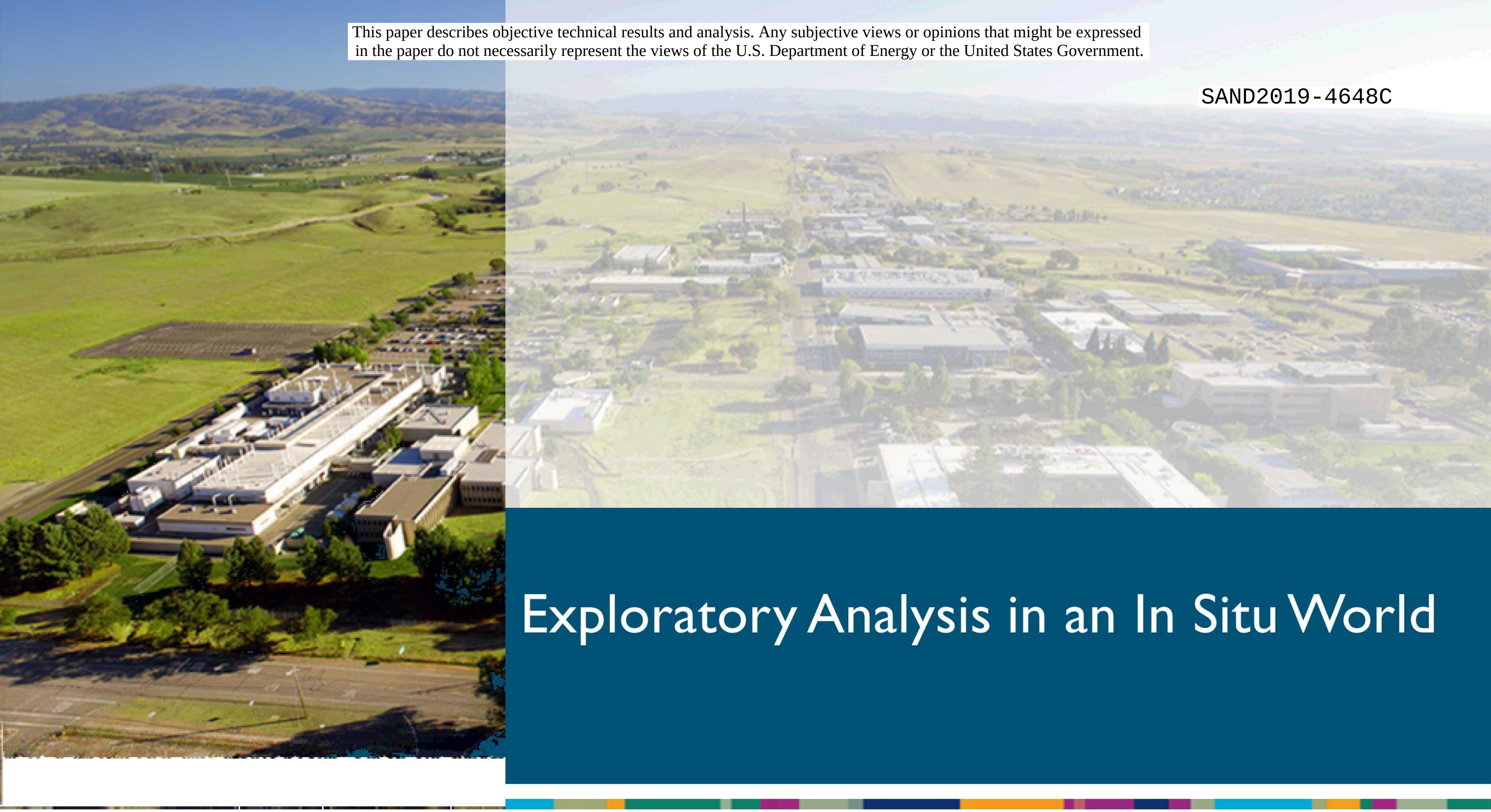
This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

SAND2019-4648C



Exploratory Analysis in an In Situ World

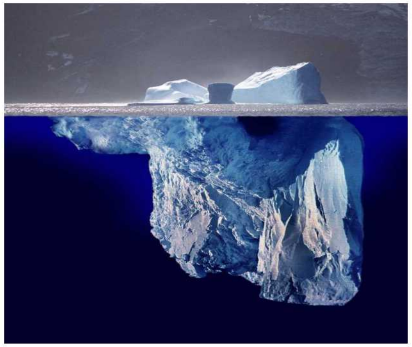
Janine C. Bennett, Sandia National Laboratories

Salishan Conference on High Speed Computing

April 25, 2019

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Bottom line up front: a single slide summary of this presentation



Exploratory analysis is a vital part of the scientific process

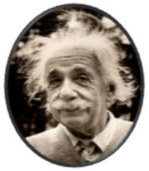
The alchemists in their search for gold discovered many other things of greater value.

Arthur Schopenhauer, German Philosopher

Challenge: Emerging HPC systems cannot output all relevant simulation data to disk

Not everything that can be counted counts, and not everything that counts can be counted.

Albert Einstein, Physicist

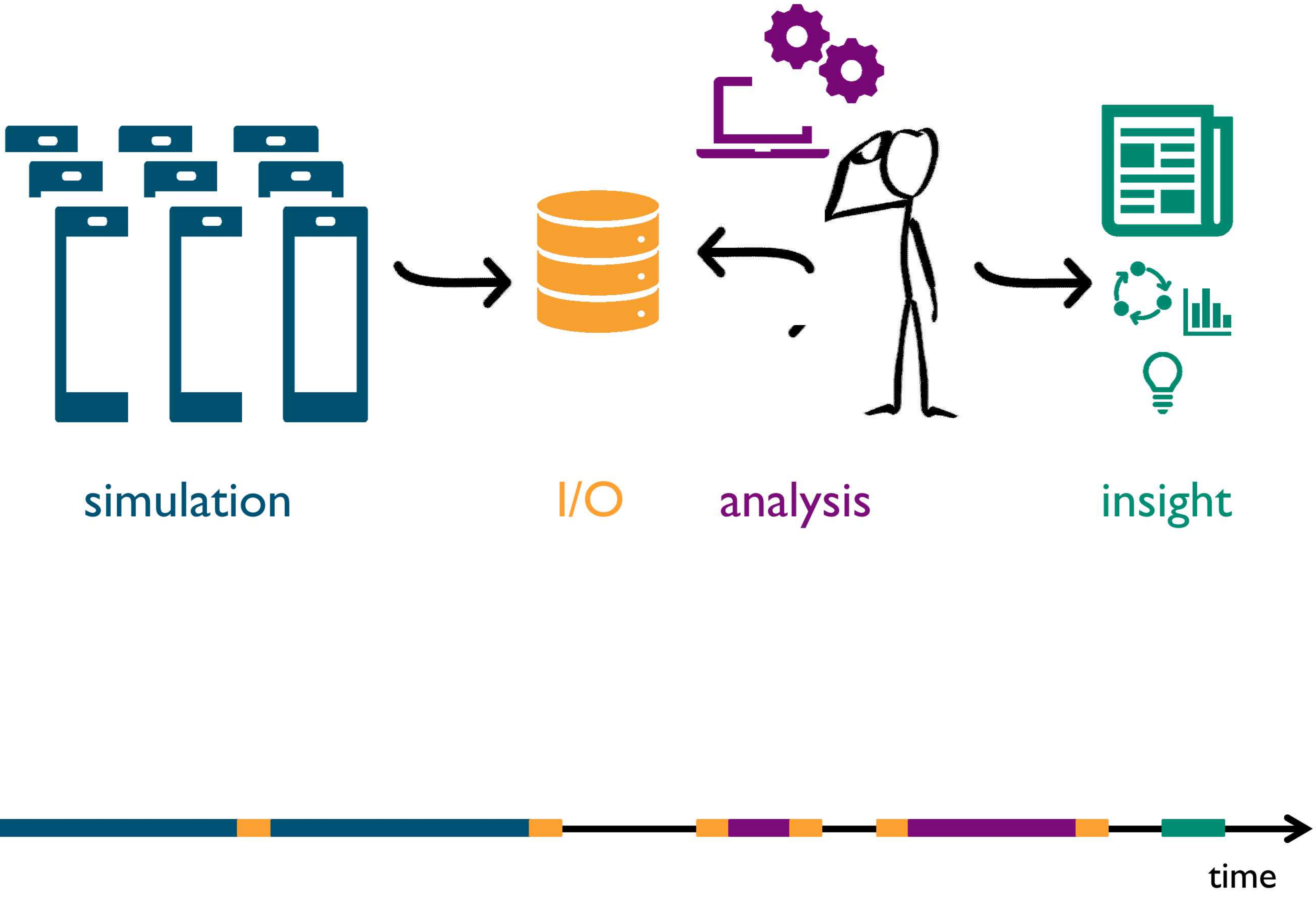


Algorithmic research will play a crucial role in maintaining our ability to perform exploratory analysis in the face of HPC system challenges

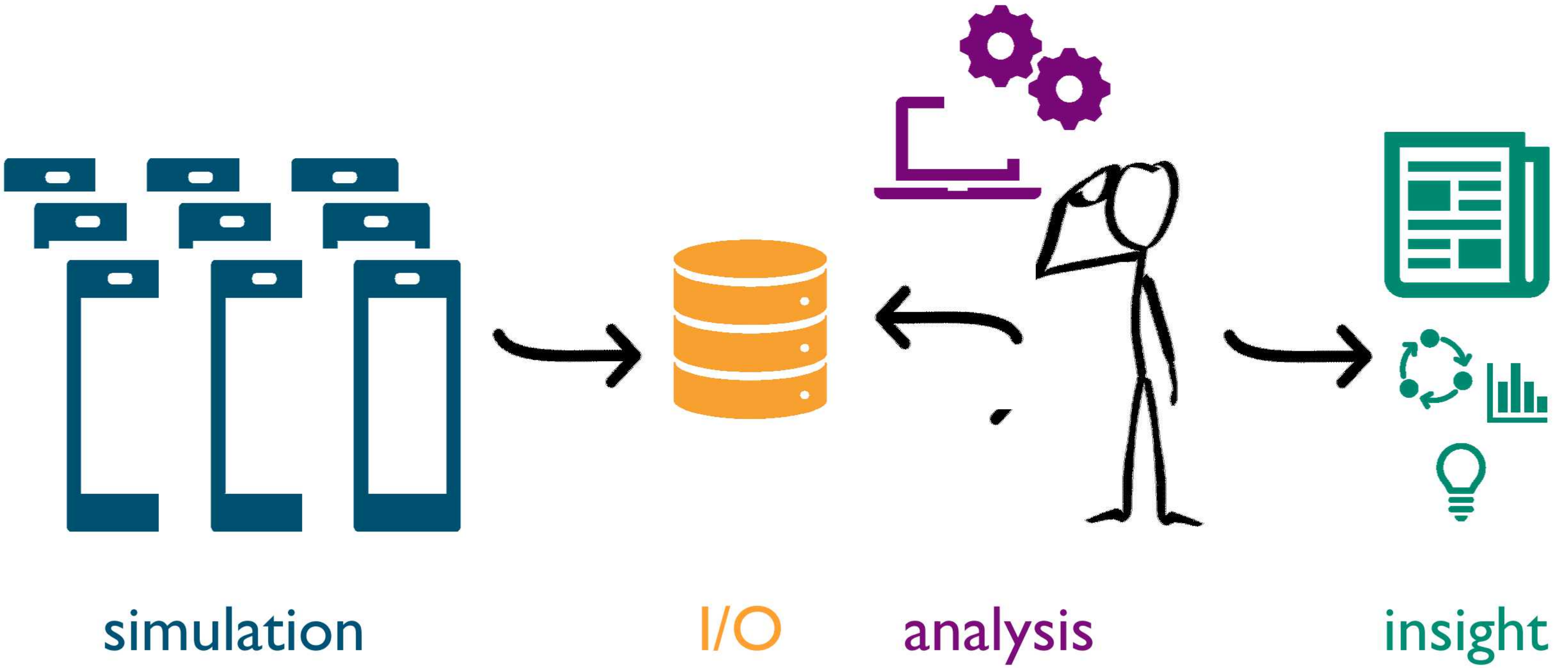
The goal is to turn data into information, and information into insight.

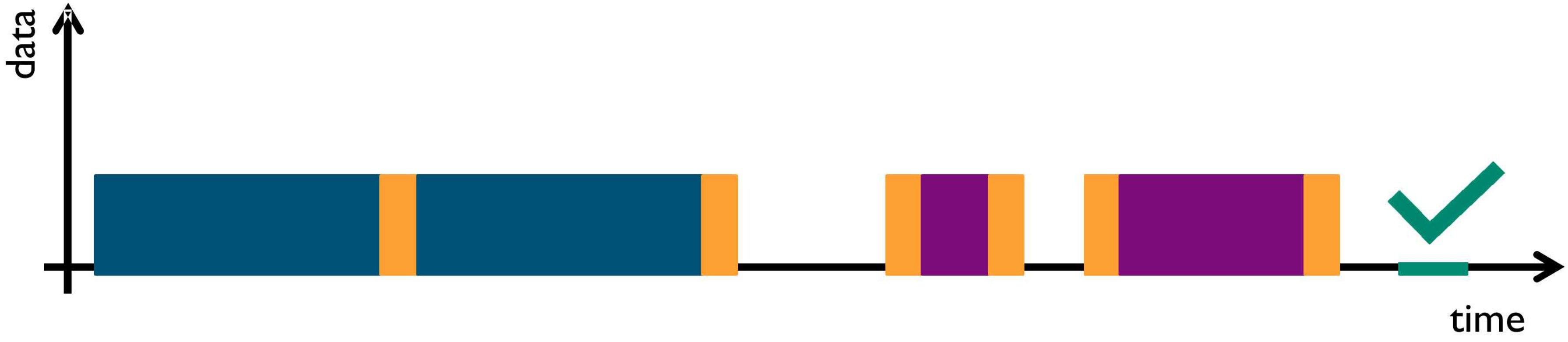
Carly Fiorina, Former CEO of HP

Exploratory analysis historically has been done post hoc - after the simulation is complete

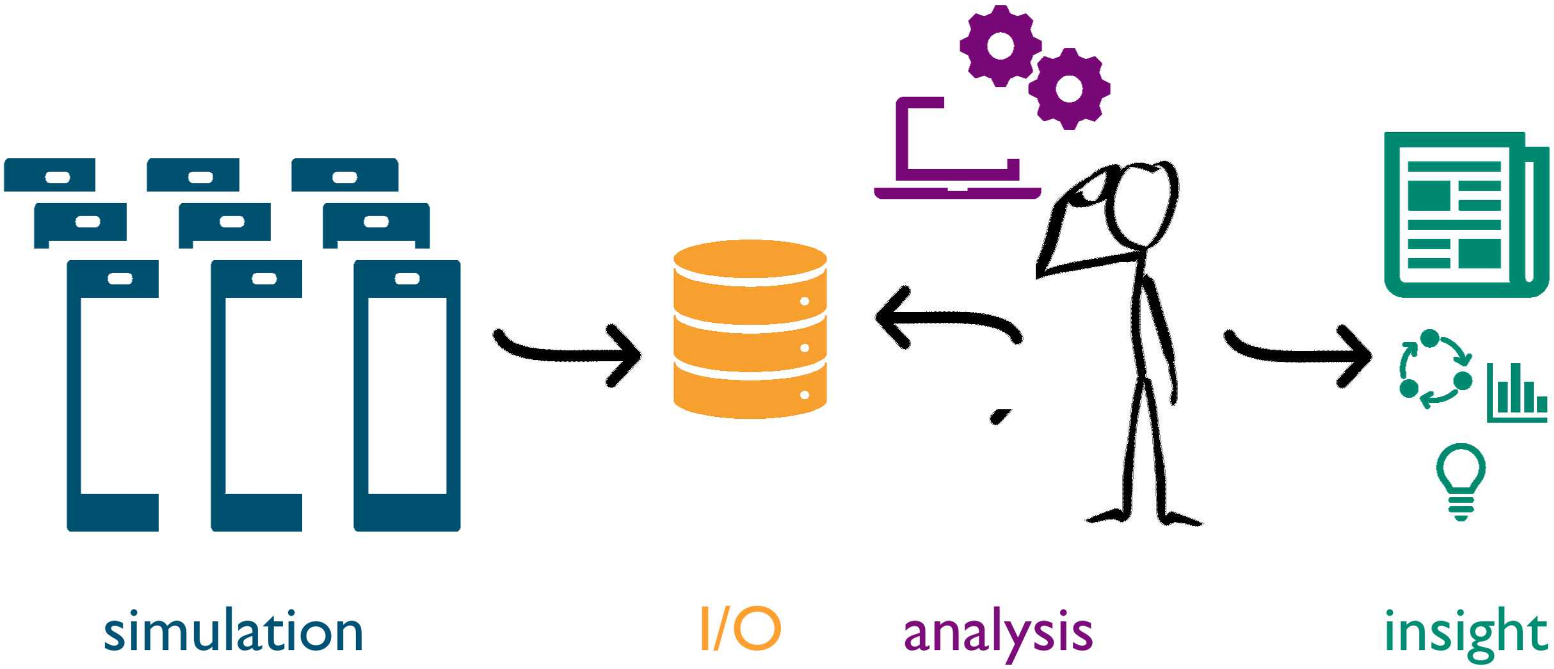


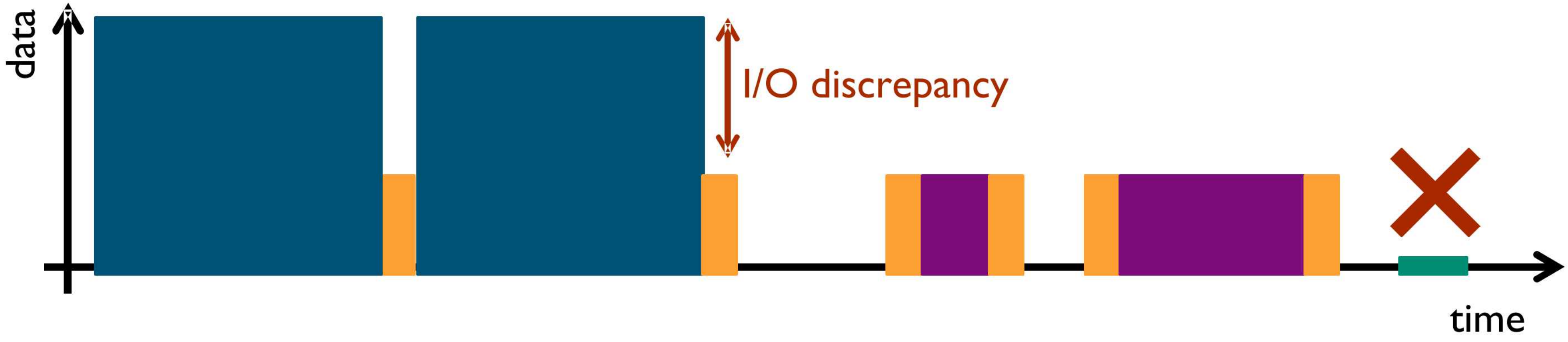
Data storage requirements for exploratory analysis could be met by previous HPC systems



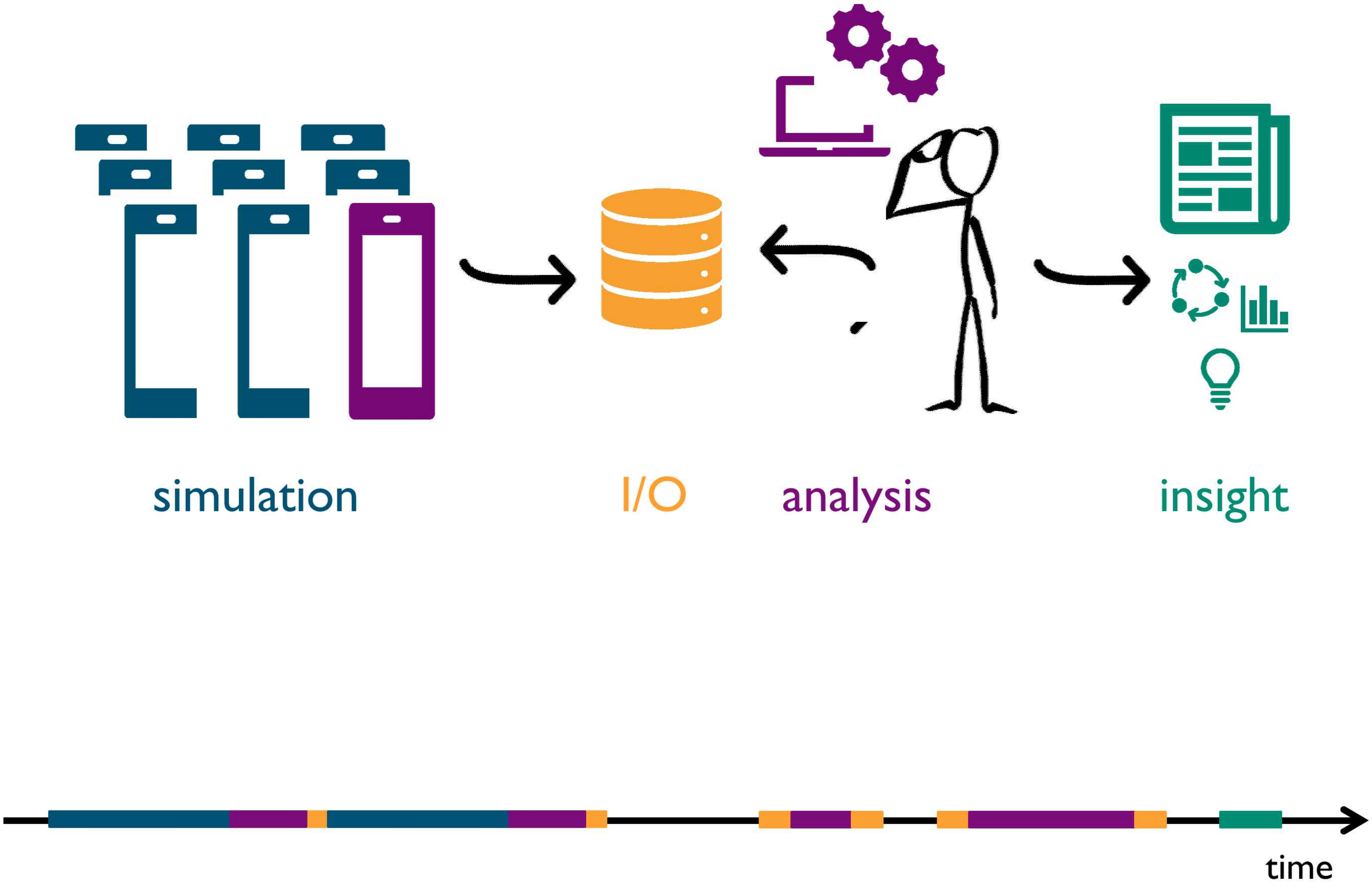


Emerging HPC systems cannot output all relevant simulation data to disk





Data I/O discrepancy is driving some of the analysis in situ - as the simulation is running

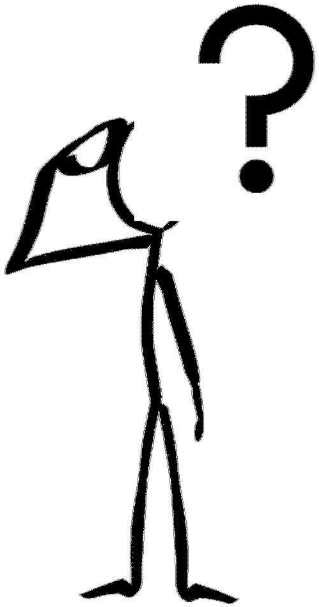


There are many algorithmic challenges posed by the shift in analysis paradigm – this presentation will explore three

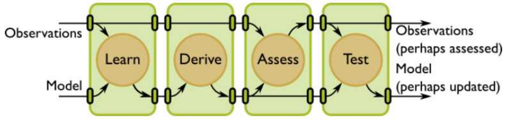
Q1: How should analysis algorithms share system resources with the application in situ?

Q2: How do analysis algorithms need to evolve to operate within a streaming regime?

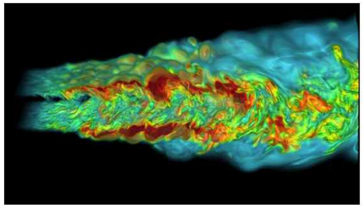
Q3: How do you enable real-time decision-making in situ with quality and reliability guarantees?



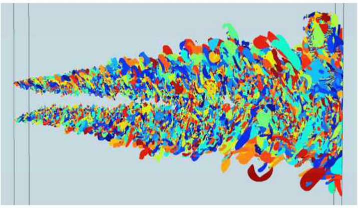
A survey categorized combustion analysis algorithms based on anticipated deployment challenges at exascale



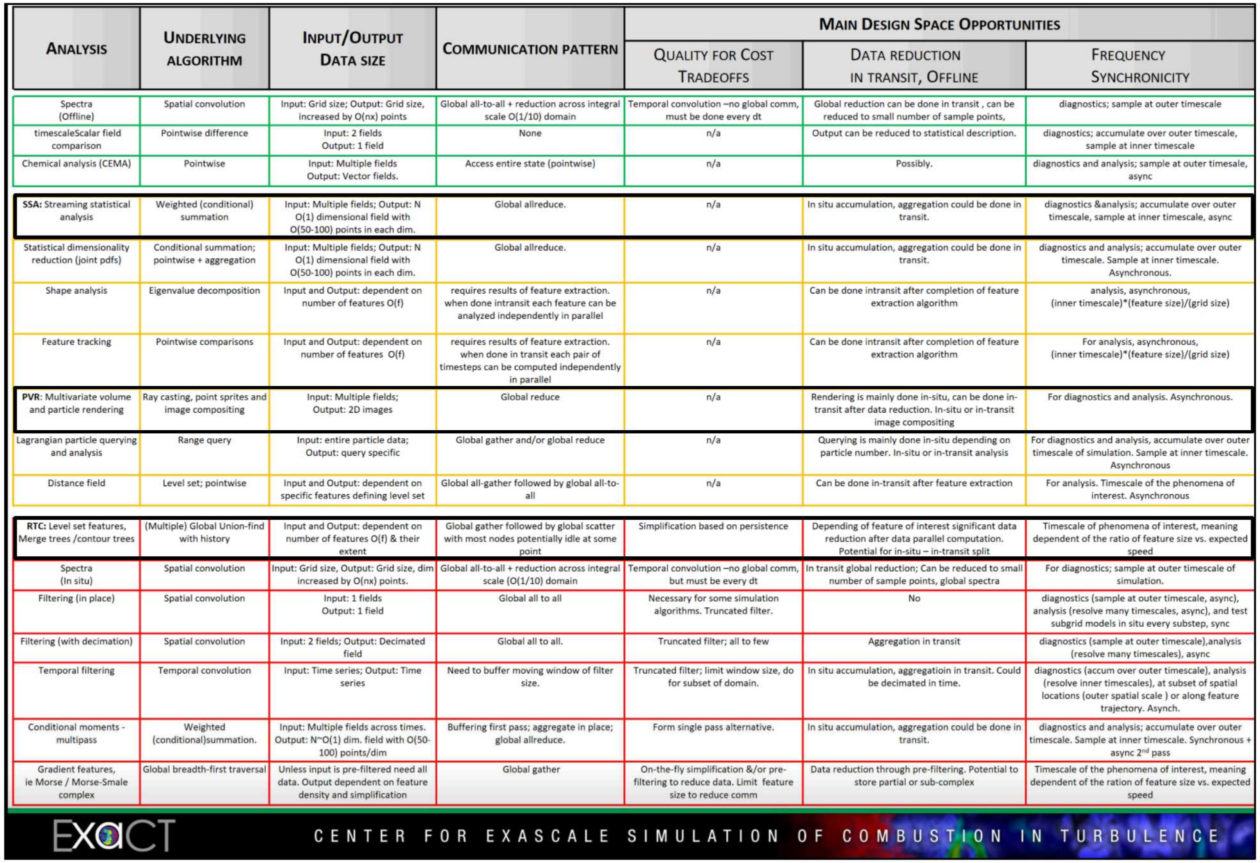
Streaming Statistics (SSA)



Volume Rendering (PVR)



Topological Analysis (RTC)

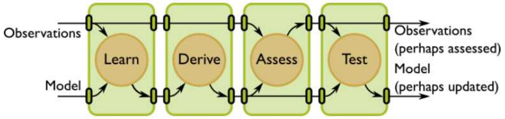


Analysis	Underlying algorithm	Input/Output Data size	Communication pattern	Main Design Space Opportunities		
				Quality for Cost Tradeoffs	Data reduction in transit, Offline	Frequency Synchronicity
Spectra (Offline)	Spatial convolution	Input: Grid size; Output: Grid size, increased by O(nx) points	Global all-to-all + reduction across integral scale O(1/10) domain	Temporal convolution –no global comm, must be done every dt	Global reduction can be done in transit , can be reduced to small number of sample points,	diagnostics; sample at outer timescale
timescaleScalar field comparison	Pointwise difference	Input: 2 fields Output: 1 field	None	n/a	Output can be reduced to statistical description.	diagnostics; accumulate over outer timescale, sample at inner timescale
Chemical analysis (CEMA)	Pointwise	Input: Multiple fields Output: Vector fields.	Access entire state (pointwise)	n/a	Possibly.	diagnostics and analysis; sample at outer timesale, async
SSA: Streaming statistical analysis	Weighted (conditional) summation	Input: Multiple fields; Output: N O(1) dimensional field with O(50-100) points in each dim.	Global allreduce.	n/a	In situ accumulation, aggregation could be done in transit.	diagnostics &analysis; accumulate over outer timescale, sample at inner timescale, async
Statistical dimensionality reduction (joint pdfs)	Conditional summation; pointwise + aggregation	Input: Multiple fields; Output: N O(1) dimensional field with O(50-100) points in each dim.	Global allreduce.	n/a	In situ accumulation, aggregation could be done in transit.	diagnostics and analysis; accumulate over outer timescale. Sample at inner timescale. Asynchronous.
Shape analysis	Eigenvalue decomposition	Input and Output: dependent on number of features O(f)	requires results of feature extraction. when done intransit each feature can be analyzed independently in parallel	n/a	Can be done intransit after completion of feature extraction algorithm	analysis, asynchronous, (inner timescale)*(feature size)/(grid size)
Feature tracking	Pointwise comparisons	Input and Output: dependent on number of features O(f)	requires results of feature extraction. when done in transit each pair of timesteps can be computed independently in parallel	n/a	Can be done intransit after completion of feature extraction algorithm	For analysis, asynchronous, (inner timescale)*(feature size)/(grid size)
PVR: Multivariate volume and particle rendering	Ray casting, point sprites and image compositing	Input: Multiple fields; Output: 2D images	Global reduce	n/a	Rendering is mainly done in-situ, can be done in-transit after data reduction. In-situ or in-transit image compositing	For diagnostics and analysis. Asynchronous.
Lagrangian particle querying and analysis	Range query	Input: entire particle data; Output: query specific	Global gather and/or global reduce	n/a	Querying is mainly done in-situ depending on particle number. In-situ or in-transit analysis	For diagnostics and analysis, accumulate over outer timescale of simulation. Sample at inner timescale. Asynchronous
Distance field	Level set; pointwise	Input and Output: dependent on specific features defining level set	Global all-gather followed by global all-to-all	n/a	Can be done in-transit after feature extraction	For analysis. Timescale of the phenomena of interest. Asynchronous
RTC: Level set features, Merge trees /contour trees	(Multiple) Global Union-find with history	Input and Output: dependent on number of features O(f) & their extent	Global gather followed by global scatter with most nodes potentially idle at some point	Simplification based on persistence	Depending of feature of interest significant data reduction after data parallel computation. Potential for in-situ – in-transit split	Timescale of phenomena of interest, meaning dependent of the ratio of feature size vs. expected speed
Spectra (In situ)	Spatial convolution	Input: Grid size, Output: Grid size, dim increased by O(nx) points.	Global all-to-all + reduction across integral scale (O(1/10) domain	Temporal convolution –no global comm, but must be every dt	In transit global reduction; Can be reduced to small number of sample points, global spectra	For diagnostics; sample at outer timescale of simulation.
Filtering (in place)	Spatial convolution	Input: 1 fields Output: 1 field	Global all to all	Necessary for some simulation algorithms. Truncated filter.	No	diagnostics (sample at outer timescale, async), analysis (resolve many timescales, async), and test subgrid models in situ every substep, sync
Filtering (with decimation)	Spatial convolution	Input: 2 fields; Output: Decimated field	Global all to all.	Truncated filter; all to few	Aggregation in transit	diagnostics (sample at outer timescale),analysis (resolve many timescales), async
Temporal filtering	Temporal convolution	Input: Time series; Output: Time series	Need to buffer moving window of filter size.	Truncated filter; limit window size, do for subset of domain.	In situ accumulation, aggregatioin in transit. Could be decimated in time.	diagnostics (accum over outer timescale), analysis (resolve inner timescales), at subset of spatial locations (outer spatial scale) or along feature trajectory. Asynch.
Conditional moments - multipass	Weighted (conditional)summation.	Input: Multiple fields across times. Output: N~O(1) dim. field with O(50-100) points/dim	Buffering first pass; aggregate in place; global allreduce.	Form single pass alternative.	In situ accumulation, aggregation could be done in transit.	diagnostics and analysis; accumulate over outer timescale. Sample at inner timescale. Synchronous + async 2nd pass
Gradient features, ie Morse / Morse-Smale complex	Global breadth-first traversal	Unless input is pre-filtered need all data. Output dependent on feature density and simplification	Global gather	On-the-fly simplification &/or pre-filtering to reduce data. Limit feature size to reduce comm	Data reduction through pre-filtering. Potential to store partial or sub-complex	Timescale of the phenomena of interest, meaning dependent of the ration of feature size vs. expected speed

ExaCT CENTER FOR EXASCALE SIMULATION OF COMBUSTION IN TURBULENCE

Q1: How should analysis algorithms share system resources with the application in situ?

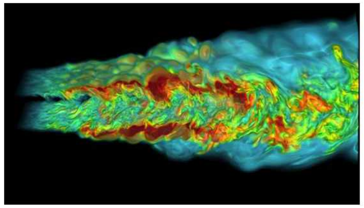
Three algorithms selected for deeper analysis spanned different analysis use cases



Streaming Statistics (SSA)

Quantitative summary of global trends in the data

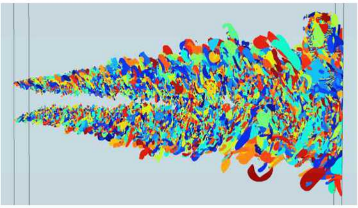
Debugging and analysis



Volume Rendering (PVR)

Qualitative visual depiction of data

Debugging and analysis



Topological Analysis (RTC)

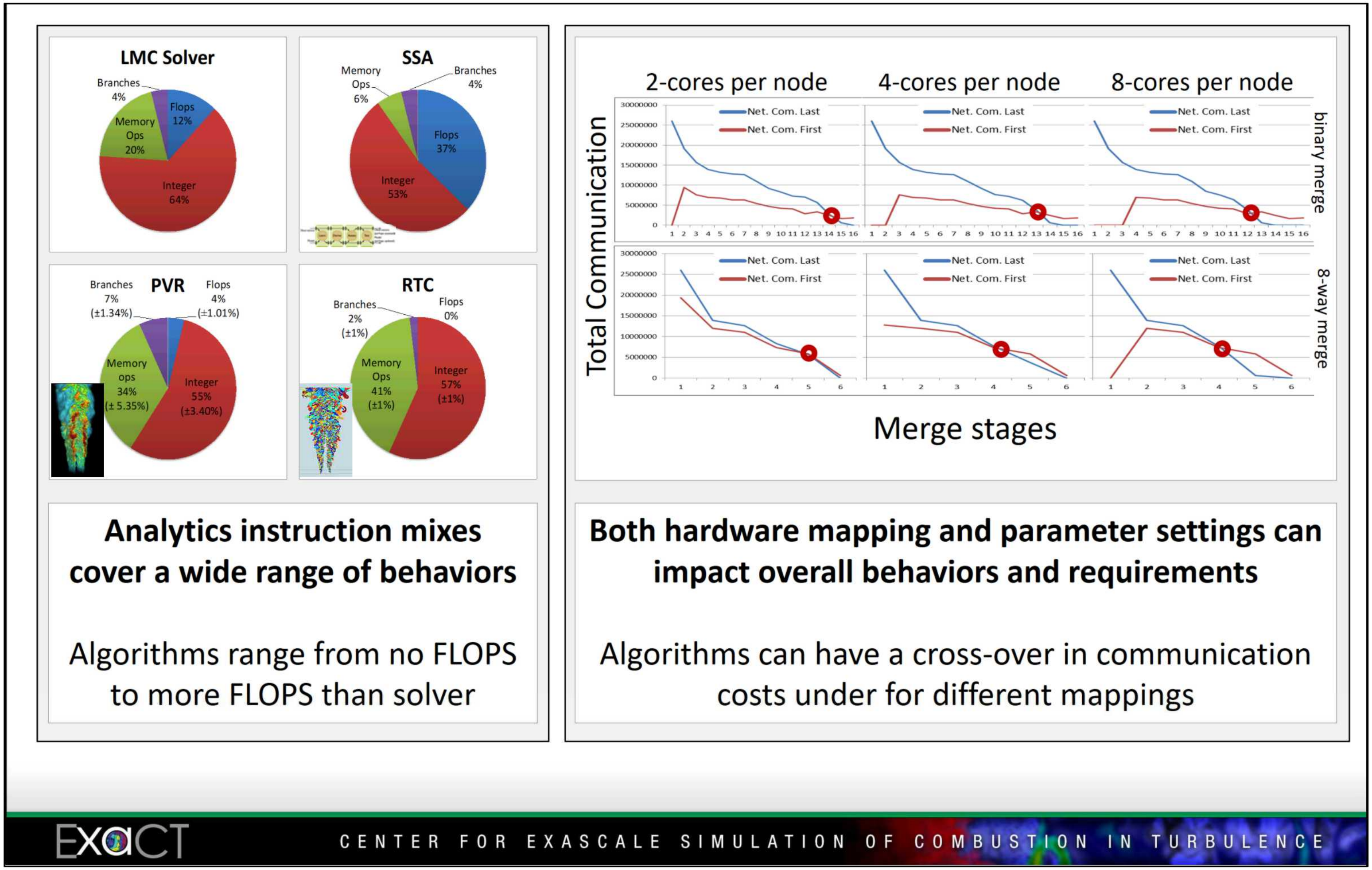
Data reduction technique enabling exploratory analysis post hoc

A complete characterization of the level-set behavior of simulation variable

Quantitative & qualitative tool to define features of interest

Q1: How should analysis algorithms share system resources with the application in situ?

Deeper investigation into the three algorithms found they had heterogenous behaviors and requirements



Analytics instruction mixes cover a wide range of behaviors

Algorithms range from no FLOPS to more FLOPS than solver

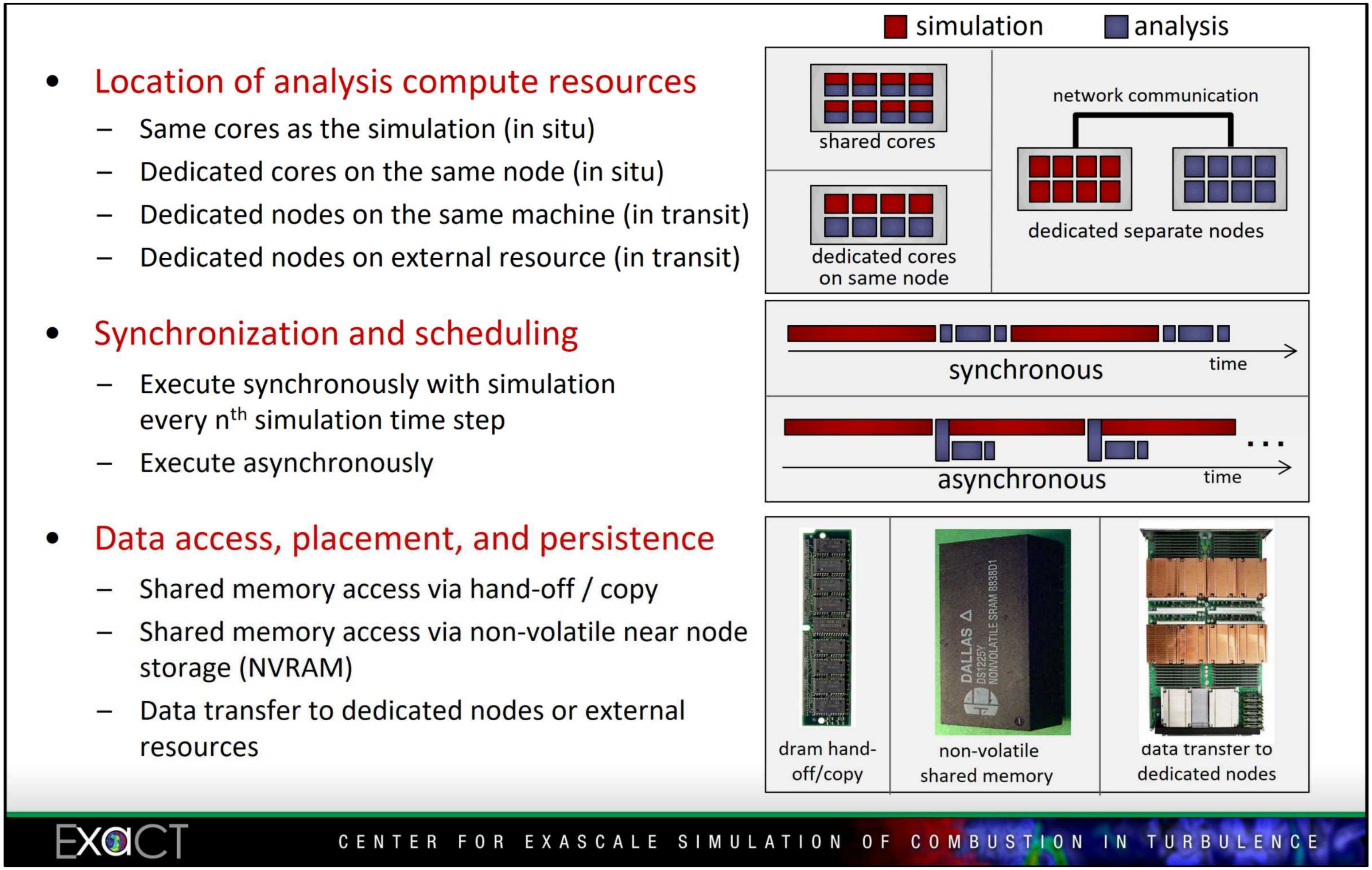
Both hardware mapping and parameter settings can impact overall behaviors and requirements

Algorithms can have a cross-over in communication costs under for different mappings

ExaCT CENTER FOR EXASCALE SIMULATION OF COMBUSTION IN TURBULENCE

Q1: How should analysis algorithms share system resources with the application in situ?

An empirical study mapping analysis tasks to hardware was considered along three axes



- Location of analysis compute resources
  - Same cores as the simulation (in situ)
  - Dedicated cores on the same node (in situ)
  - Dedicated nodes on the same machine (in transit)
  - Dedicated nodes on external resource (in transit)
- Synchronization and scheduling
  - Execute synchronously with simulation every nth simulation time step
  - Execute asynchronously
- Data access, placement, and persistence
  - Shared memory access via hand-off / copy
  - Shared memory access via non-volatile near node storage (NVRAM)
  - Data transfer to dedicated nodes or external resources

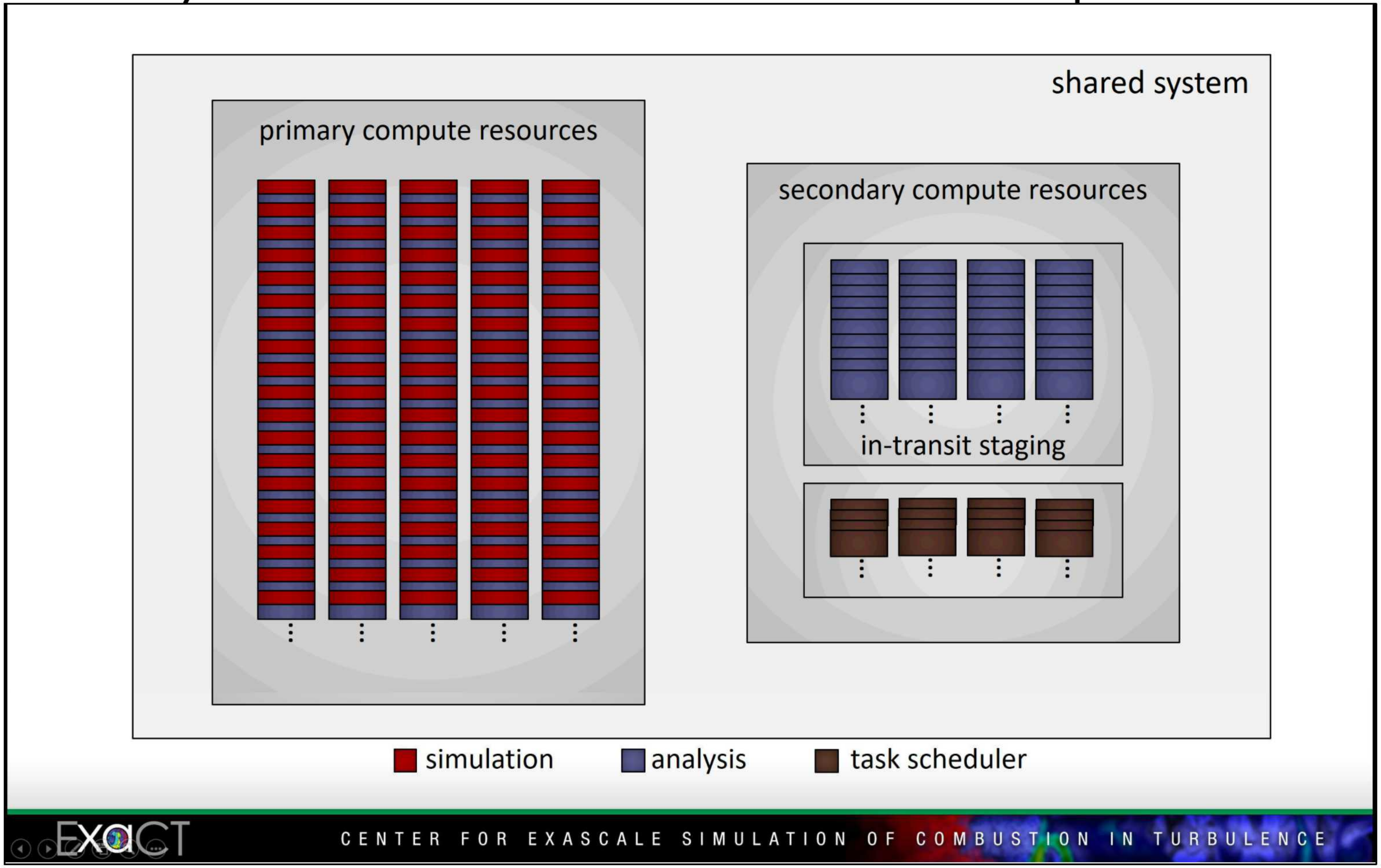
EXaCT CENTER FOR EXASCALE SIMULATION OF COMBUSTION IN TURBULENCE

[Combining In-situ and In-transit Processing to Enable Extreme-Scale Scientific Analysis](#)

Q1: How should analysis algorithms share system resources with the application in situ?

In situ and hybrid in situ + in transit variants of each algorithm were implemented and deployed

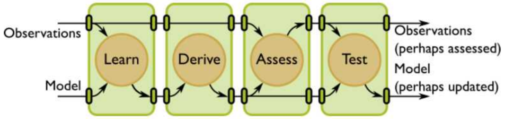
Primary resources: execute main computation and in situ computations
Secondary resources: task scheduler and in transit computations


shared system
primary compute resources
secondary compute resources
in-transit staging
🟥 simulation
🟦 analysis
🟫 task scheduler
ExaCT
CENTER FOR EXASCALE SIMULATION OF COMBUSTION IN TURBULENCE

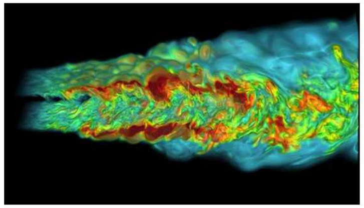
[Combining In-situ and In-transit Processing to Enable Extreme-Scale Scientific Analysis](#)

Q1: How should analysis algorithms share system resources with the application in situ?

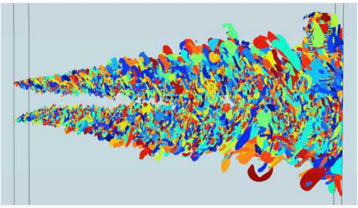
A hybrid approach to workflow mapping was found to minimize impact to the simulation



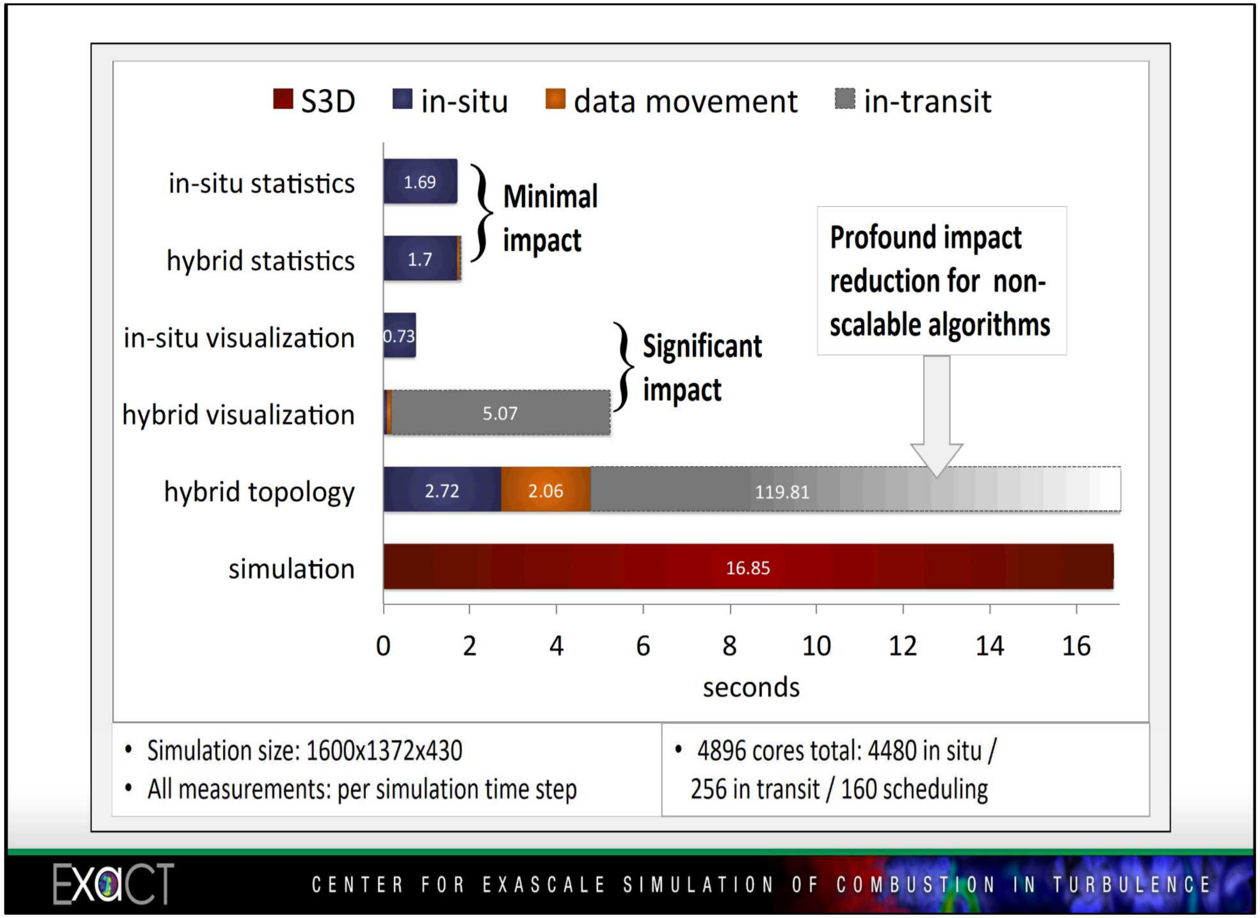
Streaming statistics (SSA)



Volume Rendering (PVR)



Topological Analysis (RTC)



Minimal impact

Significant impact

Profound impact reduction for non-scalable algorithms

- Simulation size: 1600x1372x430
- All measurements: per simulation time step

- 4896 cores total: 4480 in situ / 256 in transit / 160 scheduling

EXaCT CENTER FOR EXASCALE SIMULATION OF COMBUSTION IN TURBULENCE

Combining In-situ and In-transit Processing to Enable Extreme-Scale Scientific Analysis

Q1: How should analysis algorithms share system resources with the application in situ?

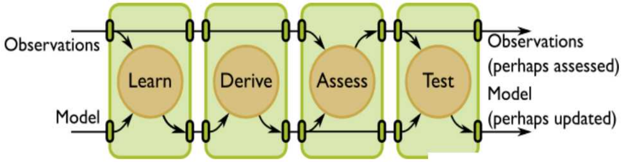
Streaming statistics leveraged previous algorithmic research - making it amenable for deployment in situ

Single-pass, numerically stable online update formulas for arbitrary-order moments

$$
\begin{aligned}
M_{p,\mathscr{S}} = {} & M_{p,\mathscr{S}_1} + M_{p,\mathscr{S}_2} \\
& + \sum_{k=1}^{p-2} \binom{k}{p} \left[\left(-\frac{n_2}{n} \right)^k M_{p-k,\mathscr{S}_1} + \left(\frac{n_1}{n} \right)^k M_{p-k,\mathscr{S}_2} \right] \delta_{2,1}^k \\
& + \left(\frac{n_1 n_2}{n} \delta_{2,1} \right)^p \left[\frac{1}{n_2^{p-1}} - \left(\frac{-1}{n_1} \right)^{p-1} \right]. \qquad \text{(III.1)}
\end{aligned}
$$

Serve as a basis for many statistics algorithms

- Correlative, Multi-Correlative Statistics
- Principal Component Analysis
- K-Means Clustering
- Sobol Indices for Sensitivity Analysis (work @ INRIA, France)



New ECP activity implementing streaming statistics within the ECP ecosystem of tools

- ASCENT (in situ library)
- VTK-M (vis kernels for emerging processor architectures)

Melissa: Large Scale In Transit Sensitivity Analysis Avoiding Intermediate Files

Design and performance of a scalable, parallel statistics toolkit

Numerically stable, single-pass, parallel statistics algorithms

Numerically stable, scalable formulas for parallel online computation of higher-order multivariate central moments with arbitrary weights

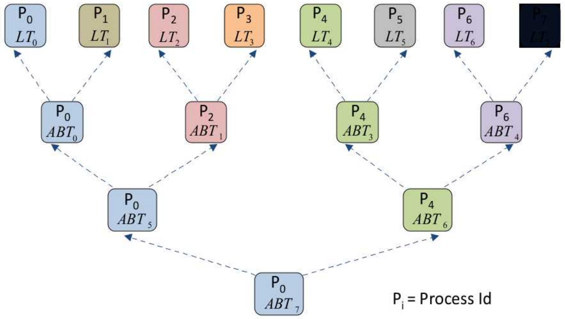
Q2: How do analysis algorithms need to evolve to operate within a streaming regime?

The ExaCT empirical study led to multi-fidelity variants of the topological analysis tools for deployment in situ

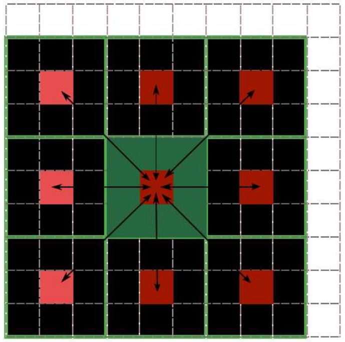
Science use cases highlighted that approximate solution was often sufficient

Refinement of requirements for streaming regime

K-way merge: Generates an exact global solution but has limited scalability



Region Growing: Scalable, local approximation that guarantees to correctly extract all features up to a predefined size



In situ Feature Extraction of Large-Scale Combustion Simulations Using Segmented Merge Trees

Q2: How do analysis algorithms need to evolve to operate within a streaming regime?

Many of the early in situ analysis tools supported analysis at prescribed frequencies

At what frequency should I/O or analysis be done?

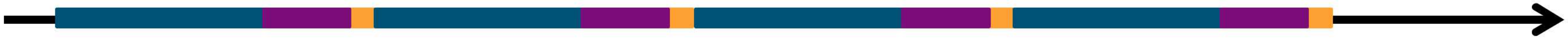
Can this decision be made in an adaptive, data-driven fashion at run-time?

- Avoid missing interesting science
- Avoid expensive analysis and I/O when simulation state is evolving slowly

Indicator: lightweight analysis to be deployed at high frequency
Trigger: return true when the indicator has met a specific property

CHALLENGE: How do we define indicators and triggers that

1) capture scientific phenomena of interest; and
2) are cost efficient enough to be deployed at high frequency



Key: simulation analysis I/O

A combustion use case highlights the challenges in developing cost-efficient triggers and indicators

Homogenous Compression Charge Ignition (HCCI)

Many small heat kernels develop slowly prior to ignition

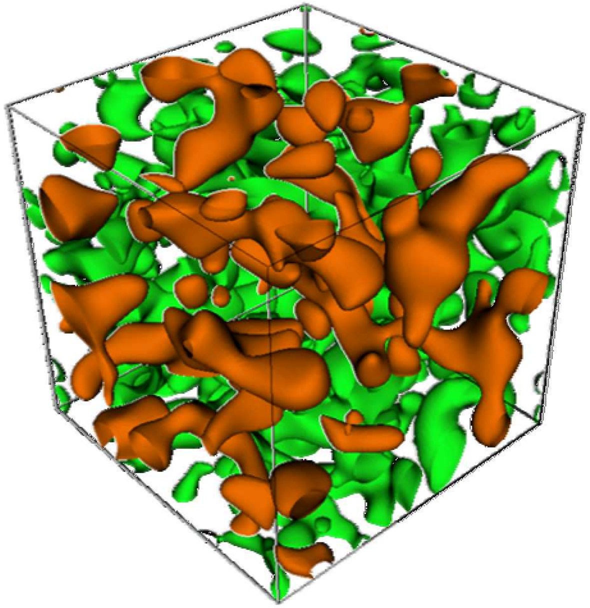
Scientists we were working with wanted:

- Coarse grid and less frequent I/O before "things get interesting"
- Refined grid and more frequent I/O afterwards

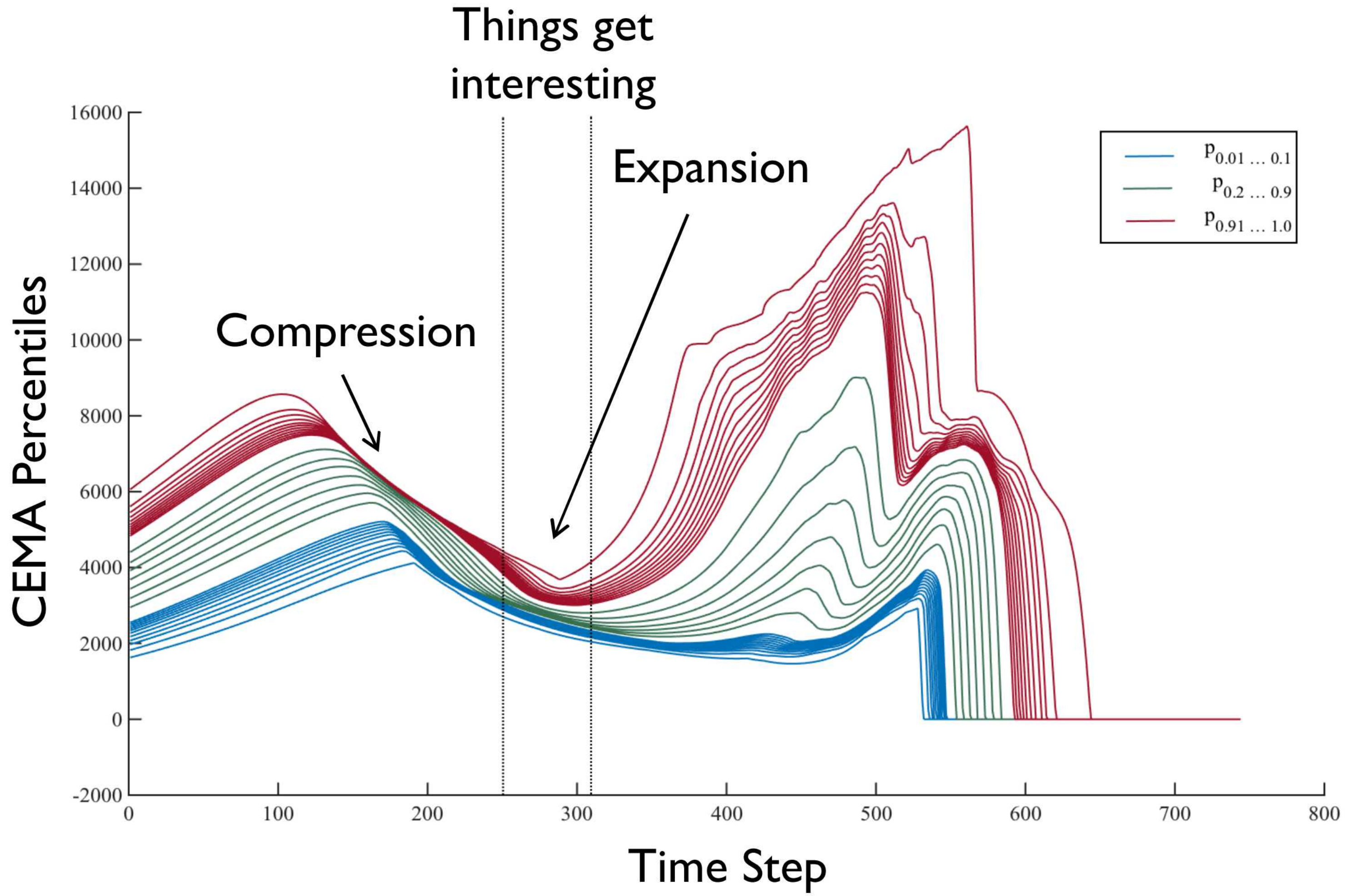
Chemical Explosive Mode Analysis (CEMA)

- Good predictor of heat release
- It tells you when "things get interesting"
- Point-wise Jacobian of chemical species
- Cost depends on complexity of simulation

In recent simulations, a full CEMA cost up to 60 times the cost of a simulation time step



“Things get interesting” when there is a compression followed by an expansion in CEMA percentile values



P_0 = min value

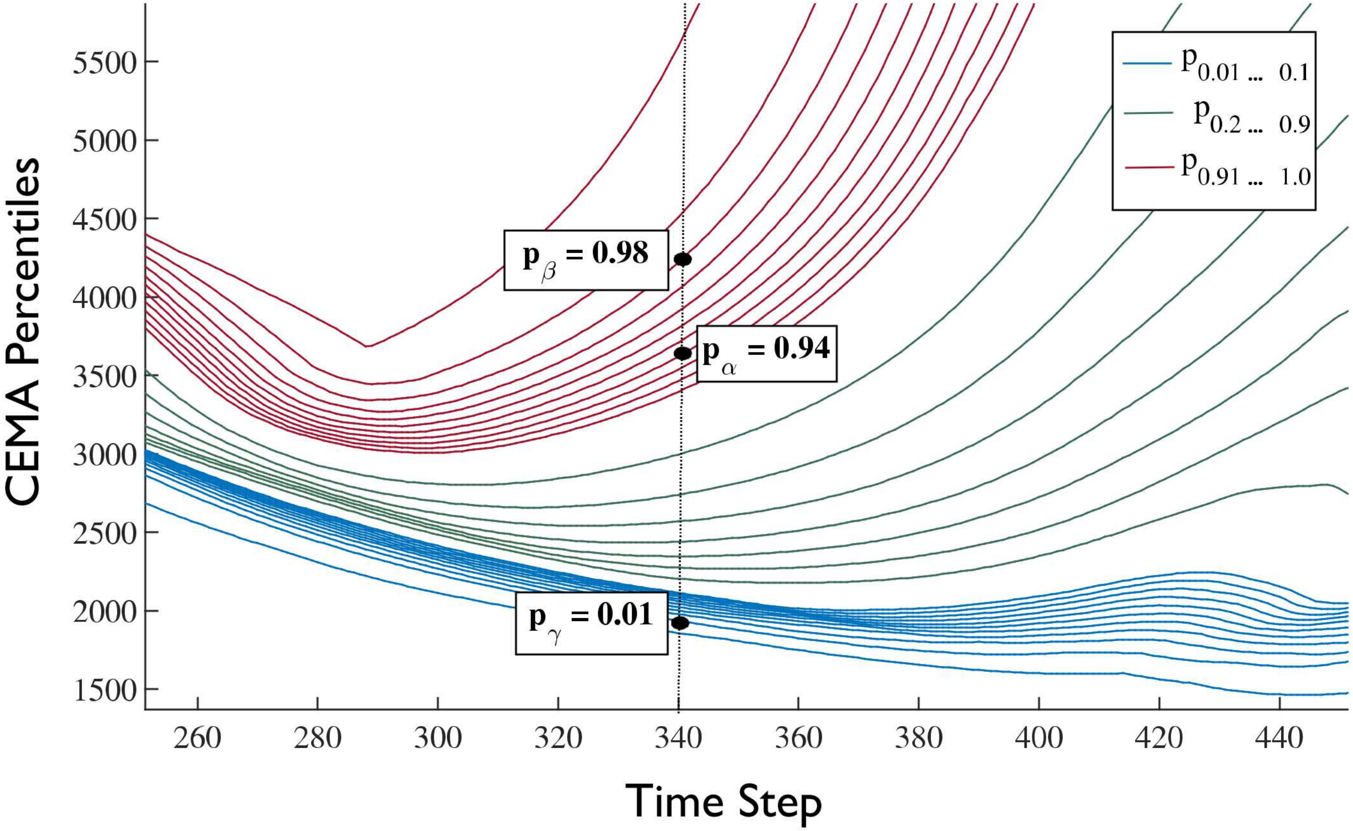
P_{100} = max value

Q3: How do you enable real-time decision-making in situ with quality and reliability guarantees?

We developed a CEMA indicator based on the variability of CEMA values in relation to the mean

$$P_{\alpha,\beta,\gamma}(t) = \frac{p_\alpha(t) - p_\gamma(t)}{p_\beta(t) - p_\gamma(t)}.$$

$$C_{\alpha,\beta}(t) = \sqrt{\frac{\mu}{N-1}\sum_{s=\alpha}^{\beta}(p_s/\mu - 1)^2} \quad \mu = \frac{1}{N}\sum_{s=\alpha}^{\beta} p_s$$

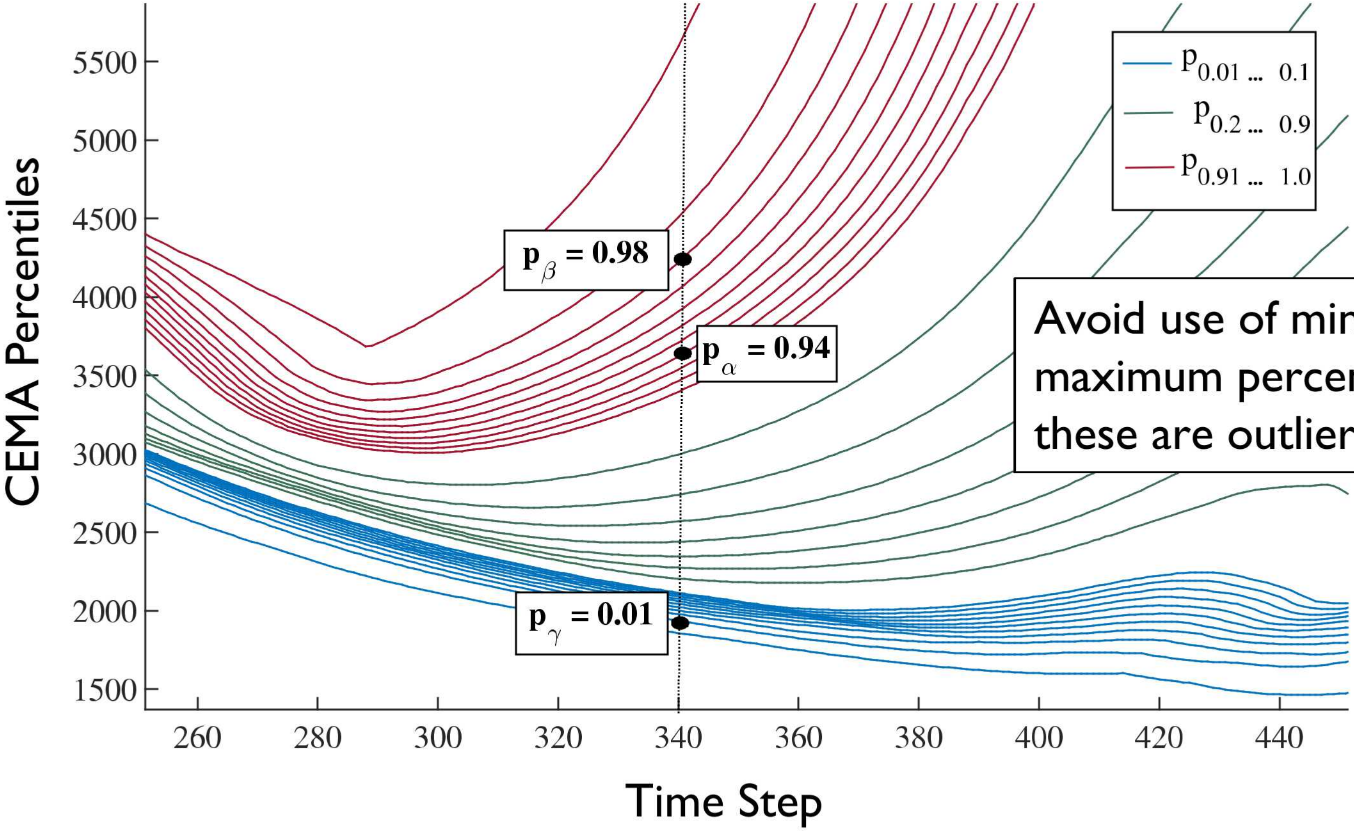


Q3: How do you enable real-time decision-making in situ with quality and reliability guarantees?

We developed a CEMA indicator based on the variability of CEMA values in relation to the mean

$P_{\alpha,\beta,\gamma}(t) =$ Range of ratios of percentile values

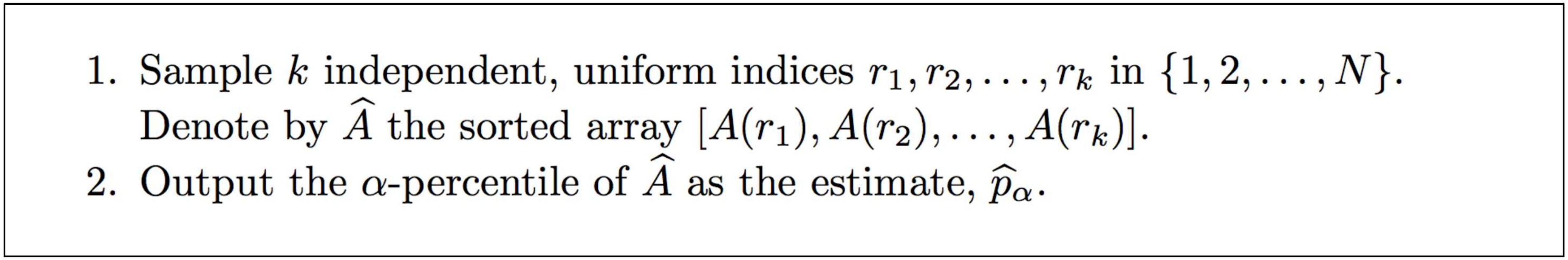
$C_{\alpha,\beta}(t) =$ Look at variability in percentile values in relation to the mean percentile value



Avoid use of minimum and maximum percentiles as these are outliers

We proposed a simple strategy for sampling percentiles that scales well

Sampling-Based Algorithm



1. Sample k independent, uniform indices $r_1, r_2, \ldots, r_k$ in $\{1, 2, \ldots, N\}$. Denote by $\widehat{A}$ the sorted array $[A(r_1), A(r_2), \ldots, A(r_k)]$.
2. Output the α-percentile of $\widehat{A}$ as the estimate, $\widehat{p}_\alpha$.

Scientists we were working with were initially hesitant to use a sampling-based approach to computing CEMA, for fear of

- Drawing the wrong conclusions
- Missing interesting science

Sublinear analysis can provide rigorous guarantees (math proofs) on the samples and accuracy required

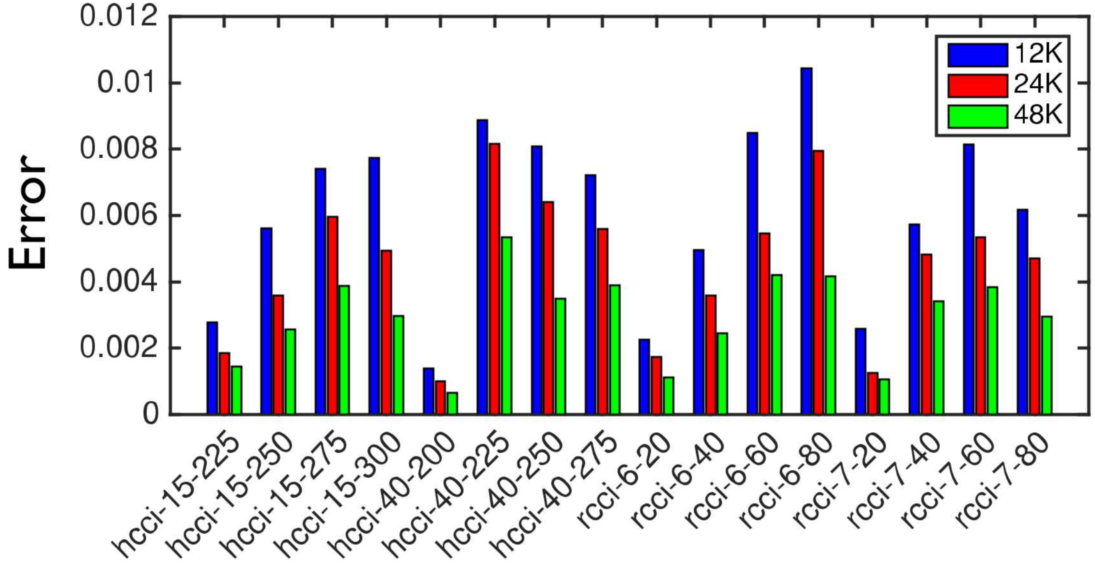
Number of samples is dependent on accuracy desired

Number of samples is NOT dependent on size of domain

Highly scalable approach

Empirical studies verified the guarantees provided by the sublinear analysis proofs

Number of Samples



Error in indicator function for various simulation runs

[Trigger Detection for Adaptive Scientific Workflows Using Percentile Sampling]

[Enabling Adaptive Scientific Workflows Via Trigger Detection]

Combined empirical and theoretical proofs were extremely useful from a user-adoption perspective

Benefits of sublinear analysis:

- Provides useful guarantees regarding accuracy and samples required for algorithms that characterize trends in data
- Enables rigorous deployment of algorithms otherwise too expensive to compute for real-time decision making
- Analysis could play a role in reproducibility initiatives when you can’t count everything

Limitations of sublinear analysis:

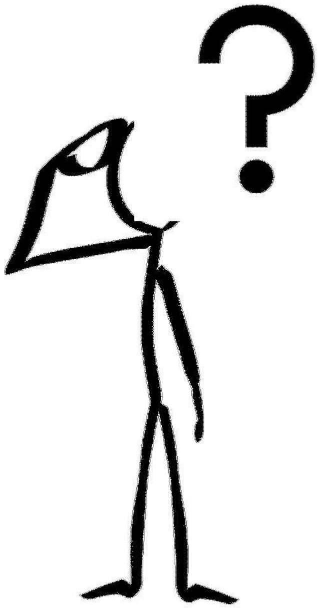
- Doesn’t apply to analysis algorithms focused on anomaly detection
- Need to team with an applied mathematician or statistician familiar with the underlying sublinear analysis techniques

There are many interesting in situ analysis challenges beyond these three

Q1: How should analysis algorithms share system resources with the application in situ?

Q2: How do analysis algorithms need to evolve to operate within a streaming regime?

Q3: How do you enable real-time decision-making in situ with quality and reliability guarantees?



An ASCR workshop on In Situ Data Management (ISDM) identified six priority research directions

1. **Pervasive ISDM:** Apply ISDM methodologies and in situ workflows at a variety of platforms and scales.
2. **In Situ Algorithms:** Redesign data analysis algorithms for the in situ paradigm.
3. **Composable ISDM:** Develop interoperable ISDM components and capabilities for an agile and sustainable programming paradigm.
4. **Co-designed ISDM:** Coordinate the development of ISDM with the underlying system software so that it is part of the software stack.
5. **Controllable ISDM:** Understand the design space of autonomous decision-making and control of in situ workflows.
6. **Transparent ISDM:** Increase confidence in reproducible science, deliver repeatable performance, and discover new data features through the provenance of ISDM.

ASCR WORKSHOP ON IN SITU DATA MANAGEMENT

Enabling Scientific Discovery from Diverse Data Sources

January 28–29, 2019

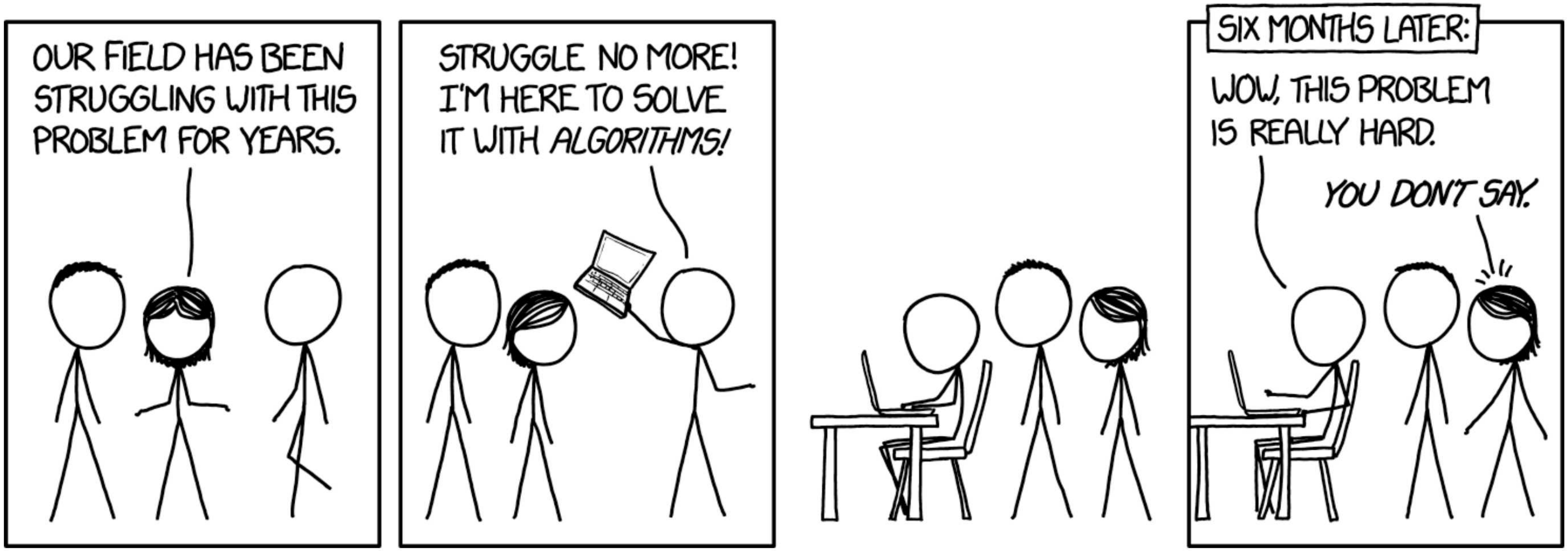
U.S. DEPARTMENT OF ENERGY

Workshop brochure available now

Full report later this year

Algorithmic research will play a crucial role, but not in isolation

- Crossing the "valley of death" - algorithms must be pulled into a sustainable ecosystem of tools
- Interdisciplinary challenges - different subject matter experts
- Guarantees on accuracy, reliability, and reproducibility in the regime where you "can't count everything"
- Use cases beyond exploratory analysis: UQ, ensembles, integration of experimental & simulation data


OUR FIELD HAS BEEN STRUGGLING WITH THIS PROBLEM FOR YEARS.
STRUGGLE NO MORE! I'M HERE TO SOLVE IT WITH ALGORITHMS!
SIX MONTHS LATER:
WOW, THIS PROBLEM IS REALLY HARD.
YOU DON'T SAY.

xkcd

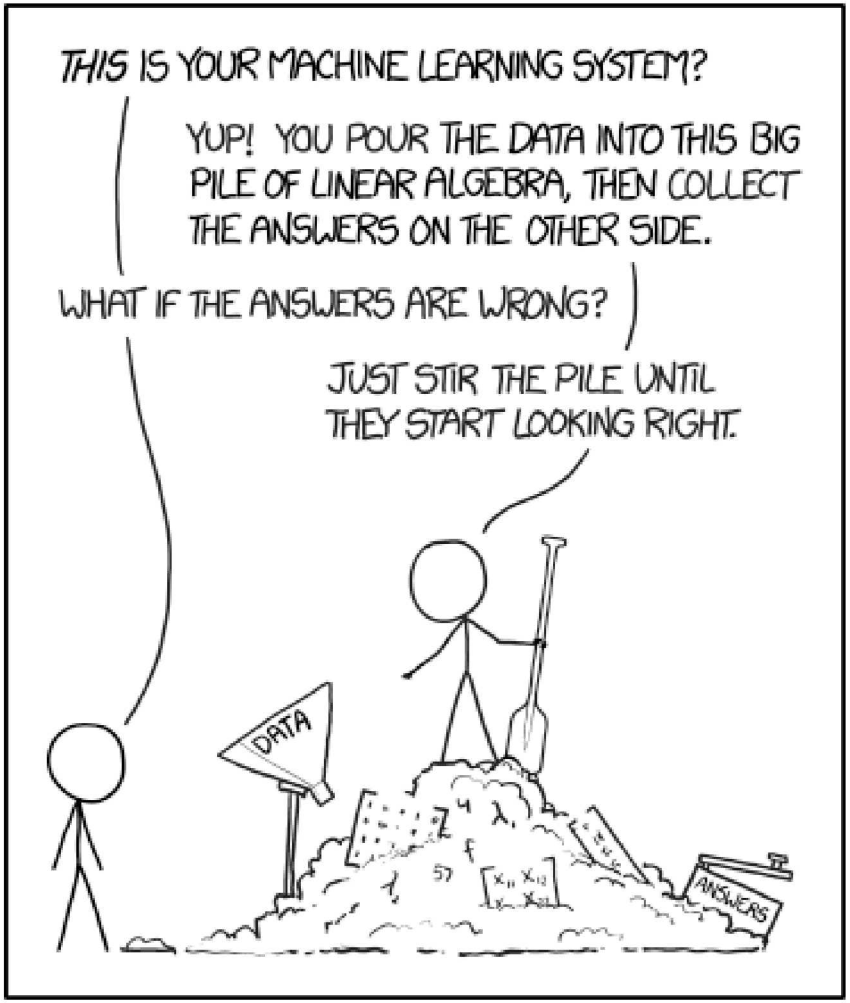
Machine learning (ML) has tremendous potential as a tool, but we do not yet know how to wield it

Potential Benefits:

- Automation and real-time decision-making
- Optimal configuration of algorithms
- Physics + data constrained reduced order models

Challenges:

- Training and updating models
- Developing first-principles constrained ML models
- Explainable and validated models
  - Is the model learning based on spurious correlations?


THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS

xkcd

Nature Communications Article: Unmasking Clever Hans predictors and assessing what machines really learn

Acknowledgements: Collaborators and Funding

ExaCT Co-design Center: Jacqueline Chen, Hasan Abbasi, Peer-Timo Bremer, Ray Grout, Attila Gyulassy, Tong Jin, Scott Klasky, Hemanth Kolla, Manish Parashar, Valerio Pascucci, Philippe Pebay, David Thompson, Hongfeng Yu, Fan Zhang

Streaming Statistics: Philippe Pebay, David Thompson, Timothy Terriberry, Ajith Mascarenhas

Topological Analysis: Peer-Timo Bremer, Valerio Pascucci, Aaditya Lange, Attila Gyulassy, Hemanth Kolla, Jacqueline Chen

Volume Rendering: Hongfeng Yu, Jacqueline Chen, Hemanth Kolla, Kwan-Liu Ma

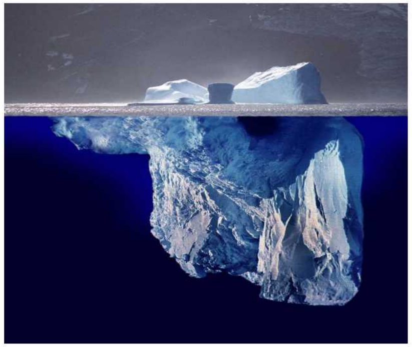
Sublinear Sampling: Seshadri Comandur, Ali Pinar, Maher Salloum, Jacqueline Chen, Hemanth Kolla

LDRD LABORATORY DIRECTED RESEARCH & DEVELOPMENT

U.S. DEPARTMENT OF ENERGY | Office of Science

ECP EXASCALE COMPUTING PROJECT

Questions?



Exploratory analysis is a vital part of the scientific process

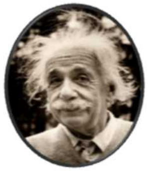
The alchemists in their search for gold discovered many other things of greater value.

Arthur Schopenhauer, German Philosopher

Challenge: Emerging HPC systems cannot output all relevant simulation data to disk

Not everything that can be counted counts, and not everything that counts can be counted.

Albert Einstein, Physicist



Algorithmic research will play a crucial role in maintaining our ability to perform exploratory analysis in the face of HPC system challenges

The goal is to turn data into information, and information into insight.

Carly Fiorina, Former CEO of HP