Federal Register Notice: 89 FR 51554, Federal Register :: Networking and Information Technology Research and Development Request for Information on Digital Twins Research and Development, June 18, 2024.

Request for Information on the National Digital Twins R&D Strategic Plan

Lawrence Berkeley National Laboratory

Response to RFI on
Digital Twins Research and Development Plan

Submitted July 28, 2024

Prepared by: Ann Almgren, John Bell, Xiaoye Sherry Li, Daniel F. Martin, James Sethian, Stefan M. Wild (POC, [redacted] Computing Sciences, Lawrence Berkeley National Laboratory

Computing Sciences at Lawrence Berkeley National Laboratory, a U.S. DOE Office of Science National Laboratory, advances computational science in the national interest. This response is based on the extensive experience and expertise of our teams in adaptive modeling of complex physical and biological systems.

This document is approved for public dissemination. The document contains no business-proprietary or confidential information. Document contents may be reused by the government in the National Digital Twins R&D Strategic Plan and associated documents without attribution.

The potential of digital twins rests on the ability to intertwine and integrate advances in experimental sensing and measurements, data analysis, real-time feedback and autonomous steering, and advanced computational capabilities to model and simulate highly complex physics and phenomena across a range of time and space scales, changing environments, and dynamic data integration. This will require advances on many crosscutting and discipline-specific fronts.

Leveraging DOE Advances in a National Digital Twin R&D Ecosystem:

The recent emergence of a host of new capabilities within DOE National Laboratories is set to profoundly advance our abilities to intimately couple experiment and simulation to provide a new framework for high-fidelity digital twins.

These new capabilities[1] include:

---

[1] See, e.g.,

- Hendrickson et al, ASCR@40: Highlights & Impacts of ASCR's Programs, doi:10.2172/1631812
- Heroux et al, ECP Software Technology Capability Assessment Report, doi:10.2172/1888898
- Siegel et al, Map Applications to Target Exascale Architecture with Machine-Specific Performance Analysis, Including Challenges and Projections, doi:10.2172/1838979
- CAMERA: Center for Advanced Mathematics for Energy Research Applications, https://camera.lbl.gov

1. High-order and robust new mathematical and computational methods across fluid mechanics, materials sciences, biological modeling, advanced chemistry, and core physics – including mechanics, quantum effects, and magneto-electric dynamics – that effectively exploit advanced high-performance computing architectures to achieve unprecedented accuracy, resolution, and fidelity;
2. Sophisticated new methods for data analysis, inversion, and edge processing directly at experimental sources to minimize latency between physical and virtual systems;
3. New statistical sampling methodologies to facilitate principled Bayesian model calibration and uncertainty quantification, and provide confidence in understanding; and
4. Powerful new AI/ML techniques to exploit existing data to rapidly interpret results, extract new information, and automatically steer coupled experiments and simulations to achieve unprecedented efficiency and accuracy in new discoveries.

These developments have been concurrent with advances in computing as well as state-of-the-art networking and data transfer to provide close-to-real time processing of information and remote access.

The ability for new methodologies to also take advantage of a DOE computing infrastructure that is second to none provides a unique combination of technical capabilities and the computing capacity to make them transformational components of a national digital twin ecosystem.

Holistic Integration of Verification, Validation, and Uncertainty Quantification (VVUQ):

A key aspect of digital twins is the bi-directional interaction between the physical and virtual system. The physical system provides data to calibrate the virtual system and sets the initial and boundary conditions for the virtual system's predictions. The virtual system provides feedback to the physical system such as control strategies, decision support or requirements for additional measurements. Developing a digital twin for a complex multiphysics system described by partial differential equations, either natural or engineered, poses a number of research challenges. One key challenge is that numerical models (whether traditional, ML-based, or hybrid) for these types of problems involve large numbers of unknown parameters. Uncertainties in the data observed from the physical system can be expressed in terms of a probability distribution. By adopting a Bayesian perspective, one can use data from the physical system to refine the distribution. Feedback from the virtual system to the physical system will be inherently probabilistic, requiring methodology for optimization under uncertainty. Additionally, for risk assessment, specialized methodology for rare-event sampling will also be required. While many of the most advanced numerical models have been developed with rigorous error bounds, advances will be needed to make these models interoperable with such Bayesian and optimization “outer loop processes.” Also needed is further theoretical and algorithmic research for error analysis, including in cases where available measurements are sparse or involve reduced-order models or evolving dynamics in latent spaces.

Sustainable Progress in an Evolving Computing Landscape:

Computing hardware – from high-performance systems to edge devices – will continue to evolve rapidly. In developing an ecosystem for digital twins, it is vital to ensure an intentionally sustainable software stack, including the numerical libraries built on top of system software. Sustaining such software often requires whole-of-community efforts[2].

Additional R&D Focus Area: Goal-Oriented and Hierarchical Models for Purpose-Driven Digital Twins:

Another feature of complex multiphysics systems is that there are a wide range of potential uses of a digital twin for the given system. For example, given a digital twin for a wind farm, one might ask: Where should sensors be placed to best improve predictions of system performance? How does the wind farm respond to extreme weather events? What service schedule plan will best keep the complex hierarchy of underlying equipment and materials operational? Answering each of these questions likely requires models of different fidelity in terms of numerical accuracy and/or what physical processes are included in the model, ranging from simple surrogates to detailed physical descriptions of the system. Always using the highest fidelity available is computationally infeasible in most cases. This suggests that an effective digital twin must include a hierarchy of models of different complexity levels. For each model within the hierarchy, one would need to characterize the specific domain of applicability and quantify the uncertainty over that domain.

Effectively exploiting a hierarchy of models of different fidelities and operating at different scales can significantly reduce the (computational, energy, etc.) costs associated with a digital twin. Although emerging HPC architectures have enabled simulations of unprecedented fidelity, directly embedding these types of simulations as part of an optimization algorithm or sampling methodology is not feasible. In addition to understanding the limitations of different models, one also needs to quantify the relationship between models. How can refinement of one model be used to improve other models in the hierarchy? What is the uncertainty associated with mapping between models? Some systems can be characterized by distinct components that can be probed independently. For these types of systems, the relationship between models is fairly straightforward. When the different processes in the system are coupled nonlinearly or when controlled experiments are not feasible, however, the relationship between models can become quite complex. Given a hierarchy of models, one can then bootstrap up from simple, inexpensive models to more models with sufficient complexity to answer a given question, avoiding performing expensive simulations in regions that are not of interest.

Digital twins represent a timely opportunity to advance science and engineering. Realizing their full potential will require taking advantage of recent advances at the DOE National Laboratories; making numerical models and simulations digital-twin-ready, with rigorous error bounds and VVUQ capabilities; developing a sustainable software stack; and addressing their specific goals and downstream uses – each with corresponding fidelity levels – to ensure their effectiveness.

---

[2] See, e.g., Consortium for the Advancement of Scientific Software, https://cass.community