Chapter III
From Enabling to Ensuring
Grid Workflows

Junwei Cao
Tsinghua University, China & Tsinghua National Laboratory for Information Science and Technology, China

Fan Zhang
Tsinghua University, China

Ke Xu
Tsinghua University, China

Lianchen Liu
Tsinghua University, China & Tsinghua National Laboratory for Information Science and Technology, China

Cheng Wu
Tsinghua University, China & Tsinghua National Laboratory for Information Science and Technology, China

ABSTRACT

Grid workflows are becoming a mainstream paradigm for implementing complex grid applications. In addition to existing grid enabling techniques, various grid ensuring techniques are emerging, e.g. workflow analysis and temporal reasoning, to probe potential pitfalls and errors and guarantee quality of services (QoS) at a design phase. A new state π calculus is proposed in this work, which not only enables flexible abstraction and management of historical grid system events, but also facilitates modeling and verification of grid workflows. Some typical patterns in grid workflows are captured and both static and dynamic formal verification issues are investigated, including structural correctness, specification satisfiability, logic satisfiability and consistency. A grid workflow modeling and verification environment, GridPiAnalyzer, is implemented using formal modeling and verification methods proposed in this work. Performance evaluation results are included using a grid workflow for gravitational wave data analysis.
INTRODUCTION

Grid Workflow QoS

Advance in technology has made collections of internet-connected computers a viable computational platform. Grids connecting geographically distributed resources have become a promising infrastructure for solving large problems. The definition of Grids has been redefined over time. Initially Grids were defined as an infrastructure to provide easy and inexpensive access to high-end computing (Foster, 1998). Then, it was refined in (Foster, 2001) as an infrastructure to share resources for collaborative problem solving. More recently, in (Foster, 2002) the Grid definition evolves into an infrastructure to virtualize resources and enable their use in a transparent fashion.

Grid workflows (Cao, 2003), a composition of various grid services according to prospective processes, have become a typical paradigm for problem solving in various e-Science domains (Yu, 2004), e.g. gravitational wave data analysis (Deelman, 2002). With increasing complexity of e-Science applications, how to implement reliable and trustworthy grid workflows according to specific scientific criteria is becoming a critical research issue. In addition to existing grid enabling techniques, e.g. job scheduling, workflow enactment and resource locating, various grid ensuring techniques are developed (Xu, 2006), e.g. data flow analysis and temporal reasoning.

Issues of quality of service (QoS) are of increasing importance to the success of those Grid-based applications. As defined by I. Foster in the three point checklist of the Grid (Foster, 2002), the Grid has to deliver to nontrivial qualities of service, relating for example to response time, throughput, availability, and security, and/or co-allocation of multiple resource types to meet complex user demands. This requirement is especially pronounced in experimental science applications such as the National Fusion Collaboratory (Keahey, 2004) and NEESgrid (Pearlman, 2004). Enabling such interactions on the Grid requires two related efforts: (1) the development of sophisticated resource management strategies and algorithms and (2) the development of protocols enabling structural negotiation for the use of those resources.

Most of existing research on grid workflow QoS is related to task scheduling. In the work described in (Spooner, 2005), application performance prediction is coupled with genetic algorithms for workflow management and scheduling with consideration of makespans and job deadlines. QoS guided min-min heuristic for grid task scheduling is also proposed in (He, 2003). Similar work can also be found in (Zhang, 2004) and (Brandic, 2005) for QoS aware grid workflow scheduling using performance prediction and optimization. In the grid standard organization, Global Grid Forum, a WS-Agreement model is proposed and defined in (Czajkowski, 2003). This provides an infrastructure to agreement-based application like (Keahey, K., Araki, T., et al. 2004) and (Zhang, 2004), within which QoS can be negotiated and obtained.

While all of above in common is that they show how task can be scheduled to improve efficiency of grid workflows, this work is dedicated to ensuring mechanisms on workflows as a whole. All of services in a workflow are guaranteed without redundancy and collision. Also how to make sure all services in a workflow is reachable and terminatable is another concern in this work. All these issues are modeled, verified and finally implemented using our environment, GridPiAnalyzer.
Related Content

Black Hole Traffic Anomaly Detections in Wireless Sensor Network
www.igi-global.com/article/black-hole-traffic-anomaly-detections-in-wireless-sensor-network/128360?camid=4v1a

Service Level Agreement (SLA) in Utility Computing Systems
www.igi-global.com/chapter/service-level-agreement-sla-utility/64489?camid=4v1a

Testing-Effort Dependent Software Reliability Model for Distributed Systems
www.igi-global.com/article/testing-effort-dependent-software-reliability/78150?camid=4v1a

Introduction
www.igi-global.com/chapter/introduction/43100?camid=4v1a