Accelerating the Design of Scientific Workflows with Simulation-Based Rapid Prototyping

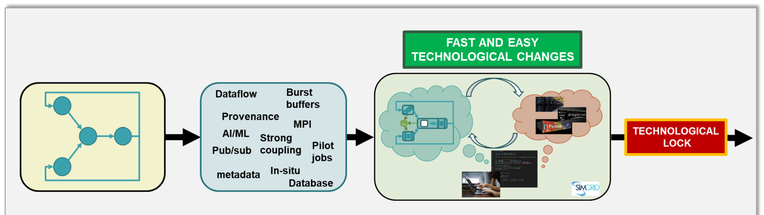
Laboratory Directed Research & Development
LDRD.ORNL.GOV

LOIS ID:11184

PROGRAM: Strategic Hire
PI: Fred Suter (suterf@ornl.gov)
CONTRIBUTORS:
Y.-C. Wong, H. Casanova, K. Mehta, J. Mc Donald, R. Ferreira da Silva

Develop a comprehensive simulation-based framework to help domain scientists prototyping their workflows and provide them with the most efficient design and implementation

Simulation-based Rapid Prototyping

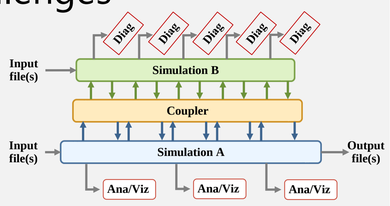


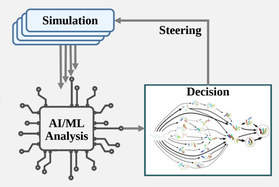
1. **Design abstract** workflow
   - Components and dependencies
2. **Identify patterns** and adapted candidate strategies
   - User engagement
   - Pencil and papers prototyping
3. **Implement** workflow **simulator**
   - **Test** on simulated infrastructure(s)
   - Fast **iterative refinement** process
   - Converge on **concrete** workflow
4. Select the **most adapted WMS**
5. **Deploy** on target infrastructure(s)

Workflow Execution Motifs

- **Modern** scientific workflows are no DAGs
- New **orchestration** and **data management** challenges

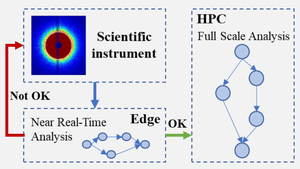
Strong Code Coupling and Analytics

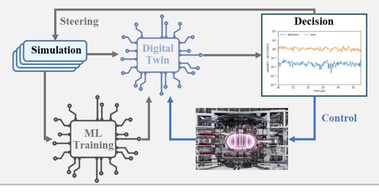




AI/ML-based Steering

Edge-to-HPC Multi-Stage Analysis

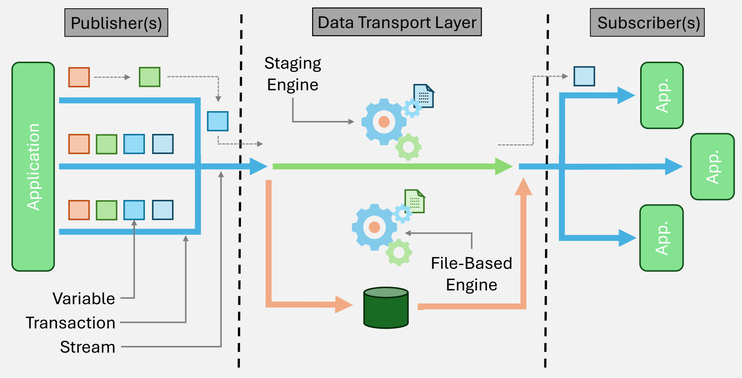




Digital Twins

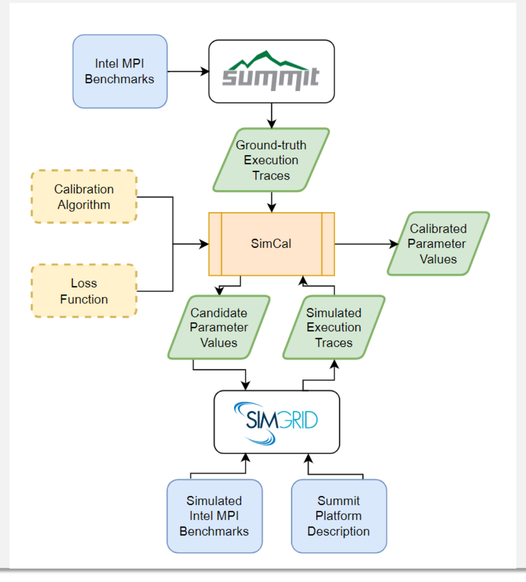
Versatile Data Transport Layer

- Generic **Publish/ Subscribe** abstraction to applications and services
- **Decouple** data movement and storage
- Enable **testing and evaluation** of
  - File-based / Streaming / In-memory / No-cost



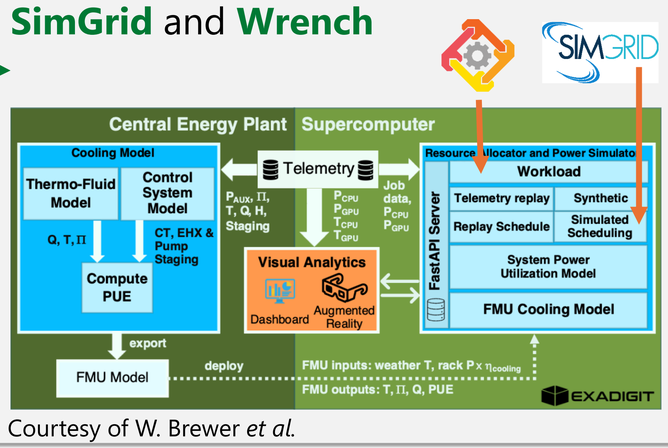
Automated Calibration of Platform Descriptions

- **Objective:** reflect reality but abstract away some real-world behaviors
- **Approach:** Calibrate simulation parameters from ground-truth real-world execution traces



Towards a Comprehensive Digital Twin of Leadership Class Facilities

- **Objective:** Gain full insight on energy-efficiency optimization
  - Requires coordination across many interoperating components
    - Applications, system software, data mgmt, hardware, data center power and cooling
- **Approach:** Couple the **ExaDigiT** digital twin of Frontier developed by NCCS to **SimGrid** and **Wrench**



Courtesy of W. Brewer *et al.*

Project Outcomes

SCAN ME

- **Software artifacts**
  - SimGrid
  - File system simulation module
  - Simulated Data Transport Layer
- **Awards**
  - Best poster award at e-science 2023
- **Collaborations**
  - NCCS ATS Section
  - University of Hawai`i at Manoā
  - KIT (Germany)
  - Inria (France)
- **Future funding opportunities**
  - Participation to the ASCR Energy-Efficient Computing for Science Workshop

Research sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy.

OAK RIDGE National Laboratory