Accelerated EuRopean clOud

AERO

An open source cloud software ecosystem for the EPI hardware

Barcelona, 09/06/2023

Funded by the European Union

AERO Motivation

- Cloud Service Providers (CSPs) offer a diverse & rich HW/SW ecosystem for deploying applications
  - Mainly driven by US companies
    - Strong **dependencies** to overseas technology providers
    - Increased **security** & **data privacy** concerns
    - **Limited flexibility** for custom designs that satisfy requirements of the European market & ecosystem

- European Processor Initiative (EPI) develops the first EU made chip and compute units (processors & accelerators)

- To achieve EU sovereignty the cloud SW ecosystem needs to follow HW developments

Funded by the European Union

AERO

Accelerated EuRopean clOud

Vision

Enable the future heterogeneous EU cloud infrastructure

AERO will **upbring** and **optimise** all components necessary to **achieve out-of-the-box heterogeneous execution** of the cloud ecosystem on the European processor. The outcome will be a set of **compilers**, **runtime systems**, **operating systems**, **system software**, and auxiliary **software deployment services**.

Accelerate the adoption of the EU cloud ecosystem

AERO will accelerate the adoption of the EU cloud ecosystem via upstreaming to open source projects, communicating and disseminating AERO's results to industry, academia, and standardization bodies.

Funded by the European Union

AERO

Accelerated EuRopean clOud

HW Platforms - Testbeds

Envisioned development & deployment platform

- SIPEARL's Rhea processor
  - ARM Neoverse V1 processors
  - PCIe support for GPUs
  - RISC-V accelerators

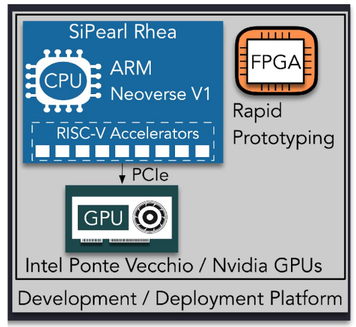
- Expected in Q1 2024

Alternative options for development

- Ampere platforms
  - ARM Neoverse N1 processors
  - PCIe support for GPUs
  - FPGA boards to test RISC-V accelerators

- AWS Graviton3 instances
  - ARM Neoverse V1 processors
  - No custom accelerators
  - Low level firmware cannot be modified

Hardware



Funded by the European Union

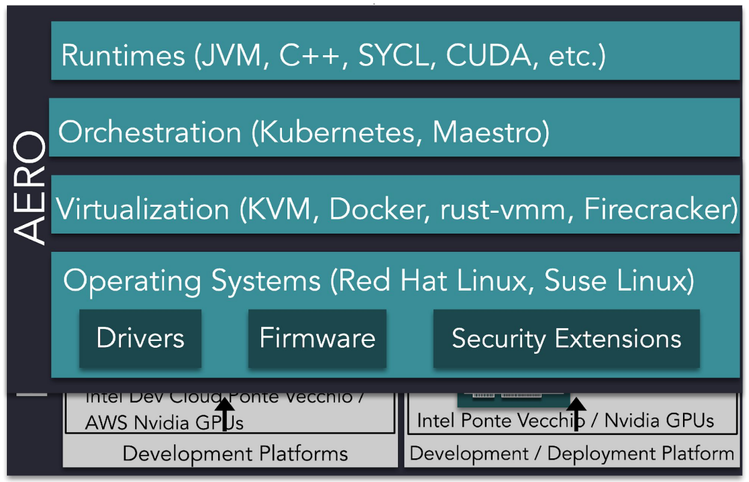
AERO

Accelerated EuRopean clOud

System Software

- Execution Runtimes
  - Optimized execution of programming languages & runtime systems serving as the backbone of software deployed/executed on cloud
    - OpenJDK, GraalVM managed programming languages (Java, Python, Scala, R, etc.)
    - TornadoVM GPU HW acceleration of managed programming languages
    - SYCL & DPC++/OneAPI HW acceleration of non-managed applications running in C/C++

- Cloud orchestration & Management Frameworks
  - Kuberneters, Maestro



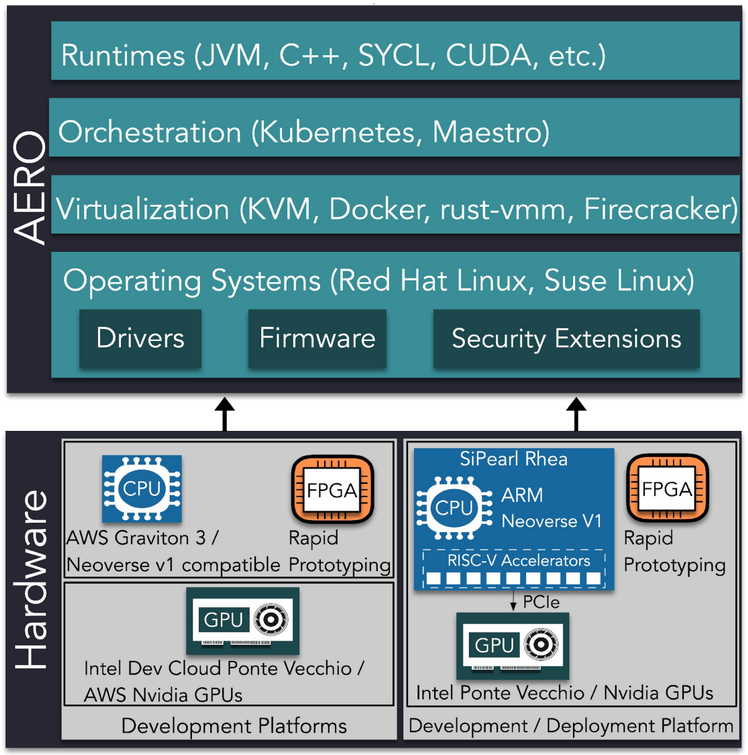
Funded by the European Union

AERO

Accelerated EuRopean clOud

System Software

- Virtualization Technologies
  - VMs (KVM), containers (Docker), microVMs (Firecracker)
  - Virtualization of GPUs & Rhea's RISC-V security elements
- Operating Systems – Drivers – Firmware
- Security extensions
  - Provide software interfaces for harnessing the underlying security IP blocks of Rhea
  - Design & prototype future post-quantum encryption algorithms



Funded by the European Union

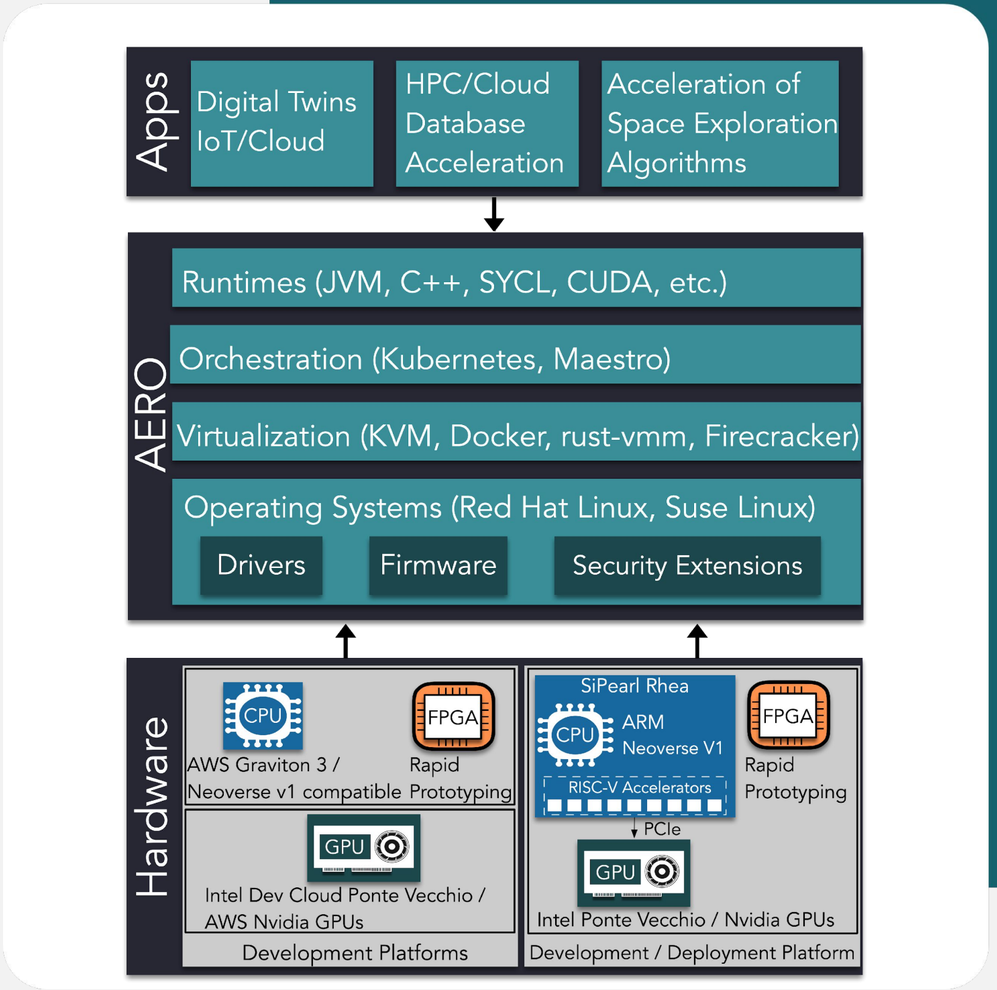
AERO

Accelerated EuRopean clOud

Application Software

- Three main pilots
  - Automotive domain (digital twins)
  - Algorithms for space exploration (analytics)
  - Database acceleration for scientific computing

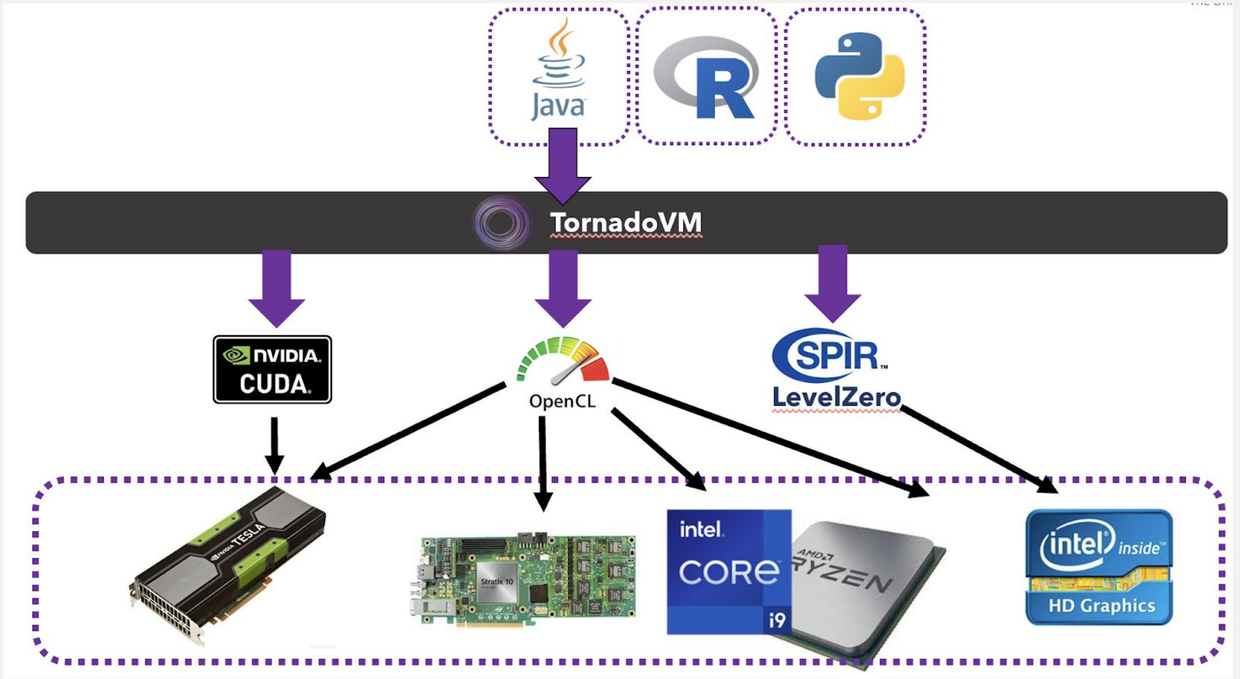
- Serverless workloads



Funded by the European Union

Leveraging the AERO stack

Accelerating Java/Cloud workloads on RISC-V



Enabling Language Frontends

- **TornadoVM**: a **Java** parallel programming framework and a **JVM plugin for** transparent **hardware acceleration** on multi-core CPU, GPUs and FPGAs.
- Java bytecodes OpenCL, PTX/CUDA, SPIR-V
- Developers select backend - TornadoVM takes care of data manipulation, partitioning, and orchestration of the execution

Funded by the European Union

Leveraging the AERO stack

Accelerating Java/Cloud workloads on RISC-V

Enabling Language Frontends

- **TornadoVM**: a **Java** parallel programming framework and a **J**
  plugin for transparent **hardware acceleration** on multi-core CP
  GPUs and FPGAs
- Java bytecodes OpenCL, PTX/CUDA, SPIR-V
- Developers select backend - TornadoVM takes care of data manipulation, partitioning, and orchestration of the execution

Enabling RISC-V Backends

- **ComputeAorta** enables implementation of open standards such as OpenCL
- It includes **tooling to convert OpenCL C and SPIR-V into target ISA** using existing LLVM backends

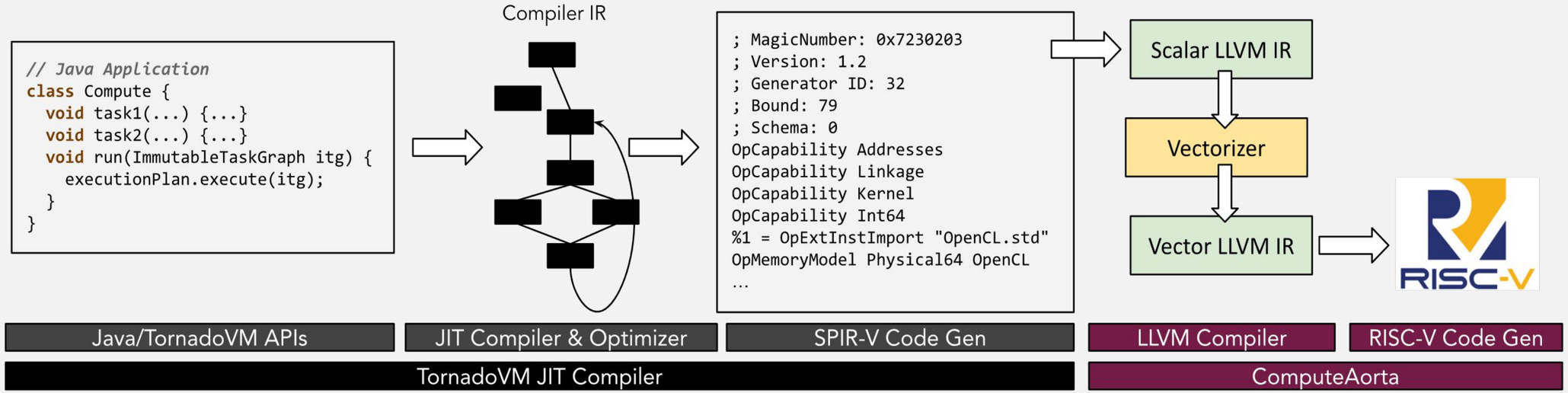
Funded by the European Union

Leveraging the AERO stack

Accelerating Java/Cloud workloads on RISC-V

Enabling Vectorization for RISC-V

- **Data-parallel** programs written in **Java** with **TornadoVM** can be accelerated using **ComputeAorta's** vector units via **RISC-V** RVV ISA instructions generated from Java **scalar code**



Funded by the European Union

Current Status

- AERO is completing its “Compatibility Report”
  - First critical technical deliverable
  - We have identified all the necessary software components that need to be part of the AERO stack
  - The compatibility report will describe which components need to be ported and/or optimized

- Waiting for access to Rhea-based testbeds
  - In the meantime keep working on Ampere platforms, AWS and ARM Fast Models

Funded by the European Union

Thank you!

Contact us:

aero-project.eu

@AERO_Project_EU

github.com/AERO-Project-EU

zenodo.org/communities/aero