



AI Factory for Science

Where traditional high-performance computing meets AI-native innovation for scientific computing.



From HPC to AI Factory

For decades, high-performance computing (HPC) has been the backbone of scientific discovery and engineering innovation. Researchers have relied on physics-based simulations to understand fluid flow, model molecular dynamics in chemistry, and predict the weather.

End of an Era

Moore's Law and Dennard scaling have ended. No longer can new generations of processors deliver twice the speed at half the energy. Scientific progress in traditional HPC simulations is slowing because of speed and energy limitations. Yet solving the world's pressing issues—from pandemics to energy generation—requires two to four orders of magnitude more performance.

Achieving this will require new methods.

AI-Augmented and AI-Native Simulation

We're now entering a new era—where AI augments and even replaces parts of the physics-based HPC simulation.

In AI-augmented workflows for fluid dynamics, machine learning can speed up meshing, improve turbulence models, and build surrogate models for faster iterations without abandoning the physics foundation.

With the latest AI-native approaches, physics-informed neural networks (PINNs) and foundation models, like NVIDIA's PhysicsNeMo, FourCastNet, and CorrDiff, deliver high-resolution predictions with statistical probabilities—achieving what was once impossible in scale, speed, and efficiency.

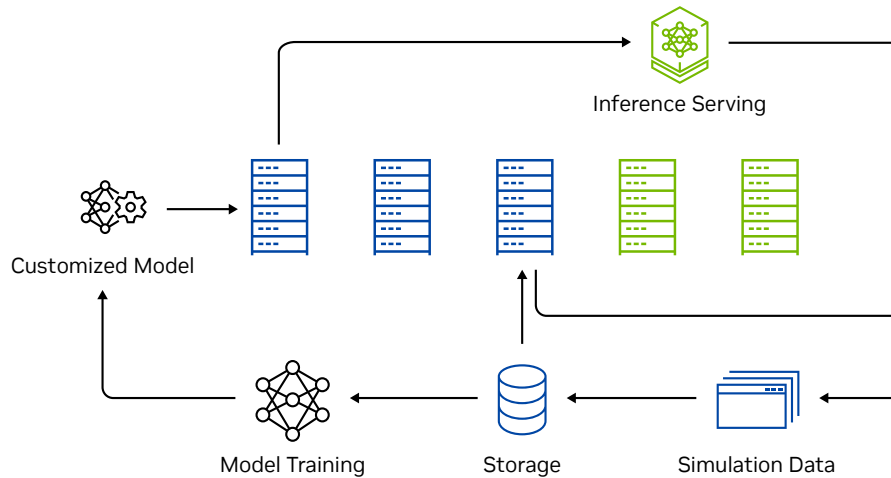
This transformation extends beyond the algorithm level, requiring scalability across the entire infrastructure.

Key Challenges for AI Workloads in Scientific Computing

- > Integration complexity
- > Computational demands
- > Data availability
- > Scalability
- > Scaling interprocessor communications
- > Storage limitations
- > Infrastructure limitations
- > Security and compliance
- > Model transparency

These challenges are addressed by robust, scalable, and flexible AI infrastructure and AI factory workflow solutions to realize the transformative promise of AI in scientific computing.

Integrated Scientific AI Factories



Enter the NVIDIA AI Factory

NVIDIA AI Factory is an end-to-end platform that unifies HPC, AI, and industrial-scale automation into a single, seamless workflow.

It spans the entire journey—from collecting and simulating data to training massive AI models to deploying real-time inference at scale.

With NVIDIA AI Factory, researchers can generate insights and deliver solutions at speeds and scales that were previously unimaginable.

How it All Fits Together

Simulation	Traditional	AI-Augmented	AI-Native	AI-Factory
Compute	CPU/GPU clusters	GPU-accelerated with AI assist	AI-optimized GPUs + PINNs	Scalable factory with Grace Blackwell + NVIDIA Networking
Workloads	Pure physics solvers	Hybrid (physics + AI)	Mostly AI, with physics constraints	Runs all, at scale
Outcome	Accurate but slow	Faster + scalable	Real-time + scalable	Enables everyone to build and deploy at industrial scale

Ready to Get Started?

To learn more about NVIDIA data center products, visit www.nvidia.com/en-us/data-center/products

© 2025 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA and the NVIDIA logo are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 4121976. AUG25

Partner
Logo

