Welcome!

European Exascale Processor & Memory Node Design

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 671578

Paving the way towards a highly energy-efficient and highly integrated compute node for the Exascale revolution: the ExaNoDe approach.

Kevin Pouget

Virtualization Engineer
Virtual Open Systems (ExaNoDe partner)

August 30th, 2017

Disclaimer: This presentation does not represent the opinion of the EC and the EC is not responsible for any use that might be made of information appearing herein.

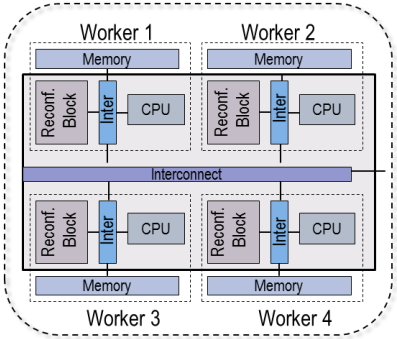
ExaNoDe technological context

ExaNeSt

Interconnect
Storage

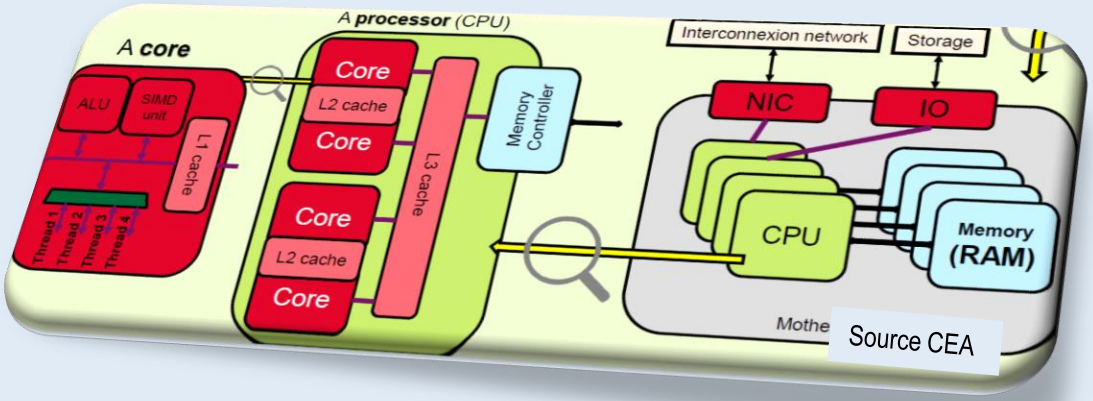
ECOSCALE

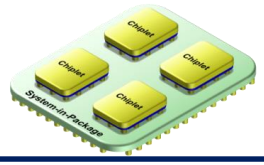
FPGA acceleration



ExaNoDe

Compute node
(HW + SW)





EURO SERVER

Architecture

(ended in Jan. 17)

ExaNoDe Project Implementation

Start: ***October 1st, 2015***
Duration: ***3 years***

Coordinator

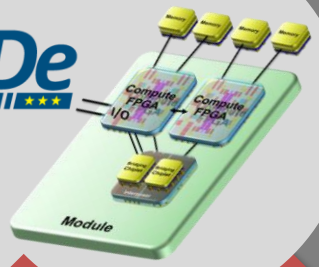
CONSORTIUM

IMPLEMENTATION

PROTOTYPE

ExaNoDe

8.6 M€



WP1 : Management

WP2 : Co-Design for Exa-scale HPC systems

WP3 : Enablement of Software Compute Node

WP4 : Compute node design and manufacture

WP5 : System Integration & Evaluation

WP6 : Dissemination and Exploitation

Mini-applications-Design guidelines

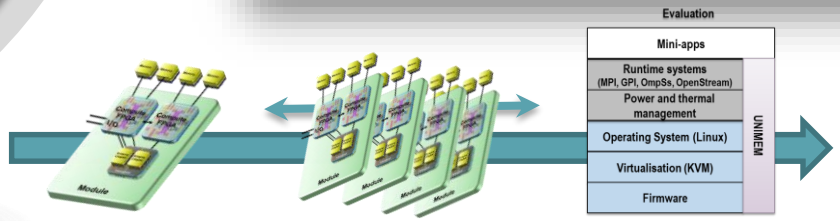
SW eco-system

HW

Measure

Evaluation

Assessment



Compute Node | HW Prototype | SW Stack

ExaNoDe at a glance

Key technologies for compute nodes towards a future **Exascale capability**

	System Architecture	Silicon Integration	Software

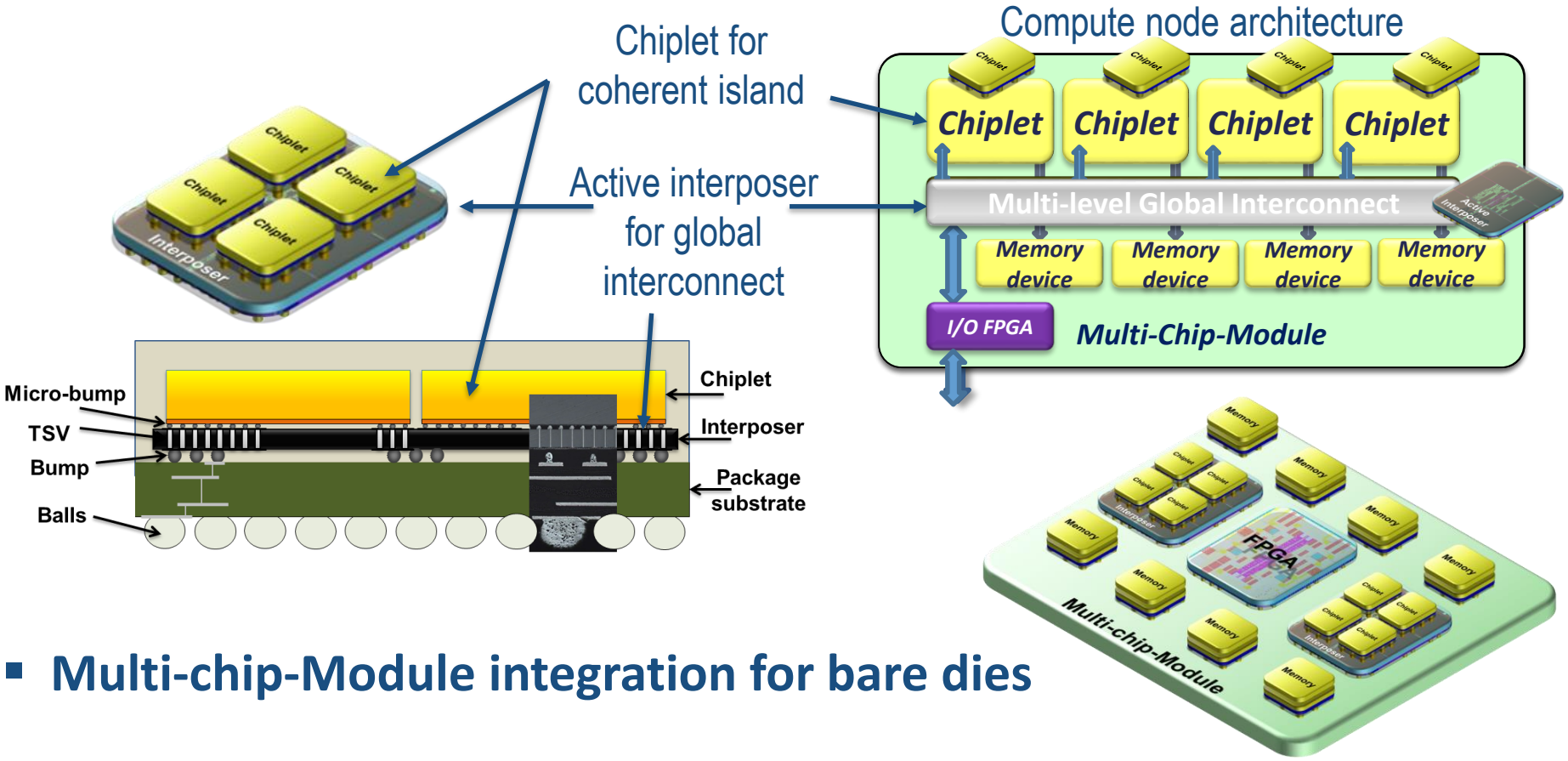
Key technologies	• ARMv8
	• Coherent islands
	• Global Address Space
		• 3D Integration:
		  • Chiplet
		  • Active Interposer
		• Multi-Chip-Module:
		  • FPGA, Memory
			• FW, OS
			• Virtualization
			• Programming models
			• Runtimes
			• Mini-apps

Exascale requirements	👍 Energy efficiency 👍 Scalability	👍 Design/manufacturing costs 👍 Heterogeneity/Specialization	👍 Co-design 👍 Scalability

All combined in an integrated prototype

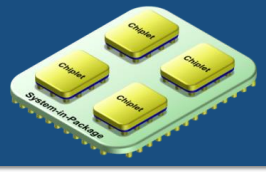
ExaNoDe Integration Approach

- 3D integration with active silicon interposer and chiplets

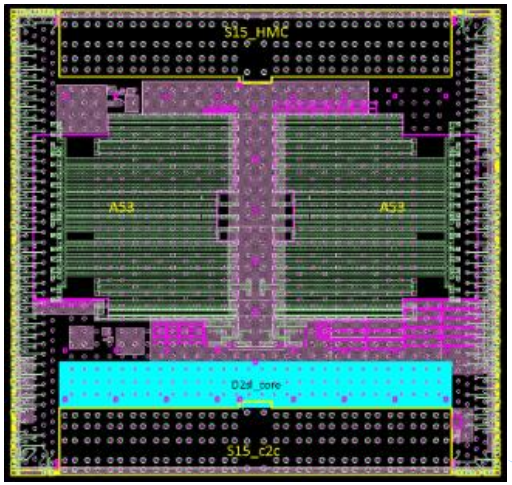


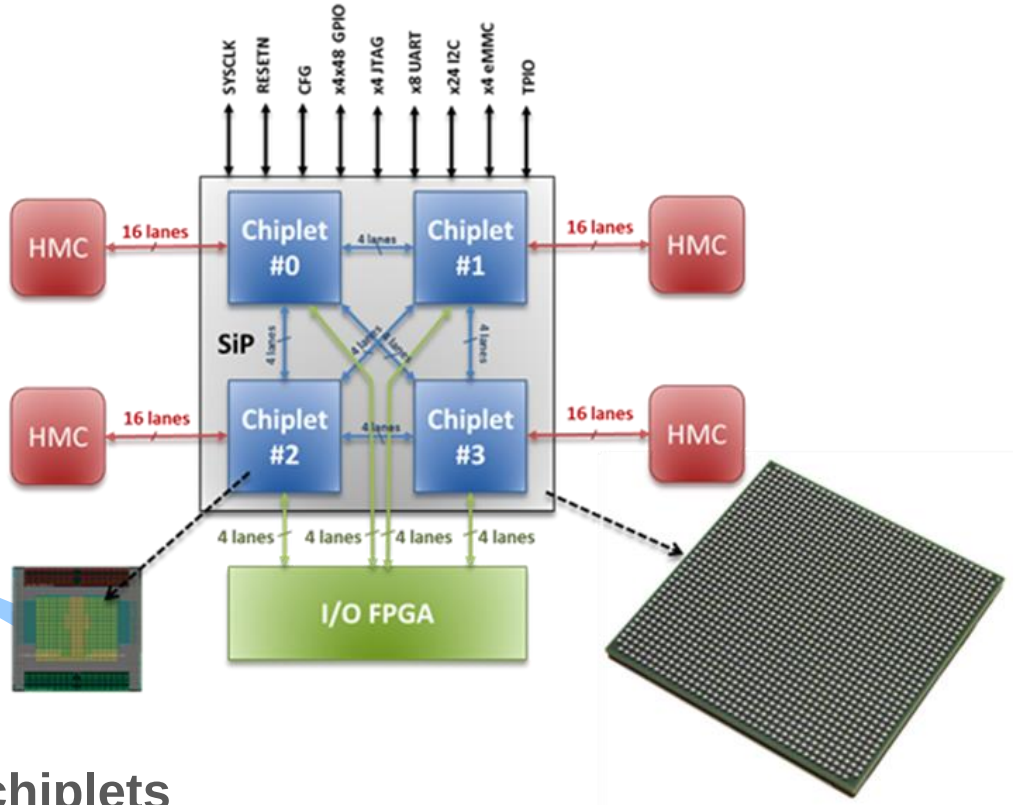
- Multi-chip-Module integration for bare dies

EUROSERVER: Package



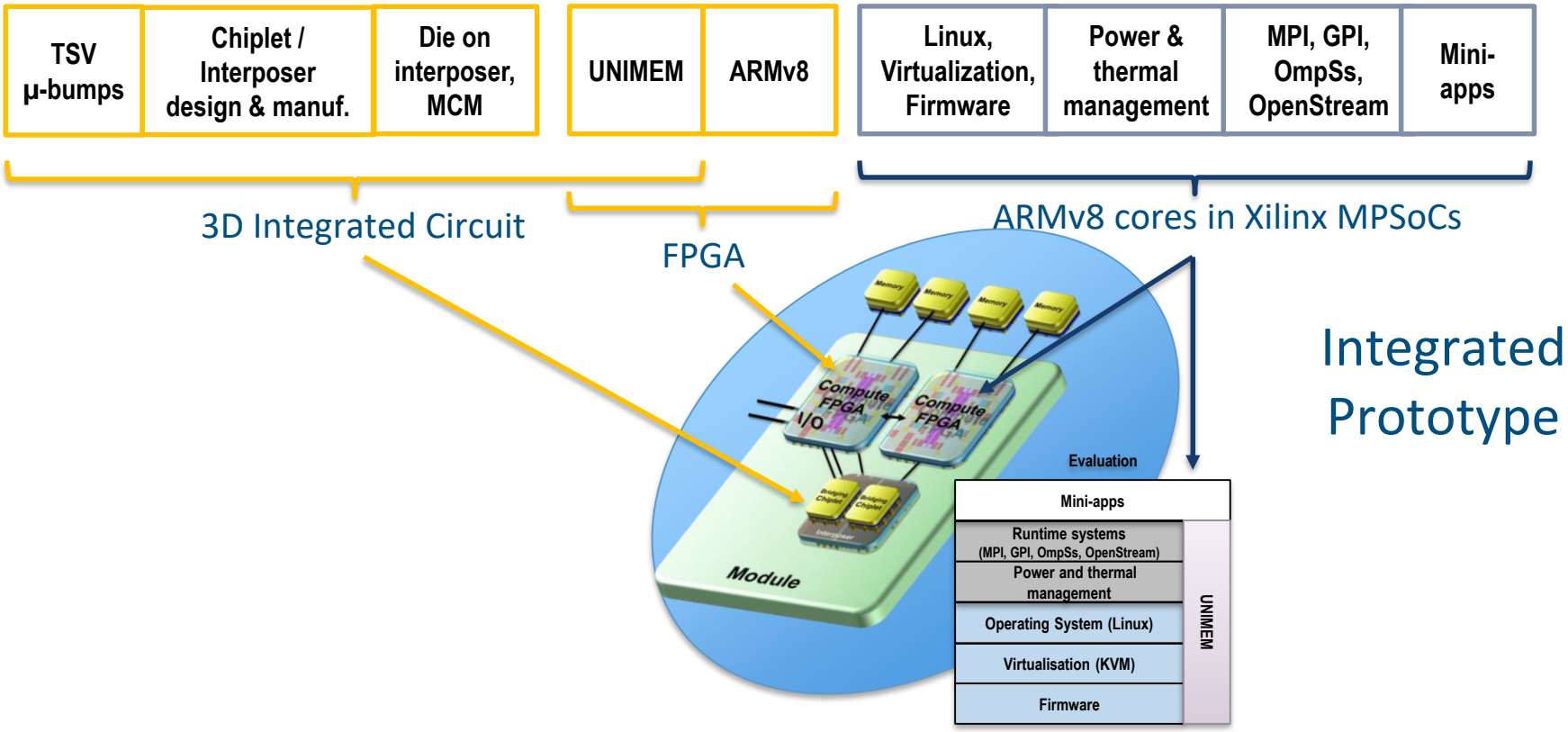
- **Multi-Core System with four chiplets in a 40x40 package**





- A package contains 4 chiplets
- Each chiplet contains 2 quad-core ARMv8 A53
- System of 32 A53 cores in a Package

ExaNoDe Hardware Prototype

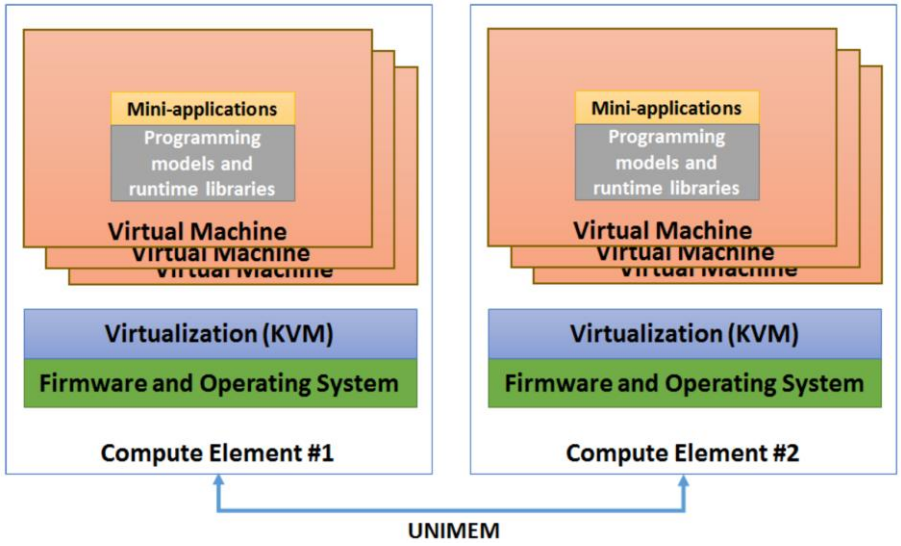


Integrated Prototype

ExaNoDe SW objectives

- **Mini-apps for co-design process**
  - Select of HPC applications to co-design the ExaNoDe architecture.
- **Software infrastructure**
  - Deploy a software ecosystem for the ARM-based compute node …
  - in conjunction with the UNIMEM system architecture.

ExaNoDe Software Architecture



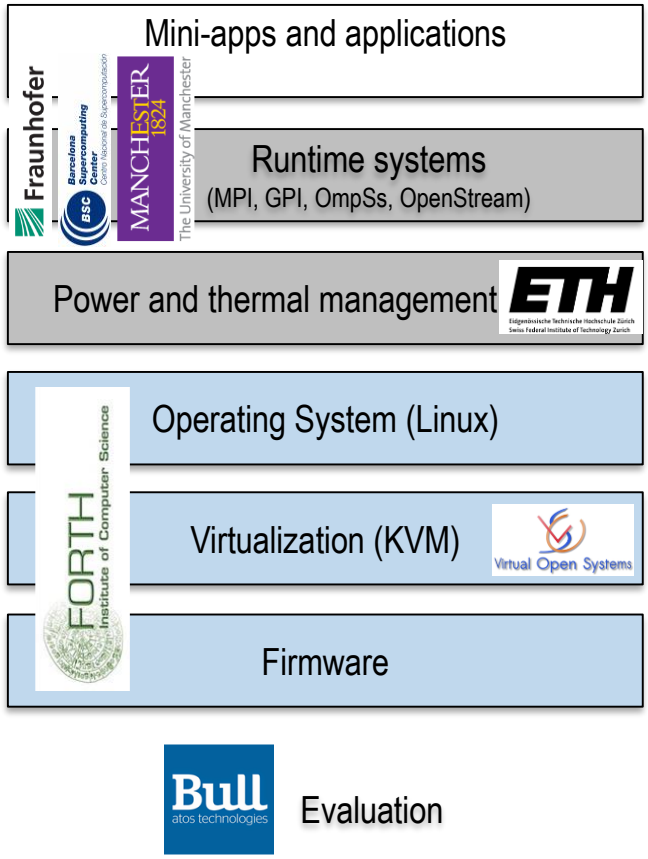
- **Evaluation**
  - Analyse and compare ExaNoDe architecture.

SW stack: deployment support

- **Deliver firmware and Numa-aware OS**
  - For multi-board and ExaNoDe prototypes

  (Unimem data movement, memory protection and integration with peripheral devices; OS interface for light RDMA operations)
- **Provide virtualization layer**

  (next slide)
- **Support programming models:**
  - Enable portable exploitation of UNIMEM (MPI, GPI, OmpSs, OpenStream)
- **Evaluate UNIMEM architecture**
  - Latency, bandwidth, memory footprint, CPU usage, …



SW stack: Virtualization Layer

Virtual Open Systems

- **For deployment ease,**
  compatibility,
  efficiency of resource usage

- **For an increased flexibility and reliability**
  - **Snapshot:** save now, restore later, maybe elsewhere
  - **Checkpoint**: periodic and incremental, to cope with HW or SW failures

- **For performance KVM + paravirtualization**
  - **API Remoting:** use accelerator APIs inside VMs
    (UNIMEM atomics and RDMA, MPI, OpenCL …)

ExaNoDe: Conclusion

- **ExaNoDe concept:**
  - **Architecture**: Many simple cores instead of few complex ones
  - **Integration**: Many simple heterogeneous chiplets instead of few complex System-on-Chip
  - **Compute node interconnect**: Active silicon interposer