Extending gLite VOs with volunteer and institutional BOINC-based desktop grids to execute parameter sweep applications

P. Kacsuk
MTA SZTAKI
European Year of Volunteering 2011

Volunteer! Make a difference

Presentation for: ISGC2011
Author: Peter Kacsuk

Source: http://europa.eu/volunteering
First great achievement of volunteer computing

- **SETI@home**: analyzes data from the Arecibo radio telescope, looking for Extra-Terrestrial Intelligence in radio signals.
  - 3.8M volunteers from 226 countries (2004)
  - 1200 CPU years/day
  - 38 TFlops sustained performance (Japanese Earth Simulator 32 TFlops)
Potential of volunteer computing

- Number of privately owned PCs:
  - currently 1.5 billion
  - grow to 2 billion by 2015
- ~1 million computers are actively participating in volunteer computing → supplying ~10PetaFLOPS of computing power
- Fastest supercomputers supply few PetaFLOPS
- Largest grids (e.g. EGI) have several hundreds of thousands of hosts.
- In the area of so-called bag of tasks, parameter sweep applications volunteer computing is competitive.
- Near-term potential of volunteer computing goes well beyond Exa-scale.
Why volunteer desktop grids are important?

Volunteer desktop grids can collect very large number of resources

http://knowledgebase.e-irg.eu
The Grid Ecosystem

Desktop grids (DGs)
- Volunteer DGs – home computers
- Organizational DGs – institutional desktops
- Inexpensive,
- Very large number of CPUs (~10K – 1M)
- Param. sweep appls

Cluster based service grids (SGs)
- EGI, OSG, etc.
- Moderately expensive,
- Moderate number of sites and CPUs
- Any appls

Supercomputer based SGs
- TeraGrid, DEISA
- Very expensive,
- Small number of sites and very large number of cores
- MPI appls

EDGeS
EU FP7 projects on desktop grids: EDGeS → EDGI and DEGISCO

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

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

DEGISCO
- Supported by
  - Disseminate EDGeS results world-wide
  - Green IT aspects

2008 - 2010

2010 - 2012
Types of Desktop Grids

- **Global (volunteer) Desktop Grid**
  - Aim is to collect resources for grand-challenge scientific problems
  - Examples:
    - SETI@home, Folding@home, Shakemovie@home, LHC@home
    - Community World Grid, IberCivis, SZTAKI Desktop Grid

- **Local (institutional) Desktop Grid**
  - Aim is to enable the quick, easy and inexpensive creation of grid for any community (company, university, etc.) to solve their own applications
  - Example:
    - SZTAKI Desktop Grid (SZDG) local version (used within EDGeS, EDGI, DEGISCO)
Local DGs in practice – University of Westminster DG system based on SZDG

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

Lifecycle of a node:
1. PCs basically used by students/staff
2. If unused, switch to Desktop Grid mode
3. No more work from DG server -> shutdown (green solution)
Results of EDGeS:
Production DG ⇒ EGI Infrastructure

- **gLite based EDGeS VO**
  - voms
  - lfc
  - bdii
  - lb
  - wms
  - myproxy

- **XtremWeb-EGI bridge (ui2)**
- **BOINC-EGi bridge (ui1)**

**XtremWeb-based DGs**
- IN2P3, Almere
  - local DGs
  - IN2P3
  - public DGs

- SZTAKI 16 cpus
- CIEMAT 20 cpus
- LIP 1080
- BIFI 36
- Sinica 6
- IPB 672
- UFCG 6
- KFKI 410

**BOINC-based DGs**
- UoW, Correlation Systems
  - local DGs
  - SZDG, Almere, Ibercivis
  - public DGs

**UoW, Correlation Systems**

**Local DGs**
- CNRS 1800 cpus

**Public DGs**
- IN2P3, Almere
- LIP 1080
- BIFI 36
- IPB 672
- UFCG 6
- KFKI 410
A typical example of the DG=>EGI direction: ViSAGE - Video Stream Analysis in a Grid Environment
By Correlation Systems Ltd. - Israel

Visage processes Image pairs over the BOINC/EDGeS grid and paints movement in yellow.

Insert video source:

Set mode to EDGeS

Video options: forward, backward, pause, play..
ViSAGE - Video Stream Analysis in a Grid Environment
By Correlation Systems Ltd. - Israel

ViSAGE - Video Stream Analysis in a Grid Environment
By Correlation Systems Ltd. - Israel

Small local DG at Israel

DG to SG bridge

EDGeS VO

Presentation for: ISGC2011
Author: Peter Kacsuk
Results of EDGeS: Production

EGI⇒DG Infrastructure

- gLite based service grid
- voms
- Ifc
- AR
- bdii
- lb
- wms
- myproxy

XtremWeb-based DGs
- IN2P3
- local DG
- IN2P3 @home
- Almere
- public DGs

BOINC-based DGs
- UoW
- local DGs
- SZDGr,
- Fundecyt,
- Almere,
- EDGeS@home
- public DGs

Presentation for: ISGC2011
Author: Peter Kacsuk
EDGeS@Home

Project Performance:
Number of users: 8215
Number of hosts: 10810
Estimated performance of last 48 hours: 2011.187 GFlop/s

About EDGeS@Home

The aim of the EDGeS@Home project is to support the execution of selected and validated scientific applications developed by the EGEE and EDGeS community.

The project currently hosts the ISDEP - Integrator of Stochastic Differential Equations in Plasmas - application at production level.

It also hosts several other applications at beta (experimental) level.

More information about the EDGeS project can be found here.

Join EDGeS@Home

- Read our rules and policies
- This project uses BOINC. If you're already running BOINC, select Attach to Project. If not, download BOINC.
- When prompted, enter http://home.edges-grid.eu/home/
- If you're running a command-line or pre-5.0 version of BOINC, create an account first.
- If you have any problems, get help here.

Returning participants

- Your account - view stats, modify preferences

Community

- Profiles
Major Application on EDGeS@home

- Fusion community application: **ISDEP - Integrator of Stochastic Differential Equations in Plasmas**
- Used in EGI before EDGeS
- EDGeS created:
  - A production quality BOINC implementation deployed on the Ibercivis and EDGeS@home Desktop Grids.
  - Application runs through the EGI=>BOINC bridge allowing EGI fusion VO users to utilise resources of EDGeS@home.
BOINC/gLite applications available @ EGI AppDB

22 applications
XtremWeb/gLite applications available @ EGI AppDB

- AUTODOCK
  Molecular docking simulations using AutoDock
- BNB.Grid
  A Generic Framework for Implementing Optimization Algorithms on Distributed Systems
- Convert
  ImageMagick Convert
- Correlizer
  DNA sequence correlations
- DART
  Audio Retrieval
- DASSP
  Digital Alias-free Signal Processing - Defining the class of optimal periodic non...
- Dummy.App
  Testing AppDB registration
- EMMIL
  E-Marketplace Model Integrated with Logistics
- SLinCA
  Scaling Laws in Cluster Aggregation
- VAST
  VisuAl and SemantIc image search

10 applications
How to extend gLite VOs with volunteer and institutional desktop grids?
Security issues when gLite VOs are extended with DGs

• In DG projects
  – Applications are trusted
  – DG server is trusted
  – Clients are not trusted

• Therefore any EGI application that is intended to run in a volunteer DG should be validated to make it trusted

• EDGeS community provides:
  – Application validation service
  – **Application repository** (AR) where trusted applications are stored
  – EDGeS bridge transfers only those applications to DGs that are stored in the AR
gLite VO → DG system via EDGeS services

- EGI BDII
- EGI WMS
- EGI LB
- EGI UI
- EGI VOMS
- X509 proxy

Submit job → Watch Get output → EGI VOMS → X509 proxy

DG servers

BOINC1

BOINC2

XWeb
gLite VO → DG system via EDGeS services

EGI BDII

EGI WMS

EGI LB

EGI UI

Submit job

Watch Get output

Submit job

Watch Get output

Submit job

Watch Get output

Submit job

Watch Get output

Submit job

Watch Get output

X509 proxy

X509 proxy

X509 proxy

X509 proxy

BOINC1

BOINC2

BOINC1

BOINC2

XWeb

XWeb

DG servers

DG servers

3G Bridge

3G Bridge

3G Bridge

3G Bridge

WS interface

WS interface

WS interface

WS interface

BOINC plugin

BOINC plugin

BOINC plugin

BOINC plugin

WS interface

WS interface

WS interface

WS interface

BOINC1

BOINC2

XWeb
gLite VO → DG system via EDGeS services

EGI BDII (with DG-CE)

 EGI WMS

 Submit job

 EGI LB

 Watch Get output

 EGI UI

 EGI VOMS

 X509 proxy

 EDGeS Bridge Services

 Modified CE

 Info provider

 Send job

 3GB WS client

 DG-LRMS

 Check EXE

 3GB WS client

 EDGeS Application Repository

 DG servers

 BOINC1

 BOINC2

 XWeb

 BOINC plugin

 WS interface

 3G Bridge

 BOINC plugin

 WS interface

 3G Bridge

 BOINC plugin

 WS interface

 3G Bridge

 XWeb plugin

 WS interface

 3G Bridge

 X509 proxy

 DG servers

 BOINC1

 BOINC2

 XWeb

 BOINC plugin

 WS interface

 3G Bridge

 BOINC plugin

 WS interface

 3G Bridge

 XWeb plugin

 WS interface

 3G Bridge

 XWeb
gLite VO → DG system via EDGeS services

- EGI BDII (with DG-CE)
- EGI WMS
- EGI LB
- EGI UI
- X509 proxy
- EGI VOMS
- EGI BDII (with DG-CE)
- EDGeS Bridge Services
  - Modified CE
  - QM WS client
  - DG-LRMS
  - EDGeS Bridge Services
    - Info provider
    - EDGeS Application Repository
- DG servers
  - BOINC plugin
  - BOINC1
  - BOINC plugin
  - BOINC2
  - XWeb
  - BS interface
  - WS interface
  - 3G Bridge
  - BS interface
  - BOINC plugin
  - BOINC plugin
  - BS interface
  - 3G Bridge
  - BS interface
  - XWeb plugin
  - BS interface
  - WS interface

Actions:
- Submit job
- Watch Get output
- Check EXE
- Send job
- Send output
- Log events
- Report resources and performance
- Get EXE
Target 1: extend your gLite VO with volunteer DG Infrastructure

Extend your gLite VO with volunteer DGs

Steps:
Connect the modified CE of the EDGeS Bridge services to your gLite VO using your BDII service
Target 2: extend your gLite VO with institutional DGs

Steps:

1. Establish e.g. univ. DGs
2. Deploy 3G Bridge on the univ. DG server
3. Ask EDGeS Bridge Service provider to connect your DG server to the EDGeS Bridge services
Plans for EDGI

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

EDGeS scope only
- for compute intensive applications
- for gLite

Extend Desktop Grids with Clouds for QoS
Companies and universities can extend their cheap but limited capacity local DG systems with pay-as-you-go additional resources on-demand.
Key issue: SUSTAINABILITY

The *International Desktop Grid Federation* (IDGF) brings together:

- Desktop Grid developers
- Desktop Grid operators
- Application developers
- Everyone else interested in Desktop Grid computing

Open membership
Founding members of IDGF

Presentation for: ISGC2011
Author: Peter Kacsuk
Join to the volunteer computing community today and offer your notebook for volunteer computing!

Volunteer! Make a difference

Source: http://europa.eu/volunteering
Thank you for your attention ...

For more information please visit the EDGeS, EDGI, DEGISCO and IDGF Websites:

http://www.edges-grid.eu/
http://edgi-project.eu
http://degisco.eu
http://desktopgridfederation.org

and/or send e-mail to me:

kacsuk@sztaki.hu