Supporting scientific and Web 2.0 communities by desktop grids

P. Kacsuk
MTA SZTAKI
(Univ. of Westminster)

The EDGeS/EDGI projects receive Community research funding

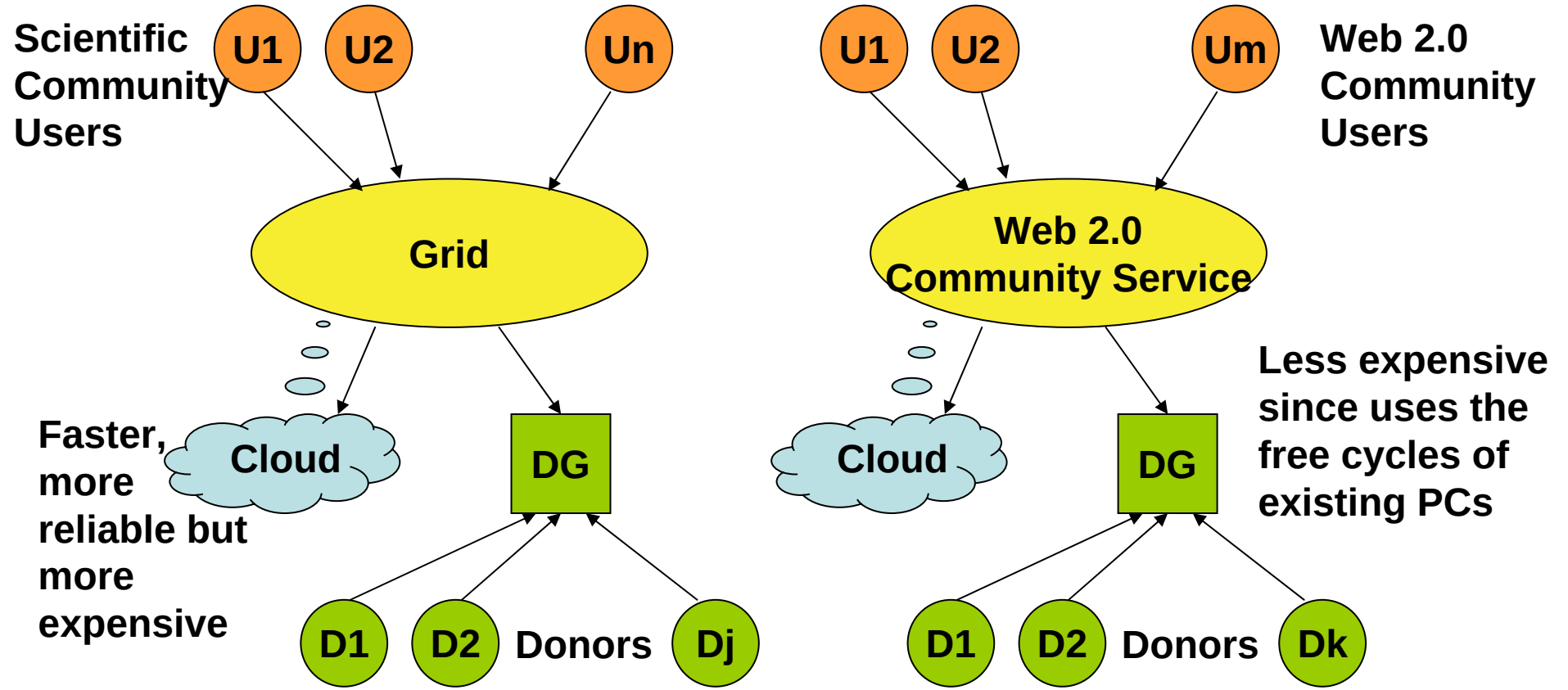
The strength of communities

- Web 2.0 services provide frameworks for communities to put together contents
- Volunteer computing (VC) is a community activity to provide desktop resources for typically scientific communities -> Desktop Grids (DGs)
- Our goal is to bring together the two communities:
  - Web 2.0 communities can help scientific projects to advertise their projects and attract volunteer donors
  - Web 2.0 communities can extend the scope of their applications by using VC technology provided for science
  - Web 2.0 service providers can increase their services without further resource investment

Goals of the EDGI and Web2Grid projects

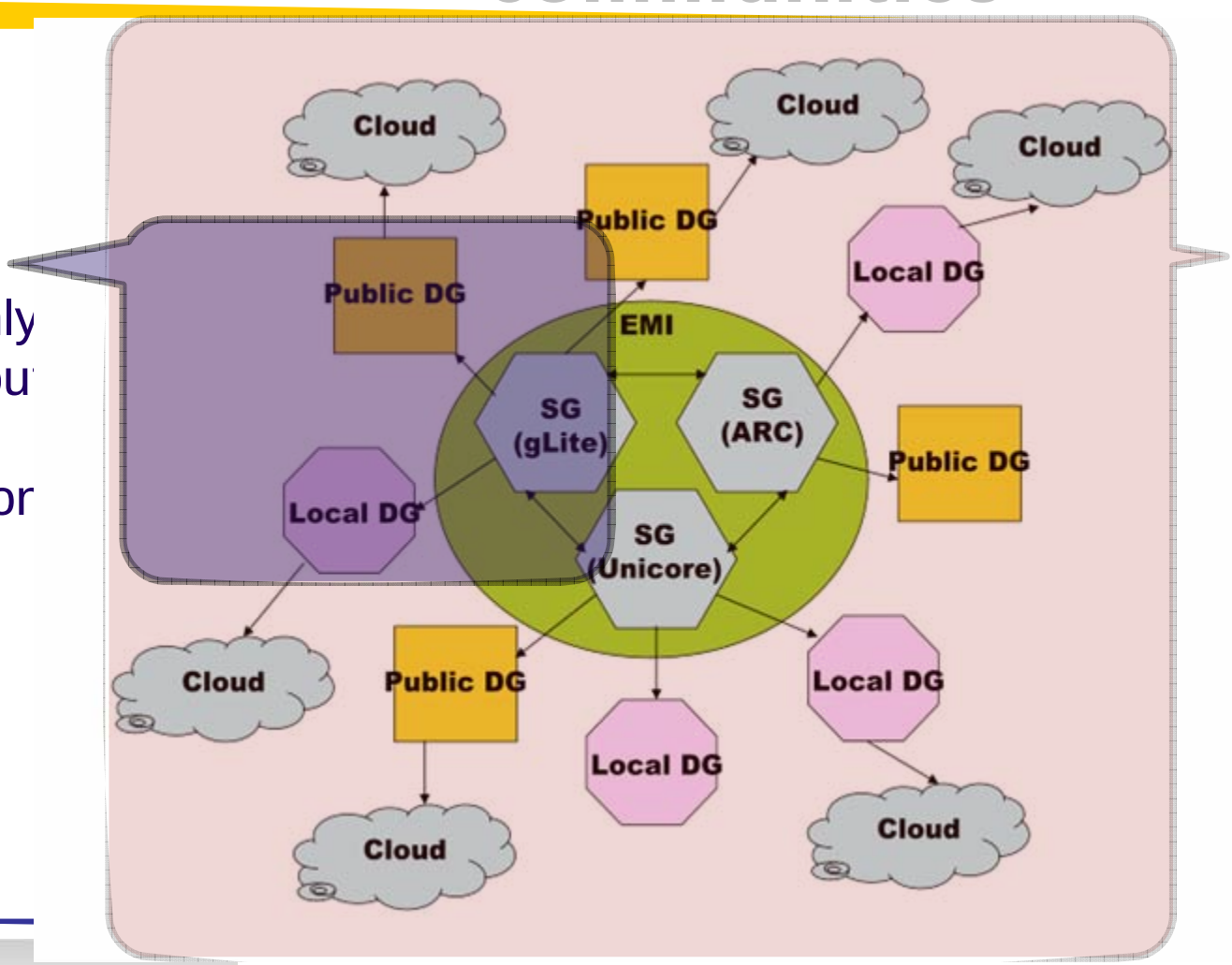
- To support both scientific and Web 2.0 communities by desktop grids
- EDGI: enables scientific communities to extend the existing service grids (SGs) with DG resources
- Web2Grid enables Web 2.0 communities to create large computing infrastructure as community effort

Possible options to extend the infrastructure for communities



EDGI: Extending the grid infrastructure for scientific communities

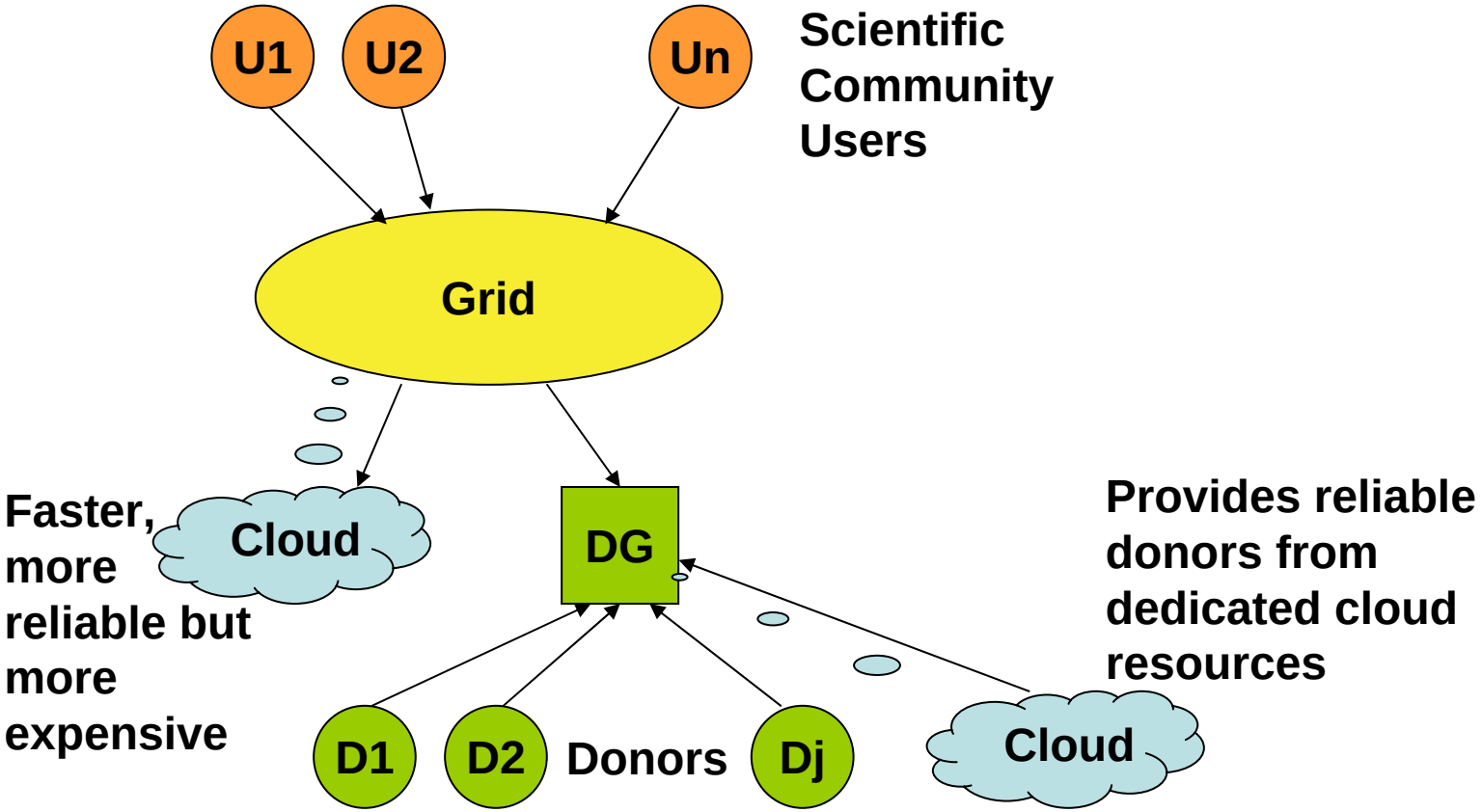
EDGeS scope only for compu intensive applicatior for EGEE (gLite)



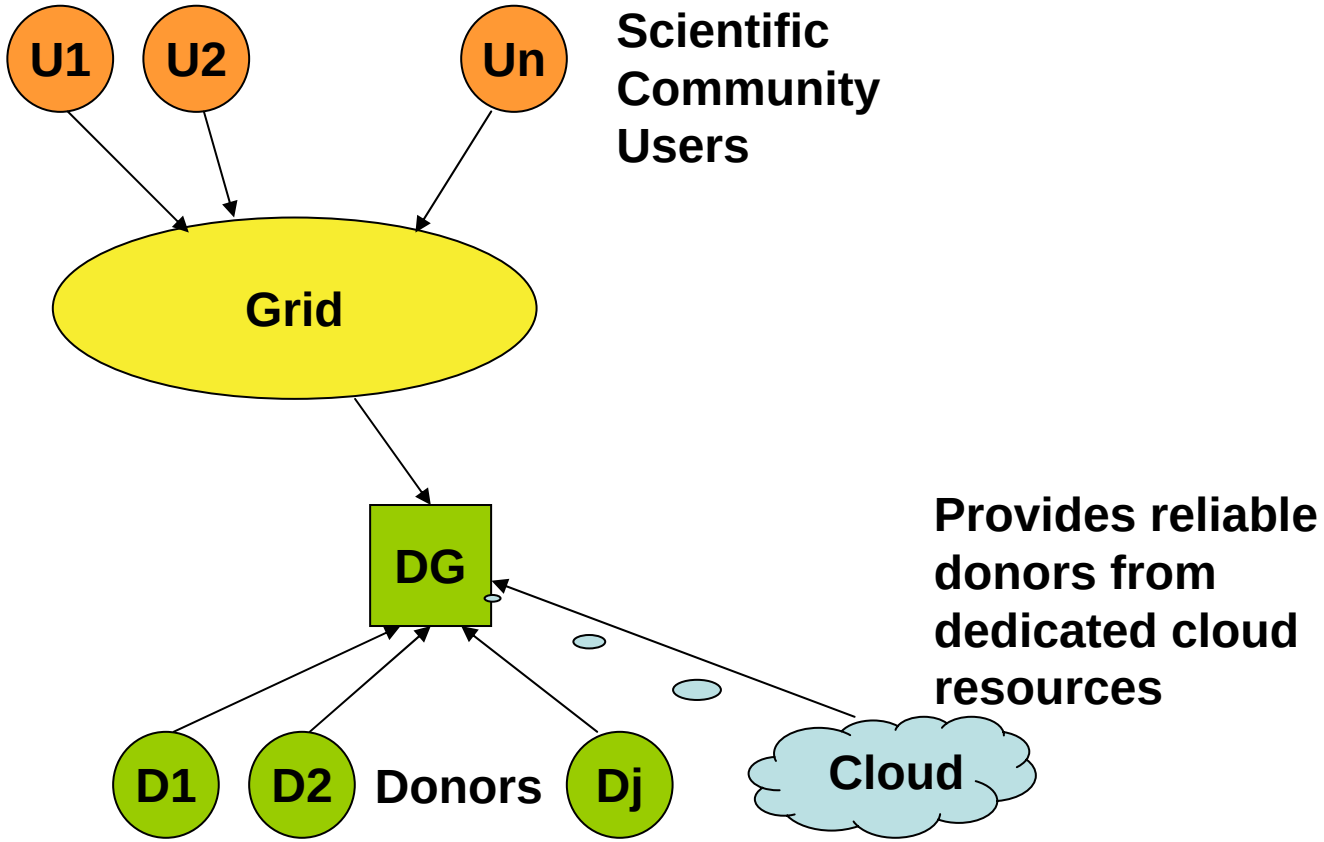
EDGI scope for both compute and data intensive applications for EMI/EGI (gLite, ARC, Unicore)

Extend Desktop Grids with Clouds for QoS

Extending the grid infrastructure for scientific communities

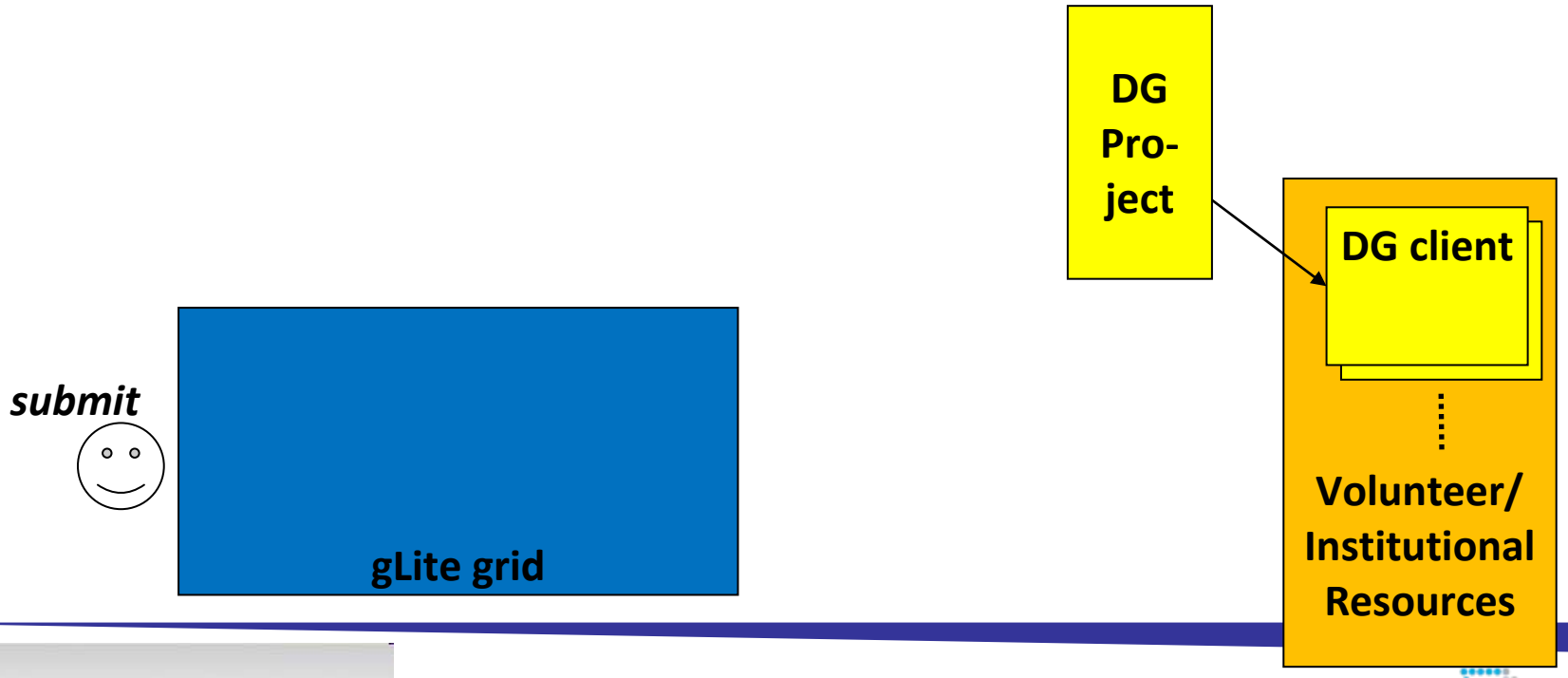


Extending the grid infrastructure for scientific communities



Steps of the technical solution

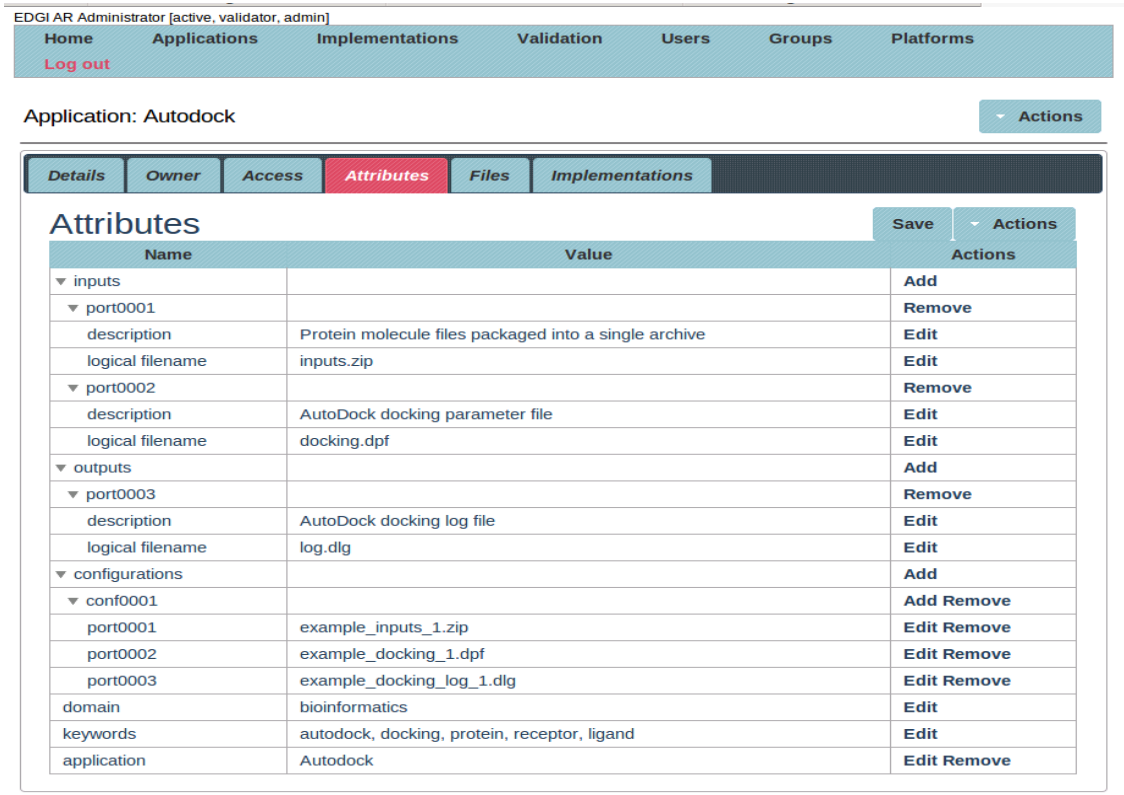
The starting state



Security issues

- SG -> DG direction
  - In DG projects
    - Applications are trusted
    - DG server is trusted
    - Clients are not trusted
  - Therefore any SG application that is intended to run in a DG should be **validated** to make it trusted
  - Therefore EDGI provides:
    - Application validation service
    - **Application repository** (AR) where trusted applications are stored
    - EDGI bridge transfers only those applications to DGs that are stored in the AR

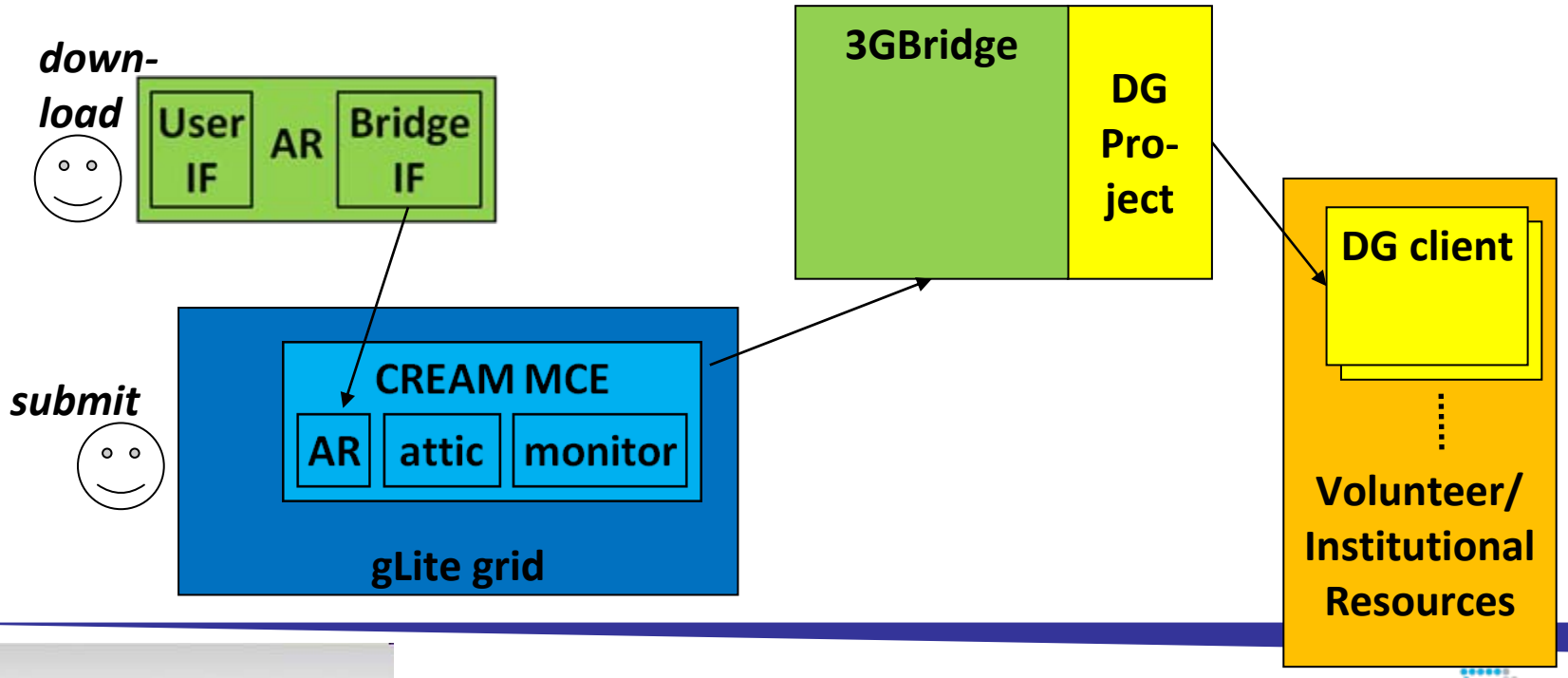
Step 1: Providing application repository



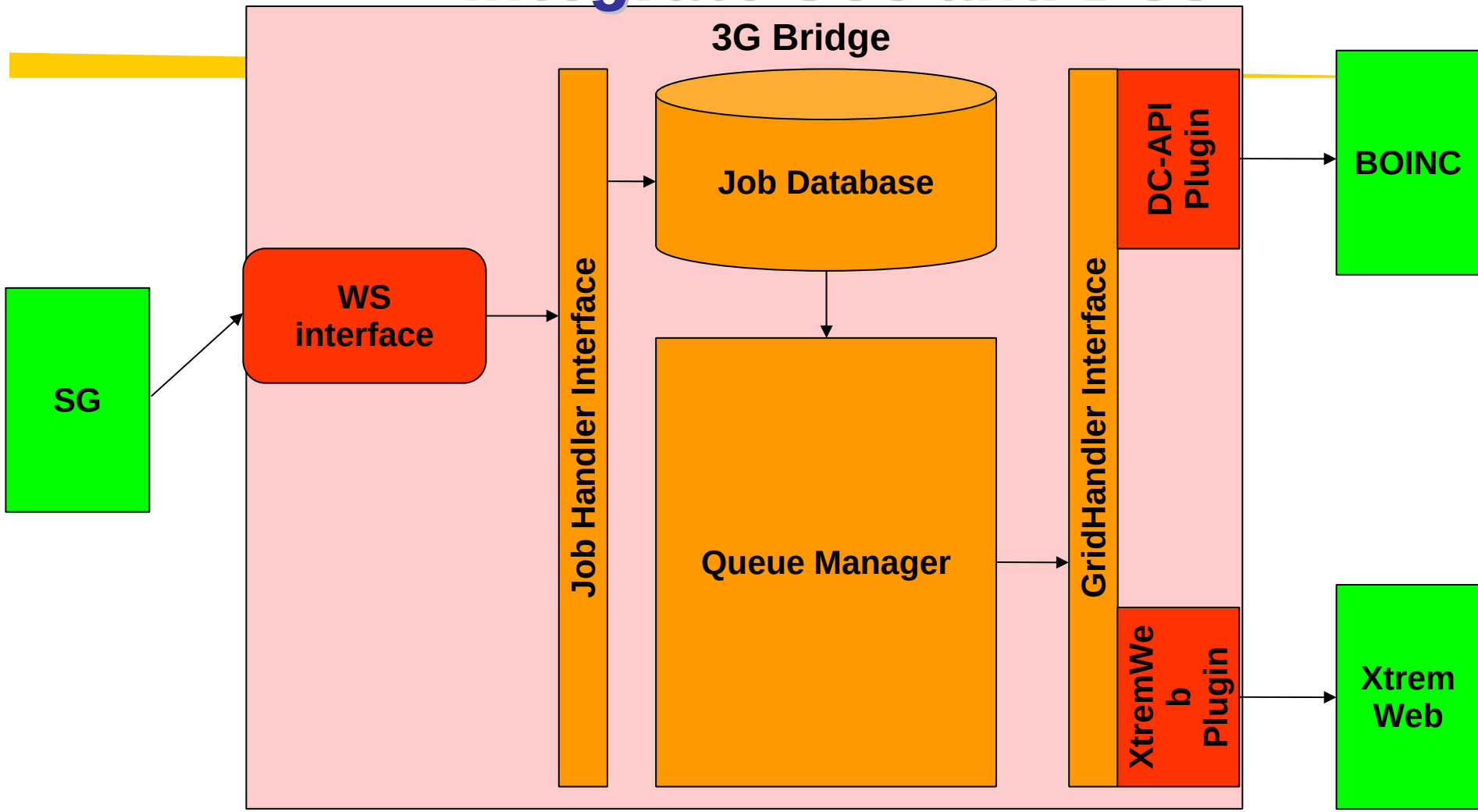
Name	Value	Actions
inputs		Add
port0001		Remove
description	Protein molecule files packaged into a single archive	Edit
logical filename	inputs.zip	Edit
port0002		Remove
description	AutoDock docking parameter file	Edit
logical filename	docking.dpf	Edit
outputs		Add
port0003		Remove
description	AutoDock docking log file	Edit
logical filename	log.dlg	Edit
configurations		Add
conf0001		Add Remove
port0001	example_inputs_1.zip	Edit Remove
port0002	example_docking_1.dpf	Edit Remove
port0003	example_docking_log_1.dlg	Edit Remove
domain	bioinformatics	Edit
keywords	autodock, docking, protein, receptor, ligand	Edit
application	Autodock	Edit Remove

- metadata templates describe applications and their implementations to support
  browse and search operations
- access without registration for all users to browse and search the repository
- API to enable repo access for modified CEs

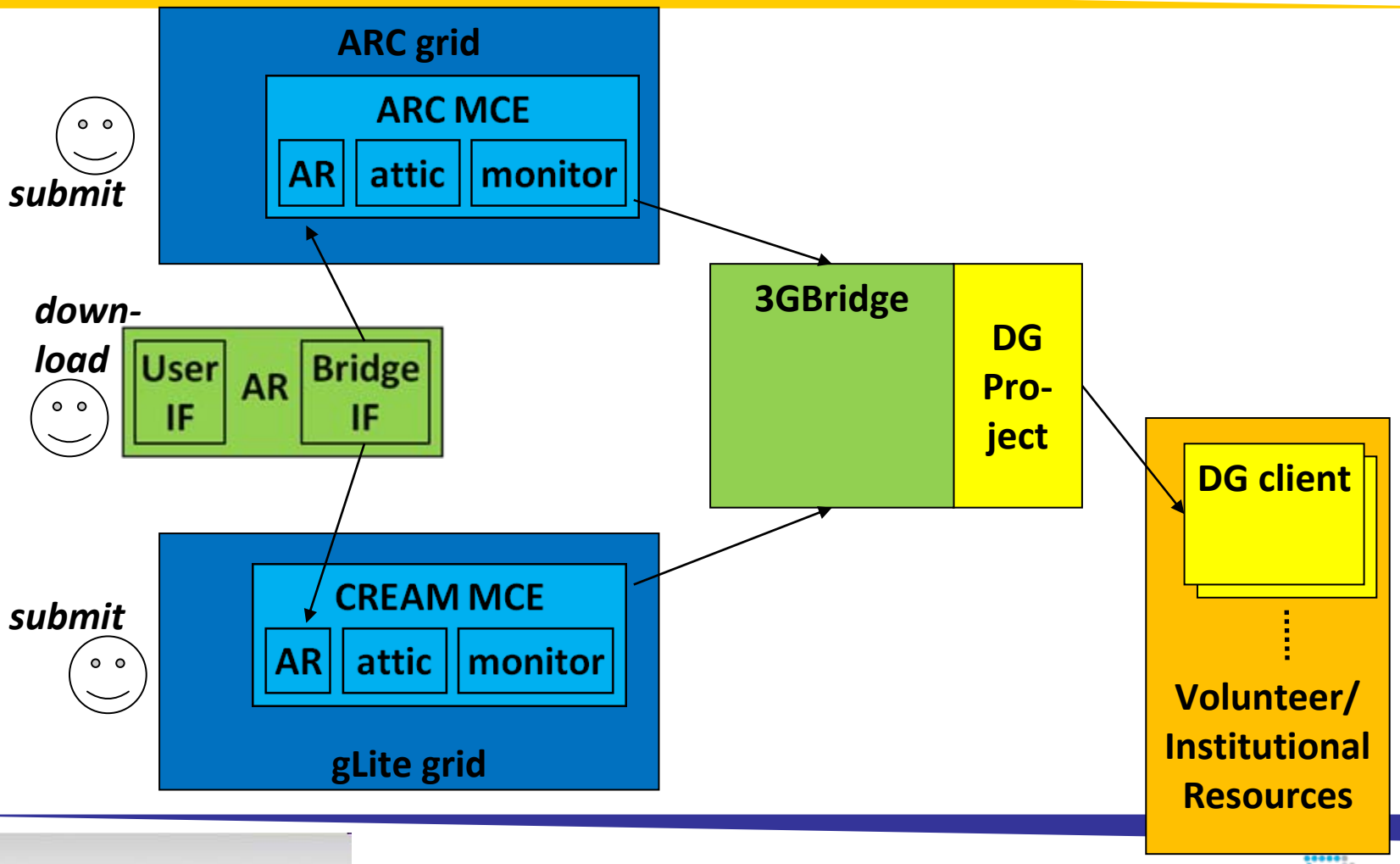
Step 2: Submit job from gLite to DG project



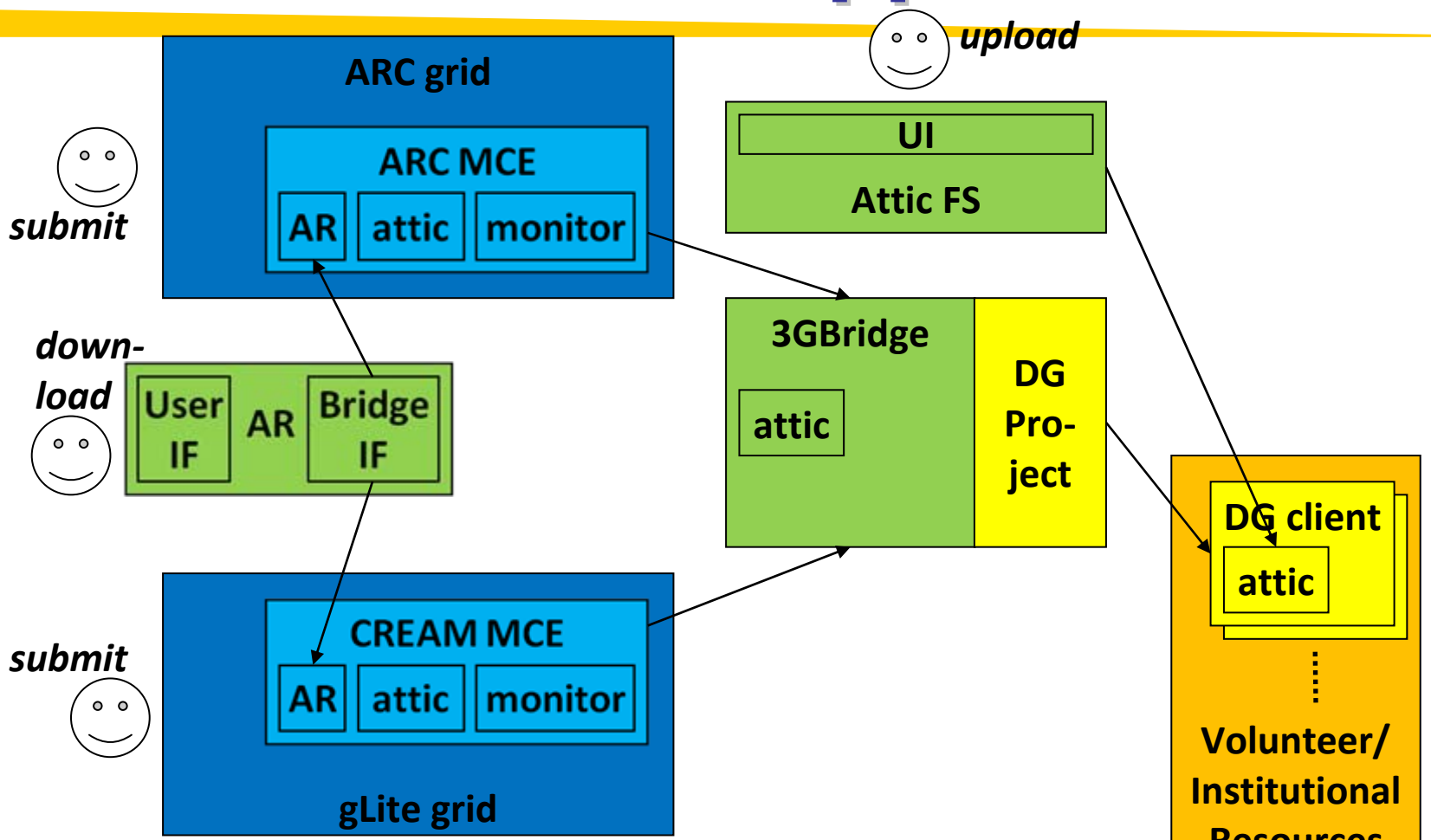
Generic Grid-Grid (3G) Bridge to integrate SGs and DGs



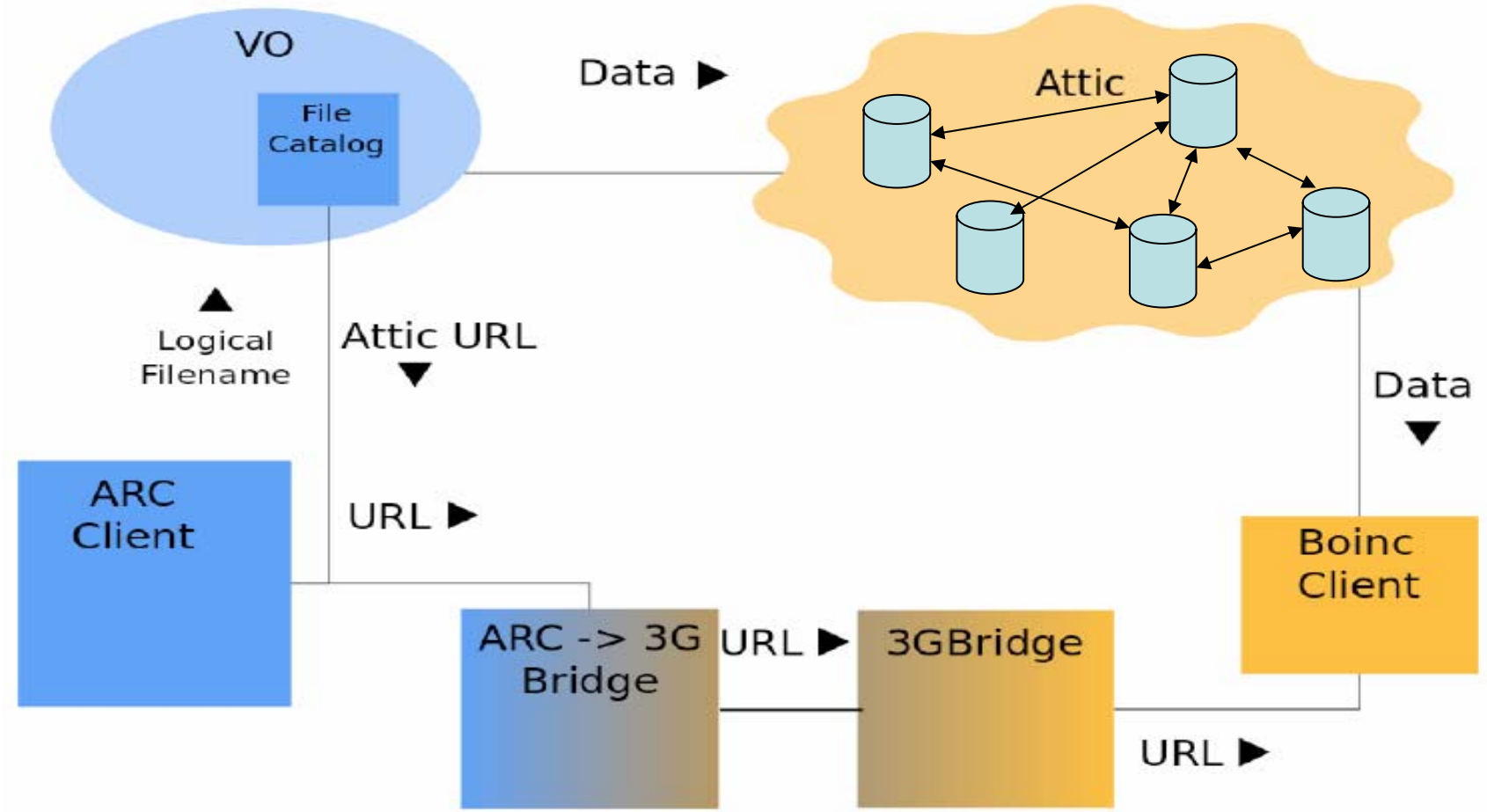
Step 3: Submit job from ARC to DG project



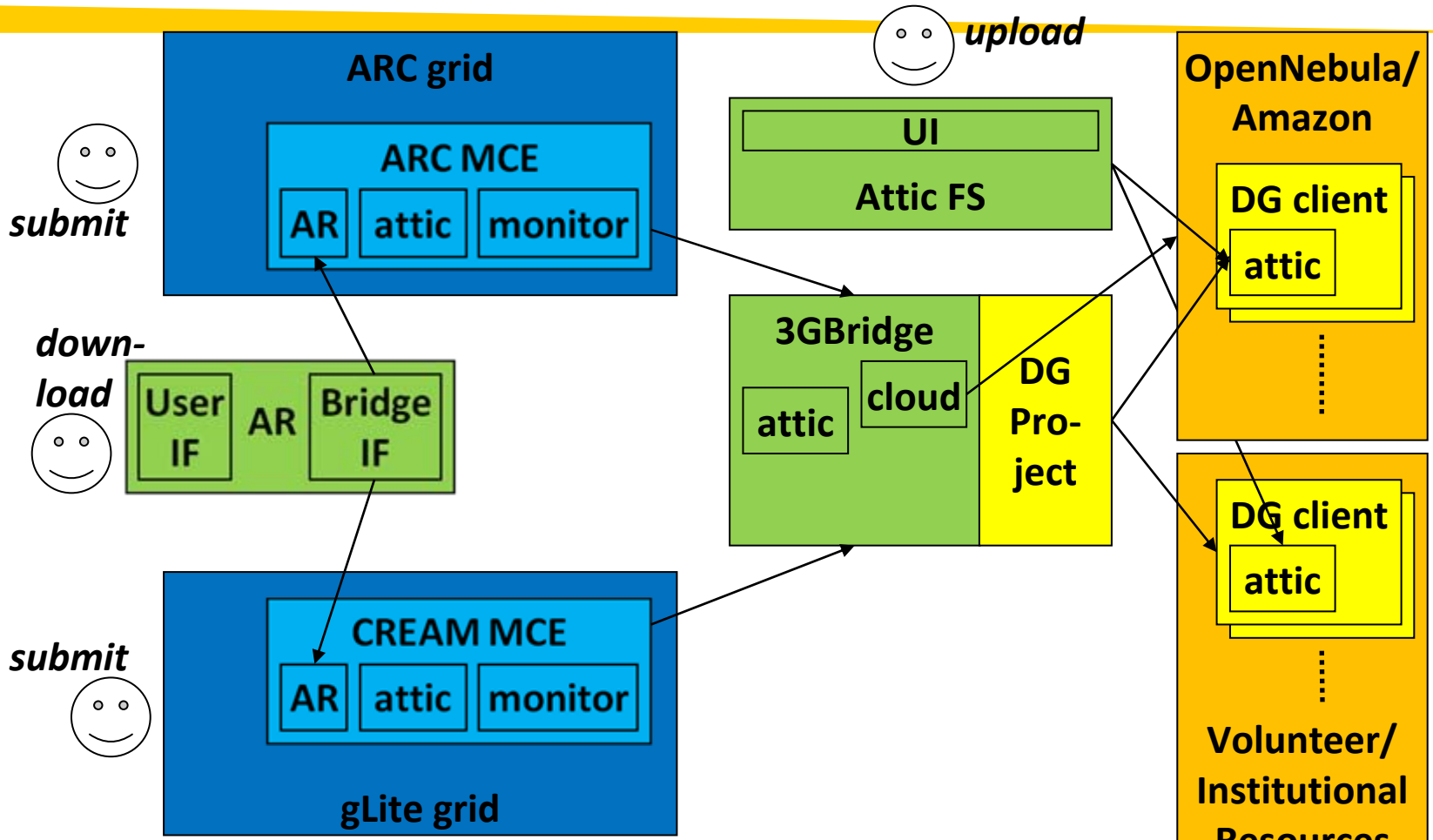
Step 4: Support for Data-intensive applications



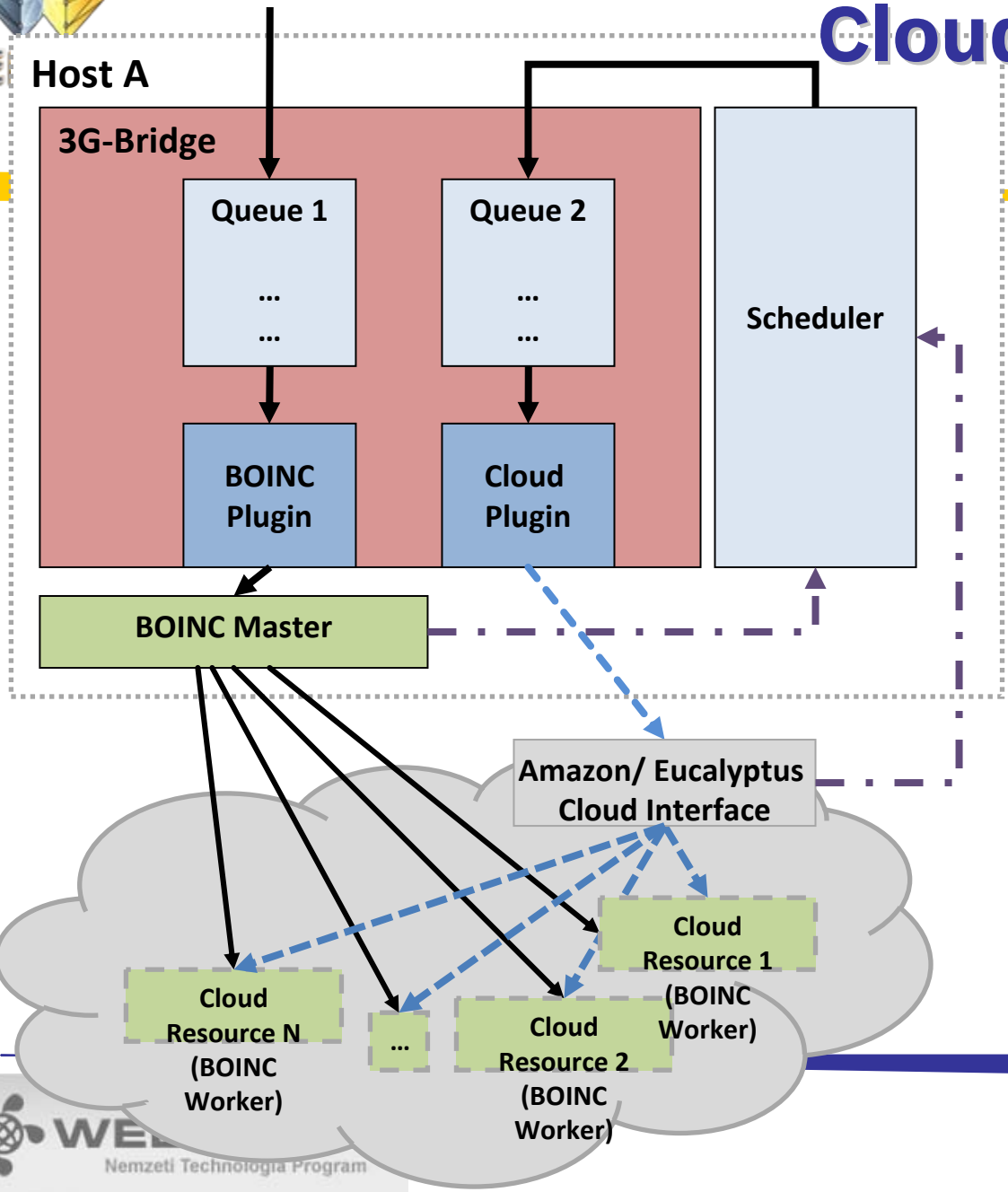
Usage of Attic P2P File System in EDGI



Step 5: QoS Support by Clouds



3G Bridge Extension for supporting Clouds

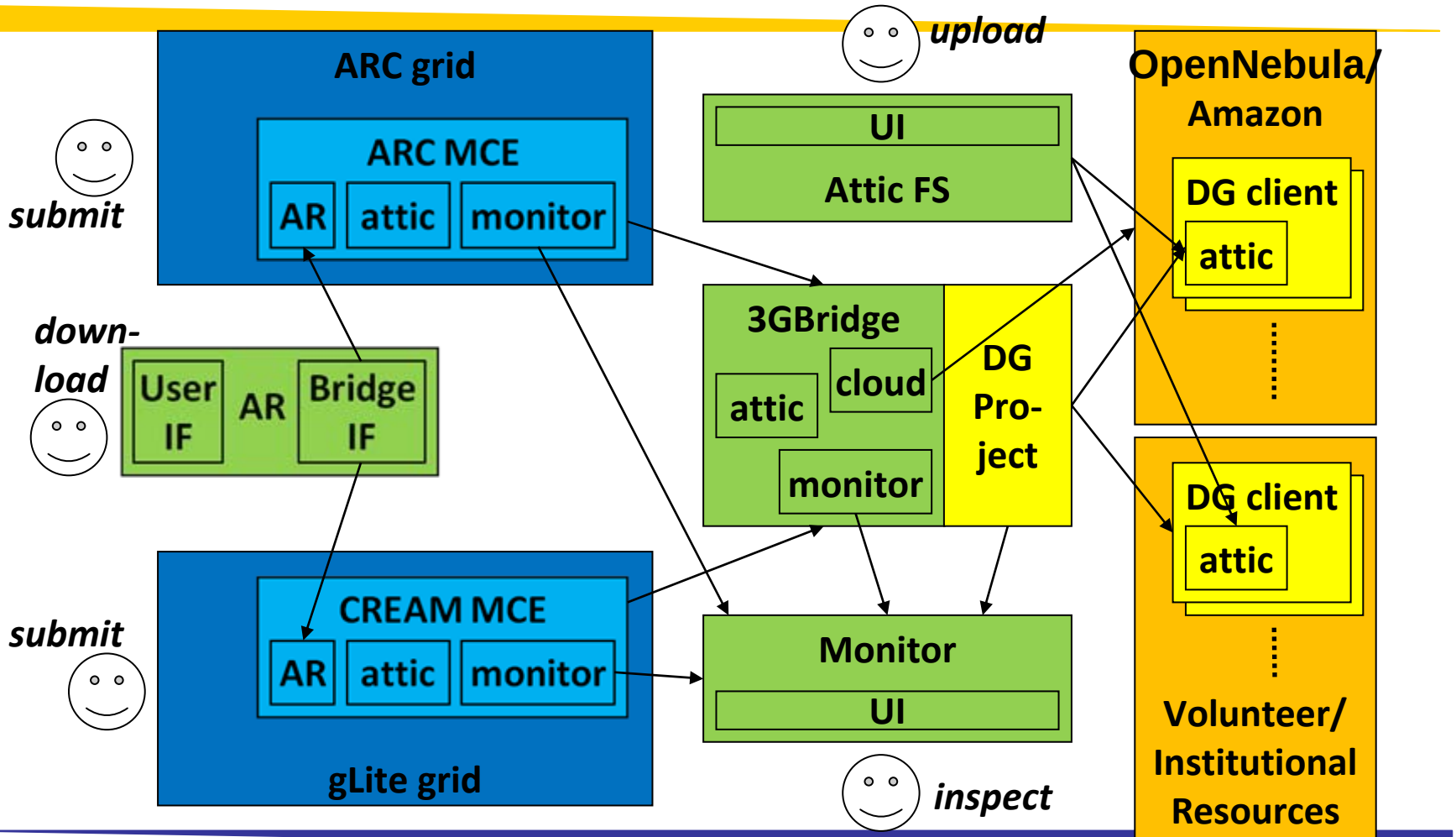


1. Job is submitted from MCE to 3G Bridge.
2. 3G Bridge submits the job to a BOINC DG using the BOINC Plugin
3. The Scheduler keeps track of the number of jobs in the BOINC queue (Queue 1) and of the number of the running Cloud Resources (workers).
4. If the BOINC DG is overloaded, the Scheduler starts new workers by submitting a job to the queue of the Cloud Plugin (Queue 2).
5. Each job in Queue 2 launches a new BOINC Worker in the Cloud.
6. If the cluster is underutilized, the Scheduler stops some workers (cloud resources) by sending cancel jobs into Queue 2.

Legend

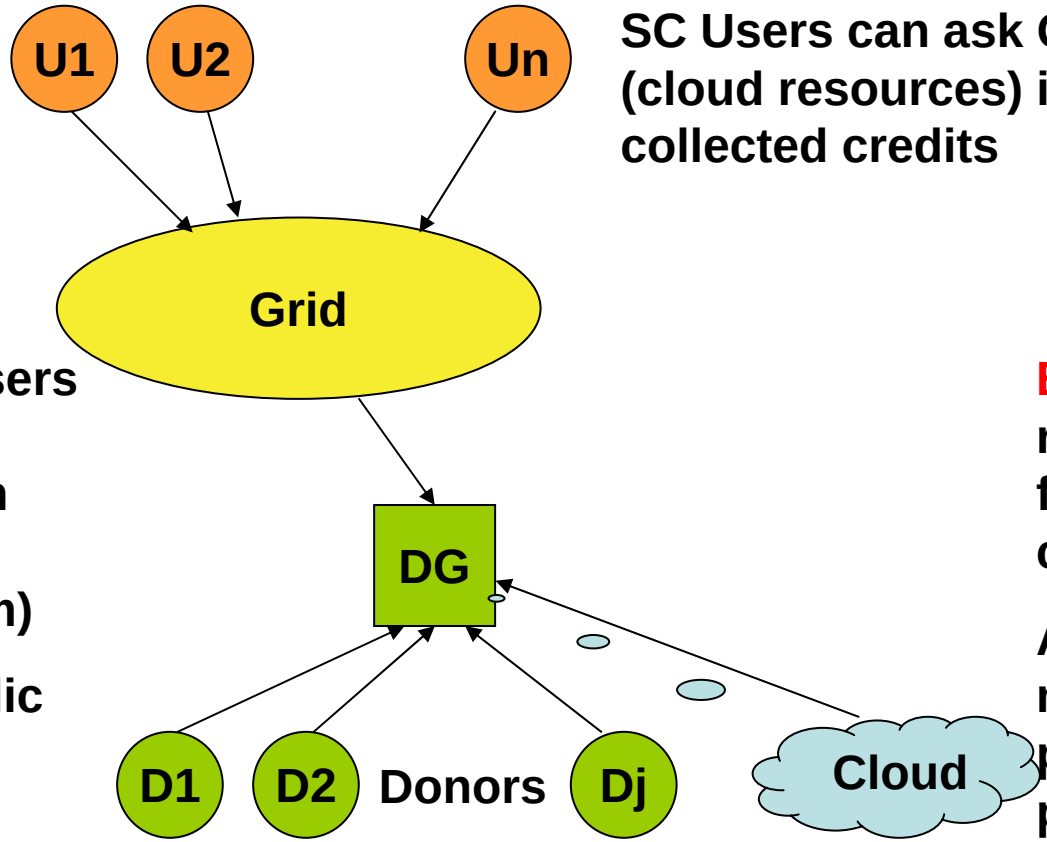
Job
Command
Information

Step 6: Monitoring




Desktop Grid Monitoring Page - Windows Internet Explorer
http://edgi.dei.uc.pt/EDGIMonitoring/faces/DG2DG.jsp
Live Search
File Edit View Favorites Tools Help
Favorites
Suggested Sites
Free Hotmail
Web Slice Gallery
Workshop Program
WebSVN - edgi - Rev 787 - /j...
Desktop Grid Monitoring ...
Page
Safety
Tools
Appl. name
Last day
Last week
Last month
PR
visivo
laserac
fusion
EDGeS@home
GFlops
RRDTOOL / TOBI OETIKER
Sat 00:00
Sat 12:00
Mon
Wed
Fri
Week 21
Week 22
Week 23
Week 24
Week 25
Internet
125%
Start
Meeting_Madeira
Desktop Grid Moni...
Cetraro2011_Kacsuk...
EDGI-JRA1-summary...
EDGID61-main-v6.do...
3:44 PM

EDGI business model



SC Users can ask QoS support (cloud resources) if they have collected credits

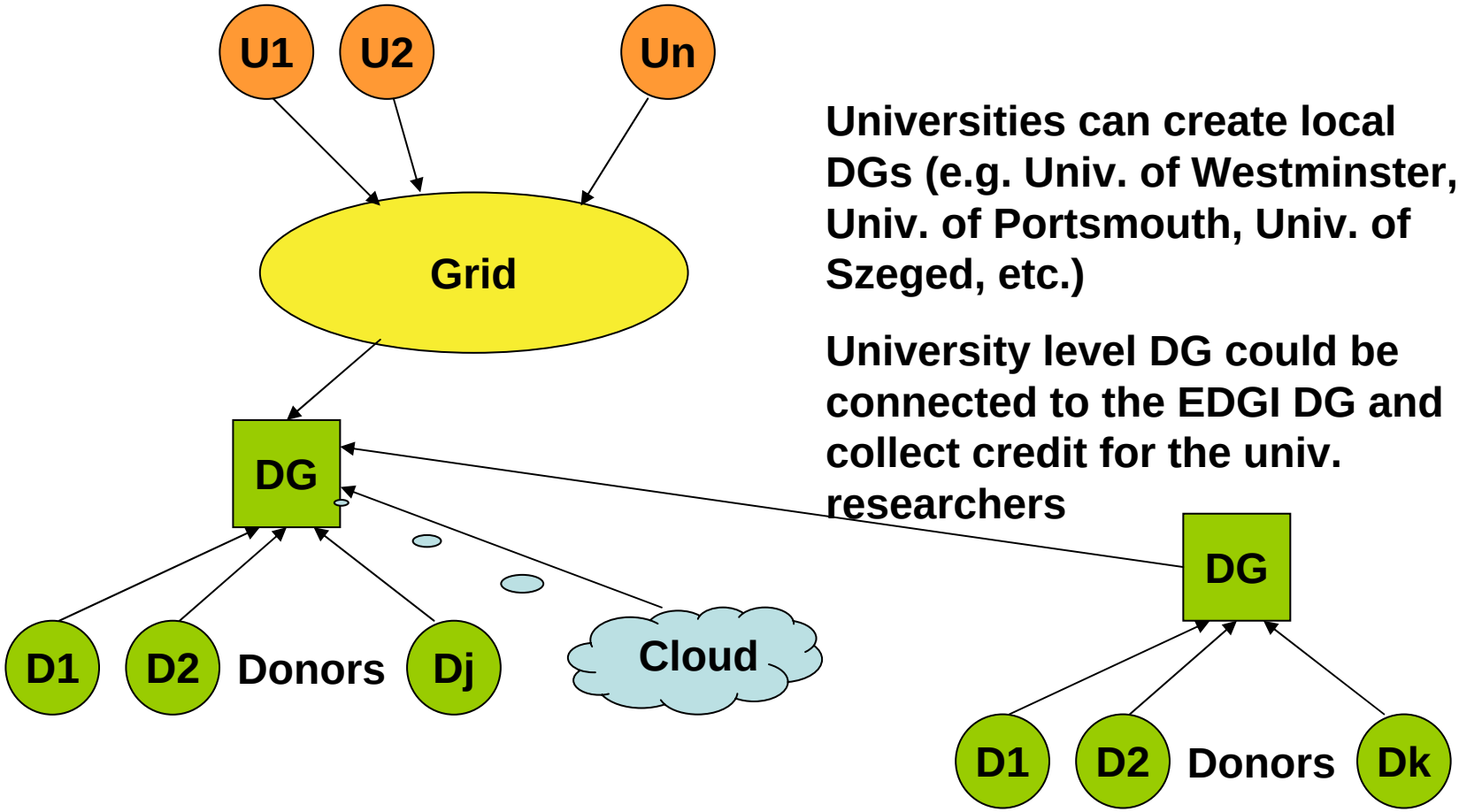
Institutes of SC users can donate DG resources and can collect credits (by BOINC mechanism)

Students and public can support universities by offering their credits

EDGI provides reliable donors from dedicated cloud resources

After EDGI, IDGF members or cloud providers can provide cloud resources

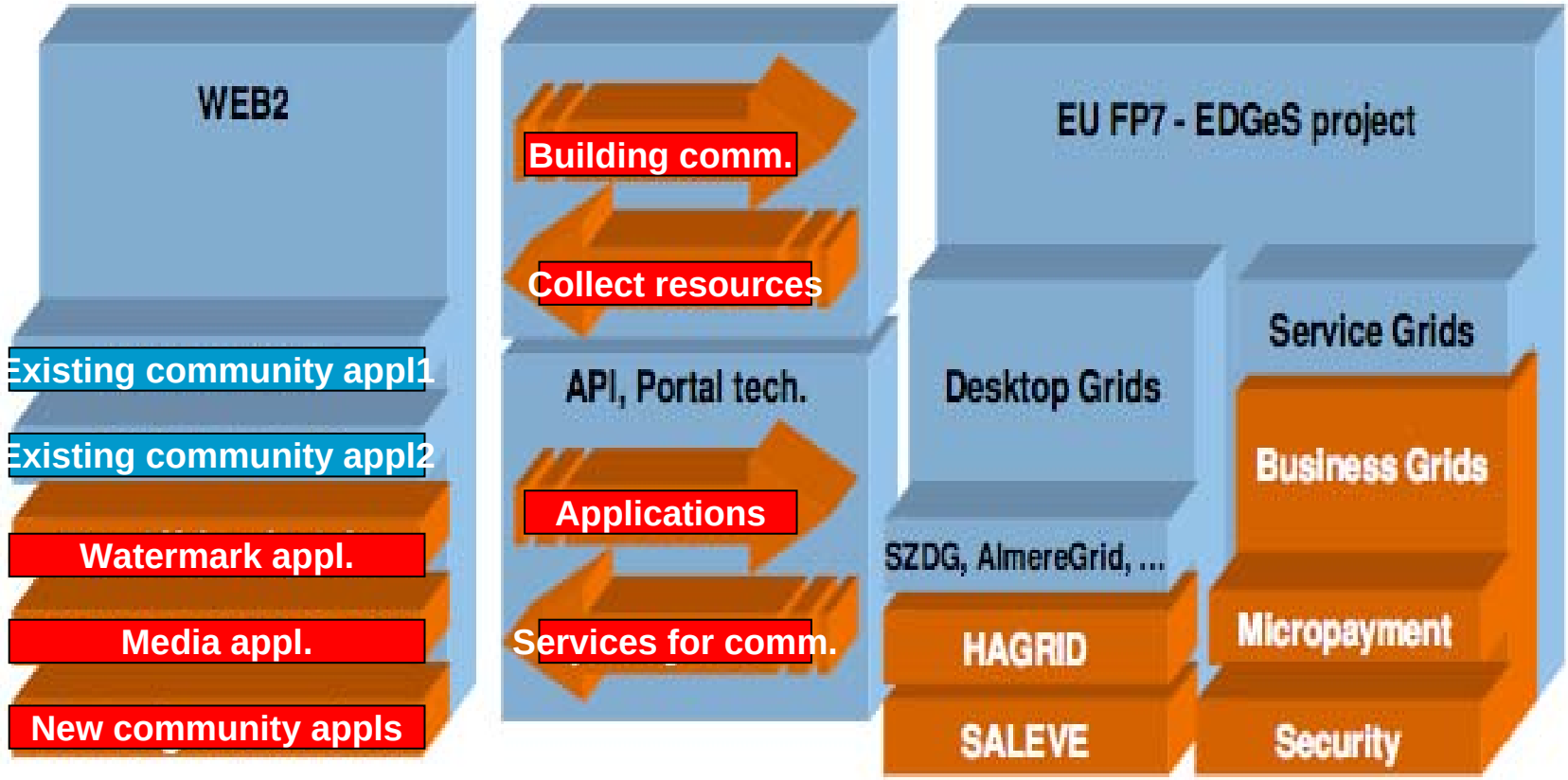
EDGI business model



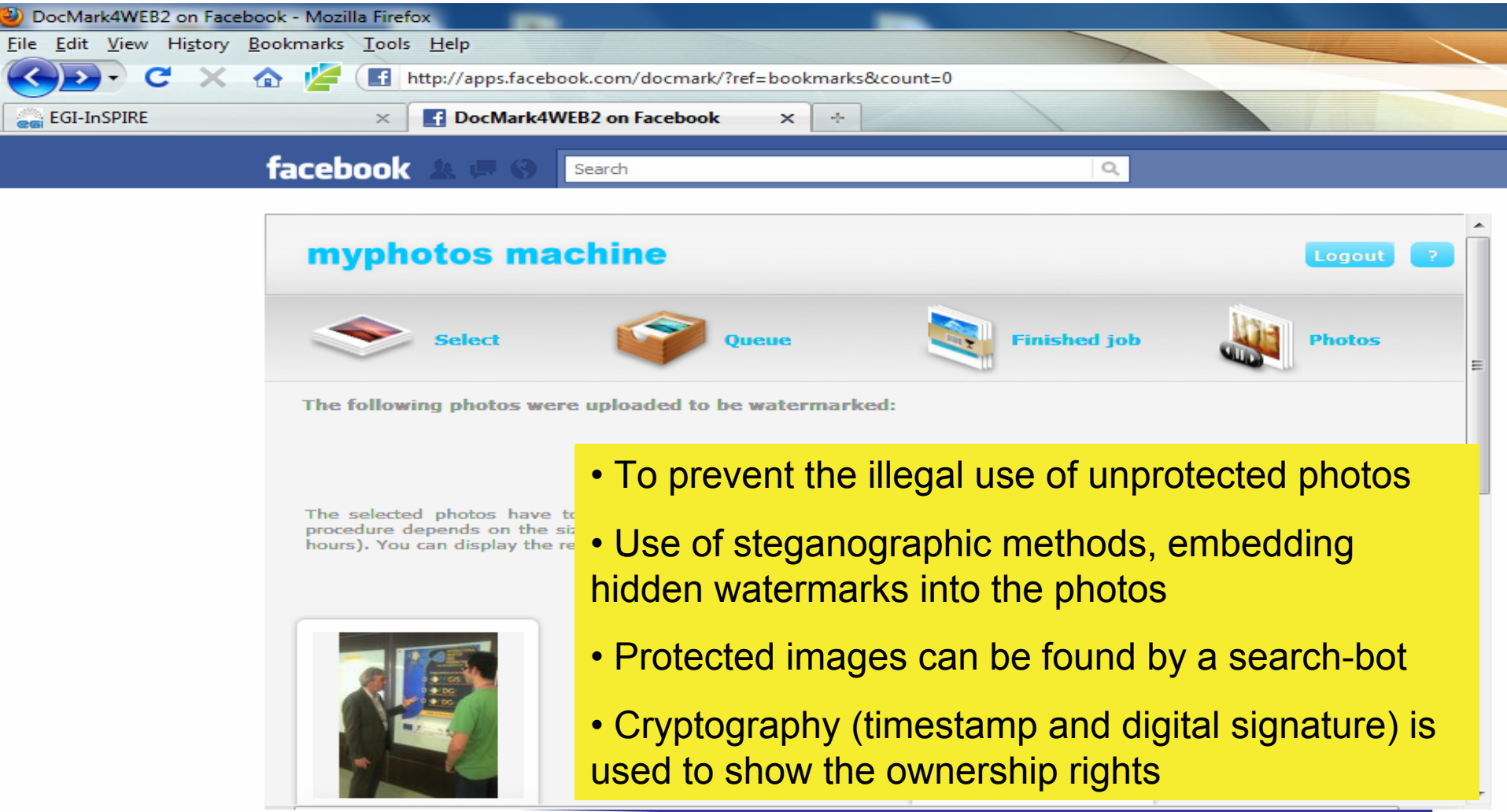
Universities can create local DGs (e.g. Univ. of Westminster, Univ. of Portsmouth, Univ. of Szeged, etc.)

University level DG could be connected to the EDGI DG and collect credit for the univ. researchers

WEB2GRID Project

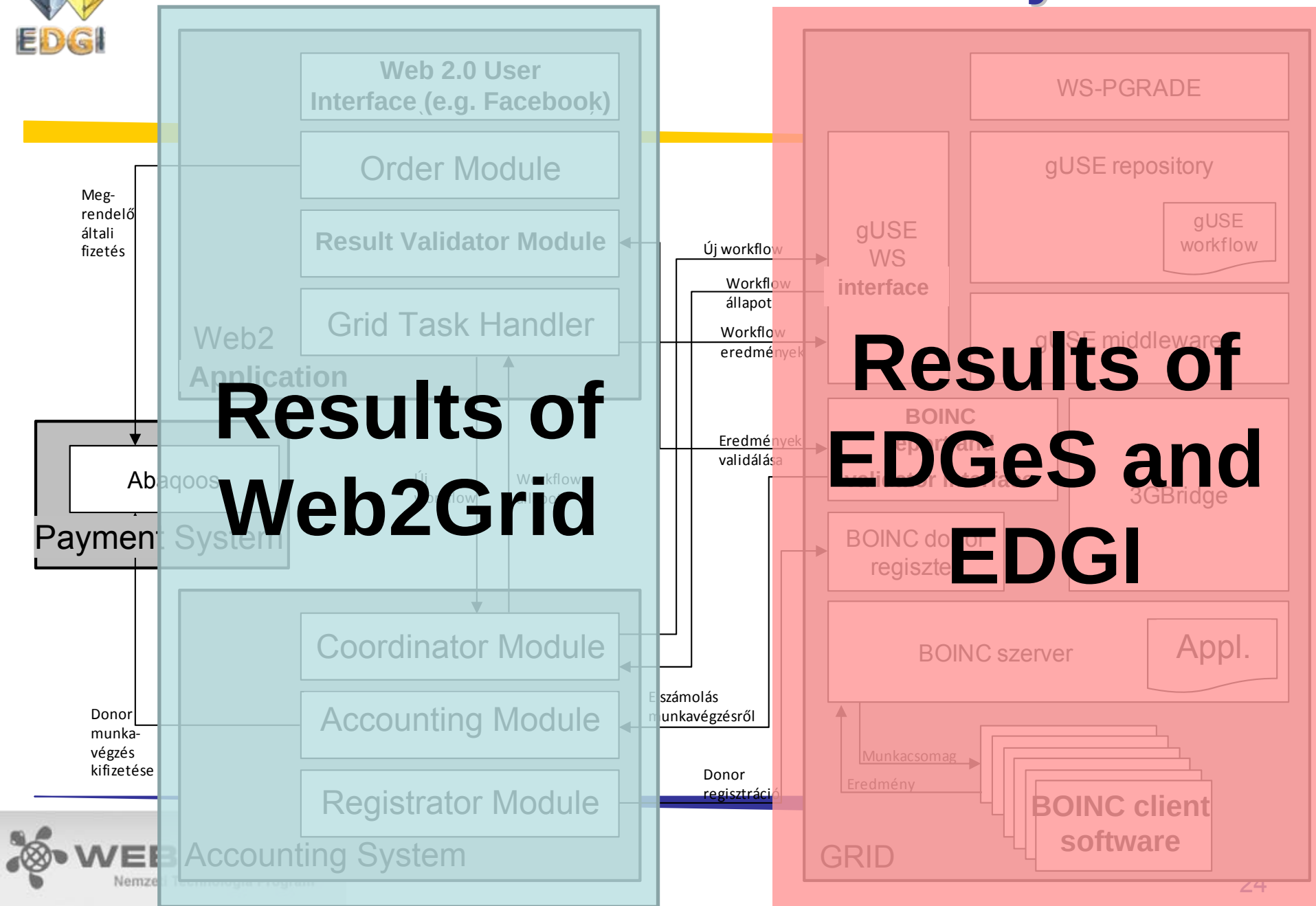


Facebook interface for watermarking photos



- To prevent the illegal use of unprotected photos
- Use of steganographic methods, embedding hidden watermarks into the photos
- Protected images can be found by a search-bot
- Cryptography (timestamp and digital signature) is used to show the ownership rights

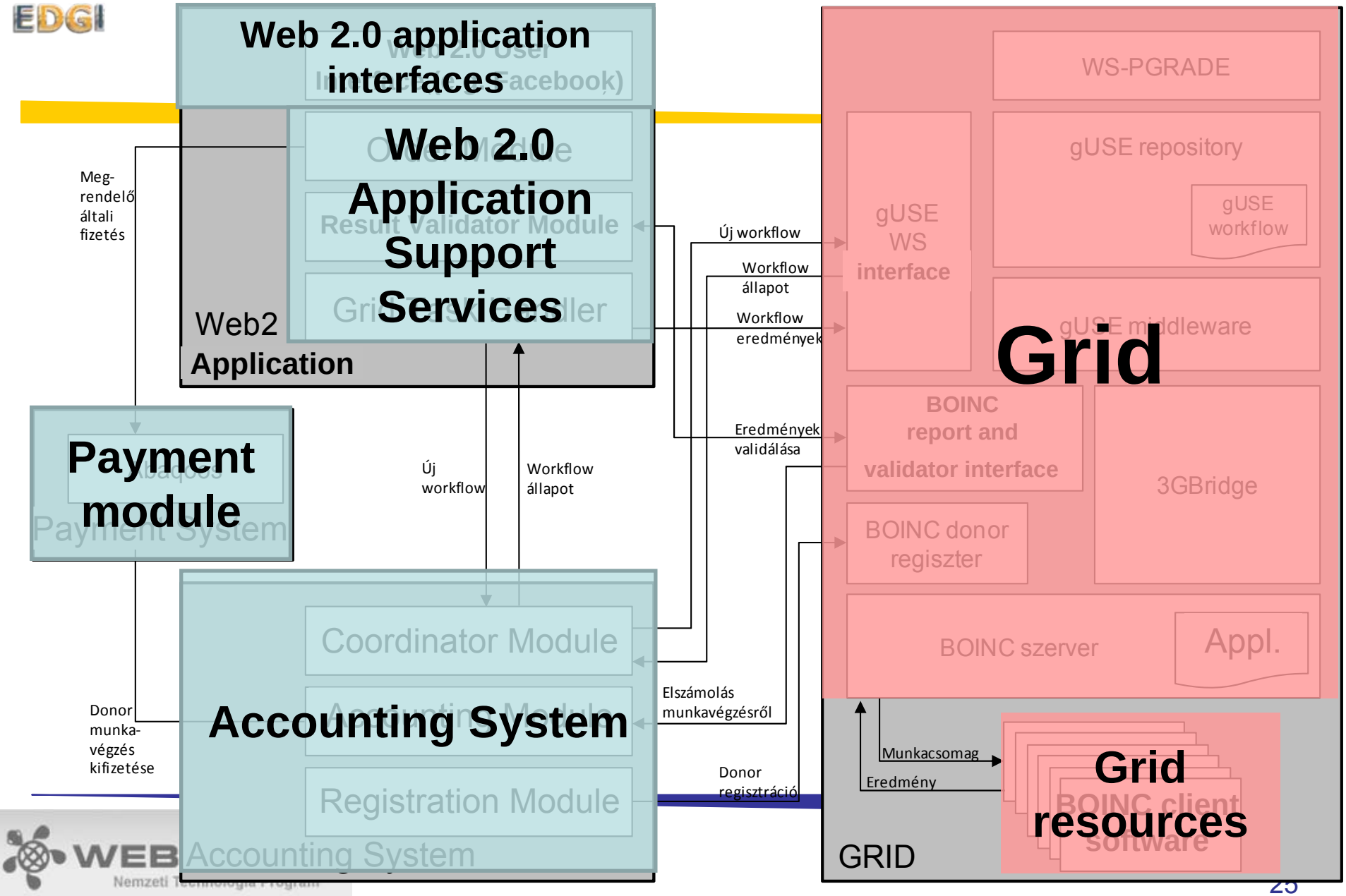
Architecture of Web2Grid system



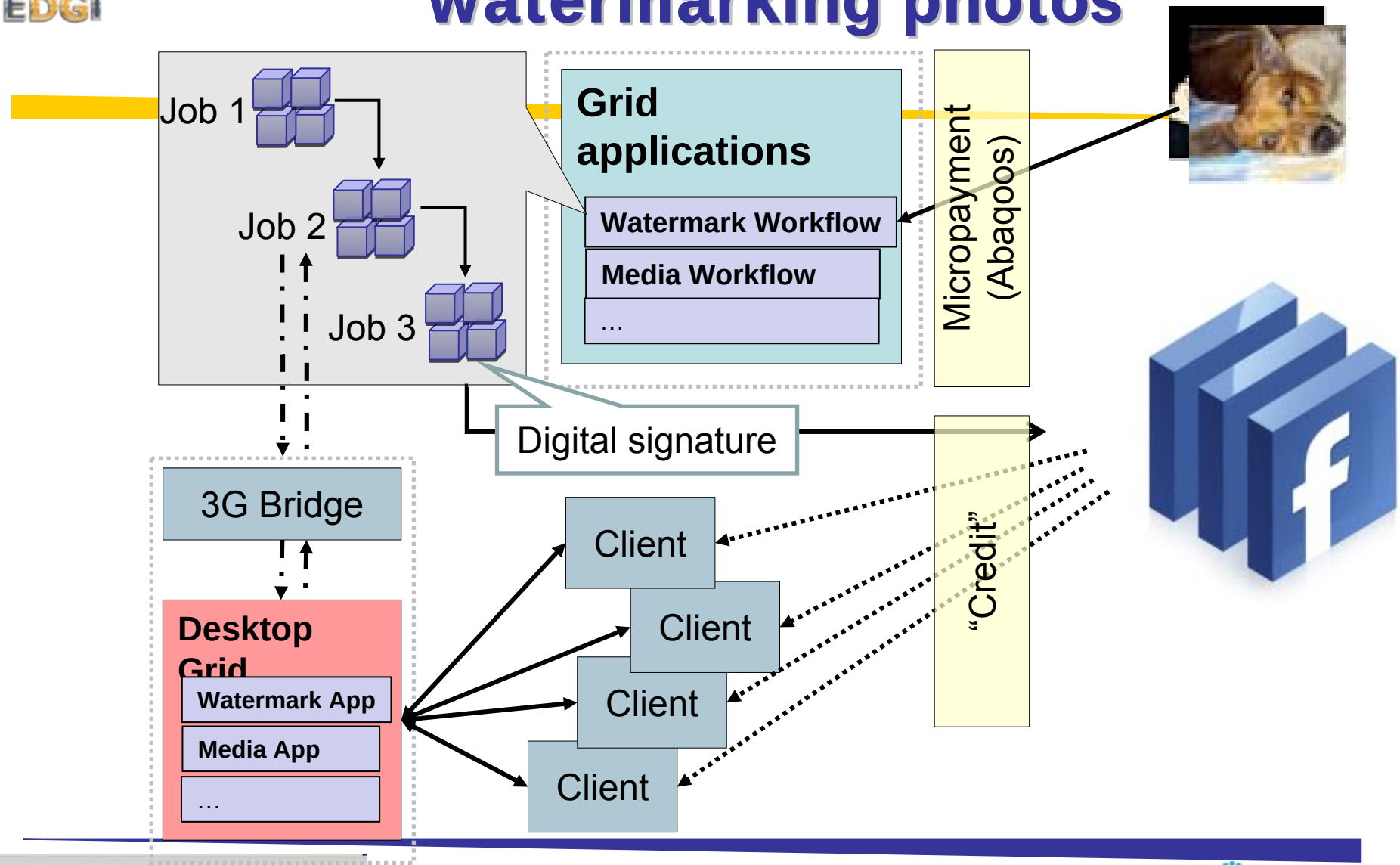
Results of Web2Grid

Results of EDGeS and EDGI

Architecture of Web2Grid system



Example Web2Grid application: watermarking photos



Possible business models

- Service provider provides the whole framework
- Supplier model:
  - Service provider invites trusted individuals to connect clients
  - Pays for the collected credits
- Volunteer model:
  - Web 2.0 community members provide the clients
  - They collect credits
  - Credits are transferred to real money and offered to charity organizations

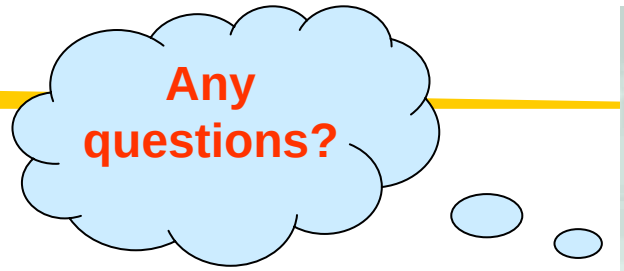
International Desktop Grid Federation (IDGF)

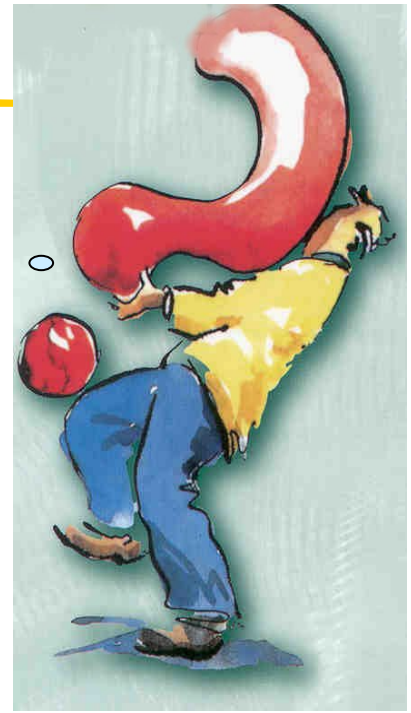
- Support those NGIs that want to extend their VO with volunteer or local DGs
- Support those universities, institutes, companies who want to set up local DG
- Run dissemination campaign to attract more volunteers to donate their computer resources
- Federation is open for organisations and individuals in research and industry. Current members:
  - >100 individuals
  - > 30 organizations (including companies)

Summary

- Volunteer DG resources can be provided for scientific and Web 2.0 communities as an inexpensive alternative to cloud resources
- EDGI and Web2Grid projects have developed the required technical solutions
- They also show possible business models
- Future task: to implement the business models

Thank you for your attention ...





For more information please visit the EDGeS and EDGI Websites:

http://www.edges-grid.eu/

http://edgi-project.eu

and/or send e-mail to me:

kacsuk@sztaki.hu