Hewlett Packard Enterprise

HPC for Meterology in the post-Exascale Age

Ilene Carpenter
Earth Sciences Segment Manager
ilene.carpenter@hpe.com

October 10, 2023

Oak Ridge National Laboratory's Frontier Supercomputer

- **74 HPE Cray EX cabinets**
- **9,408 AMD CPUs, 37,632 AMD GPUs**
- **700 petabytes of storage capacity, peak write speeds of 5 terabytes per second using Cray ClusterStor storage system**
- **HPE Slingshot networking cables providing 100 GB/s network bandwidth.**

1

Built by HPE, ORNL's Frontier supercomputer is #1 on the TOP500.

1.1 exaflops of performance.

2

Built by HPE, ORNL's TDS and full system are ranked #2 & #6 on the Green500.

62.68 gigaflops/watt power efficiency for ORNL's TDS system, **52.23 gigaflops/watt** power efficiency for full system.

1

Built by HPE, ORNL's Frontier supercomputer is #1 on the HPL-MxP list.

7.9 exaflops on the HPL-MxP benchmark (formerly HPL-AI).

Source: November 2022 Top500 release, HPL-MxP mixed-precision benchmark (formerly HPL-AI).

HPE Slingshot Wins: Span Verticals, GEOGRAPHIES, and CPU/GPUs

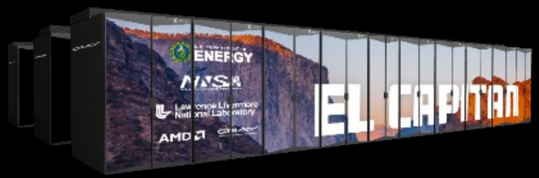
Jun. '23 Top500 w/HPE Slingshot

#1	Frontier (Oak Ridge NL)
#3	LUMI (EuroHPC/CSC)
#8	Perlmutter (LBNL/NERSC)
#12	Adastra (GENCI-CINES)
#17	Setonix-GPU (Pawsey)
#18	Discovery5 (ExxonMobil)
#19	Polaris (Argonne NL)
#30	ARCHER2 (EPSRC/U. of Edinburgh)
#33	Ghawar-1 (Saudi Aramco)
#34	Frontier TDS (Oak Ridge NL)
#59	Derecho CPU Partition (NCAR)
#61	Cactus (GDIT/NOAA)
#62	Dogwood (GDIT/NOAA)
#77	Dardel GPU (KTH Royal Inst. Of Tech)
#79	LANTA (NSTDA)
#83	Narwhal (Navy DSRC)
#101	LUMI-C (EuroHPC/CSC)
#116	rzVernal (LLNL)
#130	Derecho GPU (NCAR)
#132	Tioga (LLNL)
#153	Dardel CPU (KTH)
#156	Warhawk (Air Force Res. Lab.)
#167	Delta (NCSA)
#168	Hotlum (HPE)
#194	Tenaya (LLNL)
#197	Aspire GPU (NSSC)

.... Plus 205, 229, 233, 314, 335, 418

"Frontier"
Oak Ridge National Laboratory

"Aurora"
Argonne National Laboratory



"El Capitan"
Lawrence Livermore Nat'l Laboratory

"LUMI"
EuroHPC JU

"Crossroads"
Tri-Labs

Shaheen KAUST

"Perlmutter"
NERSC

"Setonix"
Pawsey Supercomputing Ctr, Australia

"Alps"
Swiss National Computing Center


"Fawbush" and "Miller"
ORNL (US Air Force Weather)

"Dardel"
KTH Royal Institute of Technology

"Kestrel"
National Renewable Energy Lab (NREL)

Both HPE Cray EX **and Apollo** leadership systems!
Performance with both **HPE Slingshot NIC and Industry NICs**!

Light Blue = HPE Slingshot NIC
Dark Grey = CX5 NIC (demonstrating fabric performance at scale even with standard Ethernet)

Energy Exascale Earth System Model (E3SM)

The first global cloud-resolving model to simulate a year of climate in a day

- Problem: Clouds play a critical role in Earth's climate system. Traditional models can't represent their small overturning circulation in the atmosphere properly.
- Team of researchers from Sandia National Laboratories developed **S**imple **C**loud **R**esolving **E**3SM **A**tmospheric **M**odel (**SCREAM**) model running at a global 3.25-kilometer resolution on the first exascale system in the world, Frontier.

> "We have created the first global cloud-resolving model to simulate a world's year of climate in a day. We're ushering in a new era of accuracy."
>
> **Mark Taylor**, technical lead for Gaea, ORNL

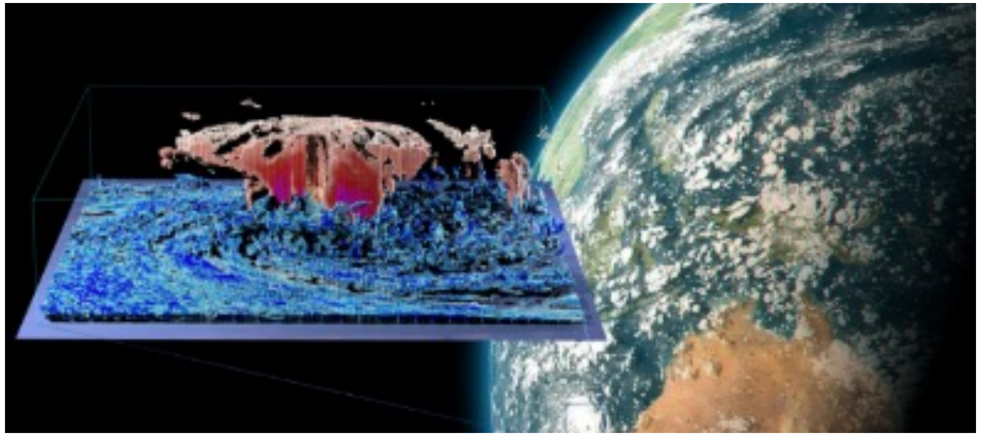
- The research will enable multiyear climate simulations with a more accurate treatment of clouds for the first time.
  - Leads to more accurate predictions of future weather and climate

Additional resources: Read the press release

About the system:

Frontier—The first Exascale system in the world (Oakridge National Laboratory)

- 9000 compute nodes
- AMD MI250X GPUs
- HPE Slingshot interconnect



Cloud predictions—A snapshot from a Simple Cloud Resolving E3SM Atmosphere Model simulation shows a tropical cyclone off the west coast of Australia. The global view displays clouds where the condensed water content is greater than 0.1 grams of water per kilogram of air. The inset shows a 3D cross section with ice mass in red and liquid cloud structure in blue. (Image by Brad Carvey.)

New supercomputer for NOAA climate research

Doubling the power of two previous systems

- Gaea C5—Fifth supercomputer for National Oceanic and Atmospheric Administration (NOAA) installed at and ran by National Climate-Computing Research Center (NCRC) at ORNL.
- C5 is almost doubling the power of previous two Gaea systems (C3 and C4) combined.

> "The power efficiency, cooling efficiency, and CPU power all increase significantly over time. We can replace all of the computational power of C3 with a single cabinet of C5, which has eight cabinets total."
>
> **Paul Peltz**, chief computational scientist of the Energy Exascale Earth System Model, Sandia National Laboratories

- Gaea powers research into the relationship between climate change and extreme weather, such as hurricanes. Unlocking the understanding of what role oceans, which cover nearly three-quarters of the globe, play in our planet's climate.

Additional resources:

Read the press release

Read fact sheet

About the system:

HPE Cray EX—10 Petaflops

- AMD CPU-based
- HPE Slingshot interconnect

Photo – courtesy of InsideHPC

Trends and conditions affecting HPC for meteorology

Hardware diversity

- In processors and accelerators
- In storage (HBM, DDR, NVRAM, SSDs, HDDs)
- CXL

Workload evolution—Rapid increase in AI/ML work

Bitwise reproducibility is an ongoing discussion. Are new approaches needed?

NCAR workshop Nov. 9-10

Data volumes continue to expand.

Supply chain constraints on GPUs

Energy supply constraints and costs

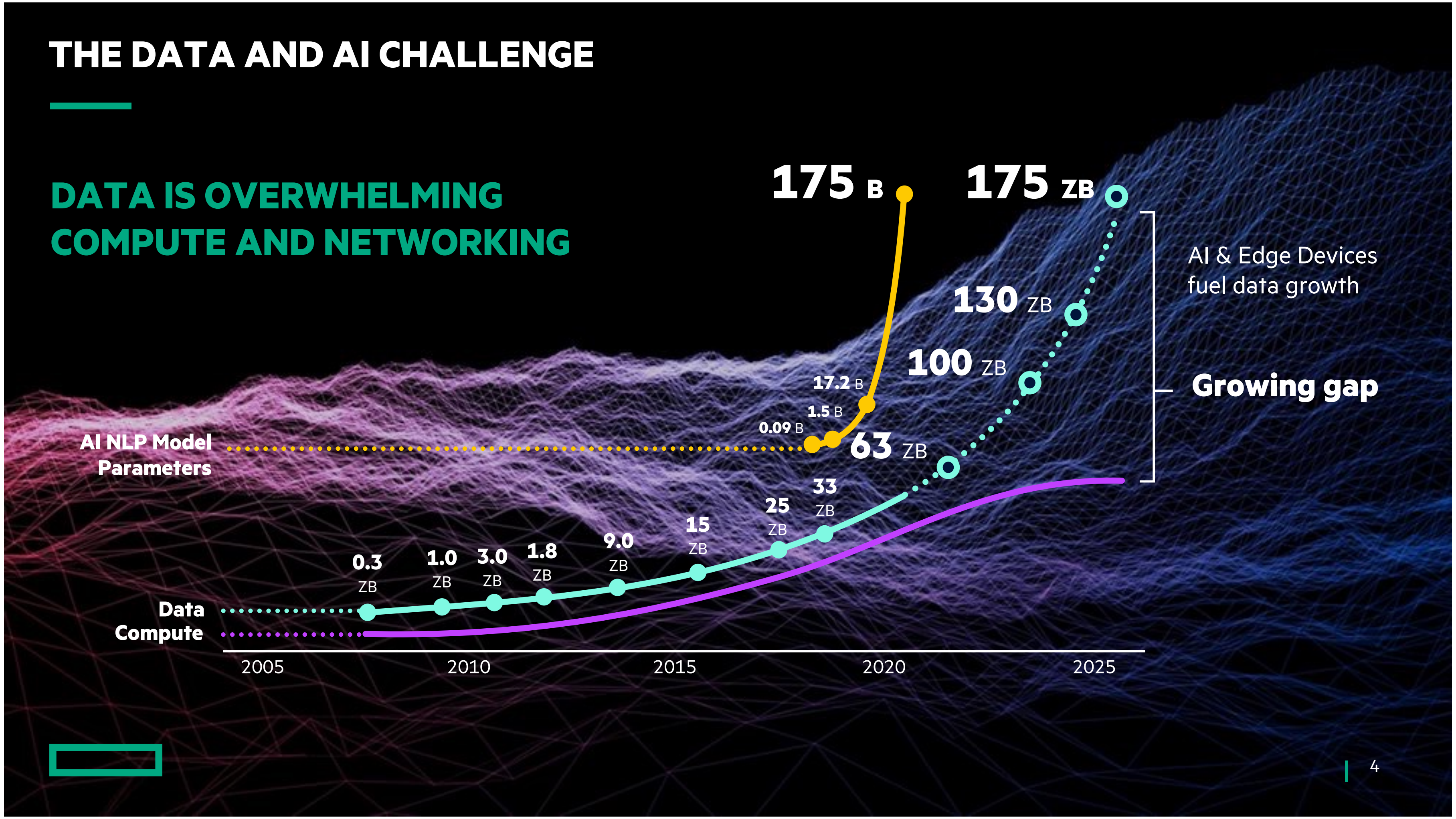
Data center challenges (space, power, cooling, building new or expanding old facilities)

New standards for data services

Hybrid Cloud to maximize value.

Ok, so what’s next?

- Not Zettascale
- Not Exascale for everyone
- Why?
  - Power, space, cost
  - Needs have changed



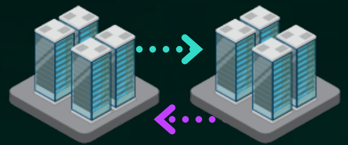
Post-Exascale Vision

Today

Exascale Supercomputer

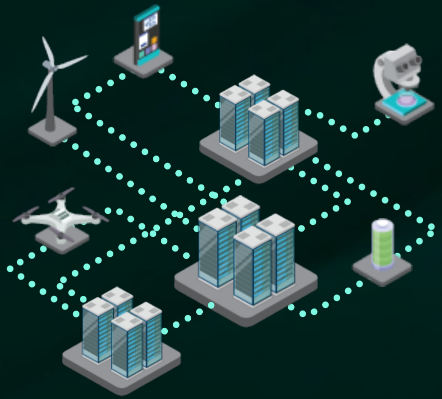
Leadership Class HPC

Productivity and agility for HPC and AI applications



Multi-dimensional, complex workflows

modeling, simulation, data analytics, and artificial intelligence



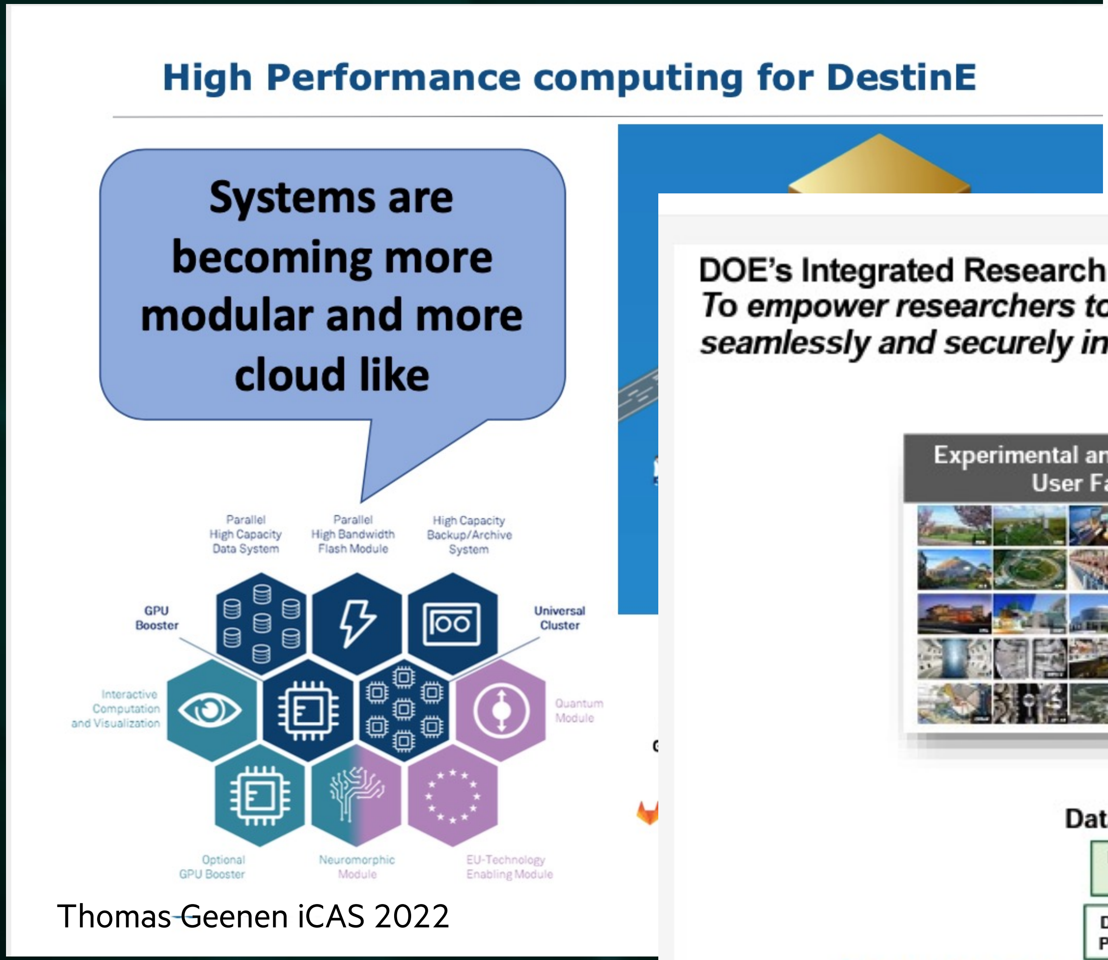
Federated Diverse Systems

Integrate, automate, and optimize workflows that span multiple locations, organizations, and vendors

World's fastest
Supercomputer

World's fastest
Workflows

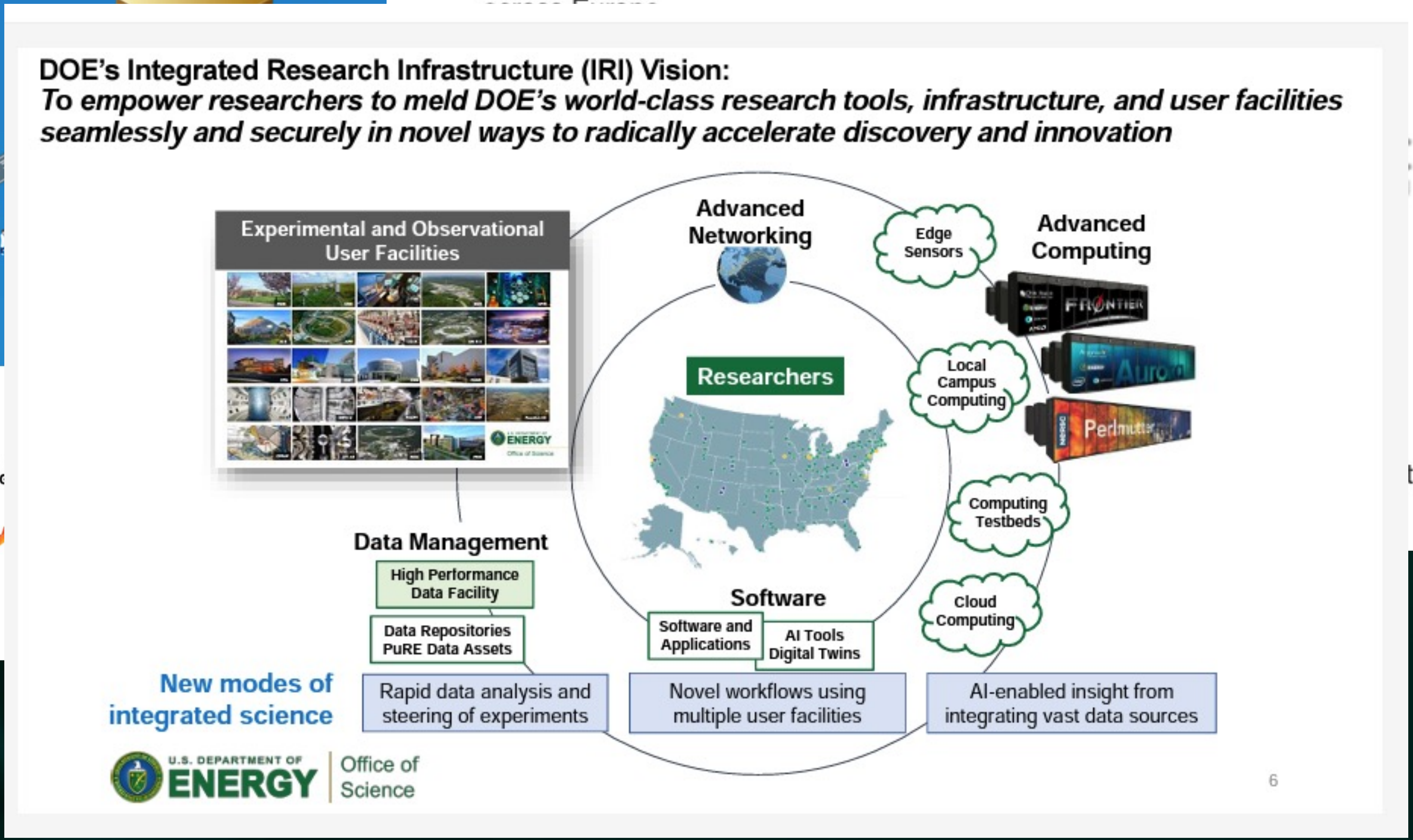
Federated HPC



EuroHPC Launches New Call for Tender Targeting Federation of Supercomputers and Quantum Computers

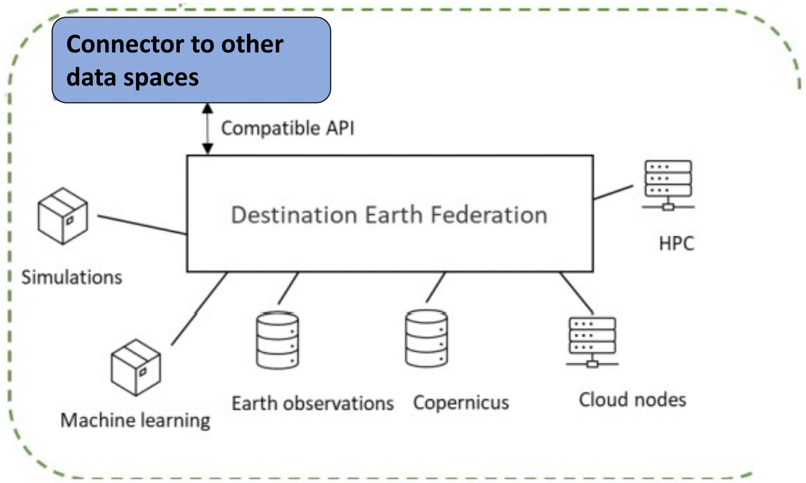
October 6, 2023

Oct. 6, 2023 — The EuroHPC JU has launched a call for tender for the deployment and operation of a platform for federating European resources and providing secure services for a wide range of public and private users



Workflows and data flows

Data handling for the Digital twin Engine



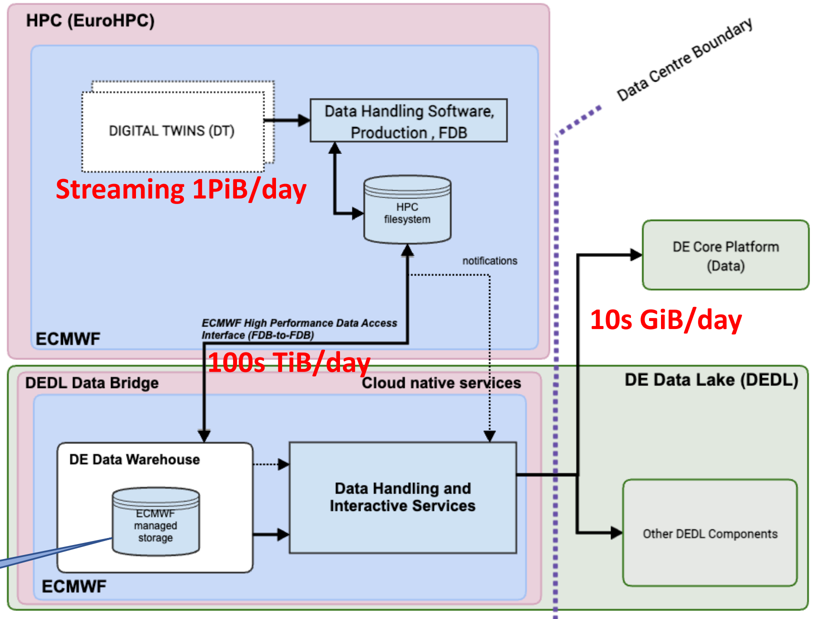
ECMWF

Data handling pipelines

Edge computing: "a part of a distributed computing topology where information processing is located close to the edge, where things and people produce or consume that information"

Streaming data
On the fly consumption (AI/ML)
Store model state not full 4D-resolution

Storage volume is **capped!**
Software solutions to reduce/window output data
Data reconstructed NOT stored



"Hot"
data & checkpoints

ECMWF

Funded by the European Union

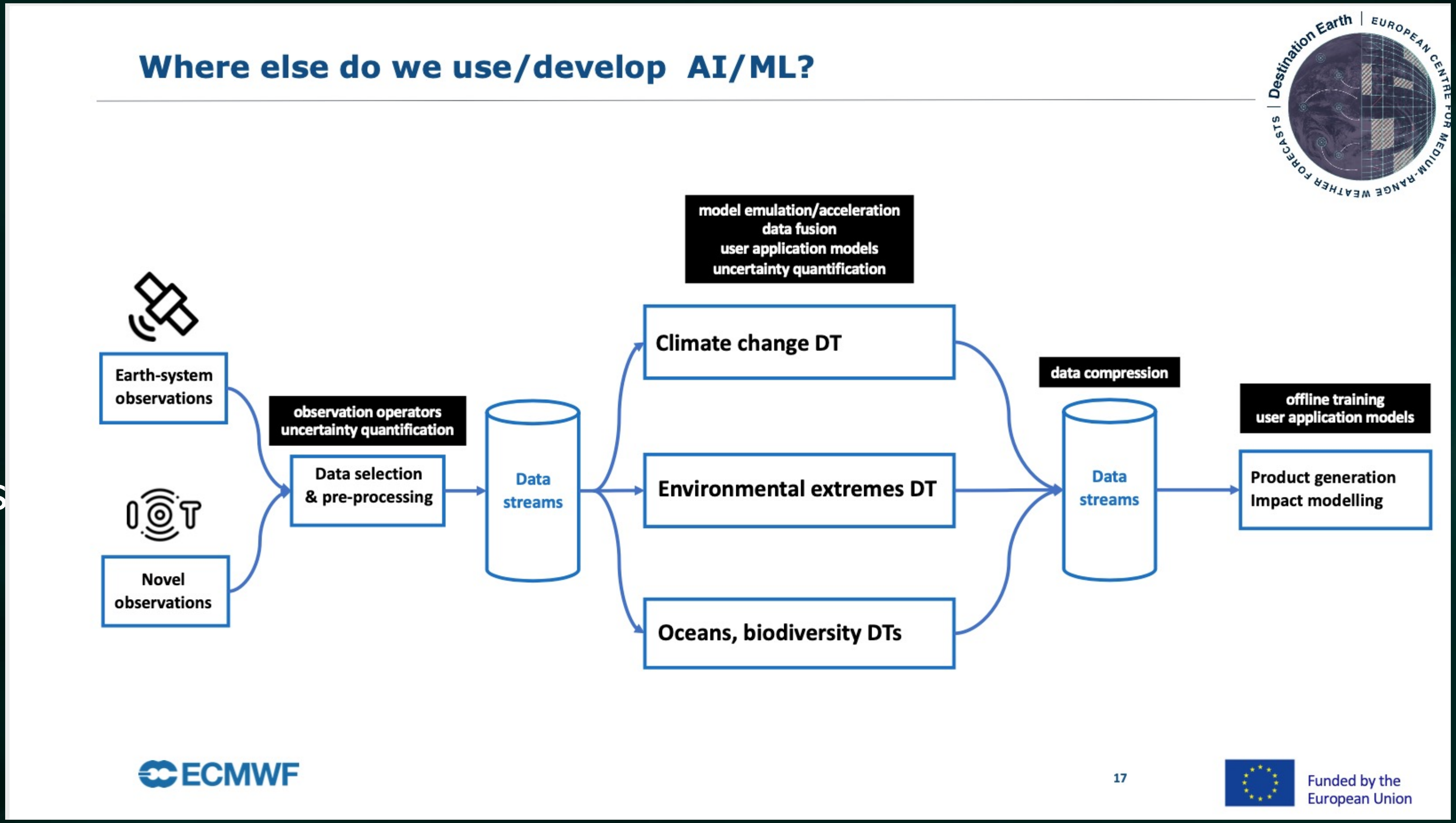
POST-EXASCALE VISION

The ”Interconnect” is the supercomputer

”Workflows” are the new applications

Changing workloads – the AI revolution

- HPC is no longer just for simulations and scientific computing
- AI training is an HPC problem
  - Need fast interconnect for models that run on multiple GPU-enabled nodes
  - GPUs are expensive – need to worry about efficient use
- AI methods are being adopted in traditional scientific computing domains
- What architecture for hybrid ModSim/AI applications?
- What architecture for generative AI models?



Thoman Geenen iCAS2022

arXiv > physics > arXiv:2202.11214

Search...

Help | Advan

Physics > Atmospheric and Oceanic Physics

[Submitted on 22 Feb 2022]

FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators

Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, T Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, Pedram Hassanzadeh, Karthik Kashinath, Animashree Anand

FourCastNet, short for Fourier Forecasting Neural Network, is a global data-driven weather forecasting model that provides accurate medium-range global predictions at 0.25° resolution. FourCastNet accurately forecasts high-resolution, fast-timescale variables suc surface wind speed, precipitation, and atmospheric water vapor. It has important implications for planning wind energy resources, p extreme accurac for large generat of rapid how dat models.

Subjects: At

Cite as: ar

(c

ht

MIT Technology Review

ARTIFICIAL INTELLIGENCE

DeepMind's AI predicts almost exactly wh it's going to rain

The firm worked with UK making short term predic

DeepMind and Google Introduce GraphCast: A Fast and Scalable Machine Learning Weather Simulator

By Tanushree Shenwai - December 30, 2022

Reddit Y in 0 SHARES

People account for the forecasted weather in every aspect of their lives, from choosing an outfit to what to do in the event of a hurricane. Forecasting over a time frame that is typically three to seven days out is referred to as medium-range forecasting. Several sectors, like agriculture, construction, travel, etc., rely on "medium-range" weather forecasts for making decisions, which are offered up to four times daily by weather bureaus like the European Centre for Medium-Range Weather Forecasts (ECMWF).

There are two primary parts to medium-range weather forecasts, both simulated using massive high-performance computing (HPC) clusters. The first part is "data assimilation," which is the method of forecasting weather conditions by analyzing current and historical

Article

Skilful precipitation nowcasting using deep generative models of radar

https://doi.org/10.1038/s41586-021-03854-z

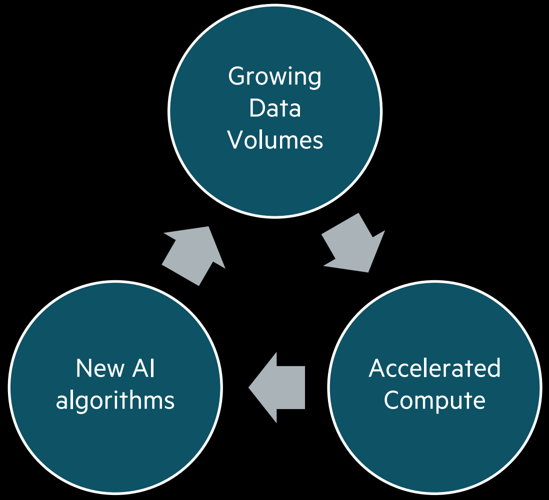
Received: 17 February 2021

Accepted: 27 July 2021

Suman Ravuri[1,5], Karel Lenc[1,5], Matthew Willson[1,5], Dmitry Kangin[2,3], Remi Lam[1], Piotr Mirowski[1], Megan Fitzsimons[2], Maria Athanassiadou[2], Sheleem Kashem[1], Sam Madge[2], Rachel Prudden[2,3], Amol Mandhane[1], Aidan Clark[1], Andrew Brock[1], Karen Simonyan[1], Raia Hadsell[1], Niall Robinson[2,3], Ellen Clancy[1], Alberto Arribas[2,4] & Shakir Mohamed[1]✉

AI AT SCALE REQUIRES NEW SOFTWARE PLATFORM

Emerging AI Mega Trends



Difficult Ecosystem Choices

DIY using hundreds of point solutions – most without commercial support

or

Adopt CSP or accelerator-vendor provided technologies that create lock-in

Few Successful Implementers

Big Tech Companies ✓
(e.g., Alphabet, Meta)

AI Native Companies ✓
(e.g., Open AI, Cruise, Aleph Alpha)

Majority of organizations ✗

Market Requirements

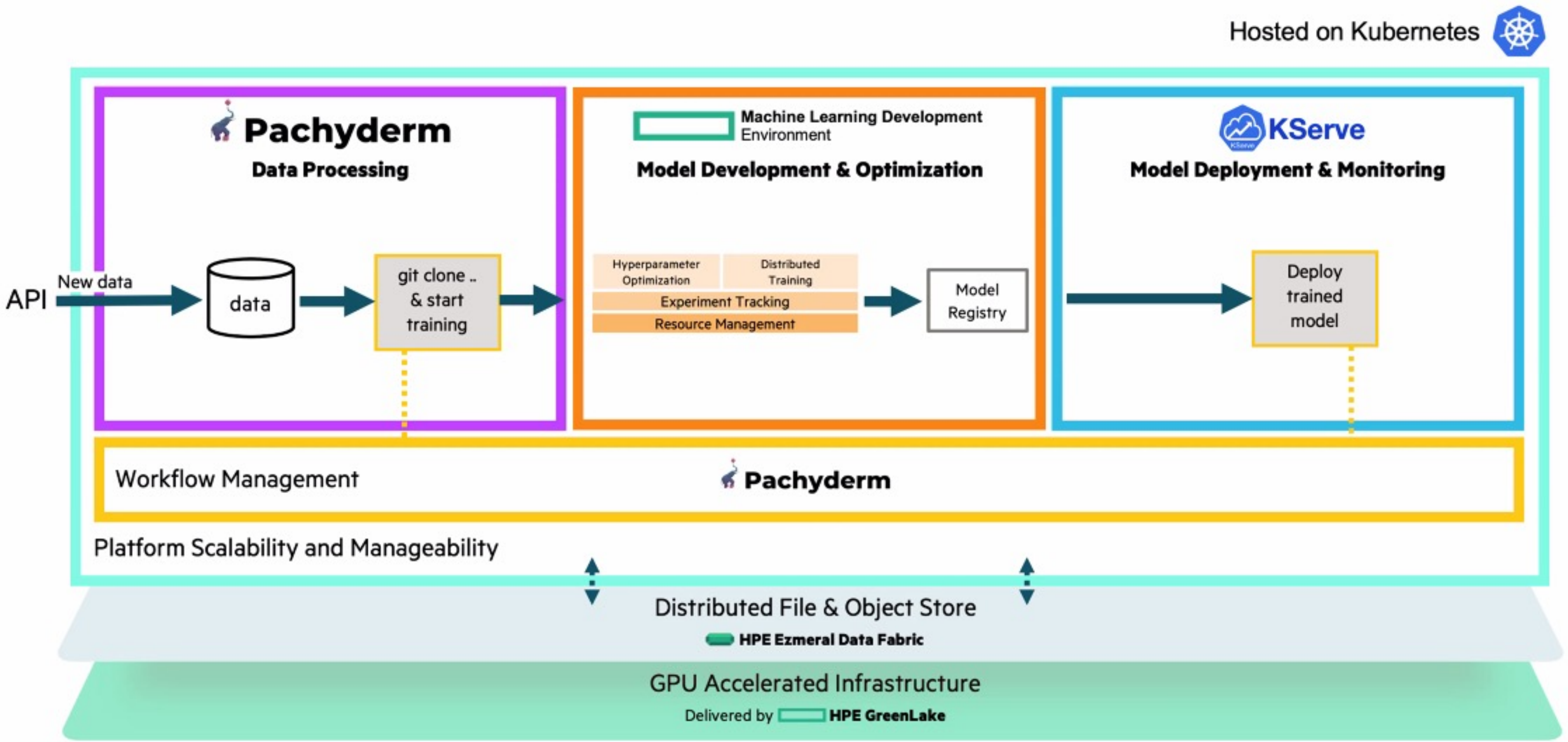
AI lifecycle software based on open technologies and built for scale

End-to-end capabilities:
- Data Acquisition & Preparation
- Development & Training
- Deployment & Inference
- Governance & Performance Management

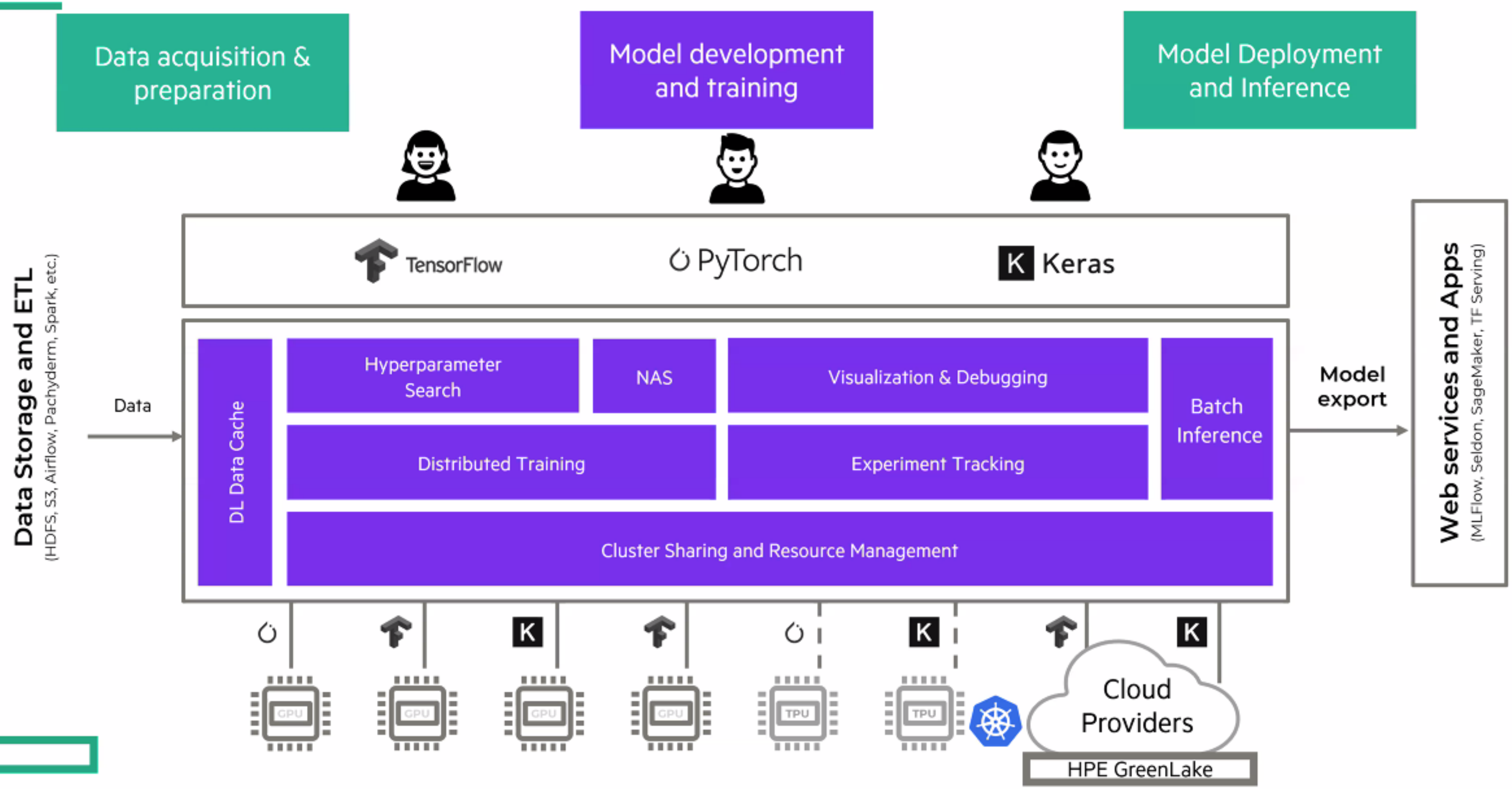
Common user experience with deployments from edge to cloud

Optimized performance across heterogenous compute

Early Innings of ML



HPE Machine Learning Development Environment built on Determined AI



Hewlett Packard Enterprise

SMARTSIM

SMARTSIM: ENHANCING HPC WORKFLOWS WITH MACHINE LEARNING

Andrew Shao, PhD | Senior HPC&AI Research Scientist, HPE
13 July 2023
andrew.shao@hpe.com

CHALLENGES IN EMBEDDING ML INSIDE AND OUTSIDE OF SIMULATIONS

Technical

- How can we call an ML model from a Fortran/C/C++ code?

- How can we scale this to the needed compute scales?
  - ML models in physics simulations tend to be small
  - Want to efficiently allocate GPU resources

These technical challenges impede scientific progress!

- Naturally focuses the community on how to replace simulations with AI
- AI to solve physics problems tends to be in idealized contexts or static data
- Scientists are asking the questions, but do not have the tools to test this practically

SMARTSIM ENABLES AI-ENHANCED SIMULATIONS

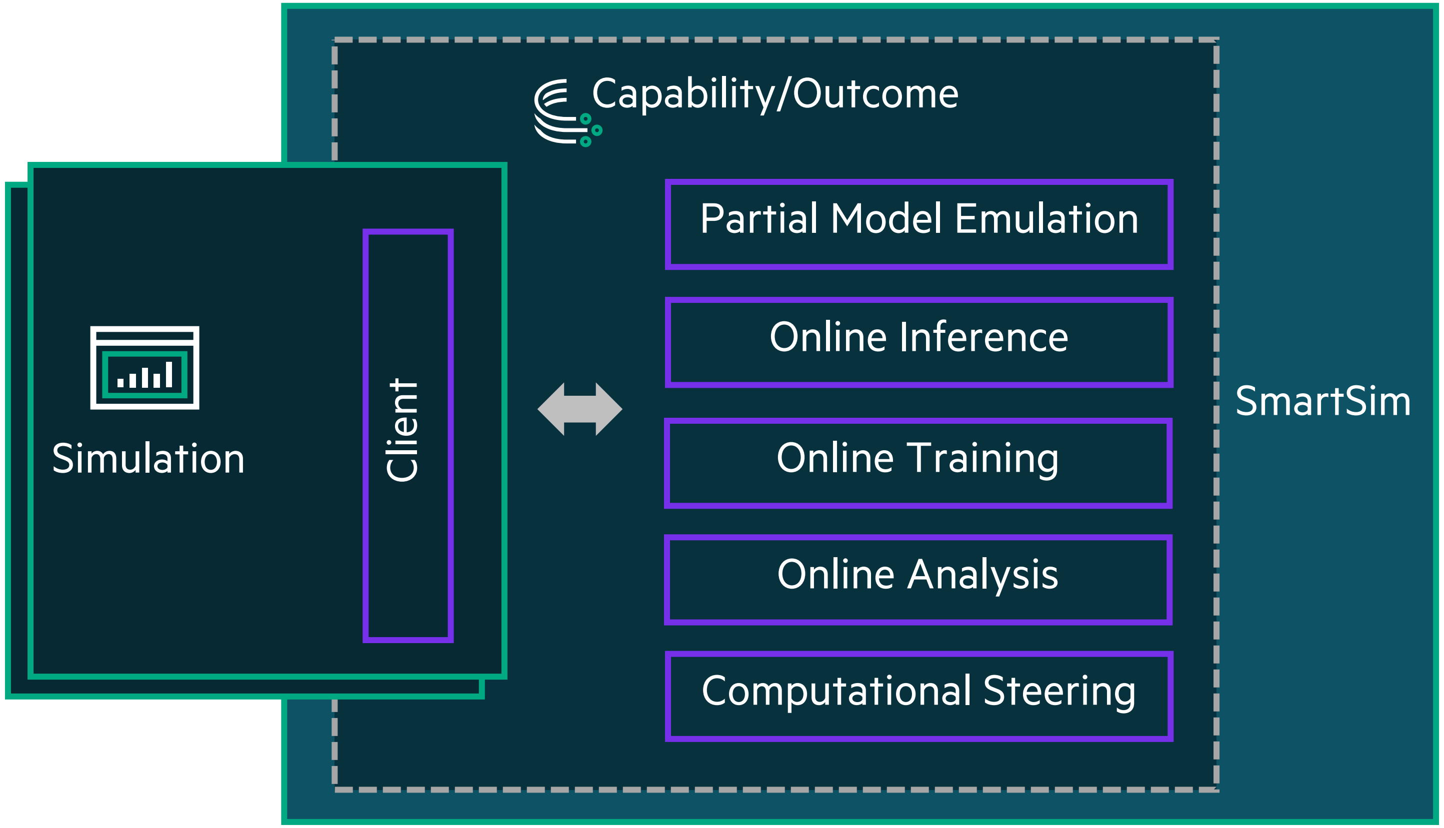
Lower the barriers to entry

- Enable Fortran, C, C++ simulations to interact with ML packages efficiently
- Rapidly prototype and iterate with new ML models
- Allow scientists to focus on building applications *not* code or infrastructure

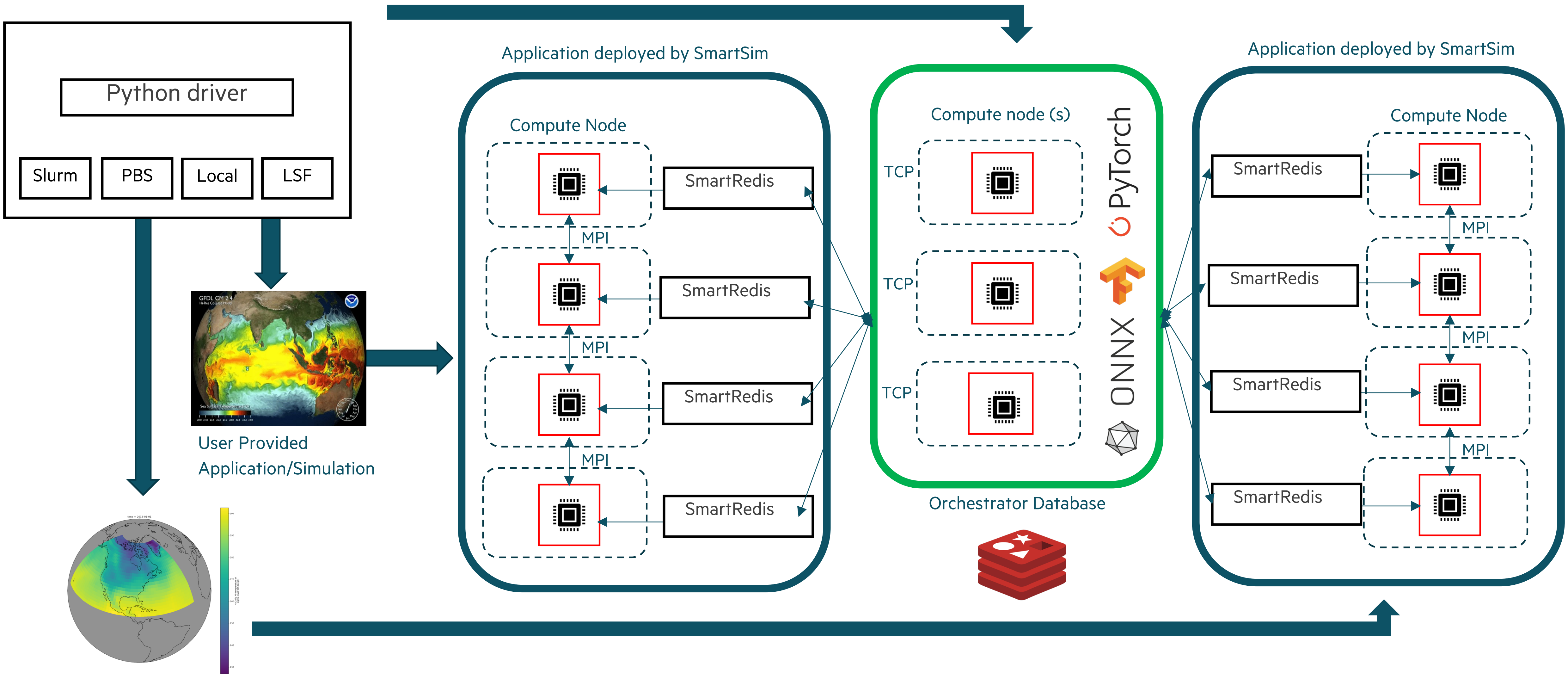
Introduce a new workflow paradigm

- Simulations as producers and consumers of data and ML models
- Data is kept in-memory and available in-flight

GOAL: Enable new tools for scientific discovery using ML methods and scientific simulations



CLUSTERED DATABASE



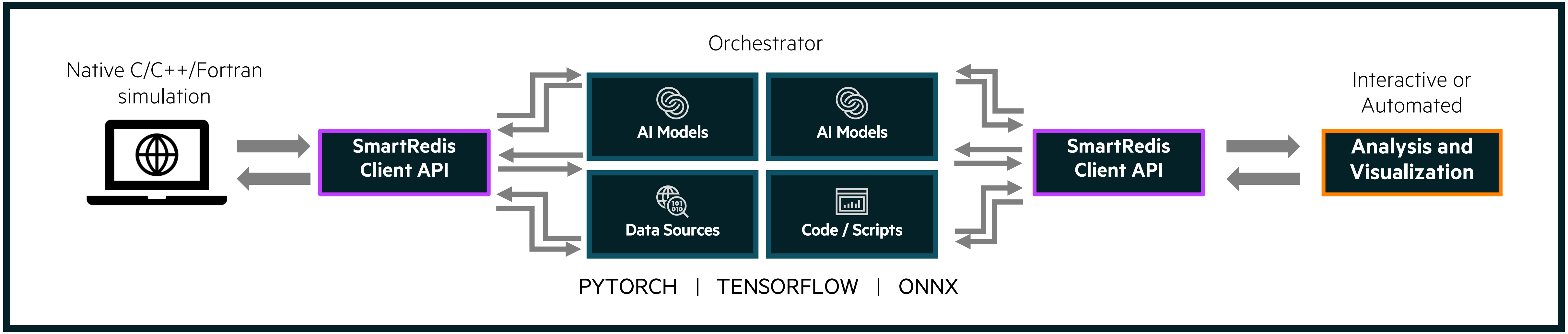
ABOUT SMARTSIM

The SmartSim open-source library bridges the divide between traditional numerical simulation and data science

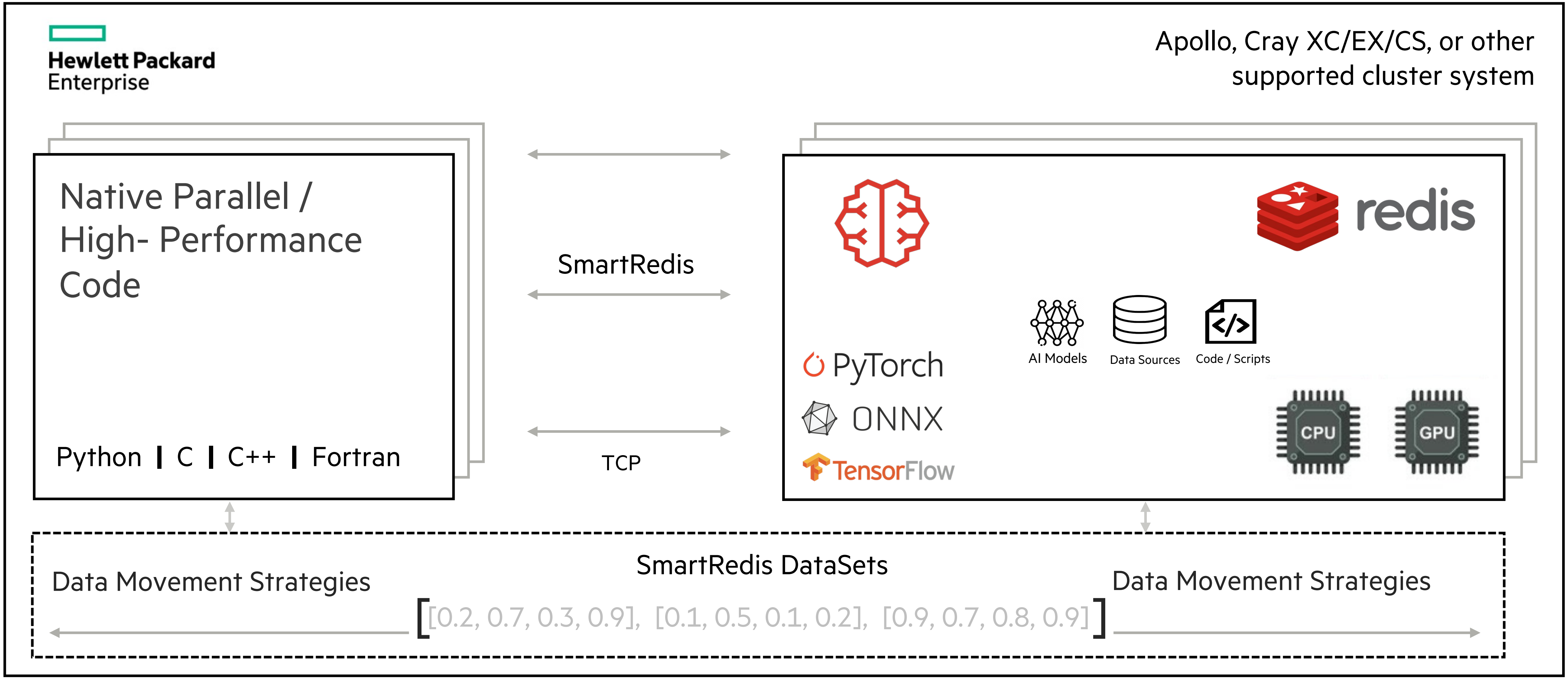
SmartSim enables simulations to be used as engines within a system, producing data, consumed by other services to create **new applications**

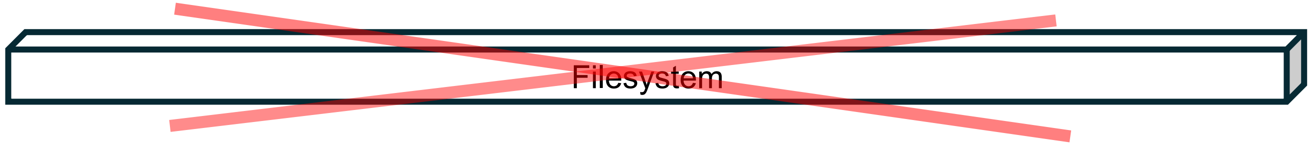
- Use Machine Learning (ML) models in existing Fortran/C/C++ simulations
- Communicate data between C, C++, Fortran, and Python applications
- Train ML models and make predictions using TensorFlow, PyTorch, and ONNX
- Analyze data streamed from HPC applications while they are running

All of these can be done without touching the filesystem, i.e., **data-in-motion**



HIGH-LEVEL ARCHITECTURE

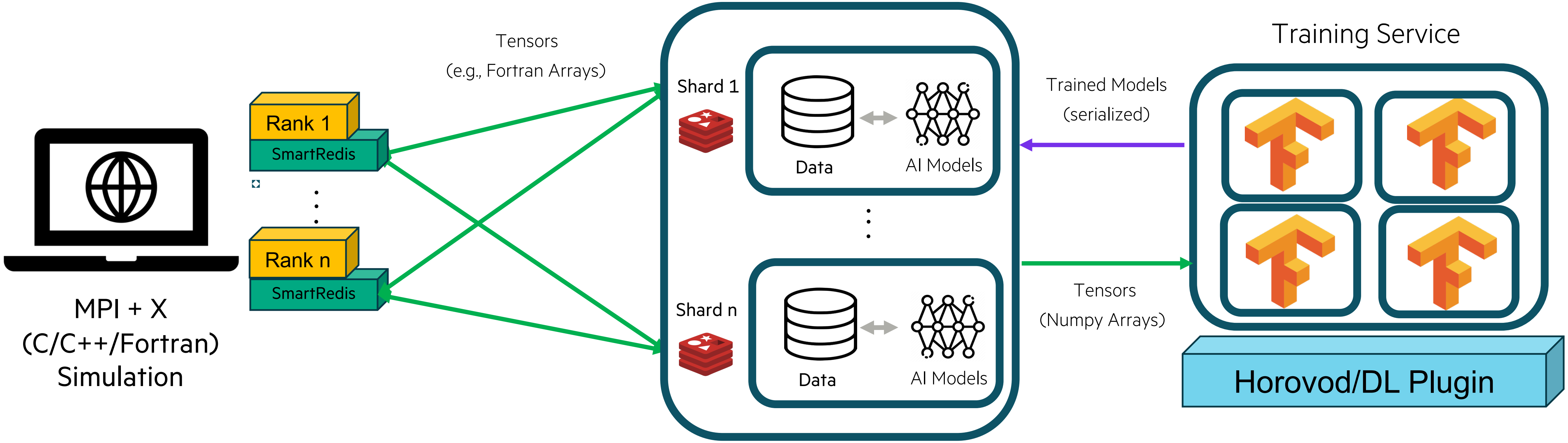




CONTINUOUS ONLINE TRAINING

I want to create machine learning models on simulation data without performing application collectives or writing to the filesystem

- SmartRedis Client can be used inside DataLoaders for TensorFlow and PyTorch to perform Stream Training
- Trained models can be checkpointed, saved and later sent to the database for online inference.
- Can be set from any language and called from any language (e.g., model set from Python, called from Fortran)

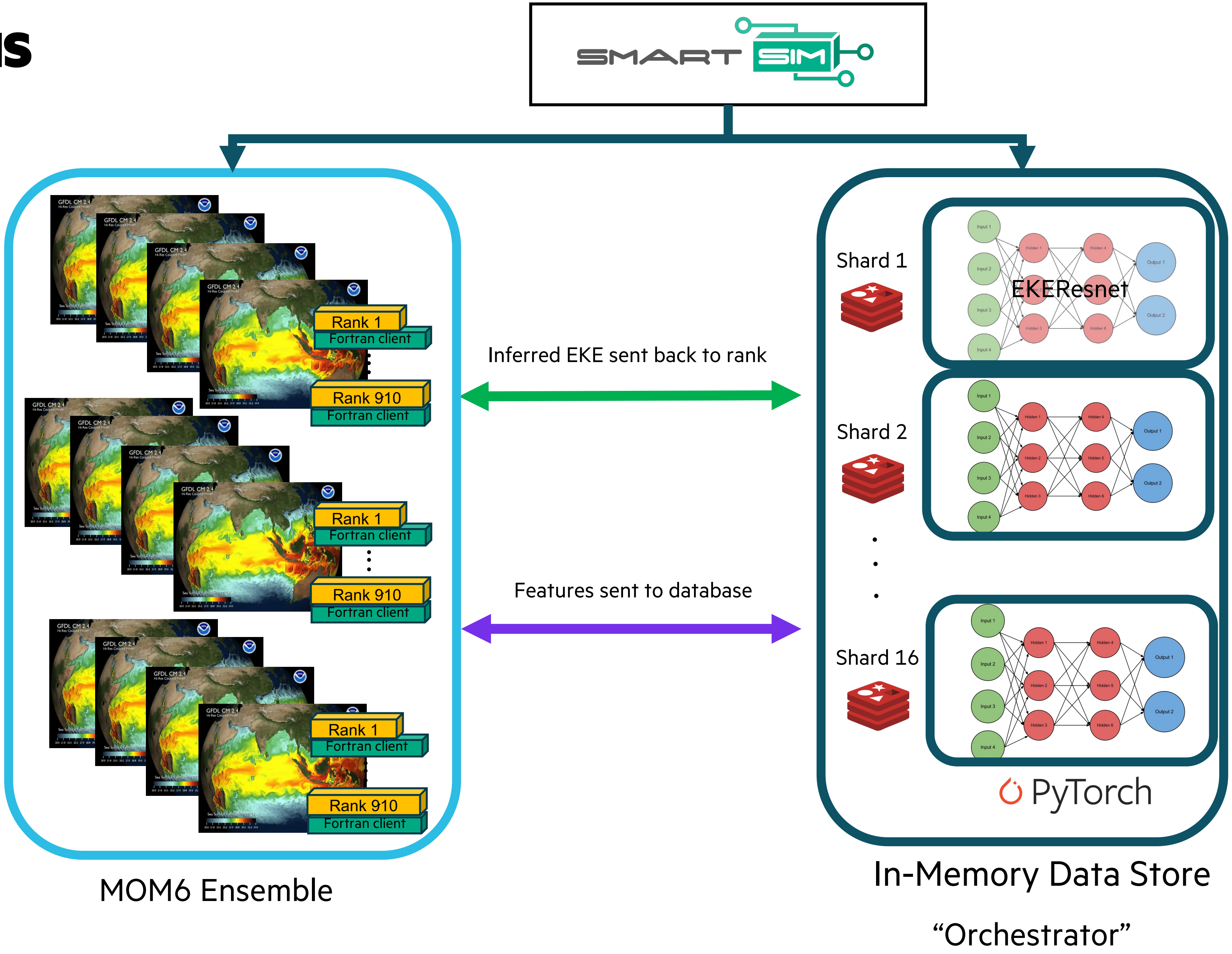


AI IN OCEAN SIMULATIONS

Goal: Improve representation of oceanic turbulence for quasi-LES simulations

Solution:
- Train machine-learning model to predict sub grid-scale turbulent kinetic energy
- Model 1: Use predictions within the model of eddy turbulence
  - 12 ensemble members
  - 10,920 CPUs, ~200 CPU-only nodes
  - 16 P100 GPUs

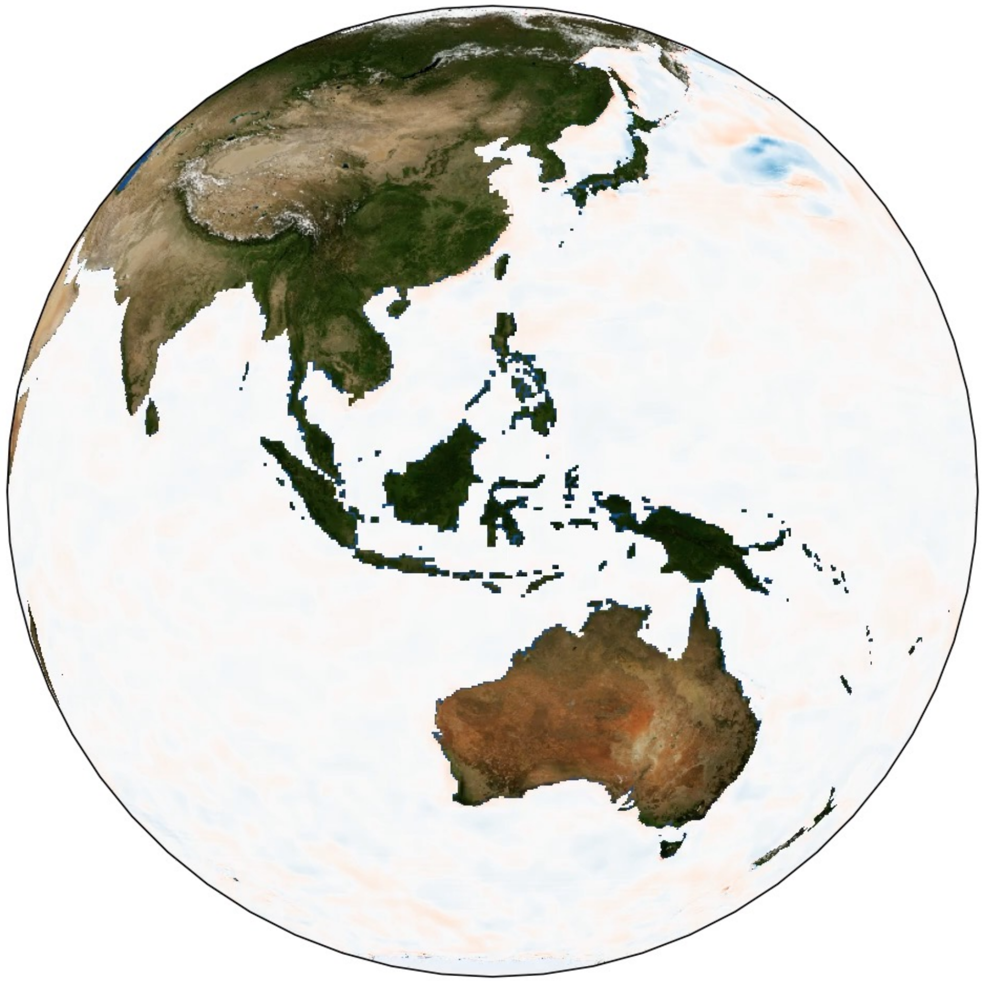
Milestone: First demonstration of AI embedded in the solver of a realistic ocean simulation



Collaboration between HPE/National Center for Atmospheric Research
Partee et al. [2022], doi.org/10.1016/j.jocs.2022.101707

DEPLOYING A SECOND TURBULENCE MODEL ON FRONTIER

- Prologue
  - Collaborators had a small, one-node ocean simulation with CNN parameterization
  - Successfully integrated SmartSim into MOM6
  - Cluster available to them did not have enough nodes to support a 'realistic' simulation
- June 12:
  - SmartSim engineers get access to allocation on Frontier
  - Install SmartSim from scratch
  - Run base case
- June 13:
  - Start doing scaling studies using SmartSim ensembles to understand how to efficiently use Frontier Resources
- June 14:
  - Run one year of realistic simulation on 50 nodes of Frontier
- June 15
  - Run application on 672 nodes of Frontier (~20,000 CPUs and 5,000 AMD GPUs)



DOCUMENTATION

www.craylabs.org

SmartSim 0.4.1 documentation

Search the docs ...

Versions

Contents

Library Design

Introduction

SmartSim enables scientists to utilize machine learning inside traditional HPC workloads

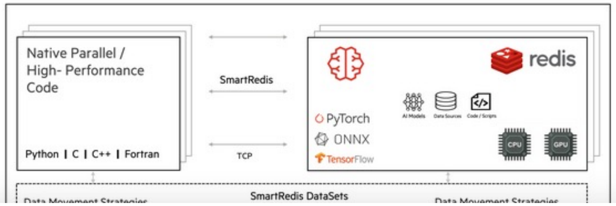
SmartSim provides this capability by

1. Automating the deployment of HPC workloads and distributed, in-memory storage (Redis).
2. Making TensorFlow, Pytorch, and ONNX callable from Fortran, C, and C++ simulations.
3. Providing flexible data communication and formats for hierarchical data, enabling online analysis, visualization, and processing of simulation data.

The main goal of SmartSim is to provide scientists a flexible, easy to use method for interacting at runtime with the data generated by simulation. The type of interaction is completely up to the user.

- Embed calls to machine learning models inside a simulation
- Create hooks to manually or programmatically steer a simulation
- Visualize the progression of a simulation integration from a Jupyter notebook

The figure below shows the architecture of SmartSim for a given use case. SmartSim can create, configure and launch workloads (called a `Model`), as well as groups of workloads (`Ensembles`). The data communication between a workload and in-memory storage is handled by the SmartRedis clients, available in Fortran, C, C++, and Python.



Theme by the Executable Book Project

Contents

SmartSim API

Experiment

`Experiment.__init__`(name[, exp_path, launcher])	Initialize an Experiment instance
`Experiment.start`(*args[, block, summary, ...])	Start passed instances using Experiment launcher
`Experiment.stop`(*args)	Stop specific instances launched by this `Experiment`
`Experiment.create_ensemble`(name[, params, ...])	Create an `Ensemble` of `Model` instances
`Experiment.create_model`(name, run_settings)	Create a general purpose `Model`
`Experiment.create_database`([port, db_nodes, ...])	Initialize an Orchestrator database
`Experiment.create_run_settings`(exe[, ...])	Create a `RunSettings` instance.
`Experiment.create_batch_settings`([nodes, ...])	Create a `BatchSettings` instance
`Experiment.generate`(*args[, tag, overwrite])	Generate the file structure for an `Experiment`
`Experiment.poll`([interval, verbose, ...])	Monitor jobs through logging to stdout.
`Experiment.finished`(entity)	Query if a job has completed.
`Experiment.get_status`(*args)	Query the status of launched instances
`Experiment.reconnect_orchestrator`(checkpoint)	Reconnect to a running `Orchestrator`
`Experiment.summary`([format])	Return a summary of the `Experiment`

```
class Experiment(name, exp_path=None, launcher='local')
```
[source]

Bases: `object`

Experiments are the Python user interface for SmartSim.

Experiment is a factory class that creates stages of a workflow and manages their execution.

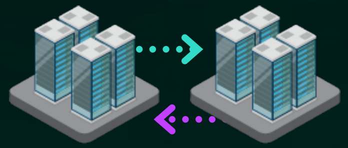
Post-Exascale Vision

Today

Exascale Supercomputer

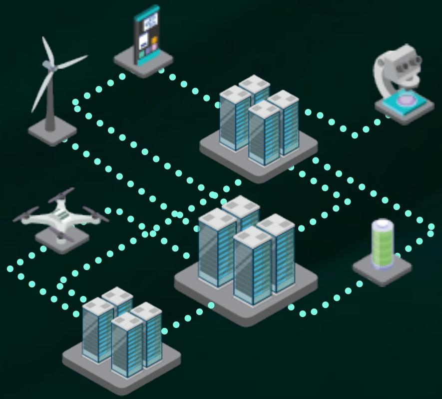
Leadership Class HPC

Productivity and agility for HPC and AI applications



Multi-dimensional, complex workflows

modeling, simulation, data analytics, and artificial intelligence



Federated Diverse Systems

Integrate, automate, and optimize workflows that span multiple locations, organizations, and vendors

World's fastest **Supercomputer**

World's fastest **Workflows**

Thank you

Ilene.Carpenter@hpe.com