Karlsruhe Institute of Technology

OpenGL-based Data Analysis in Virtualized Self-Service Environments

Viktor Mauch
Matthias Bonn, Suren Chilingaryan, Andreas Kopmann, Wolfgang Mexner, Doris Ressmann

Steinbuch Centre for Computing (SCC), KIT

Contents

- Introduction
- Motivation
- Concept
- Implementation
- Conclusion & Outlook

Introduction

- *Current analysis workflow for data produced by accelerator labs like ANKA has potential for improvements concerning **flexibility and comfort.** Current drawbacks are:*
  - transfer, access and storage of huge data sets, **up to several 100 Gigabyte**
  - required amount of **CPU and RAM resources**
  - OpenGL/DirectX capable **high-end GPUs** for visual output, eventually also CUDA or OpenCL are necessary
  - expensive **workstation licenses** for commercial analysis software

use of standard end-user devices is not practical for data analysis
operation of dedicated workstations for scientists is inevitable

Motivation

Therefore, we operate computing rooms with several high-end workstations for data analysis.

- **However, users have to …**
  - rent time ranges for workstation usage
  - be present locally
  - contact an IT administrator for additional software (because of guest limitations)
  - get along with higher background noise
  - copy their data sets to the local hard drive
- **Furthermore, …**
  - room place has to be reserved
  - administrators have to provide support for additional hardware instances

automatic deployment of virtual machines (VMs) for users could improve this situation

Concept - Goal

Federal Ministry of Education and Research

Karlsruhe Institute of Technology

ASTOR Project - “Arthropod Structure revealed by ultra-fast Tomography and Online Reconstruction”

WP2: New Analysis-as-a-Service approach based on automatic deployment for VMs

- **Users should be able to …**
  - manage their data sets and VMs via a **web service**
  - request a **VM on-demand** with a corresponding configuration concerning CPU/GPU/RAM resources and software packages
  - connect to the VM from **any location via a network** remote protocol which provides support for rendering and display OpenGL/DirectX visualization features. 
- **Virtual Machines should …**
  - automatically provide **user specific network mounts** to the corresponding data sets over fast interconnect technologies (10 GE / InfiniBand). 
  - contain **all relevant software packages** for analysis, activated via license servers
  - have direct access to **dedicated high-end GPU** resources

Concept – OpenGL capable remote connections

There are just a few complete solutions suites available for offering the provision of virtualized workstations for professional visualization applications:

- **Citrix XenDesktop** via **HDX** protocol
- **Microsoft Hyper-V** via **RemoteFX** protocol, since WinServer 2012 R2
- **VMware vSphere & Horizon View** via **PCoIP** protocol

All solutions are based on shared/dedicated GPU resources within the VMs:

	GRID K1	GRID K2
Number of GPUs	4 x entry Kepler GPUs	2 x high-end Kepler GPUs
Total NVIDIA CUDA cores	768	3072
Total memory size	16 GB DDR3	8 GB GDDR5
Max power	130 W	225 W
Board width	Dual slot	Dual slot
Display IO	None	None
Aux power	6-pin connector	8-pin connector
PCIe	x16	x16
PCIe generation	Gen3 (Gen2 compatible)	Gen3 (Gen2 compatible)
Technical Specifications	GRID K1 Board Specifications	GRID K2 Board Specifications

NVidia GRID K2
with 2 High-End Kepler Cores

Concept – Fast network Interconnect within VMs

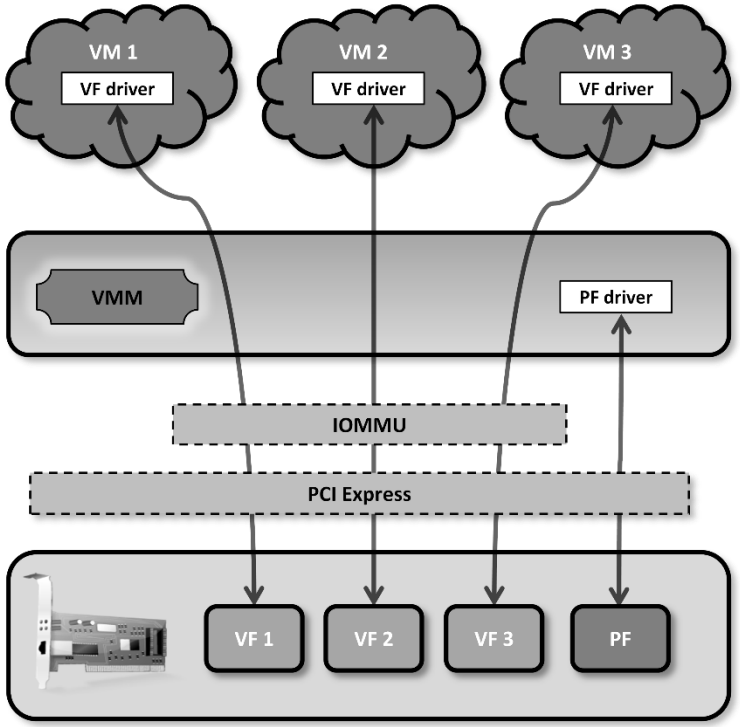
To ensure data processing on remote network mounts,
fast network technologies within VMs are necessary

- 10 Gbit/s Ethernet
- 56 Gbit/s InfiniBand

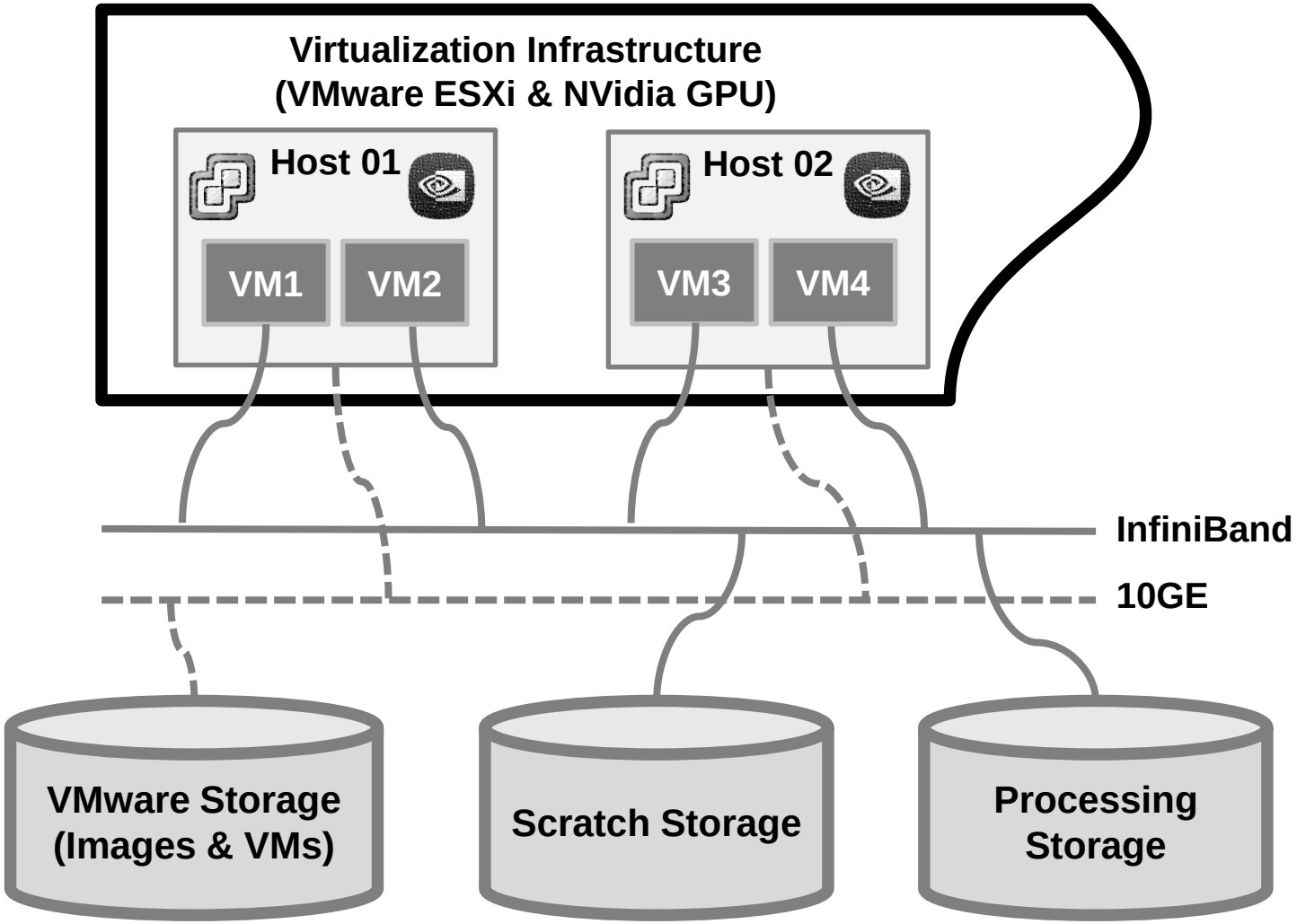
Hardware virtualized approach provides
near-native network performance within a VM:

Single Root – I/O Virtualization (SR-IOV)

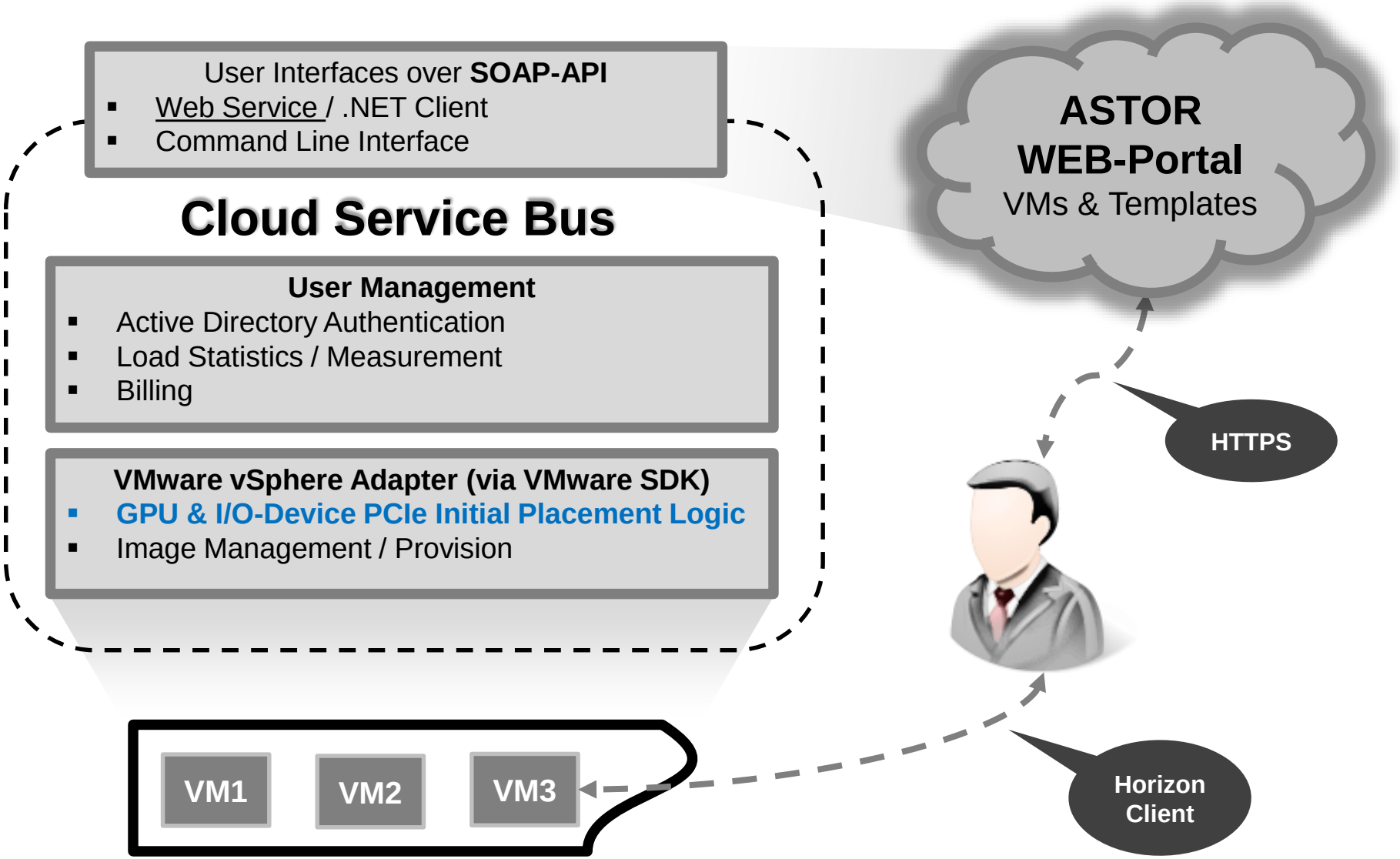
- A single PCIe network I/O device appears as multiple separate “Virtual Functions” (VFs)
- Each VF can be allocated to a VM via PCI passthrough
- Bypass of Hypervisor (Host-CPU) compared with software-based I/O virtualization



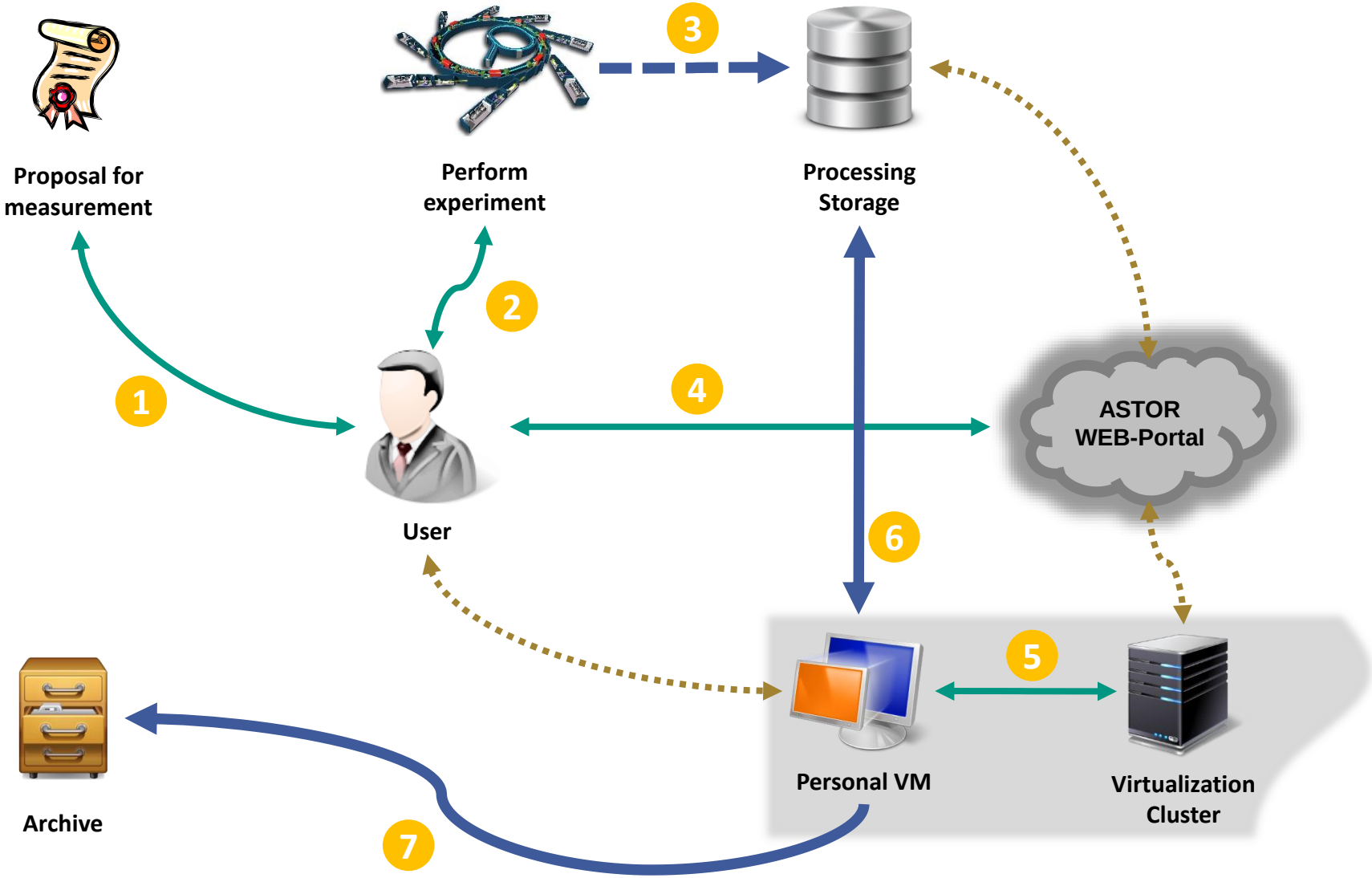
Implementation - Hardware



Implementation – Management Software



Concept - Analysis-as-a-Service Vision



Conclusion & Outlook

- **What we already have:**
  - Initial Placement Logic for VMs with dedicated PCIe devices (GPU and I/O) within a multi-host environment, important for IaaS Cloud Service
  - Web-Portal with Authentication via Active Directory and local accounts
  - On-demand deployment of VMs with dedicated GPUs via a web portal triggered by users
  - Provision of pre-configured VM templates with the corresponding analysis software packages

- **What we will develop/provide in future:**
  - Integration of data set management into the ASTOR Portal
  - Support of Shibboleth and other federated authentication technologies
  - Evaluation of noSQL databases / alternative storage technologies for processing concerning performance and cost-benefit ratio

- Although we are just in the middle of the project funding period, our new analysis implementation of a novel analysis workflow is already a significate improvement compared to traditional strategy with dedicated analysis workstations.