RESEARCH
COMPUTING
CENTRE

TBD

David Abramson

Director, Research Computing Centre

RESEARCH
COMPUTING
CENTRE

Workflows for Science and Engineering

David Abramson

Director, Research Computing Centre

Introduction

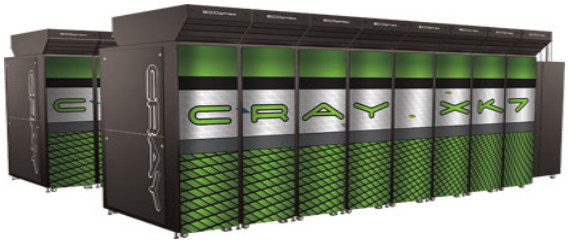
RESEARCH COMPUTING CENTRE

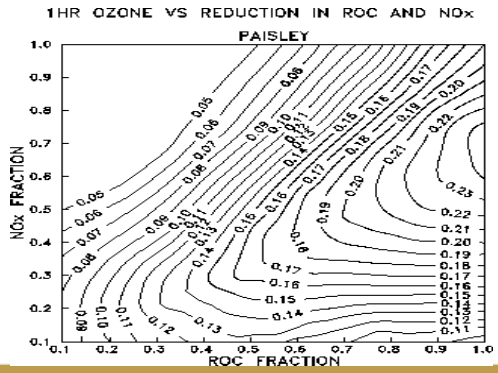
- Some Science and Engineering Optimization Problems
- The Nimrod tool family
- Scientific Workflows
- Interacting with designs
- Towards data intensive workflows
- Conclusions

Some Science and Engineering Optimization Problems

Air pollution
Cope, Victorian EPA (in 1990)

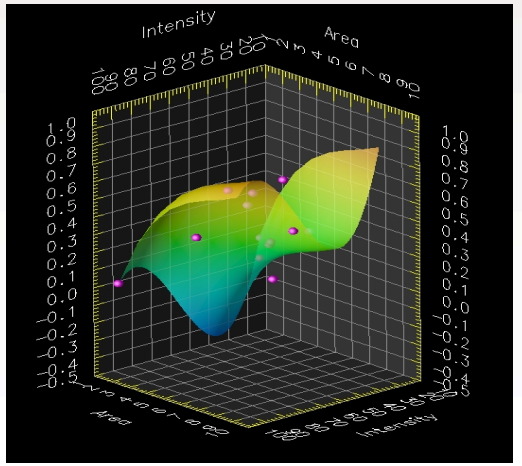
RESEARCH COMPUTING CENTRE

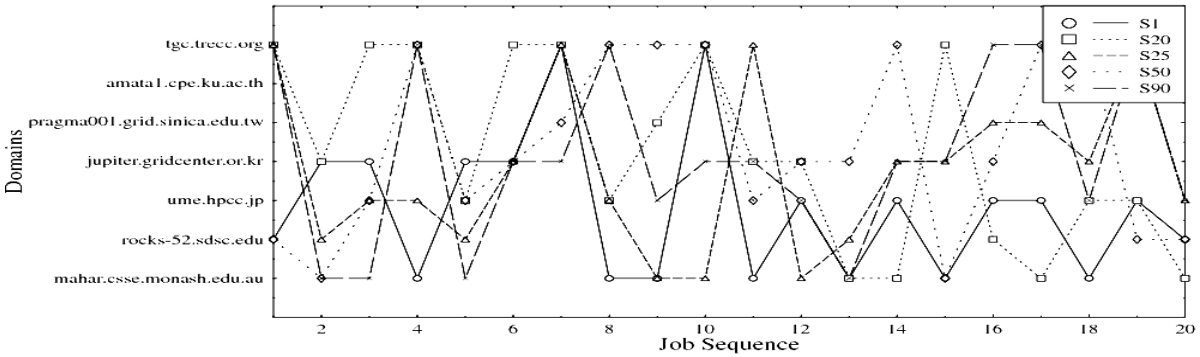


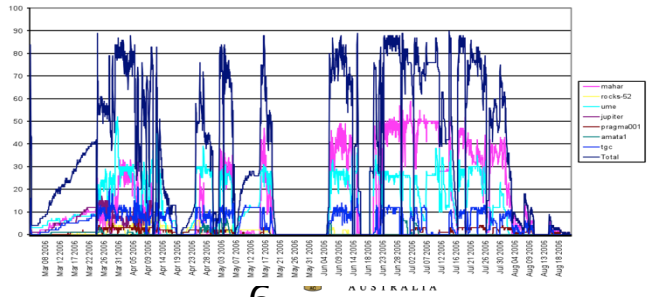


Wildfires
Lynch, Beringer, Uotila Monash U, AU

RESEARCH COMPUTING CENTRE



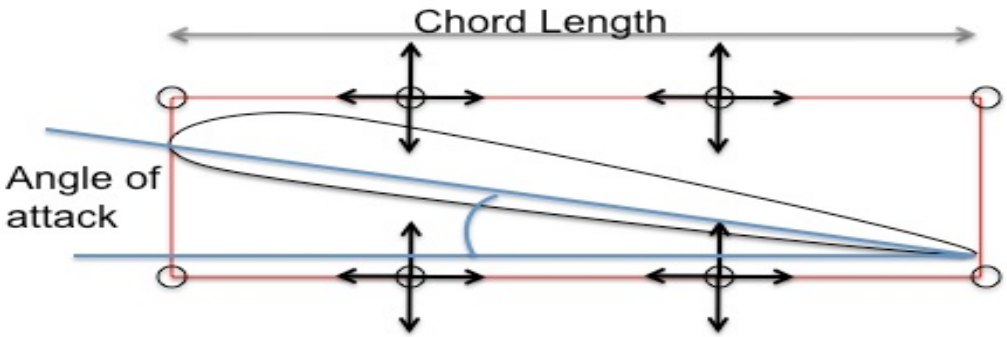




Aerofoil Design
Kipouros, Cambridge, UK

RESEARCH COMPUTING CENTRE

- 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

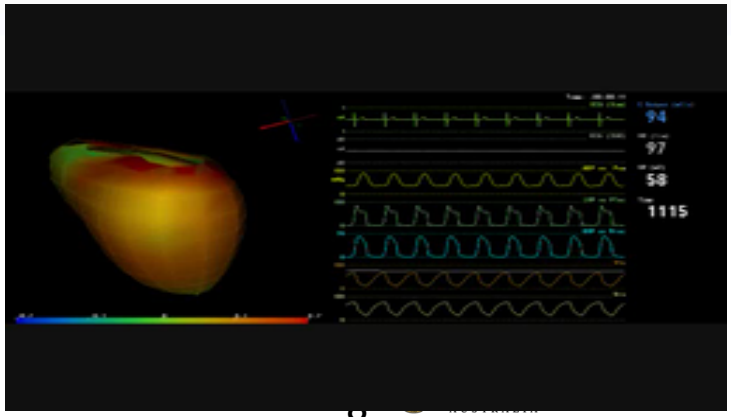


THE UNIVERSITY OF QUEENSLAND AUSTRALIA

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

RESEARCH COMPUTING CENTRE

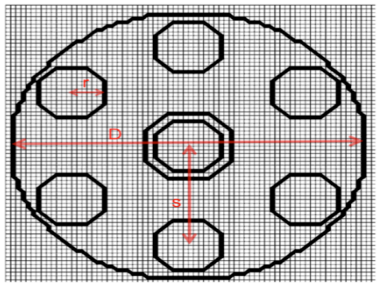
- Heart disease still leading cause of death
- Understanding the underlying physiological mechanisms is cheaper and faster when experimental studies are performed together with mathematical models & computer simulations
- Studying pathologies
- Developing & Testing drugs



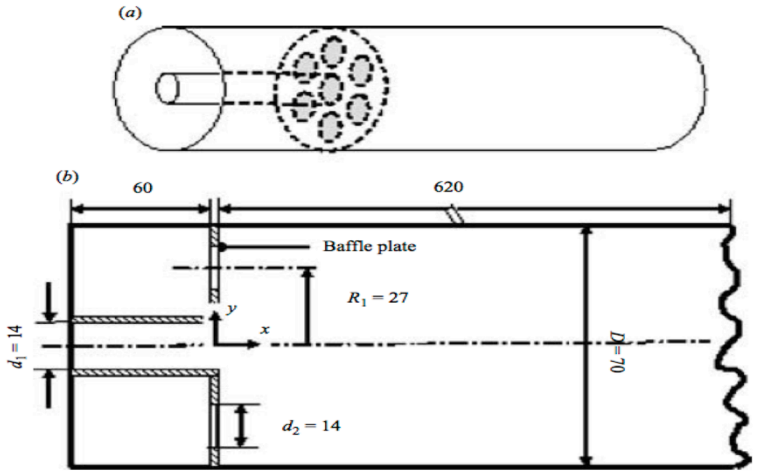
Micromixer optimization
Kipouros, Cambridge, UK

RESEARCH COMPUTING CENTRE

- Microfluidics
  - 10-9 to 10–18 litres amounts of fluids
  - Gaining importance in various fields
- Micromixer deals with mixing fluids in the smallest scale
- Active vs. passive

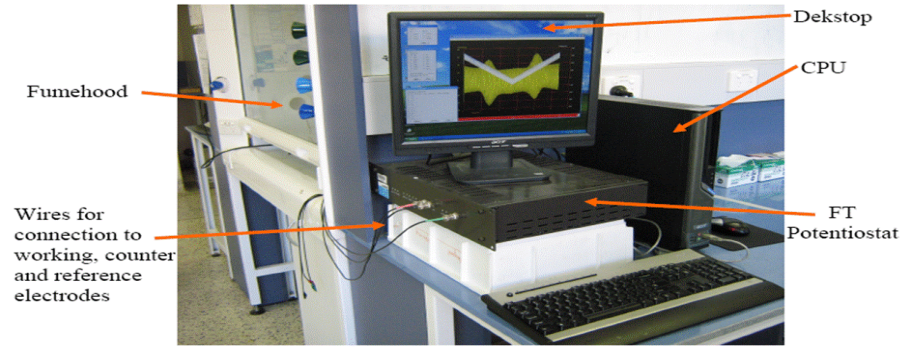


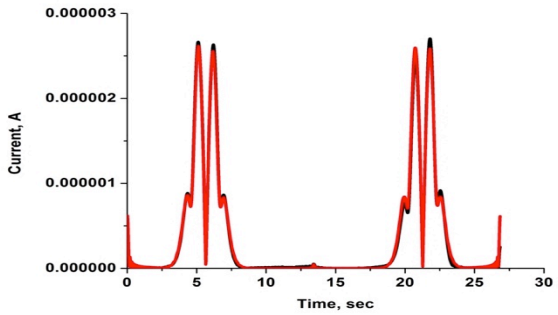
s >= 2*r + 1
s <= 34 - r

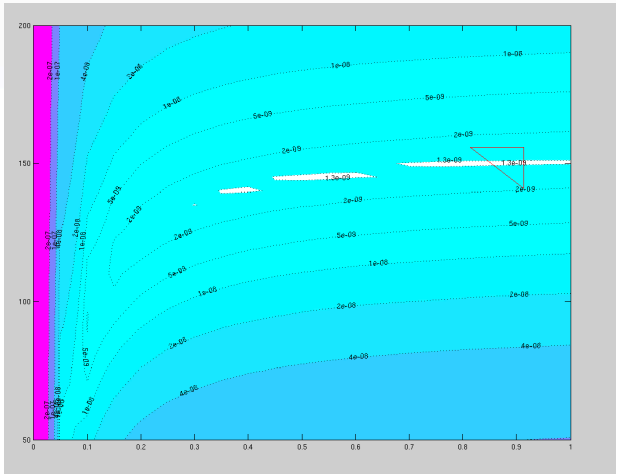


THE UNIVERSITY OF QUEENSLAND AUSTRALIA

Electro-chemistry
Bond, Gavaghan: Monash, Oxford





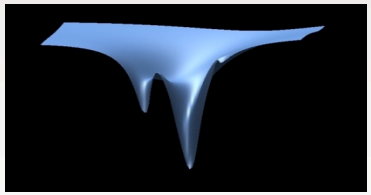


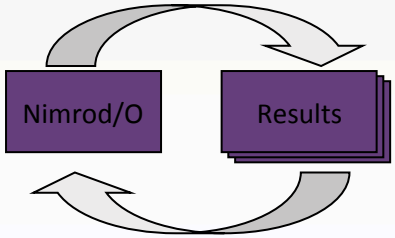
The Nimrod Tools Family

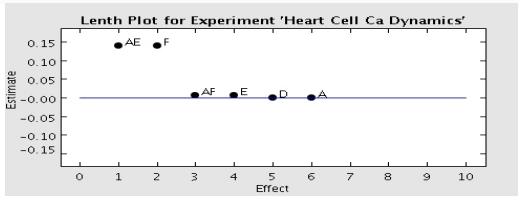
Nimrod supporting “real” science and engineering

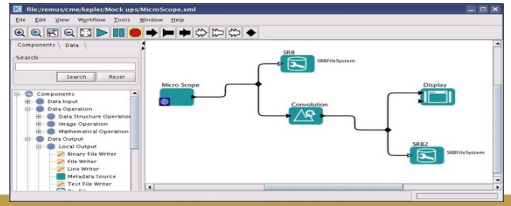
RESEARCH COMPUTING CENTRE

- 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)









Legacy Nimrod family

Plan File

Nimrod Portal

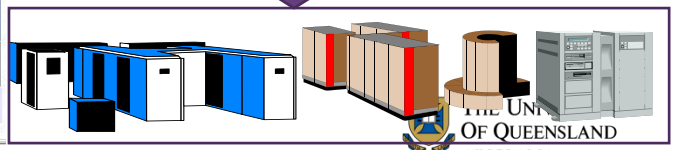
Nimrod/O

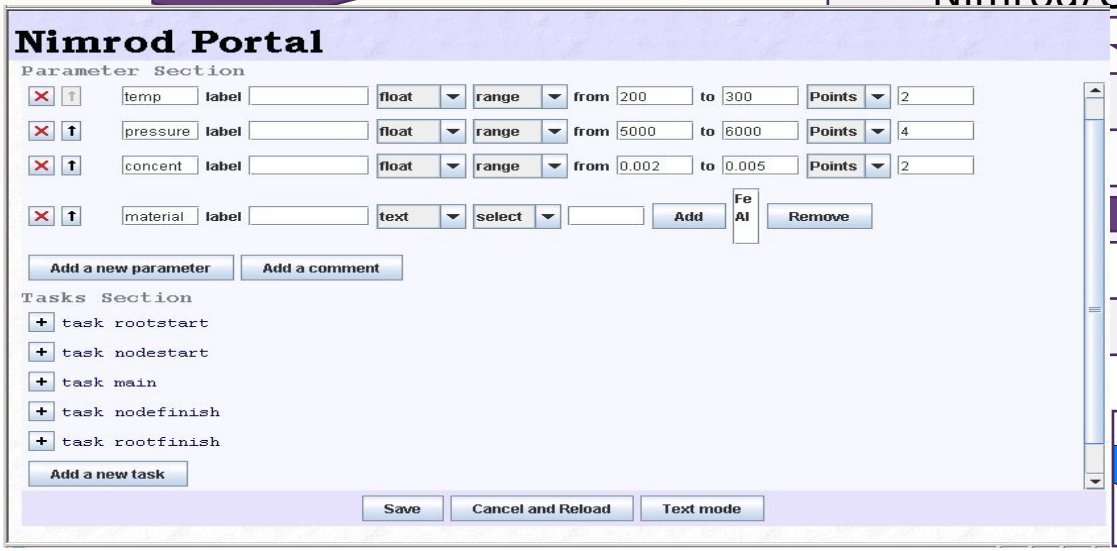
Nimrod/E

Nimrod/G

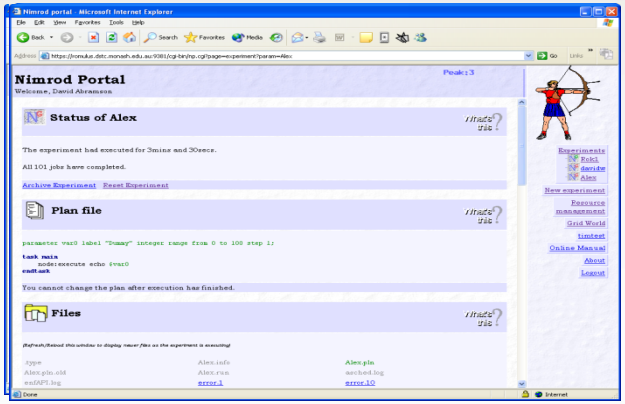
Actuators

Grid Middleware

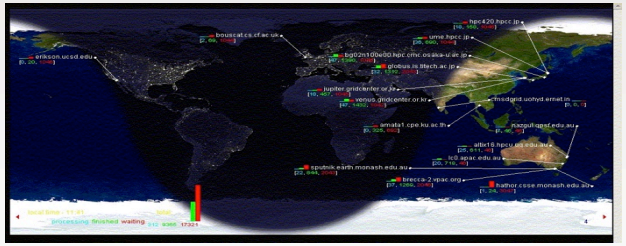




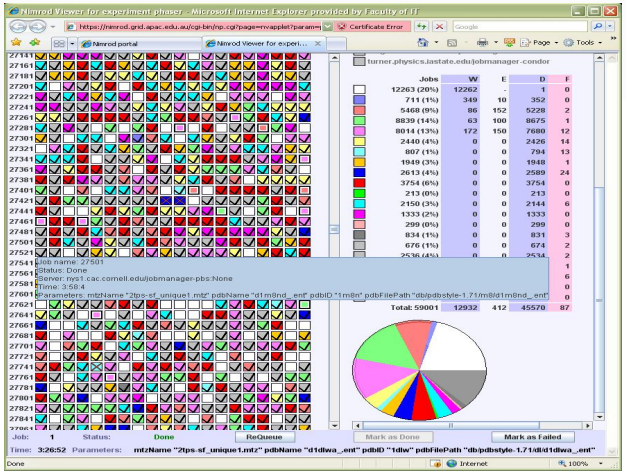
Nimrod Development Cycle



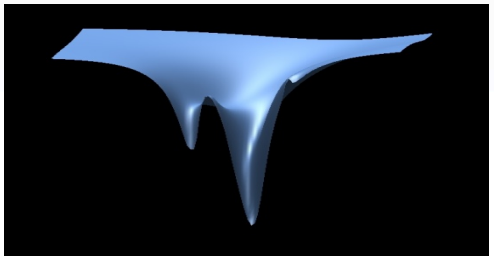
Prepare Jobs using Portal



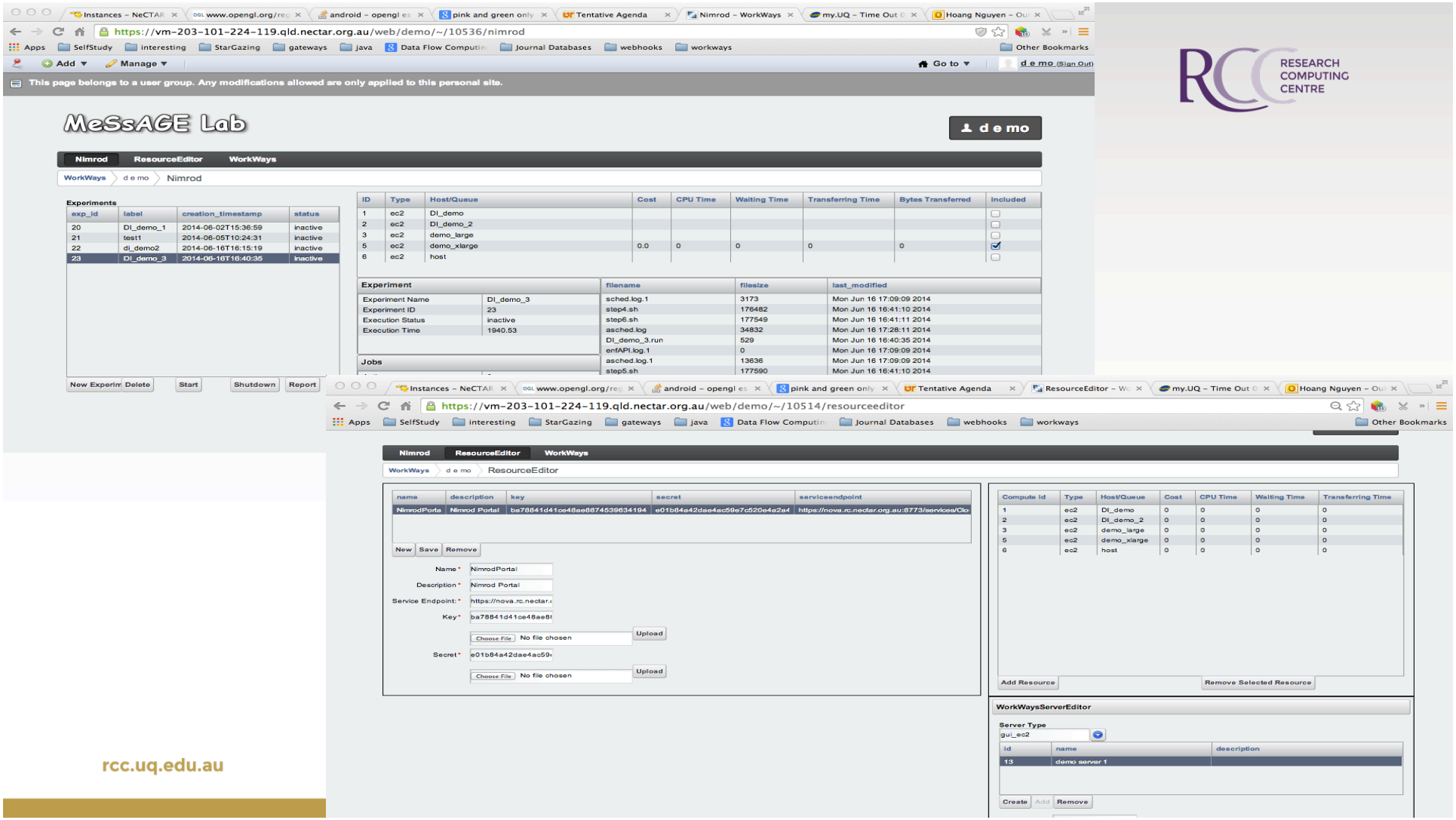
Sent to available machines



Jobs Scheduled Executed Dynamically

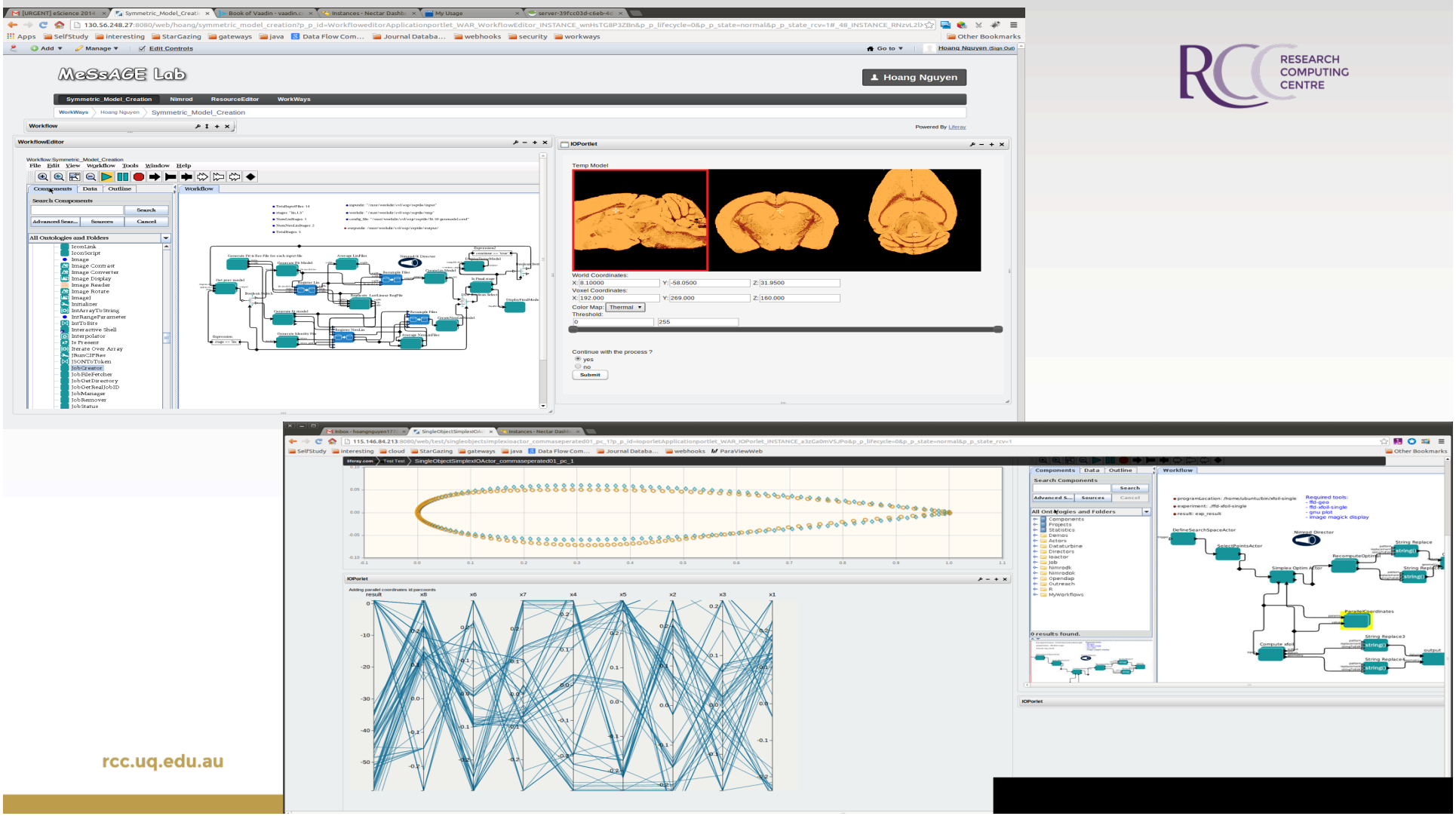


Results displayed & interpreted



RESEARCH COMPUTING CENTRE

rcc.uq.edu.au

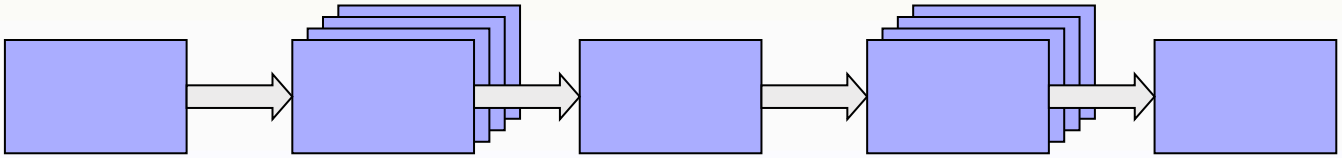

MeSsAGE Lab
Hoang Nguyen
Symmetric_Model_Creation
Nimrod
ResourceEditor
WorkWays
WorkflowEditor
IOPortlet
Temp Model
World Coordinates:
Voxel Coordinates:
Color Map: Thermal
Threshold:
Continue with the process ?
yes
no
Submit
RESEARCH COMPUTING CENTRE
rcc.uq.edu.au

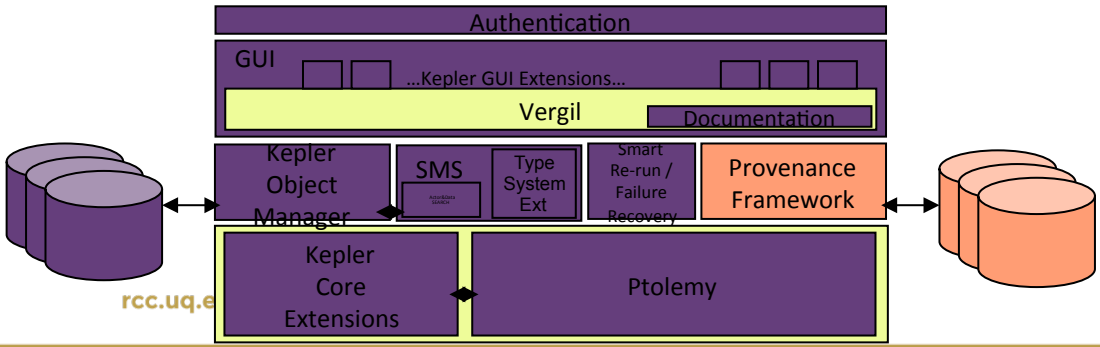
Scientific Workflows and Nimrod

Nimrod/K Workflows

RESEARCH COMPUTING CENTRE

- Nimrod/K integrates Kepler with
  - Massively parallel execution mechanism
  - Special purpose function of Nimrod/G/O/E
  - General purpose workflows from Kepler

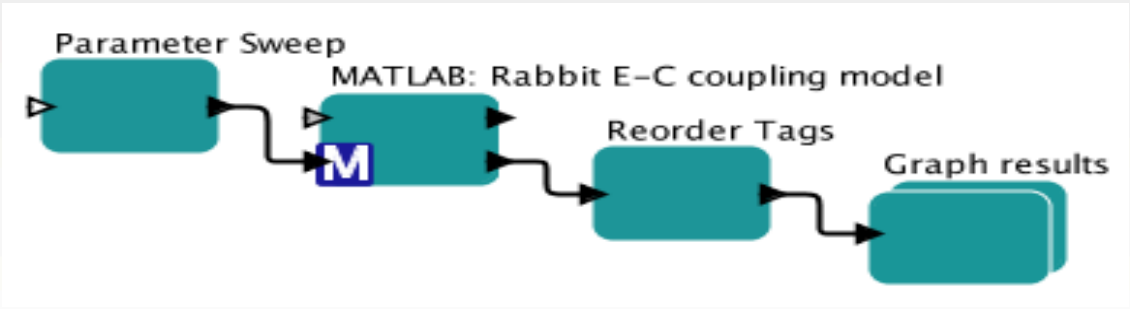


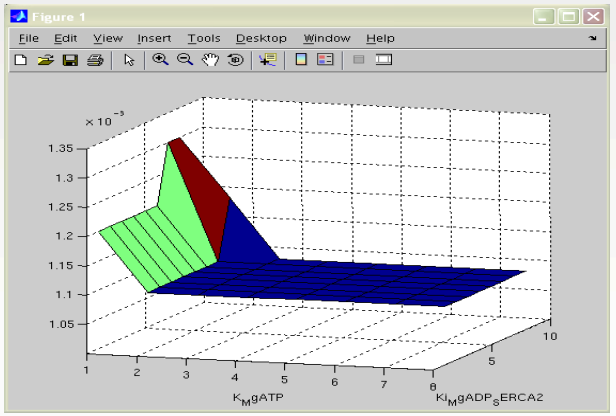


Workflow Threading

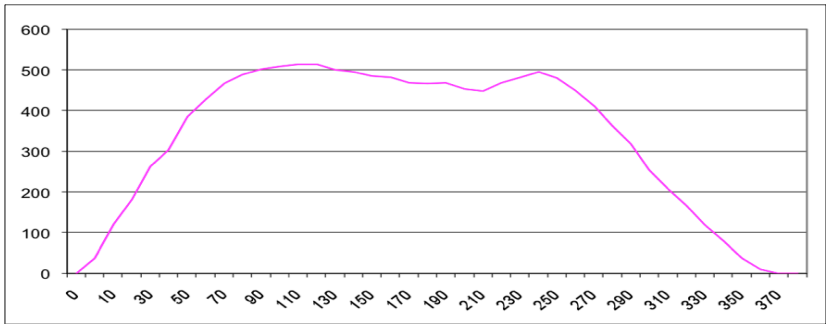
- Nimrod parameter combinations can be viewed as threads
- Multi-threaded workflows allow independent sequences in a workflow to run concurrently
  - This might be the whole workflow, or part of the workflow
- Tokens in different threads do not interact with each other in the workflow

Complete Parameter Sweep

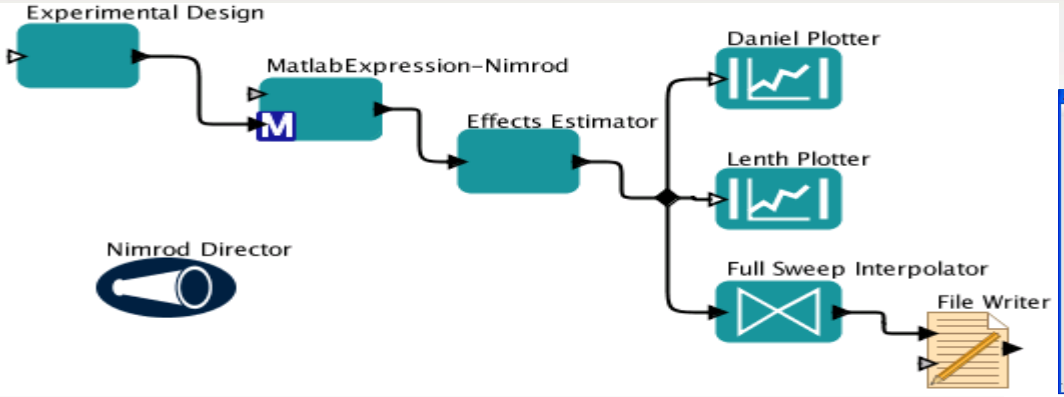


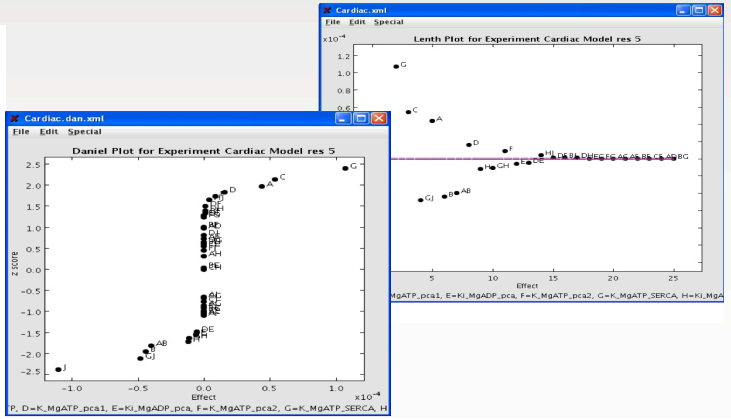


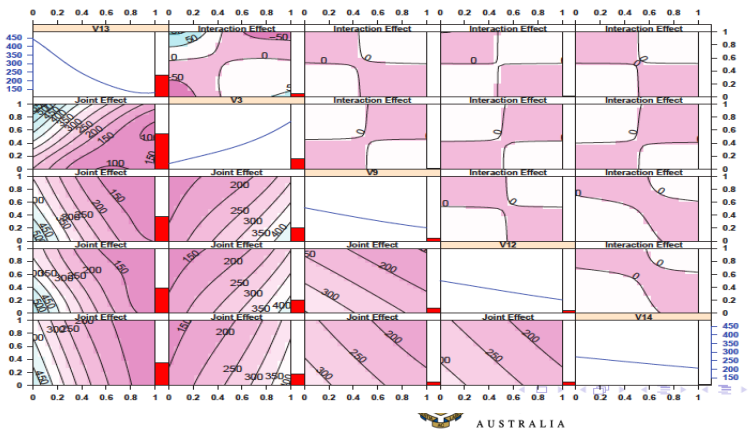
- Using a MATLAB actor provided by Kepler
- Local spawn
  - Multiple thread ran concurrently on a computer with 8 cores (2 x quads)
  - Workflow execution was just under 8 times faster
- Remote Spawn
  - 100's – 1000's of remote processes



Nimrod/EK Actors



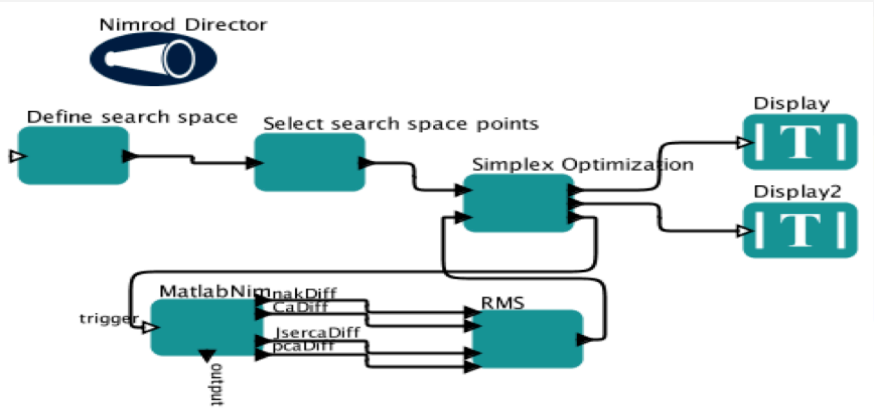




- Actors for generating and analyzing designs
- Leverage concurrent infrastructure

Nimrod/OK Workflows

- Nimrod/K supports parallel execution
- General template for search
  - Built from key components
- Can mix and match optimization algorithms



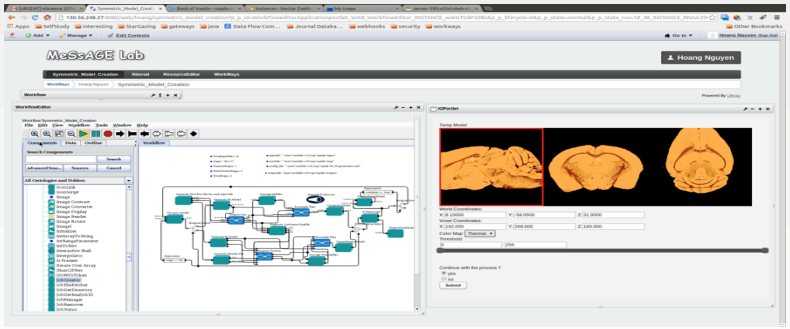
Interacting with Workflows

Workways (Hoang Nguyen)

WorkWays

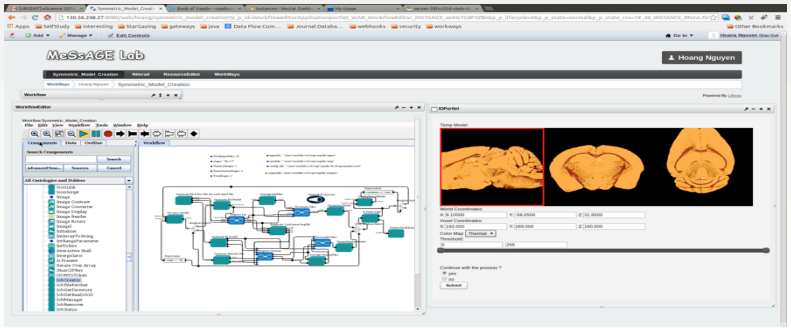
RESEARCH COMPUTING CENTRE

- Ease of use of Science Gateway
- Workflows as service
- IO through portlets
- Extensibility
  - Different IO mechanisms, protocols and topolgy
  - Different UI clients
- Currently Kepler as the workflow engine



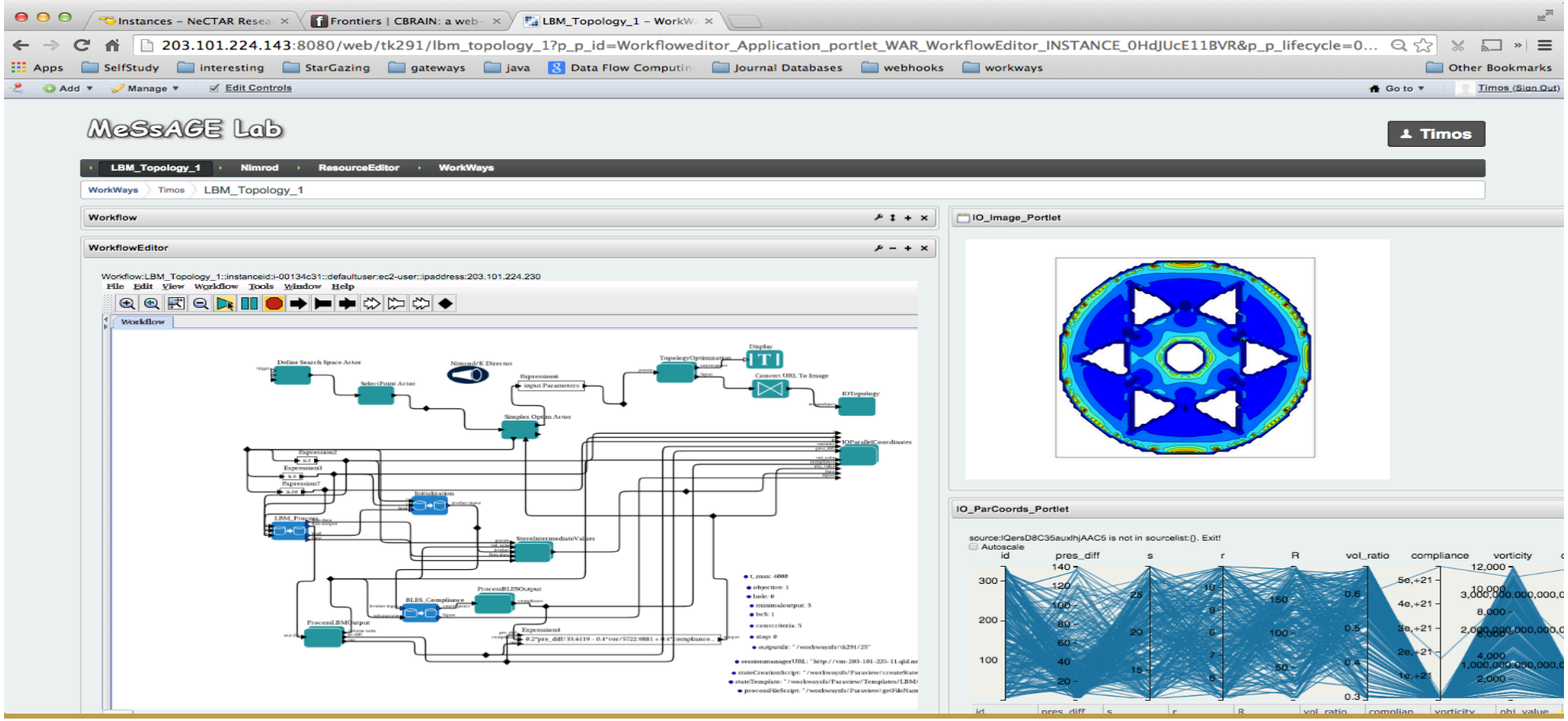
WorkWays

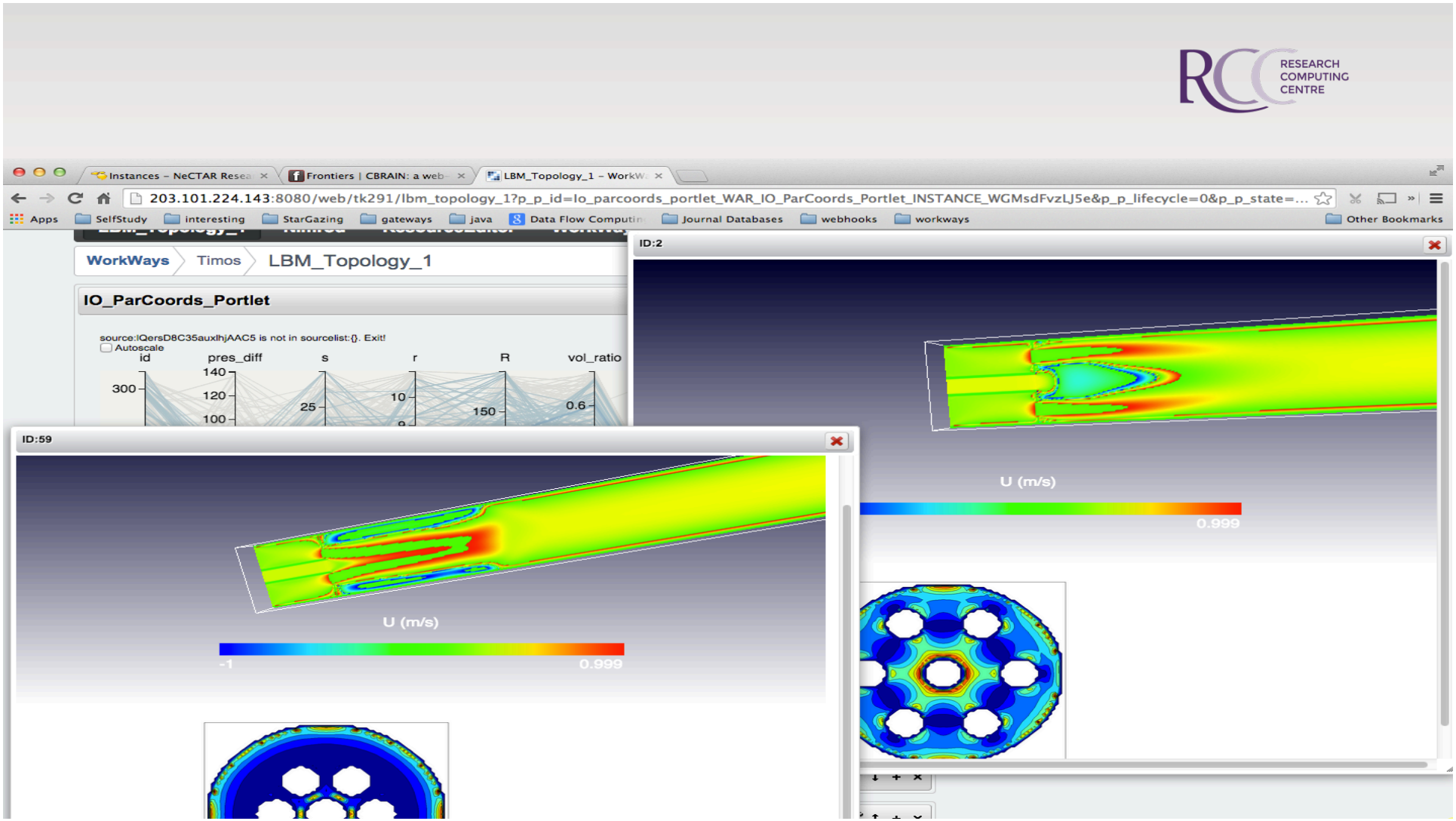
- Leverage various existing technologies
  - Kepler workflow
  - Nimrod family toolkit
  - Liferay portal
- Virtual desktops
- AAF
- Various Web-based visualization tools
  - Parallel coordinates
  - Para-view Web

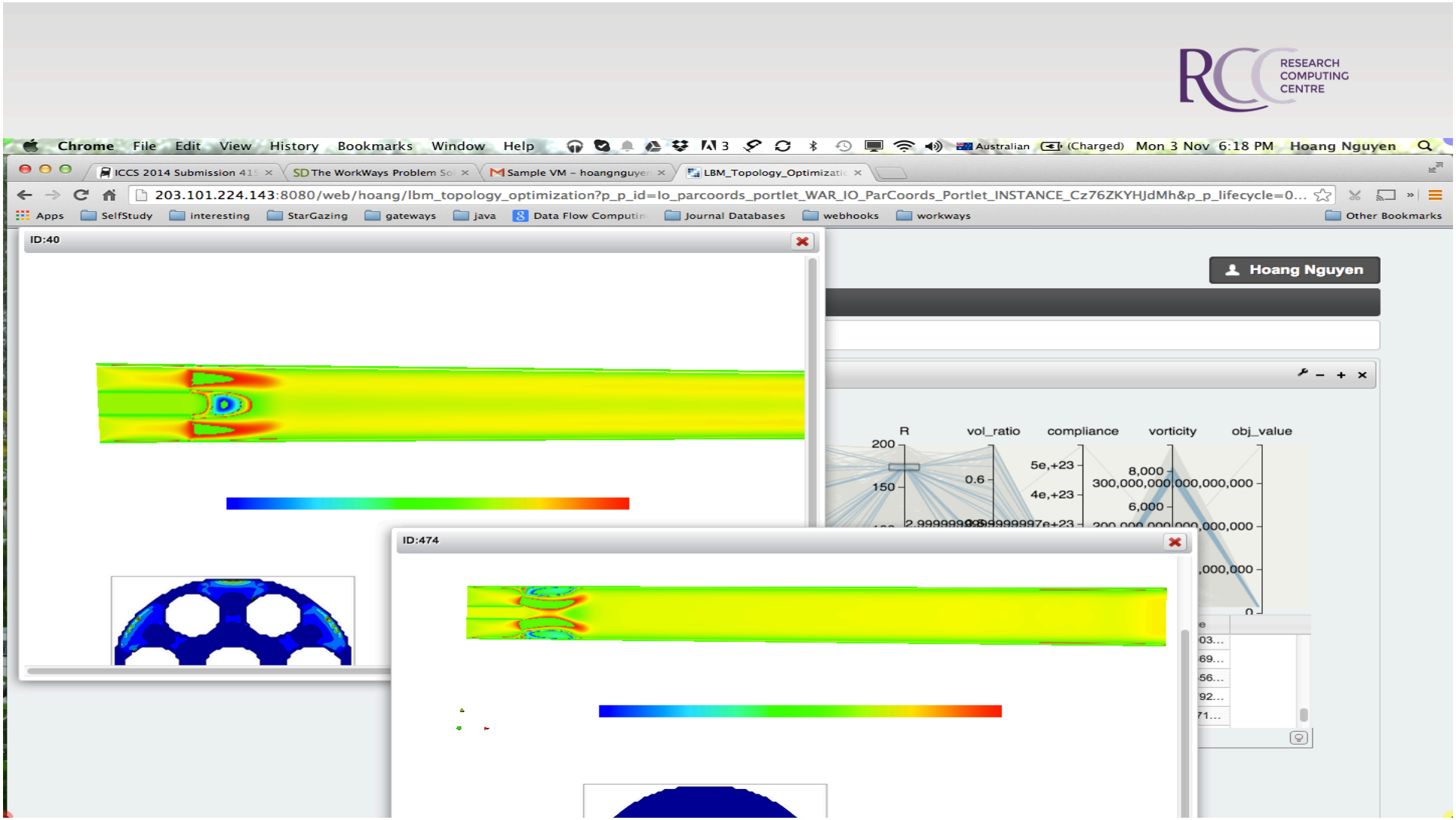


Micro-mixer design

RESEARCH COMPUTING CENTRE

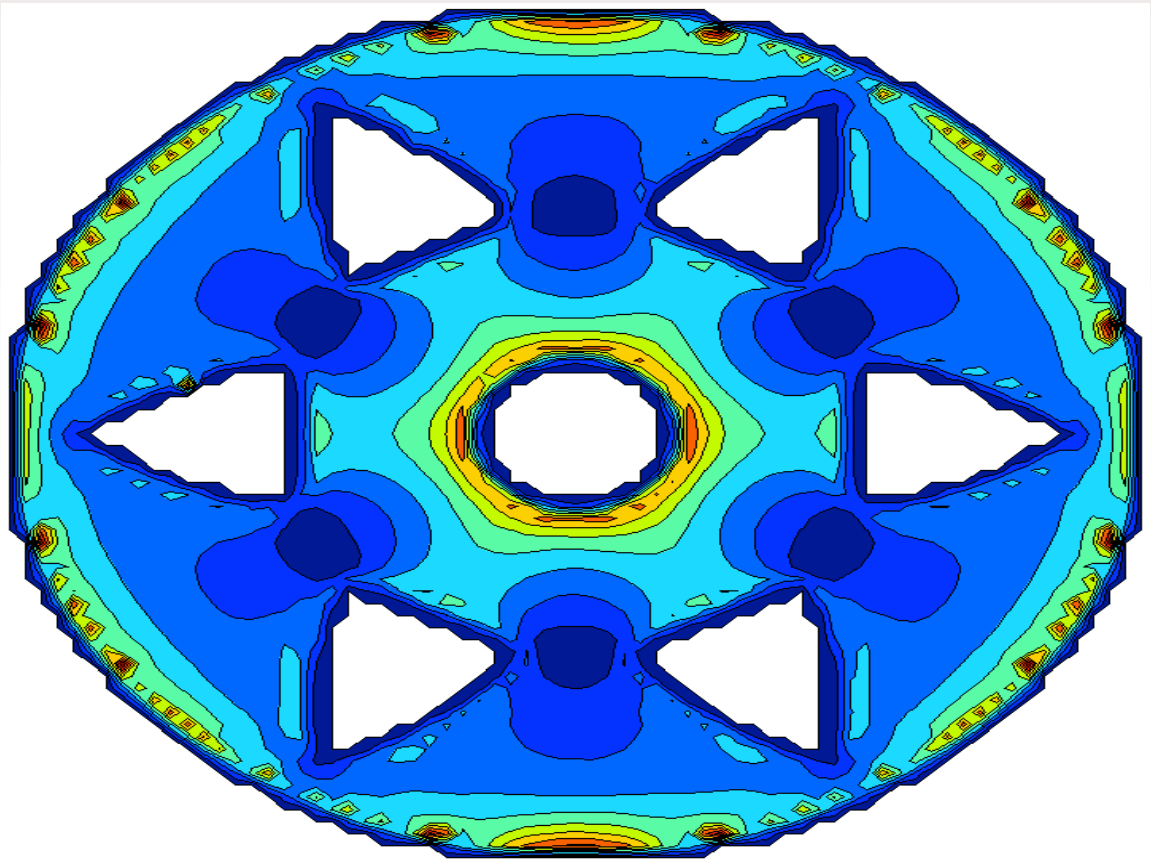



RESEARCH
COMPUTING
CENTRE
Instances – NeCTAR Rese
Frontiers | CBRAIN: a web
LBM_Topology_1 – WorkW
203.101.224.143:8080/web/tk291/lbm_topology_1?p_p_id=io_parcoords_portlet_WAR_IO_ParCoords_Portlet_INSTANCE_WGMsdFvzLJ5e&p_p_lifecycle=0&p_p_state=...
Apps
SelfStudy
interesting
StarGazing
gateways
java
Data Flow Computin
Journal Databases
webhooks
workways
Other Bookmarks
WorkWays
Timos
LBM_Topology_1
IO_ParCoords_Portlet
source:IQersD8C35auxlhjAAC5 is not in sourcelist:{}. Exit!
Autoscale
id
pres_diff
s
r
R
vol_ratio
300
140
120
100
25
10
150
0.6
ID:2
U (m/s)
0.999
ID:59
U (m/s)
-1
0.999

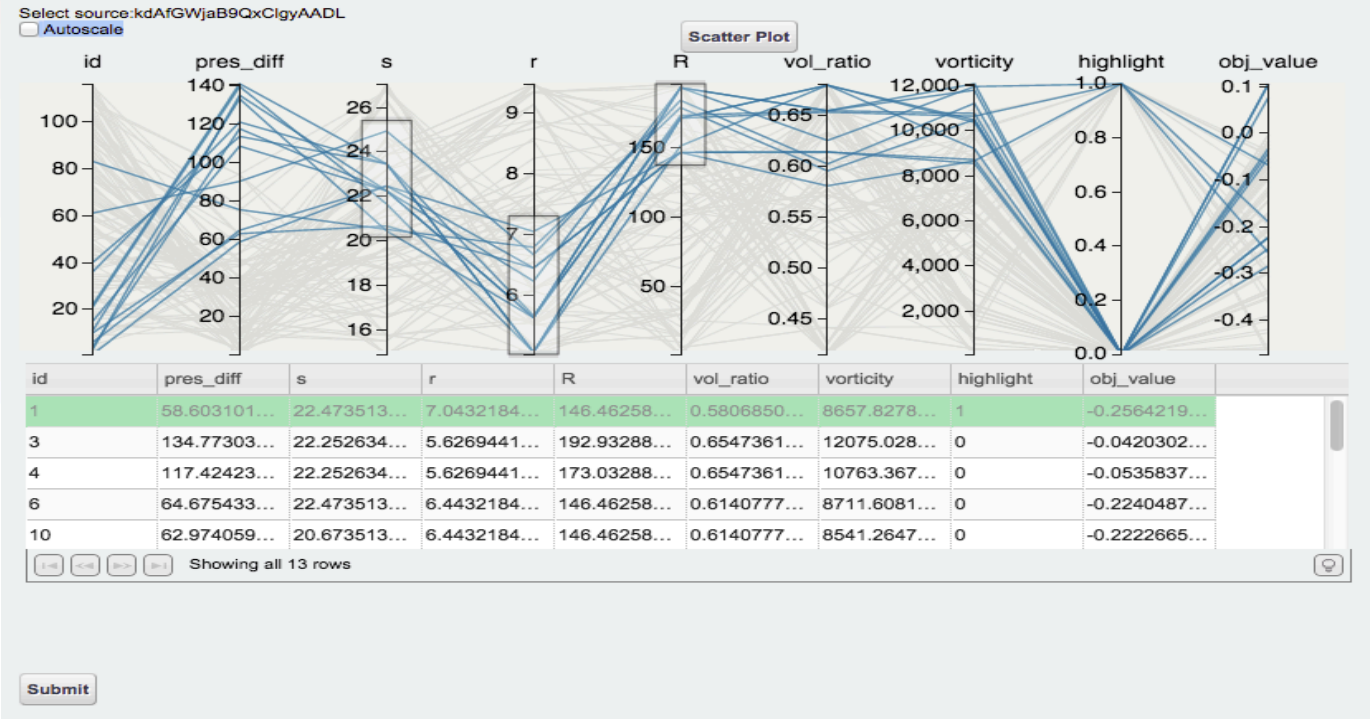

RESEARCH COMPUTING CENTRE
Chrome File Edit View History Bookmarks Window Help
Australian (Charged) Mon 3 Nov 6:18 PM Hoang Nguyen
ICCS 2014 Submission 41
The WorkWays Problem
Sample VM – hoangnguyen
LBM_Topology_Optimizatic
203.101.224.143:8080/web/hoang/lbm_topology_optimization?p_p_id=lo_parcoords_portlet_WAR_IO_ParCoords_Portlet_INSTANCE_Cz76ZKYHJdMh&p_p_lifecycle=0...
Apps SelfStudy interesting StarGazing gateways java Data Flow Computin Journal Databases webhooks workways Other Bookmarks
ID:40
Hoang Nguyen
R
vol_ratio
compliance
vorticity
obj_value
200
150
0.6
5e,+23
4e,+23
8,000
6,000
0
ID:474

Optimal Topology

RESEARCH COMPUTING CENTRE

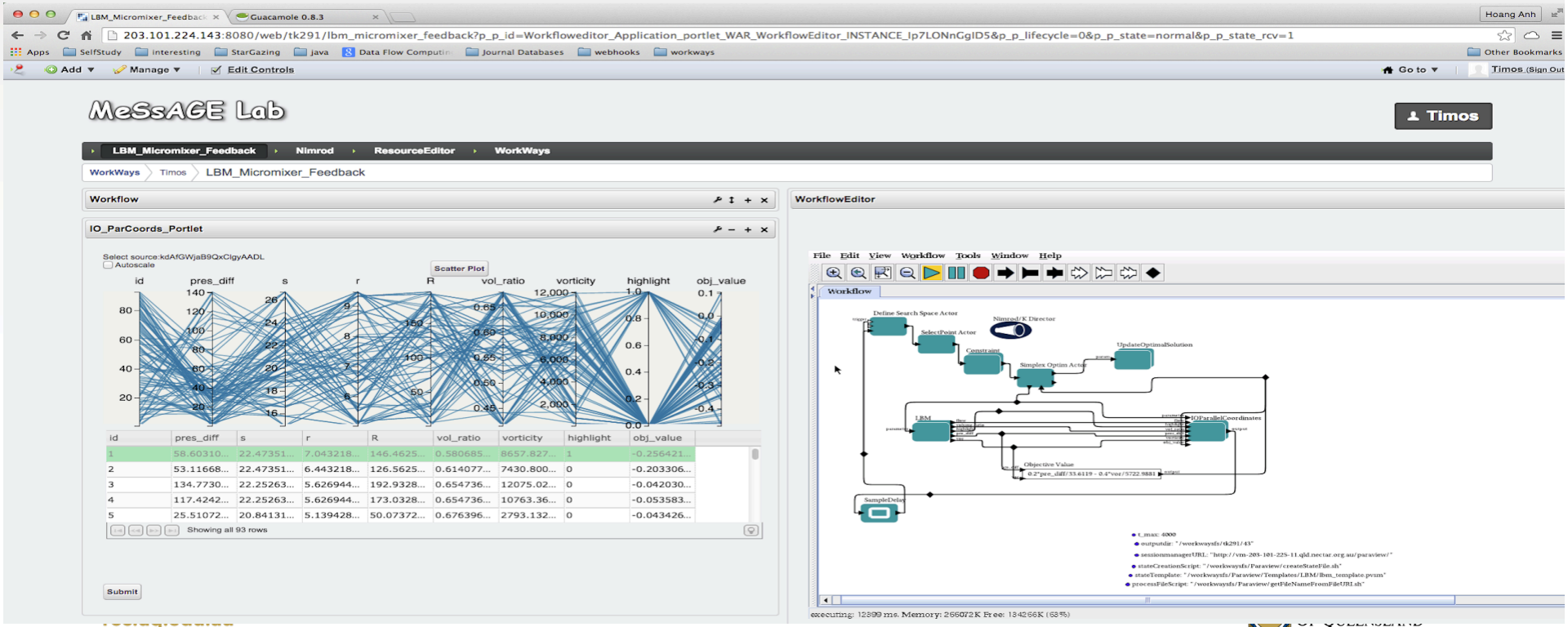


Parallel Coordinates

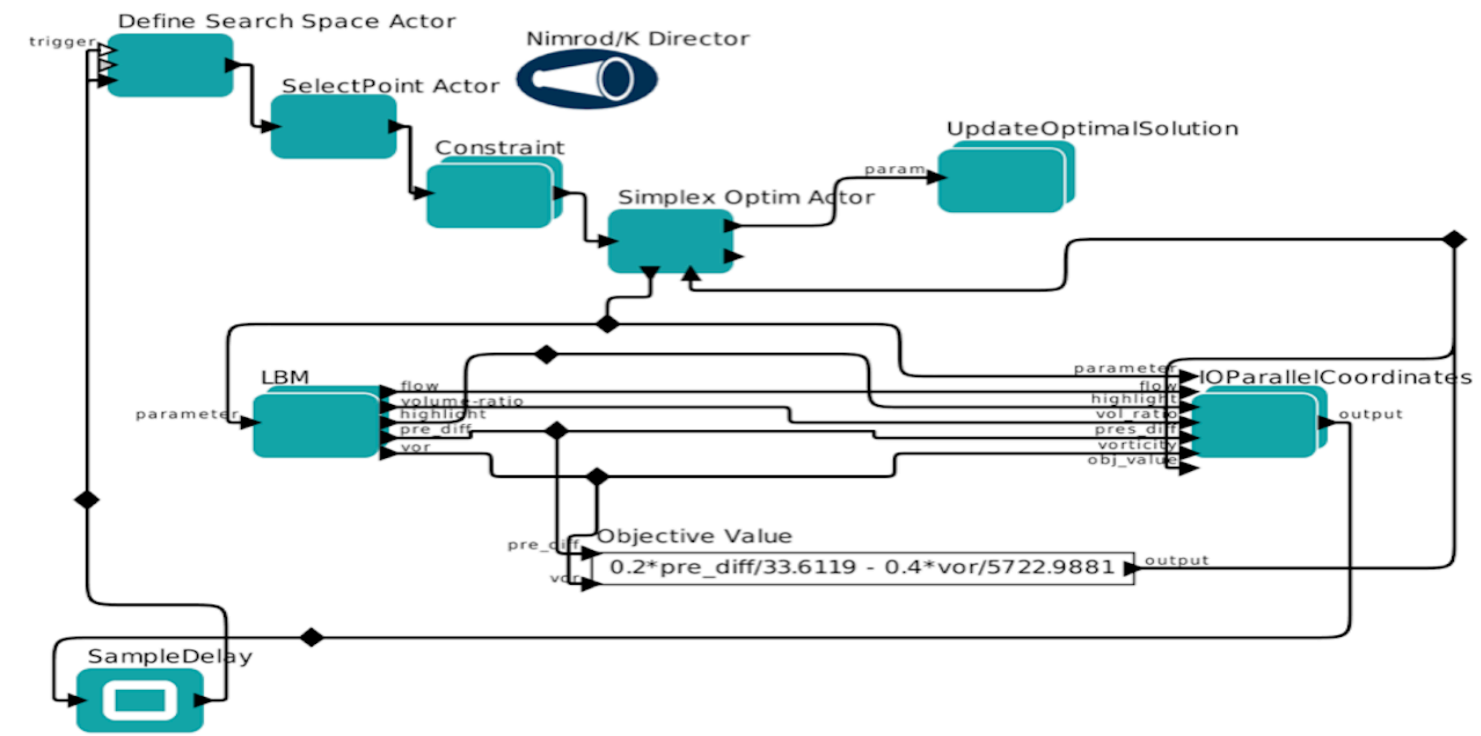


id	pres_diff	s	r	R	vol_ratio	vorticity	highlight	obj_value
1	58.603101...	22.473513...	7.0432184...	146.46258...	0.5806850...	8657.8278...	1	-0.2564219...
3	134.77303...	22.252634...	5.6269441...	192.93288...	0.6547361...	12075.028...	0	-0.0420302...
4	117.42423...	22.252634...	5.6269441...	173.03288...	0.6547361...	10763.367...	0	-0.0535837...
6	64.675433...	22.473513...	6.4432184...	146.46258...	0.6140777...	8711.6081...	0	-0.2240487...
10	62.974059...	20.673513...	6.4432184...	146.46258...	0.6140777...	8541.2647...	0	-0.2222665...

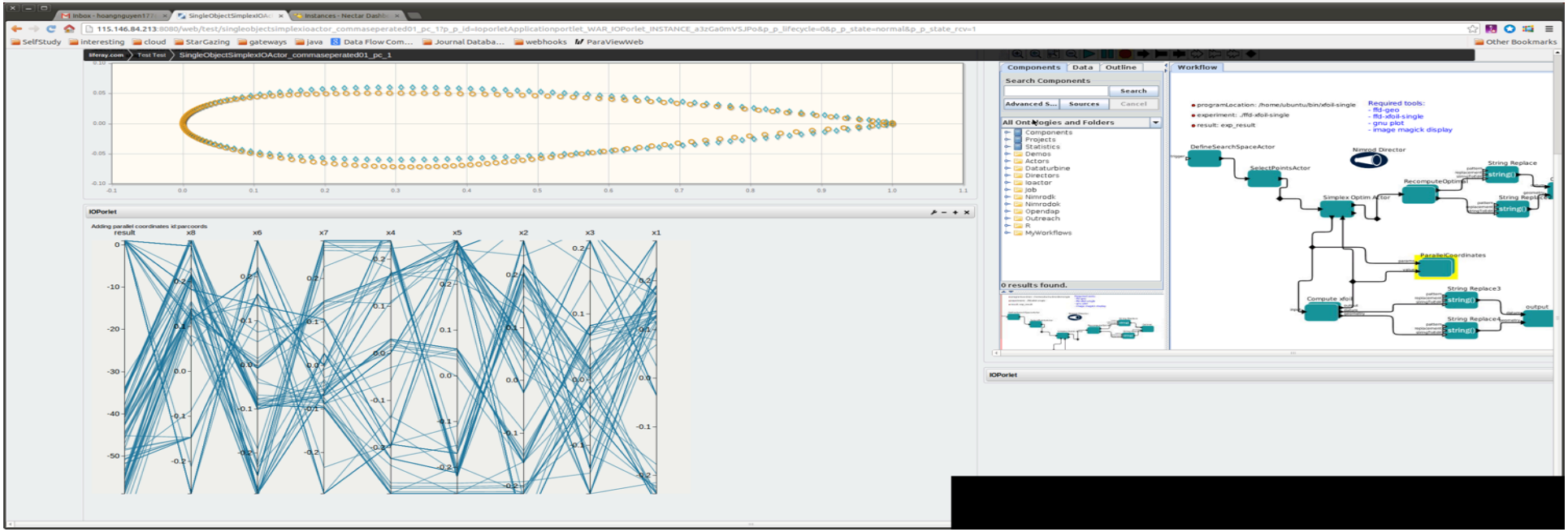
Micromixer Optimization



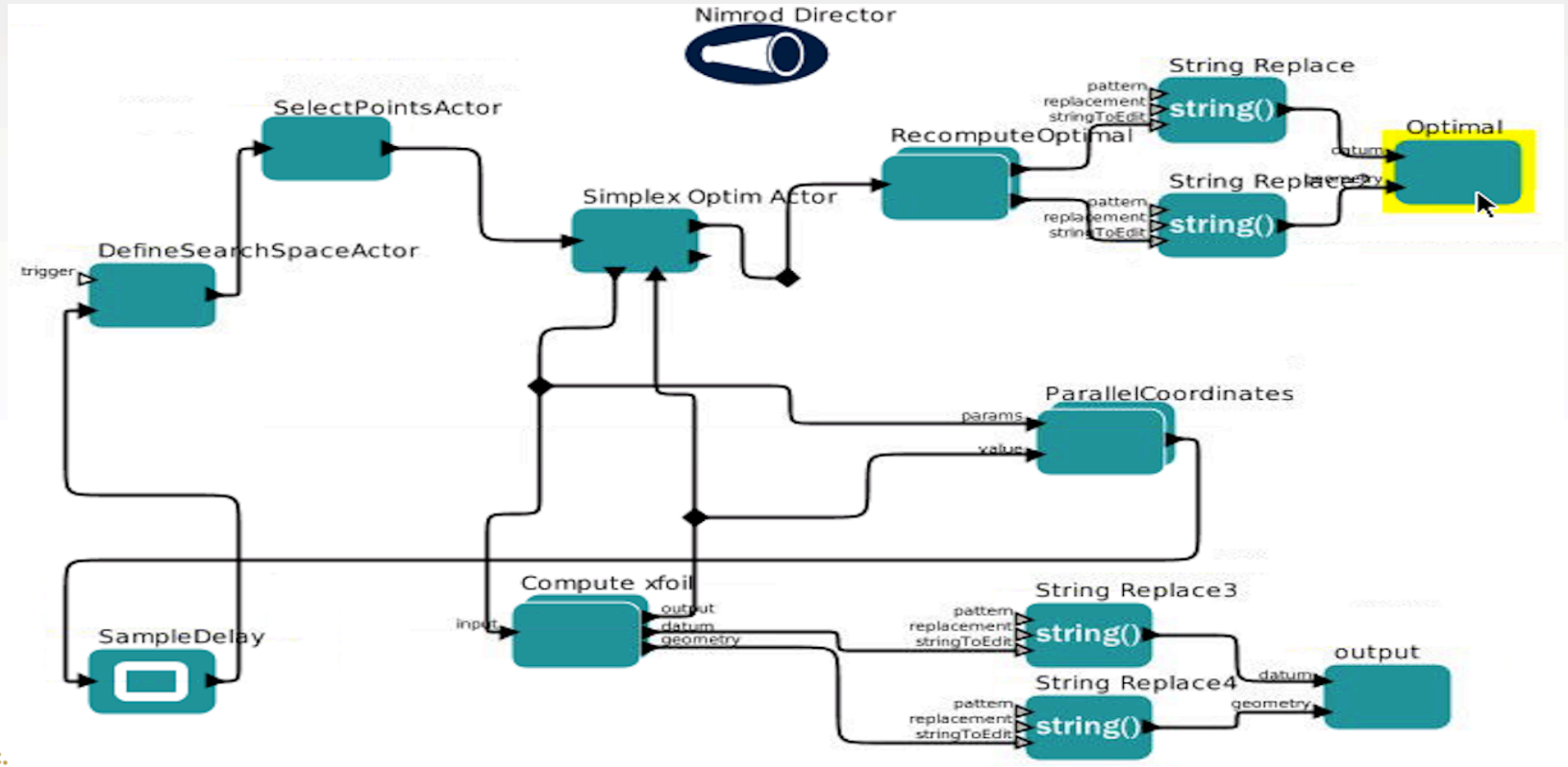
Micromixer optimization Workflow



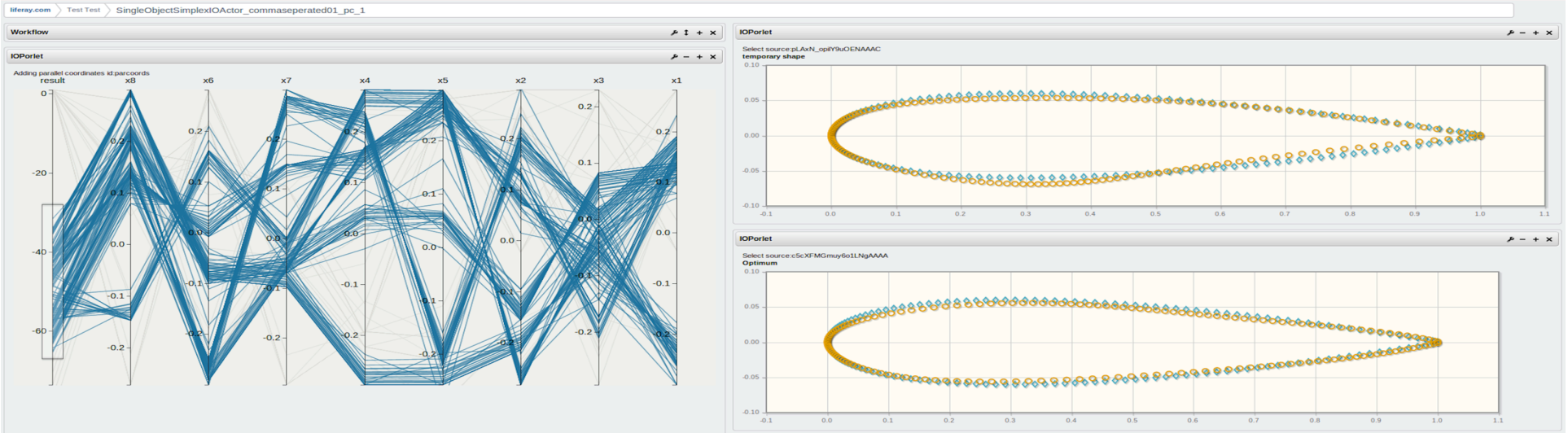
Airfoil Design



2D airfoil optimization



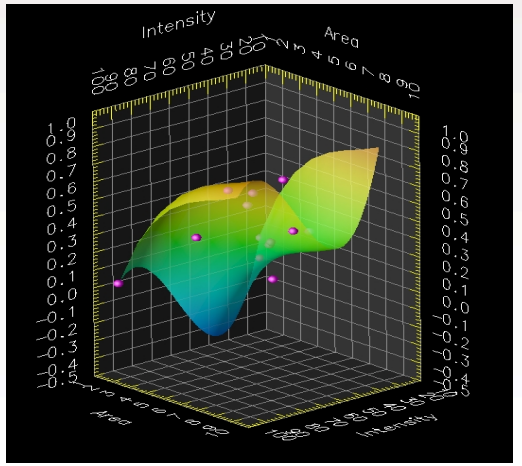
2D airfoil optimization

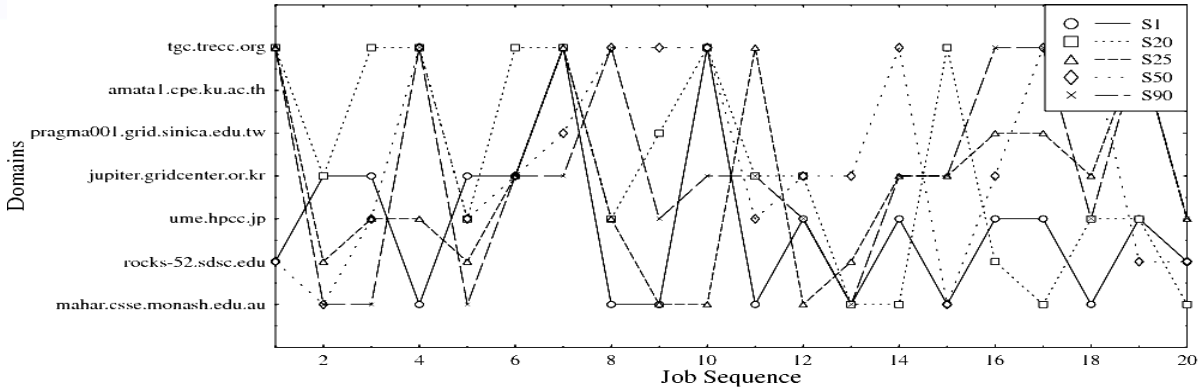


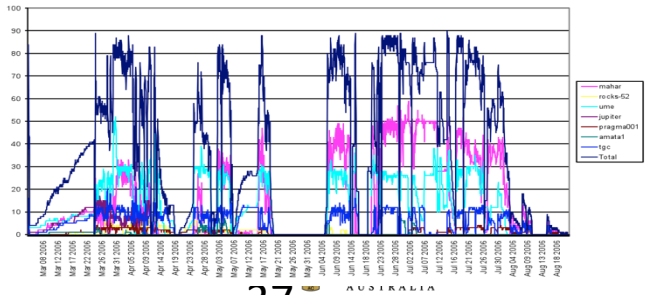
Data Intensive Workflows

Wildfires
Lynch, Beringer, Uotila Monash U, AU

RESEARCH COMPUTING CENTRE







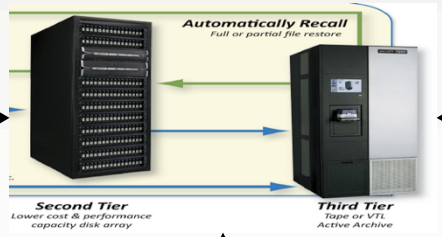
Use Case: Microscopy

RESEARCH COMPUTING CENTRE

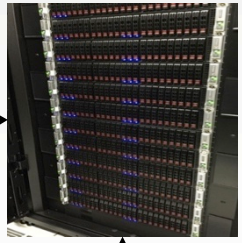
Capture

Pre-process

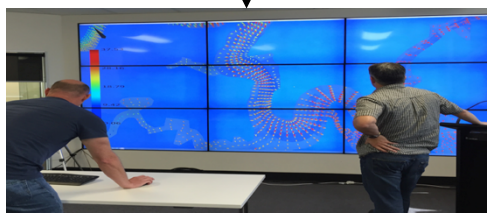
Store



Analyze



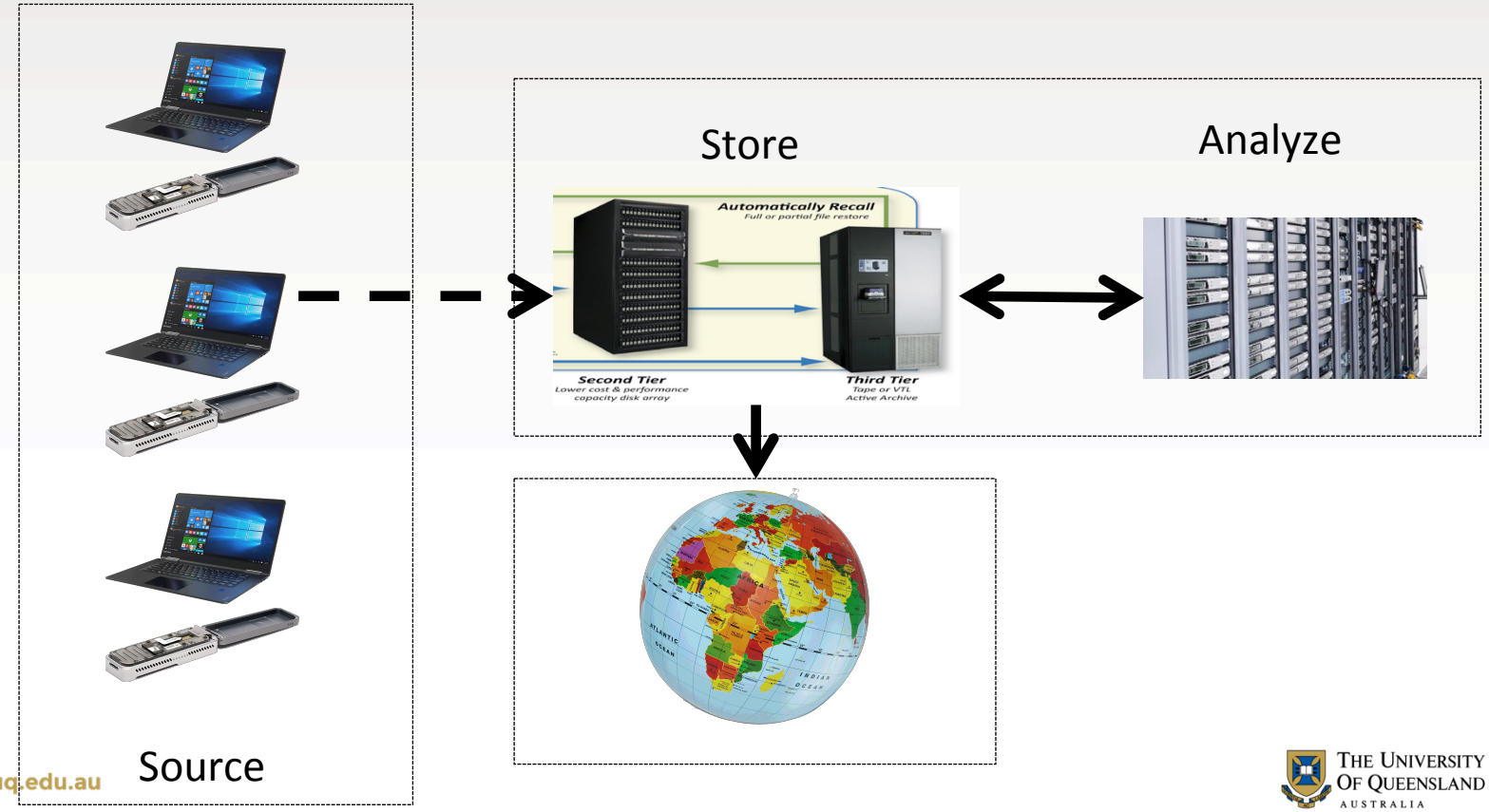
Publish & Share



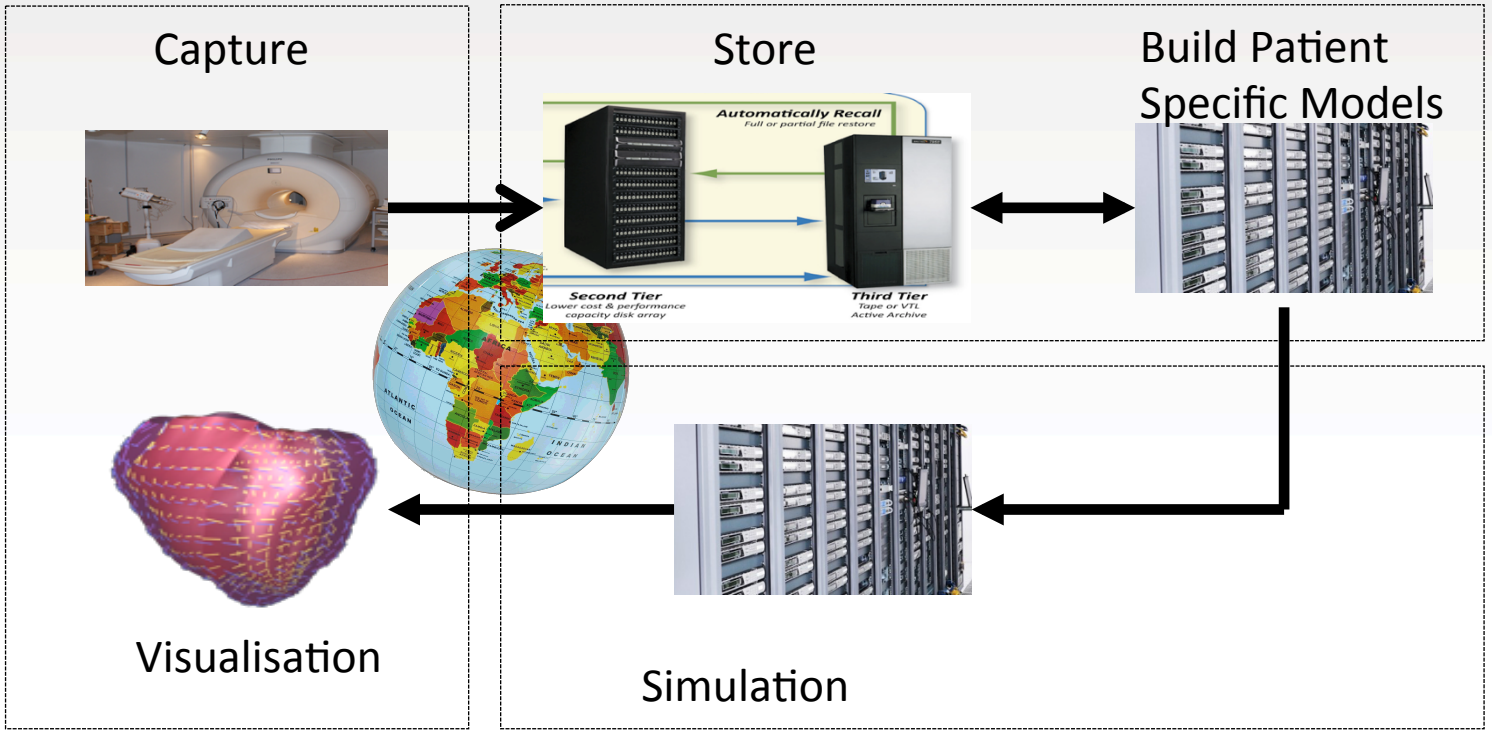
Interpret

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

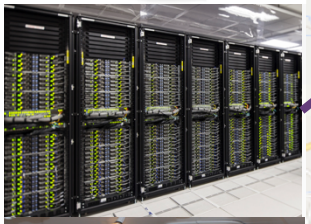
Use Case: Personal Genomics

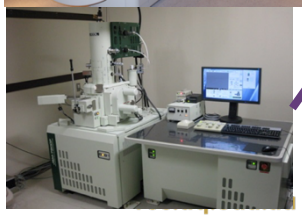


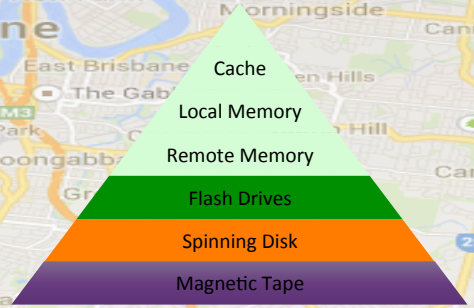
Use Case: Cardiac Science



The MeDiCI landscape

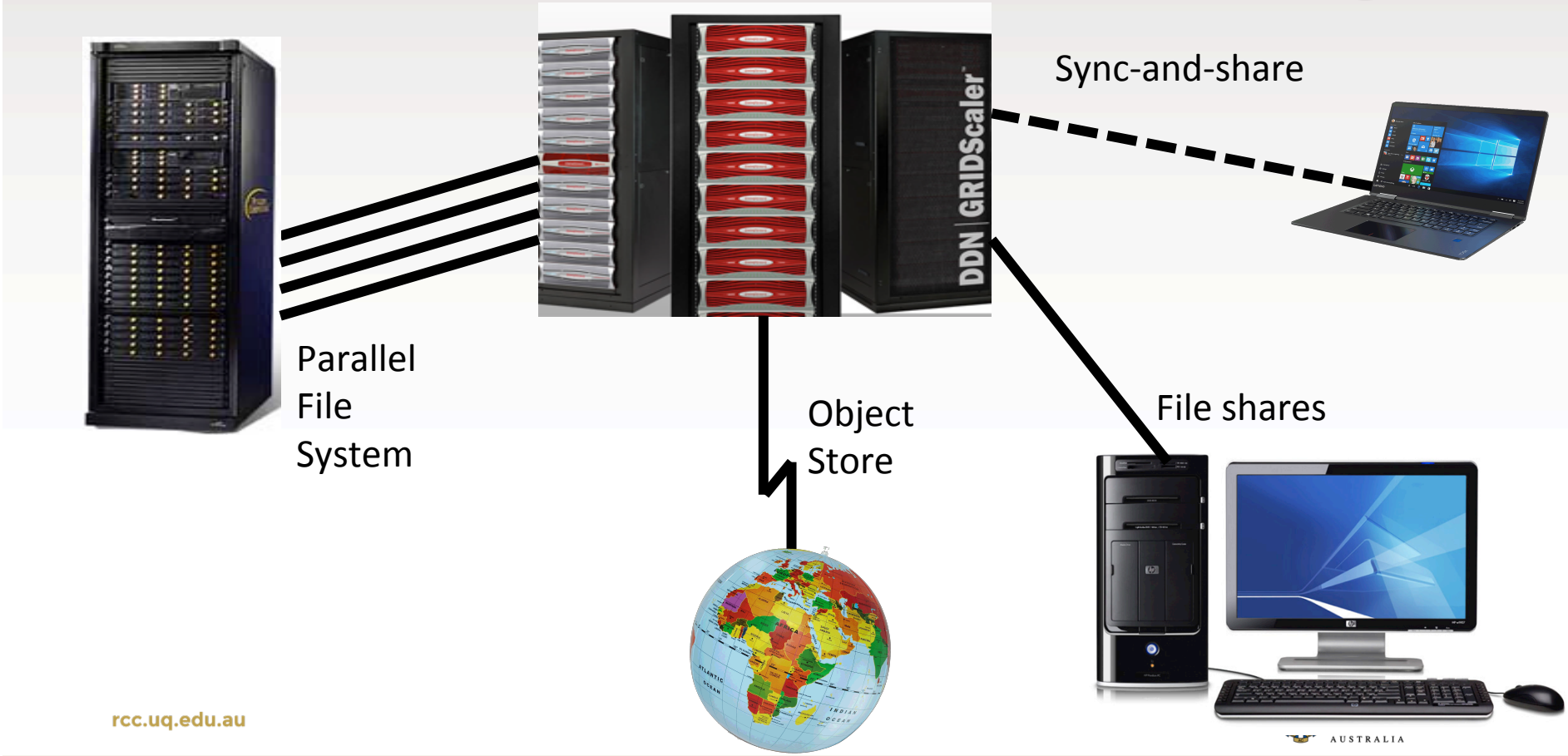






Remote Data Centre

MeDiCI unifies data access



Conclusions

RESEARCH COMPUTING CENTRE

- Workflows are useful for scripting complex computational science and engineering problems
- Conceptually easy to add optimization
- User interaction requires new workflow actors
- Integration to a Science Gateway allows very powerful workflows to be exposed to wider communities.
- Placing a common file system under the workflow engine seems to be promising