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

Mościcki, J.T.

Citation for published version (APA):

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.
# Table of Contents

1 Motivation and research objectives ................................................. 1
  1.1 Distributed applications: common patterns and characteristics .......... 2
  1.2 Infrastructures for scientific computing ..................................... 8
  1.3 Higher-level middleware systems ............................................. 9
  1.4 User requirements ................................................................ 13
  1.5 The research objectives and roadmap ....................................... 15

2 Dynamics of large computing grids ................................................ 19
  2.1 EGEE – world’s largest computing and data Grid ......................... 19
  2.2 Grid as an infrastructure ....................................................... 22
  2.3 Grid as a task processing system ............................................. 27
  2.4 Summary ............................................................................. 39

3 Analysis and modeling of task processing with late binding on the Grid ... 41
  3.1 Introduction .......................................................................... 41
  3.2 Task processing model .......................................................... 42
  3.3 Distribution of job queuing time .............................................. 44
  3.4 Simulation of task processing models ....................................... 48
  3.5 Summary ............................................................................. 57

4 Development of the User-level Overlay ........................................... 59
  4.1 Vision .................................................................................. 60
  4.2 Functional breakdown and architecture ...................................... 62
  4.3 DIANE and Ganga software packages ...................................... 63
  4.4 Operation of the User-level Overlay ....................................... 64
  4.5 The DIANE task coordination framework .................................. 66
  4.6 The Ganga resource access API and user interface ..................... 73
# TABLE OF CONTENTS

4.7 Heuristic resource selection ........................................ 80
4.8 Adaptive workload balancing ....................................... 85
4.9 Summary ..................................................................... 89

5 User-level Overlay in action ............................................. 91
5.1 Monte Carlo simulation with Geant4 toolkit ...................... 92
5.2 Workflows for medical imaging simulations ...................... 99
5.3 Data processing for ATLAS and LHCb experiments ............. 102
5.4 Massive molecular docking for Avian Flu ......................... 103
5.5 Other examples of using DIANE/Ganga overlay .................. 105
5.6 Summary ..................................................................... 106

6 Capability computing case study: ITU broadcasting planning .... 109
6.1 Introduction .................................................................. 109
6.2 Broadcasting planning process ...................................... 110
6.3 Compatibility analysis ................................................ 111
6.4 Implementation of grid-based analysis system for the RRC06 .... 113
6.5 Analysis of task processing .......................................... 115
6.6 Summary ..................................................................... 120

7 Capacity computing case study: LatticeQCD simulation .......... 121
7.1 Introduction .................................................................. 121
7.2 Problem to be solved .................................................. 122
7.3 Simulation model ....................................................... 123
7.4 Implementation and operation of the simulation system ........ 125
7.5 Task scheduling and prioritization ................................... 130
7.6 Analysis of adaptive resource selection ............................ 137
7.7 Exploiting low-level parallelism for finer lattices ................. 139
7.8 Summary ..................................................................... 140

8 Conclusions and future work ............................................. 143
8.1 Grid dynamics and its consequences for task processing ........ 143
8.2 Contributions of this work .......................................... 144
8.3 Open issues .................................................................. 146
8.4 Future work .................................................................. 147
8.5 Postscriptum ............................................................. 148

Bibliography ..................................................................... 164

Summary .......................................................................... 165

Nederlandse samenvatting .................................................. 167

Streszczenie po polsku ....................................................... 169

Publications .................................................................... 171
TABLE OF CONTENTS

Acknowledgments 175
Index 177