NVIDIA SIMNET™ 20.12: AN AI-ACCELERATED SIMULATION TOOLKIT

Jay Chen, Data Scientist, April 2021 (presented to NCHC)

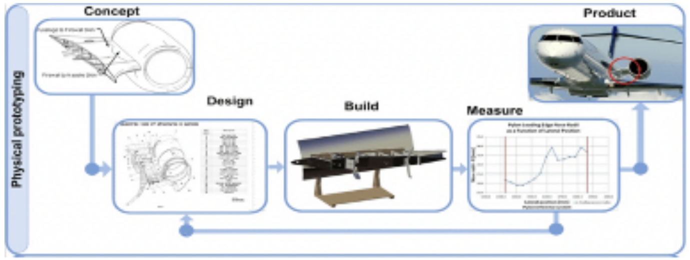
AGENDA

- AI in Science and Engineering - Motivation
- AI in Engineering - Taxonomy
- Data Driven Neural Networks (DDNNs)
- SimNet: Physics Informed Neural Networks (PINNs)
- SimNet Results
- SimNet Product
- More Information

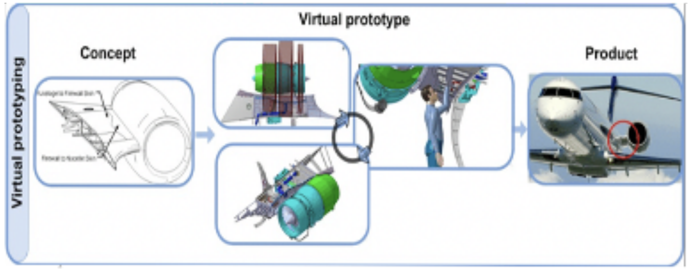
AI in Science & Engineering: Motivation

Why SimNet?

Physical Prototyping



Traditional Simulations



AI based Techniques

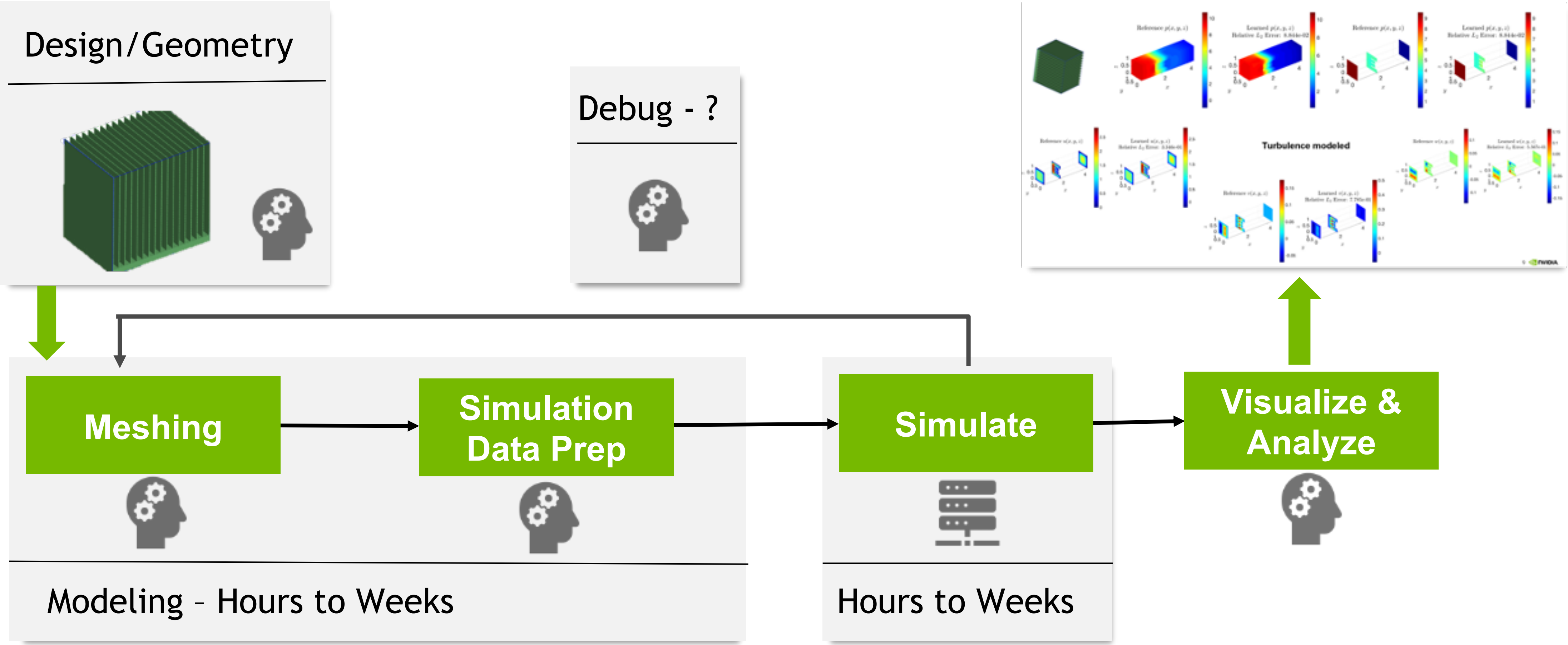
Physical Prototyping is iterative,
- time consuming, very costly
- not optimized for material and characteristics

Traditional numerical solvers work on one problem at a time making design process
- Time consuming
- Not real-time simulations
- No data assimilation, inverse problems

Data driven NN require data, are
- oblivious to physics laws
- not generalizable

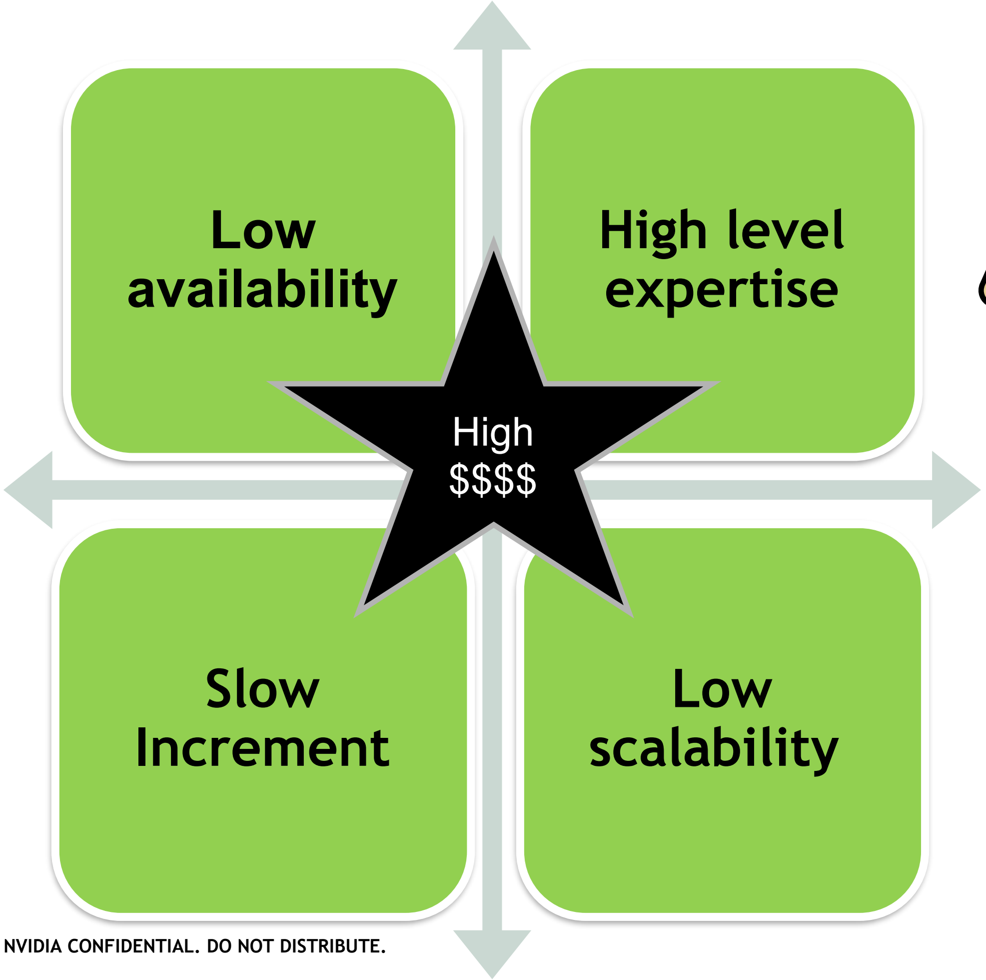
TRADITIONAL SIMULATION WORKFLOW

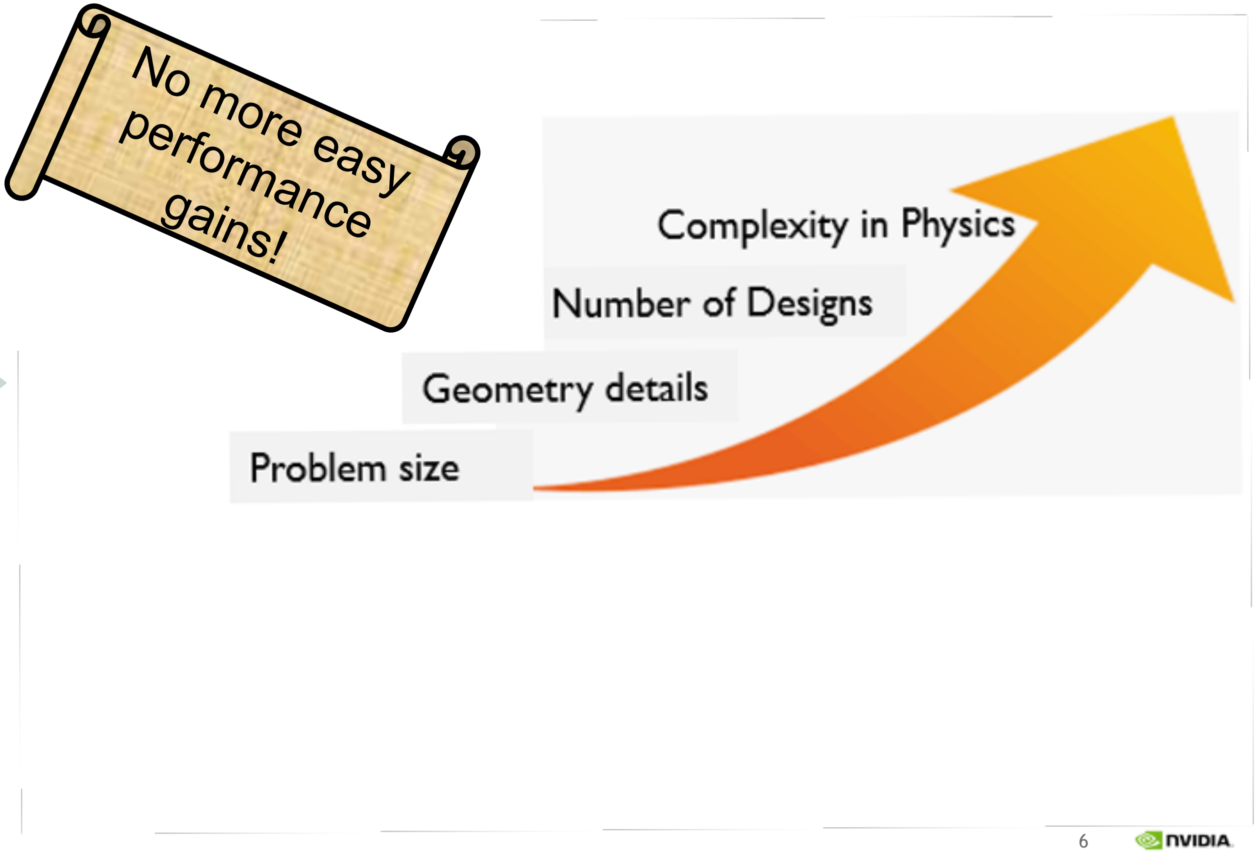
Human in-the-loop



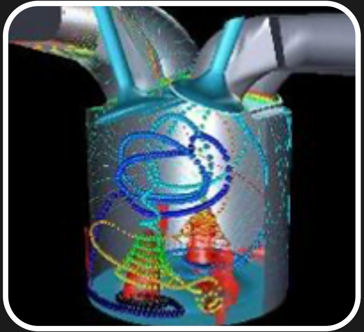
SATURATING PERFORMANCE IN TRADITIONAL HPC

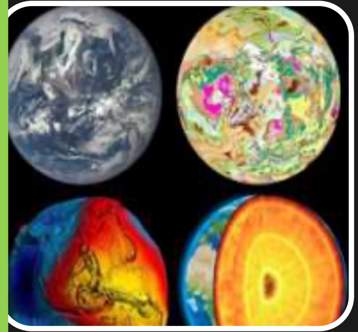
Simulations are getting larger & more complex

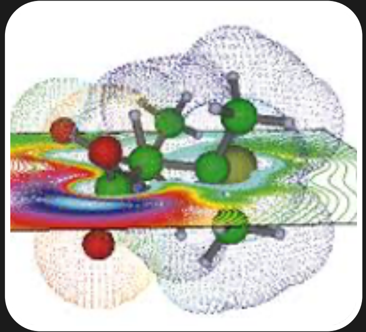




AI POWERED COMPUTATIONAL DOMAINS







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,

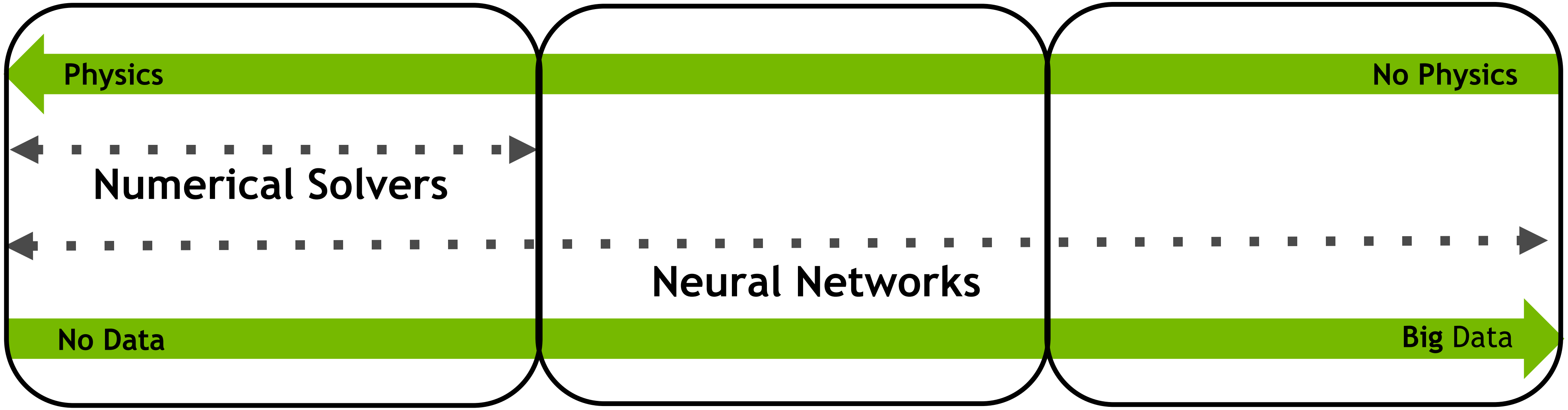
Molecular Dynamics

Process/Product Design, Manufacturing, Testing, In-Service

AI in Engineering: Taxonomy

AI IN COMPUTATIONAL SCIENCES

Primary Driver: Data vs. Physics



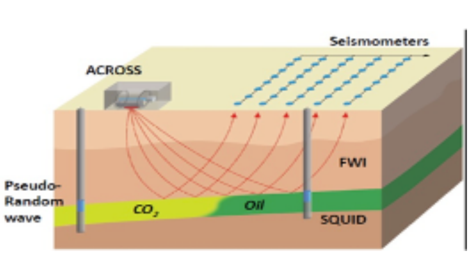
Forward Solution | **Data Assimilation** | **Data-Driven Solution**

AI BENEFITS IN ENGINEERING

Primary Driver

Physics & Data

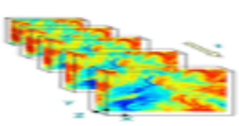
Inverse & Data Assimilation Problems



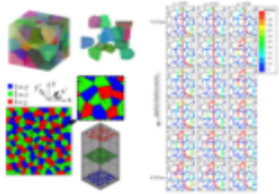
Full Wave Inversion

Aneurysm

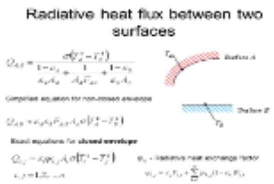
Improved Physics & Predictions



Turbulence



Micro-mechanical Material Model

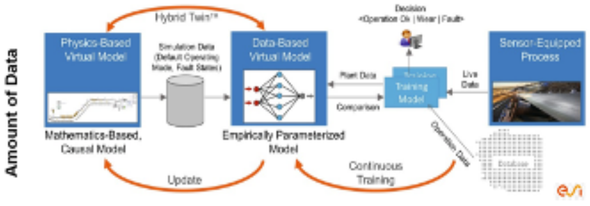


Radiation

Physics

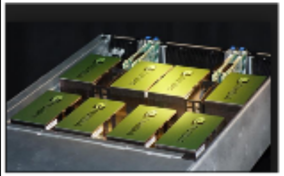
Real Time Simulations

Robotics

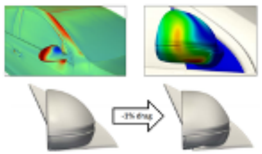


Digital Twin - IoT/Simulation

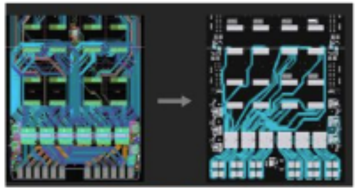
Design Space Exploration



Heat Sink



Aerodynamics



Vias on a PCB

Small to Medium

Large

Training Time

AI BENEFITS IN ENGINEERING

Primary Driver

	Small to Medium	*Large*
Physics & Data	**Enrich the Sensor or Imaging data with Physics** Discovering hidden physics by assimilating data from observation with calculations	**Higher Accuracy & Predictability** Replace solver empirical models by NN trained on highly accurate physics results or experiments Predict step n+1000 result from step n
Physics	**Improved Time to Decision** Getting rapid results for a single model using inferencing in a trained NN	**Best Design** Fast simulation surrogate enables increased sampling of many models for optimization

Training Time

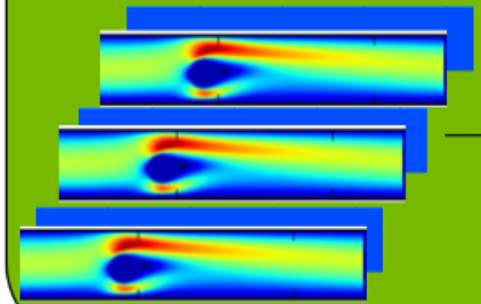
SOLVING PDES WITH NEURAL NETWORKS

A **Data Driven Neural Network**
requires training data

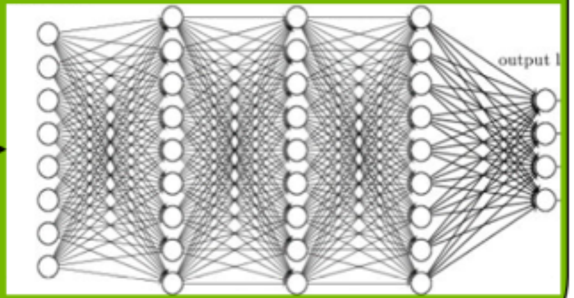
A **Physics Driven Neural Network**
solver does **NOT** require training data

Supervised Data Driven Approach

Simulation Dataset



Neural Network



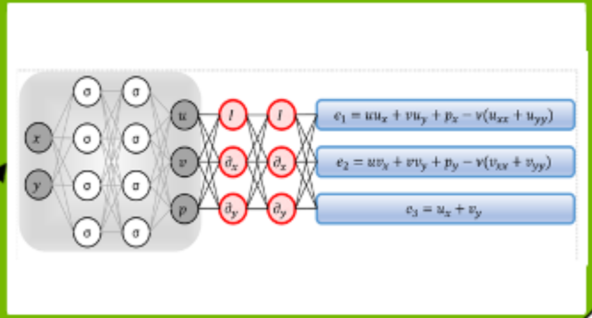
N layers

Unsupervised Physics Driven Approach

PDE and Boundary Conditions

$$0 = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$$
$$0 = u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{\partial p}{\partial x} - \nu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)$$
$$0 = u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \frac{\partial p}{\partial y} - \nu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right)$$

Neural Network



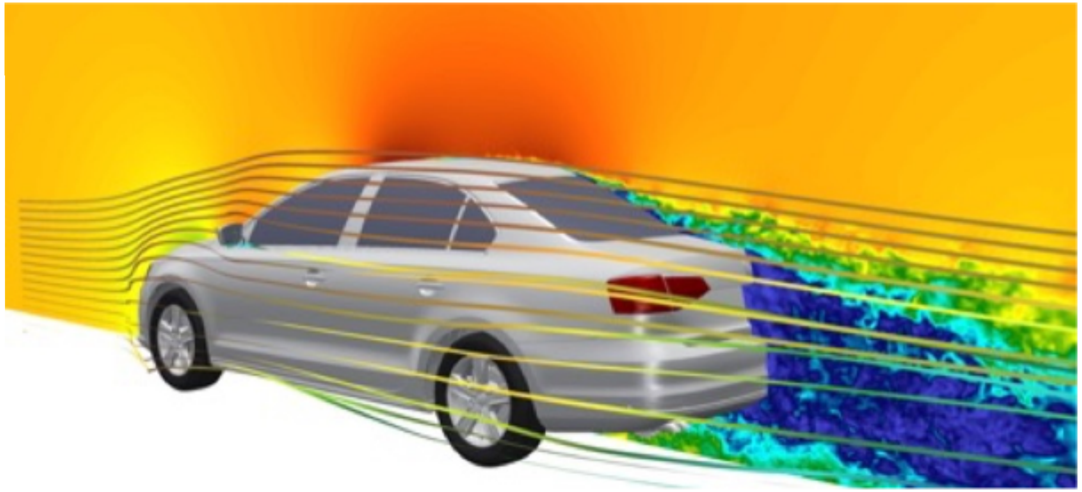
m*N layers (for m^{th} order PDE)

Data Driven Neural Networks (DDNNs)

DATA DRIVEN METHODS

Pros
Not dependent on Physics

Cons
No physics awareness; Generalization ability may be limited
Need to generate a lot of simulations (accuracy dependent on the simulation code)
Not very efficient for complex 3D geometries/curved surfaces
Interpolation/extrapolation errors



SimNet : Physics Informed Neural Networks (PINNs)

Physics-informed neural networks (PINNs)

$$\mathbf{W}^* = \arg\min_{\mathbf{W}} \mathcal{L}_{phy}(\mathbf{W}),$$

$$\mathcal{L}_{phy}(\mathbf{W}) = \underbrace{\lambda_e \left\|\mathscr{F}(\mathbf{u}^p, p^p)\right\|_{\Omega_f}}_{\text{Equation loss}} + \underbrace{\lambda_b \left\|\mathcal{B}(\mathbf{x}, \mathbf{u}^p, p^p, \boldsymbol{\theta})\right\|_{\partial\Omega_f}}_{\text{Boundary loss}} + \underbrace{\lambda_d \left\|\mathbf{f}^d(t, \mathbf{x}, \boldsymbol{\theta}) - \mathbf{f}^p(t, \mathbf{x}, \boldsymbol{\theta}; \mathbf{W})\right\|}_{\text{Data loss}},$$

Late '90s — Solving simple ODEs and PDEs (Lee et al. 1990, Lagaris et al. 2000)

Most recently — revived due to recent advances in AI and machine learning.

- Seminal work of PINN for solving forward/inverse PDEs (Karniadakis and co-workers 2018, 2019 ...)
- Applications of PINN for subsurface flows (Wang et al 2020, JoH), VIV (Raissi et al. 2020, JFM), turbulent flows (Raissi et al. 2019 PRF, Jin et al. 2020 arXiv), cardiovascular systems (Sun et al. 2019 CMAME, Raissi et al. 2020 Science, Kissas et al. 2020 CMAME, Costabal et al. 2020, Frontiers), metamaterial design (Fang et al. 2019 IEEE, Liu et al. JMD, Chen et al. arXiv.), and others
- Extended for multi-fidelity datasets (Meng and Karniadakis 2020 CMAME) and system identification problems (Tartakovsky et al. 2018, Berg and Nystrom 2019 JCP, Lu et al. 2019 arXiv)
- Bayesian formulation (Sun and Wang 2019 TCML, Yang et al. 2020 arXiv, Yang et al. 2019 JCP, Zhang et al. 2019 JCP)
- Variational formulation (Zhang et al. 2020 JCP, Samaniego et al. 2020 CMAME, Mehr et al. 2020 arXiv, Kharazmi et al. 2019 arXiv), Leverage domain decomposition (Li et al. 2019, Kharazmi et al. 2020 arXiv)
- Convergence analysis (Jagtap et al. 2020 JCP, S. Wang et al. 2020, arxiv ....)
- Surrogate modeling and stochastic PDEs (Sun et al. 2019, Karumuri et al. Zhu et al. 2019, Nicholos et al. 2020)

DeepXDE

build passing · docs passing · code quality A · pypi package 0.9.1 · downloads 38k · Anaconda Cloud 0.9.1 · downloads 41k · license Apache 2.0

DeepXDE is a deep learning library on top of TensorFlow. Use DeepXDE if you need a deep learning library that

- solves forward and inverse partial differential equations (PDEs) via physics-informed neural network (PINN),
- solves forward and inverse integro-differential equations (IDEs) via PINN,
- solves forward and inverse fractional partial differential equations (fPDEs) via fractional PINN (fPINN),
- approximates functions from multi-fidelity data via multi-fidelity NN (MFNN),
- approximates nonlinear operators via deep operator network (DeepONet),
- approximates functions from a dataset with/without constraints.

Documentation: ReadTheDocs, Extended abstract, Short paper, Full paper, Slides, Video

SciANN: Neural Networks for Scientific Computations

SciANN is a Keras wrapper for scientific computations and physics-informed deep learning.

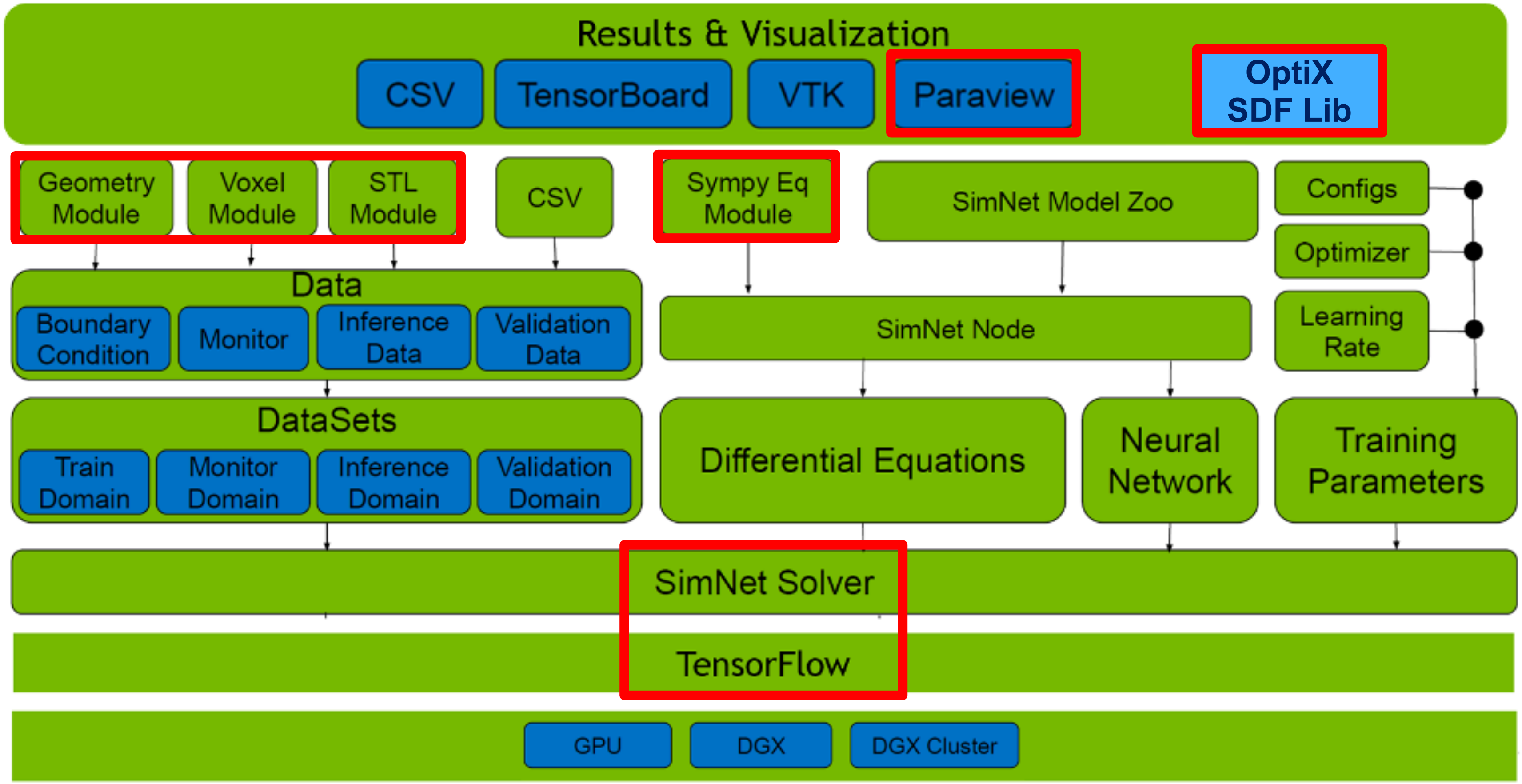
NeuralPDE

chat on gitter · CI passing · coverage unknown · codecov 75% · docs stable · docs dev

NeuralPDE.jl is a solver package which consists of neural network solvers for partial differential equations using scientific machine learning (SciML) techniques such as physics-informed neural networks (PINNs) and deep BSDE solvers. This package utilizes deep neural networks and neural stochastic differential equations to solve high-dimensional PDEs at a greatly reduced cost and greatly increased generality compared with classical methods.

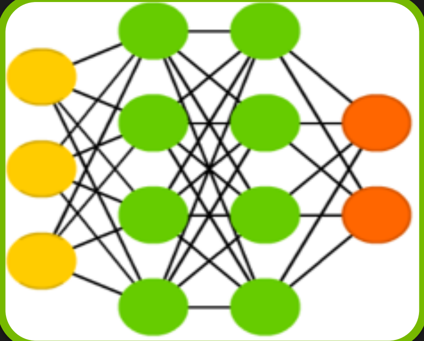
SIMNET

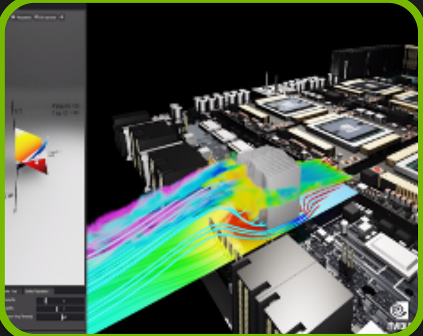
Product Architecture



SIMNET

AI-accelerated Physics Simulation Toolkit





Solve larger problems faster *with* **XLA, AMP and TF32** support, and Multi-GPU, Multi-Node implementation

Several advanced networks model Multiple Physics in ***Forward, Inverse and Data Assimilation simulations*** with accuracy & convergence

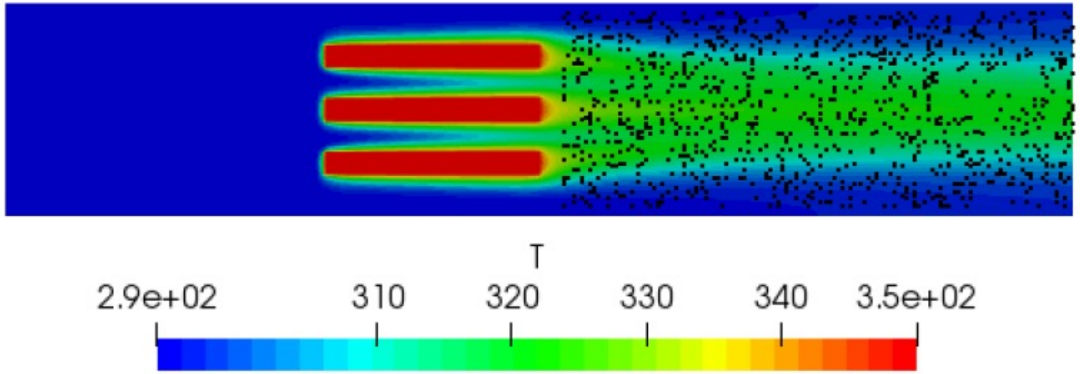
Parameterized system representation to solve ***multiple scenarios simultaneously***

APIs for implementing new Physics, Geometry, and Domains and detailed ***User Guide examples***

SimNet Results

INVERSE PROBLEMS

Finding Unknown Coefficients of a PDE: 2D Heat Sink

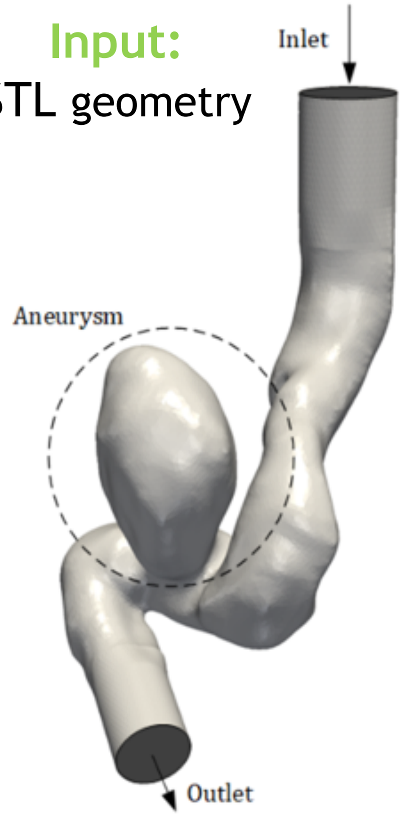


Property	OpenFOAM (True)	SimNet (Predicted)
Kinematic Viscosity (m^2/s)	1.00×10^{-2}	1.03×10^{-2}
Thermal Diffusivity (m^2/s)	2.00×10^{-3}	2.19×10^{-3}

TRANSIENT: MEDICAL IMAGING OF AN ICA

SDF LIBRARY (NVIDIA OptiX)

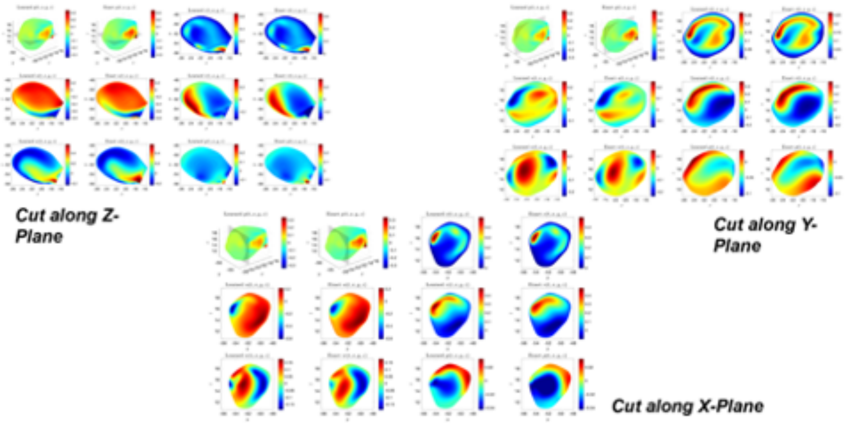
Input:
STL geometry



Intracranial Cerebral Aneurysm

SimNet

Nvidia SDF library → **Output:** Point Cloud, location relative to surface, SDF & SDF derivatives for SimNet → **NN Solver**

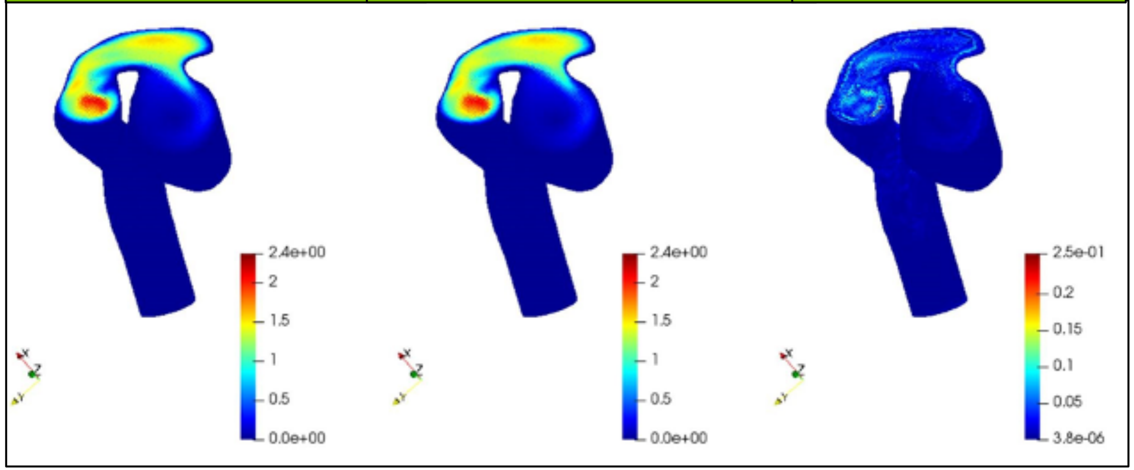


https://www.youtube.com/watch?v=QjY_8xFjsgE

ICA – COMPARISON BETWEEN SIMNET & OPENFOAM

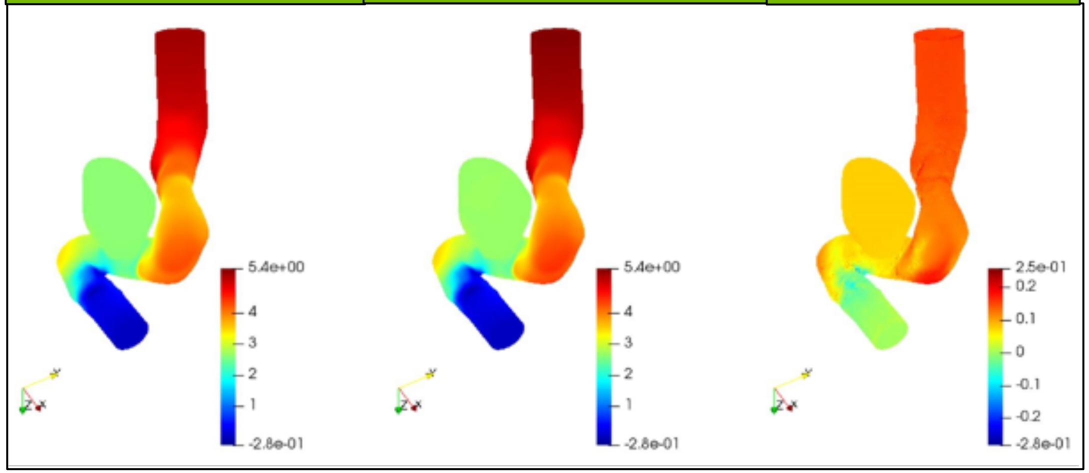
FORWARD SIMULATION

Simnet	OpenFOAM	Difference



Velocities

Simnet	OpenFOAM	Difference

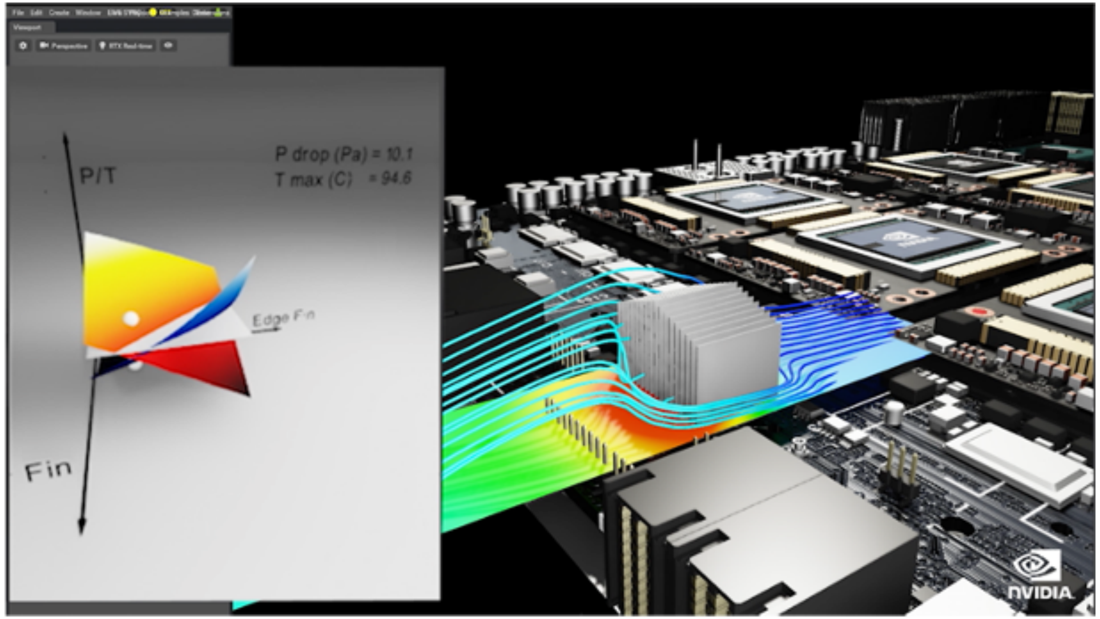


Pressures

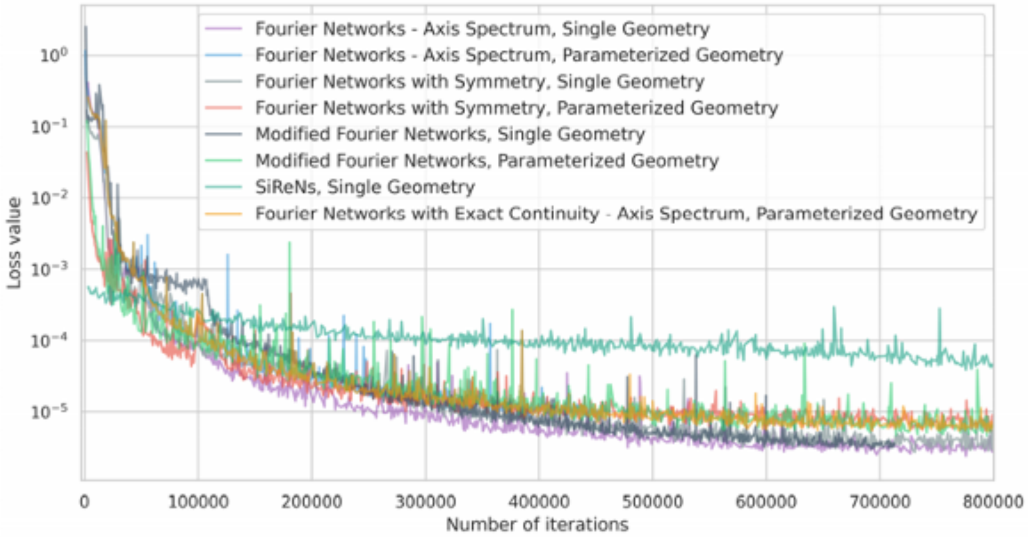
FPGA HEAT SINK

Multi-Physics Application: Fluids + Heat Transfer

https://www.youtube.com/watch?v=Oq2Mpi5pF1w

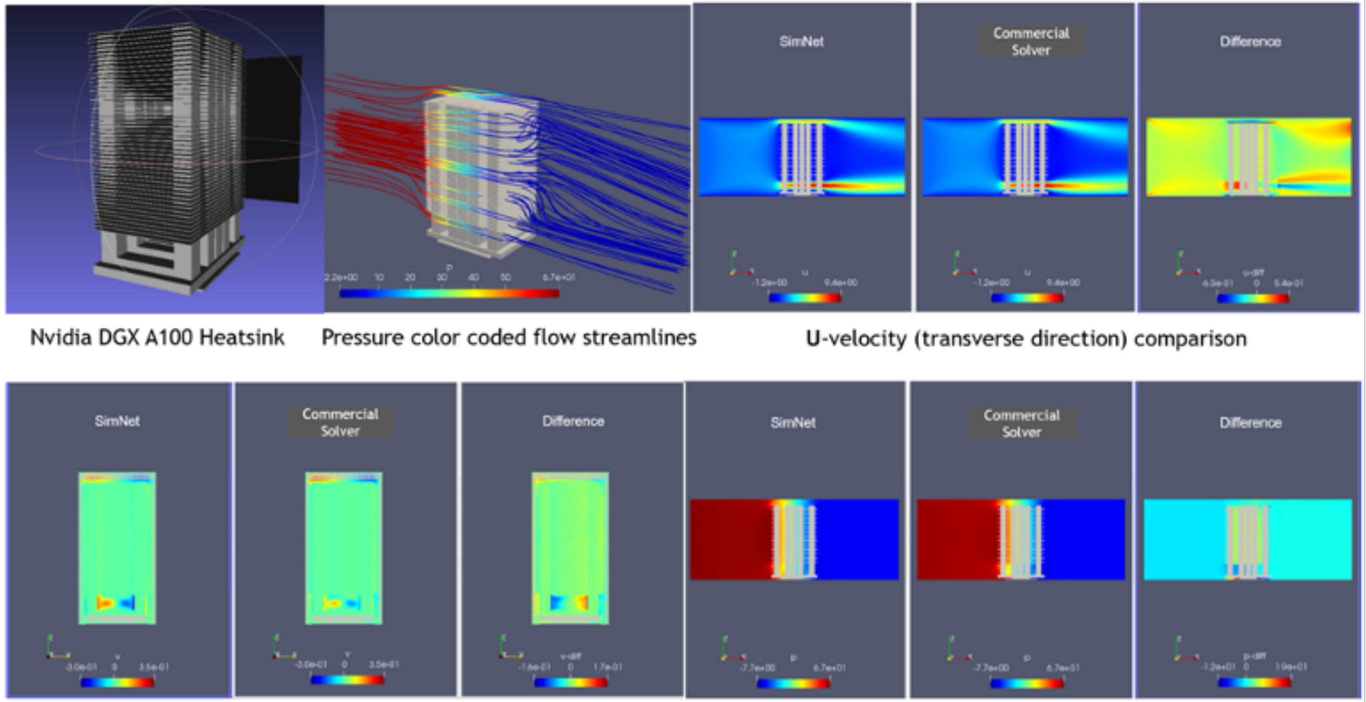


Turbulent Flow (Re=13,239.6)		
	Temperature	Pressure Drop
SimNet - Fourier Network	73.5 °C	25.7 Pa
SimNet - SiReN	72.0 °C	29.7 Pa
Commercial CFD Solver	72.4 °C	24.0 Pa



A100 NVSWITCH HEAT SINK

Multi-Physics Application: Fluids + Heat Transfer



Nvidia DGX A100 Heatsink

Pressure color coded flow streamlines

U-velocity (transverse direction) comparison

V-velocity (transverse direction) comparison

Pressure comparison

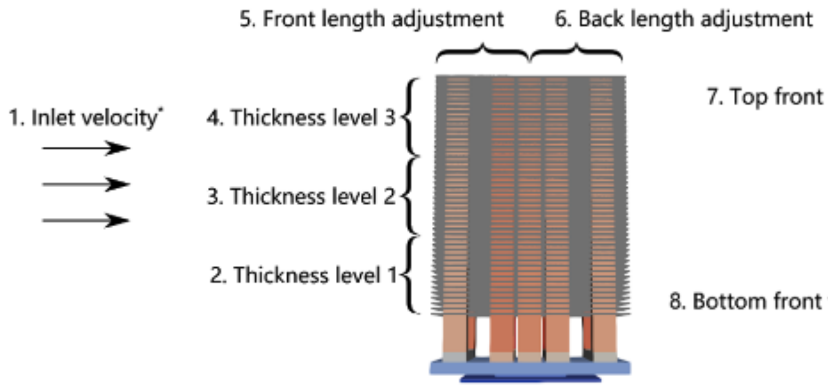
Turbulent Flow (Re=19,000)		
	Temperature	Pressure Drop
SimNet - Fourier Network	43.1 °C	4.05
OpenFOAM (method 1)	41.6 °C	3.56
OpenFOAM (method 2)	41.6 °C	4.58

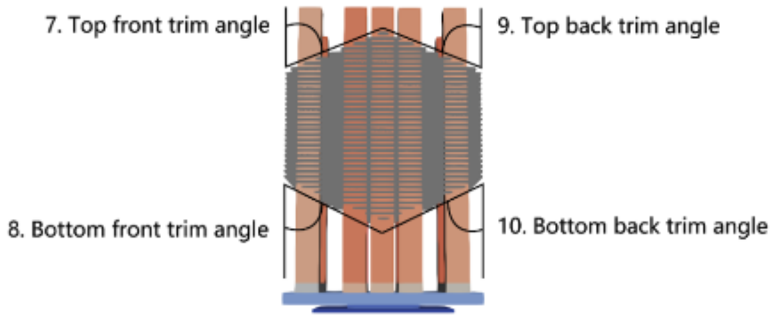
Computational Times (10 parameters, 3 values per parameter)	
SimNet	1000 V100 GPU hrs.
Traditional Solver (OpenFOAM) 59,049 separate runs (26 wall hours on 12 CPU cores)	18.4 Million CPU hrs.

PARAMETERIZED A100 NVSWITCH HEAT SINK

Optimization with 10 Design Parameters (3^10~=60,000)

Limerock Design Variables

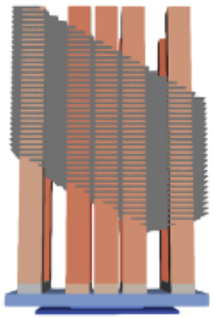




*Inlet velocity is not a design variable. It will be used for robust design optimization of Limerock in future.

Limerock Optimal Design

1. Inlet velocity: 5.7 m/s
2. Thickness level 1: 0.0031
3. Thickness level 2: 0.0044
4. Thickness level 3: 0.0030
5. Front length adjustment: -0.0124
6. Back length adjustment: 0.0025
7. Top front trim angle: 0.0223 rad
8. Bottom front trim angle: 0.5197 rad
9. Top back trim angle: 0.5147 rad
10. Bottom back trim angle: 0.2217 rad

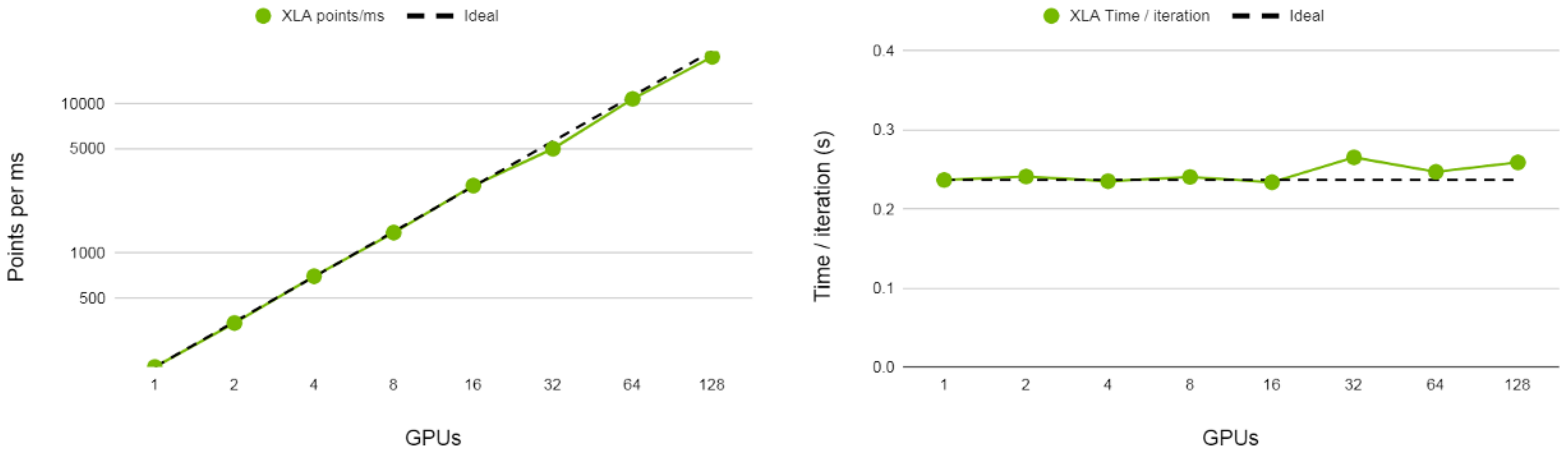


Max allowed pressure drop: 2.59
Number of random design evaluations: 4M

Peak temperature: 38.25 degC
Pressure drop: 2.5896

Multi-GPU/Multi-Node

Scalable Performance (16 DGX1 - 128 V100 GPUs)

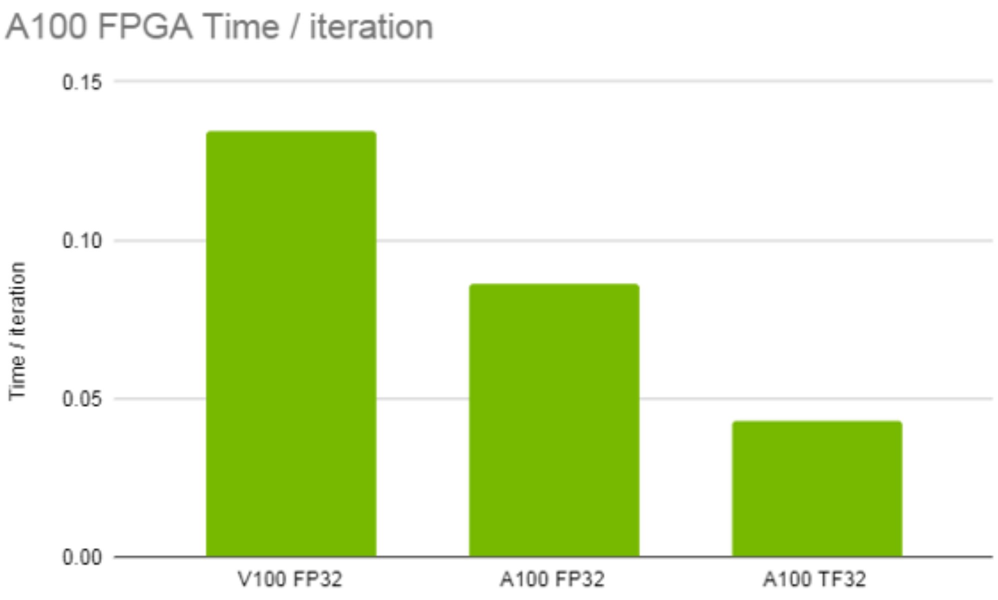


(a) Points per ms

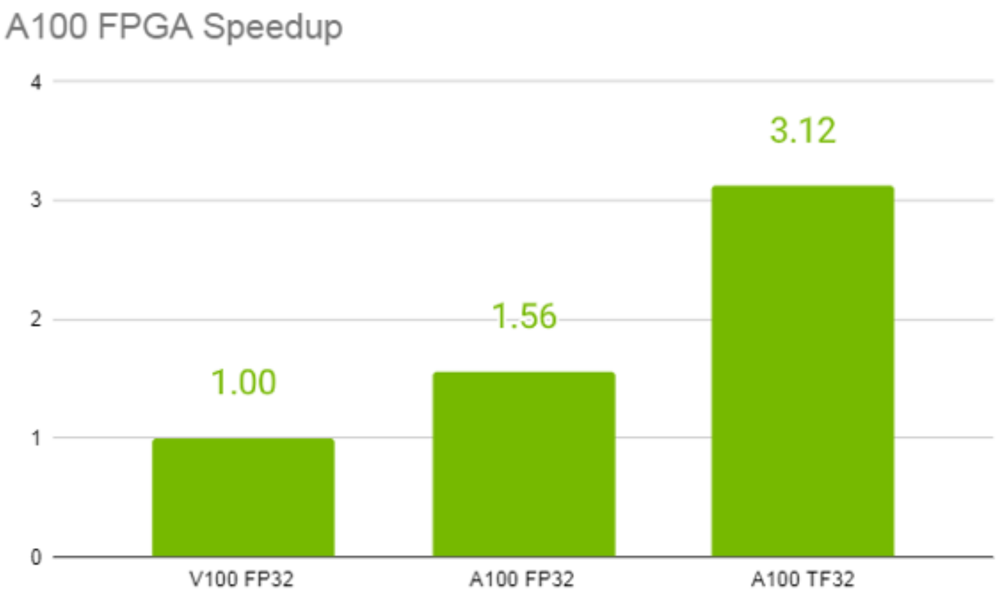
(b) Time per iteration

A100 Performance

TF32 vs FP32 & A100 vs V100



(a) Time per iteration



(b) Speed-up

SIMNET ROADMAP

(Content & release dates are subject to change)

V0.1 (May '20)

- PINNs for multi-Physics:
  - CFD & Heat Transfer
  - Integral Losses
  - SDF weighting
- APIs:
  - Domain, Geometry, Physics
- Parametrized Geometry
- Multi-GPU & Multi-Node
- AMP and XLA support
- Code documentation
- User Guide
- **Use Case:** FPGA (laminar flow)

V0.2 (Sept '20)

- STL geometry integration
- Improved accuracy:
  - Turbulence
  - Exact mass-balance
- New PINNs architecture I:
  - Fourier Features
  - Perdikaris, DGM
  - SiReNs
- Global & Local learning rates
- A100/TF32 support
- Adaptive activation function
- Halton Sequences
- **Use Case:** FPGA (turbulent flow)

20.12 (Dec '20)

- Ordinary differential equations
- Integral method (variational form)
- Linear Elasticity for solids
- XLA for Fourier Features Network
- New PINNs architecture II:
  - Highway networks
  - Discontinuous Galerkin
- Learning rate improvements for strong scaling
- **Use Cases:**
  - Digital Twin
  - Aneurysm
  - Multi-parameter training

- Conjugate Heat Transfer
- Freq. domain Electromagnetics
- Acoustics
- GPU based CSG library
- Mitchell-Beltrami: Porous Medium
- New PINNs architecture III:
  - Temporal: Convolutional-LSTM
- Omniverse visualization I
- Linear elastic fracture mechanics
- **Use Cases:**
  - 2D FWI to predict velocity model
  - Heat Sinks
  - GTC demo

- Eigenvalues
- Transient Electromagnetics
- Forward wave propagation with multiple sources on layered velocity model
- Uncertainty Quantification
- New PINNs architecture III:
  - Hybrid
  - metaSDF
- Omniverse visualization II
- Automated testing infra.
- **Use Cases:**
  - Automotive Aerodynamics
  - PCB Signal Delay

- Nonlinear Solids
- Transient Electromagnetics
- Design Optimization
- Elastic wave propagation
- Compressible Flows
- New PINNs architecture III:
  - Hyper-networks
  - meshSDF
- Omniverse visualization
- Automated testing infra.
- **Use Cases:**
  - 3D FWI
  - Data Center Cooling

Early Access

https://developer.nvidia.com/nvidia-aerial-sdk-early-access-program

NVIDIA Aerial SDK - Early Access Program

Our focus is to provide the SDK to customers who aim to build a RAN system. This is an NDA program available to download the L1 implementation of 5G stack. Please register or log in using your company email credentials to help us evaluate and grant access. We thank you for your patience as we ramp up this program.

The **NVIDIA Aerial SDK** helps developers building telecommunications applications build the most programmable, scalable and energy efficient software-defined 5G networks.

If you are a developer building **Virtual Radio Access Networks (vRAN)** and are interested in evaluating our SDK please request access to the NVIDIA Aerial early access program by clicking the "Join now" button below.

Who is it for: Aerial SDK selected customers for developing cloud native vRAN solutions for various deployment scenarios and applications.

- Telco vendors looking to develop network stack over vRAN L1 layer
- NEP customers to collaborate on network infrastructure with transmission equipment, CPE and other system integration.
- Higher education and research community focusing on advanced topics in 5G stack, NFV, SDN, Open RAN, etc.
- Government agencies looking to optimize signal processing using GPU compute integrated with Mellanox SmartNIC offering.

Please review the license agreement before joining the program. If you accept the terms please join below.

Join now

NVIDIA GTC 21

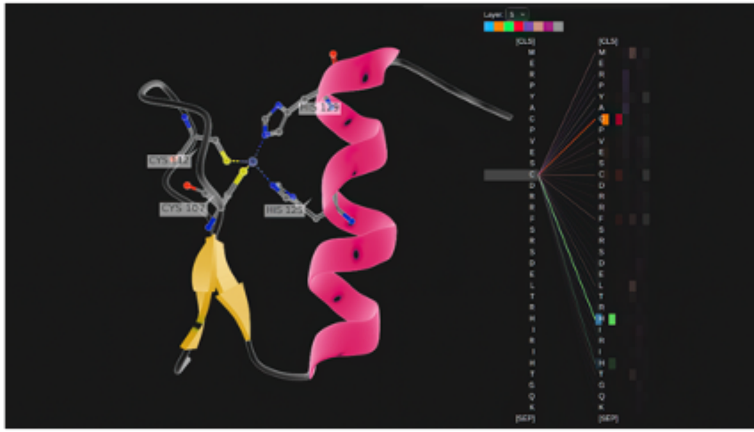
THE CONFERENCE FOR AI INNOVATORS, TECHNOLOGISTS, AND CREATIVES

Join us at GTC 2021 on April 12 - 16 for the latest in AI, HPC, healthcare, game developing, networking, and more.

NVIDIA’s GTC brings together a global community of developers, researchers, engineers, and innovators to experience global innovation and collaboration.

Don’t miss out on the exclusive GTC keynote by Jensen Huang on **April 12**, available to everyone.

Visit www.nvidia.com/gtc to learn more and be notified when registration opens.




NVIDIA