

IDGF-SP

International Desktop Grid Federation - Support Project

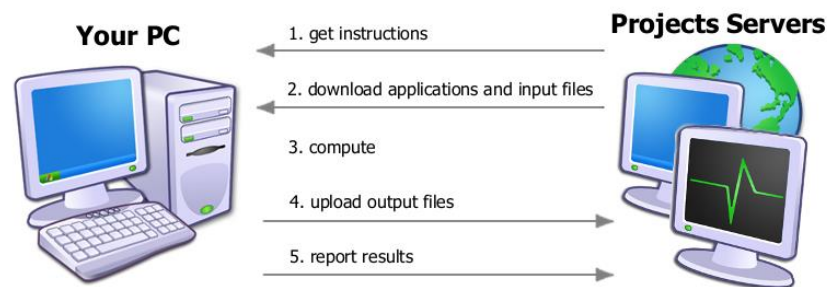
Crowd computing to support EGI scientists

József Kovács, Róbert Lovas
{josef.kovacs,robert.lovas}@sztaki.mta.hu
MTA SZTAKI – LPDS

IDGF-SP is to be supported by the FP7 Capacities Programme under contract nr RI-312297 .

Volunteer and Crowd Computing

- **Volunteer computing** is an arrangement in which people (**volunteers**) provide computing resources to **projects**, which use the resources to do distributed computing and/or storage.
- The most widespread middleware to implement volunteer computing is **BOINC**.
- More than 1 billion PCs -> tablets, smart phones, game consoles
- **Crowd computing**: volunteer, but focusing on utilising **human intellectual capabilities** (e.g. recognition)



IDGF-SP core production infrastructure



Supported VOs

vlemmed
fusion
gilda
hungrid
seegrid
edgiprod.vo.edgi-grid.eu
chem.vo.ibergrid.eu

compchem
gaussian
trgrida
trgridb
biomed
enmr.eu
inaf



meta job

single job

Application Repository-

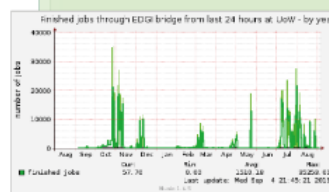


Desktop Grid Servers

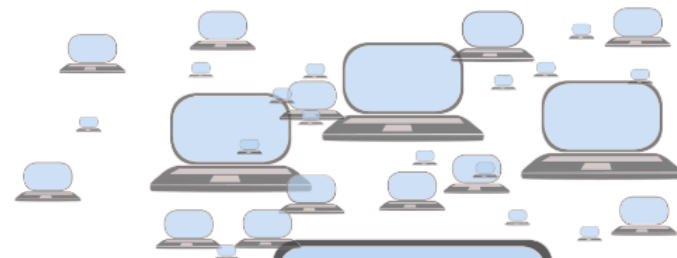


Applications

GBAC



Volunteer Resources



ABC@home
AlmereGrid
Charity Engine
EDGeS@home
EDGIDemo
SZTAKI Desktop Grid
Westminster Campus DG

Clouds

Amazon (on demand)
Westminster Cloud
SZTAKI Cloud
LPDS Cloud

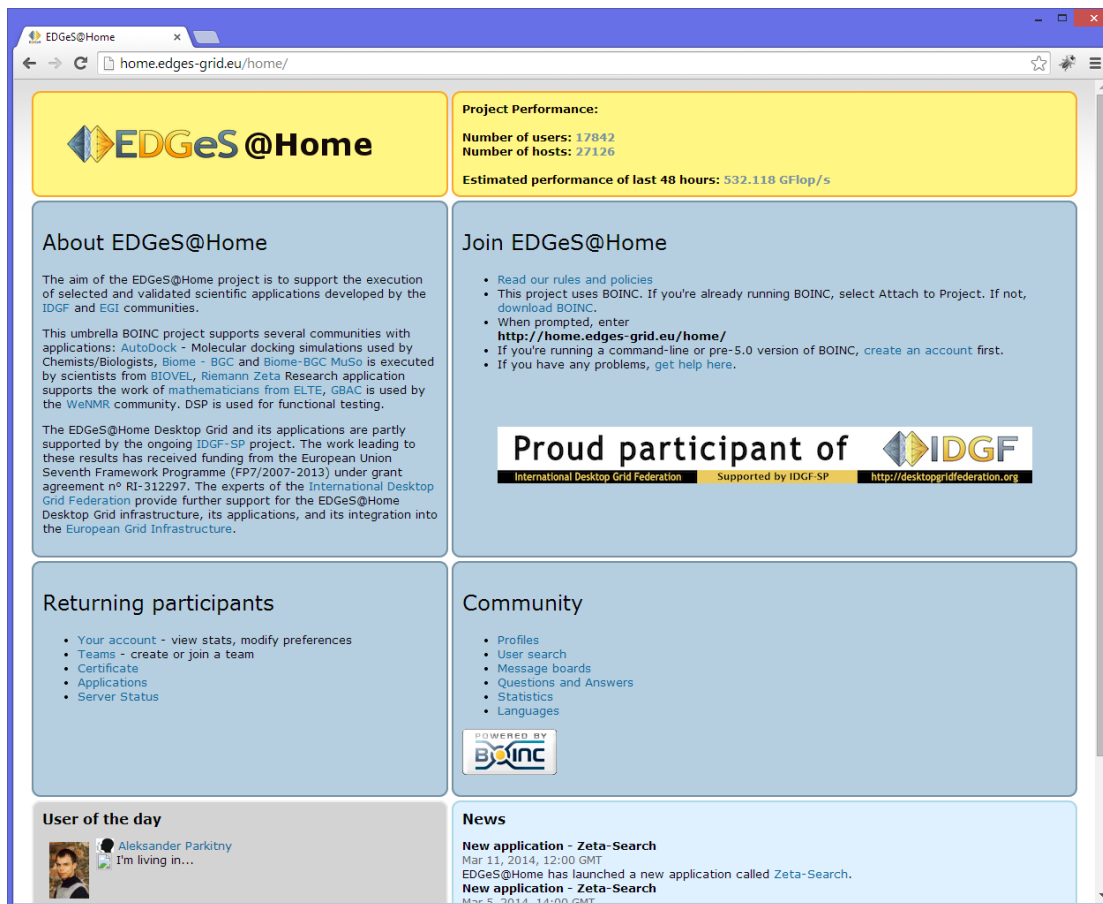
BOINC Client



AutoDock Portal



Applications at EDGeS@home for EGI scientists



The screenshot shows the EDGeS@home website interface. At the top, there's a navigation bar with the EDGeS@home logo and a link to home.edges-grid.eu/home/. Below this, the main content area is divided into several sections:

- Project Performance:** Number of users: 17842, Number of hosts: 27126, Estimated performance of last 48 hours: 532.118 GFlop/s.
- About EDGeS@Home:** A section describing the project's aim to support the execution of selected and validated scientific applications developed by the IDGF and EGI communities. It mentions the umbrella BOINC project and its support for various scientific communities.
- Join EDGeS@Home:** A section with links to read rules and policies, download BOINC, and create an account. It also provides a link to the project's website.
- Proud participant of:** A banner showing EDGeS@home as a proud participant of the IDGF (International Desktop Grid Federation), supported by IDGF-SP.
- Returning participants:** A section with links to view account stats, create or join a team, and view server status.
- Community:** A section with links to profiles, user search, message boards, questions and answers, statistics, and languages.
- User of the day:** A section featuring a user profile (Aleksander Parkitny) and a message.
- News:** A section with news items, including a new application called Zeta-Search.

- **CNS [GBAC] (WeNMR)**
- **DIRAC [GBAC] (DIRAC)**
- **BBGC/MUSO (BIOVEL)**
- **Autodock (publicly available)**
- **Zeta-search (ELTE, Hungary)**
- **LinAlgOpt [GBAC] (Pannon, Hungary)**

Promoting Desktop Grids Virtual Team in EGI

■ General project info

- Start (planned) : November 2013
- Duration : 6 months + **3 months extension**
- Leader: Robert Lovas
- URL: <http://go.egi.eu/vtdg>

- *“EGI-InSPIRE is ideally placed to **join** together the new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and **desktop grids**, for the benefit of user communities within the European Research Area.”* --- EGI-InSPIRE Project fact sheet

■ Motivation

- Despite the on-going work and the benefits of the integrated Desktop Grid DCI resources available for EGI users, only **a part of the EGI community is aware** of the latest Desktop Grid related achievements, and a fraction of the EGI users and infrastructure operators take their advantages in everyday practice.
- In order to fill this gap; a part of the targeted objectives are technical ones but there is stronger focus on the human aspects and **promotion** – training, networking and support activities.

Monitoring and accounting

Performance of bridges - x

desktopgridfederation.org/pe

Start Members News Doc

International Desktop Grid Federation > Intern

Navigation

Technology

- Infrastructures
 - Status
 - Overview Bridged infrastructures
 - Bridge services: qLite to Desktop Grid
 - Overview
 - Available applications
 - Performance of bridges
 - Bridge service status
 - Bridge services: Desktop Grid to qLite
 - Cloud services for Desktop Grids
- Applications
- Green Crowd Computing

EGI Accounting Portal --> x

accounting.egi.eu/egi.php?ExecutingSite=SZTAKI&query=njobs&startYear=2014&startMonth=1&endYear=2014&endMonth=5

EGI ACCOUNTING PORTAL

GLOBAL View VO MANAGER View VO MEMBER View SITE ADMIN View USER View REPORTS METRICS PORTAL LINKS

Hierarchical Tree EGI View --> Production

http://accounting.egi.eu

Data to graph: Number of jobs Total number of jobs run

Period: Start year: 2014 Start month: 1 End year: 2014 End month: 5

Groupings: Show data for: VO as a function of: DATE

VO Groups: ☐ LHC ☐ TOP 10 ☐ ALL ☐ Custom

Chart: Type: GROUP BAR Scale: LINEAR

dteam VO: ☒ Exclude dteam and ops VOs jobs information

Local Jobs: ☒ Grid Jobs Only ☐ Grid Jobs and Local Jobs ☐ Local Jobs Only

Refresh

SZTAKI Total number of jobs by VO and DATE.
VOs. January 2014 - May 2014.

Developed by CESGA

The following table shows the distribution of Total number of jobs grouped by VO and DATE.

Total number of jobs run by VO and DATE (Excluded dteam and ops VOs)							
VO	Jan 14	Feb 14	Mar 14	Apr 14	May 14	Total	%
biomed	533	15	19	90	497	1,154	8.57%
edgiprod.vo.edgi-grid.eu	23	6	5	0	0	34	0.25%
enmr.eu	671	1,535	1,252	2,913	662	7,033	52.25%
vo.aginfra.eu	0	1	5	0	0	6	0.04%
vo.edges-grid.eu	3,877	0	0	1,357	0	5,234	38.88%
Total	5,104	1,557	1,281	4,360	1,159	13,461	
Percentage	37.92%	11.57%	9.52%	32.39%	8.61%		

Click here for a CSV dump of this table

Click here for a Extended CSV dump of this table

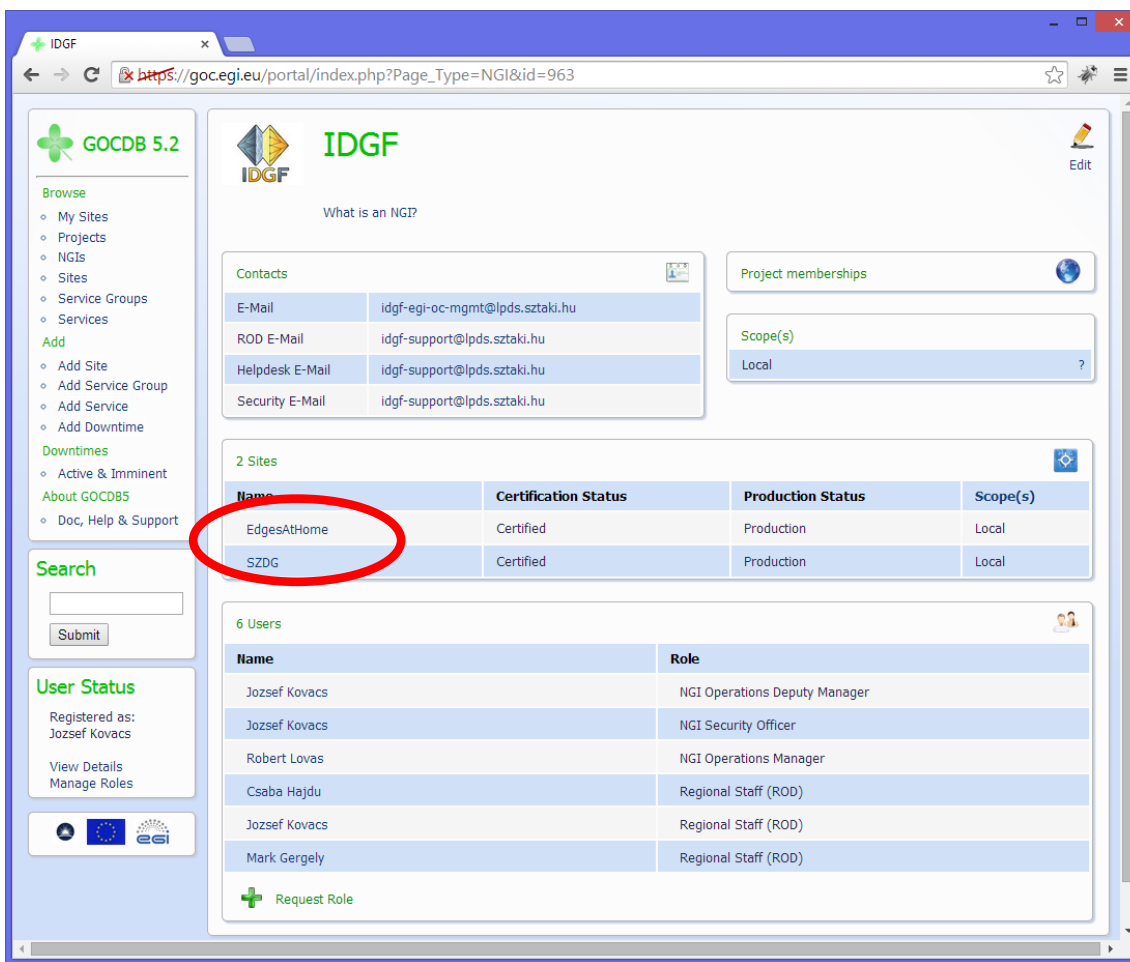
Click here for XML encoded data

European Grid Infrastructure

(C) CESGA 2012

http://desktopgridfederation.org/infrastructures

IDGF (Regional) Operation Center



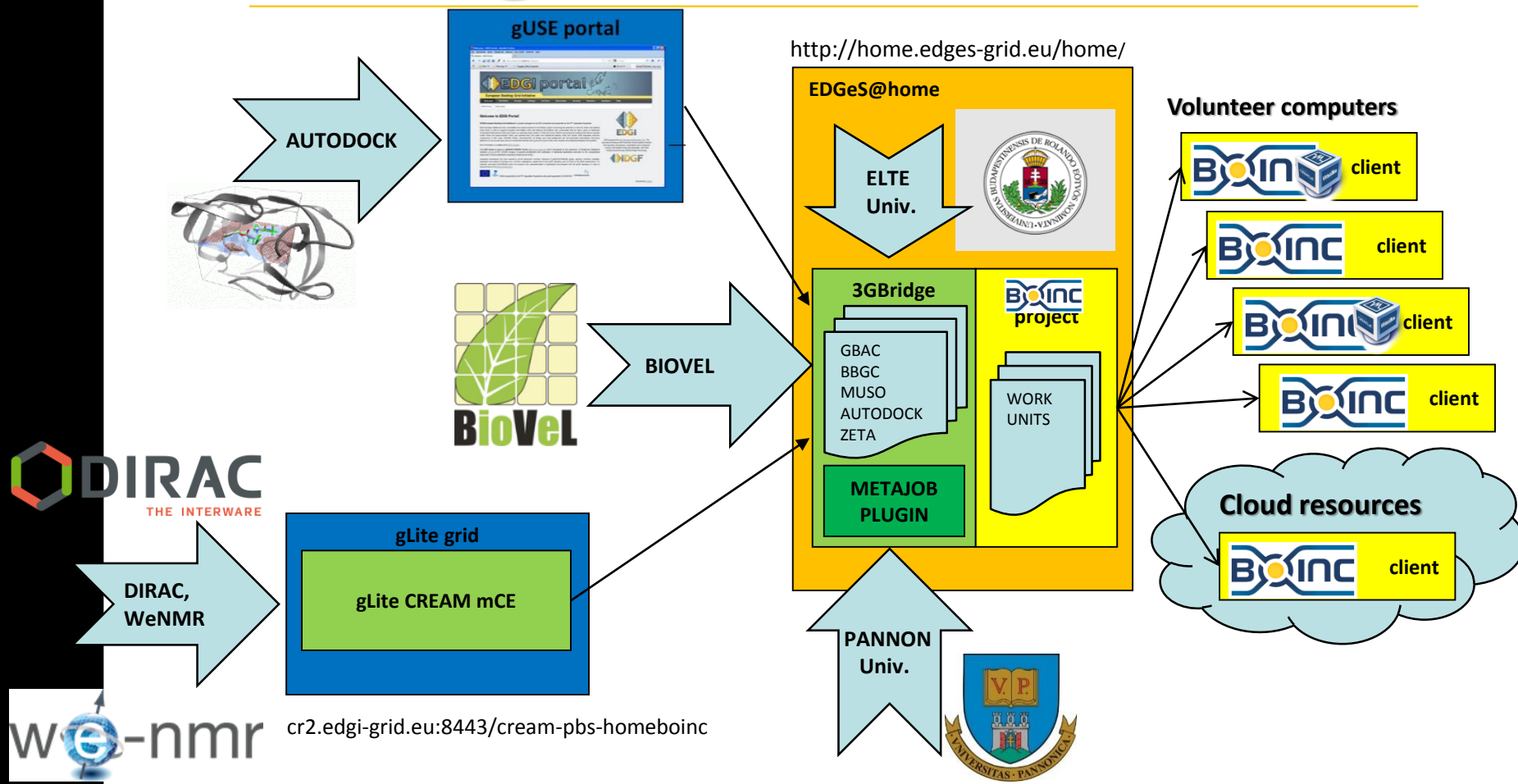
The screenshot displays the IDGF (Regional) Operation Center web interface. The main content area includes sections for Contacts, Project memberships, Scope(s), 2 Sites, and 6 Users. The '2 Sites' section is highlighted with a red circle, showing a table with columns: Name, Certification Status, Production Status, and Scope(s). The '6 Users' section shows a table with columns: Name and Role.

Name	Certification Status	Production Status	Scope(s)
EdgesAtHome	Certified	Production	Local
SZDG	Certified	Production	Local

Name	Role
Jozsef Kovacs	NGI Operations Deputy Manager
Jozsef Kovacs	NGI Security Officer
Robert Lovas	NGI Operations Manager
Csaba Hajdu	Regional Staff (ROD)
Jozsef Kovacs	Regional Staff (ROD)
Mark Gergely	Regional Staff (ROD)

- IDGF OC has been established
- IDGF OC to collect Desktop Grid resources for EGI
- Sites are being set-up
- One site represent a DG server
- EDGeS@home & SZDG

Overview of various utilisation of EDGeS@home volunteer resources

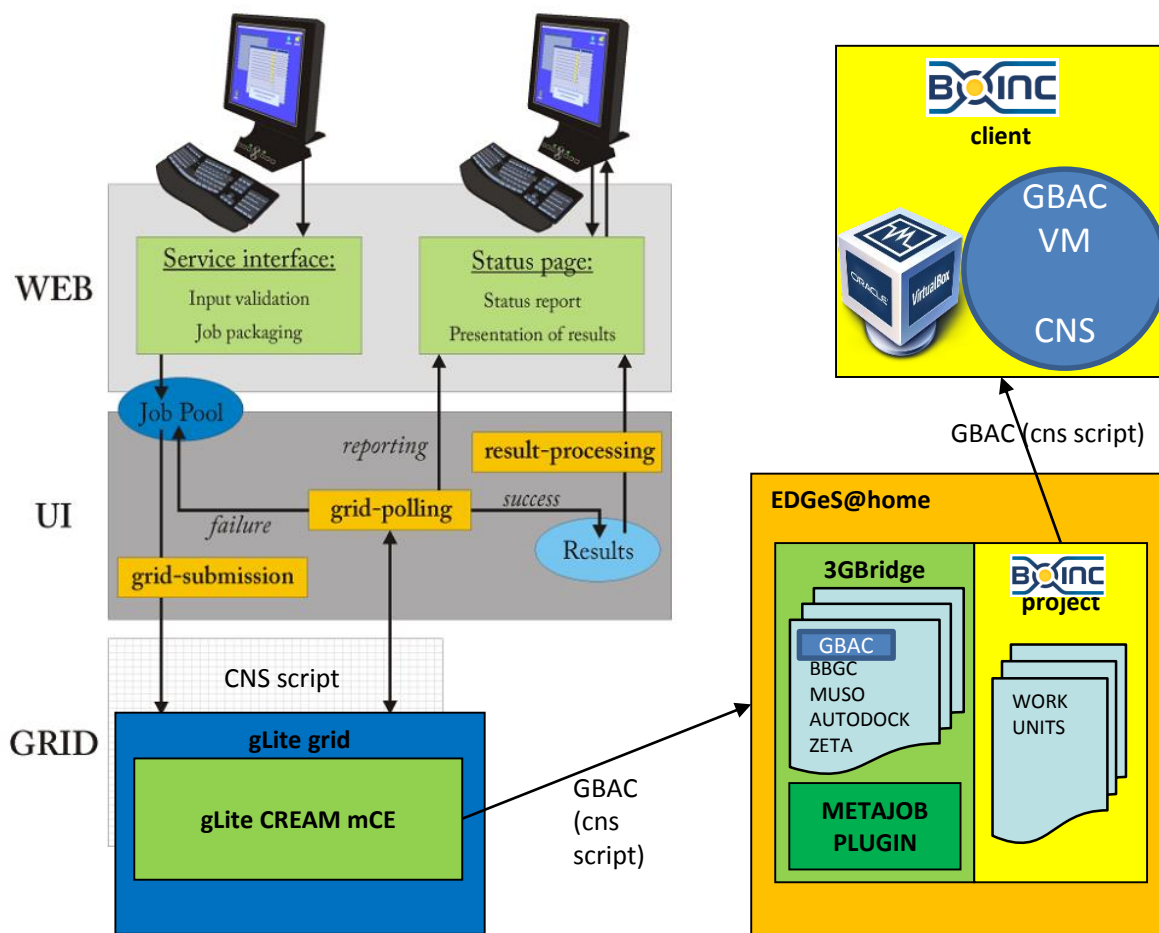


FP7 WeNMR



- WeNMR (<https://www.wenmr.eu/>) is a **Virtual Research Community** supported by EGI, the largest one within the **life science area**. WeNMR aims at bringing together complementary research teams in the structural biology and life science area into a virtual research community at a worldwide level and provide them with a platform integrating and streamlining the computational approaches necessary for NMR and SAXS data analysis and structural modeling.
- WeNMR has a **portal and workflow system** that links and integrates into a variety of computing infrastructures. An important WeNMR application has been ported to Desktop Grids. Also **EDGeS@home** infrastructure has been integrated into WeNMR, it is in production phase.
- The coordinator of WeNMR project has joint the IDGF initiated **“Promoting Desktop Grid” Virtual Team** in the framework of EGI-InSPIRE project.

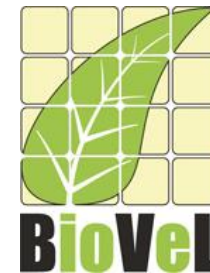
WeNMR HADDOCK portal: modeling of biomolecular complexes



Haddock portal is sending 1 out of 10 jobs to **EDGeS@home** desktop Grid through the modified computing element (Bridge).

Utilisation of volunteer resources with **minimal effort**: no porting, no windows version development

FP7 BioVEL



- BioVel is a virtual e-laboratory that supports research on **biodiversity** issues using large amounts of data from cross-disciplinary sources. BioVel offers the possibility to use computerised "workflows" (series of data analysis steps) to process data, be that from one's own research and/or from existing sources.
- BioVel applications have been ported to Desktop Grids. BioVEL announced collaboration with IDGF-SP partner SZTAKI and support for [EDGES@HOME](http://www.edgesathome.eu). IDGF became one of the "Friends of BioVEL".
- BioVel : <http://www.biovel.eu/>
- Portal: <http://ecos.okologia.mta.hu/bbgcdb/>

Research

Updates on our Workflows

Achievements for ecosystem modeling workflows

By Ferenc Horváth, Zoltán Barcza, Péter Ittles and Dóra Krasser

The Biome-BGC team has achieved considerable scientific and technological breakthroughs recently.



The cooperation with OpenNESS has resulted in a broad range of quantitative ecosystem service indicators (ESI) based on ecosystem simulation and the first

Biome-BGC ESI workflow on the BioVel Portal. Using these indicators (such as annual wood production; yearly maximum of total above ground biomass of grasslands —as biomass provisioning ESI; energy absorption by evapotranspiration —as micro and regional scale climate regulation ESI; damping of ecosystem daily water outflow —as hydrological cycle and water flow maintenance ESI, etc.) in real ecosystem simulations, different climate or land use change scenarios can be evaluated. These new developments were demonstrated at INTECOL 2013.

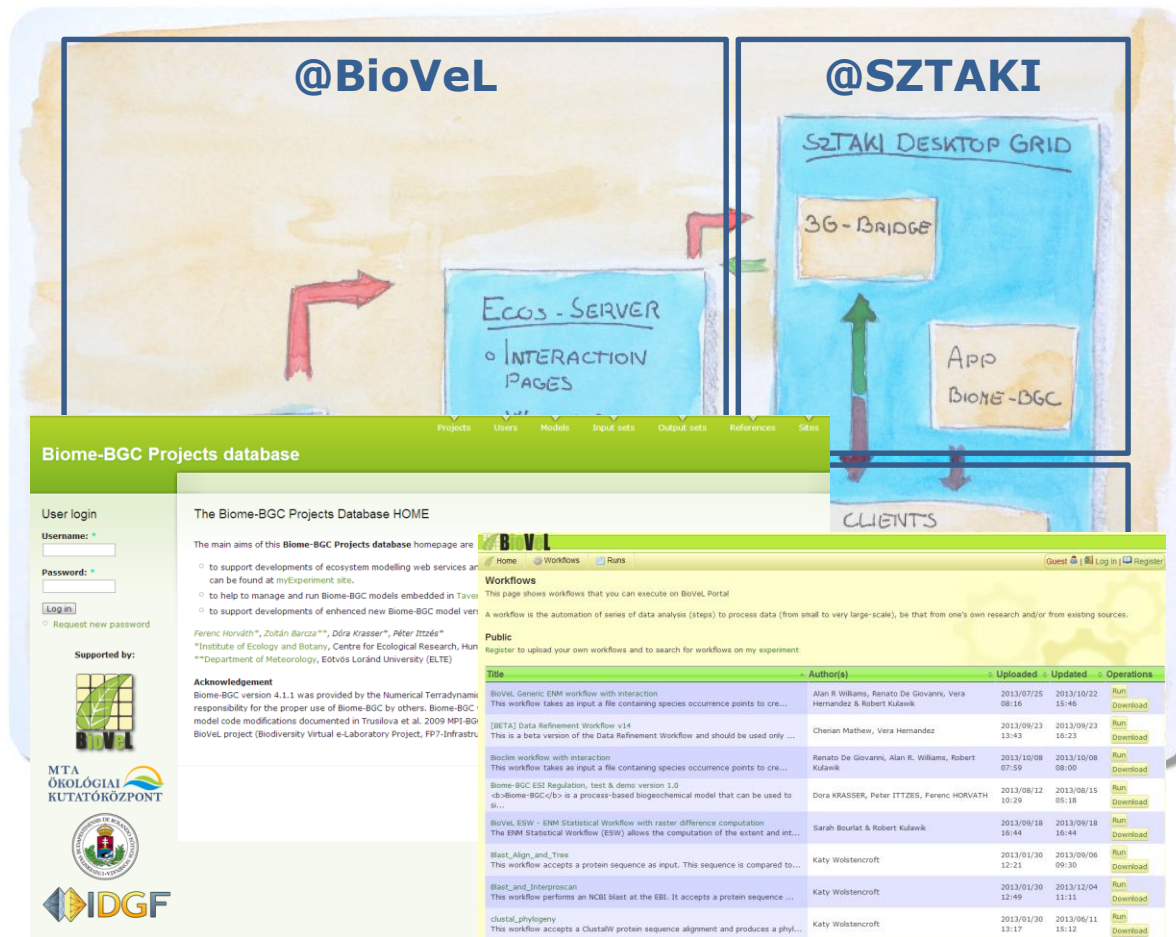
We are also proud to report that the full integration of Biome-BGC based web services and BOINC desktop grid technology was reached, and the 'Biome-BGC Ecosystem Modelling' application is running on the public EDGES@Home at beta level (see figure). Although single run simulations can be executed on the ECOS server, Monte Carlo experiment based sensitivity analysis, data-model harmonization and spatial extension workflows are computationally very demanding.

We also had a very fruitful collaboration with the team of the SZTAKI PERL and we welcome them as "Friends of BioVel." (see page 4)



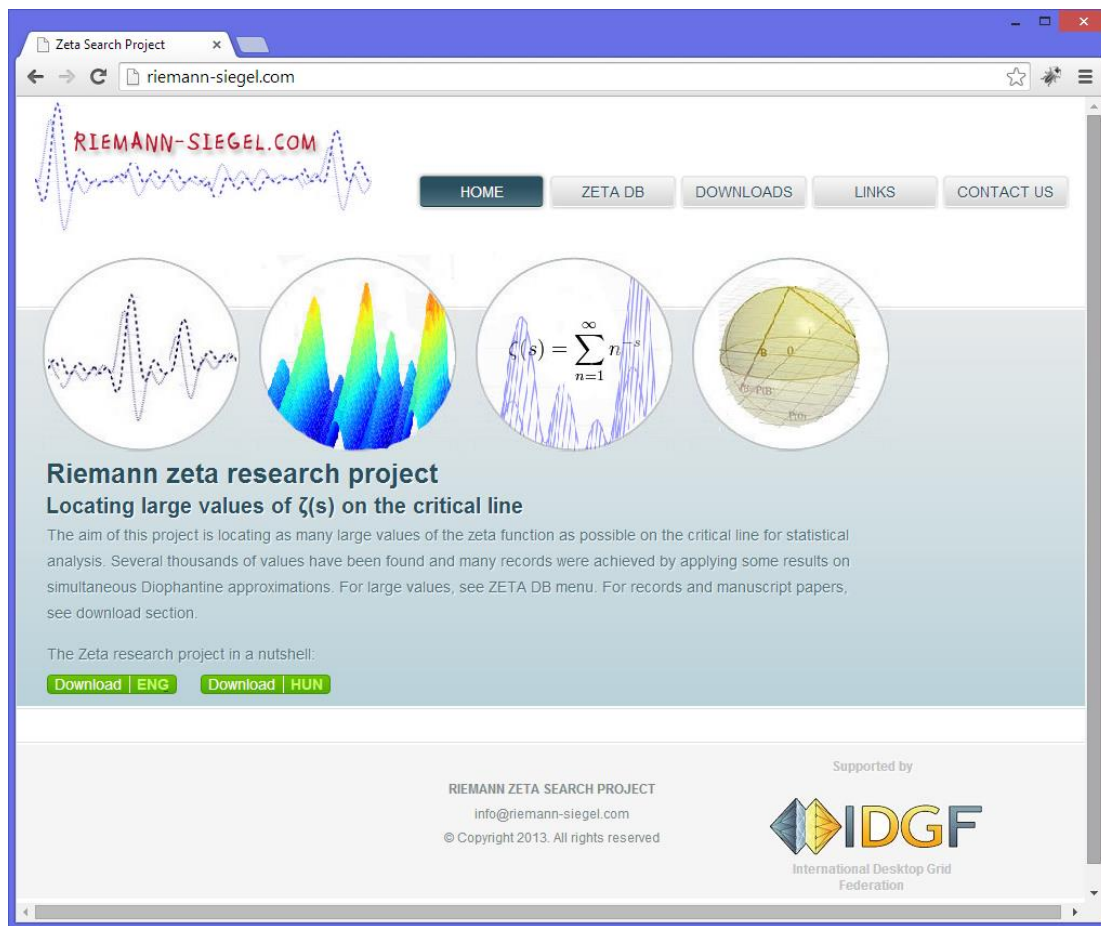
BioVel, Newsletter No. 4, Autumn 2013 11

BIOVEL: ecosystem modeling on E@H



- Ecosystem modelling: **BBGC** and **BBGC-Muso**
- Both apps are driven by the **BioVeL Web-Service** component
- Jobs are generated by **Taverna workflows**
- Taverna workflows are started by users through **BIOVEL portal**

Zeta-search application by ELTE

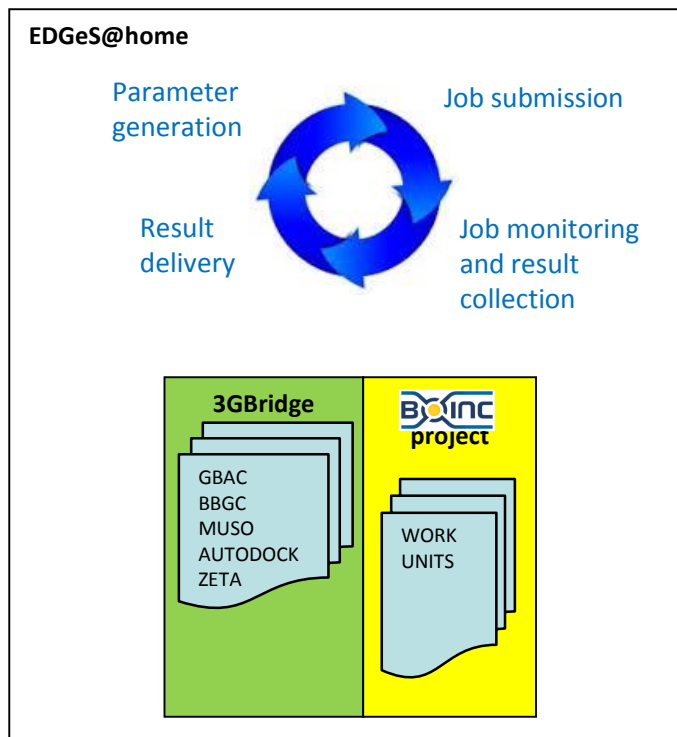


Zeta-search at E@H and SZDG: locating many values where $Z(t)$ (Riemann-Siegel formula) is large in order to get a better understanding of the behavior of the **distribution of primes**, scans the numbers towards infinite.

Application by the Eotvos Lorand University, Hungary

Zeta-search execution on EDGeS@home and SZDG

- A huge **parameter space** is defined
- **Parameter generator** is executed as part of the infrastructure
- Jobs are automatically generated and **submitted**
- Results are **collected** (preprocessed if needed) and **sent** to the application owner



- Zeta-search is an easily programmable parameter sweep application potentially **running for years**
- It generates jobs in the **range of millions**

<http://riemann-siegel.com/>

FP7 SOCIENTIZE + KOPI

- One of the main aims of the FP7 SOCIENTIZE project (<http://www.socientize.eu>) is to **promote the usage** of science infrastructures (dedicated and external resources), including professional and amateur scientists, and also sets up a network where infrastructure providers and researchers recruit volunteers from a **general public to perform science at home**.
- Some of the experiments from SOCIENTIZE are based on the *PyBossa* (pybossa.com) platform. PyBossa is an open source platform for crowd-sourcing online (volunteer) assistance **to perform tasks that require human cognition, knowledge or intelligence** (e.g. image classification, transcription, information location etc.)
- The **KOPI Online Plagiarism Search Portal** (<http://kopi.sztaki.hu>) is an open service that enables (among others) the **users to check for identical or similar contents between their own documents** and the files uploaded by other authors or available in Wikipedia, where KOPI uses the **SZTAKI Desktop Grid** to pre-process the immense database of Wikipedia.

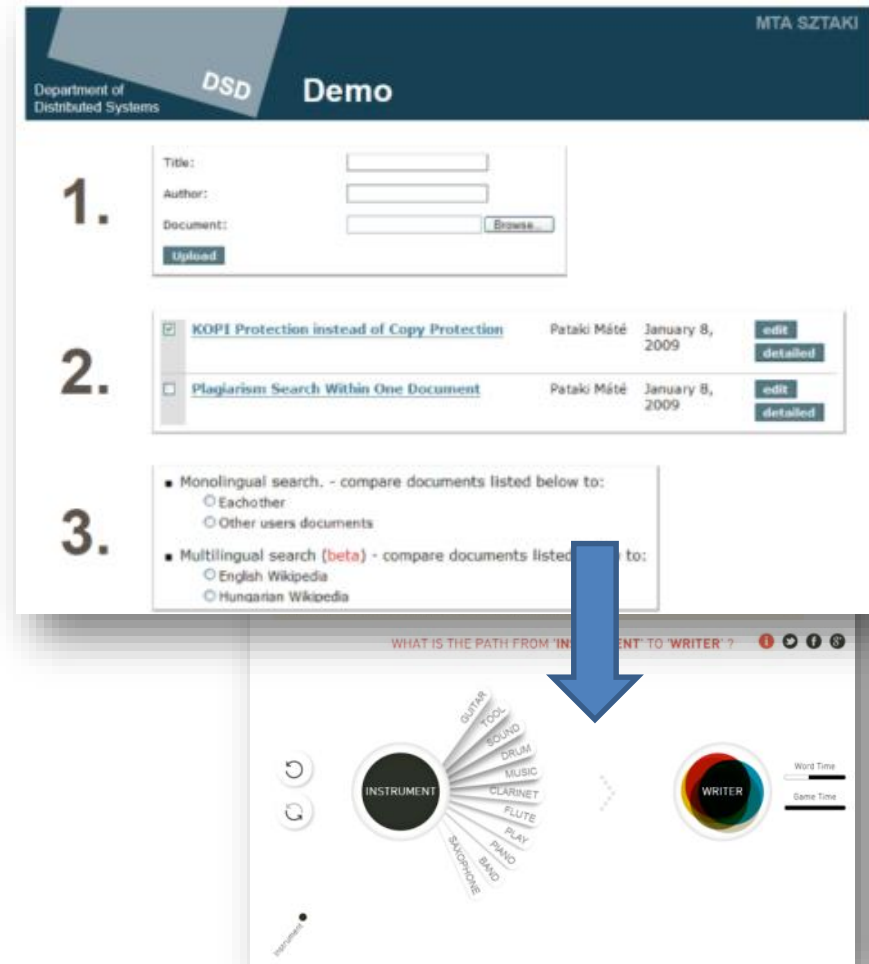
Conclusion

Volunteers (citizen scientist) can be classified in two ways:

- **Beneficiaries** (plagiarism service, improving teaching)
- **Active contributors** (uploaded documents, computer capacity, human intellectual effort)

Scientists

- are able to **harness the power of the general public**, and the
- citizens will get **more** information and **insight** view about the **science**



Thank you for the attention!

Project websites:

<http://idgf-sp.eu>

<http://desktopgridfederation.eu>

<http://doc.desktopgrid.hu>



József Kovács, Róbert Lovas

{josef.kovacs,robert.lovas}@sztaki.mta.hu

Acknowledgement:

- IDGF-SP EU support project (RI- 312297)