A SYSTEMS ARCHITECTURE FOR IMBEDDING PREDICTIVE CLIMATE SIMULATION IN THE DIGITAL ECONOMY

Mark Seager
Intel Fellow, CTO for the HPC Ecosystem
RackScale Design Group

Presented to iCAS14, Annecy France
September 11, 2017

Scientific simulation at scale is qualitatively different. We do not yet understand the full implications of this technological development

By being on the leading edge at the 10 teraFLOP/s level with a toehold into the 100 teraFLOP/s → petaFLOP/s computing era, we are fundamentally changing the nature and scope of the scientific method.

- Edsger Dijkstra: “A quantitative difference is also a qualitative difference, if the quantitative difference is greater than an order of magnitude.

- A quantitative example in transportation
  - 1 Mi/Hr is the speed of a baby crawling
  - 10 Mi/Hr is the speed of a top marathon runner
  - 100 Mi/Hr is the speed of a fast automobile
  - 1,000 Mi/Hr is the speed of a fast jet
- Qualitative ramifications of this transportation example
  - Driving allows people to go to places they could not reach on foot.
  - Flying allows people to go to places they could not reach in time.

From LLNL Jan 2004

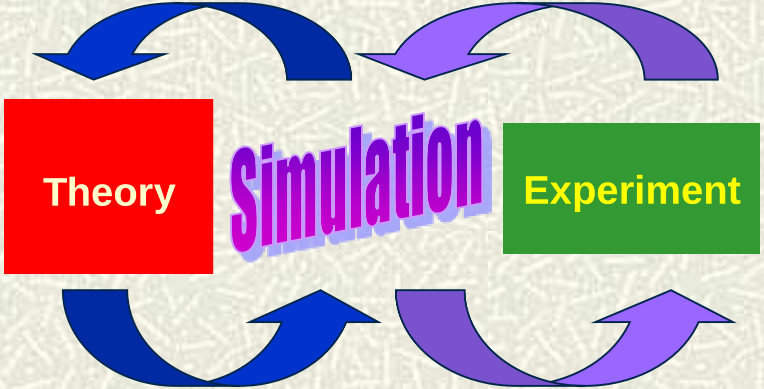
Simulation has become the critical integrating element between theory and experiment

Predictive simulation ENABLES

- **Detailed predictive assessment of complex models for overarching physical problems**
- **Design of experiments**
- **Impact assessment of policy choices**
- **Elimination of costly physical prototypes**

Predictive simulation REQUIRES

- Verification and validation of complex models (**experiment**)
- Development of science based models (**theory**)
- **Databases** of physical properties and catalogues of scientific data
- **Petascale simulation environments**



Revolution in the making: BlueGene/L at LLNL

When was the last major change in the Scientific Method? In 1609, 401* YAG

- The telescope was one of the central instruments of what has been called the Scientific Revolution of the seventeenth century.
- It revealed hitherto unsuspected phenomena in the heavens and had a profound influence on the controversy between followers of the traditional geocentric astronomy and cosmology and those who favored the heliocentric system of Copernicus.
- **It was the first extension of one of man's senses**, and demonstrated that ordinary observers could see things that the great Aristotle had not dreamed of.
- **It therefore helped shift authority in the observation of nature from men to instruments. In short, it was the prototype of modern scientific instruments.**

http://galileo.rice.edu/sci/instruments/telescope.html

* Jan 9, 2007

Industrial Revolution : A period of unprecedented technological and economic development

You Are He

4TH INDUSTRIAL REVOLUTION

COAL & STEAM	ELECTRIFICATION	COMPUTE & COMMUNICATIONS
1760's	1860's	1960's

innovation ACROSS ALL INDUSTRIES

“DIGITAL FUSION”

Blending of Traditional and Digital Business Models


Smart Agriculture


3D Printing in Healthcare


Wearables (Industrial/Lifestyle)


Robotic Surgery


Autonomous Cars


Human Augmentation

BY 2020...

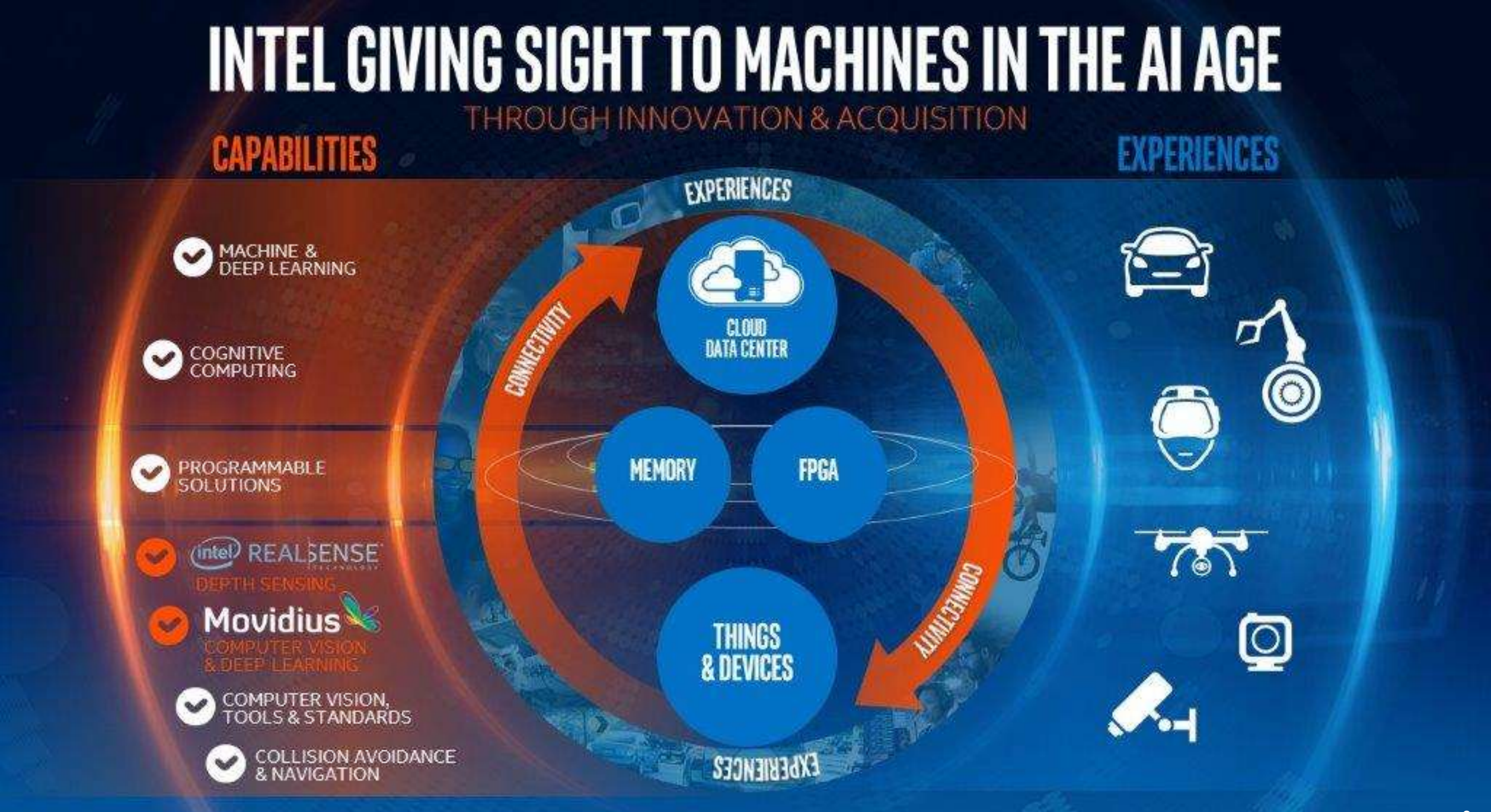
AVG INTERNET USER	~1.5 GB * TRAFFIC PER DAY
SMART HOSPITAL	3,000 GB * PER DAY
AUTONOMOUS AUTOMOBILE	4,000 GB * PER CAR... EACH
AIRPLANE DATA	40,000 GB * PER DAY
SMART FACTORY	1,000,000 GB * PER DAY

A FLOOD OF DATA IS COMING

Data Center traffic is DOUBLING every 18 months**

** Source: Intel Investor Day Feb 9, 2017

* Sources:
http://www.cisco.com/c/en/us/solutions/service-provider/vni-network-traffic-forecast/infographic.html
http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html
https://datafloq.com/read/self-driving-cars-create-2-petabytes-data-annually/172



INTEL GIVING SIGHT TO MACHINES IN THE AI AGE
THROUGH INNOVATION & ACQUISITION
CAPABILITIES
EXPERIENCES
MACHINE & DEEP LEARNING
COGNITIVE COMPUTING
PROGRAMMABLE SOLUTIONS
intel REALSENSE
DEPTH SENSING
Movidius
COMPUTER VISION & DEEP LEARNING
COMPUTER VISION, TOOLS & STANDARDS
COLLISION AVOIDANCE & NAVIGATION
EXPERIENCES
CONNECTIVITY
CLOUD DATA CENTER
MEMORY
FPGA
THINGS & DEVICES
CONNECTIVITY
EXPERIENCES

Other emerging predictive simulation usage models are driven by coupled Big Data, Computational Science and AI

Government & Research

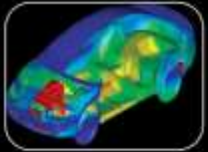
"My goal is simple. It is complete understanding of the universe, why it is as it is, and why it exists at all"

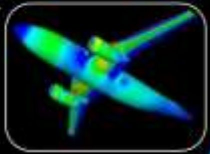
Stephen Hawking

Basic Science

Commerce & Industry

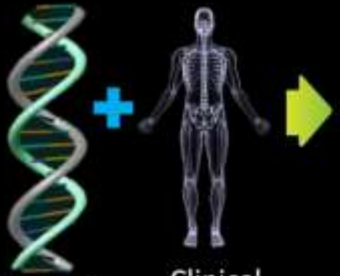
Better Products
Digital Twin
Faster Time to Market
Reduced R&D
New Business Models
Data Services





Business Transformation

New Users & New Uses



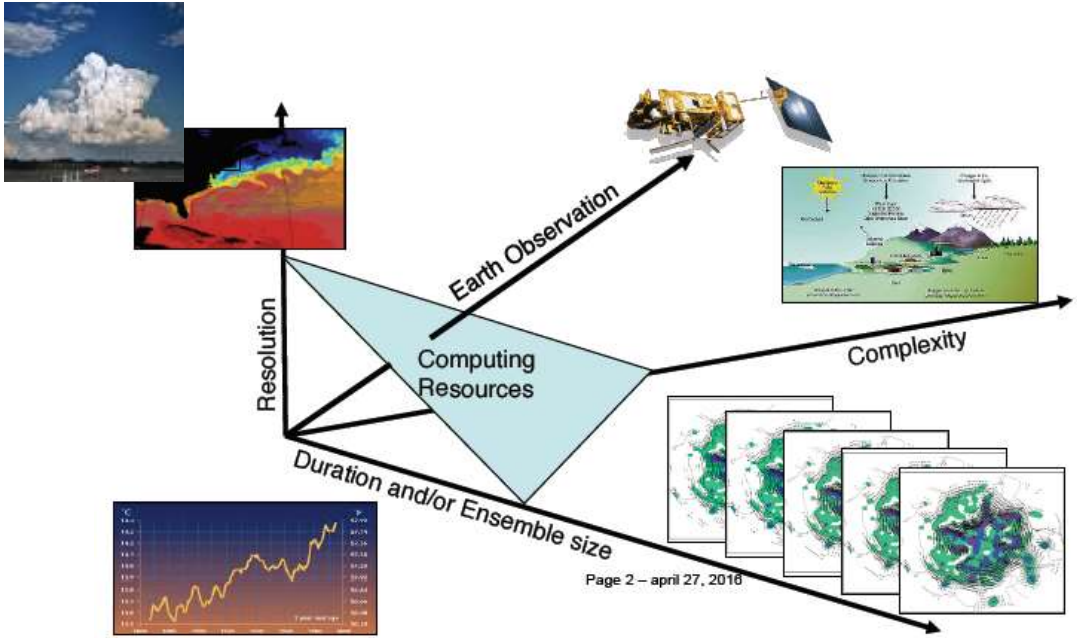
From Diagnosis to personalized treatments quickly

Data-Driven Discovery

Transform data into useful knowledge

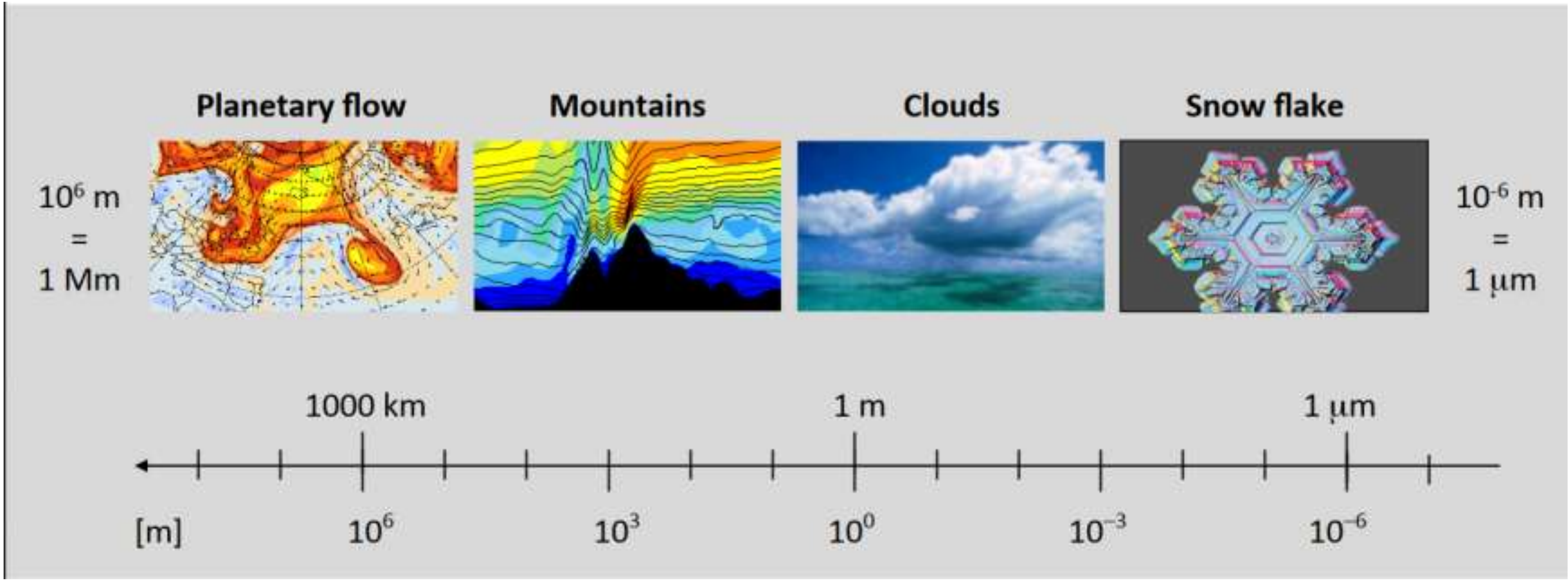
Source: IDC: Worldwide Technical Computing Server 2013–2017 Forecast; Other brands, names, and images are the property of their respective owners.

Climate models are complex in a number of dimensions



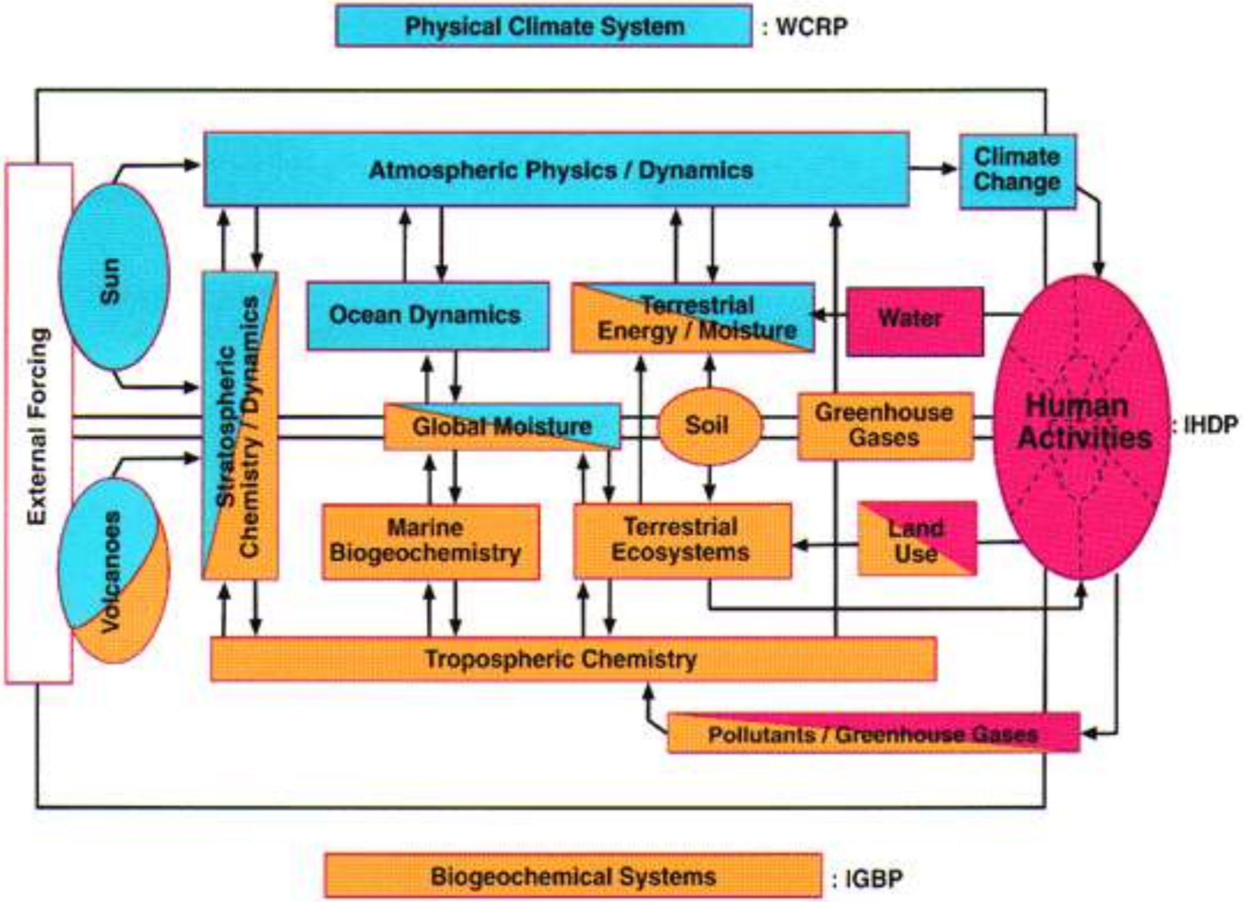
Climate Models have to accurately simulate multiple-physical phenomena at wide range of scales

Multi-physics, Multi-Scale



O. Furher. http://on-demand.gputechconf.com/gtc/2013/presentations/S3417-GPU-Accelerated-Operational-Weather-Forecasting.pdf

Climate Application is a large set of tightly coupled Science Packages



Climate Model Characteristics

Very large, complex codes and workflows

- >1m LOC common
- Runtime measured in months for “Simulated Years Per Day”
- Multiple components coupled in parallel
  - Amdahl's Law & load balancing critical

Data movement critical limiting factor

- Memory Bandwidth limited
- High speed, Low Latency MPI communications
- Moderate IO, but critical for analysis and visualization

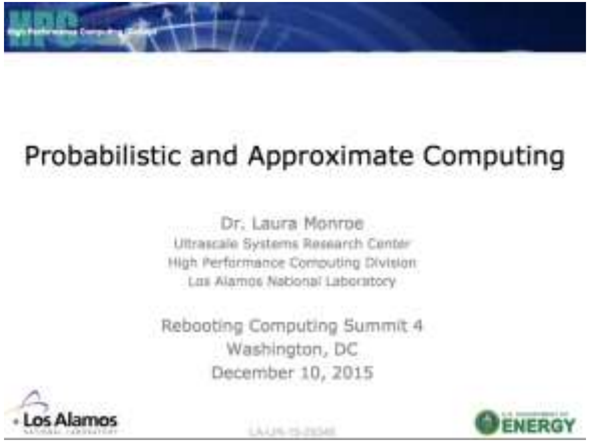
Knowledge Gaps in Physics-based models

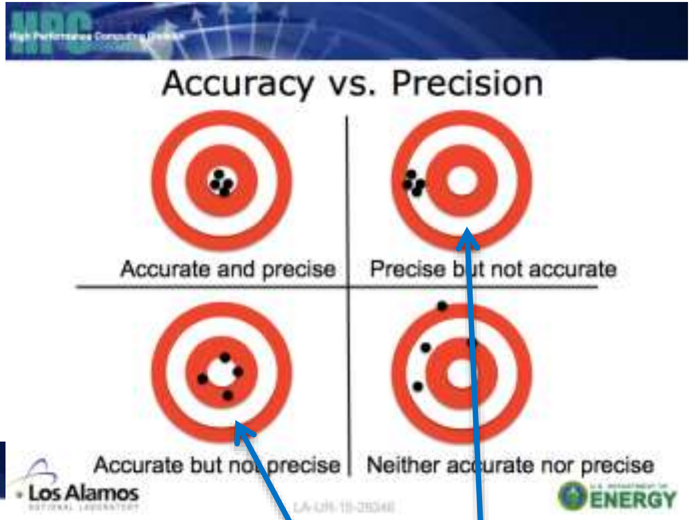
Inadequate understanding of key climatic processes

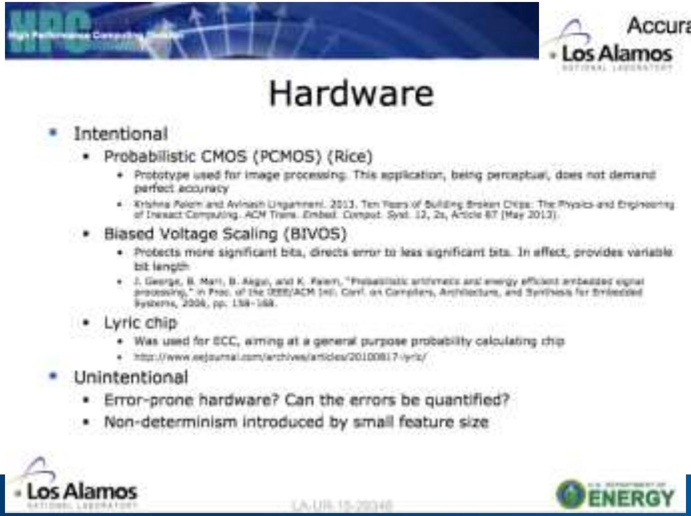
- Cloud physics, precipitation extremes, etc.

Parameterized approximations for handling knowledge gaps lead to overly complex models

- Poor predictive performance
- High uncertainty
- Difficult to explain/interpret







Where we are now.

Where next gen models need to be in < 10 years time.

Why Do Black-box ML Methods Fail?

Scientific problems are often under-constrained

- Small number of samples, large number of variables
- High-quality climate observations only available for the recent past (40 to 100 years)
- Standard methods for assessing and ensuring generalizability of machine learning models break down
- Huge number of samples is critical to success of methods such as deep learning

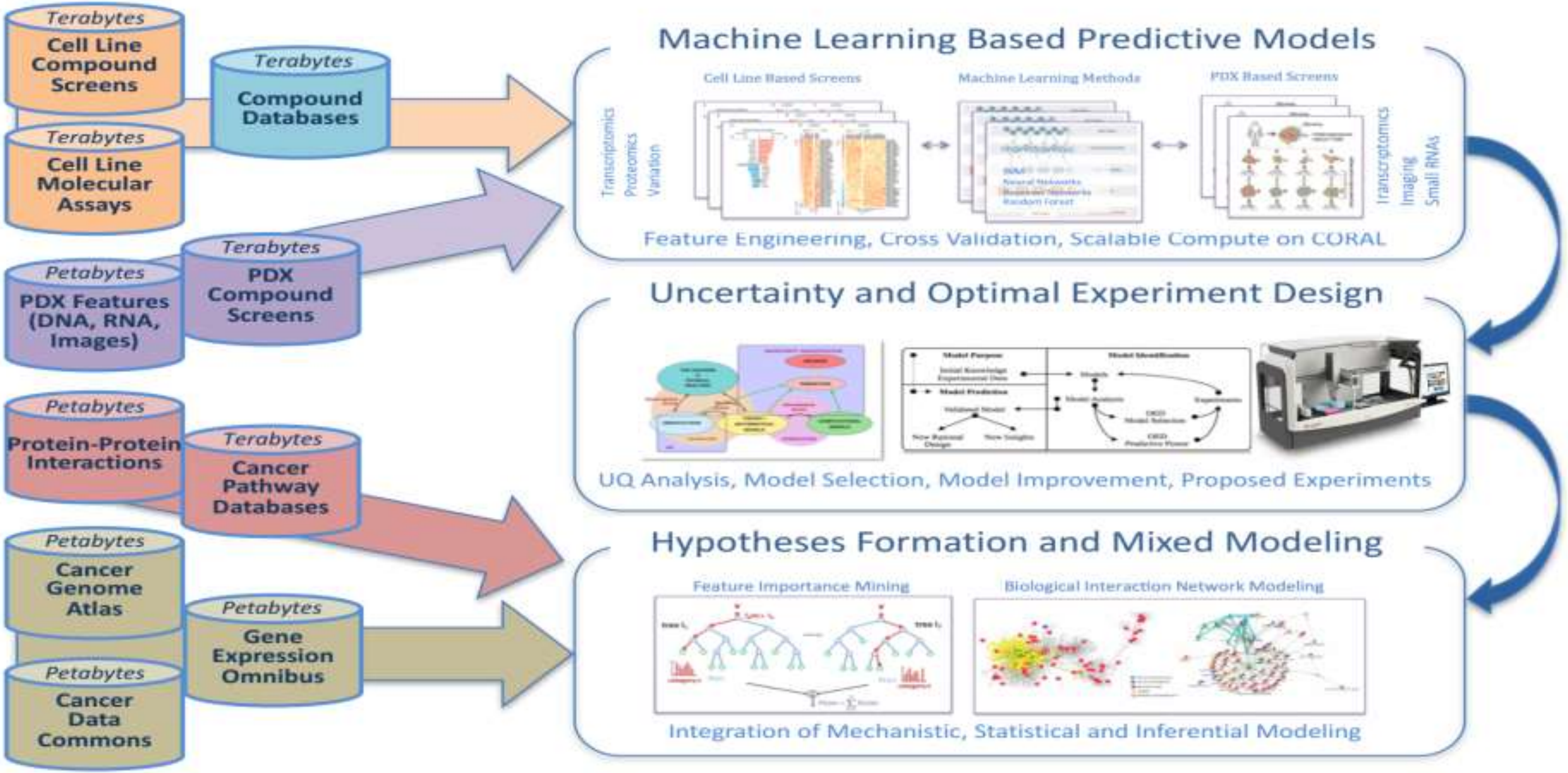
Advances in data science methods are needed to accelerate scientific discovery

SCIENTIFIC AMERICAN.

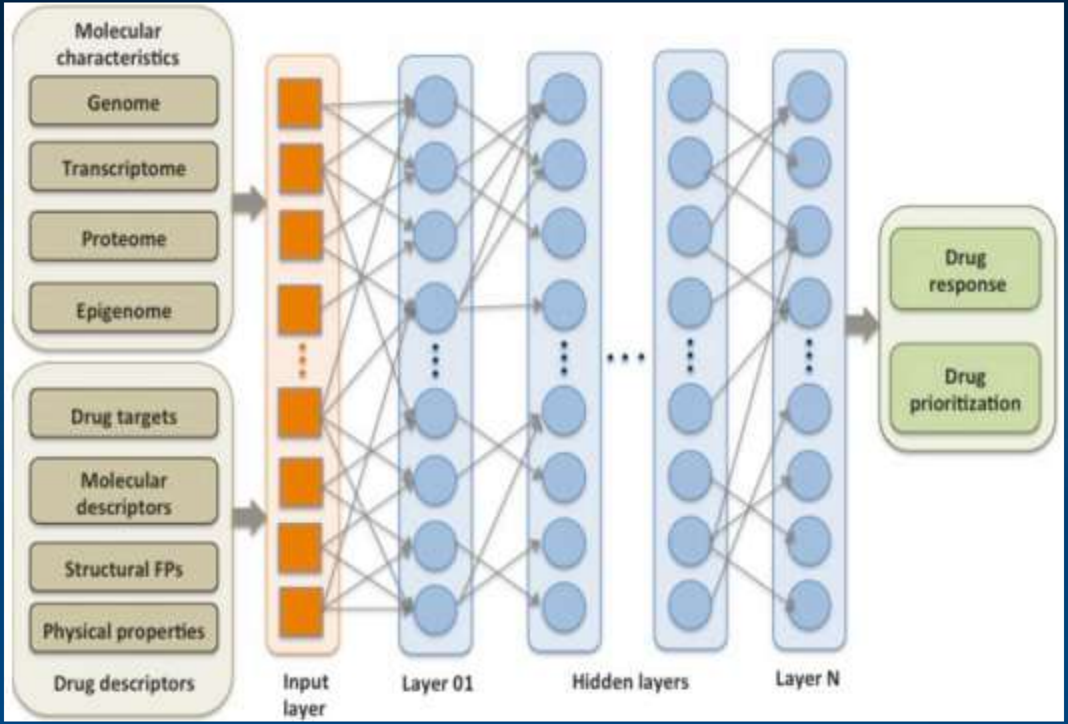
Big Data Needs a Big Theory to Go with It

- Geoffrey West 2013

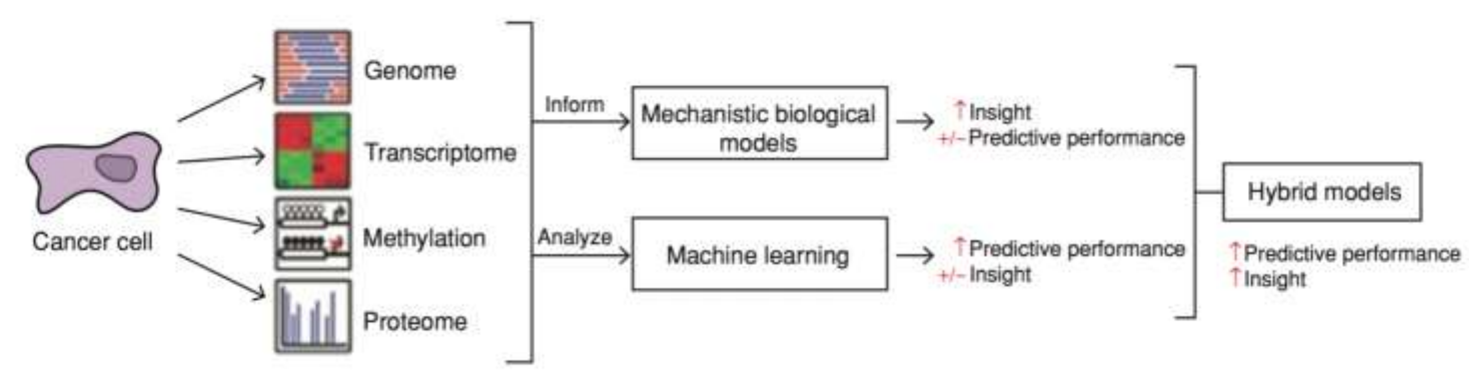
Pilot 1: Predictive Models for Pre-Clinical Screening



HOW DEEP LEARNING CAN PLAY A ROLE



HYBRID MODELS ARE NEEDED IN CANCER RESEARCH



Applications of Machine Learning in Climate

1. Evaluation and refinement of models

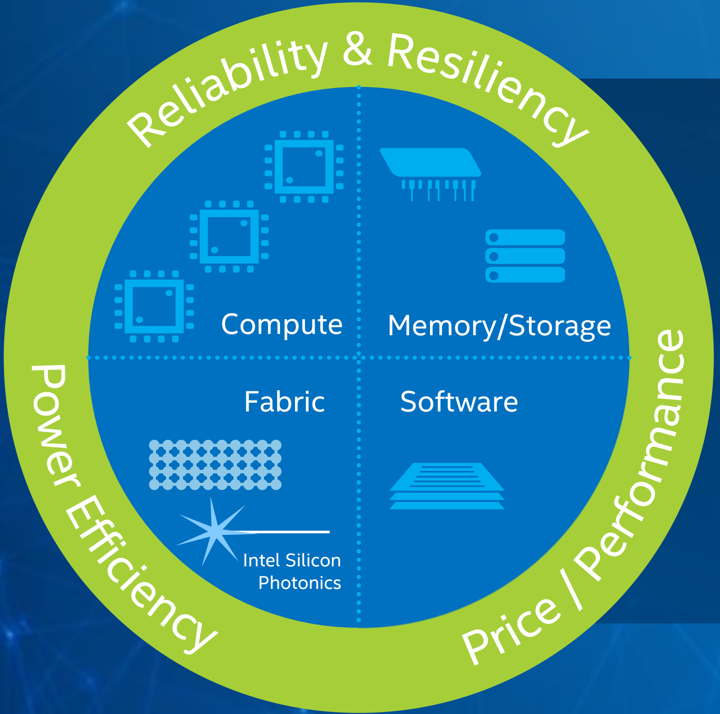
- Evaluate models based on their ability to capture key climatic processes (e.g., teleconnections) that are extracted using ML algorithms
- Significantly speed up model refinement cycle by providing quick diagnostics

2. Design of hybrid-physics-data models

- Replace/enhance individual components of physics-based models to address knowledge gaps and improve predictive performance

Intel® Scalable System Framework

One Framework for Multiple Complex Workflows



Small Clusters Through Peta and Exascale

Compute and Data-Centric Computing

Standards-Based Programmability

IA and HPC Ecosystem Enabling

On-Premise and Cloud-Based

Intel® Xeon® Processors
Intel® Xeon Phi™ Coprocessors
Intel® Xeon Phi™ Processors

Intel® True Scale Fabric
Intel® Omni-Path Fabric
Intel® Ethernet

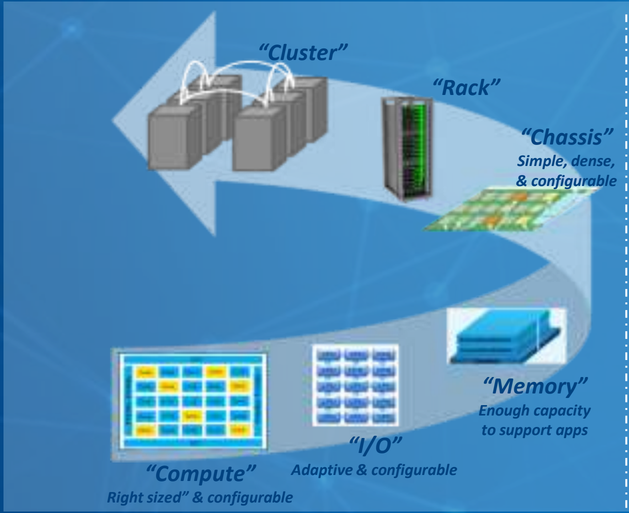
Intel® SSDs
Intel® Lustre*-based Solutions
Intel® Silicon Photonics Technology

Intel® Sofware Tools
Intel® Cluster Ready Program

SSF: Enabling Configurability & Scalability

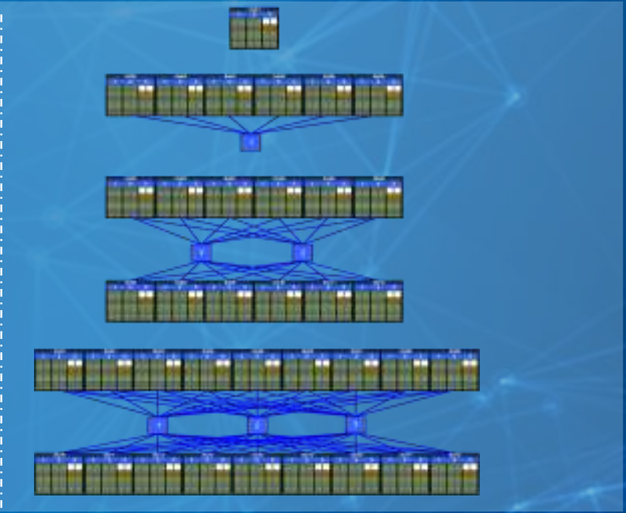
from components to racks to clusters

SSF Path To Exascale



- Xeon or Xeon-Phi – based on workload needs
- Compute flexibly aggregated
- Lowest latency compute to compute interconnect

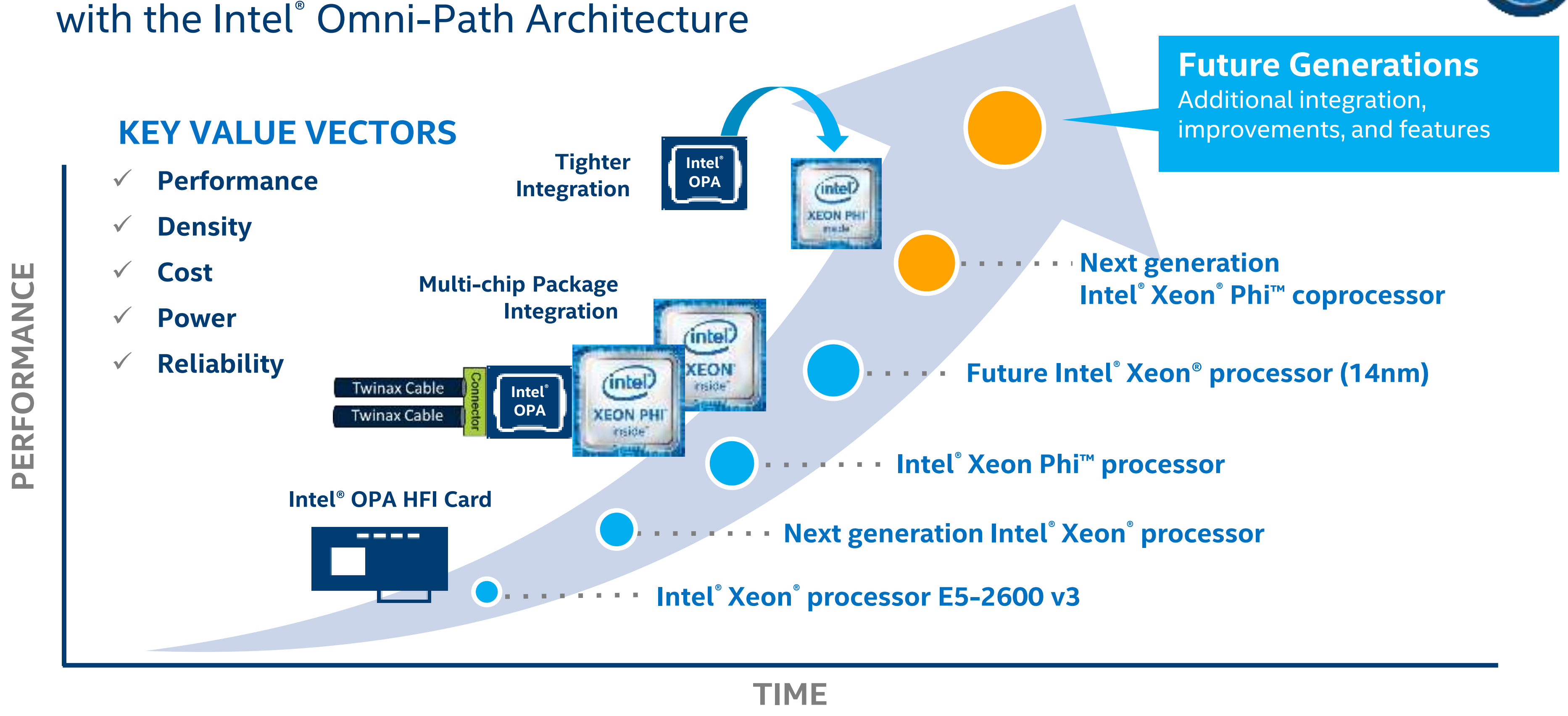
SSF for Scalable Clusters



- I/O Topologies for best performance
- Configurable I/O bandwidth director switch
- Burst buffer to decouple storage from I/O

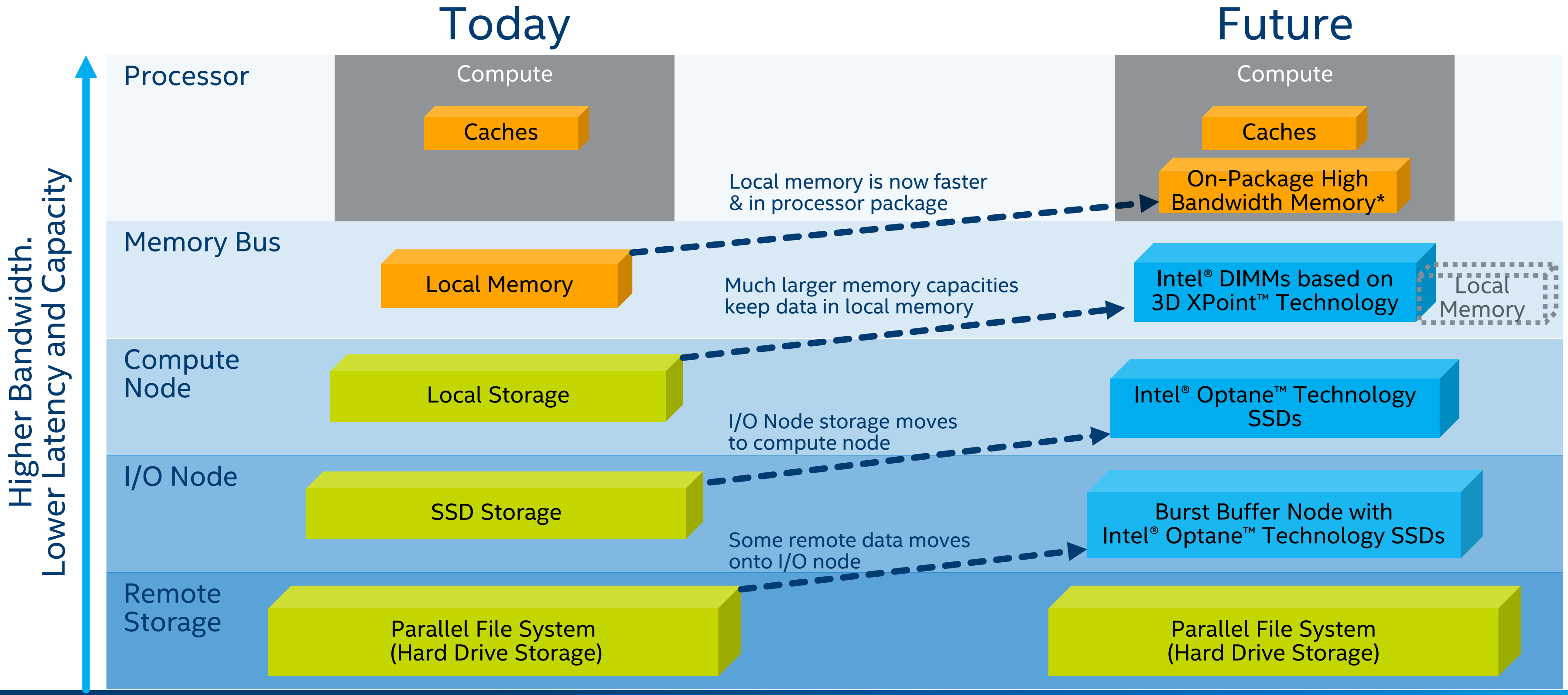
CPU-Fabric Integration

with the Intel® Omni-Path Architecture



Tighter System-Level Integration

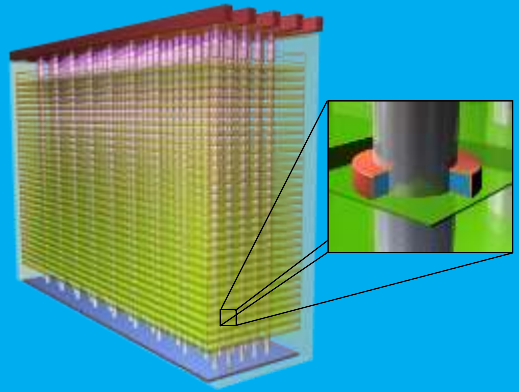
Innovative Memory-Storage Hierarchy



*cache, memory or hybrid mode

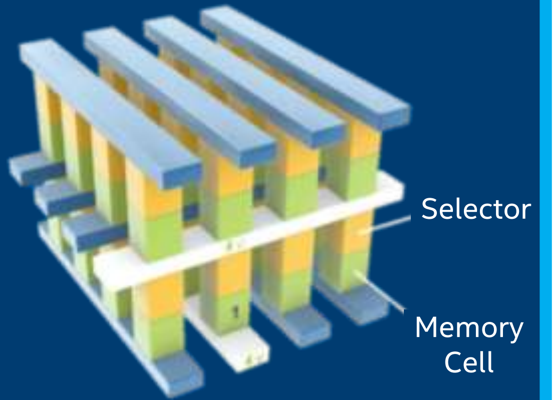
NAND Flash and 3DXPoint Enable New solutions for Data Intensive IO bottleneck

3D MLC and TLC NAND



- Disrupts HDD and block IO software
- Improves BW 10x and IOPs 1,000x, NVMe IO interface for Climate Warm Data Tier

3D XPoint™ Technology



- New memory class storage devices
- Improves BW 1,000x, 100,000x IOPs, endurance 100x and memory interface for Climate Warm Hot Tier

Intel® Software Solutions

Intel® Software Defined Visualization

Low Cost

No Dedicated Viz Cluster

Excellent Performance

Less Data Movement, I/O
Invest Power, Space, Budget in Greater Compute Capability

High Fidelity

Work with Larger Data Sets – Not Constrained by GPU Memory

HPC System Software Stack

An Open Community Effort

Broad Range of Ecosystem Partners
Open Source Availability

Benefits the Entire HPC Ecosystem

Accelerate Application Development
Turnkey to Customizable

Intel® Parallel Studio

Faster Code

Boost Application Performance on Current and Next-Gen CPUs

Create Code Faster

Utilizing a Toolset that Simplifies Creating Fast and Reliable Parallel Code

Open Software Available Today!

Bringing Your Data into Focus

Intel-Supported Software Defined Visualization (SDVis)

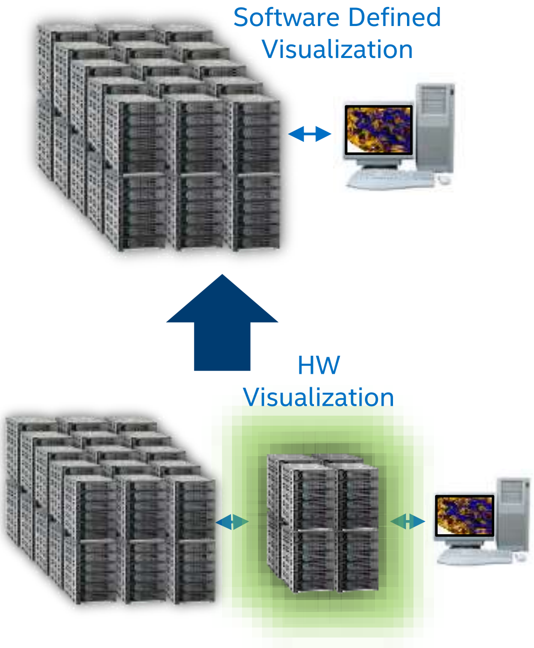
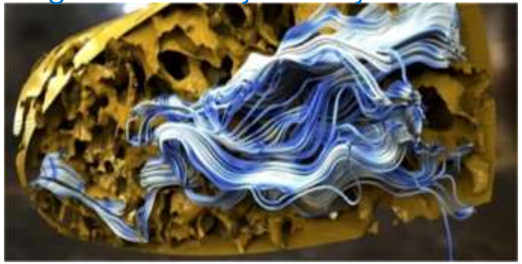
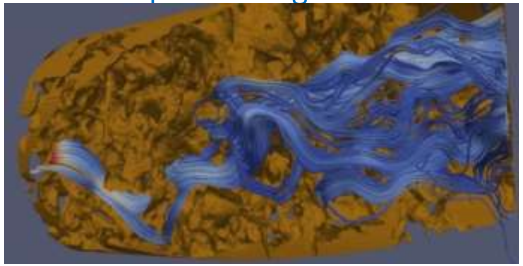


Image Rendered by OSPRay



Standard OpenGL Image



Embree
- CPU Optimized Ray Tracing Algorithms
- 'Tool kit' for Building Ray Tracings Apps
- Broadly Adopted by 3rd Party ISVs
- Web Site: http://embree.github.io

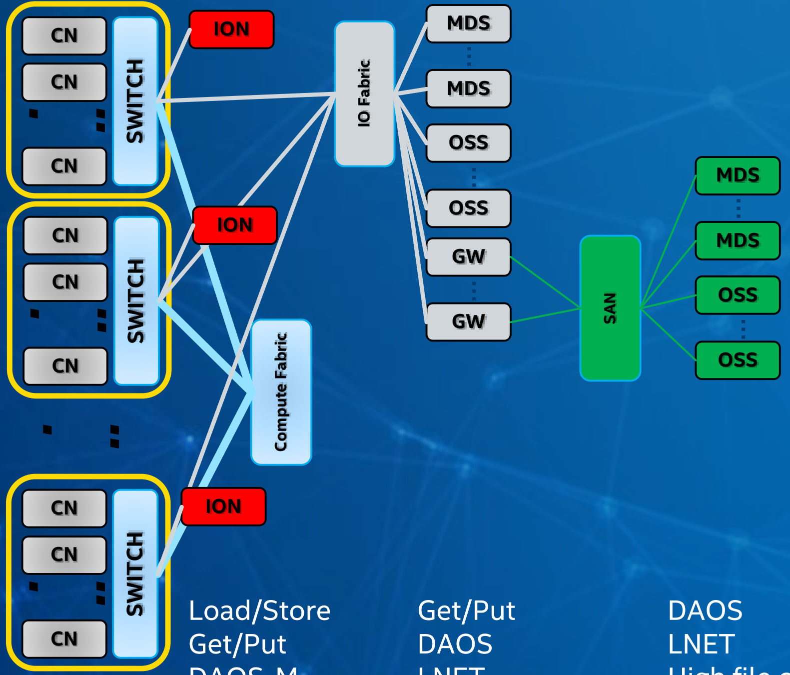
OSPRay[1]
- Rendering Engine Based on Embree
- Library and API Designed to Ease the Creation of Visualization Software
- Web Site: http://ospray.org

OpenSWR[1]
- Rasterization Visualization on CPUs
- Good Enough to Replace HW GPU
- Supports ParaView, Visit, VTK
- Web Site: http://openswr.org

[1] Currently available in alpha

New storage paradigm for data intensive systems

HOT	WARM	COLD
Crystal Ridge	3D NAND	SSD+HDD
~200 TB/s	~10-50 TB/s	~1-2 TB/s
~15 PB	~150 PB	~250 PB

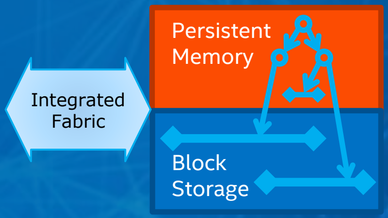


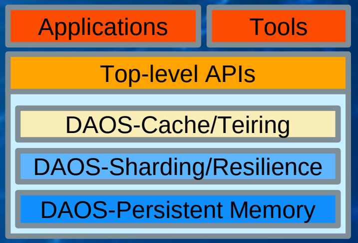
Load/Store	Get/Put	DAOS
Get/Put	DAOS	LNET
DAOS-M	LNET	High file create
Extreme BW+LL	High BW+IOPS	

SSF Enables HPC+HPDA workloads

- System components can be configured to match workload requirements
- Enables new access methodologies (DAOS) to create new generation applications
- Incremental improvements to Lustre to provide enhanced performance for existing applications

Distributed Asynchronous Object Storage





intel®

experience what's inside™