



Funded by
the European Union



OpenCUBE: Open-Source Cloud-Based Services on EPI Systems

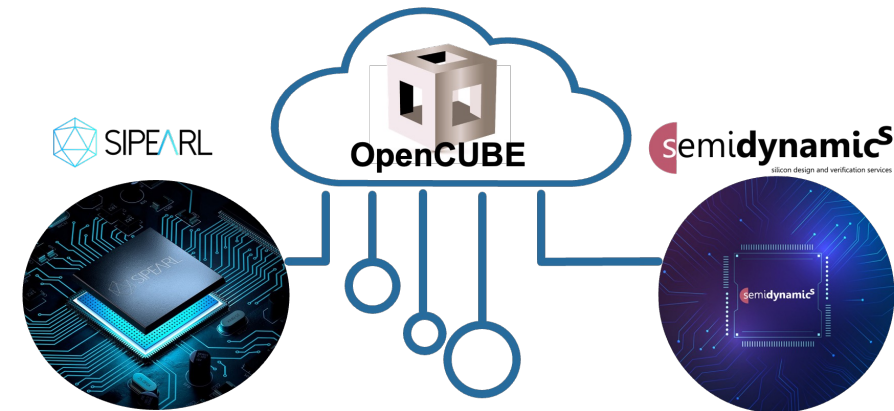
Ivy Peng, Assistant Professor, Department of Computer Science, KTH

ivybopeng@kth.se

Design, Implement and Validate a full cloud stack

The OpenCUBE project will provide open-source cloud services on a European Processor Initiative (EPI)-based computing blueprint

- Deploy a HW infrastructure hosting EPI processor + RISC-V accelerator
- Target the converged computing continuum
- Enable industrial & consumer cloud workloads
- Design with energy-efficiency at all levels





The OpenCUBE Consortium

Coordinator: KTH Royal Institute of Technology, Sweden

Start: 01/01/2023

Duration: 36 months

Partners:

- Semidynamics
- Sipearl
- ECMWF
- TUM
- HPE



Funded by
the European Union

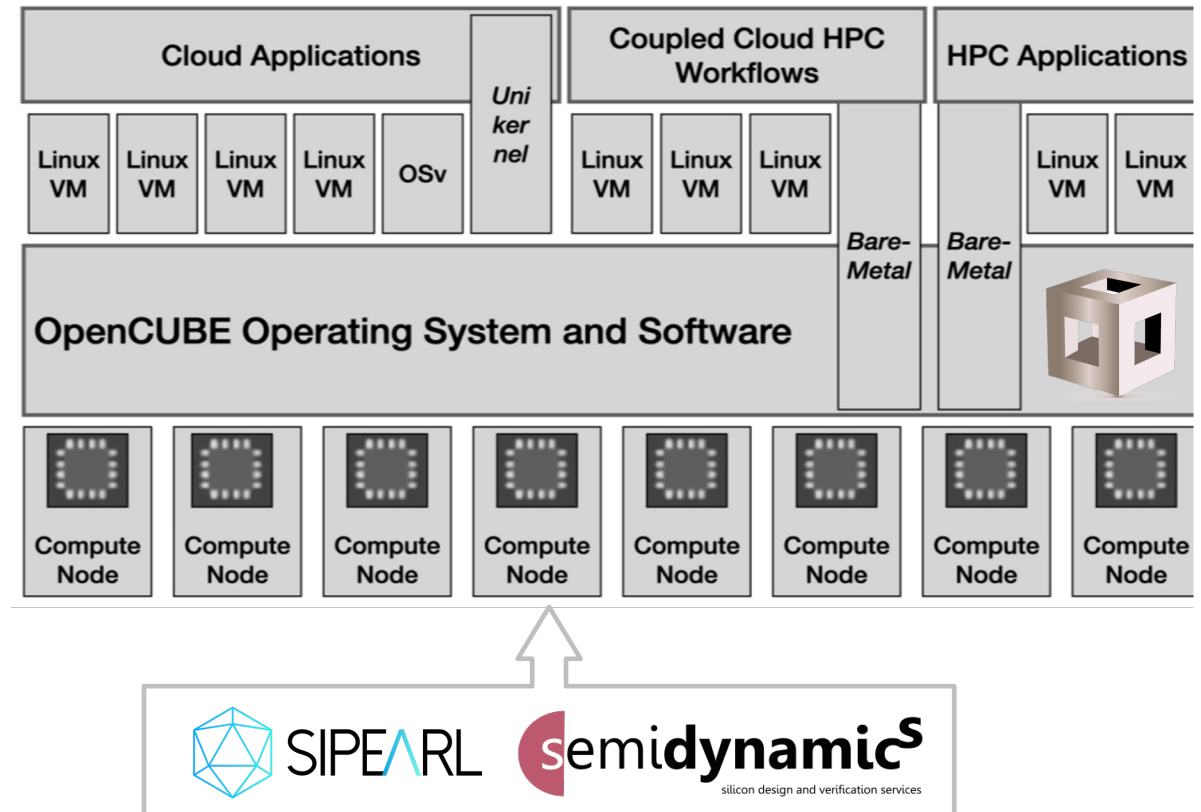


OpenCUBE

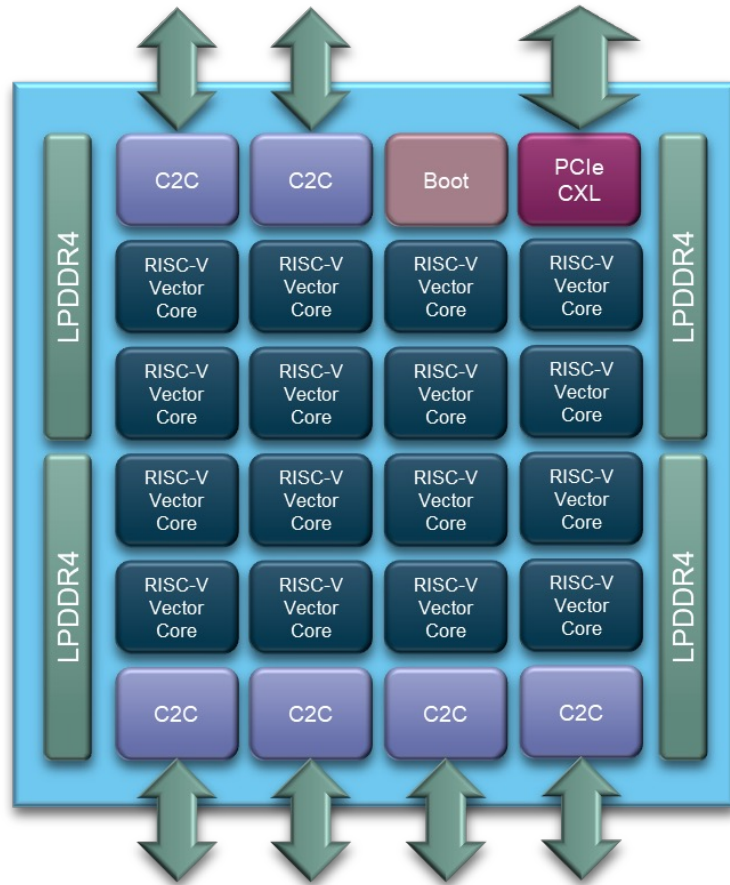
OpenCUBE HW Infrastructure

Lead by Sipearl

- Compute node: Sipearl Rhea processor + Semidynamics RISC-V accelerator
- High-performance ethernet network from HPE slingshot
- Fabric-attached Memory: for memory expansion and data staging between HPC and cloud partitions



The RISC-V accelerator expected in OpenCUBE



This would be an output from RISER project – great outreach!

Pilot SoC:

- PCIe connection with CXL.io/mem
- Linux capable

Atrevido Architecture:

- RISC-V64GCV + A, Half Precision, CMO and B extensions
- SV48/SV39 virtual memory modes
- Coherent memory model based on AMBA5 CHI
- Gazzillion misses™ technology for high bandwidth

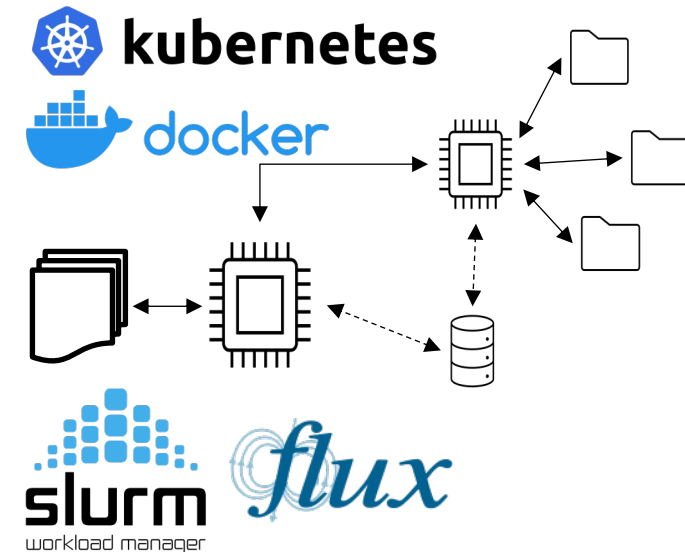
VPU architecture

- VLEN = 16384b
- DLEN = 512b
- Compliant with RISC-V Vector 1.0

OpenCUBE Full Software Stack - I

Use open-source component and open APIs

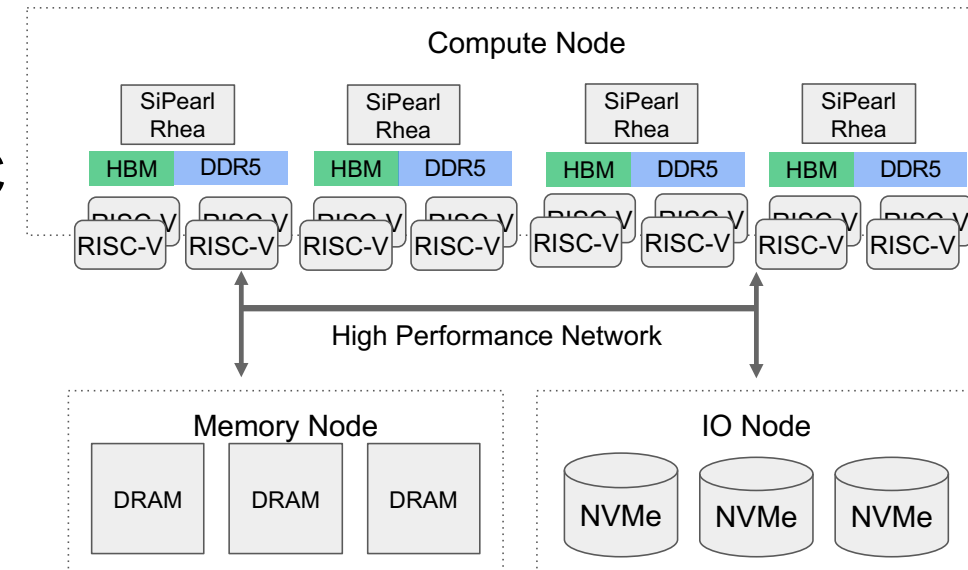
- **OS:** provide a base system supporting requirements of native cloud and HPC
- **Resource management and scheduling:** exploiting open-source framework like Kubernetes and extending for ARM-based and RISC-V hardware resources
- **Performance tools:** extension to open-source efforts with hardware performance counter, monitoring API, and system monitoring interface



OpenCUBE Full Software Stack - II

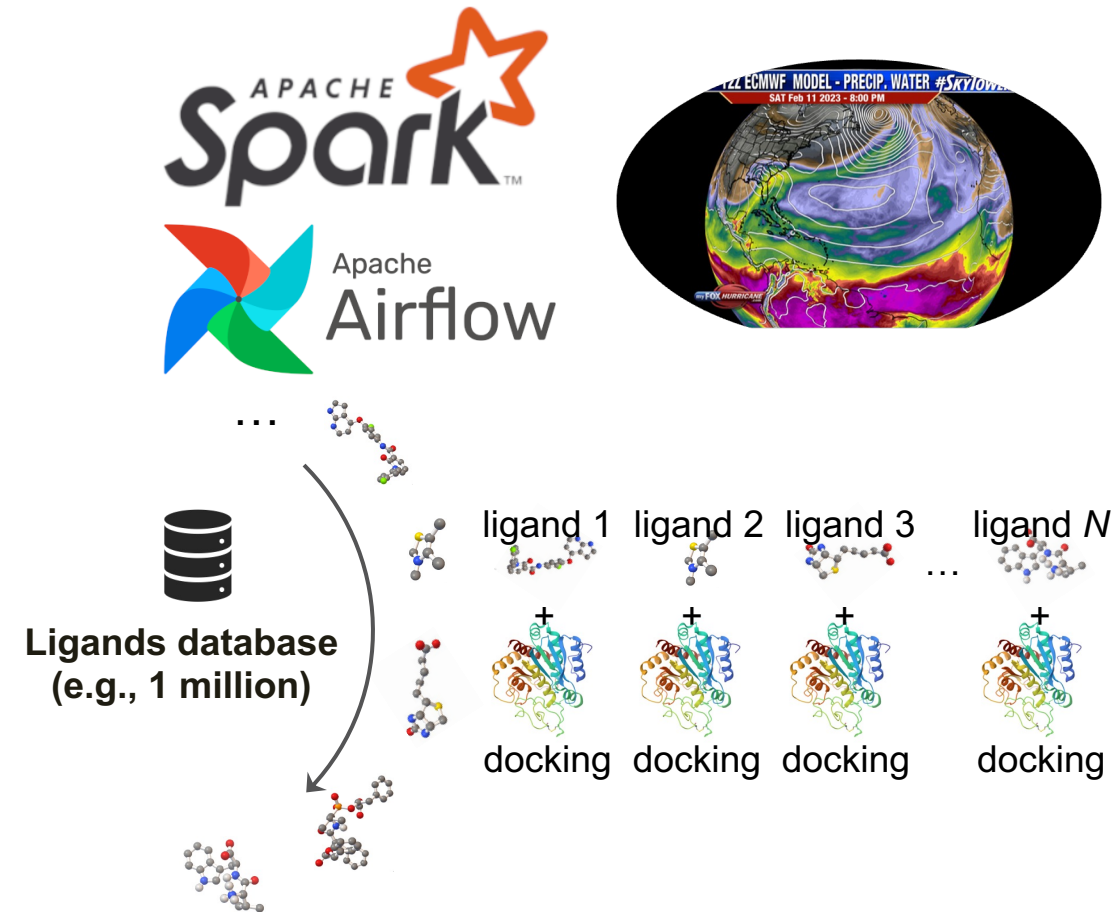
Develop middleware services for the heterogeneous DC architecture

- **Heterogeneous memory HBM+DRAM:** memory tiering support for cloud and HPC workloads
- **Fast and hierarchical storage:** integrate OpenFAM in popular key-value stores for data staging
- **Network software:** extension to Kubernetes CNI for optimized ethernet communication and support compute-in-network



OpenCUBE Driver Workloads

- Native cloud workloads, e.g., Spark, Airflow, etc
- Global weather forecast workflow
 - Simulations assimilated with observation data
- Virtual screening for rapid drug discovery
 - Elastic on-demand molecular docking workflow
- Machine learning integrated particle classification workflow



[1] P. Bauer, A. Thorpe, and G. Brunet. The quiet revolution of numerical weather prediction. Nature 525, no.7567: 47-55, 2015.

[2] Markidis, S., & Lapenta, G. (2010). Multi-scale simulations of plasma with iPIC3D. Mathematics and Computers in Simulation, 80(7), 1509-1519.

[3] Goodsell, D. S., Sanner, M. F., Olson, A. J., & Forli, S. (2021). The AutoDock suite at 30. Protein Science, 30(1), 31-43.

Thank you!

Follow Us at



horizon-opencube.eu



[@opencube_EU](https://twitter.com/opencube_EU)