



UvA-DARE (Digital Academic Repository)

Understanding and mastering dynamics in computing grids: processing moldable tasks with user-level overlay

Mościcki, J.T.

Publication date
2011

[Link to publication](#)

Citation for published version (APA):

Mościcki, J. T. (2011). *Understanding and mastering dynamics in computing grids: processing moldable tasks with user-level overlay*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Summary

Scientific communities are using a growing number of distributed systems, from local batch systems, community-specific services and supercomputers to general-purpose, global grid infrastructures. Increasing the research capabilities for science is the *raison d'être* of such infrastructures which provide access to diversified computational, storage and data resources at large scales. Grids are rather chaotic, highly heterogeneous, decentralized systems where unpredictable workloads, component failures and variability of execution environments are commonplace. Understanding and mastering the heterogeneity and dynamics of such distributed systems is prohibitive for end users if they are not supported by appropriate methods and tools. The time cost to learn and use the interfaces and idiosyncrasies of different distributed environments is another challenge.

Obtaining more reliable application execution times and boosting parallel speedup are important to increase the research capabilities of scientific communities. Late binding is one of techniques to achieve these goals because the majority of jobs which are in production in grids and supercomputers are moldable. Moldable jobs may use variable number of resources and be more flexibly partitioned than classical, rigid parallel jobs. Moldable job application examples include Monte Carlo simulations, parameter sweeps, directed acyclic graphs and workflows, data-parallel analysis algorithms and many more.

We analyze spatial and temporal dynamics and study the performance variations in large, loosely coupled distributed systems such as the EGEE Grid – the largest Grid infrastructure to date. We develop a mathematical description of task processing in the Grid, where system parameters are taken as random variables with empirical distributions. We analyze the Quality of Service indicators such as variance of makespan to qualitatively compare late and early-binding task processing models. Using a continuous approximation we analytically demonstrate that properties of the late-binding model allow to reduce the makespan distribution according to fundamental laws of statistics. To analyze the discrete cases and more complex parameters, including the communication overheads, we use Monte Carlo simulation. We identify that under certain conditions late binding allows to achieve speedups which are often greater than an

order of magnitude compared to early binding.

We describe the principles guiding the development of a lightweight, User-level Overlay which exploits late binding to achieve an improved Quality of Service in unreliable and unpredictable distributed environments. Our strategy is based on loosely coupled, user-space tools, where the DIANE scheduler manages task allocation in a pool of worker nodes which is asynchronously created and managed by the GANGA interface. This approach makes it easy (1) to create resource selection mechanisms such as the heuristic-based worker agent factory, and (2) to plug-in adaptive workload-balancing algorithms for task scheduling. Other key features include an ability to interface to a wide range of distributed systems; an ability to extend and customize the system with application-specific scheduling and processing methods; ease of use and uniform interface to heterogeneous job management systems.

Using real-life applications in the EGEE Grid, local batch systems and dedicated clusters we demonstrate new and improved capabilities which are provided by the GANGA/DIANE User-level Overlay above generic middleware stack. These capabilities include efficient short-deadline computing, increased dependability, autonomous large-scale operations, efficient parameter sweeps, man-in-the loop scenarios, automated DAGs/workflows and semi-interactivity.

We present two case-studies of capacity and capability computing with the User-level Overlay. We show how a large number of tasks with a short deadline was coordinated on the Grid to improve dependability of locally available resources for the International Telecommunication Union Regional Radio Conference 2006. Then we describe how task prioritization and resource selection was implemented for the Lattice QCD simulations for the QCD thermodynamics studies in the context of heavy-ion collisions experiments (LHC, RHIC).

This work is a contribution to the debate if Quality of Service in grids may be efficiently implemented at the application level. We demonstrated that it is indeed possible (1) by giving a theoretical explanation of the effects of late binding on key task processing metrics, and (2) by showing examples of applications which successfully applied the User-level Overlay.