The University Of Queensland AUSTRALIA | CREATE CHANGE

Translational Research in Computer Science

David Abramson
Research Computing Centre
University of Queensland, Brisbane, Australia

Introduction

- The role of a French Chateaux and red wine
- Translational Research
  - Medicine
  - Computer Science
- Why is it the right time for Translational Computer Science?
- Some exemplars
- Important issues
- Laboratory Scale Matters
- Don’t forget the research

The role of a French chateaux and red wine … with Manish Parashar

SC19

Invited Talks

Case Studies in Translational Computer Science

- The Pegasus workflow management system: Translational computer science in practice
  Ewa Deelman, Rafael Ferreira da Silva, Karan Vahi, Mats Rynge, ... Miron Livny
- The virtual assay software for human in silico drug trials to augment drug cardiac testing
  Elisa Passini, Xin Zhou, Cristian Trovato, Oliver J Britton, ... Blanca Rodriguez
- Translational research in the MPICH project
  William Gropp, Rajeev Thakur, Pavan Balaji
- The MVAPICH project: Transforming research into high-performance MPI library for HPC community
  Dhabaleswar Kumar Panda, Hari Subramoni, Ching-Hsiang Chu, Mohammadreza Bayatpour
- Translating the grid: How a translational approach shaped the development of grid computing
  Ian Foster, Carl Kesselman
- Building Cyberinfrastructure for Translational Impact: The WIFIRE Example
  Ilkay Altintas
- Computational analysis of cardiac structure and function in congenital heart disease: Translating discoveries to clinical strategies
  Nickolas Forsch, Sachin Govil, James C Perry, Sanjeet Hegde, ... Andrew D McCulloch
- Translating novel HPC techniques into efficient geoscience solutions
  Lin Gan, Haohuan Fu, Guangwen Yang
- Principles, technologies, and time: The translational journey of the HTCondor-CE
  Brian Bockelman, Miron Livny, Brian Lin, Francesco Prelz
- Translational process: Mathematical software perspective
  Jack Dongarra, Mark Gates, Piotr Luszczek, Stanimire Tomov
- Translational Computer Science at the Scientific Computing and Imaging Institute
  Chris Johnson

Background: Translational Medicine

- An “interdisciplinary branch of the biomedical field supported by three main pillars:
  - Benchside, Bedside and Community.
  - Combine 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 (driver).
  - 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

Traditional Computer Science Research

- Can range from
  - theory to practice
  - "soft" to "hard"
- Theory
  - Might be some fundamental advance in theory of computation
  - New algorithms (ways of solving problems)
  - Leads to better "science"
- Practice
  - New solutions (i.e. practical algorithms)
  - Applying new technologies
  - Leads to better "engineering"

- Theories and experimental results
  - Papers published in reviewed journals, conferences
- Patents
  - Published by the patents office
- Prototypes
  - Computer programs
  - Hardware

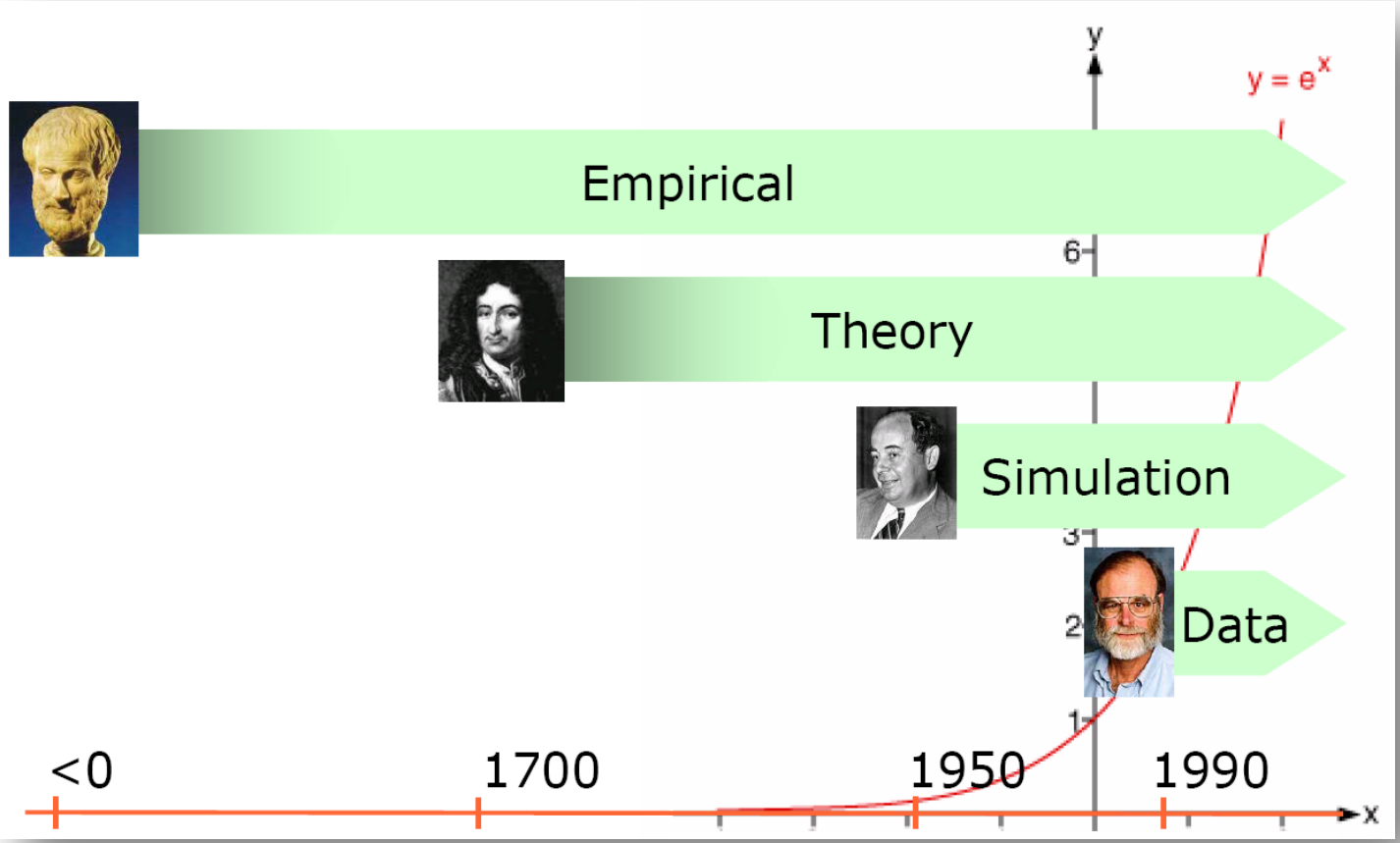
Translational Computer Science

- In TM, translation relies on
  - Taking research from the laboratory Bench to the Bedside
  - More recent refinements involve Community
    - healthy populations, patients and medical practitioners.
- In TCS, translation relies on
  - Taking research from the laboratory Laboratory to the 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

Computer Science is increasingly important

- Systems Biology and Health
  - Human Genome Project, Protein function, Virtual Physiological Human, Blue Brain
- Engineering
  - Aerospace, civilian, automotive, domestic, …
- Environment
  - Climate, weather, pollution, ….
- Chemistry
  - Drug design, novel pathways, …
- Physics
  - Particle Physics, Xray treatment, Astrophysics, ….
- Business
  - Manpower planning, Logistics, Resource allocation

Potential for translation has never been higher



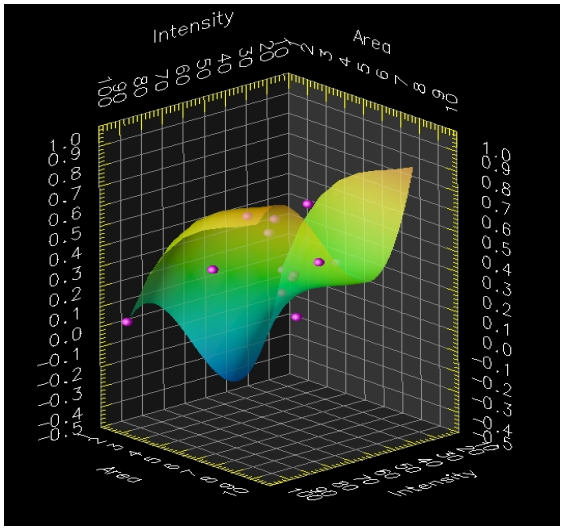
Slide: Ian Foster

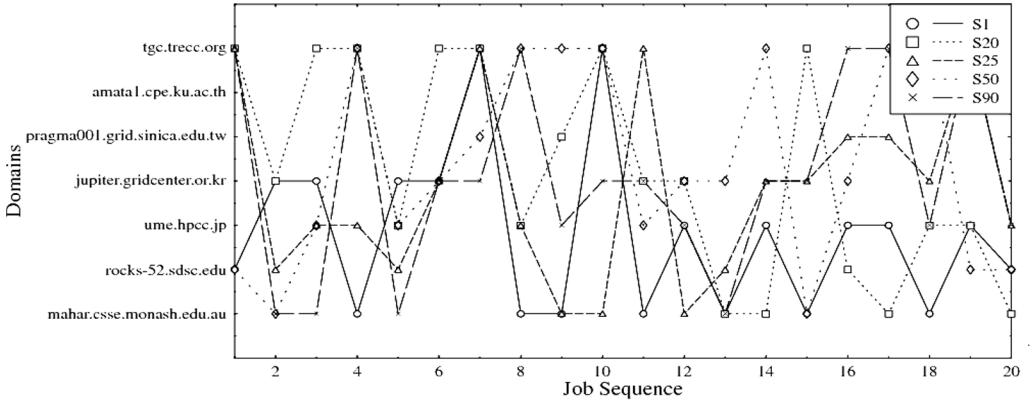
Some exemplar applications

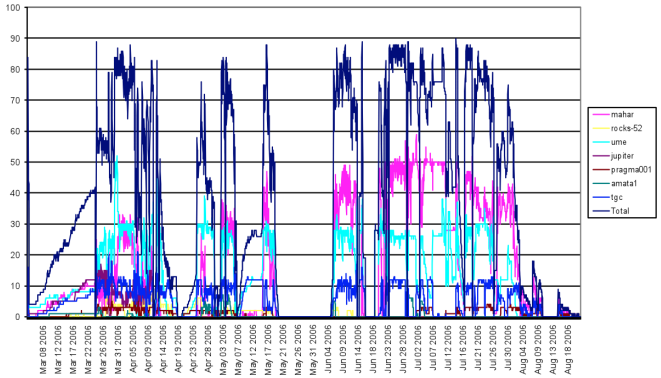
A personal perspective

Environmental Science

Lynch, Beringer, Uotila Monash U, AU







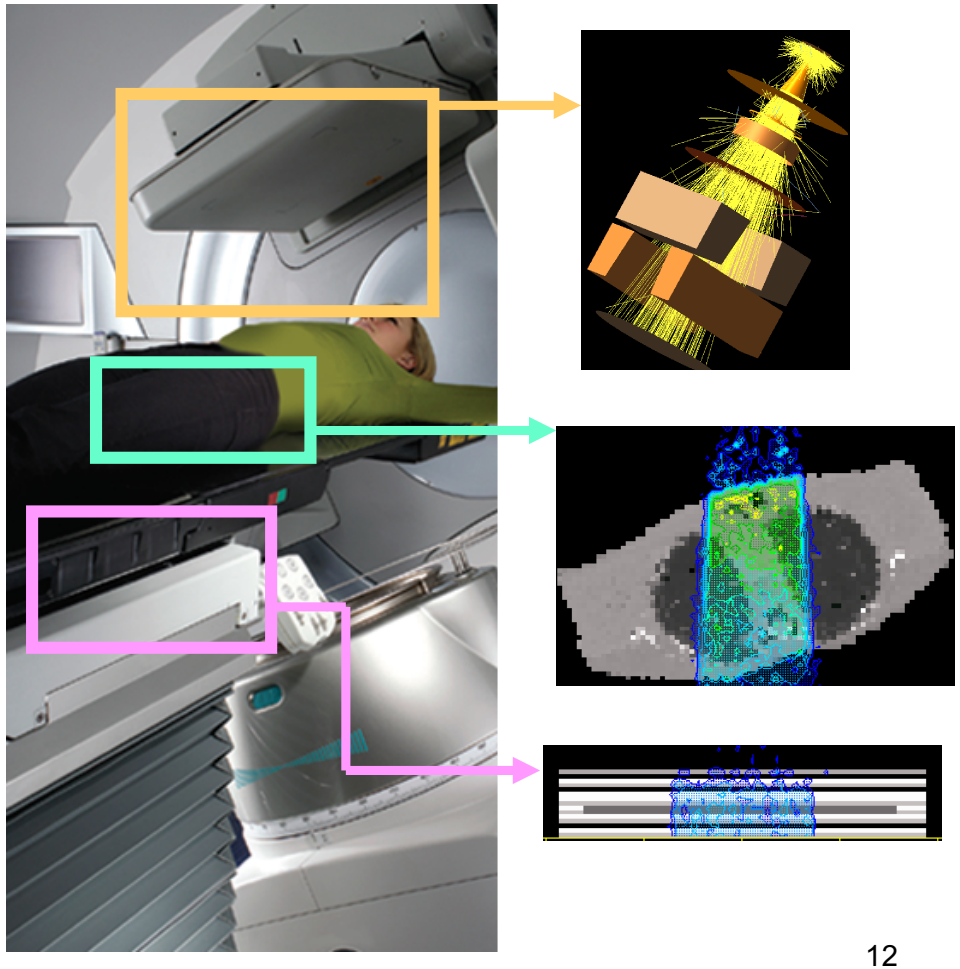
Radiotherapy planning

Giddy, Chin, Lewis, Welsh e-Science Centre, UK

BEAMnrc

EGS

DOSXYZnrc

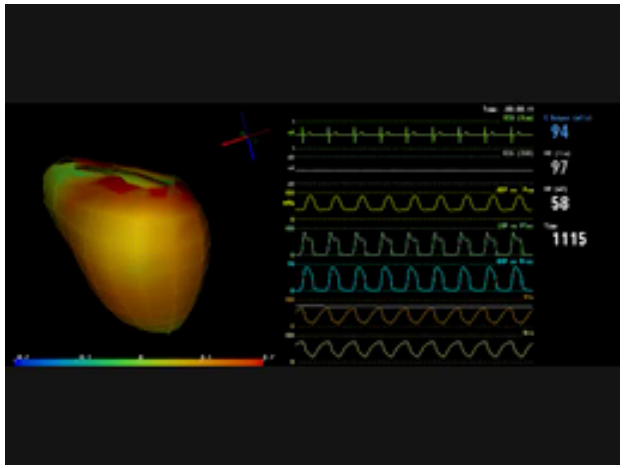


www.utsouthwestern.edu/.../270177SynergyS.2.bmp

Cardiac Science

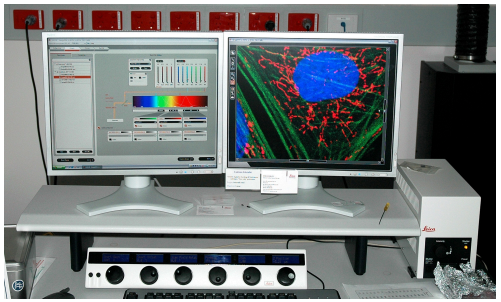
Sher, Gavaghan, Rodriguez, Oxford
Mcculloch, Mihaylova, Kerckhoffs, UCSD

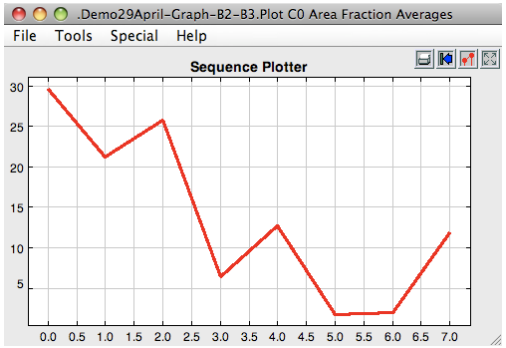
- Heart disease still leading cause of death
- Understanding the underlying physiological mechanisms is cheaper and faster when experimental studies are performed together with mathematical models & computer simulations
- Studying pathologies
- Developing & Testing drugs

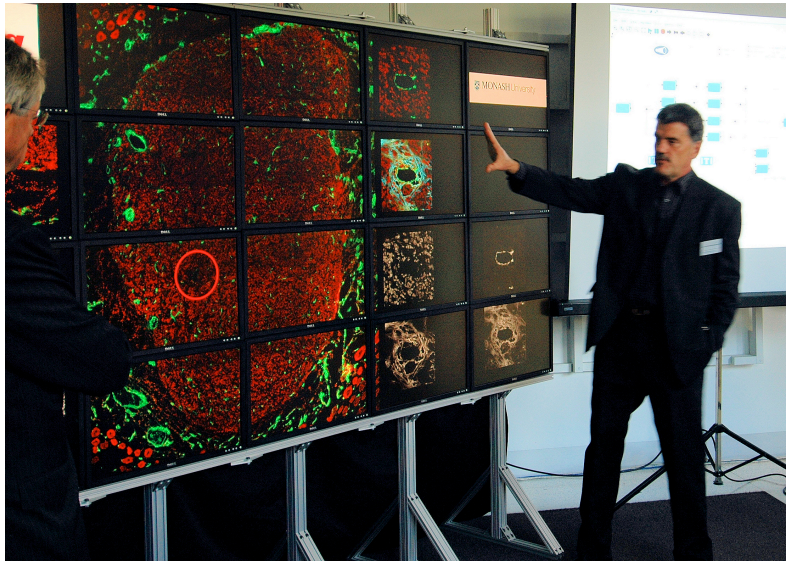


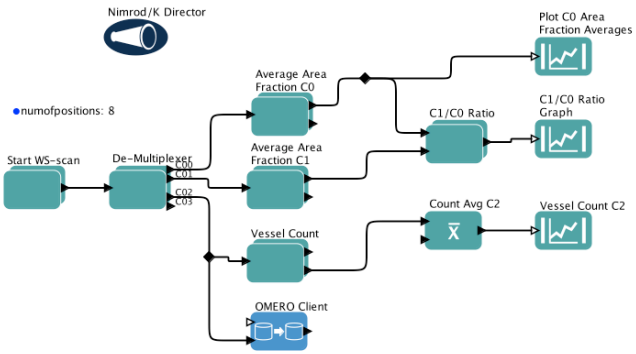
Cancer Imaging and Therapy

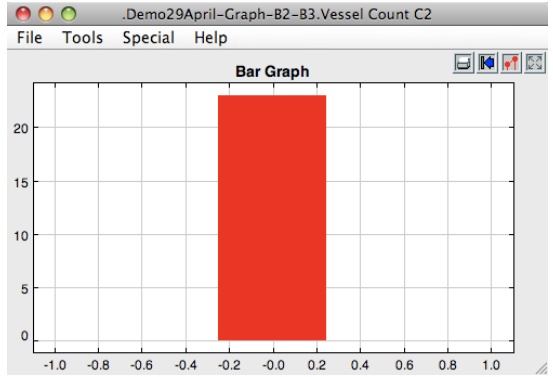
Martin Lackmann, Mary Vail







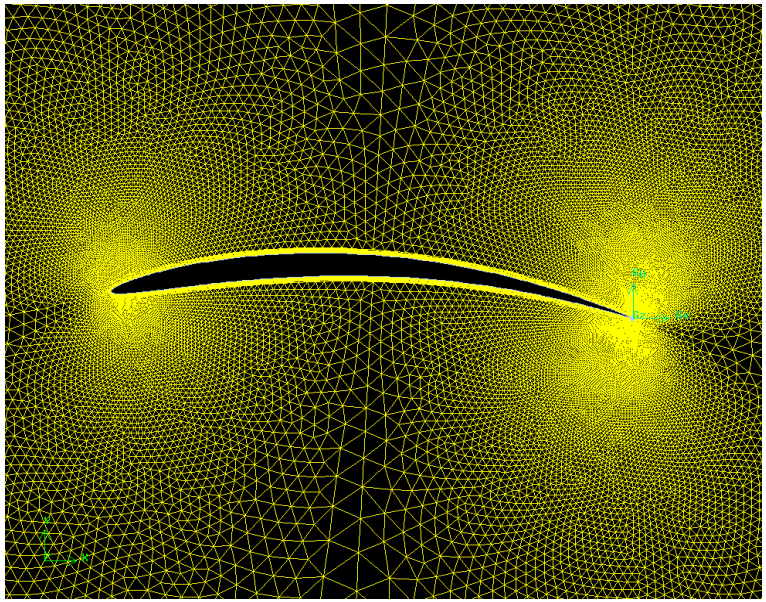


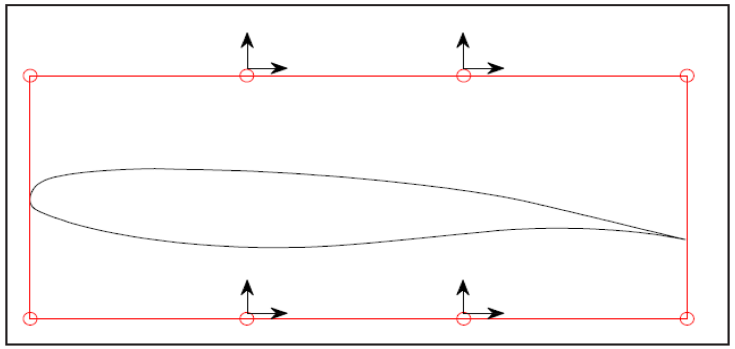


Aerodynamic Design

Kipouros, Cambridge, UK

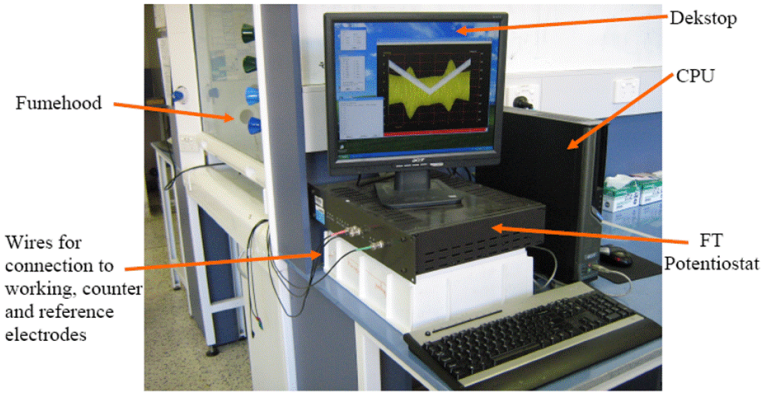
- Geometry management using Free Form Deformation – 8 design variables
- Evaluation of the aerodynamic characteristics, Cl, Cd, and Cm coefficients using Xfoil
- Investigation of the lift to drag trade-off subject to hard geometrical constraints to the thickness of the airfoil at 25% and 50% of the chord (in order to maintain practical significance to the design problem)

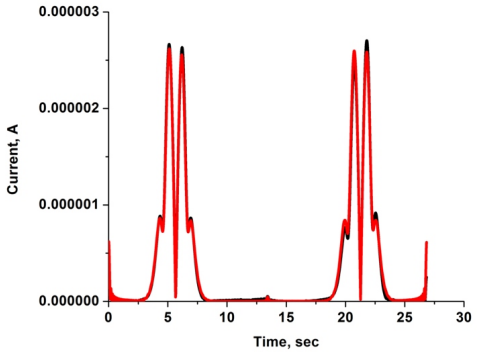


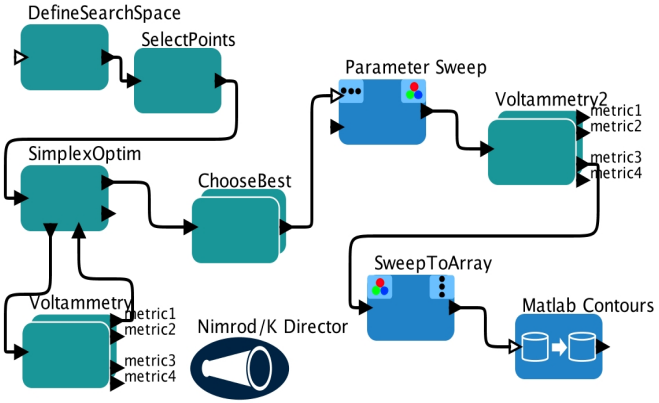


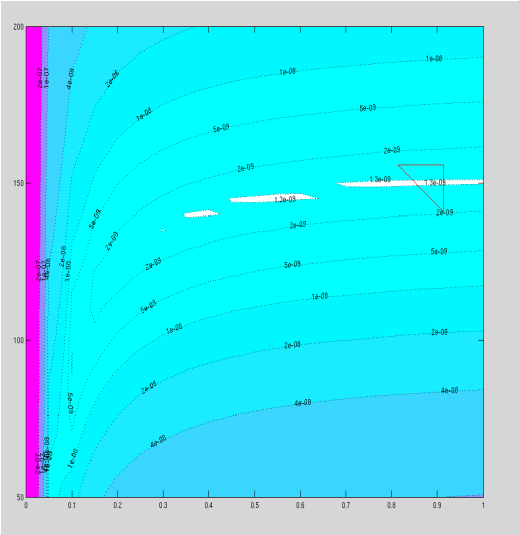
Electro-chemistry

Bond, Gavaghan: Monash, Oxford



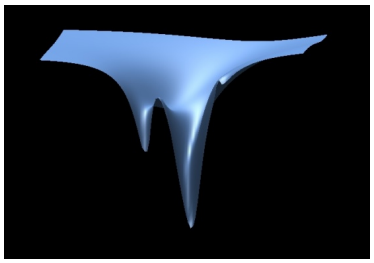


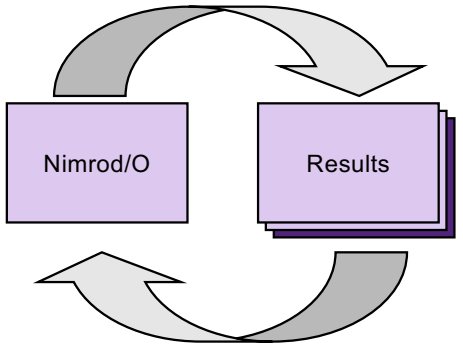


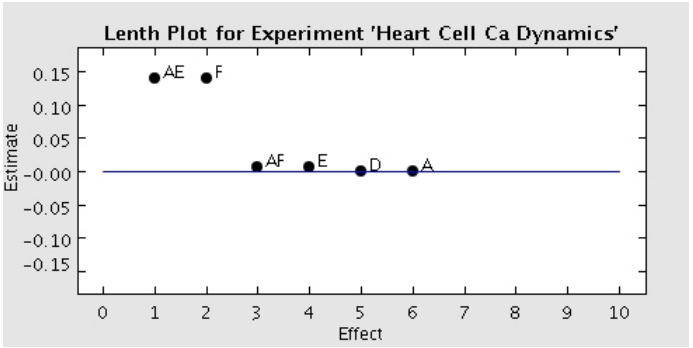


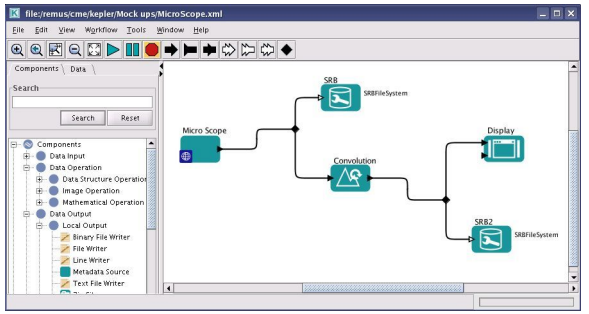
Nimrod supporting “real” science

- A full parameter sweep is the cross product of all the parameters (Nimrod/G)
- An optimization run minimizes some output metric and returns parameter combinations that do this (Nimrod/O)
- Design of experiments limits number of combinations (Nimrod/E)
- Workflows (Nimrod/K)







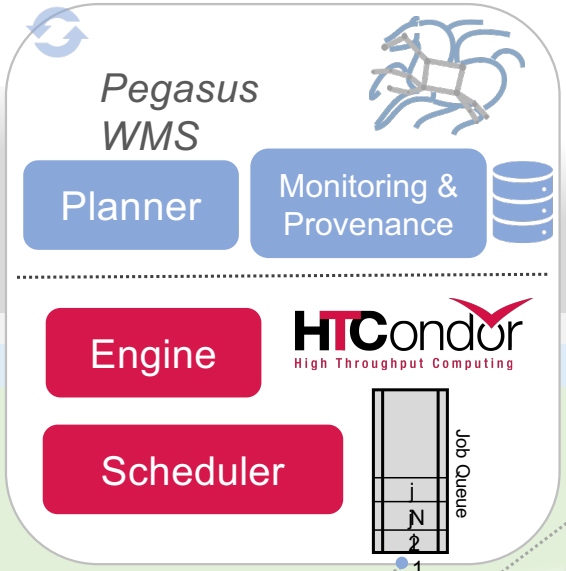


Pegasus Workflow Management System

Ewa Deelman - University of Southern California - PI - deelman@isi.edu

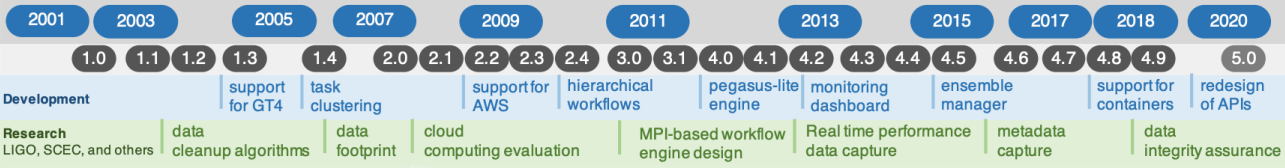
Funded Under National Science Foundation Grant #1664162

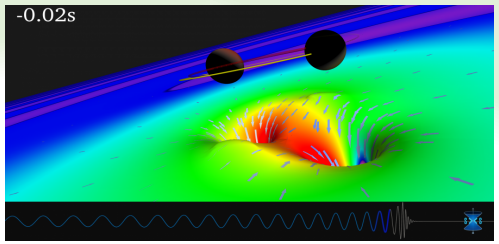
pegasus.isi.edu



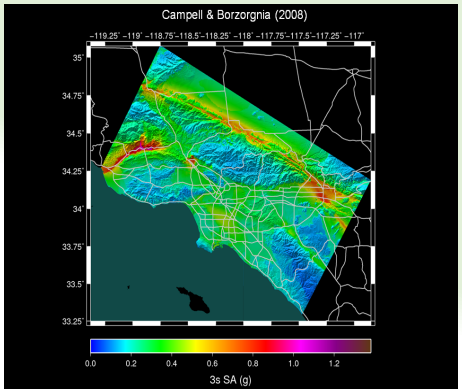
End To End Workflow Management and Execution

- Develop ***portable scientific workflows*** in Python, Java, and R
- Compile workflows to be ***run on heterogenous resources***
- Monitor and debug workflow execution via ***CLI and web-based tools***
- Recover from failures with built-in ***fault tolerance mechanisms***
- ***Regular release schedule*** incorporating latest research and development

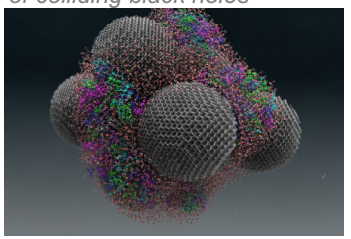




LIGO observation of colliding black holes



Hazard map indicating maximum amount of shaking at a particular geographic location generated from SCEC's CyberShake Pegasus workflow

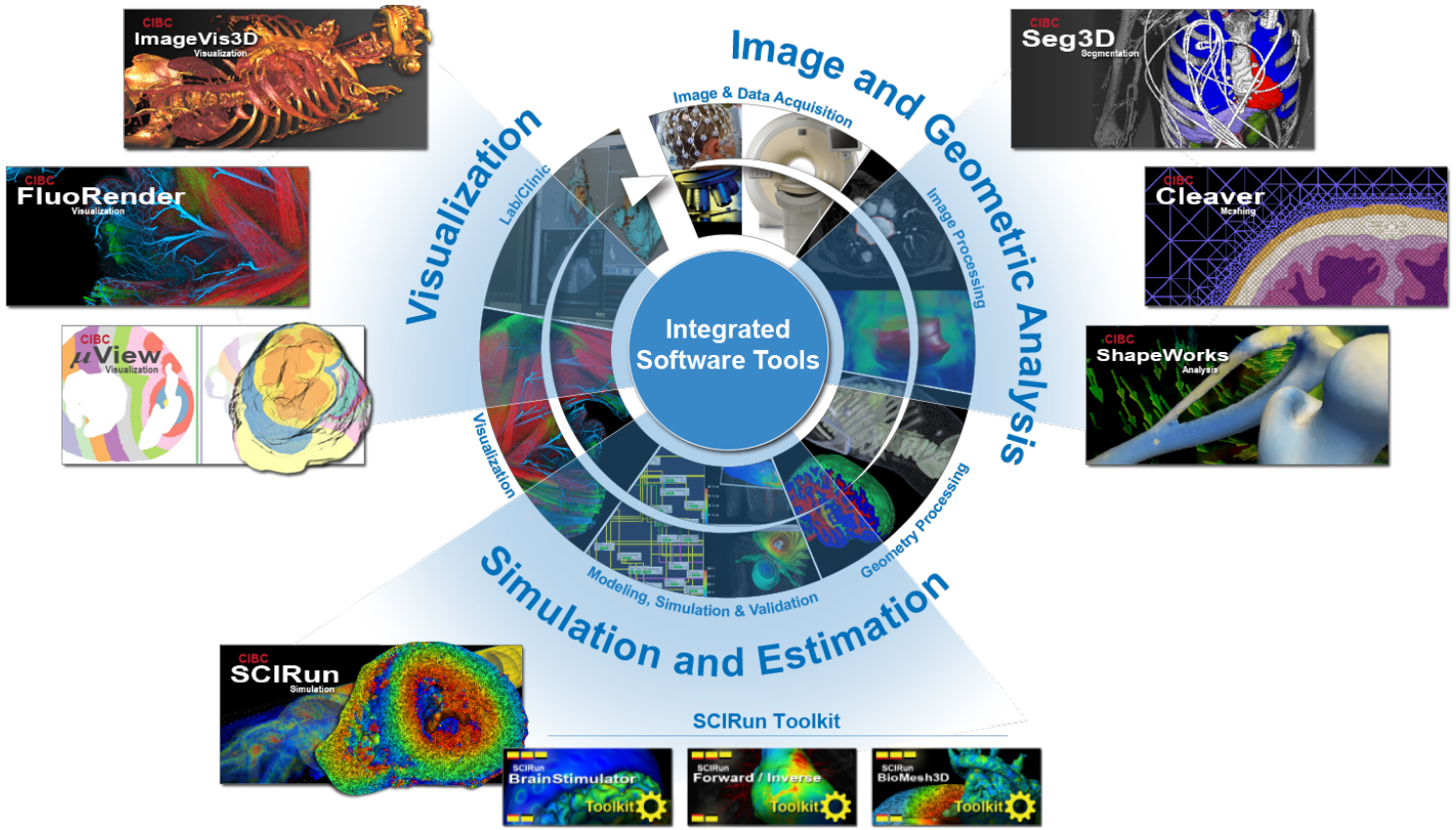


Visualization of water on nanodiamond spheres from

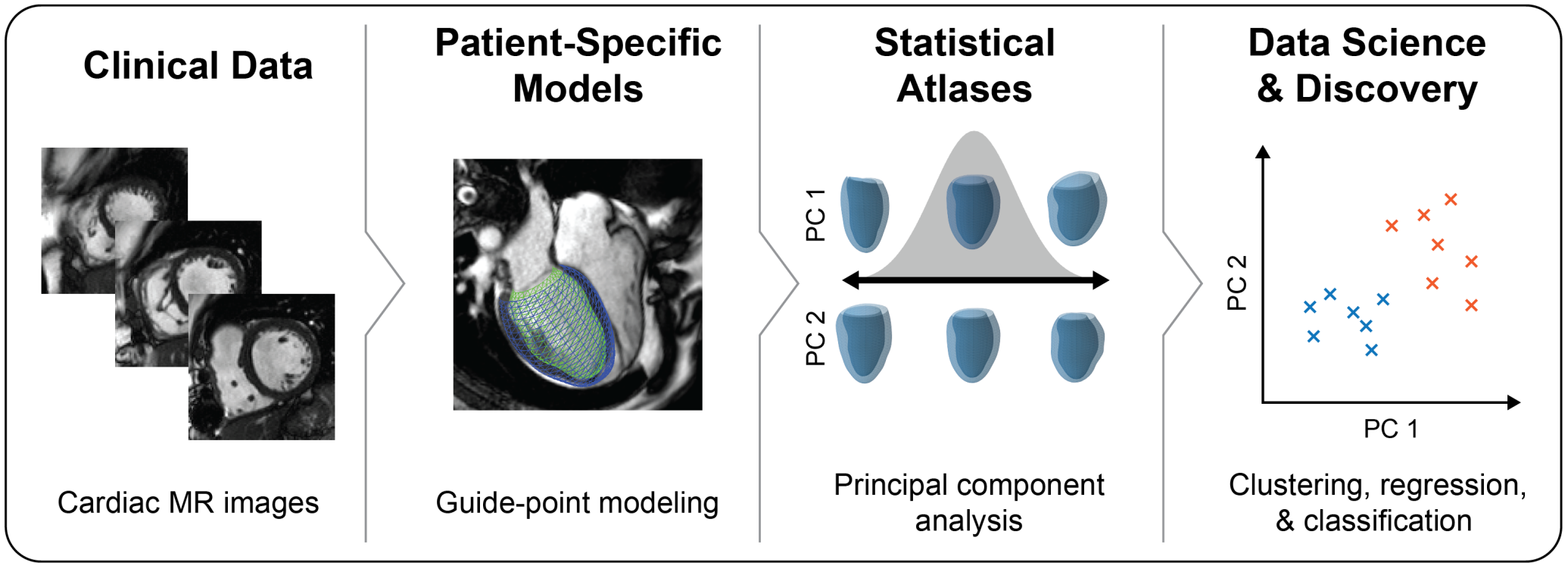
Pegasus In Practice

- Laser Interferometer Gravitational Wave Observatory (LIGO) develops large scale analysis pipelines used for gravitational wave detection.
- Southern California Earthquake Center (SCEC) CyberShake project generates hazard maps using hierarchical workflows .
- Oak Ridge National Lab (ORNL) conducted studies on tRNA and nanodiamonds to improve drug delivery design principles.

CIBC Center Vision - Software



Statistical atlases of cardiac geometry enable the discovery of novel biomarkers: McCulloch et al



Forsch et al., *Journal of Computational Science,* 2020

Thrust Research Areas: Dongara et al

INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY of TENNESSEE

- Numerical Linear Algebra Algorithms and Software
- EISPACK, LINPACK, BLAS, LAPACK, ScaLAPACK, PBLAS, Templates, ATLAS
- PLASMA (Manycore + distributed)
- MAGMA (Accelerators; Intel, Nvidia, AMD,…)
- Batched BLAS
- SLATE (Designed for the Exascale Systems)

- Heterogeneous Distributed/Cloud Computing
- MPI
- NetSolve, FT-MPI, Open-MPI

- Performance Evaluation
- Linpack Benchmark, Top500, PAPI, HPCC, HPCG, HPL-AI

TCS Important issues

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
4. PhD programs don't allocate time and resources to translation
5. Traditional publication venues don't value translation
6. There are a lack of exemplars

Translation is not commercialization

- Commercialisation almost always occurs after the research has been completed,
  - almost never funded as part of the original research proposal.
- Commercialisation implies a financial angle that has little to do with the research per-se.

Use of Open Source

- Helps with distribution of a software system, but doesn't intrinsically drive impact
- No direct link between the way the software is used, and the research program. Thus, there is no explicit feedback from lessons learned in the adoption into the research itself.
- More focussed on producing software that is maintained in a sustainable way, by building a distributed workforce.

Funding bodies don't typically support translation

- Evaluation criteria typically focus on the quality of the investigator team, the project quality and innovation, the feasibility and the benefit.
- Translation is not usually highlighted as a desirable property, thus a proposal might be marked down for including translational activities.
- A budget that allocates resources to items such as a community trial, software distribution, software maintenance, may be pruned back to the basic research program.

PhD timelines don't support translation

- Typical PhD projects in computer science follow a very standard and often rigid template.
  - Students engage in a project of interest to them
  - Execute a plan much like any other research project.
  - Milestones and deliverables include software prototypes, experiments and tests, producing publication outputs along with possibly software and data artefacts.
  - At the very least, a PhD student needs to produce a thesis.
- TR adds complexity by requiring a translation phase,
  - might extend the timeline beyond that of current PhD programs.

Traditional publication venues don’t value translation

- Many editorial boards would argue translation is secondary to their scope,
- More focussed on primary research outcomes in computer science
- Many translational research projects are interdisciplinary,
  - Outcomes might not align well with the journal’s primary focus.
- Most journals do not publish failures.

Lack of exemplars

- Numerous examples of computer science research being commercialised and adopted
- Few examples of successful translational research projects
- Changing the culture in an organisation is difficult because people don't know what a good TCS project looks like.

Funding

- Currently ad hoc funding.
- Sustained funding programs and mechanisms focussed on fostering and nurturing TCS
- More money alone not solution
- Need to *build* translation into the research plan
- Funding must be used to experiment with solutions and prototypes.
- TCS typically involves substantial interaction with end users.
- Additional travel, user engagement, and provisioning of computing resources
- Translation process feeds back into the research,
  - may be a loop of research and translation rather than a linear waterfall style of workflow
- Should be free to report on both research successes, but also translation **success** or **failures**.

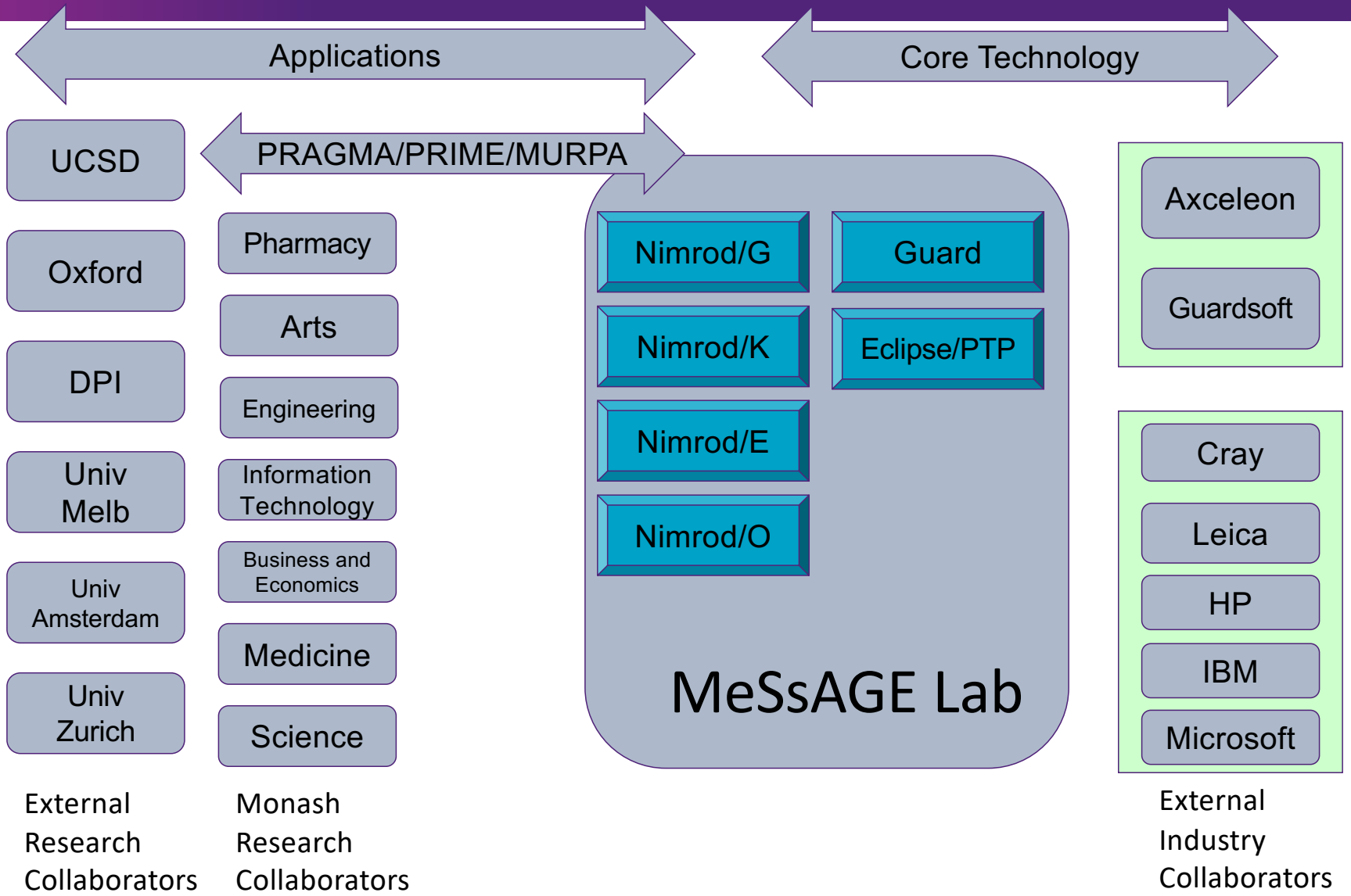
Venues, metrics and reward structures

- Traditional publications are not well suited to TCS.
  - Drawing on TM, new journals have been created that explicitly target translational medicine.
  - New set of similarly scoped journals and conferences.
  - Metrics, recognitions and rewards structures, especially in the academic community.
    - software and data and to track their use, citations and impact are a step in the right direction
  - Metrics that report uptake of their work, 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.
  - doctoral training centers 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

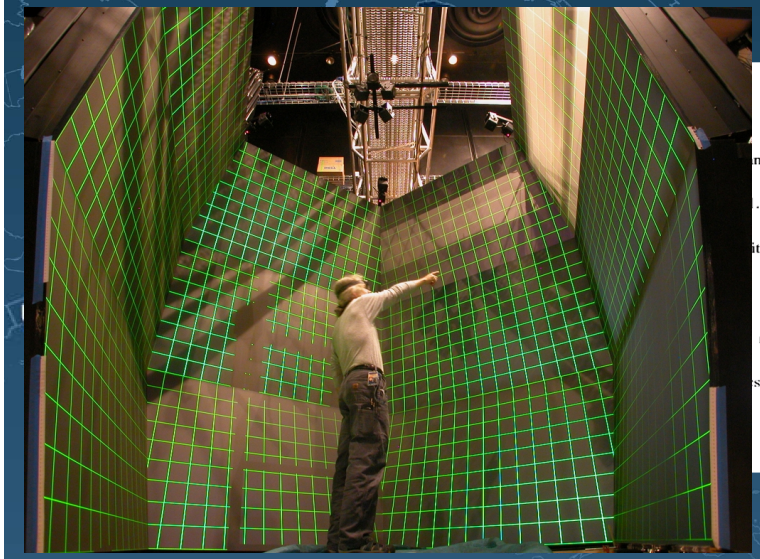
Laboratory Scale Matters

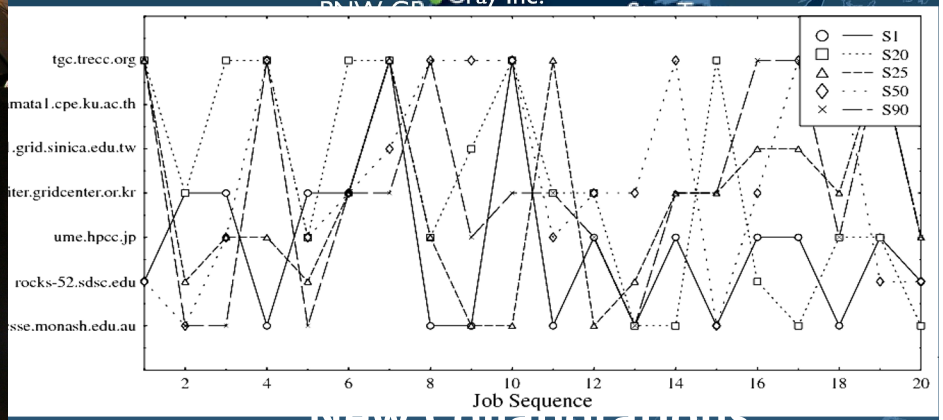

Applications
Core Technology
UCSD
PRAGMA/PRIME/MURPA
Oxford
DPI
Univ Melb
Univ Amsterdam
Univ Zurich
Pharmacy
Arts
Engineering
Information Technology
Business and Economics
Medicine
Science
Nimrod/G
Guard
Nimrod/K
Eclipse/PTP
Nimrod/E
Nimrod/O
MeSsAGE Lab
Axceleon
Guardsoft
Cray
Leica
HP
IBM
Microsoft
External Research Collaborators
Monash Research Collaborators
External Industry Collaborators


New techniques and technologies allow us to both ask, and answer, bold new questions.
This is a significant shift in the way research is conducted
Researchers need to embrace the new ways of working
RCC is a key partner in this journey

PRAGMA

A Practical Collaborative Framework





Cray Inc.

New Collaborations

Work with Science Teams to Advance Grid Technologies and Improve the Underlying Infrastructure

BeSTGRID

In the Pacific Rim and Globally

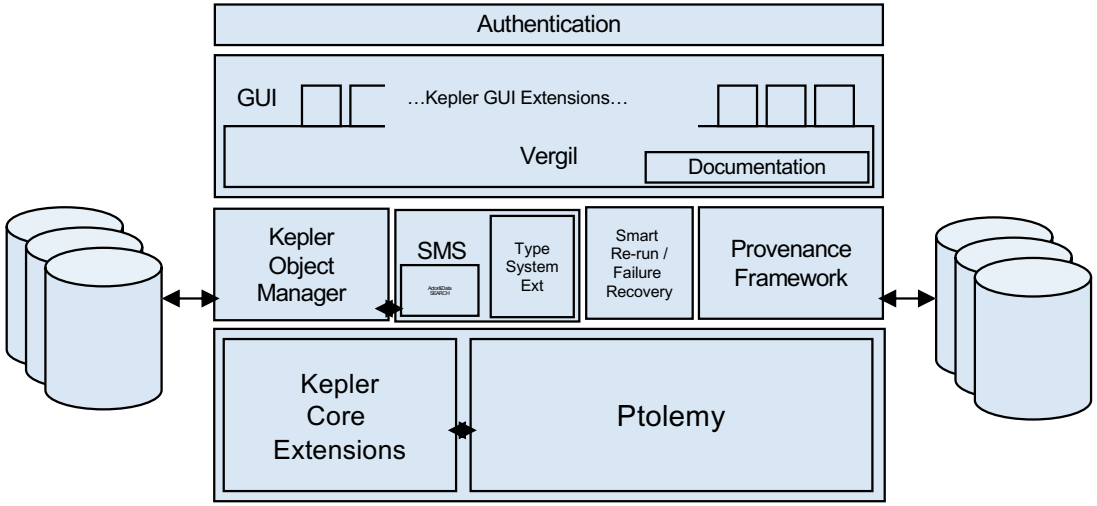
http:/

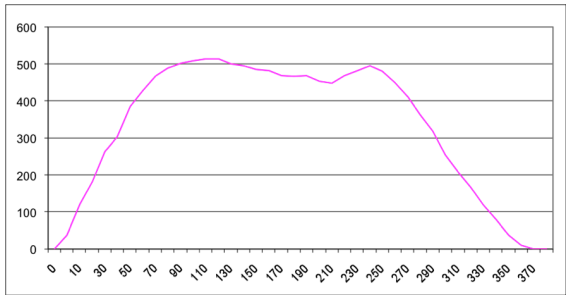
Computer Science Research

Don’t forget the research!

But are there computer science outcomes?

- Workflows
  - Complex experiments
  - Instrument control
  - Optimization
- Parallelism
- Grid and Cloud infrastructure
- Science gateways
- New optimization algorithms
- Scalable Parallel Debugging





THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
CREATE CHANGE

Thank you and Questions

David Abramson| Professor and Director
Research Computing Centre
David.Abramson@uq.edu.au
07 0000 000

facebook.com/uniofqld
Instagram.com/uniofqld
twitter.com/RCCUQ
facebook.com/rccuq

rcc.uq.edu.au