Chapter 12
On-Demand Visualization on Scalable Shared Infrastructure

Huadong Liu
University of Tennessee, USA

Jinzhu Gao
University of The Pacific, USA

Jian Huang
University of Tennessee, USA

Micah Beck
University of Tennessee, USA

Terry Moore
University of Tennessee, USA

ABSTRACT

The emergence of high-resolution simulation, where simulation outputs have grown to terascale levels and beyond, raises major new challenges for the visualization community, which is serving computational scientists who want adequate visualization services provided to them on-demand. Many existing algorithms for parallel visualization were not designed to operate optimally on time-shared parallel systems or on heterogeneous systems. They are usually optimized for systems that are homogeneous and have been reserved for exclusive use. This chapter explores the possibility of developing parallel visualization algorithms that can use distributed, heterogeneous processors to visualize cutting edge simulation datasets. The authors study how to effectively support multiple concurrent users operating on the same large dataset, with each focusing on a dynamically varying subset of the data. From a system design point of view, they observe that a distributed cache offers various advantages, including improved scalability. They develop basic scheduling mechanisms that were able to achieve fault-tolerance and load-balancing, optimal use of resources, and flow-control using system-level back-off, while still enforcing deadline driven (i.e. time-critical) visualization.

INTRODUCTION

The emergence of high-resolution simulation, where simulation outputs have grown to terascale levels and beyond, raises major new challenges for the visualization community, which is serving computational scientists who want adequate visualization services provided to them on-demand. For one thing, visualizing such massive datasets inevitably requires large-scale parallelism, but parallel systems of adequate size are still too scarce a resource for widespread routine use. This practical bottleneck is exacerbated by the fact that many existing algorithms for parallel visualization were not designed to operate optimally on time-shared parallel systems or on heterogeneous systems. They are usually optimized for systems that are homogeneous and have been reserved for exclusive use. Few parallel visualization algorithms now available can effectively utilize aggregated, network accessible computing resources to serve the data-intensive and on-demand visualization needs of a group of concurrent users.

To ameliorate this situation and start supporting the changing demands of simultaneous users, we need to develop parallel visualization algorithms that do not assume that the underlying processors are homogeneous or always available. Such algorithms should work well even when a parallel visualization task obtains different performances from different processors or when overloaded processors appear temporarily unavailable.

Weakening other assumptions could enable some even greater advantages. In particular, instead of assuming that all available processors are connected by a system-area network, we could develop algorithms that assume the wide-area Internet as the interconnect. Algorithms so designed could recruit all processing resources available in the distributed environment and obtain a system that scales beyond the conventional boundaries of administrative domains. Moreover, with the standard Internet providing universal access, geographically separated users could use such an inherently distributed parallel system as a shared infrastructure for collaborative data sharing and data-intensive visualization. It would be unnecessary to provision dedicated clusters and create local replicas of large data objects at each separate site.

In this chapter, we present our research to meet such new needs. We developed a test-bed consisting of 100 networked computers, none of which were specially provisioned for visualization. While all the distributed processors involved are heterogeneous, freely available, and independent of each other, our work shows that they can be formed into a generic, shared infrastructure on which 10 concurrent users can visualize and interact with a 128 time-step simulation dataset totaling 250 GB.

Our contributions include the following. First, we designed middleware that combines loosely coupled, distributed resources into a common execution environment that dynamically discovers the most efficient computing resources available in the system. Using a novel scheme of data replication and distributed caching, it also reduces runtime data movement by leveraging the data access patterns of volume visualization.

Second, for parallel visualization algorithms operating in the master-worker model, we devised a novel scheduling algorithm that runs on a user's local computer, i.e. on the client, and orchestrates a parallel visualization run on distributed heterogeneous processors. The scheduling algorithm implements mechanisms for performance and fault-tolerance. In addition, the scheduler is designed with a robust protocol of distributed flow-control to maintain the scalability and stability of a large-scale shared system. The flow-control mechanism involves a two-level back-off scheme, i.e. system-level back-off as well as application level back-off. The system level back-off (less aggressive task assignment) regulates every client's use of the system to ensure fairness of resource utilization and at the same time avoid overloading the overall distributed system. The application

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/demand-visualization-scalable-shared-infrastructure/62831

Related Content

From Enabling to Ensuring Grid Workflows
Junwei Cao, Fan Zhang, Ke Xuand Lianchen Liu (2009). *Quantitative Quality of Service for Grid Computing: Applications for Heterogeneity, Large-Scale Distribution, and Dynamic Environments (pp. 46-73).*
www.irma-international.org/chapter/enabling-ensuring-grid-workflows/28270

Harnessing the Cloud for Mobile Social Networking Applications
Juwel Rana, Josef Hallberg, Kåre Synnesand Johan Kristiansson (2010). *International Journal of Grid and High Performance Computing (pp. 1-11).*
www.irma-international.org/article/harnessing-cloud-mobile-social-networking/43880

Grid Data Handling
Alexandru Costan (2012). *Computational and Data Grids: Principles, Applications and Design (pp. 112-139).*
www.irma-international.org/chapter/grid-data-handling/58743

Bio-Inspired Grid Resource Management
Vineet R. Khareand Frank Zhigang Wang (2009). *Handbook of Research on Grid Technologies and Utility Computing: Concepts for Managing Large-Scale Applications (pp. 118-125).*
www.irma-international.org/chapter/bio-inspired-grid-resource-management/20514

The Impact of Virtualization on High Performance Computing Clustering in the Cloud
Ouidad Achahbarand Mohamed Riduan Abid (2015). *International Journal of Distributed Systems and Technologies (pp. 65-81).*
www.irma-international.org/article/the-impact-of-virtualization-on-high-performance-computing-clustering-in-the-cloud/136253