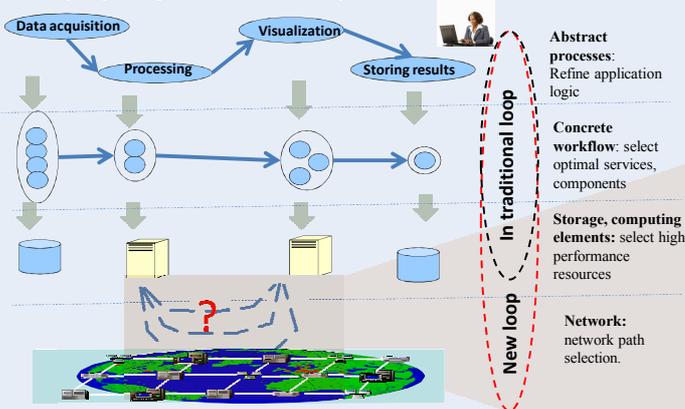


An agent based planner for including network QoS in scientific workflows

Zhiming Zhao Paola Grosso Ralph Koning Jeroen van der Ham Cees de Laat
 System and Network Engineering research group, Informatics Institute, University of Amsterdam
 Science Park 904, 1098XH, Amsterdam, the Netherlands
 {Z.Zhao | P.Grosso | R.Koning | vdHam | C.T.A.M.Delaat}@uva.nl

Background

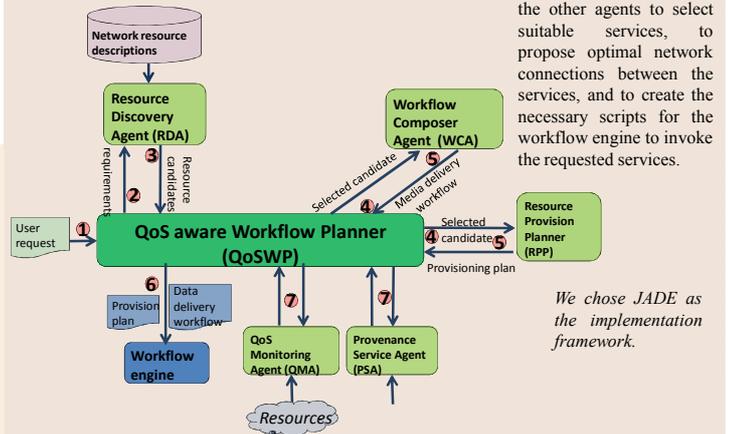
The quality of the network services has so far rarely been considered in composing and executing scientific workflows. Currently, scientific applications tune the execution quality of workflows neglecting network resources, and by selecting only optimal software services and computing resources. One reason is that IP-based networks provide few possibilities for workflow systems to manage the service quality, and limit or prevent bandwidth reservation or network paths selection. We see nonetheless a strong need from scientific applications, and network operators, to include the network quality management in the workflow systems.



Novel network infrastructures open up new possibilities in network tuning at the application level. In this position paper, we discuss our vision on this issue and propose an agent based solution to include network resources in the loop of workflow composition, scheduling and execution when advanced network services are available. We present the first prototype of our approach in the context of the CineGrid project.

Network aware Workflow QoS planner:

An agent based solution. The NeWQoSPlanner architecture consists of six agents: a Resource Discovery Agent (RDA), a Workflow Composition Agent (WCA), a Resource Provisioning Planner (RPP), a QoS Monitor Agent (QMA), a Provenance Service Agent (PSA), and a coordination agent called QoS aware workflow planner (QoSWP).



The QoSWP coordinates the other agents to select suitable services, to propose optimal network connections between the services, and to create the necessary scripts for the workflow engine to invoke the requested services.

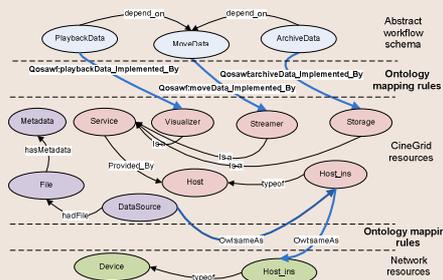
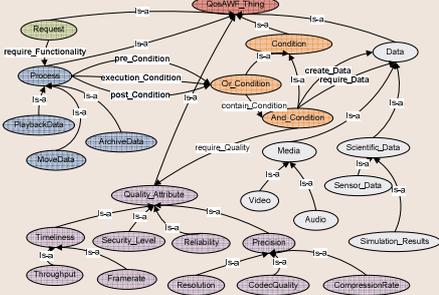
We chose JADE as the implementation framework.

Being compliant to the FIPA protocol, Jade provides a standard architecture for scheduling agent activities, which makes the inclusion of high level functionality easy, e.g., adding a Prolog module for activity reasoning. The ontology enabled agent communication between agents promotes seamless integration between the semantic network descriptions and communication messages.

Network resource selection

QoS Abstract workflow process

The input to the RDA are the requirements for data related processes which are needed by the high level workflow. Based on the experience of early work, we propose an ontology for describing abstract workflow process *qosawf.owl*. It defines the basic concepts of workflow processes, pre/post/execution conditions of the process, media data, and quality attributes.

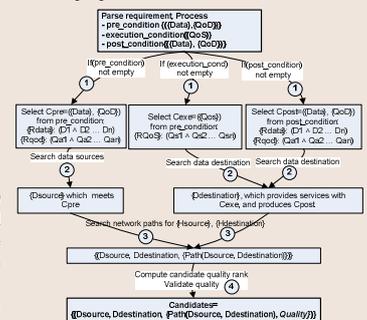


Ontology integration

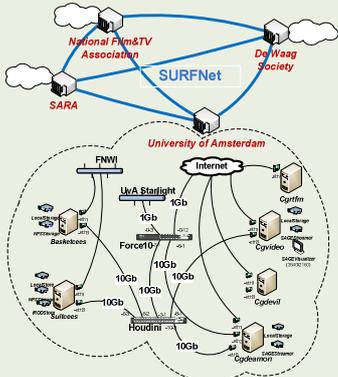
The Owl provides three build-in properties to map ontologies: *owl:sameAs* between instances, *owl:equivalentClass* between classes, and *owl:equivalentProperty* between properties. The CineGrid resources are integrated with the network level resources via property *owl:sameAs*. The mapping between abstract workflow and the CineGrid resources is via property: *qosawf:implemented_By*, which contains sub properties for each specific process.

Network resource selection

The resource discovery agent 1) parses the input description, 2) searches suitable grid resources which meet the requirements for being the data sources and destination, 3) looks for optimal network paths between them, and 4) computes the quality of resource candidates and proposes solutions.



A use case: Quality guaranteed digital media delivery on demand



The goal of the *digital media delivery on demand* use case is to retrieve media material from the infrastructure, and request quality guaranteed connections to deliver the data to qualified nodes for further processing, such as playback or visualization.



The use case is from the context of CineGrid. Four locations in Amsterdam host CineGrid resources and are connected via dedicated and configurable circuits provided by SURFnet. The portal uses the RDA agent as the back end to search network resources. The RDA receives the user requirements and network resources from the QoSWP. The RDA parses the given abstract workflow and searches the resource description; it returns results in the form of (storage host, visualization host, path, quality rank).

Summary

From the research, we can conclude:

- Including network resources in the workflow lifecycle is crucial for optimizing the quality of workflow processes which involve large quantity data movement;
- Semantic technologies play an important role in modeling QoS attributes and mapping quality description between different layers of resources in workflow system;
- Agent technology is suitable for decomposing the intelligence for discovering and selecting distributed network resources.

References

- [1] Z.Zhao et al., *An agent based planner for including network QoS in scientific workflows*, ABC:MI Oct. 18-20, 2010, Wisla, Poland.
- [2] Z.Zhao et al., *Network resource selection for data transfer processes in scientific workflow*, WORKS, SuperComputing 2010, USA.
- [3] <http://cinegrid.uvalight.nl/portal/>
- [4] <http://cinegrid.uvalight.nl/owl/qosawf.owl>