HPC in Industry

Brendan McGinty
NCSA Industry Director

ILLINOIS
NCSA | National Center for Supercomputing Applications

NCSA goal: 6 months ahead of competition

Industry Dedicated

- Technical Teams
- HPC Resources
- Business Leadership and Project Management

Tradition

- Industry as a part of NCSA’s mission for > 30 years

Culture

- Work at industrial pace with NDAs
- Deliver on time and under budget

Largest Industrial HPC Program in the World

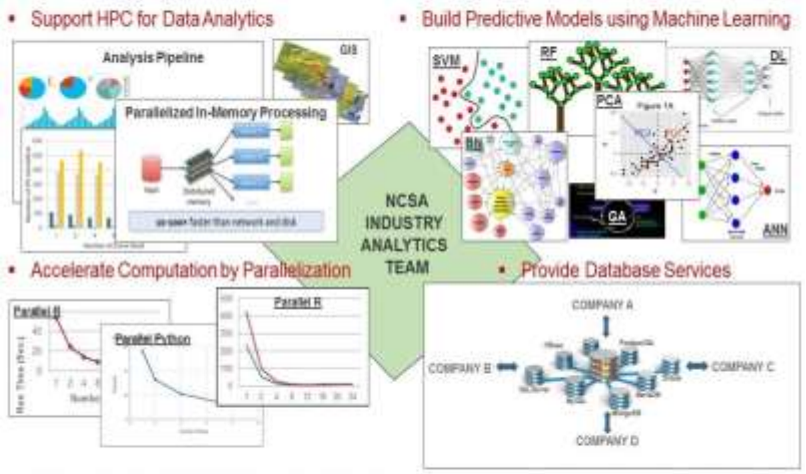
NCSA Industry Technical Team Expertise

- Modeling and Simulation
- Bioinformatics and Genomics
- “Big” Data Analytics, GIS, and AI
- Code Profiling and Optimization
- Rapid User Support and Domain/HPC Training
- Cyber Infrastructure and Security
- Visualization
- Much more at NCSA and the University of Illinois

Industrial Data Analytics Group

Expertise

- Data analysis and management on massive scale
- In-memory data analytics
- Industrial applications of machine learning
- Converge of artificial intelligence with high-performance computing
- Geographic Information System (GIS) for spatial data

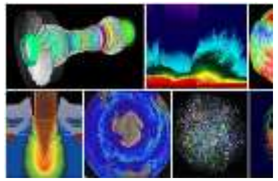


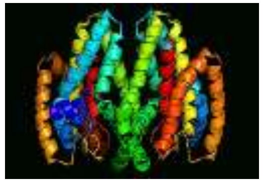
Current Projects

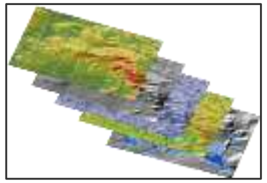
- **Pharmaceutical:** Drug discovery, side effect analysis
- **Agriculture:** Predicting crop production, crop disease
- **Energy**: Manufacturing and operational optimizations using AI
- **Insurance:** Analyzing insurance policies, cost control
- **Engineering:** Re-engineering workflows, simulating operations
- **Manufacturing:** Smart manufacturing, quality assurance, supply-demand control
- **Finance:** Decision-making on loans, predicting customer's credibility

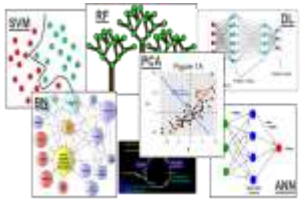
Dora Cai
Technical Program Manager
doracai@Illinois.edu

*Forge – The HPC Environment for Industry









- **Latest and best**
  - Computing (Intel/Skylake 192-256 GB)
  - In-memory “big” data analytics (SPARK)
  - GPU driven AI technologies (V100)
- **99% uptime and live upgrades**
- **Development and production workhorse**
- **Rapid user support and advanced consulting**
- **Built exclusively for Industry’s applications and workflows**

National Petascale Computing Facility

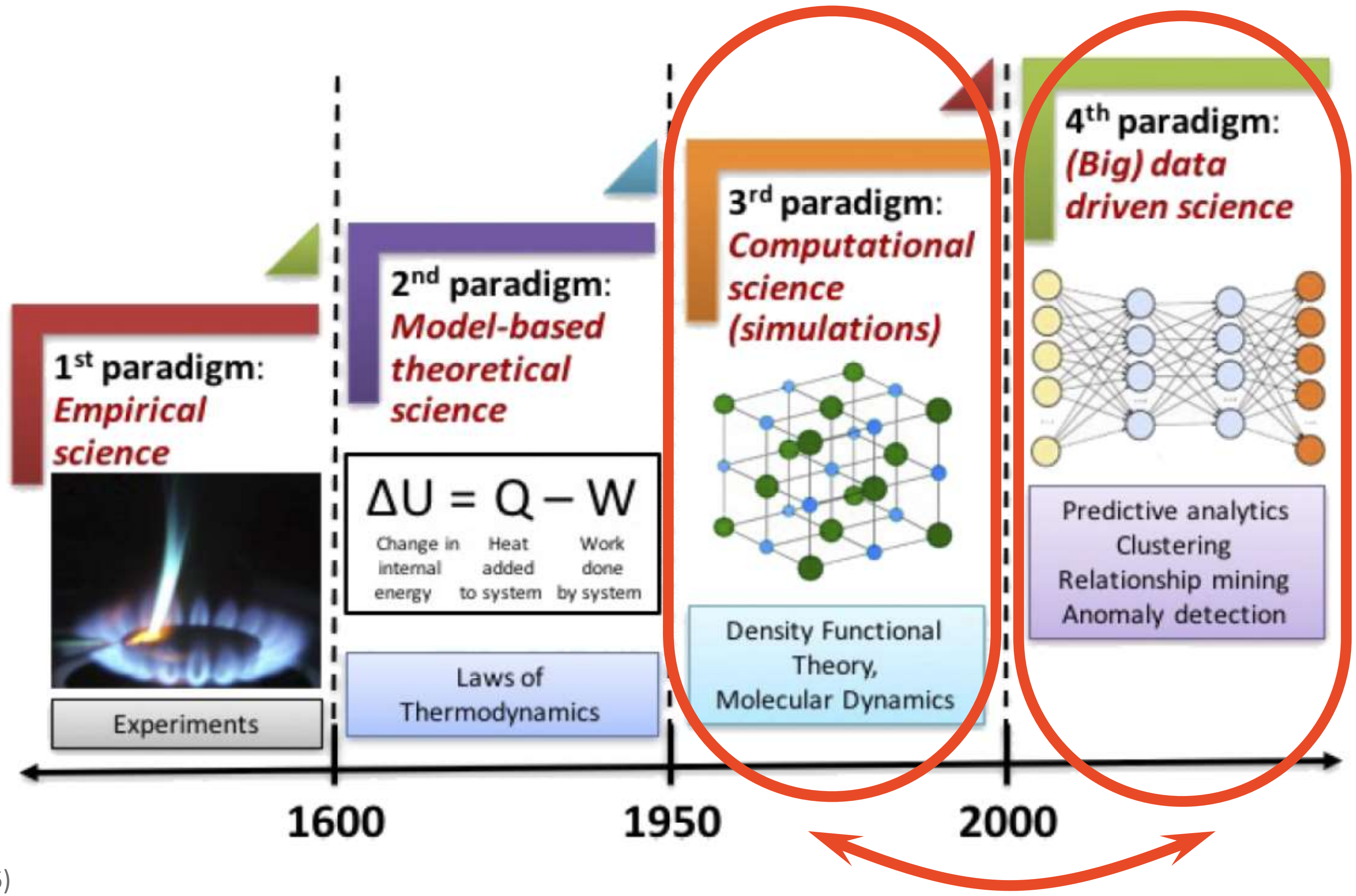
World-Class Data Center

- Dept. of Energy-like security
- 88,000 sqft
- 25 MW of power; LEED Gold
- 400+ Gb/sec bandwidth

Hosting Benefits to Industry

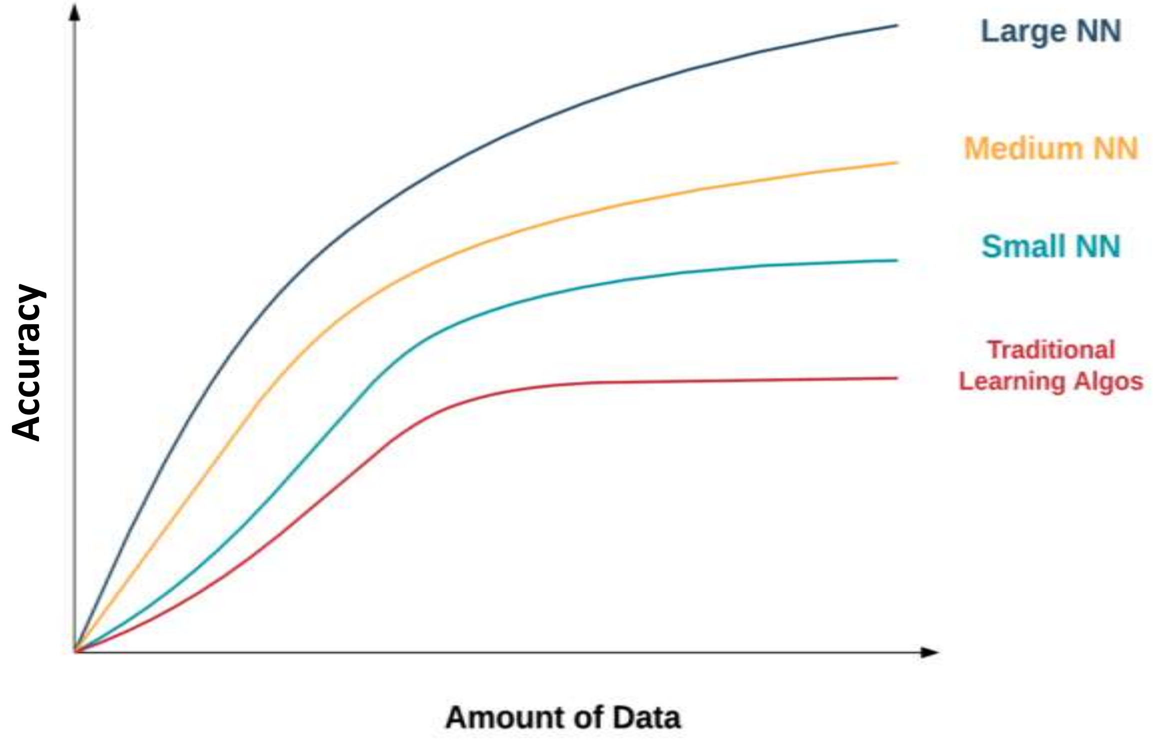
- Low-cost power & cooling
- 24/7/365 Help Desk
- Adjacent to and aligned with UIUC Research Park

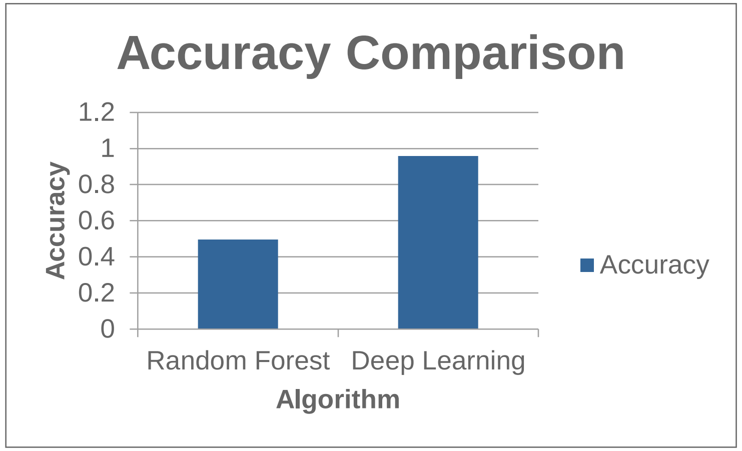
Four Paradigms in Science and Engineering

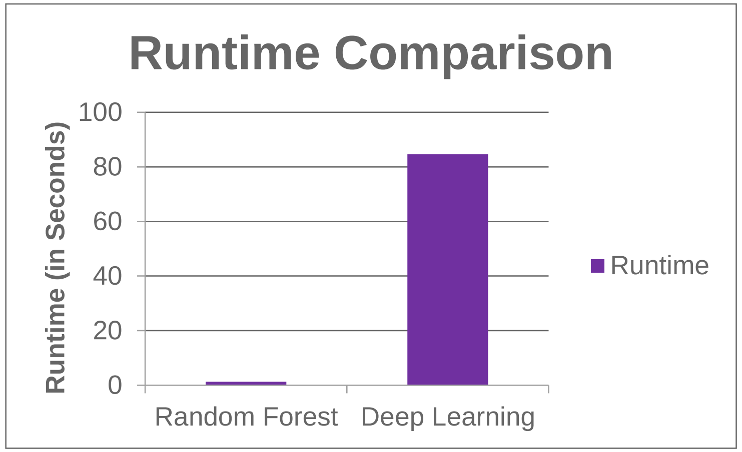


APL Materials 4, 053208 (2016)

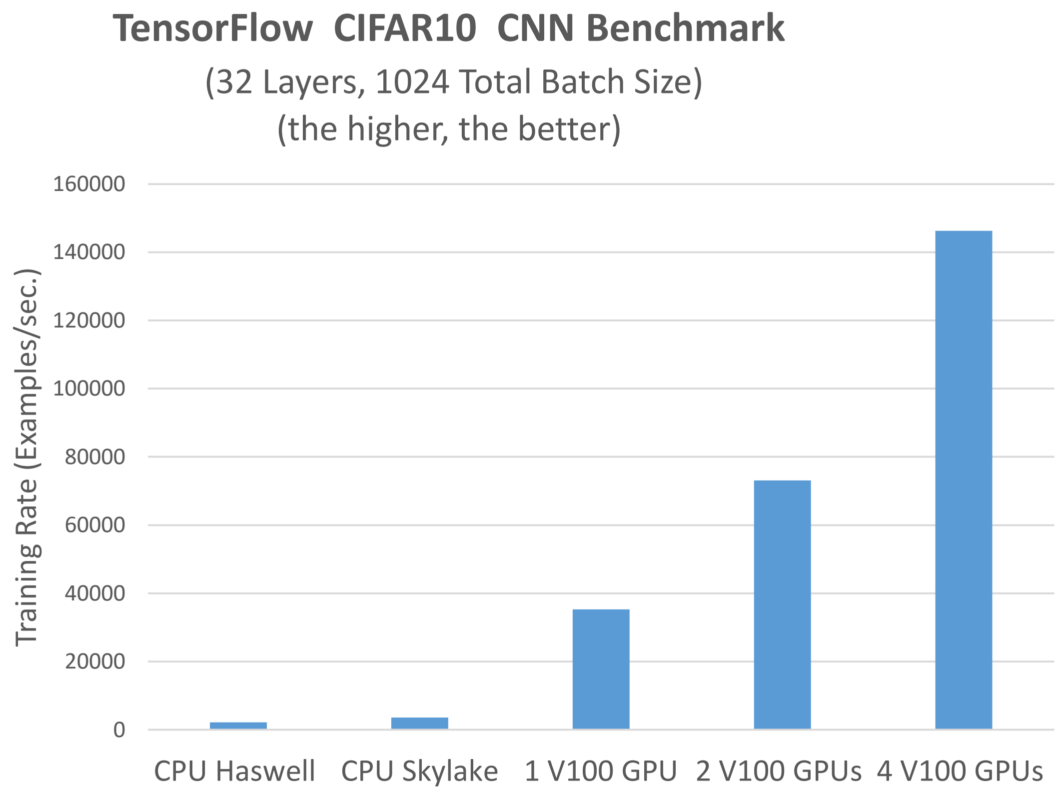
Big Data and HPC Driven Deep Learning

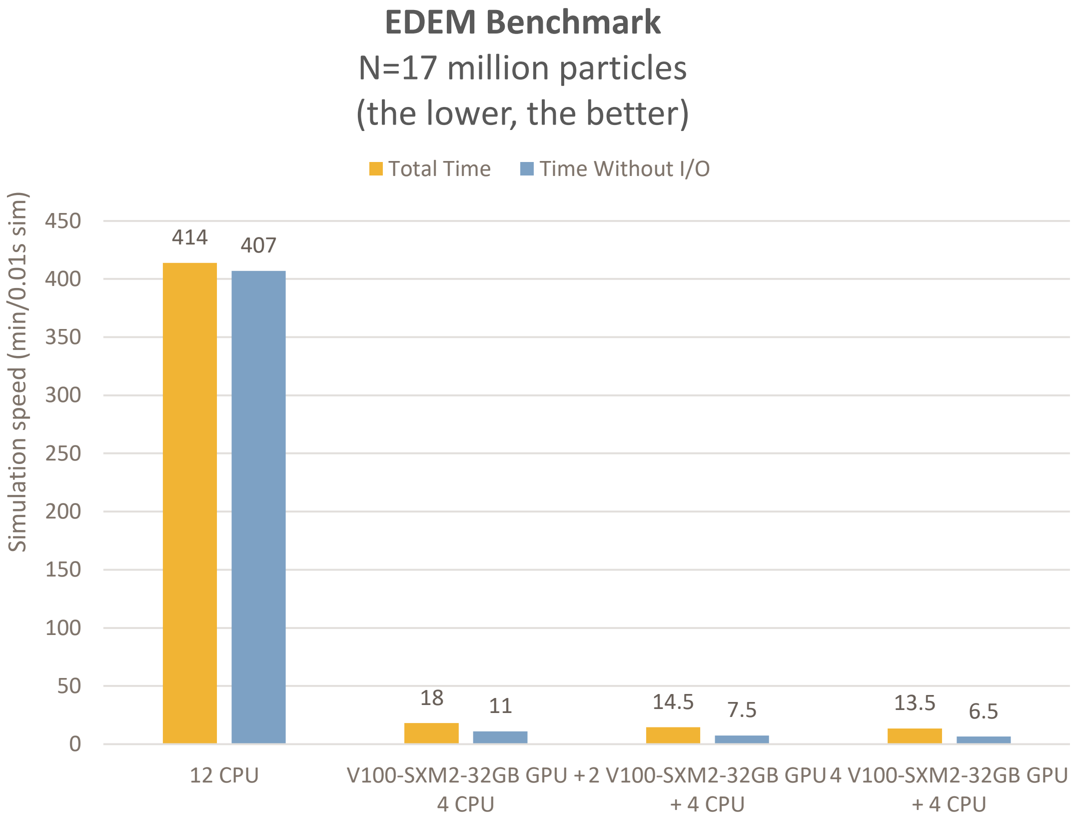




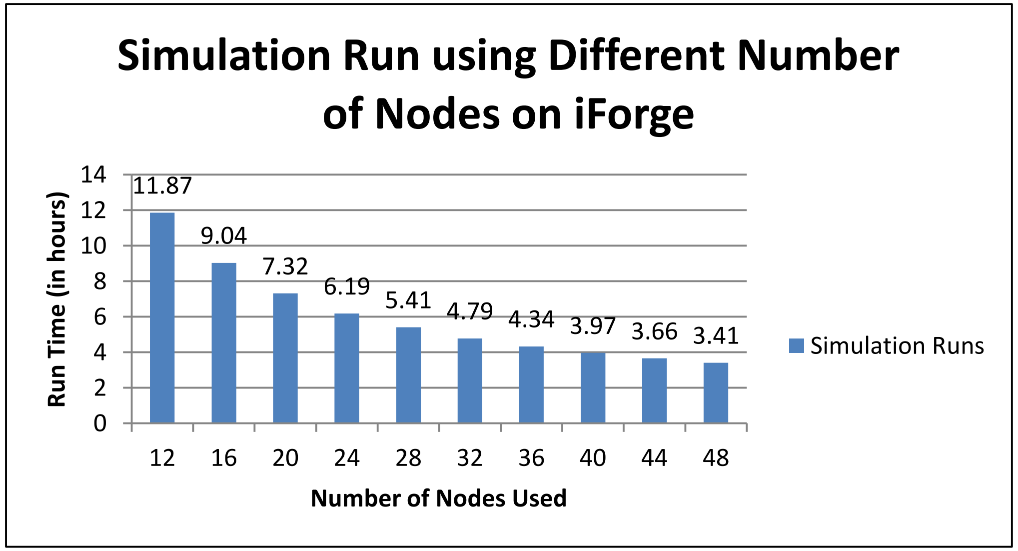


New iForge V100 GPUs Enable Significantly Faster DL Training and DEM Modeling





Large Scale Statistical HPC Analysis in Agriculture



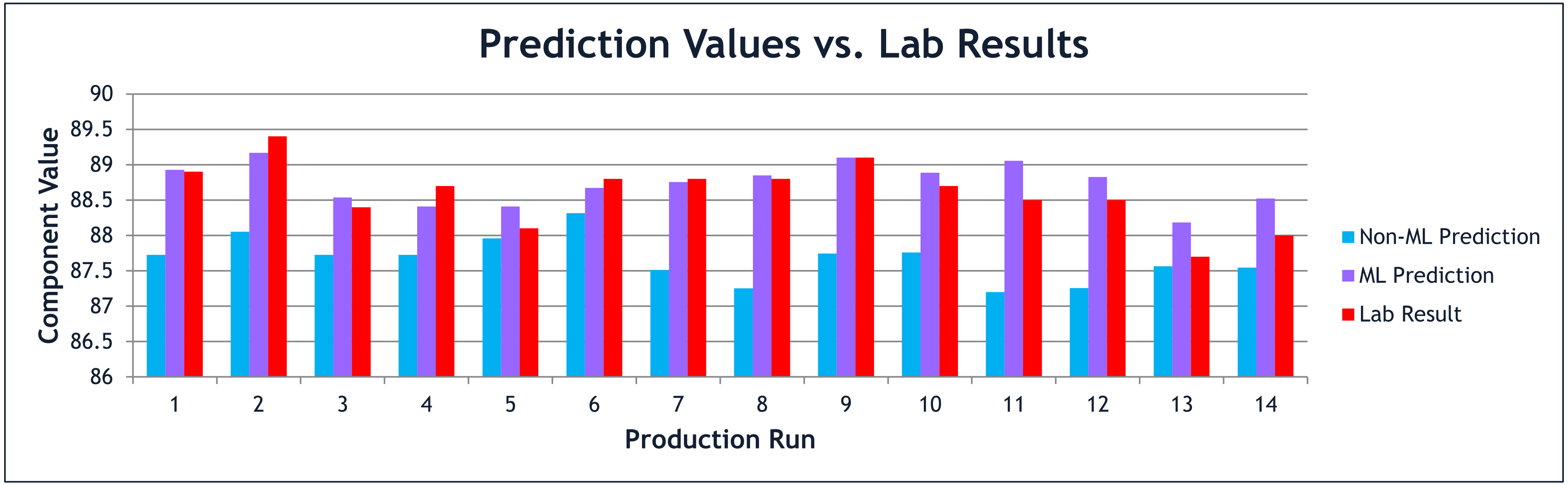
- Power statistical analysis uses massive data collected from farm field trials to allow an agriculture partner of NCSA to assess quality of their experimental designs

- NCSA has developed an efficient and scalable implementation in **R** to perform massive simulation using multi-node parallelization and variable instantiation techniques

- Our new implementation decreases the size of the program from over 50,000 lines to less than 100 lines, reduces the processing time for a simulation with over 70,000 cases from **175 days** (@partner) **to less than 3.5 hours)** (@HPC/iForge)

Courtesy of Dr. Dora Cai and an Industrial Partner of NCSA

Reduce Product Cost using AI Predictive Models

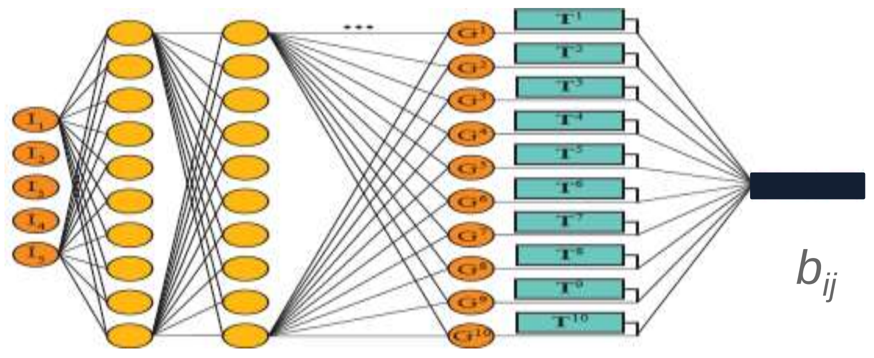


Courtesy of Dr. Dora Cai and an Industrial Partner of NCSA

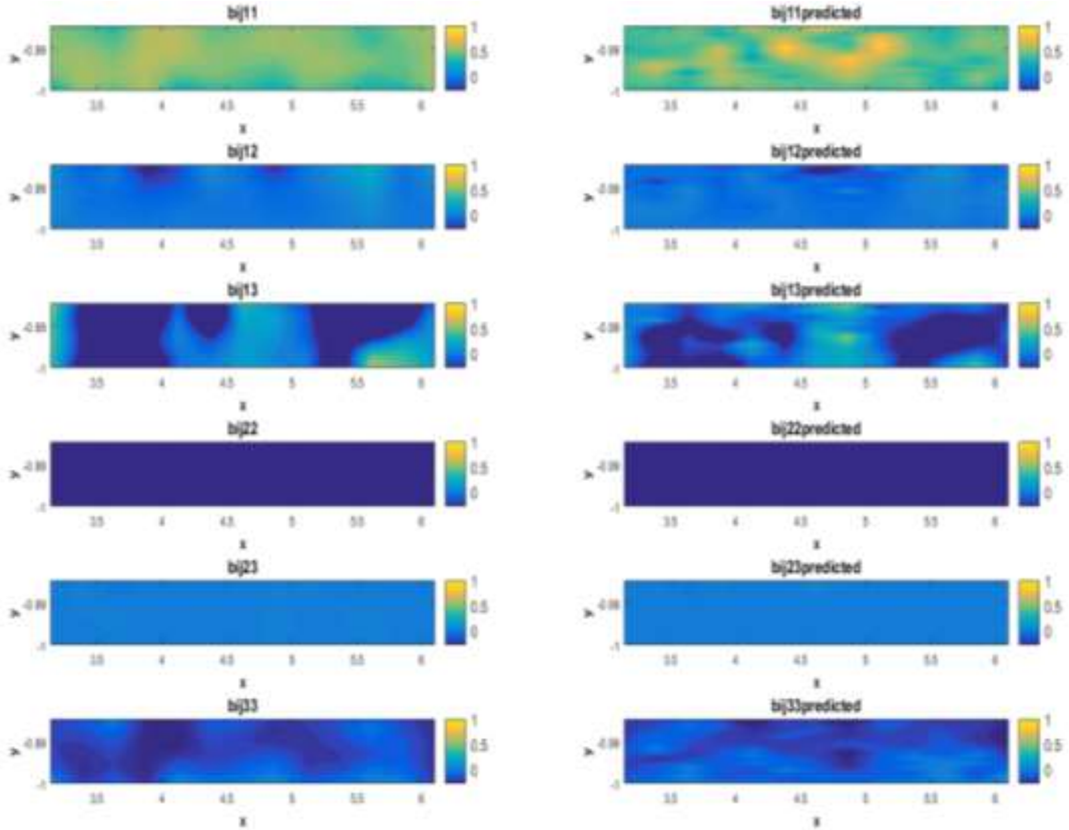
- NCSA corporate partner
- Optimize "recipe" using Machine Learning predictive models
- Make the predicted values closer to the real lab test results
- Reduce *Mean Absolute Errors* (**MAE**) from 0.73 to 0.43
- **ROI: USD$18M** annually by reducing the product cost

AI Accelerated Turbulence Modeling

$$b_{ij} = -\frac{\nu_t S_{ij}}{k} + C_1 \frac{\nu_t}{\tilde{\epsilon}} \left(2 S_{ik} S_{kj} - \frac{2}{3} S_{kl} S_{kl} \delta_{ij} \right) + C_2 \frac{\nu_t}{\tilde{\epsilon}} (2 R_{ik} S_{kj} + 2 R_{jk} S_{ki}) + C_3 \frac{\nu_t}{\tilde{\epsilon}} \left(2 R_{ik} R_{jk} - \frac{2}{3} R_{kl} R_{kl} \delta_{ij} \right)$$



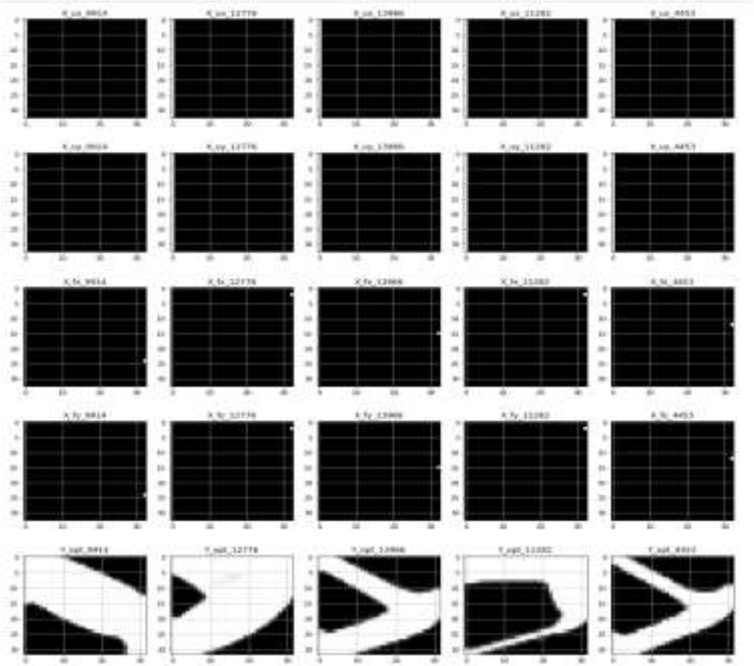
Fully connected neural network using invariant tensor from velocity gradient **to predict DNS quality of Reynolds stress under the cost of k-ε** !

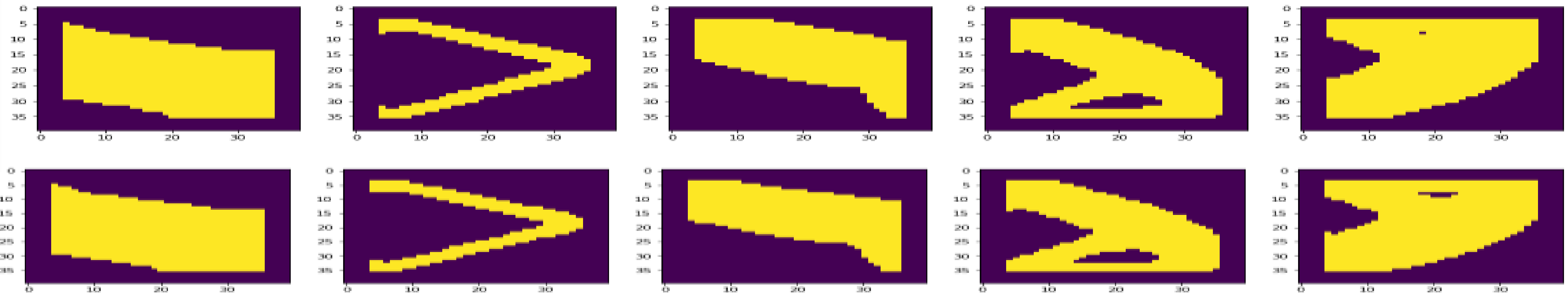


Courtesy of Dr. Shirui Luo and an Industrial Partner of NCSA

AI Accelerated Topological Optimization

1. Setup the problem geometry in Abaqus/CAE
2. Generate FEM mesh of 1. in Abaqus/CAE
3. Loop over parameters, generate training data
   - 3.1 Apply BCs and Loads
   - 3.2 Run Abaqus + Tosca
   - 3.3 Save Optimal Solution
   - 3.3 Build a multichannel image representation of 1,2,and 3.1
   - 3.4 Build a gray scale image of 3.3
4. Train Deep Neural Network model using data generated in 3.3 and 3.4
5. Evaluate accuracy
6. Release model



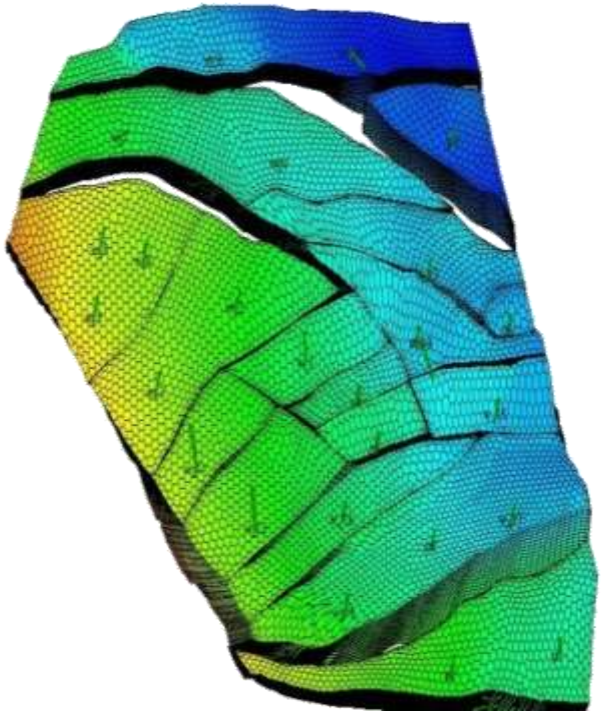


Ground Truth

Prediction

Courtesy of Diab Abueidda and Dr. Nahil Sobh

Massively Parallel Modeling & ROI



ExxonMobil

- Reservoir simulation models the complex subsurface flows of fluids in oil and natural gas reservoirs
- Previous runtime: 3.5 months on prem
- Optimized: 10 minutes on Blue Waters
- 716800 MPI processes, a sector world record in 2017 for degree of parallelism

- **Minimized costs and environmental impact**
- **ROI: USD$1+B**

Best Practices - for NCSA Industry, anyway…

- It's about the talent first, compute second
- Be consultative – and listen first!
- Provide custom solutions – no two companies are alike
- Respect their time, or lack thereof
- Leverage resources
  - University (faculty, students, research park, et al.)
  - Collaborators/vendors/partners
  - Other centers (NLs, international SCs)
- There is enough opportunity to go around!
- Never settle: keep pushing, testing, being fearless

HPC in Industry

Brendan McGinty
NCSA Industry Director
bmcginty@illinois.edu
+1 217-722-3430