AI FOR SCIENCE FRAMEWORK: MODULUS

戴志翔 zhixiangd@nvidia.com
NVIDIA SA Team

AGENDA

WHAT IS MODULUS

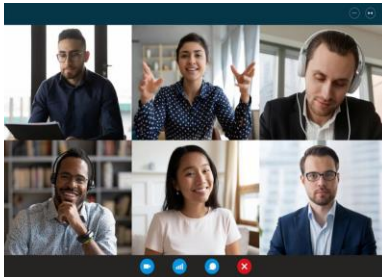
AI IS ALREADY TRANSFORMING EVERY INDUSTRY

Demand for Fast and Easy Deployment Greater than Ever

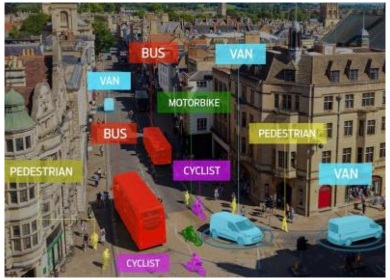
CREDIT CARD FRAUD
1.1B Credit Transactions / Day

CONTACT CENTER AI
500M Calls / Day

MEETING TRANSCRIPTION
15B Meeting Minutes / Day

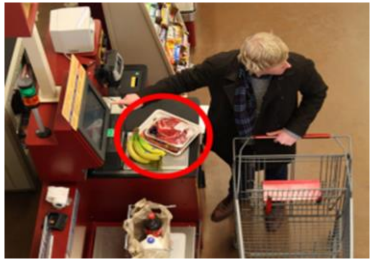


PUBLIC SAFETY
> 1B Smart City Cameras Deployed



PRODUCT RECOMMENDATIONS
300M E-commerce Visitors / Day

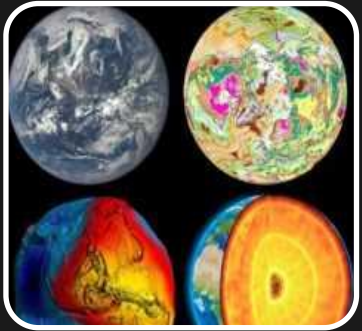
RETAIL ASSET PROTECTION
$275M Inventory Loss / Day

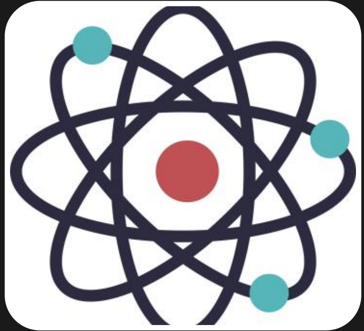
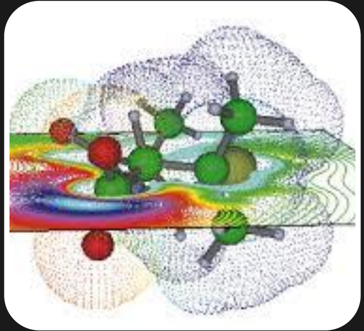


MEDICAL IMAGING
10M Diagnostic Scans / Day

INDUSTRIAL INSPECTION
94M Vision Sensors Installed by 2025

AI POWERED COMPUTATIONAL DOMAINS

Computational Eng.

Solid & Fluid Mechanics,

Electromagnetics,

Thermal, Acoustics,

Optics, Electrical,

Multi-body Dynamics,

Design Materials,

Systems

Earth Sciences

Climate Modeling,

Weather Modeling,

Ocean Modeling,

Seismic Interpretation

Life Sciences

Genomics,

Proteomics

Computational Physics

Particle Science,

Astrophysics

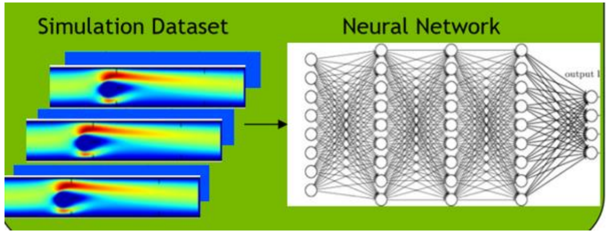
Computational Chemistry

Quantum Chemistry,

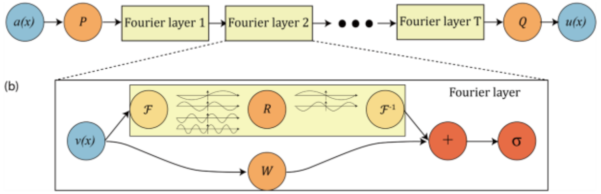
Molecular Dynamics

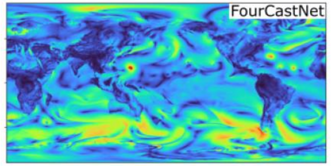
Physics-ML categorization

Neural Network

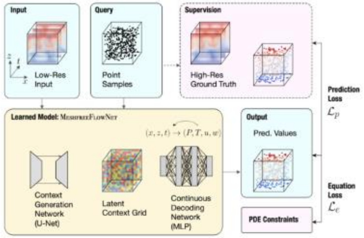


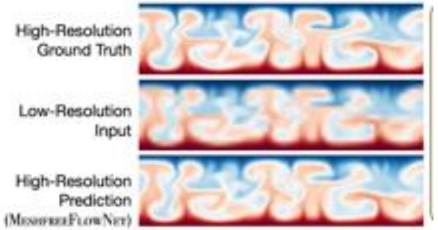
Neural Operator



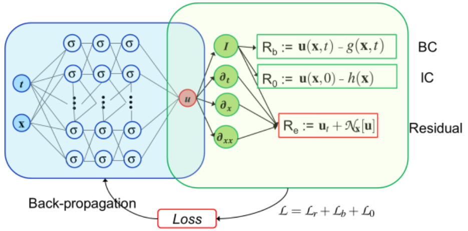


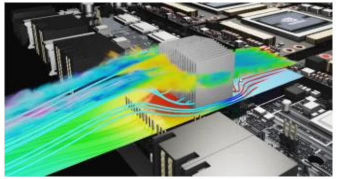
Physics constrained





Physics-Informed





Data

Physics

SimNet > Modulus

SimNet

A Neural Network Based Partial Differential Equation Solver

NVIDIA Modulus

A Framework for Developing Physics Machine Learning Neural Network Models

PyTorch/TensorFlow: AI

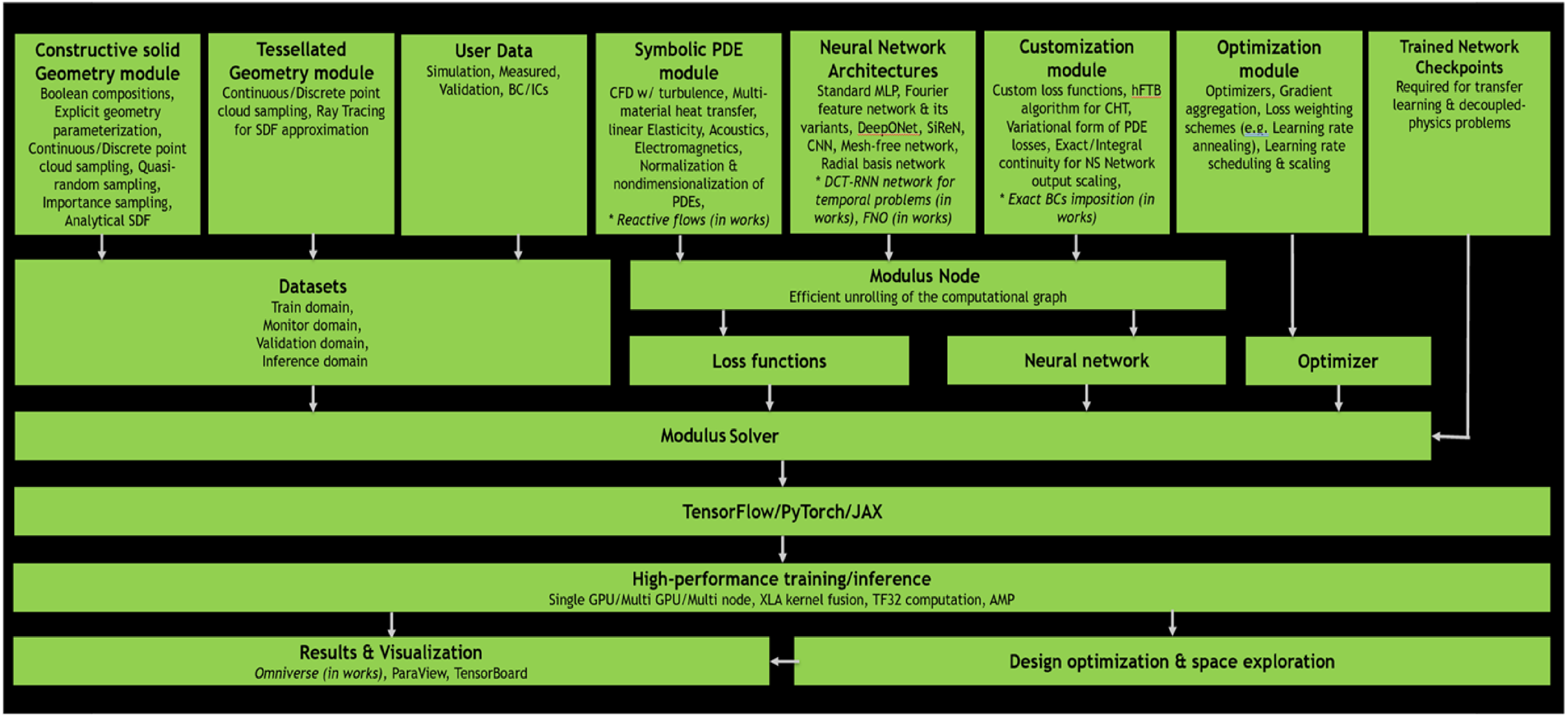
Modulus: AI For Science

History & Current Version:

- EA in June 2020
- First GA in June 2021, v21.06
- EA in April 2022 – v. 22.03
- Latest release in July 2022 – v. 22.07
- Next scheduled release 22.09 (Sept. 2022)

PRODUCT OVERVIEW | DOCUMENTATION | DOWNLOAD

Modulus



Constructive solid Geometry module
Boolean compositions, Explicit geometry parameterization, Continuous/Discrete point cloud sampling, Quasi-random sampling, Importance sampling, Analytical SDF
Tessellated Geometry module
Continuous/Discrete point cloud sampling, Ray Tracing for SDF approximation
User Data
Simulation, Measured, Validation, BC/ICs
Symbolic PDE module
CFD w/ turbulence, Multi-material heat transfer, linear Elasticity, Acoustics, Electromagnetics, Normalization & nondimensionalization of PDEs,
* Reactive flows (in works)
Neural Network Architectures
Standard MLP, Fourier feature network & its variants, DeepONet, SiReN, CNN, Mesh-free network, Radial basis network
* DCT-RNN network for temporal problems (in works), FNO (in works)
Customization module
Custom loss functions, hFTB algorithm for CHT, Variational form of PDE losses, Exact/Integral continuity for NS Network output scaling,
* Exact BCs imposition (in works)
Optimization module
Optimizers, Gradient aggregation, Loss weighting schemes (e.g. Learning rate annealing), Learning rate scheduling & scaling
Trained Network Checkpoints
Required for transfer learning & decoupled-physics problems
Datasets
Train domain, Monitor domain, Validation domain, Inference domain
Modulus Node
Efficient unrolling of the computational graph
Loss functions
Neural network
Optimizer
Modulus Solver
TensorFlow/PyTorch/JAX
High-performance training/inference
Single GPU/Multi GPU/Multi node, XLA kernel fusion, TF32 computation, AMP
Results & Visualization
Omniverse (in works), ParaView, TensorBoard
Design optimization & space exploration

Framework Features

Novel NN architectures

Physics-Informed neural Networks :

- Fully Connected (FC)
- Fourier Feature Network
- Sinusoidal Representation Network
- …

Neural Operators:

- Fourier Neural Operator (FNO)
- Adaptive Fourier Neural Operator (AFNO)
- FourCastNet
- DeepONet
- …

Physics Informed Neural Operators:

- PINO (PINN+FNO)
- physics informed DeepONet
- …

others:

- Pix2Pix Net
- Super Resolution Net
- …

Bringing novel AI architectures that have demonstrated success for engineering and science problems

Using reference examples to show which of the curated architectures are applicable to certain class of problems

PHYSICS-INFORMED NEURAL NETWORKS

Physics-Informed neural Networks (PINN)

物理信息神经网络

$$\mathbf{u}_t + \mathcal{N}_{\mathbf{x}}[\mathbf{u}] = 0, \quad \mathbf{x} \in \Omega, \ t \in [0,T],$$
$$\mathbf{u}(\mathbf{x},0) = h(\mathbf{x}), \quad \mathbf{x} \in \Omega,$$
$$\mathbf{u}(\mathbf{x},t) = g(\mathbf{x},t), \quad \mathbf{x} \in \partial\Omega, \ t \in [0,T],$$

$\mathcal{N}_{\mathbf{x}}$ is the differential operator

$\mathbf{x}$ and t are spatial and temporal coordinates

Ω and $\partial\Omega$ are domain & boundary

$\mathbf{u}(\mathbf{x},t)$ is approximated by a neural network: INPUT $(\mathbf{x},t)$ → OUTPUT $\mathbf{u}_{NN}(\mathbf{x},t)$

loss function:

$$\mathcal{L} = \mathcal{L}_r + \mathcal{L}_b + \mathcal{L}_0,$$

$$\mathcal{L}_r = \frac{1}{N_r}\sum_{i=1}^{N_r} \left|\mathbf{u}_t(\mathbf{x}^i,t^i) + \mathcal{N}_{\mathbf{x}}[\mathbf{u}(\mathbf{x}^i,t^i)]\right|^2$$

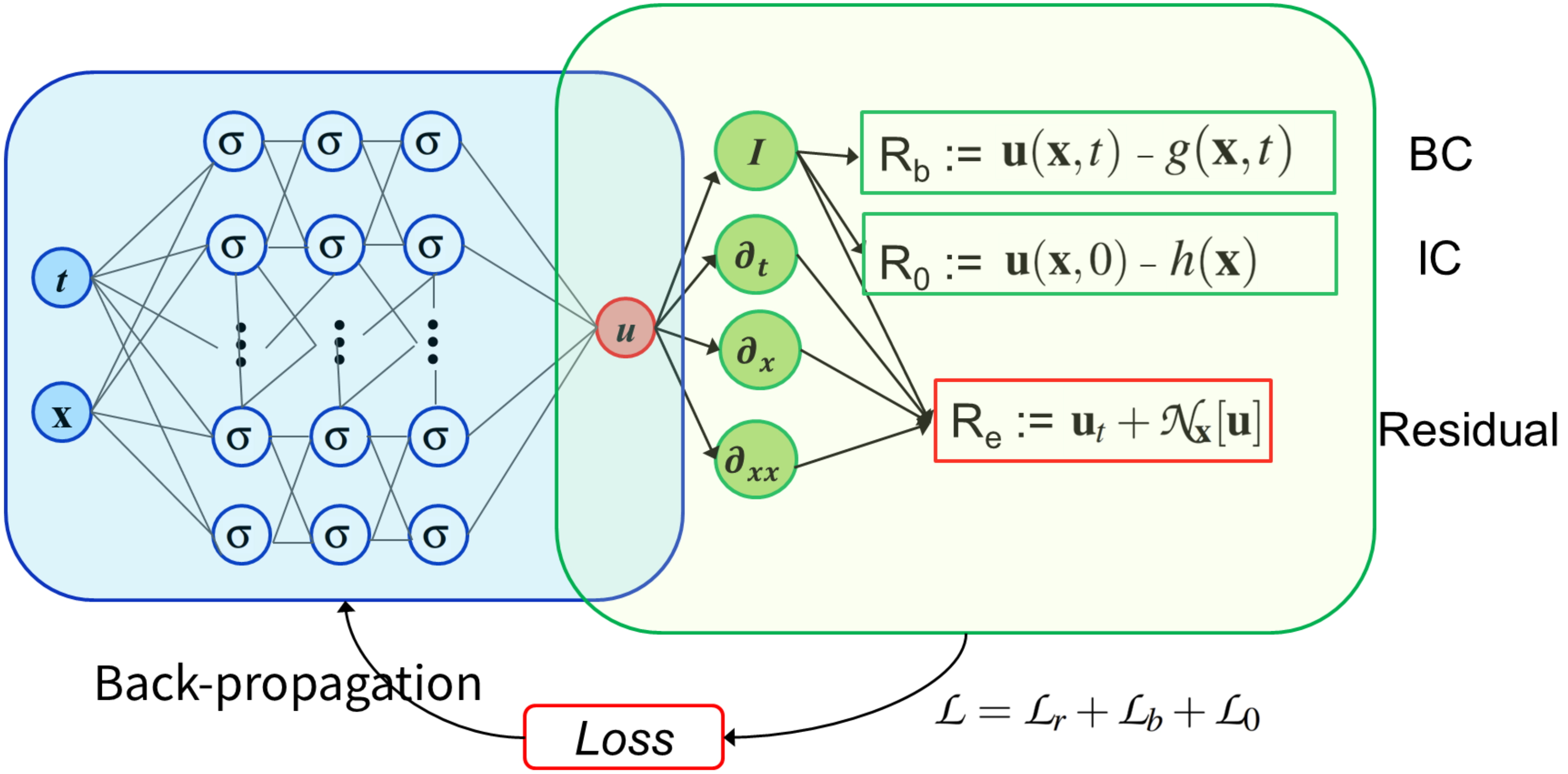
$$\mathcal{L}_b = \frac{1}{N_b}\sum_{i=1}^{N_{ub}} \left|\mathbf{u}(\mathbf{x}^i,t^i) - g^i\right|^2$$

$$\mathcal{L}_0 = \frac{1}{N_0}\sum_{i=1}^{N_0} \left|\mathbf{u}(\mathbf{x}^i,t^i) - h^i\right|^2.$$

$\mathcal{L}_r$, $\mathcal{L}_b$ and $\mathcal{L}_0$ penalize the residuals of governing eqs, BC and IC

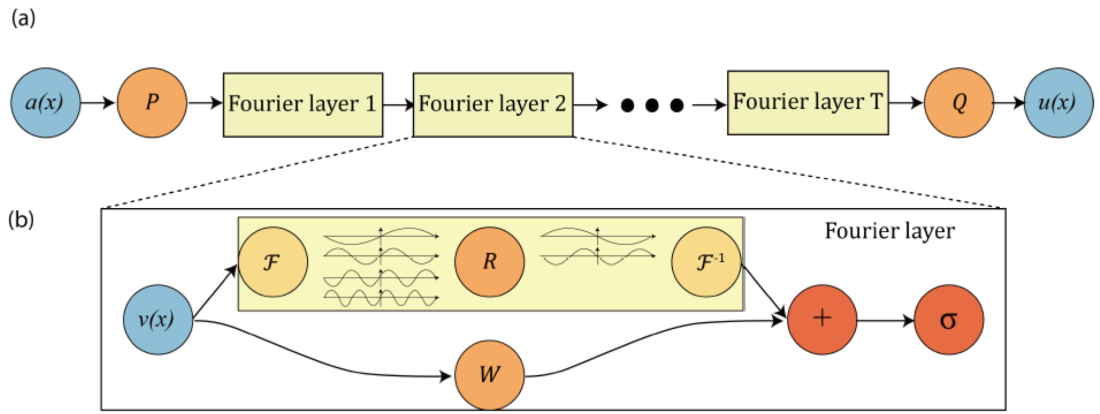
N_r, N_b and N_0 are the number of data points for different terms

Schematic of a PINN framework



NEURAL OPERATOR

Fourier Neural Operator



The Fourier layer consists of three steps:

- Fourier transform
- Linear transform on the lower Fourier modes *R*
- Inverse Fourier transform

$$v_{t+1}(x) = \sigma(Wv_t(x) + \mathcal{F}^{-1}(R_\phi \mathcal{F}(v_t)))$$

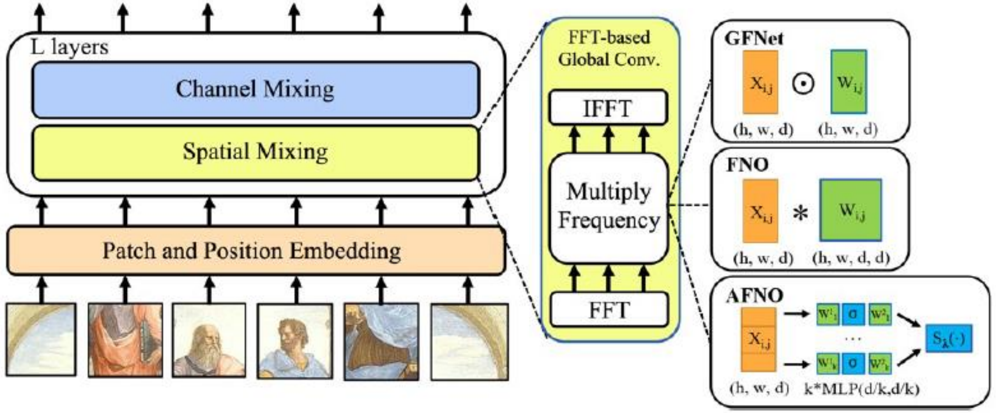
Fast. FFT has complexity O(n log n).

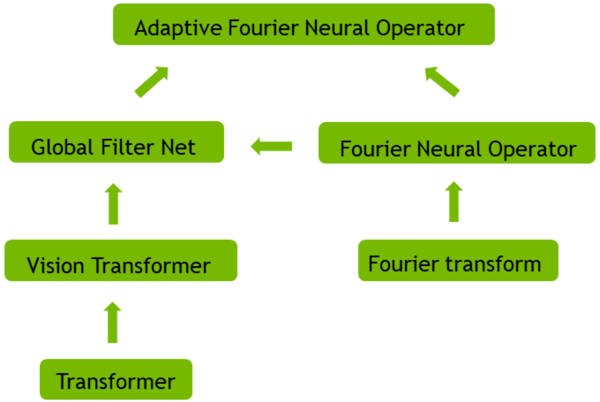
$$k_\phi * v_t = \mathcal{F}^{-1}(\mathcal{F}(k_\phi)\mathcal{F}(v_t))$$

Efficient. More efficient to represent PDE in Fourier space.

$$\mathcal{F}(\frac{d}{dx}f(x)) = 2\pi i\xi\mathcal{F}(f)$$

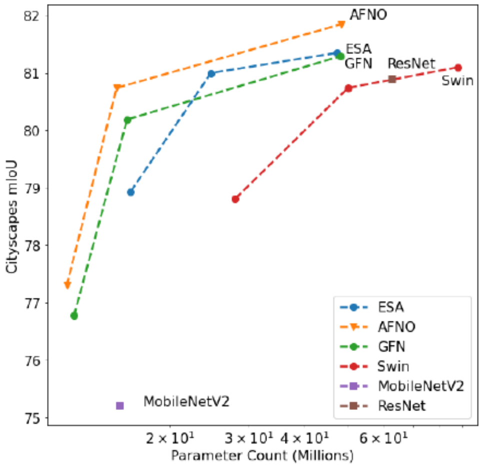
Adaptive Fourier Neural Operator





AFNO: Three steps to improve scalability and robustness

- Block-Diagonal Structure on W
- Adaptive Weight Sharing
- Adaptive Soft-Thresholding



Models	Complexity (FLOPs)	Parameter Count
Self-Attention	$N^2d + 3Nd^2$	$3d^2$
GFN	$Nd + Nd \log N$	Nd
FNO	$Nd^2 + Nd \log N$	Nd^2
AFNO (ours)	$Nd^2/k + Nd \log N$	$(1 + 4/k)d^2 + 4d$

Adaptive Fourier Neural Operators: Efficient Token Mixers for Transformers https://arxiv.org/abs/2111.13587

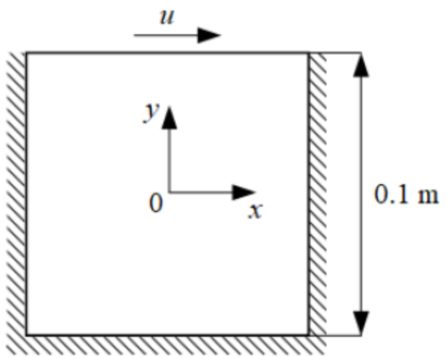
APPLICATION AND VERIFICATION

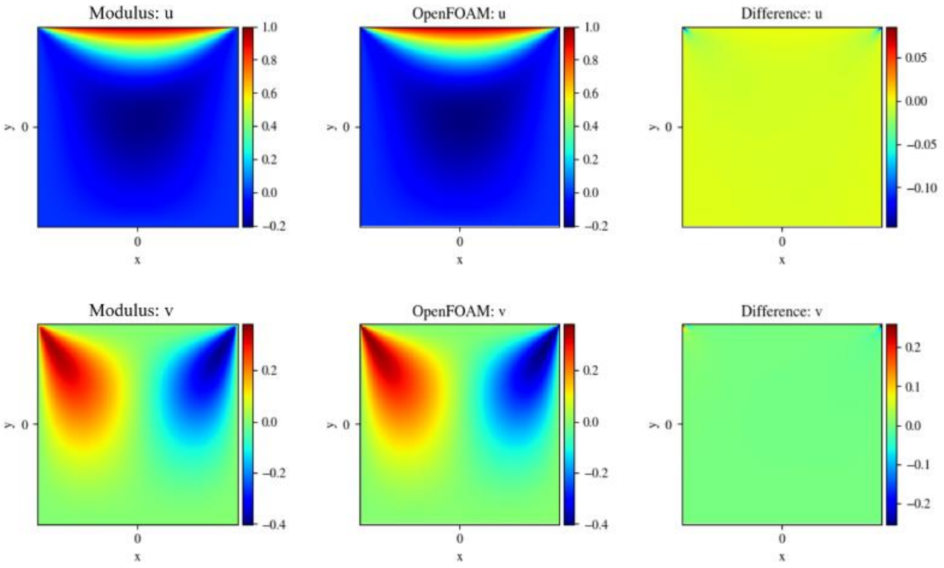
Lid Driven Cavity Flow

CFD: Navier-Stokes Equations

$$\frac{d\rho}{dt} + \rho\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) = 0$$

$$\rho\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}\right) = \rho f_x - \frac{\partial p}{\partial x} + \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right)$$
$$\rho\left(\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z}\right) = \rho f_y - \frac{\partial p}{\partial y} + \mu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right)$$
$$\rho\left(\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z}\right) = \rho f_z - \frac{\partial p}{\partial z} + \mu\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right)$$





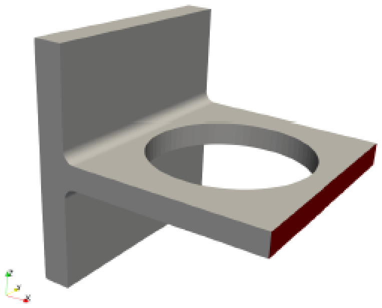
Linear Elasticity

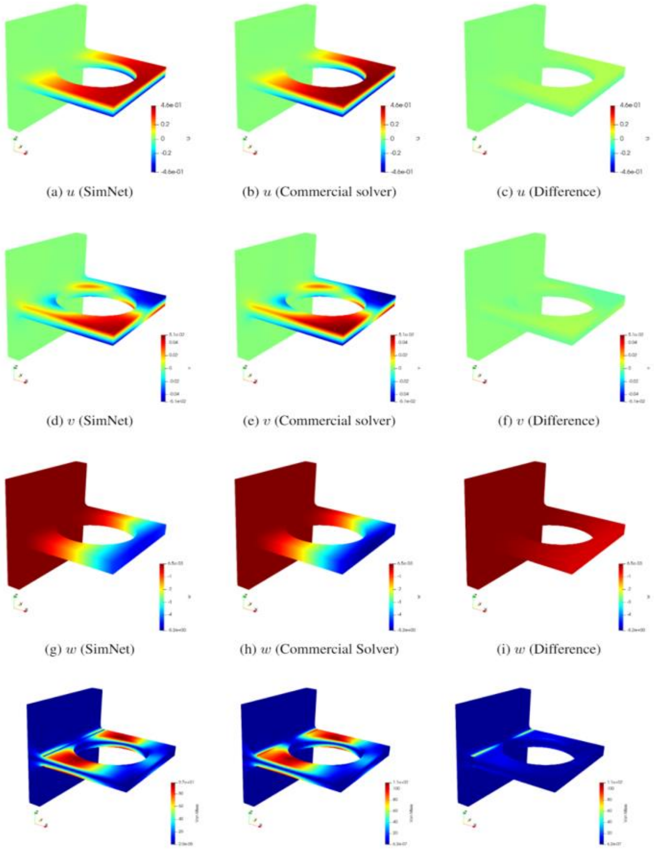
Solid Mechanics: Hooke's Elasticity

Equilibrium: $$\sigma_{ji,j} + f_i = 0,$$

Stress-Strain: $$\sigma_{ij} = \lambda \epsilon_{kk} \delta_{ij} + 2\mu \epsilon_{ij},$$

Strain-Displacement: $$\epsilon_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right).$$





(a) u (SimNet) (b) u (Commercial solver) (c) u (Difference)

(d) v (SimNet) (e) v (Commercial solver) (f) v (Difference)

(g) w (SimNet) (h) w (Commercial Solver) (i) w (Difference)

(j) σ_{VM} (SimNet) (k) σ_{VM} (Commercial Solver) (l) σ_{VM} (Difference)

Seismic Problems

Solve acoustic wave propagation

2D acoustic wave equation:

$$m\frac{\mathrm{d}^2 u(t,x,y)}{\mathrm{d}t^2} - \Delta u(t,x,y) + \eta\frac{\mathrm{d}u(t,x,y)}{\mathrm{d}t} = q(t,x,y;x_s,y_s)$$

Single Ricker wavelet source location in the middle

Constant and Layered Velocity models

Initial conditions: *u(x,y,0) = 0*

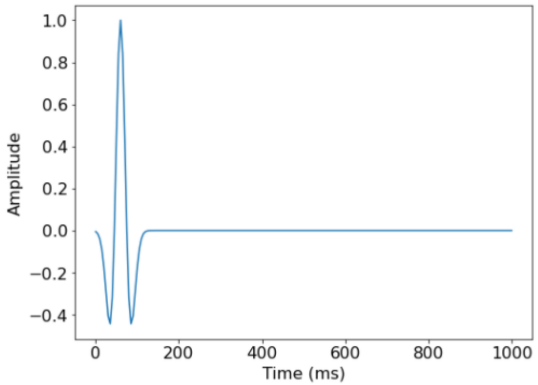
Boundary conditions: Open BC

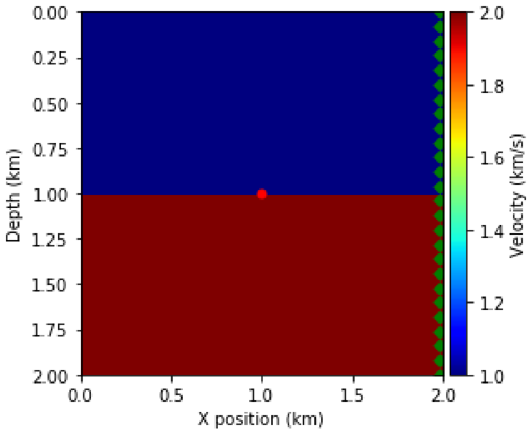
Training Loss function = physics loss + Boundary loss

$$\mathcal{L}(\theta;\mathcal{T}) = w_f\mathcal{L}_f(\theta;\mathcal{T}_f) + w_b\mathcal{L}_b(\theta;\mathcal{T}_b),$$

$$\mathcal{L}_f(\theta;\mathcal{T}_f) = \frac{1}{|\mathcal{T}_f|}\sum_{\mathrm{x}\in\mathcal{T}_f}\left\|f\left(\mathrm{x};\frac{\partial\hat{u}}{\partial x_1},\ldots,\frac{\partial\hat{u}}{\partial x_d};\frac{\partial^2\hat{u}}{\partial x_1\partial x_1},\ldots,\frac{\partial^2\hat{u}}{\partial x_1\partial x_d};\ldots;\boldsymbol{\lambda}\right)\right\|_2^2,$$

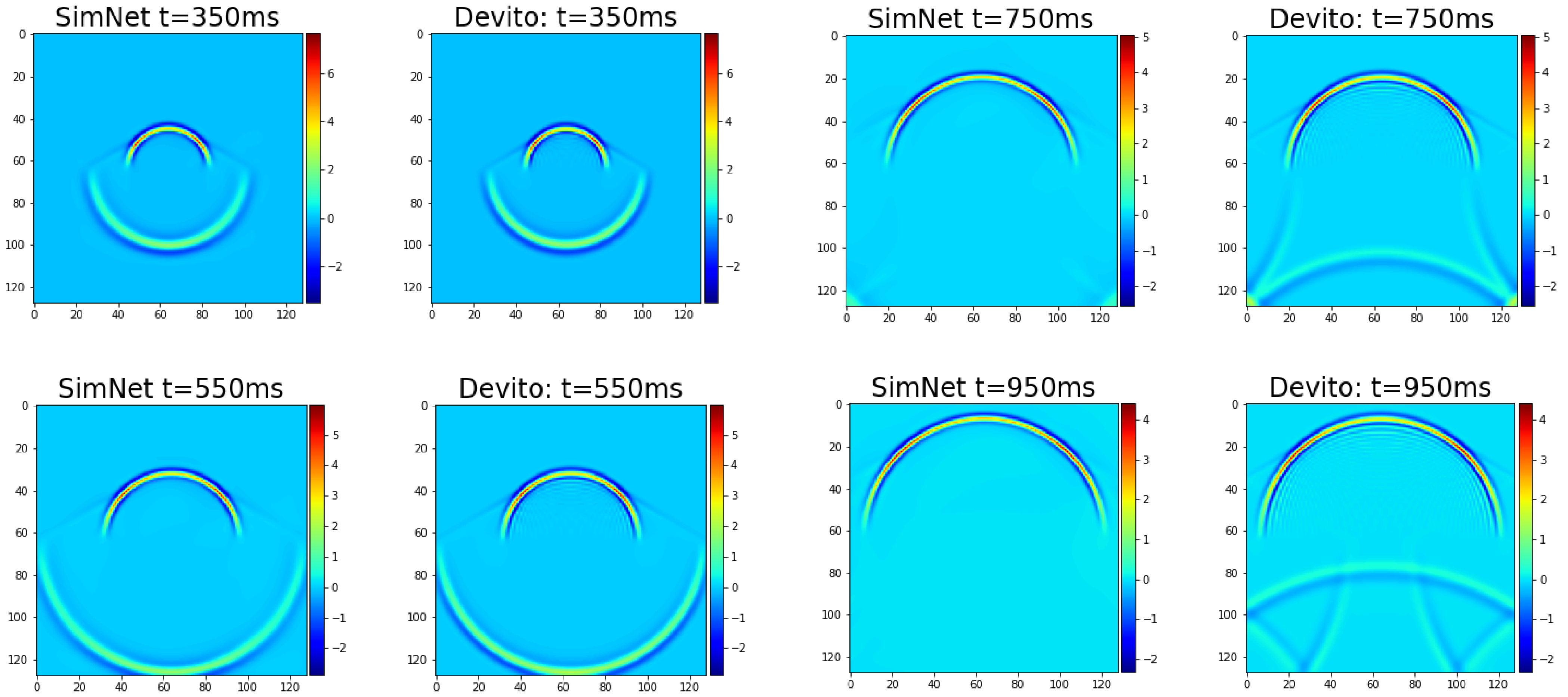
$$\mathcal{L}_b(\theta;\mathcal{T}_b) = \frac{1}{|\mathcal{T}_b|}\sum_{\mathrm{x}\in\mathcal{T}_b}\|\mathcal{B}(\hat{u},\mathrm{x})\|_2^2,$$





Seismic Problems

Wave propagation with layered velocity



2D Waveguide Cavity

Electromagnetics: Maxwell’s Equations

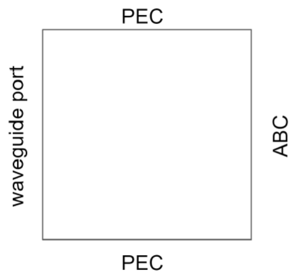


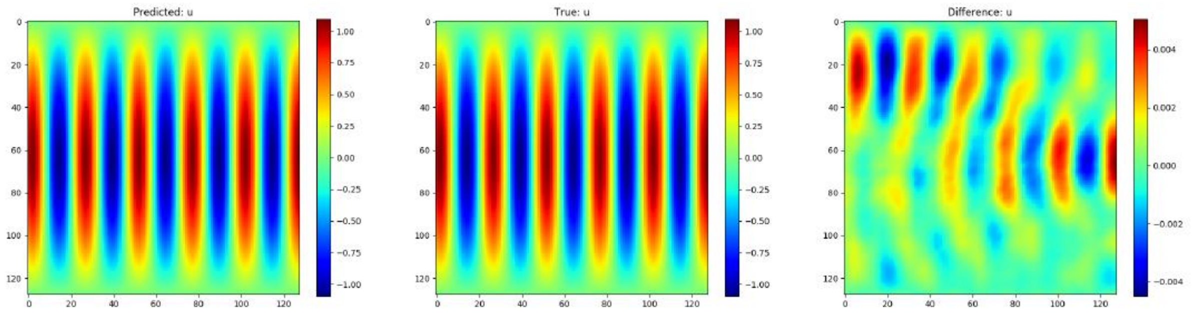
Figure 42: Domain of 2D waveguide

In this example we will solve this waveguide problem by transverse-magnetic (TM_z) mode, so that our unknown variable is $E_z(x, y)$. The governing equation in Ω is

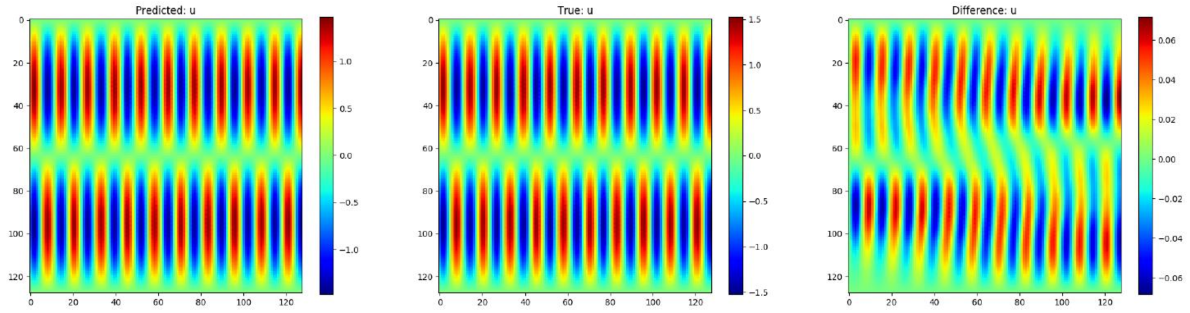
$$\Delta E_z(x, y) + k^2 E_z(x, y) = 0,$$

where k is the wavenumber. Notice in 2D scalar case, the PEC and ABC will be simplified in the following form, respectively:

$$E_z(x, y) = 0 \text{ on top and bottom boundaries, } \frac{\partial E_z}{\partial y} = 0 \text{ on right boundary.}$$



Wave Number = 16, Eigen Mode = 1



Wave Number = 32, Eigen Mode = 2

DIGITAL TWINS

HRSG FLUID ACCELERATED CORROSION SIMULATION — SIEMENS ENERGY

Use Case

- Detecting and predicting point of corrosion in heat recovery steam generators (HRSGs)

Challenges

- Using standard simulation to detect corrosion, it took SE at least couple of weeks, and the overall process took 14-16 weeks for every HRSG unit.

Solution

- Using NVIDIA Modulus Physics-Informed Neural Network, SE simulates the corrosive effects of heat, water and other conditions on metal over time to fine-tune maintenance needs.
- SE can replicate and deploy HRSG plant digital twins worldwide with NVIDIA Omniverse.

NVIDIA Solution Stack

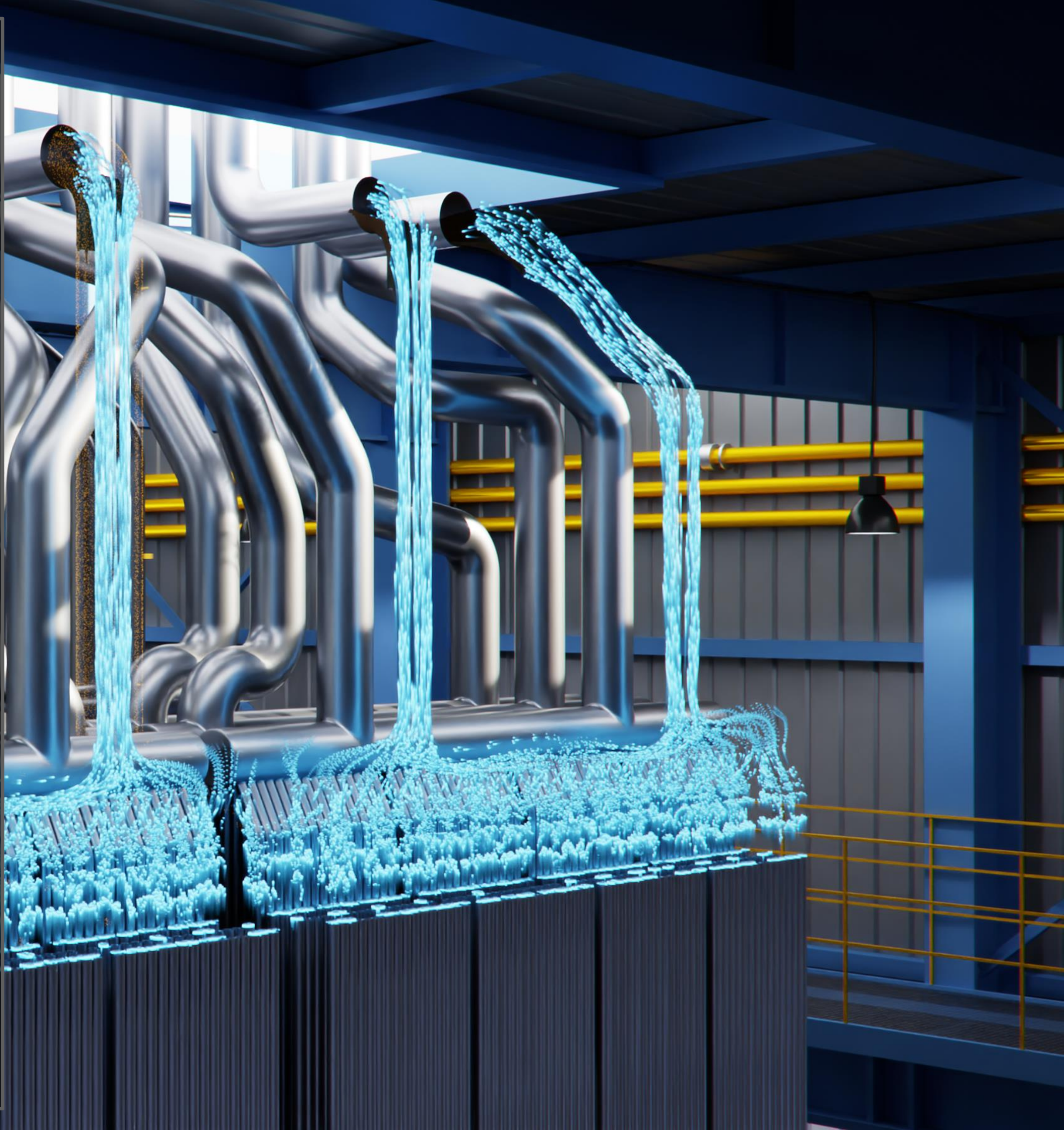
- Hardware: NVIDIA V100 & A100 Tensor Core GPUs
- Software: NVIDIA Modulus, NVIDIA Omniverse

Outcome

- 10,000X speed-up and inference in seconds can reduce downtime by 70%, saving the industry $1.7 billion annually

Link to Demo

NVIDIA SIEMENS energy



WIND TURBINE WAKE OPTIMIZATION — SIEMENS GAMESA

Use Case

- Developing optimal engineering wake models to optimize wind farm layouts
- Simulating the effect that a turbine might have on another when placed in close proximity

Challenges

- Generating high-fidelity simulation data from Reynolds-averaged Navier-Stokes (RANS) or Large Eddy Simulations (LES) can take over a month to run, even on a 100-CPU cluster.

Solution

- NVIDIA Omniverse and Modulus enable accurate, high-fidelity simulations of the wake of the turbines, using low-resolution simulations as inputs and applying super resolution using AI.

NVIDIA Solution Stack

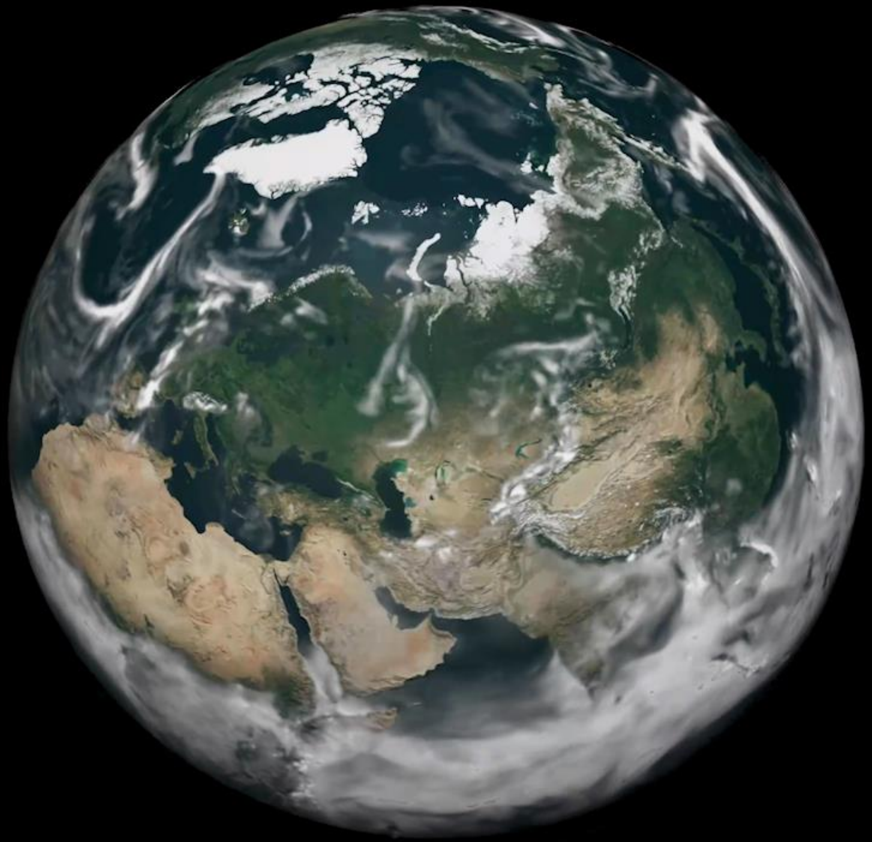
- Hardware: NVIDIA A100, A40, RTX 8000 GPUs
- Software: NVIDIA Omniverse, NVIDIA Modulus

Outcome

- Approximately 4,000X speedup for high-fidelity simulation
- Optimizing wind farm layouts in real-time increases overall production while reducing loads and operating costs.

Demo

NVIDIA SIEMENS Gamesa



ACCELERATING EXTREME WEATHER PREDICTION WITH FourCastNet IN NVIDIA MODULUS

Use Case

- Climate change is making storms both stronger and less predictable, leading to more fires, floods, heatwaves, mudslides, and droughts.
- Predicting global weather patterns and extreme weather events, like atmospheric rivers, is important to quantify any catastrophic event with confidence.

Challenges

- To develop the best strategies for mitigation and adaptation, we need climate models that can predict the climate in different regions of the globe over decades.

Solution

- NVIDIA has created a physics-ML model that emulates the dynamics of global weather patterns and predicts extreme weather events, like atmospheric rivers, with unprecedented speed and accuracy.

NVIDIA Solution Stack

- Hardware: NVIDIA A100
- Software: NVIDIA Omniverse, NVIDIA Modulus

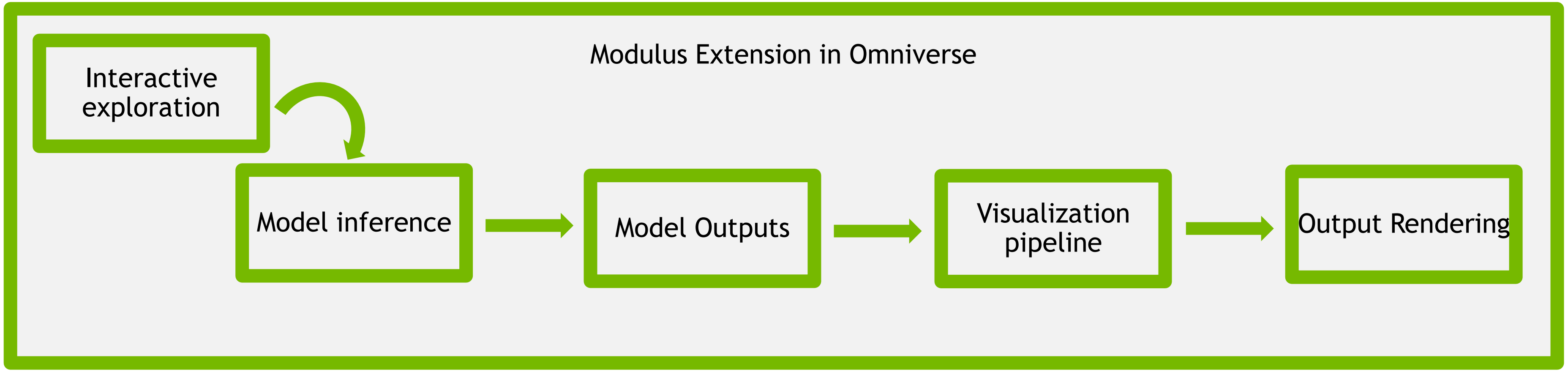
Outcome

- Powered by the Fourier Neural Operator, this GPU-accelerated AI-enabled digital twin, called FourCastNet, is trained on 10 TB of Earth system data.
- Using this data, together with NVIDIA Modulus and Omniverse, we are able to forecast the precise path of catastrophic atmospheric rivers a full week in advance.

Demo

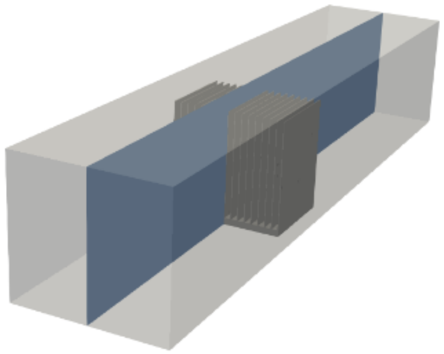
NVIDIA

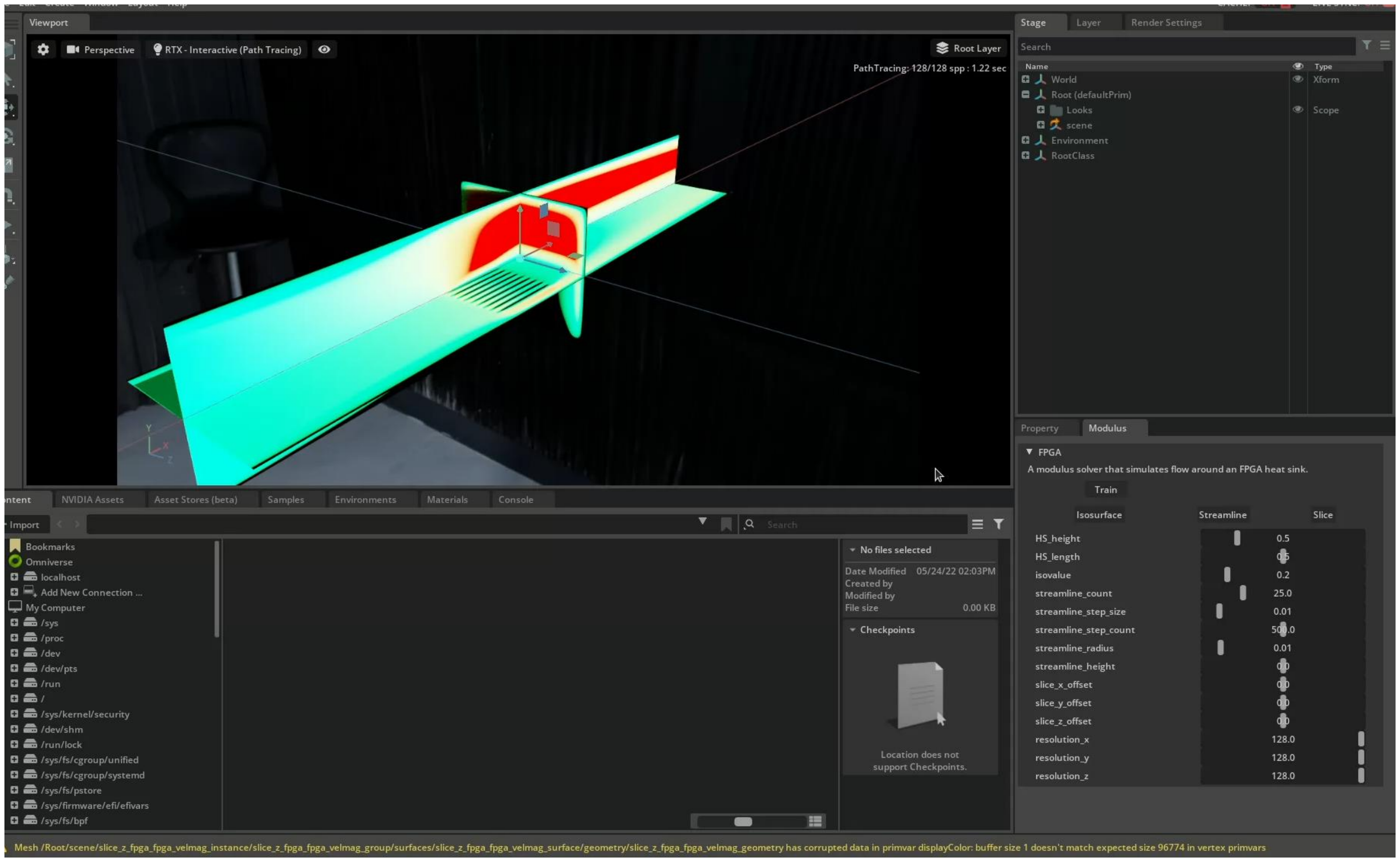
Modulus - Omniverse integration



- **Modulus training and inference workflows are python API based and the resulting trained model outputs are brought in as scenarios into OV using this extension.**
- **What does Modulus Omniverse extension do?**
  - enables importing outputs of Modulus trained model into a visualization pipeline for common output scenarios ex: streamlines, iso-surface
  - provides an interface that enables interactive exploration of design variables/parameters to infer new system behavior

Modulus – Omniverse integration





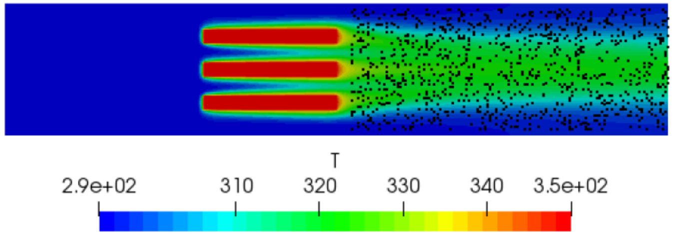
Common visualization modes such as Isosurface, Streamlines, Slices are available.

Each mode will populate the geometry, which will be updated as you change parameters.

ADVANTAGES

Inverse Problem

Finding Unknown Coefficients of a PDE: Heat Sink



Fluid Heat Convection:

$$0 = \nabla \cdot (D_{fluid} \nabla \theta_{fluid}) - \nabla \cdot (U \theta_{fluid}) \qquad D_{fluid} = \frac{k_{fluid}}{\rho_{fluid} c_{pfluid}}$$

Solid Heat Conduction:

$$0 = \nabla \cdot (k_{solid} \nabla \theta_{solid}) \qquad D_{solid} = \frac{k_{solid}}{\rho_{solid} c_{psolid}}$$

Interface Conditions:

$$\theta_{solid} = \theta_{fluid}$$

$$k_{solid}(N \cdot \nabla \theta_{solid}) = k_{fluid}(N \cdot \nabla \theta_{fluid})$$

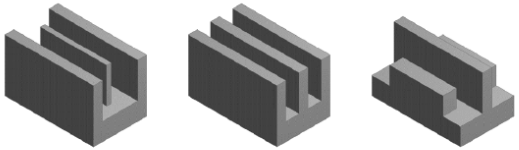
Results:

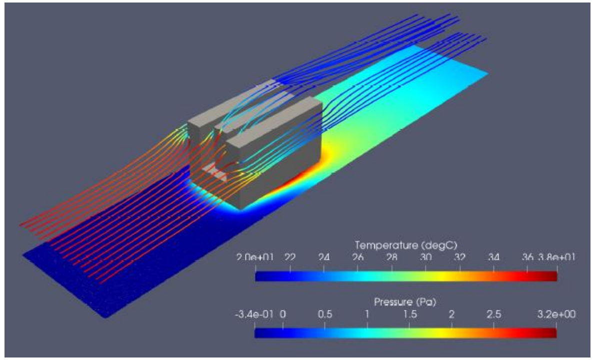
Property	OpenFOAM (True)	SimNet (Predicted)
Kinematic Viscosity (m^2/s)	1.00×10^{-2}	1.03×10^{-2}
Thermal Diffusivity (m^2/s)	2.00×10^{-3}	2.19×10^{-3}

Parameterized Simulation: 3D heat sink

Multi Physics Application: Fluids + Heat Transfer

$$
\begin{aligned}
h_{centralfin} &= (0.0, 0.6),\\
h_{sidefins} &= (0.0, 0.6),\\
l_{centralfin} &= (0.5, 1.0)\\
l_{sidefins} &= (0.5, 1.0)\\
t_{centralfin} &= (0.05, 0.15)\\
t_{sidefins} &= (0.05, 0.15)
\end{aligned}
$$



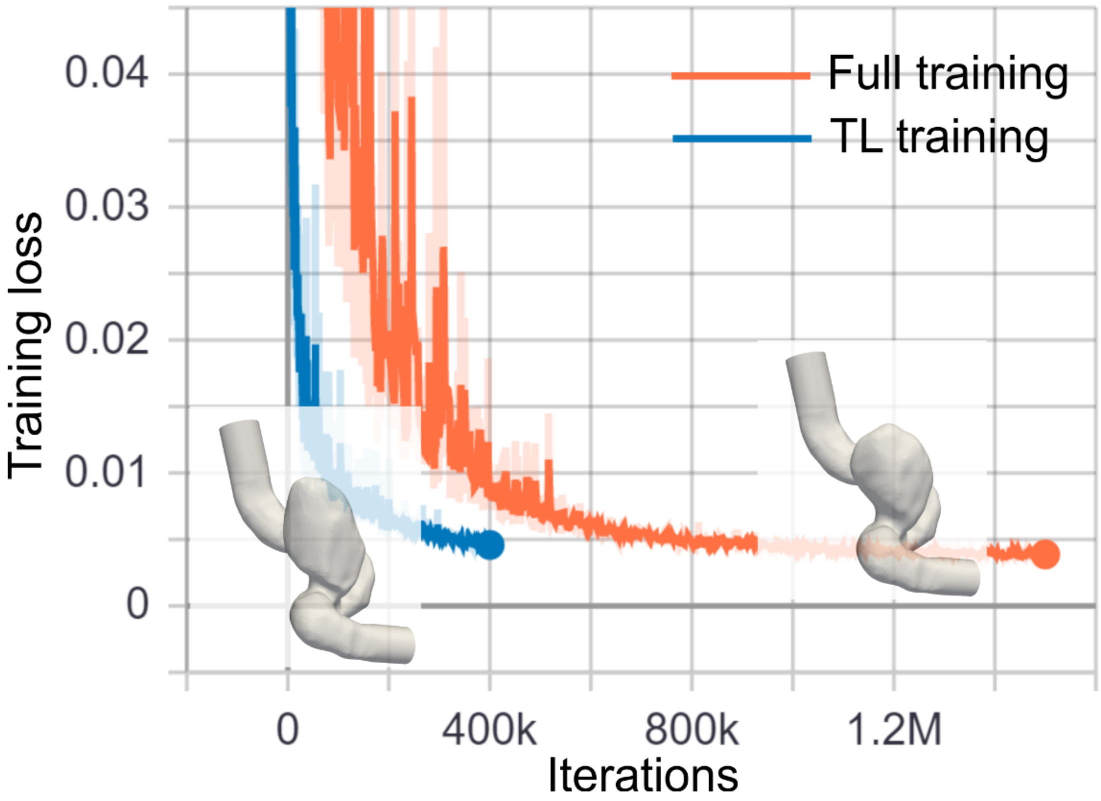


	Variable/Function	Description
minimize	$Peak\ Temperature$	Minimize the peak temperature at the source chip
with respect to	$h_{centralfin}$, $h_{sidefins}$, $l_{centralfin}$, $l_{sidefins}$, $t_{centralfin}$, $t_{sidefins}$	Geometric Design Variables of the Heat Sink
subject to	$Pressure\ drop < 2.5$	Limit on the pressure drop. Maximum value of the pressure drop that can be provided by the cooling system

Solver	OpenFOAM	SimNet
Compute Time (hrs)	4099	120

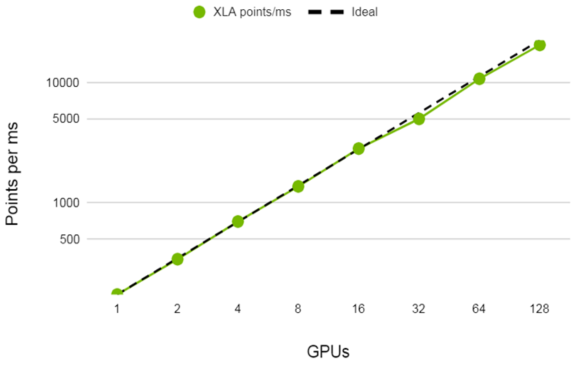
Transfer Learning: Blood Flow

Acceleration Of Training

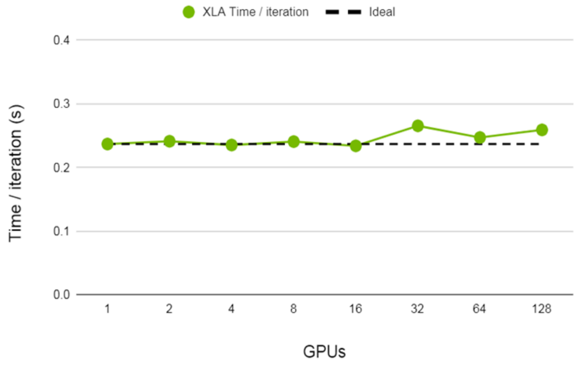


Multi-GPU/Node Performance

Scalability

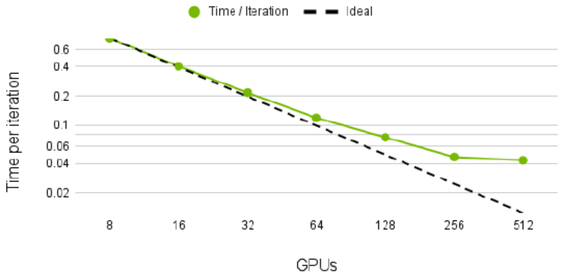


(a) Points per ms

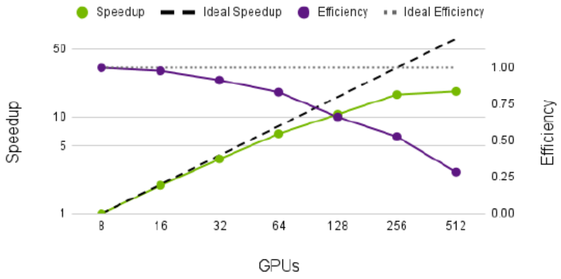


(b) Time per iteration

weak scaling plots for flow over an annular ring problem



(a) Time per iteration



(b) Speedup (green, left axis) and scaling efficiency (purple, right axis)

strong scaling results for the Taylor-Green vortex problem

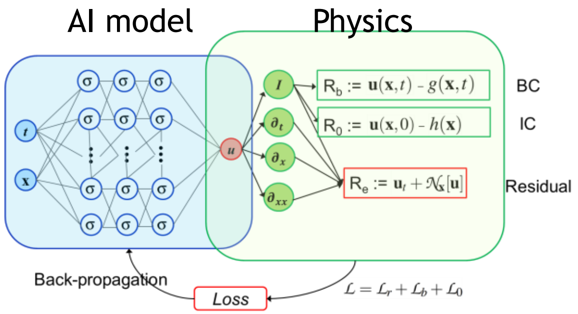
CONCLUSION

NVIDIA MODULUS

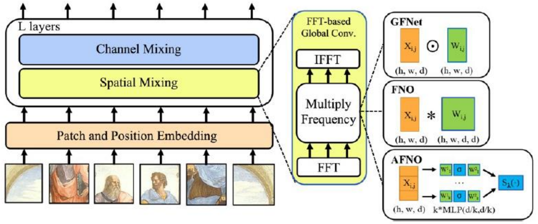
Framework for developing physics machine learning neural network models

What’s Modulus?

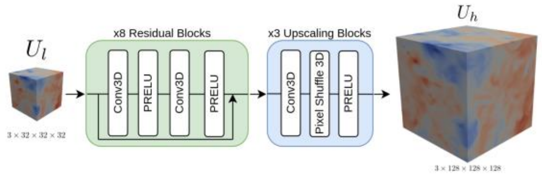
a Framework for developing physics-ML models (Similar to TensorFlow/PyTorch)



Physics-Informed neural Networks



Neural Operator

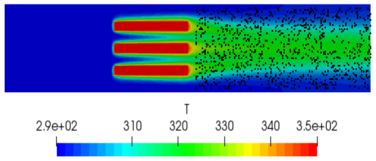


Super Resolution Network

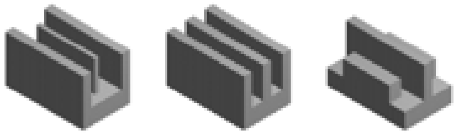
Where can Modulus be used?

CFD, Thermal, Solids, Acoustics(Seismic Wave), Electromagnetics(Maxwell’s Equation),Weather Forecast, Super Resolution

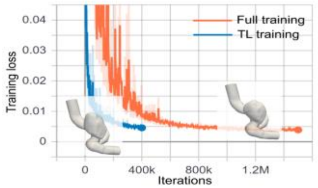
What are the advantages?



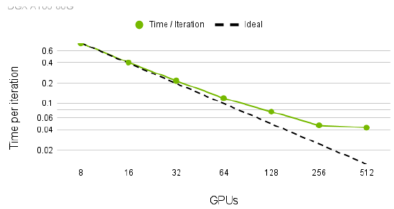
Inverse Problem



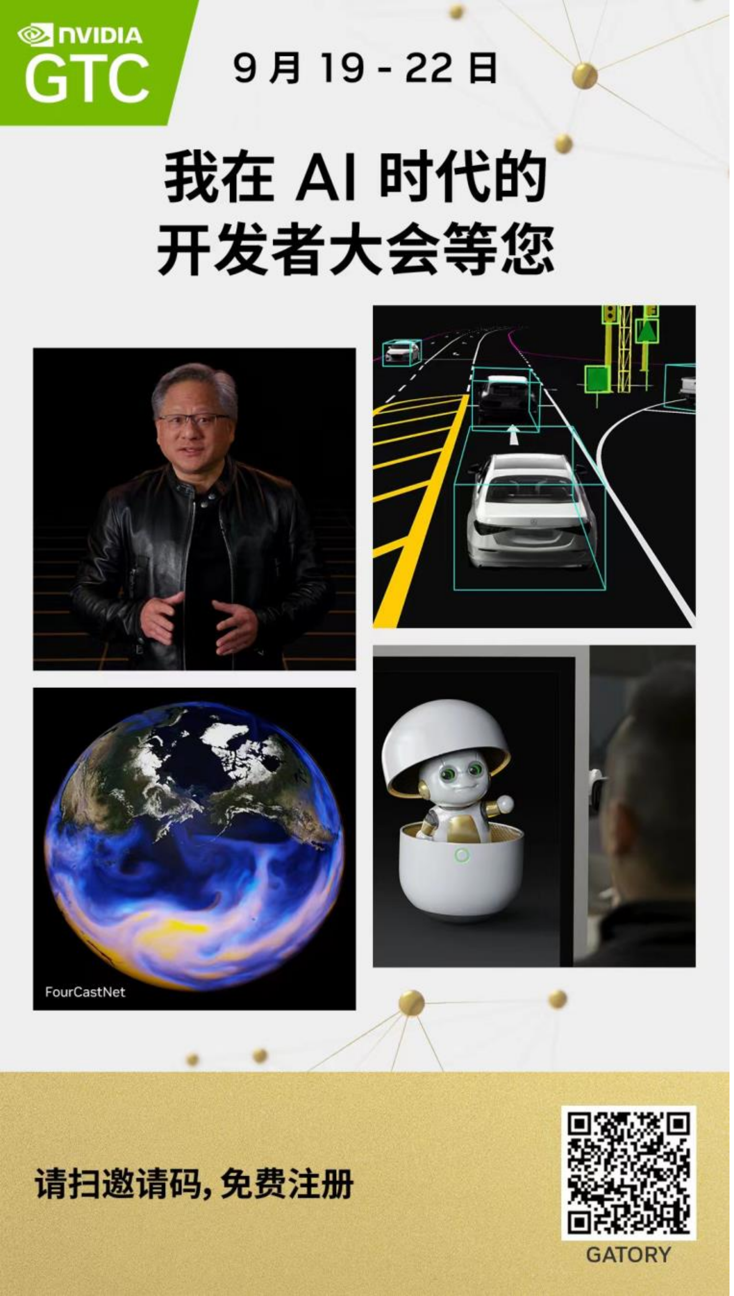
Parameterized Simulations



Transfer Learning



Multi GPUs



Modulus in GTC 2022

NVIDIA's Earth-2: Progress and Challenges Towards Building a Digital Twin of the Earth for Weather and Climate [A41326]

Accelerating a New Path to Innovation, Efficiency, and Discovery [A41237]

Journey Toward Zero-Carbon Emissions Leveraging AI for Scientific Digital Twins [A41224]

MODULUS学习群

该二维码7天内(8月24日前)有效，重新进入将更新

https://github.com/openhackathons-org/gpubootcamp/tree/master/hpc_ai/PINN

THANKS