High performance scientific computations and services on BalticGrid and LitGrid

Dr. Algimantas Juozapavičius

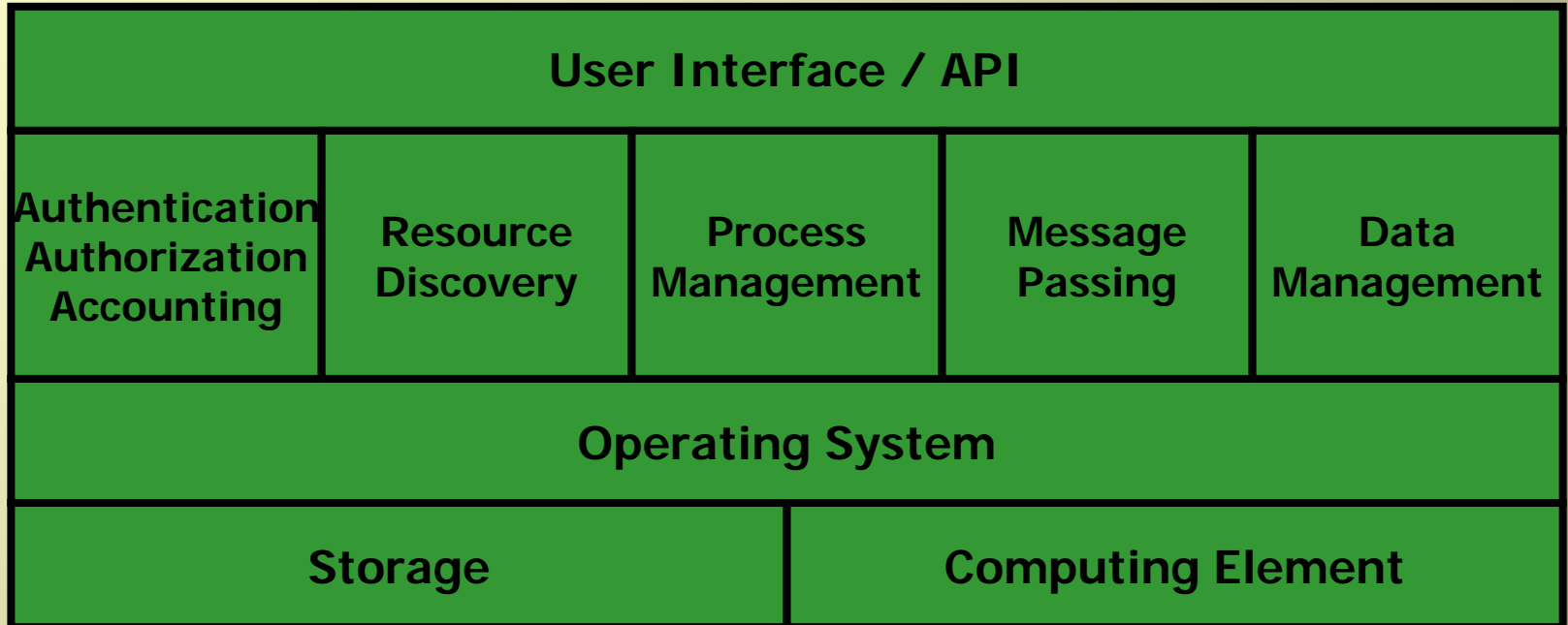Vilnius University

algimantas.juozapavicius@maf.vu.lt

Content

- **real life of grids**
- **EGEE and other grids**
- **BalticGrid infrastructure**
- **LitGrid infrastructure**
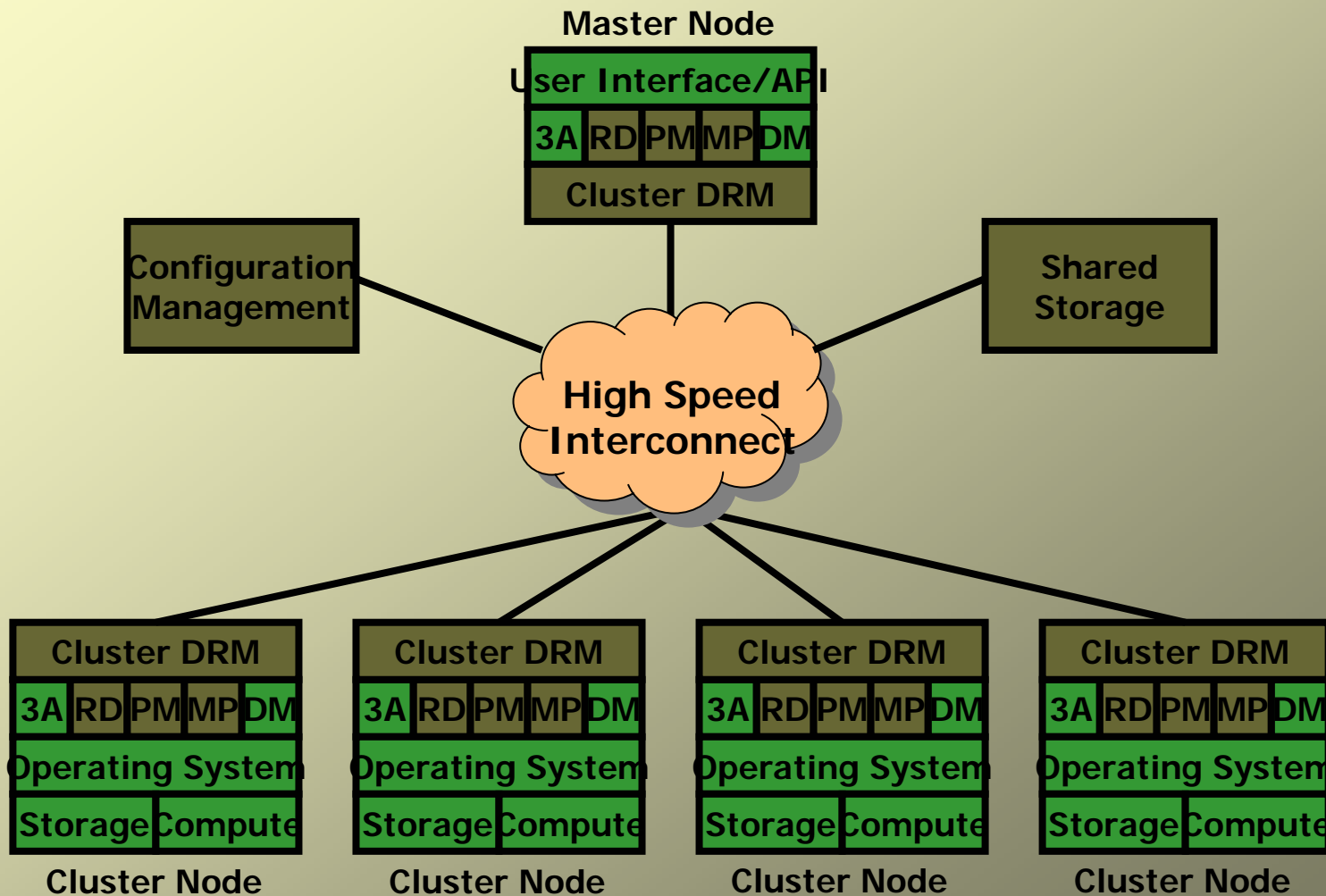- **SIG services**
- **financial application**

The single system model:
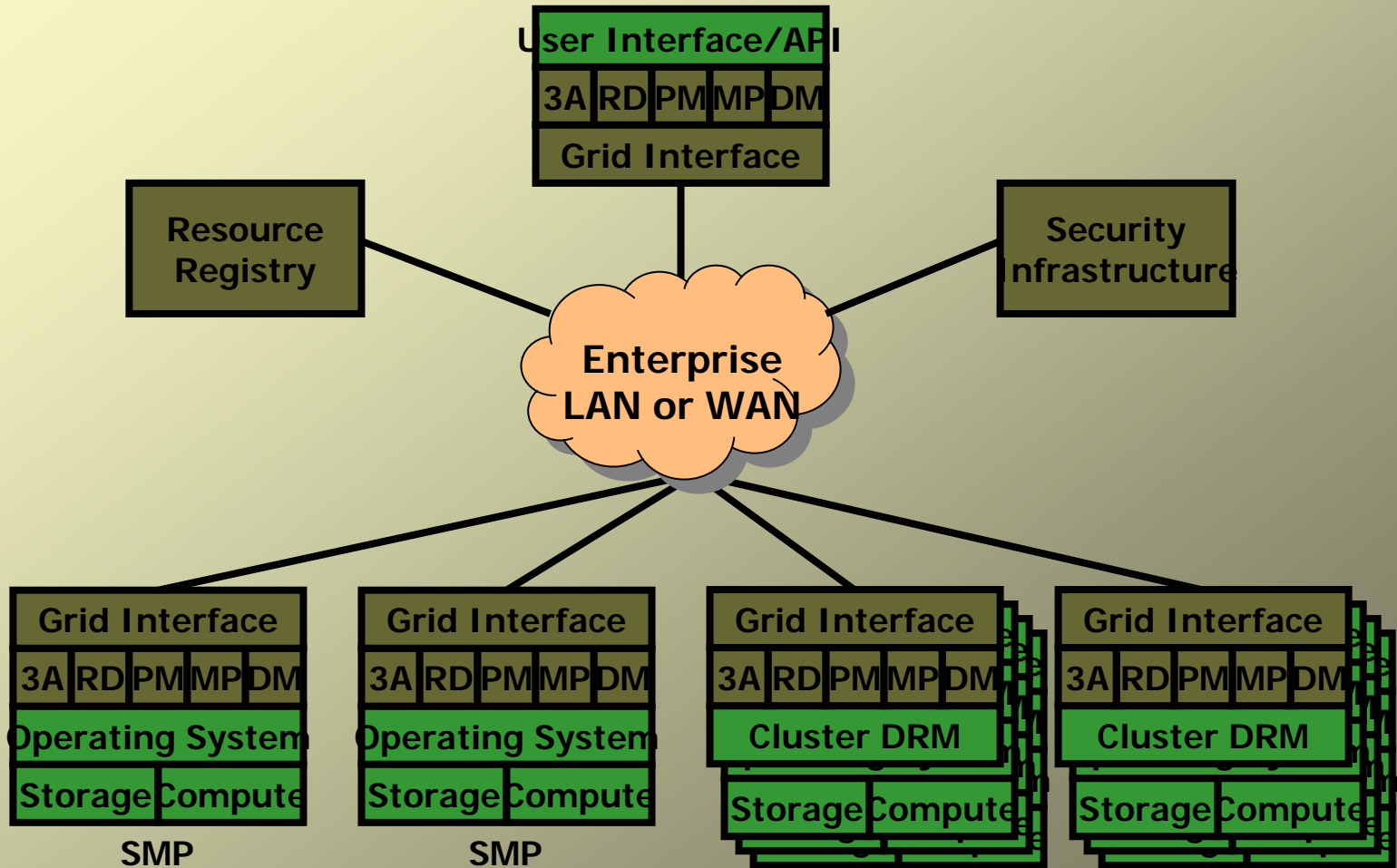
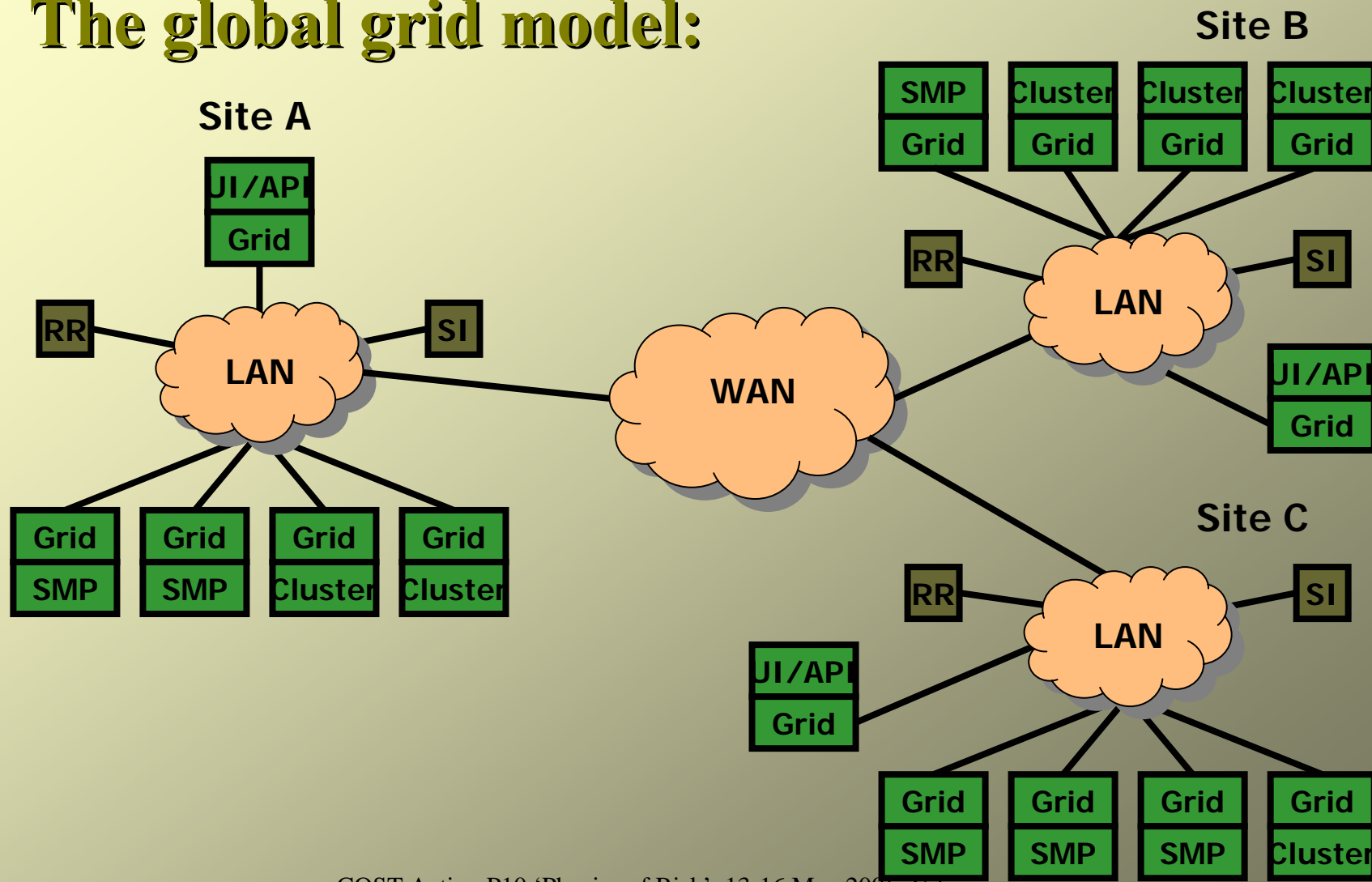| User Interface / API				
Authentication Authorization Accounting	Resource Discovery	Process Management	Message Passing	Data Management
Operating System				
Storage		Computing Element		

The cluster model:

Master Node

User Interface/API

| 3A | RD | PM | MP | DM |

Cluster DRM

Configuration Management

Shared Storage

High Speed Interconnect

Cluster DRM

| 3A | RD | PM | MP | DM |

Operating System

| Storage | Compute |

Cluster Node

Cluster DRM

| 3A | RD | PM | MP | DM |

Operating System

| Storage | Compute |

Cluster Node

Cluster DRM

| 3A | RD | PM | MP | DM |

Operating System

| Storage | Compute |

Cluster Node

Cluster DRM

| 3A | RD | PM | MP | DM |

Operating System

| Storage | Compute |

Cluster Node

COST Action P10 'Physics of Risk', 13-16 May 2006, Vilnius

The enterprise grid model:

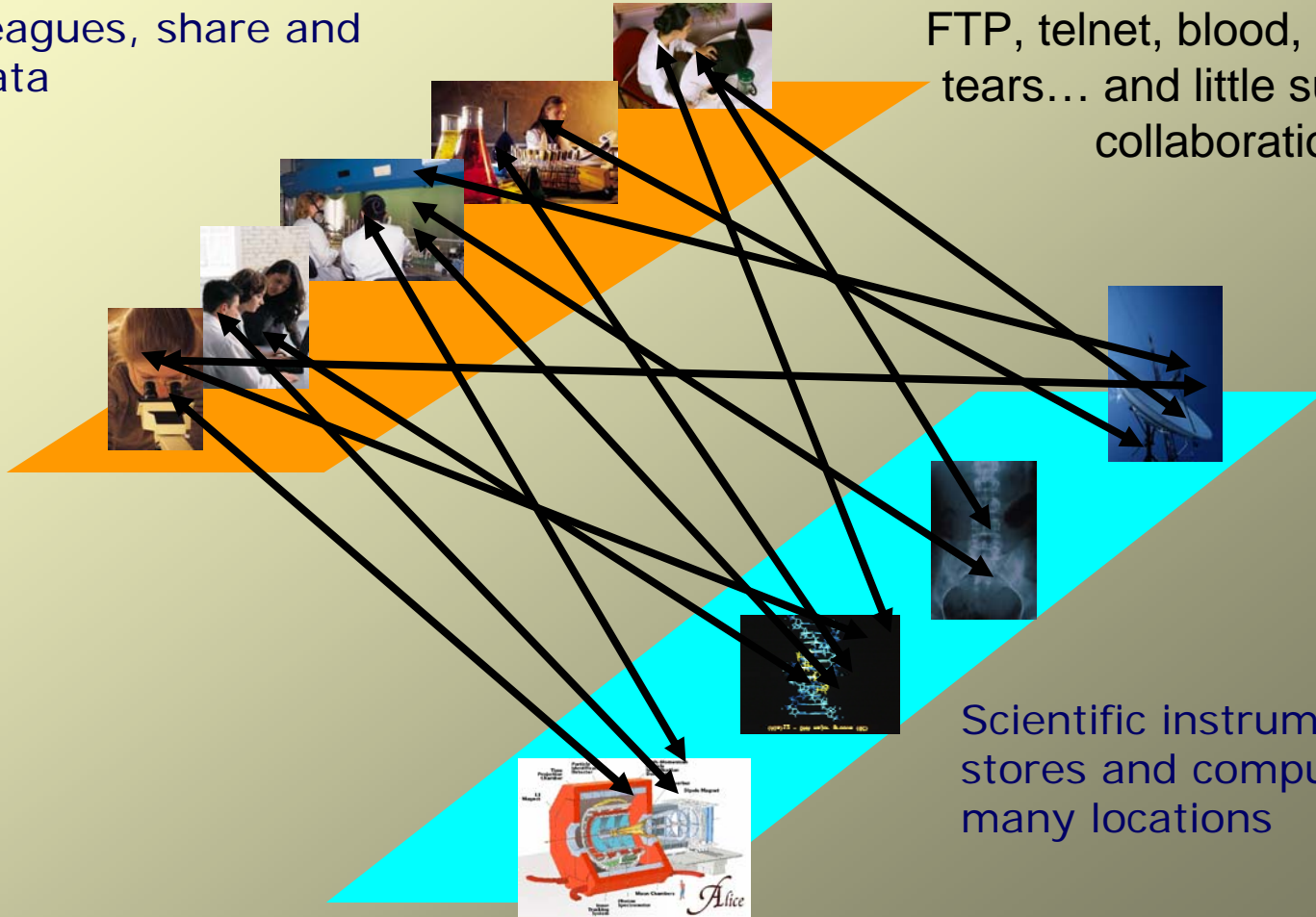COST Action P10 'Physics of Risk', 13-16 May 2006, Vilnius

The global grid model:

Before Grids:

Researchers in many locations need to interact with colleagues, share and access data

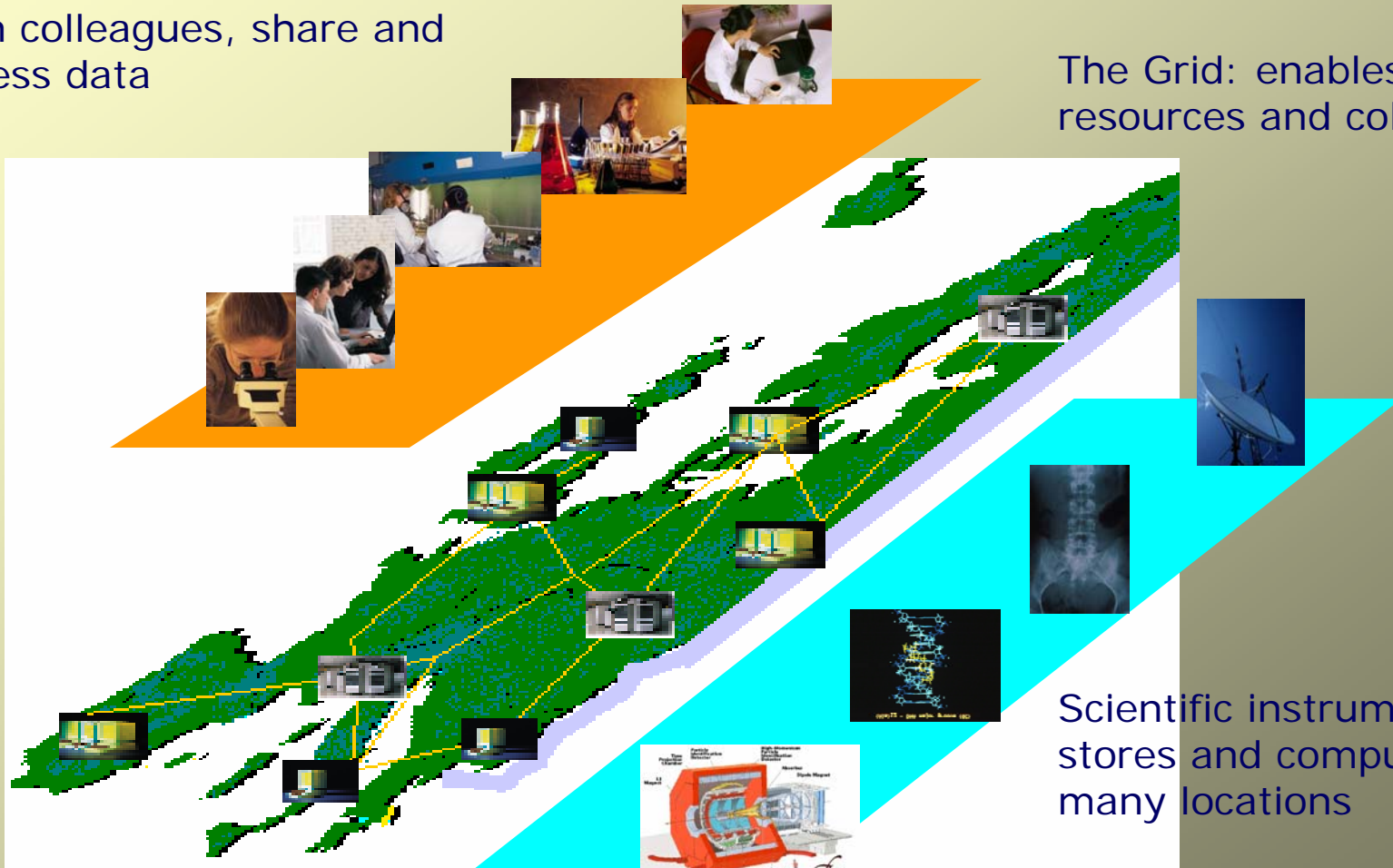FTP, telnet, blood, sweat and tears… and little support for collaboration

Scientific instruments, data stores and computers in many locations

COST Action P10 'Physics of Risk', 13-16 May 2006, Vilnius

The grid vision:

Researchers in many
locations need to interact
with colleagues, share and
access data

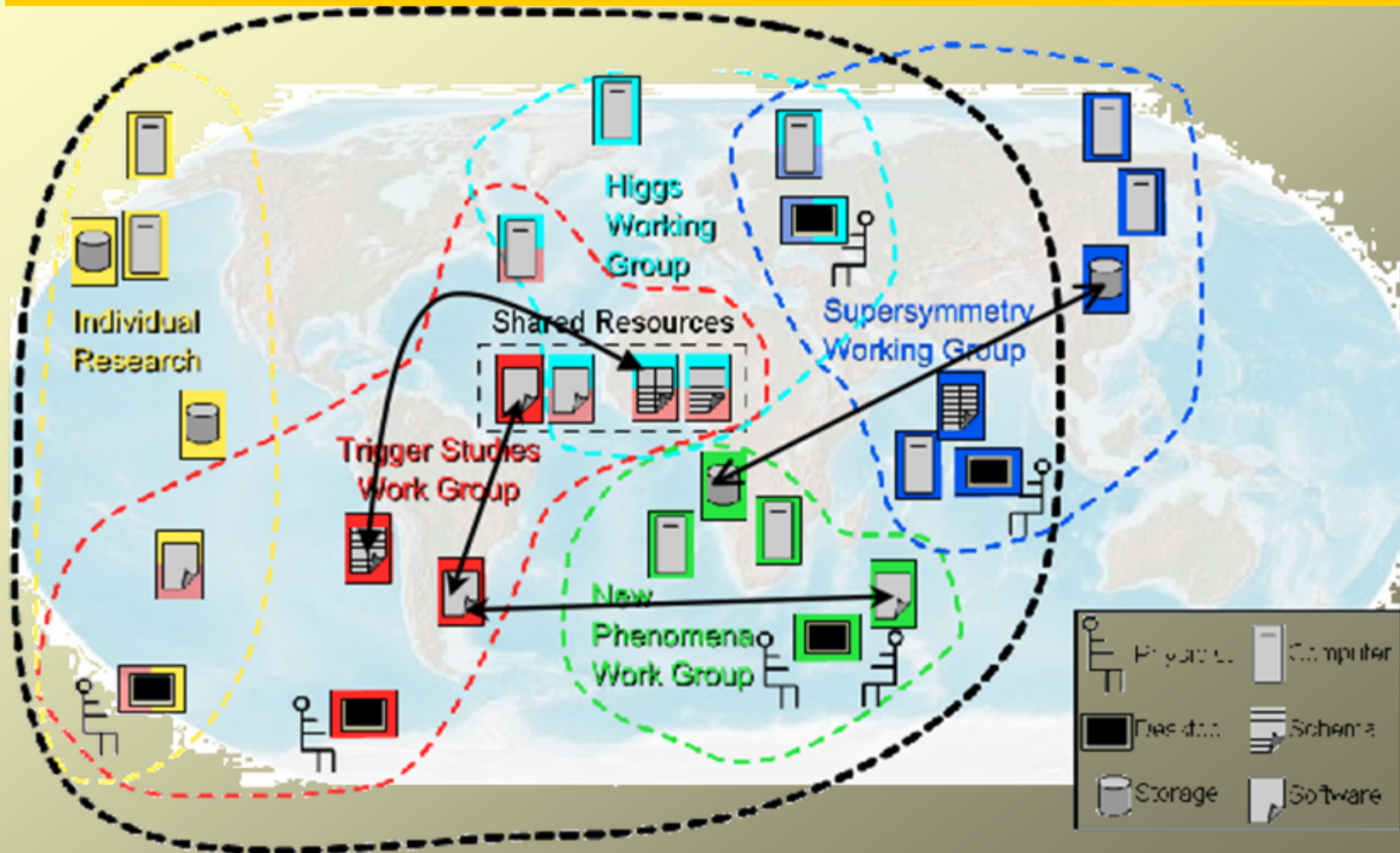The Grid: enables sharing of
resources and collaboration

Scientific instruments, data
stores and computers in
many locations

The emergence of global knowledge communities:

COST Action P10 'Physics of Risk', 13-16 May 2006, Vilnius

Grid applications:

- **High Energy Physics** *(analysing the results from particle collisions)*
- **Medical/Healthcare** *(imaging, diagnosis and treatment)*
- **Bioinformatics** *(study of the human genome and proteome to understand genetic diseases)*
- **Nanotechnology** *(design of new materials from the molecular scale)*
- **Engineering** *(design optimization, simulation, failure analysis and remote Instrument access and control)*
- **Natural Resources and the Environment** *(weather forecasting, earth observation, modeling and prediction of complex systems)*
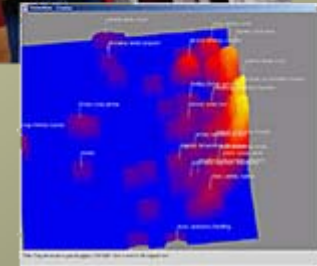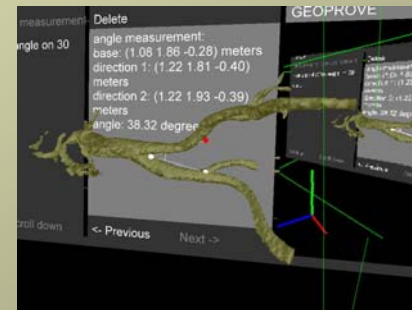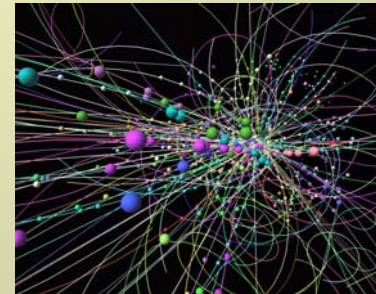
Grid projects:

Many Grid development efforts — all over the world

- NASA Information Power Grid
- DOE Science Grid
- NSF National Virtual Observatory
- NSF GriPhyN
- DOE Particle Physics Data Grid
- NSF TeraGrid
- DOE ASCI Grid
- DOE Earth Systems Grid
- DARPA CoABS Grid
- NEESGrid
- DOH BIRN
- NSF iVDGL

- DataGrid (CERN, ...)
- EuroGrid (Unicore)
- DataTag   (CERN,…)
- Astrophysical Virtual Observatory
- GRIP (Globus/Unicore)
- GRIA (Industrial applications)
- GridLab (Cactus Toolkit)
- CrossGrid (Infrastructure Components)
- EGSO (Solar Physics)

- UK – OGSA-DAI, RealityGrid, GeoDise, Comb-e-Chem, DiscoveryNet, DAME, AstroGrid, GridPP, MyGrid, GOLD, eDiamond, Integrative Biology, …
- Netherlands – VLAM, PolderGrid
- Germany – UNICORE, Grid proposal
- France – Grid funding approved
- Italy – INFN Grid
- Eire – Grid proposals
- Switzerland - Network/Grid proposal
- Hungary – DemoGrid, Grid proposal
- Norway, Sweden - NorduGrid

EGEE – what does it deliver?
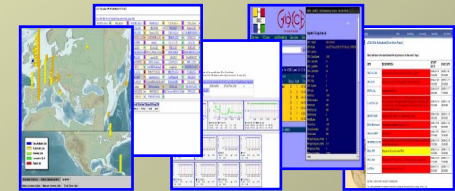
- Infrastructure operation
  - **Continuous monitoring of grid services in distributed global infrastructure**
  - **Automated site configuration/management**
- Middleware
  - **Production quality middleware distributed under business friendly open source licence**
- User Support - *managed process from first contact through to production usage*
  - **Training**
  - **Documentation**
  - **Expertise in grid-enabling applications**
  - **Networking events (User Forum, conferences etc.)**
- Future
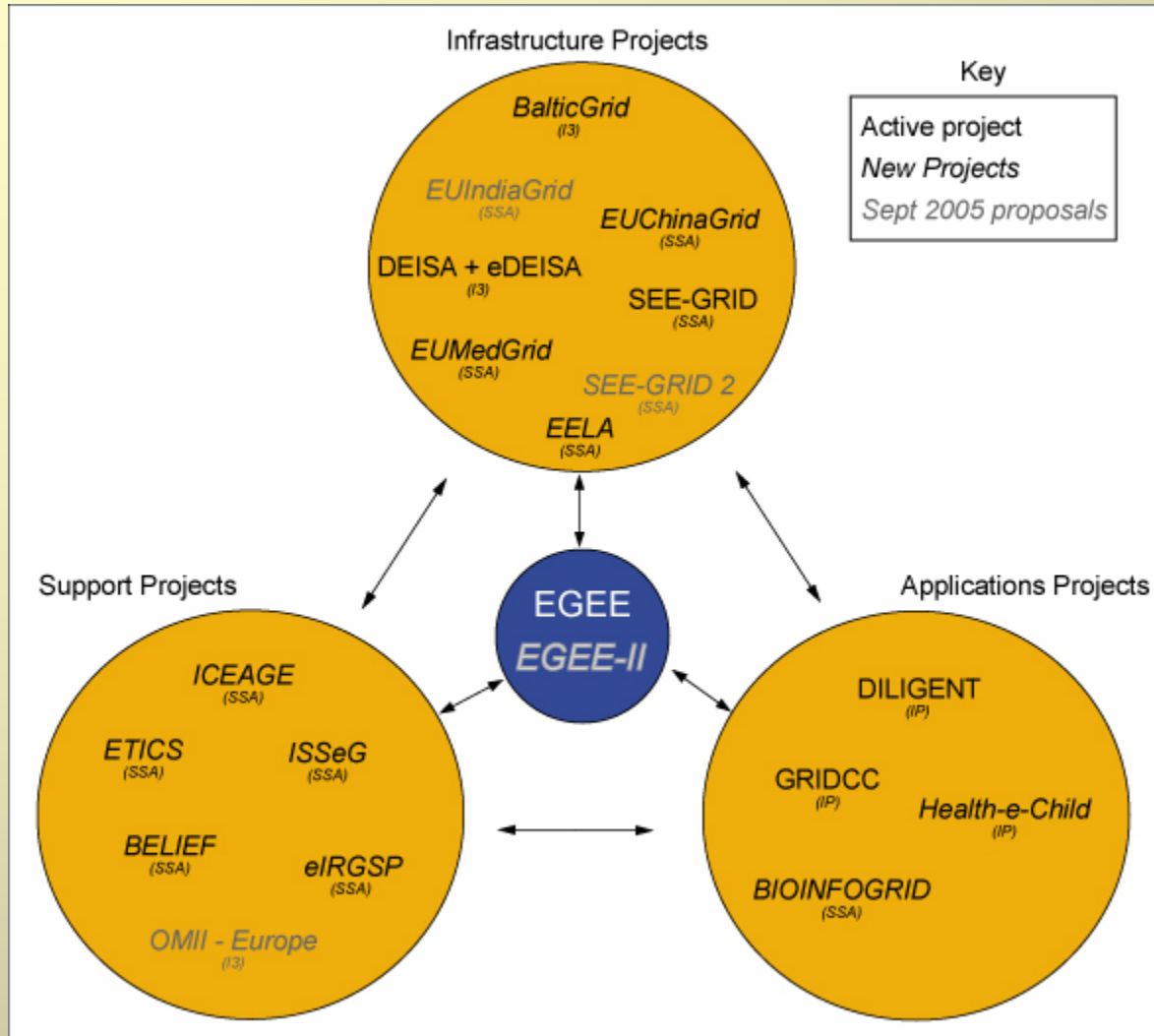 - **Expand on interoperability with related infrastructures**

Expertise & Resources

- More than 90 partners

- 32 countries

- 12 federations

→ Major and national Grid projects in Europe, USA, Asia

+ 27 countries through related projects:

  – BalticGrid

  – SEE-GRID

  – EUMedGrid

  – EUChinaGrid

  – EELA

Related EU projects

Infrastructure Projects

BalticGrid (I3)

EUIndiaGrid (SSA)

EUChinaGrid (SSA)

DEISA + eDEISA (I3)

SEE-GRID (SSA)

EUMedGrid (SSA)

SEE-GRID 2 (SSA)

EELA (SSA)

Key

Active project
New Projects
Sept 2005 proposals

EGEE
EGEE-II

Support Projects

ICEAGE (SSA)

ETICS (SSA)

ISSeG (SSA)

BELIEF (SSA)

eIRGSP (SSA)

OMII - Europe (I3)

Applications Projects

DILIGENT (IP)

GRIDCC (IP)

Health-e-Child (IP)

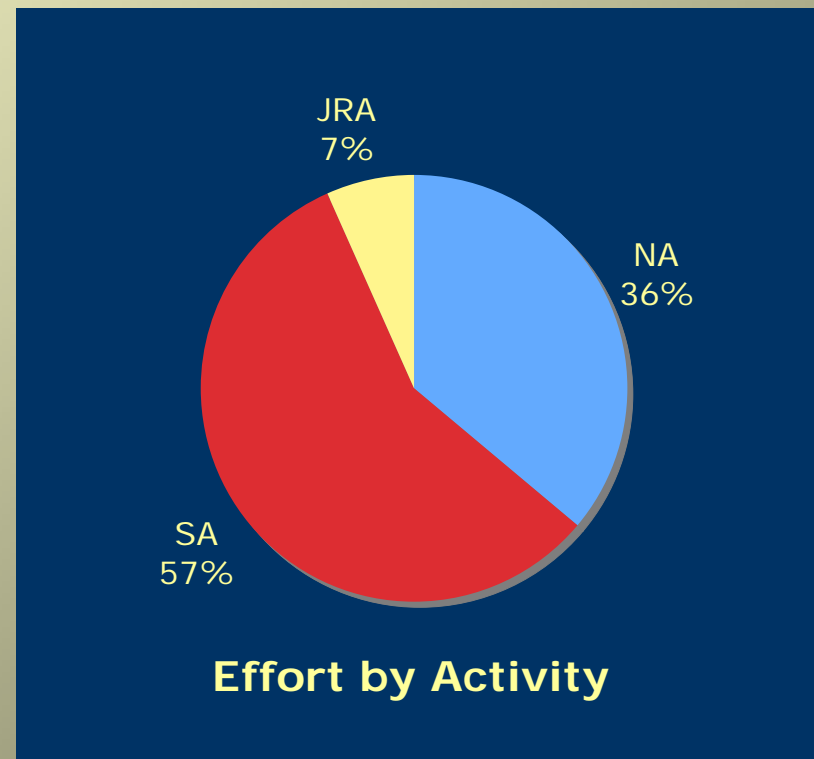**BIOINFOGRID** (SSA)

COST Action P10 'Physics of Risk', 13-16 May 2006, Vilnius

BalticGrid in one slide:

- **partners:**
  - **10 leading institutions in six countries in the Baltic Region and CERN**

- **budget:**
  - **3.0 M€ over 30 months**

- **coordinator:**
  - **KTH PDC, Stockholm**

- **compute resources:**
  - **17 resource centres**

- **start:**
  - **1 November 2005**

Effort by Activity

SA - Specific Service Activities
NA - Networking Activities
JRA - Joint Research Activities

BalticGrid objectives:

- **sustainable integration of the research and educational computing and communication infrastructure in the Baltic States into the European Grid infrastructure**

- **enable the formation of effective research collaborations in the Baltic States, within Europe and beyond**

- **enabling an efficient sharing of unique instruments and data, as for instance environmental data related to the Baltic Sea**

BalticGrid approach:

- **maximize use of human resources for Grid development and deployment by implementing the Baltic Grid as an extension of EGEE**
  - Grid operations coordinated with the EGEE North European Regional Operating Centre

- **assure manageable co-existence with other European Grids**
  - NorduGrid, DEISA, CrossGrid, …

- **engage the Baltic States in Grid related policy and standards activities**

Eight activities:

- **Networking activities**
  - **NA1: Management of the I3**
  - **NA2: Education, Training, Dissemination and Outreach**
  - **NA3: Application Identification and Support**
  - **NA4: Policy and Standards Development**

- **Specific Service activities**
  - **SA1: Baltic Grid Operation**
  - **SA2: Network Resource Provisioning**

- **Joint Research activity**
  - **JRA1: Account Service Level Agreements Markets and Dynamic Account Management**

Relationship to EGEE and other grids:

- **Collaborating or related grid projects**
  - **EGEE**
    - **Infrastructure, Operations**
  - **ICEAGE**
    - **Education and Training**
  - **BELIEF**
    - **Dissemination and Outreach**
  - **eIRGSP**
    - **Policy and Standards Development**

Relationship to EGEE and other projects:

- **Middleware**
  - **gLite**
- **Infrastructure**
  - **EGEE**
- **Applications**
  - **Application Pilots**
  - **Regional Special Interest Groups**
- **ETDO**
  - **Own activities as well as join(t) activities with EGEE(-II) and ICEAGE**
- **Policy Issues**
  - **NA4: Policy and Standards Development**
    - **eIRG**
    - **GGF**
    - **EUGridPMA**

The Lithuanian Grid project:

The Lithuanian Grid project (LitGrid) aims to:

- develop the research and education computing and communication infrastructure in Lithuania

- integrate this infrastructure into the emerging European Grid and Baltic Grid infrastructure

- bring the knowledge in grid technologies and use of grids in Lithuania to a level comparable to that in EU member states with a longer experience in the development, deployment and operation of grids,

- further engage Lithuania in policy and standards setting activities

How the LitGrid started:

- **the list of potential and valuable applications for grid computing was compiled by Lithuanian scientists and specialists**

- **than the Lithuanian Science and Studies Foundation (VMSF) took the initiative to issue a call for grid projects**

- **the proposal for LitGrid program was prepared and submitted to VMSF, accepted in July 2005**

- **the LitGrid program initially included 10 partners, universities and research institutes, as well as business partner**

- **now, for the second year of the program we have 12 partners, covering all suitable universities and research institutes (it means, having computing potential and interest o compute) of Lithuania**

Contribution of partners – clusters for LitGrid:

- **10 clusters**

- **about 200 CPUs, about 10 TBs of storage**

- **CPU architecture: x86 (Intel-32) or Ia64 (Itanium2) or Sparc**

LitGrid aims to develop products and services:

- **DP1 – LitGrid operational infrastructure and support**

- **DP2 – the development of parallel and distributed algorithms, grid deployment of datasets, other HPC procedures for research applications**

- **DP3 – the education and dissemination services for LitGrid**

- **DP4 – the group of software developers for analysis, design and implementation of algorithms for the grid infrastructure**

- **DP5 – the administration and management services for LitGrid**

Research Projects:

- **Transport Modelling System - Mobile Solutions of Public Information and Transportation Systems**

- **Mathematical Modeling of Heterogenous Processes and Systems Interaction**

- **Mathematical Modelling of Lithuanian Economy**

Applications, summarized:

- **Modeling of heterogeneous systems and processes in biology and biochemistry**

- **Topics in material science (GAMESS group)**

- **Energy spectra and hyperfine structure of atoms and ions**

- **Modeling of free and twosided surfaces**

- **Powder dynamics**

- **Embeded systems**

- **Analysis of biological macromolecules**

- **3D visual data mining**

BalticGrid applications, comparing to LitGrid:

- **High-Energy Physics (statistical data analysis, Monte Carlo simulation)**

- **Material Sciences (atomic and molecular structures, solid state surfaces, photon and electron interactions)**

- **Bioinformatics (sequence pattern discovery, modeling of biosensors)**

- **Special Interest Groups** *(SIG: Baltic Sea Eco-System Modelling, Text Annotation Service, Text-to-Speech Service, Stellar Spectra Computation, Atomic and Nuclear Computation, Modelling of Heterogeneous Processes)*

- **Application Support (application adaptation experts, deployment of Migrating desktop)**

SIG – e-services:

- **Access Grid is an internet-based model for group-to-group communication: large-scale distributed meetings, collaborative work sessions, seminars, lectures, tutorials, and training**

- **BalticGrid will add repository, and other tools**

Financial application – real-time service:

- grids are quite attractive for large scale financial applications like dynamic financial services (applications must complete within strict deadlines)

- these complex tasks need an amount of computing power that is unfeasible to keep in house and a plenty of 'headroom' in resource availability

- grids can deliver the amounts of power needed

- application is developed on EGRID facility for a test case - an application that estimates the sensitivities of a set of stocks to specific risk factors

- implemented solutions allocate grid resources in advance to eliminate latency due to the job submission mechanism (as soon as we get enough resources allocated we can interact with them in real time)

Financial application – real-time service:

- **the booking mechanism:**
  - **an early submission of a bunch of jobs is run for securing the availability of WN at a given time**
  - **each pooled node executes a program regularly checking a host**
  - **the contacted host enrolls this WN for the user's program, as soon as the user executes that program**
  - **when the execution terminates the results are available in real time without any delay introduced by WMS**
  - **the WNs remain booked, and so are ready to be enrolled again for other program executions**
- **this approach, where the WN asks to be enrolled in a computation thereby acting as a client, is needed because the WN cannot be reached directly from the UI**

LitGrid program: www.litgrid.lt

BalticGrid: www.balticgrid.org

EGEE: www.eu-egee.org

thank you for attention,

looking forward for comments and suggestions

sincerely Algimantas Juozapavicius