h3-Open-BDEC: Innovative Software Platform for Scientific Computing in the Exascale Era by Integrations of (Simulation + Data + Learning)

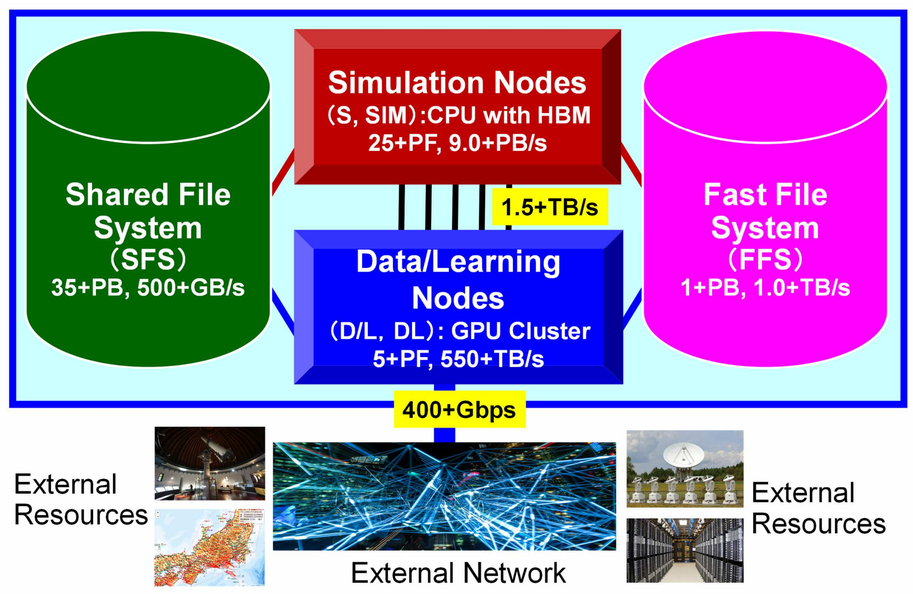
T. Iwashita (Hokkaido U.), K. Nakajima (U. Tokyo/RIKEN), T. Shimokawabe (U. Tokyo), H. Nagao (U.Tokyo), T. Ogita (TWCU), T. Katagiri (Nagoya U.), H. Yashiro (NIES), H. Matsuba (U.Tokyo)

Purpose & Background

- In this study, we propose an innovative method of computational science for sustainable promotion in scientific discovery by supercomputers towards the **Exascale Era/ Society 5.0** (Super Smart/Human-Centered Society, the Cabinet Office of Japan) by integration of **(Simulation + Data + Learning (S+D+L))**, where ideas of data science and machine learning are introduced to computational science
- This project is supported by Japanese Government through JSPS Grant-in-Aid for Scientific Research (S) during FY.2019-2023 (PI: Kengo Nakajima, The University of Tokyo) (Total Budget: 152.7M JPY= 1.41M USD)

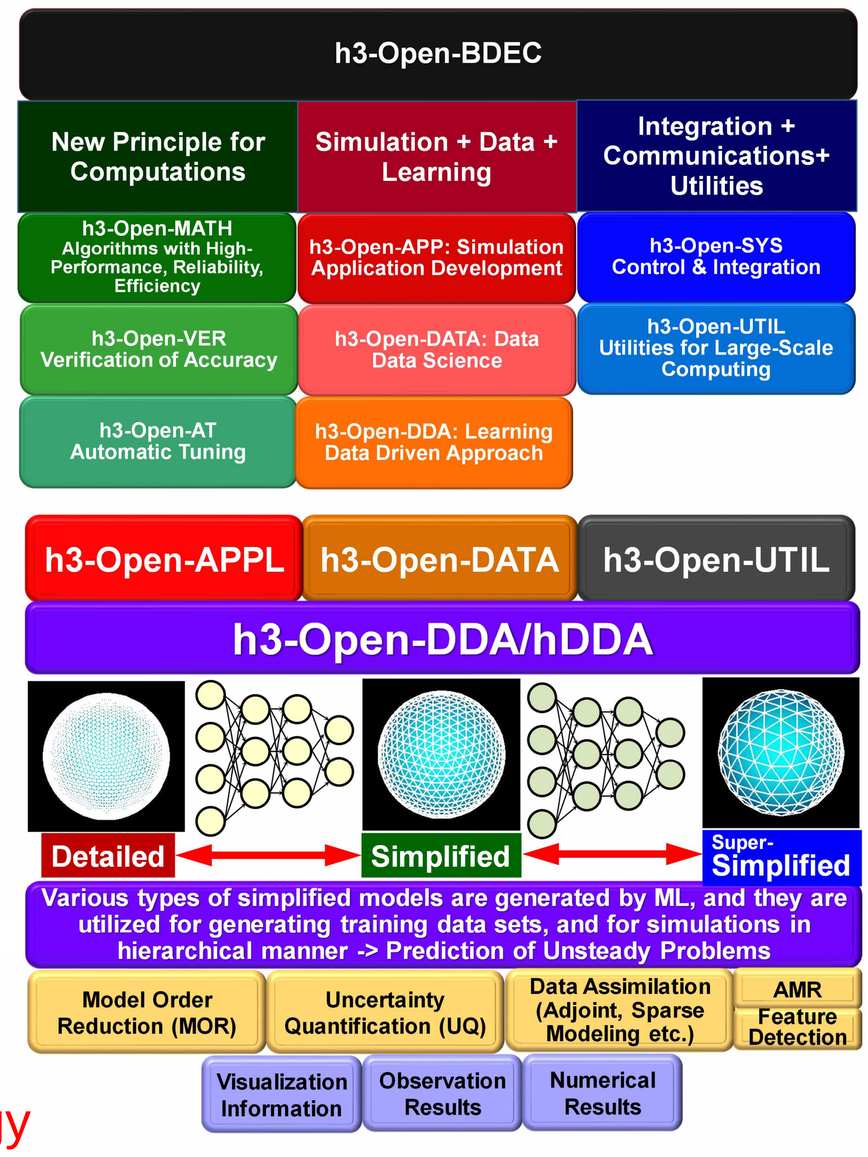
BDEC System (Big Data & Extreme Computing)

- The BDEC System (Big Data & Extreme Computing, 30+PF, 9.55+PB/sec) is a platform for the integration of (S+D+L), and it will be introduced to the Information Technology Center, the Tokyo University in Spring 2021.
- **Simulation Nodes (SIM, 25+PF, Manycore CPU with HBM) for CSE, Data/Learning Nodes (DL, 5+PF, GPU Cluster) for Data Analytics and AI/ML workloads**
  - –Some of the DL nodes are connected to external resources (e.g. data storage, servers, sensor networks, and etc.) directly through an external network (e.g., SINET, Japan).
  - –Hierarchical, Hybrid, Heterogeneous (h3) system



h3-Open-BDEC: Innovative Software Platform for Integration of (S+D+L)

- In this study, we consider the BDEC as the platform for integration of (S+D+L), develop an innovative software platform “h3-Open-BDEC” for integration of (S+D+L), and evaluate the effects of integration of (S+D+L) on the BDEC.
- The h3-Open-BDEC is designed for extracting the max. performance of the supercomputers with minimum energy consumption focusing on:
  - – (1) Innovative method for numerical analysis with high-performance/high-reliability/power-saving based on the new principle of computing by adaptive precision, accuracy verification and automatic tuning, and
  - –(2) Hierarchical Data Driven Approach (hDDA) based on machine learning.
- **Data Driven Approach (DDA)**: Technique of AI/ML is introduced for predicting the results of simulations with different parameters.
  - –DDA generally requires $O(10^3\text{-}10^4)$ runs for generation of training data.
- ***h*DDA (Hierarchical DDA)** : Simplified models with coarser meshes for efficient training are constructed automatically by machine learning with Feature Detection, MOR, UQ, Sparse Modeling and AMR.
- Possible applications on the BDEC with h3-Open-BDEC are combined simulations/data assimilations in climate/weather and real-time disaster simulations, such as flood, earthquake and tsunami.
- h3-Open-BDEC (Adaptive Precision + *h*DDA) enables saving total energy consumption for simulations to 10% of that by conventional methods



The h3-Open-BDEC is the 1st innovative software platform for integration of (S+D+L) on Exascale systems, where computational scientists can achieve such integration without supports by other experts in data analytics and AI/ML.

–Source codes and documents are open to public for various kinds of computational environments.

Contact: Kengo Nakajima <nakajima@cc.u-tokyo.ac.jp>
URL: http://nkl.cc.u-tokyo.ac.jp/