OpenGL-BASED DATA ANALYSIS IN VIRTUALIZED SELF-SERVICE ENVIRONMENTS

V. Mauch, M. Bonn, S. Chilingaryan, A. Kopmann, W. Mexner, D. Ressmann
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Abstract

Modern data analysis applications for 2D/3D data samples require complex visual output features which are often based on OpenGL, a multi-platform API for rendering vector graphics. They demand special computing workstations with a corresponding CPU and GPU power, enough main memory and fast network interconnects for a performant remote data access. For this reason, users depend heavily on available free workstations, both temporally and locally. The provision of virtual machines (VMs) accessible via a remote connection could avoid this inflexibility. However, the automatic deployment, operation and remote access of OpenGL-capable VMs with professional visualization applications is a non-trivial task.

In this paper, we discuss a concept for a flexible analysis infrastructure that will be part in the project ASTOR, which is the abbreviation for "Arthropod Structure revealed by ultrafast Tomography and Online Reconstruction". We present an Analysis-as-a-Service (AaaS) approach based on the on-demand allocation of VMs with dedicated GPU cores and a corresponding analysis environment to provide a cloud-like analysis service for scientific users.

INTRODUCTION

Particle accelerators like ANKA [1] provide synchrotron radiation for the investigation of solid material and biological samples. X-ray imaging techniques produce large data sets of up to several 100 Gigabytes. The subsequent processing of the data requires special analysis applications whose features and performance depend heavily on available CPU, GPU and memory resources. Typical analysis application like Amira [2] or VG Studio MAX [3] provide visual output based on OpenGL [4], a multi-platform API for rendering vector graphics. Up to now, several conditions concerning the user analysis process have complicated a flexible analysis workflow:

- transfer, access and storage of huge data sets
- required amount of computing and memory resources
- OpenGL/DirectX capable GPUs for visual output, eventually also CUDA [5] or OpenCL [6] are necessary
- expensive workstation licenses for commercial analysis software

All aspects mentioned above prevent the use of standard end-user devices for analysis and result in the operation of dedicated workstations for scientists. The required professional workstations need to provide a high-throughput network connection to the data set storage, licensed analysis software and hardware setup. This scenario implies further disadvantages for users. They have to rent free time ranges for the workstation usage. Therefore they have to be present in special computing rooms within their institution shared with other users. Due to the operation of powerful hardware combined with an increased heat generation, inside the computing room one can often sense a high background noise. Furthermore, the workstations are typically configured with limited guest account privileges, which is why users have to consult an administrator to install additional software packages for a special purpose.

To improve this situation, one part of the project ASTOR deals with the automatic deployment of virtual machines (VMs) with remote connections supporting the rendering and display of OpenGL features. Scientists should be able to use a web portal with an overview of their available data sets and allocate virtual resources for their analysis on-demand. In this paper we discuss an Analysis-as-a-Service (AaaS) concept and present a first prototype implementation for the synchrotron community.

CONCEPT

The replacement of static localized stand-alone workstation towards a flexible analysis infrastructure concept for scientists is achievable with an Infrastructure-as-a-Service (IaaS) approach. Besides the provision of virtualized analysis environments, a suitable analysis workflow has to be defined which also considers the previous data detection and recording, data set access and the final result archiving. The complete analysis workflow could be considered as a novel Analysis-as-a-Service (AaaS) concept, defining a cloud-like service for individually customized analysis processes.

The main part of this concept is the intelligent integration of virtualized analysis environments. However, until a few years ago, the provision of remote connections to VMs supporting OpenGL or other complex visual APIs was impossible, as most virtualization solutions only came up with simplified graphic interfaces within a guest system. Furthermore the remote access protocol is a key aspect with regard to lossless data compression and smooth transfer of visual information. It must ensure a low-latency user interaction without any disturbing lag effects, as most analysis processes require several hours of work. Currently, there are just a few complete solution suites available offering the provision of virtualized workstations for professional visualization applications:

- Citrix XenDesktop [7]
- Microsoft RemoteFX [8]
- VMware Horizon View [9]

All of them are based on special GPU resources like the NVidia GRID Technology [10]. The allocation of the GPU resources is possible via a special software layer which allows GPU sharing between several VMs. To take advantage of all possible GPU-based features within a VM, it is usual to grant direct access to a dedicated GPU core via PCI passthrough.

A fast network interconnect between VMs and the data storage systems is another critical factor. The remote processing of large data sets in the order of up to 100 GB via network mount points requires a network bandwidth starting at ~10 Gbit/s. In this regard software virtualized network devices are reaching their limits. Providing near-native performance of high-speed network interconnects like 10 Gbit/s Ethernet or even 56 Gbit/s InfiniBand within VMs requires hardware virtualized approaches like Single Root - I/O Virtualization (SR-IOV) [11]. This PCI specification extension allows a single PCI Express (PCIe) I/O device to appear as multiple, separate devices — called Virtual Functions (VF) —a kind of a "light weighted" virtual PCIe device. Each VF can be allocated to one VM via PCI passthrough.

To ensure the best possible user experience, the provision of the virtualized environments has to be integrated within the IT systems of the corresponding synchrotron institution. Users should be able to authenticate with their existing federated accounts towards a web service providing their data sets and analysis infrastructure. New created VMs should contain the user-specific network mount points for the data set access. Pre-installed proprietary analysis software has to be activated automatically by contacting existing predefined license servers. Non-used allocated resources have to be reassigned to other users. In this case a mechanism is needed which "motivates" inactive user to release their unused resources, e.g. shut down and/or delete the VM.

In the next section, we will describe our prototypic implementation based on the basic conditions mentioned above.

IMPLEMENTATION

One goal of the ASTOR project foresees the provision of virtualized workstations for the execution of analysis software like Amira [2]. Our first implementation for this purpose is based on the VMware Horizon View [9] virtualization solution combined with the vSphere v5.5 hypervisor. We use NVidia GRID K2 GPU [10] resources, which are allocated to the VMs via *Virtualized Dedicated Graphics Acceleration (vDGA)* [12] using PCI passthrough. Each VM gets physical access to a dedicated high-end Kepler core with access to 4GB GDDR5 memory. Combined with the PCoIP [13] remote protocol, all important visualization APIs like OpenGL 4.1x or DirectX 11 are supported. Carefree interactive usage without annoying time delays is already possible from home office with a common DSL connection about 16 Mbit/s and better.

Access and processing of data sets should take place on remote network-attached storage systems, which requires a fast network interconnect. Therefore we will provide Infini-

Band support within VMs. Each VM gets direct access to a hardware-virtualized InfiniBand VF provided by an SR-IOV capable InfiniBand Host Channel Adapter. Deployment and management of running VMs and templates is performant enough with a NFS shared storage system accessible via 10 Gbit/s Ethernet.

All VMware resources are managed by our software layer called *Cloud Service Bus*, which was originally developed for the transparent provision of public and private IaaS cloud resources [14]. It utilizes VMware SDK and the Amazon Web Services (AWS) SDK for the control of the corresponding VMs/templates and appends value-added services. A unified SOAP-API provides all common functionalities.

Currently VMware does not provide any intelligent schedulers for initial placement supporting VMs with allocated PCIe devices within a multi-node cluster. Furthermore, allocation and release of PCIe resources during different VM/cluster states has to be triggered manually. Therefore we extended our management software functionality with an initial placement mechanism for VMs with PCIe devices. The placement decision of new created and existing VMs within the cluster is based on free available PCIe, RAM and CPU resources. This is an important feature for an automated VM deployment as cluster load balancing via live-migration of VMs[1] is not possible with direct hardware access to PCIe devices.

Our management software provides further essential features for the integration of IaaS resources into a holistic analysis workflow. Several prepared VM templates are available with different analysis software packages. Users are able to define their specific VM hardware configuration concerning their needs on-demand. Customization mechanisms within the Cloud Service Bus allow to pre-configure new created VMs with corresponding specific network mount points, license server information and more, as available VM templates are cleaned up by all personalized data.

Users can authenticate themselves with their existing Active Directory (AD) account towards the Cloud Service Bus. The complete resource usage and load statistics is monitored for general liability and billing. These informations are also accessible by the users for their personal purposes. It is possible to define billing algorithms based on the CPU, RAM and DISK consumption respectively wall time reservation of a VM. To prevent resource allocation by inactive users, credits can be assigned. In the case of exhausted credits, VMs will be shut down after a warning email.

Based on the available SOAP-API, we developed a crossplatform web service called YVAINE [15], with a special focus on a user friendly and intuitive design. Users should create and manage their resources with a few mouse clicks only. This service is the basis for the ASTOR web portal and will be combined with an overview concerning the users' data sets. The simple choice of an existing data set will create a specific VM with all necessary software packages, network

---

[1] Also known as VMware vMotion and used by the VMware Distributed Resources Scheduler (DRS) within a multi-note virtualization environment for load balancing.

mounts and required computing resources and provide the corresponding access credentials to the users. Figure 1. illustrates the architecture of our current Analysis-as-a-Service workflow implementation.
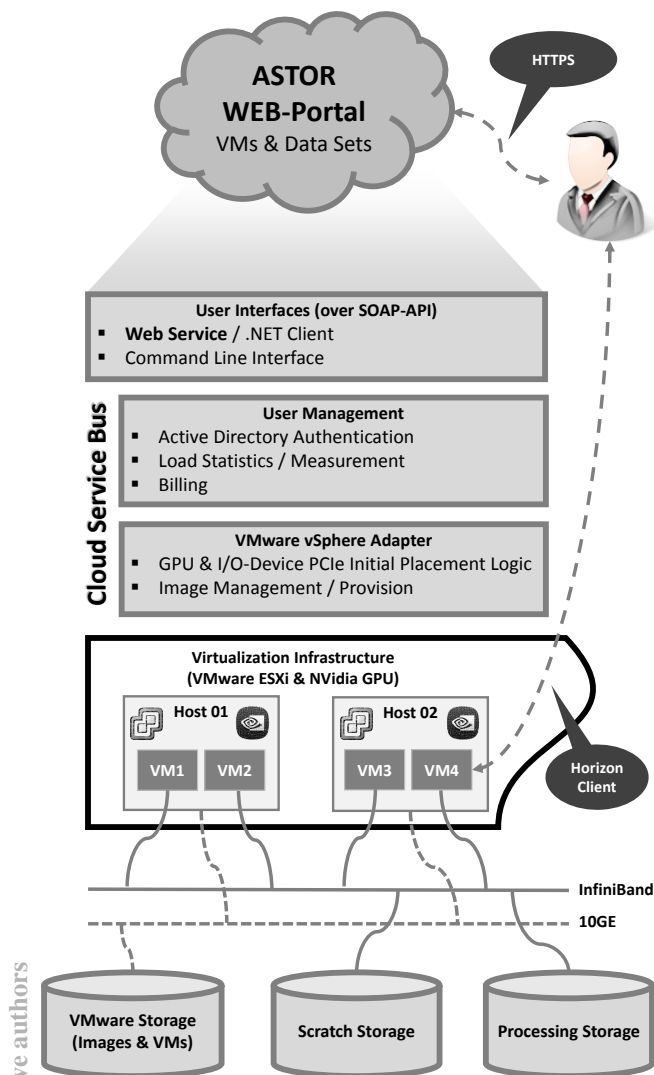
Analysis-as-a-Service (AaaS)

Figure 1: Analysis-as-a-Service Architecture. Users login to the ASTOR web portal, choose their specific data sets and get access credentials to their specific virtual machines for analysis, best suited to their needs.

CONCLUSION & OUTLOOK

The first prototype of the ASTOR "Analysis-as-a-Service" infrastructure is already a significant improvement compared to the traditional strategy with dedicated analysis workstations. Scientific users are flexible concerning their scheduled time and get professional high-end resources and fast access to their large data sets for data analysis at arbitrary locations.

Future plans foresee the support of more federated account solutions like Shibboleth. Furthermore we will improve the initial placement mechanisms based on the past and expected resource consumption of running VMs. The evaluation of alternative storage solutions like noSQL databases is another research aspect within the ASTOR project.

ACKNOWLEDGMENT

REFERENCES

[1] ANKA – the Synchrotron Radiation Facility at the Karlsruhe Institute of Technology (KIT)
    `http://www.anka.kit.edu/`

[2] Amira 3D Software for Life Sciences
    `http://www.fei.com/software/amira-3d-for-life-sciences/`

[3] VG Studio MAX
    `http://www.volumegraphics.com/produkte/vgstudio-max/`

[4] OpenGL - The Industry's Foundation for High Performance Graphics
    `http://www.opengl.org/`

[5] CUDA – Parallel Programming and Computing Platform
    `http://www.nvidia.com/object/cuda_home_new.html`

[6] OpenCL – The open standard for parallel programming of heterogeneous systems
    `http://www.khronos.org/opencl/`

[7] Citrix XenDestkop
    `http://www.citrix.com/products/xendesktop/overview.html`

[8] Microsoft RemoteFX
    `http://technet.microsoft.com/en-us/library/ff817578(v=ws.10).aspx`

[9] VMware Horizon View
    `http://www.vmware.com/products/horizon-view`

[10] NVidia GRID - Graphics-Accelerated Virtualization
    `http://www.nvidia.com/object/grid-technology.html`

[11] Single Root - I/O Virtualization (SR-IOV)
    `https://www.pcisig.com/specifications/iov/single_root/`

[12] Graphics Acceleration in View Virtual Desktops
    `http://www.vmware.com/files/pdf/techpaper/vmware-horizon-view-graphics-acceleration-deployment.pdf`

[13] PC over IP remote protocol
    `http://www.teradici.com/pcoip-technology`

[14] B2B- Infrastructure Integration via Cloud Service Bus
    iX Kompakt: IT-Management in der Cloud, 04-2012, Pages 48–51, September, 2012

[15] YVAINE, SCC News 01/2013, Karlsruhe, Germany
    `http://www.scc.kit.edu/publikationen/scc-news.php`