Chapter 6
OpenGL® API-Based Analysis of Large Datasets in a Cloud Environment

Wolfgang Mexner
Karlsruhe Institute of Technology (KIT), Germany

Matthias Bonn
Karlsruhe Institute of Technology (KIT), Germany

Andreas Kopmann
Karlsruhe Institute of Technology (KIT), Germany

Viktor Mauch
Karlsruhe Institute of Technology (KIT), Germany

Doris Ressmann
Karlsruhe Institute of Technology (KIT), Germany

Suren A. Chilingaryan
Karlsruhe Institute of Technology (KIT), Germany

Nicholas Tan Jerome
Karlsruhe Institute of Technology (KIT), Germany

Thomas van de Kamp
Karlsruhe Institute of Technology (KIT), Germany

Vincent Heuveline
Heidelberg University, Germany

Philipp Lösel
Heidelberg University, Germany

Sebastian Schmelzle
Technische Universität Darmstadt (TUD), Germany

Michael Heethoff
Technische Universität Darmstadt (TUD), Germany

DOI: 10.4018/978-1-5225-2785-5.ch006

ABSTRACT

Modern applications for analysing 2D/3D data require complex visual output features which are often based on the multi-platform OpenGL® API for rendering vector graphics. Instead of providing classical workstations, the provision of powerful virtual machines (VMs) with GPU support in a scientific cloud with direct access to high performance storage is an efficient and cost effective solution. However, the automatic deployment, operation and remote access of OpenGL® API-capable VMs with professional visualization applications is a non-trivial task. In this chapter the authors demonstrate the concept of such a flexible cloud-like analysis infrastructure within the framework of the project ASTOR. The authors present an Analysis-as-a-Service (AaaS) approach based on VMware™-ESX for on demand allocation of VMs with dedicated GPU cores and up to 256 GByte RAM per machine.

INTRODUCTION

Due to the ability of X-rays to penetrate materials, they are highly appropriate to visualize internal structures of opaque objects. Moreover, X-ray-computed tomography provides the opportunity to visualize internal structures of optically dense materials in 3D. The intensity of X-rays emitted by synchrotron light sources is several orders of magnitudes higher than from laboratory sources and provide brilliant and partially coherent radiation for fast imaging with synchrotron radiation (Cloetens, Bolle, Ludwig, Baruchel, & Schlenke, 2001). The application of synchrotron-based X-ray-micro-tomography for biological samples was the onset of a new era of morphological research on millimetre-sized animals like small arthropods (e.g. Heethoff & Norton, 2009; van de Kamp, Vagovič, Baumbach, & Riedel, 2011; Schmelzle, Norton, & Heethoff,2015). In recent years, new setups enabled unrivalled opportunities of high-throughput measurements, 3D/4D-tomographic imaging of dynamic systems, and even living organisms (dos Santos Rolo, 2014). Online data evaluation became possible by the usage of advanced graphic processors for scientific computing (Chilingaryan, Kopmann, Mirone, dos Santos Rolo, & Vogelgesang, 2011). These new technologies, however, result in large amounts of data: currently up to 100 GByte per volume summing up to 15 TByte/day. Technical limitations are reached regarding data acquisition, storage, and organization. Analysis of tomographic data is usually time-consuming and many analysis steps relies on commercial applications that provide visual output based on the OpenGL® API for Microsoft operating systems (Mauch et al., 2014). An example for frequently used commercial applications are AMIRA™ and VG Studio MAX™, which are costly and might not be available at all user's home institutions.