



QosCosGrid & EGEE

The Industrial Sustainability of gLite

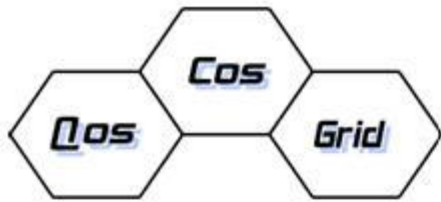
- Facilitating Industry Use Cases



Bernhard Schott

bschott@platform.com

Date: 5th March 2009



QosCosGrid: 11 partners from 9 countries



Israel Institute of
Technology (TECH)
Israel



University of Ulster (UU), United
Kingdom

Cranfield
UNIVERSITY
School of Management

Cranfield University (CU),
United Kingdom



Universitat Pompeu Fabra (PFU) Spain



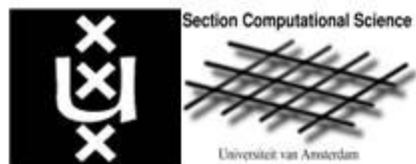
INRIA, France



Poznan Supercomputing and
Networking Centre (PSNC), Poland



Collegium Budapest
(ColBud) Hungary



Section Computational Science
University of
Amsterdam (UvA),
Netherlands

Platform
Accelerating Intelligence™

Platform Computing Sarl, France

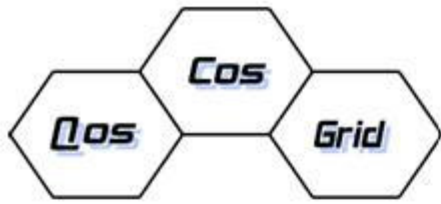


AITIA International
Inc., Hungary



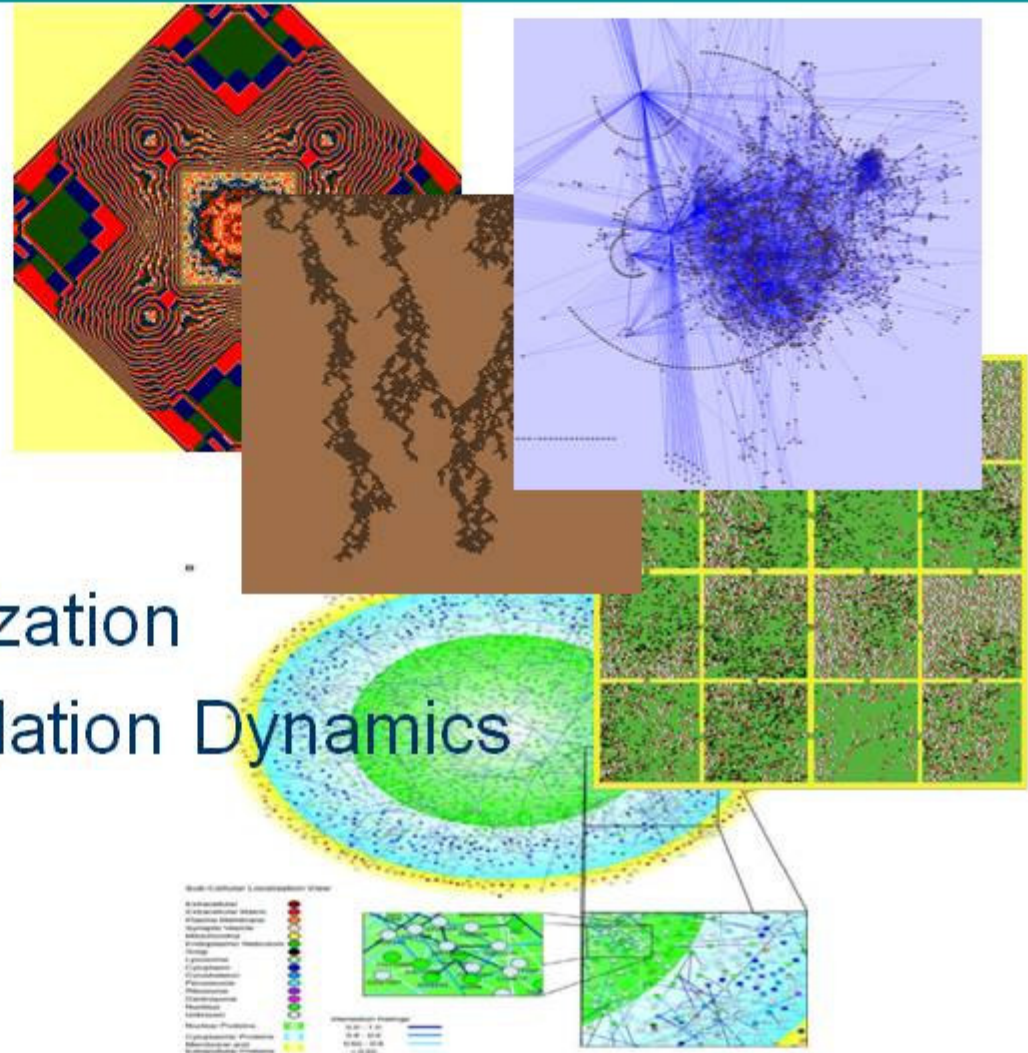
Information Society
Technologies

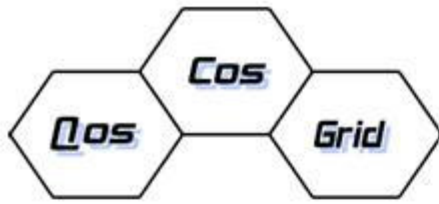
This work was supported by the EC grant FP6-2005-IST-5 033883 for the QosCosGrid project



Complex Systems Simulations (Examples)

- N-Particle Systems
- Protein Interactions
- Metabolic Pathways
- Financial Markets
- Market Research
- Supply Chain Optimization
- Ecological and Population Dynamics
- Stellar Systems

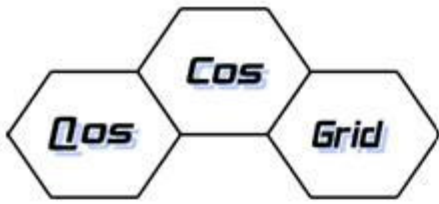




- N-Particle Systems
- Protein Interactions
- Metabolic Pathways
- Financial Markets
- Market Research
- Supply Chain Optimization
- Ecological and Population Dynamics
- Stellar Systems

*Large scale
Nontrivially
parallel*

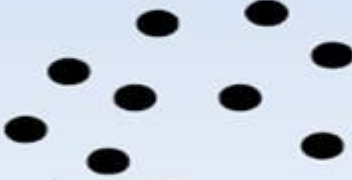

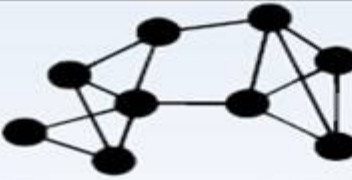
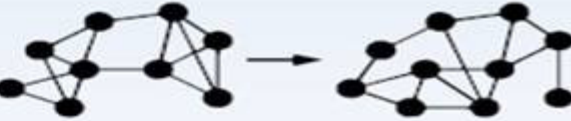






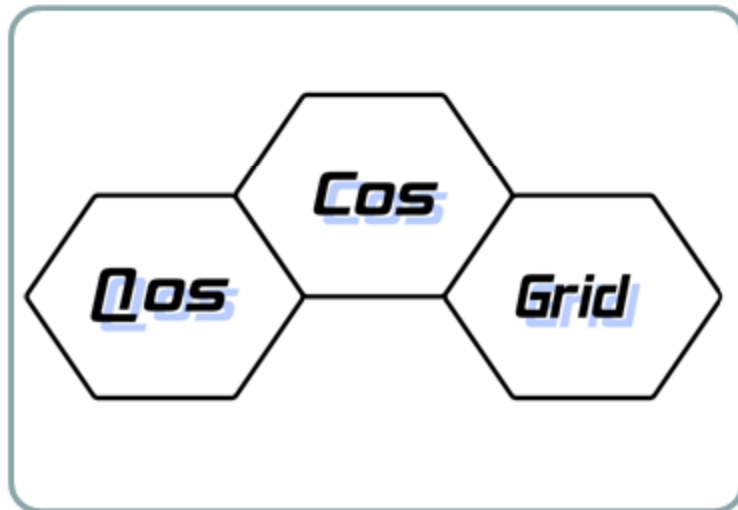
Communication Minimization

Categorization

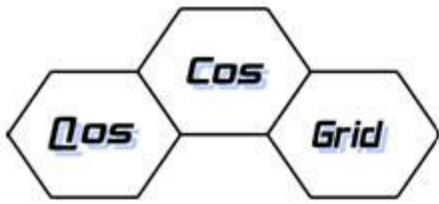
Categorization in various templates (T) based on CSS structure:

Special Cases	 Independent Components T0	 Unknown Topology T5
	Static	Dynamic
Network	 Fixed graph T1	 Dynamic graph T2
Spatial	 Cellular Automata T3	 Mobile Agents T4

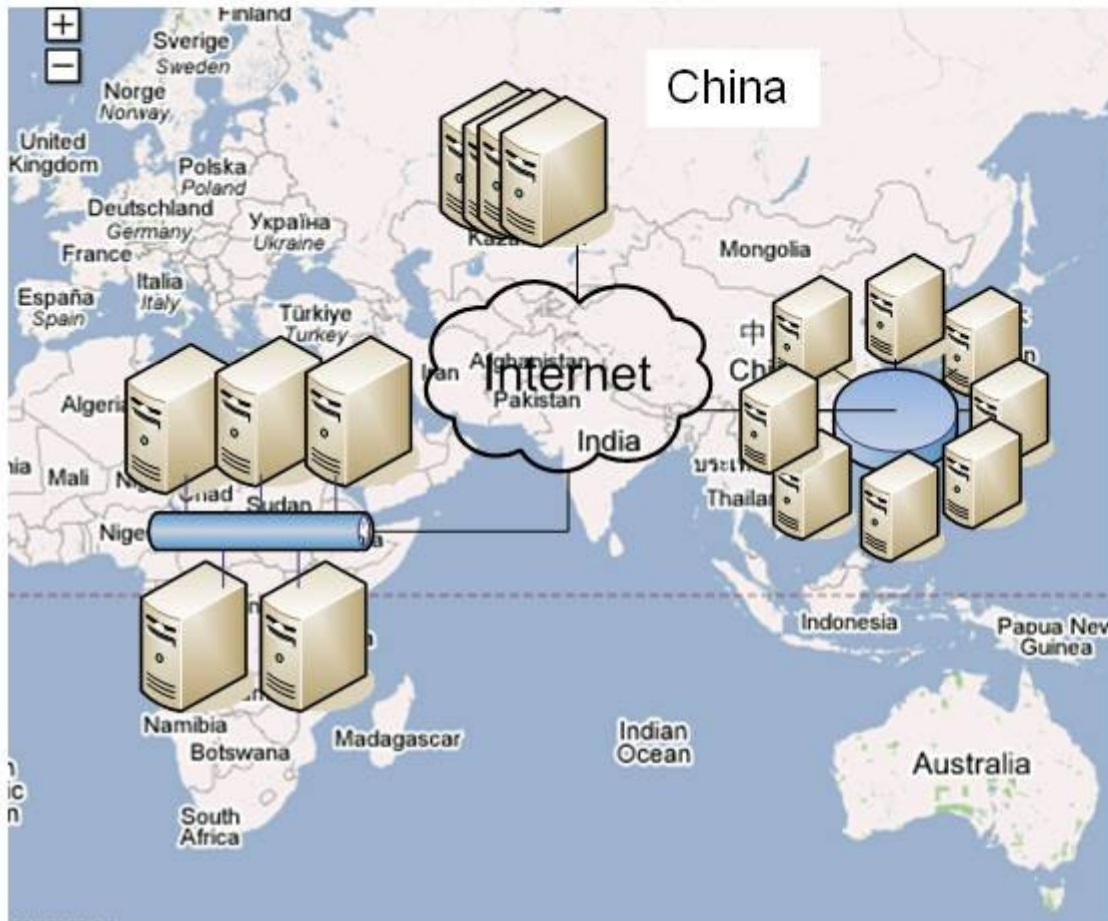




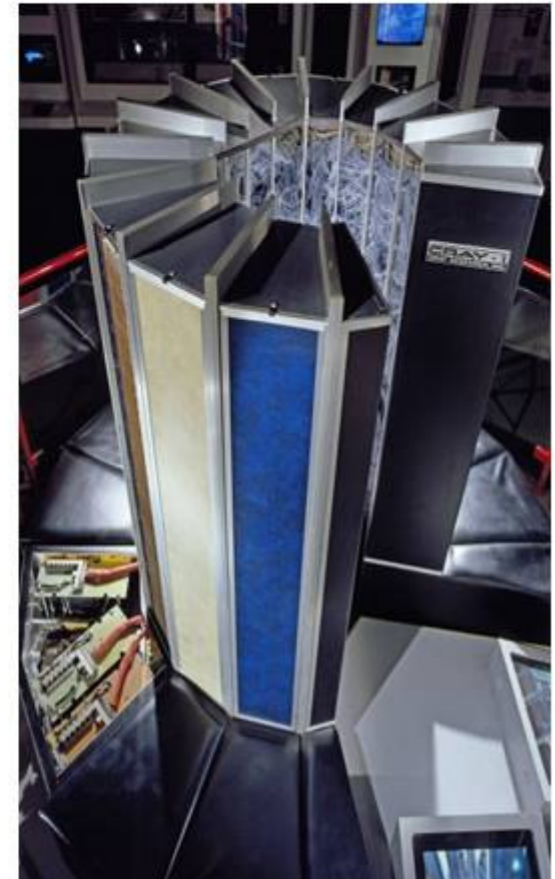
QosCosGrid:
formation of a Grid
on request

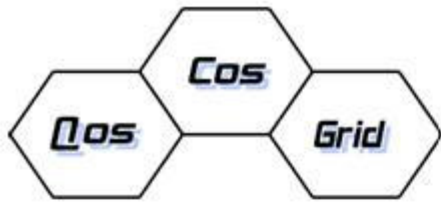


Vision: Grid \approx virtual supercomputer

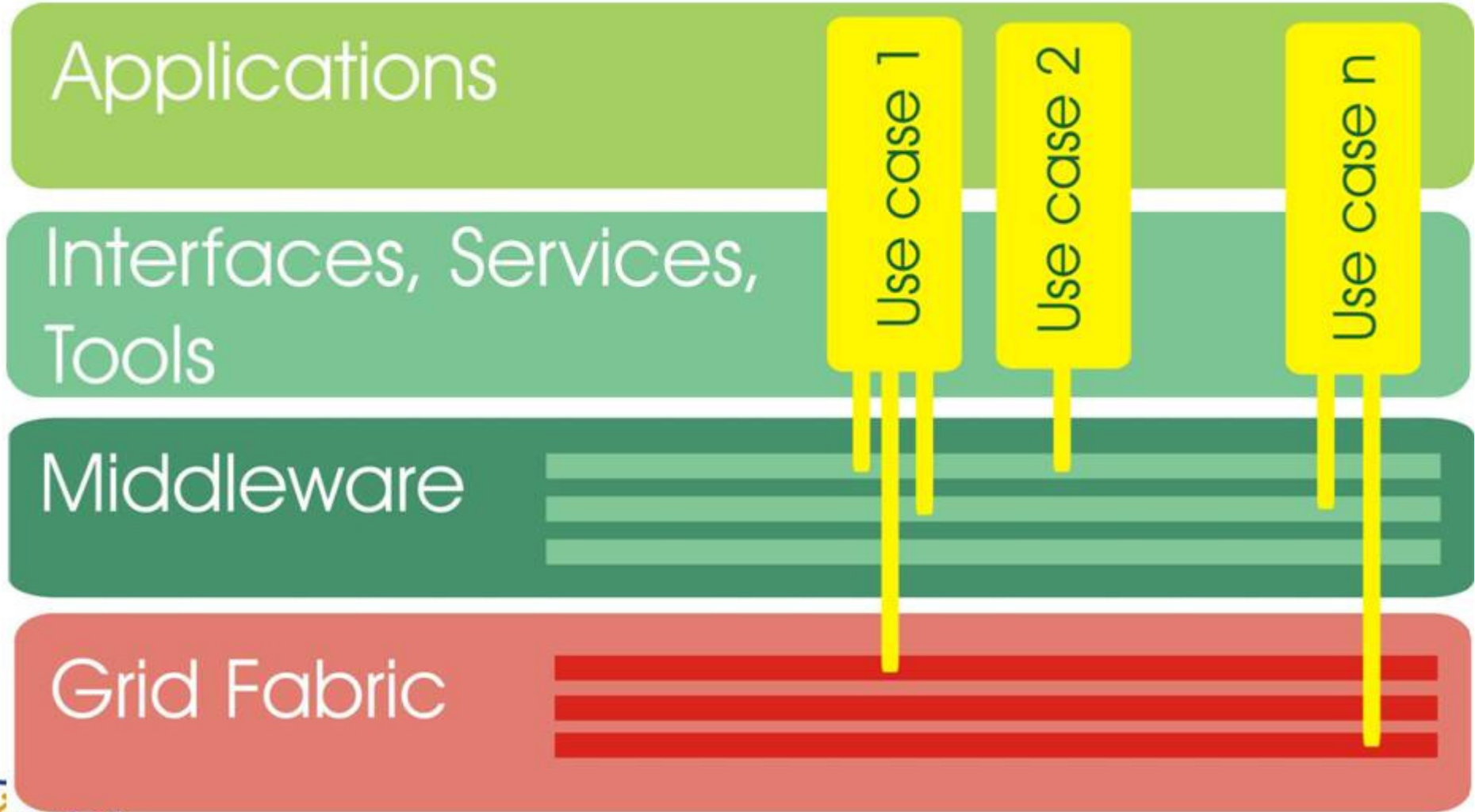


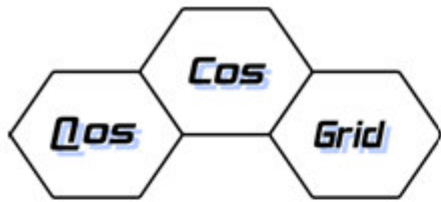
\approx



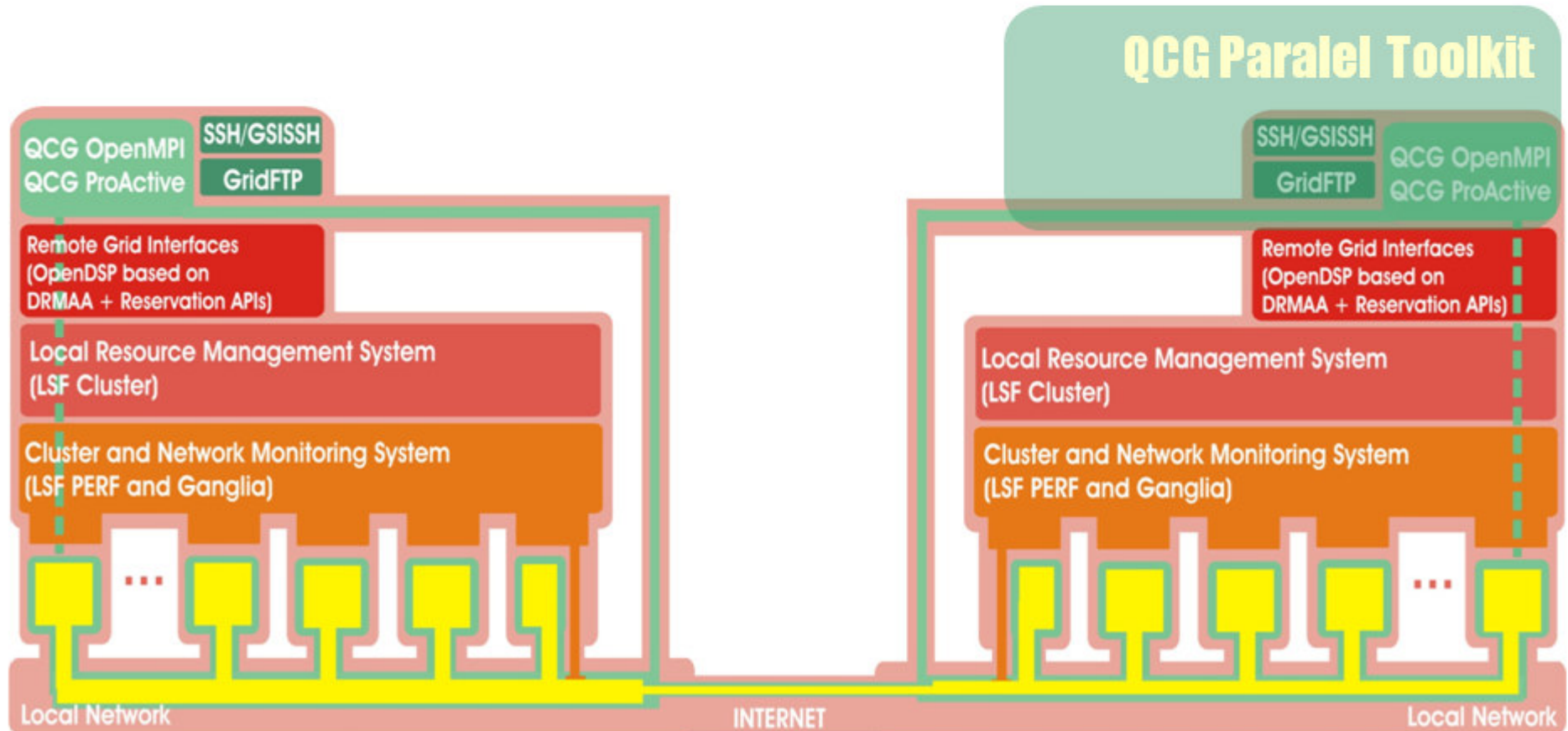


Architecture Overview





ADC structure

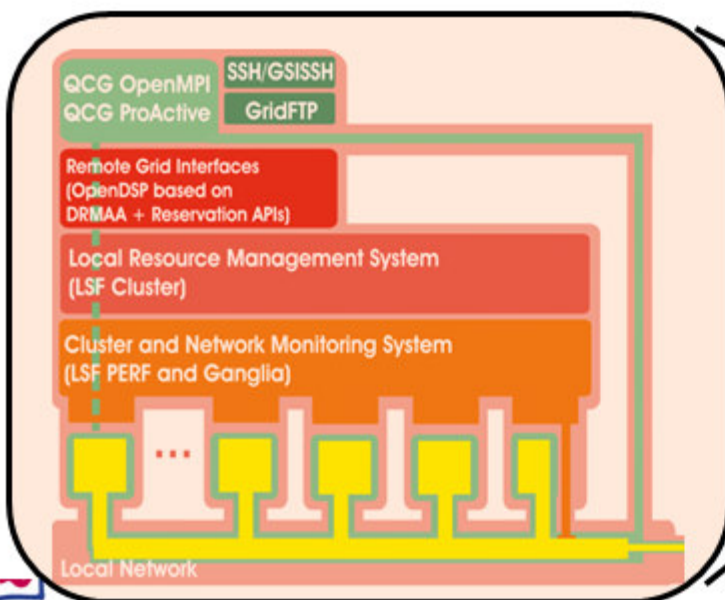
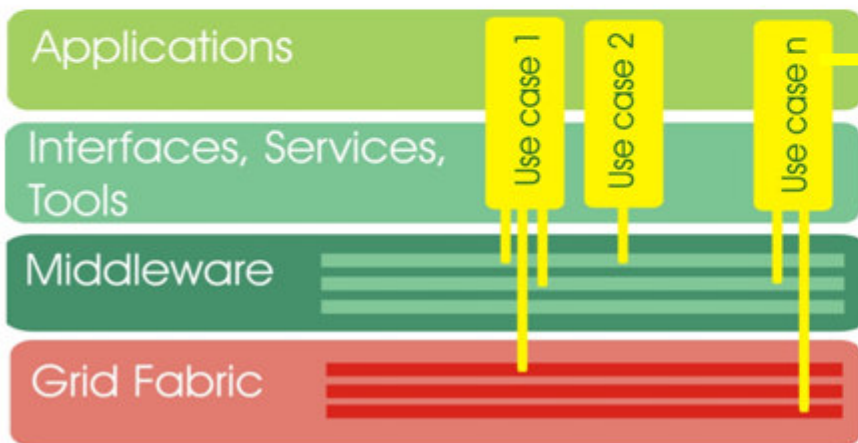


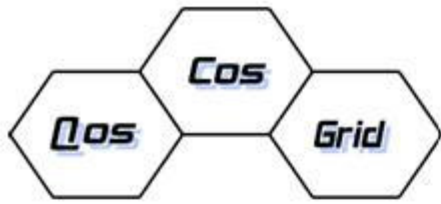
Legend



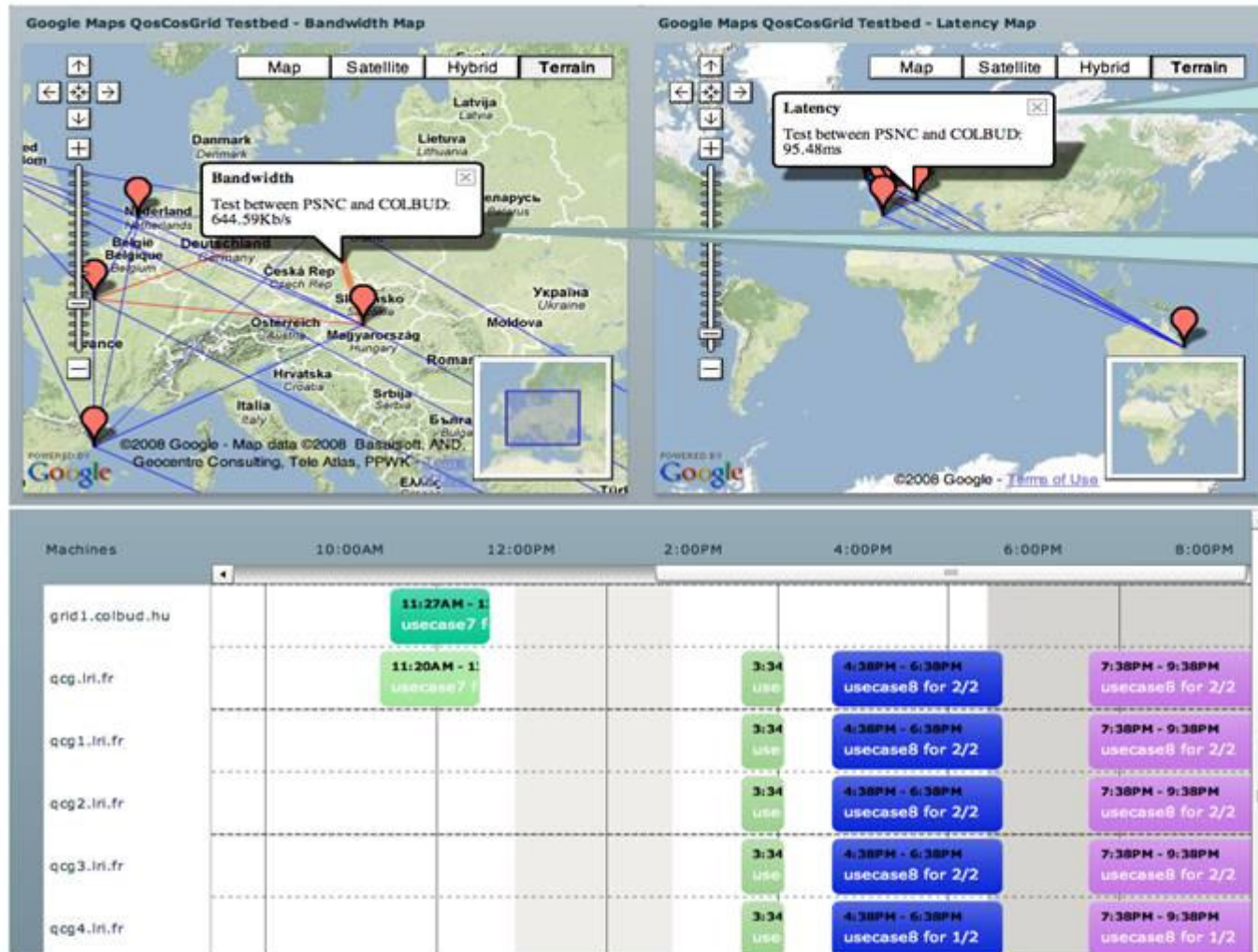


Dynamic Formation of a Grid





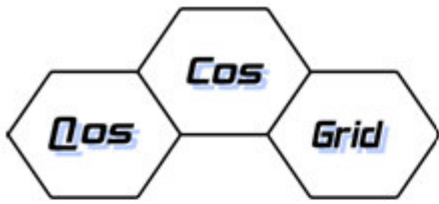
Web based QosCosGrid Gateway – comfortable user access



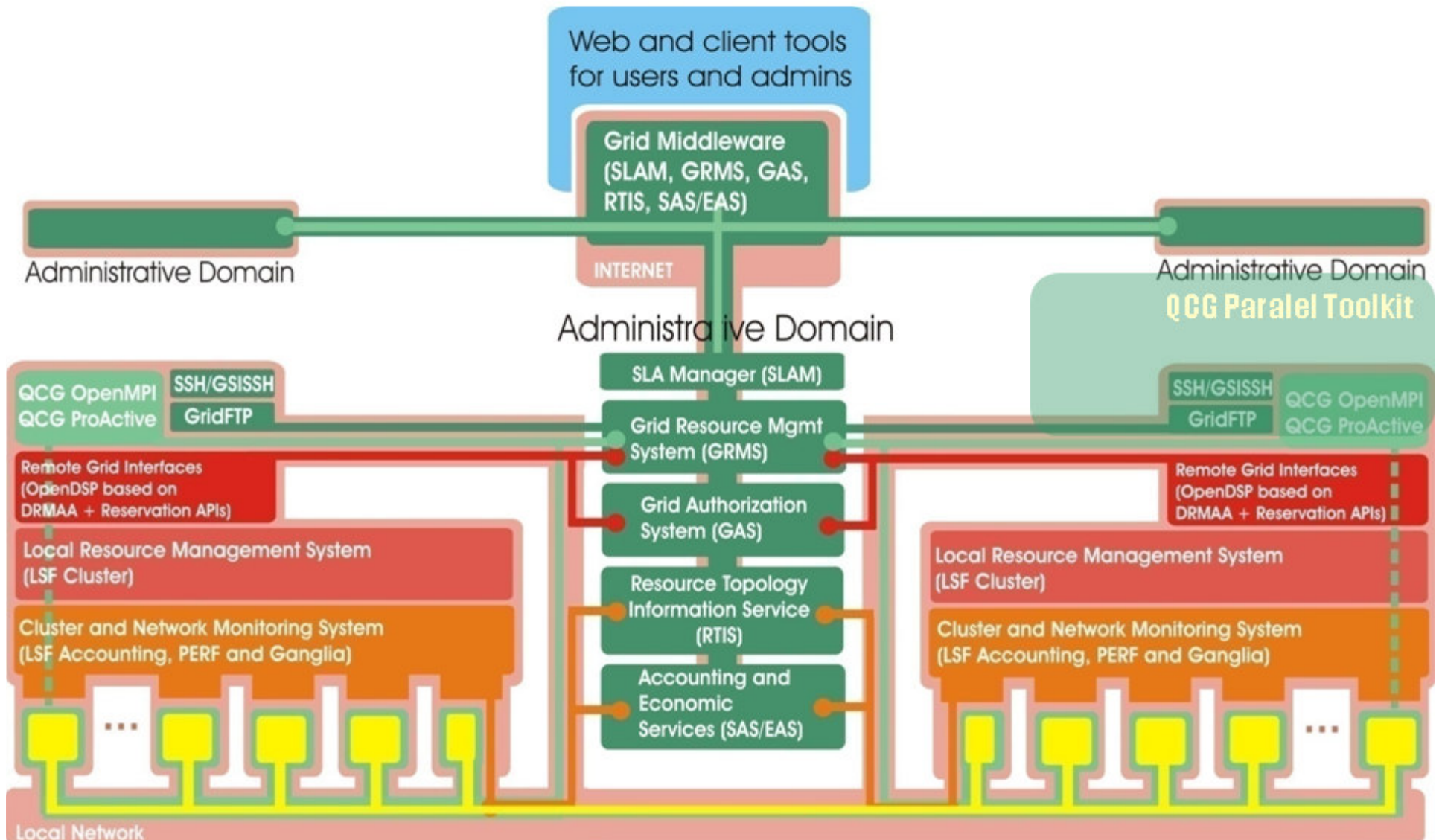
Live info on inter-site latencies

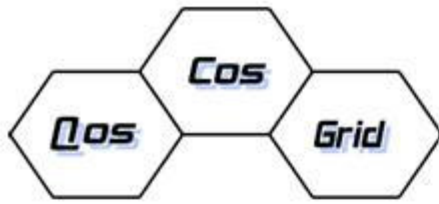
Live info on inter-site bandwidth

Visualized resource reservation and allocation plan



QosCosGrid Implementation

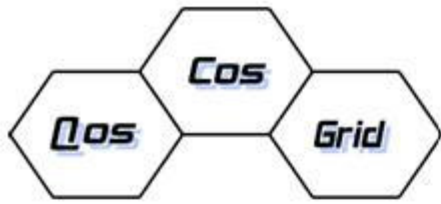




QosCosGrid workflow

- On-demand resource acquisition and formation of application specific Grids: QosCosGrid
 - QosCosGrid (= Quasi Opportunistic Supercomputing for Complex Systems on the Grid) Complex Systems applications use the QosCosGrid-Toolbox to parallelize their workload in order to use distributed resources.
 - Towards the QosCosGrid-Broker they express requirements and behavior by xml “job profile”
 - The Broker acquires resources on-demand and form an application specific Grid, creates RTG = resource topology graph
 - RTG is used to map the application to the resources, placing MPI communicators at the right place

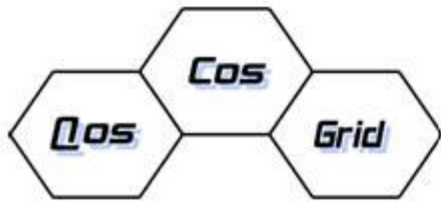




QCG Job profile

- The QCG Job Profile document is inspired by an existing XML-based job description language supported by one of the main components of the QCG middleware called GRMS.
- End users can describe topology and resource requirements, in particular:
 - required aggregations and hierarchies of resources (computing nodes, clusters, sub-clusters, storage elements etc.),
 - required resource properties (operating system, memory, number of CPUs, speed of the CPU on a resource),
 - required network and connection properties (bandwidth, latency and capacity),
 - required applications and licenses available at destination computing resources.

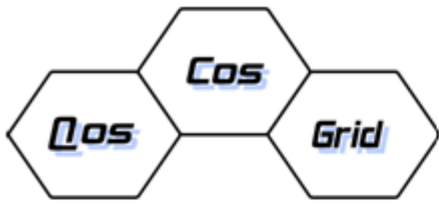




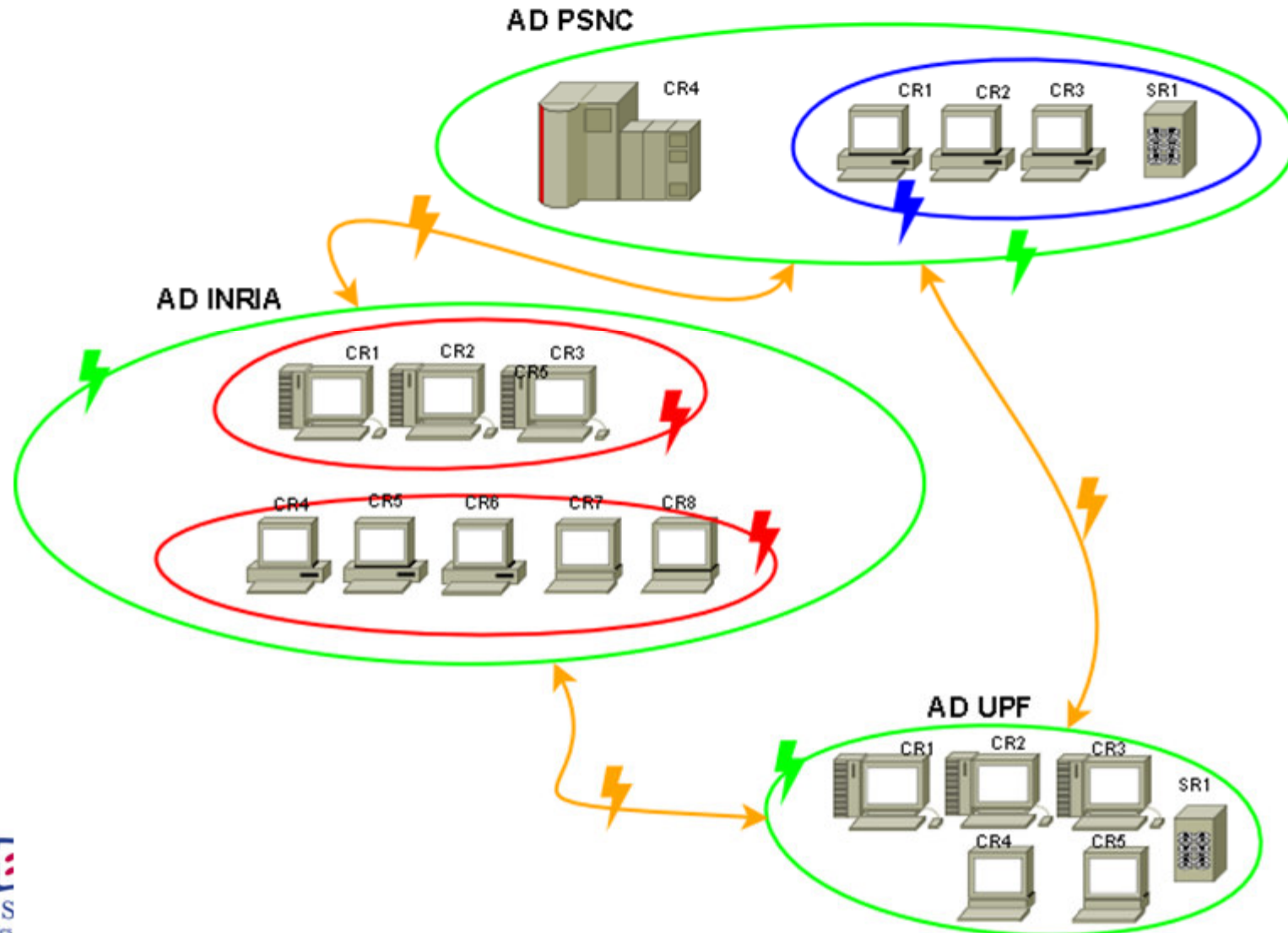
QCG Resource Description Model

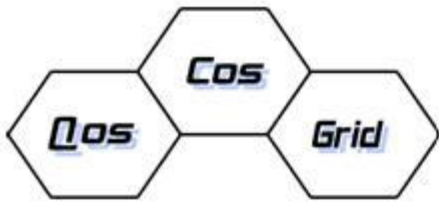
- RTG (Resource Topology Graph)
- A common XML resource description language
- Provide description of:
 - Resources, tasks, processes
 - Topology
 - Communication properties
- Serves as a “bridge” between the various system components
- Used to describe, publish, evaluate, reserve and monitor heterogeneous resources across the QosCos Grid
- Supplementary Java implementation:
 - Functional behavior and logic
 - XML to Java objects marshaling/un-marshaling
 - Specialized types of RTG objects, according to the middleware requirements (i.e. Resource advertisement, Meta-scheduling, SLA's, Monitoring, and User requirements.)



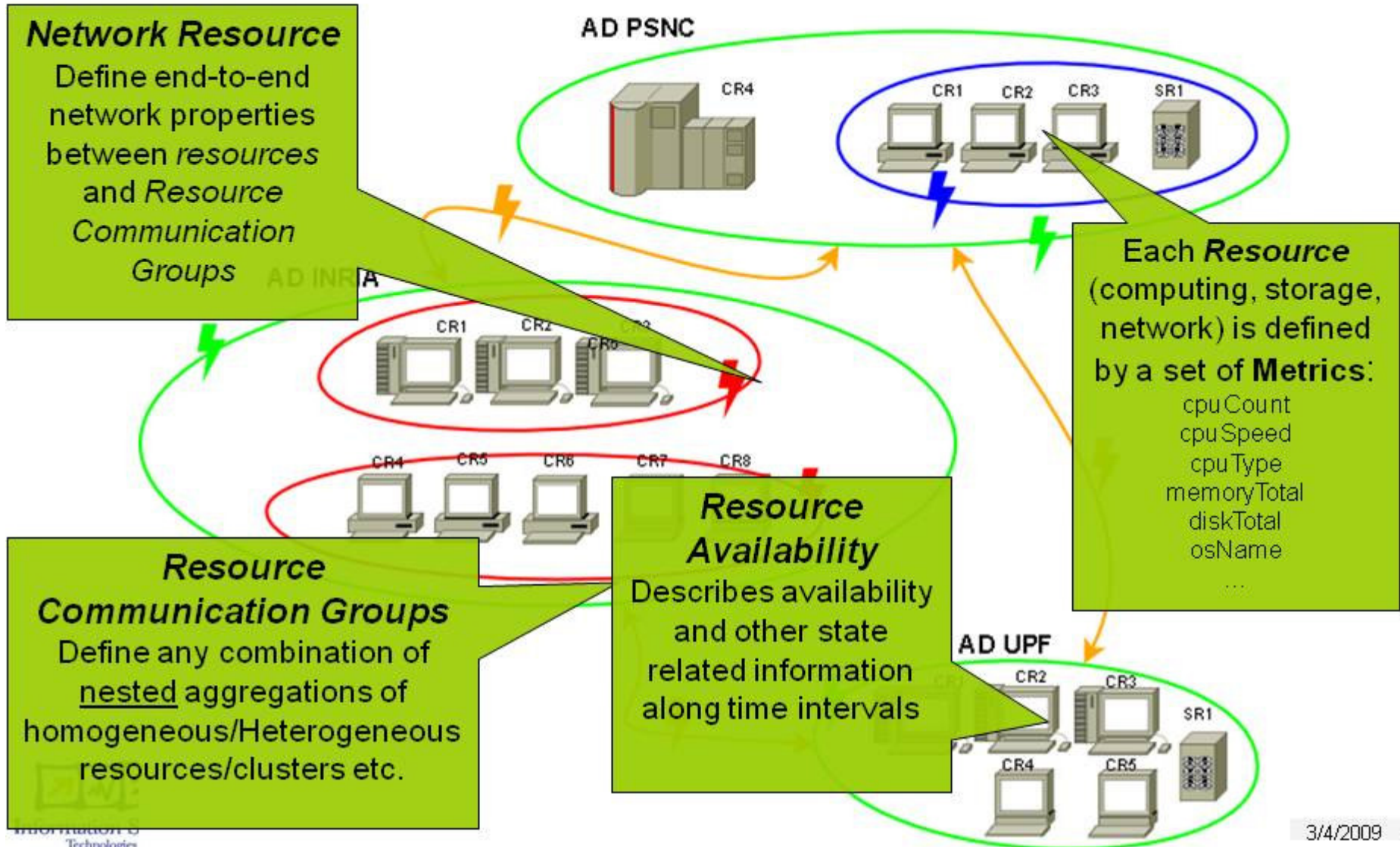


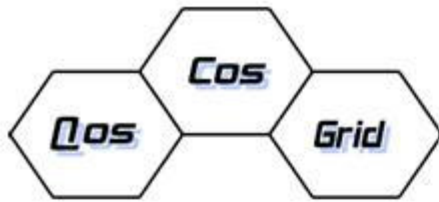
QCG RTG – Resource Description View



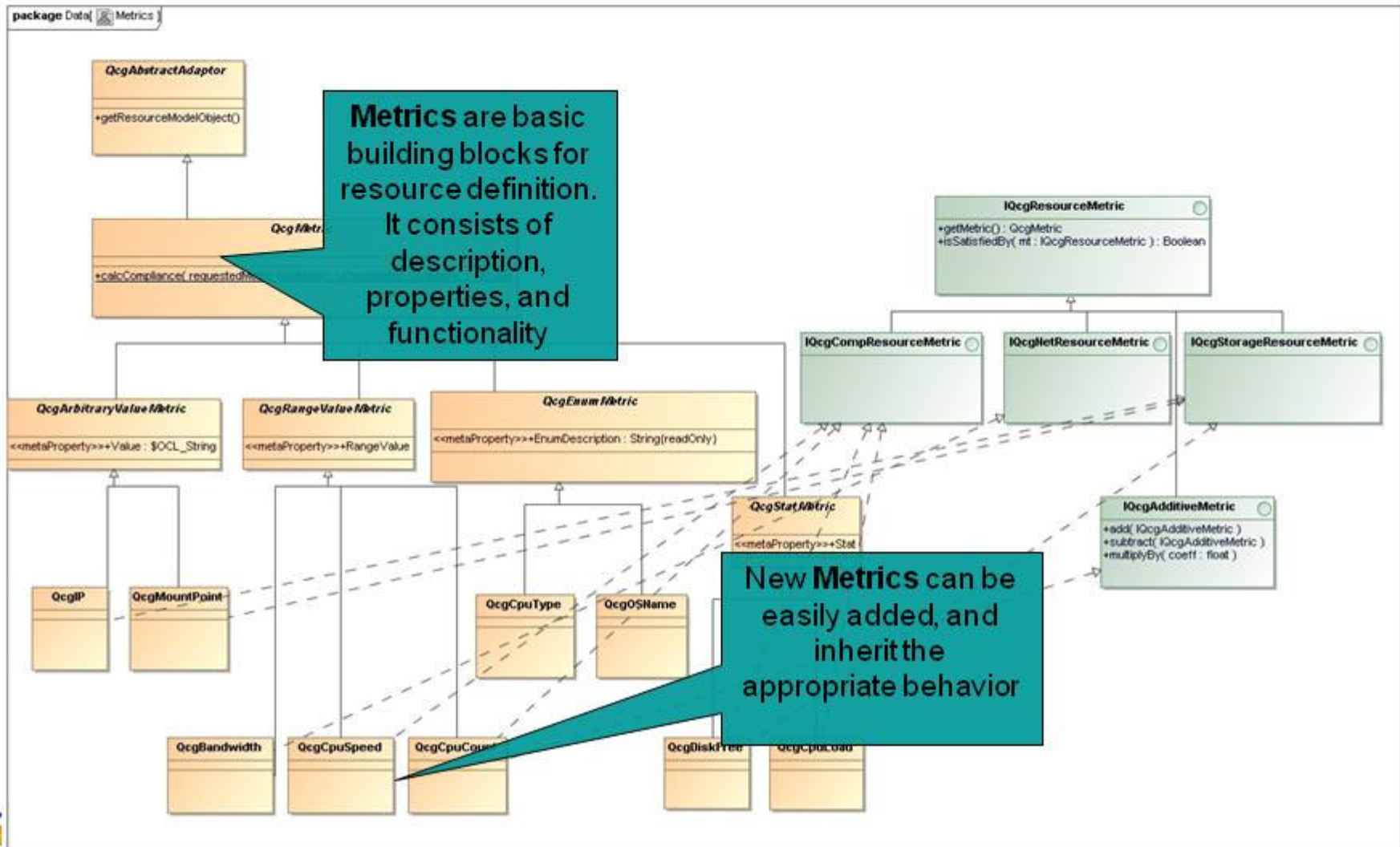


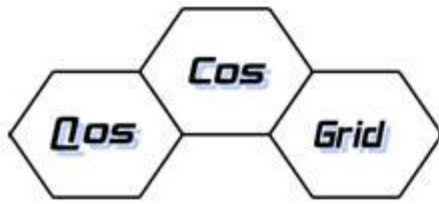
QCG RTG – Resource Description View





QCG RTG - Metrics





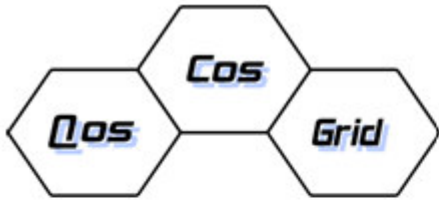
Resource Co-Allocation Protocol

- A simple, robust 2-phase commit alike protocol for planning based co-allocation of resource across multiple ADs

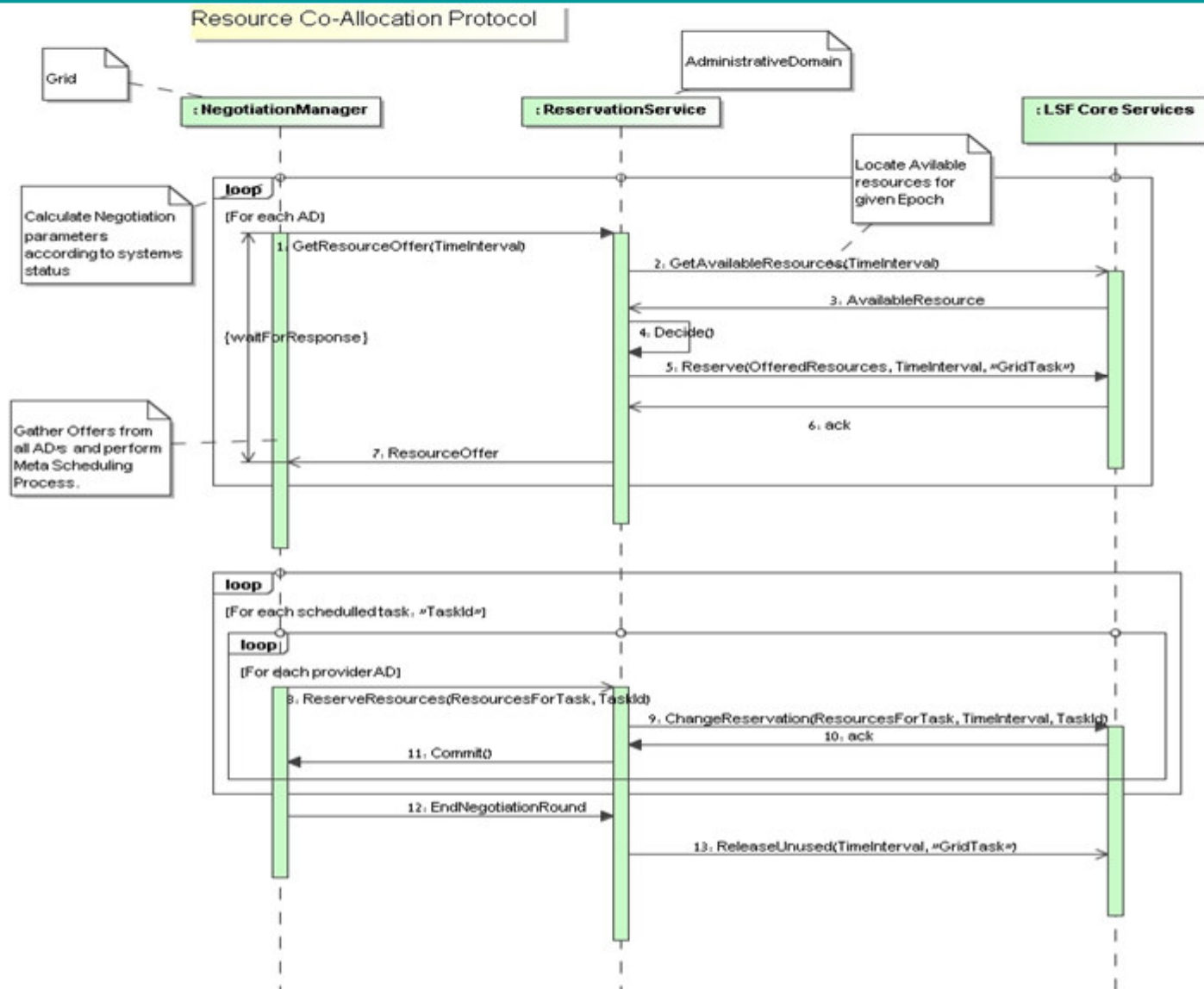
Highlights:

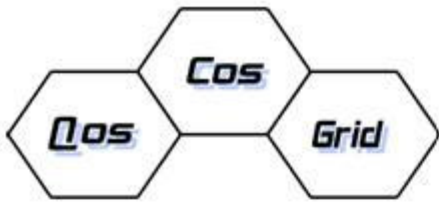
- Simple negotiation strategy, prevent “bargaining” which would prolong the number of negotiation rounds.
- Global guaranteed resource pricing methodology, allows each AD to estimate theirs potential profit.
- Each AD decides what to contribute, and provides a guaranteed “Resource Offer” accordingly.
- Support co-allocation scenarios, requiring a coordinated booking from different ADs.



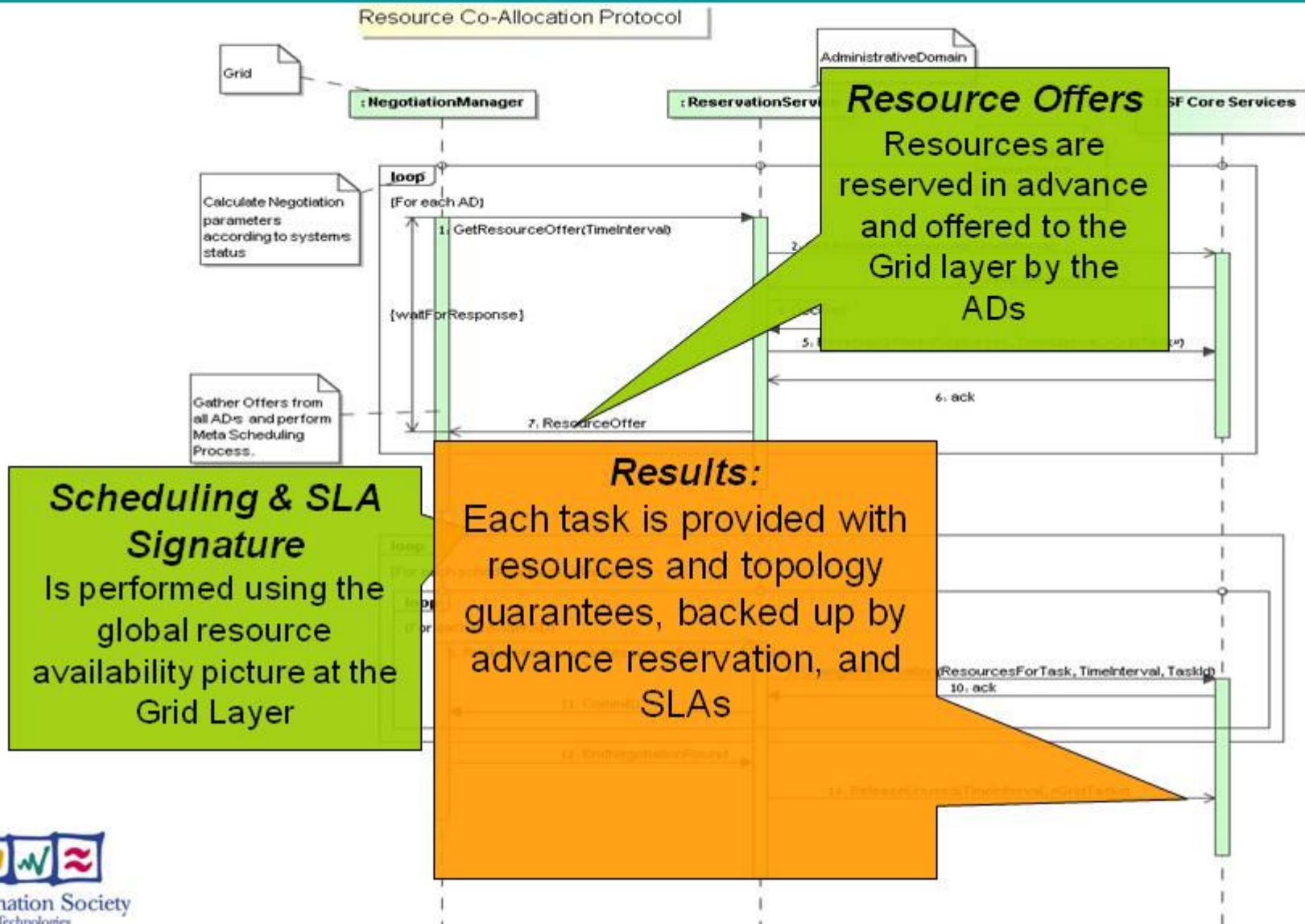


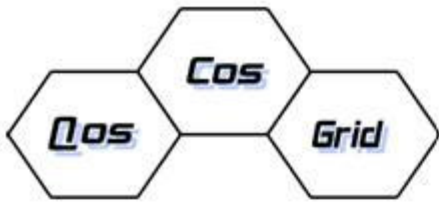
Resource Co-Allocation Protocol



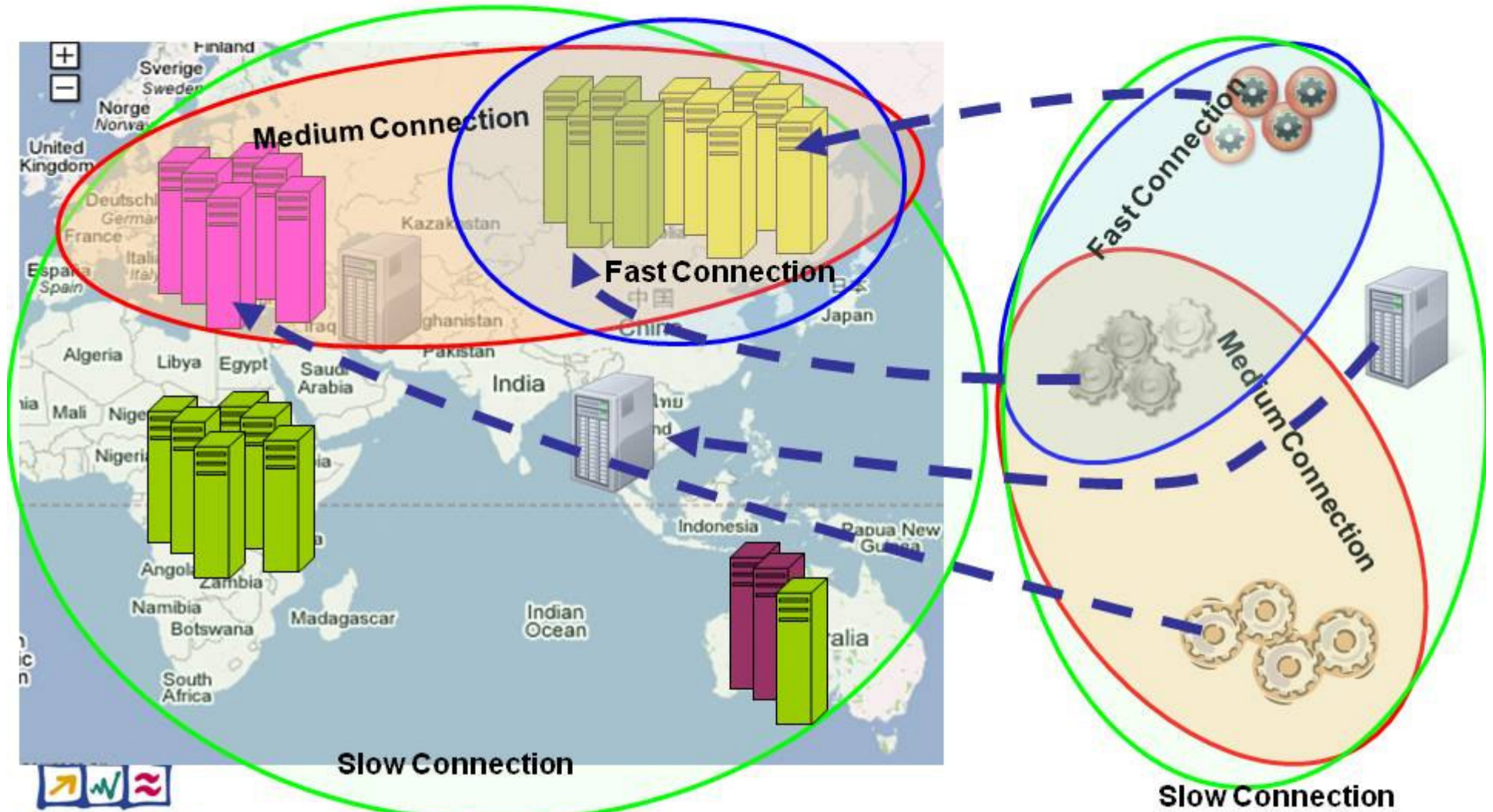


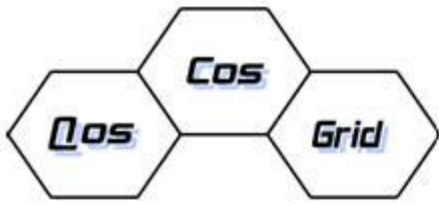
Resource Co-Allocation Protocol





QosCosGrid Meta Scheduler

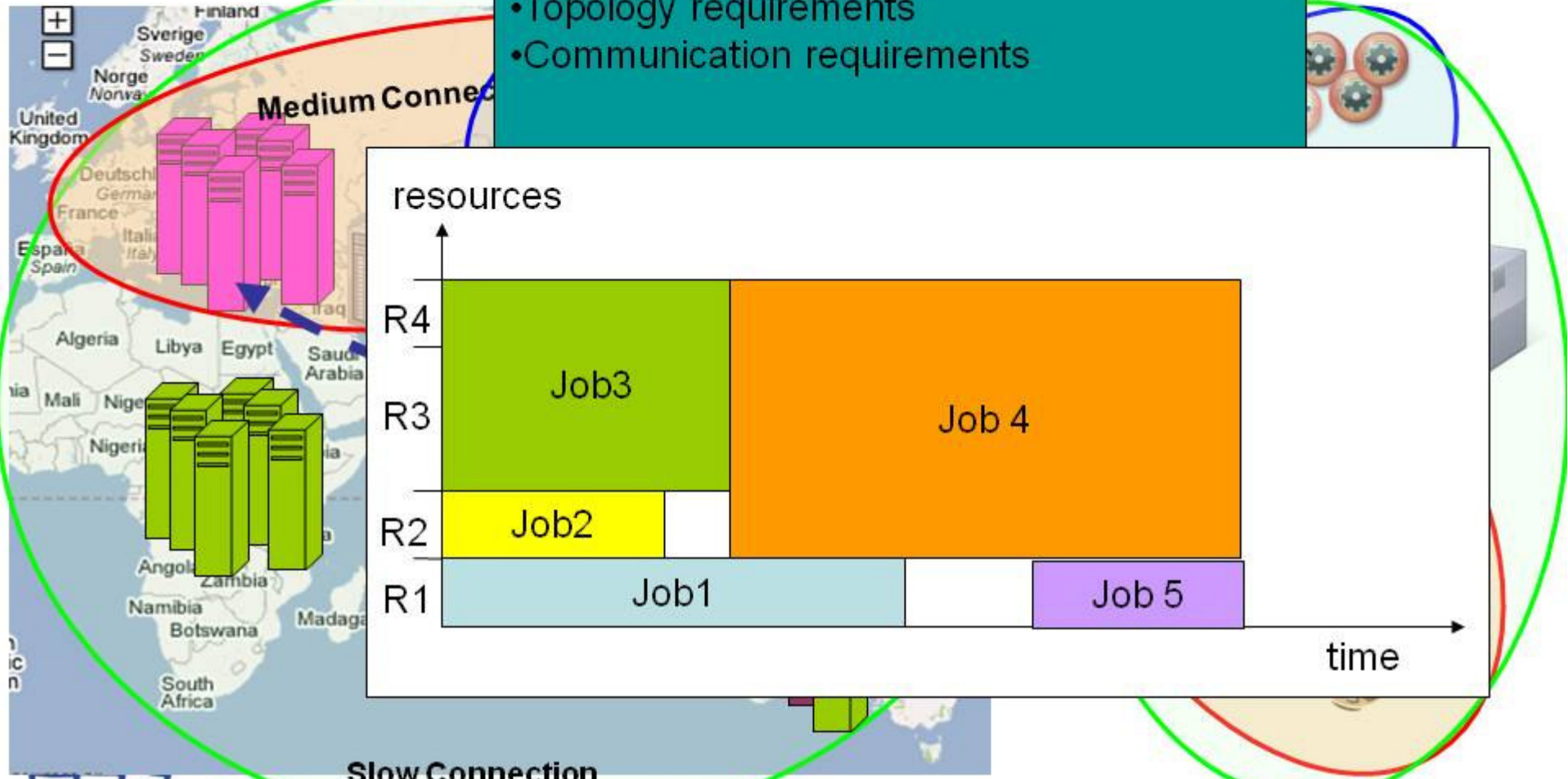


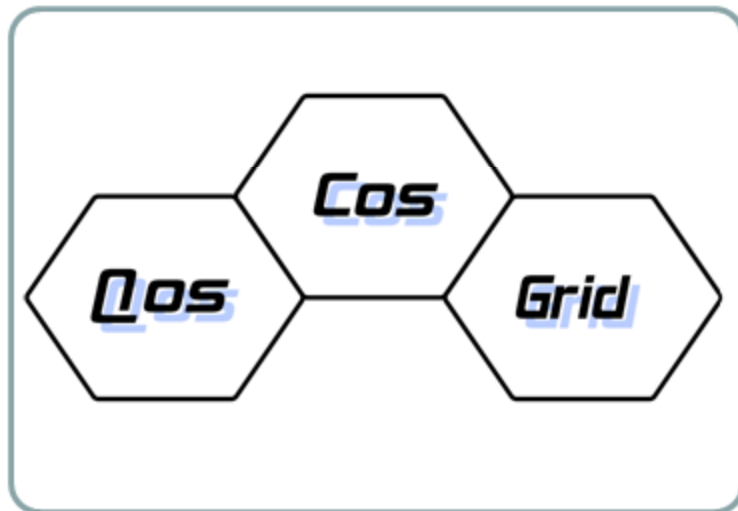


QosCosGrid Meta Scheduler

To provide a resource allocation for the user's tasks, constrained by:

- Resource requirements
- Topology requirements
- Communication requirements

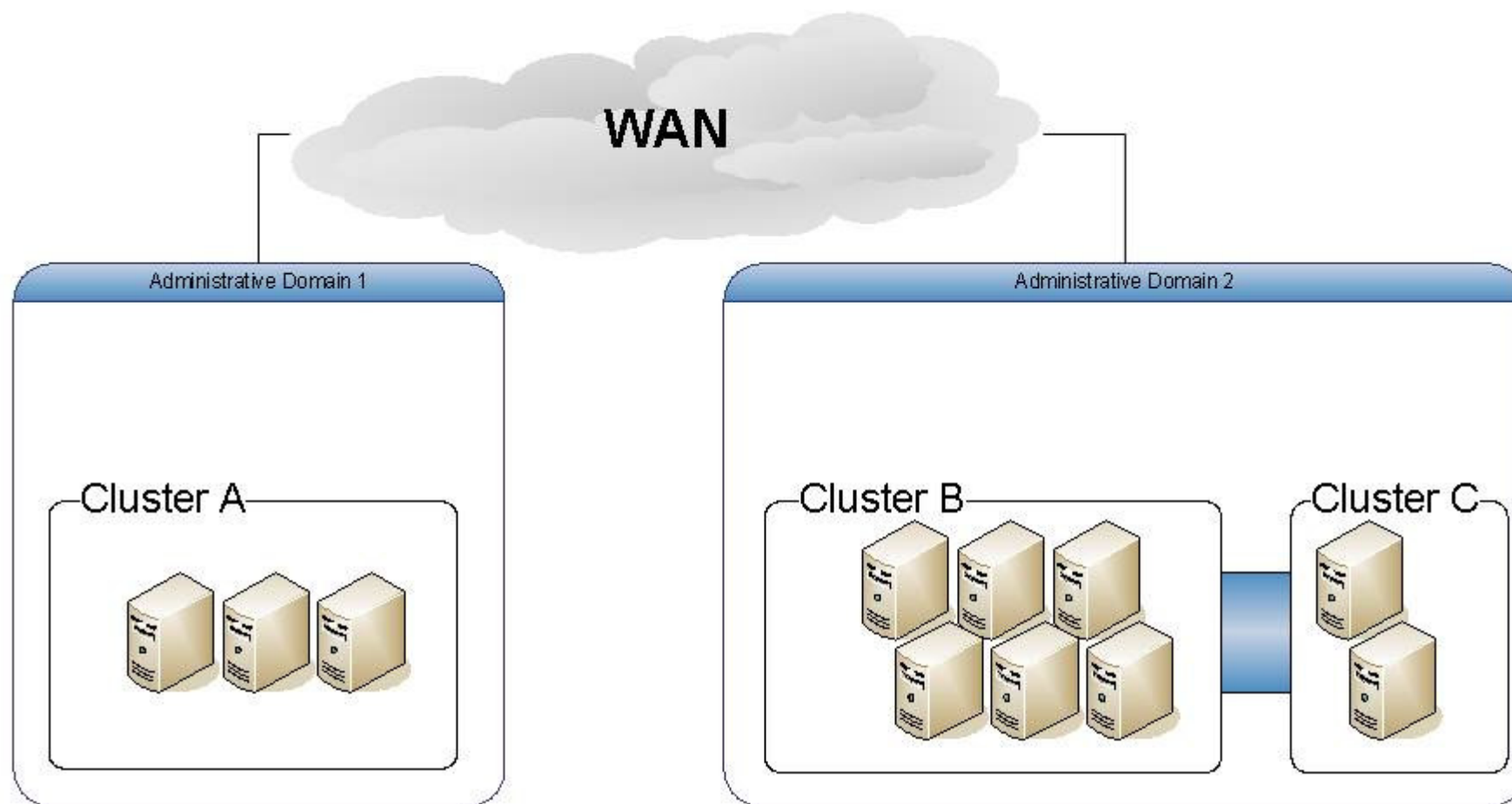


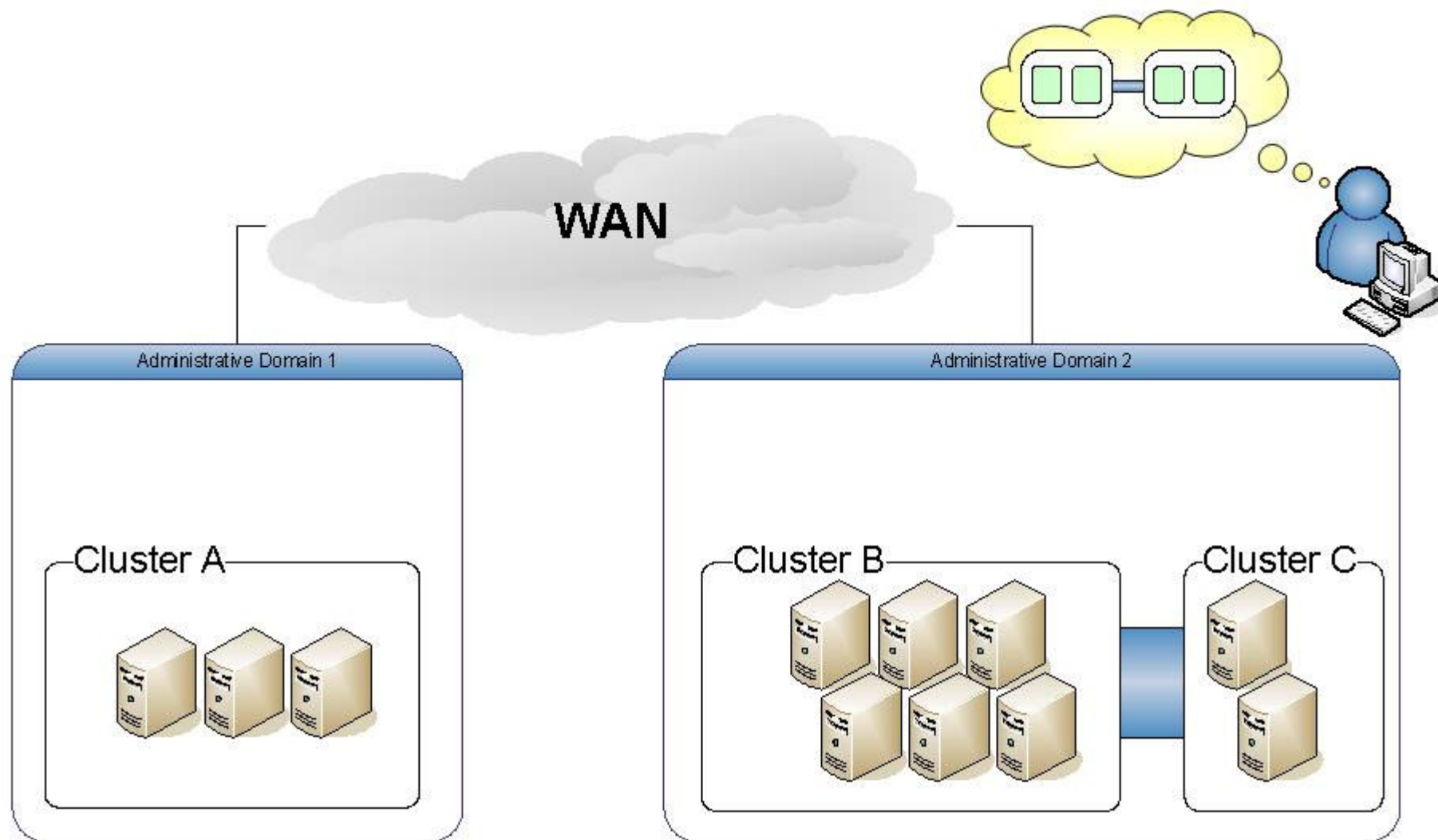


Topology aware QCG-OpenMPI Integrations

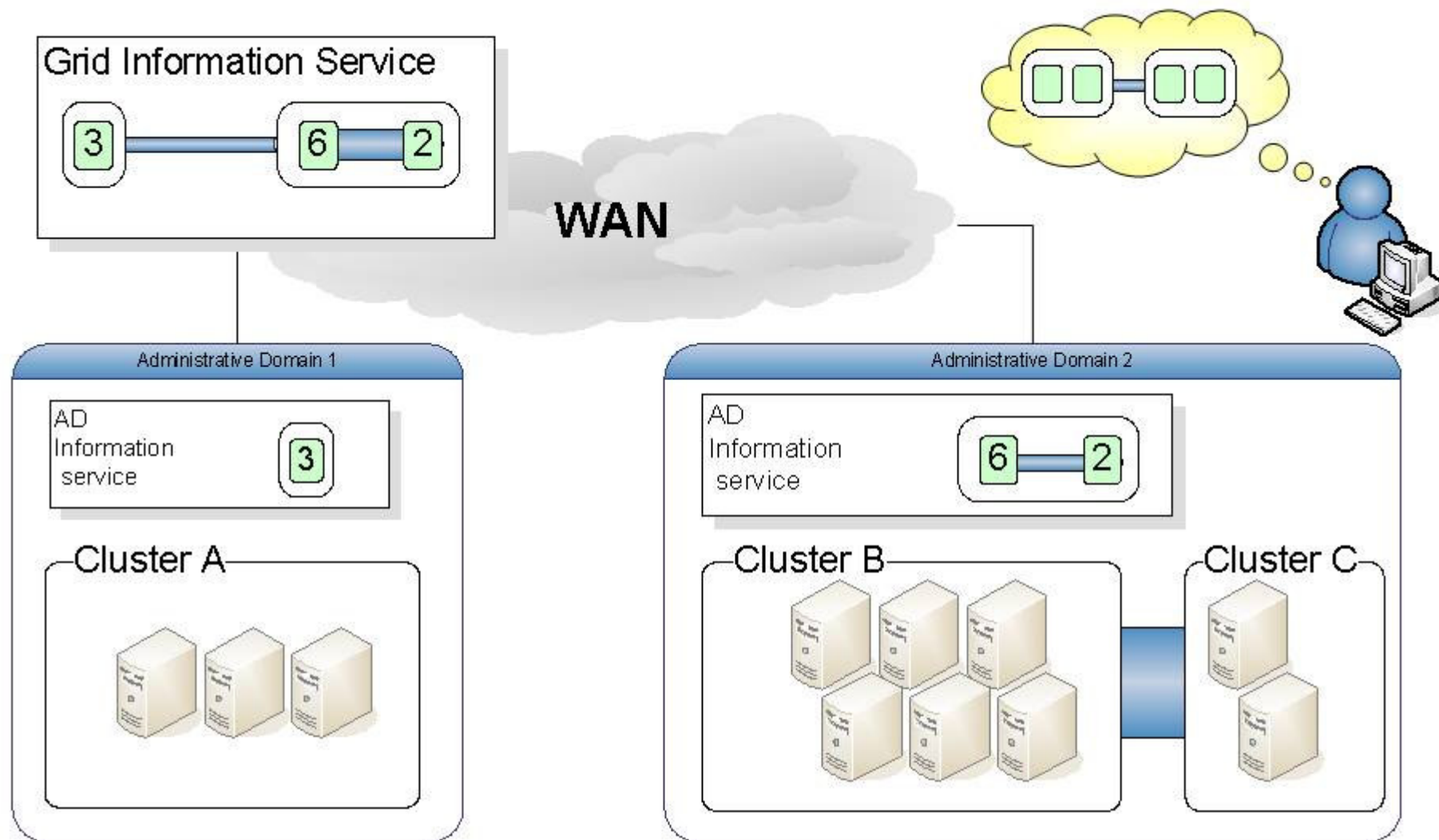


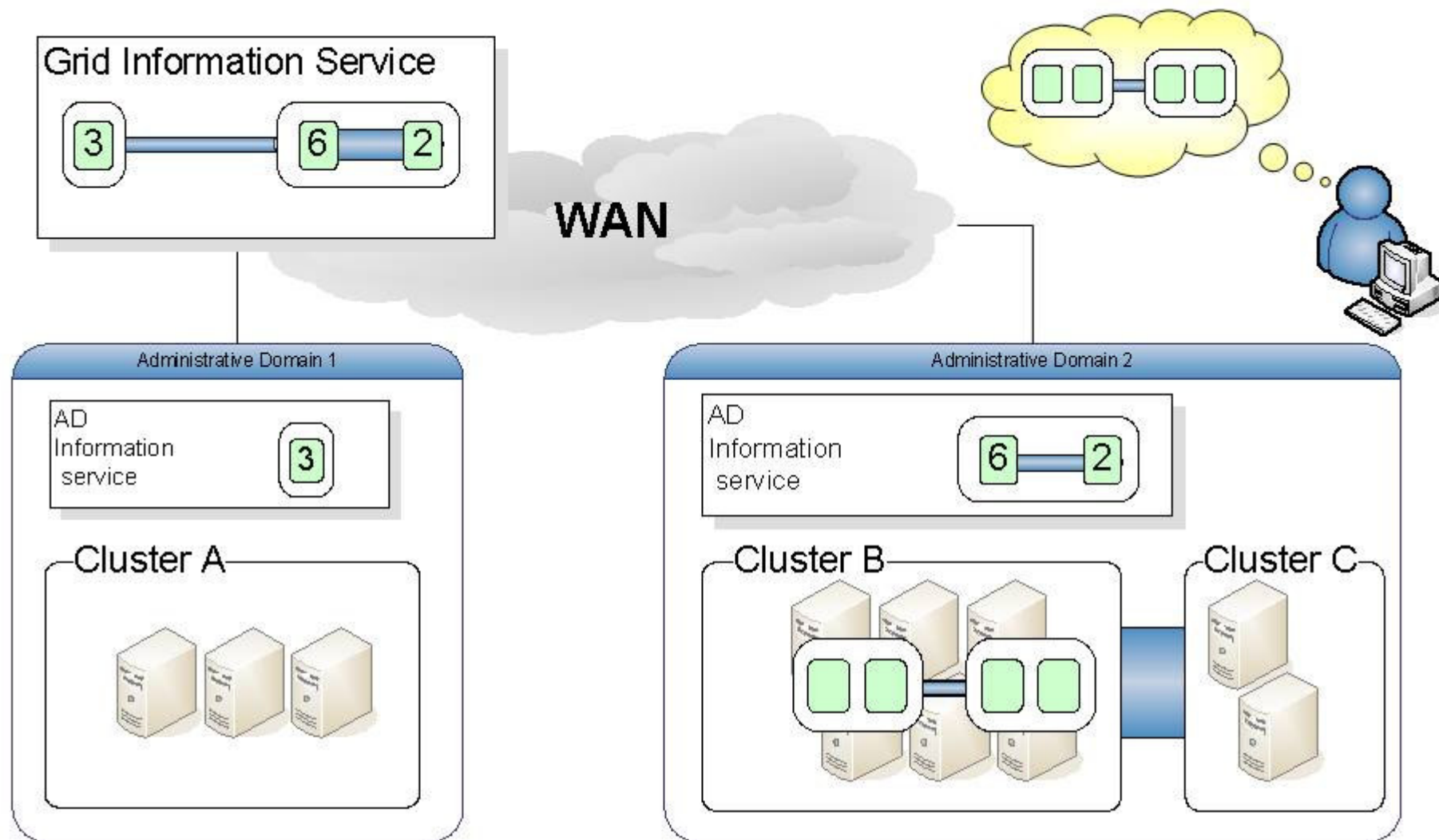
Example QosCos Grid

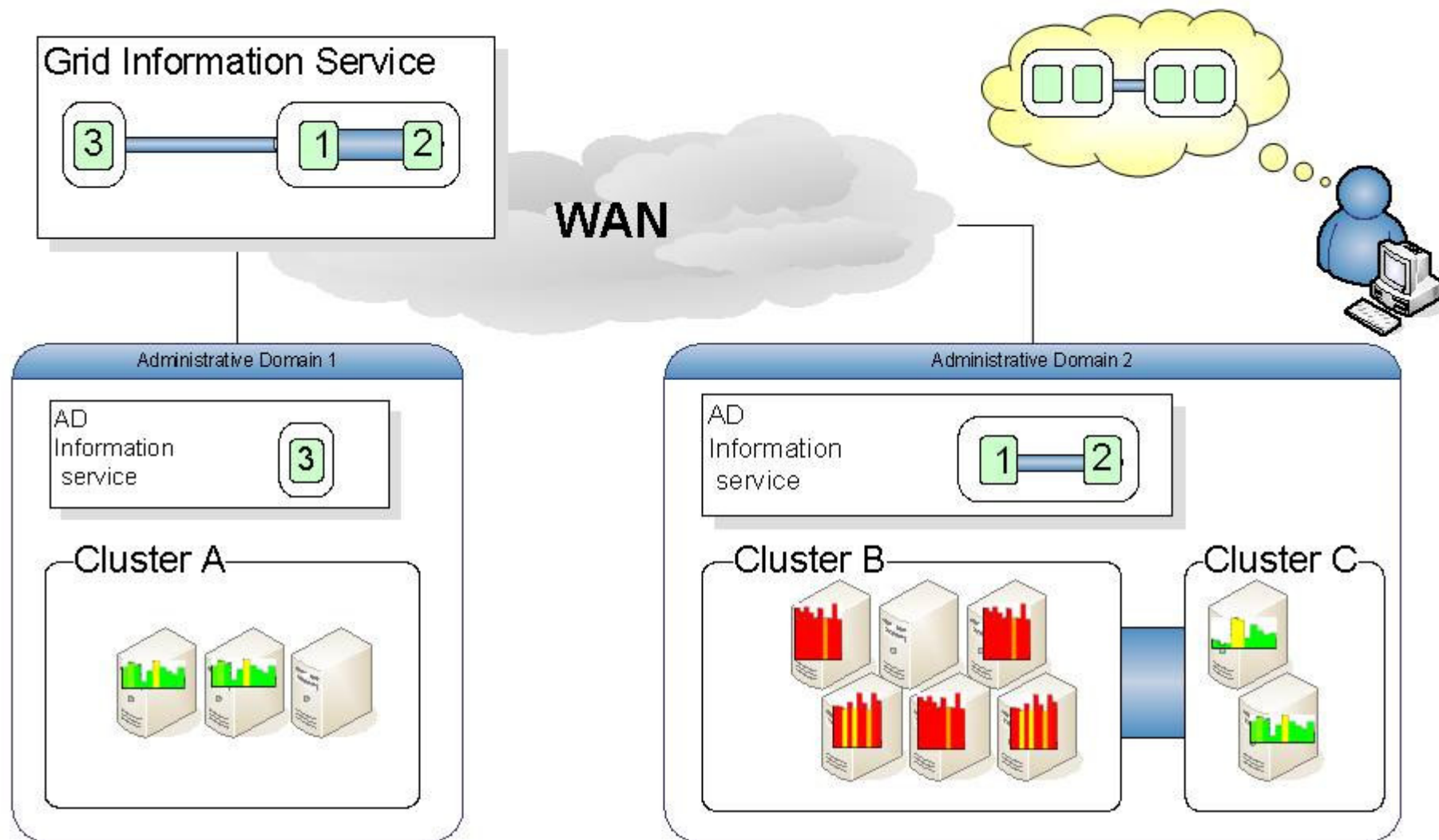




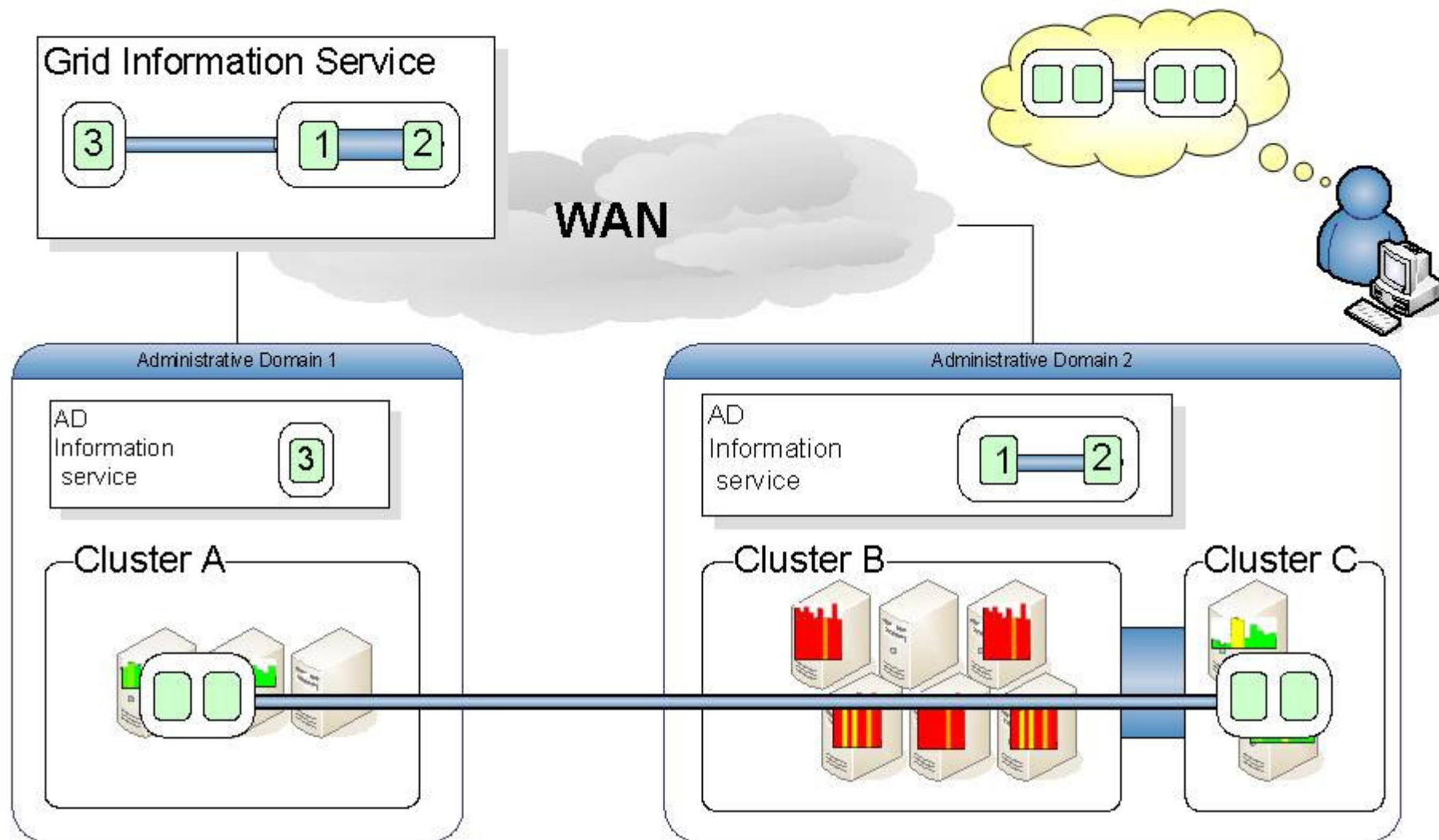
Job scheduling

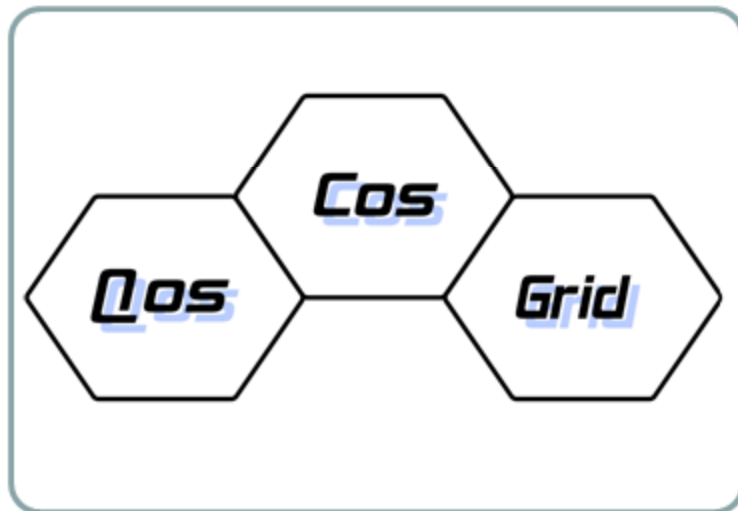






Job scheduling





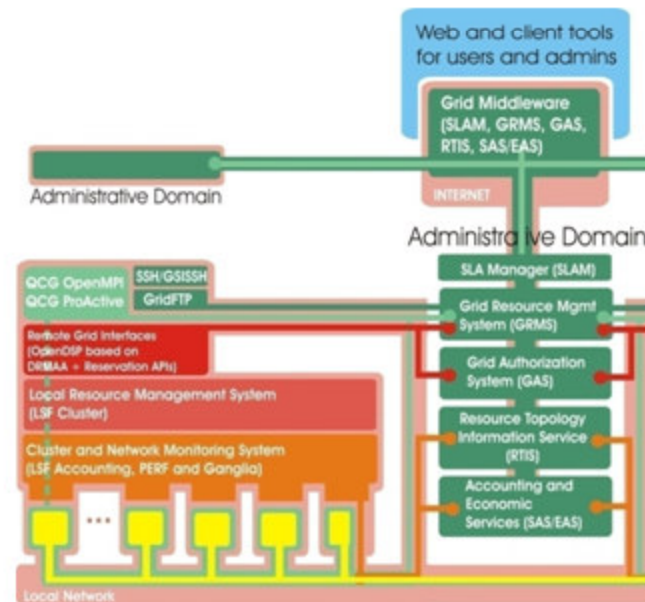
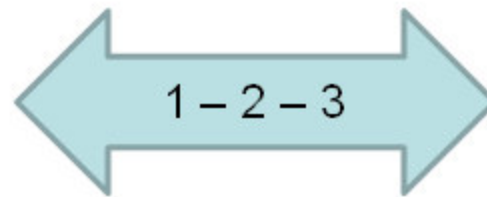
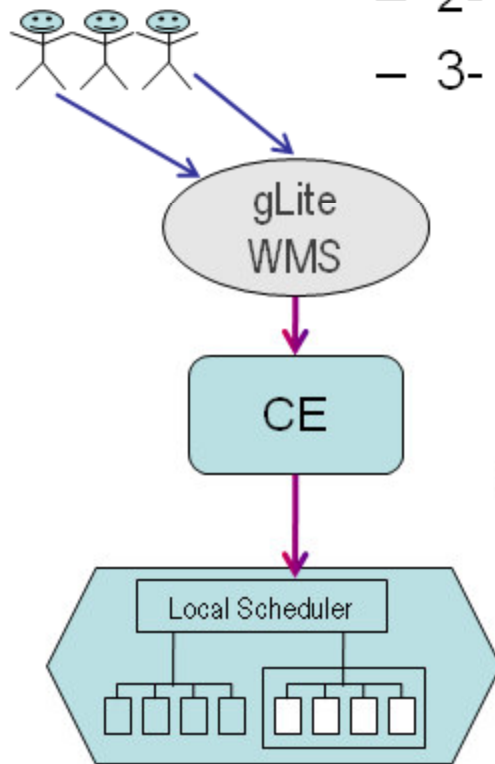
QosCosGrid & EGEE:
Run Complex Systems
via gLite



gLite integration ↔ QosCosGrid

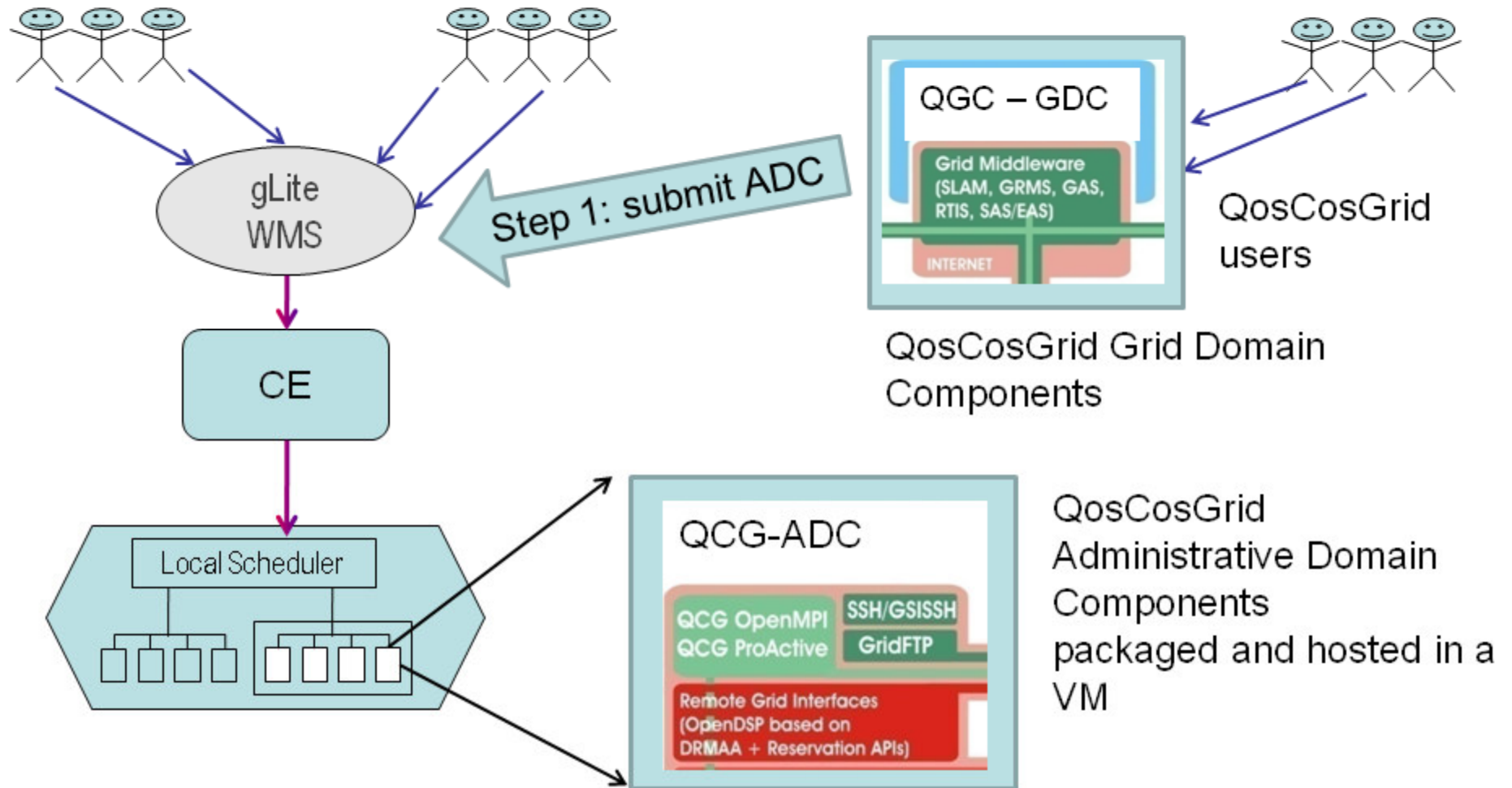
- QosCosGrid – gLite integration 1-2-3

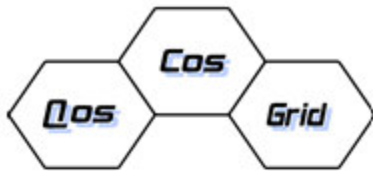
- 1- run job “QosCosGrid-ADC”
- 2- Connect QCG-ADC with LRMS and run MPI jobs
- 3- integration with monitoring services, WMS, CE (?), ..



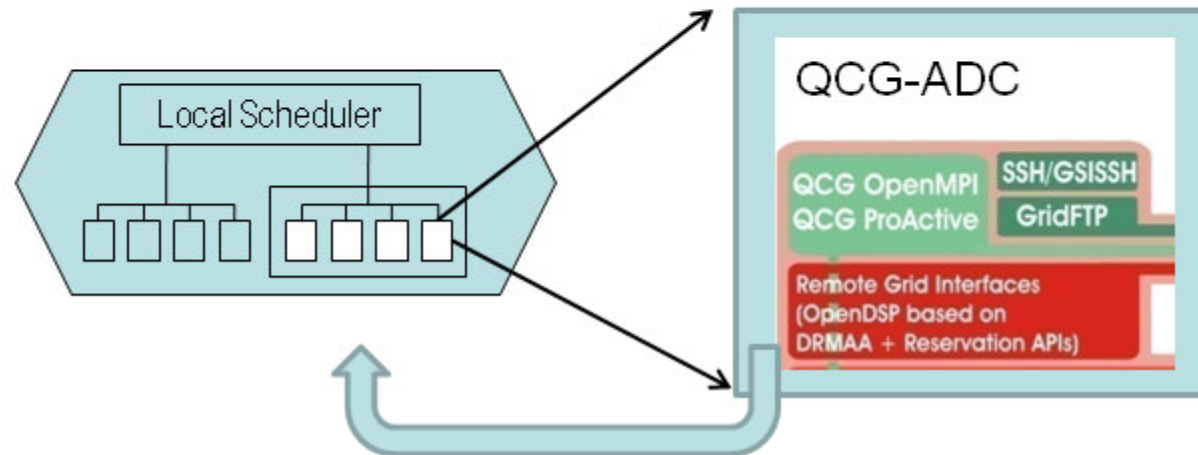


Step 1: Start “QosCosGrid-ADC”





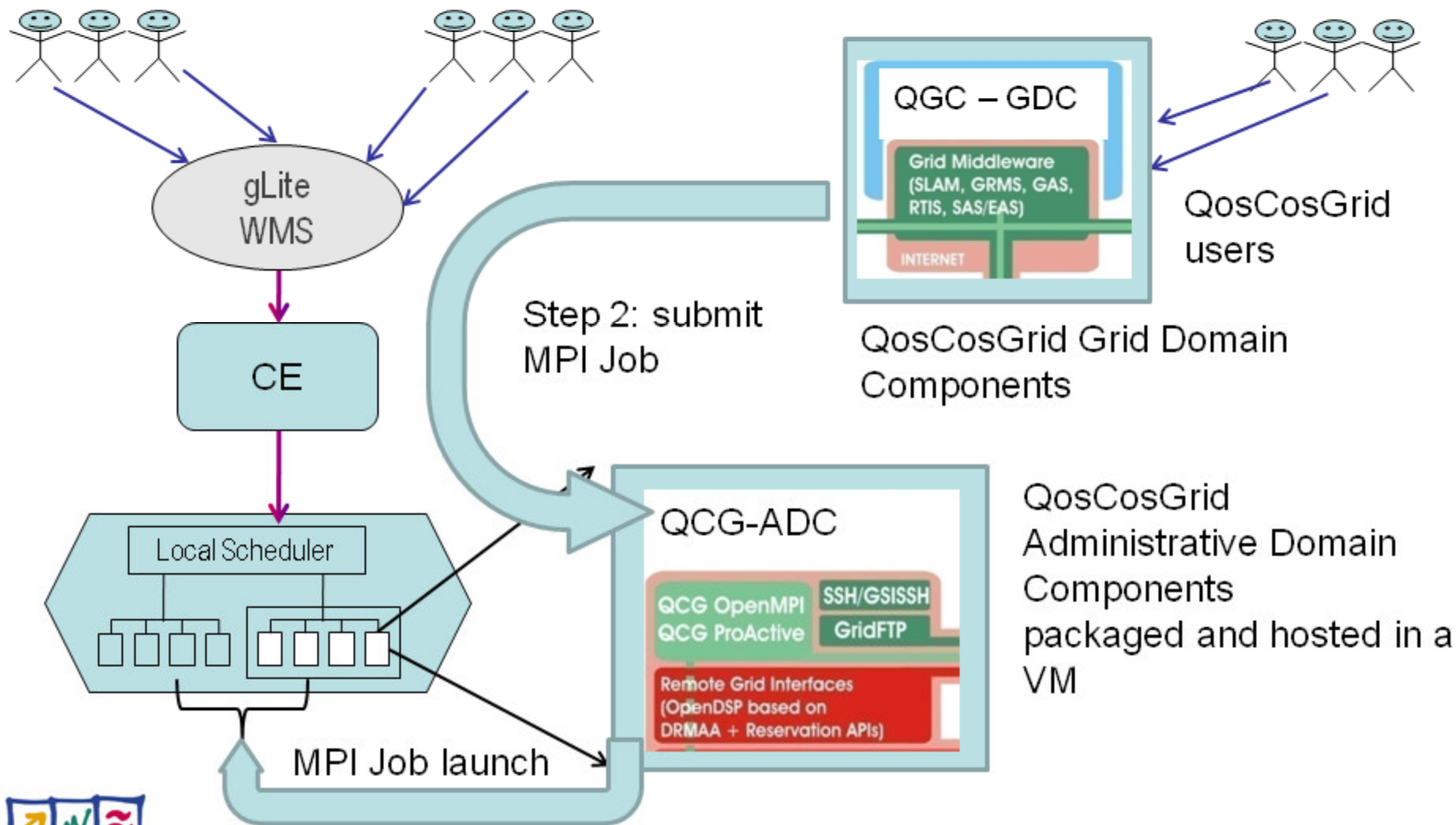
Step2: connect QCG-ADC ⇔ LRMS



- Submit Adv.Resv. and resource requirements
 - This may need admin preparation
- Receive resource allocation fix QoS SLA
- Interface (APIs) to monitoring and resource usage accounting
- Autodeploy QCG-OpenMPI
- QosCosGrid ADC connect to QosCosGrid GDC
- Ready to run Parallel Job



Step2: start parallel job

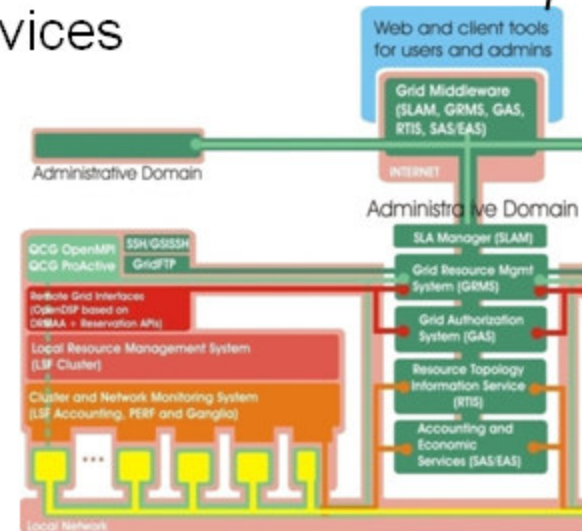
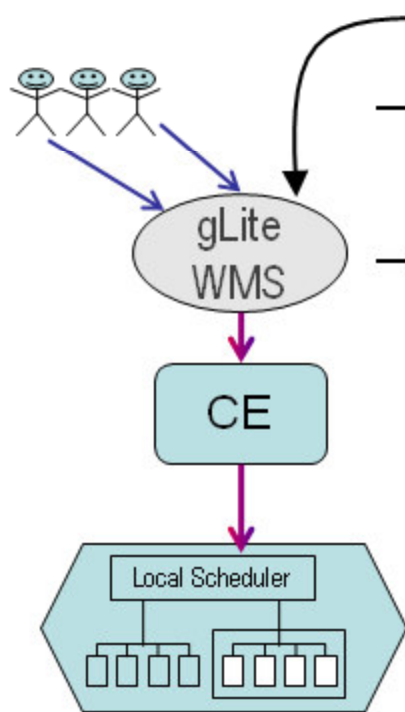


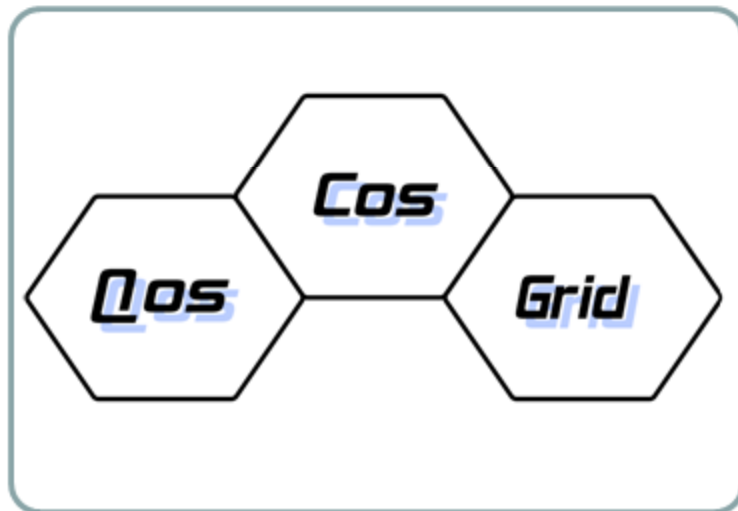


Step 3: gLite integration for workload management & monitoring

- QosCosGrid – gLite integration for workload management and monitoring – TO DO / To Decide

- 1- run job “QosCosGrid-ADC”
 - WMS “awareness” for this particular job type
- 2- Connect QCG-ADC with LRMS and run MPI jobs
 - Certificates handling / multi-node authorization
- 3- integration with monitoring services
 - Reporting on usage planned
 - Reporting on progress
 - ...
 - Open for more ideas





Sustenance for
QosCosGrid:
Working with industry
And research



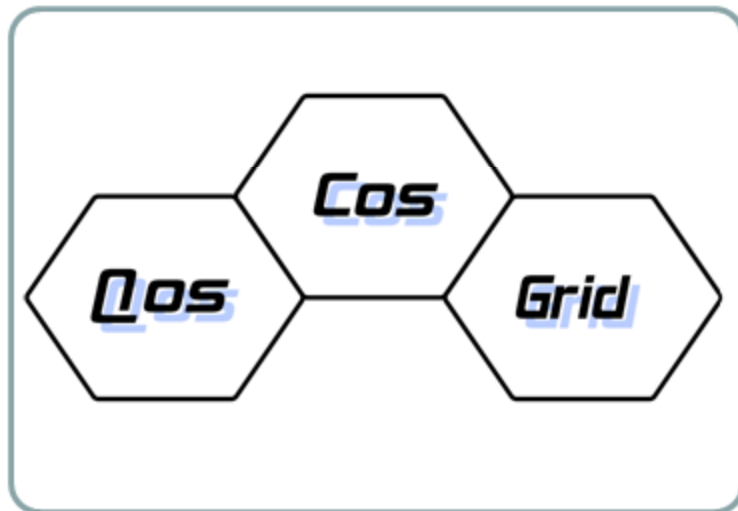
Sustenance: continue after project end

- QosCosGrid has actively pursued industry opportunities from the very start of the project.
- All industry (and even academia) contacts requested a trustable sustenance plan as prerequisite before engaging.
 - Actually, to “open source” the technology was appreciated, but not regarded as sufficient sustenance solution.
- How to live on after project end?
 - Example success story: UNICORE FORUM
 - Launched in 1998
 - Sustained solution availability ever since.



Sustenance: after project end

- Solution: Foundation of a charitable association
 - Persistent contact point for user and technologists
 - Support forum, maintenance, documentation, ...
 - Partner in new research proposals
 - Funded by fees and donations (Tax advantage!)
- Start date: end of March 2009
- Starting with QosCosGrid Grid Technology, but:
 - Open for all distributed computing techniques!
- ... why not gLite?



Open for all distributed computing technologies:
an invitation for gLite



Open for Grid business

- Publicly funded ICT research projects are practically disabled by regulations to work with industry
 - Have you ever tried to deliver value to industry and in return receive money? Ask your legal department!
 - How about commercial workload on GEANT?
 - No go!
 - Could you sign a support contract with an industry user?

How to pursue industry engagements at all?



Open for Grid business

- A charitable association is free to work with industry and create revenue – e.g. to be used to support research.
- Industry and commerce interested in Grid Technology can work with such a charitable association – leveraging know-how from Grid research projects like QosCosGrid and ... potentially ... EGEE ?



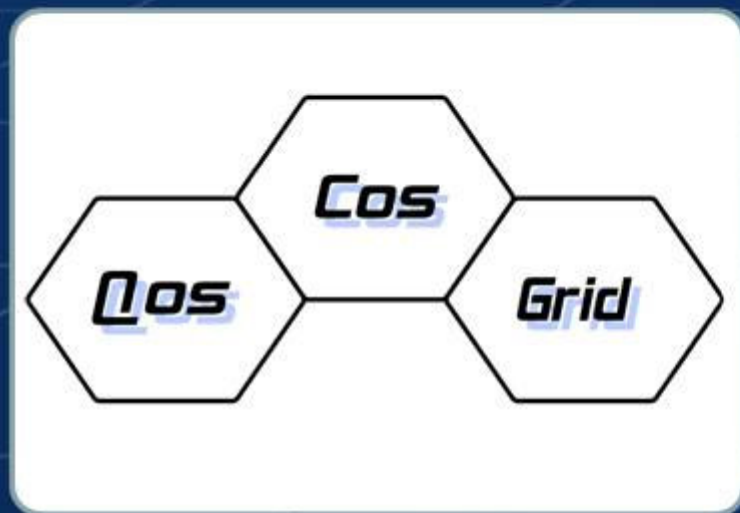


Summary

- [QosCosGrid](#) is a research project providing Grid solutions for parallel and complex-systems workload
- At EGEE'08, the concept how QosCosGrid enhances gLite / EGEE capabilities, adding new user groups and workload types was presented.
- QosCosGrid charitable association plans to include and embrace more technologies, accelerate Grid uptake in academia and **industry** – commercial environments.
- How about facilitating joint gLite/QosCosGrid deployments for industry?
- Looking forward to discuss and share with you!



Bernhard Schott
QosCosGrid Exploitation Manager
EU-Research Program Manager
Platform Computing
Mobile: +49 (0) 171 6915 405
Email: bschott@platform.com
Skype: [bernhard_schott](#)



Thank you !