REliable Power and time-ConstraInts-aware Predictive management of heterogeneous Exascale systems

Politecnico di Milano (POLIMI), Universitat Politècnica de València (UPV), Centro Regionale Information Communication Technology scrl (CeRICT), École Polytechnique Fédèrale de Lausanne (EPFL), Barcelona Supercomputing Center (BSC), Poznan Supercomputing and Networking Center (PSNC), IBT Solutions S.r.l. (IBTS), Centre Hospitalier Universitaire Vaudois (CHUV)
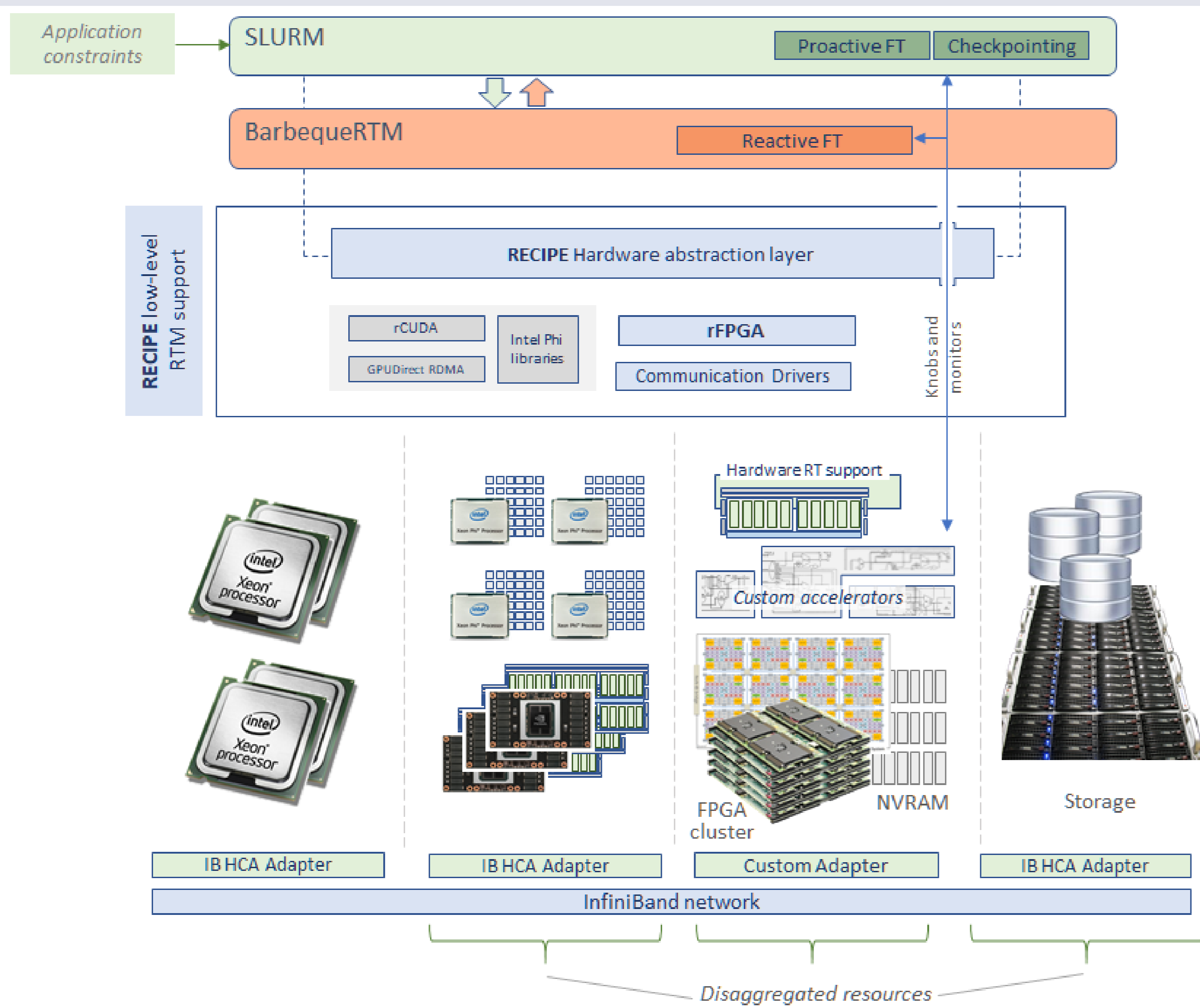
Challenges in High Performance Computing

- ► Emerging classes of applications with diverse QoS constraints
- ► Need for powerful but less energy-hungry supercomputers
- ► Exascale HPC infrastructures complexity (deep heterogeneity)
- ► Increasing likelihood of transient and long-term faults

RECIPE reconfigurable accelerators

- ► Disaggregation of HW resources through SW API and HW interface
- ► Full custom HDL implementation, library based designs, etc.
- ► Software programmed accelerators: nu+ vector core, PEAK SMP

RECIPE tools to manage heterogeneous resources in future HPC

- ► *Runtime resource management* infrastructure to
  - ► optimize energy efficiency, minimize thermal hotspots
  - ► enforce application time constraints
  - ► ensure reliability for both time-critical and throughput-oriented computation
- ► *Predictive reliability* methodology to
  - ► enforce QoS guarantees
  - ► deal with transient and long-term hardware failures
  - ► include thermal, timing and reliability models
- ► Integration layers supporting *resource disaggregation*
  - ► resource manager – application
  - ► resource manager – deeply heterogeneous architecture

Hardware Software Stack

Deeply heterogeneous accelerators resulting from the EU FET-HPC MANGO project coupled with state-of-the-art heterogeneous resources
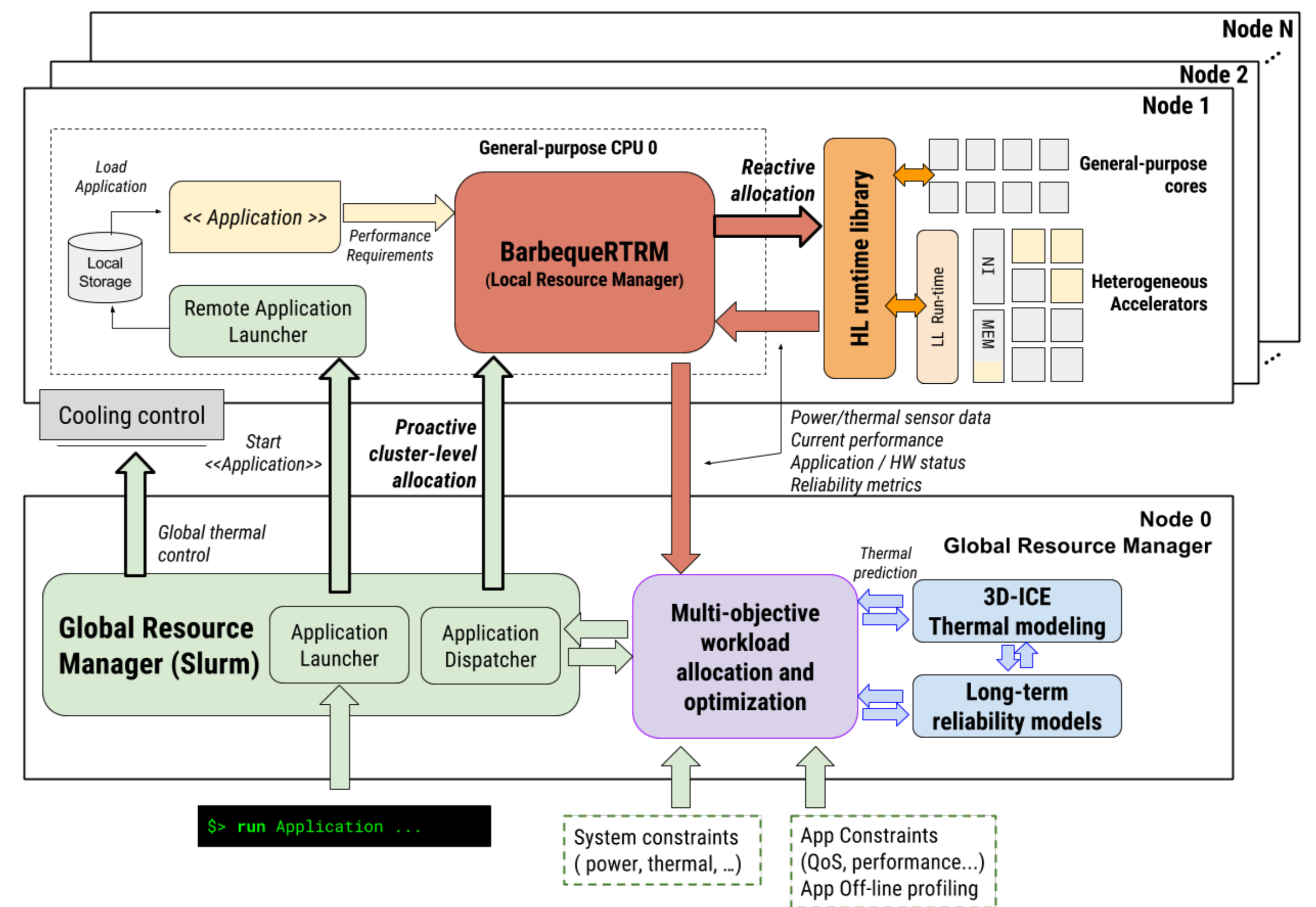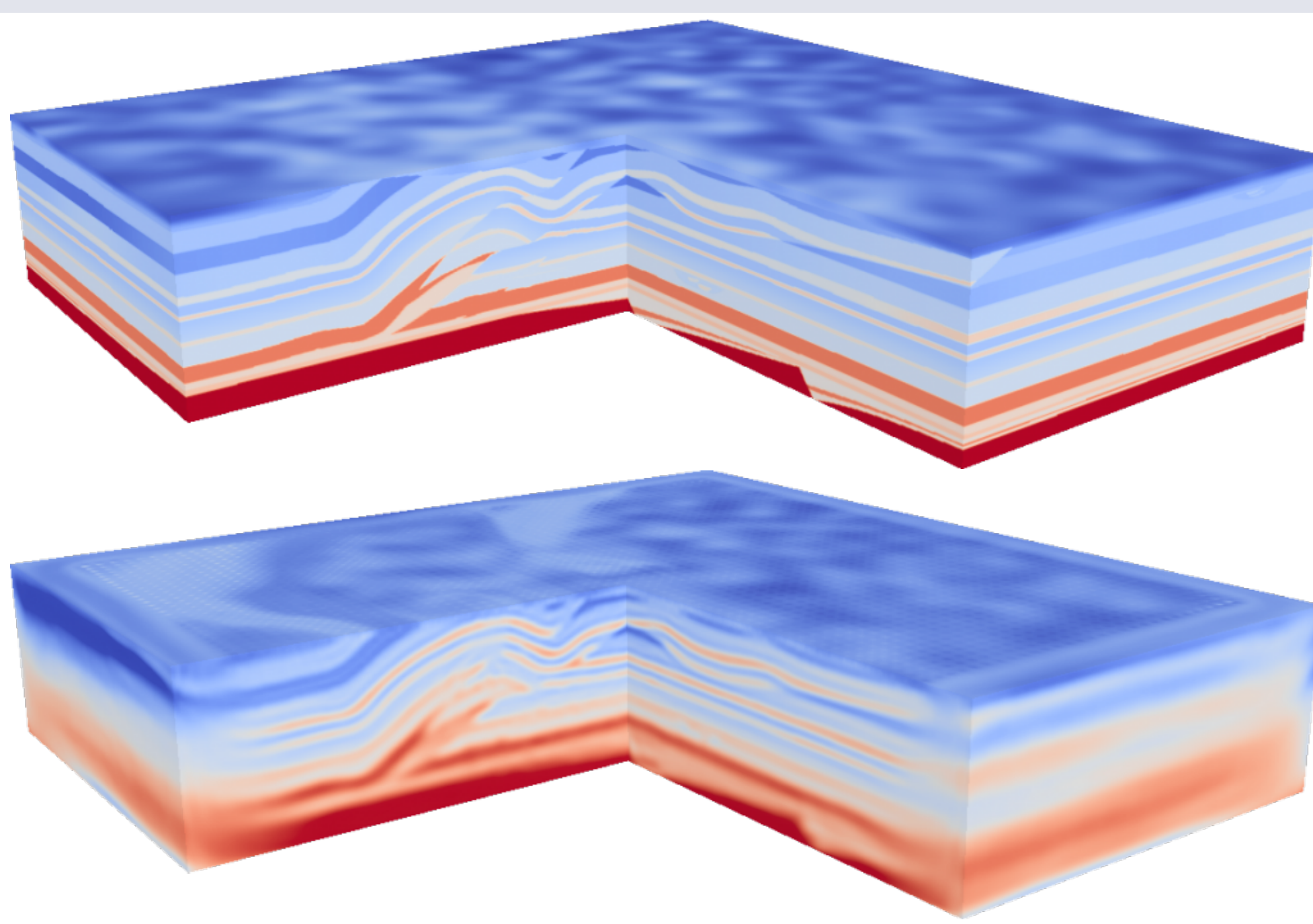

Resource Management Infrastructure

Supporting hierarchical resource management of deeply heterogeneous resources
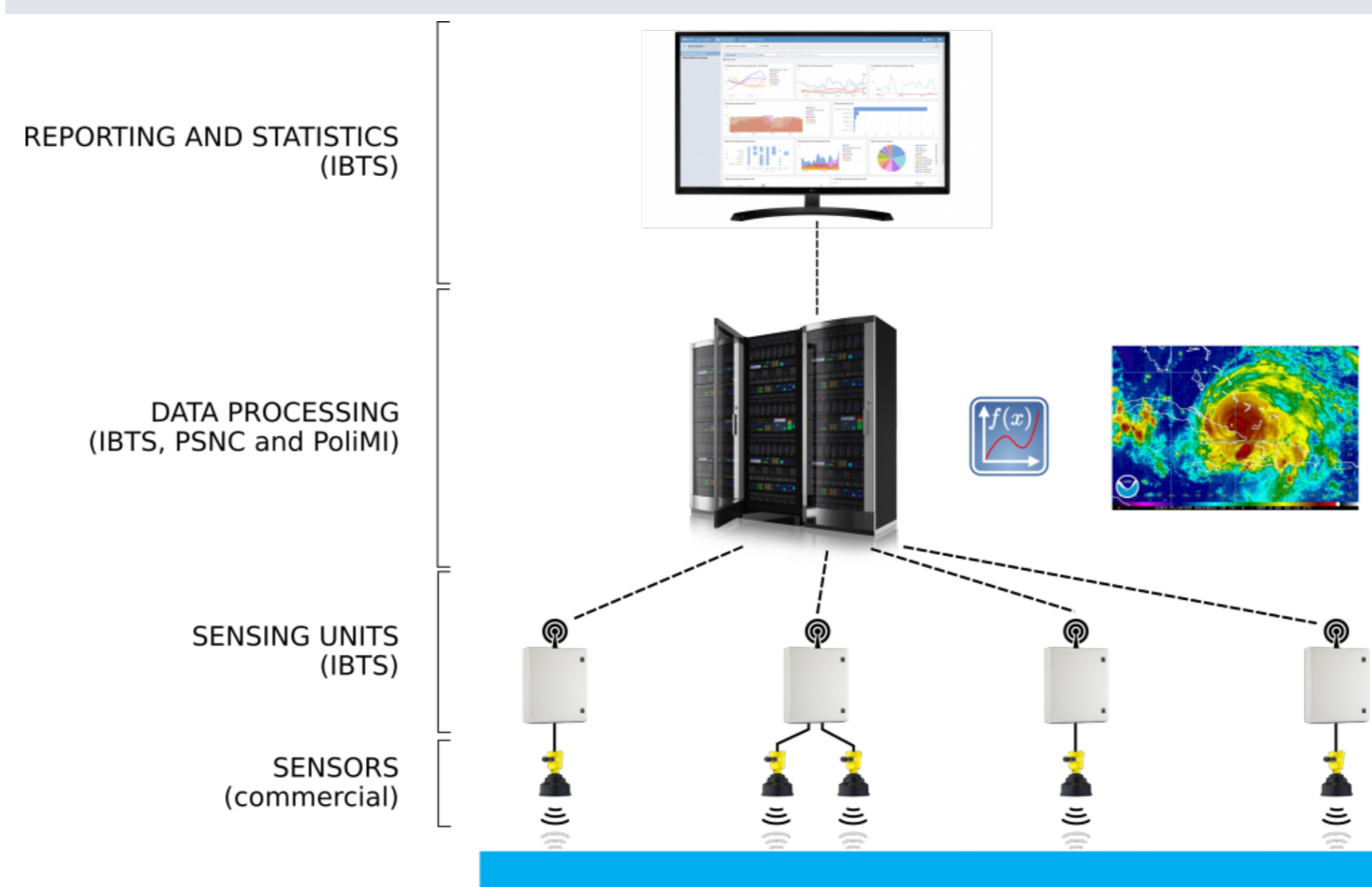

Use Case 1: Geophysical Exploration

- ► Analysis of subsoil properties to locate potential oil/gas reservoirs
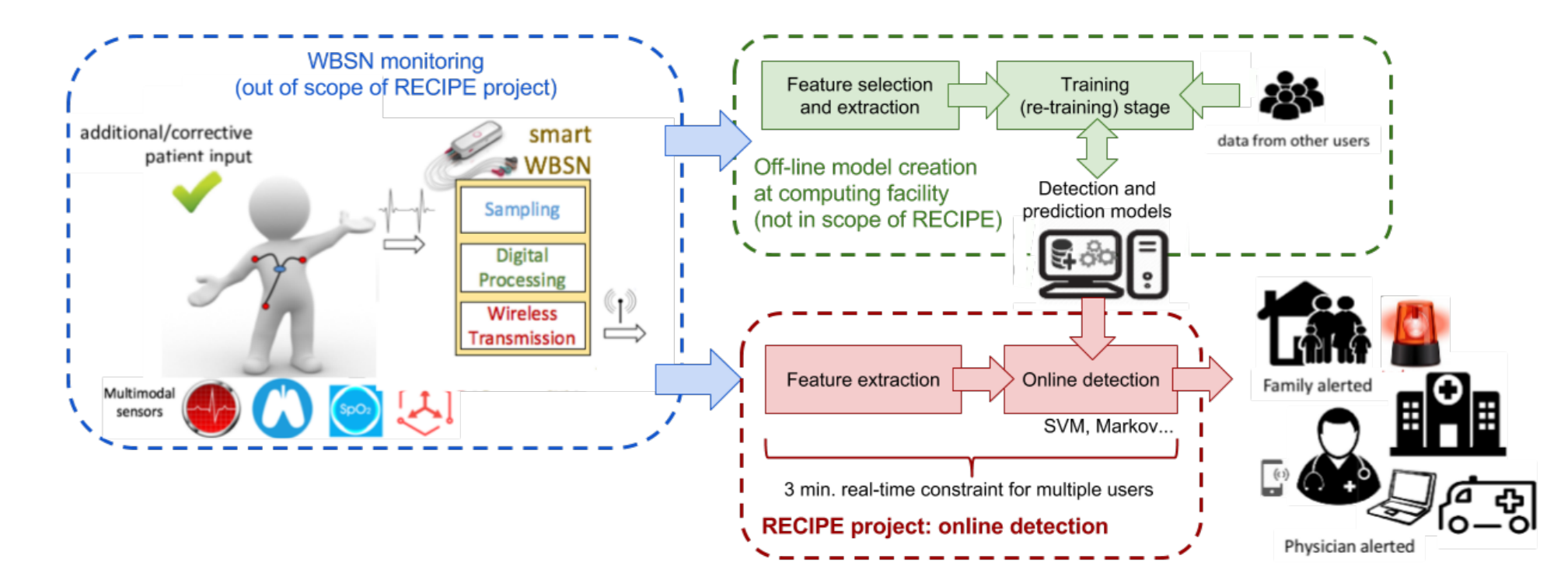- ► Fast waveform inversion mini-app

Use Case 2: Climate & Renewable Energies

- ► Prediction of renewable energy production from wind farms (24h forecast)
- ► Analysis of sensor data for detection/prevention of flood events (within 30 minutes of event);

Use Case 3: Biomedical machine learning and data analytics

- ► Monitor biomedical sensor data
- ► Detect epileptic seizures within 1 minute from onset
- ► Allow timely intervention from medical teams