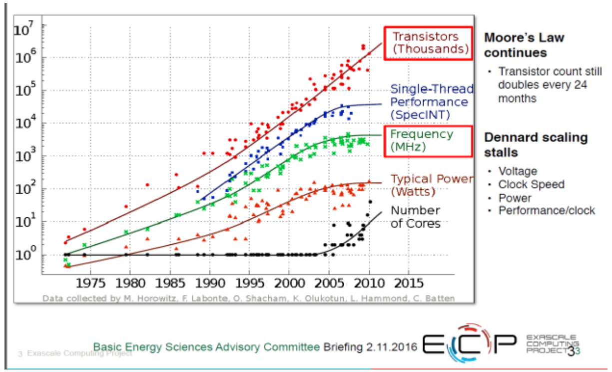
HPC EXASCALE + AI

Introducing a new era for innovation

FACTORS DRIVING HISTORIC CHANGES IN HPC



End of Dennard Scaling places a cap on single threaded performance

Increasing application performance will require fine grain parallel code with significant computational intensity

AI and Data Science emerging as important new components of scientific discovery

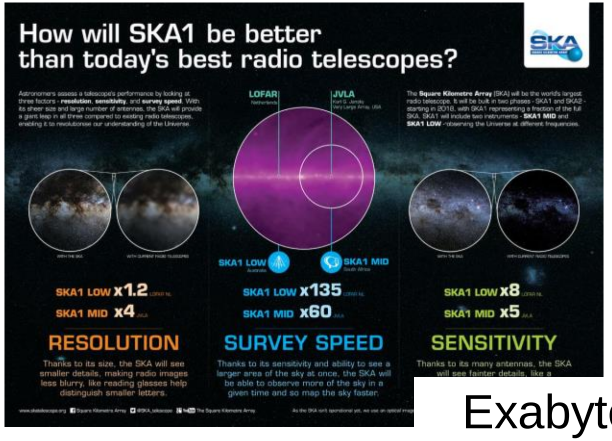
Dramatic improvements in accuracy, completeness and response time yield increased insight from huge volumes of data

Cloud based usage models, in-situ execution and visualization emerging as new workflows critical to the science process and productivity

Tight coupling of interactive simulation, visualization, data analysis/AI

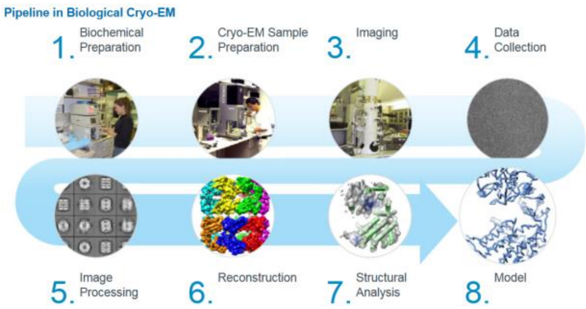
THE EX FACTOR IN THE EXASCALE ERA

Multiple EXperiments Coming or Upgrading In the Next 10 Years

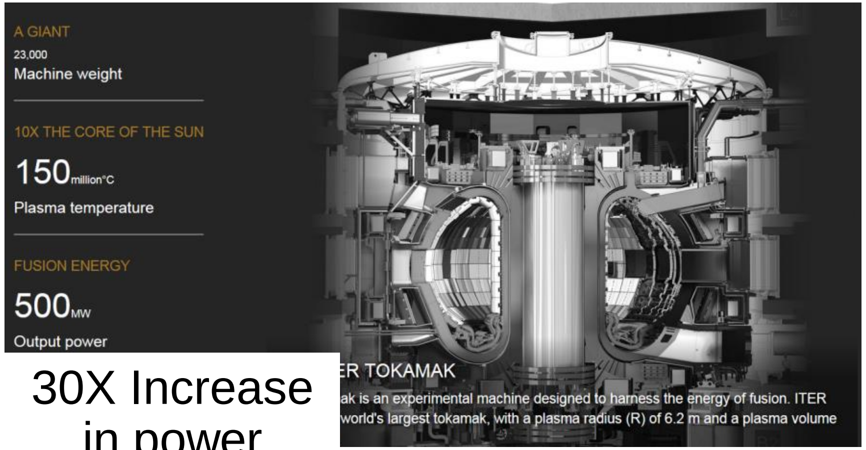


Exabyte/Day

15 TB/Day

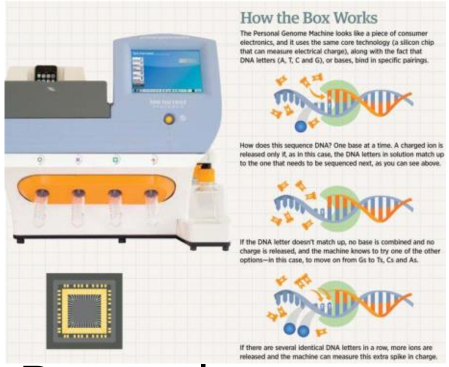


Cryo EM



30X Increase in power

10X Increase in Data Volume



Personal Genomics

NVIDIA

THE POTENTIAL OF EXASCALE HPC + AI

HPC	AI
+40 years of Algorithms based on first principles theory Proven statistical models for accurate results in multiple science domains	New methods to improve predictive accuracy, insight into new phenomena and response time with previously unmanageable data sets

Commercially viable fusion energy

Understanding the Origins of the Universe

Clinically Viable Precision Medicine

Improve/validate the Standard Model of Physics

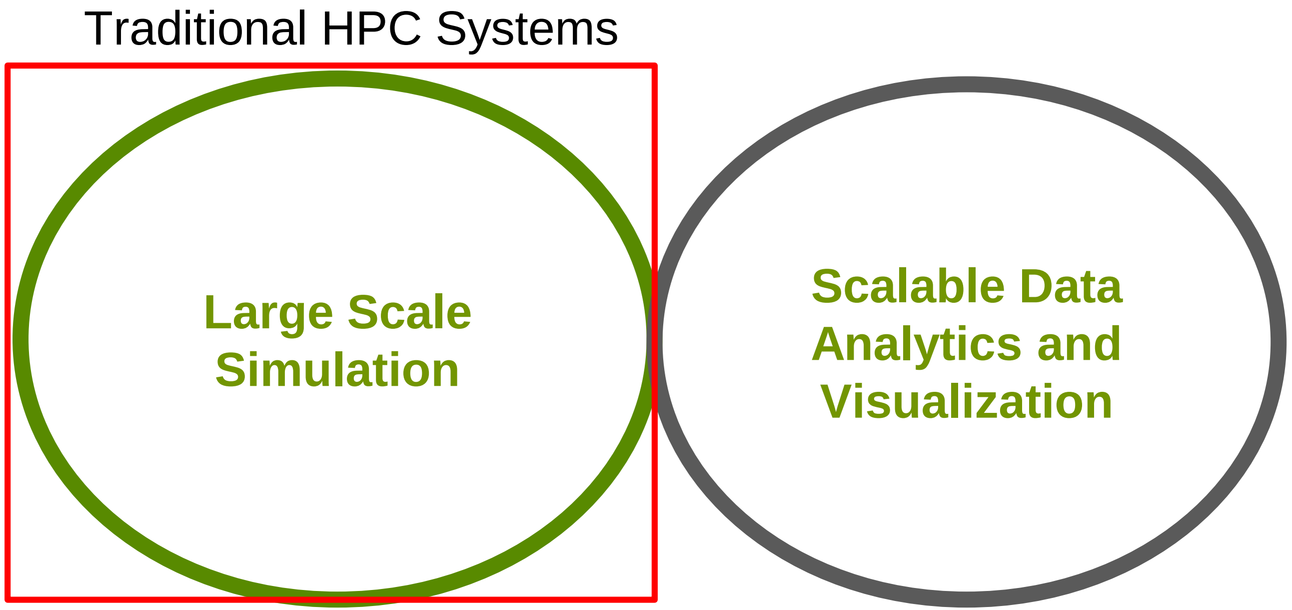
Climate/Weather forecasts with ultra high fidelity

*

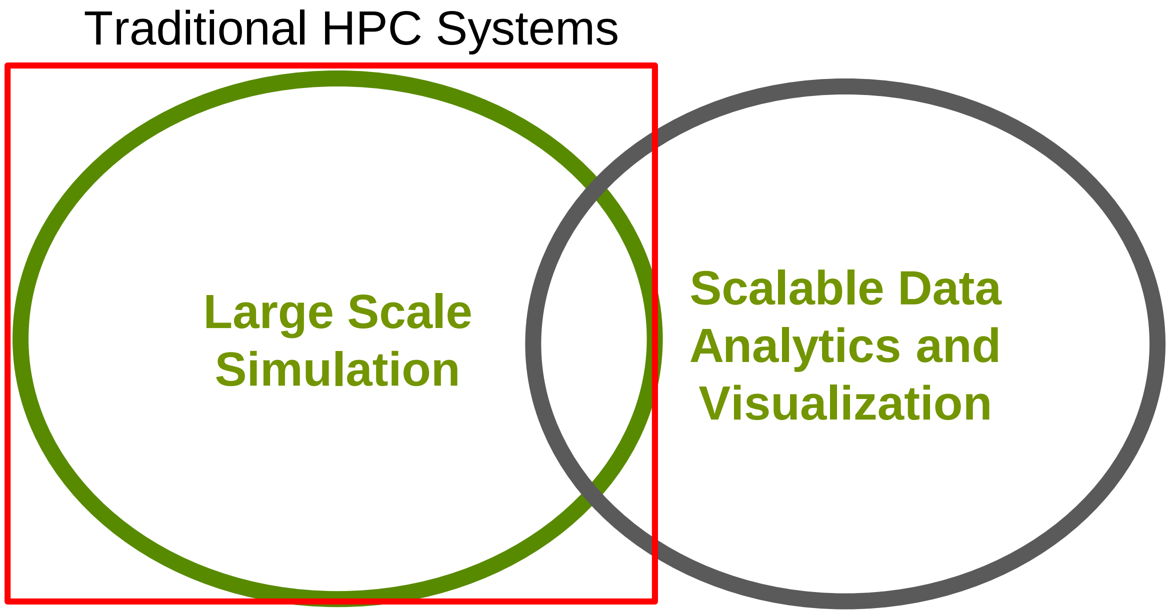
*

*

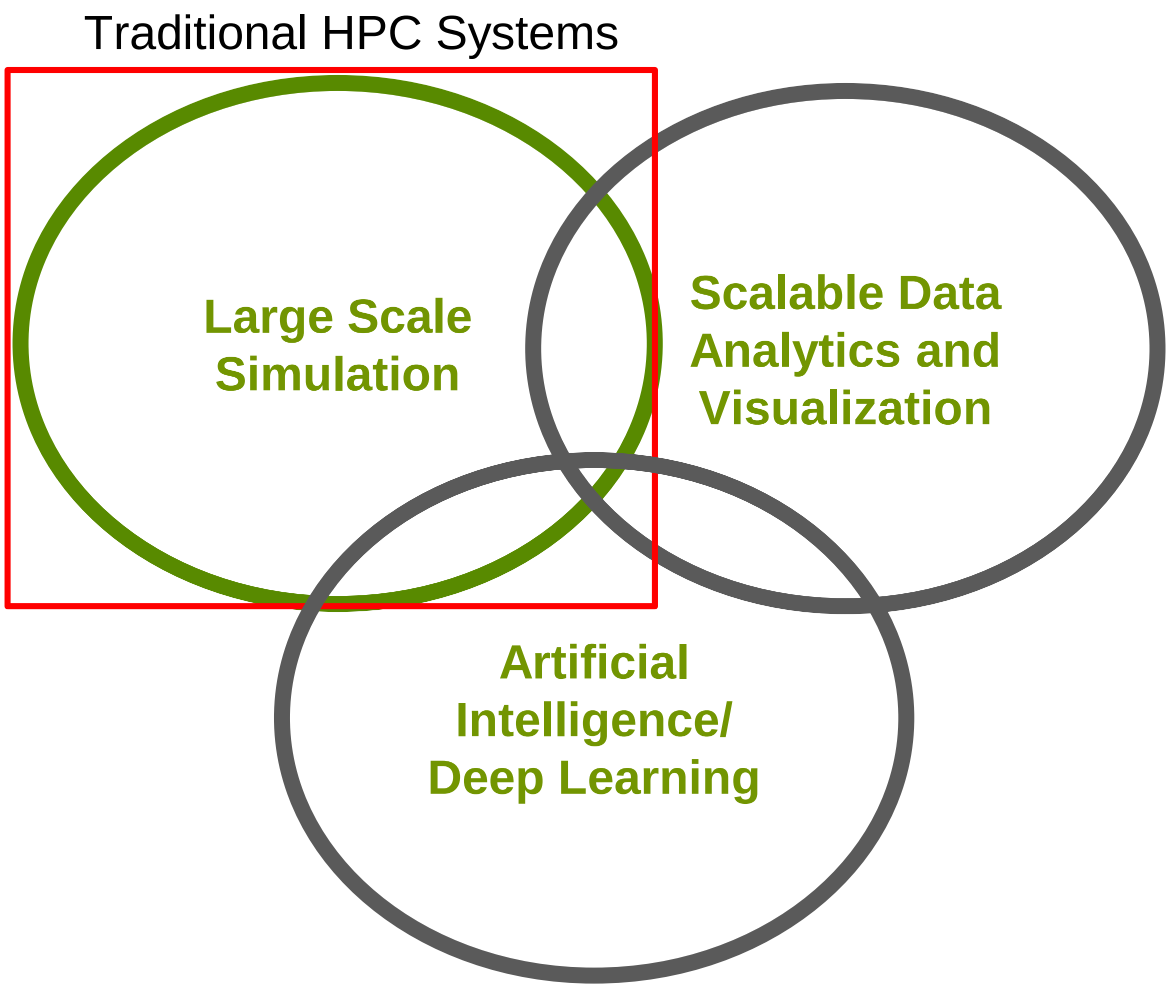
TRADITIONAL HPC METHOD



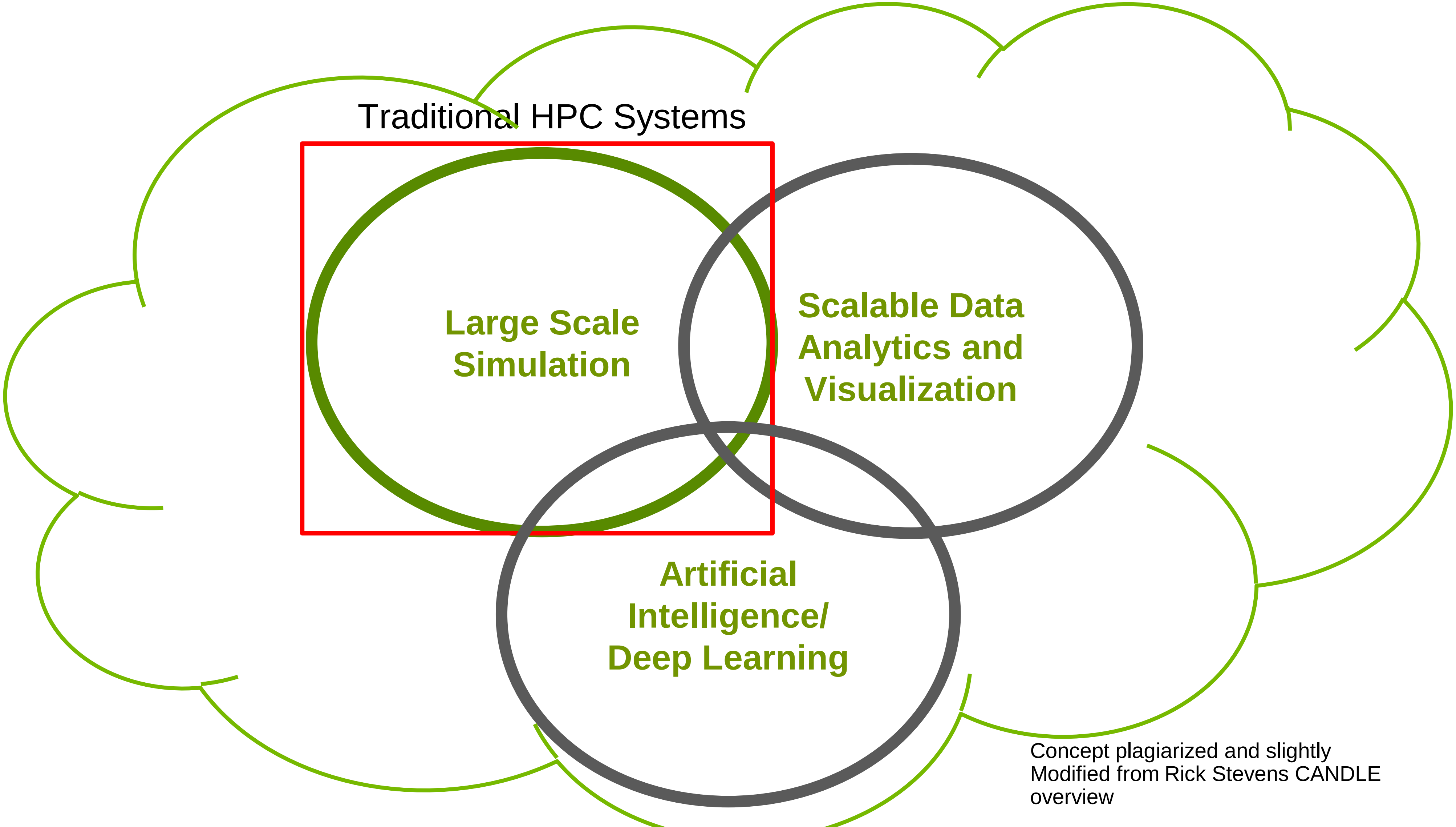
TRADITIONAL HPC METHOD



TRADITIONAL HPC METHOD



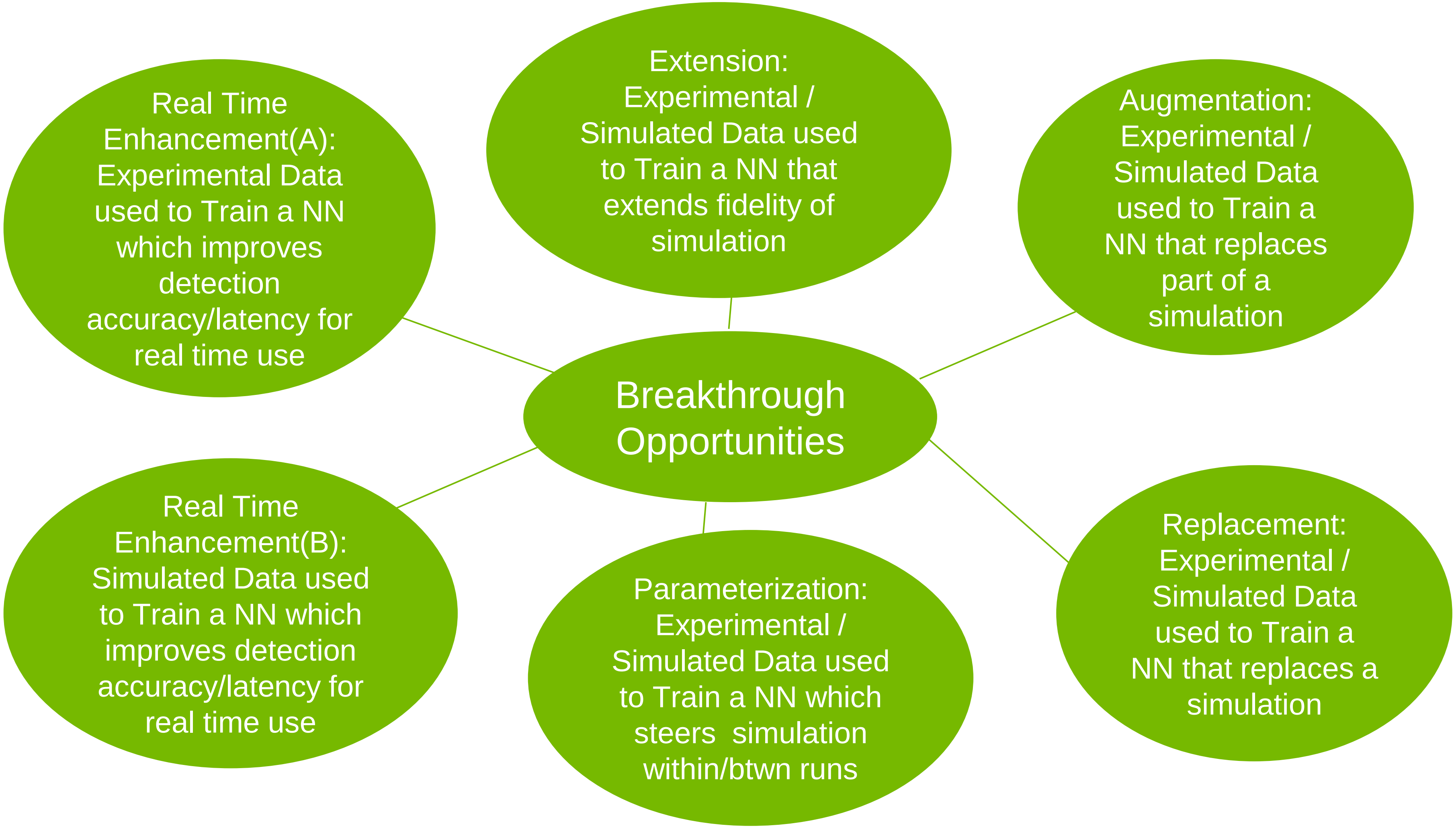
TRADITIONAL HPC METHOD



Concept plagiarized and slightly Modified from Rick Stevens CANDLE overview

TAXONOMY

Examples of HPC + AI Convergence



Despite the latest development in computational power, there is still a large gap in linking relativistic theoretical models to observations.
Max Plank Institute

MULTI-MESSENGER ASTROPHYSICS

Background

The aLIGO (Advanced Laser Interferometer Gravitational Wave Observatory) experiment successfully discovered signals proving Einstein's theory of General Relativity and the existence of cosmic Gravitational Waves. While this discovery was by itself extraordinary it is seen to be highly desirable to combine multiple observational data sources to obtain a richer understanding of the phenomena.

Challenge

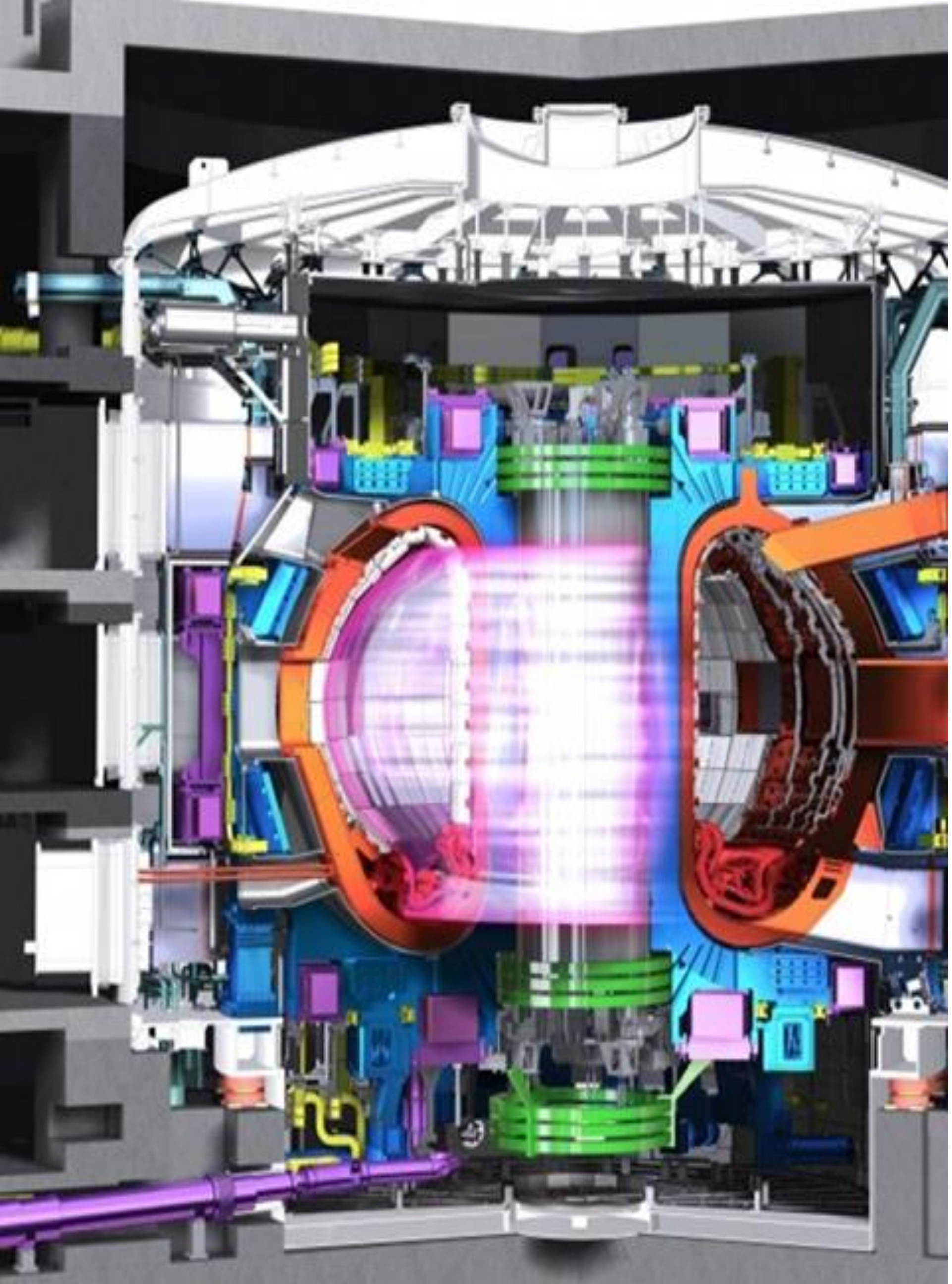
The initial a LIGO discoveries were successfully completed using classic data analytics. The processing pipeline used hundreds of CPU's where the bulk of the detection processing was done offline. Here the latency is far outside the range needed to activate resources, such as the Large Synaptic Space survey Telescope (LSST) which observe phenomena in the electromagnetic spectrum in time to "see" what aLIGO can "hear".

Solution

A DNN was developed and trained using a data set derived from the CACTUS simulation using the Enistein Toolkit. The DNN was shown to produce better accuracy with latencies 1000x better than the original CPU based waveform detection.

Impact

Faster and more accurate detection of gravitational waves with the potential to steer other observational data sources.



Predicting Disruptions in Fusion Reactor using DL

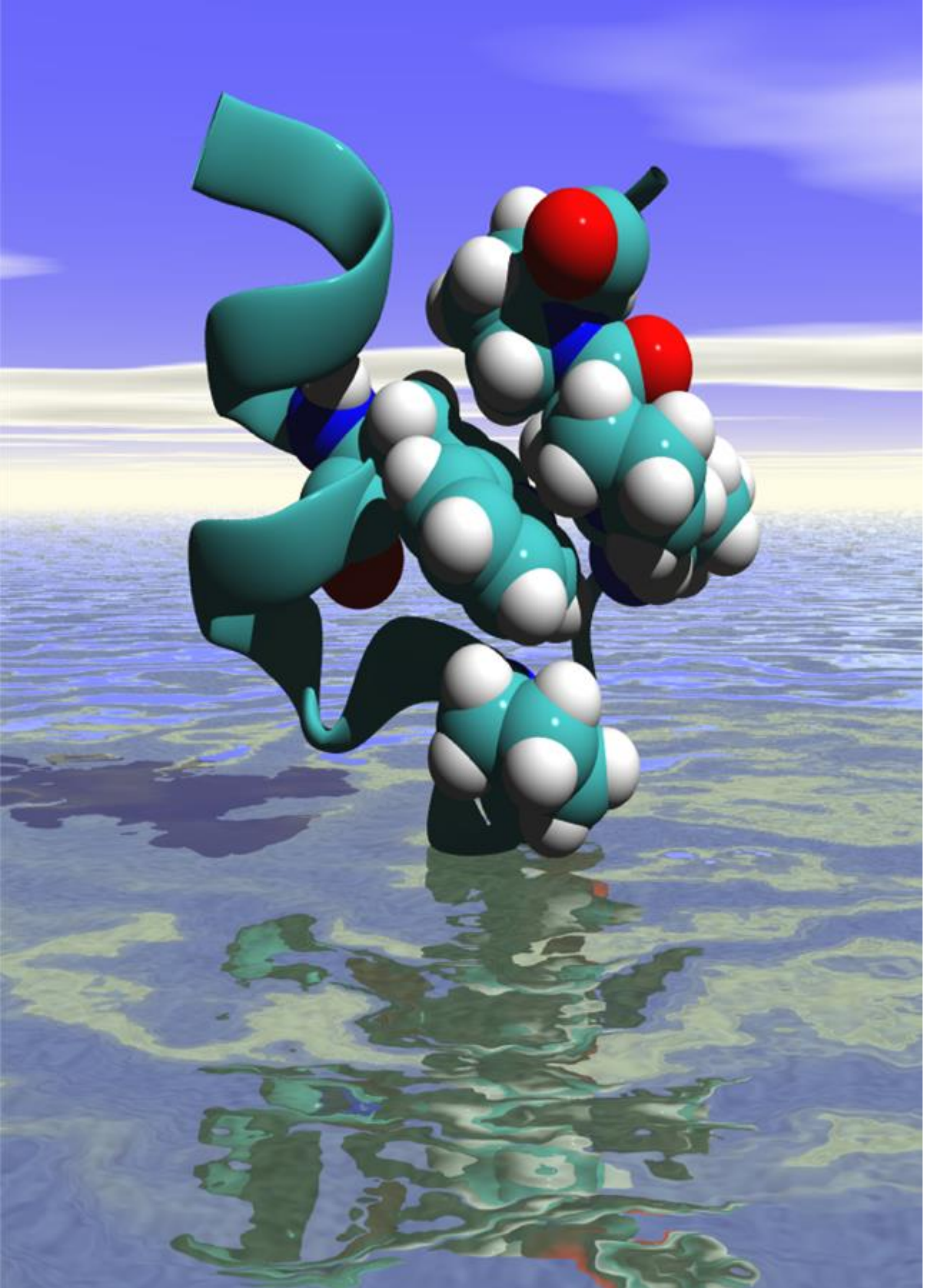
Background
Grand challenge of fusion energy offers mankind changing opportunity to provide clean, safe energy for millions of years. ITER is a $25B international investment in a fusion reactor.

Challenge
Fusion is highly sensitive, any disruption to conditions can cause reaction to stop suddenly. Challenge is to predict when a disruption will occur to prevent damage to ITER and to steer the reaction to continue to produce power. Traditional simulation and ML approaches don’t deliver accurate enough results.

Solution
DL network called FRNN using Theano exceeds today's best accuracy results. It scales to 200 Tesla K20s, and with more GPUs, can deliver higher accuracy. Goal is to reach 95% accuracy.

Impact
Vision is to operate ITER with FRNN, operating and steering experiments in real-time to minimize damage and down-time.

PRINCETON UNIVERSITY



AI Quantum Breakthrough

Background

Developing a new drug costs $2.5B and takes 10-15 years. Quantum chemistry (QC) simulations are important to accurately screen millions of potential drugs to a few most promising drug candidates.

Challenge

QC simulation is computationally expensive so researchers use approximations, compromising on accuracy. To screen 10M drug candidates, it takes 5 years to compute on CPUs.

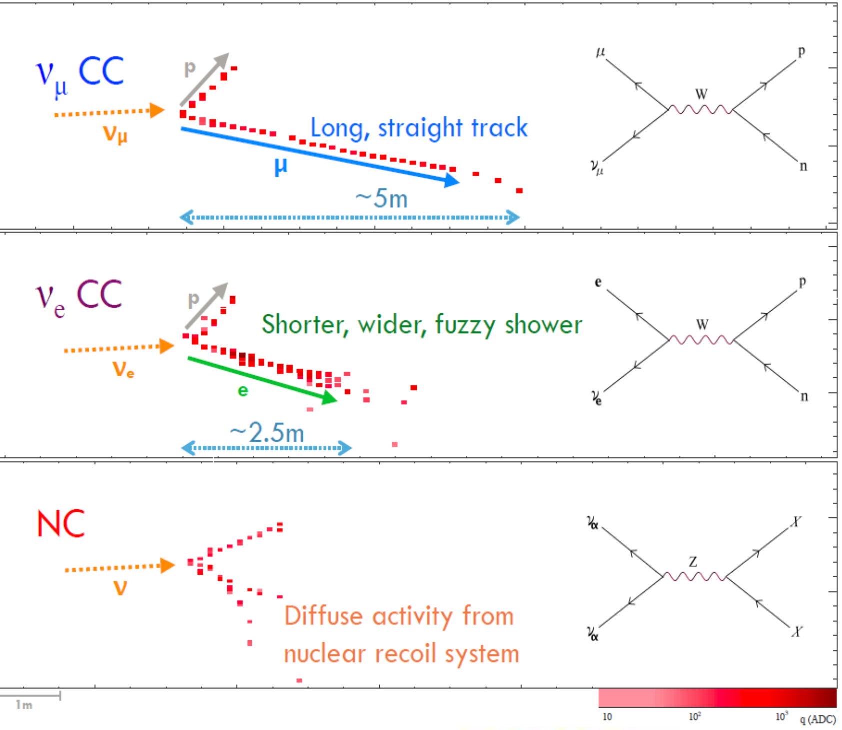
Solution

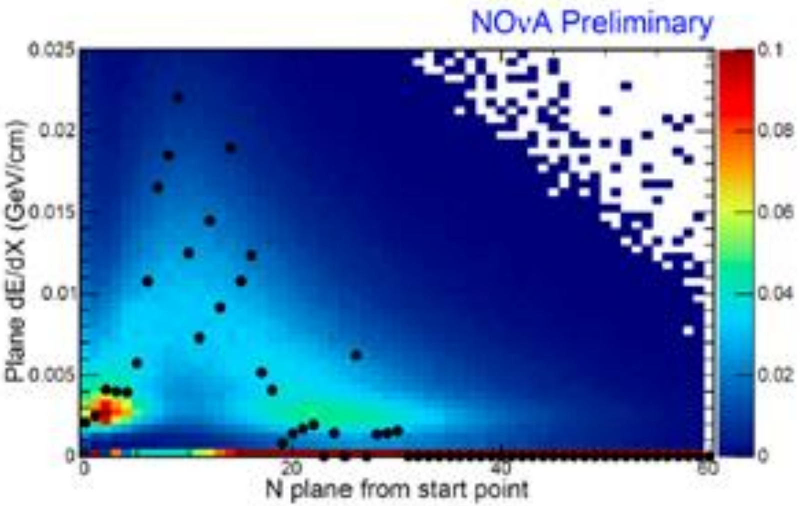
Researchers at the University of Florida and the University of North Carolina leveraged GPU deep learning to develop ANAKIN-ME, to reproduce molecular energy surfaces with super speed (microseconds versus several minutes), extremely high (DFT) accuracy, and at 1-10/millionths of the cost of current computational methods.

Impact

Faster, more accurate screening at far lower cost

UF UNIVERSITY of FLORIDA





FINDING THE "GHOST PARTICLE" WITH AI

Background

The NoVA experiment managed by Fermi lab comprises 200 scientists at 40 institutions in 7 countries. The goal is to track neutrino's, which are often referred to as the "Ghost Particle", and detect oscillation which is used to better understand how this super abundant, and elusive particle interacts with matter.

Challenge

The experiment is built underground and is comprised of a main injector beam and two large detector apparatus located 50 miles apart. The near detector is 215 Tons and the Far detector is 15,000 Tons. The experiment can be thought of as a 30 Mn pound detector that takes 2 Mn pictures per second.

The detectability of the current experiment is proportional to the size of the detectors, so increasing the "visibility" is complex and costly.

Solution

A DNN was developed and trained using a data set derived from multiple HPC simulations including GENIE and GEANT using 2 K40 GPU's. the CVN was basd on convolutional neural networks used for image processing

Impact

The result was an overall improvement of 33%, where the optimized CVN signal-detection-optimized efficiency of 49% is a significant gain over the efficiency of 35% quoted in prior art. This would net to a 10Mn pound increase the physical detector

Forecasting Fog at Zurich Airport

WORK IN PROGRESS

Background
Unexpected fog can cause an airport to cancel or delay flights, sometimes having global effects in flight planning.

Challenge
While the weather forecasting model at MeteoSwiss work at a 2km x 2km resolution, runways at Zurich airport is less than 2km. So human forecasters sift through huge simulated data with 40 parameters, like wind, pressure, temperature, to predict visibility at the airport.

Solution
MeteoSwiss is investigating the use of deep learning to forecast type of fog and visibility at sub-km scale at Zurich airport.

Impact

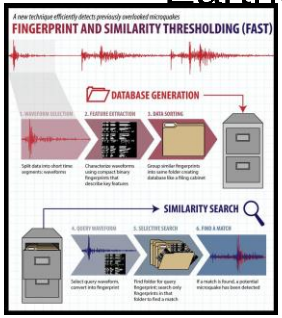
MeteoSwiss

Earthquake Prediction

WORK IN PROGRESS

Multiple Examples of AI for earthquake prediction are underway

Shaazam for Earthquakes



SCIENTIFIC AMERICAN.

COMPUTING

Can Artificial Intelligence Predict Earthquakes?

The ability to forecast temblors would be a tectonic shift in seismology. But is it a pipe dream? A seismologist is conducting machine-learning experiments to find out

PRINCETON UNIVERSITY

SCEC an NSF+USGS center

Stanford University

THERE WILL BE NO REASON TO ASK

WHY DOES HPC + AI MATTER?

NVIDIA

COMBINING THE STRENGTHS OF HPC AND AI

HPC	AI
+40 years of Algorithms based on first principles theory Proven statistical models for accurate results in multiple science domains	New methods to improve predictive accuracy, insight into new phenomena and response time with previously unnavigable data sets
Develop training data sets using first principal models Apply Bayesian regression methods to expedite/ensure training accuracy Incorporate AI models in semi-empirical style applications to improve throughput Validate new findings from AI	Train inference models to improve accuracy and comprehend more of the physical parameter space Implement inference models with real time interactivity Analyze data sets that are simply intractable with classic statistical models Control and manage complex scientific experiments or apparatus