EVEREST

DESIGN ENVIRONMENT
FOR EXTREME-SCALE BIG DATA ANALYTICS
ON HETEROGENEOUS PLATFORMS

START DATE: OCTOBER 1, 2020 | END DATE: SEPTEMBER 30, 2023 | GRANT NUMBER: 957269 | EU CONTRIBUTION: 5'037'372,50 €
TOPIC: ICT-51-2020 – BIG DATA TECHNOLOGIES AND EXTREME-SCALE ANALYTICS | CALL: H2020-ICT-2020-1 | FUNDING SCHEME: RIA – RESEARCH AND INNOVATION ACTION

PROJECT COORDINATOR: CHRISTOPH HAGLEITNER (IBM ZURICH) | **SCIENTIFIC COORDINATOR:** CHRISTIAN PILATO (POLITECNICO DI MILANO)

Concept and approach

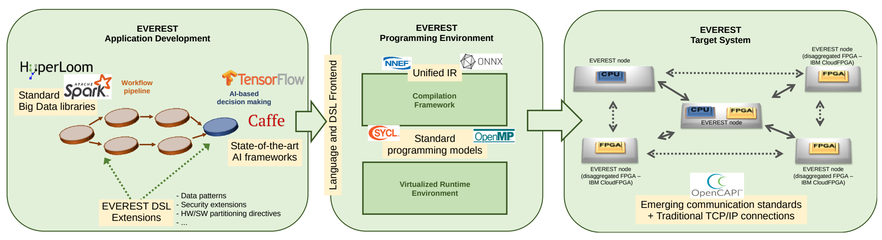
EVEREST focuses on **High Performance Big Data Analytics (HPDA) applications.**

- Future Big Data systems will be data-driven.
- Complex heterogeneous and reconfigurable architectures are difficult to program.

The EVEREST project aims at developing a holistic approach for co-designing computation and communication in a heterogeneous, distributed, scalable, and secure system for HPDA.

Main features:

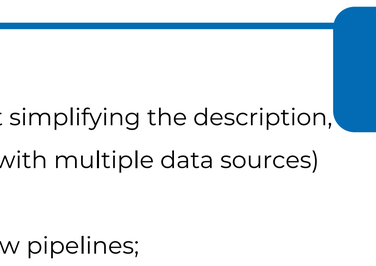
- **data-driven design approach;**
- combination of **compiler transformations, high-level synthesis, and memory management;**
- efficient monitoring of the execution with a **virtualisation-based environment.**



EVEREST proposes a **design environment** that combines state-of-the-art, stable programming models, and emerging communication standards with novel and **dedicated domain-specific extensions.**

Application use cases

Weather-based renewable-energy prediction



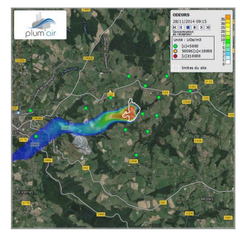
EVEREST will enrich the output of an **ensemble of meteorological simulations** at cloud permitting grid spacing (2–4 km) with **AI post-processing techniques** on grid data.

GOAL: achieve **better accuracy of forecast products** for the day-ahead energy market, intraday energy market, and next continuous energy trading market.

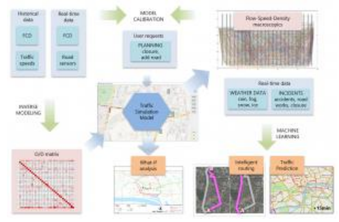
Air-quality monitoring for industrial sites

EVEREST will execute an **ensemble of meteorological simulations** at cloud permitting grid spacing (2–4 km) and apply **machine learning (ML) approaches with local meteorological measurements.**

GOAL: improve the **quality of the day-to-day forecasts** combined with an emission forecast from the industrial site to **take proper actions.**



Real-time traffic modelling in smart cities



EVEREST will extend the **computation of precise 3D traffic models** based on the processing of several data sources such as historical and real-time floating car data (FCD) of moving vehicles and weather data (precipitation, temperature).

GOAL: computation of **traffic prediction and truly intelligent routing** for global traffic optimisation flow using various machine learning techniques.

Data allocation, storage and communication

Data-driven approach: specific analysis of data requirements:

- access patterns;
- allocation policies;
- communication primitives.

To **co-optimise computation and data movements** based on the characteristics of the data.

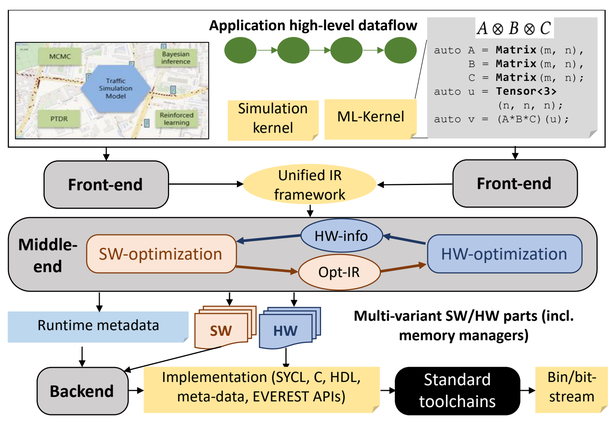
Compilation framework

The EVEREST compilation framework aims at simplifying the description, optimisation, and implementation of HPDA (with multiple data sources) onto FPGA-based distributed architectures:

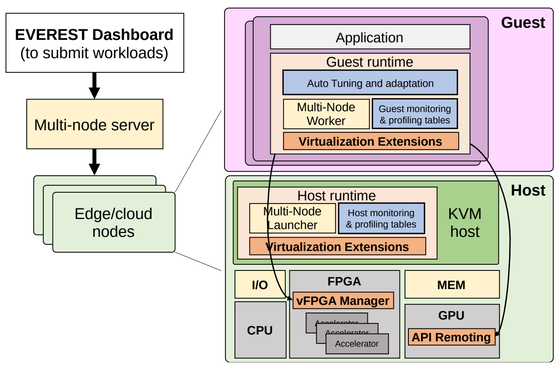
- **high-level libraries** to describe the workflow pipelines;
- **domain-specific extensions** to abstract functionality and data properties;
- integration with **existing AI libraries and frameworks**, and **communication libraries.**

For facilitating hardware/software co-design and monitoring, EVEREST offers the generation of **optimised code variants**, reconfigurable accelerators with the use of **high-level synthesis**, and the creation of **specialised memory managers.**

Programming environment: EVEREST system development kit



Virtualised runtime environment



The virtualisation environment dynamically adapts the underlying hardware based on workload conditions and the availability of the different hardware resources:

- **Distributed runtime support** for workload distribution and load balancing.
- **Auto-tuning approaches** to change the application parameters, enabling the implementation of a dynamic hardware-software adaptation layer.
- **Virtualisation support and hypervisor extensions** to expose hardware configurable parameters directly to the applications inside the Virtual Machines.

Together with security mechanisms in federated systems it creates a **complete data protection layer.**

Data security

Use of **federated computing resources** and secure data management policies.

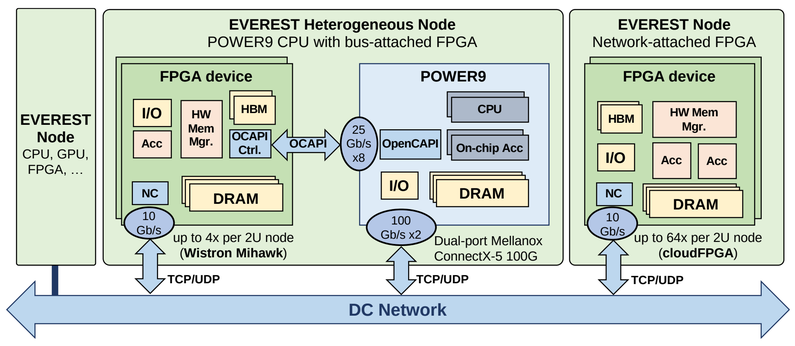
Development of **language extensions, libraries, and synthesizable components** for enforcing security and privacy of the data.

Strong focus on **programmability** (simplifying application development and compilation), **optimisation** (combining design and runtime approaches), and **interoperability** (relying on standardised API and interfaces).

Target system

EVEREST is targeting FPGA-based distributed architectures to accelerate extreme-scale Big Data applications:

- **multi-node hardware system** potentially organised in **federated data centres:**
  - CPU-based data centres, network-attached FPGA cards, cloud-based infrastructures;
- EVEREST SDK built on top of **existing communication libraries and orchestration infrastructures;**
- **heterogeneous resources** to study **application portability.**



EVEREST CONSORTIUM

IBM · Politecnico Milano 1863 · Università della Svizzera italiana · Technische Universität Dresden · cima Research Foundation · VSB TUO IT4I · Virtual Open Systems · Duferco Energia · Numtech · Sygic