



Running Multiscale applications with QosCosGrid



MAPPER
1st Seasonal
School

Poznan Supercomputing and Networking Center

Motivations



- Tightly coupled multiscale applications are composed of many single scale models that may have:
 - different hardware requirements (e.g. GPU, MPP, SMP),
 - or different software requirements (e.g. Palabos, CPMD, LAMMPS).
 - These application frequently must be run in parallel, but this is not supported by existing middleware in EGI and PRACE.
 - We have developed new core middleware services, QosCosGrid, which provides these capabilities, and can be deployed on EGI and PRACE resources.

The Tools



- **QCG-Computing**
 - Basic Execution Service (BES) supports advance reservation.
- QCG-Coordinator
 - Supports QCG-Computing in cross-cluster execution of jobs
- QCG-Notification
 - Notification capabilities based on WS-Notification
- **QCG-Broker**
 - Resource management and brokering service
- **QCG-Client**
 - Text-based client for QosCosGrid
- QCG-Icon
 - Lightweight desktop client for QosCosGrid
- QCG-Tools
 - Various elements extending the QosCosGrid stack
- QCG-Nagios
 - Nagios probes for the QosCosGrid stack

Challenges

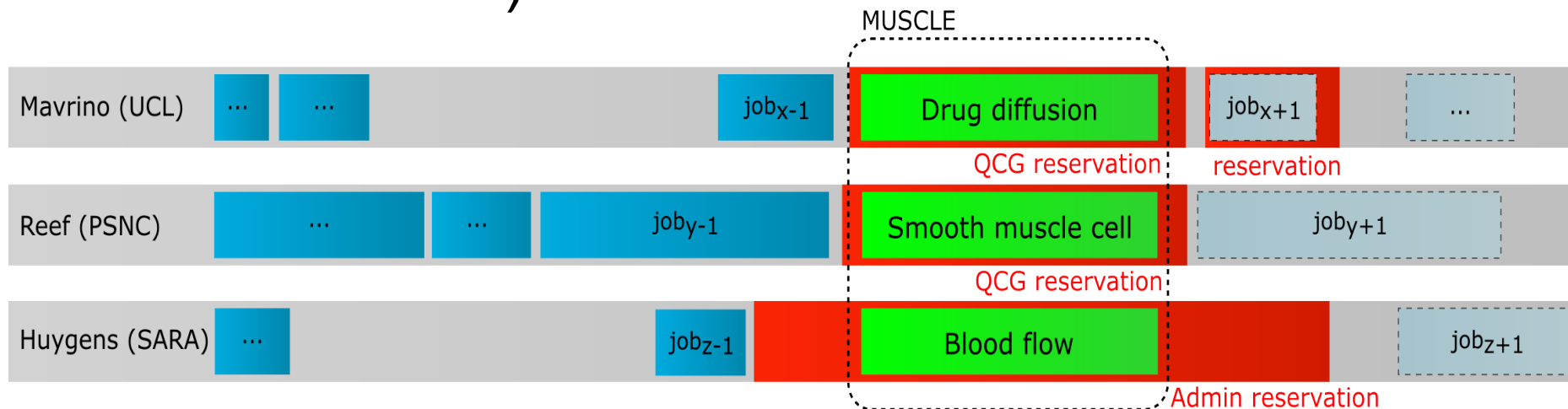


- Running multiscale application in cross-cluster environment requires addressing the following issues:
 - co-allocation of heterogeneous resources,
 - coordination of application spawning at multiple sites,
 - enabling connectivity between firewalled resources and between private IP domains.

Resources Co-allocation



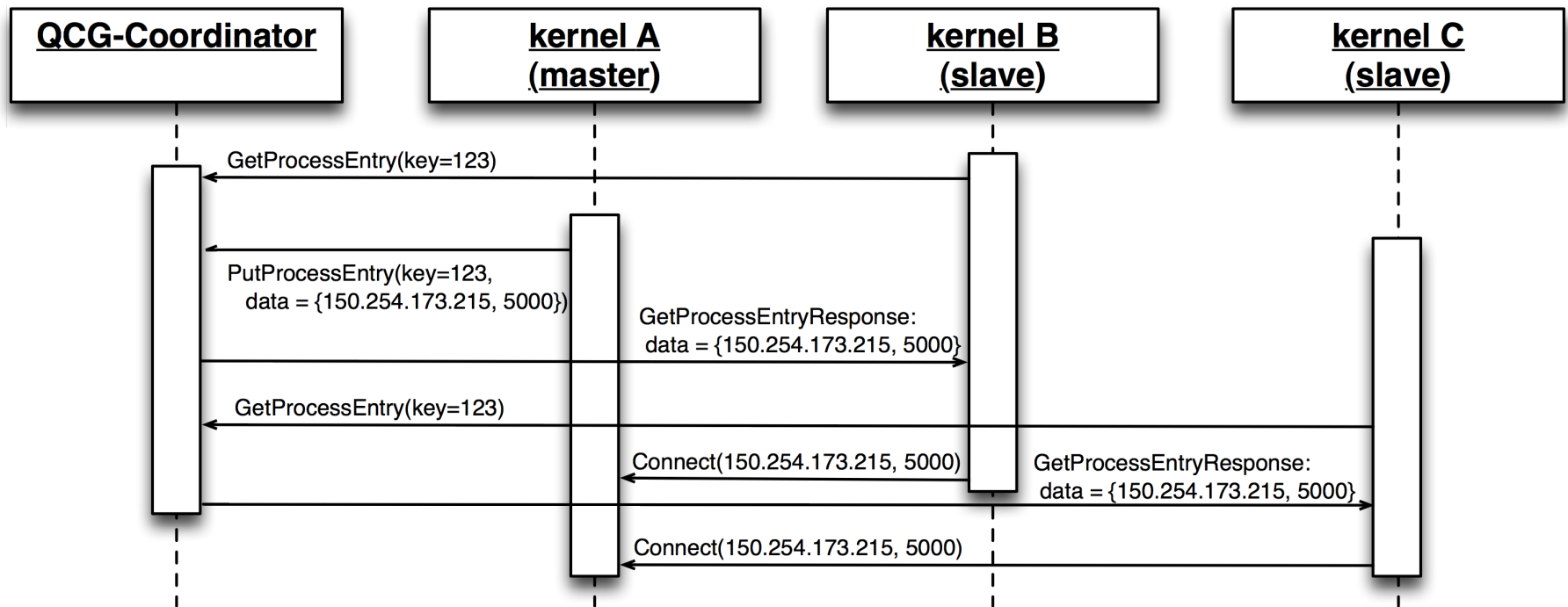
- Based on Advance Reservation mechanism,
- Process managed by QCG-Broker,
- Reservations can be created on demand (using QCG-Computing) or manually (by an administrator).



Application Spawning



- Problem: Kernels are started independently
- Solution: External service: QCG-Coordinator

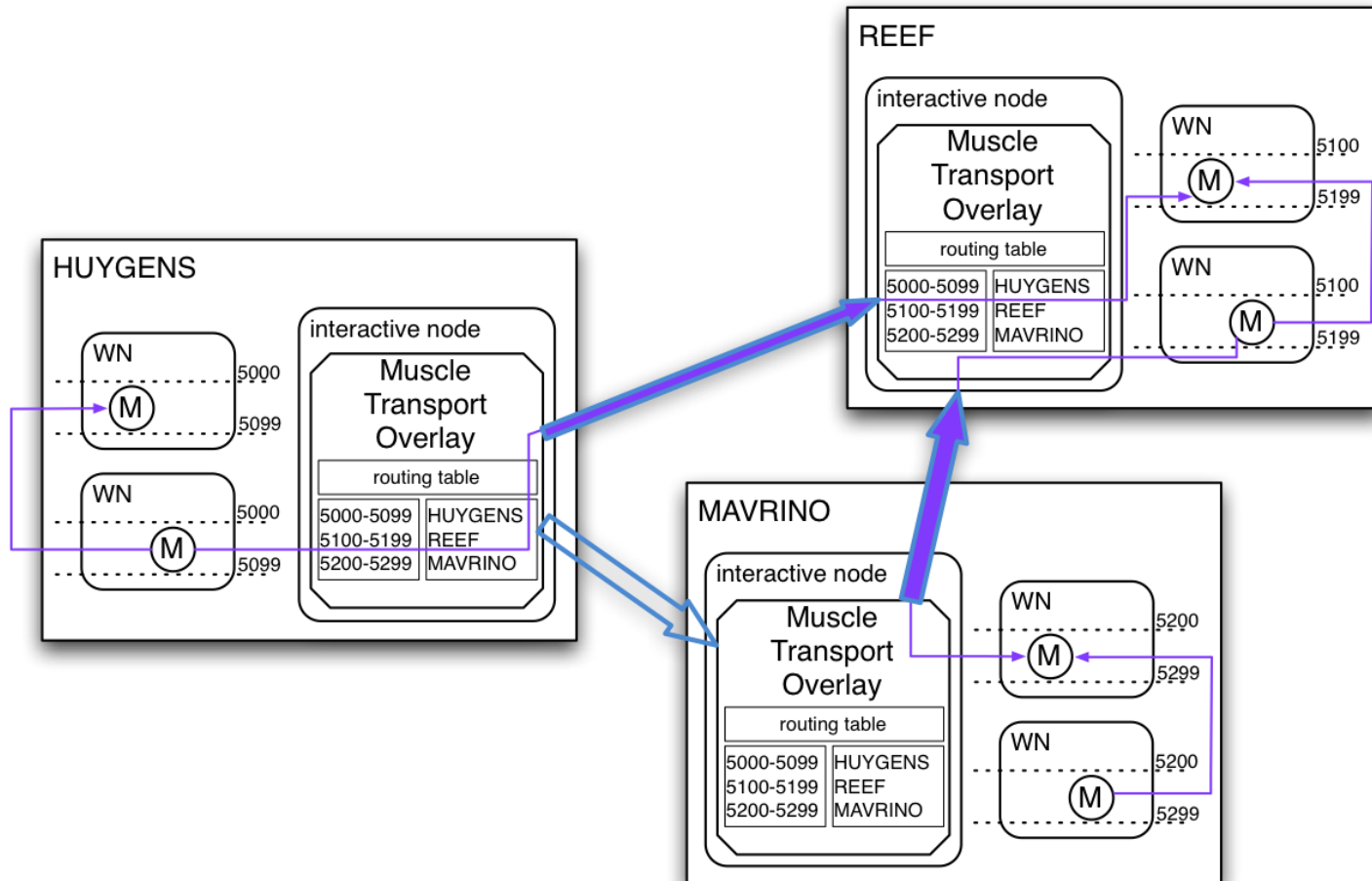


Connectivity (I)



- Many clusters use private IP addresses for their back-end compute nodes.
- Some sites also restrict outgoing traffic.
- Solution:
 - Use predefined port ranges.
 - Deploy Muscle Transport Overlays (MTOs) user-space daemons on the interactive nodes.

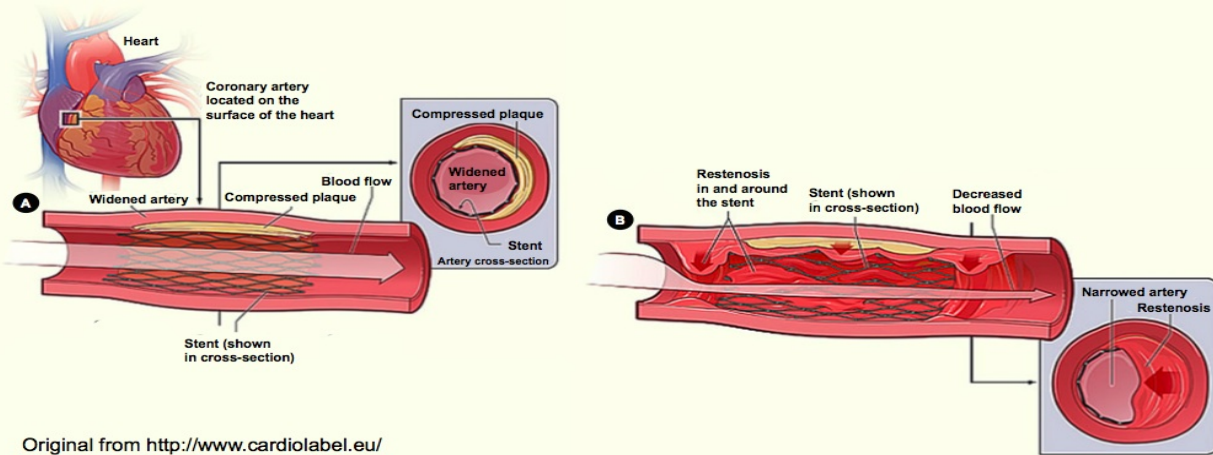
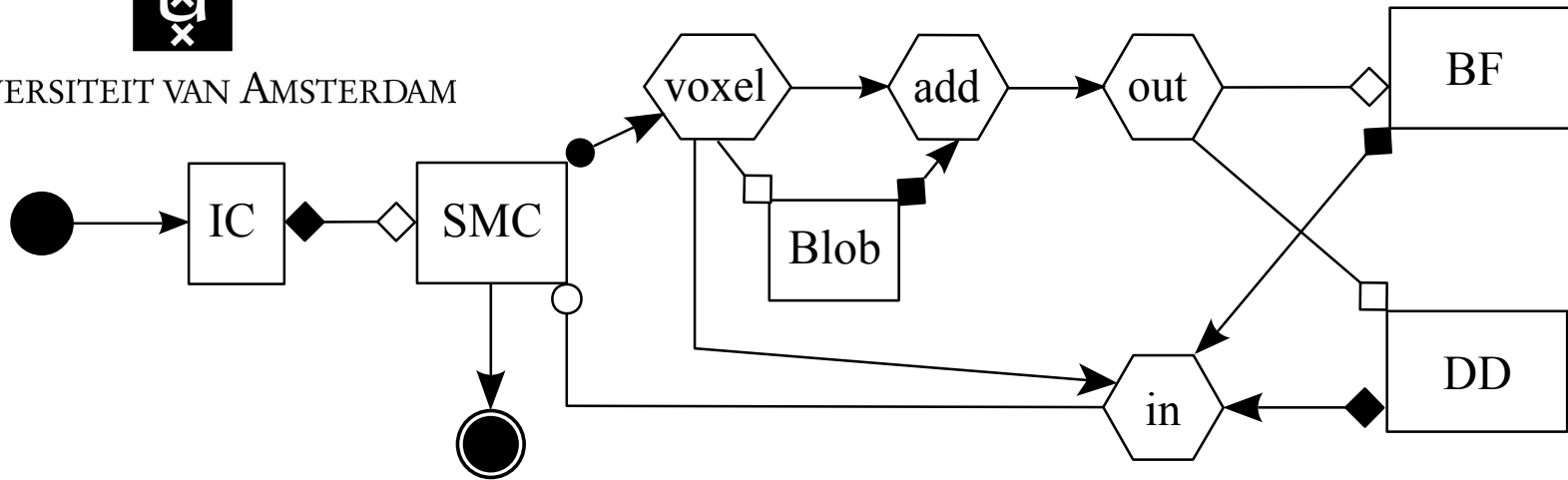
Connectivity - MTO (II)



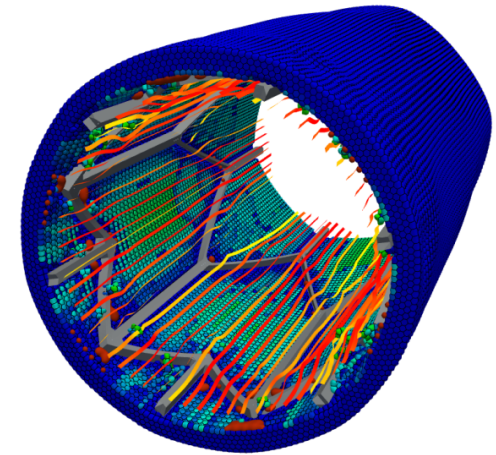
Application: In-stent Restenosis 3D



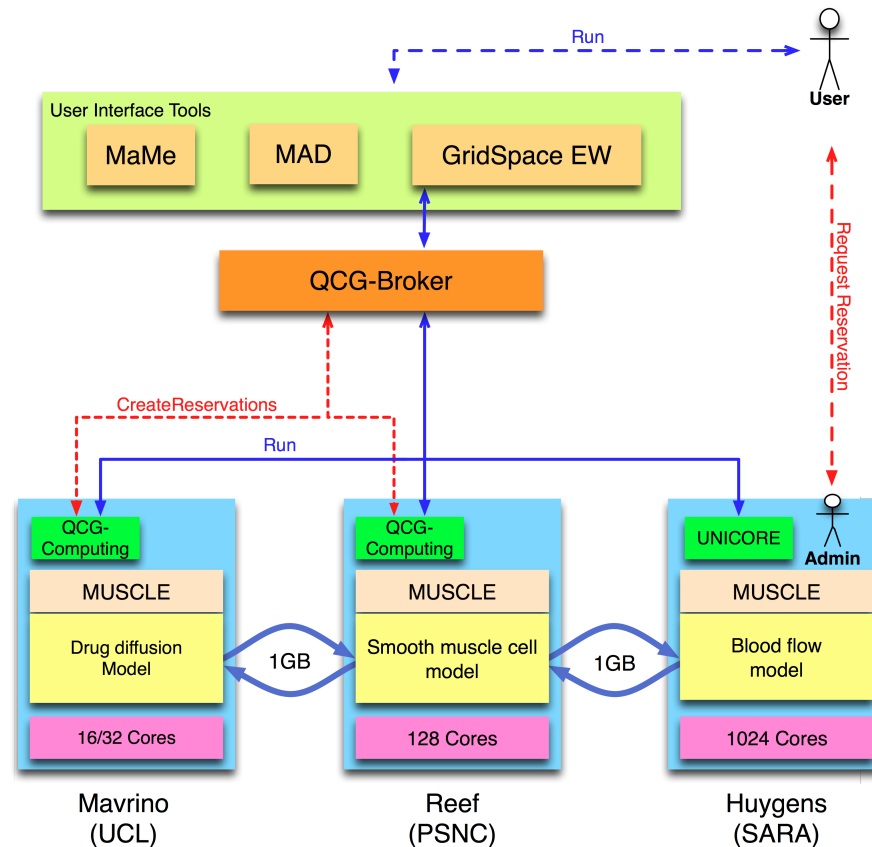
UNIVERSITEIT VAN AMSTERDAM



Original from <http://www.cardiolabel.eu/>



Example Run – 1st MAPPER Review



- This demo integrated resources provided by EGI, PRACE and local infrastructures.

QCG-Broker - JobProfile



```
<topology>
  <processes processesId="BF">
    <processesCount>
      <value>64</value>
    </processesCount>
    <candidateHosts>
      <hostName>huygens.sara.nl</hostName>
    </candidateHosts>
    <reservation type="LOCAL">p6012.huygens.sara.nl.537.r</reservation>
  </processes>
  <processes processesId="SMC:collector:distributor:Blob:voxelizer:thrombusMapper">
    <processesCount>
      <value>1</value>
    </processesCount>
    <candidateHosts>
      <hostName>reef.man.poznan.pl</hostName>
    </candidateHosts>
  </processes>
  <processes processesId="DD">
    <processesCount>
      <value>4</value>
    </processesCount>
    <candidateHosts>
      <hostName>mavrino.chem.ucl.ac.uk</hostName>
    </candidateHosts>
  </processes>
</topology>
```

Resource Requirements

Advance Reservation created by administrator

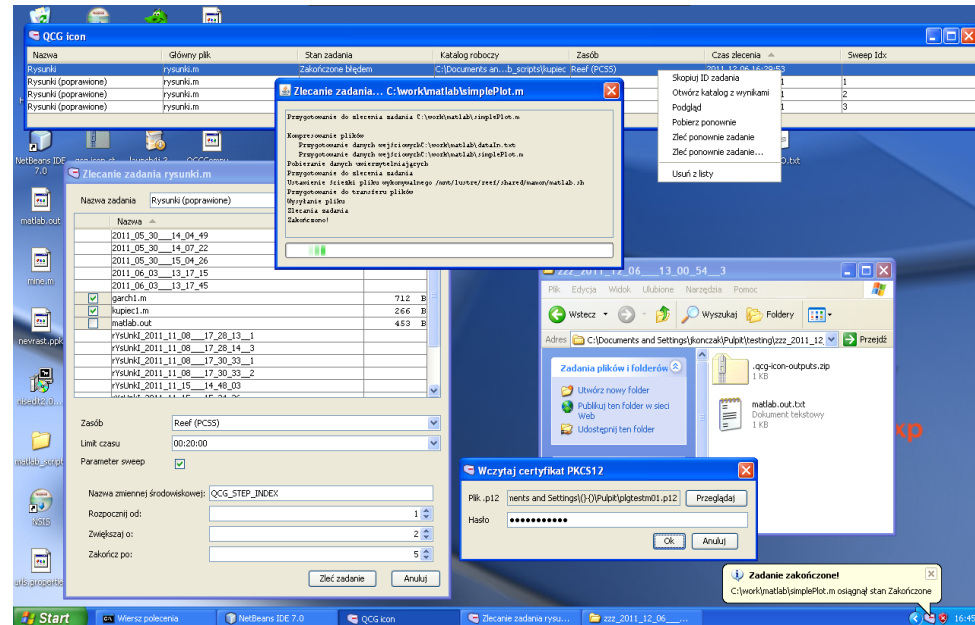
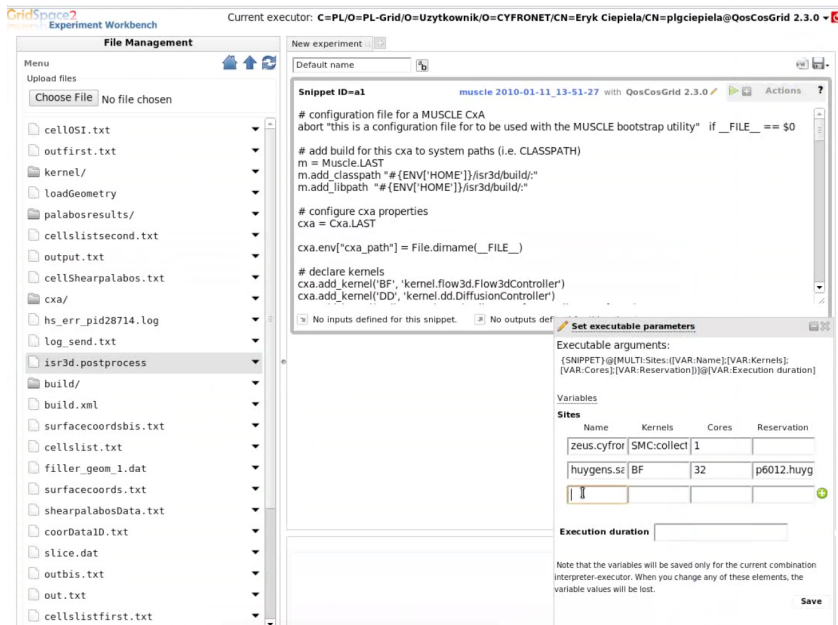
Kernels List

Target System

User Tools



- Command line clients (QCG-Client).
- Desktop client (QCG-Icon).
- QosCosGrid Executor in GridSpace EW.



DRMAA



- Distributed Resource Management Application API version 1.0 (more than 8 implementations),
- Exploited both by QCG and GridSpace,
- DRMAA 2.0 standard already released (Advance Reservation support),
- HPC-BASH

```
module load openmpi

#pragma hpc-bash parallel for
for ((file_num=$first; file_num<=$last; file_num=$file_num+1))
do
    cd $BASE_DIR/MODEL$file_num/
    #pragma hpc-bash batch-job walltime(00:30:00)
    $CACTUS_HOME/exe/cactus_SandT SandTank.par
done
cd $BASE_DIR/
```

QosCosGrid in EGI



- October 2012 – PSNC signs Memorandum of Understanding with EGI.eu

European Grid Infrastructure

towards a sustainable infrastructure

INFRASTRUCTURE • SERVICES • HOW DO I...? • CASE STUDIES • NEWS & MEDIA • ABOUT

HOME > NEWS & MEDIA > NEWSFEED > NEW MIDDLEWARE FOR NEW COMMUNITIES

New middleware for new communities



EGI.eu has signed an agreement with the Poznan Supercomputing and Networking Centre (PSNC) to integrate the QosCosGrid middleware stack into the European Grid Infrastructure (EGI).

Based in Poznan, PSNC is responsible for the development and management of the national optical research network, high-performance computing and various e-Science services and applications in Poland.

1 November 2012
Sara Coelho

qcg-offer - Look before you leap



- User: provides detailed view of free resources
- Administrator: helps with cluster diagnostics

HYDRA:
Summary:

Metric Name	nodes/cores	share
Total Resources:	282/5340	100%/100%
Up Resources:	256/4744	90%/88%
Free Resources:	107/2248	37%/42% (FreeNodes=66x12, 19x16, 16x48, 6x64)
PartFree Resources:	148/2878	52%/53% (AvgFreeCoresPerNode=19)
Reserved Resources:	2/24	0%/0% (Utilization=0%)

GALERA:
Summary:

Metric Name	nodes/cores	share
Total Resources:	196/2532	100%/100%
Up Resources:	167/2184	85%/86%
Free Resources:	6/72	3%/02% (FreeNodes=6x12)
PartFree Resources:	8/86	4%/03% (AvgFreeCoresPerNode=10)
Reserved Resources:	7/264	3%/10% (Utilization=0%)

QCG Community Modules



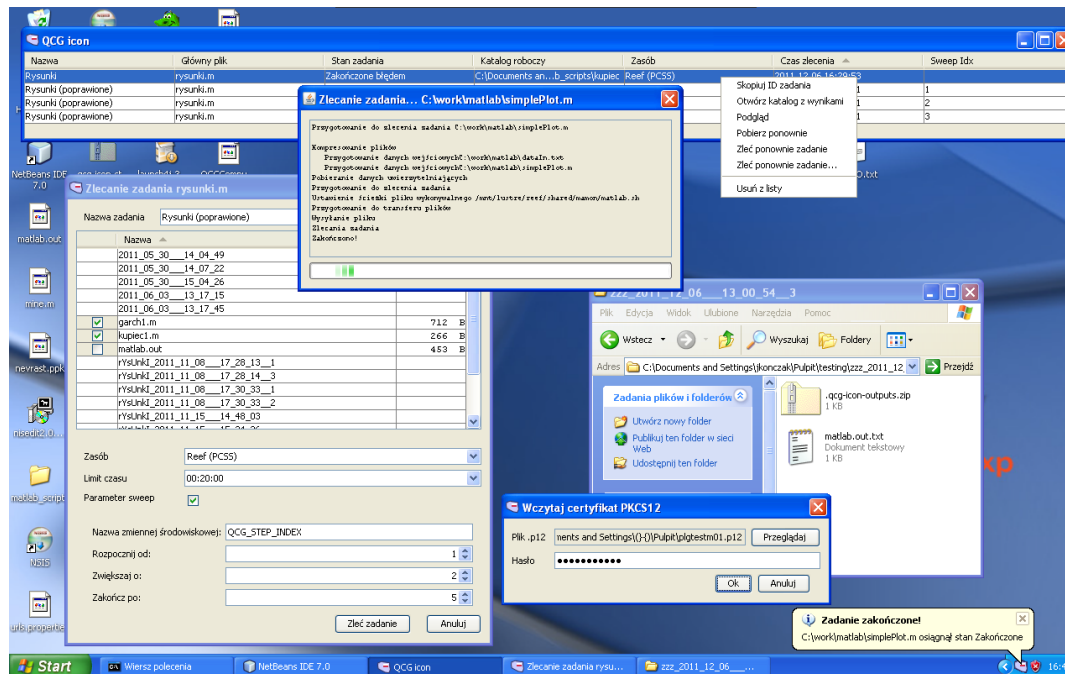
- user-space **environment modules** providing unified and coherent access to self-installed application among many clusters.

```
$ qcg-module-create -g plggmapper muscle/2.0
Creating module for the group: 'plggmapper'
Using module name: muscle
Using module version: 2.0
...
module-whatism "muscle, 2.0"
#prepend-path LD_LIBRARY_PATH your-lib-path
#prepend-path PATH your-bin-path
...
$ module load muscle/2.0
$ muscle2 --version
```


End Users tools: QCG-Icon



- Java based desktop application
- integrates tightly with the system (systray, context menu)
- Platforms: Windows, Mac OSX and Linux
- demo movies: <http://www.qoscosgrid.org/trac/qcg-icon>



Further Reading



- QosCosGrid homepage:
<http://www.qoscosgrid.org/>
- MAPPER Seasonal School Hands On:
<http://www.mapper-project.eu/web/guest/wiki/-/wiki/Main/QosCosGrid+tutorial+for+1st+seasonal+school>
- DRMAA & HPC-BASH:
<http://apps.man.poznan.pl/trac/pbs-drmaa/wiki/>
- DRMAA 2.0:
<http://www.ogf.org/documents/GFD.194.pdf>



Mariusz Mamoński, Tomasz Piontek



{mamonski,piontek}@man.poznan.p

|