A FIWARE based digital twin for high performance computing centers

Jonathan Kriwet
RWTH Aachen University
Aachen, Germany
jonathan.kriwet@eonerc.rwth-aachen.de

Christian Wassermann
RWTH Aachen University, IT Center
Aachen, Germany
wassermann@itc.rwth-aachen.de

Rita Streblow
RWTH Aachen University
Aachen, Germany
rstreblow@eonerc.rwth-aachen.de

Dirk Müller
RWTH Aachen University
Aachen, Germany
dmueller@eonerc.rwth-aachen.de

Abstract

This paper presents a modular digital twin framework for High-Performance Computing (HPC) centers, to enable the optimization of their energy efficiency. The framework combines a FIWARE-based IoT platform with a hybrid modeling approach. The methodology is demonstrated at a German HPC center, where the digital twin is able to predict the Power Usage Effectiveness with an error of less than 2.5 % while also identifying several operational inefficiencies.

Keywords

Digital Twin, High-Performance Computing, FIWARE, Cooling Infrastructure, Data Center Optimization, Hybrid Modeling

ACM Reference Format:
Jonathan Kriwet, Christian Wassermann, Rita Streblow, and Dirk Müller. 2025. A FIWARE based digital twin for high performance computing centers. In *The 16th ACM International Conference on Future and Sustainable Energy Systems (E-ENERGY '25), June 17–20, 2025, Rotterdam, Netherlands.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3679240.3734677

1 Introduction

High-Performance Computing (HPC) centers are essential for scientific advancement, but their increasing energy demands present significant environmental and economic challenges. Studies have shown that integrated optimization of both computing and cooling infrastructure can significantly improve HPC centers energy efficiency [3]. Digital twins have emerged as powerful tools for such system-wide optimization [2]. The FIWARE-Framework has proven successful for digital twin implementations in various domains, including intelligent transportation [4] and urban mobility [1]. To the best of the authors' knowledge, this is the first application of FIWARE for a digital twin in the HPC domain.

This paper presents a novel hybrid approach to creating a digital twin for an HPC center that combines data-based models with physics-based Modelica models, implemented through a modular FIWARE-based architecture with online learning capabilities.

E-ENERGY '25, Rotterdam, Netherlands

ACM ISBN 979-8-4007-1125-1/25/06
https://doi.org/10.1145/3679240.3734677

2 Methodology

As illustrated in Figure 1, the digital twin has a modular architecture, whose components can be categorized into six elements: IoT platform, HPC model, energy infrastructure model, communication agents, standardized data model, and operations dashboard.

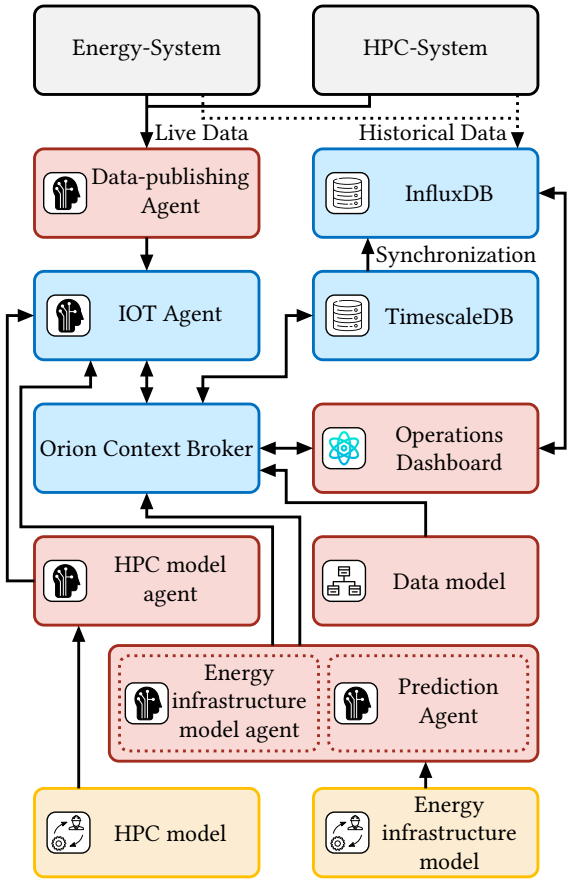

Figure 1: The structure of the digital twin. Blue components are existing FIWARE components. Red components were developed for the communication processes. The yellow components are the models. The two subsystems under consideration are shown in gray.

At the core of the framework lies the FIWARE-based IoT platform, which manages component communication and data handling. Central to this platform is the Context Broker, maintaining real-time context data, while a dual-database system using TimescaleDB and InfluxDB ensures both comprehensive backup and efficient historical data access.

For predicting cluster power consumption with the HPC model, a linear regression approach based on aggregated utilization metrics from power distribution units is used. To be able to account for system changes, an online learning strategy with weekly model retraining using rolling windows of data is implemented.

The energy infrastructure model is physics-based and done in Modelica. It includes all components of the cooling system, such as cooling distribution units, side coolers, heat exchangers, cooling towers and a connection to district cooling. The model is calibrated using a genetic algorithm and measurement data from the HPC center. Due to the complexity of the system, several simplifications were necessary, including a simplified representation of the hydraulics, uniform heat exchanger flow distribution, and simplified air-cooled component representation. Since the optimization efforts focus primarily on improving the Power Usage Effectiveness (PUE), for which pump powers are needed, which are not immediately modelled correctly due to the simplified hydraulics, a separate machine learning approach is developed. This approach predicts all missing power values, using simulated data from both the HPC and energy infrastructure models to avoid error propagation, while also employing an online learning strategy with weekly model retraining and automatic feature selection.

The end user can interact with the digital twin through an operations dashboard in the form of a web application. With this dashboard, users can both visualize data and perform scenario analyses, to evaluate the impact of different setpoints on the system and identify potential optimization opportunities.

3 Use Case and Results

The digital twin is Docker-containerized, primarily comprising FIWARE components for the IoT Platform with all other components except the Modelica-based model implemented in Python. It is deployed for a German HPC center, undergoing a system transition, where a new cluster is brought online while the old one is gradually shut down. Two separate datasets are used for the training/calibration of the models and the validation. During development, the digital twin already yielded valuable insights through modeling deviations that indicated potential issues in the real system:

- District cooling was activating unnecessarily during periods when the cooling towers provided sufficient cooling power.
- Contrary to expectations, cooling water flow through the shutdown cluster increased as shutdown progressed, rather than decreased.
- Temperature sensor calibration issues were identified through constant deviations from setpoints.

The digital twin is able to achieve high prediction accuracy for all key operational variables, resulting in a Mean Absolute Percentage Error (MAPE) of 2.1 % for the PUE during the validation period, shown in Figure 2. Due to the mentioned error in the real system,

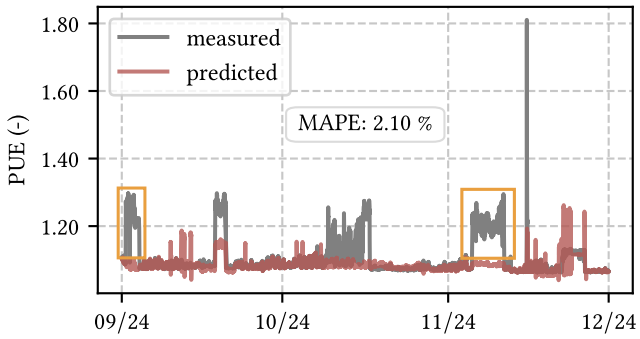

Figure 2: Predicted and measured PUE, including the MAPE. The grey line shows the measured values, while the red line represents the predicted values. Exemplary areas in which the integration of district cooling does not work in orange.

excluding the district cooling from the calculation of the PUE reduces the error in the same period to 0.41 %. The online learning approach demonstrates its strength in capturing and adapting to the gradual shutdown of the old cluster. Finally, in a simulated live operation, it was possible to identify scenarios, where a change in setpoints would lead to a reduction of the PUE, resulting in a more efficient operation.

4 Conclusion and Outlook

The digital twin framework presented in this paper could be successfully implemented at an example HPC center, demonstrating good prediction accuracies and showcasing the ability to improve system efficiency during the simulated live operation. For an actual live operation, an adjustment needs to be made to the HPC model so that the utilization metrics can be predicted via the scheduled jobs. This development is currently ongoing. Nevertheless, the digital twin was already able to identify several real system inefficiencies, indicating its potential in continuous system optimization, even beyond its primary prediction functionality.

Acknowledgments

We gratefully acknowledge the financial support by the German Federal Ministry of Education and Research (BMBF), promotional reference 16ME0614K.

References

[1] Martin Bauer, Flavio Cirillo, Jonathan Fürst, Gürkan Solmaz, and Ernö Kovacs. 2021. Urban Digital Twins – A FIWARE-based model. *at - Automatisierungstechnik* 69, 12 (Dec. 2021), 1106–1115. doi:10.1515/auto-2021-0083

[2] Xuan Liu, He Gan, Ying Luo, Yangquan Chen, and Liang Gao. 2023. Digital-Twin-Based Real-Time Optimization for a Fractional Order Controller for Industrial Robots. *Fractal and Fractional* 7, 2 (Feb. 2023), 167. doi:10.3390/fractalfract7020167

[3] Yongyi Ran, Han Hu, Yonggang Wen, and Xin Zhou. 2023. Optimizing Energy Efficiency for Data Center via Parameterized Deep Reinforcement Learning. *IEEE Transactions on Services Computing* 16, 2 (March 2023), 1310–1323. doi:10.1109/TSC.2022.3184835

[4] Franca Rocco Di Torrepadula, Alessandra Somma, Alessandra De Benedictis, and Nicola Mazzocca. 2024. Smart Ecosystems and Digital Twins: an architectural perspective and a FIWARE-based solution. *IEEE Software* (2024), 2–9. doi:10.1109/MS.2024.3518752