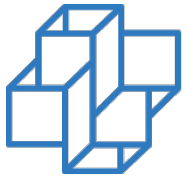
Scientific Data Management @ Hoscar (Group 2)

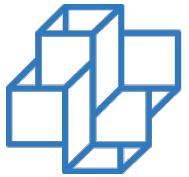


Laboratório Nacional de Computação Científica

DEXL LAB
EXTREME DATA LAB

DEXL LAB
EXTREME DATA LAB

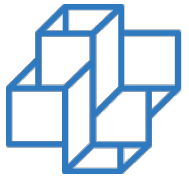
Introduction



- The HOSCAR project re-enforced the productive collaboration among LNCC-COPPE/UFRJ – Zenith INRIA
  - Various Projects
    - Saravá (2009-2011)
      - Coordinated by M. Mattoso (COPPE) P. Valduriez (INRIA)
    - Dataluge (2011-2013) UFRJ-INRIA
      - Coordinated by M. Mattoso (COPPE) P. Valduriez (INRIA)
    - MUSIC (2014-2016)
      - Coordinated Fabio Porto (LNCC), E. Pacitti (INRIA)

DEXL LAB
EXTREME DATA LAB

Hoscar's objectives - 2012
Group 2: Scientific Data Management



Processing of very large datasets.

- This topic deals with the management of very large datasets that are manipulated (accessed and produced) by data-centric scientific workflows in HPC environments.
- Addressing the very scale of the datasets requires new scalable parallel data management techniques as well as scalable data-aware scheduling strategies.
- Furthermore, getting data in and out HPC environments from the scientists' own environment is a major challenge.
- Finally, it is important to provide support for data provenance, a key function that records critical metadata about experiments to help scientists understanding results or reusing some workflow parts.

Current results:

- New CS techniques on these 4 issues
- joint publications among CS-BR & CS-FR
- validations with scientific applications from BR partners

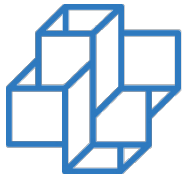
Challenges for HPC4E:

- Traversal actions between groups

http://www-sop.inria.fr/hoscar/pmwiki-2.2.40/pmwiki.php/Main/Objectives

DEXL LAB
EXTREME DATA LAB

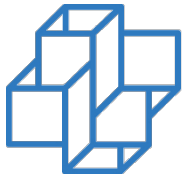
Publications: LNCC



- Multidimensional array support for simulation output analysis (SimDB)
  - Lustosa, H., Porto, F., Valduriez, P., Managing Simulation Data with Multidimensional Arrays, SBBD 2015
- Uncertainty representation and inference (UpsilonDB)
  - Gonçalves, B., Porto, F.; Managing Scientific hypothesis as data with support of predictive analytics, Ieee Computing in Science & Eng., 17(5), Sept., 2015
- Astronomy data integration (NACluster)
  - Freire, V. P., Porto, F., Macedo, J.F., Akbarinia, R., NACluster : A Non- supervised Clustering Algorithm for Matching Catalogues, IEEE eScience workshop, 2014

DEXL LAB
EXTREME DATA LAB

Publications: COPPE-UFRJ



- Algebraic workflow engine (2012-2013)
- E. S. Ogasawara, J. Dias, V. Silva, F. S. Chirigati, D. de Oliveira, F. Porto, P. Valduriez, and M. Mattoso. Chiron: a parallel engine for algebraic scientific workflows. ***Concurrency and Computation:*** *Practice and Experience*, 25(16):2327–2341, 2013.
- Chirigati, F S ; Sousa, V. ; Ogasawara, E. ; Oliveira, D. ; Dias, J. ; Porto, F. ; Valduriez, P. ; Mattoso, Marta . Evaluating Parameter Sweep Workflows in High Performance Computing. In: **Int Workshop on Scalable Workflow Enactment Engines and Technologies (SWEET'12)**, 2012, Phoenix. SIGMOD.
- Dynamic Loops in workflow execution (2013-2014)
- J. Dias, G. Guerra, F. Rochinha, A. Coutinho, P. Valduriez, M. Mattoso. Data-Centric Iteration in Dynamic Workflows. ***Future Generation Computer Systems***, Vol. 4, 114-126, 2015.
- J. Dias, E. S. Ogasawara, D. de Oliveira, F. Porto, P. Valduriez, and M. Mattoso. Algebraic Dataflows for
  Big Data Analysis. **IEEE Bigdata Conference** 2013
- Querying *in-situ* raw data files (2014-2015)
- V. Silva, D. de Oliveira, P. Valduriez, M. Mattoso. Analyzing Related Raw Data Files through Dataflows. ***Concurrency and Computation: Practice and Experience***, 2015
- R. Souza, V. Silva, D. de Oliveira, P. Valduriez, A. Lima, M. Mattoso. Parallel Execution of Workflows Driven by a Distributed Database Management System **IEEE/ACM SuperComputing Conference** *(SC15)*, 2015.

DEXL LAB
EXTREME DATA LAB

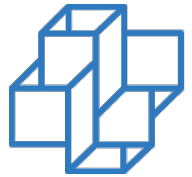
Publications: COPPE-UFRJ

- Multi-site cloud workflow execution (2014-2015)

- J. Liu, V. Silva, E. Pacitti, P. Valduriez, M. Mattoso, "Scientific Workflow Partitioning in Multisite Cloud". In: **3rd Workshop on Big Data Management in Clouds**, Proc. of the Europar 2014
- J. Liu, V. Silva, E. Pacitti, P. Valduriez, M. Mattoso. Parallelization of Scientific Workflows in the Cloud. ***Journal of Grid Computing***
- J. Liu, E. Pacitti, P. Valduriez, D. de Oliveira, M. Mattoso. Workflow Scheduling in Multisite Cloud. draft paper, 2015.

DEXL LAB
EXTREME DATA LAB

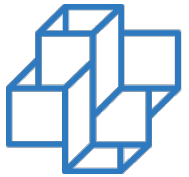
Students Interchange



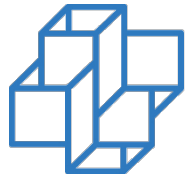
- COPPE-INRIA
  - Jonas Dias – BR→ FR (PhD, COPPE 2013, supervised: M Mattoso & P Valduriez)
  - Vitor Silva – BR→ FR (MSc, COPPE 2014, internship at INRIA Oct-Dec 2013)
  - Ji Liu – FR→ BR (PhD ongoing, INRIA, supervision: M Mattoso, E Pacitti, P Valduriez)
- LNCC-INRIA
  - Daniel Gaspar BR→ FR (PhD, LNCC , ongoing Internship INRIA 2015-2016, supervision: F. Porto, R. Akbarinia)

DEXL LAB
EXTREME DATA LAB

Opportunities



- To extend the collaboration with other groups willing to improve their data management and "out of core" computations



Following Talks in Group 2

- Tuesday 22nd
  - Supporting efficient data analysis on numerical simulation output
    - Fabio Porto (LNCC)
  - Integrating Big Data and relational data with a functional SQL-like query language
    - Patrick Valduriez (INRIA – Zenith)
  - Analyzing related raw data files through dataflows
    - Marta Mattoso (COPPE-UFRJ)
- Thursday 24th
  - FPHadoop: efficient execution of parallel jobs over skewed data
    - Reza Akbarinia INRIA Zenith

Data Management @ HOSCAR

Fabio Porto - fporto@lncc.br

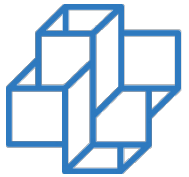
DEXL Lab - LNCC

HOSCAR 2015

Collaboration with COPPE-UFRJ / INRIA Zenith

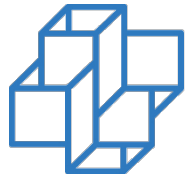
Ministério da
Ciência, Tecnologia
e Inovação

GOVERNO FEDERAL
BRASIL
PAÍS RICO É PAÍS SEM POBREZA



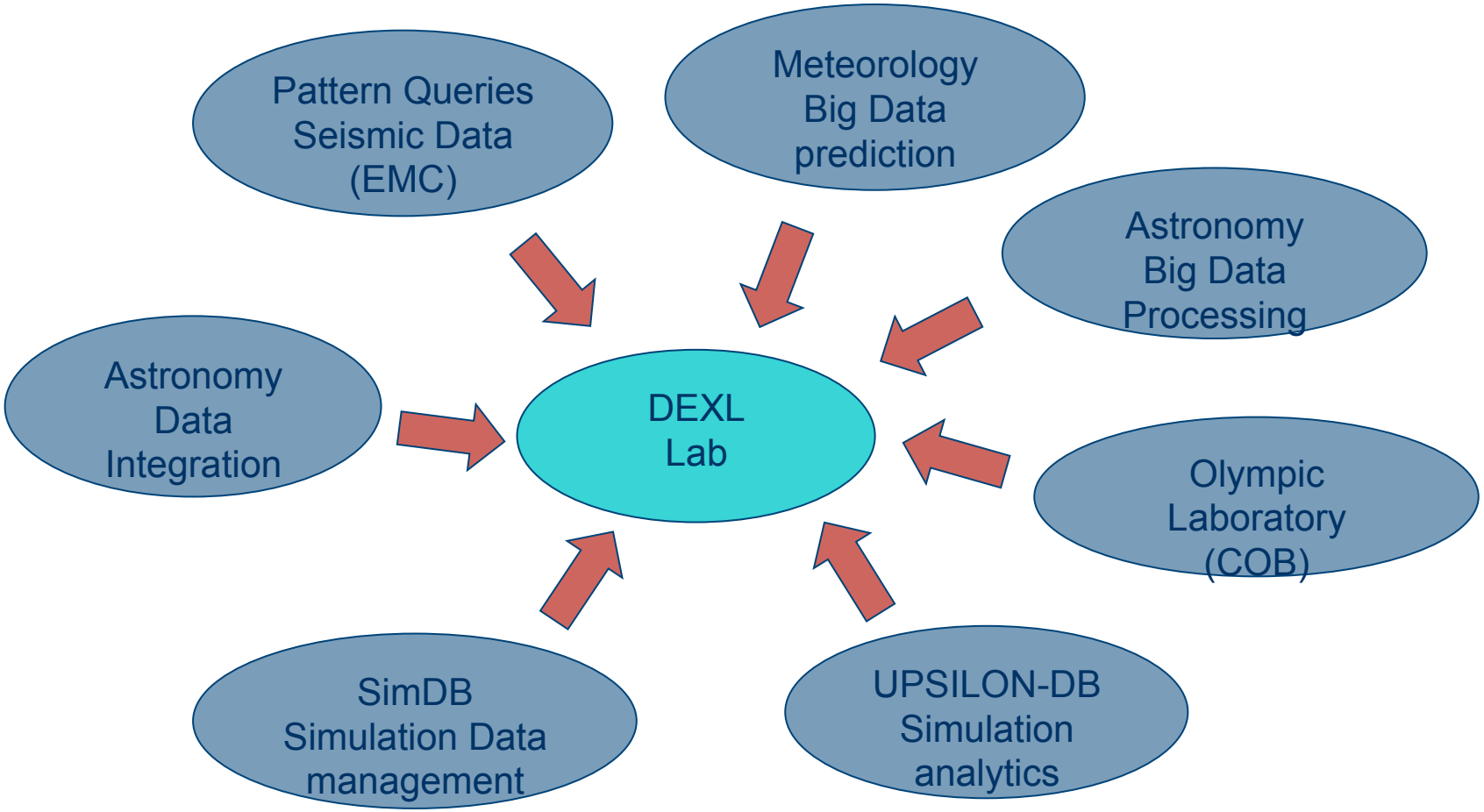
Agenda

- Introduction
- The Challenge of Scientific Big Data
- Using Multidimensional array database in support of natural phenomena data analytics
- Final Comments



INTRODUCTION

Big Data at LNCC



DEXL LAB
EXTREME DATA LAB

The effect of Data Deluge

- "Scientists are spending most of their time manipulating, organizing, finding and moving data, instead of researching. And it's going to get worse"
  - Office Science of Data Management challenge - DoE

Topic of the Decade

DEXL LAB
EXTREME DATA LAB

- Funding agencies requires:
- Data curation
- Reproducibility
- Scientists want
  - Data publication
  - Data re-analysis
  - Predictive capacity & Error analysis

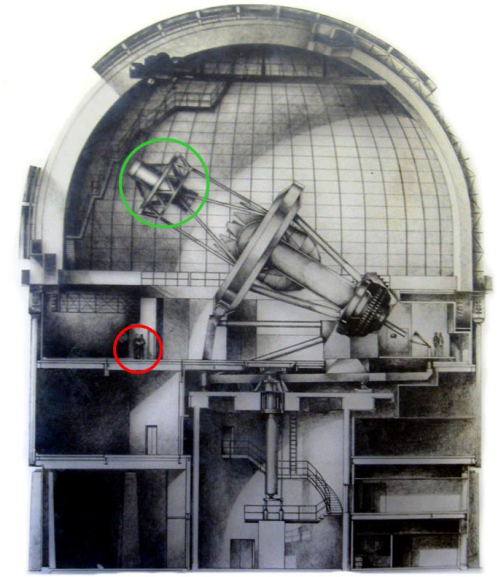
Sept. 2015

Sources of Data

LSST - telescope

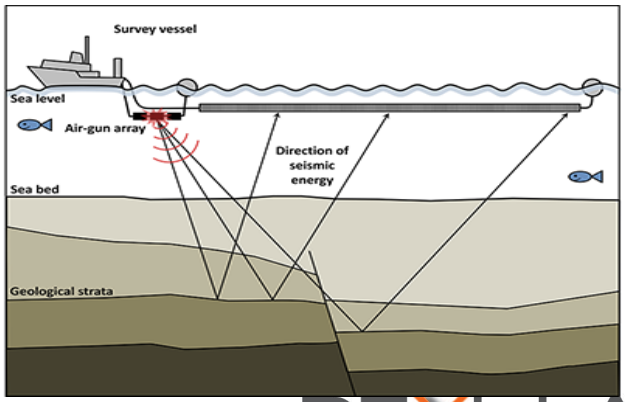
Illumina – HISeq2000

Blanco Telescope



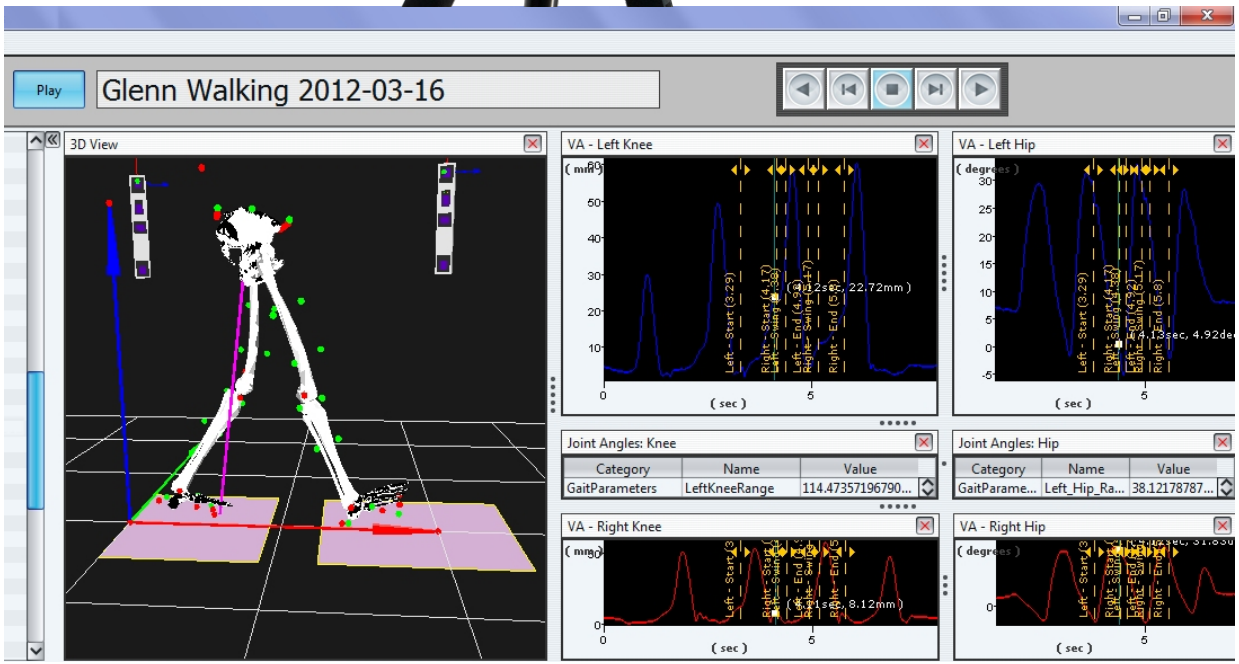
Waters – Mass Spectrometry

Seismic Hydrophones

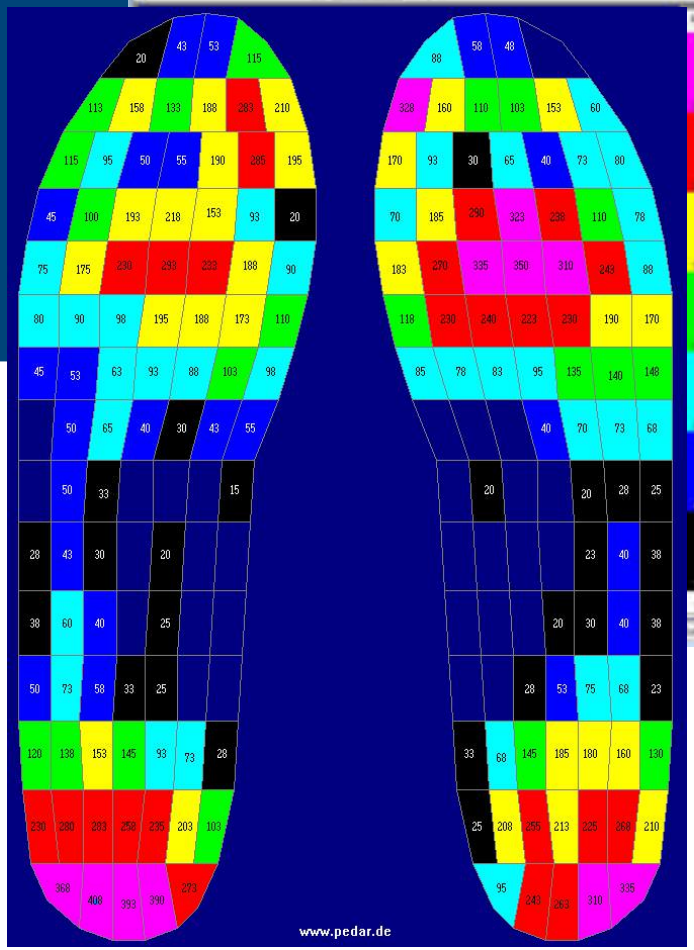


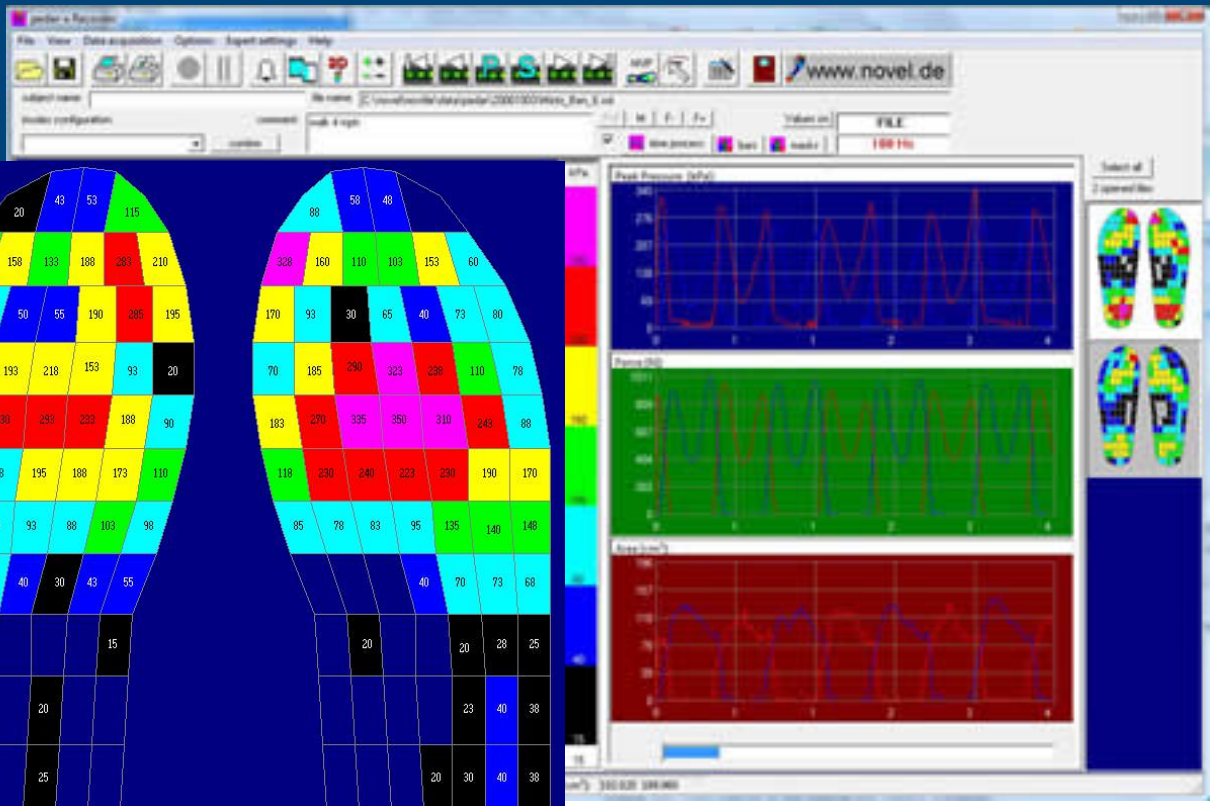
DEXL LAB
EXTREME DATA LAB

codamotion


Play
Glenn Walking 2012-03-16
3D View
VA - Left Knee
VA - Left Hip
Joint Angles: Knee
Category
Name
Value
GaitParameters
LeftKneeRange
114.47357196790...
Joint Angles: Hip
GaitParame...
Left_Hip_Ra...
38.12178787...
VA - Right Knee
VA - Right Hip
(sec)

novel.de





From Raw data to tailored Data

cleaning

Integration

Transformation

Raw

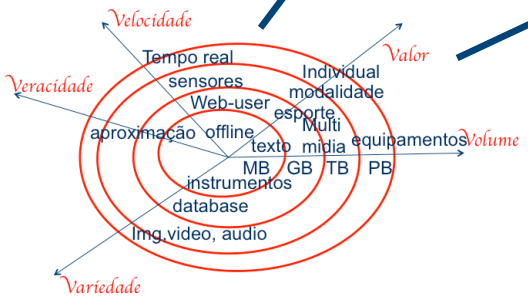
Tier:0

Intermediary

Tier:1

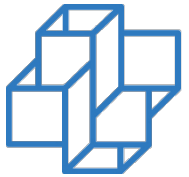
Tailored

Tier:2



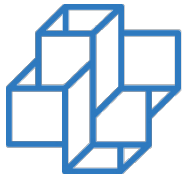
DEXL LAB
EXTREME DATA LAB

Our Point is



- The case of natural phenomena simulation data requirements is similar:
  - High-volume
    - Multidimensional space-time
  - Non-traditional data formats: SEG-Y, HDF5, FITS,…
  - Complemented with meta-dada (variety)
    - Simulation provenance
    - Model metadata
    - Equations, Hypotheses
  - Uncertain (veracity)

Objective

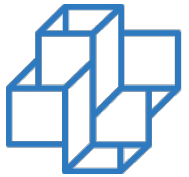


- To provide scientific applications the data efficiency found in business applications
  - High-level language for data analysis
  - Data Partitioning and Parallel data access
  - Data representation
  - Efficient execution of queries and dataflows

From Raw Data to Database

Simulation Data

SIMDB



- Data management support for Numerical Simulation
  - Quality assessment
    - ✓ Simulation output data analysis
    - Uncertainty representation
  - Efficient input parameter access
  - Big Data based predictions

Mean SQR error

Simulation analysis

$$e(\bar{x}) = \frac{\sum_{t=0}^{n} [F(\bar{x}, t) - G(\bar{x}, t)]^2}{n}$$

Query in SQL

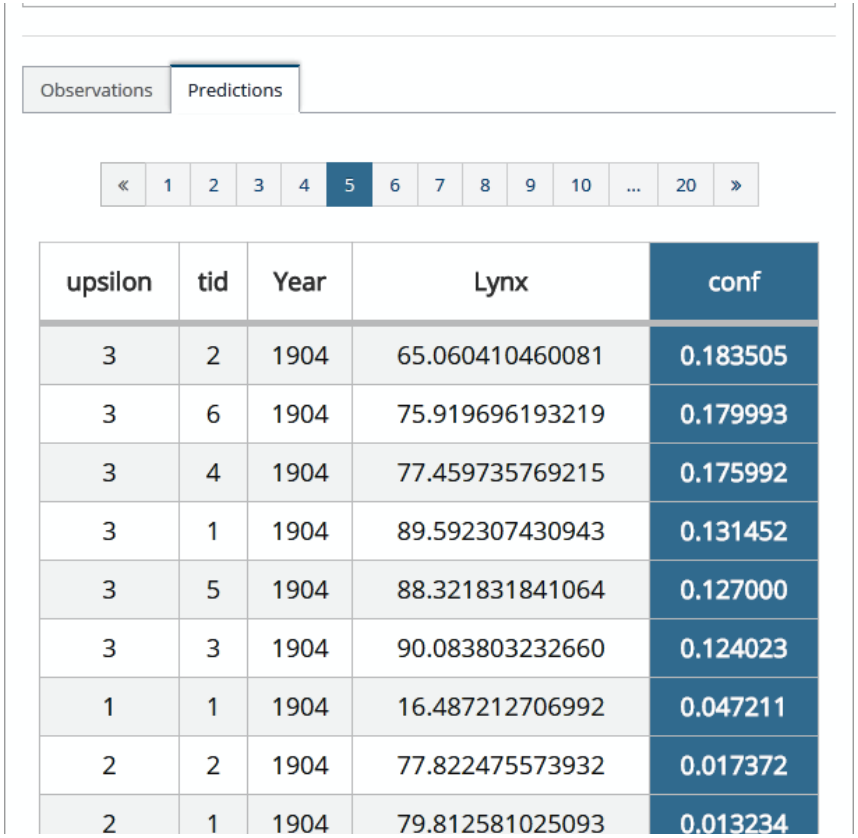
```
SELECT SUM(SQR(F.value  - G.value))/N
FROM F
INNER JOIN G ON F.x = G.x
                AND F.y = G.y
                AND F.z = G.z
                AND F.t = G.t
GROUP BY F.x, F.y, F.z;
```

UpsilonDB - HOSCAR 2014/ Gramado

- Data management support for Numerical Simulation
  - Quality assessment
    - simulation output data analysis
    - ✓ Uncertainty representation

      UPSILON-DB
  - Efficient input parameter access
    - MHM Method (opportunity)

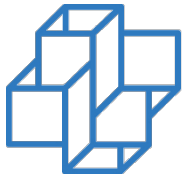
Uncertainty Representation



upsilon	tid	Year	Lynx	conf
3	2	1904	65.060410460081	0.183505
3	6	1904	75.919696193219	0.179993
3	4	1904	77.459735769215	0.175992
3	1	1904	89.592307430943	0.131452
3	5	1904	88.321831841064	0.127000
3	3	1904	90.083803232660	0.124023
1	1	1904	16.487212706992	0.047211
2	2	1904	77.822475573932	0.017372
2	1	1904	79.812581025093	0.013234

Gonçalves, B., Porto, F.; Managing Scientific hypothesis as data with support of predictive analytics, Ieee Computing in Science & Eng., 17(5), Sept., 2015

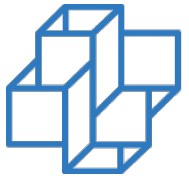
DEXL LAB
EXTREME DATA LAB



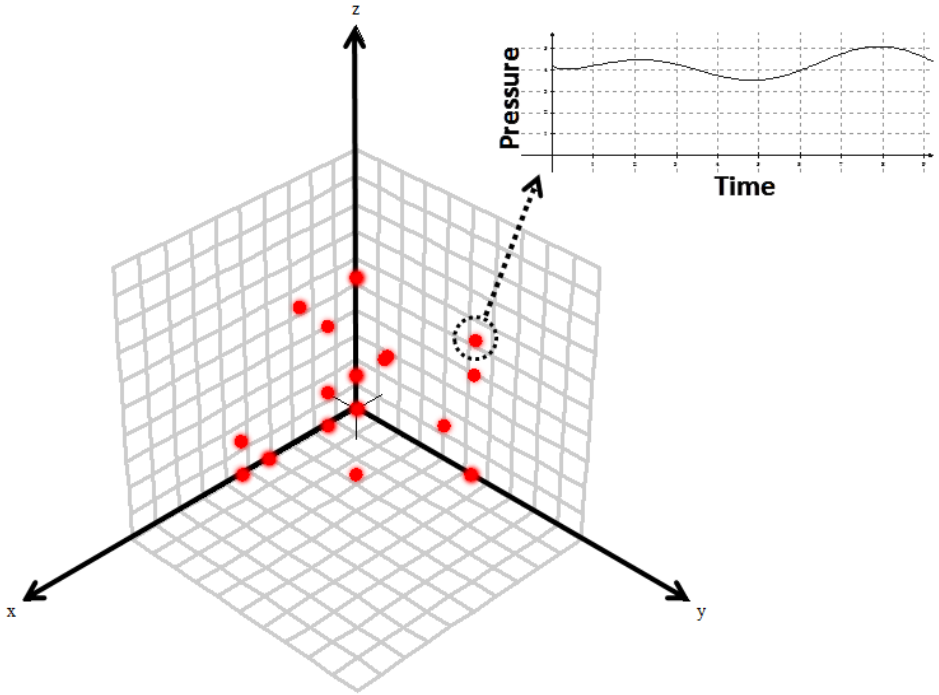
Hermano Lustosa, Fabio Porto, Patrick Valduriez

SIMDB

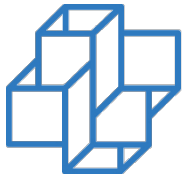
Multidimensional Arrays Data Model (SciDB)



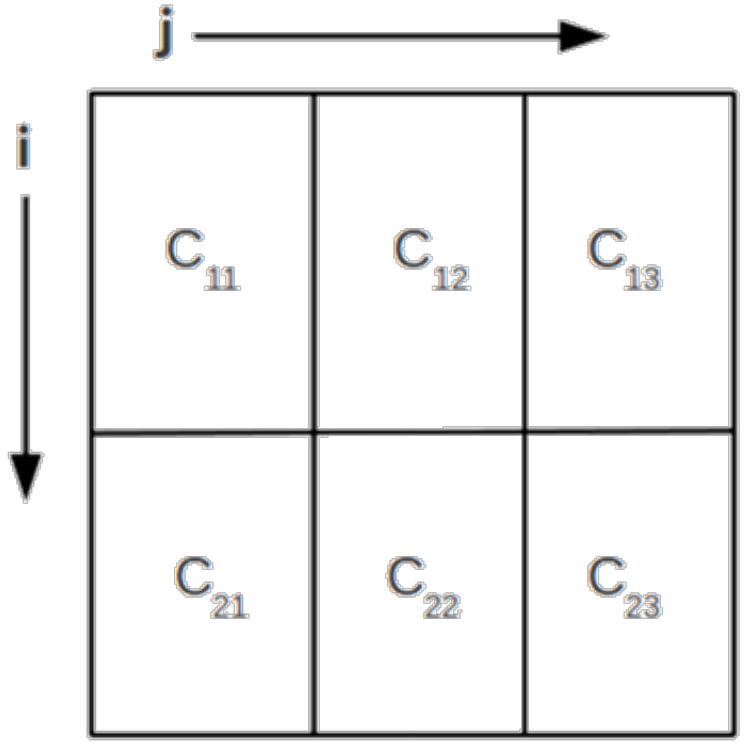
- Arrays:
  - Data Model: A=(D,O,Att,*fi*)
  - Dimensions Indexes -> Cell with Attributes
  - Query languages : AQL, AFL



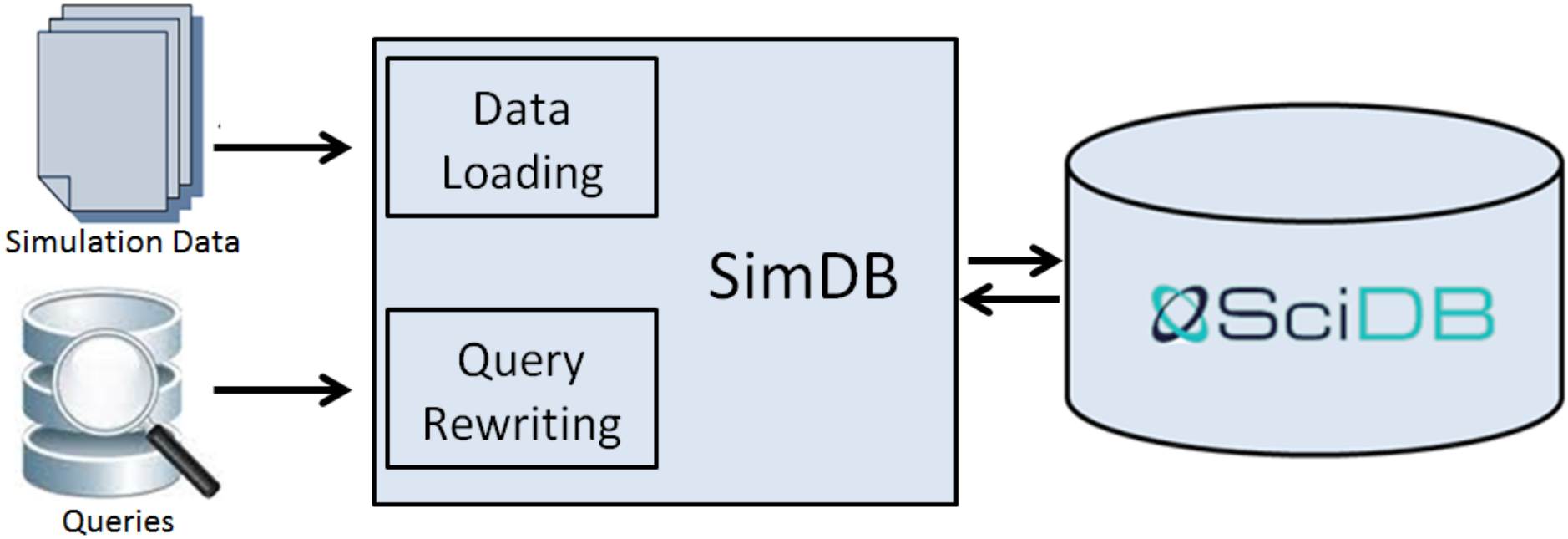
Multidimensional Arrays Data Model (SciDB)



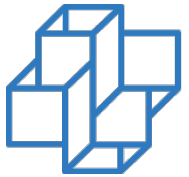
- Data allocation in "chunks"
  - A dimension-based array partitioning



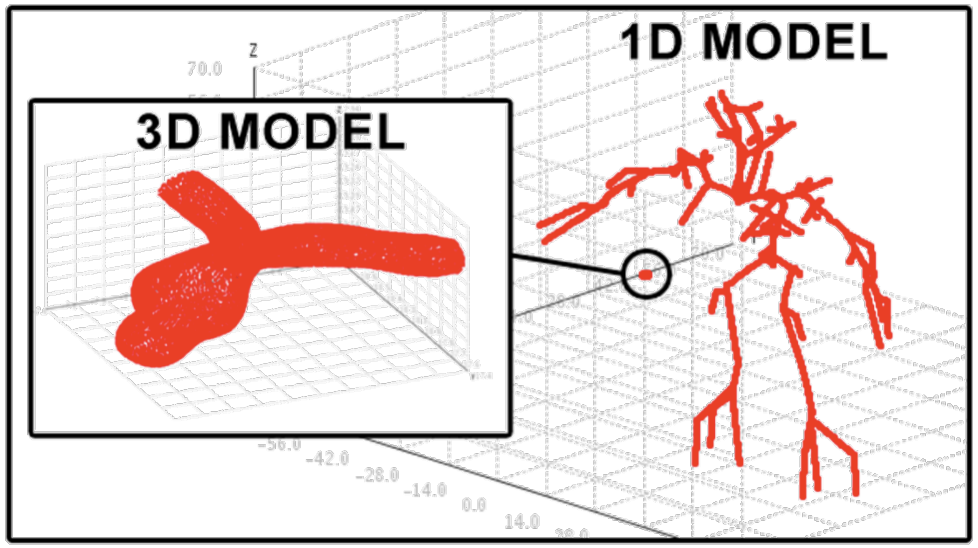
SimDB General View



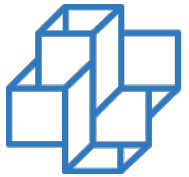
Use Case – Cardio-Vascular System



- Hemolab 3D-1D coupled model
- Mesh with 90k vertexes
- 240 time steps per simulation
- Total of 430M points

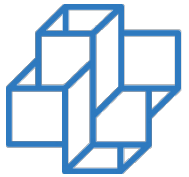


At first Glance



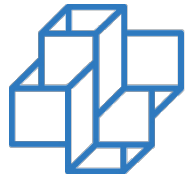
- Multidimensional Array model seems a perfect bet for natural phenomena data representation:
  - Space, Time, Simulation
  - Data Partitioning and Parallel query execution
  - Query Language
- A natural fit for multidimensional data
  - **But….**

The Problem



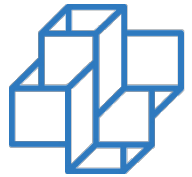
- Coordinate systems with indexes as real numbers
  - **But** Arrays indexes are integers
- Coordinates need to be mapped into integers indexes
  - Without compromising data sparcity

Designing Efficient arrays



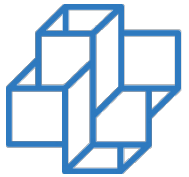
- Integer dimensions are straight forwardly mapped to array coordinates
- Non-integer dimensions
  - Mapping techniques
  - Drastic effect on performance

Mapping Techniques (SPACE PRESERVING)



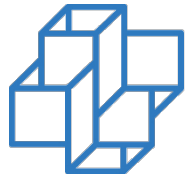
- **Precision Elimination**
  - Multiply each coordinate value by 10^n
- Pros
  - Cheap
  - Range queries are easy to write
- Cons
  - Generate very sparse arrays
  - Hard to define chunk partitioning

Mapping Techniques (NON SPACE PRESERVING)

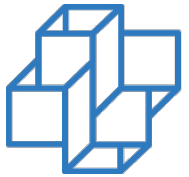


- **Histograms**
  - Create Equi-depth Histograms on each dimension
  - Attribute Indexes to coordinate values based on the histograms
- Pros
  - Distributes data more evenly
  - Easier to define regular chunk partitioning
- Cons
  - Requires queries to be translated
  - Requires mappings to be explicitly stored

Mapping Techniques

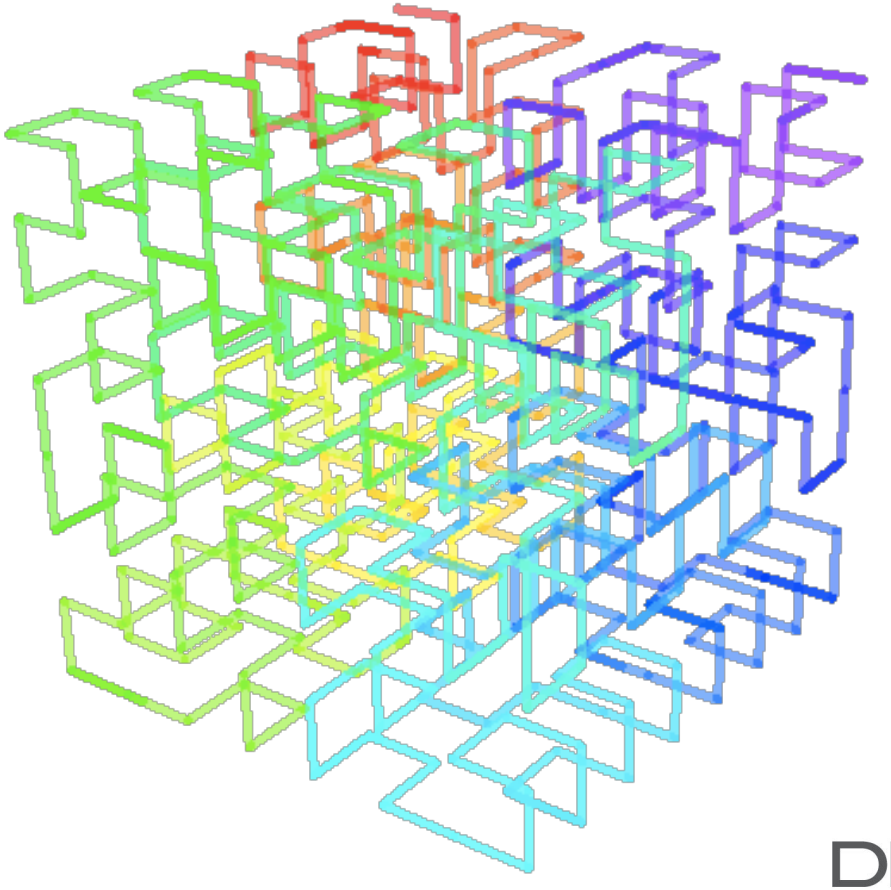


- Space Filling Curve
  - Map 3D space to 1 dimension
  - Use Hilbert Function
- Pros
  - Creates a dense array
- Cons
  - Range queries require a costly join with an auxiliary array containing original data

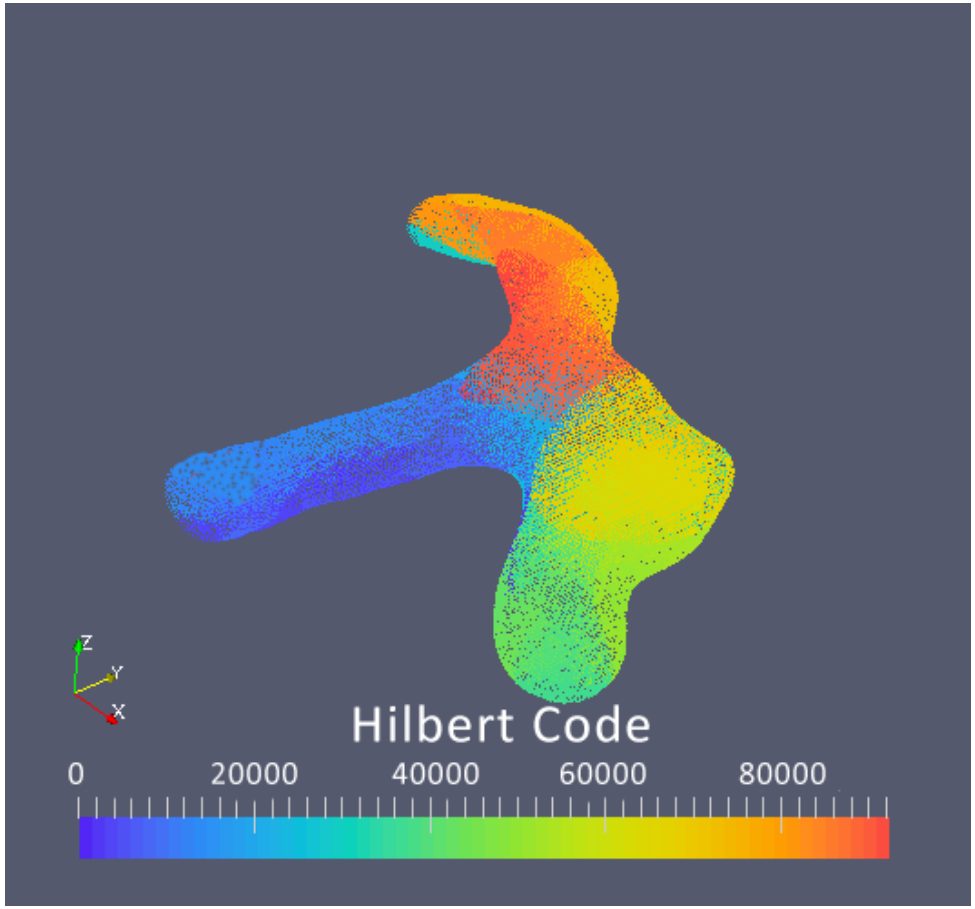


Mapping Techniques

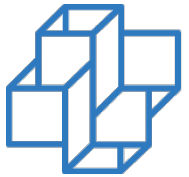
- 3D Hilbert Curve example



Hilbert representation of an aneurism

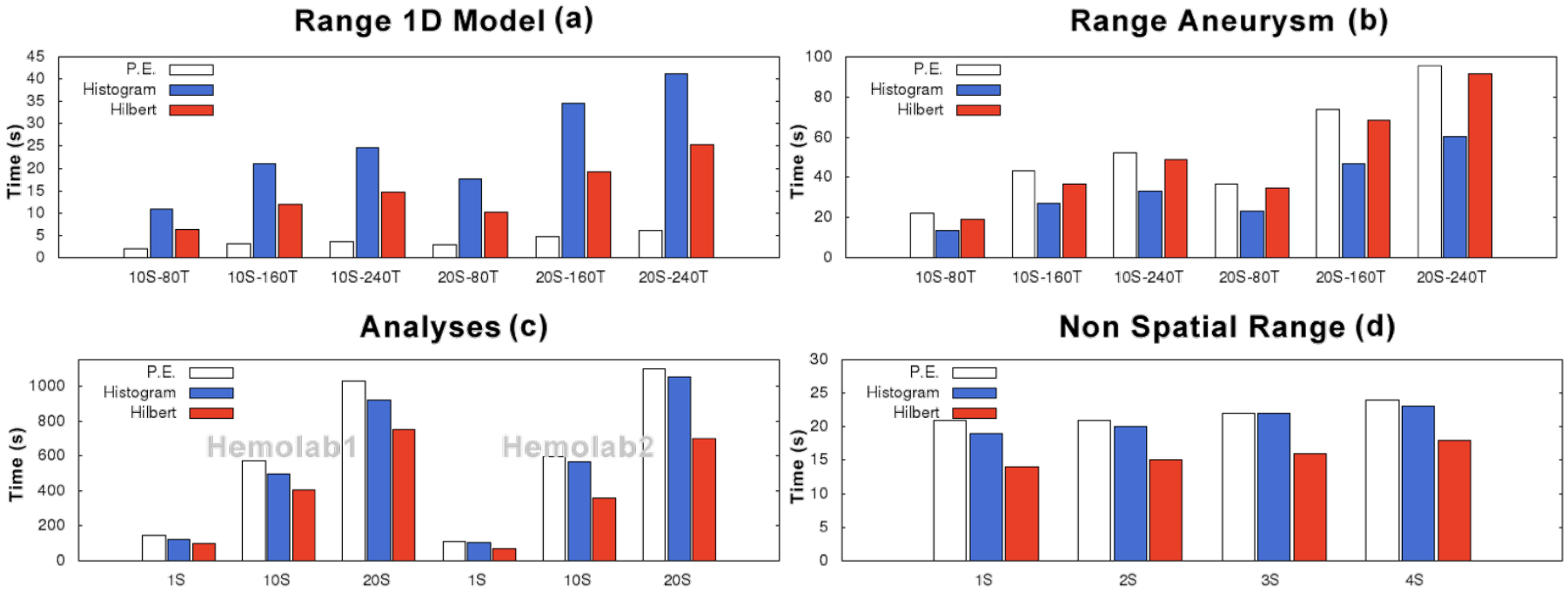


Evaluation

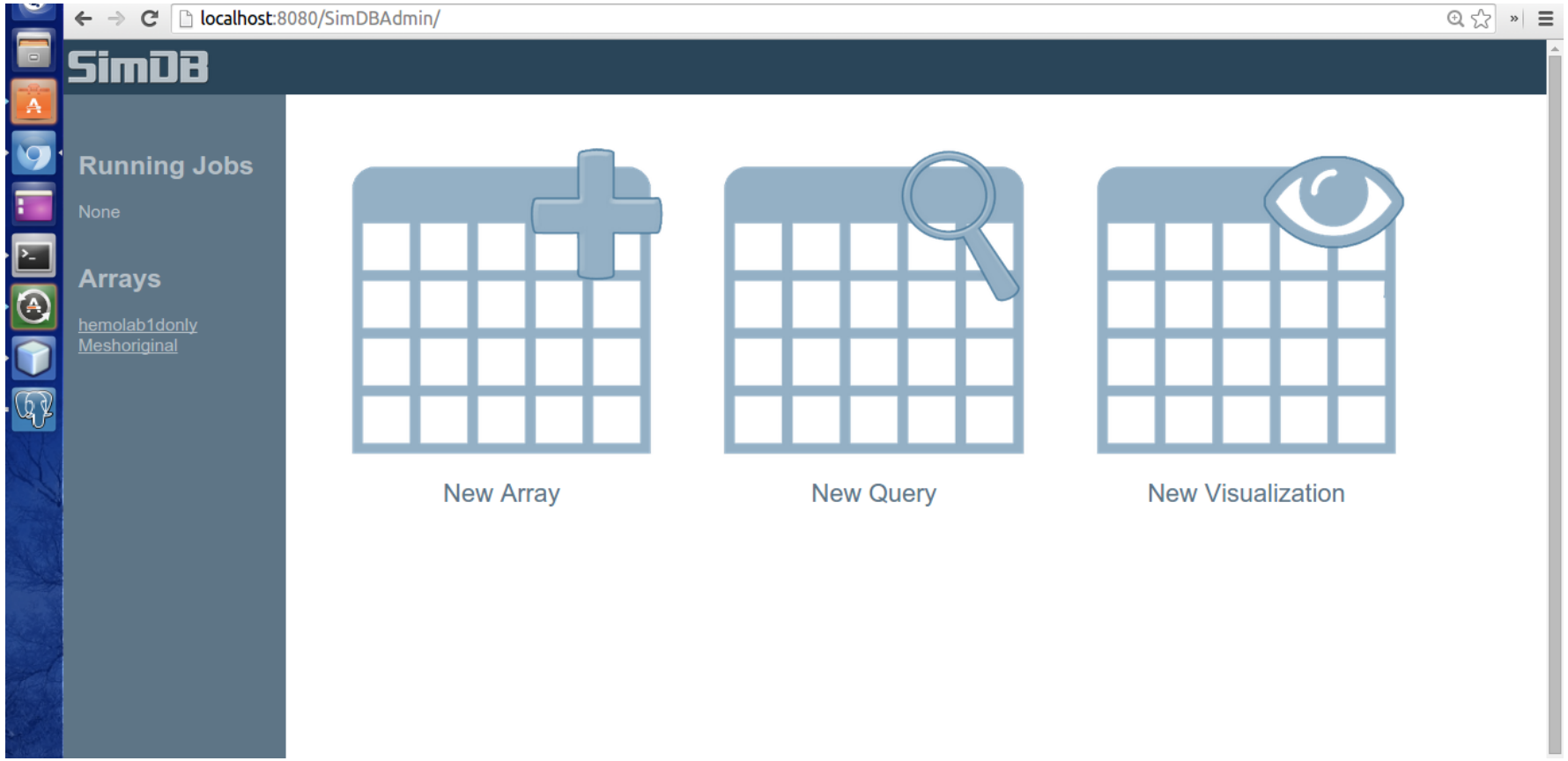


- Spatial Range Queries
  - Aneurysm (3D model)
  - Entire 1D model
- Non Spatial Range
- Real Life Analyses
  - Hemolab 1: Mean square error between data and a reference function
  - Hemolab 2: Time step in which pressure reaches its maximum value on each point

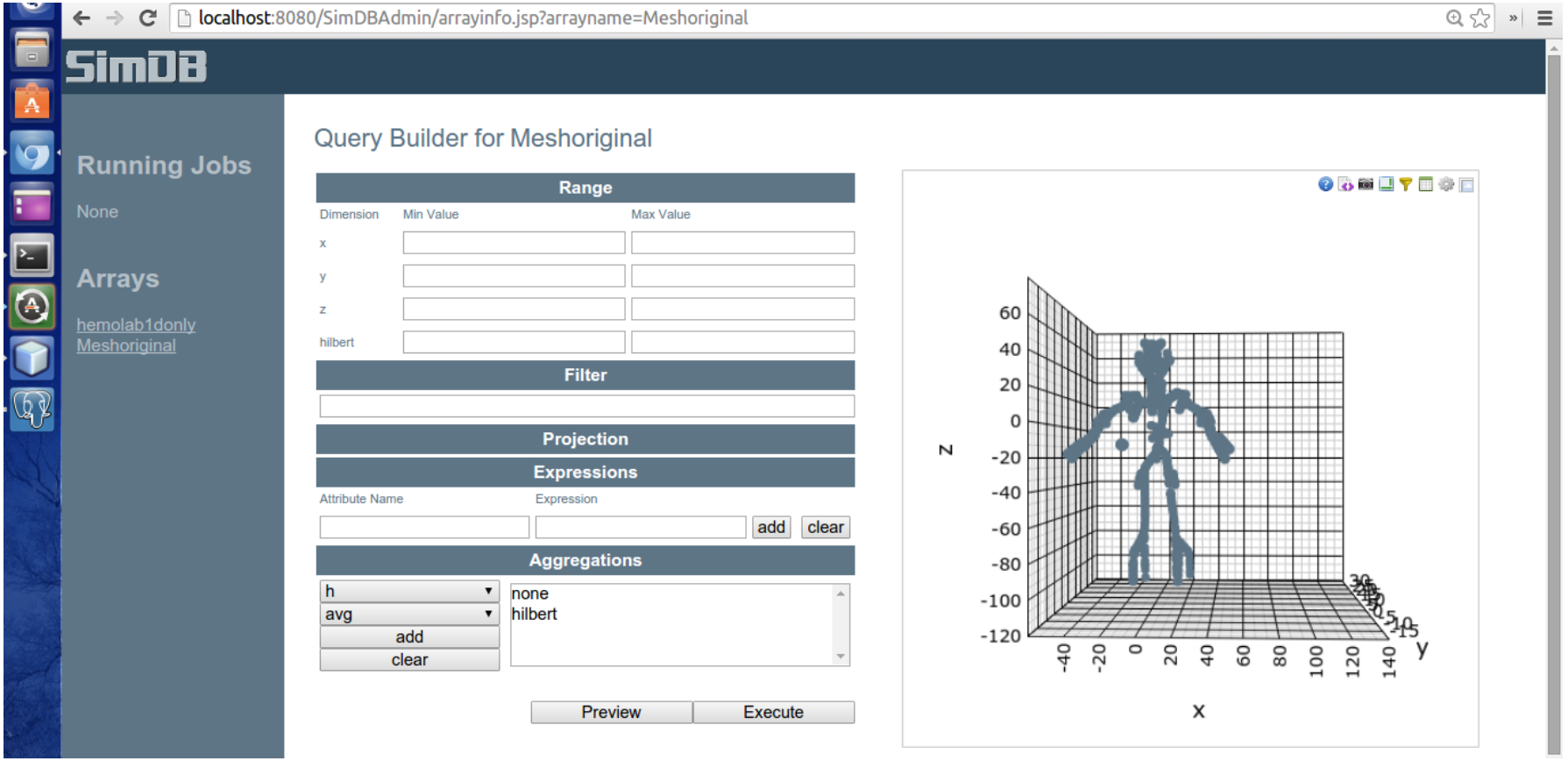
Evaluation


Range 1D Model (a)
Time (s)
P.E.
Histogram
Hilbert
10S-80T
10S-160T
10S-240T
20S-80T
20S-160T
20S-240T
Range Aneurysm (b)
Analyses (c)
Hemolab1
Hemolab2
1S
10S
20S
Non Spatial Range (d)
2S
3S
4S

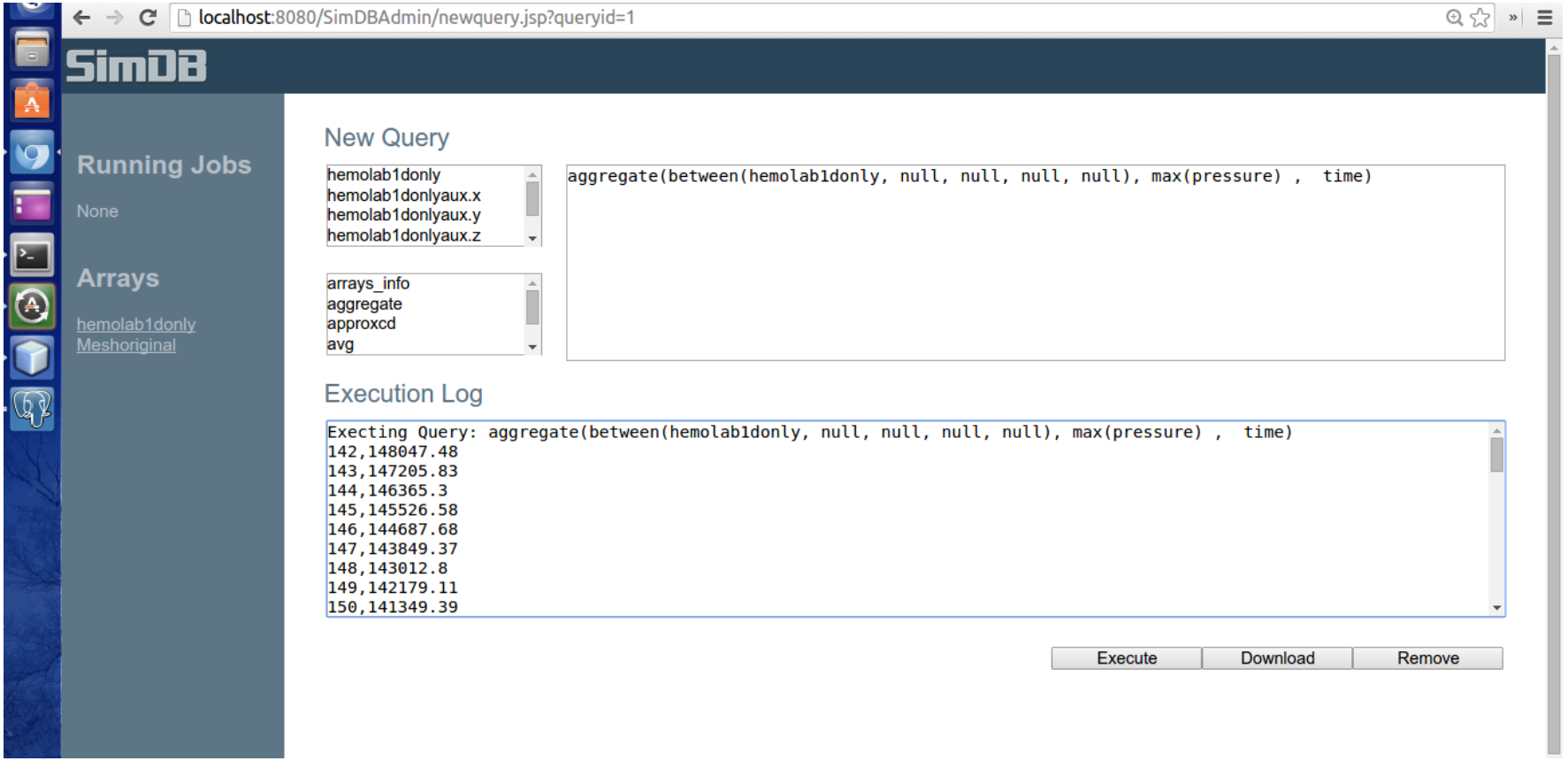
The SimDB Prototype



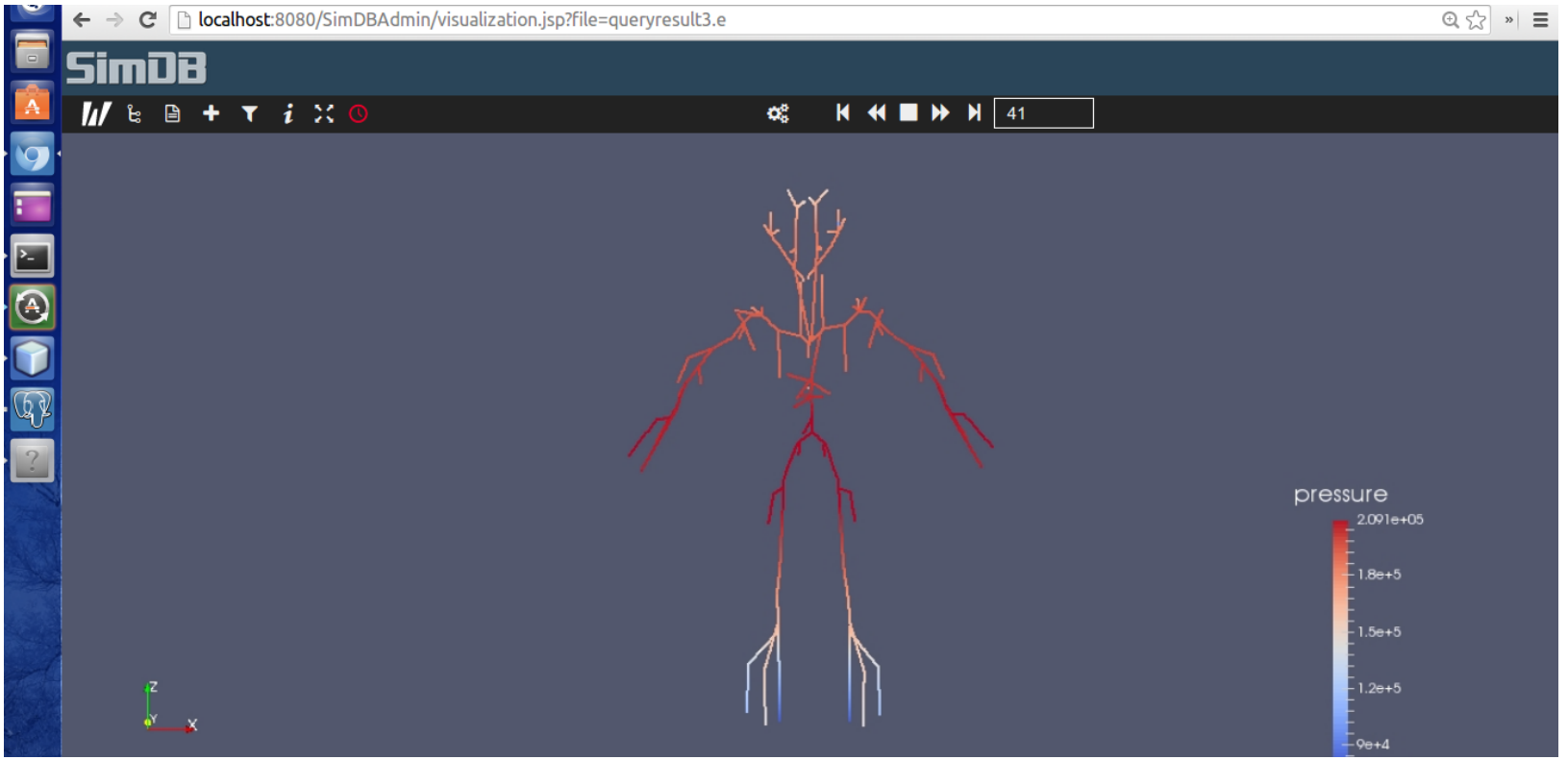
Query Formulation



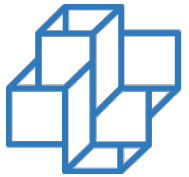
Query Execution



Visualization on Paraview



Final Comments



- We aim at providing data management quality to simulation data
  - High-level language for data analysis
  - Supporting Input and output data access
  - Metadata describing the process and its entities
  - Data publication and analysis
  - Data Visualization


thank you
danke
謝謝
ngiyabonga
teşekkür ederim
gracias
merci
obrigado
dziękuję
bedankt
go raibh maith agat
mochchakkeram
dank je
tapadh leat
спасибо
рахмат
Баярлалаа
faafetai lava
kiitos
dankie
dhanyavad
nanni
nantri
mauruuru
hvala
gràcie
bayarlalaa
köszönöm
enkosi
vinaka
спасибі
blagodaram
misaotra
matondo
paldies
grazzi
mahalo
хвала
asante
manana
obrigada
tenkí
murakoze
chokrane
mamnun
djiere dieuf
tau
дякую
akun
dankon
aciü
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
sukriya
kop khun krap
sagolun
sobodi
dèkuji
mèsi
didi madloba
najis tuke
kam sah hamnida
rahmat
terima kasih
তোমাকে ধন্যবাদ
감사합니다
ありがとう
tanemirt
rahmet
grazie
xièxie
ευχαριστώ
arigatô
takk
dakujem
trugarez
diolch
dhanyavadagalu
shukriya
mercé
мерси