GRAAP-WG

Grid Resource Allocation Agreement Protocol Working Group

GGF7

Tokyo, Japan

March, 2003

Chairs

Volker Sander
Forschungszentrum Jülich GmbH, Germany
v.sander@fz-juelich.de

Jon MacLaren
Manchester University, UK
jon.maclaren@man.ac.uk

Wolfgang Ziegler
Fraunhofer–Institute for Algorithms and Scientific Computing, Germany
Wolfgang.Ziegler@scai.fhg.de

Agenda

1. Charter and clarifications. Discuss web-page clarifications about the group and its purpose - anything more required?
2. Links to other groups. Anything to report from our existing links? Find people to link to the groups we don't have people for yet. Are there any groups to add to the list? Should we name the link people on the web site?
3. Finalize definition of Advance Reservation.
4. State-of-the-art. Discuss final document/ additions.
5. Discussion of requirements (using use-cases document).

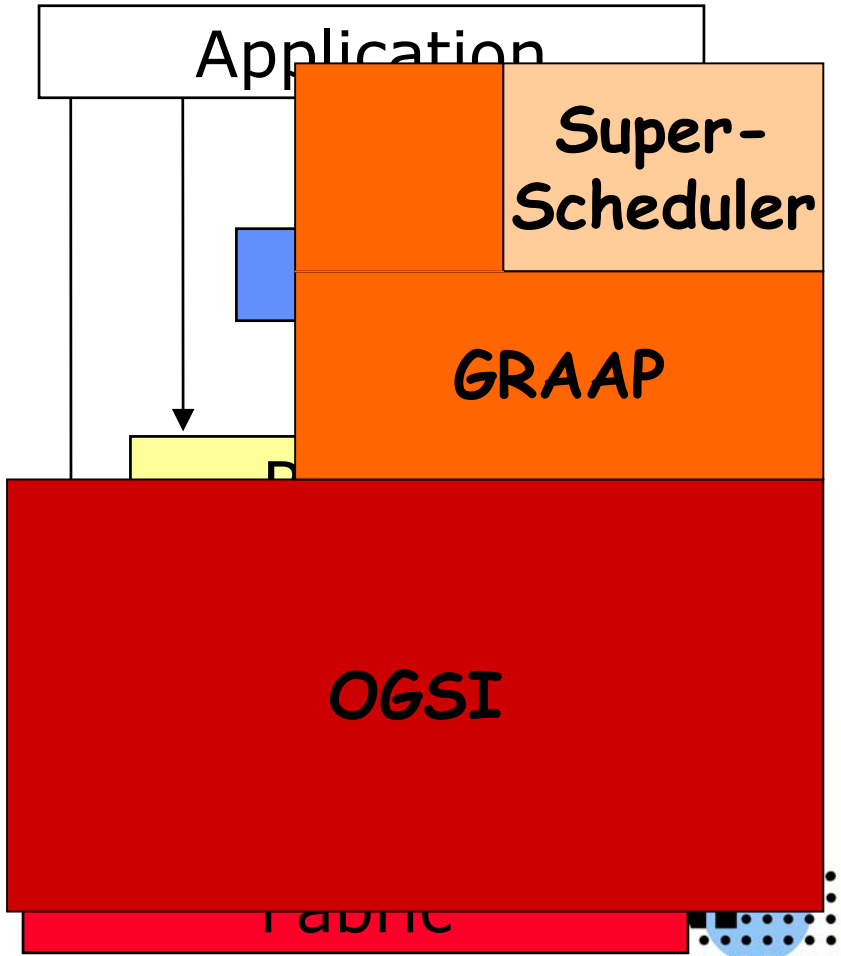
Layered Grid Architecture

"Managing multiple resources": ubiquitous infrastructure services

"Sharing single resources": negotiating access, controlling use

"Talking to things": communication (Internet protocols) & security

"Controlling things locally": Access to, & control of, resources



1.1. Charter: Revised Milestones

Milestones (including past):

End of May:	First draft of the charter ready **(done, charter approved)**
GGF-5:	Discussion of the charter, SchedWD 12.2, SNAP; next Steps **(done)**
GGF-6:	Grid RAA Protocol: Discussion of Use Cases to procure requirements **(done)**
GGF-7:	Grid RAA Protocol: Description of Requirements ready
GGF-8:	Grid RAA Protocol: Description of Operations ready
GGF-9:	Description of Leverage/Interaction with other Grid Service Standards
GGF-10:	Grid RAA Protocol: First Description of Bindings ready
GGF-11:	Final Grid RAA Protocol specification ready

1.2. Charter: Clarifications of Scope

To clarify the purpose of the group, we have added the following clarification to our Web Page:

- We are not working on a co-scheduler or super-scheduler, although the protocol could be used (as an enabling technology) by a co-scheduler.
- For now, GRAAP-WG will consider the reservation of timeslots on computational resources as the key use of the protocol. We hope, in time, that the protocol can be used for other resources, such as networks, but from a discussion at GGF6 we realized that there are complex issues regarding network reservation (do you also include the reservation of ports for communication, or allow the protocol to reconfigure firewalls?). Although we will try and be as general as possible, we cannot take all possible uses for the protocol into account.
- The group is not currently working on a Resource Description Language, and will not be for the near future. Our protocol will need to communicate in some RDL (or perhaps several), but we will wait and see if a suitable RDL emerges from the community. In the meantime, we may well produce some requirements for a suitable RDL.

2. Other GGF Groups to Link To

Existing Groups, who we have links with:

Open Grid Service Architecture (OGSA-WG)	Karl Czajkowski
Grid Protocol Architecture (GPA-WG)	John Brooke
Grid Security Infrastructure (GSI-WG)	Karl Czajkowski
CIM-based Grid Schema (CGS-WG)	Jon MacLaren
Scheduling Dictionary (SD-WG)	Wolfgang Ziegler

Existing Groups, who we want links with:

- Distributed Resource Management Application API (DRMAA-WG)
- New Productivity Initiative (NPI-WG)

Questions:

- Are there any more groups to add to the list?
  - JSDL
  - CIM based Grid Schema
  - Scheduling Architecture
- Do we have volunteers to fill the gaps?
- Do the link people have anything specific they'd like to report?
- Do we need (or want) to name the link people on the web-site?

3. Definition of Advance Reservation

The definition has being discussed at some length on the mailing list, and is currently:

- "An advance reservation is a possibly limited or restricted delegation of a particular resource capability over a defined time interval, obtained by the requester from the resource owner through a negotiation process."

Discussion?

4. State-of-the-Art: Current state

Document now in HTML format at:

http://people.man.ac.uk/~zzcgujm/GGF/sched-graap-2.0.html

A PDF of the document's current state is at:

- http://www.gridforum.org/Meetings/ggf7/drafts/sched-graap-2.0.pdf

Document is now almost finished, it just needs some tidying up.

Key changes this time:

- Definition of Advance Reservation now matches the current GRAAP-WG definition.
- Added contributions on PBS, Sun GridEngine, Condor. Thanks to al who contributed to this!

Anything else to add to this document?

Any other schedulers?

If not, will tidy up, pass round group, then submit to document process.

Use cases Document

Karl Czajkowski (karlcz@isi.edu)
Stephen Pickles (Stephen.Pickles@man.ac.uk)
Jim Pruyne (pruyne@hpl.hp.com)
Volker Sander (v.sander@fz-juelich.de)

Scope of the Document

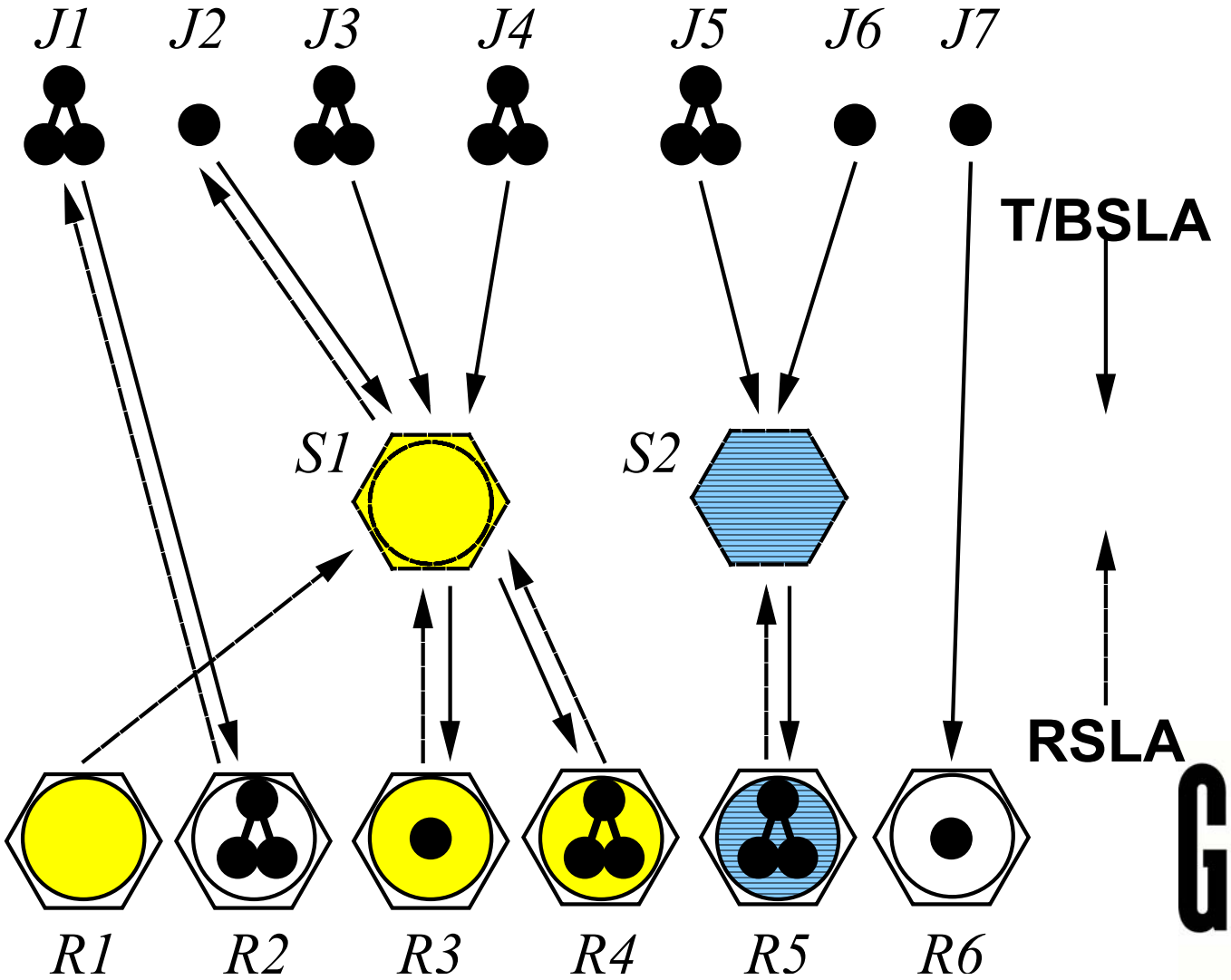
Description of Use Scenarios for GRAAP

Build a base for a common understanding

- GRAAP is about to establish service level agreements
  - An SLA allows clients to understand what to expect from resources without requiring detailed knowledge of competing workloads or resource owners' policies
  - SLAs are useful to model advance reservations
- Services might be composed of different levels
  - The acquisition of a service by some end-user requires the transitive access to all agreements
  - Resource SLA (RSLA)
    - A promise of resource availability
      - *Client must utilize promise in subsequent SLAs*
    - Advance Reservation is an RSLA
  - Task SLA (TSLA)
    - A promise to perform a task
      - *Complex task requirements*
      - *May reference an RSLA*
  - Binding SLA (BSLA)
    - Binds a resource capability to a TSLA
      - *May reference an RSLA (i.e. a reservation)*
      - *May be created lazily to provision the task*

Derive Protocol Requirements

Scope of the Document



Scenarios

The very long running service or application

- A user has developed a service that they wish to provide for the grid community at large (e.g. Netlib).
- Requirements :
  - Must be able to specify indeterminate run-length or completion time

The variable demand long running service

- Demand varies, so resource demand varies
  - Experiments might produce peaks of data which have to be processed
- Requirements:
  - Must be able to specify “demand profile.”. Often Periodical
    - Deterministic
    - Probabilistic
  - Peak demand, Average Demand

The configurable application

- An application that can be configured relative to its resource availability
  - Constraints about acceptable service levels
- Requirements:
  - Must be able to specify alternatives.
    - Alternative selection requires user’s criteria
    - Scheduler should be able to provide this information based on constraints

Scenarios

- **The "Templated" Application**
  - Application developer and application user are distinct
    - Developer provides "template" describing app. demands to be filled in by user based on their needs, environment
    - E.g. A proprietary scientific application
  - Protocol Requirements
    - Introduce "free variables" in to specification
      - Ideally templates are "pre-registered" and named by the developer
      - User refines the level of service within the templated space
    - Composition of existing SLAs
- **Monitoring of resource utilization and changing of reservation parameters**
  - Progress of an application indicates a changed requirement
  - Protocol Requirements
    - Monitoring of existing SLAs
    - Re-negotiations (following the model of GFD-E.5)

Scenarios

- **File Transfer Scenario**
  - Applications are often fragmented into phases which might be executed in different hosting environments
    - Pre- and Postprocessing might require staging operations
    - Staging should be coordinated with further operations
  - Protocol Requirements
    - Synchronization of Phases
    - Deadline File Staging
- **Co-Allocation of Several Resources**
  - Coordinated mapping of complex Workflows
  - Protocol Requirements
    - Interfaces to various resource management systems
    - Composition of SLAs
    - Re-negotiation of SLAs

Scenarios

- **The "bottom-feeder" application**
  - Acquisition of unused resources
    - HTC
    - SETI@Home or a distributed.net program
  - Protocol Requirements
    - Propagation of Policy information
    - Notification about changes of the resource level
- **Complex Workflows**
  - Emphasizes the aspects of synchronization
  - Question: Synchronization by BSLA and/or TSLA

A use case from the RealityGrid project (1/9)

ntroduction (1/3)

The RealityGrid project (http://www.realitygrid.org) aims to predict the realistic behaviour of matter based on the properties of the microscopic components using diverse simulation methods (Lattice Boltzmann, Molecular Dynamics and Monte Carlo) spanning many time and length scales and the discovery of new materials through integrated experiments. A central theme of RealityGrid is the facilitation of distributed and collaborative exploration of parameter space through computational steering and on-line, high-end visualization.

A use case from the RealityGrid project

mmediate Requirements for Advance Reservation

- The most pressing requirement for advance reservation in RealityGrid arises out of the need to co-allocate (or co-schedule)
  (a) multiple processors to run a parallel simulation code and
  (b) multiple graphics pipes and processors on the visualization system
- Co-allocation may be required now (either by a RealityGrid developer or by a scientist engaged in routine investigations) or at some more distant time in the future (for a scheduled collaborative session). We expect advance reservation to subsume both co-allocation scenarios.
- The visualization resources (b) will usually be located on a different system to the computational resources (a).
- The two sets of resources ((a) and (b)) will often be located on systems owned and administered by different organisations, and the administration teams within the two organisations, if aware of each other's existence at all, are unlikely to have established comprehensive Service Level Agreements.

5. Requirements?

- **Service requesters can make and relinquish advance reservations**
  - Infinite / Indefinite lifetime possible
  - Specification of a variable demand profile
- **Service requesters must be able to re-negotiate advance reservations, both before the reservation begins, and during the reservation's active time**
  - Malleable reservations
  - Reducing the Service Level should succeed
- **Resource owners must be able to initiate re-negotiation (or even cancellation) of advance reservation**
  - Both parties of an SLA can initiate renegotiation
  - Note that the related adaptation step is an optional feature which does not have to be supported by all clients.
- **Re-negotiations should be possible for active reservations (both user and owner initiated)**

5. Requirements?

- **The protocol should be robust**
  - Life Time of Reservations
  - Soft-State, i.e. an SLA has to be refreshed
  - Monitoring and Feedback mechanisms
- **The protocol should operate at multiple stages, each of representing the negotiation for different type of agreements**
  - Composition of SLAs
  - Might be heterogeneous
  - The protocol should support the composition of existing agreements
- **The protocol should be able to cope with different granularities, from course to fine**
  - Follows from composition
- **Propagation of policy information**
  - SLA policies, i.e. constraints and priorities

5. Requirements: What RDL do we rely on?

- **The GRAAP protocol will carry a description of the resources that are being negotiated, specified in some Resource Description Language or RDL**
  - We want to model SLAs
  - The language should be the same for service providers and service requesters
- **We want GRAAP to be RDL-neutral, i.e. to be able to handle any well-defined where the resource description can be encoded as some sort of byte-stream, e.g. string**
  - Abstraction
- **We also want to be able to use and test GRAAP**
  - So do we need to write a simple exemplar RDL, e.g. for blocks of supercomputer time?
  - Or can we wait for a while to see if the Grid community will do this for us?
- **Must stay within the scope of our group…**
  - Abstraction might help here again
  - We need to communicate our requirements

5. Discussion of Requirements