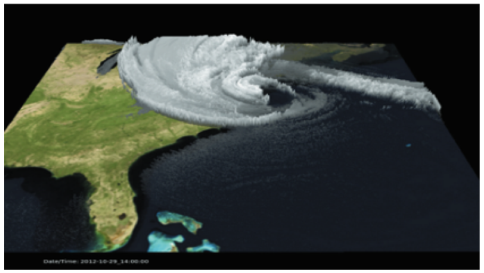
Converged Simulation and AI Workload Trends in Earth Sciences

Per Nyberg

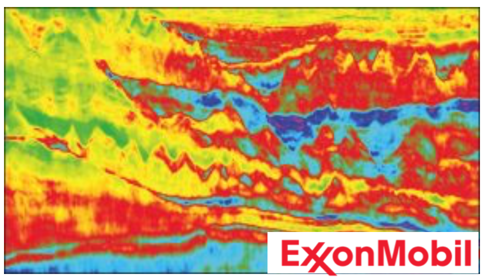
VP, Artificial Intelligence

Breakthrough Simulation Capabilities Provided by Powerful Algorithms and Supercomputers

CRAY



- Largest ever storm prediction model
- Over 4 billion points used to simulate the landfall of Hurricane Sandy
- Urban scale grid resolution of 500m (compared to standard 3km)
- Enables the research to understanding fine grained properties of hurricanes and other meteorological phenomena



- 3D seismic imaging has been the industry standard and the most accurate method for locating fossil fuel deposits
- Full wavefield inversion (FWI) has been a breakthrough capability in the industry addressing the algorithmic and computational challenges of 3D seismic
- FWI is able to leverage all of the available data to produce a full digital model of the subsurface that includes the geological and geophysical properties

TU/e

KU LEUVEN

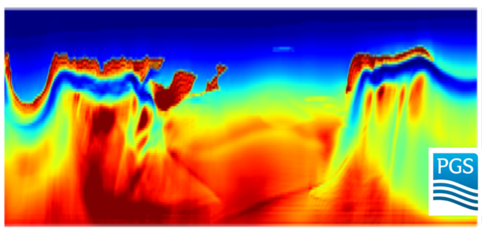
ANSYS

- Computational fluid dynamics (CFD) is used to predict flow, turbulence, heat transfer and reactions for industrial applications ranging from air flow over an aircraft wing to combustion in a furnace and from blood flow
- The largest aerodynamics sports simulation ever performed modeled a full peloton of 121 cyclists, comprising three billion cells.
- Proved that it is about four times easier to pedal in the middle of the peloton than if you are cycling by yourself.

Machine Learning is Now Augmenting These Capabilities

CRAY

- Deep learning methods are being applied to improving meteorological products including local nowcasts for weather, wind power yield assessments and air quality analysis.
- Using current and historical weather data, neural networks are being trained to optimize the orientation of wind-powered generators for maximum efficiency.



- Machine learning optimization techniques such as regularization and steering are being applied to Full Waveform Inversion (FWI) seismic imaging.
- Compared with manual velocity model determination and tuning, FWI with ML converges more quickly and efficiently.

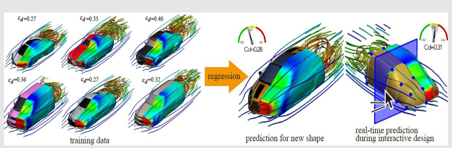


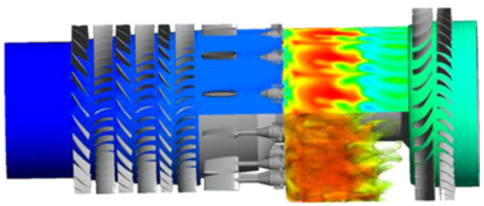
Image source: Nobuyuki Umetani, Univ of Tokyo

- Deep Learning is being applied to simulation data to explore high-dimensional spaces that typically require substantial user interaction.
- The design of experiments can be based on neural networks trained on the outputs of previous simulations, resulting in improvements in human and computer resource utilization.

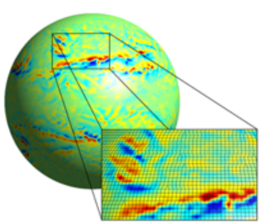
How are Individual Domains Evolving ?

SIMULATION	COGNITIVE SIMULATION	ARTIFICIAL INTELLIGENCE

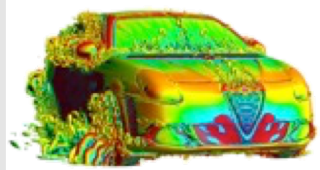
Combustion Modeling

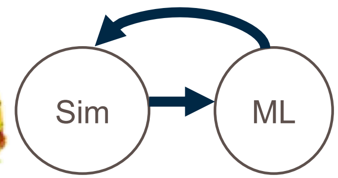


Numerical Weather Prediction

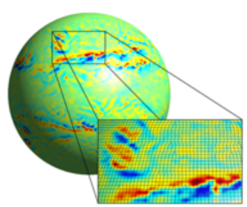


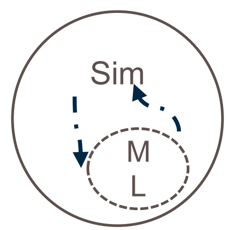
Deep CFD
"Learning on the outside"





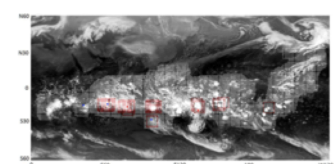
Machine Learning Based Physics Emulators in Numerical Weather Prediction Models
"Learning on the inside"



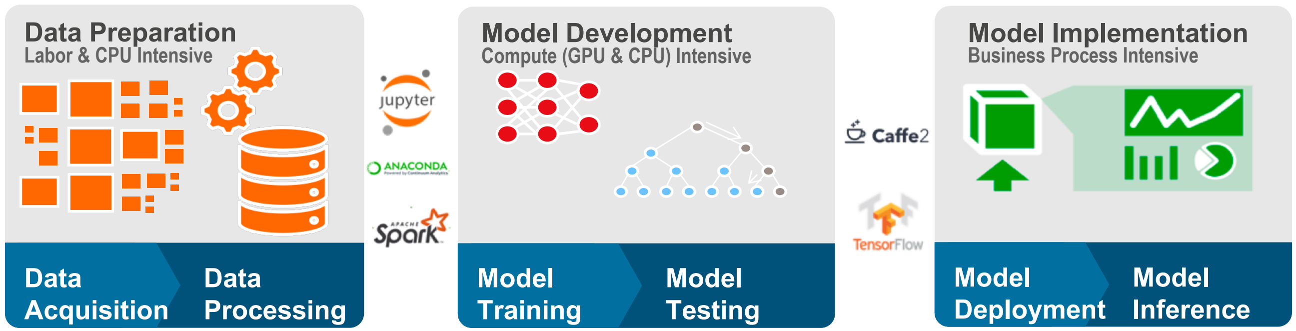


Object Detection

Researching Precursors of Tropical Cyclones Using Deep Learning



Observation #1 - AI Workflows are Evolving Rapidly with Increasing Technology Diversity



- **Compute Requirements (By Stage and Task):**
  - Data preparation
  - Machine Learning Training
  - Deep Learning Training
  - Inference

- **Data Requirements (Size, Type, Performance):**
  - Large volume data stores
  - High-bandwidth
  - Fast IOPs
  - Direct connectivity to external storage
  - HDF5 ~ Key-Value Stores

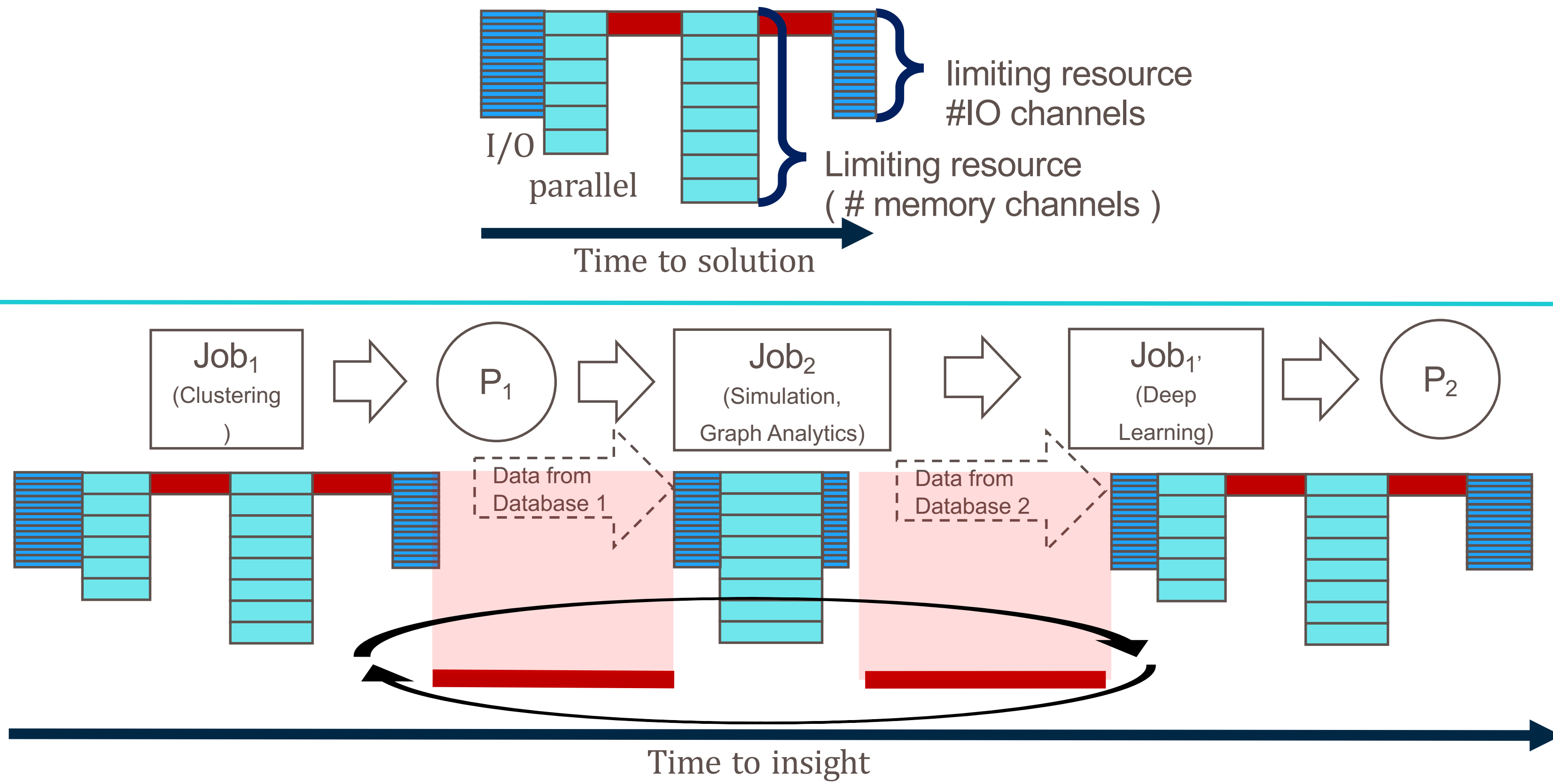
- **Performance Requirements (By Stage and Task):**
  - Scaling
  - Throughput

- **User-productivity**
  - Containers
  - Collaborative notebooks
  - Open-source frameworks
  - High productivity interpreted languages

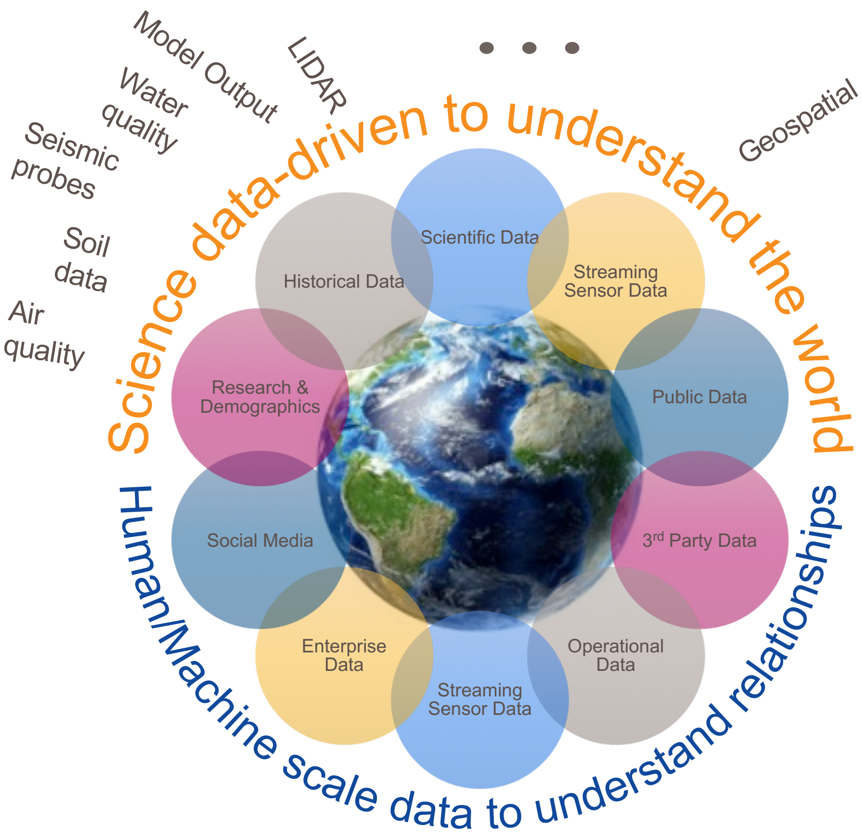
- **Training and Inference Processor Landscape**
  - Intel, AMD, ARM, NVIDIA
  - Google TPU v2
  - Graphcore
  - Habana
  - 30+ startups….

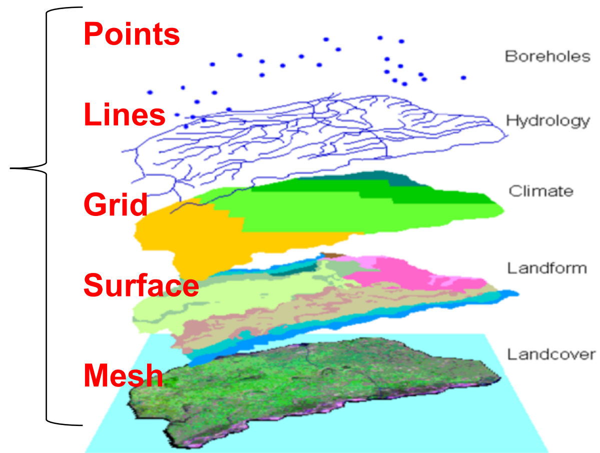
- **Storage Technology Landscape**
  - Tape
  - HDD
  - SSD-SATA/SAS
  - SSD-NVMe
  - Flash
  - On-node vs. off-node – Datawarp vs. NVMe-over-Fabrics

Observation #2 - Time to Insight: Performance Emphasis Will Move From Optimized "Codes" to "Workflows"



Observation #3 - Development of Domain and Use-Case Specific Multi-Models





	AI Today	Multi-Model Future
Model	CNN, RNN, LSTM, GAN etc.	Domain-specific
Baseline	Humans, Other ML algorithms	Theory, Science
Use Case	Speech, Image interpretation	Computational Steering Proxy models
Figure of Merit	Time-to-accuracy, Model-size	+ Interpretability, Feasibility
Model Design	Transfer/Incremental Learning	Hyper-parameter optimization
Model Testing	A/B Testing on a cadence	Statistical rigor

Observation #4 - AI for Science: Increasing Mingling of Simulation and AI Methodologies

Model Complexity

Domain-Specific Models
- Fluid Dynamics
- Schrodinger Equation
- Full-Wave Inversion
- Integration with user-facilities

Combinatorial Patterns
- Search for meta-stable states
- Search for particles
- Search for N-way correlations
- Search in a high-dimensional feature space

Emerging Use-cases

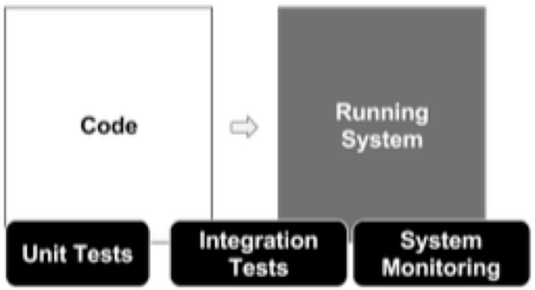
Model Repurposing
- Deep Learning for X
- Feature engineering
- Predictive modeling
- Hypothesis creation with interpretation

Workflows and Automation
- Smarter initialization for simulations
- Computational steering
- Hybrid models of theory and AI

Computational Complexity

Observation #5 - Convergence Does Not Mean Sameness

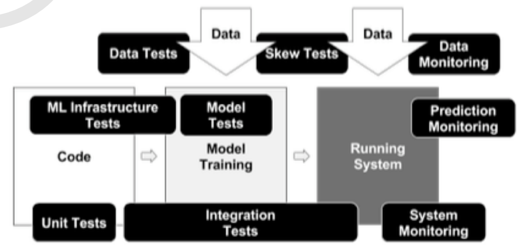
SIMULATION



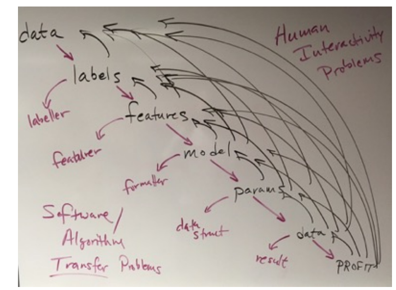
Traditional System Testing and Monitoring

- Simulation codes are universal, long lasting and evolve more predictably
- Focused on making sure the code mirrors the physics
- Architecture specific optimizations are long lasting
- Tightly integrated processor-memory-interconnect & network storage

ARTIFICIAL INTELLIGENCE



ML-Based System Testing and Monitoring



- Workflows are bespoke: Use case and data-specific
- Code can be disposable
- ML-based system behaviour is not easily specified in advance - depends on dynamic qualities of the data and on various model configuration choices
- Maximize data movement-- scan/sort/stream all the data all the time

Credit: “The ML Test Score: A Rubric for ML Production Readiness and Technical Debt Reduction” Eric Breck et al, Google, Inc.

Converged Simulation and AI Workload Trends

CRAY

1. AI Workflows are Evolving Rapidly with Increasing Technology Diversity
2. Time to Insight: Performance emphasis will move from optimized “Codes” to “Workflows”
3. Development of domain and use-case specific multi-models
4. AI for Science: Increasing Mingling of Simulation and AI Methodologies
5. Convergence does not mean sameness

Heterogeneous environments that will not settle for some time

Change in how we think about performance and efficiency– both machine and human

Growth in computational demand and complexity

Simulation and AI methodologies will be inseparable

Simulation and AI - the unique characteristics of each need to be addressed

Goal is harmony where both simulation and AI optimally co-exist within the same environment

Thank You.

Per Nyberg
VP, Artificial Intelligence

CRAY™