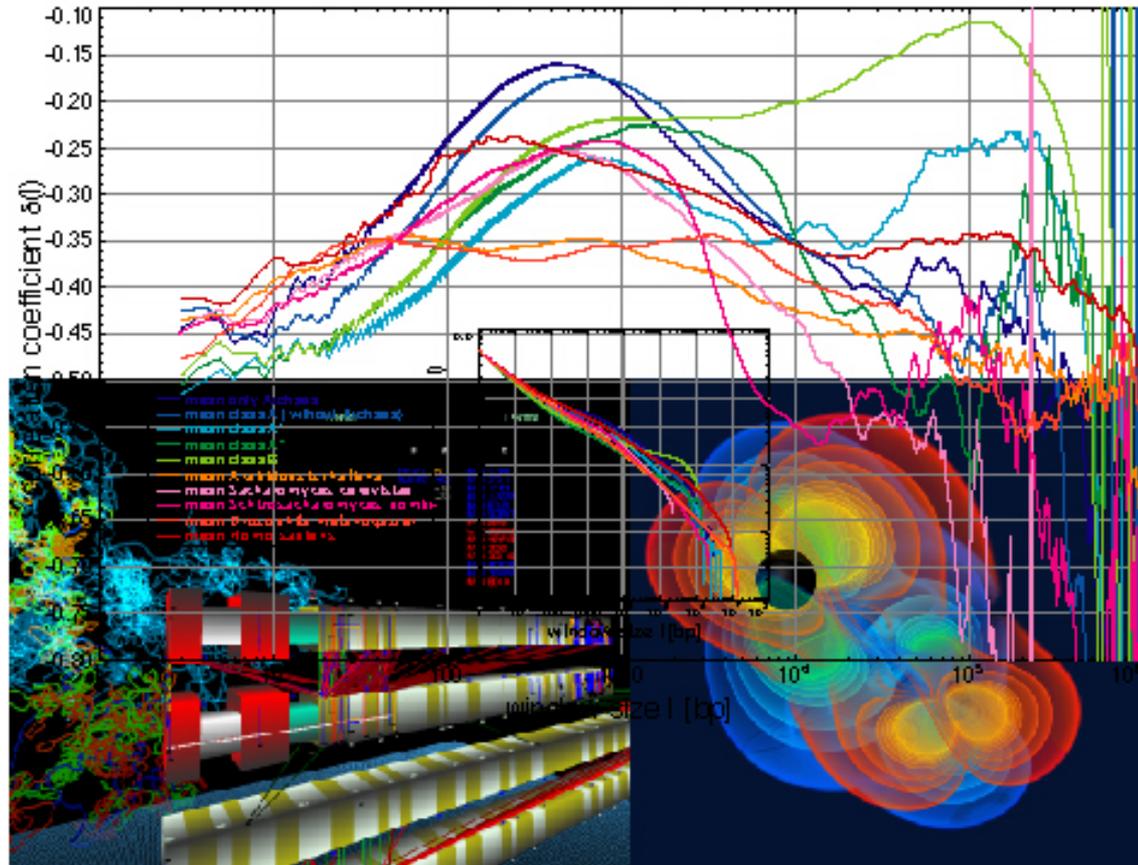# *Virtual Research Infrastructure for Genom- and LISA-Research in Europe*

Einstein@Home & LISA
Alexander Beck-Ratzka (AEI)

System Biology & Genom Analysen
Tobias A. Knoch (ErasmusMC/
BioQuant/DKFZ)

Agenda

- Grid Player AEI
  - Einstein@Home
  - Grid Application Toolkit (GAT)
  - GridSphere
- Grid Player ErasmusMC/Bioquant/DKFZ
  - Project/Job/Applikations Types
  - Security
  - Statistics
- Suggestions for a Research Infrastructure for Genom- and LISA research

Agenda

- Grid Player AEI
  - Einstein@Home
  - Grid Application Toolkit (GAT)
  - GridSphere
- Grid Player ErasmusMC/Bioquant/DKFZ
  - Project/Job/Applikations Types
  - Security
  - Statistics
- Suggestions for a Research Infrastructure for Genom- and LISA research

Einstein@Home

- Einstein@Home is the most successful use case in D-Grid:
  - Executables available for nearly all architectures,
  - Based on the BOINC Framework.
  - This makes Einstein@Home a perfect candidate for a heterogeneous Grid.

- Before the execution on a Grid host an Einstein@Home Client must be deployed there.
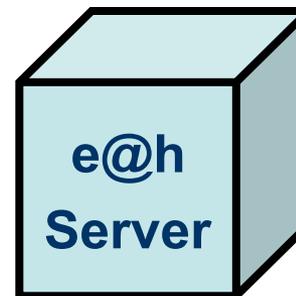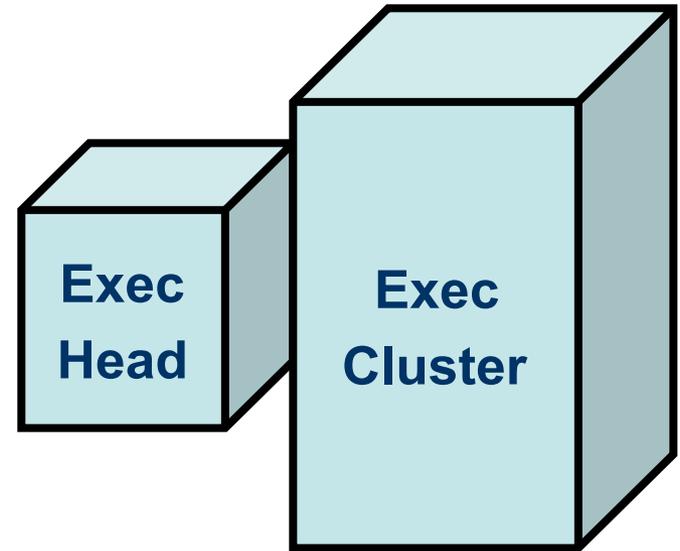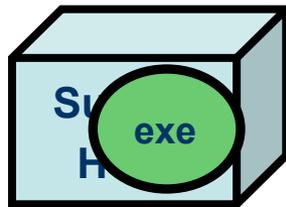
e@h – Deployment (1)

- Automatic Deployment.
- Execution of e@h in the Grid requires the installation of:
  - The GEO600 framework (Perl)
  - A BOINC client for e@h
  - A MySQL client (statistics)
  - appropriate Perl modules.
- Deployment realized as Globus job in a Globus Grid

e@h – how it works

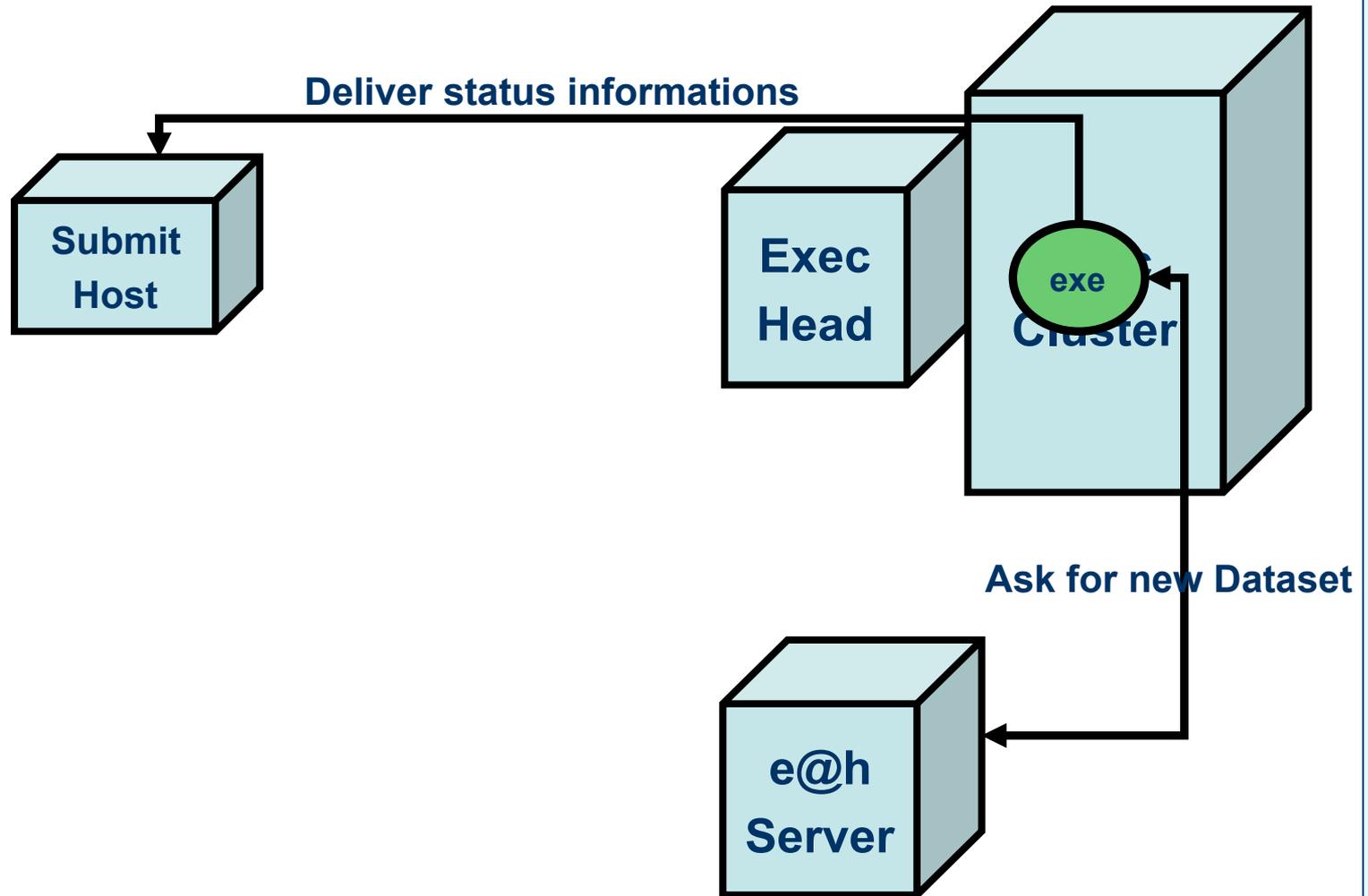- One single Einstein@Home as worker node.
- Every Einstein@Home job is identified with a unique task ID.
- Einstein@Home Client asks for dataset to be nalayzed at Einstein@Home server.
- Results are re-transmitted to the server.
- Reaching the walltime limits the checkpoints are stored at the execution host, to serve as starting point for the next job.
- Submission via Perl scripts.

e@h schematic view
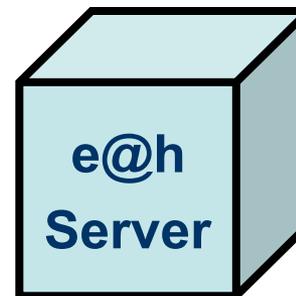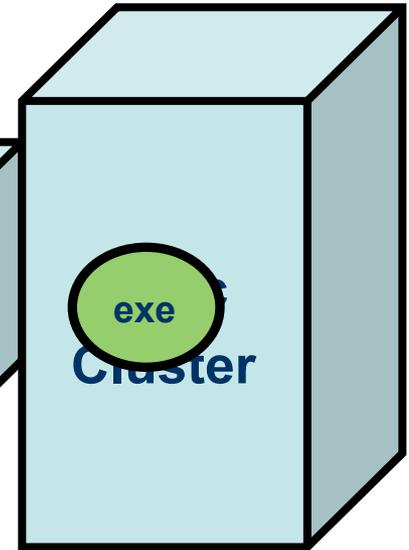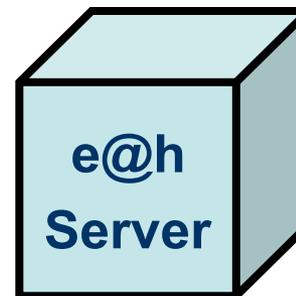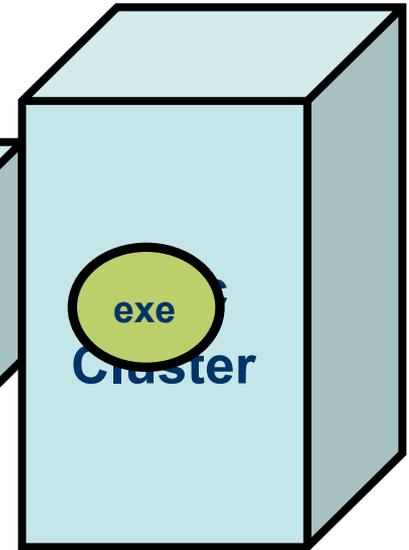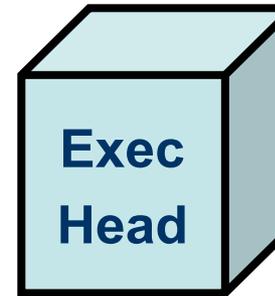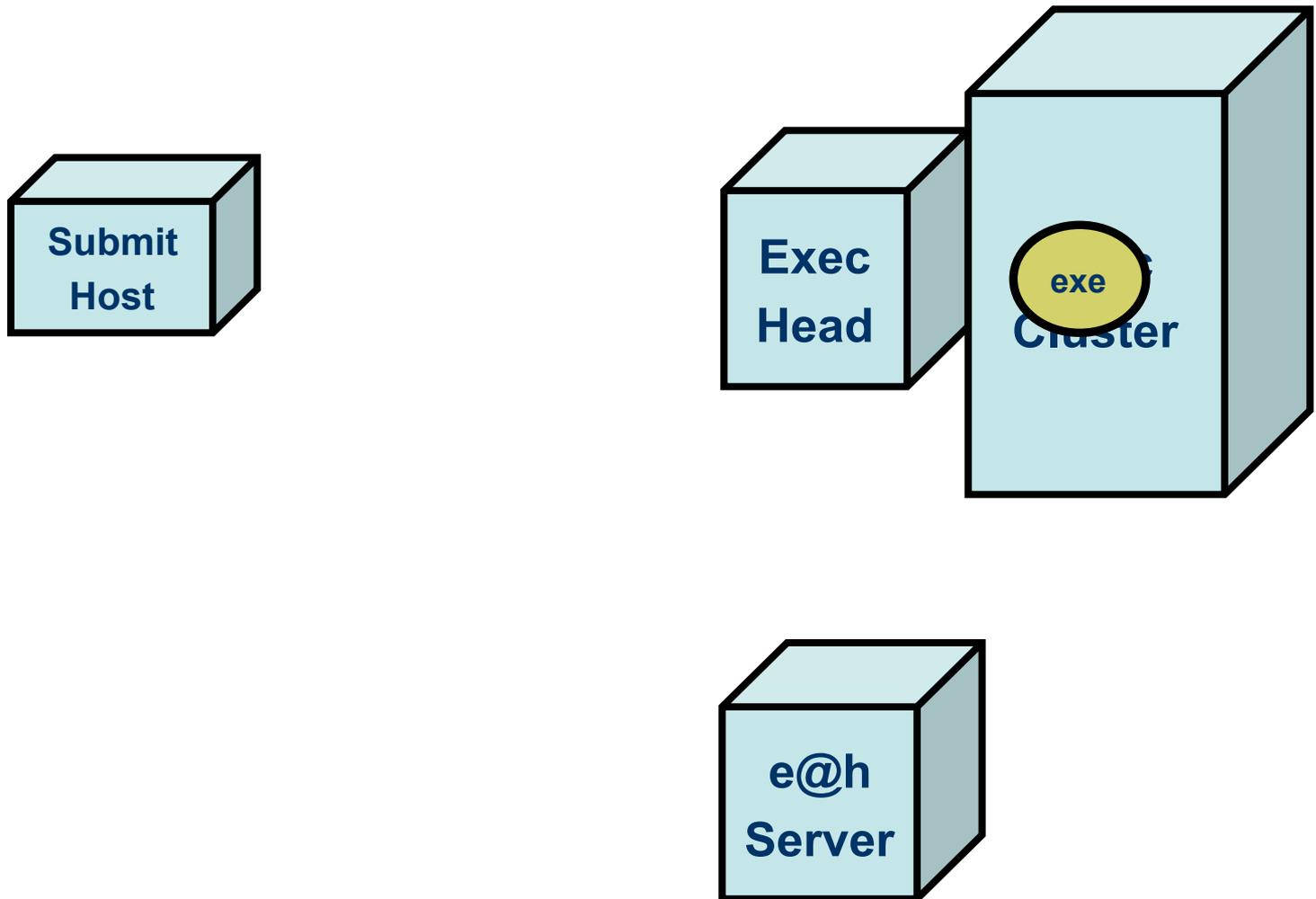
e@h schematic view

Deliver status informations

Submit Host

Exec Head

Cluster

exe

Ask for new Dataset

e@h Server

e@h schematic view

e@h schematic view

e@h schematic view

e@h schematic view

e@h schematic view

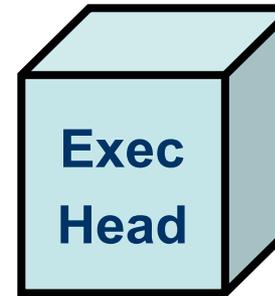Submit
Host

Exec
Head

exe

Cluster

e@h
Server

e@h schematic view

e@h schematic view

Submit Host

Exec Head

cp data

Cluster

Exec Filesystem

e@h Server
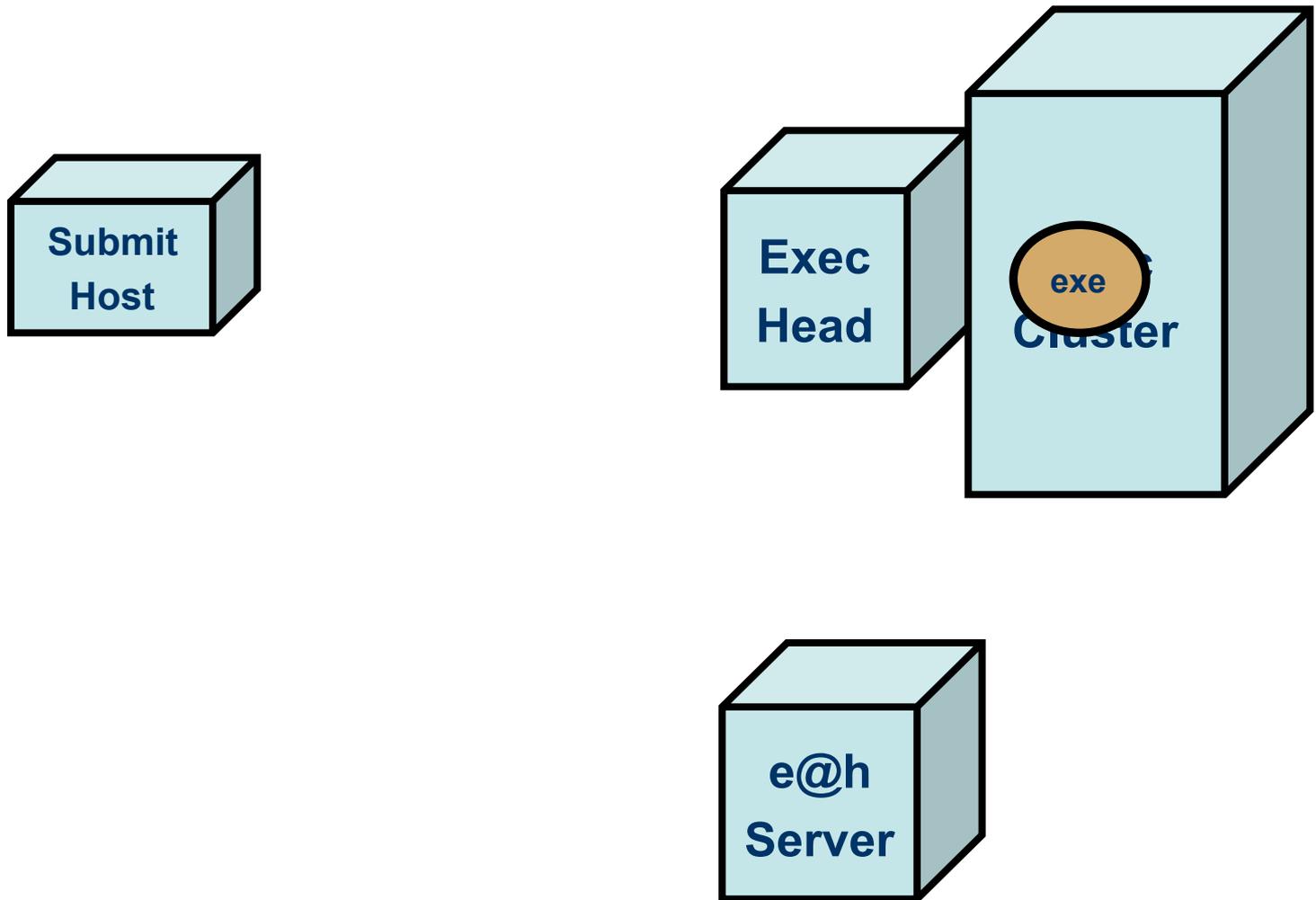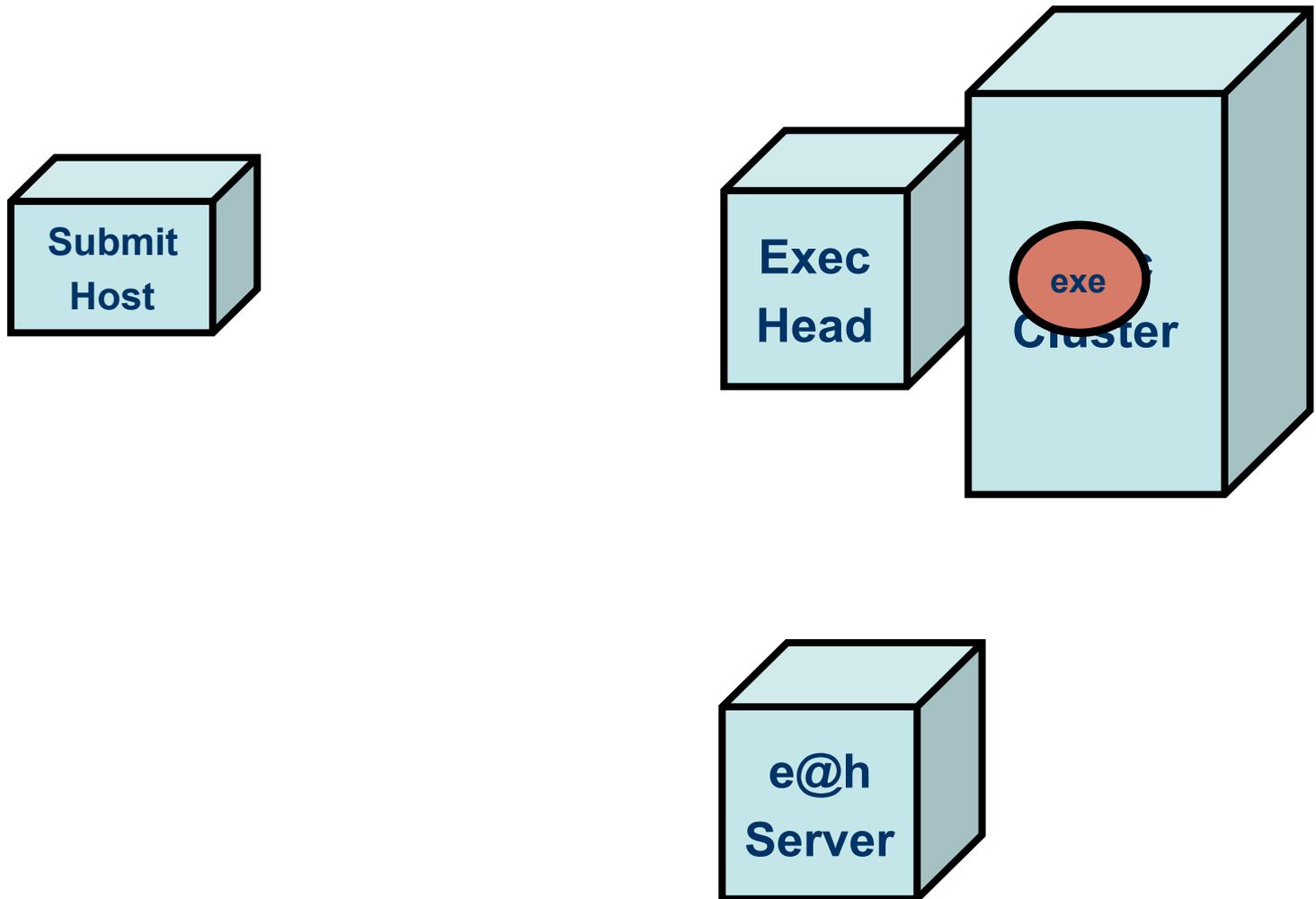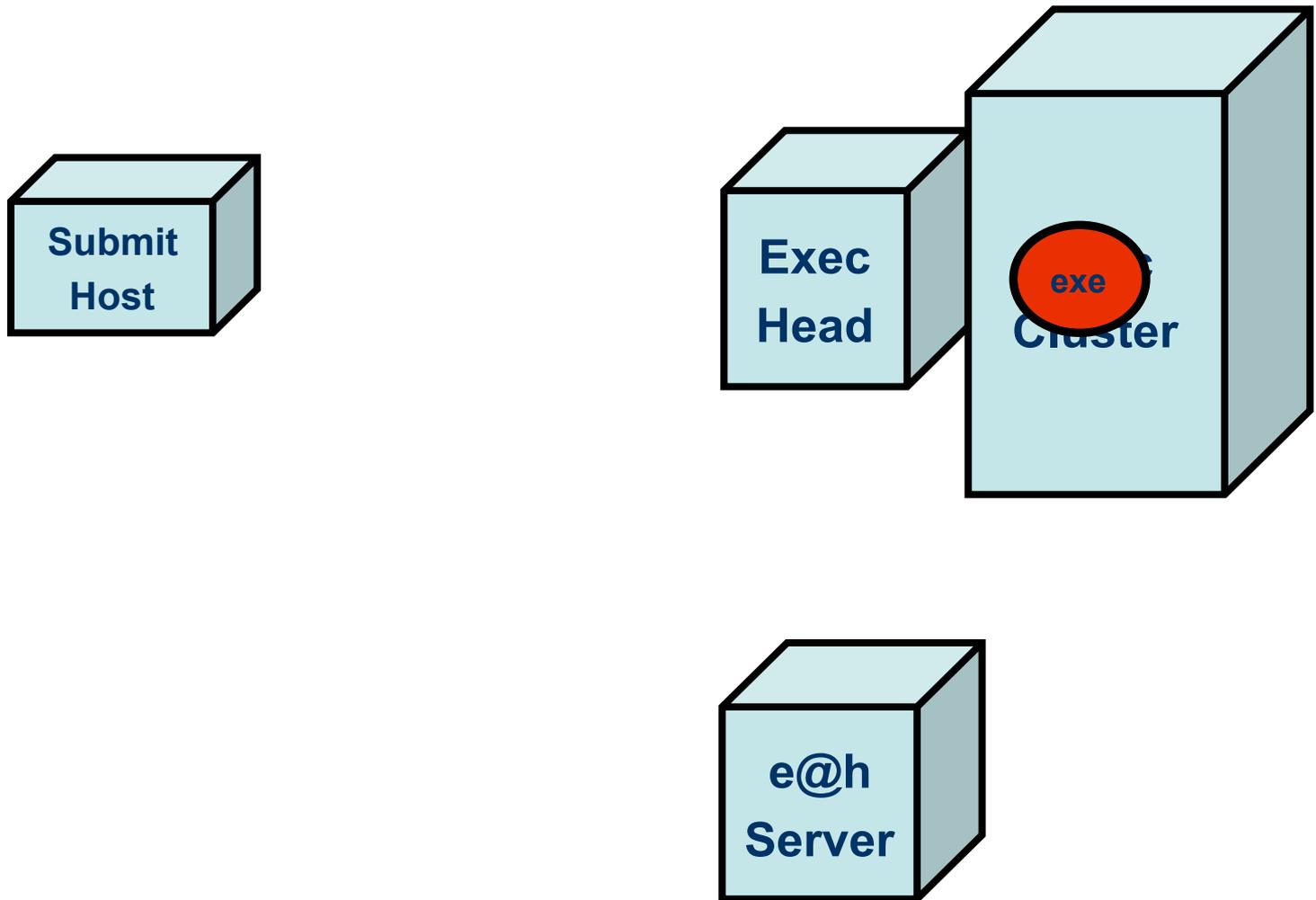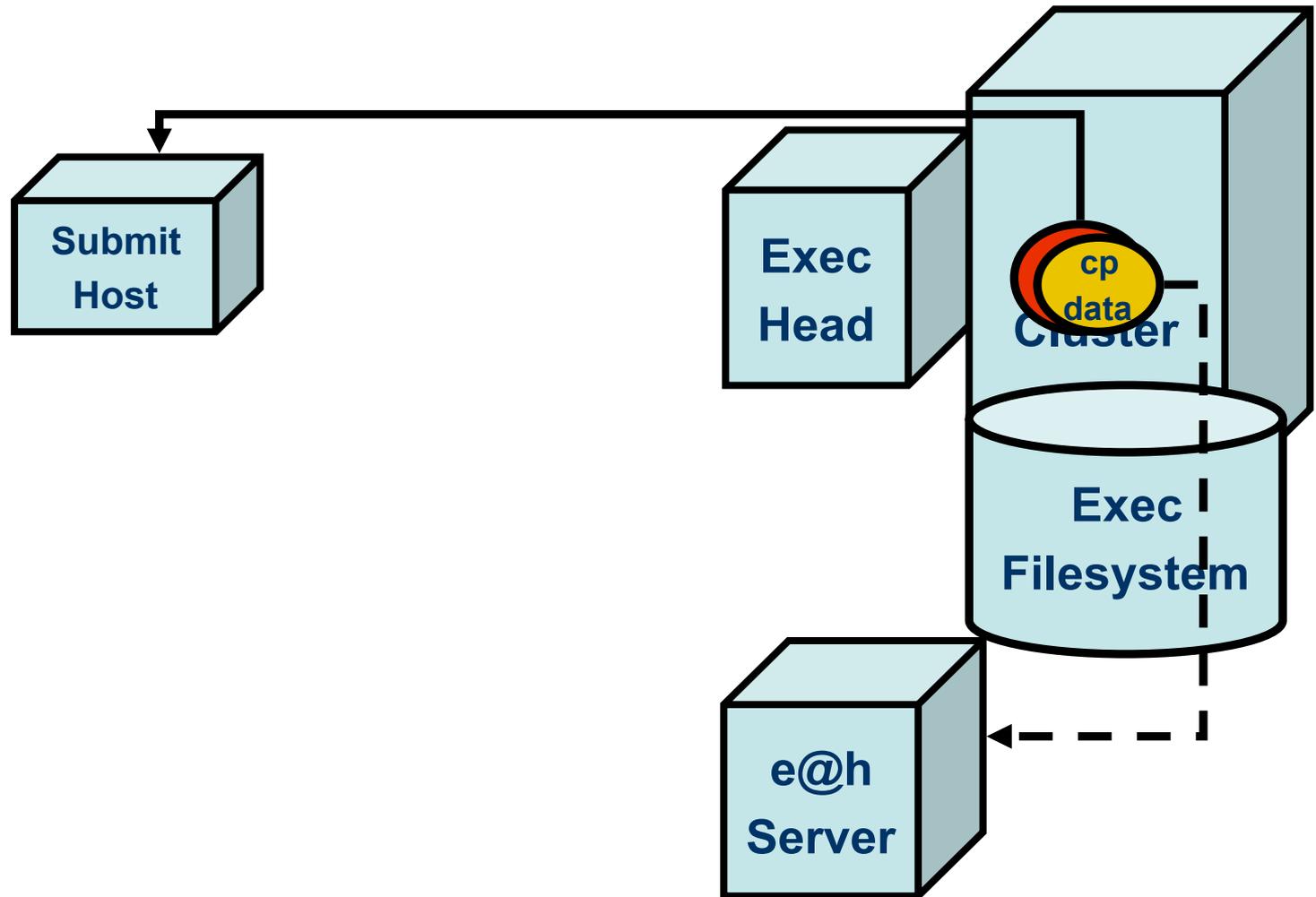
e@h schematic view
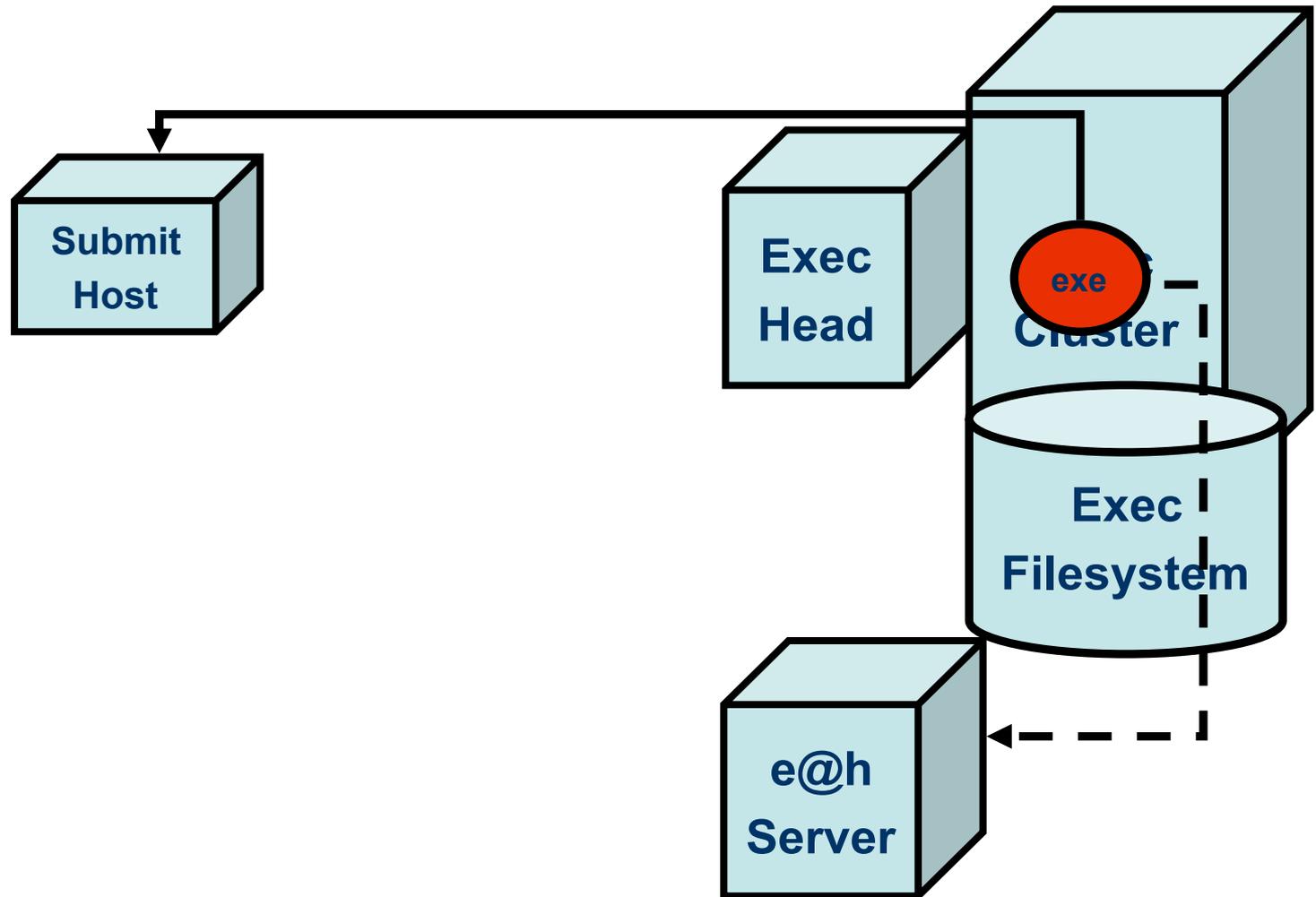
e@h – error handling

- "Inspect" skripts check for fixable errors.

- Ceck the status of all Einstein@Home tasks on a Grid host.

- If an error can be fixed automatically the necessary actions are performed.

e@h – reaching production level

- Continous restart of e@h jobs and Inspect jobs via cron.

- configuration datasets allow to set
  - walltime,
  - amount of Jobs on a resource,
  - Storage location of the checkpoints.

- A statistics portlet enable a fast overview, whether larger problems have occured; Emails containing the output the inspect jobs are sent to the submitter (cron-Feature).

e@h Statistiken

Consumed CPU time GEO600 September 2009

e@h Statistiken
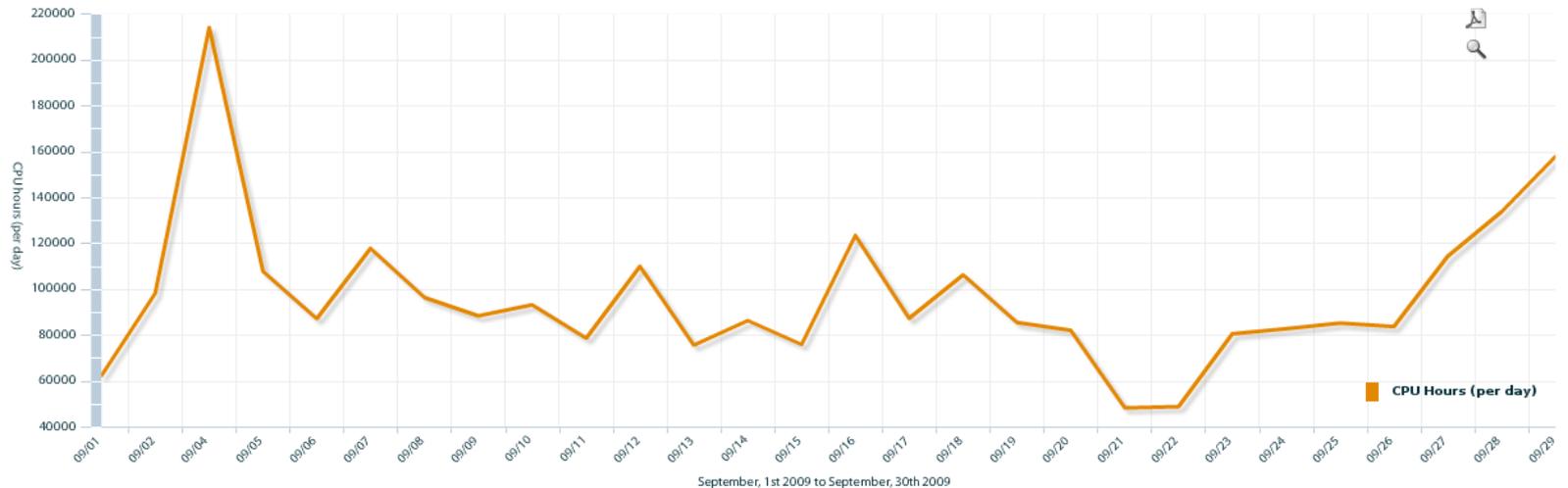
Consumed CPU time GEO600 September 2009

GEO6...
Chart o...

Einstein@Home:
The first of two success stories in D-Grid!

GAT

- A simple and unique API for Grid access
- Connect the different Grid Middlewares via adaptors
- Local adaptors enable the development of the program logic without having access to the Grid
- GAT offers Grid client functionality without any Grid Middleware; no installation of gLite, Globus or Unicore required.
- API groups for
  - File and FileStream operations
  - Logical File management
  - AdvertService management
  - Job management

File Copy Globus RFT

```java
package org.globus.ogsa.gui;

import java.io.BufferedReader;
import java.io.File;
import java.io.FileReader;
import java.net.URL;
import java.util.Date;
import java.util.Vector;
import javax.xml.rpc.Stub;
import org.apache.axis.message.MessageElement;
import org.apache.axis.utils.XMLUtils;
import org.globus.*;
import org.gridforum.ogsi.*;
import org.gridforum.ogsi.holders.TerminationTimeTypeHolder;
import org.w3c.dom.Document;
import org.w3c.dom.Element;

public class RFTClient {
 public static void copy (String source_url, String target_url) {
  try {
   File requestFile     = new File (source_url);
   BufferedReader reader = null;
   try {
    reader = new BufferedReader (new FileReader (requestFile));
   } catch (java.io.FileNotFoundException fnfe) { }
   Vector requestData = new Vector ();
   requestData.add (target_url);
   TransferType[] transfers1    = new TransferType[transferCount];
   RFTOptionsType multirftOptions = new RFTOptionsType ();

   multirftOptions.setBinary          (Boolean.valueOf  (
       (String)requestData.elementAt (0)).booleanValue ());
   multirftOptions.setBlockSize       (Integer.valueOf  (
       (String)requestData.elementAt (1)).intValue      ());
   multirftOptions.setTcpBufferSize   (Integer.valueOf  (
       (String)requestData.elementAt (2)).intValue      ());
   multirftOptions.setNotpt           (Boolean.valueOf  (
       (String)requestData.elementAt (3)).booleanValue ());
   multirftOptions.setParallelStreams (Integer.valueOf  (
       (String)requestData.elementAt (4)).intValue      ());
   multirftOptions.setDcau(Boolean.valueOf(
        (String)requestData.elementAt (5)).booleanValue ());

  int i = 7;
  for (int j = 0; j < transfers1.length; j++)
  {
    transfers1[j] = new TransferType ();

    transfers1[j].setTransferId     (j);
    transfers1[j].setSourceUrl      ((String)requestData.elementAt (i++));
    transfers1[j].setDestinationUrl ((String)requestData.elementAt (i++));
    transfers1[j].setRftOptions     (multirftOptions);
  }
```

```java
TransferRequestType transferRequest = new TransferRequestType ();
transferRequest.setTransferArray (transfers1);

int concurrency = Integer.valueOf
                  ((String)requestData.elementAt(6)).intValue ();

if (concurrency > transfers1.length)
{
  System.out.println ("Concurrency should be less than the number"
                      "of transfers in the request");
  System.exit (0);
}
transferRequest.setConcurrency (concurrency);

TransferRequestElement requestElement = new TransferRequestElement ();
requestElement.setTransferRequest (transferRequest);

ExtensibilityType extension = new ExtensibilityType ();
extension = AnyHelper.getExtensibility (requestElement);

OGSIServiceGridLocator factoryService = new OGSIServiceGridLocator ();
Factory factory = factoryService.getFactoryPort (new URL (source_url));
GridServiceFactory gridFactory = new GridServiceFactory (factory);

LocatorType locator = gridFactory.createService (extension);
System.out.println ("Created an instance of Multi-RFT");

MultiFileRFTDefinitionServiceGridLocator loc
              = new MultiFileRFTDefinitionServiceGridLocator();
RFTPortType rftPort = loc.getMultiFileRFTDefinitionPort (locator);
  ((Stub)rftPort)._setProperty (Constants.AUTHORIZATION,
                                NoAuthorization.getInstance());
  ((Stub)rftPort)._setProperty (GSIConstants.GSI_MODE,
                                GSIConstants.GSI_MODE_FULL_DELEG);
  ((Stub)rftPort)._setProperty (Constants.GSI_SEC_CONV,
                                Constants.SIGNATURE);
  ((Stub)rftPort)._setProperty (Constants.GRIM_POLICY_HANDLER,
                                new IgnoreProxyPolicyHandler ());

  int requestid = rftPort.start ();
  System.out.println ("Request id: " + requestid);

 }
 catch (Exception e)
 {
  System.err.println (MessageUtils.toString (e));
 }
}
```

File Copy GAT

```java
import org.gridlab.gat.GAT;
import org.gridlab.gat.GATContext;
import org.gridlab.gat.URI;
import org.gridlab.gat.io.File;;

public class RemoteCopy {
        public static void main(String[] args) throws Exception {
                GATContext context = new GATContext();
                URI src = new URI("any://sourcehost.grid.org//home/myfile"]);
                URI dest = new URI("any://deshost.grid.org//home/mydestfile");
                File file = GAT.createFile(context, src);    // create file object
                file.copy(dest);                                // and copy it
                GAT.end();
        }
}
```

File Copy GAT

in

GAT is easy!
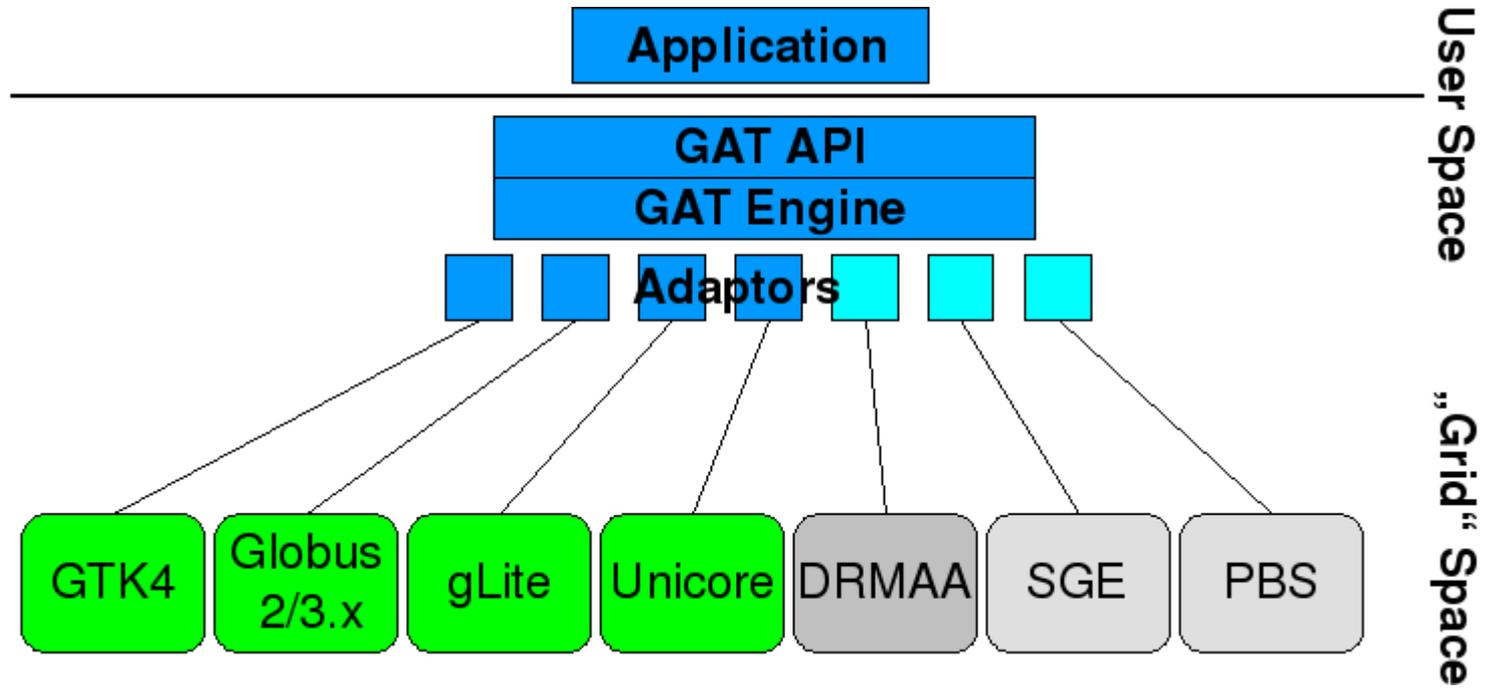
GAT Architecture

Example: ProC MPA Garching
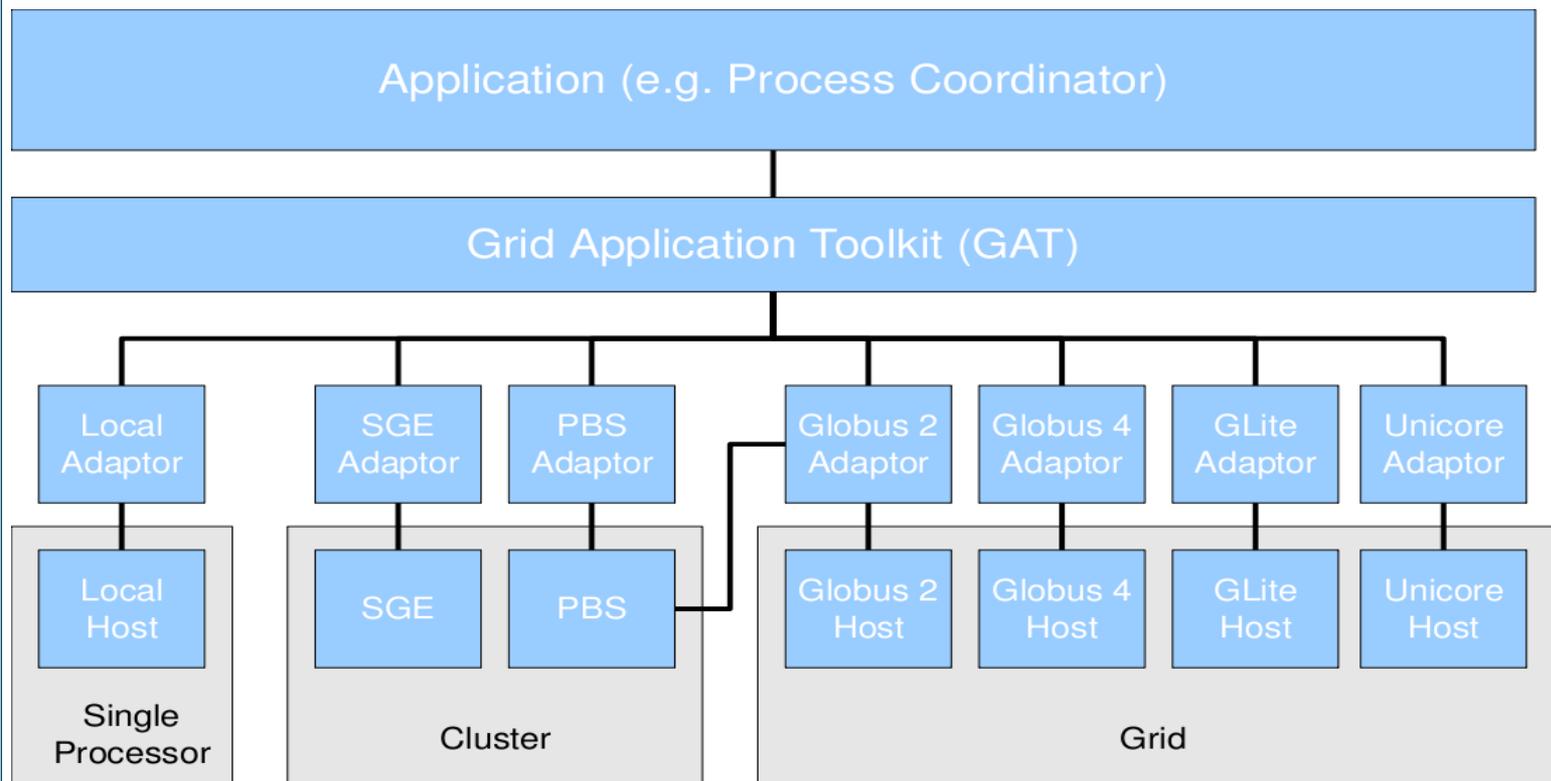
GridSphere

A Portal solution for the Grid

GridSphere – Features (I)

- Portlet API passed Sun TCK and is 100% JSR 168 compliant, JSR 286 compliance under development
- Support for the easy development and integration of new portlet applications
  - Template project generation and build scripts makes life easier for developers
- Higher-level model for building complex portlets using visual beans and the GridSphere User Interface (UI) tag library
- Flexible XML based portal presentation description can be easily modified to create customized portal layouts
- Built-in support for Role Based Access Control (RBAC) separating users into guests, users, admins and super users
- Sophisticated portlet service model that allows for creation and reusability of new business logic with support for persistence of data

GridSphere – Features (II)

- Persistence of data provided using Hibernate OQL for database support
  - Supports many databases including hsqlDB, MySQL, Postgres,
- Prototype Junit/Cactus unit tests for complete server side testing of portlet services including the generation of test reports
- GridSphere core portlets:
  - Login, Logout, Locale settings
  - Profile personalization and Layout customization
  - Administration portlets for creation of users, groups, portlet management and portal layout customization
- Localization support in the Portlet API implementation and portlets support French, English, German, Czech, Polish, Hungarian and Italian, Arabic and Chinese!
- Pluggable authentication modules allows admin to select among supported modules, or develop new ones: Hashed
- Shibboleth (single sign on) available.

What makes GridSphere different?

- Already many other OS portals out there:
  - Jetspeed2, uPortal, StringBeans, Exo, Liferay, Jboss
- Lightweight: no EJB, based on popular, robust libraries e.g. Hibernate for persistence
- Ability to add support for new authentication schemes with pluggable auth modules descriptor
- Visual UI tags and beans makes presentation development much easier
- Support for the Grid!!
  - GridPortlets offered as add-on webapp
  - Provides Library and collection of portlets for Credential management, job execution, data transfer (gridftp)

Gatlet: Gridsphere uses GAT

- Grid portlets are using the GAT-API for acessing the Grid

- Grid access to gLite, Globus and Unicore available!

- Changes in middlewares only require changes in GAT adaptors, but not in the Grid portlets!

Gatlet: BWGrid architecture
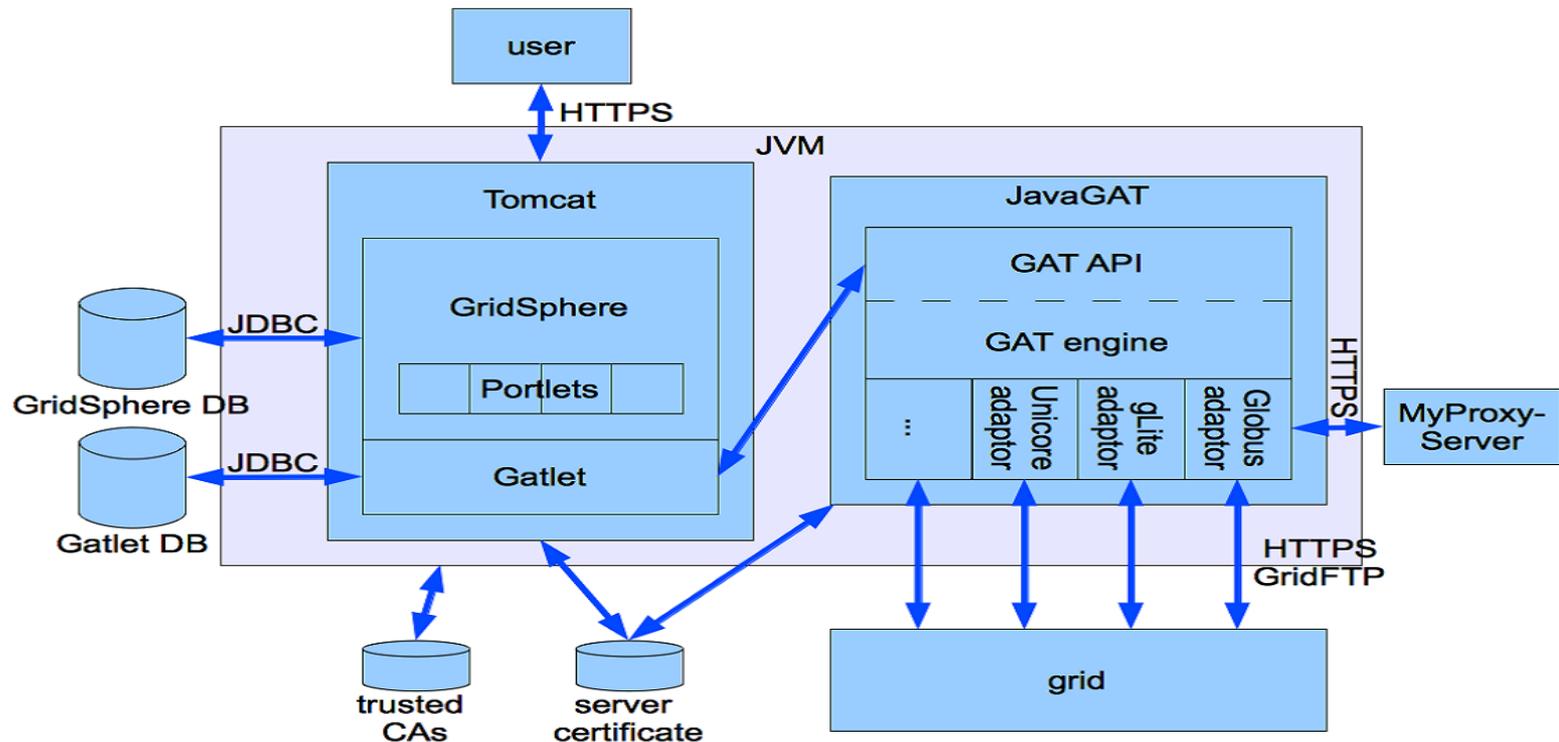
Agenda

- Grid Player AEI
  - Einstein@Home
  - Grid Application Toolkit (GAT)
  - GridSphere
- Grid Player ErasmusMC/Bioquant/DKFZ
  - Project/Job/Applikations Types
  - Security
  - Statistics
- Suggestions for a Research Infrastructure for Genom- and LISA research

Genom & System Biologie Project/Job/
Application Types

- ## Project types:
    - Online visualization
    - Pipelines in a (near-)online frame
    - Backfilling long-term data mining
    - Large-scale parallel analyses and simulation

<= all this across several Grid infrastructures =>

- ## Performance Requirements:
    - Instant highest performance.
    - Highets reliability.
    - Highest trackability, stability and error handling.
    - Highest instantaneous availability and reliability.

Extreme conditions at Real Life conditions !

Genom & System Biology
Security & Statistics

- ## Security aspects:
    - Data transfer always encoded and compressed.
    - Production environment always secure.
    - In case of error always no local data leftovers.
    - Controlling and error managment always needed.
    - Grid-Prozesse, Accounting, Billing always „TOP SECRET" !

- ## Statistics:
    - World wide about 100 users
    - In several Grid >300000 CPUh per *DAY*.
    - Pro Jahr > 20 Publikationen.
    - Scientific & industrial cooperations steadily increasing

+++ … +++ … +++ … +++ Rauch et al., **Differentiation** 76(1), 66-82, 2008 +++ Jhunjhunwala et al., *Cell 133(2), 265-279, 2008* +++ **Knoch** et al., **Stud. Health. Technol. Inform.** 147, 105-116, 2009 +++ **Knoch** et al., **Stud. Health. Technol. Inform.** 147, 269-276, 2009 +++ **Knoch** et al., **Eur. Biophys. J.** 38(6), 757-779, 2009 +++ Estrada et al., **Bioinformatics** 25(20), 2750-2752, 2009 +++ de Wit et al., **Mol. Cell. Biol**. 30(1), 2009 +++ **Kepper** et al., **Stud. Health. Technol. Inform.** 159, 171-180, 2010 +++ **Kepper** et al., **Stud. Health. Technol. Inform.** 159, 264-271, 2010 +++ … +++ … +++

Genom & System Biology
Security & Statistics

- ## Security aspects:
  - Data transfer always encoded and compressed.
  - Production environment always secure.
  - In case of error always no local data leftovers.

Genom & System Biology:
The second of two
sucess stories
in D-Grid!

++ ... *9, 2008* ... **orm.** 147, 2750-2752, 2009 +++ de Wit et al., **Mol. Cell. Biol**. 30(1), 2009 +++ **Kepper** et al., **Stud. Health. Technol. Inform.** 159, 171-180, 2010 +++ **Kepper** et al., **Stud. Health. Technol. Inform.** 159, 264-271, 2010 +++ … +++ … +++

Agenda

- Grid Player AEI
  - Einstein@Home
  - Grid Application Toolkit (GAT)
  - GridSphere
- Grid Player ErasmusMC/Bioquant/DKFZ
  - Project/Job/Applikations Types
  - Security
  - Statistics
- Suggestions for a Research Infrastructure for Genom- and LISA research

Virtual Research Infrastructure

- First round: Access for System Biology, Genom and LISA research groups

- Portal as „Entry Point" for scientists

- Modular approach, to allow later on other research groups to adapt there own virtual research infrastructure

- Integrating CPUs und and storage resources into a Grid (usage of available Grid technologies)

Offer of special services over portlets: more general view (I)

- **Digital Library**
  - A main portlet offers access to well known digital libaries via sub portlets.

- **Conference Room**
  - The main portlet offers several conference access point over separate portlets. We need to check, which web based conference access points are already available.

- **Discussion Portlet**
  - Main portlet offers several "discussion portlets".
  - scientific forums in context to LISA. Sys Bio and Genom, but also access to more general forums.

Offer of special services over portlets: more general view (II)

- **Groupware Portlet**
  - Offer groupware functionalities
  - Appointment calendar
  - Meeting scheduler
- **Working Portlet**
  - A portlet for research works around LISA.
  - Portlets for the generation of workflows,
  - Portlets for viewing, selecting, and adding data sets to data repositories
  - Portlets for providing new results to other colleagues
  - Acces to several grids
  - A main grid portlet shows the available grids (and VOs), sub portlets offer access to these grids
  - Grid jobs on GAT / SAGA (Gatlet), in order to have a middleware independent access to a grid.

Offer of special services over portlets: some suggesestions

- Portlet for the generic submission and monitoring for „Taskfarming-Jobs", and other jobs
  - Access to the Grid via GATLET
  - Selection of resources, executable, amount in queue, amount in executing, storage location checkpoints, etc...
- Portlet for searching and viewing of datasets
- Portlet as virtual information exchange for LISA-, Genom research-, and SysBio informations
- Portlets for generating workflows (not only Grid)
- Portlets for streaming of videos
- Portlet for video conferencing, including a virtual white board (DFN services might be used)
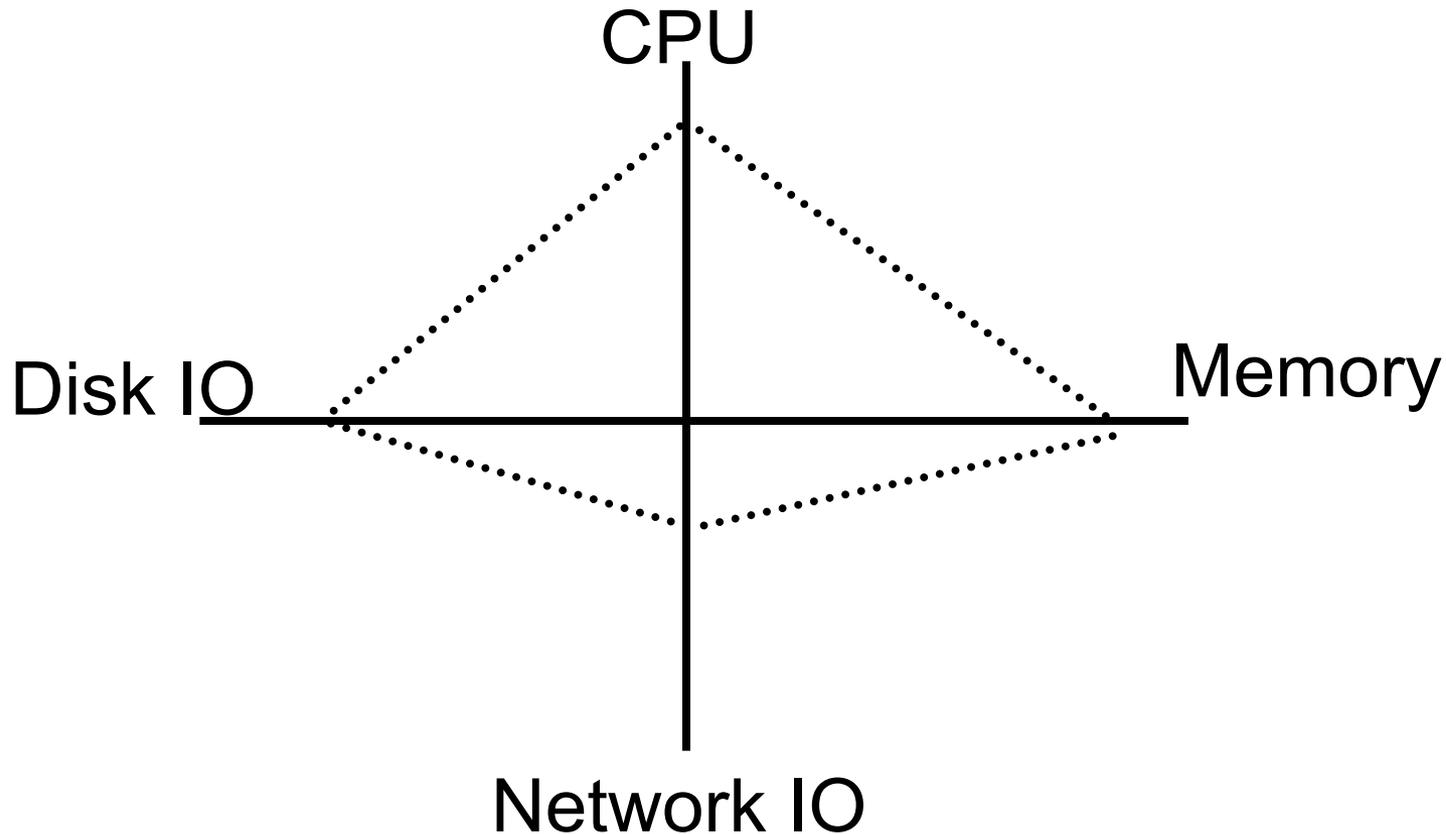
A very special one...

„Nice to have" – some Benchmarks

Benchmarking-Adaptor

- ## User shall get Benchmarking informations for his application
  - Which resource is the best suited?
  - Informations to CPU usage (incl. speed), Memory, Network-IO, Disk-IO

- ## Requires good contact to the Admins of the Grid resources; we have this conact in Germany

Benchmarking graphs

Q & A

A Virtual Research Infrastructure

for

Genome and Lisa Research

in

Europe.

Beck-Ratzka, A. & **Knoch, T. A.**

Laboratoire Astro Particule et Cosmologie (APC), University Diderot-Paris VII, Paris, France, 24th September, 2010.

Abstract

In *VWCISC* we propose a unique working and collaboration environment for international scientific communities. Science is increasingly collaborative at an international level, which enables projects to build teams of specialists best suited to the goals of the research effort, but which also puts heavy burdens on the scientists to communicate with one another. Email and international travel are primary collaboration tools, but increasingly specialist software is being developed to widen the collaboration channel. A number of software collaboration tools have been developed during the last 10 years, within the Grid and Cloud paradigms, that offer scientists easier and more functional ways to collaborate, and in Europe there has been very substantial funding for these efforts.

But most scientific communities outside the field of computer science have not taken up these tools; they still rely on travel and email almost exclusively. A big reason is that there is no simple way to access the many different collaboration tools or to integrate them into the work of a project. Individually very useful, the tools still have high support overheads if they are to be employed in international projects. A simple example is the use of Grid-based tools: there is no universal Grid infrastructure software, and even within Europe tools will have to run on Globus in some nations, on gLite in others, and so on.

The aim of our proposal, therefore, will be to offer EU scientific projects a general environment for international collaboration that solves the most difficult technical issues and leads to a system that looks identical to users in different countries, that can be installed on their local Grid protocols, that requires only a small amount of local support, that offers within this consistent environment a number of basic collaboration tools that many different projects will find useful, and that additionally offers a simple way of incorporating new tools specific to different projects.

The vehicle for this environment will be a web Portal. Such an environment can be made universally accessible and will have a familiar interface. Collaborations that wish to use it will find that their members have a short learning curve and can discover resources intuitively. We will develop a pilot for this Portal that will serve two very different European scientific communities: the SysBio/Genome community and the LISA gravitational-wave research community. By showing how useful the Portal can be for these two projects we will demonstrate that it is a tool that should be useful for many others.

In its final release, the Portal will offer an environment for collaboration that features easy infrastructure integration and infrastructure use. It should be also usable for other disciplines later on. The infrastructure will consist of data, software, and digital library repositories, and of compute resources connected into Grids. All the components of the infrastructure should be easily connectable by portlets within the portal.

Our solution must allow users to connect new infrastructures easily. So it will be a framework, which contains an administration tool that permits connections to new Grids, digital libraries, software and data repositories, etc... However, the main collaboration tools as

- access to digital Libraries,

- access to a conference room, with the ability for video conferencing and a virtual white board,

- access to an information exchange desk. Here we are not only thinking of different scientific discussion forums,

- access to a "Groupware". Which connects at least personal calendars and meeting scheduler,

should be available as templates, where users themselves can supply necessary new configuration data (e.g. the connection data for video conferencing).

We have selected the two research fields of the LISA and Gravitational Wave Research on the one hand, and System Biology and Genom research on the other, because both are real international communities, and importantly because they are currently the most successful grid users within Europe. Concerning the access to remote data, compute resources and repositories, these communities have different needs, but they are interested in a solution which can be used in production mode. Because we already use the Grid in production mode, we are aware of the problems of Grid computing, and we have solved them successfully. From our experience we know what is required, in order to have an easy and successful access to the Grid.

Corresponding author email contact: TA.Knoch@taknoch.org

Keywords:

Genome, genomics, genome organization, genome architecture, structural sequencing, architectural sequencing, systems genomics, coevolution, holistic genetics, genome mechanics, genome statistical mechanics, genomic uncertainty principle, genome function, genetics, gene regulation, replication, transcription, repair, homologous recombination, simultaneous co-transfection, cell division, mitosis, metaphase, interphase, cell nucleus, nuclear structure, nuclear organization, chromatin density distribution, nuclear morphology, chromosome territories, subchromosomal domains, chromatin loop aggregates, chromatin rosettes, chromatin loops, chromatin fibre, chromatin density, persistence length, spatial distance measurement, histones, H1.0, H2A, H2B, H3, H4, mH2A1.2, DNA sequence, complete sequenced genomes, molecular transport, obstructed diffusion, anomalous diffusion, percolation, long-range correlations, fractal analysis, scaling analysis, exact yard-stick dimension, box-counting dimension, lacunarity dimension, local nuclear dimension, nuclear diffuseness, parallel super computing, grid computing, volunteer computing, Brownian Dynamics, Monte Carlo, fluorescence in situ hybridization, chromatin cross-linking, chromosome conformation capture (3C), selective high-resolution high-throughput chromosome interaction capture (T2C), confocal laser scanning microscopy, fluorescence correlation spectroscopy, super resolution microscopy, spatial precision distance microscopy, auto-fluorescent proteins, CFP, GFP, YFP, DsRed, fusion protein, in vivo labelling, information browser, visual data base access, holistic viewing system, integrative data management, extreme visualization, three-dimensional virtual environment, virtual paper tool.

Literature References

Knoch, T. A. Dreidimensionale Organisation von Chromosomen-Domänen in Simulation und Experiment. (Three-dimensional organization of chromosome domains in simulation and experiment.) *Diploma Thesis*, Faculty for Physics and Astronomy, Ruperto-Carola University, Heidelberg, Germany, 1998, and TAK Press, Tobias A. Knoch, Mannheim, Germany, ISBN 3-00-010685-5 and ISBN 978-3-00-010685-9 (soft cover, 2rd ed.), ISBN 3-00-035857-9 and ISBN 978-3-00-0358857-0 (hard cover, 2rd ed.), ISBN 3-00-035858-7, and ISBN 978-3-00-035858-6 (DVD, 2rd ed.), 1998.

Knoch, T. A., Münkel, C. & Langowski, J. Three-dimensional organization of chromosome territories and the human cell nucleus - about the structure of a self replicating nano fabrication site. *Foresight Institute - Article Archive*, Foresight Institute, Palo Alto, *CA, USA,* http://www. foresight.org, 1- 6, 1998.

Knoch, T. A., Münkel, C. & Langowski, J. Three-Dimensional Organization of Chromosome Territories and the Human Interphase Nucleus. *High Performance Scientific Supercomputing*, editor Wilfried Juling, Scientific Supercomputing Center (SSC) Karlsruhe, University of Karlsruhe (TH), 27- 29, 1999.

Knoch, T. A., Münkel, C. & Langowski, J. Three-dimensional organization of chromosome territories in the human interphase nucleus. *High Performance Computing in Science and Engineering 1999*, editors Krause, E. & Jäger, W., High-Performance Computing Center (HLRS) Stuttgart, University of Stuttgart, Springer Berlin-Heidelberg-New York, ISBN 3-540-66504-8, 229-238, 2000.

Bestvater, F., **Knoch, T. A.**, Langowski, J. & Spiess, E. GFP-Walking: Artificial construct conversions caused by simultaneous cotransfection. *BioTechniques* 32(4), 844-854, 2002.

Knoch, T. A. (editor), Backes, M., Baumgärtner, V., Eysel, G., Fehrenbach, H., Göker, M., Hampl, J., Hampl, U., Hartmann, D., Hitzelberger, H., Nambena, J., Rehberg, U., Schmidt, S., Weber, A., & Weidemann, T. Humanökologische Perspectiven Wechsel - Festschrift zu Ehren des 70. Geburtstags von Prof. Dr. Kurt Egger. Human Ecology Working Group, Ruperto-Carola University of Heidelberg, Heidelberg, Germany, 2002.

Knoch, T. A. Approaching the three-dimensional organization of the human genome: structural-, scaling- and dynamic properties in the simulation of interphase chromosomes and cell nuclei, long- range correlations in complete genomes, *in vivo* quantification of the chromatin distribution, construct conversions in simultaneous co-transfections. *Dissertation*, Ruperto-Carola University, Heidelberg, Germany, and TAK†Press, Tobias A. Knoch, Mannheim, Germany, ISBN 3-00-009959-X and ISBN 978-3-00-009959-5 (soft cover, 3rd ed.), ISBN 3-00-009960-3 and ISBN 978-3-00-009960-1 (hard cover, 3rd ed.), ISBN 3-00-035856-9 and ISBN 978-3-00-010685-9 (DVD, 3rd ed.) 2002.

Knoch, T. A. Towards a holistic understanding of the human genome by determination and integration of its sequential and three-dimensional organization. *High Performance Computing in Science and Engineering 2003*, editors Krause, E., Jäger, W. & Resch, M., High-Performance Computing Center (HLRS) Stuttgart, University of Stuttgart, Springer Berlin-Heidelberg-New York, ISBN 3- 540-40850-9, 421-440, 2003.

Wachsmuth, M., Weidemann, T., Müller, G., Urs W. Hoffmann-Rohrer, **Knoch, T. A.**, Waldeck, W. & Langowski, J. Analyzing intracellular binding and diffusion with continuous fluorescence photobleaching. *Biophys. J.* 84(5), 3353-3363, 2003.

Weidemann, T., Wachsmuth, M., **Knoch, T. A.**, Müller, G., Waldeck, W. & Langowski, J. Counting nucleosomes in living cells with a combination of fluorescence correlation spectroscopy and confocal imaging. *J. Mol. Biol.* 334(2), 229-240, 2003.

Fejes Tóth, K., **Knoch, T. A.**, Wachsmuth, M., Frank-Stöhr, M., Stöhr, M., Bacher, C. P., Müller, G. & Rippe, K. Trichostatin A induced histone acetylation causes decondensation of interphase chromatin. *J. Cell Science* 177, 4277-4287, 2004.

Ermler, S., Krunic, D., **Knoch, T. A.**, Moshir, S., Mai, S., Greulich-Bode, K. M. & Boukamp, P. Cell cycle-dependent 3D distribution of telomeres and telomere repeat-binding factor 2 (TRF2) in HaCaT and HaCaT-myc cells. *Europ. J. Cell Biol.* 83(11-12), 681-690, 2004.

Kost, C., Gama de Oliveira, E., **Knoch, T. A.** & Wirth, R. Spatio-temporal permanence and plasticity of foraging trails in young and mature leaf-cutting ant colonies (*Atta spp.*). *J. Trop. Ecol.* 21(6), 677- 688, 2005.

Winnefeld, M., Grewenig, A., Schnölzer, M., Spring, H., **Knoch, T. A.**, Gan, E. C., Rommelaere, J. & Cziepluch, C. Human SGT interacts with BAG-6/Bat-3/Scythe and cells with reduced levels of either protein display persistence of few misaligned chromosomes and mitotic arrest. *Exp. Cell Res.* 312, 2500-2514, 2006.

Sax, U., Weisbecker, A., Falkner, J., Viezens, F., Yassene, M., Hartung, M., Bart, J., Krefting, D., **Knoch, T. A.** & Semler, S. Grid-basierte Services für die elektronische Patientenakte der Zukunft. *E- HEALTH-COM - Magazin für Gesundheitstelematik und Telemedizin*, 4(2), 61-63, 2007.

de Zeeuw, L. V., **Knoch, T. A.,** van den Berg, J. & Grosveld, F. G. Erasmus Computing Grid - Het bouwen van een 20 TeraFLOP virtuele supercomputer. *NIOC proceedings 2007 - het perspective of lange termijn.* editor Frederik, H. NIOC, Amsterdam, The Netherlands, 52-59, 2007.

Rauch, J., **Knoch, T. A.**, Solovei, I., Teller, K. Stein, S., Buiting, K., Horsthemke, B., Langowski, J., Cremer, T., Hausmann, M. & Cremer, C. Lightoptical precision measurements of the Prader- Willi/Angelman Syndrome imprinting locus in human cell nuclei indicate maximum condensation changes in the few hundred nanometer range. *Differentiation* 76(1), 66-82, 2008.

Sax, U., Weisbecker, A., Falkner, J., Viezens, F., Mohammed, Y., Hartung, M., Bart, J., Krefting, D., **Knoch, T. A.** & Semler, S. C. Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die EPA der Zukunft. *Telemedizinführer Deutschland 2008*, editor Jäckel, A. Deutsches Medizinforum, Minerva KG, Darmstadt, ISBN 3-937948-06-6, ISBN-13 9783937948065, 47-51, 2008.

Drägestein, K. A., van Capellen, W. A., van Haren, J. Tsibidis, G. D., Akhmanova, A., **Knoch, T. A.**, Grosveld, F. G. & Galjart, N. Dynamic behavior of GFP-CLIP-170 reveals fast protein turnover on microtubule plus ends. *J. Cell Biol. 180(4)*, 729-737, 2008.

Jhunjhunwala, S., van Zelm, M. C., Peak, M. M., Cutchin, S., Riblet, R., van Dongen, J. J. M., Grosveld, F. G., **Knoch, T. A.**[+] & Murre, C.[+] The 3D-structure of the Immunoglobulin Heavy Chain Locus: implications for long-range genomic interactions. *Cell 133(2)*, 265-279, 2008.

Krefting, D., Bart, J., Beronov, K., Dzhimova, O., Falkner, J., Hartung, M., Hoheisel, A., **Knoch, T. A.**, Lingner, T., Mohammed, Y., Peter, K., Rahm, E., Sax, U., Sommerfeld, D., Steinke, T., Tolxdorff, T., Vossberg, M., Viezens, F. & Weisbecker, A. MediGRID - Towards a user friendly secured grid infrastructure. *Future Generation Computer Systems 25(3)*, 326-336, 2008.

Knoch, T. A., Lesnussa, M., Kepper, F. N., Eussen, H. B., & Grosveld, F. G. The GLOBE 3D Genome Platform - Towards a novel system-biological paper tool to integrate the huge complexity of genome organization and function. *Stud. Health. Technol. Inform. 147*, 105-116, 2009.

Knoch, T. A., Baumgärtner, V., de Zeeuw, L. V., Grosveld, F. G., & Egger, K. e-Human Grid Ecology: Understanding and approaching the Inverse Tragedy of the Commons in the e-Grid Society. *Stud. Health. Technol. Inform. 147*, 269-276, 2009.

Dickmann, F., Kaspar, M., Löhnardt, B., **Knoch, T. A.,** & Sax, U. Perspectives of MediGRID. *Stud. Health. Technol. Inform. 147*, 173-182, 2009.

Knoch, T. A., Göcker, M., Lohner, R., Abuseiris, A. & Grosveld, F. G. Fine-structured multi-scaling long-range correlations in completely sequenced genomes - features, origin and classification. *Eur. Biophys. J. 38(6)*, 757-779, 2009.

Dickmann, F., Kaspar, M., Löhnhardt, B., Kepper, N., Viezens, F., Hertel, F., Lesnussa, M., Mohammed, Y., Thiel, A., Steinke, T., Bernarding, J., Krefting, D., **Knoch, T. A.** & Sax, U. Visualization in health-grid environments: a novel service and business approach. *LNCS 5745*, 150-159, 2009.

Dickmann, F., Kaspar, M., Löhnhardt, B., Kepper, N., Viezens, F., Hertel, F., Lesnussa, M., Mohammed, Y., Thiel, A., Steinke, T., Bernarding, J., Krefting, D., **Knoch, T. A.** & Sax, U. Visualization in health-grid environments: a novel service and business approach. *Grid economics and business models - GECON 2009 Proceedings, 6th international workshop, Delft, The Netherlands.* editors Altmann, J., Buyya, R. & Rana, O. F., GECON 2009, LNCS 5745, Springer-Verlag Berlin Heidelberg, ISBN 978-3-642-03863-1, 150-159, 2009.

Estrada, K.[*], Abuseiris, A.[*], Grosveld, F. G., Uitterlinden, A. G., **Knoch, T. A.**[+] & Rivadeneira, F.[+] GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics 25(20), 2750-2752*, 2009.

Kepper, N., Schmitt, E., Lesnussa, M., Weiland, Y., Eussen, H. B., Grosveld, F. G., Hausmann, M. & **Knoch T. A.,** Visualization, Analysis, and Design of COMBO-FISH Probes in the Grid-Based GLOBE 3D Genome Platform. *Stud. Health Technol. Inform. 159*, 171-180, 2010.

Kepper, N., Ettig, R., Dickmann, F., Stehr, R., Grosveld, F. G., Wedemann, G. & **Knoch, T. A.** Parallel high-performance grid computing: capabilities and opportunities of a novel demanding service and business class allowing highest resource efficiency. *Stud. Health Technol. Inform. 159*, 264-271, 2010.

Skrowny, D., Dickmann, F., Löhnhardt, B., **Knoch, T. A.** & Sax, U. Development of an information platform for new grid users in the biomedical field. *Stud. Health Technol. Inform. 159*, 277-282, 2010.

Knoch, T. A., Baumgärtner, V., Grosveld, F. G. & Egger, K. Approaching the internalization challenge of grid technologies into e-Society by e-Human "Grid" Ecology. *Economics of Grids, Clouds, Systems, and Services – GECON 2010 Proceedings*, 7th International Workshop, Ischia, Italy, editors Altman, J., & Rana, O. F., Lecture Notes in Computer Science (LNCS) 6296, Springer Berlin Heidelberg New York, ISSN 0302-9743, ISBN-10 3-642-15680-0, ISBN-13 978-3-642-15680-9, 116-128, 2010.

Dickmann, F., Brodhun, M., Falkner, J., **Knoch, T. A.** & Sax, U. Technology transfer of dynamic IT outsourcing requires security measures in SLAs. *Economics of Grids, Clouds, Systems, and Services – GECON 2010 Proceedings*, 7th International Workshop, Ischia, Italy, editors Altman, J., & Rana, O. F., Lecture Notes in Computer Science (LNCS) 6296, Springer Berlin Heidelberg New York, ISSN 0302-9743, ISBN-10 3-642-15680-0, ISBN-13 978-3-642-15680-9, 1-115, 2010.

Knoch, T. A. Sustained Renewability: approached by systems theory and human ecology. *Renewable Energy 2*, editors M. Nayeripour & M. Keshti, Intech, ISBN 978-953-307-573-0, 21-48, 2011.

Kolovos, P., **Knoch, T. A.**, F. G. Grosveld, P. R. Cook, & Papantonis, A. Enhancers and silencers: an integrated and simple model for their function. *Epigenetics and Chromatin 5(1)*, 1-8, 2012.

Dickmann, F., Falkner, J., Gunia, W., Hampe, J., Hausmann, M., Herrmann, A., Kepper, N., **Knoch, T. A.**, Lauterbach, S., Lippert, J., Peter, K., Schmitt, E., Schwardmann, U., Solodenko, J., Sommerfeld, D., Steinke, T., Weisbecker, A. & Sax, U. Solutions for Biomedical Grid Computing - Case Studies from the D-Grid Project Services@MediGRID. *JOCS 3(5)*, 280-297, 2012.

Estrada, K.[*], Abuseiris, A.[*], Grosveld, F. G., Uitterlinden, A. G., **Knoch, T. A.**[+] & Rivadeneira, F.[+] GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Dissection of the complex genetic architecture of human stature and osteoporosis.* cummulative dissertation, editor Estrada K., Erasmus Medical Center, Erasmus University Rotterdam, Rotterdam, The Netherlands, ISBN 978-94-6169-246-7, 25-30, 1st June 2012.

van de Corput, M. P. C., de Boer, E., **Knoch, T. A.**, van Cappellen, W. A., Quintanilla, A., Ferrand, L., & Grosveld, F. G. Super-resolution imaging reveals 3D folding dynamics of the b-globin locus upon gene activation. *J. Cell Sci. 125*, 4630-4639, 2012.