Enable the use of BOINC for new communities - the IDGF experience

Hannover, 17.08.2011

Peter Kacsuk, MTA SZTAKI, kacsuk@sztaki.hu

EDGI is supported by the FP7 Capacities Programme under contract nr RI-261556.
IDGF (International Desktop Grid Federation) objectives

- Build a community of those people who are interested in any form of desktop grids including
  - Different technology developers (BOINC, Condor, XtremWeb, OurGrid, SZDG, etc.)
  - Desktop grid operators and system admins
  - Application developers who want to port applications to DG systems
  - End-users who want to run applications on DG systems
  - Service grid providers who would like to extend their resources with DG resources
  - Scientists who want to build campus grid
The history of establishment of IDGF

**EDGeS**
- **DG<->SG integration:**
  - gLite → BOINC, XtremWeb
  - BOINC, XtremWeb → gLite
- **Compute intensive applications**

2008 - 2010

**EDGI**
- **ARC, Unicore, Clouds**
- **QoS with Clouds**
- **Data intensive apps**
- **SG->DG direction support**

**DEGISCO**
- **Disseminates and supports the use of EDGeS results worldwide**
- **Green IT aspects**

2010 - 2012

Joint establishment of IDGF
Goals of these EU DG projects and IDGF

To enable new and mass usage scenarios for DG systems

- Extend service grids with volunteer and institutional desktop grids (EDGeS, EDGI)
- Extend DGs with cloud resources on-demand
- Extend desktop grids with service grids (EDGeS, DEGISCO)
- Enable the flexible use of a DG system by large user communities (portal access to BOINC and SZDG)
Extend service grids with volunteer and institutional desktop grids (EDGeS, EDGI)
Scope of EDGeS, DEGISCO and EDGI

EDGeS and DEGISCO scope: compute intensive applications for gLite

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

Extend Desktop Grids with Clouds for QoS
Typical current business model for volunteer desktop grids

BOINC project runs 1 appl that runs for years

DG project administrator defines the input data to be processed by the appl
EDGI business model for extending SG VOs with desktop grids

• BOINC project runs several apps

• An appl runs if a user gives input data

• Many end-users can access the project via a service grid VO UI

• end-users define the input data to be processed by the apps
Extend DGs with cloud resources on-demand (EDGI)
EDGI business model for volunteer desktop grids

- BOINC project runs several apps
- An app runs if a user gives input data
- Many end-users can access the project via a service grid VO UI
- End-users define the input data to be processed by the apps

EDGI provides cloud resources for QoS
Architecture of extending DGs with cloud resources on-demand

See details in:
Motivation for QoS

Desktop Grid
- Variable amount of resources, volatility, unpredictability, unannounced departure.

Low QoS compare to classical DCI
- Tail Effect
- We define QoS as a level of confidence in Bag of Task (BoT) execution:

Question: how do we provide QoS to users given the dynamism and volatility of the computing resources?
- Intrinsic approach: improve DG scheduler for QoS ability
- Extrinsic approach: provide additional dedicated computing resources

Solution:
- Providing cloud resources for the tail phase
- SpeQuloS
SpeQuloS: a middleware for QoS to Desktop Grids
EDGI architecture extended with clouds
Current production solution of integrating BOINC DG with OpenStack based on 3G Bridge: Wmin local DG

1. New Cavendish Street 576 nodes
2. Marylebone Campus 559 nodes
3. Regent Street 395 nodes
4. Wells Street 31 nodes
5. Little Tichfield Street 66 nodes
6. Harrow Campus 254 nodes

Total: 1881 nodes
Attracting more participants for volunteer desktop grids

- EDGI introduces a new architecture where desktop grids (BOINC, XtremWeb) can be extended with **cloud resources** for QoS
- However, in order to get cloud resources the user needs credits
- In order to collect credit either the user or her organization should **provide resources for volunteer DGs belonging to IDGF**
- If you joined to IDGF those who would like to earn more credit for cloud resources will support your DG system
Enable the flexible use of a DG system by portal access (EDGeS, EDGI)
Enable the flexible use of a DG system

- BOINC project runs several apps
- An app runs if a user gives input data

- We provide a service that can be used by many scientists via a web interface (e.g. WS-PGRADE portal)
- end-users define the input data to be processed by the apps (e.g. renderfarm.fi)
Extend desktop grids with service grids (EDGeS, DEGISCO)
Extending desktop grids with SG VO resources (DEGISCO business model)

BOINC project runs 1 appl that runs for years

DG project administrator defines the input data to be processed by the appl

The SG VO maintained by DEGISCO (and later by IDGF) provides over 2000 computers to support a VDG
# Desktop Grid VO Statistics

<table>
<thead>
<tr>
<th>Name</th>
<th>CPUs</th>
<th>Online Storage Space (GB)</th>
<th>Grid Jobs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Physical</td>
<td>Logical</td>
<td>SI2000</td>
</tr>
<tr>
<td>AEGIS01-IPB-SCL</td>
<td>176</td>
<td>704</td>
<td>1,689,600</td>
</tr>
<tr>
<td>BIFI</td>
<td>64</td>
<td>384</td>
<td>523,392</td>
</tr>
<tr>
<td>GRIF</td>
<td>1,602</td>
<td>7,411</td>
<td>15,590,608</td>
</tr>
<tr>
<td>MY-UPM-BIRUNI-01</td>
<td>86</td>
<td>344</td>
<td>4,851,088</td>
</tr>
<tr>
<td>NCG-INGRID-PT</td>
<td>312</td>
<td>1,248</td>
<td>2,131,584</td>
</tr>
<tr>
<td>RU-ISA-CGTDC</td>
<td>4</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>SZTAKI</td>
<td>14</td>
<td>34</td>
<td>44,288</td>
</tr>
<tr>
<td>TW-eScience</td>
<td>2</td>
<td>8</td>
<td>21,520</td>
</tr>
<tr>
<td>UA-KNU</td>
<td>24</td>
<td>80</td>
<td>164,000</td>
</tr>
<tr>
<td>UFCG-LSD</td>
<td>8</td>
<td>1</td>
<td>1,323</td>
</tr>
<tr>
<td>UFRJ-IF</td>
<td>62</td>
<td>244</td>
<td>478,012</td>
</tr>
<tr>
<td>Total</td>
<td>2,354</td>
<td>10,474</td>
<td>25,495,415</td>
</tr>
</tbody>
</table>
Benefit for new volunteer BOINC systems

- It takes several years until a new volunteer BOINC system can collect substantial number of resources
- IDGF will give access to the large set of SG VO resources that were created in the EDGeS and later in the DEGISCO project
- It contains over 2000 computers
<table>
<thead>
<tr>
<th>Computer ID</th>
<th>Rank</th>
<th>Owner</th>
<th>Recent average credit</th>
<th>Total credit</th>
<th>CPU type</th>
<th>Operating system</th>
</tr>
</thead>
<tbody>
<tr>
<td>55228</td>
<td>1</td>
<td>EDGeS User</td>
<td>27,952.72</td>
<td>1,517,720.99</td>
<td>GenuineIntel, Intel(R) Xeon(R) CPU E5420 @ 2.50GHz [Family 6 Model 23 Stepping 6]</td>
<td>Linux, 2.6.18.8-xen-3.2.1-2mdv</td>
</tr>
<tr>
<td>33460</td>
<td>2</td>
<td>Beyond</td>
<td>21,688.39</td>
<td>10,805,735.19</td>
<td>AuthenticAMD, AMD Phenom(tm) II X6 1055T Processor [Family 16 Model 10 Stepping 0]</td>
<td>Microsoft Windows 7 Ultimate x64 Edition, Service Pack 1, (06.01.7601.00)</td>
</tr>
<tr>
<td>45366</td>
<td>3</td>
<td>Anonymous</td>
<td>17,209.57</td>
<td>4,730,666.92</td>
<td>AMD Phenom(tm) II X6 1090T Processor [Family 16 Model 10 Stepping 0]</td>
<td>Microsoft Windows 7 x64 Edition, Service Pack 1, (06.01.7601.00)</td>
</tr>
<tr>
<td>48437</td>
<td>4</td>
<td>ritterm</td>
<td>14,540.89</td>
<td>912,200.28</td>
<td>AMD Phenom(tm) II X6 1090T Processor [Family 16 Model 10 Stepping 0]</td>
<td>Microsoft Windows 7 Home Premium x64 Edition, Service Pack 1, (06.01.7601.00)</td>
</tr>
<tr>
<td>55604</td>
<td>5</td>
<td>SazanEyes</td>
<td>14,287.79</td>
<td>533,547.39</td>
<td>GenuineIntel Genuine Intel(R) CPU 000 @ 2.27GHz [Family 6 Model 44 Stepping 0]</td>
<td>Microsoft Windows 7 Home Premium x64 Edition, Service Pack 1, (06.01.7601.00)</td>
</tr>
</tbody>
</table>
Examples of volunteer BOINC systems supported by the DEGISCO business model

- SLinCA@home (Ukraina)
- Yoyo@home (Germany)
- IberCivis (Spain and Portugal)
- CAS@home (China)
- Optima@home (Russia)
- Poem@home (Germany)
- Univ. of Westminster Campus Grid (UK)
IDGF
International Desktop Grid Federation
Get trained

Next IDGF training sessions and tutorials
- Tutorials & workshops

IDGF Desktop Grid training programme
- IDGF training course overview

Other training opportunities

Find documentation

Introductory documentation
- Road map: Desktop Grids for eScience

Technical documentation
- Technical Wiki (for IDGF members only)
- Download section (slides, documents)

Connect

You can connect your computer or infrastructure to (other) Desktop Grids or eScience infrastructures. Looking to run applications? Could be the one you want to use is already ported and certified.

Infrastructures
- Infrastructures

Available ported application
- Applications ported and available

Get certified

IDGF is working on a certification service for:
- Consultants
- Trainers
- Applications

Find experts

Chapters and working groups
- Expert groups
- Regional chapters: European Chapter.
Benefit for technology and application developer organizations of the European Chapter

- IDGF will investigate the possibility of participating in FP7 and FP8 EU projects
- If this is possible, to take IDGF as partner in EU projects will be attractive since IDGF will represent a large community (see the HealthGrid experience)
- Then the work allocated for IDGF in these projects would be distributed among the member organizations of the European Chapter
Conclusions

- Joining IDGF is worth for both SG and DG community members
- DG community members can extend their DG with SG resources
- SG community members can extend their SG VO with DG resources
- Both communities can get application porting support
- IDGF can significantly impact the community and can attract additional financial resources