Karlsruhe Institute of Technology

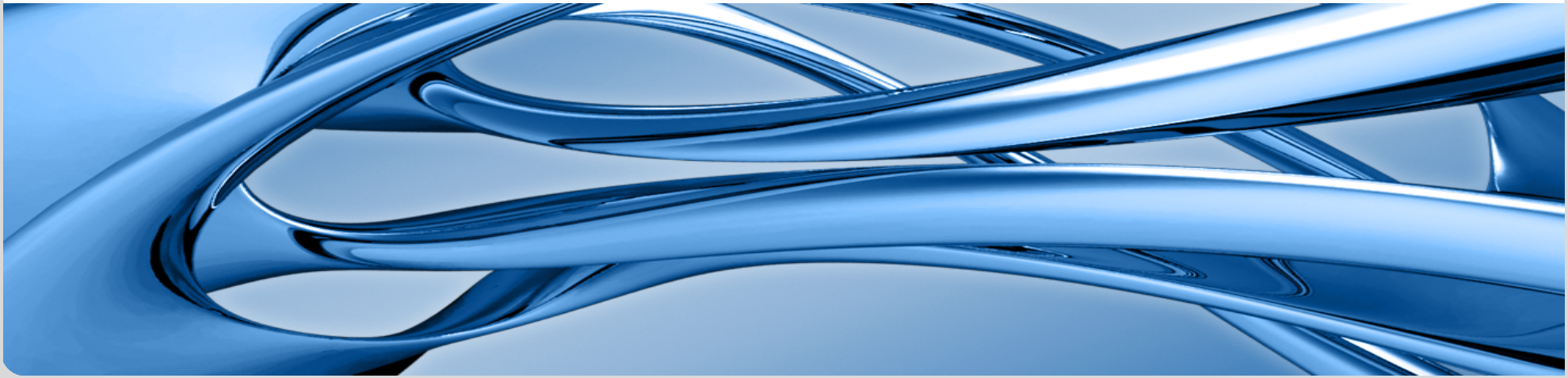
The Large Scale Data Facility

Data Intensive Computing for scientific Experiments

A. García[1], S. Bourov[1], A. Hammad[1], V. Hartmann[2], T. Jejkal[2], P. Neuberger[2], R .Stotzka[2], J. van Wezel[1], B. Neumair[1], A. Streit[1]

[1] Steinbuch Centre for Computing, KIT
[2] Institute for Data Processing and Electronics, KIT

PDSEC/IPDPS 2011, Anchorage | May 20, 2011



Executive summary

- 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 the Roadmap and Outlook

What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
  ⇐ Moore's law (cheaper IT)!
  - remember the Large Hadron Collider @ CERN? it's "small" today!
  ⇒ experiments need storage
  ⇒ need computationally intensive services
  ⇒ need sophisticated data analysis workflows

- Old data is very valuable
  - for reprocessing
  - to analyse change in time
  - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data
  ⇒ administration and accessibility greatly increases data value
  ⇒ single big DB with scientific data is more valuable than many small ones
- Data is used by large virtual communities!
  ⇒ communication and simple access to data is critical

What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
  ⇐ Moore's law (cheaper IT)!
  - remember the Large Hadron Collider @ CERN? it's "small" today!

  ⇒ experiments need storage

  ⇒ need computationally intensive services

  ⇒ need sophisticated data analysis workflows

- Old data is very valuable
  - for reprocessing
  - to analyse change in time
  - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data

  ⇒ administration and accessibility greatly increases data value

  ⇒ single big DB with scientific data is more valuable than many small ones
- Data is used by large virtual communities!

  ⇒ communication and simple access to data is critical

What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
  ⇐ Moore's law (cheaper IT)!
  - remember the Large Hadron Collider @ CERN? it's "small" today!
  ⇒ experiments need storage
  ⇒ need computationally intensive services
  ⇒ need sophisticated data analysis workflows

- Old data is very valuable
  - for reprocessing
  - to analyse change in time
  - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data
  ⇒ administration and accessibility greatly increases data value
  ⇒ single big DB with scientific data is more valuable than many small ones
- Data is used by large virtual communities!
  ⇒ communication and simple access to data is critical

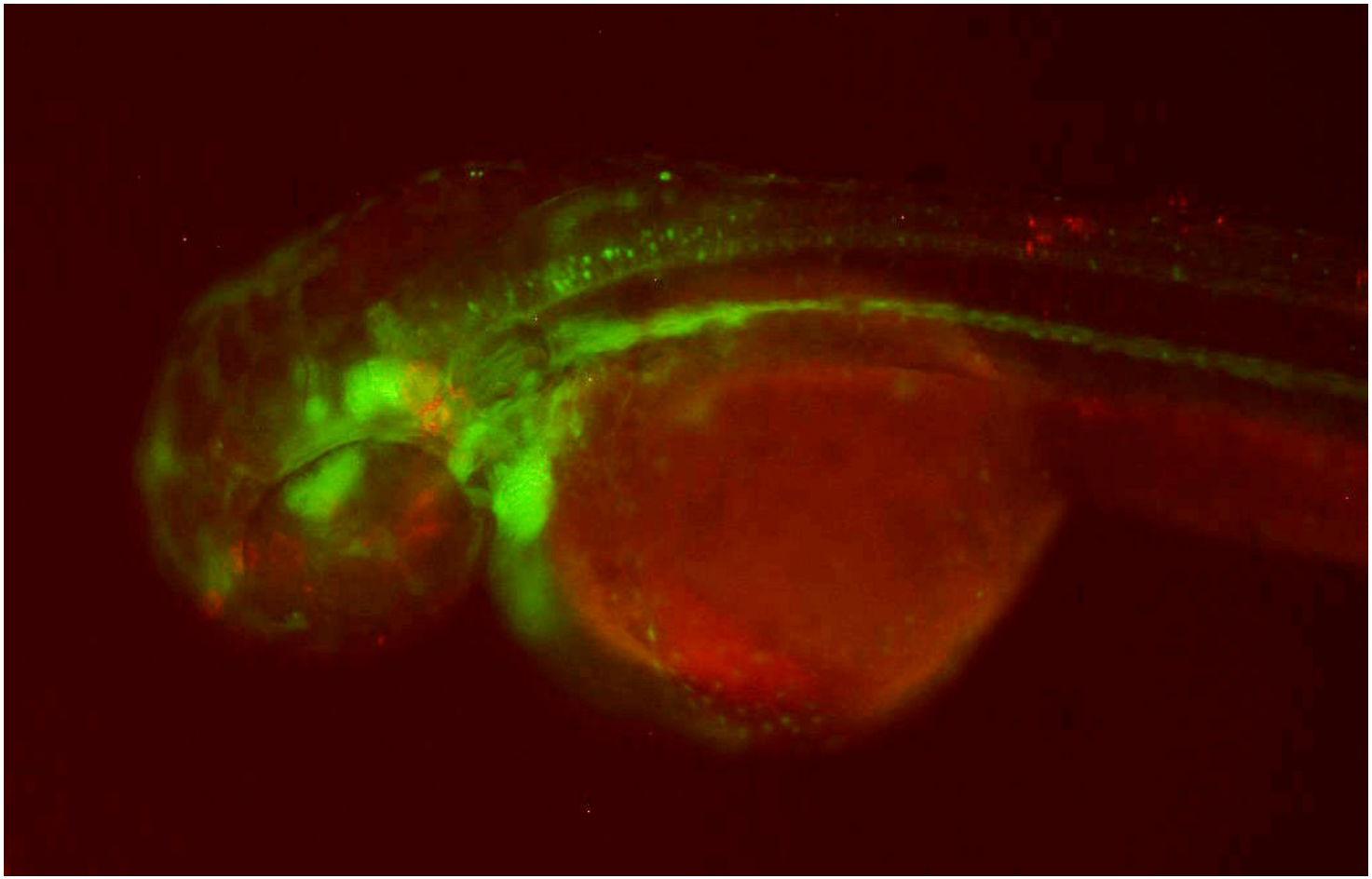
What is the data challenge?

Science produces data!

- Experiments getting exponentially more data
  ⇐ Moore's law (cheaper IT)!
  - remember the Large Hadron Collider @ CERN? it's "small" today!
  ⇒ experiments need storage
  ⇒ need computationally intensive services
  ⇒ need sophisticated data analysis workflows

- Old data is very valuable
  - for reprocessing
  - to analyse change in time
  - for analysis by other scientists, in other contexts
- Invisible (not-found, no-metadata) data is lost data
  ⇒ administration and accessibility greatly increases data value
  ⇒ single big DB with scientific data is more valuable than many small ones
- Data is used by large virtual communities!
  ⇒ communication and simple 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

  $\Rightarrow$ Zebra fishes' embryonal development reconstruction

  $\Rightarrow$ 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, wavelength, ...)
  - $\approx$200k images per day, 2 TB/day
  - Estimated: 1+ PB/year in 2012,
    6 PB/year in 2014
  - Raw data must be heavily analysed

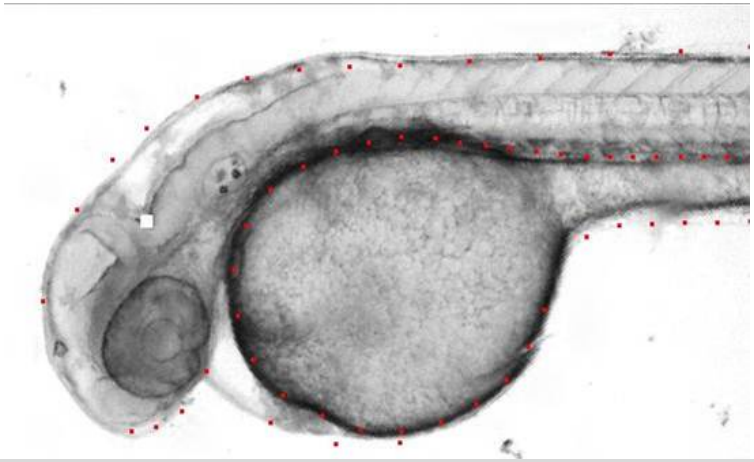
Why do experiments produce so much data?

- Institute of Toxicology and Genetics @ KIT

  $\Rightarrow$ Zebra fishes' embryonal development reconstruction

  $\Rightarrow$ 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, wavelength, ...)
  - $\approx$200k images per day, 2 TB/day
  - Estimated: 1+ PB/year in 2012,
    6 PB/year in 2014
  - Raw data must be heavily analysed



The Large Scale Data Facility Project

aka: LSDF

Started end of 2009 at KIT

- involving several institutes
- tight cooperation with BioQuant of Univ. Heidelberg

- to address the needs of Data Intensive Science
  - providing large scale storage
  - open protocols and APIs for access to data and metadata
  - transparent access over background storage and technology changes
  - added value services and tools for processing data
  - development and deployment of community specific services

The Large Scale Data Facility Project

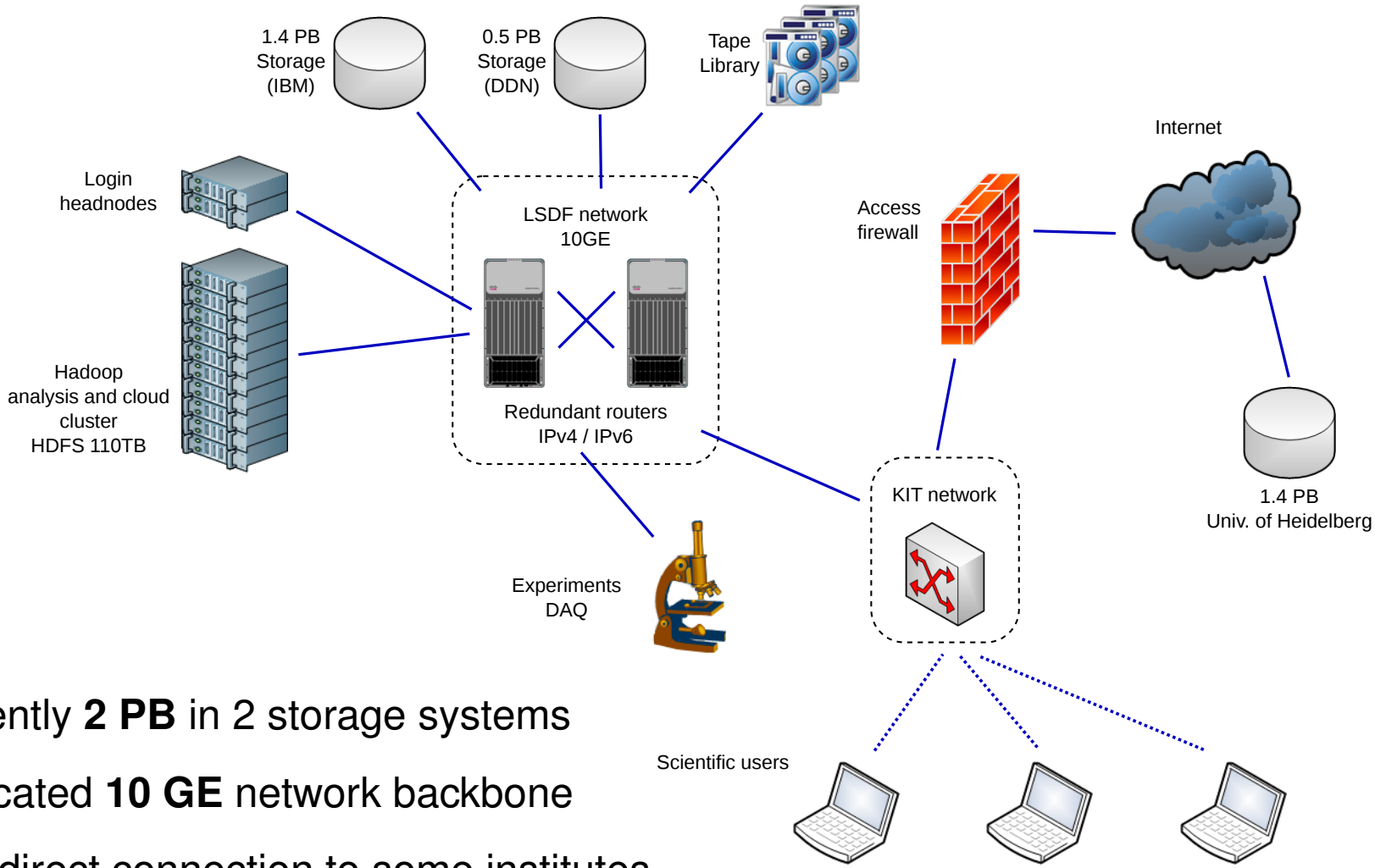
aka: LSDF

Started end of 2009 at KIT

- involving several institutes
- tight cooperation with BioQuant of Univ. Heidelberg

- to address the needs of Data Intensive Science
  - providing large scale storage
  - open protocols and APIs for access to data and metadata
  - transparent access over background storage and technology changes
  - added value services and tools for processing data
  - development and deployment of community specific services

What infrastructure are we talking about?



- currently **2 PB** in 2 storage systems
- dedicated **10 GE** network backbone
- with direct connection to some institutes
- tape backend for archive and backup

How to deal with so much data?

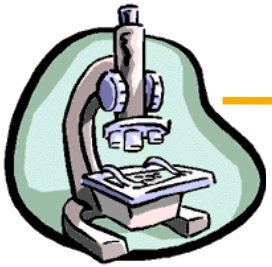
- Metadata is essential
  - Needs to be stored and kept up to date with data
  - Metadata schema is highly project-dependent

  $\Rightarrow$ we use a project metadata DB

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

$\Rightarrow$ we use a project metadata DB



Experiment

DATA
- write once, read many
- persistent

BASIC METADATA
- write once, read many
- persistent

METADATA 1
- basic metadata
- processing A metadata
- results A

METADATA 2
- basic metadata
- processing B1 metadata
- results B1
- processing B2 metadata
- results B2

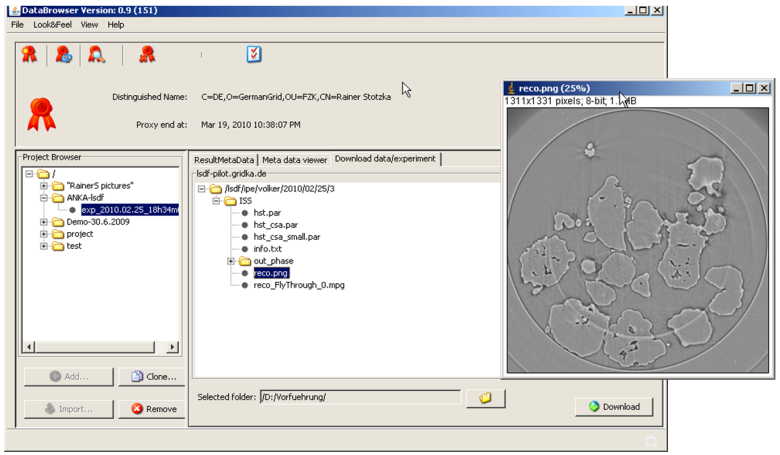
METADATA N
- basic metadata
- processing N metadata
- results N

Which access APIs and tools?

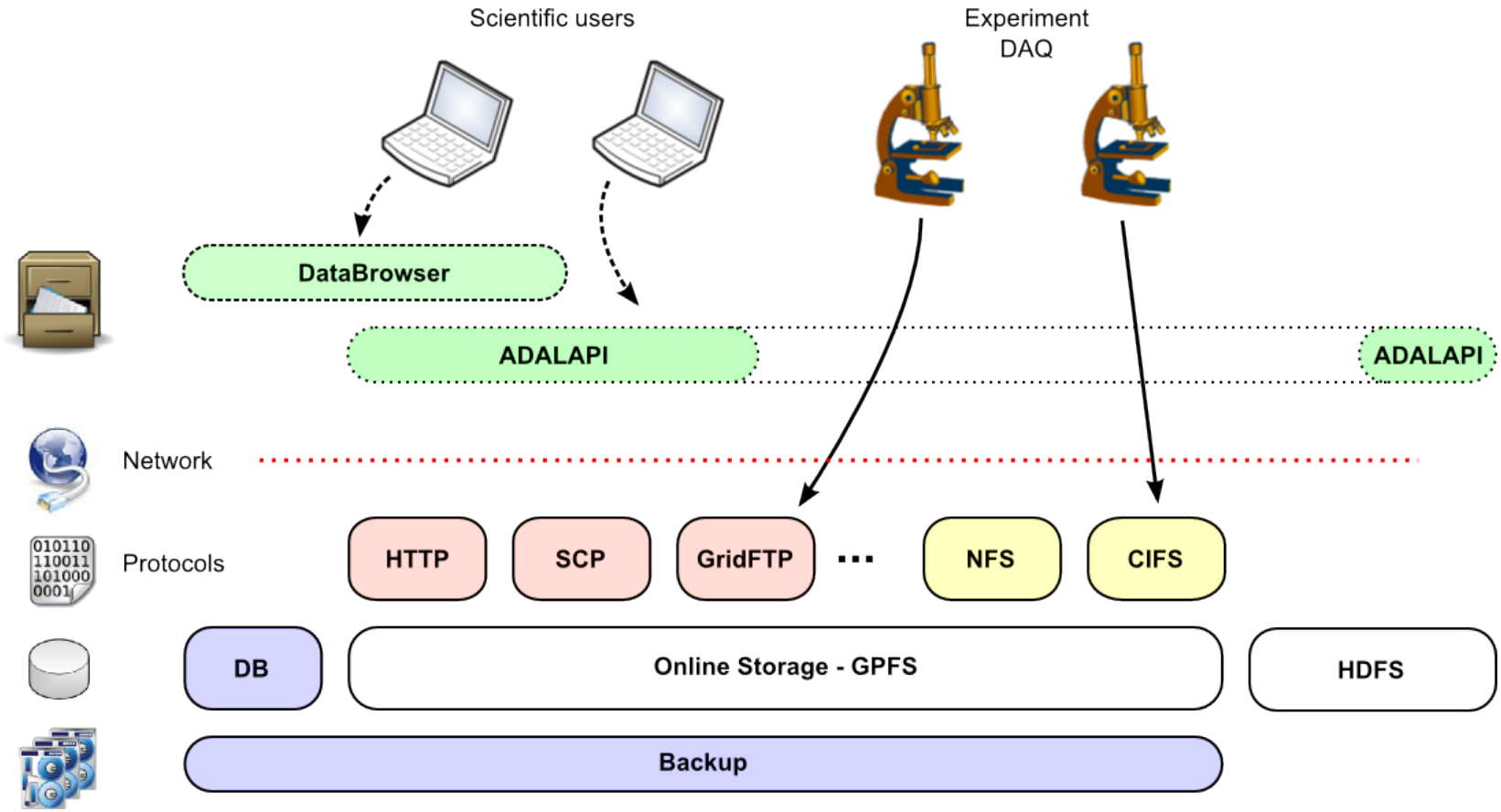
- Hardware and software choices limit the access protocols and APIs
  - $\Rightarrow$ not all components accessible through all methods
  - $\Rightarrow$ need a unified access layer
    - Abstract Data Access Layer, low-level interface to LSDF
    - $\Rightarrow$ extensible to support new backends, authentication mechanisms
- For end-users: DataBrowser
  - graphical tool for exploring and managing the LSDF data
  - based on ADAL-API
  - connects to the meta-data repository
  - will be available as web GUI

Which access APIs and tools?

- Hardware and software choices limit the access protocols and APIs
  - ⇒ not all components accessible through all methods
  - ⇒ 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
  - based on ADAL-API
  - connects to the meta-data repository
  - will be available as web GUI



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

      on commodity hardware
- Cloud environment OpenNebula
    - users can deploy own dedicated

      data-processing VMs

      (customized environment!)
    - reliable, highly flexible,

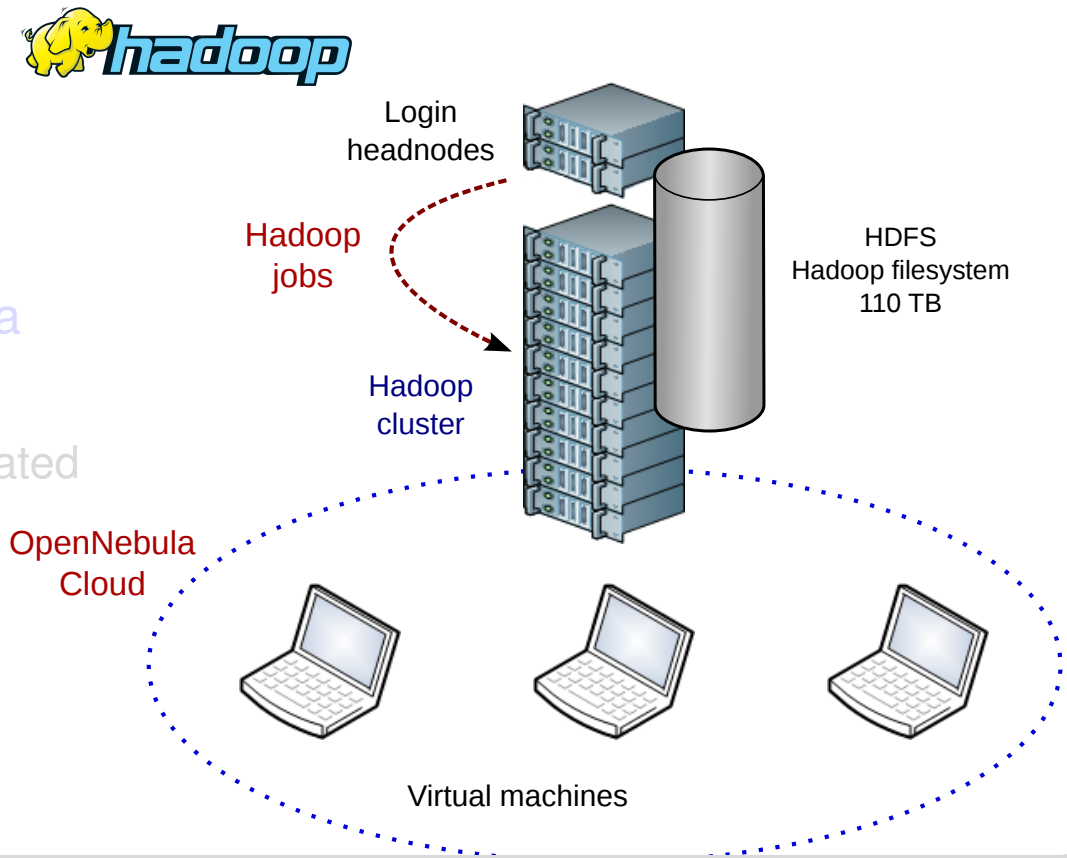
      and very fast to deploy

Can we process the data?

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

⇒ dedicated 60 nodes cluster

- Hadoop environment
  + 110 TB Hadoop filesystem
    - extreme scalability
      on commodity hardware
- Cloud environment OpenNebula
  - users can deploy own dedicated data-processing VMs (customized environment!)
  - reliable, highly flexible, and very fast to deploy

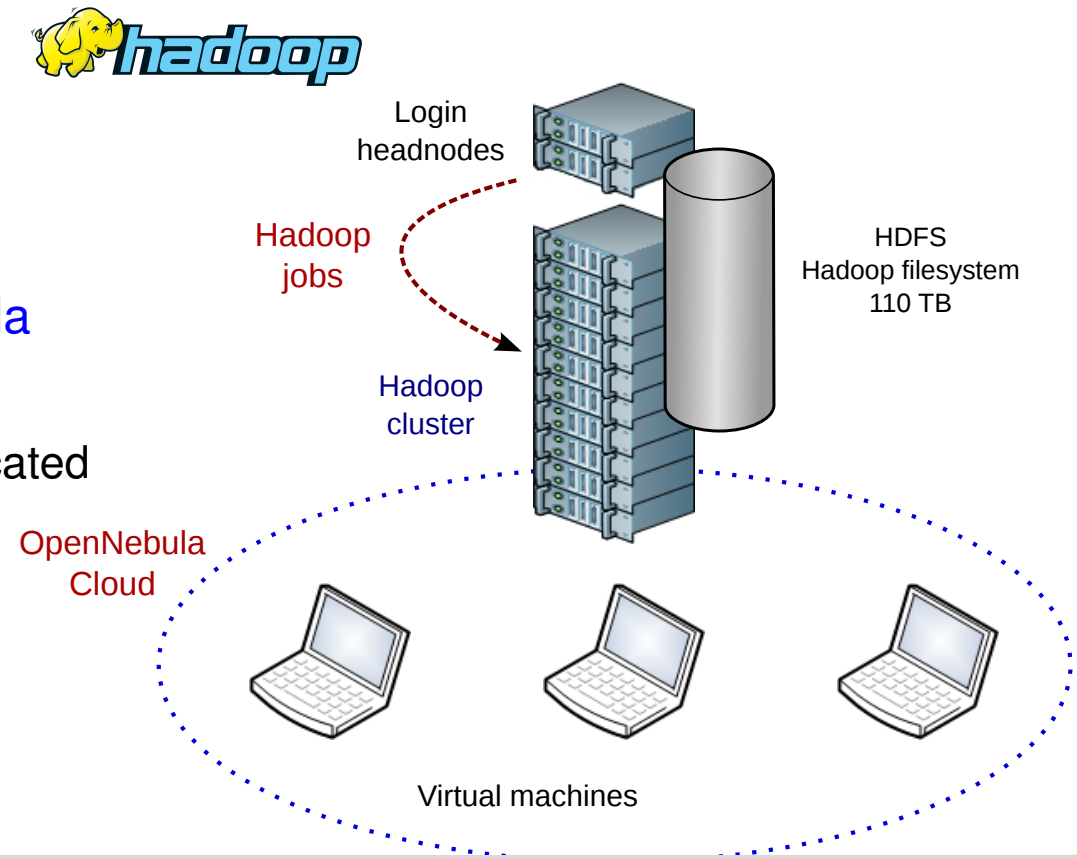


Can we process the data?

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

⇒ dedicated 60 nodes cluster

- Hadoop environment
  + 110 TB Hadoop filesystem
    - extreme scalability
      on commodity hardware
- Cloud environment OpenNebula
  OpenNebula.org
    - users can deploy own dedicated
      data-processing VMs
      (customized environment!)
    - reliable, highly flexible,
      and very fast to deploy

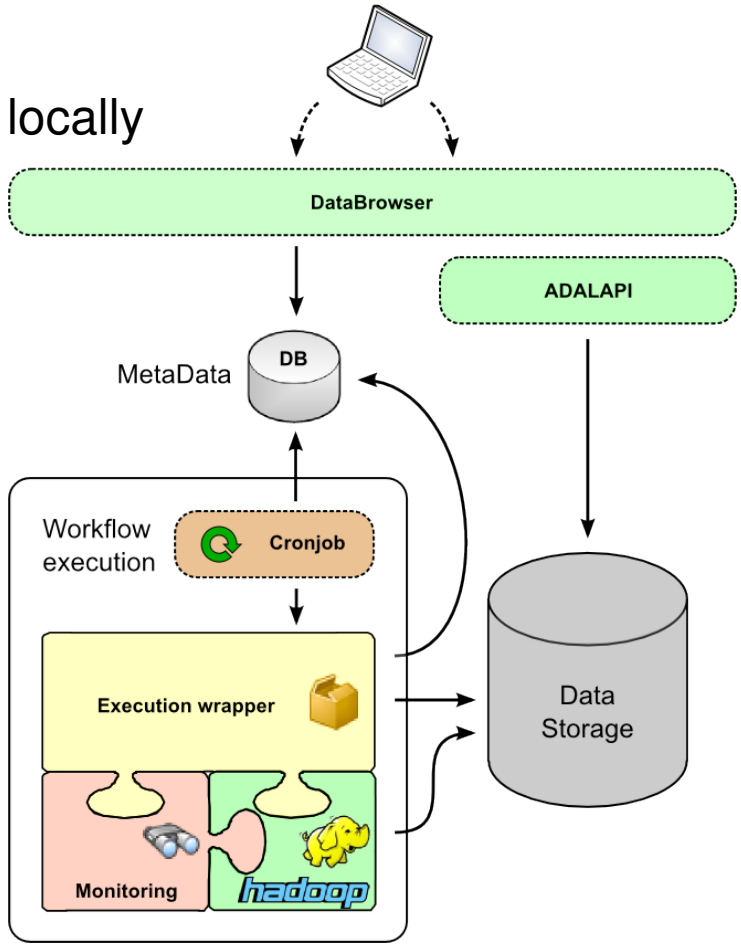


Data processing at LSDF

- Experiments should be able to process data locally
  $\Rightarrow$ 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

Data processing at LSDF

- Experiments should be able to process data locally
  $\Rightarrow$ 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

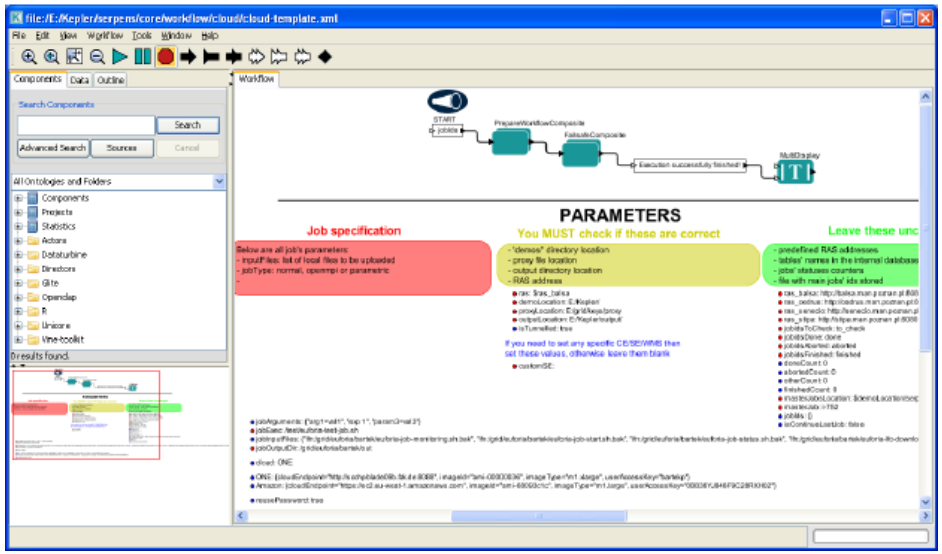


How to deal with data?

- With dedicated Hadoop applications
  - DNA sequencing and reconstruction using Hadoop tools
  - 3D Biomedical data visualization processing 1 TB dataset in 20min
- With Cloud instances, *if customized SW environment is required*
  - Integrated with the Kepler workflow orquestrator
  - user-friendly interface

How to deal with data?

- With dedicated Hadoop applications
  - DNA sequencing and reconstruction using Hadoop tools
  - 3D Biomedical data visualization processing 1 TB dataset in 20min
- With Cloud instances, *if customized SW environment is required*
  - Integrated with the Kepler workflow orquestrator
  - user-friendly interface



What's ahead?

- Improved storage, network capacity: 6 PB in 2012

Investigate and deploy new technologies

- Data management system iRODS (ongoing)
- Object Storage

- Additional communities integrated in 2011
  - KATRIN experiment, neutrino mass
  - Meteorology and climate research ("archival" quality)
  - Geophysics
- Expanding project to offer more community tailored support

Added-value services

- working with experiments towards integrated data-management workflow
  - KATRIN experiment
  - ANKA synchrotron radiation source

What's ahead?

- Improved storage, network capacity: 6 PB in 2012

Investigate and deploy new technologies

- Data management system iRODS (ongoing)
- Object Storage

- Additional communities integrated in 2011
  - KATRIN experiment, neutrino mass
  - Meteorology and climate research ("archival" quality)
  - Geophysics
- Expanding project to offer more community tailored support

Added-value services

- working with experiments towards integrated data-management workflow
  - KATRIN experiment
  - ANKA synchrotron radiation source

What's ahead?

- Improved storage, network capacity: 6 PB in 2012

Investigate and deploy new technologies

- Data management system iRODS (ongoing)
- Object Storage

- Additional communities integrated in 2011
  - KATRIN experiment, neutrino mass
  - Meteorology and climate research ("archival" quality)
  - Geophysics
- Expanding project to offer more community tailored support

Added-value services

- working with experiments towards integrated data-management workflow
  - KATRIN experiment
  - ANKA synchrotron radiation source

Concluding remarks

- Infrastructure 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

  $\Rightarrow$ Open for new partnerships, international collaborations

Thanks for listening!

Karlsruhe Institute of Technology

Questions?