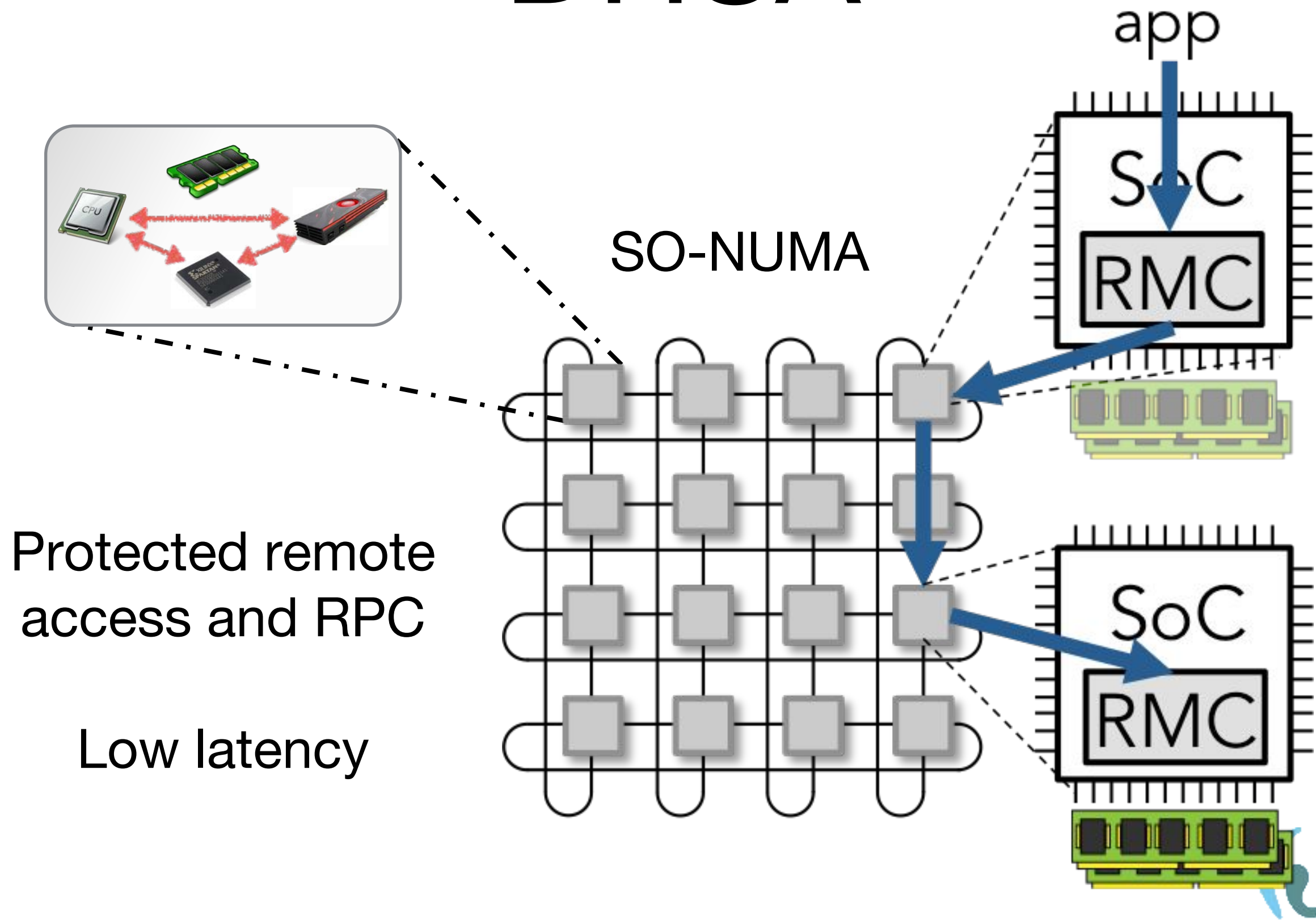
HSA + FPGA

Single memory address space



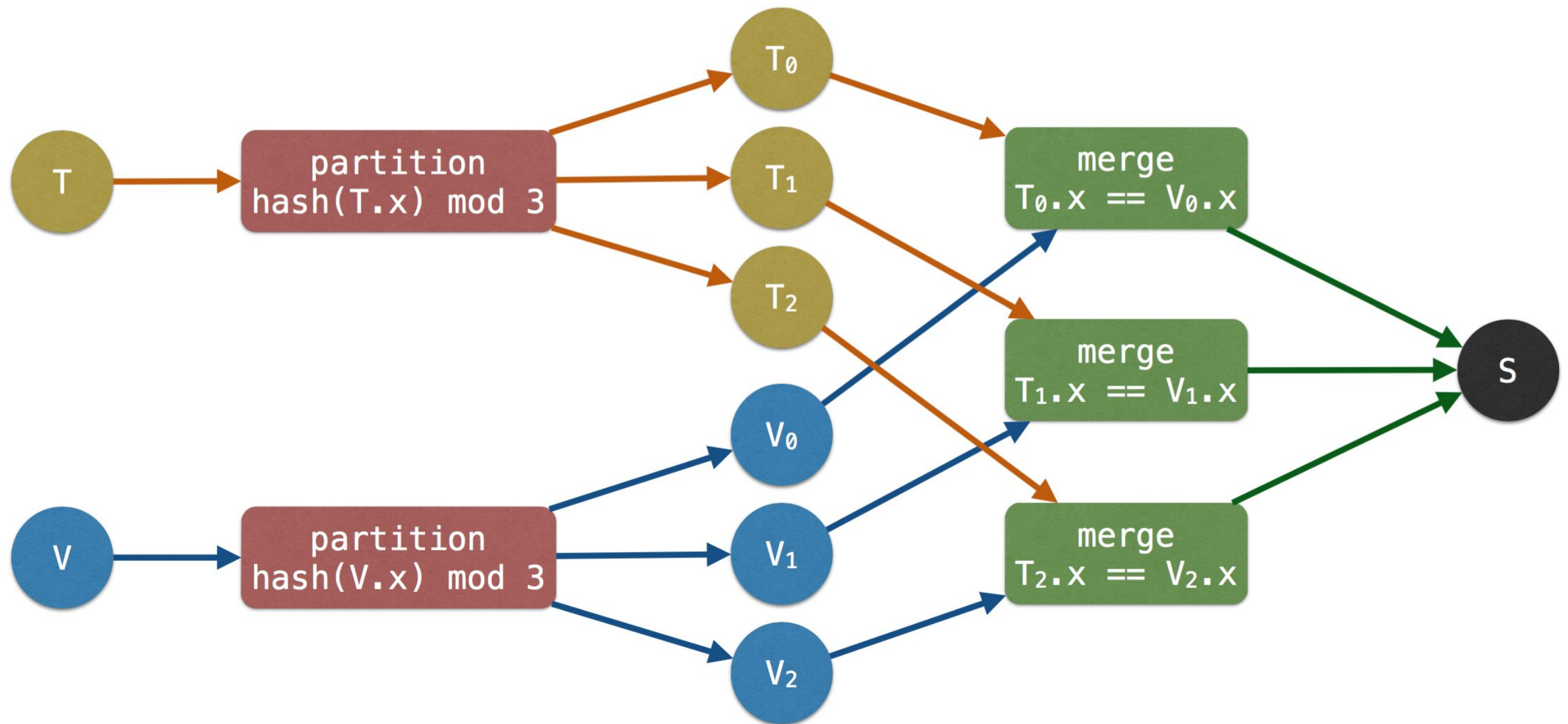
Distributed Heterogeneous Systems Architecture

DHSA



Simple, high level programming interface

DHSA Application Layer



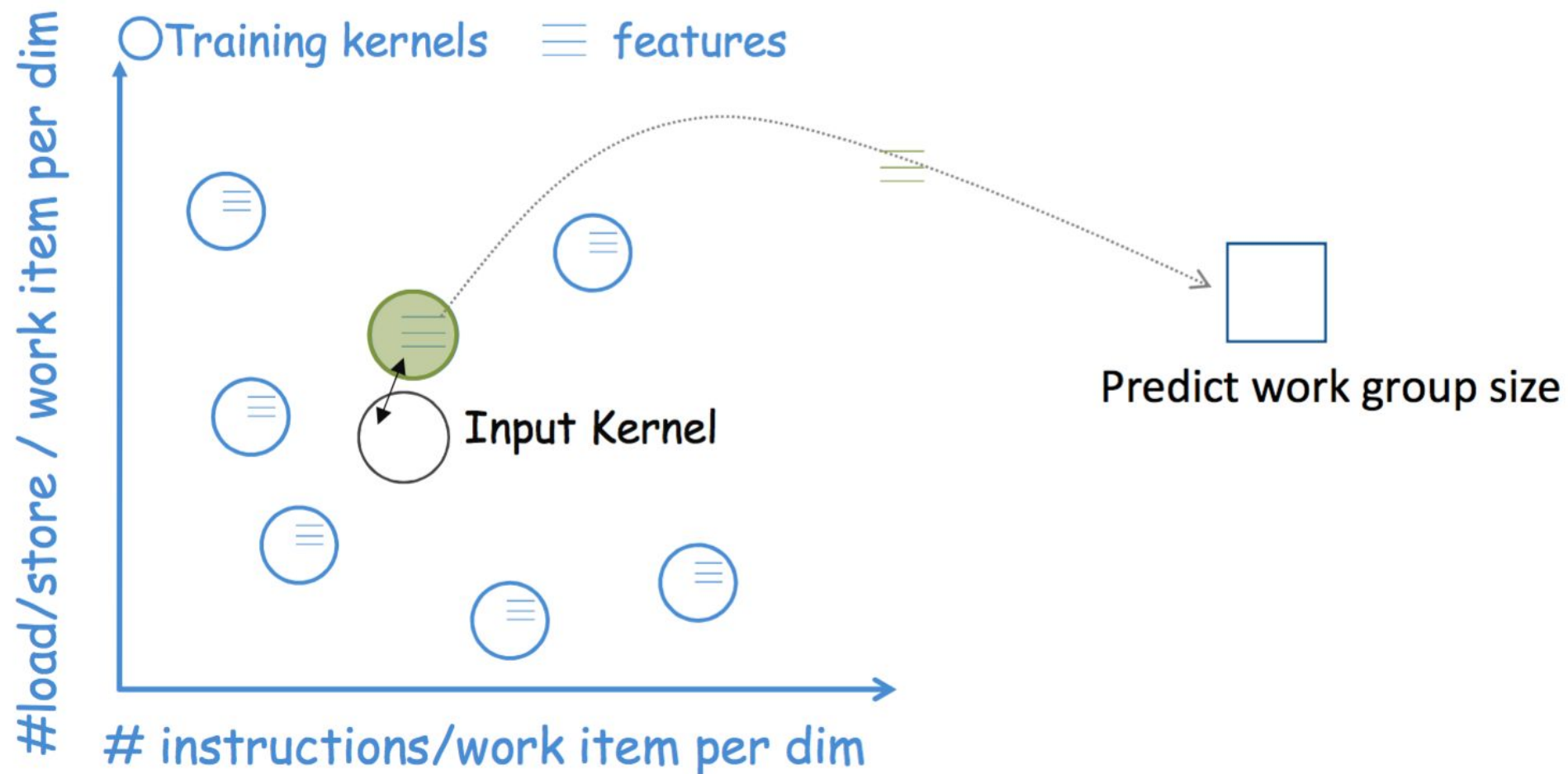
Energy Accounting



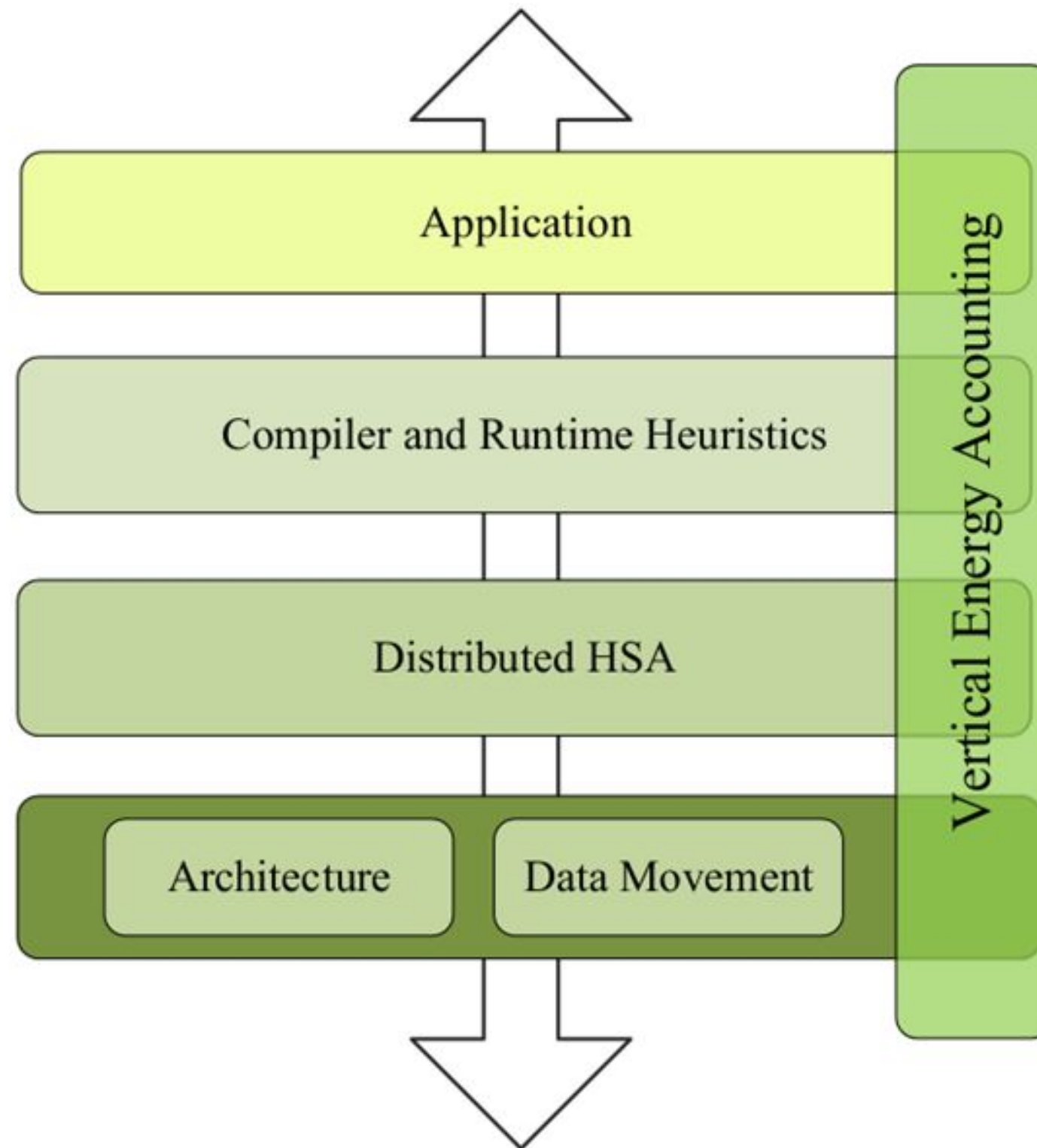
```
QModelIndex start;  
if (currentIndex().isValid())  
    start = currentIndex();  
else  
    start = d->model->index(0, 0, d->root);  
  
bool skipRow = false;  
bool keyboardTimeWasValid = d->keyboardInputTime.isValid();  
qint64 keyboardInputTimeElapsed = d->keyboardInputTime.restart();  
if (search.isEmpty() || !keyboardTimeWasValid  
    || keyboardInputTimeElapsed > QApplication::keyboardInputInterval()) {  
    d->keyboardInput = search;  
    skipRow = currentIndex().isValid(); //if it is not valid we should reali  
} else {  
    d->keyboardInput += search;  
}
```



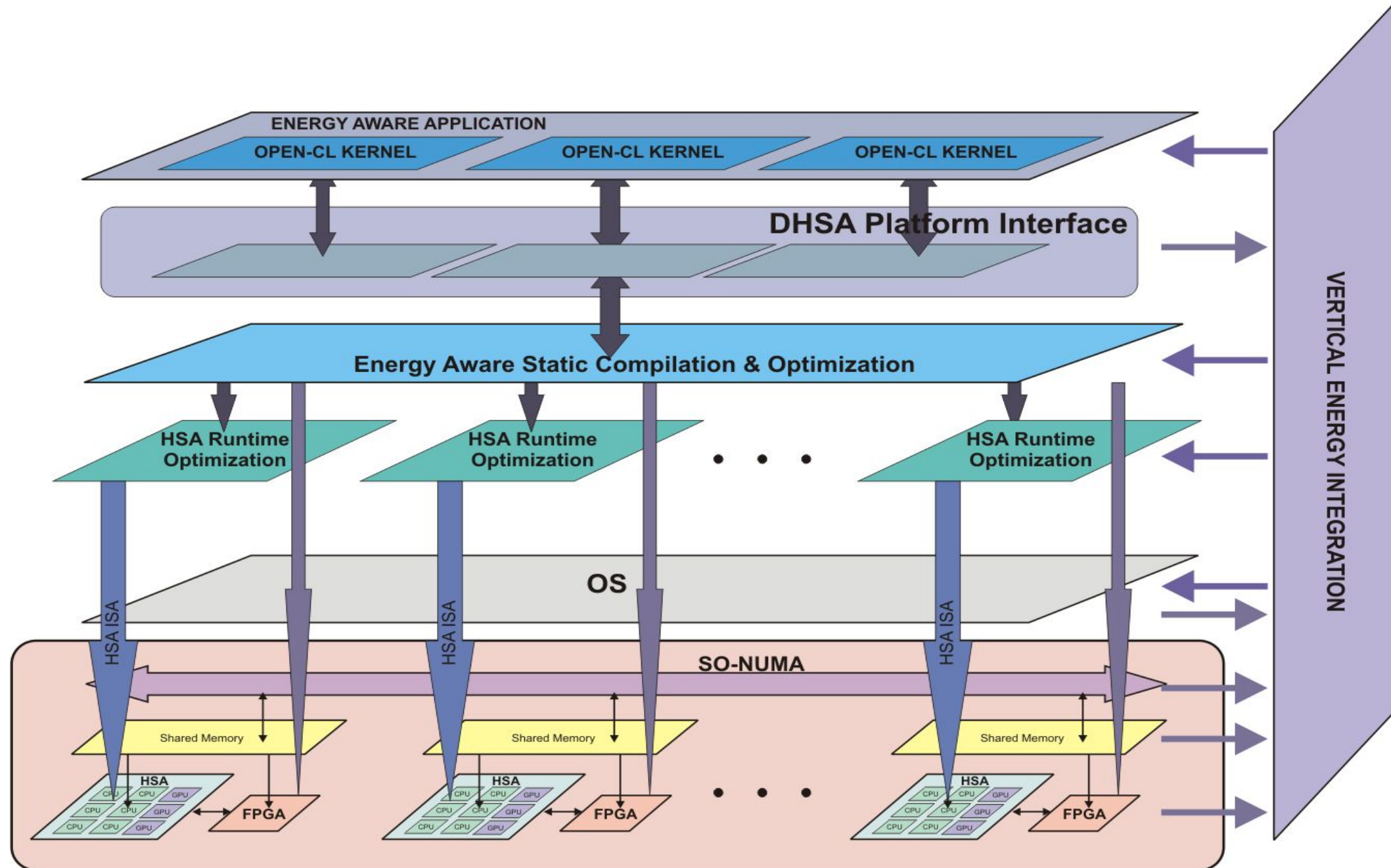
Energy Optimisation



Project organization



Project organization

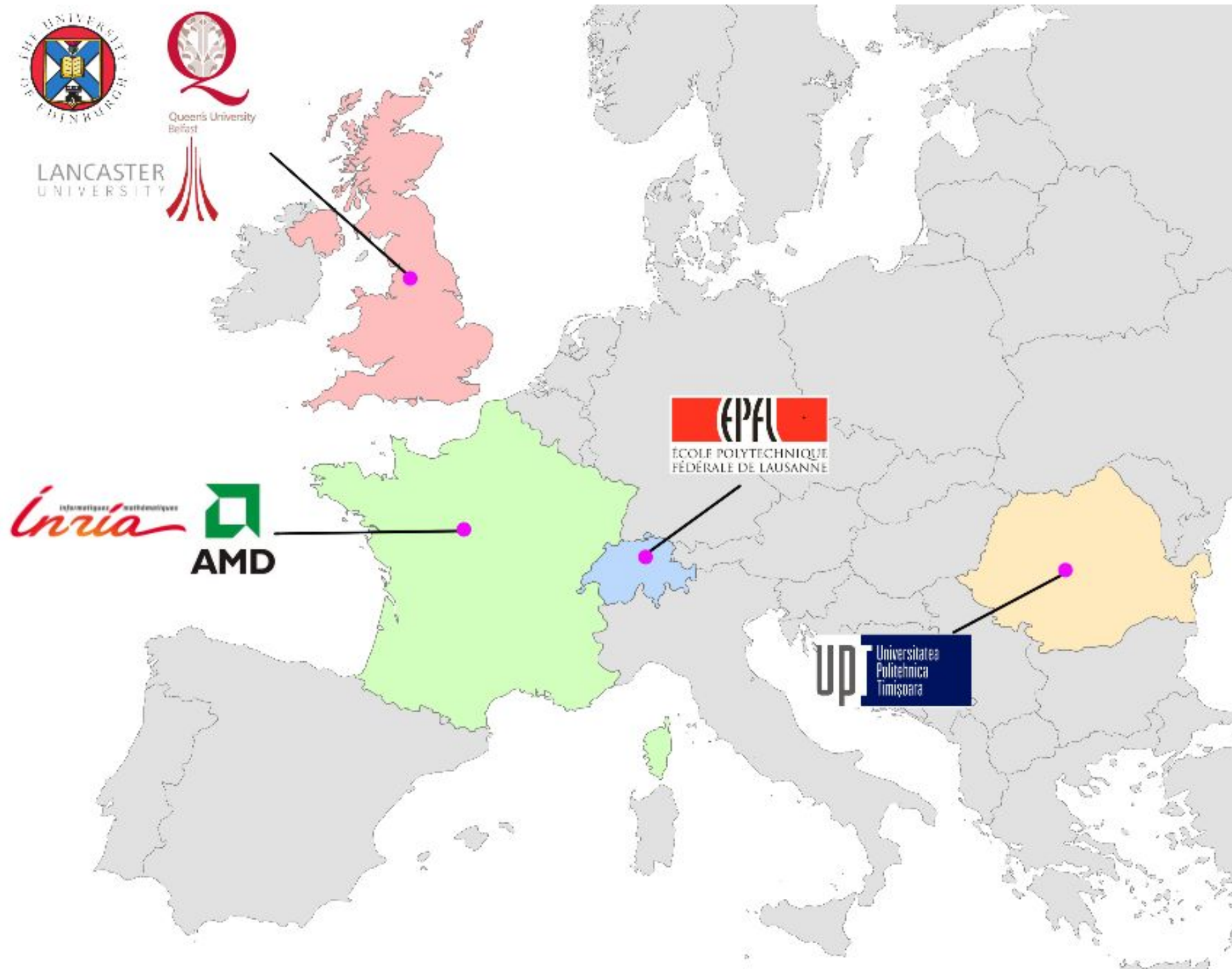


Project Partners

- 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

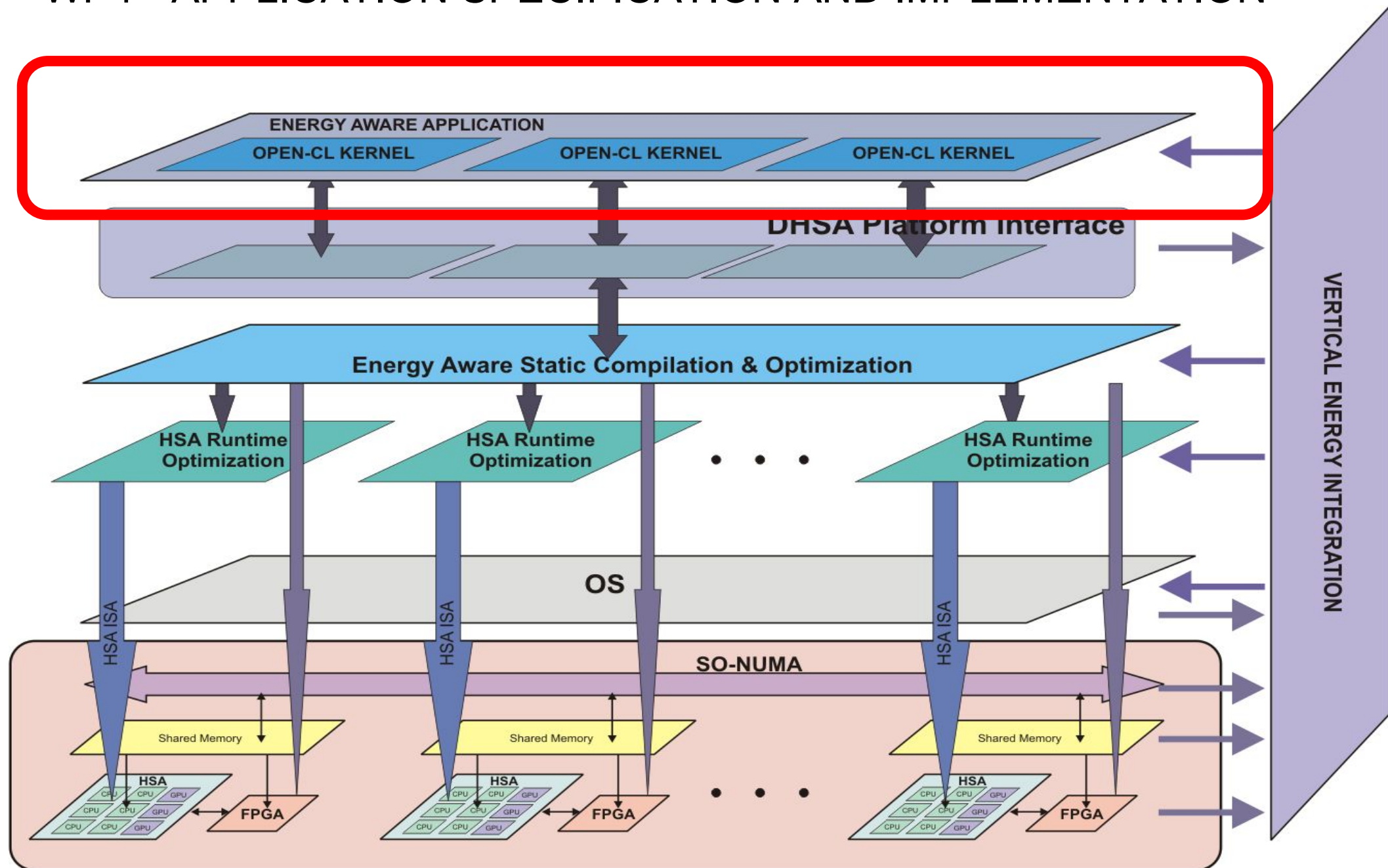


Project Partners



Project organization

WP4 - APPLICATION SPECIFICATION AND IMPLEMENTATION



University of Edinburgh

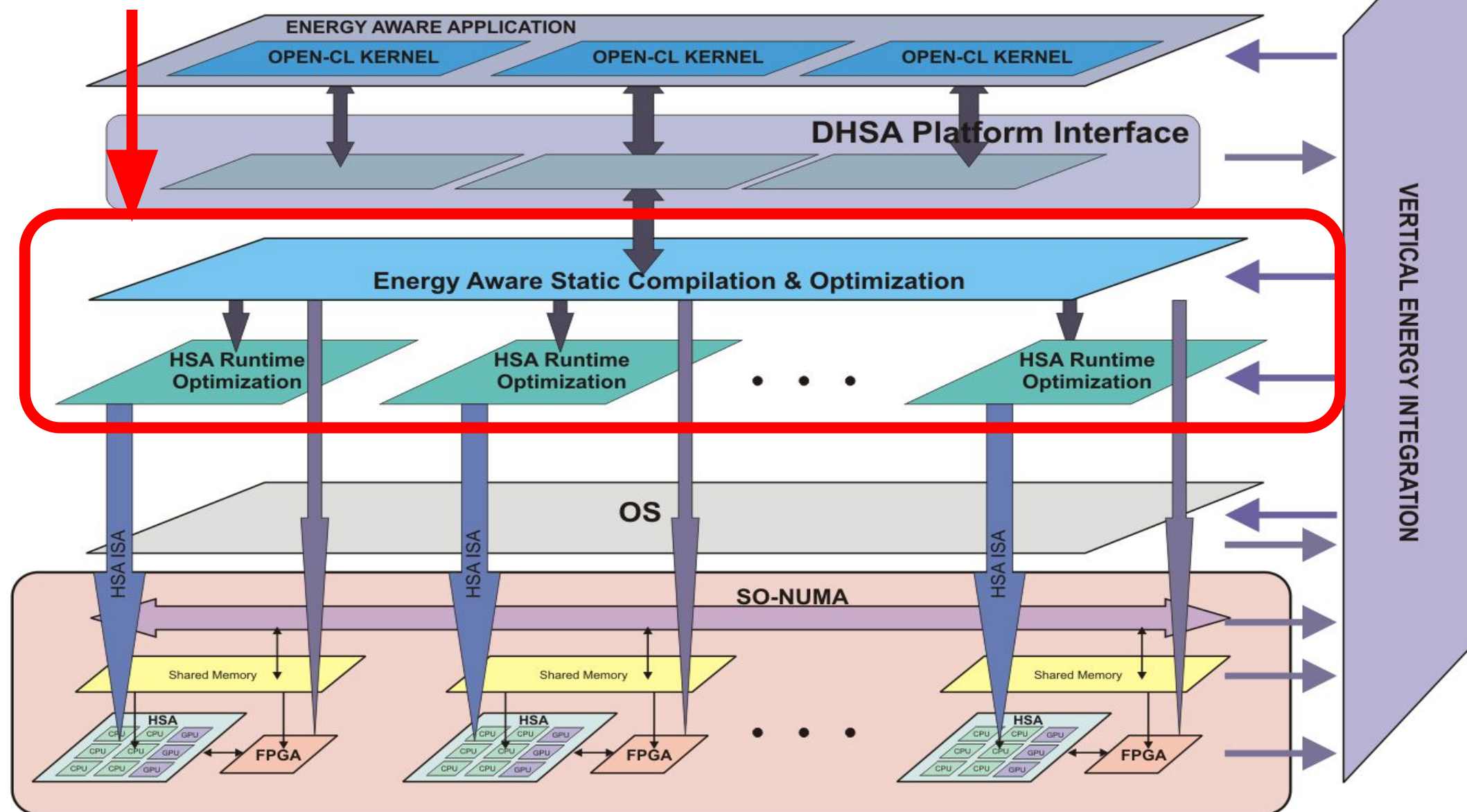
WP 4 Application Specification and Implementation

Added Value

- Database and Distribution expertise
- Distributed Heterogeneous Big Data APIs
- Benchmarks as fast as PostgreSQL

Project organization

WP3 - PORTABLE AND ENERGY AWARE COMPILER AND HEURISTIC OPTIMIZATION





University of Lancaster

WP 3 Portable and Energy-Aware Compiler and Runtime Heuristics

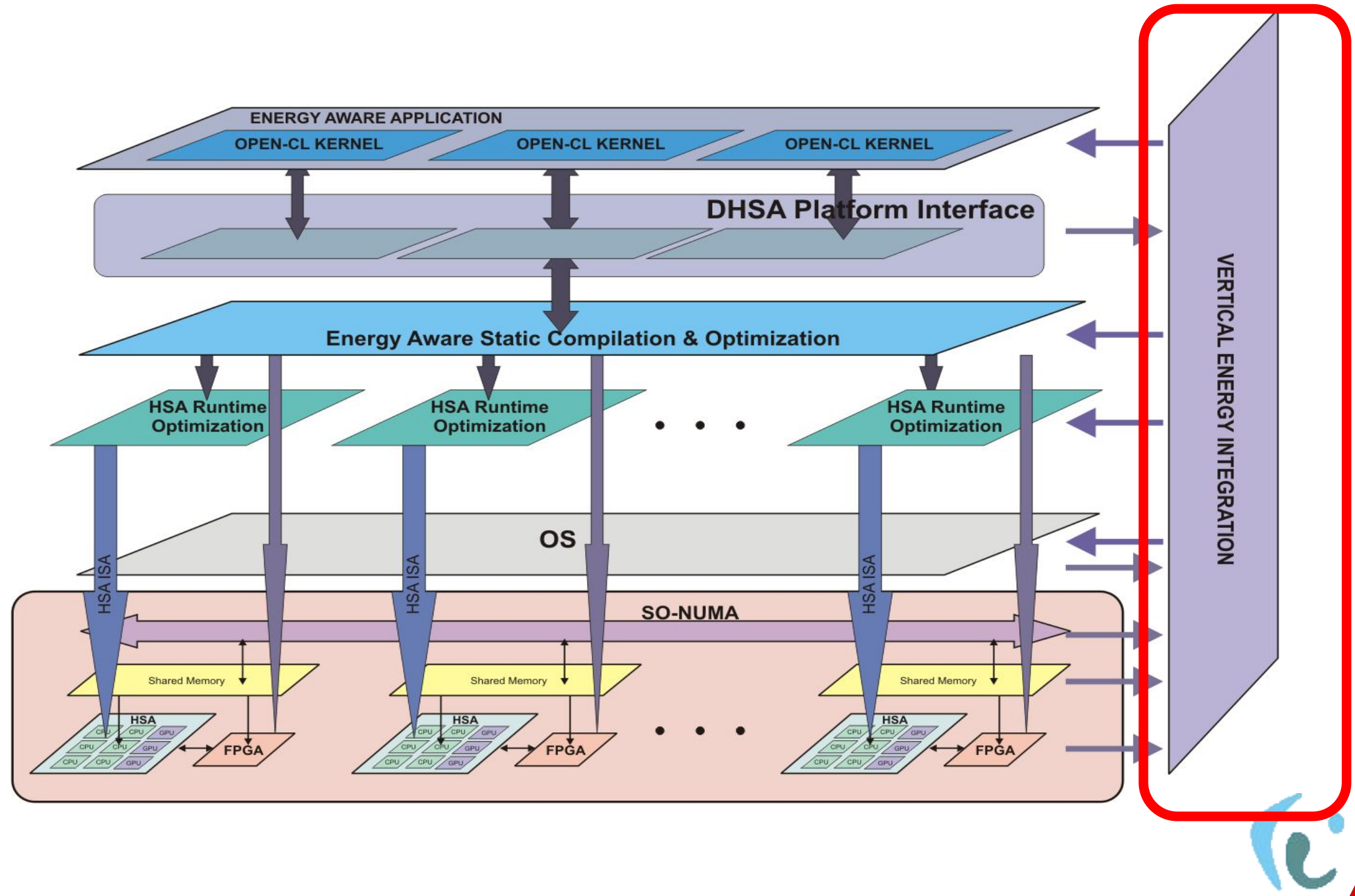
Added Value

- LLVM based, HSAIL auto-tuning for energy
- Machine learned heuristics
- Achieves 22% energy reduction



Project organization

WP2 - VERTICAL ENERGY AND PERFORMANCE FRAMEWORK





Queen's University Belfast

WP 2 Vertical EPA (Energy and Performance Accounting)
Framework

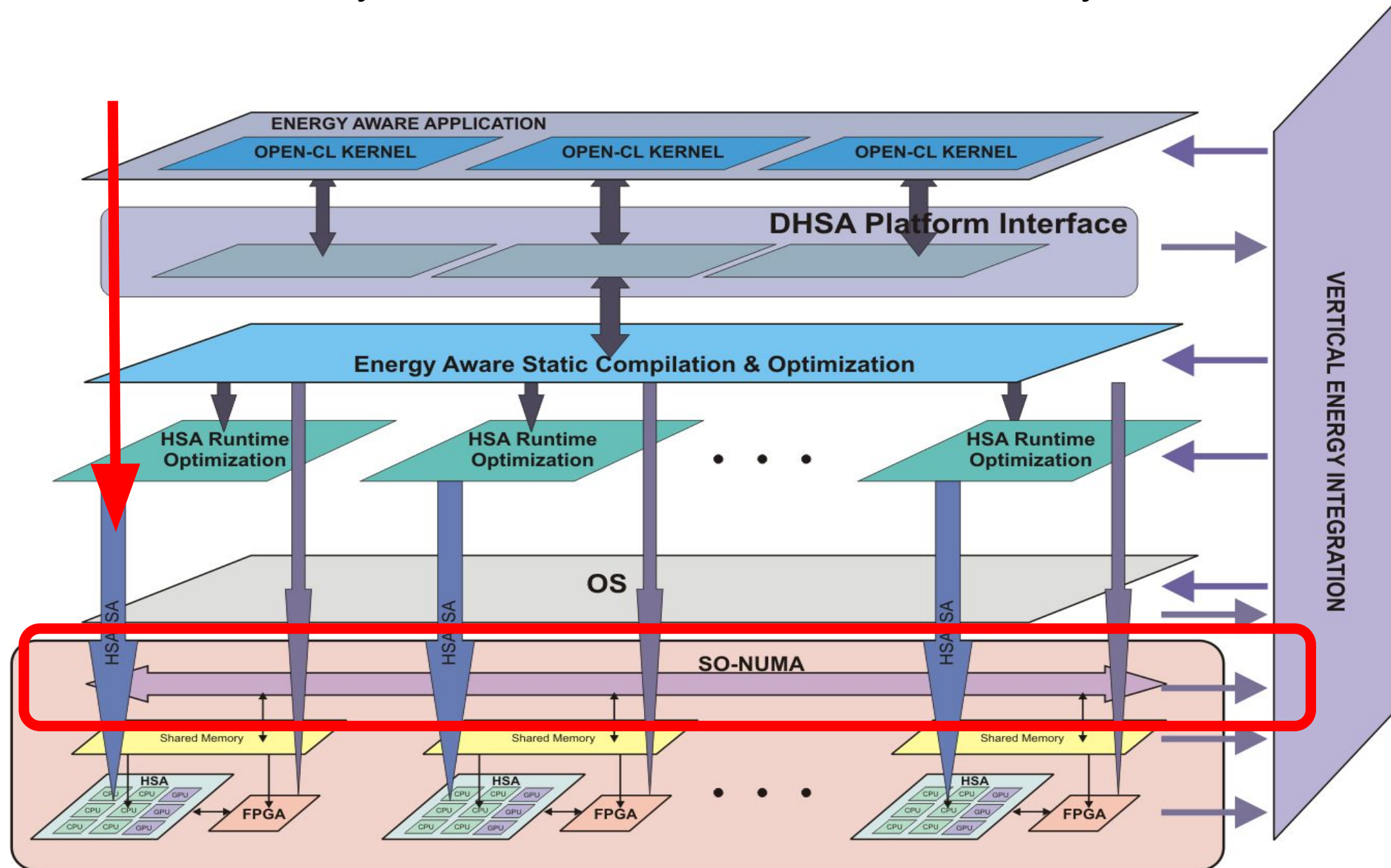
Added Value

- Cross stack energy APIs developed
- Multicore Energy Accounting
- GPGPU Energy Accounting
- VM Energy Accounting



Project organization

WP5- Low-Latency Data centre-Scale Distributed Memory





Ecole Poly-technique Fedeerale de Lausanne



WP 5 Low-Latency Datacenter-Scale Distributed Memory

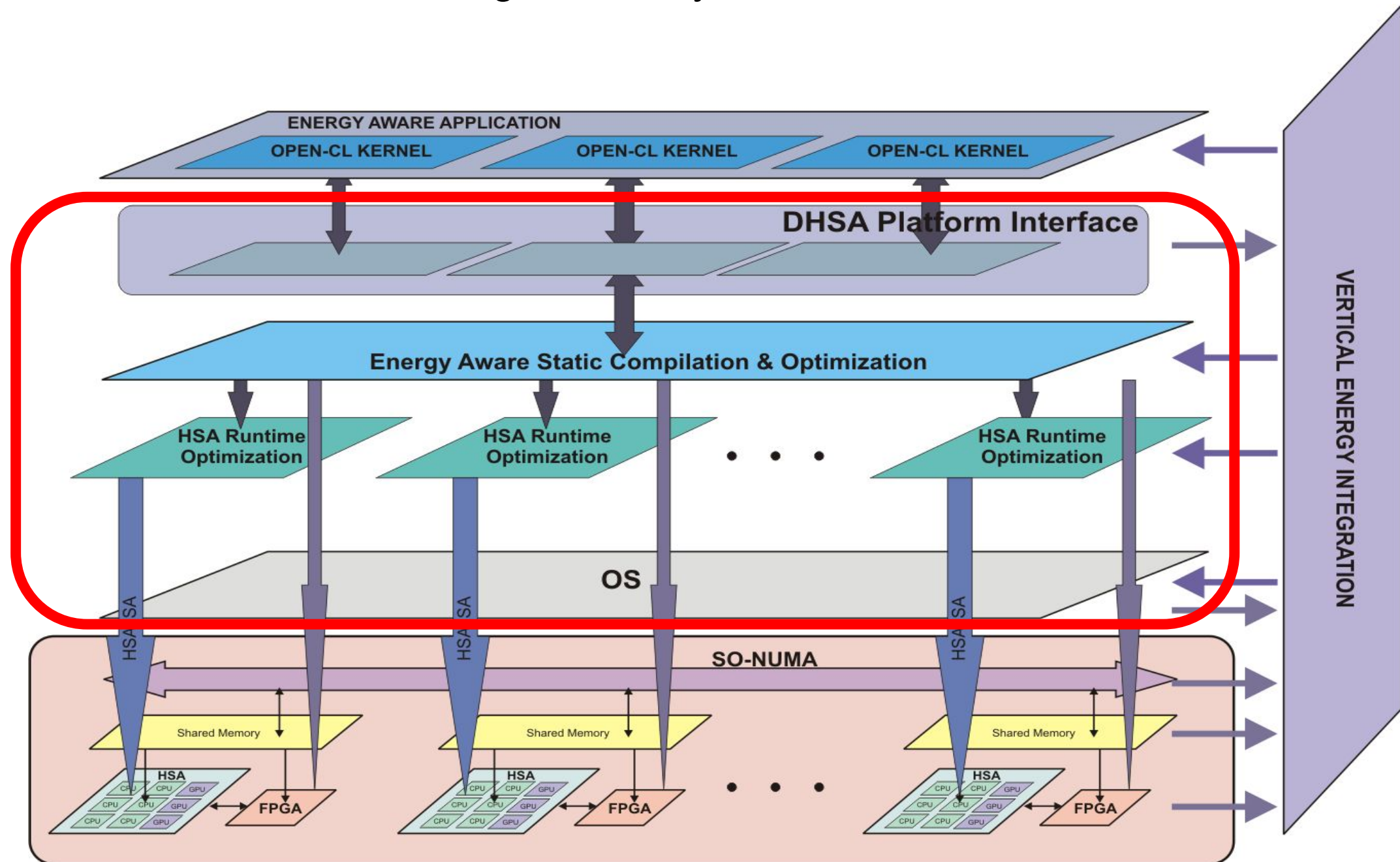
Added Value

- RPC enhanced SO-NUMA design
- SO-NUMA fabric (licensed)
- In memory accelerator design
- Architectural support for optimistic concurrency



Project organization

WP6- Distributed Heterogeneous System Architecture



INRIA & AMD

WP 6 Distributed Heterogeneous Systems Architecture

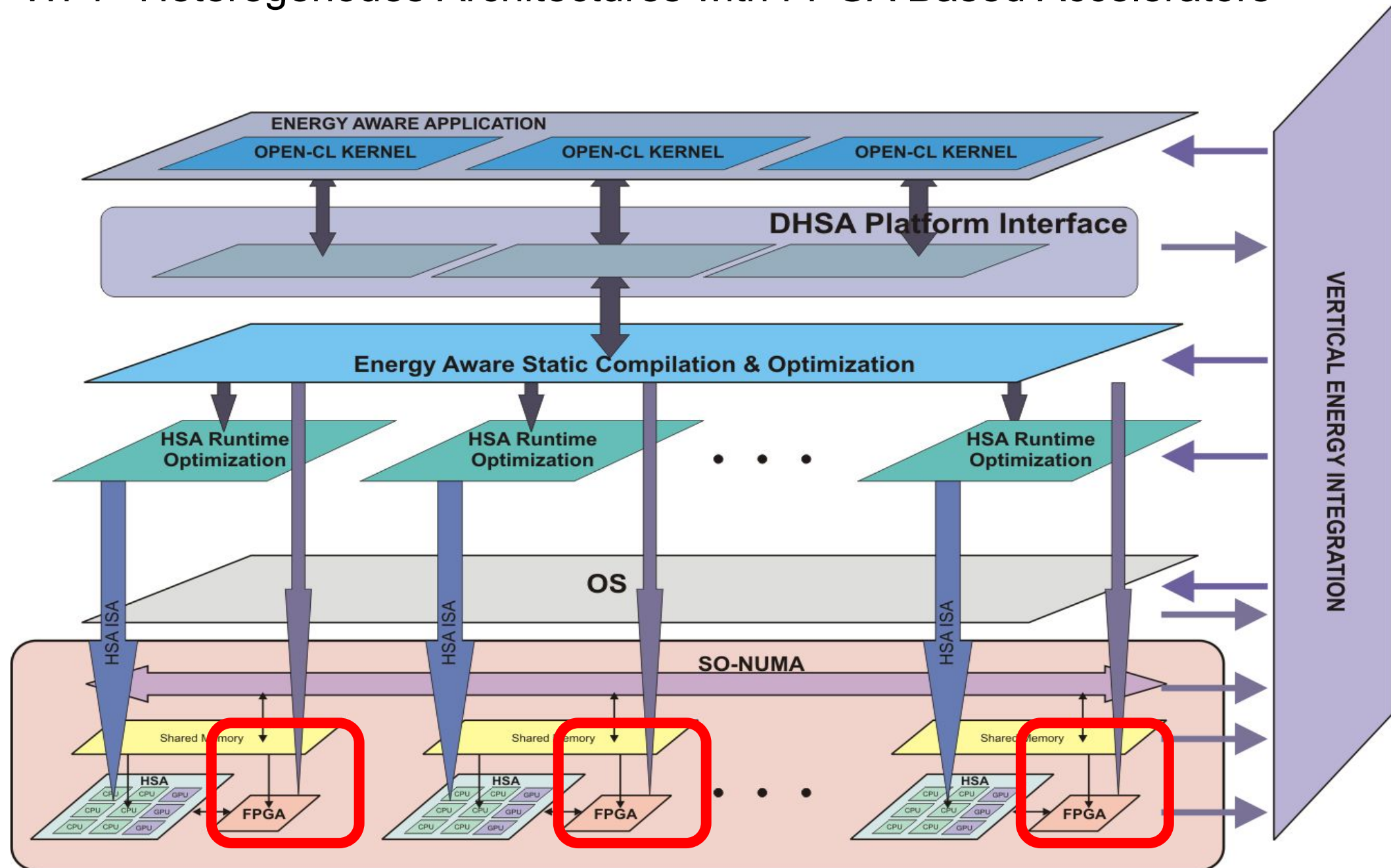
Added Value

- AMD Founding member of HSA Foundation
- HSA platform enhanced with power profiling API
- OpenCL energy measurement extension
- DHSA runtime proposal



Project organization

WP7- Heterogeneous Architectures with FPGA Based Accelerators



WP 7 Heterogeneous Architectures with FPGA-based Application Accelerators

Added Value

- FPGA Energy Accounting
- FPGA benchmark characterisation
- FPGA HSA integration prototyping

Scientific Results

This year

- 9 papers
(5 published, 4 under submission)

1. Towards Collaborative Performance Tuning of Algorithmic Skeletons (HLPGPU'16)
2. Iterative Compilation on Mobile Devices. (ADAPT'16)
3. Autotuning OpenCL Workgroup Size for Stencil Patterns (ADAPT'16)
4. Power Capping: What Works, What Does Not (ICPADS'15)
5. Performance and Fault Tolerance of Preconditioned Iterative Solvers on Low-Power ARM Architectures (ERPP'15)
6. Configurable FPGA Architecture for Hardware-Software Merge Sorting (MIXDES'16)*
7. Highly Parametrizable FPGA Architecture for K-Means Clustering (HEART'16)*
8. Minimizing the cost of iterative compilation with active learning (PACT'16)*
9. SABRes: Fast atomic remote reads for rack-scale in-memory computing (MICRO'16)

*



Valorisation

Prototypes

- Open source big data API
- Open source HSAIL auto-tuning
- Open source energy accounting tool
- HSA/OpenCL energy extension

Partnerships

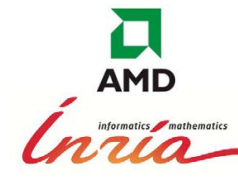
- ARM funded PhD
- INRIA part of HSA Foundation



Conclusion

Energy becoming first class concern

- Collaborative interdisciplinary problem
 - Needs vertical integration
 - First to compute energy cost of DC computation
-
- 1st year deliverables complete
 - 9 papers
 - 4 prototypes



Energy Reduction

100%

80%

60%

40%

20%

0%

-20%

Backprop.K2

bfs.K1

hotspot.K1

leukocyte.K2

leukocyte.K3

kmeans.K2

lud.K2

lud.K3

leukocyte.K1

Backprop.K1

kmeans.K1

lud.K1

lud.K4

Geo-mean



Outline

- Scientific background
- Key challenges and potential impact of the project
- Presentation of each partner of the consortium with their specificity and added value
- Main scientific results, dissemination and other output
- Sustainability/Valorisation

