DIGITAL TWINS FOR SCIENCE



SCIENTIFIC COMPUTING IS EVOLVING Thought Leaders Mapping Out the Opportunity and Constraints Given Current Market Reality

COMPUTING PERVADES ALL aspects of societ ways once imagined by only a few. Within s and engineering, computing has often bee the third paradigm, complementing theor experiment, with big data and artificial int (AI) often called the fourth paradigm, "Spa data analysis and disciplinary and multidi modeling, scientific computing systems h ever larger and more complex, and today's scientific computing systems rival global facilities in cost and complexity. However well in the land of scientific computing. In the initial decades of digital computi government investments and the insight designing and deploying supercomputer shaped the next generation of mainstream consumer computing products. 82 COMMUNICATIONS OF THE ACM | FEBRUARY 2023 | VOL.ES | NO. 2

BY DANIEL REED, DENNIS IN

Trections of the the teel, dennis Gan of teel, den

IN NATIONAL ENGLICIONAL ENGLICONAL E

Charting a Path in a Shifting Technical and Geopolitical Landscape Post-Exascale Computing for the National Nuclear Security Administration

AI FOR SCIENCE

RICK STEVENS VALERIE TAYLOR Argonne National Laboratory July 22-23, 2019

JEFF NICHOLS **ARTHUR BARNEY MACC/**

Oak Ridge National Laboratory August 21-23, 2019

KATHERINE YELICK **DAVID BROWN**

Lawrence Berkeley National Laboratory September 11-12, 2019.

Foundational Research Gaps and Future Directions for Digital Twins



A DEEPER LOOK AT THE NEW HPC ADAPTING TO THE CONSTRAINTS, MARKET REALITIES AND EXPANDING OPPORTUNITY



ACM HPC Forecast **Reed and Dongarra**

Semiconductor Constraints limit the potential increase in scale for legacy algorithms

OVERARCHING FINDING: The combination of increasing demands for computing with the technology and market challenges in HPC requires an intentional and thorough reevaluation of algorithms, software development, system design, computing platform acquisition, and workforce development.

- New Algorithms offer potential for dramatic increase in model scale and reduction in latency
- Cloud Economics have changed the supply chain ecosystem

Charting a New Path Post Exascale Computing NNSA



EDGE







TRANSITION TO POST EXASCALE ERA

	EXPERIMENTS SIMULATION Viz	EDGE		
FEATURE	TERA THROUGH EXASCALE			
USAGE	BATCH & MOSTLY LOCAL TO A SITE			
WORKLOAD	SINGLE SIMULATION/ENSEMBLE	WORKFLOW COMPRI		
EXPERIMENTS	OFFLINE DATA ANALYSIS FOR EXPERIMENTS			
DIGITAL TWINS	IN-SITU VISUALIZATION OFFLNE	INTERACTIVE_C		
QUANTUM COMPUTING	SIMULATION			
PROGRAMMING MODELS	FORTRAN, C++, MPI, OPENMP	STANDARD PA		
SYSTEM CONFIGURATION	MONOLITHIC			
CLOUD	GRID	BURST CAPABILIT		



POST EXASCALE

INTERACTIVE & DISTRIBUTED WITH MULTIPLE SITES

ISED OF SIMULATION ENSEMBLES, AI TRAINING AND INFERENCE, LIVE DATA ANALYTICS

MIX OF REAL-TIME ANALYSIS, STEERING AND OFFLINE

COMBINATION OF SIMULATION AND OBSERVATIONAL DATA

SIMULATION PREPARING FOR A HYBRID MODEL

ARALLELISM SUPPORT IN FORTRAN, C++, MPI, OPENMP, OPENACC, PYTHON, JULIA, PYTORCH, TENSORFLOW

MODULAR

FIES, FASTER REFRESH CYCLE, ACCESS TO LATEST TECHNOLOGY AT SCALE



DIGITAL TWIN FOR SCIENCE? A Relatively New Modeling Concept That is Just Emerging for Science

Digital Twin Definition from Wikipedia

Digital model of an intended or actual real-world physical product, system, or process (a physical twin) that serves as the effectively indistinguishable digital counterpart of it for practical purposes.

The concept and model of the digital twin was first publicly introduced in 2002 by Michael Grieves, at a Society of Manufacturing Engineers conference as the conceptual model underlying Product Life Cycle Management

National Academy of Science Report: Foundational Research Gaps and Future Directions:

A digital twin is a set of virtual information constructs that mimics the structure, context, and behavior of a natural, engineered, or social system (or system-of-systems), is dynamically updated with data from its physical twin, has a predictive capability, and informs decisions that realize value.

The bidirectional interaction between the virtual and the physical is central to the digital twin.

<u>A Digital Twin for Science</u>

A digital model with sufficient scale and accuracy to be indistinguishable from the real-world physical object, system or process that is dynamically updated with data from it's physical twin, where the model has sufficient fidelity to meet the requisite accuracy with time to solution to inform decisions that improve the operation of the asset or process being modeled



ML ENABLES REQUISITE ACCURACY/TIME TO SOLUTION Allows the Model to be Indistinguishable and Interactive



Log computational cost



Exact solution





ML POTENTIAL TO BRIDGE MODELS WITH CONTROL



LL

Log computational cost





EXAMPLES OF DIGITAL TWINS FOR SCIENCE Collaborate with the Global Research Community to Pursue Science Discovery that Benefits Mankind



Towards Real time Fusion Reactor Design Generative AI to Predict Disruption



Real Time Multi-Messenger Astrophysics



Large Hardon Collider



Earthquake Model with Machine Learning



Genome Scale LLMs for Covid







Destination Earth



FUSION DIGITAL TWIN WITH CONVENTIONAL SIMULATION



https://www.youtube.com/watch?v=tJgR1TSBD0k



FUSION DIGITAL TWIN WORKFLOW WITH SGTC SURROGATE File-Based Prototype





SUPERFACILITY WITH NERSC AND GA DIII-D FUSION REACTOR Couple EFIT Simulation with Experiment operation



Current Conventional Model with the Twin





GA Ion Orbiter DIII-D Vacuum Vessel Impacts Postprocessing/Reprojection pipeline -> ParaView





GA Ion Orbiter DIII-D Vacuum Vessel Impacts Postprocessing pipeline -> ParaView -> Omniverse



								7	LIVE	CACHE:	OFF 皆		
Stage Lights	Stage	Layer	Rende	r Settings									
	Search												
Frame time: 9.23 ms	Name (Old to	New)							۲	Туре			
ed, 40.8 GiB available	🗖 🙏 Root (defaultPrin	n)						۲	Xform			
d, 211.3 GiB available	Actors								۲	Xform			
12/512 spp : 6.31 sec	TableToPoints1_Surface								۲	Xform			
1280x720	- ŭ	0200121	HPH SA	n_i S I HCD a	ligned	remesh30	stl Surface		۲	Xform	icer		
		Mesh0	ieom_1			,	202001000		۲	Mesh			
	🗖 🙏 Lo	oks											
	TableToPoints1_Surface_Mat_0									Material			
	Image: Contract of the second seco									Material			
	Aluminum_Scratched_Dark_Matte									Material			
		defaultLig	ht						3 0	DistantLigh	nt		
	🗖 🙏 Enviro	nment							۲	Xform			
	📜 🤘 sk	y.							۲	DomeLight			
and the owner where the party is not													
	Property												
	+ Add defaultLight												
	Prim Path /Root/Lights/defaultLight												
	▼ Transform	r i											
	Translate	5	<	0.0		Y	0.0	- z		0.0			
	D			245.0			0.0						
	Rotate *		2	315.0	- 1		0.0	4		0.0			
	Scale 🖼	2	<	1.0		M	1.0	Z		1.0			
	▼ Light												
	-												
	▼ Main					_							
🌣 🗉	Color			R 1.	0	G	1.0	В	1.0				
torm	Enable Color Temperature										ě.		
torm	Color Temperature 6500.0												
	Intensity 0.3						0.3						
	Exposure						125						
	exposure						13.5						
	Normalize F	ower											
	Angle 1.0												
	Diffuse Multiplier 1.0												
	Charles M.	ultislier											

GA Ion Orbiter DIII-D Vacuum Vessel Impacts





in the last



GA PLAN: DIGITAL TWIN ENABLED CONTROL ROOM

CONVERGED WORKFLOW FOR GYROKINETIC FUSION Current StellFoundary Project Funded by DOE SCIDAC 2924 in 2023

indicate the physics calculations, while the off-diagonal parallelograms show the data passed between codes.

ALGORITHMS EVOLVING AT UNPRECEDENTED PACE FourCastNet High Resolution for Data-Driven Weather Models

Comparison of resolutions for data-driven weather models since 2018 (Dueben & Bauer)

SOTA evolving rapidly Recent Pre-print Kang Chen et al (2023) extend forecast to 10 days with 0.25° resolution using "cross modal Transformer"

Weyn et al. (2019), 2.5° N.H only, 72x36, 2.6k pixels, ConvLSTM

WeatherBench, Rasp et al. (2020). 5.625°, 64x32, 2K pixels, CNN

Deuben & Bauer (2018), 6°, 60x30, 1.8K pixels, MLP

FourCastNet, Pathak et al. (2022), 0.25°, ~1,000,000 Pixels, ViT+AFNO

GNN, Keisler et al. (2022), 1°, 64,000 Pixels, Graph Neural Networks

DLWP, Weyn et al. (2020). 2°, 16K pixels, Deep CNN on Cubesphere/(2021) ResNet

33

DIGITAL TWIN WITH EARTH2 https://www.youtube.com/watch?v=8cQoYcbUG_M

Our Vision: Smart "Factory" for Probing & Designing **Complex Biological Systems**

- Accelerate the discovery process
- Elevate human creativity to higher level goals
- Democratize biological systems design approaches
- Unbiased data collection and evaluation

Link AI driven workflow with self-driving laboratory

Courtesy: Ian Foster, ANL/ UChicago

Agents run on HPC/AI resources

Candidates for experimental evaluation

Self-driving lab performs experiments

Embodied Agent for Automated Lab Code Generation Reasoning: Based on the information provided, it seems like the next logical step Candidate Code Memory of Tasks

would be to prepare the master mix for the PCR reaction. This involves combining various reagents in specific volumes to create the master mix solution.

Task: Prepare the master mix for the PCR reaction.

Useful Programs:

```
def PCR_Master_Mix(labware_info, protocolContext):
*/
Input: labware_info --> json-str
Pass in a variable labware_info that contains labware information and
quantities used
*/
```

Human:

```
labware_info = {"number_of_samples":96,
```

```
"left_pipette":"flex_8channel_1000","mastermix_volume":18,"DNA_volume":2}
```


Feedback to define additional experiments

Set of peptides as input

Agents run on **HPC/AI resources**

Candidates for experimental evaluation

Self-driving lab performs experiments

Stop

NVIDIA PLATFORM EVOLVING TO MEET THE CHALLENGE SIMULATION

EXPERIMENTS **/SENSORS**

TELESCOPES

VIRTUAL DESIGN AND CONTROL

HPC*AI

Physical System

Actual Sensor

Virtual Sensor

Physics-based Models

Physics-based Models

 $\bullet \bullet \bullet$

OMNIVERSE: PLATFORM FOR BUILDING DIGITAL TWINS

NUCLEUS

Digital Twin Data

Harvesting

(AI) Surrogate

 $\bullet \bullet \bullet$

(AI) Surrogate

Virtual Actor

NUCLEUS

Source of truth

ADVANCED TOOLS AND TECHNOLOGIES

Foundational Platform Components

CONNECT

KIT

Coupling

Application API User experience

SIMULATION

Virtual Actor

RTX RENDERER

Virtual Sensor

DATA AGGREGATION AND COORDINATION VIA USD

UNIVERSAL SCENE DESCRIPTION The "HTML" of 3D Virtual Worlds

- Developed by Pixar
- Foundation for NVIDIA Omniverse
- Open-sourced API and file framework for complex scene graphs
- Easily extensible, simplifies interchange of assets between industry software
- Introduces novel concept of layering
- Enables simultaneous collaboration for large teams in different department working on the same scene
- Originated in M&E, now becoming a standard across industries including AEC, Manufacturing, Product Design, Robotics

BACKUP

EXAMPLE OF DIGITAL TWIN FOR ASTROPHYSICS Moonwalker Digital Twin of the South Pole of The Moon https://www.youtube.com/watch?v=E0Rz0ZbwhJY

SUPERFACILITY WITH GA AND NERSC/ANL: VISION FOR FUTURE

DIII-D National Fusion Facility

ITER

FUTURE POST Exascale

Using foundation models to predict SARS-CoV-2 evolution **PRE-TRAINING**

Foundation model(s) trained on 110 million PATRIC sequences

PREDICTION WORKFLOW

Diffusion model to get hierarchy of gene organization (generation)

> **Generated SARS-**CoV-2 genomes

Science goal: Understand how to "trap" the SARS-CoV-2 replication-transcription complex (RTC) using Low Res CryoEM

Cryo-EM maps Protein Data Bank dividual protei parts Initial best Atomistic (refined) Novel conformational (1)(6)Cryo-EM models guess states Global conformation al fluctuations 5) Atomistically correct subdomain orientations Interface potentials Ensemble Continuum Simulations (FFEA)

All-atom Ensemble MD Simulations (NAMD)

Hierarchical Al Methods for computational steering

GenSLMs finetuned on SARS-CoV-2 genomes can distinguish variants

