

## Minutes from the Workflow Management Research Group, October 8, 2003

Proposed Chairs: Ewa Deelman, Jim Blythe, Miron Livny

Ewa Deelman, ongoing workflow management work (see slides)

Key points of the talk

- GriPhyN project; virtual data materialized on demand
- Seamless delivery to user/application regardless of form in which data exists
- Mapping abstract to concrete workflows
- Pegasus system, maps from abstract to concrete workflow
- What makes grid workflows different from business workflows?
- Large amounts of data; data transferred among workflow activities
- Nature of execution environment: dynamic, prone to faults
- Workflow refinements; modify part or entire workflow

Jim Blythe, Planning services for workflow execution (see slides)

1. User specifies workflow outline; user wants complete control of key aspects of workflow but do not want to specify all aspects of a large workflow  
Want flexible language for user's constraints; consider existing languages
2. System completes the user's workflow  
Need to reason about entire workflow
3. Execution and monitoring  
Environment may be highly dynamic; monitor during execution  
Capability to do repair during execution

Issues:

- Are existing workflow languages expressive enough?
- What QoS can workflow service provide?
- Can workflows be shared between users and groups?

Miron Livny

- Interested in issues involved in executing workflows, delegation of workflows from one layer to another
- Expressing rules for workflow
- Problems with monitoring workflow, generating repairs
- Can think of high-level programming languages as workflow languages
- Biggest challenge may be to define what workflow is
- Right now includes too much
- Need to carefully define what we are doing, perhaps partition into multiple groups

Comment Darren Pulsipher: don't limit group to data-oriented workflows only

Other organizations are concerned with control flow as well

Too restrictive to make this distinction

Miron: getting the data to the right place is the problem; distinguishing the amount of data is probably not useful; is a dataflow problem

Marian Bubak, Grid Application Workflow Composition (see slides)

CrossGrid architecture

Abstract workflow description

Automatic Flow composer: includes a composer and a registry  
Need for optimization  
Semi-automatic workflow composition: component registry, distributed workflow composition system, GUI portal  
Performance monitoring : standard monitoring interfaces for grid services

Anthony Mayer, ICENI experiences (see slides)

Workflow: aggregating tasks  
“programming the grid”, composition of services, ordering execution or communication similar but not semantically the same  
Batch workflows; also interactive with overlapping design and execution  
ICENI Stack; component-based system  
Separation of concerns: meaning, behavior, implementation  
Composition of services: relations represent communication channels (dataflow)  
Spatial composition allows dynamic interaction  
Infer temporal ordering/workflow from spatial view

Charter discussion

Do we need to exist? Is there interest in research group?  
Who is interested in actively taking part?  
Is charter appropriate?

Workflow in the OGSA context: seems more appropriate for a working group rather than a research group

Scientific workflows: should not be specifically state in charter; industrial workflows have basically the same problems and we should not limit the scope of this group

Suggestion: start with looking at use cases from different domains

Question: what needs to be standardized in the future? APIs or protocols  
A remote site needs to be able to interpret a workflow description  
Or develop standard interfaces for workflow systems  
Research group will not develop these standards but may eventually spin off working groups to work on these standards  
State this as part of the set of goals  
Another eventual standardization would be OGSA

Comment: this charter is very broad, includes almost everything

Miron: should we have a distributed workflow activity or does it make sense to bring all the activity under one roof?  
If we take a distributed approach, we need to partition these into discrete pieces

Dennis Gannon: Several groups already doing workflow

GCE group: focus on workflow specification; service composition in portal environments  
Semantic grid group: issues of workflow composition at the metadata/semantic level  
Life sciences: focusing on requirements for their community

Ewa: this group focuses on lower level; have a workflow specification; map into grid environment; discover resources; perform workflow repair

Suggestion: collect use cases; later decide which ones you are going to work on  
Many business use cases: look at MPI group

Interfaces that this group may care about:  
Also look at JSDL group: submit jobs  
DRMAA group also

Suggestion: this is an opportunity to tease out commonalities with existing groups, refactor

Comment: focus on ultimate plans for standardization for this group  
Language composition? Protocol exchange?

One-sentence description of this group: focus on workflow refinement and execution  
Might have multiple, distributed refiners

Working on a standard architecture and framework

Difference: other groups have the users explicitly selecting resources  
Want to automate resource selection

One-day workshop on workflow is planned at GGF10  
Look at overall set of issues, tease out the pieces of the puzzle

This group does not propose to do creation of the original workflow; semantic grid and GCE do workflow composition

Comment: if you limit yourself to refinement, then that is just implementation not research  
Others disagree: think that having a clear scope is a good thing

Miron: do we need something to bring all these workflow pieces together, has the overall big picture?  
Nobody has workflow in the top level definition

Ewa: suggests joint sessions among groups to address cross-cutting issues

Dennis Gannon: there will be overlap and spillover among the groups

Carol Goble: semantic grid group is not in the least interested in workflow enactment  
Want to define workflows, monitor them, perhaps change them, but \*not\* enact them  
Life sciences group want to build workflows to do life science; want to use the state of the art in workflow languages and engines  
There is a role for this group if they are interested in enactment of workflows  
What is the relationship to many other workflow activities, e.g., Oasis  
Will part of the role of this group be to survey existing activity? DAGMan, BPEL, Petri language

Dave Berry, OGSA Working Group  
Useful to monitor what is going on in W3C and Oasis

Gregor in GCE group is working on an overview of workflow environments

Proposed milestones/volunteers:

Workflow Specification Languages documents: Darrin Pulsipher

Developing Use Case Scenarios: Dietmar Erwin

Mapping Out Issues in Workflow Management: Ali Anjomshoaa

[wfm-rg@ggf.org](mailto:wfm-rg@ggf.org) mailing list