Experiences in Translational Computer Science Research: Data-Management for Extreme Science

Manish Parashar*

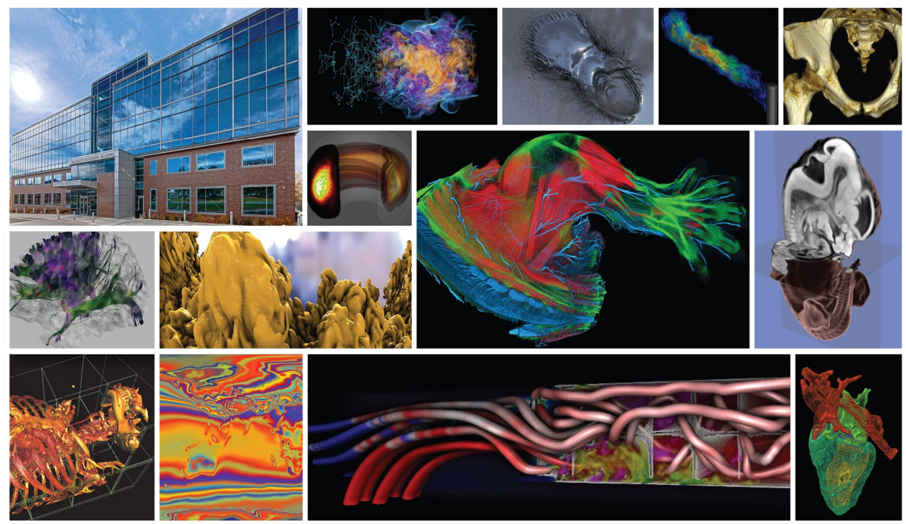
Director, Scientific Computing & Imaging (SCI) Institute & Chair in Computational Science and Engineering

Presidential Professor, School of Computing

University of Utah

PPAM 2022, Gdansk, Poland

September 14, 2022



***with David Abramson**

THE UNIVERSITY OF UTAH®

www.sci.utah.edu

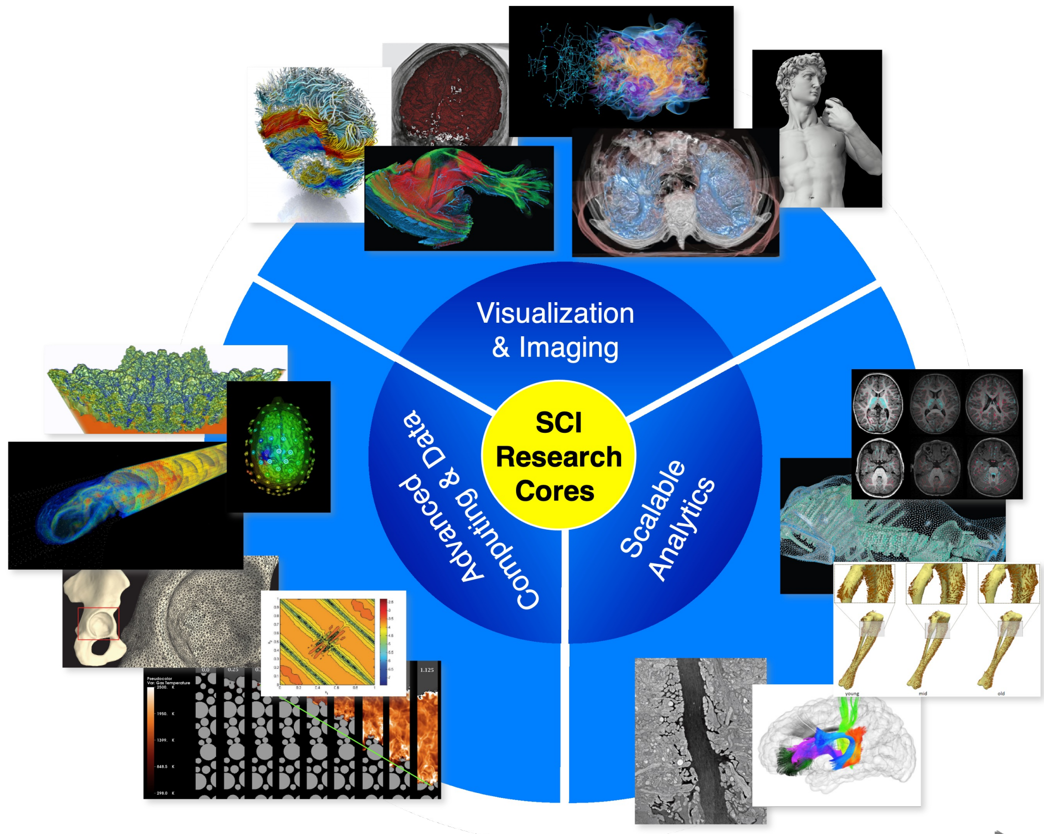
Outline

- Introducing Translational Computer Science (TCS)
- Experiences in TCS Research: Data-Management for Extreme Science
- Important Issues for TCS
- Conclusion

Scientific Computing & Imaging (SCI) Institute

Goal: Transformation of science and society through translational research and innovation in computer, computational and data science

- Multidisciplinary, convergent, collaborative
- Simulation, imaging, visualization, data management/analytics, advanced computing
- Software/system development and distribution

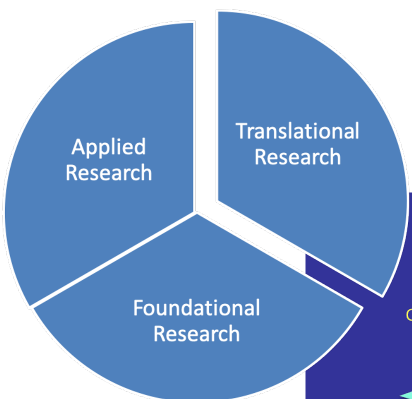


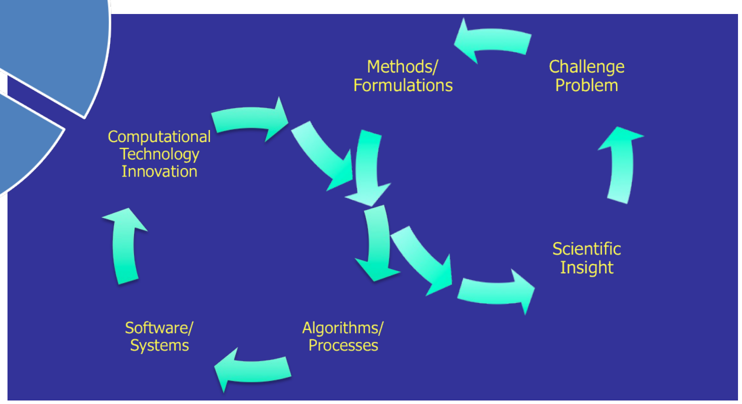
THE UNIVERSITY OF UTAH®

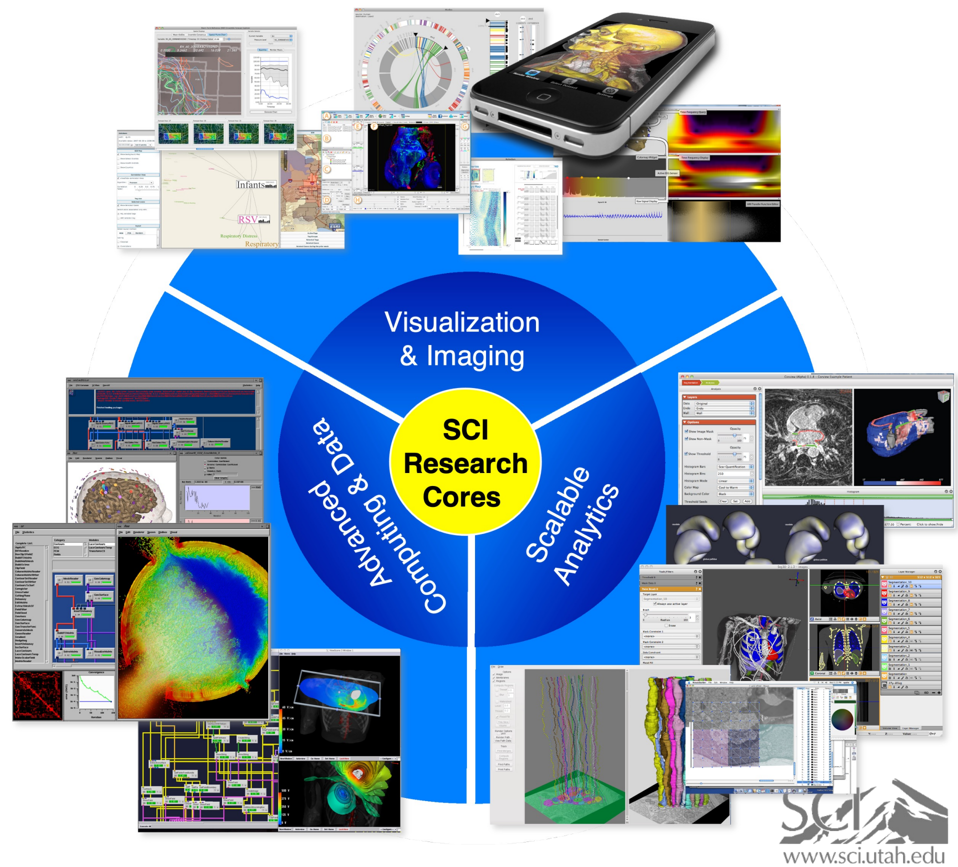
www.sci.utah.edu

Scientific Computing & Imaging (SCI) Institute

Goal: Transformation of science and society through *translational research and innovation* in computer, computational and data science







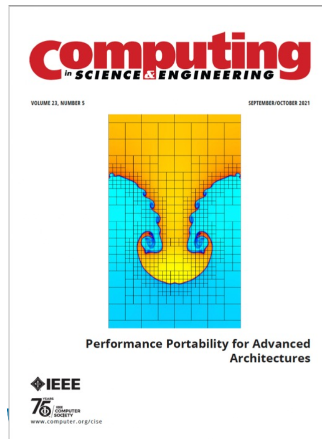
THE UNIVERSITY OF UTAH®

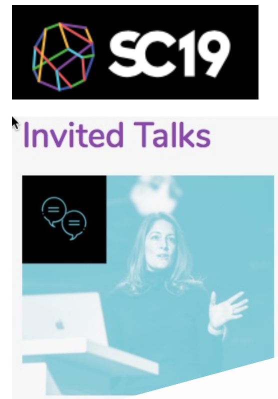
Outline

- Introducing Translational Computer Science (TCS)
- Experiences in TCS Research: Data-Management for Extreme Science
- Important Issues for TCS
- Conclusion

The role of a French chateaux and red wine … with David Abramson

Workshop on CCDSC, Lyon France, 2018
Jack Dongarra, Bernard Tourancheau





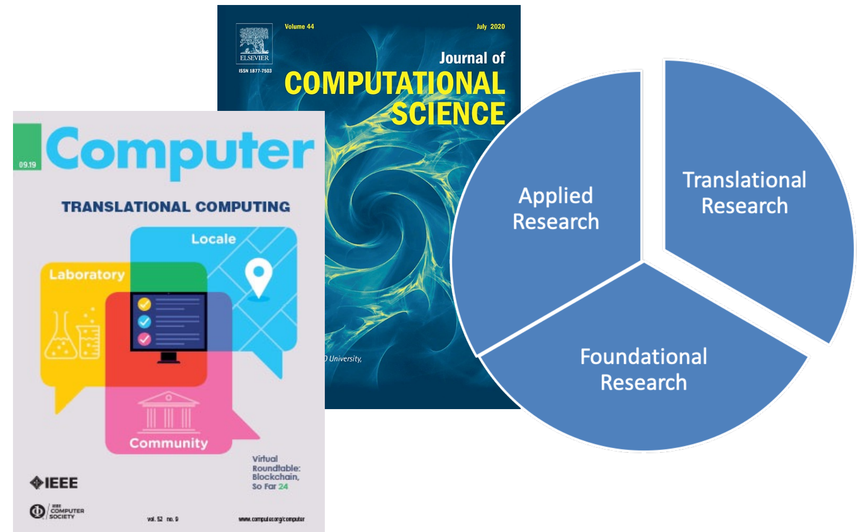
Background: Translational Medicine

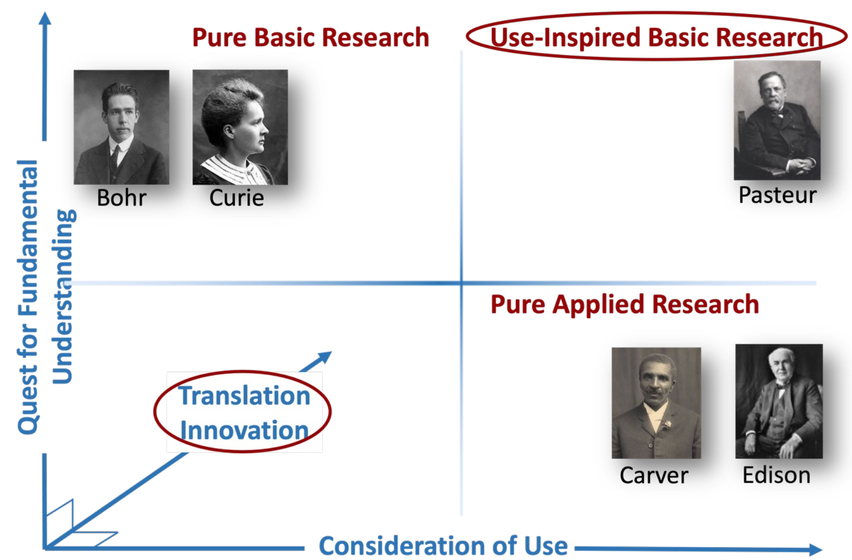
- An interdisciplinary branch of the biomedical field supported by three main pillars:
  - Benchside, Bedside and Community
  - Combines disciplines, resources, expertise, and techniques within these pillars to promote enhancements in prevention, diagnosis, and therapies
- Differs subtly from applied biomedical research, in which a research problem has a potential real-world application
  - Findings are applied as a specific phase of the research plan
  - This not only demonstrates applicability and practicality, but also generates tangible outcomes
- Now well understood and has become a de-facto standard for much of biomedical research
- Intrinsically helps generate outcomes because the research is applied as part of the original plan, as opposed to being an afterthought once the project has completed

Translational Computer Science (TCS)

TCS refers to research that bridges foundational and use-inspired (and applied) research with the *delivery and deployment of its outcomes* to the target community and supports essential *bi-direction interplays* where delivery and deployment process informs the research

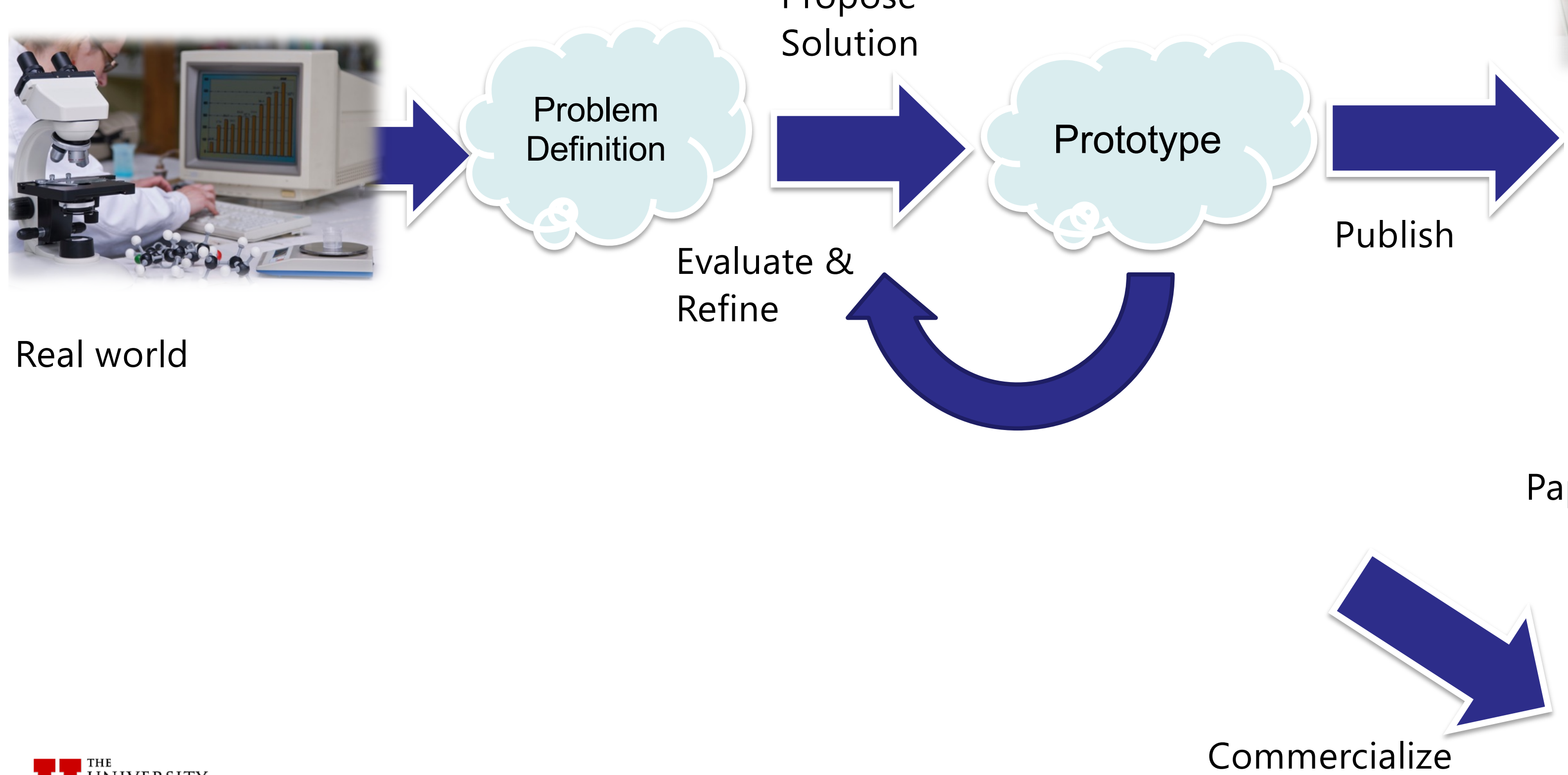
- Motivated by the growing importance of computing and data across all of science and society
- Aimed at accelerating the impact of computer, computational and data science
- Inspired by the definition and impact of Translational Medicine
- Focused on taking research from the *Laboratory* to the *Locale* to the *Community*
  - Laboratory, Locale might be physical or virtual
  - Community: Users and early adopters who work with the technology, and can include public bodies that would help in the evaluation





So how does this differ from traditional research pipeline?

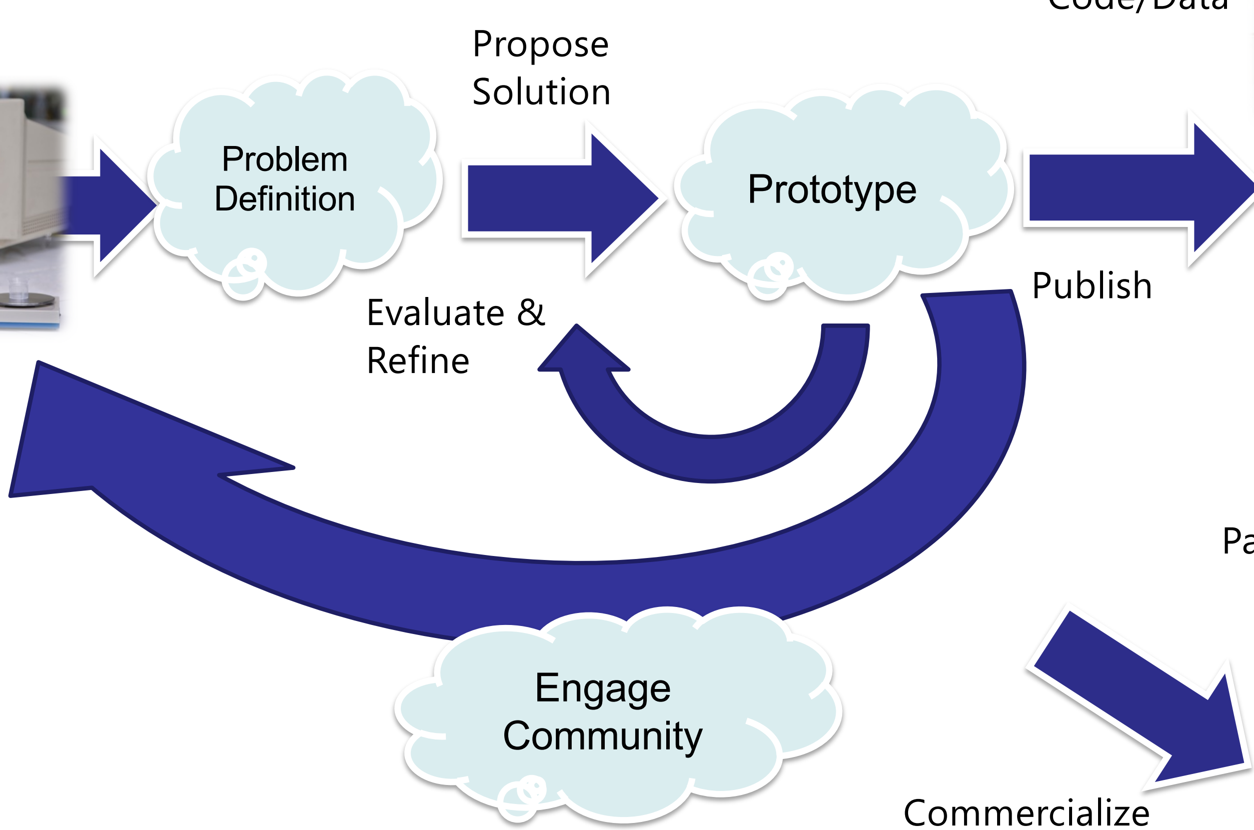
Typical CS research workflow



THE UNIVERSITY OF UTAH®

SCI www.sci.utah.edu

Translational CS research workflow



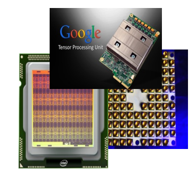
Outline

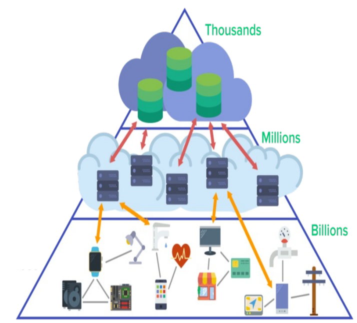
- Introducing Translational Computer Science (TCS)
- Experiences in TCS Research: Data-Management for Extreme Science
- Important Issues for TCS
- Conclusion

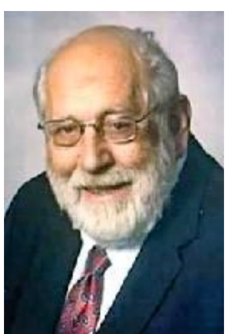
Science transformed by extreme scales, pervasive compute, data

- Exascale, and beyond
- Disruptive innovations => new opportunities
- Novel paradigms / a computing continuum
- Unconventional software stacks

Caffe
Caffe2
Chainer
MATLAB
Microsoft Cognitive Toolkit
mxnet
PaddlePaddle
PyTorch
TensorFlow
torch
Wolfram Language



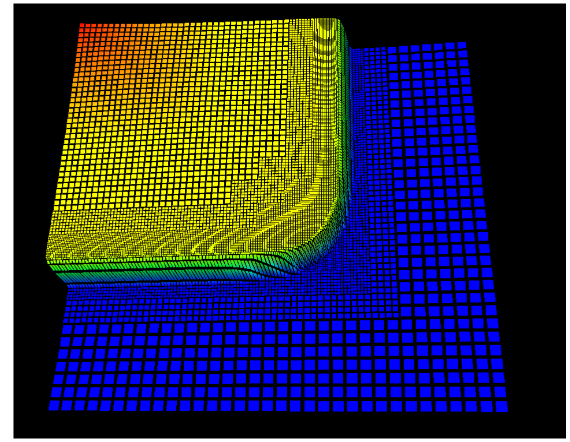




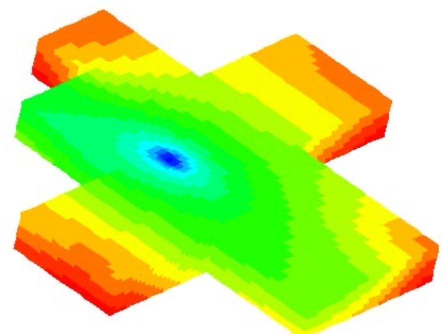
"A supercomputer is a device for turning compute-bound problems into I/O-bound problems" – Ken E. Batcher

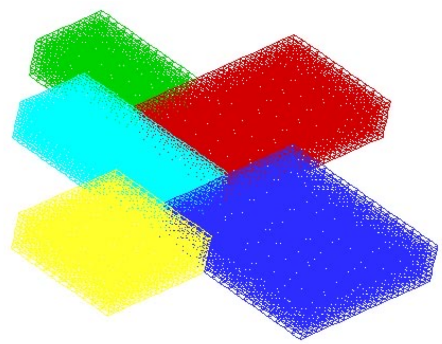
Data Management for Extreme Science: Part I – An Accidental Translationist

- Dynamic, adaptive formulations



- Coupled models and codes

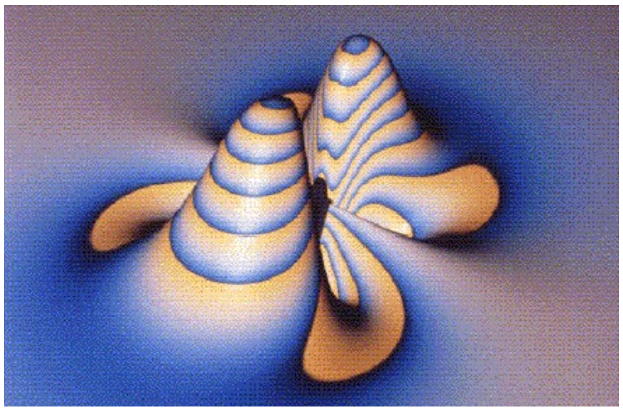


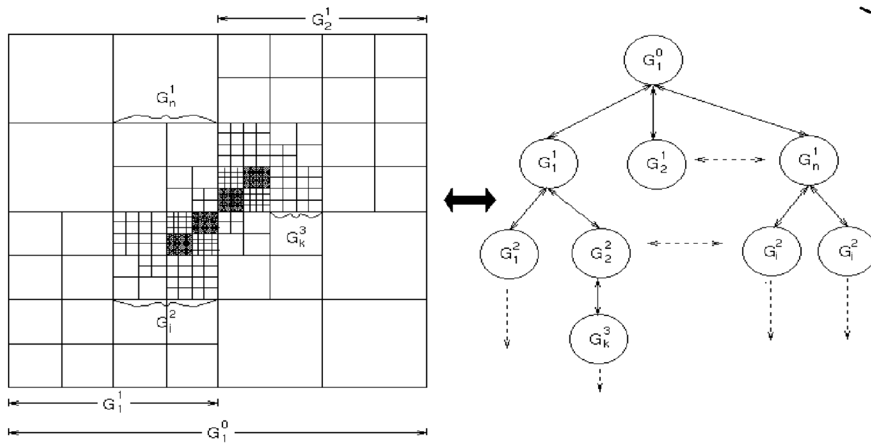


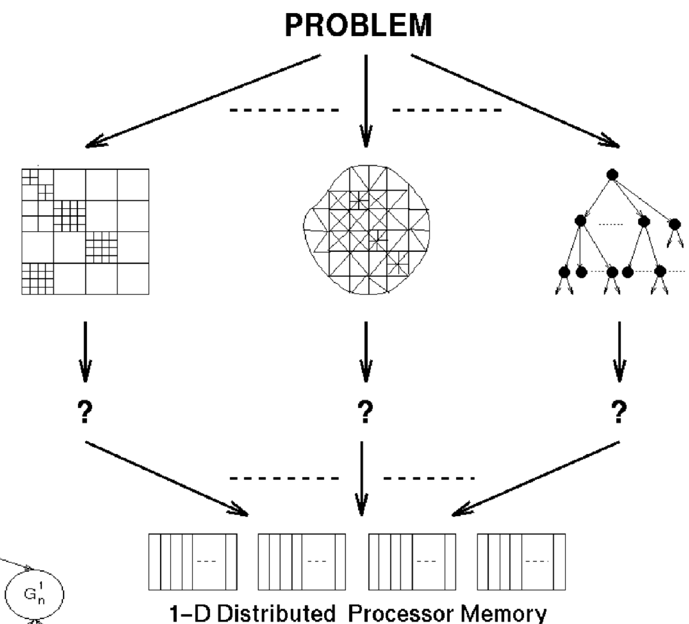
- In-situ workflows/in-transit processing

- End-to-end workflows

Binary Black Holes, Structured Adaptive Mesh Refinement, & Dynamic Parallel Applications







THE UNIVERSITY OF UTAH®

SCI
www.sci.utah.edu

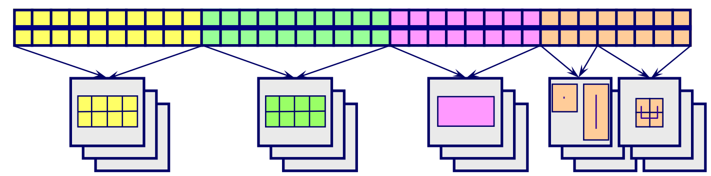
Data Management for Dynamically Adaptive Applications

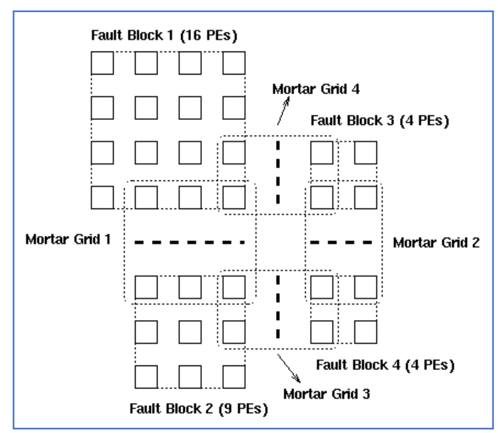
- Semantically Specialized DSM
  - Regular access semantics to dynamic, distributed data
  - Encapsulates distribution, communications
- Hierarchical Distributed Dynamic Array / Distributed Shared Objects
  - Hierarchical Index-Space + Extendible Hashing
- Adaptive Run-time Management
  - Application and system sensitive management
  - Policy-based adaptations
- Software Deployed:
  - GrACE/DAGH, MACE, ArMADA

SYSTEMS ENGINEERING FOR HIGH PERFORMANCE COMPUTING SOFTWARE: THE HDDA/DAGH INFRASTRUCTURE FOR IMPLEMENTATION OF PARALLEL STRUCTURED ADAPTIVE MESH

MANISH PARASHAR* AND JAMES C. BROWNE†

Abstract. This paper defines, describes and illustrates a systems engineering process for development of software systems implementing high performance computing applications. The example which drives the creation of this process is development of a flexible and extendible program development infrastructure for parallel structured adaptive meshes, the HDDA/DAGH package. The fundamental systems engineering principles used (hierarchical abstractions based on separation of concerns) are well-known but are not commonly applied in the context of high performance computing software. Application of these principles will be seen to enable implementation of an infrastructure which combines breadth of applicability and portability with high performance.

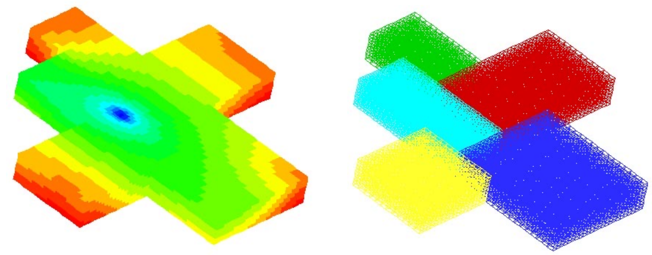




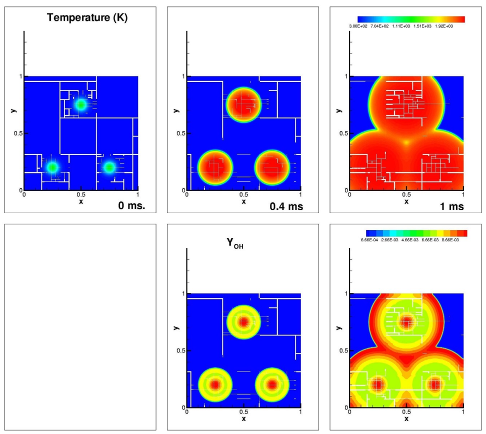
THE UNIVERSITY OF UTAH®

www.sci.utah.edu

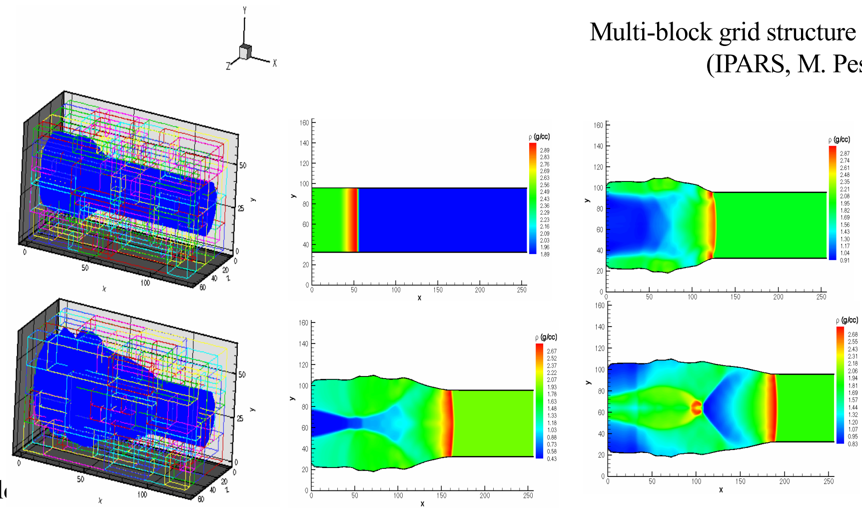
Translations Impacts



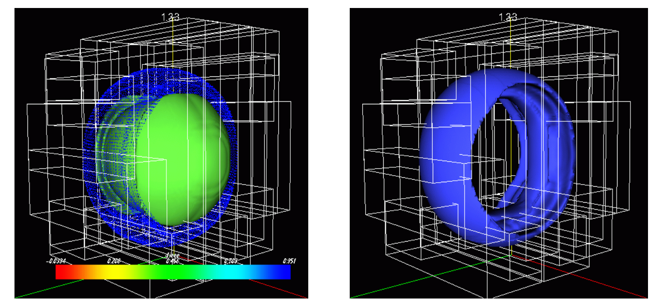
Multi-block grid structure and oil concentrations contours
(IPARS, M. Peszynska, UT Austin)



Mixture of H2 and Air in stoichiometric proportions with a non-uniform temperature fiel
(GrACE + CCA, Jaideep Ray, SNL, Livermore)



Richtmyer-Meshkov - detonation in a deforming tube - 3 levels. Z=0 plane visualized on the right (VTF + GrACE, R. Samtaney, CIT)



Blast wave in the presence of a uniform magnetic field) – 3 levels of refinement. (Zeus + GrACE + Cactus, P. Li, NCSA, UCSD)

Richtmyer Meshkov (3D) R. Samtaney, CIT

1024x128x128, 3 levels, 2K PE's
Time: ~ 15% Memory: ~25%

Translational Impacts

PHYSICAL REVIEW LETTERS

Highlights Recent Accepted Collections Authors Referees Search Press Abo

Stable Characteristic Evolution of Generic Three-Dimensional Single-Black-Hole Spacetimes

Featured in Physics

Boosted Three-Dimensional Black-Hole Evolutions with Singularity Excision

Gravitational Wave Extraction and Outer Boundary Conditions by Perturbative Matching

SC97 · Technical Paper

TOC Authors Sessions Abstracts PostScript

A Common Data Management Infrastructure for Adaptive Algorithms for PDE Solutions

UNIVERSITY OF UTAH

Journal of Physics: Conference Series

OPEN ACCESS

Toward a first-principles integrated simulation of tokamak edge plasmas

Journal of Computational Chemistry

Asynchronous replica exchange for molecular simulations

Journal of Physics: Condensed Matter

PAPER

Nucleosome positioning and composition modulate *in silico* chromatin flexibility

Towards Dynamic Data-Driven Management of the Ruby Gulch Waste Repository

SPE 37979

A New Generation EOS Compositional Reservoir Simulator: Part I – Formulation and Discretization.

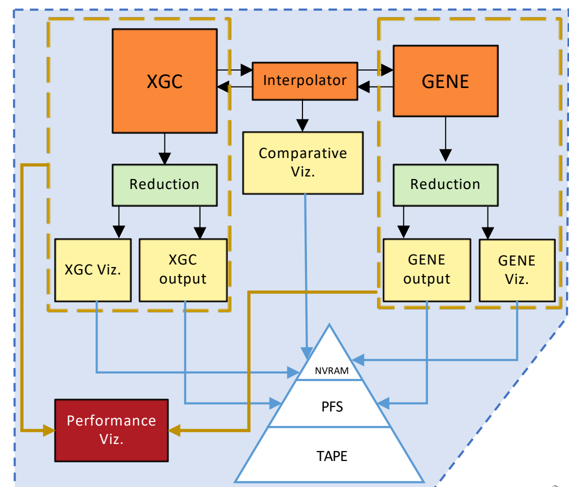
Variable Resolution Topographic Mapping of Ancient Fluvial Landscapes in Australia

Lost in translation… What happened to DAGH/GrACE, MACE?

- Lost in translation…could not be sustained
  - Translation not part of the research plan
    - Time, effort not allocated for user support, addressing user feedback, new requirements, etc.
  - Lack of funding/support for translation
    - No funding to hire developers/programmers
    - PhD student theses research does not focus on translational work
  - Academic career-path required reprioritization
    - Academic structures don't value translation
- But
  - Effective, had rapid uptake, and impact
  - Conceptual framework (and some code) persists, used by other frameworks and influences codebases
- Many lessons learnt that informed subsequent work

Data management for Extreme Science: Part II

- Dynamic, adaptive formulations
- Coupled models and codes
- In-situ workflows/in-transit processing
- End-to-end workflows



THE UNIVERSITY OF UTAH®

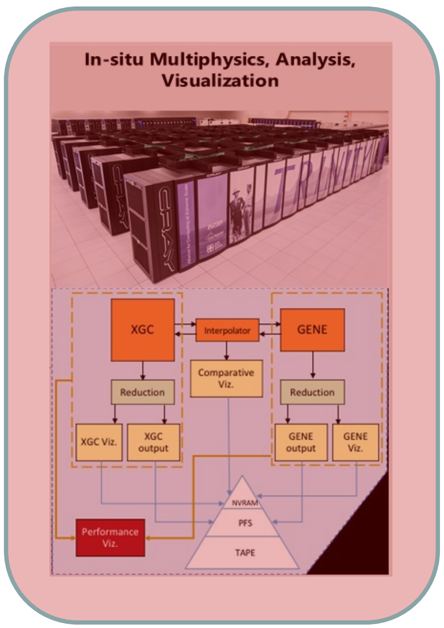
www.sci.utah.edu

Extreme-Scale, End-to-End Science Workflows

Plasma Disruption Analysis Fusion Workflow

- The goal is to implement a plasma visualization diagnostics system that enables early prediction of anomalies while the tokamak operates
- Data is a time-series of matrix images obtained via direct 2D/3D visualizations of the tokamak plasma
- Analytics include structural analysis and blob detection, filtering & visualization

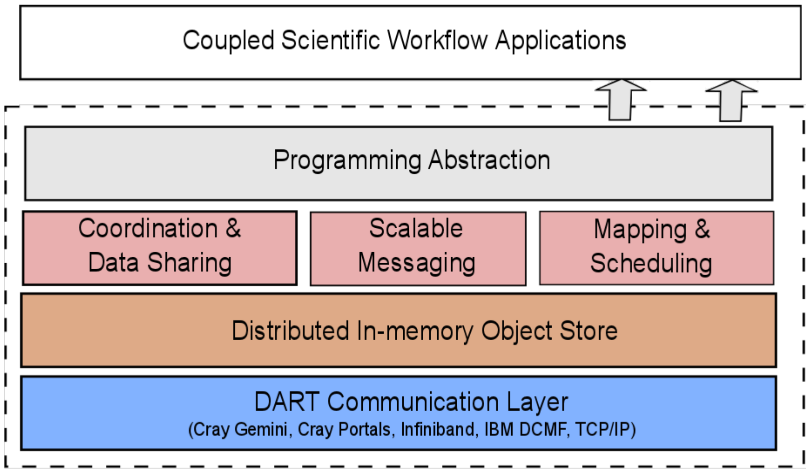
Rethinking Extreme-Scale Workflows: In-situ Workflows, In-transit Processing

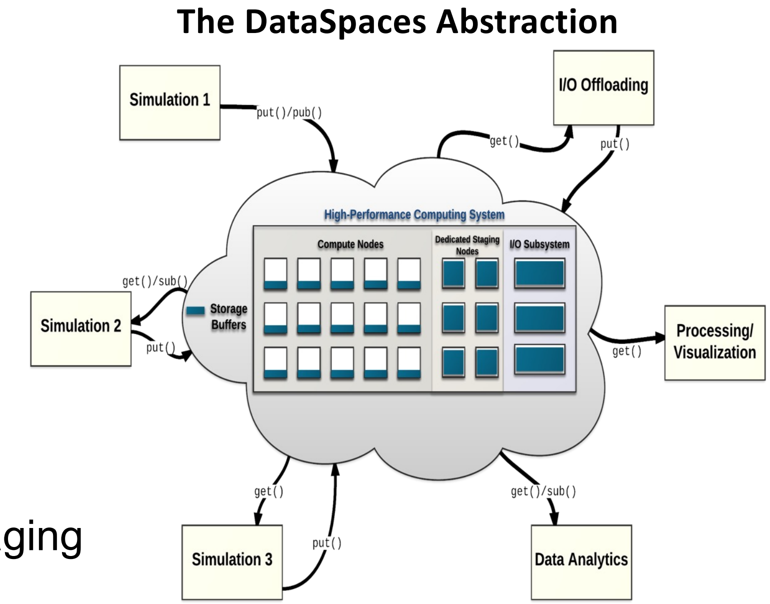


Plasma Disruption Analysis Fusion Workflow

- The goal is to implement a plasma visualization diagnostics system that enables early prediction of anomalies while the tokamak operates
- Data is a time-series of matrix images obtained via direct 2D/3D visualizations of the tokamak plasma
- Analytics include structural analysis and blob detection, filtering & visualization

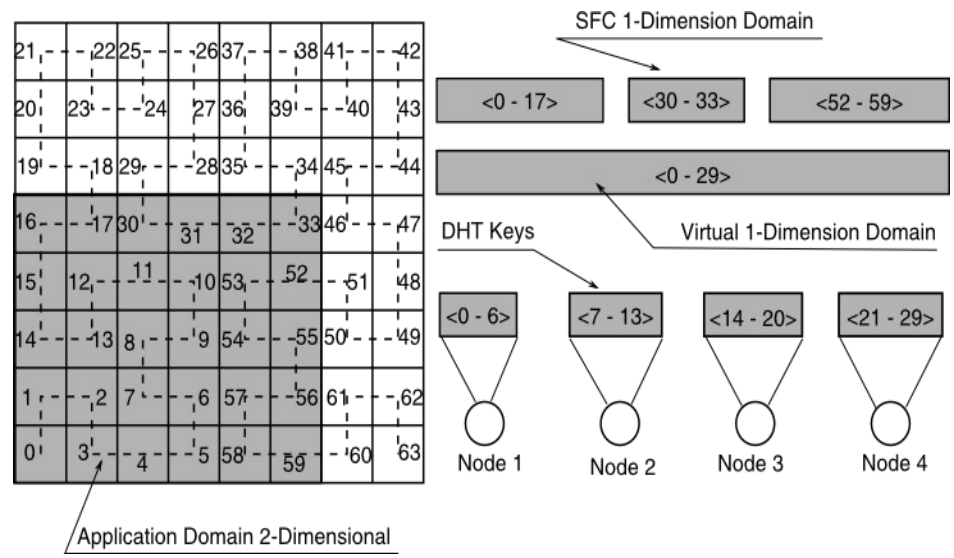
DataSpaces: Data Staging Service for In-Situ Workflows





- Virtual shared-space programming abstraction
  - Simple API for coordination, interaction and messaging
- Distributed, associative, in-memory object store
  - Online data indexing, flexible querying
- Cross-layer runtime management
  - Hybrid in-situ/in-transit execution
- High-throughput/low-latency asynchronous data transport

DataSpaces : Indexing + DHT

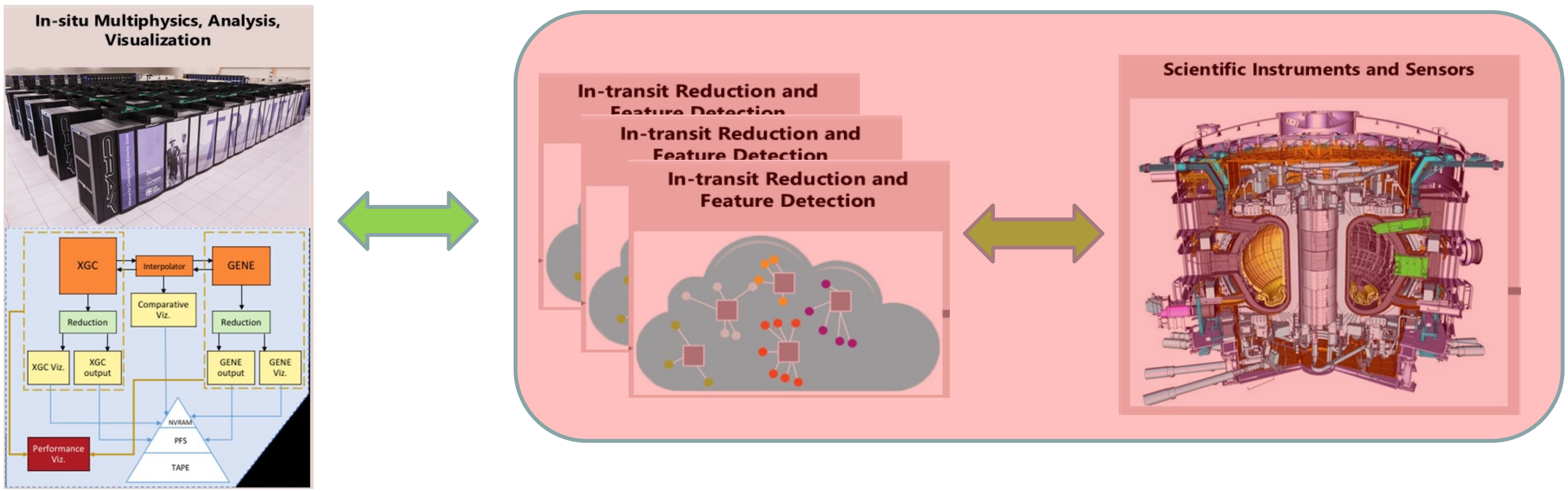


- Dynamically constructed structured overlay using hybrid staging cores
- Index constructed online using SFC mappings and applications attributes
  - E.g., application domain, data, field values of interest, etc.
- DHT used to maintain meta-data information
  - E.g., geometric descriptors for the shared data, FastBit indices, etc.
- Data objects load-balanced separately across staging cores

The SFC maps the global domain to a set of intervals

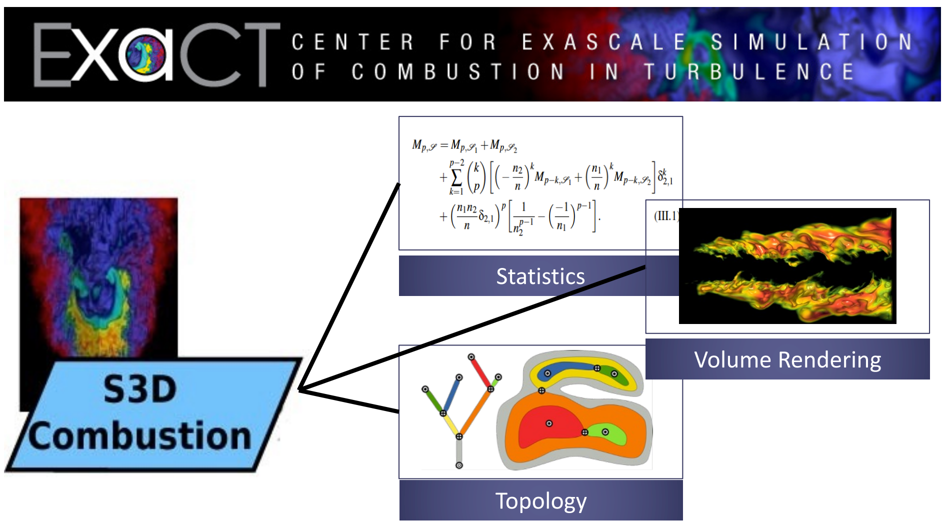
- Intervals are non-contiguous and can lead to meta-data load imbalance
- Second mapping compacts the intervals into a contiguous virtual interval
- Split the contiguous interval equally to the DataSpaces nodes

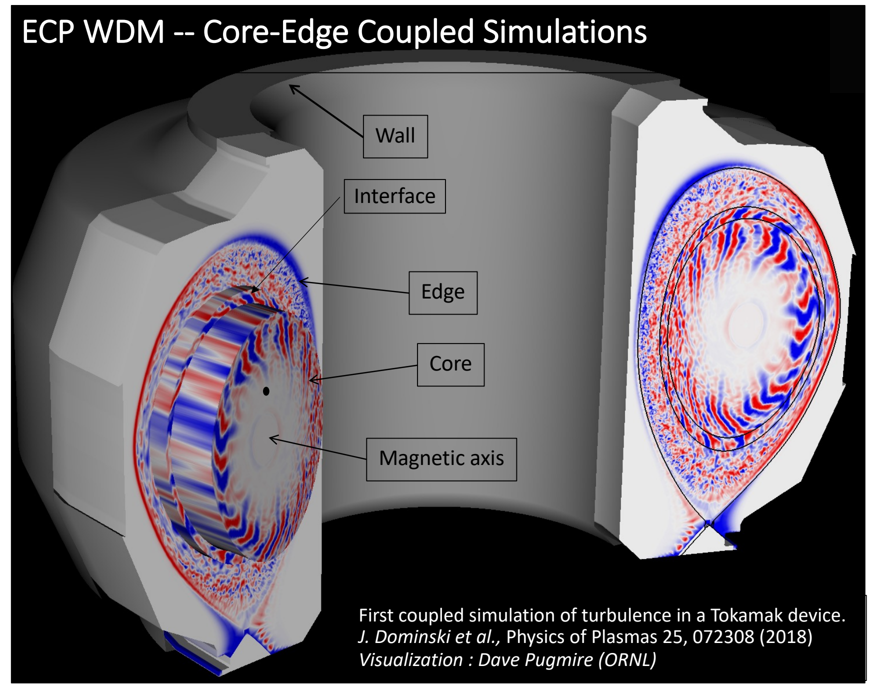
Rethinking Extreme-Scale Workflows: In-situ Workflows, In-transit Processing



- A. Zamani, D. Balouek-Thomert, J.J. Villalobos, I. Rodero, M. Parashar, An edge-aware autonomic runtime for data streaming and in-transit processing, FGCS, vol. 110, pp. 107-118, 2020.
- M. Aktas, J. Diaz-Montes, I. Rodero, M. Parashar, WA-Dataspaces: Exploring the Data Staging Abstractions for Wide-Area Distributed Scientific Workflows. ICPP 2017, pp. 251-260, IEEE.

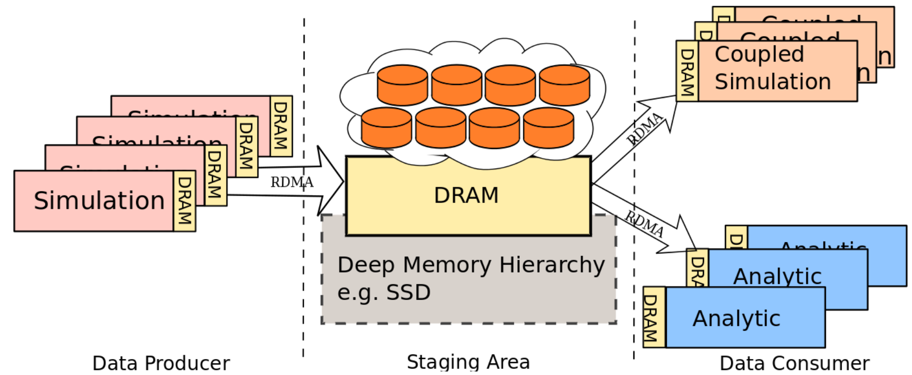
Translational Impacts

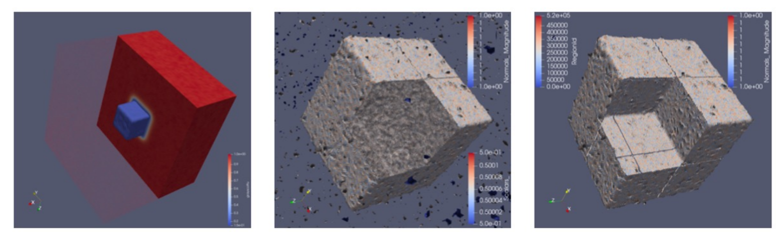




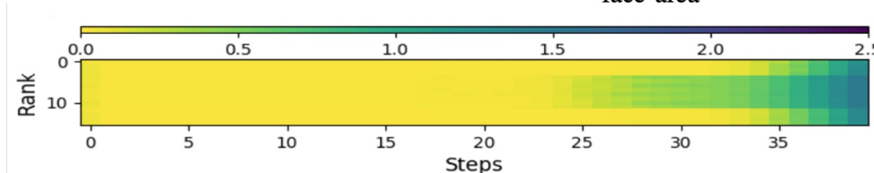
Feedback in Translation: Data Movement, Placement is Important

- Storage systems getting deeper and more complex
- Access latencies can severely impact performance
- Data-placement/movement is dependent on the data
- Staging services shared across multiple workflows
- *Autonomic, “smart” runtime management needed*





(a) Initial data (b) Isosurface (c) The largest isosurface area

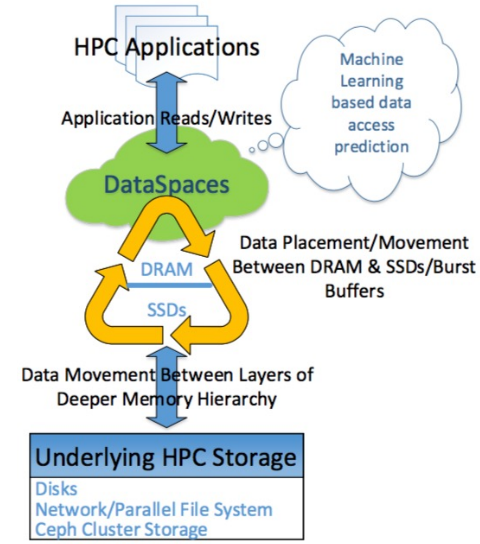


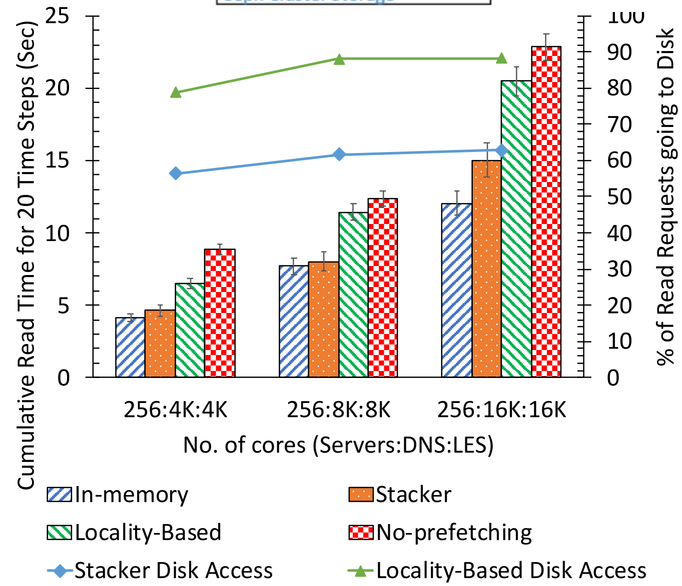
THE UNIVERSITY OF UTAH®

Stacker: ML-Guided Data Movement across the Staging Hierarchy [SC18]

- **Goal:** Reduce the number of read requests to disk using intelligent prefetching; extend in-memory data staging across DRAM, SSDs, burst buffers
- **Approach:** Data prefetching across the multi-tier staging area using an n-gram model to learn and predict the next data access request
  - Learn inter-application data access patterns, enabling support for multiple concurrent workflows
  - No a priori knowledge of application access pattern or user inputs/hints assumed
- **Results:** Read performance improved by ~40% as compared to no prefetching

Pradeep Subedi



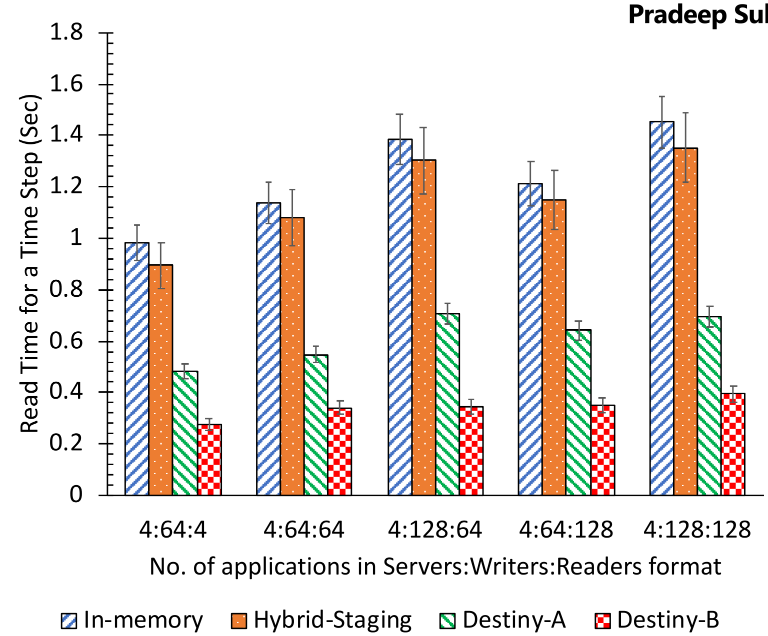


THE UNIVERSITY OF UTAH®

Destiny: Anticipatory Delivery of In-Staging Data [Cluster 19]

Pradeep Subedi

- **Goal:** Reduce data access costs for in-staging data
- **Approach:** Anticipate data accesses using an n-gram ML model, and proactively package and move data close to the consumer
  - Amortize expensive data discovery and assembly operations in data staging
- **Results:** Read response time reduced by up to 75% and 53% for co-located application processes and processes residing in separate nodes respectively

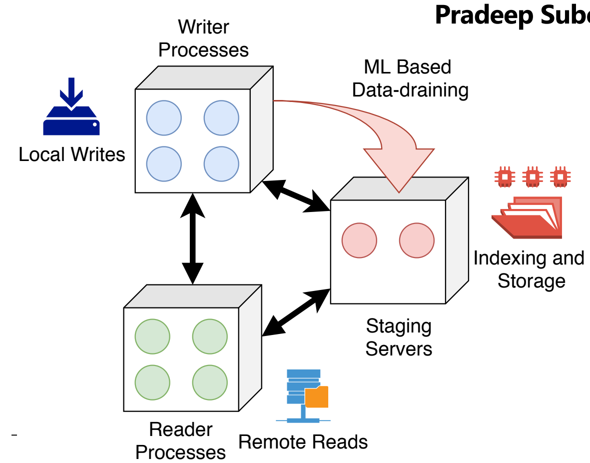


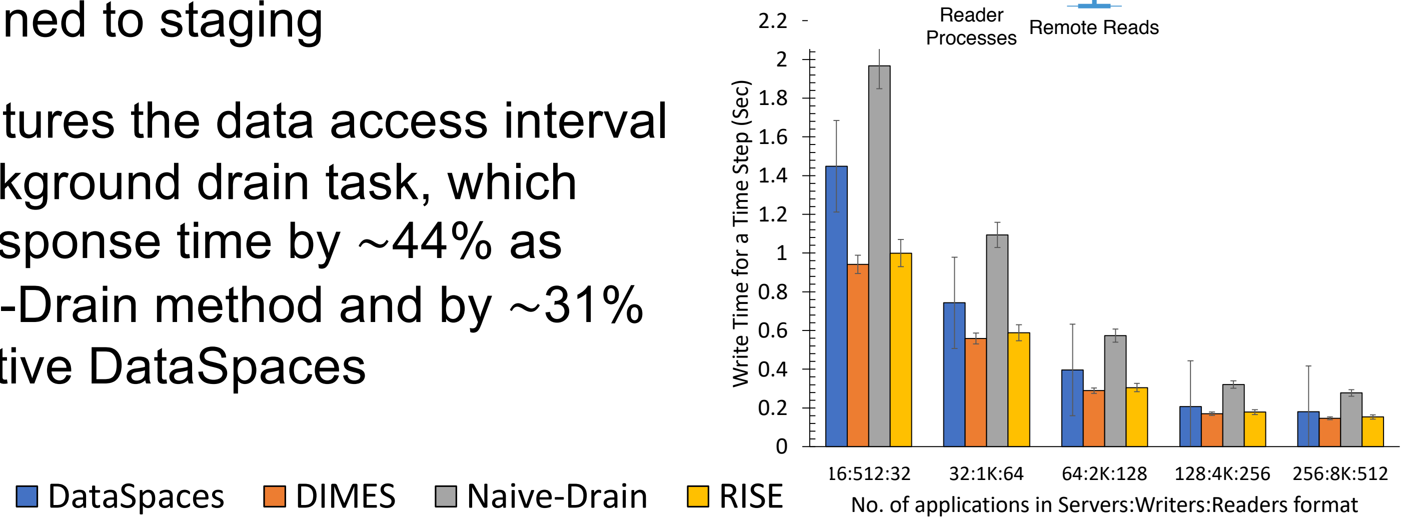
Hybrid-Staging represents just exposing data through shared memory. *Destiny-A* represents applications residing on separate nodes than staging and *Destiny-B* represents collocated staging and application processes.

RISE: Reducing I/O Contention for Staging-based In-situ Workflows [Cluster 21]

Pradeep Subedi

- **Goal:** Reduce data write access costs using "smart" data offloading
- **Approach:** Models the interactions across staging-based coupled applications to predict periods of server idle-time; determine suitable data objects to drained to staging
- **Results:** RISE captures the data access interval and schedules background drain task, which reduce the write-response time by ~44% as compared to Naive-Drain method and by ~31% as compared to native DataSpaces





THE UNIVERSITY OF UTAH®

Outline

- Introducing Translational Computer Science (TCS)
- Experiences in TCS Research: Data-Management for Extreme Science
- Important Issues for TCS
- Conclusion

TCS Roadblocks

1. In computer science, translation is often confused with commercialization
2. Open-source techniques are often confused for translation
3. Funding agencies typically don't provide support for translation
   a) Resources to sustain and maintain research artifacts (software, data) are essential.
4. PhD programs don't allocate time and resources to translation
   a) Translation is not considered as an acceptable these topic
5. Traditional academic structures, publication venues don't value translation
6. There are a lack of exemplars

COVER FEATURE TRANSLATIONAL COMPUTING

Translational Research in Computer Science

David Abramson, University of Queensland
Manish Parashar, Rutgers University

There are benefits to formalizing translational computer science (TCS) to complement traditional modes of computer science research, as has been done for translational medicine. TCS has the potential to accelerate the impact of computer science research overall.

Computing and data are transforming science and society and impacting every aspect of our lives and our environment. They have become integral to the research process and have been key catalysts, complementing theory and experimentation, for recent advances and breakthroughs. Unprecedented levels of instrumentation and the exponential growth of digital data sources, coupled with pervasive computing, provide new opportunities for discovery and innovation. Examples exist in diverse application areas, from managing extreme weather events to optimizing everyday processes and improving the quality of life.

Recently, there have been dramatic changes in the nature of research-application workflows. These are driven, in part, by the greater availability and increasing scales of experimental and observational data, the ability to model phenomena more holistically, and the desire for near-real-time data processing and actuation. There have also been dramatic changes in the technology and resource landscape, complementing this change in application workflows. This includes changes in processor price and performance, the emergence of novel storage technologies, the proliferation of accelerators, the deployment of high-bandwidth/low-latency networks, and the increasing capabilities available at the network edge. In addition, through investments by research agencies worldwide, local and national cyberinfrastructure capabilities have grown in capacity and capability

Digital Object Identifier 10.1109/MC.2019.2925650
Date of publication: 27 August 2019

Some thoughts for the path forward…

Funding

- Currently TCS funding is ad hoc => Sustained funding programs and mechanisms focused on fostering and nurturing TCS
- More money alone not solution => Need to build translation into the research plan
- Funding must support experimenting with solutions and prototypes (e.g., Transition to Practice)
- TCS typically involves substantial interaction with end users => Additional funding for travel, user engagement, provisioning of development resources
- Translation process feeds back into the research => May require a loop of research and translation rather than a linear workflow
- Research should be free (encouraged) to report on both research successes, but also translation success or failures

Venues, metrics and reward structures

- Traditional publications are not well suited to TCS => Similar to TM, establish new journals that explicitly target translation
- Metrics, recognitions and rewards structures are lacking, especially in the academic community
  - Mechanisms to track their use, citations and impact are a step in the right direction
  - Metrics that report uptake, and measure how many of these have resulted in successful translation
  - Integrate metrics into promotion processes

Education and Training

- Integration of translational approaches and methodologies into more formal computer science curricula
- New materials and mechanisms for providing translational skills to practitioners, in both computer and other disciplines
  - Institutes/centres have been established that encourage and enable trans-disciplinary research
- Extreme example: A PhD could be entirely devoted to the translation of work performed by another researcher, with no original research on the background IP per se

Outline

- Introducing Translational Computer Science (TCS)
  - Exemplars of TCS Research
- Experiences in TCS Research: Data-Management for Extreme Science
- Important Issues for TCS
- Conclusion

Conclusion

- CS research, innovations are transforming science and society
- TCS complements traditional CS research models (foundational, use-inspired, applied) and can accelerate and amplify the impact of computer science research
- There are benefits to formalizing TCS to complement traditional modes of computer science research
  - Several issues: Funding models, reward structures, publication venues, education and training, etc.

Call for Papers:

IEEE CiSE Department: Case Studies in Translational Computer Science

Editors:

Manish Parashar, Scientific Computing and Imaging Institute, University of Utah

David Abramson, Centre for Research Computing, The University of Queensland

Description: Our new CISE department explores how findings in fundamental research in computer, computational, and data science translate to technologies, solutions, or practice for the benefit of science and engineering, and society. Specifically, each department article will highlight impactful translational research examples in which research has successfully moved from the laboratory to the field and to the community. The goal is improving understanding of underlying approaches, exploring challenges and lessons learned, with the overarching aim to formulate translational research processes that are broadly applicable.

THE UNIVERSITY OF UTAH®

SCI www.sci.utah.edu

Thank you!

Manish Parashar

Email: manish.parashar@utah.edu

WWW: manishparashar.org / dataspaces.org

Acknowledging:

- David Abramson
- Pradeep Subedi, Philip Davis, Daniel Balouek-Thomert, Zhe Wang, Bo Zhang
- Many students, postdocs and collaborators

THE UNIVERSITY OF UTAH®

www.sci.utah.edu