



# **MANGO: Exploring manycore architectures for QoS-aware HPC**

PEGPUM Workshop, January 24, 2017, Stockholm

**Alessandro Cilaro**

CeRICT/University of Naples Federico II

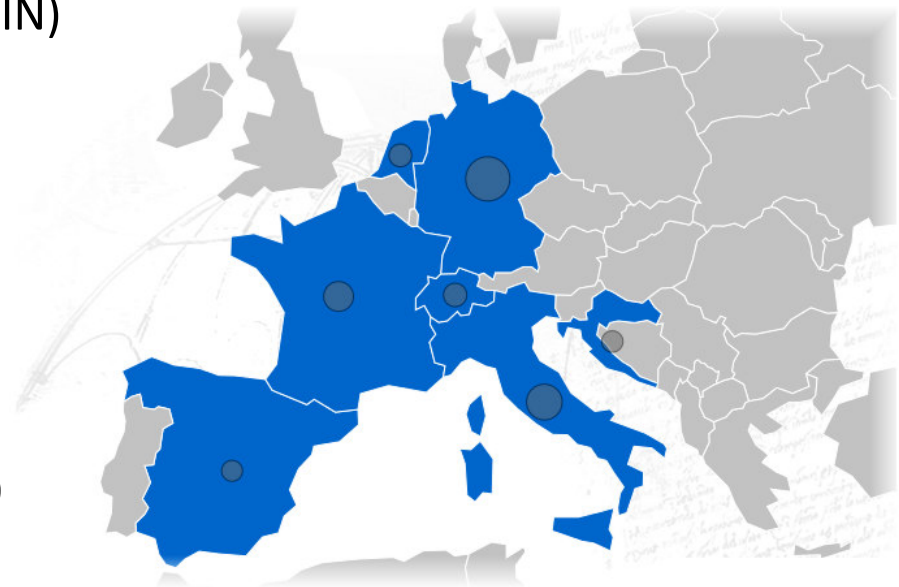
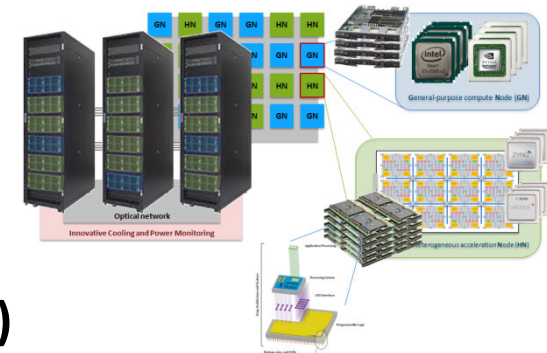


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 671668

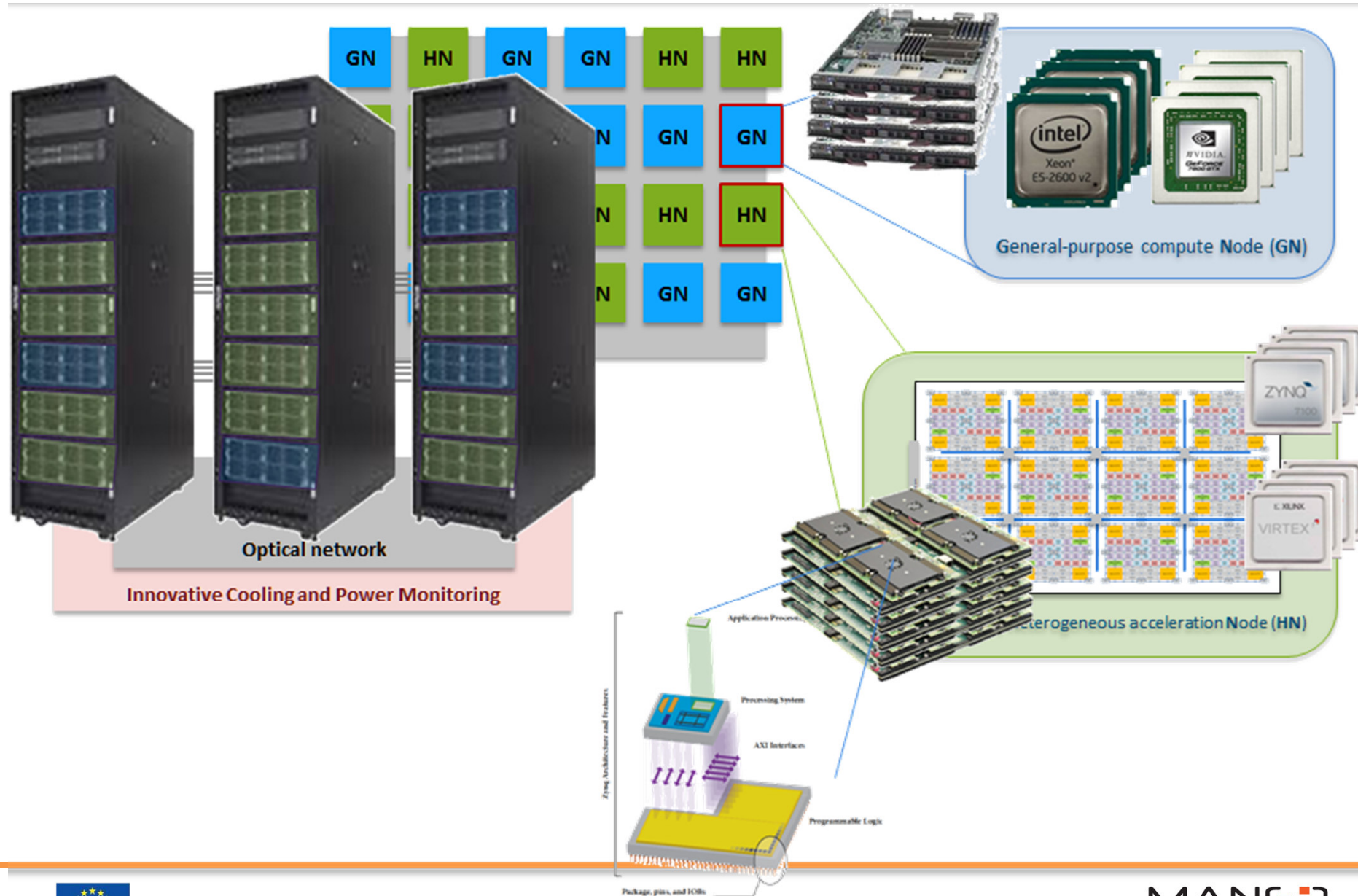


# The MANGO project and consortium

- **MANGO: *exploring Manycore Architectures for Next-GeneratiOn HPC systems***
- started Oct. 2015, budget  $\approx$  6M€
- One of the 19 large projects selected for exploring innovative HPC solutions (H2020-FETHPC 2014 call)
- Universitat Politècnica de València (SPAIN)
- CeRICT / University of Naples (ITALY)
- Politecnico di Milano (ITALY)
- Zagreb University (CROATIA)
- Pro Design GmbH (GERMANY)
- Thales Communication & Security (FRANCE)
- EPFL (SWITZERLAND)
- Philips Medical Systems (NETHERLAND)
- Eaton Industries SAS (FRANCE)

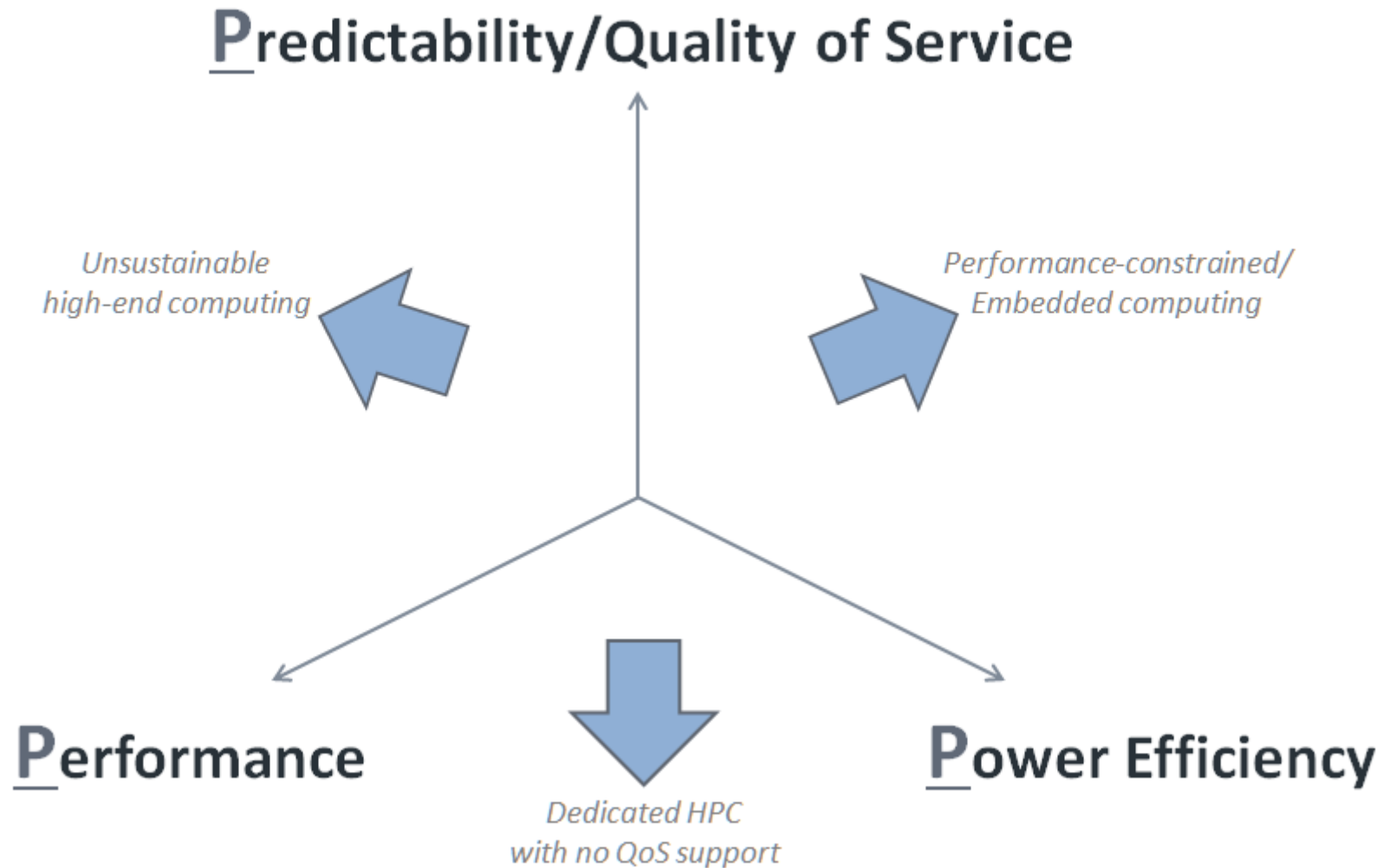


# MANGO: the big picture



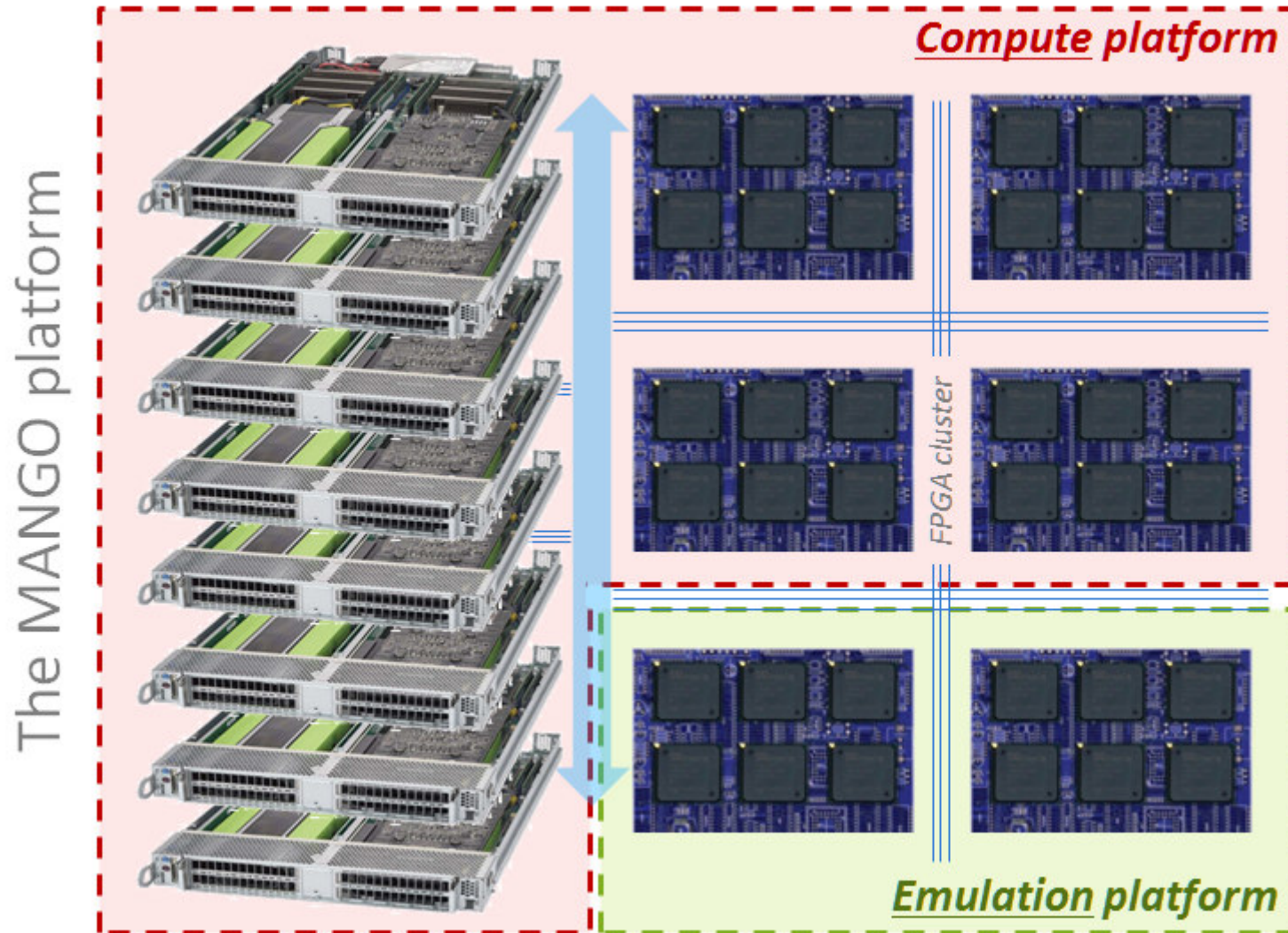
# MANGO: exploring the PPP design space

---

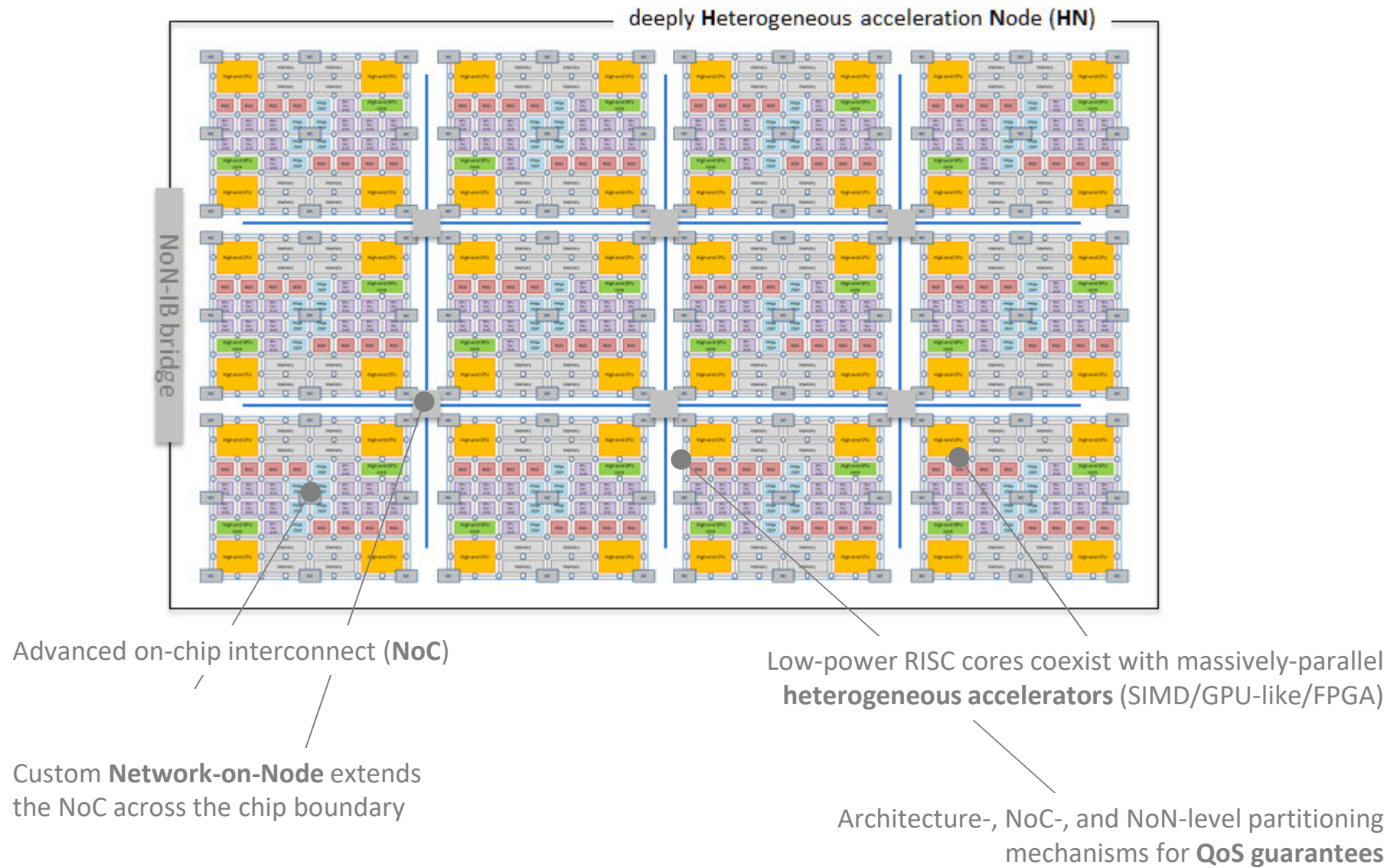




# The MANGO platform

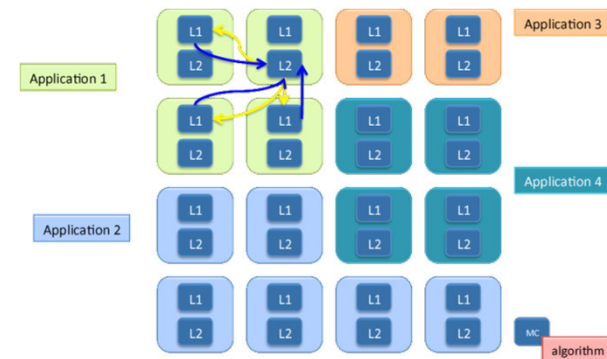


# Deeply Heterogeneous acceleration Node (HN)



# Heterogeneous Nodes: interconnect

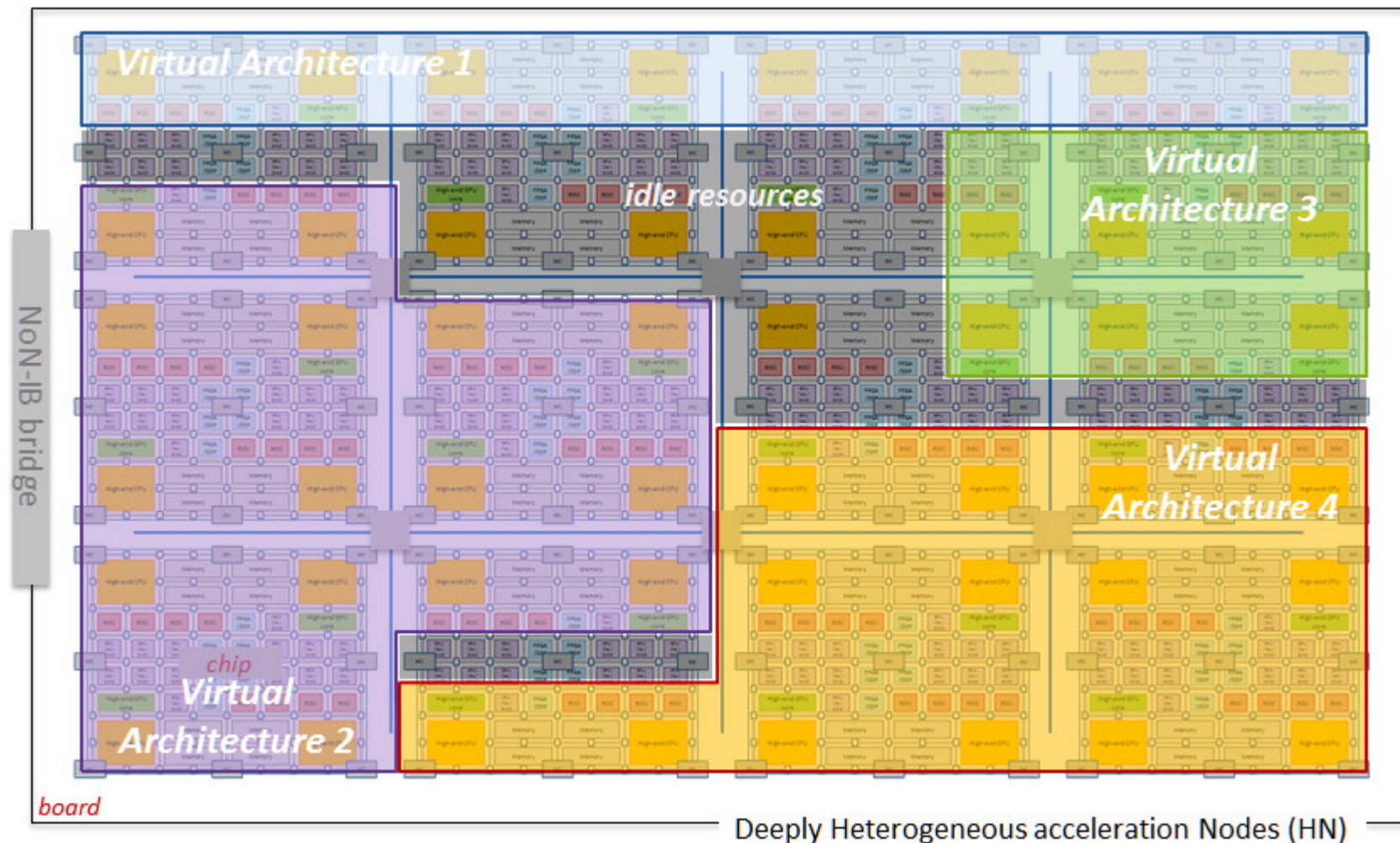
- Builds on an advanced NoC developed by UPV
  - **PEAK** (*Partition-Enabled Architecture for Kilocores*) architecture
- Provides:
  - Built-in partitionability
  - On-the-fly reconfigurability
  - Co-development of interconnect and memory hierarchy
  - QoS capabilities
  - Adaptive routing





# Heterogeneous Nodes: virtualization support

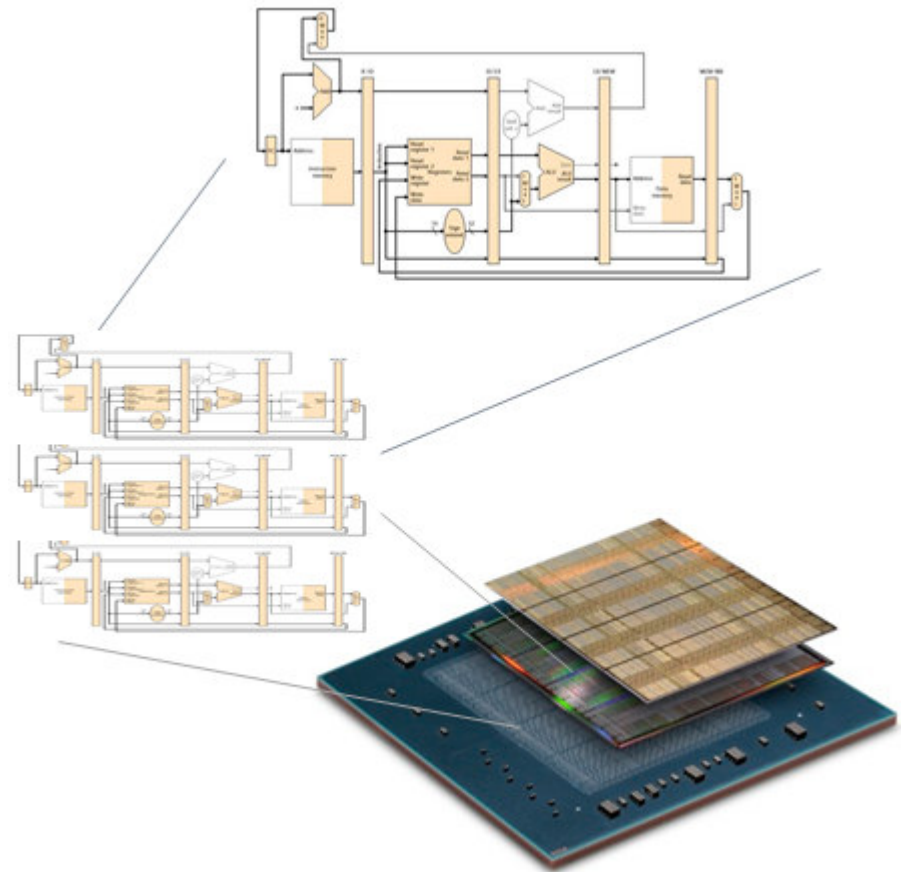
- HN: fine-grained partitioning capabilities





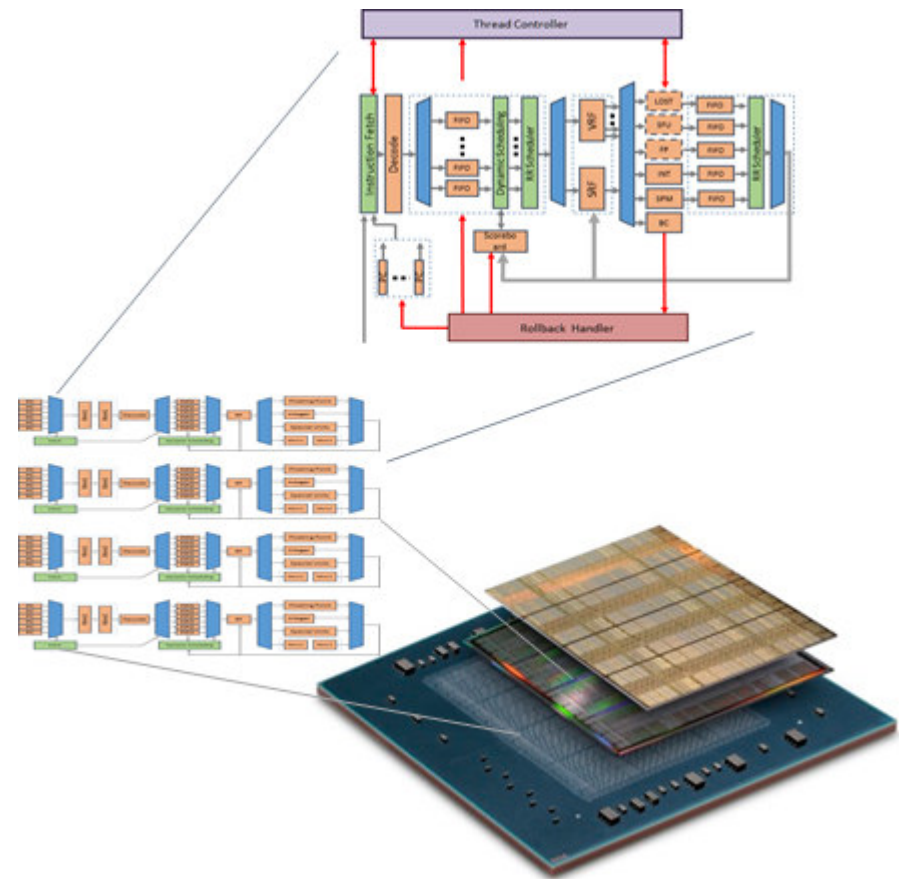
# The MANGO **RISC** general-purpose core

- MIPS architecture
- Compatible with GCC compiler
- Out-of-order datapath planned for future development
- Fully integrated in PEAK



# The MANGO GPU-like accelerator

- $\nu+$  (or **Nu+**) GPU-like core
- Open source, software-programmable HDL design
- Fits both MANGO perspectives:
  - FPGA-based **emulation**
  - FPGA-based **computation**: provides a parametrized, programmable *overlay*

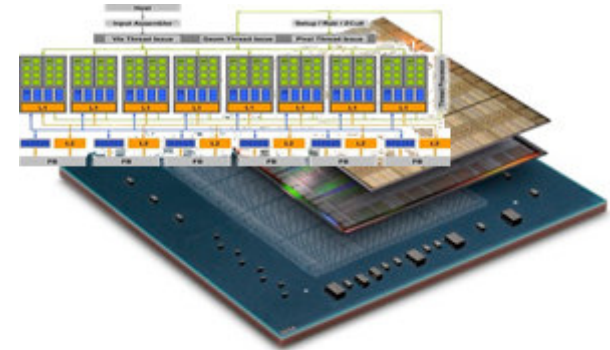


$\nu+$  (pronounce: nu:pls) GPU-like core

# The MANGO GPU-like accelerator

---

- Match recent trends
  - also including FPGA and SoC manufacturers
- Enable higher power-efficiency
- Provide an effective answer to programmability issues
  - support for high-level languages and models, like OpenCL
- MANGO: pursue *deep customization* of GPU-like cores
  - driven by applications
  - tailor architecture to specific workloads



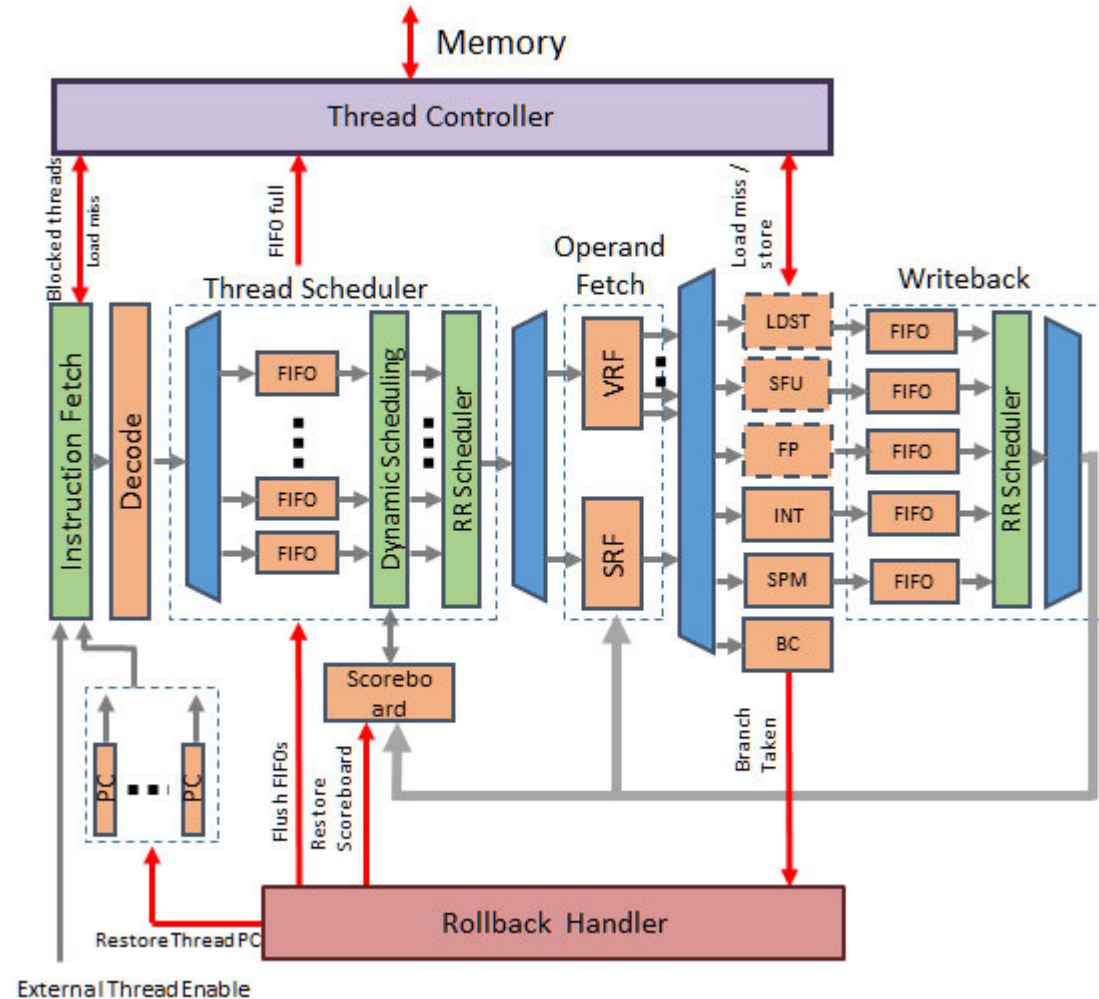
OpenCL





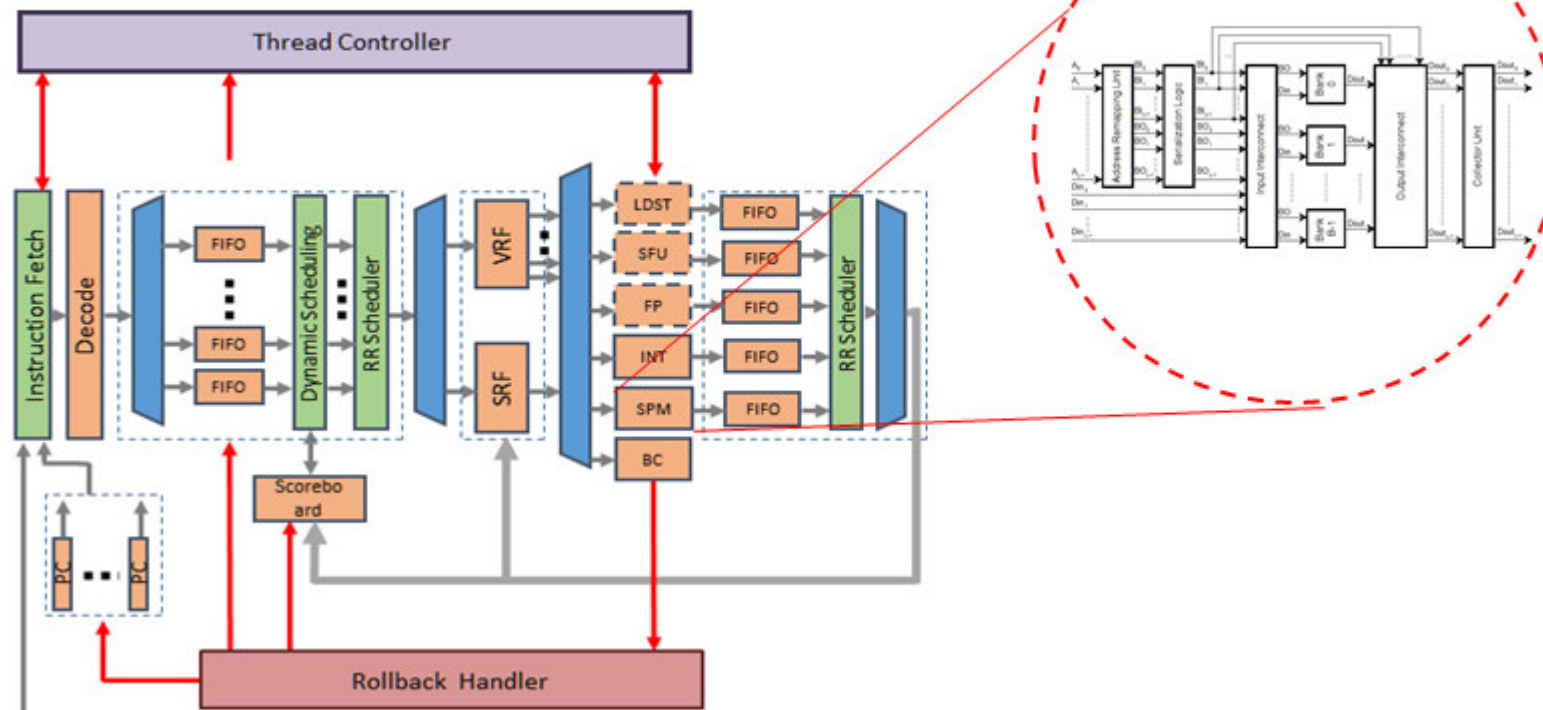
# Nu+ current microarchitecture

- You can configure:
  - Number of cores
  - Number of Threads
  - Number of hw lanes
  - Number of registers per Thread
  - Cache set-size
  - Number of ways
  - Number of 32-bit words in each line
  - SPM parameters:
    - Number/size of banks
    - Type of partitioning
  - Etc..
- Also developed an LLVM compiler backend and initial programming tools



# Nu+ Scratch-Pad Memory

- An example of application-driven customization



A. Cilaro, M. Gagliardi, C. Donnarumma, "A Configurable Shared Scratchpad Memory for GPU-like Processors", *Procs. of the International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, Springer, pp 3-14, 2016

# Nu+ Scratch-Pad Memory: evaluation

- kernels from benchmark suites (PolyBench)
- rewrote each of those kernels to increase the kernel memory access parallelism
- Used first an ad-hoc cycle-accurate emulator
- repeated the experiment for different *remapping functions* identified for the specific kernel as well as for a variable number of banks

Cyclic mapping				Block mapping				Generalized Cyclic mapping			
Bank0	Bank1	Bank2	Bank3	Bank0	Bank1	Bank2	Bank3	Bank0	Bank1	Bank2	Bank3
0x00	0x04	0x08	0x0c	0x00	0x10	0x20	0x30	0x00	0x04	0x08	0x0c
0x10	0x14	0x18	0x1c	0x04	0x14	0x24	0x34	0x1c	0x10	0x14	0x18
0x20	0x24	0x28	0x2c	0x08	0x18	0x28	0x38	0x28	0x2c	0x20	0x24
0x30	0x34	0x38	0x3c	0x0c	0x1c	0x2c	0x3c	0x34	0x38	0x3c	0x30

A. Cilaro, M. Gagliardi, C. Donnarumma, "A Configurable Shared Scratchpad Memory for GPU-like Processors", *Procs. of the International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, Springer*, pp 3-14, 2016





# Nu+ Scratch-Pad Memory: evaluation

- Results:
  - Matrix Multiplication
  - 5 × 5 Mean Filter
- Number of conflicts got by varying:
  - Number of lanes
  - Number of banks
  - Mapping strategy

Lanes	Banks	Remapping factor				
		No Remap	1	2	4	8
4	16	262146	131072	262146	262146	262146
	32	262146	0	0	131072	262146
	64	262146	0	0	0	0
	128	262146	0	0	0	0
	256	131072	0	0	0	0
	512	0	0	0	0	0
	1024	0	0	0	0	0
8	16	183505	131073	183505	183505	183505
	32	183505	0	65536	131073	183505
	64	183505	0	0	0	65536
	128	183505	0	0	0	0
	256	131073	0	0	0	0
	512	65536	0	0	0	0
	1024	0	0	0	0	0

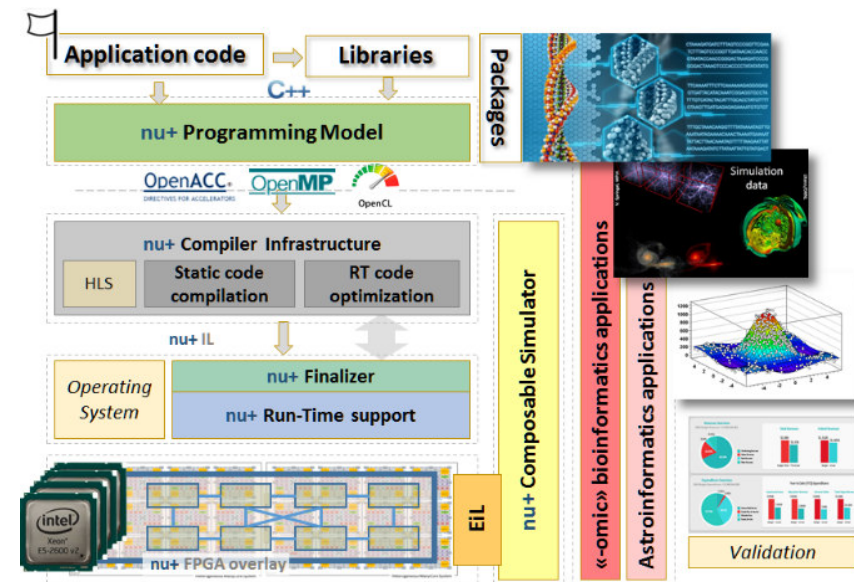
Lanes	Banks	Remapping factor				
		No Remap	1	2	4	8
16	16	109230	91756	109230	109230	109230
	32	109230	32768	65538	91756	109230
	64	109230	0	0	32768	65538
	128	109230	0	0	0	0
	256	91756	0	0	0	0
	512	65538	0	0	0	0
	1024	32768	0	0	0	0
32	16	61696	58256	61696	61696	61696
	32	59768	32769	45878	54615	59768
	64	59768	0	16384	32769	45878
	128	59768	0	0	0	16384
	256	54615	0	0	0	0
	512	45878	0	0	0	0
	1024	32769	0	0	0	0

A. Cilaro, M. Gagliardi, C. Donnarumma, "A Configurable Shared Scratchpad Memory for GPU-like Processors", *Procs. of the International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, Springer, pp 3-14, 2016



# The *Nu+* long-term objectives

- Develop a whole ecosystem of customizable FPGA-based *overlay* solutions
  - parameterized hardware cores
  - toolchain
  - software libraries
  - ambitious applications from the scientific computing and big data domains
- Ideally matches emerging compute technologies
  - TFLOPS-grade FPGAs + high-end CPUs
  - e.g. new Intel MCP solutions



# The *Nu+* GPU-like accelerator

---



integrated in the H2020-FETHPC-2014 European Project MANGO: *exploring Manycore Architectures for Next-GeneratiOn HPC systems*, project ID: 671668



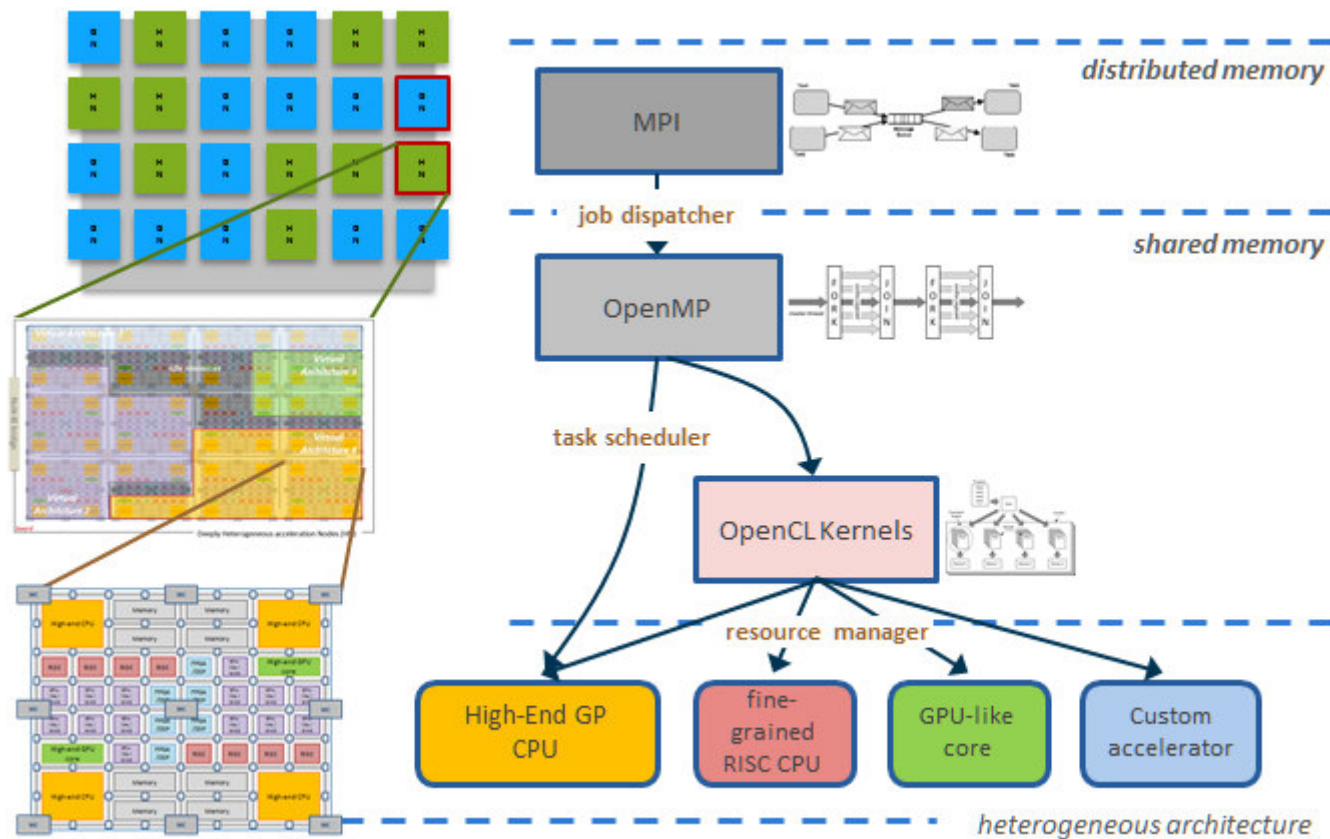
chosen to participate in the Intel "*Hardware Accelerator Research Program*" (Nov 2016)

- More details at <http://nuplus.hol.es> (*temporary address, will be changed*)
- Contacts: Alessandro Cilaro, [acilaro@unina.it](mailto:acilaro@unina.it)



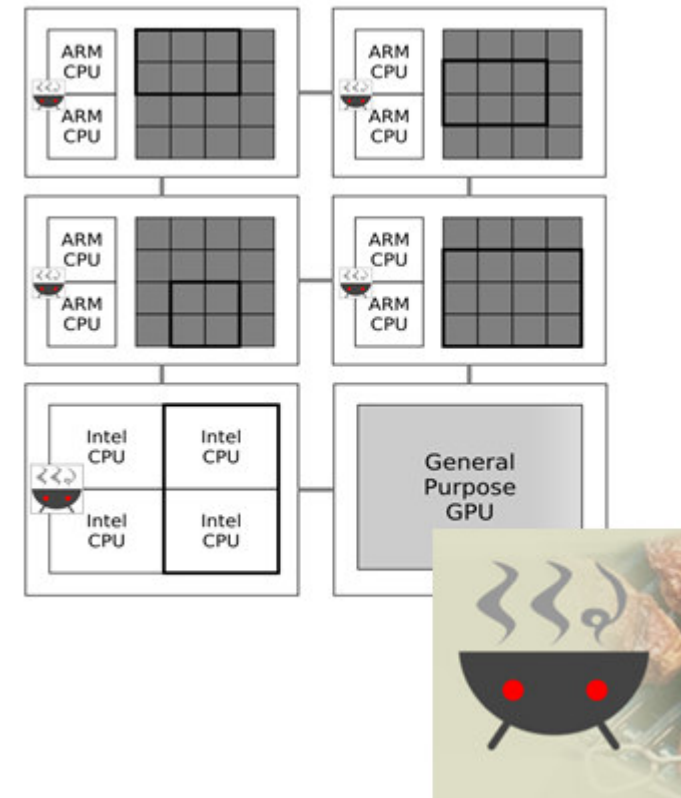


# Back to MANGO: Programming challenges



# Run-time power and thermal management

- Platform monitors and knobs
  - define which monitors and knobs to use, and what granularity
  - Hardware vs. Software
- Will take a proxy-based approach:
  - Identify performance counters
  - Determine models for each subsystem based on benchmarking
- Power models will be essential for the Run-Time Resource Manager
- MANGO will rely on Barbeque, developed by Politecnico di Milano
  - Extended to multi-node systems
  - Will rely on a hierarchical Master/Slave organization



# Run-time power and thermal management

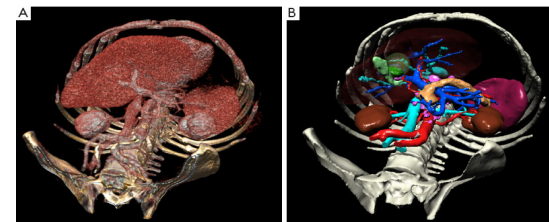
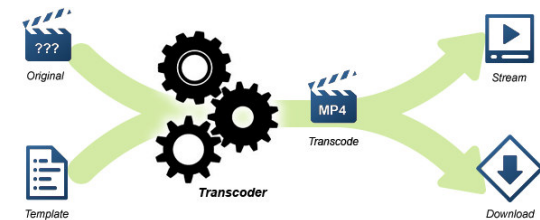
---

- Fine-grained monitoring of energy, temperature, and power in servers and racks (led by EPFL)
  - fast calculation of power/thermal figures of servers under highly dynamic workload behaviors
  - system-wide multi-objective optimization
  - hierarchical runtime manager, exploiting both OS and hypervisor levels to tune the system knobs
- Optimization of the mechanical cooling part
  - use two-phase cooling at rack level
  - novel passive thermosyphon (gravity-driven) cooling technology
  - microfluidic fuels cells combined with the liquid cooling technology

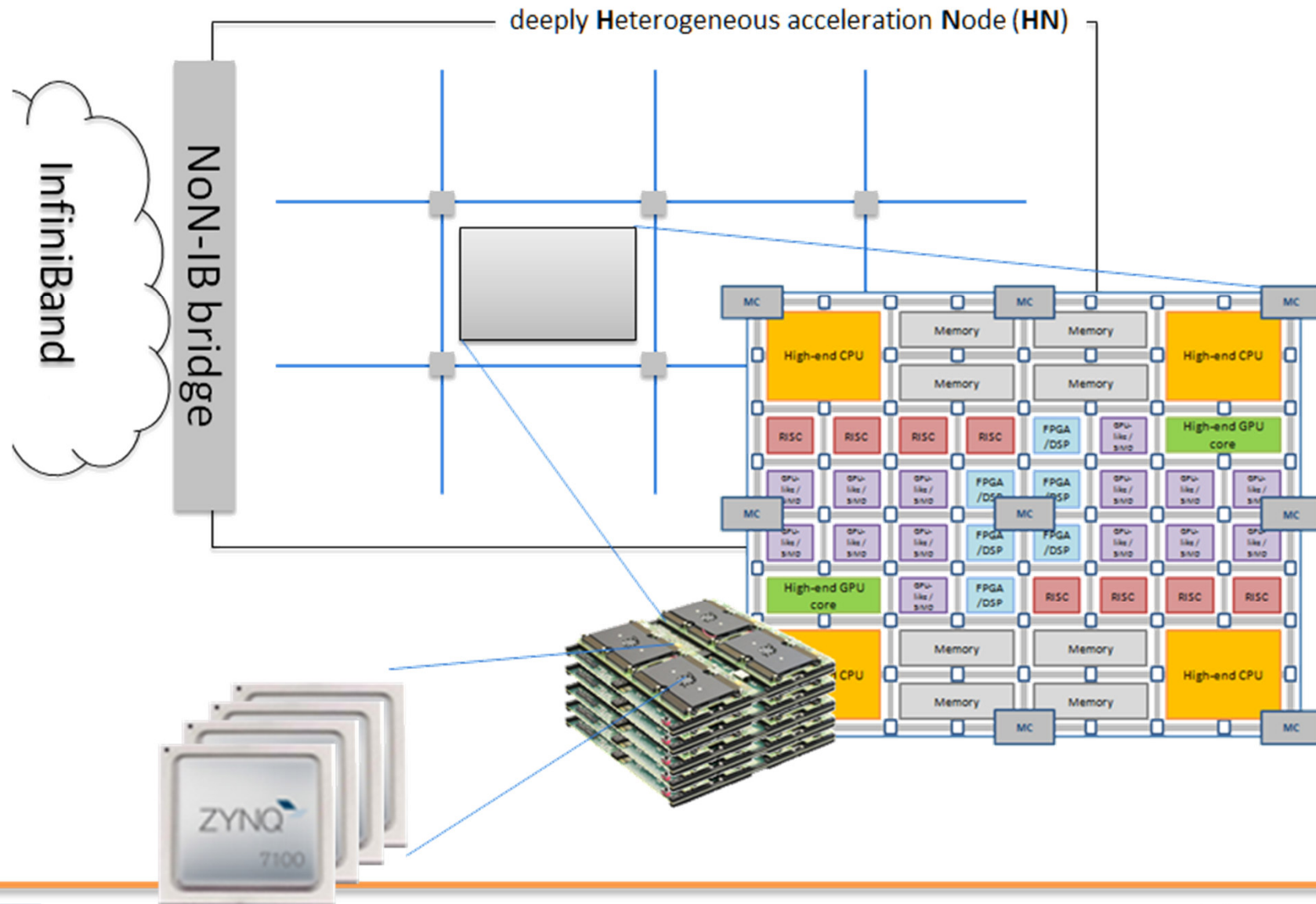


# MANGO applications

- Chose applications with stringent QoS and high-performance requirements:
  - *Video transcoding*
  - *Medical imaging*
  - *Sensor data processing*
  - *Security-related and cryptographic operations*



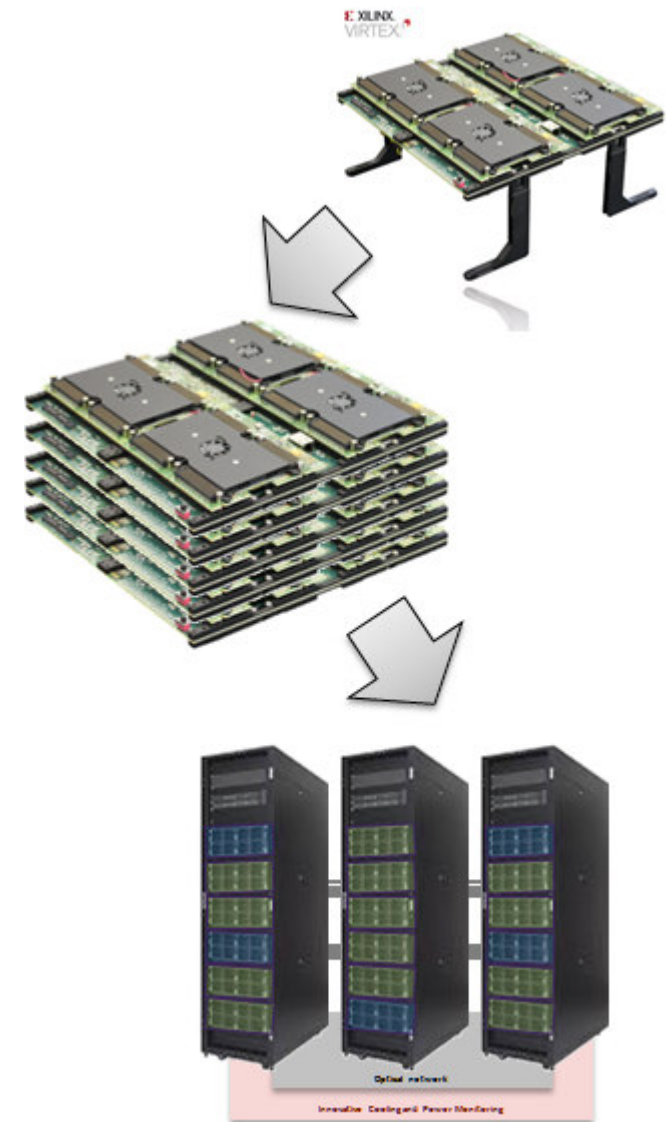
# MANGO Heterogeneous Nodes: prototyping





# The MANGO platform roadmap

- Phase 1 – *Stand-alone single-board emulator*
  - Pro-Design proFPGA quad V7 Prototyping system
- Phase 2 – *Dedicated chassis*
  - standard connectivity and form factor
- Phase 3 – *Rack assembly*
  - rack collecting up to 16 blades
  - high-end CPUs, e.g. Intel Xeon chips, and GPUs +
  - 64 HN nodes



# Stand-alone single-board emulator

---

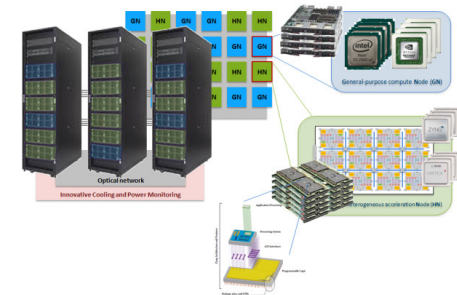
- Pro-Design proFPGA quad V7 Prototyping system
  - Scalable up to 48 M ASIC gates capacity on one board
  - Modular with up to 4 x Xilinx Virtex XC7V2000T FPGAs, or Zynq-7000, or memory modules
  - Up to 4336 signals for I/O and inter FPGA connection
  - Up to 32 individually adjustable voltage regions
  - Up to 1.8 Gbps/12.5 Gbps point to point speed



**prodesign**  
electronic gmbh

# Conclusions

- **MANGO: *exploring Manycore Architectures for Next-GeneratiOn HPC systems***



- Universitat Politècnica de València (SPAIN)
- CeRICT / University of Naples (ITALY)
- Politecnico di Milano (ITALY)
- Zagreb University (CROATIA)
- Pro Design GmbH (GERMANY)
- Thales Communication & Security (FRANCE)
- EPFL (SWITZERLAND)
- Philips Medical Systems (NETHERLAND)
- Eaton Industries SAS (FRANCE)

