
Welcome to
Large-scale Computing
meets
Big Data

Workshop I: Big Data Meets Large-Scale Computing

Part of the Long Program Science at Extreme Scales: Where Big Data Meets Large-Scale Computing

SEPTEMBER 24 - 28, 2018

Overview

Increasingly large data sets are being ingested and produced by simulations. What experience from large-scale simulation is transferable to big data applications? Conversely, what new optimal algorithms will emerge that are motivated by data-intensive applications being pushed to large scales? How will they enrich traditional simulation? As long as the software stacks, production facilities, and even developer and user communities remain separate, many opportunities for mutual enhancement will be unrealized.

- **Benefits of *in situ* convergence of simulation, analytics, and machine learning**

- **Evolving requirements of high-performance analytics and simulation**

- **Scalable hierarchical algorithms for analytics and simulation**

- **Detecting and exploiting data sparsity**

- **Open problems**

Big Data Meets Large-scale Computing: Workshop 1 themes

Benefits of *in situ* convergence of simulation, analytics, and machine learning

- **Keyes: convergence overview**
- **Costa: architectural convergence**
- **Asch: model inversion & data assimilation**
- **Perdikaris: physics-informed learning**

Evolving requirements of high-performance analytics and simulation

- **Varoquaux: ML and SP with massive data**
- **Szalay: instruments for massive data**
- **Johnson: visualization for massive data**
- **Pascucci: workflows for massive data**
- **Stoica: scalable distributed AI**

Big Data Meets Large-scale Computing: Workshop 1 themes

Scalable hierarchical algorithms for analytics and simulation

- **Peherstorfer: multi-fidelity models for MC**
- **Genton: surrogates in climate models**
- **Li: hierarchical matrices for KRR**
- **Martinsson: randomized matrix algorithms**

Detecting and exploiting data sparsity

- **Bungartz: sparse grids in HPC and big data**
- **Griebel: sparse grids and manifold learning**
- **Pflüger: sparse grids and high-dim DM**

Open problems

- Candes(1): hypothesis generation from data
- Candes(2): non-convex optimization
- Candes(3): finding replicable selections
- Charikar: importance sampling in high dim
- Meila: manifold learning in high dim
- Ghattas: design of experiments in high dim
- Yokota: second-order optimization in DL

Program context– looking outwards

Four workshops

- Big Data Meets Large-Scale Computing
- HPC and Data Science for Scientific Discovery
- HPC for Computationally and Data-intensive Problems
- New Architectures and Algorithms

Big Data Meets Large-scale Computing: Workshop 1 themes

Benefits of *in situ* convergence of simulation, analytics, and machine learning

- **steering in high-dimensional parameter space**
- **smart data compression**
- **data-driven modeling (e.g., refinement of empirical models through learning)**
- **physics-based "regularization" of analytics**
- **simulation as a source of training data**
- **machine learning to impute missing data**

Evolving requirements of high-performance analytics and simulation

Scalable hierarchical algorithms for analytics and simulation

Detecting and exploiting data sparsity

Open problems

My purpose

Advocate convergence of big data and large-scale computing

- one aspect of broad scientific agenda for these two fields

Both fields have their own momentum and are encountering their own limitations

Will provide background motivation and point to four recent community reports

Coming from simulation side…

A "big picture" talk

My lecture is "big picture"

My algorithmic interests were already presented by Jeff Hittinger in jointly authored tutorials lectures on 14 Sep

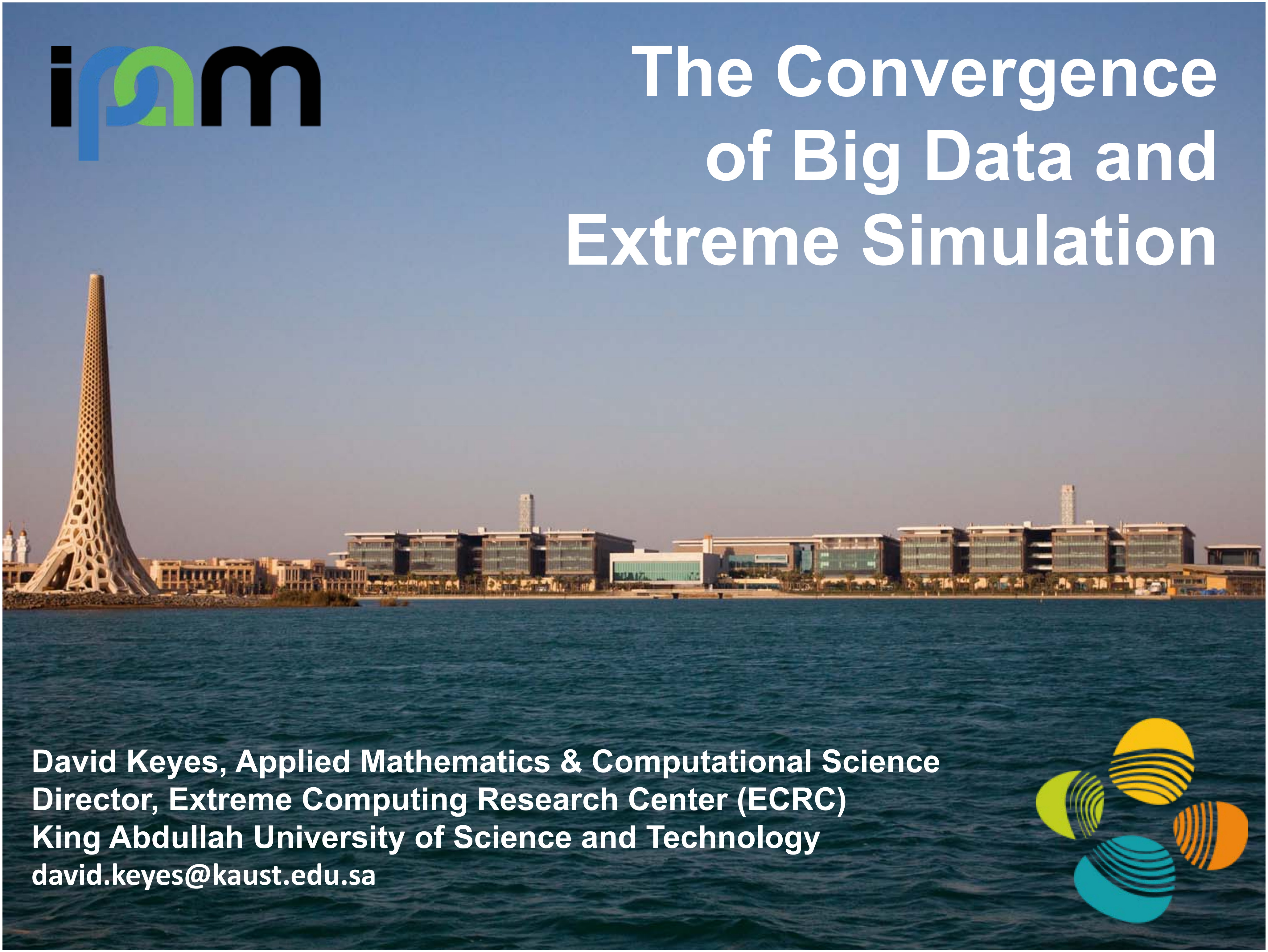
pdfs, mp3 linked

- ***Build It and They Will Come: How Hardware Influences Large-Scale Simulation***
- ***High-Performance Numerical Algorithms for Large-Scale Simulation***

Cannot resist (if time permits) calling attention back to points that Jeff mentioned on directions for algorithms that

- **benefit extreme simulation**
- **are conjectured to benefit big data**

The Convergence of Big Data and Extreme Simulation



David Keyes, Applied Mathematics & Computational Science
Director, Extreme Computing Research Center (ECRC)
King Abdullah University of Science and Technology
david.keyes@kaust.edu.sa

Conclusions up front

Many motivations exist to bring together large-scale simulation and big data analytics ("convergence")

Should be combined *in situ*

- **pipelining between simulation and analytics through disk files with sequential applications leaves too many benefits "on the table"**

Many hurdles to convergence

- **but ultimately, this will not be a "forced marriage"**

Scientists and engineers may be minority users of "big data" (today and perhaps forever) but can become leaders in the "big data" community

- **by harnessing high performance computing**
- **being pathfinders for other applications, once again!**

Converged already?

- **Artificial Intelligence**
  - *top-down, deductive, laws/rules:* **Simulation**
  - *bottom-up, Inductive, history/ examples:* **Analytics & Learning**
    - *predict data points:* **Regression**
      - **Linear**
      - **Nonlinear, Max likelihood**
    - *predict categories:* **Classification & Clustering**
      - *supervised labeled data:* **Classification**
        - **Bayesian**
        - **Decision tree**
        - **Neural networks & Deep leaning**
      - *unsupervised unlabeled data:* **Clustering, Dimension reduction & Density estimation**
        - **K-means**

c/o E.-L. Goh, HP

Simulation and analytics: virtuous cycle

Both include both models and data

- simulation uses a model (mathematical) to produce data
- analytics uses data to produce a model (statistical)

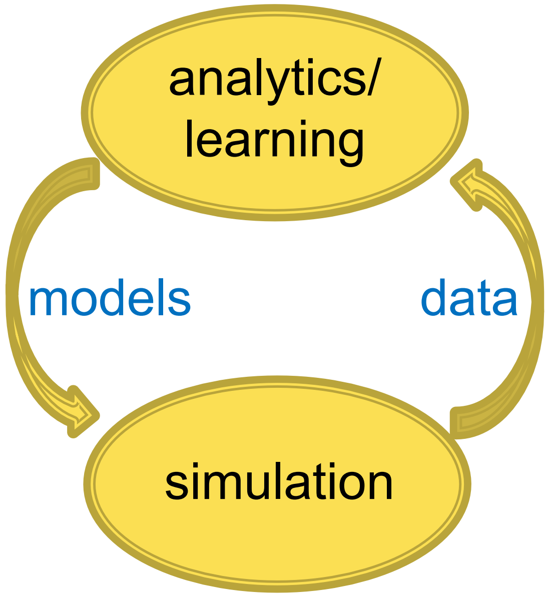
Models generated by analytics can be used in simulation

- not the only source of models, of course

Data generated by simulation can be used in analytics

- not the only source of data, of course

A virtuous cycle can be set up

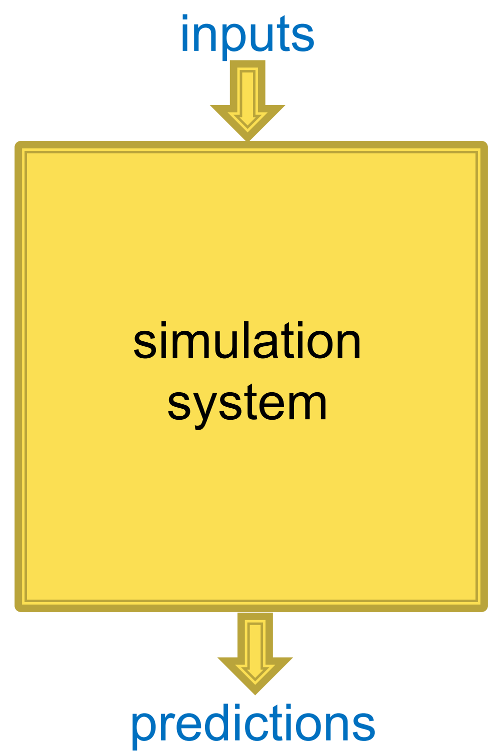


c/o A. Raies, KAUST

Simulation and learning: difference

- **Primary novelty in machine-based "intelligence" is the learning part**

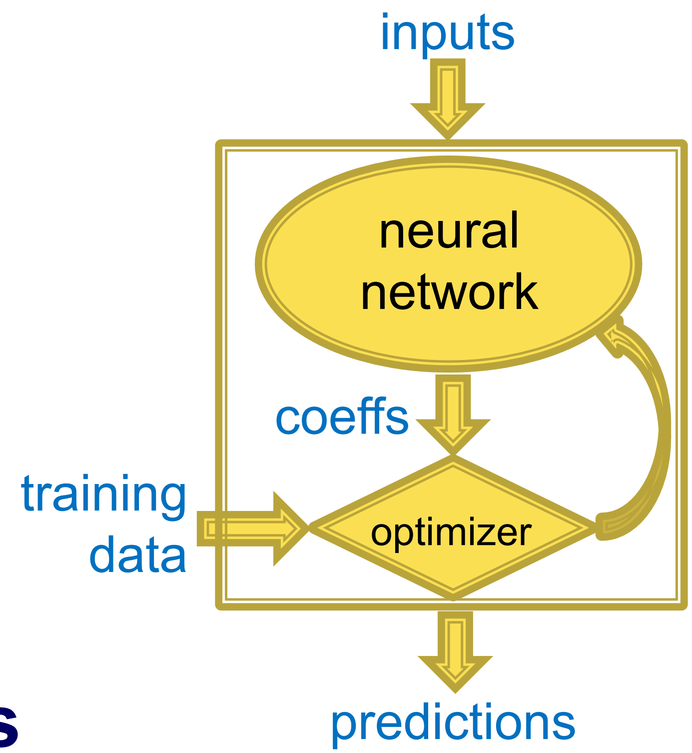
- **A simulation system is historically a fixed, human-engineered code that does not improve with the flow of data through it**



Simulation and learning: difference

Machine learning systems improve as they ingest data

- **make inferences and decisions on their own**
- **actually generate the model**



Of course, as with a child, when provided with information, a machine may learn incorrect rules and make incorrect decisions

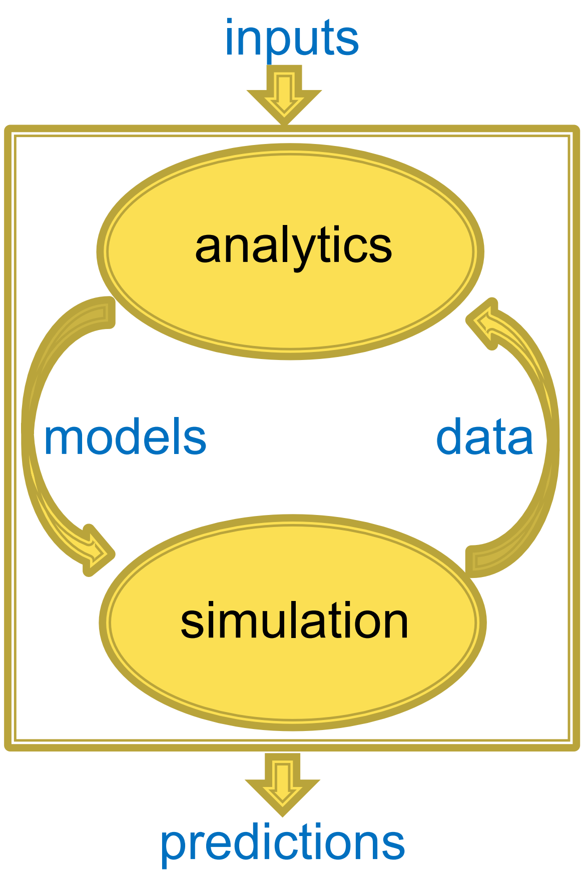
- **in scientific contexts, we have extra recourse**

An *in situ* converged system

Including learning in the simulation loop can enhance the predictivity of the simulation

Including both simulation data and observational data in the learning loop can enhance the learning

Ultimately a win-win marriage



But now, a tale of two communities…

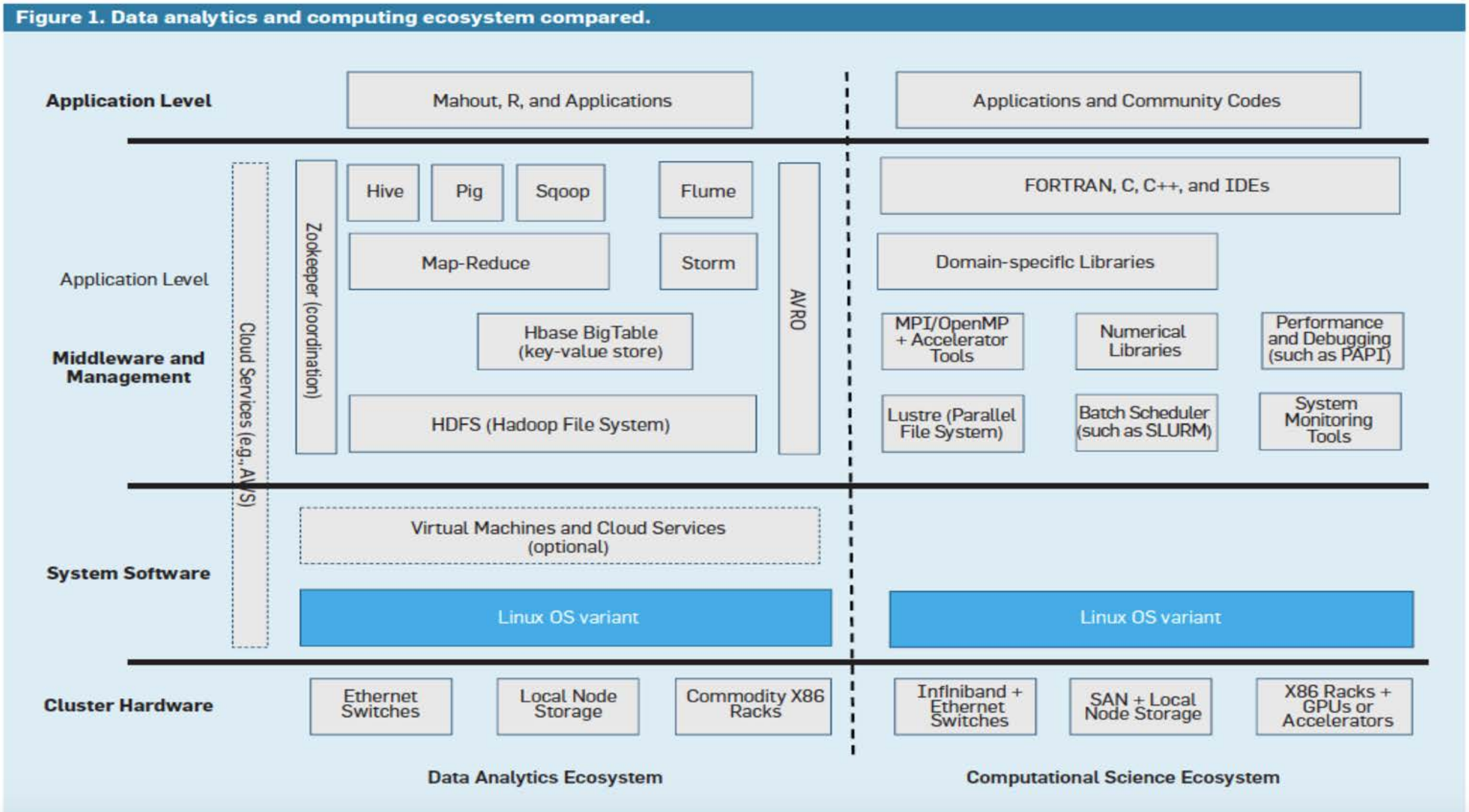
HPC: high performance computing

- grew up around Moore's Law multiplied by massive parallelism
- predictive on par with experiments (e.g., Nobel prizes in chemistry)
- recognized for policy support (e.g., nuclear weapons, climate treaties)
- recognized for decision support (e.g., oil drilling, therapy planning)

HDA: high-end data analytics

- grew up around open-source tools (e.g., Hadoop, TensorFlow) from online service providers
- created trillion-$ markets in analyzing human preferences
- now dictating the design of computer architecture (e.g., NVIDIA tensor cores, Intel A21)
- transforming university curricula and national investments
- taking on scientific data, evolving as it goes

… two *divergent* communities



Exascale Computing and Big Data

D. Reed & J. Dongarra, *Comm. ACM* **58**:56-68, July 2015

…divergent not just in software stacks

Data ownership

HPC: ***generally*** **private** **HDA:** ***often*** **curated by community**

Data access

HPC: bulk access, fixed **HDA: fine-grained access, elastic**

Data storage

HPC: local, temporary **HDA: cloud-based, persistent**

…divergent not just in software stacks

Scheduling policies

HPC: batch — **HDA:** interactive

HPC: exclusive space — **HDA:** shared space

Community premiums

HPC: capability, reliability — **HDA:** capacity, resilience

Hardware infrastructure

HPC: "fork-lift upgrades" — **HDA:** incremental upgrades

Pressure on HPC

Vendors, even those facing the lucrative call for exascale systems by government (>$1B projects in Japan, China, USA) must leverage their technology developments for the much larger data science markets

This includes preoccupation with lower precision floating point

Fortunately, there are critical cross-cutting concerns

- energy efficiency
- limited memory per core
- limited memory bandwidth per core

Pressure on HDA

Since the beginning of the big data age, data has been moved over "stateless" networks

- routing is based on address bits in the data packets
- no system-wide coordination of data sets or buffering

Workarounds coped with volume but are now creaking

- ftp mirror sites, web-caching (e.g., Akamai)

Solutions for buffering massive scientific data sets from the "edge" …

- seismic arrays, satellite networks, telescopes, scanning electron microscopes, beamlines, sensors, drones, etc.

…will be useful for the "fog" environments of the big data "cloud"

CS&E report

"HPC supercomputers and cloud data centers […] face challenges […] of extreme scalability, fault tolerance, cost of data movement, and power management. The advent of big data has spearheaded new large-scale distributed computing technologies and parallel programming models such as MapReduce, Hadoop, Spark, and Pregel, which offer innovative approaches to scalable high-throughput computing, with a focus on data locality and fault tolerance. […]"

"In many applications, the need for distributed computing arises from the sheer volume of the data. […] The growing levels of parallelism in computer architectures require software in distributed machine learning systems such as TensorFlow to be highly parallel. […] Economy-of-scale pressures will contribute to a convergence of technologies for computing at large scale."

"Research and Education in CS&E" report (2018)

Research and Education in CS&E report

SIAM REVIEW
Vol. 60, No. 3, pp. 707–754

© 2018 Society for Industrial and Applied Mathematics

Research and Education in Computational Science and Engineering*

Officers of the SIAM Activity Group on Computational Science and Engineering (SIAG/CSE), 2013–2014:

"Rüde report" (2018)

SIAM REVIEW
Vol. 43, No. 1, pp. 163–177

© 2001 Society for Industrial and Applied Mathematics

Graduate Education in Computational Science and Engineering*

SIAM Working Group on CSE Education†

"Petzold report" (2001)

Convergence matrix

	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

Table 1 from "Pathways to Convergence" report (2018)

Convergence report

International Journal of
HIGH PERFORMANCE
COMPUTING APPLICATIONS

Research Paper

The International Journal of High
Performance Computing Applications
2018, Vol. 32(4) 435–479

DOI: 10.1177/1094342018778123
journals.sagepub.com/home/hpc
SAGE

Big data and extreme-scale computing: Pathways to Convergence-Toward a shaping strategy for a future software and data ecosystem for scientific inquiry

M Asch, T Moore, R Badia, M Beck, P Beckman, T Bidot, F Bodin, F Cappello, A Choudhary, B de Supinski, E Deelman, J Dongarra, A Dubey, G Fox, H Fu, S Girona, W Gropp, M Heroux, Y Ishikawa, K Keahey, D Keyes, W Kramer, J-F Lavignon, Y Lu, S Matsuoka, B Mohr, D Reed, S Requena, J Saltz, T Schulthess, R Stevens, M Swany, A Szalay, W Tang, G Varoquaux, J-P Vilotte, R Wisniewski, Z Xu and I Zacharov

10^{18} BDEC
BIG DATA AND EXTREME-SCALE COMPUTING

(URL in last slide)

“Pathways to Convergence” report (2018)

There is no single big data type

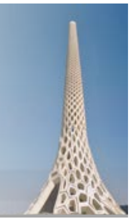
In scientific big data, different approaches may be natural for three different categories:

- data arriving from edge devices (often in real time, e.g., beamlines) that is never centralized but processed on the fly
- federated multi-source data (e.g., bioinformatics) intended for "permanent" archive
- combinations of data retrieved from archival source and dynamic data from a simulation (e.g., assimilation in climate/weather)

"Pathways" report addresses these challenges in customized sections

	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

Theory-Guided data science



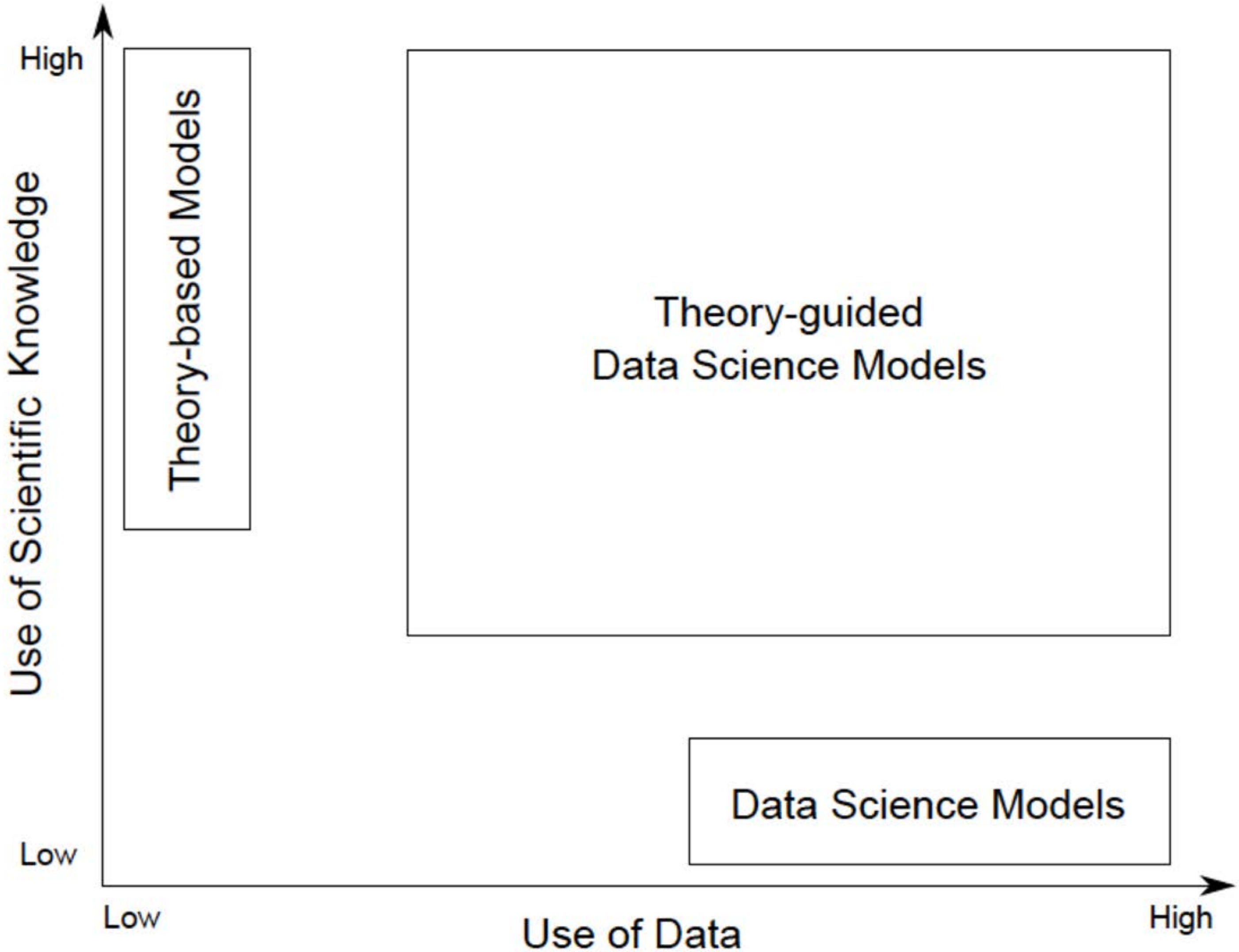


Figure 1 from "Theory Guided Data Science" report (2017)

Theory-Guided Data Science report

86 references, including many examples from biology, chemistry, earth science and engineering, may be found in:

2318 IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 29, NO. 10, OCTOBER 2017

Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data

Anuj Karpatne, Gowtham Atluri, James H. Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar

Abstract—Data science models, although successful in a number of commercial domains, have had limited applicability in scientific problems involving complex physical phenomena. Theory-guided data science (TGDS) is an emerging paradigm that aims to leverage the wealth of scientific knowledge for improving the effectiveness of data science models in enabling scientific discovery. The overarching vision of TGDS is to introduce scientific consistency as an essential component for learning generalizable models. Further, by producing scientifically interpretable models, TGDS aims to advance our scientific understanding by discovering novel domain insights. Indeed, the paradigm of TGDS has started to gain prominence in a number of scientific disciplines such as turbulence modeling, material discovery, quantum chemistry, bio-medical science, bio-marker discovery, climate science, and hydrology. In this paper, we formally conceptualize the paradigm of TGDS and present a taxonomy of research themes in TGDS. We describe several approaches for integrating domain knowledge in different research themes using illustrative examples from different disciplines. We also highlight some of the promising avenues of novel research for realizing the full potential of theory-guided data science.

Index Terms—Data science, knowledge discovery, domain knowledge, scientific theory, physical consistency, interpretability

“Theory-Guided Data Science” report (2017)

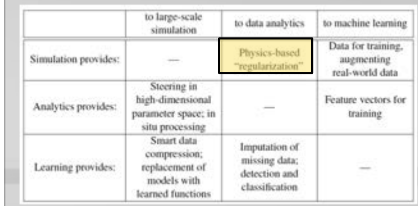
	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

Theory-Guided data science

"Data science models, although successful in a number of commercial domains, have had limited applicability in scientific problems involving complex physical phenomena. Theory-guided data science (TGDS) is an emerging paradigm that aims to leverage the wealth of scientific knowledge for improving the effectiveness of data science models in enabling scientific discovery. The overarching vision of TGDS is to introduce scientific consistency as an essential component for learning generalizable models."

"Further, by producing scientifically interpretable models, TGDS aims to advance our scientific understanding by discovering novel domain insights. Indeed, the paradigm of TGDS has started to gain prominence in a number of scientific disciplines such as turbulence modeling, material discovery, quantum chemistry, bio-medical science, bio-marker discovery, climate science, and hydrology."

"Theory Guided Data Science" report (2017)



	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

Theory-Guided data science

Each point in the (generally) high-dimensional space below represents a model; three families of increasing complexity are depicted.

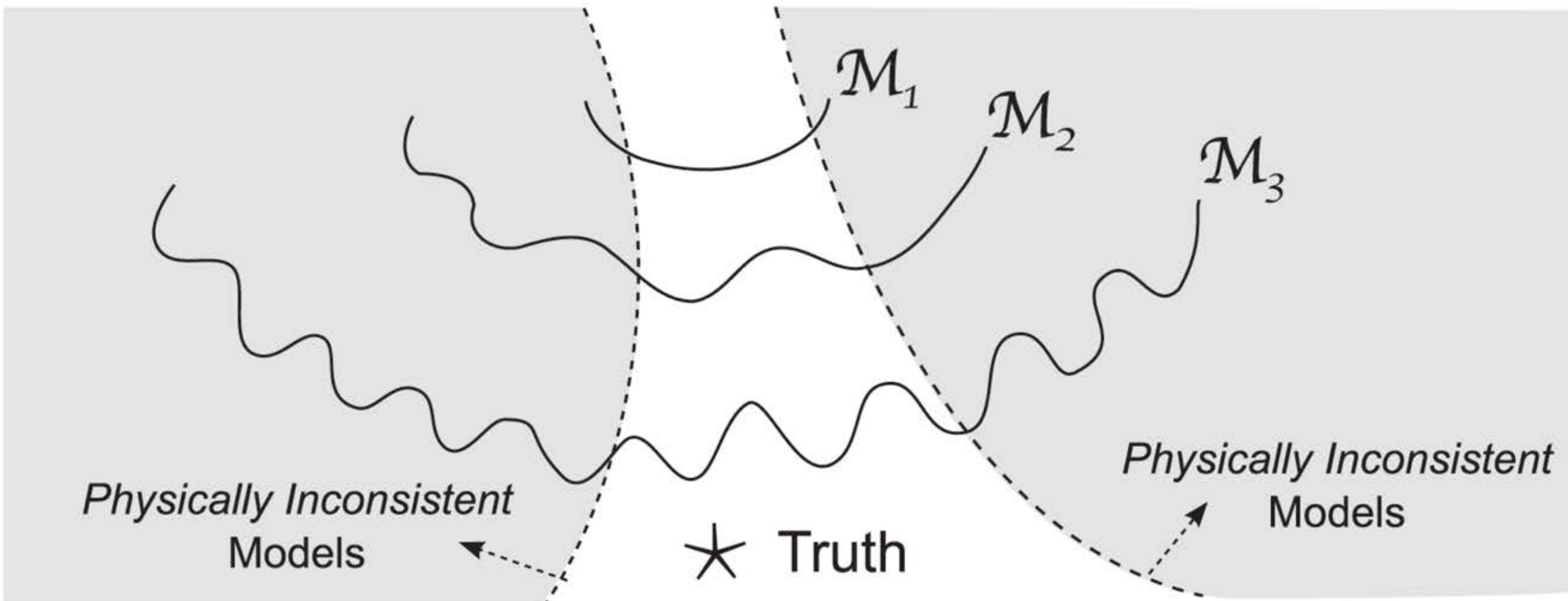
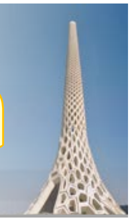


Fig. 2. Scientific knowledge can help in reducing the model variance by removing physically inconsistent solutions, without likely affecting their bias.

"Theory Guided Data Science" report (2017)

	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

CNNs for binding prediction



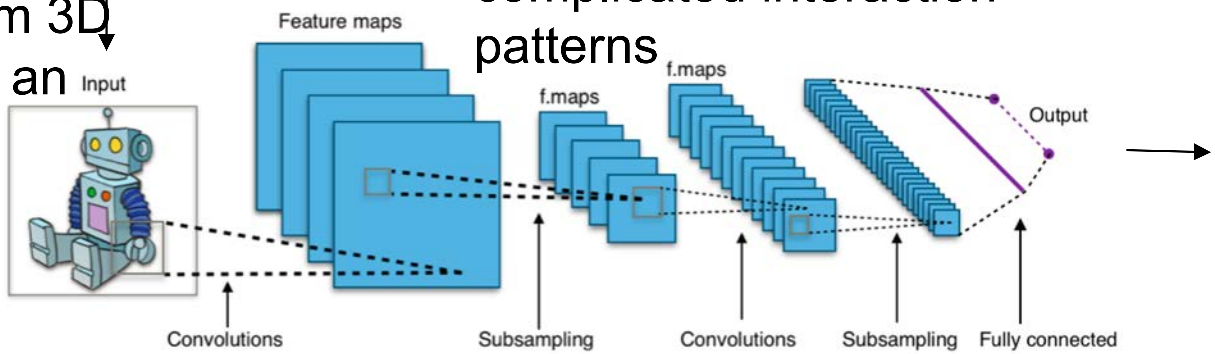
Major Histo-Compatibility Complex (MHC)

Peptide

The generation of the images from molecular dynamics software creates a more favorable training set than in classification of standard images because differences due to rotation, illumination, etc. can be removed.

Transform 3D model to an image

Train the CNN to capture complicated interaction patterns



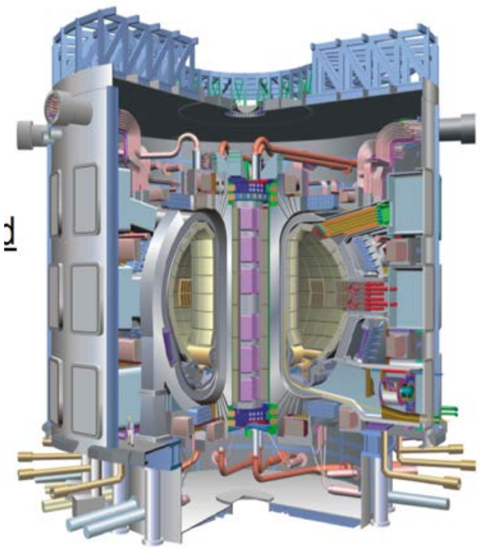
Estimate binding probability

This campaign has led to success in the Critical Assessment of PRediction of Interactions (CAPRI) competition for protein docking, now in its 42nd round

c/o M. Ignatov, SUNY Stony Brook

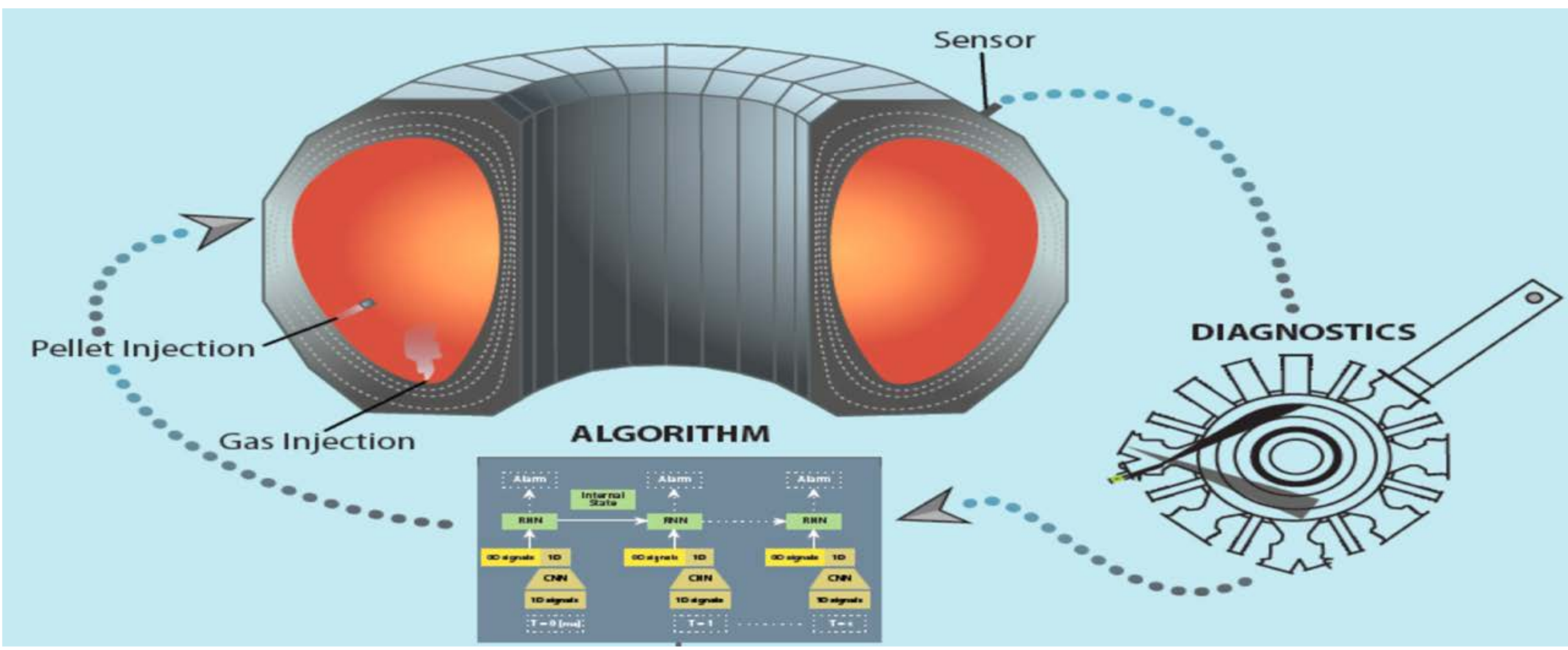
	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

DL for disruption mitigation



Predicting and mitigating disruptions (fluid instabilities) of the magnetic bottle that contains the 300,000,000 Celsius burning plasma cannot be done in real time with simulations

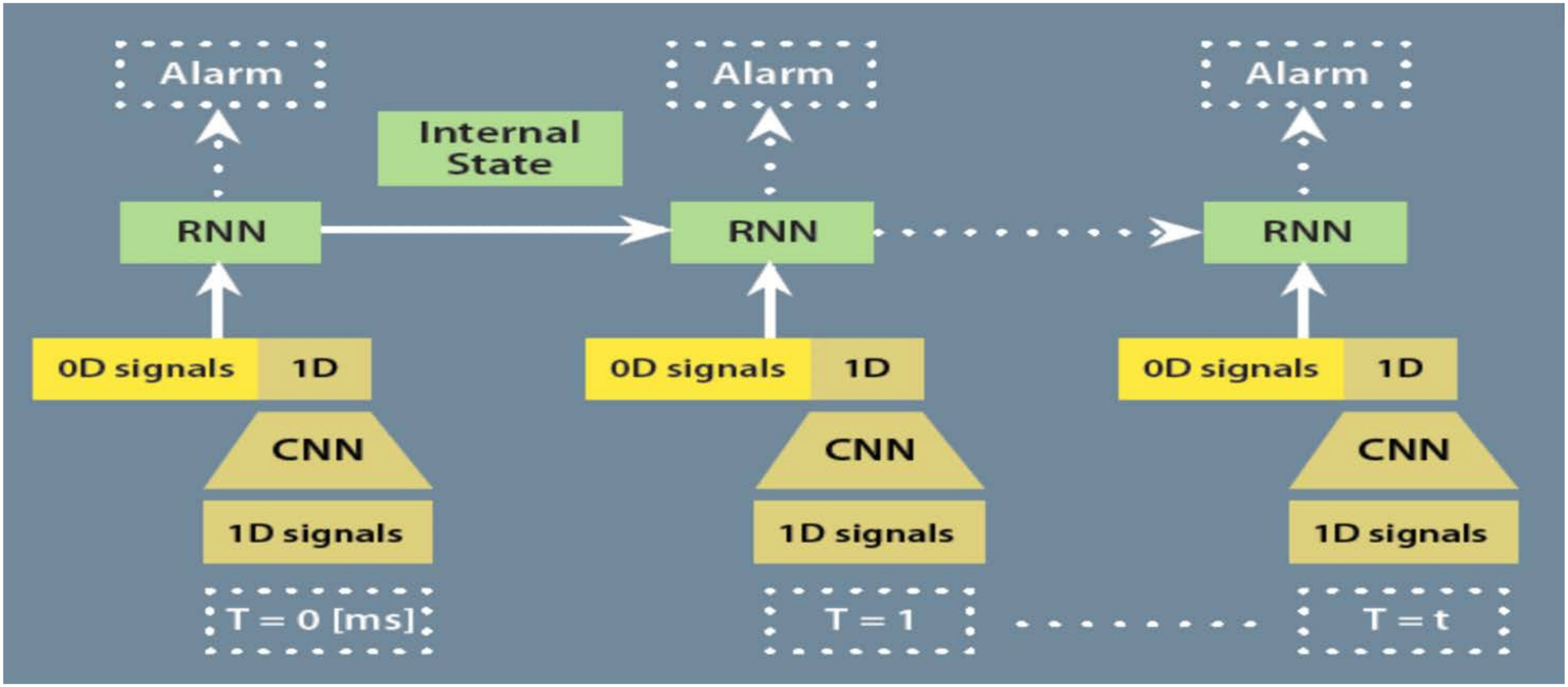
30ms warning is needed for effective control



c/o W. Tang, Princeton Plasma Physics Lab

	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

DL for disruption mitigation



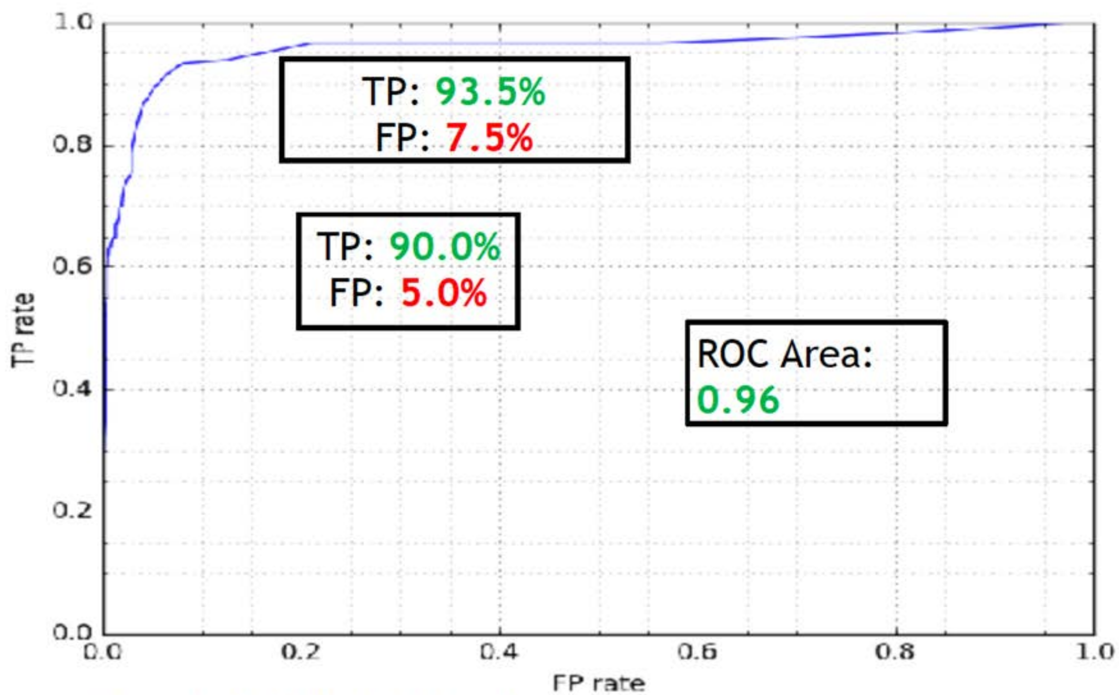
This campaign won the NVIDIA Global Impact Award at the 2018 GPU Technology Conference

c/o W. Tang, Princeton Plasma Physics Lab

FRNN Code PERFORMANCE: ROC CURVES

JET ITER-like Wall Cases @30ms before Disruption

Performance Tradeoff: Tune True Positives (good: correctly caught disruption) **vs. False Positives** (bad: safe shot incorrectly labeled disruptive).

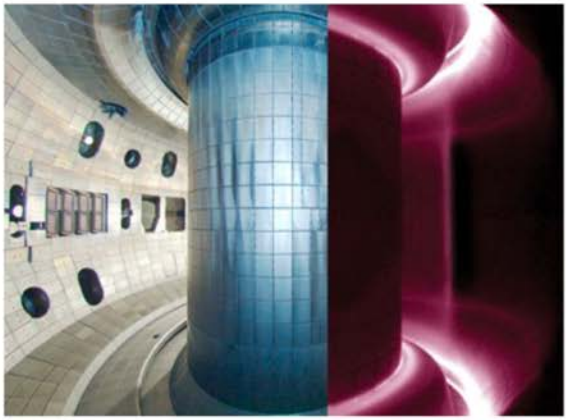


Data (~50 GB), 0D signals:

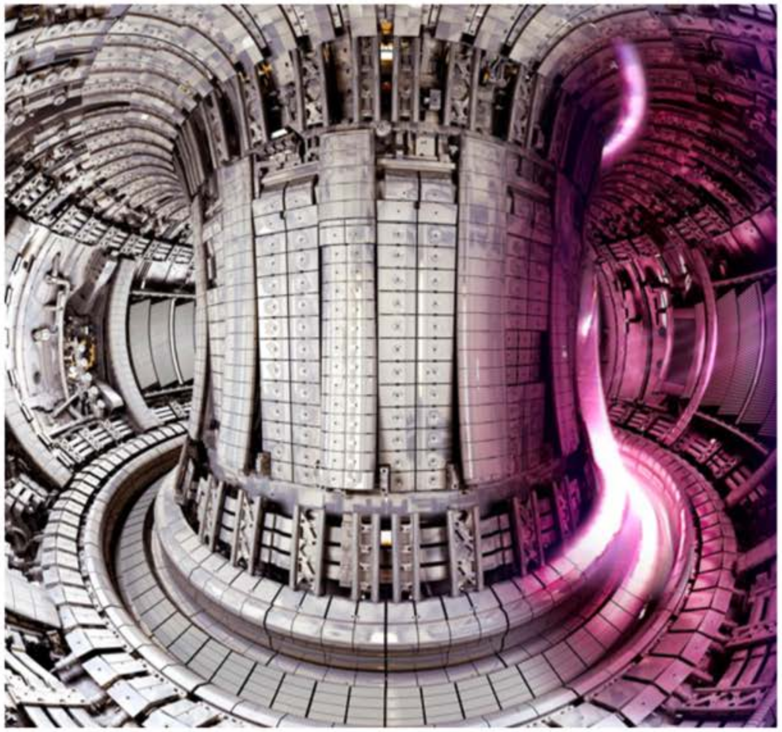
- Training: on **4100 shots** from **JET C-Wall** campaigns
- Testing **1200 shots** from **Jet ILW** campaigns
- **All shots used**, no signal filtering or removal of shots

	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

Cross-machine prediction



Train (DIII-D)



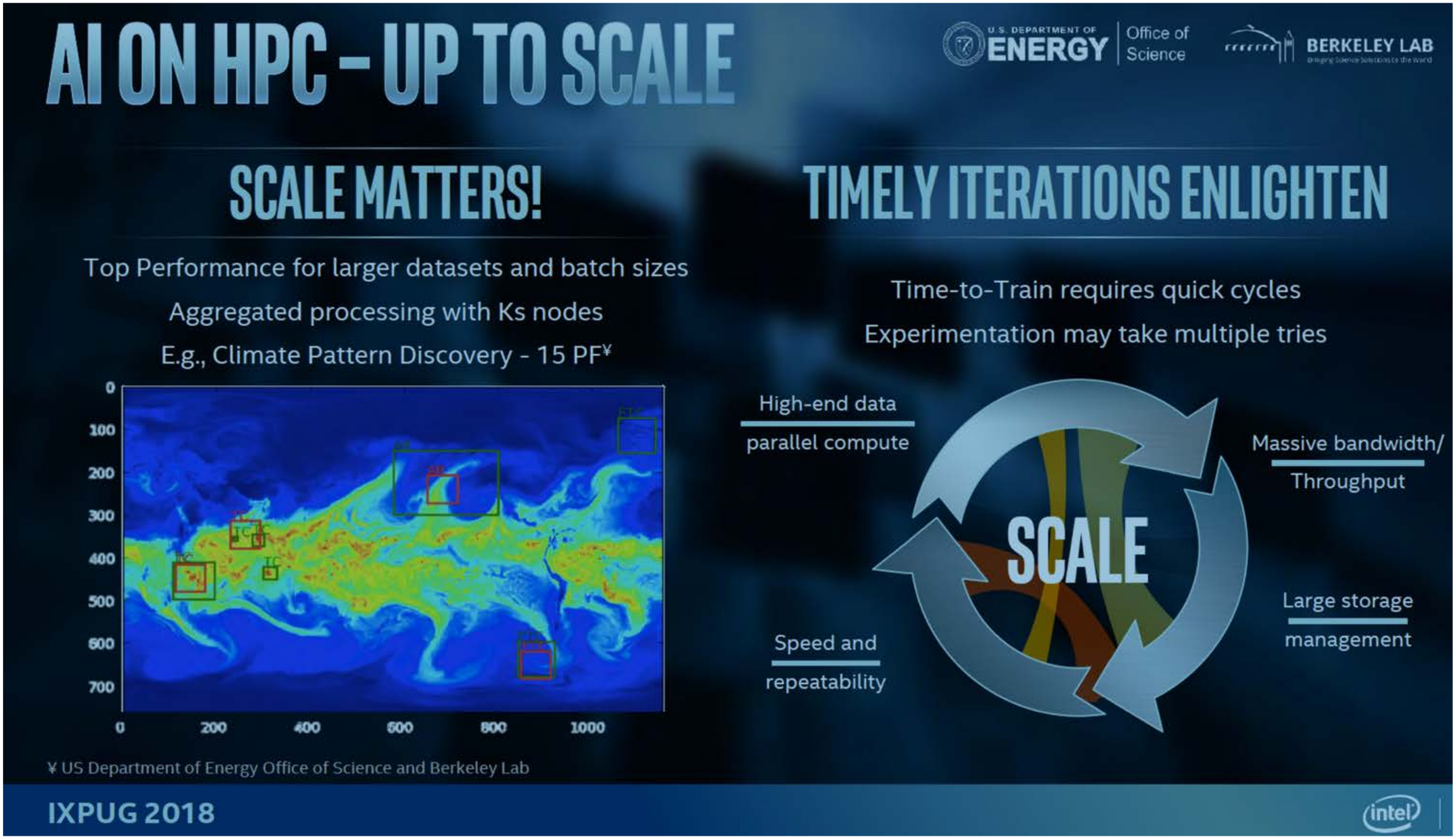
Test (JET)

FRNN 1D	**0.836**
FRNN 0D	0.817
XGBoost	0.616

c/o W. Tang, Princeton Plasma Physics Lab

	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

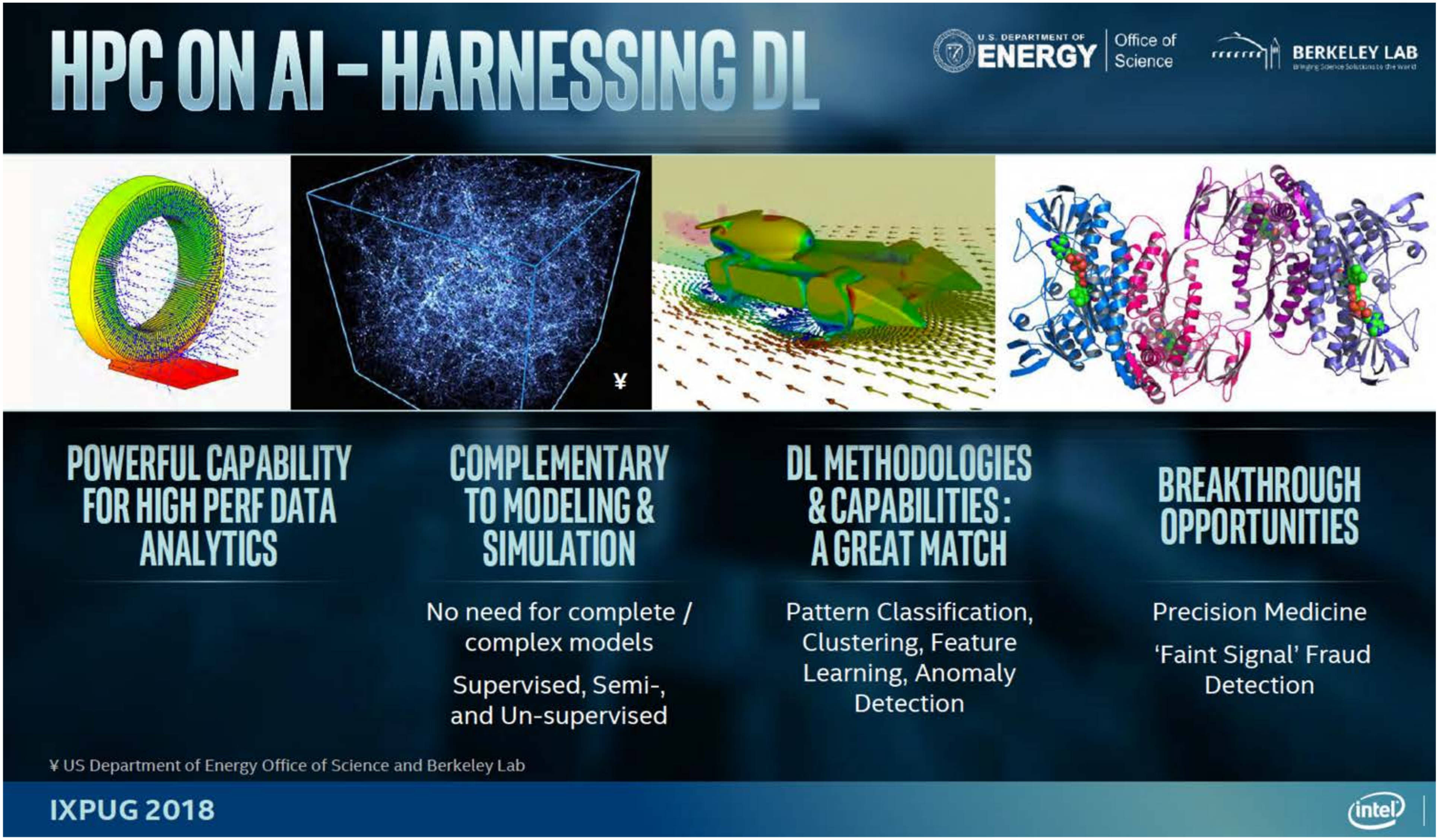
Top-row advances



c/o A. Gara, Intel

	to large-scale simulation	to data analytics	to machine learning
Simulation provides:	—	Physics-based "regularization"	Data for training, augmenting real-world data
Analytics provides:	Steering in high-dimensional parameter space; in situ processing	—	Feature vectors for training
Learning provides:	Smart data compression; replacement of models with learned functions	Imputation of missing data; detection and classification	—

Left-column advances



A new instrument is emerging

"Nothing tends so much to the advancement of knowledge as the application of a new instrument. The native intellectual powers of men in different times are not so much the causes of the different success of their labors, as the peculiar nature of the means and artificial resources in their possession."

—Humphrey Davy (1778-1829)

Inventor of electrochemistry (1802)

Discoverer of K, Na, Mg, Ca, Sr, Ba, B, Cl (1807-1810)

Scientific method "on steroids"

SC18

Dallas, TX | hpc inspires.

The International Conference for High Performance Computing, Networking, Storage, and Analysis

The "steroids" are high performance computing technologies

SC18 Invited Speakers

Tuesday	Speakers	Topics
10:30-11:15	Chris Johnson, U Utah	Scientific Visualization
11:15-12:00	Steve Furber, U Manchester	Neuromorphic Computing
3:30-4:15	Margaret Martonosi, Princeton	Computer Architecture for Mobile Computing
4:15-5:00	Bryan Catanzaro, NVIDIA	Machine Learning
Wednesday		
10:30-11:15	Doug Kothe, ORNL	US DOE's Exascale Computing Program
11:15-12:00	Depei Qian, Xi'an Jiaotong U	China's Exascale Computing Program
3:30-4:15	Satoshi Sekiguchi, AIST	Japanese Program in Artificial Intelligence
4:15-5:00	Mary-Anne Piette, LBNL	HPC Modeling of Urban Systems
Thursday		
8:30-9:15	Matthias Troyer, Microsoft	Quantum computing
9:15-10:00	Cecilia Aragon, U Washington	Enabling humans to explore and gain insight from vast data sets
10:30-11:15	Pete Beckman, ANL	Internet of Things
11:15-12:00	Padma Raghavan, Vanderbilt	Energy Efficiency and Linear Algebra

Dallas, TX | hpc inspires.

Gordon Bell Finalists

Tsuyoshi Ichimura, U Tokyo
A Fast Scalable Implicit Solver for Nonlinear Time-evolution Earthquake City Problem on Low-ordered Unstructured Finite Elements with Artificial Intelligence

Wenguang Chen, Tsinghua U
ShenTu: Processing Multi-Trillion Edge Graphs on Millions of Cores in Seconds

Andre Walker-Loud, LBNL
Simulating the weak death of the neutron in a femtoscale universe with near-exascale computing

Prabhat, LBNL
Towards Exascale Deep Learning: Analysis of Extreme Weather Patterns at 263 PF/s

Robert Patton, ORNL
151-PFlops Deep Learning for Electron Microscopy: From Learning Physics to Atomic Manipulation

Daniel Jacobson, ORNL
Attacking the Opioid Epidemic: Determining the Epistatic and Pleiotropic Genetic Architectures for Chronic Pain and Opioid Addiction

Motivations for convergence

Scientific and engineering advances

- tune physical parameters in simulations for predictive performance
- tune algorithmic parameters of simulations for execution performance
- filter out nonphysical candidates in learning
- provide data for learning

Economy of data center operations

- obviate I/O
- obviate computation!

Development of a competitive workforce

- leaders in adopting disruptive tools have advantages in capability and in recruiting

Economy of data center operations

Reduce the time burden of I/O

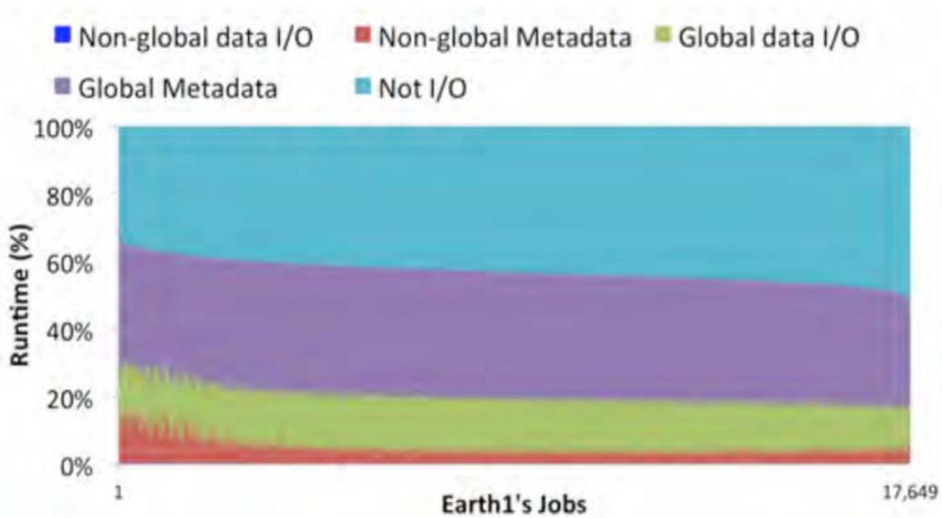


Figure 4: Breakdown of total run time for each Earth1 job.

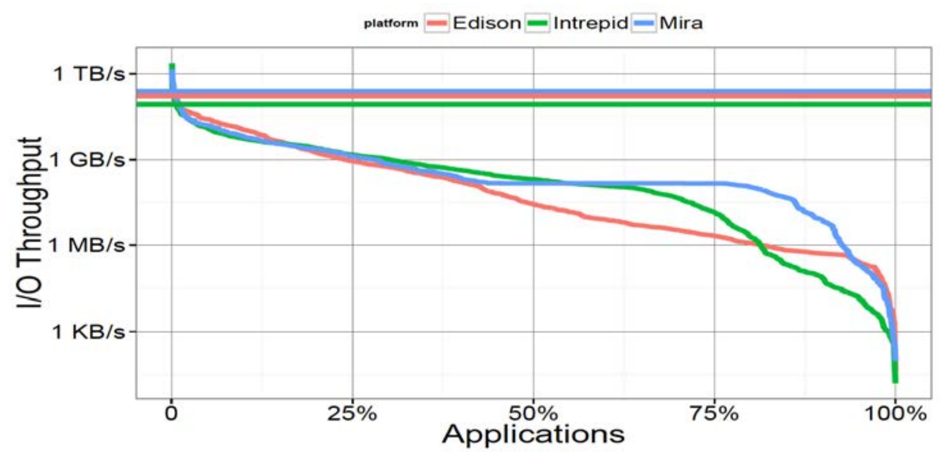
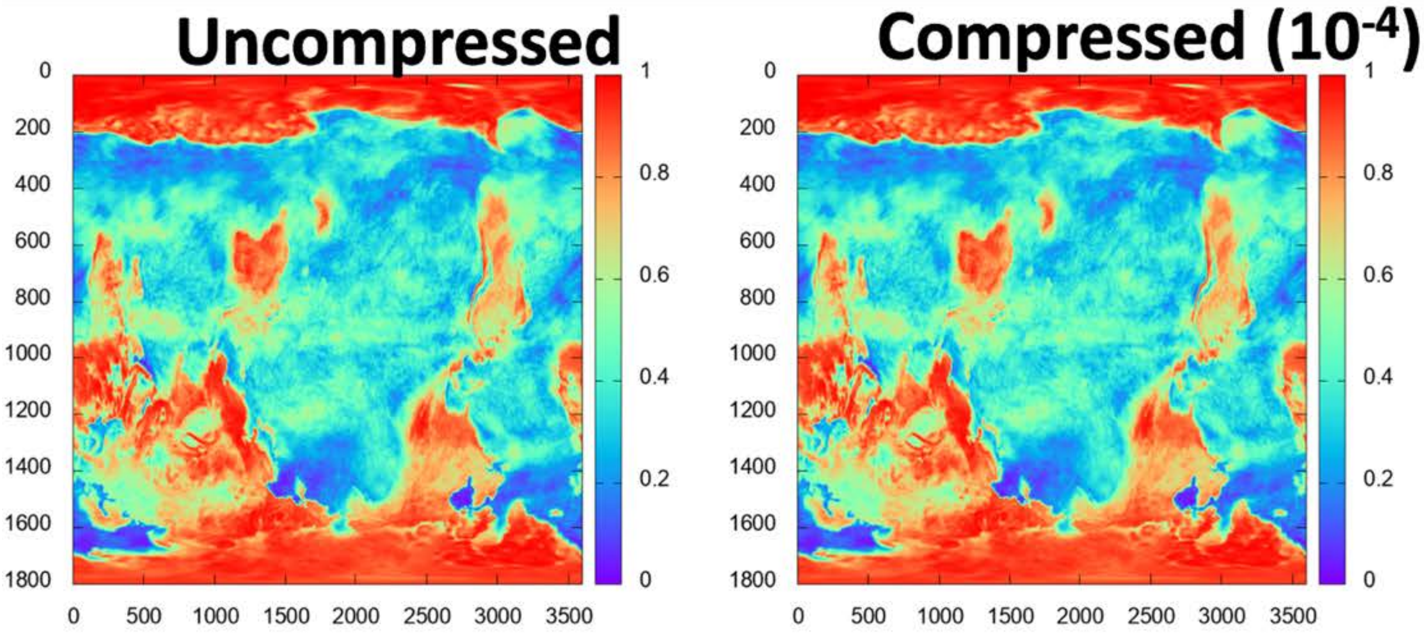


Figure 6: Maximum I/O throughput of each app across all its jobs on a platform, and platform peak I/O throughput.

c/o W. Gropp, UIUC

Economy of data center operations

Reduce the space burden of I/O

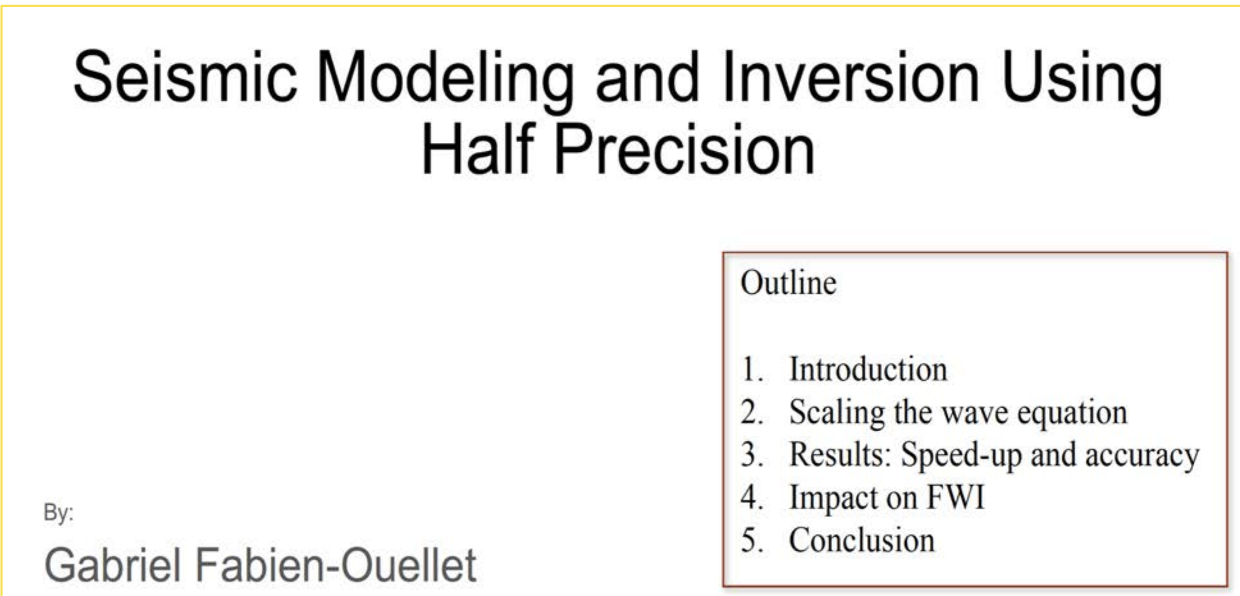


SZ Compression factor: 6.4

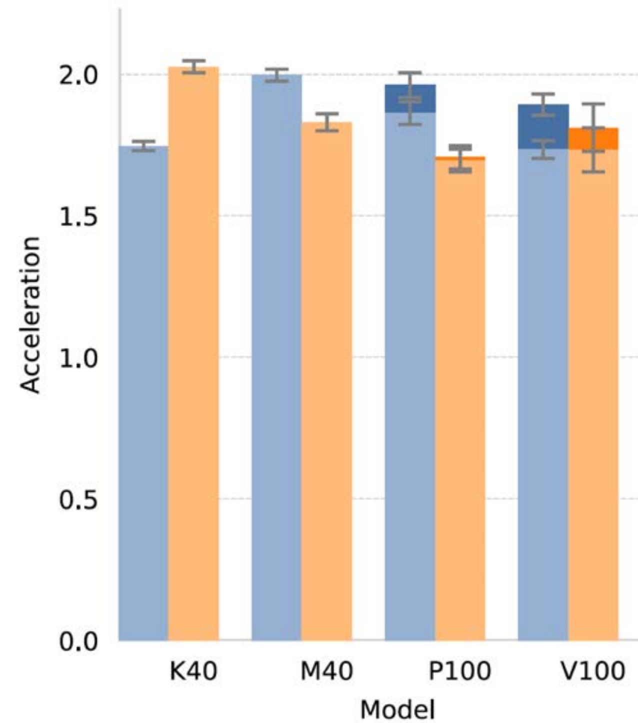
(1.4 with GZIP)

c/o F. Cappello, Argonne

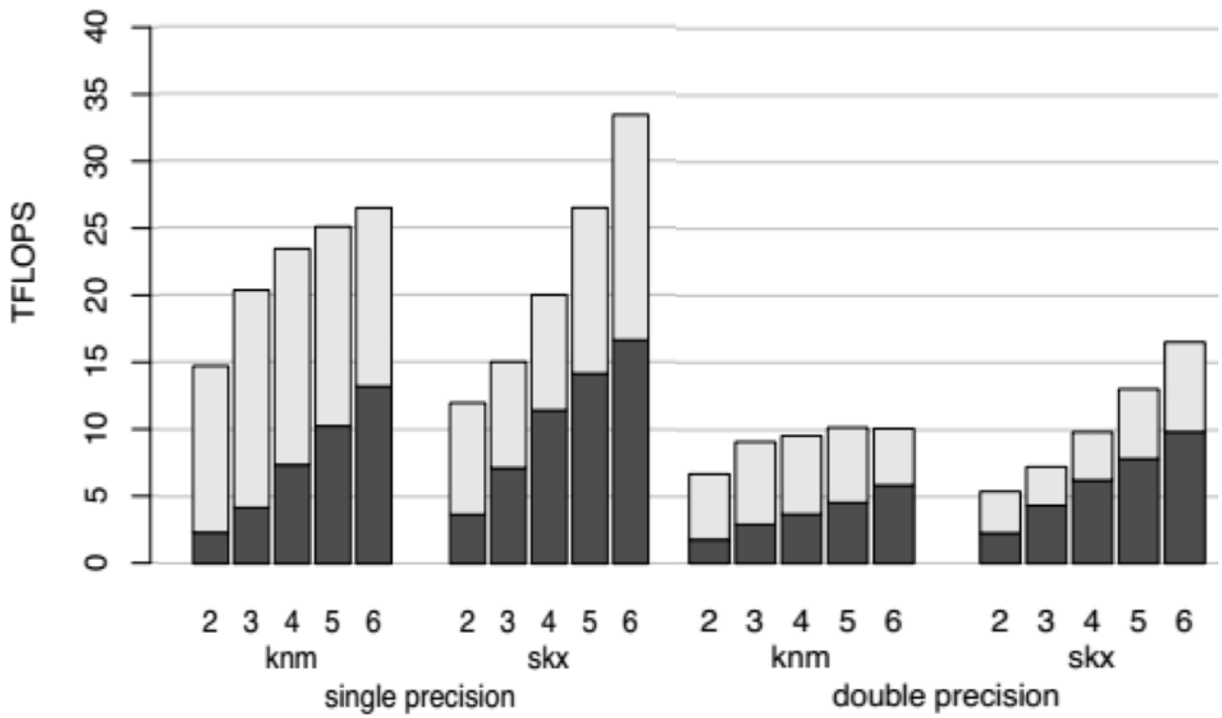
Bonus: rethinking HPC in HDA datatypes



FP16 over FP32



Bonus: rethinking HPC in HDA datatypes



Development of competitive workforce

CALL FOR PROPOSALS: A21 ESP DATA, LEARNING PROJECTS

- **CFP 10 January 2018**
  - Deadline 8 April 2018
- Selections June 2018
  - 5 Data projects
  - 5 Learning projects
- Two-year funded ALCF postdoc
- Cross-cutting proposals targeting the convergence of simulation, data and learning are very much encouraged.

DATA

- Experimental/observational data
  - Image analysis
  - Multidimensional structure discovery
- Complex and interactive workflows
- On-demand HPC
- Persistent data techniques
  - Object store
  - Databases
- Streaming/real-time data
- Uncertainty quantification
- Statistical methods
- Graph analytics

LEARNING

- Deep learning
- Machine learning steering simulations
  - Parameter scans
  - Materials design
  - Observational signatures
- Data-driven models and refinement for science using ML/DL
- Hyperparameter optimization
- Pattern recognition
- Reduced model derivation
- Bridging gaps in theory

12

Argonne NATIONAL LABORATORY

Tang's tokamak disruption detection project is one of those selected

c/o D. Martin, Argonne

Roadmap report

INTERNATIONAL EXASCALE SOFTWARE PROJECT ROADMAP 1.0

THE INTERNATIONAL JOURNAL of HIGH PERFORMANCE COMPUTING APPLICATIONS

The International Exascale Software Project roadmap

The International Journal of High Performance Computing Applications
000(00) 1–58

DOI: 10.1177/1094342010391989
hpc.sagepub.com
SAGE

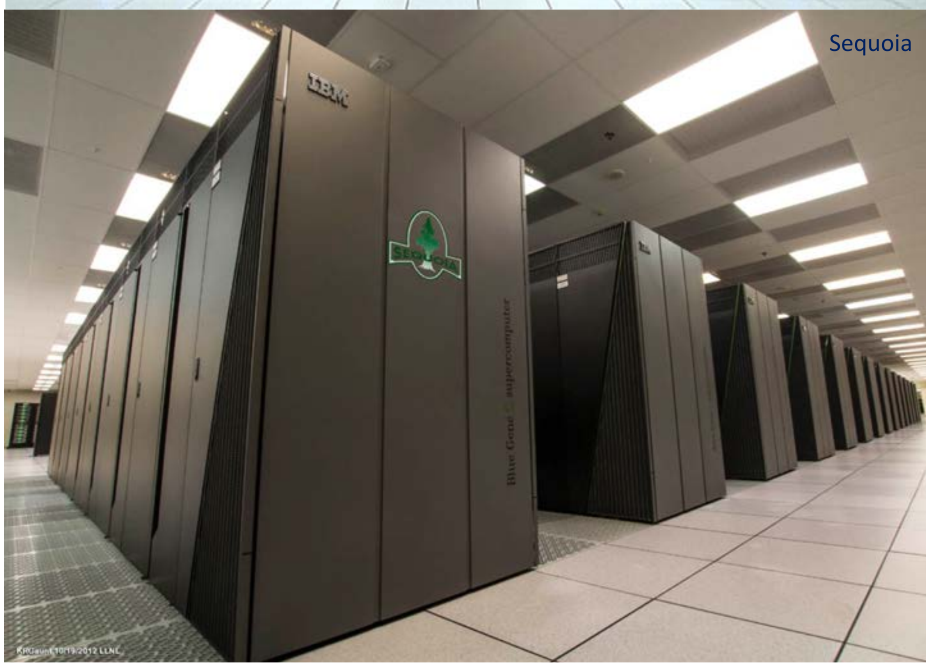
Jack Dongarra, Pete Beckman, Terry Moore, Patrick Aerts, Giovanni Aloisio, Jean-Claude Andre, David Barkai, Jean-Yves Berthou, Taisuke Boku, Bertrand Braunschweig, Franck Cappello, Barbara Chapman, Xuebin Chi, Alok Choudhary, Sudip Dosanjh, Thom Dunning, Sandro Fiore, Al Geist, Bill Gropp, Robert Harrison, Mark Hereld, Michael Heroux, Adolfy Hoisie, Koh Hotta, Zhong Jin, Yutaka Ishikawa, Fred Johnson, Sanjay Kale, Richard Kenway, David Keyes, Bill Kramer, Jesus Labarta, Alain Lichnewsky, Thomas Lippert, Bob Lucas, Barney Maccabe, Satoshi Matsuoka, Paul Messina, Peter Michielse, Bernd Mohr, Matthias S. Mueller, Wolfgang E. Nagel, Hiroshi Nakashima, Michael E Papka, Dan Reed, Mitsuhisa Sato, Ed Seidel, John Shalf, David Skinner, Marc Snir, Thomas Sterling, Rick Stevens, Fred Streitz, Bob Sugar, Shinji Sumimoto, William Tang, John Taylor, Rajeev Thakur, Anne Trefethen, Mateo Valero, Aad van der Steen, Jeffrey Vetter, Peg Williams, Robert Wisniewski and Kathy Yelick

“Exascale Software Roadmap” report

The familiar

Taihu Light

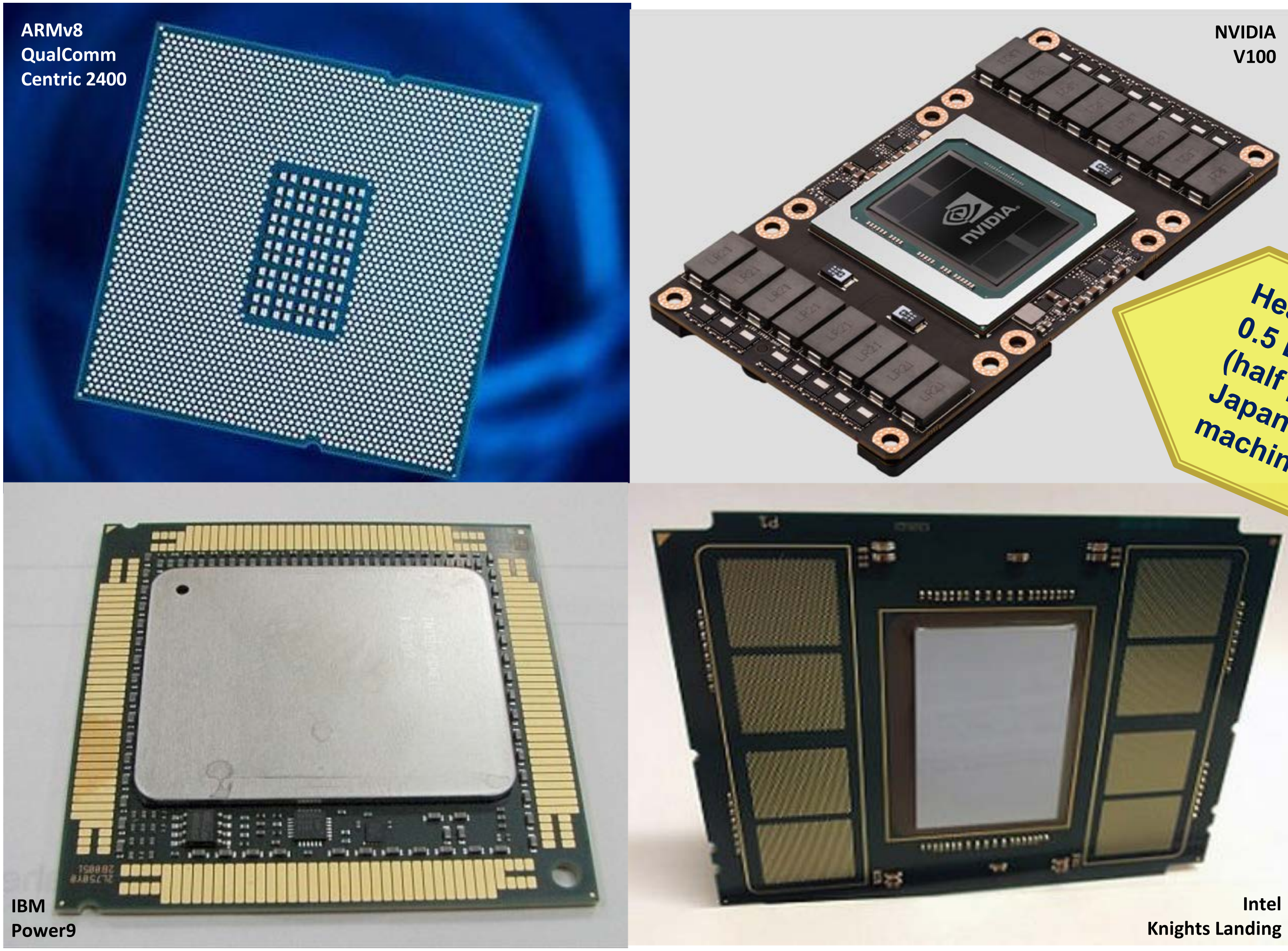
Shaheen



Sequoia

K

The challenge



Heart of the 0.5 Exaflop/s (half precision) Japanese AI machine

Architectural imperatives for algorithms

Reduce synchrony

- in frequency or span or both
- cannot afford to synchronize a billion imbalanced cores

Reside "high" on the memory hierarchy

- as close as possible to the processing elements
- latency to DRAM may be a thousand cycles
- moving data is orders of magnitude more costly in energy than computing

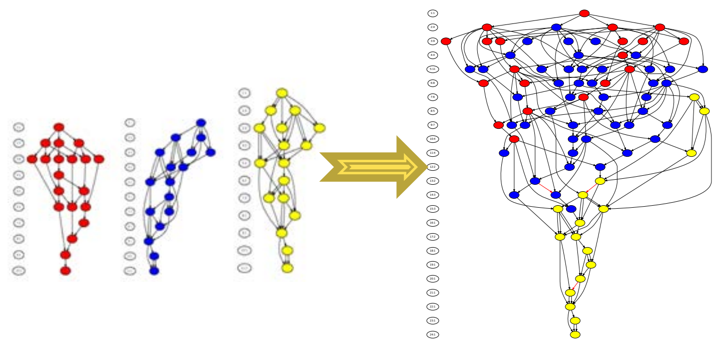
Increase SIMT/SIMD-style shared-memory concurrency

- one instruction can trigger 8 (AVX 512) to 64 (tensor core) operations

Exascale algorithmic strategies

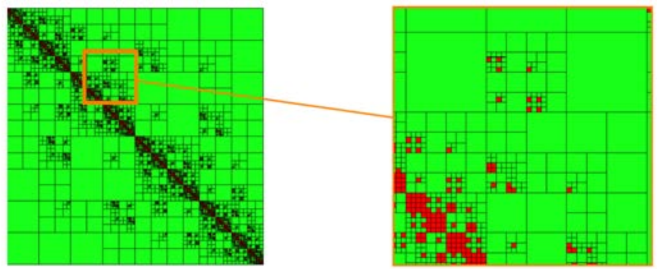
Employ dynamic runtime systems based on directed acyclic task graphs (DAGs)

- e.g., ADLB, Argo, Charm++, HPX, Legion, OmpSs, Quark, STAPL, StarPU



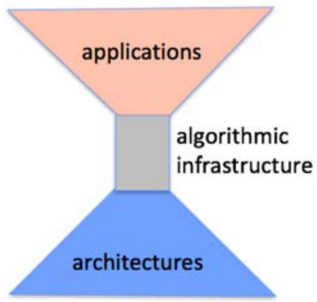
Exploit hierarchical low-rank data sparsity

- meet "curse of dimensionality" with "blessing of low rank"



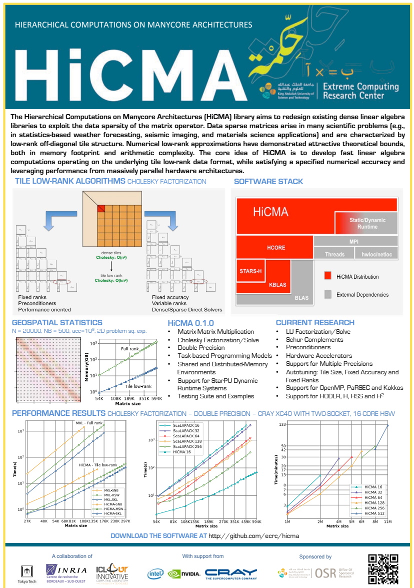
Code to the architecture, but present an abstract API

- "hourglass model" of IP/TCP for processors

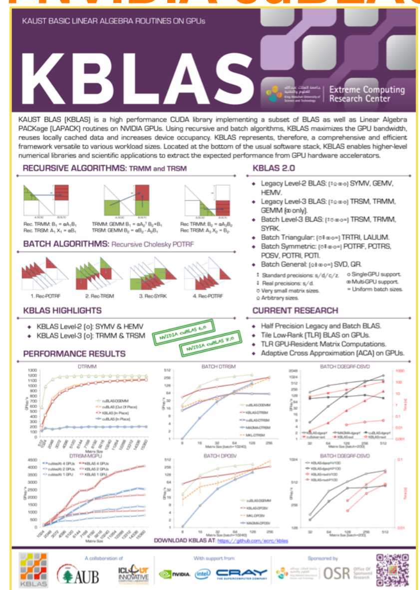


One institution's response

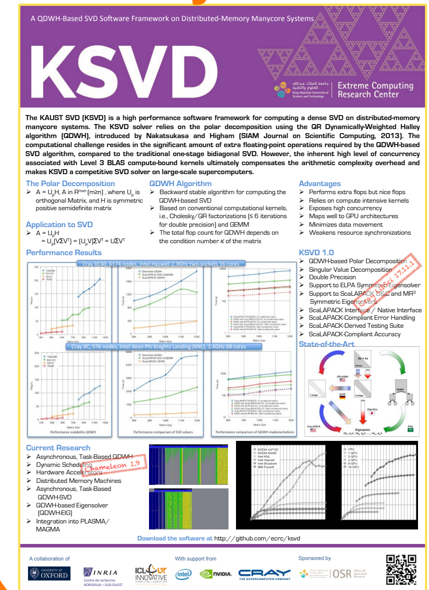
Some open source software released by KAUST's ECRC @ github



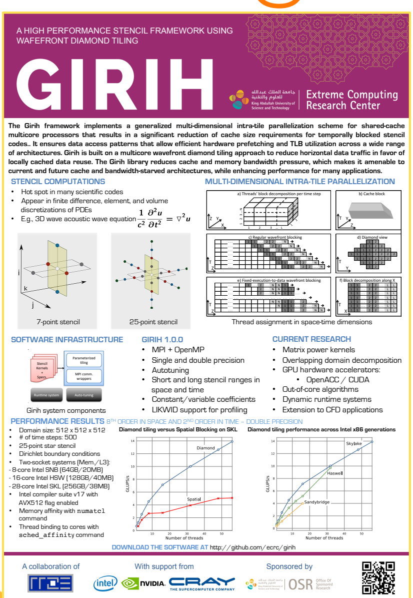
in NVIDIA cuBLAS

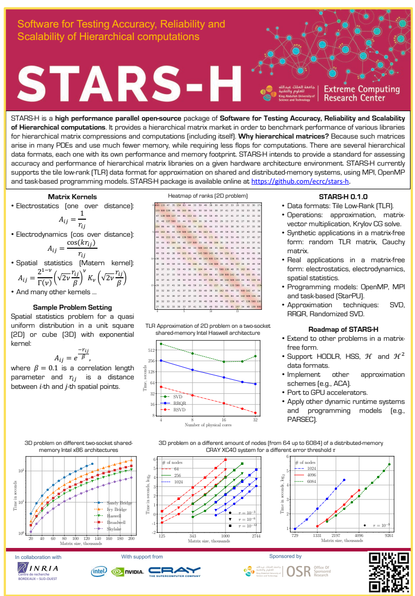


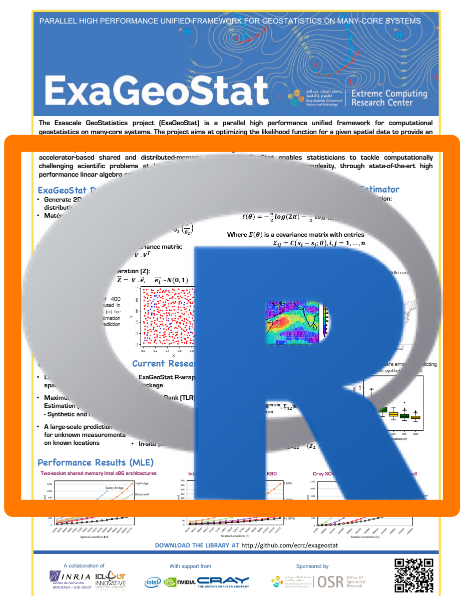
in Cray LibSci



further dev @ Intel







Taskification based on DAGs

Advantages

- remove artifactual synchronizations in the form of subroutine boundaries
- remove artifactual orderings in the form of pre-scheduled loops
- expose more concurrency

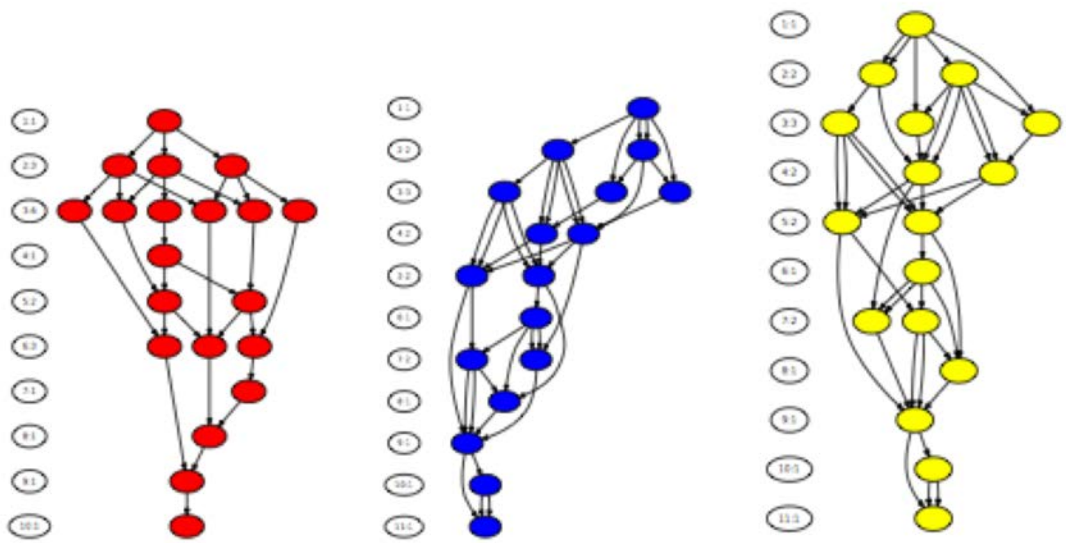
Disadvantages

- pay overhead of managing task graph
- potentially lose some memory locality

Reduce over-ordering synch w dataflow

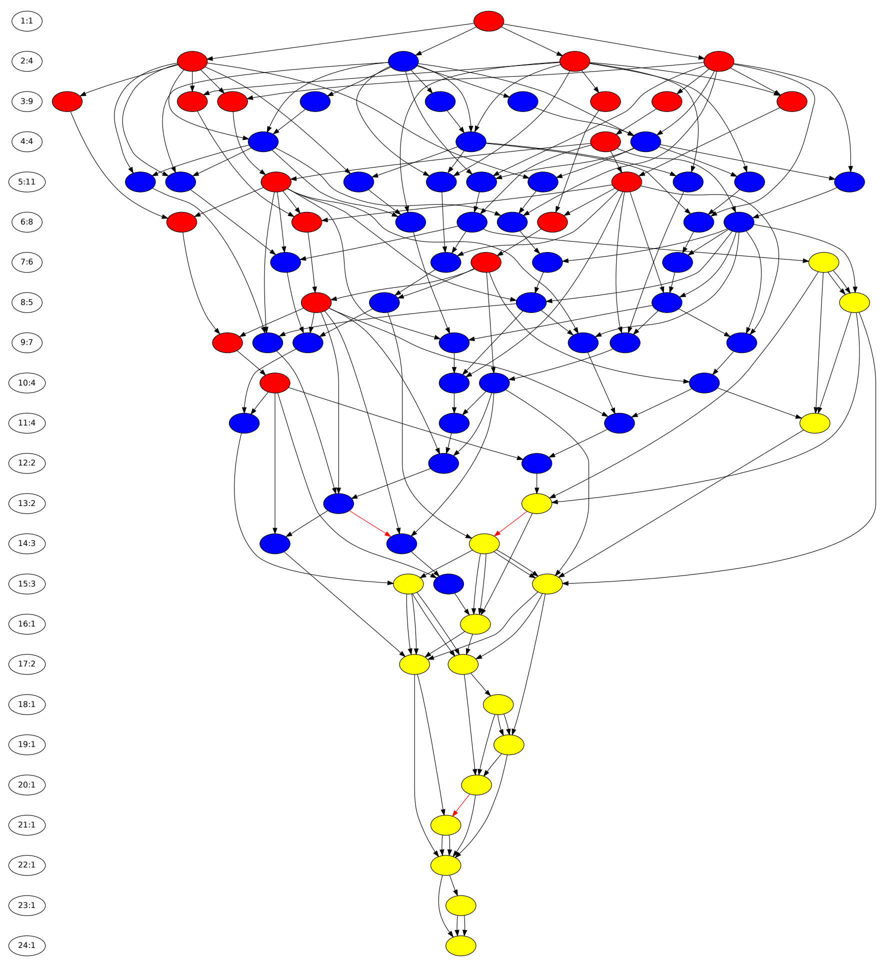
$$Ax = \lambda Bx$$

Operation	Explanation	LAPACK routine name
❶ $B = L \times L^T$	Cholesky factorization	POTRF
❷ $C = L^{-1} \times A \times L^{-T}$	application of triangular factors	SYGST or HEGST
❸ $T = Q^T \times C \times Q$	tridiagonal reduction	SYEVD or HEEVD
❹ $Tx = \lambda x$	QR iteration	STERF



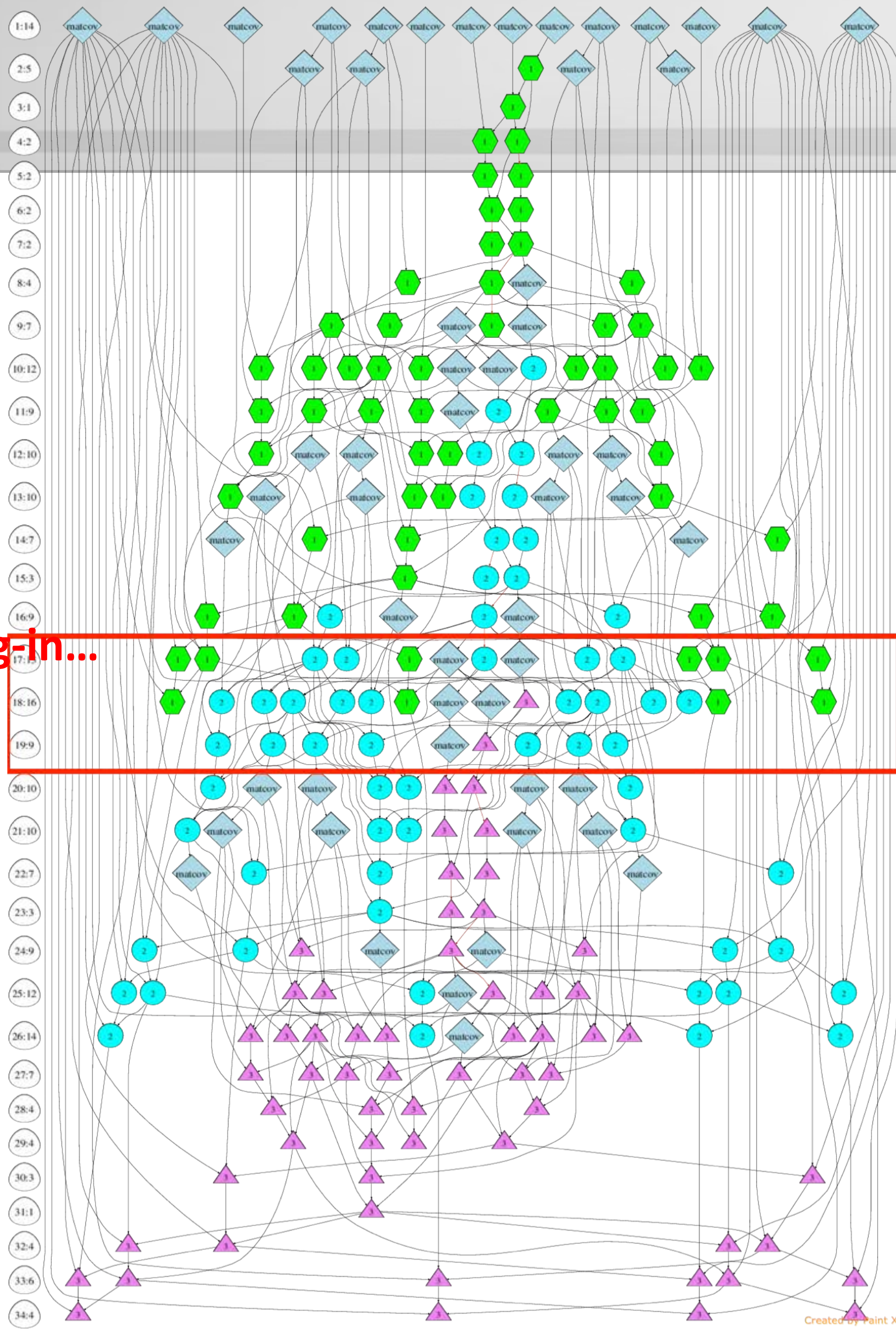
Replace loop nests & subroutine calls w DAGs

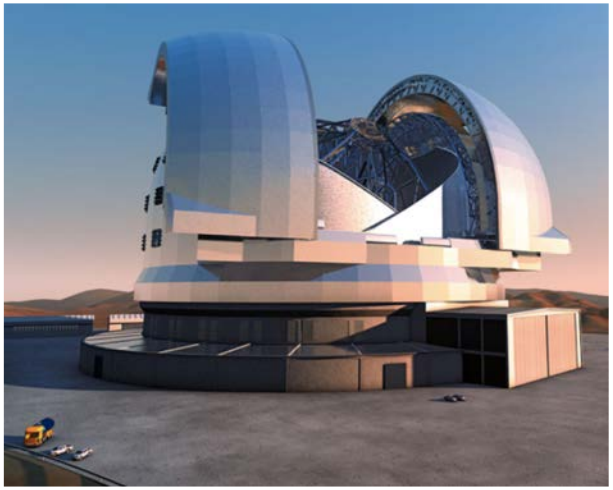
- Diagram shows a dataflow ordering of the steps of a 4 × 4 symmetric generalized eigensolver
- Nodes are tasks, color-coded by type, and edges are data dependencies
- Time is vertically downward
- Wide is good; short is good



Overlapped Loops

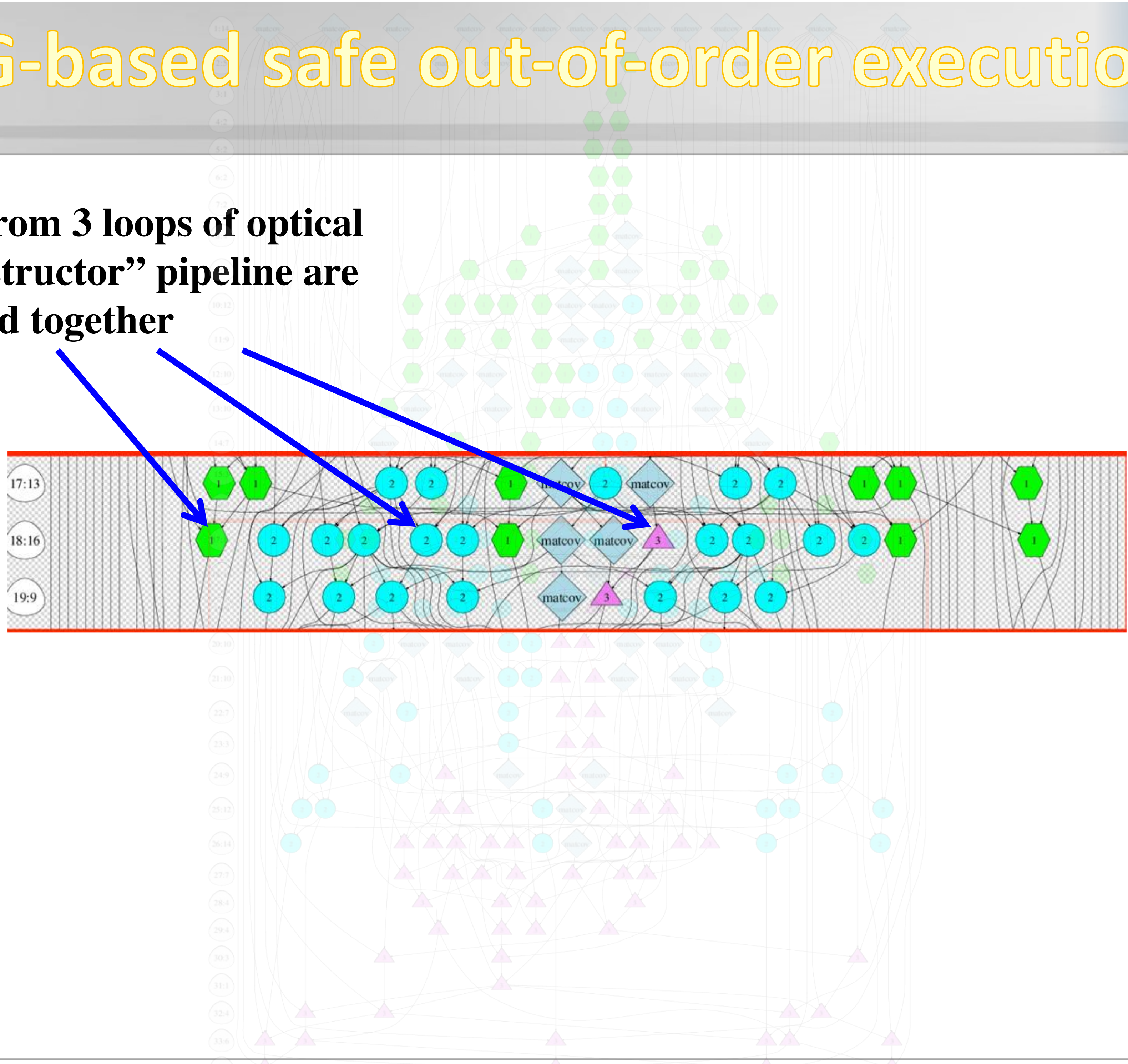
Green, blue and magenta symbols represent tasks in separate loop bodies with dependences from an adaptive optics computation





DAG-based safe out-of-order execution

Tasks from 3 loops of optical "reconstructor" pipeline are executed together



c/o H. Ltaief (KAUST) & D. Gratadour (OdP)

Hierarchically low-rank operators

Advantages

- shrink memory footprints to live higher on the memory hierarchy
  (higher means quicker access)
- reduce operation counts
- tune work to accuracy requirements
  (e.g., preconditioner versus solver)

Disadvantages

- must pay cost of compression
- not all operators compress well

Key tool: hierarchical matrices

- [Hackbusch, 1999] **: off-diagonal blocks of typical differential and integral operators have low effective rank**
- **By exploiting low rank, *k* , memory requirements and operation counts approach optimal in matrix dimension *n*:**
  - **polynomial in *k***
  - **lin-log in *n***
  - **constants carry the day**
- **Such hierarchical representations navigate a compromise**
  - **fewer blocks of larger rank ("weak admissibility") or**
  - **more blocks of smaller rank ("strong admissibility")**

Parallel universes of NLA

Flat

* Global indices *

do i {
  do j {
    for (i,j) in S do op
  }
}

Hierarchical

* Local indices *

<u>for matrix blocks (k,l)</u>

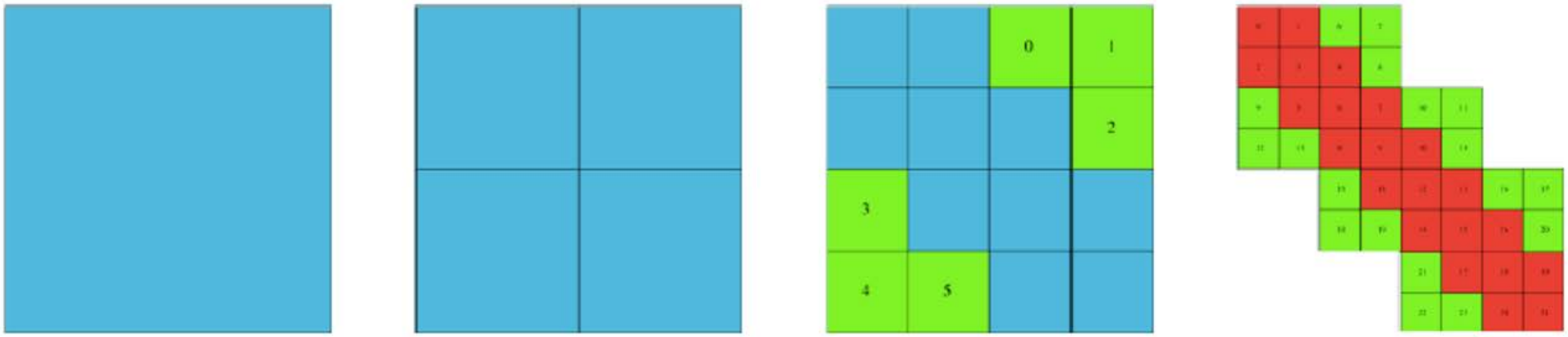
do i {
  do j {
    for (i,j) in $S_{k,l}$ do op
  }
}

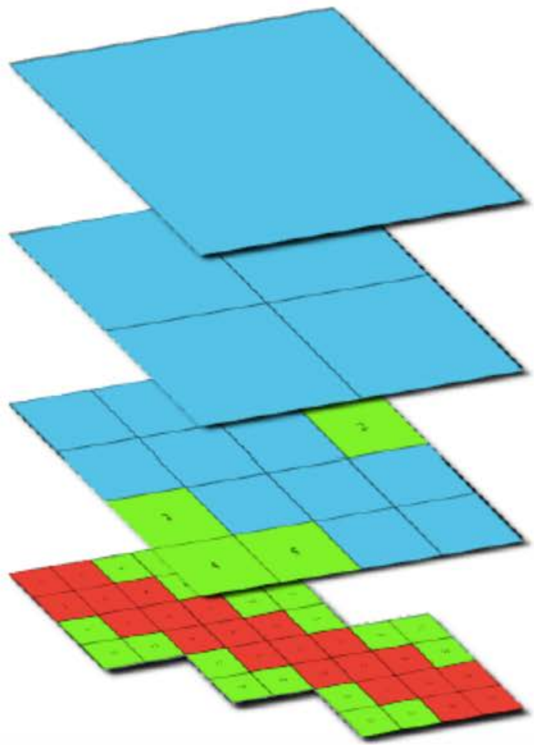
Example: 1D Laplacian

$$
A = \left[\begin{array}{ccc|cccc}
2 & -1 & & & & & \\
-1 & 2 & -1 & & & & \\
 & -1 & 2 & -1 & & & \\
\hline
 & & -1 & 2 & -1 & & \\
 & & & -1 & 2 & -1 & \\
 & & & & -1 & 2 & -1 \\
 & & & & & -1 & 2
\end{array}\right]
\qquad \longleftrightarrow \qquad
\boxed{= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 0 \end{bmatrix}}
$$

$$
A^{-1} = \frac{1}{8} \times \left[\begin{array}{ccc|cccc}
7 & 6 & 5 & 4 & 3 & 2 & 1 \\
6 & 12 & 10 & 8 & 6 & 4 & 2 \\
5 & 10 & 15 & 12 & 9 & 6 & 3 \\
\hline
4 & 8 & 12 & 16 & 12 & 8 & 4 \\
3 & 6 & 9 & 12 & 15 & 10 & 5 \\
2 & 4 & 6 & 8 & 10 & 12 & 6 \\
1 & 2 & 3 & 4 & 5 & 6 & 7
\end{array}\right]
\qquad \longleftrightarrow \qquad
\boxed{= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 4 & 3 & 2 & 1 \end{bmatrix}}
$$

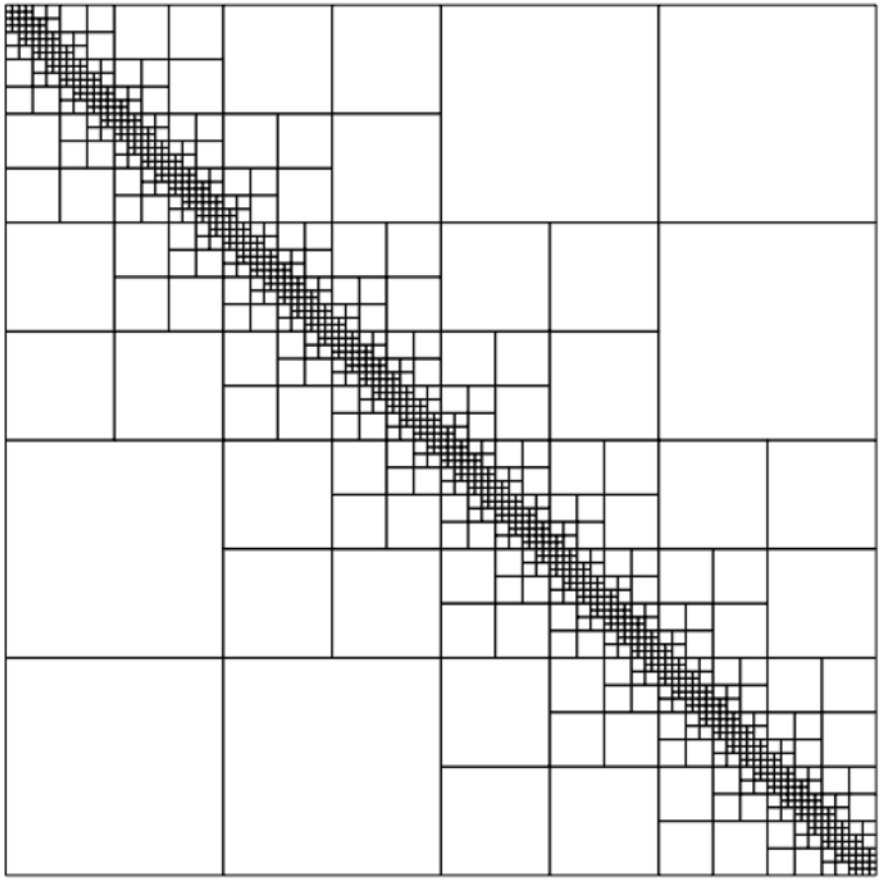
Recursive construction of an *H*-matrix



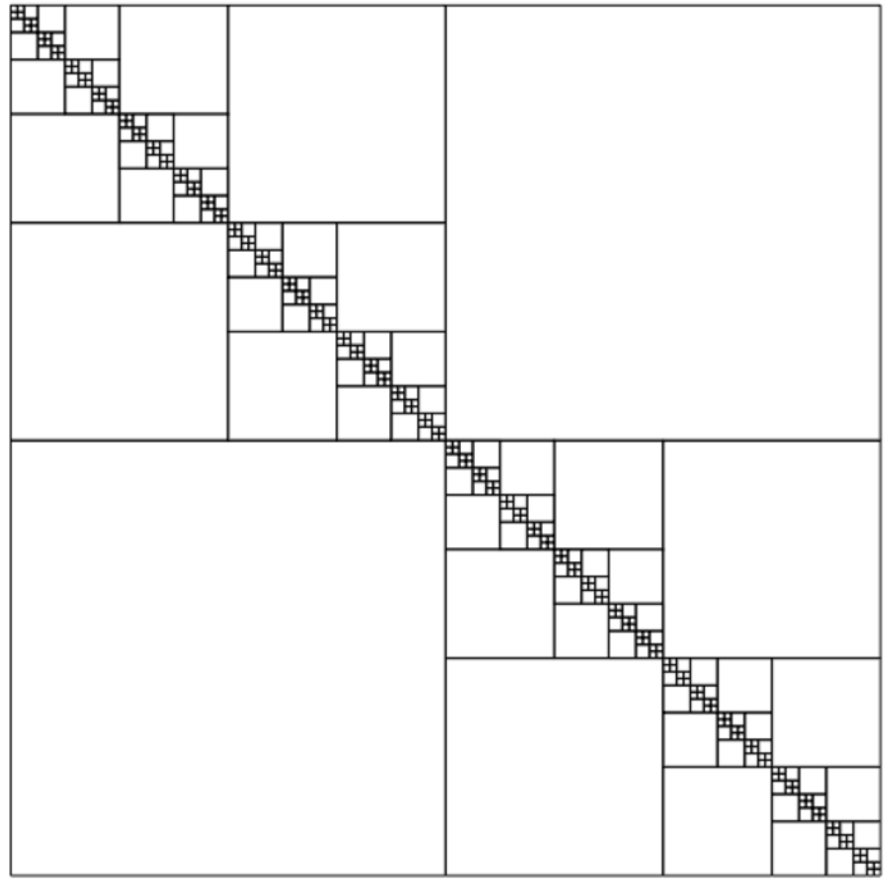


c/o W. Boukaram & G. Turkiyyah (KAUST)

"Standard (strong)" vs. "weak" admissibility



strong admissibility



weak admissibility

after Hackbusch, et al., 2003

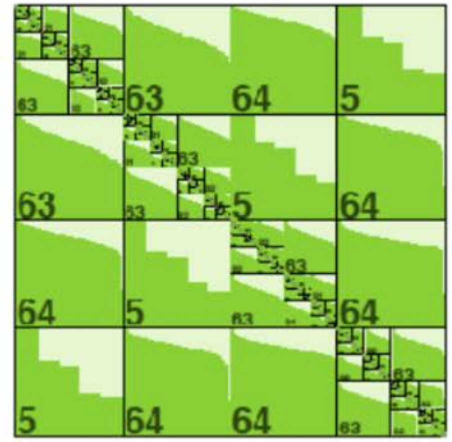
Hierarchically low-rank renaissance

Replace dense linear algebra

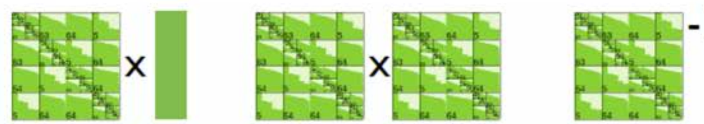
Compute : $\mathcal{O}(N^3) \longrightarrow \mathcal{O}(k^a N \log^b N)$

Memory : $\mathcal{O}(N^2) \longrightarrow \mathcal{O}(kN)$

Hierarchical off-diagonal blocks
Approximated with rank k
a and b are small constants



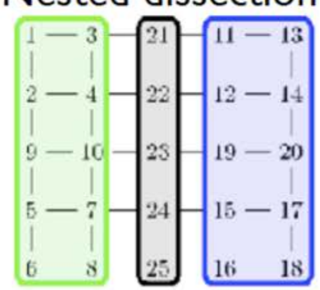
Augment sparse linear algebra

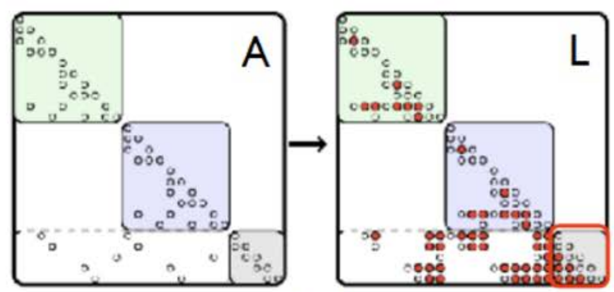


Sparse direct solvers

Schur complement (frontal matrix) is dense but numerically low-rank

Nested dissection





Schur complement

Iterative solvers

Use small k to precondition

Less sensitive to matrix condition than multigrid

Coding to the architecture

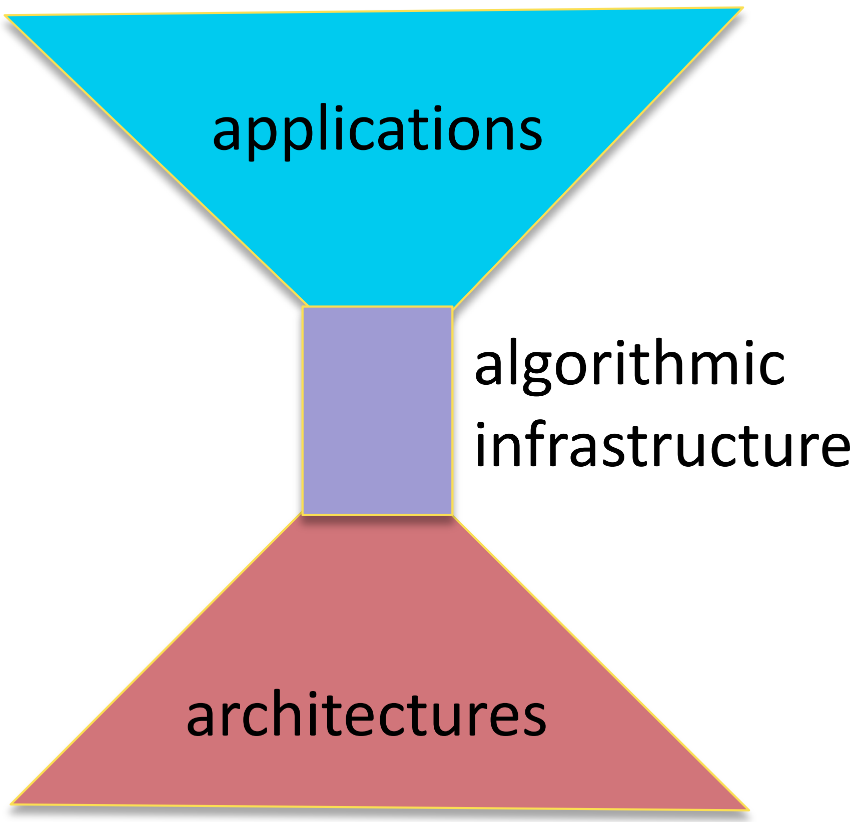
Advantages

- tiling and recursive subdivision create large numbers of small problems suitable for batched operations on GPUs and MICs
  - reduce call overheads
  - polyalgorithmic approach based on block size
- non-temporal stores, coalesced memory accesses, double-buffering, etc. reduce sensitivity to memory

Disadvantages

- code is more complex
- code is architecture-specific at the bottom

Hourglass model for algorithms



Closing observations

"Convergence" began as an architectural imperative due to market size, but flourishes as a stimulus to both simulation science and data science

However, the two distinct ecosystems require blending

In standalone modes, architectures, operations, software, and data characteristics often strongly contrast

This must be overcome since standalone mode may not be competitive

Closing observations

"Convergence" began as an architectural imperative due to market size, but flourishes as a stimulus to both simulation science and data science

However, the two distinct ecosystems require blending

In standalone modes, architectures, operations, software, and data characteristics often strongly contrast

This must be overcome since standalone mode may not be competitive

Early BDEC workshop slide

on the left

on the right

Comparing Architecture

Big Data	Extreme Computing
? Cost in memory and interconnect bandwidth	***Significant Cost*** in memory and interconnect bandwidth
Little Cost for resilient hardware in data storage	***Significant Cost*** in resilient hardware in shared file system
Little Cost for hardware to support system-wide resilience	***Significant Cost*** in resilience hardware to reduce whole-system MTTI
Significant Cost: ***increased aggregate IOPs***	Significant Cost: ***cutting-edge CPU performance features***
Often trades performance for capacity	***Often trades capacity for performance***

Comparing Operations

Big Data	Extreme Computing
Continuous access to long-lived "services" created by science community	***Periodic access*** to compute resources via job submitted to scheduler and queue
Time-shared access to elastic resources	***Space-shared*** compute resources for exclusive access during jobs
New hardware capacity ***purchased incrementally***	New tightly integrated system ***purchased every 4 years***
Users charged for all resources (storage, cpu, networking)	***Users charged for CPU hours, storage and networking is free***

Comparing Software

Big Data	Extreme Computing
Software responds to elastic resource demands	After allocation, ***resources static until termination***
Data access often ***fine-grained***	Data access is ***large bulk*** (aggregated) requests
Services are resilient to fault	***Applications restart after fault***
Often ***customized*** programming models	Widely ***standardized*** programming models
Libraries help ***move computation to storage***	Libraries help ***move data to CPUs***
Users routinely deploy their own services	***Users almost never deploy customized services***

Comparing Data

Scientific Big Data	Extreme Computing
Inputs ***arrive continuously***, streaming workflows	Inputs ***arrive infrequently***, buffering carefully managed
Data is ***unrepeatable*** snapshot in time	Data often ***reproducible*** (repeat simulation)
Data generated by sensors (***error: from measurement***)	Data generated from simulation (***error: from simulation***)
Data rate ***limited by sensors***	Data rate ***limited by platform***
Data often ***shared and curated*** by community	Data ***often private***
Often ***unstructured***	***Semi-structured***

Closing thoughts

HPC hardware technology "trickle down" benefits

- "Petascale in the machine room means terascale on the node." [Petaflops Working Group, 1990s]
- Extrapolating: exascale on the machine room floor means petascale under the desk.

HDA software technology "trickle back" benefits

- "Google is living a few years in the future and sends the rest of us messages." [Doug Cutting, Hadoop founder]

Conclusions (recap)

Many motivations exist to bring together large-scale simulation and big data analytics ("convergence")

Should be combined *in situ*

- pipelining between simulation and analytics through disk files with sequential applications leaves too many benefits "on the table"

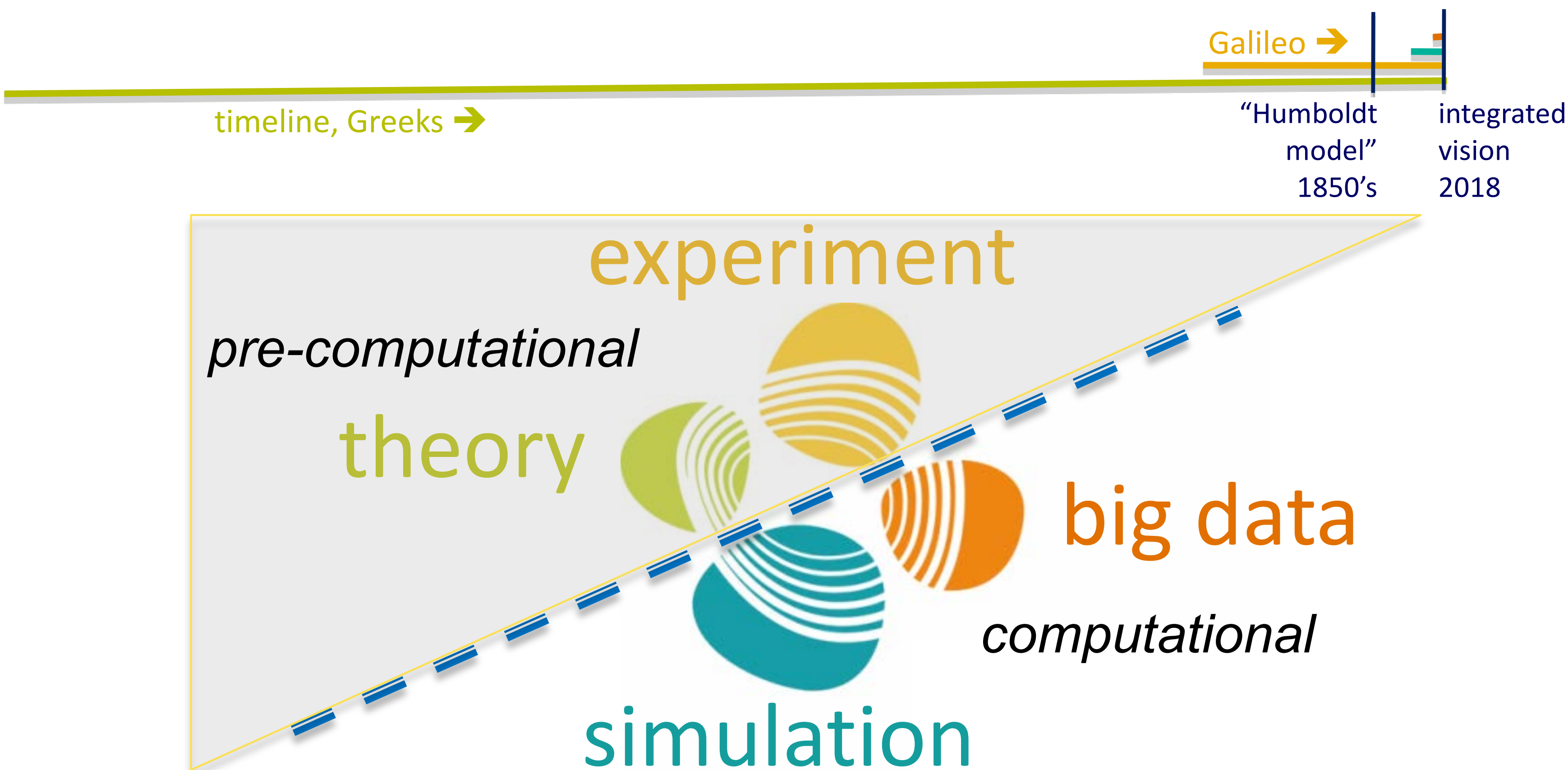
Many hurdles to convergence

- but ultimately, this will not be a "forced marriage"

Scientists and engineers may be minority users of "big data" (today and perhaps forever) but can become leaders in the "big data" community

- by harnessing high performance computing
- being pathfinders for other applications, once again!

Four ways of knowing



Conclusions as a haiku

Models from physics

Or processed observations?

Better together!


Thank you!
شكرا
جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Follow-up: community reports

http://www.exascale.org/bdec/sites/www.exascale.org.bdec/files/whitepapers/bdec2017pathways.pdf

- "Big Data and Extreme-scale Computing: Pathways to Convergence," M. Asch, *et al., Int. J. High Perf. Comput. Applics., 2018*

http://www.exascale.org/mediawiki/images/2/20/IESP-roadmap.pdf

- "The International Exascale Software Roadmap," J. Dongarra, *et al., Int. J. High Perf. Comput. Applics.,* 2011

https://arxiv.org/abs/1610.02608

- "Research and Education in Computational Science and Engineering," U. Rüde*, et al., SIAM Review, 2018*

https://arxiv.org/abs/1610.02608

- "Theory-guided Data Science," A. Karpatne*, et al., IEEE Trans. Knowledge and Data Engineering, 2017*