

Data Management and Analysis at the Large Scale Data Facility

A. García¹, S. Bourov¹, A. Hammad¹, T. Jejkal², J. Otte³, S. Pfeiffer⁴, T. Schenker⁴, C. Schmidt⁴, J. van Wezel¹, B. Neumair¹, A. Streit¹

¹ Steinbuch Centre for Computing, KIT

² Institute for Data Processing and Electronics, KIT

³ Institute of Toxicology and Genetics, KIT

⁴ Institute for Applied Computer Science, KIT

ICDIM 2011, Melbourne | September 27, 2011

KARLSRUHE INSTITUTE OF TECHNOLOGY

- Many experiments have a data-management problem!
- This project aims at improving the situation, with adapted infrastructure and services
- Data Intensive Computing workflows are critical for the value of the data
- We present a concrete data intensive analysis use case at the LSDF

What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
 - ⇒ experiments need storage
 - ⇒ need sophisticated **data analysis workflows**
- Old data is very valuable
 - to analyse change in time
 - for reprocessing
 - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data
 - ⇒ adequate **meta data** greatly increases data value
 - ⇒ **single big scientific DB is more valuable than many small ones**
- Data is used by large virtual communities!
 - ⇒ access to data is critical



What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
 - ⇒ experiments need storage
 - ⇒ need sophisticated **data analysis workflows**
- Old data is very valuable
 - to analyse change in time
 - for reprocessing
 - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data
 - ⇒ adequate **meta data** greatly increases data value
 - ⇒ **single big scientific DB is more valuable than many small ones**
- Data is used by large virtual communities!
 - ⇒ access to data is critical



What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
 - ⇒ experiments need storage
 - ⇒ need sophisticated **data analysis workflows**
- Old data is very valuable
 - to analyse change in time
 - for reprocessing
 - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data
 - ⇒ adequate **meta data** greatly increases data value
 - ⇒ **single big scientific DB is more valuable than many small ones**
- Data is used by large virtual communities!
 - ⇒ access to data is critical



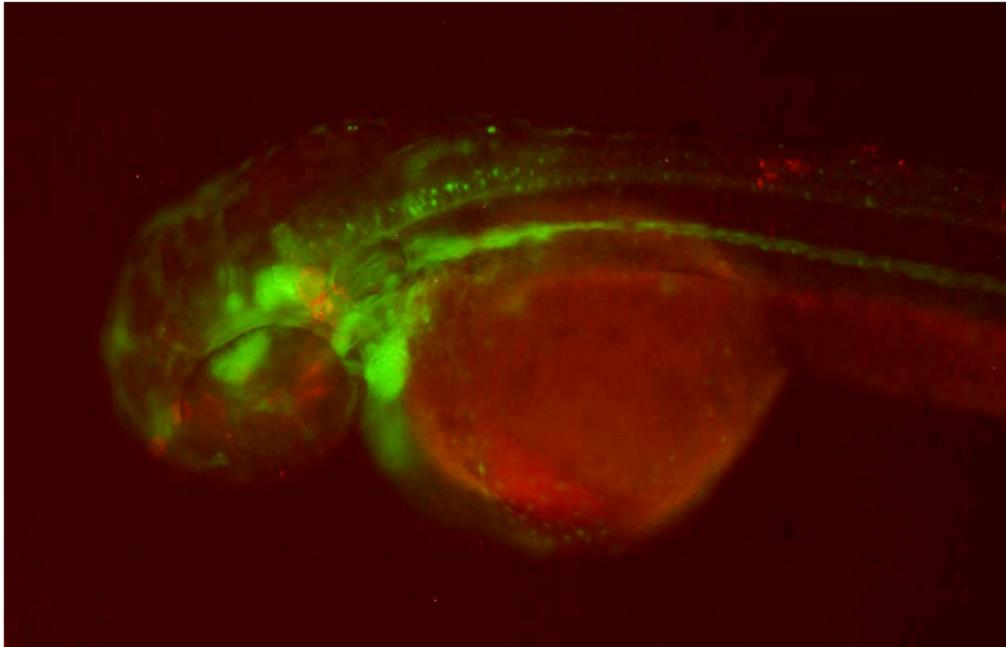
What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
 - ⇒ experiments need storage
 - ⇒ need sophisticated **data analysis workflows**
- Old data is very valuable
 - to analyse change in time
 - for reprocessing
 - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data
 - ⇒ adequate **meta data** greatly increases data value
 - ⇒ **single big scientific DB is more valuable than many small ones**
- Data is used by large virtual communities!
 - ⇒ access to data is critical



Why do experiments produce so much data?



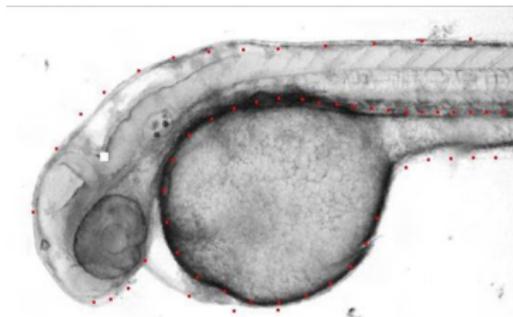
Zebrafish embryo, raw picture, 4MB (24 per fish)

Why do experiments produce so much data?

- Institute of Toxicology and Genetics @ KIT
 - ⇒ Zebra fishes' embryonal development reconstruction
 - ⇒ Toxicological studies of drugs
 - High Throughput Microscopy
 - fully automated microscopes
 - robot moves object to microscope
 - can potentially run 24*7
 - produce high resolution images (4 MB each)
 - over varying parameters (focus point, cell markers, ...)
 - ≈200k images per day, ≈ 2 TB/day
 - Estimated: 1+ PB/year in 2012,
6 PB/year in 2014
 - Raw data must be heavily analysed

Why do experiments produce so much data?

- Institute of Toxicology and Genetics @ KIT
 - ⇒ Zebra fishes' embryonal development reconstruction
 - ⇒ Toxicological studies of drugs
 - High Throughput Microscopy
 - fully automated microscopes
 - robot moves object to microscope
 - can potentially run 24*7
 - produce high resolution images (4 MB each)
 - over varying parameters (focus point, cell markers, ...)
 - ≈200k images per day, ≈ 2 TB/day
 - Estimated: 1+ PB/year in 2012,
6 PB/year in 2014
 - Raw data must be heavily analysed



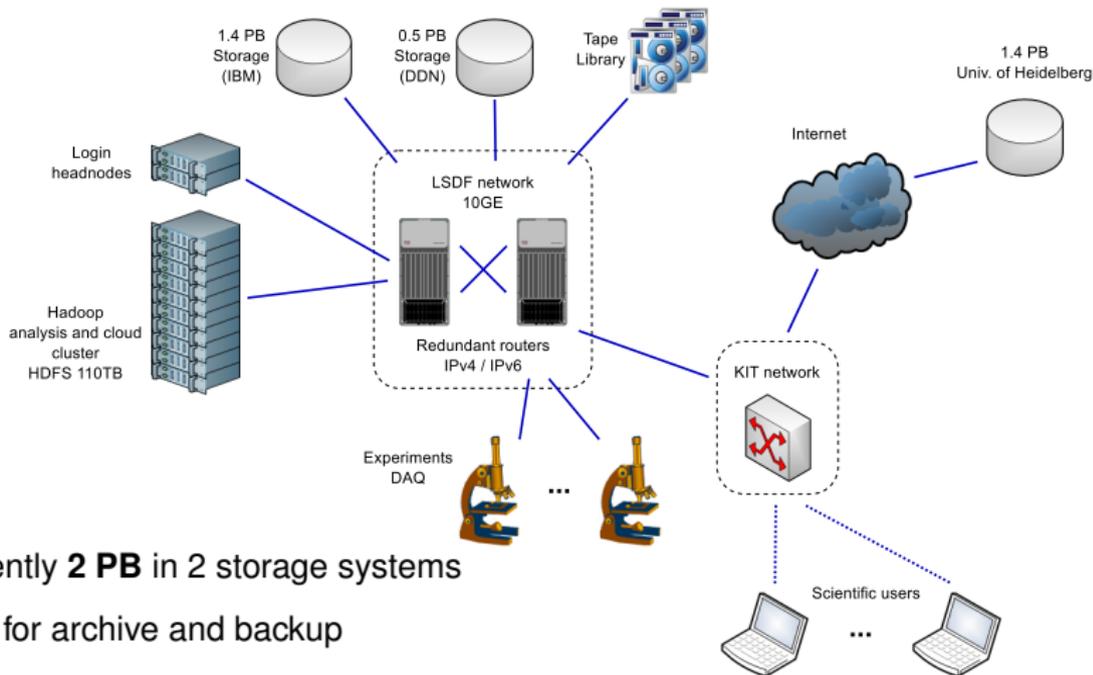
LSDF started end of 2009 at KIT

- involving several institutes
 - in cooperation with BioQuant of Univ. Heidelberg
-
- to address the needs of Data Intensive Science:
 - providing **large scale storage** and data processing
 - open protocols and APIs for access to data and **metadata**
 - **added value services** for **community specific needs** (multi-disciplinary)

LSDF started end of 2009 at KIT

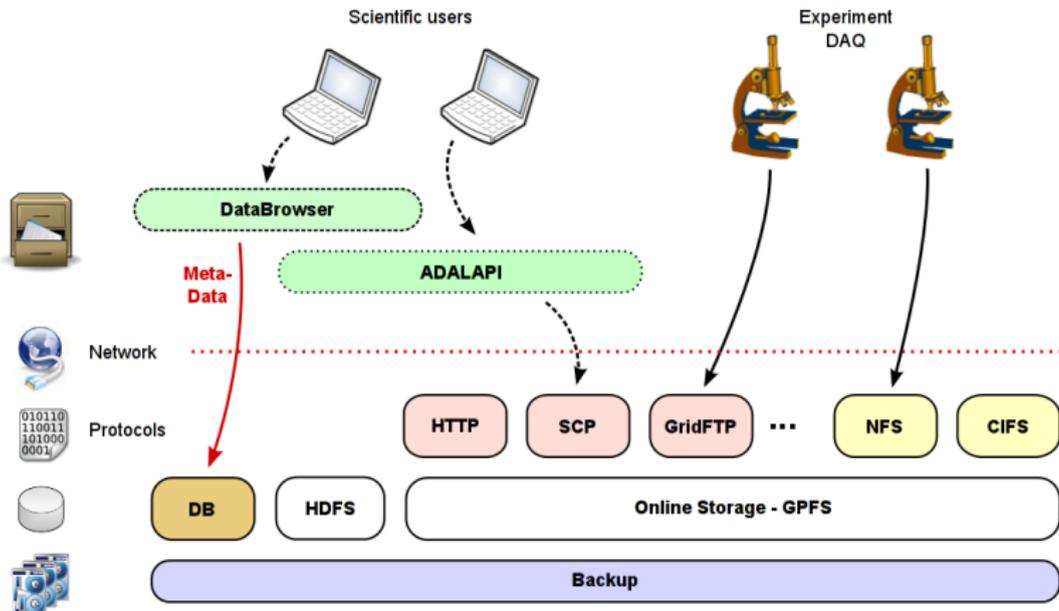
- involving several institutes
 - in cooperation with BioQuant of Univ. Heidelberg
-
- to address the needs of Data Intensive Science:
 - providing **large scale storage** and data processing
 - open protocols and APIs for access to data and **metadata**
 - **added value services** for **community specific needs** (multi-disciplinary)

What infrastructure are we talking about?



- Currently **2 PB** in 2 storage systems
- Tape for archive and backup
- Dedicated **10 GE** network backbone
- Direct network connection to some institutes

The current architecture



How to deal with so much data?

- Metadata is essential

- Needs to be stored and kept up to date with data
- Metadata schema is highly project-dependent

⇒ use a project metadata DB

Example:

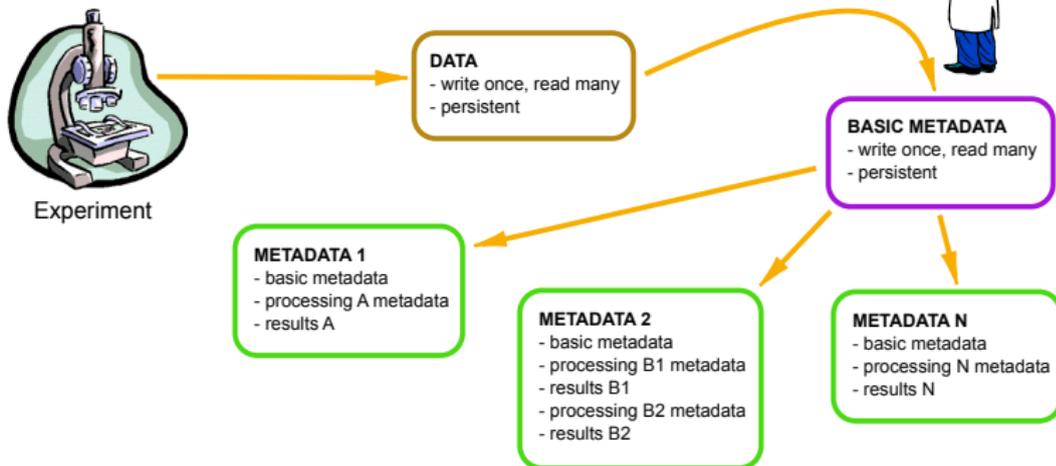
How to deal with so much data?

■ Metadata is essential

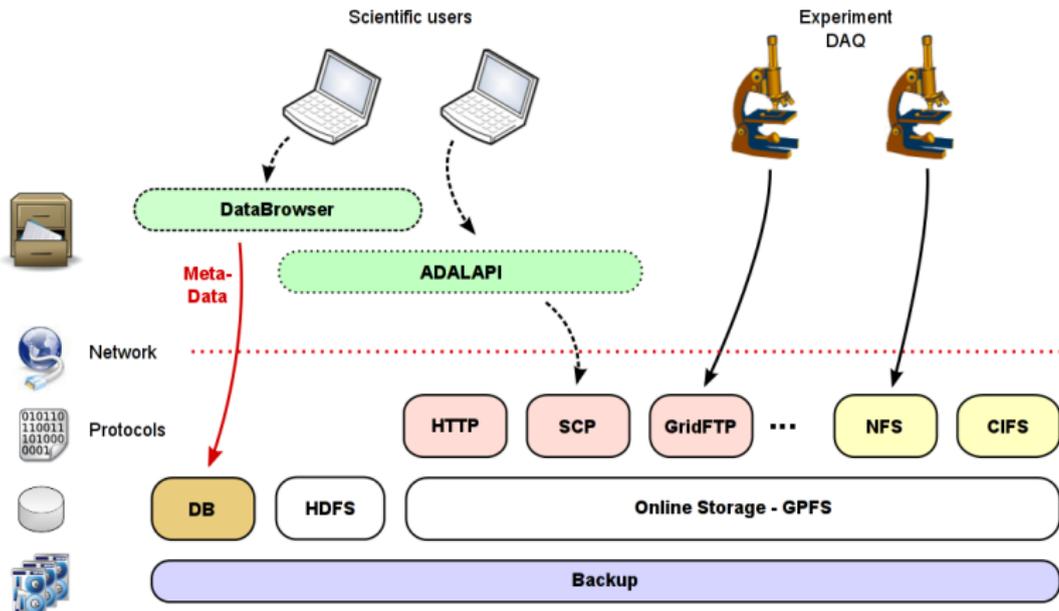
- Needs to be stored and kept up to date with data
- Metadata schema is highly project-dependent

⇒ use a project metadata DB

Example:



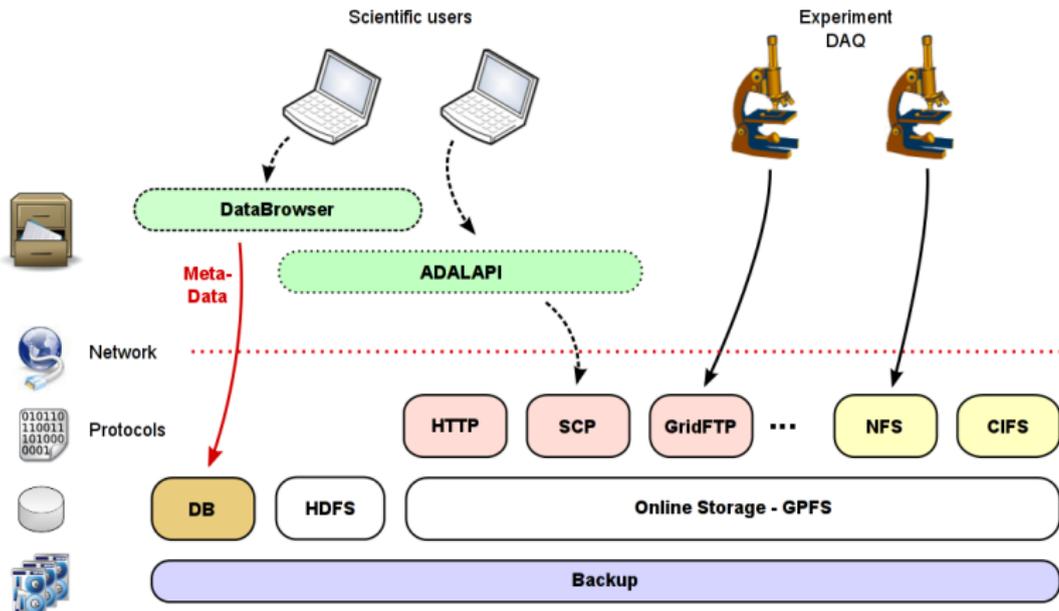
The current architecture



- Different components accessible through different protocols and APIs
 - ⇒ need a **unified access layer**
 - **Abstract Data Access Layer**, low-level interface to LSDF
 - ⇒ extensible to support new backends, authentication mechanisms
- For end-users: **DataBrowser**
 - graphical tool for exploring and managing the LSDF data
 - also library for meta-data access
 - based on **ADALAPI**
 - connects to the meta-data repository
 - will be available as web GUI

R. Stotzka et. al., Perspective of the Large Scale Data Facility supporting nuclear fusion applications, Proceedings PDP 2011.

The current architecture



Can we process the data?

- Data has to be processed!
- Exascale \Rightarrow **bring computing to the data!!**
(15 days to transfer 1 PB over ideal 10Gb/s link)

\Rightarrow dedicated 60 nodes cluster

- **Hadoop environment**
 - + 110 TB Hadoop filesystem
 - extreme scalability
- **Cloud environment OpenNebula**
 - users can deploy own dedicated data-processing VMs
 - reliable, highly flexible, and very fast to deploy

Can we process the data?

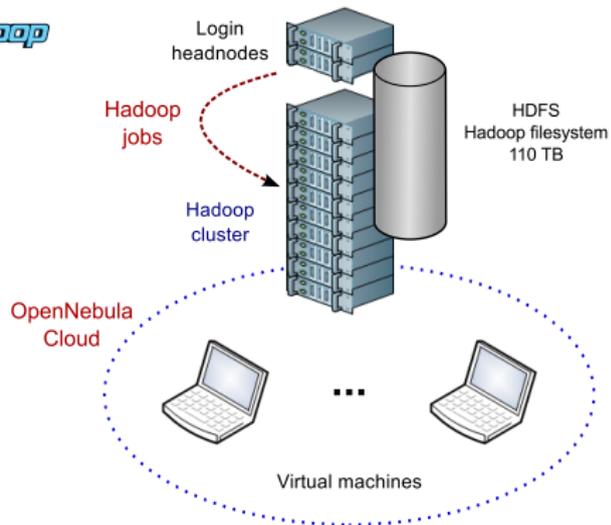
- Data has to be processed!
- Exascale \Rightarrow **bring computing to the data!!**
(15 days to transfer 1 PB over ideal 10Gb/s link)

\Rightarrow dedicated 60 nodes cluster

- **Hadoop environment** 
+ 110 TB Hadoop filesystem
 - extreme scalability

- **Cloud environment OpenNebula**

- users can deploy own dedicated data-processing VMs
- reliable, highly flexible, and very fast to deploy



Can we process the data?

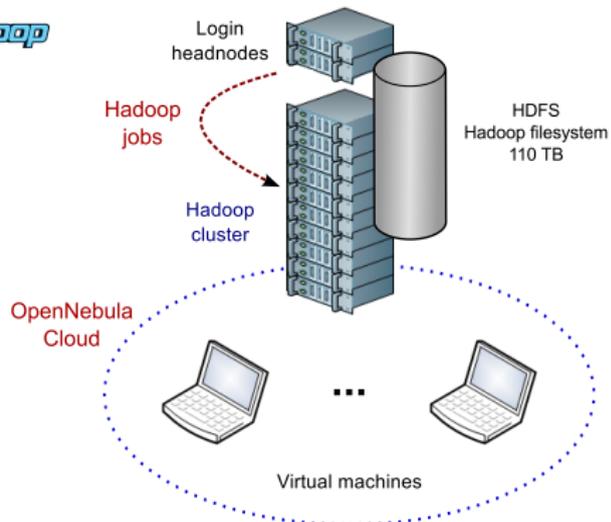
- Data has to be processed!
- Exascale \Rightarrow **bring computing to the data!!**
(15 days to transfer 1 PB over ideal 10Gb/s link)

\Rightarrow dedicated 60 nodes cluster

- **Hadoop environment** 
+ 110 TB Hadoop filesystem
 - extreme scalability

- **Cloud environment OpenNebula**
OpenNebula.org

- users can deploy own dedicated data-processing VMs
- reliable, highly flexible, and very fast to deploy

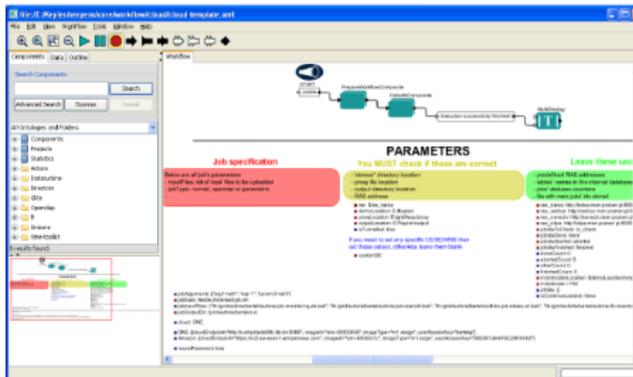


- With Cloud instances, *if customized SW environment is required*
 - Integrated with the Kepler workflow orchestrator
 - user-friendly interface

- With workflows in the Hadoop cluster

How to deal with data?

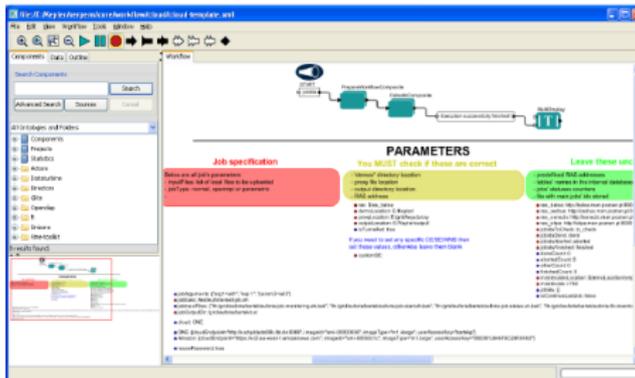
- With Cloud instances, *if customized SW environment is required*
 - Integrated with the Kepler workflow orchestrator
 - user-friendly interface



- With workflows in the Hadoop cluster

How to deal with data?

- With Cloud instances, *if customized SW environment is required*
 - Integrated with the Kepler workflow orchestrator
 - user-friendly interface



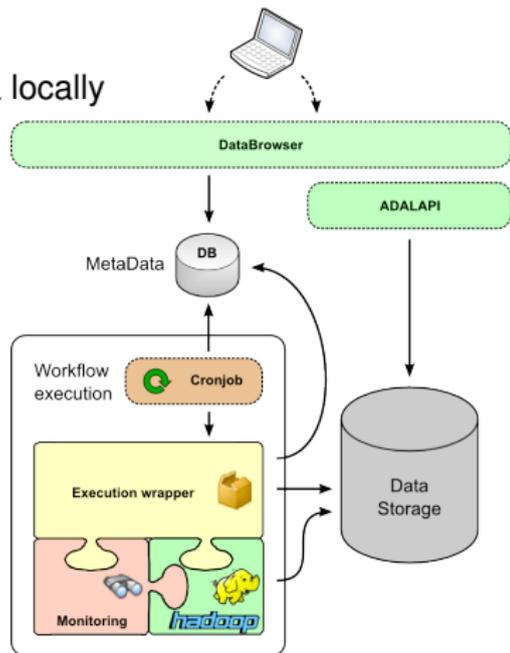
- With workflows in the Hadoop cluster

- Experiments should be able to process data locally
⇒ [help the users automate the workflows](#)
- Allow tagging data and triggering execution via [DataBrowser](#)
- Data from finished workflows stored and tagged in DB

- used for zebrafish microscopy data

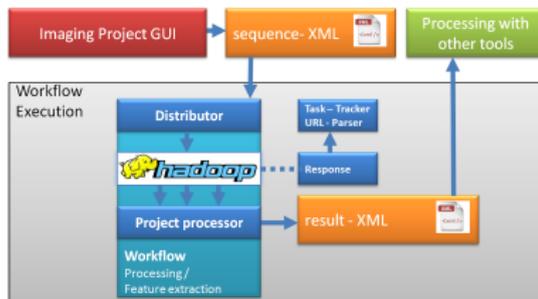
T. Jejkal et al., LAMBDA – The LSDF Execution Framework for Data Intensive Applications, Proceedings PDP 2012.

- Experiments should be able to process data locally
⇒ help the users automate the workflows
- Allow tagging data and triggering execution via **DataBrowser**
- Data from finished workflows stored and tagged in DB
- used for zebrafish microscopy data



T. Jejkal et al., LAMBDA – The LSDF Execution Framework for Data Intensive Applications, Proceedings PDP 2012.

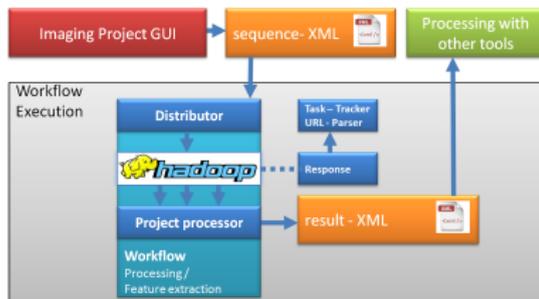
- Zebra-fish data sets
- Data imported with DataBrowser, **meta-data** added
- **ImagingProjectGUI** to define image analysis chain
 - based on operators from different libraries (VTK, ITK, Matlab,...)
 - creates **sequence-XML**, XML description of operators' chain
- User tags data for processing
 - using DataBrowser
 - associating sequence-XML
- Cron job starts execution based on tags
 - projectProcessor
- Executed in Hadoop cluster
 - **result-XML** stored
- Speedup of 60-120 depending on operators
 - I/O bound analysis



S. Pfeiffer, M. Mai, and J. Callies, *On the Computational Benefit of Tensor Separation for High-Dimensional Discrete Convolutions, Multidimensional Systems and Signal Processing 31, 2010, pp. 1-25.*

J.C. Otte, S. Pfeiffer, T. Schenker, and C. Schmidt, *Concept of multidimensional image analysis with flexible operators and cluster-computing, 6th European Molecular Imaging Meeting, Leiden, The Netherlands, 2011.*

- Zebra-fish data sets
- Data imported with DataBrowser, **meta-data** added
- **ImagingProjectGUI** to define image analysis chain
 - based on operators from different libraries (VTK, ITK, Matlab,...)
 - creates **sequence-XML**, XML description of operators' chain
- User tags data for processing
 - using DataBrowser
 - associating sequence-XML
- Cron job starts execution based on tags
 - projectProcessor
- Executed in Hadoop cluster
 - **result-XML** stored
- Speedup of 60-120 depending on operators
 - I/O bound analysis



S. Pfeiffer, M. Mai, and J. Callies, *On the Computational Benefit of Tensor Separation for High-Dimensional Discrete Convolutions*, *Multidimensional Systems and Signal Processing* 31, 2010, pp. 1-25.

J.C. Otte, S. Pfeiffer, T. Schenker, and C. Schmidt, *Concept of multidimensional image analysis with flexible operators and cluster-computing*, *6th European Molecular Imaging Meeting*, Leiden, The Netherlands, 2011.

- Improved storage, network capacity: 6 PB in 2012

Investigate and deploy new technologies

- Data management system iRODS
- Object Storage
- Additional communities being integrated
 - Meteorology and climate research (“archival” quality)
 - Geophysics, seismology data
 - KATRIN experiment, neutrino mass
 - DARIAH, arts and humanities

Project expansion

- to offer more community tailored support

- Improved storage, network capacity: 6 PB in 2012

Investigate and deploy new technologies

- Data management system iRODS
- Object Storage

- Additional communities being integrated
 - Meteorology and climate research (“archival” quality)
 - Geophysics, seismology data
 - KATRIN experiment, neutrino mass
 - DARIAH, arts and humanities

Project expansion

- to offer more community tailored support

- Improved storage, network capacity: 6 PB in 2012

Investigate and deploy new technologies

- Data management system iRODS
- Object Storage

- Additional communities being integrated
 - Meteorology and climate research (“archival” quality)
 - Geophysics, seismology data
 - KATRIN experiment, neutrino mass
 - DARIAH, arts and humanities

Project expansion

- to offer more community tailored support

- Infrastructure, 2 PB and storage services up and running
- First software tools available
- Experimental data being stored and processed
- Many scientific communities interested and getting involved

Focus on users, added value services

- Can't just "store files"
 - Training for new tools, data management workflows
-
- Same problem at most (all?) research institutions
 - ⇒ Open for new partnerships, international collaborations

Thanks for listening!

Questions?