Proceedings of the International Multiconference on
Computer Science and Information Technology pp. 231–238
ISBN 978-83-60810-27-9
ISSN 1896-7094

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

***Abstract*—Advanced network infrastructure plays an important role in the e-Science environment to provide high quality connections between largely distributed data sensors, and computing and storage elements. However, 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.

***Index Terms*—Quality of Service, Semantic web, Advanced Network, Scientific Workflow, e-Science**

I. Introduction

ADVANCED network infrastructure plays an important role in the e-Science environment to provide high quality connections between largely distributed data sensors, and computing and storage elements; by enabling large scale data movement between different devices and nodes, applications which require extremely large scale data movement can be enabled. However, The quality of the network connections and services has rarely been taken into account by current scientific workflow management systems, because 1) traditional IP-based networks provide limited reservation capability for workflow engines; 2) the existing e-Science applications assume available network connections as non-changeable services, and seek customized solutions at software level to optimise computing processes and data storage; and 3) the existing applications mainly consider the functionality of the e-Science services, and limited support has been provided for including (network) quality requirements for the services in the composition, enactment and execution of workflows.

Still advanced network services are essential for many scientific applications, where data transmission delays become a bottleneck of the global performance of the application [1], [2] . Without the quality guarantee at the network level, the global workflow quality requirements can not be assured. Several strategies have been tried to improve the workflow performance in these cases, such as caching data in a closer location to the computing element [3], [4], or reducing the load of computing tasks by reusing the previous computed results [5]. However, for applications that require data streams from remote data sensors, such as in [6], those solutions will not be sufficient, and advanced network services are crucial to accelerate large data transfer between the distributed application components.

The recent emergence of advanced network infrastructures for e-Science enables tuning of network performance at the application level. For instance, the hybrid network architecture [7] can provide connections on different layers based on the same physical fibre; this allows applications to reserve dedicated circuits for transferring very large quantity data. By including network resources in the scheduling loop, the high level application gets an extra opportunity to optimize execution and improve performance.

These new opportunities come with some challenges. Not all network infrastructures provide the network services for reserving specific connections or allocating network bandwidth; the service invocation in different network domains is often proprietary and not easily extensible, and makes request for network service provisioning across sites difficult [8]; scheduling network resources requires knowledge on the current state of the network, which implies the existence of a sophisticated monitoring system.

In this paper, we discuss the challenges in including network resources in scientific workflows, and propose an agent based architecture to accomplish this goal and realize application level quality tuning of network resources. We apply this in the context of an ongoing project, CineGrid [9]. The paper is organized as follows; first, we will review the state of the art in this field, and then present our solution and the first prototype of our system.

II. Resource QoS in scientific workflow

The quality of a service is often expressed as set of quality attributes. Sabata presented a QoS taxonomy in which QoS attributes are grouped into *metrics* and *polices* [10]. The metrics include quantifiable attributes for measuring application performance (namely performance metrics), security (security level) and importance, while the polices refer to the qualitative attributes for specifying service level and management polices.

A QoS attribute is typically represented as a triple (attribute name, value, the range of the value), and timeliness and accuracy are the most often cited performance metrics. A description language such as QML [11] aims at a generic model for specifying different categories of quality attributes.

Since the development of services are highly application dependent, and the quality description models are often proposed by different society, semantic web based technologies have been widely used for describing QoS attributes [12]–[14], and to integrate different QoS models [15]. While composing a workflow, services at different abstraction level will naturally have different model on the quality; a semantic model for matching these quality attributes and mapping them will be essential.

From the perspective of a scientific workflow life cycle QoS is relevant to four aspects: composition, selection, execution control and provenance. In the following sections we briefly review the existing work on each issue.

A. QoS aware workflow composition

A workflow composition process that is QoS-aware must: 1) compose a service of the highest quality and 2) determine the quality of the composition process itself. The first goal is achieved by computing the global quality starting from the QoS attributes of constituting services [16]. Graph reduction is a widely used approach [17]; a pre-defined set of logic patterns define certain reduction rules which can be used to simplify the logical dependencies among constituting services. From the reduction rules, the quality parameters are computed; for instance the computing time of two sequentially connected services is computed as the sum of the quality of each of them, the computing time of a two parallel services is computed as the maximal one from them. The second goal requires modelling the quality attributes of the semantic links between services, the composition quality of the workflow can then be evaluated by the semantic fit and the reliability of the selected service in the workflow [16].

B. QoS aware service selection

Searching for suitable services from available resources is a basic procedure in composing a workflow. QoS aware service selection implies two steps: properly formulating the requirements and selecting resources that meet these requirements. Rosenberg proposed a QoS enabled description language, the Vienna composition language (VCL) [18], to specify an abstract flow for workflow composition. The VCL defines an abstract workflow as four parts: feature definition, feature constraints, global constraints and the business protocol (the desired workflow language). The feature constraints and global constraints include both functional constraints and QoS attributes. The problem of resource selection has been formulated differently. A commonly used formulation is *shortest path finding in a weighted graph*, in which the available services are represented as a directed graph according to the service types, and the graph nodes are labelled by the quality attributes of the service [19]. Well known shortest path finding algorithms include Bellman-Ford and Dijkstra's. These algorithms exhibit optimal performance because of their greedy search strategy and avoid backtracking operations during the search; however, the minimal cost path found by the algorithms is often not the most optimal solution if there are multiple constraints on the quality attributes. Therefore, the problem has also been formulated as a multi constraint optimal path problem [20], or multi objective optimization problem. Ant colony optimization (ACO) is a meta heuristic search approach proposed in [21], [22] for discovering minimum cost path in a graph, and for solving NP-hard combinatorial optimization problems. Fang et al, [23] applied ACO in service selection and proposed a multi objective ACO approach which can simultaneously optimize several objectives. Genetic algorithm in searching optimal paths, and constraint programming or Integer programming methods are also widely used for the multi objective optimization problem.

C. QoS aware workflow execution

Workflow execution is the mapping of workflow processes to underlying computing resources and the scheduling of the execution sequence. Task based scheduling is a straightforward approach, in which the workflow tasks are submitted to the local manager of the computing infrastructure. Several researchers have instead proposed a workflow level scheduling that takes into account future task performance [24]; this approach will achieve higher performance and better resource utilization than only using local resource managers. Multi objective optimizations are widely used to formulate the problem of QoS aware scheduling. Avanes proposed a constraint programming based approach to search for best match between workflow requirements and the available computing resources [25]. The basic idea is to describe the quality requirements and resource dependencies as constraints by partitioning the workflow into different parts based on its patterns and their QoS requirements. One of the contributions from Avanes work is that the network dynamics has been also included in the procedure of constraint resolving. Resource provision plays an important role to improve the fault tolerance and the performance of the workflow [26]. Basically, provisioning can be either static or dynamic. Advance reservation is a typical static provisioning mechanism, and several batch based schedulers support it. Based on the high level quality requirements, the workflow engine reserves computing resources and time slots from the Grid resource manager. One of the disadvantages of static provisioning is its overhead on the total cost for computing the workflow. To improve this, Raicu et al. [27] proposed multi level scheduling strategies, in which the application level scheduler is able to interact with the low level resource manager to tune the requirements at runtime. This approach introduces a dynamic component in the provisioning process.

D. QoS and provenance

The provenance service tracks the events occurred in the workflow execution, and allows scientists to trace the evolution of data computed in the workflow and to obtain insights in the experiment processes. Moreover, provenance data can also be used to debug errors of the workflow execution and optimize the workflow design. The Open Provenance Model (OPM) [28] emerges as a standard model to represent workflow provenance information. Including QoS information of the workflow processes and the execution in the provenance model allows scientists to analyze the quality of the services and the workflow scheduling. In [29], the provenance service is provided by a QoS aware middleware, which records the changes of the service quality as events. Evaluating trust and reliability of the provenance data itself has also been discussed in the literature [30]. However, research on the provenance model which includes the QoS information of the workflow processes is still in its very early stage.

III. Problem and challenges: network QoS and scientific workflows

Our research interest focuses on the inclusions of the network quality of service in the high level e-Science workflows. Workflow systems have been recognized as an important tool to manage the invocation of lower level network services and the security related routines [31]. In SC09, the VLAM workflow system [32] demonstrated allocation of workflow modules over specialized network connections. Several e-Science environments already recognized the importance of the network resources reservation in the context of workflow scheduling, for instance in [33]. When including network resources QoS in a scientific workflow, we have to solve several issues:

1) We need a representation and mapping mechanism between the user level requirements and the quality model of underlying network resources. When mapping high level workflow processes to low level resources, the abstract quality descriptions of the workflow also have to be translated to the proper quality constraints on the underlying resources.
2) A good network resource monitoring system is essential in using QoS in the system design, execution and evaluation [34]. A workflow description composed from carefully selected resources or services does not imply that the actual execution will achieve the desired performance. The quality guarantee of the workflow execution is a dynamically adapting procedure between the workflow engine and the actual state of the resources involved in the workflow execution, unless one can guarantee a priori that all the resources perform at their best state. A problem is that monitoring quality metrics of the network resources in the workflow context is also in its initial state. Truong proposed a monitoring framework for Grid services, which monitors the resources from four different levels: machine, network path, middleware and application [34], and derives the QoS value for the reliability and availability of the machines, network paths and middleware from the monitoring.
3) We need to optimize the negotiation between the workflow engine and the underlying network resource manager, for both resource provisioning and tuning the scheduling of workflow (or replacing workflow components). This is made more difficult by the ad-hoc APIs adopted by the different network providers, which prevent a generic solution.
4) In a scientific workflow we need to schedule multi level resources to support real time applications. Solving multi-level constraints between heterogeneous resources requires different optimization searching technologies in the problem space and it is time consuming compared to the state updates of the underlying resources. For instance, discovering suitable network paths often involves not only searching connected intra-domain paths between network devices, but also inter-domain exchange of topologies. Further, the evolution of the underlying network state make scheduling more difficult.
5) We need to adapt the provenance model to include the (network) quality of service information, and to let the logging facility obtain such information and record it in the database. The current provenance model mainly focuses on the evolution of the data, and uses certain partial relation to indicate the dependencies of the data process; it does not explicitly include the quality of actual services when the data was produced. Such information can be very important for improving the scheduling strategies utilised by the workflow engine.

We propose a QoS ware planner which covers the life-cycle of workflow not only at composition, but also scheduling and execution. Our system tackles the issues we just summarised above.

IV. Design requirements and system design

We had two alternatives when we looked at the inclusions of QoS aware functionalities in a scientific workflow system: 1) re-engineer the functional components of an existing workflow system to include the QoS support, or 2) consider the existing workflow systems as legacy systems, and provide QoS support as plugged components to the system. Each alternative exhibits advantages and disadvantages. In the context of our research, we chose the second approach; one of the motivations is that generic functional components can be encapsulated as reusable tools which can serve different specific scientific workflow systems to get QoS support.

A. Design requirements

We can highlight three scenarios where network QoS support can be applied: QoS aware resource selection, resource provisioning and quality assured workflow execution. The designed system thus needs to meet the following functional requirements:

1) The system must include QoS aware resource discovery and selection of network resources. Network resources and the quality attributes are necessarily described, and a search tool is provided to check the suitable resources based on the input requirements.
2) The system should be able to generate a resource provisioning plan for selected resources based on the input requirements. The plan is made based on the provisioning services that the available network infrastructure provides.
3) The system should be able to generate workflows handling large data movement between network resources with guaranteed data transfer quality, and wrap the generated workflow as a service which can be executed standalone or included in a third party workflow.
4) At runtime, the system should provide monitoring services to track the actual quality of the network resources. It should also provide interfaces for third party workflows to invoke during their provenance procedure to record all the runtime information.

B. Agent based technologies

The Agent Oriented (AO) methodology complements the object and component oriented methods with knowledge related notions to manage system complexity [35], and emerges as an important modelling and engineering approach for constructing complex systems, such as workflow management systems. The concept of *agents* originated in the mid-1950s as *a 'soft robot' living and doing its business within the computer's world* [36]. Wooldridge distinguished three types of agent architectures: deliberative, reactive and hybrid [37]. The difference between the deliberative and reactive architectures is that the former incorporates a detailed and accurate symbolic description of the external world and uses sophisticated logic to reason about the activities, while the latter one only implements a stimulus-reaction scheme. Reactive architectures are easier to implement but lack a subtle reasoning capability. Hybrids of the two schemes are commonly used. During the past two decades, agent based models, in particular reactive models, have been applied as an advanced technology in modelling and constructing complex system. Agent frameworks, such as FIPA [38], abstract the structure of basic agents and define standardized communication languages to represent interactions between agents, which facilitate the implementation of agent based applications.

JADE (Java Agent DEvelopment Framework) is a free software and distributed by Telecom Italy [38]. Fully implemented in java, Jade realizes a FIPA compliant multi agent middleware. In our project, a number of reasons motivate us to choose JADE as the implementation framework. First, the Jade platform can be distributed across machines and the configuration can be controlled via a remote GUI. The Java language makes the development portable; the Jade framework allows agents move from one machine to another at runtime. Moreover, 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. Finally, the ontology enabled agent communication between agents promotes seamless integration between the semantic network description, QoS aware searching modules, underlying models of workflow descriptions, and other necessary functional components of our system.

C. An agent based QoS workflow planner

We propose an agent based architecture, composed of a *QoS aware workflow planner (QoSWP)* and five more agents: *a Resource Discovery Agent (RDA), a Workflow Composition Agent (WCA), a Resource Provisioning Planner (RPP), a QoS Monitor Agent (QMA)* and *a Provenance Service Agent (PSA)*. Fig. 1 illustrates a conceptual schema of our agent system.

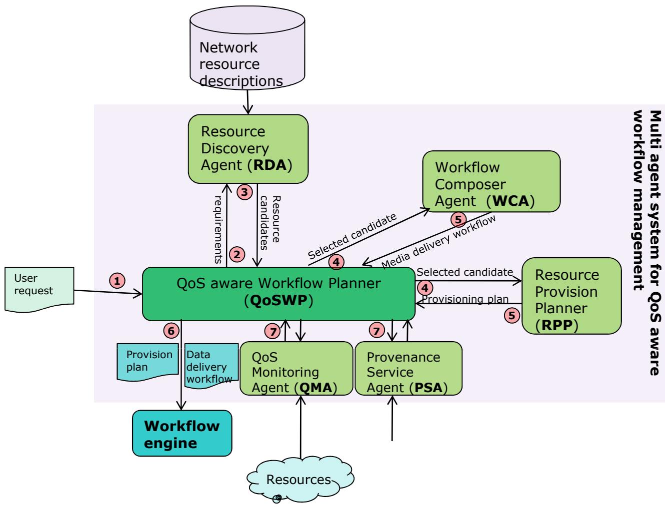

Fig. 1. An agent based solution to adaptive QoS aware workflow management.

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. A typical use case scenario will illustrate the role of each component (see Fig.1). The QoSWP receives the request for data process services and the service requirements from the user (step1). After that, the RDA reads the description of the resources and the network topologies from the registry, and searches suitable data sources and destinations, and network paths between them (step2). The RDA returns a list of qualified candidates, and sorts them based on the quality metrics of each candidate (step3). From the candidates, the QoSWP selects the best one, and request WCA and RPP to generate a resource provisioning plan and a data transfer workflow (step4 and step5), both of which will be executed by the workflow engine (step6). At run time, the QMA monitors the actual state of the resources and checks whether the global quality required by the workflow is satisfied (step7). Based on the states updated by the QMA, the QoSWP decides whether the resources of the workflow should be adapted. The provenance service records events in the resources provisioning, allocation, and combine the actual state of the quality attributes with the log data (step7).

V. Prototype and use case

The research is conducted in the context of CineGrid. An important mission of the CineGrid project is to provide a dedicated network environment to connect distributed parties from different domains to share large quantities of very-high-quality digital media, such as the high definition video material used in the movie industry.

We are prototyping our ideas using a small portion of the CineGrid infrastructure as a test bed. Four locations in Amsterdam host CineGrid resources and are connected via dedicated and configurable circuits provided by SURFnet [39]. The results reach beyond the workflow field, and they can be beneficial to understand how advanced network connections enhance the digital media delivery in the academic and education context.

In this section, we will demonstrate how the designed architecture works in a use case, and discuss the technical considerations to prototype the system.

A. A CineGrid use case: QoS guaranteed high quality media on demand

We are focusing on a *digital media delivery on demand* use case: the goal 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. Using the proposed agent framework, the use case will be prototyped as follows:

1) The user uses the schema provided by the system to describe the name and properties of the media, and to specify the quality requirements for visualizing the data.
2) The QoSWP parses the user input and creates queries for the RDA to look for data sources of the media;
3) Based on the input requirements, the RDA looks for the data repositories which contains the required media, and the visualization devices which meet the required playback quality. Then the RDA looks for all possible network paths between the sources and the visualization devices.
4) The RDA returns a list of candidates in the form of (source, destination, path) triplets, and the candidates are ordered based on the quality they provide. The QoSWP selects the best candidate from the list and send it to the RPP and WCA to make resource provisioning plan, and to create a workflow which can deliver the media from the source to the visualisation device, and to play it back in the visualization device.
5) To help RPP and WCA make the provision plan and the workflow compliant to a specific workflow engine, the QoSWP also explicit tells the RPP and WCA what language of the third party engine will use.
6) After receiving the scripts generated by the RPP and WCA, the QoSWP sends them to the third party engine to execute the provisioning plan and the delivery plan.

B. Semantic resource description

The CineGrid community uses semantic web technologies to describe the services, devices and the network topology. The UvA team in the project have developed two ontologies. The Network Description Language (NDL) [40] models the different levels of a network infrastructure: physical, domain, capability, layer and topology[1]. The CineGrid Description Language (CDL) [9] describes the services and resources on top of the network infrastructure[2]. Fig. 2 shows concepts defined in the CDL. Using these two languages, we have described the resources in the research test bed. Four locations (UvA, SARA, De Waag and the Dutch Film and TV institute) in Amsterdam are connected with up to two dedicated switchable 1Gbit/s links, which can be dynamically changed between locations using the openDRAC [41] network provisioning software used by SURFnet.

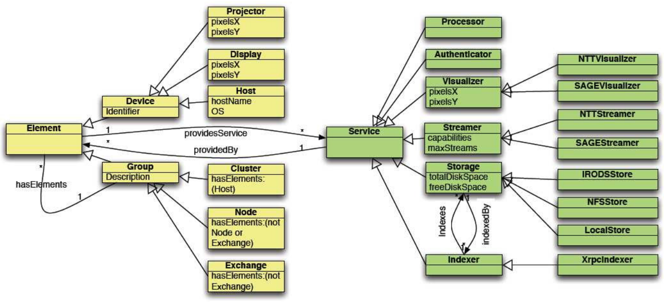

Fig. 2. The schema of the CineGrid description language.

C. QoSAWF: an abstract workflow description language

Based on the experience of early work [18], [42]–[44], we propose an ontology for describing abstract workflows (*qosawf.owl*). It defines the basic concepts of workflow processes, pre/post/execution conditions of the process, media data, and quality attributes, as shown in Fig 3. The description of the user request is described as an object of the *Request* class, and a *Request* consists of one or more *Processes* which can be accessed via the *request_Functionality* property. A *Process* class uses *pre_Condition* and *post_Condition* to indicate the requirements for *Data* the process requires and generates, and the quality for the required data. The *Process* class also uses *execution_Condition* to indicate the service quality for the process. In the current definition, *Data* contains two specific types: *Media* and *Scientific_Data*. And the service quality is modelled as set of *Quality_Attributes*. Based on the QoS taxonomy defined in [10], *Quality_attribute* can more specifically be *Precision*, *Timeliness*, *Reliability* and *Security_Level*. In our case, where the pre and post conditions consist of requirements for data and the data quality, *and_Condition* and *or_Condition* are the two most important types. Using the above ontology, a user is able to formulate a request for obtaining and playing

[1] Available at: http://cinegrid.uvalight.nl/owl/ndl-domain.owl and http://cinegrid.uvalight.nl/owl/ndl-topology.owl

[2] Available at: http://cinegrid.uvalight.nl/owl/cdl/2.0

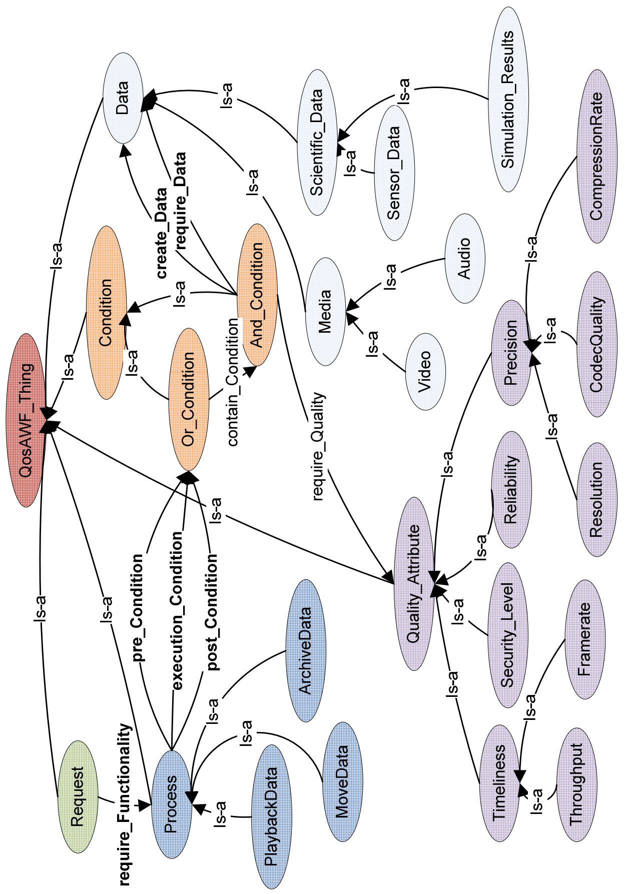

Fig. 3. The concepts defined in the qosawf schema.

back a specific video material with a minimal resolution and frame rate. Fig. 4 shows the graphical representation.

D. First prototype

We have compared different options to realise the resource search mechanism; we have evaluated several Query languages (RQL, RDQL, N3, Versa, SeRQL, SPARQL) and Rule languages (SWRL, Prolog/RDF lib, JESS etc.). We have finally chosen the RDF library of SWI-Prolog; its triple based manipulation interface is easy for the high level language we use to implement the agents (Java); it is also easy to access the runtime state of the triples. Finally, the Prolog language provides effective solutions to realise graph path findings.

The FIPA [38] standards provide a suitable architecture to implement distributed agents in our system. The Agent Communication Language (ACL) allows agents to exchange messages using an explicitly defined semantic schema, which allows seamless integration between agents and remote Ontology knowledge bases.

In the current prototype, the RDA receives the URI of 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).

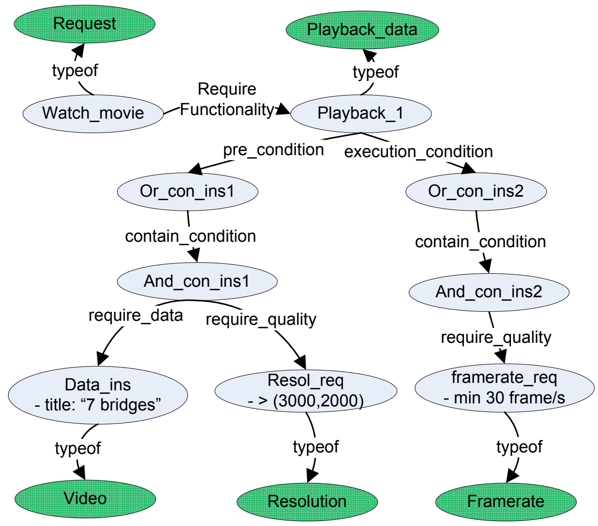

Fig. 4. An example of abstract workflow description.

VI. Summarizing remarks

In this paper, we have reviewed the state of the art of the QoS aware workflow management. We have discussed different technologies for realizing QoS aware resource discovery, workflow composition, scheduling and adaptation. In the review, we can summarize several issues:

1) Semantic technologies play an important role in modelling QoS attributes and mapping quality description between different layers of resources in workflow system.
2) QoS aware service selection has been tackled using different solutions: minimal cost path finding in a weighted graph, multi objective optimization or integer (constraint) programming. For selections with constraints of different QoS attributes, multi objective optimization is a better solutions.
3) The QoS of network resources have been tried in workflow system in different ways: using workflow to manage network services, for interactively creating network connections from the workflow, or scheduling part of the workflow in the resources which have special quality connections. These solutions focuses on runtime part of the network resource allocation, which is often tightly coupled with the specific network service.

Based on the existing work, we have proposed an agent based solution to include QoS aware network resources planning in the loop of workflow composition, provisioning and execution. We have proposed an RDF based schema for describing the data and QoS requirements in an abstract workflow.

The research is conducted in the context of the CineGrid project; however, the results of the research aim to be generic and applicable for any e-Science applications which suffer from the bottleneck of transferring large quantity data over not optimal networks. We will develop the QoSWP as a generic service for legacy workflow engines to support QoS aware network resource selection, provisioning and invocation of network services for transferring data. Moreover we will define a provenance model which includes the causality relations between scientific data and the quality of the utilised resources. This will allow scientists to investigate the dynamics of the underlying resources allocation and to optimize the workflow scheduling strategies. The work described in the paper is still in its early stage. The research will proceed as follows. Firstly, we will prototype the RDA to enhance an early version of the CineGrid portal developed in the UvA for QoS aware network resources discovery. Secondly, we will realise the WCA and RPP; we will use the VLEWF-Bus system [45] as the first test case for the underlying third party workflow engine. Thirdly, we will focus on the QMA and PSA. The OPM will be used as starting model to develop PSA.

ACKNOWLEDGMENT

We would like to thank the National e-Science Research Center and the Dutch national and education network SURFnet, through the GigaPort Research on Network (RoN) project, for sponsoring this research.

REFERENCES

[1] Osamu Tatebe, Youhei Morita, Satoshi Matsuoka, Noriyuki Soda, and Satoshi Sekiguchi. Grid datafarm architecture for petascale data intensive computing. *Cluster Computing and the Grid, IEEE International Symposium on*, 0:102, 2002.

[2] T. Lavian, J. Mambretti, D. Cutrell, H. Cohen, S. Merrill, R. Durairaj, P. Daspit, I. Monga, S. Naiksatam, S. Figueira, D. Gutierrez, D. Hoang, and F. Travostino. Dwdm-ram: a data intensive grid service architecture enabled by dynamic optical networks. *Cluster Computing and the Grid, IEEE International Symposium on*, 0:762–764, 2004.

[3] Khaled Almi'ani, Javid Taheri, and Anastasios Viglas. A data caching approach for sensor applications. *Parallel and Distributed Computing Applications and Technologies, International Conference on*, 0:88–93, 2009.

[4] David Chiu and Gagan Agrawal. Hierarchical caches for grid workflows. *Cluster Computing and the Grid, IEEE International Symposium on*, 0:228–235, 2009.

[5] Ewa Deelman, James Blythe, Yolanda Gil, Carl Kesselman, Scott Koranda, Albert Lazzarini, Gaurang Mehta, Maria Alessandra Papa, and Karan Vahi. Pegasus and the pulsar search: From metadata to execution on the grid. In *PPAM*, pages 821–830, 2003.

[6] Robert Morris, Jennifer Dungan, and Petr Votava. A workflow model for earth observation sensor webs. *Space Mission Challenges for Information Technology, IEEE International Conference on*, 0:69–76, 2009.

[7] Cees de Laat, Erik Radius, and Steven Wallace. The rationale of the current optical networking initiatives. *Future Generation Computer Systems*, 19(6):999 – 1008, 2003. 3rd biennial International Grid applications-driven testbed even t, Amsterdam, The Netherlands, 23-26 September 2002.

[8] Zhenhai Duan, Zhi-Li Zhang, and Yiwei Thomas Hou. Service overlay networks: Slas, qos and bandwidth provisioning. *Network Protocols, IEEE International Conference on*, 0:334, 2002.

[9] Cinegrid initiative, http://www.cinegrid.org.

[10] Bikash Sabata, Saurav Chatterjee, Michael Davis, Jaroslaw J. Sydir, and Thomas F. Lawrence. Taxomomy of qos specifications. *Object-Oriented Real-Time Dependable Systems, IEEE International Workshop on*, 0:100, 1997.

[11] S. Frolund and J. Koisten. Qml: A language for quality of service specification, 1998.

[12] Ioannis V. Papaioannou, Dimitrios T. Tsesmetzis, Ioanna G. Roussaki, and Miltiades E. Anagnostou. A qos ontology language for web-services. In *In Proc. of 20th Intl. conf. on Advanced Information Networking and Applications (AINA*, page 2006, 2006.

[13] Glen Dobson, Russell Lock, and Ian Sommerville. Qosont: a qos ontology for service-centric systems. In *31st EUROMICRO Conference on Software Engineering and Advanced Applications*, pages 80–87, 2005.

[14] Glen Dobson and Alfonso Snchez-Maci. Towards unified qos/sla ontologies. *IEEE Services Computing Workshops*, 0:169–174, 2006.

[15] Tian Gramm Naumowicz, M. Tian, A. Gramm, T. Naumowicz, H. Ritter, and J. Schiller. A concept for qos integration in web services. In *In 1st Web Services Quality Workshop (WQW2003) at WISE*, pages 149–155. IEEE Computer Society, 2003.

[16] Freddy Lecue and Nikolay Mehandjiev. Towards scalability of quality driven semantic web service composition. *Web Services, IEEE International Conference on*, 0:469–476, 2009.

[17] Jorge Cardoso, John Miller, Amit Sheth, and Jonathan Arnold. Quality of service for workflows and web service processes. *Journal of Web Semantics*, 1:281–308, 2002.

[18] Florian Rosenberg, Philipp Leitner, Anton Michlmayr, Predrag Celikovic, and Schahram Dustdar. Towards composition as a service - a quality of service driven approach. *Data Engineering, International Conference on*, 0:1733–1740, 2009.

[19] Yingqiu Li, Minghua Chen, Tao Wen, and Lei Sun. Quality driven web services composition based on an extended layered graph. *Computer Science and Software Engineering, International Conference on*, 3:153–156, 2008.

[20] Jia Yu, Michael Kirley, and Rajkumar Buyya. Multi-objective planning for workflow execution on grids. *Grid Computing, IEEE/ACM International Workshop on*, 0:10–17, 2007.

[21] Youmei Li and Zongben Xu. An ant colony optimization heuristic for solving maximum independent set problems. *Computational Intelligence and Multimedia Applications, International Conference on*, 0:206, 2003.

[22] In Alaya, Christine Solnon, and Khaled Ghdira. Ant colony optimization for multi-objective optimization problems. *Tools with Artificial Intelligence, IEEE International Conference on*, 1:450–457, 2007.

[23] Fang Qiqing, Peng Xiaoming, Liu Qinghua, and Hu Yahui. A global qos optimizing web services selection algorithm based on moaco for dynamic web service composition. *Information Technology and Applications, International Forum on*, 1:37–42, 2009.

[24] Fumiko Harada, Toshimitsu Ushio, and Yukikazu Nakamoto. Adaptive resource allocation control for fair qos management. *IEEE Transactions on Computers*, 56:344–357, 2007.

[25] Artin Avanes and Johann-Christoph Freytag. Adaptive workflow scheduling under resource allocation constraints and network dynamics. *Proc. VLDB Endow.*, 1(2):1631–1637, 2008.

[26] Gideon Juve and Ewa Deelman. Resource provisioning options for large-scale scientific workflows. In *ESCIENCE '08: Proceedings of the 2008 Fourth IEEE International Conference on eScience*, pages 608–613, Washington, DC, USA, 2008. IEEE Computer Society.

[27] Ioan Raicu, Yong Zhao, Catalin Dumitrescu, Ian Foster, and Mike Wilde. Falkon: a fast and light-weight task execution framework. In *SC '07: Proceedings of the 2007 ACM/IEEE conference on Supercomputing*, pages 1–12, New York, NY, USA, 2007. ACM.

[28] Luc Moreau, Juliana Freire, Joe Futrelle, Robert E. Mcgrath, Jim Myers, and Patrick Paulson. The open provenance model: An overview. pages 323–326, 2008.

[29] Anton Michlmayr, Florian Rosenberg, Philipp Leitner, and Schahram Dustdar. Service provenance in qos-aware web service runtimes. *Web Services, IEEE International Conference on*, 0:115–122, 2009.

[30] Shrija Rajbhandari, Arnaud Contes, Omer F. Rana, Vikas Deora, and Ian Wootten. Trust assessment using provenance in service oriented applications. *Enterprise Distributed Object Computing Workshops, International Conference on*, 0:65, 2006.

[31] Y. Demchenko, L. Gommans, C. de Laat, A. Taal, A. Wan, and O. Mulmo. Using workflow for dynamic security context management in grid-based applications. *Grid Computing, IEEE/ACM International Workshop on*, 0:72–79, 2006.

[32] Adam S. Z. Belloum, David L. Groep, Zeger W. Hendrikse, Bob L. O. Hertzberger, Vladimir Korkhov, Cees T. A. M. de Laat, and Dmitry Vasunin. Vlam-g: a grid-based virtual laboratory. *Future Gener. Comput. Syst.*, 19(2):209–217, 2003.

[33] Christoph Barz, Markus Pilz, and Andr Wichmann. Temporal routing metrics for networks with advance reservations. *Cluster Computing and the Grid, IEEE International Symposium on*, 0:710–715, 2008.

[34] Hong-Linh Truong, Robert Samborski, and Thomas Fahringer. Towards

a framework for monitoring and analyzing qos metrics of grid services. *e-Science and Grid Computing, International Conference on*, 0:65, 2006.

[35] Philippe Massonet, Yves Deville, and Cèdric Néve. From aose methodology to agent implementation. In *Proceedings of the first international joint conference on Autonomous agents and multi agent systems*, pages 27–34. ACM Press, 2002.

[36] A. Kay. Computer software. *Scientific American*, 251(3):53–59, 1984.

[37] M. Wooldridge and N. Jenings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152, 1995.

[38] Fabio Bellifemine, Agostino Poggi, and Giovanni Rimassa. JADE: a FIPA2000 compliant agent development environment. In *Proceedings of the fifth international conference on Autonomous agents*, pages 216–217. ACM Press, 2001.

[39] The SurfNet. The surfnet homepage. In *http://www.surfnet.nl/*, 2002.

[40] Jeroen van de Ham. *A semantic model for complex computer networks the Network Description Language*. PhD thesis, University of Amsterdam, 2010.

[41] The opendrac network provisioning software. http://www.opendrac.org.

[42] Doina Caragea and Tanveer Syeda-Mahmood. Semantic api matching for automatic service composition. In *WWW Alt. '04: Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*, pages 436–437, New York, NY, USA, 2004. ACM Press.

[43] Matthias Klusch, Benedikt Fries, and Katia Sycara. Automated semantic web service discovery with owls-mx. In *AAMAS '06: Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 915–922, New York, NY, USA, 2006. ACM.

[44] Marian Bubak, Tomasz Gubala, Michal Kapalka, Maciej Malawski, and Katarzyna Rycerz. Workflow composer and service registry for grid applications. *Future Gener. Comput. Syst.*, 21(1):79–86, 2005.

[45] Zhiming Zhao, Suresh Booms, Adam Belloum, Cees de Laat, and Bob Hertzberger. Vle-wfbus: a scientific workflow bus for multi e-science domains. In *Proceedings of the 2nd IEEE International conference on e-Science and Grid computing*, pages 11–19, Amsterdam, the Netherlands, December 4- December 6 2006. IEEE Computer Society Press.