

Virtualization of services

Bonn 08.12.2011

Oliver Oberst

IEKP



Benefits of Virtualization

- Easier access to non-HEP resources
 - Enables use of private and public Clouds (HTCaaS, see e.g. ROCED)
 - Shared resources at universities (see ViBatch)
- Legacy support
 - Different software versions (SL4, SL5)
 - “Cern software appliance”: CERNVM
- Temporary setup of environments
 - Test-deployment of infrastructures and development environments
 - Teaching, Schools (Gridka School, Statistics WS, ...)

Virtualization within HGF-A

■ **Building on Experience** within HGF-Alliance

- Dynamic-Workernode-Virtualization (Alliance project)
- Virtualization for Schools and teaching (Theory-, Analysis-, Computing-WS,...)
- Widespread usage of service virtualization (server consolidation) within the alliance

■ **Current research:**

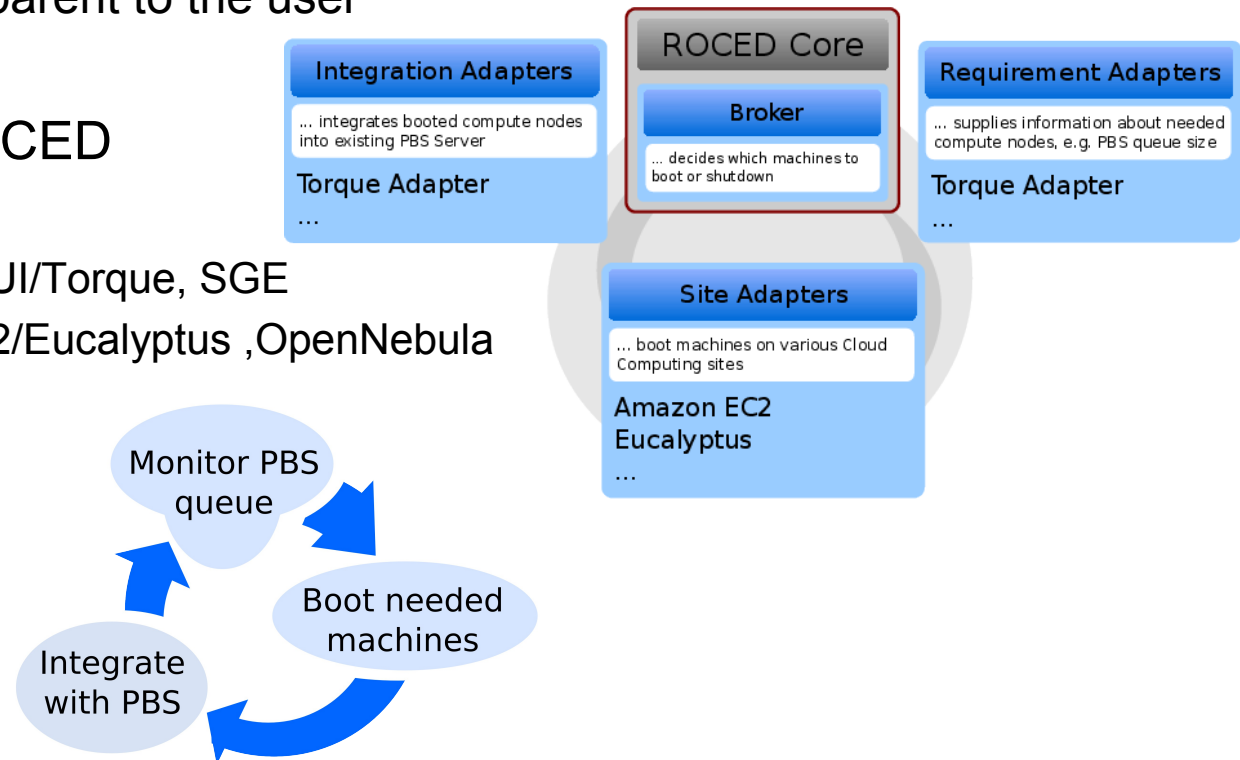
- Combination of Batch-to-Cloud Scheduler ROCED and Dynamic-Workernode-Virtualization ViBatch
 - Test setup running, (Results: CHEP 2012 Abstract submitted)

Cloud Project at KIT (ROCED)

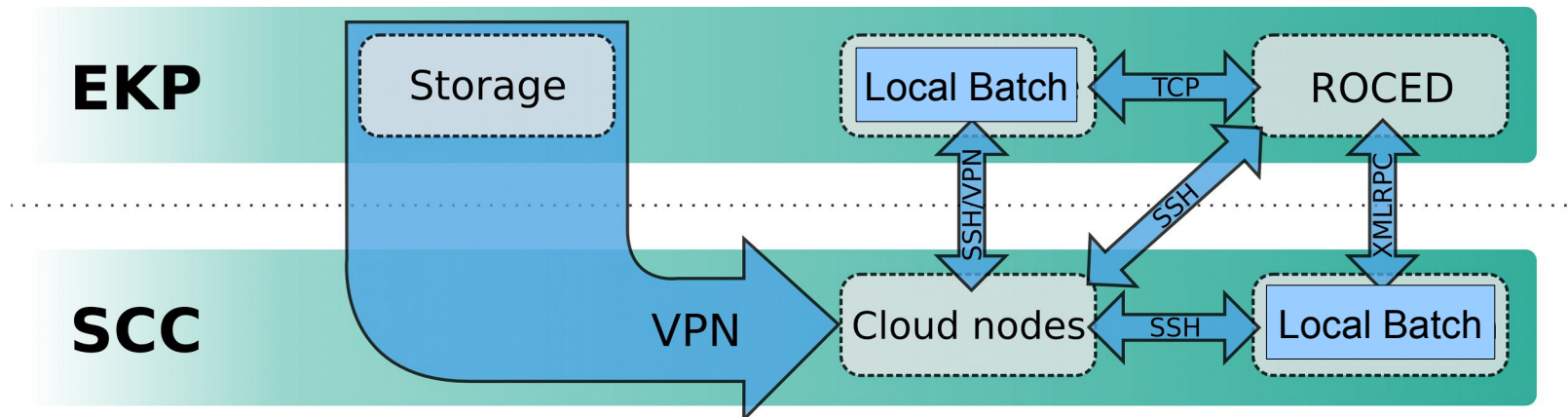
- Extension of local batch server infrastructures to Cloud resources
 - Dynamically, depending on the occupancy of the local infrastructure
 - Cost calculation, able to choose from different Cloud providers based on current cost
 - Completely transparent to the user

■ Implementation: ROCED

- Compatible with
 - Batch-sys: MAUI/Torque, SGE
 - Cloud API: EC2/Eucalyptus ,OpenNebula



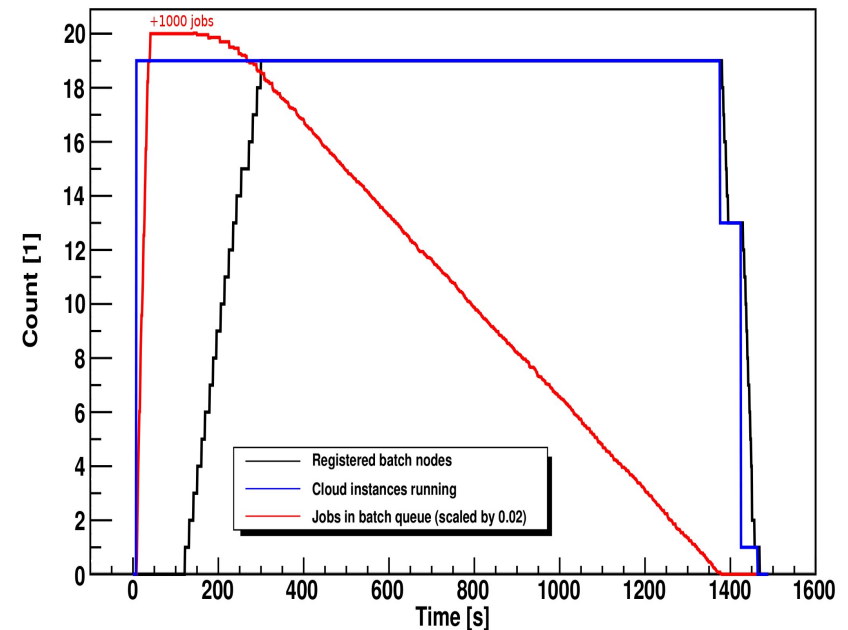
KIT ROCED Setup



- ROCED runs on the same machine as the local batch system.
- Local batch system communicates with its nodes and users via TCP.
- Commands to the OpenNebula host are sent via XMLRPC call.
- The Communication between the Cloud nodes, ROCED and the Cloud Server are done via SSH.
- No modifications to the firewall (besides VPN tunnel) needed.

HTCaaS Cloud Project at KIT (ROCED) II

- Recent changes:
 - OpenNebula Adapter
 - In use at KIT SCC Cloud
 - Daemonization
 - Configuration file infrastructure rewritten
 - “Towards public release”
 - Integrated monitoring
 - Cloud WN usage plots extended
 - In development: web based monitoring output



Dynamic Virtualization Project at KIT

The ViBatch Idea

■ Isolated Computing Cluster

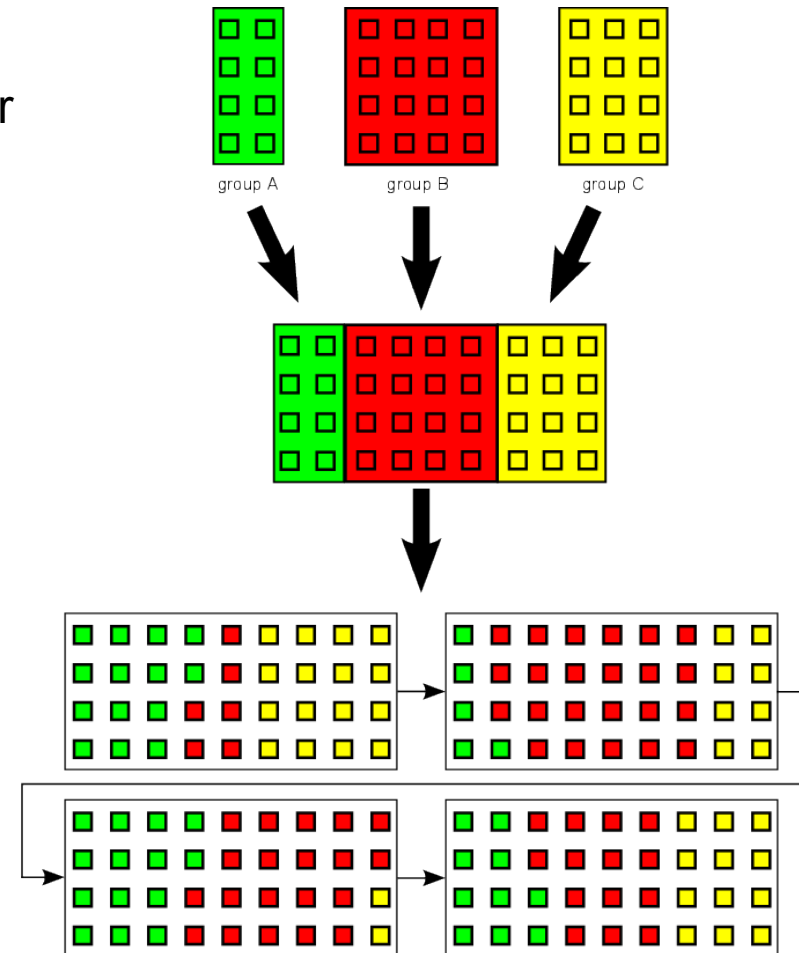
- Each group/institution has sep. cluster
 - ➔ Administration overhead
 - ➔ Can not cover peak loads

■ Shared Computing Cluster

- All groups share one cluster
 - ➔ Setup compromise not always possible
 - ➔ Load-balancing by fair-share

■ Dynamic Partitioned Cluster

- Configure cluster in real-time with VMs
 - ➔ Allows any software/OS configuration
 - ➔ Virtualization layer hidden
 - ➔ Load-balancing by fair-share



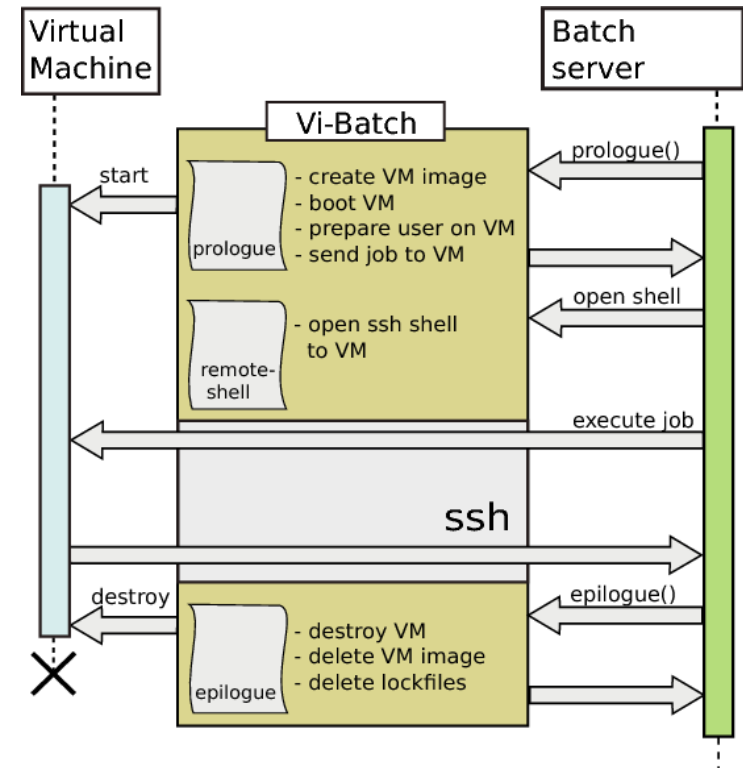
Dynamic Virtualization Project at KIT: Implementation

■ ViBatch

- Wrapper script (Prologue/Epilogue) around the actual computing job inside the batch system
- Virtualization system hidden from the users
- Technical Details:
 - Runs with MAUI/Torque
 - Easy portable to other batch systems
 - Currently uses SLC 5 VM images with CERNVMFS installed (CMS Software included!)
- During first development phase strong cooperation with DESY within this project (Desy implementation `vmimagemanager.py`)
- **In production usage for ~1 year now.**

<https://ekptrac.physik.uni-karlsruhe.de/trac/BatchVirt>

ViBatch workflow:



Current Development: Fusion of ViBatch+ROCED

■ Idea: Fusion of Dynamic Virtualization and Batch-to-Cloud Extension

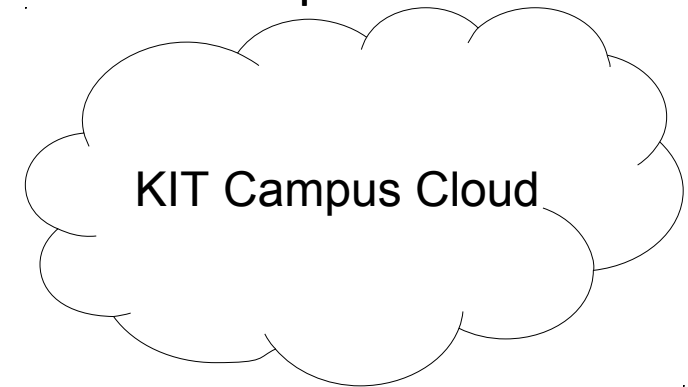
- Extension of local dynamically virtualized resources within private and public Clouds (cloudburst)

■ Current Status:

- First test setup running.
 - Cloud nodes setup manually at “remote” KIT Cloud site
 - SLC5 worker nodes using CERNVMFS
 - Gathering performance/reliability statistics

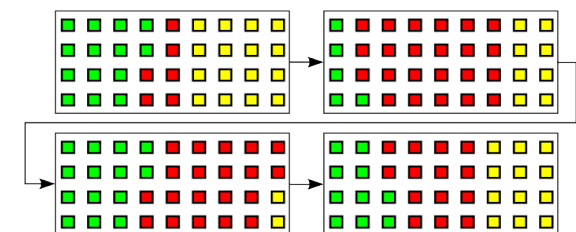
■ To be implemented:

- Central ViBatch and ROCED VM deployment
 - Deployment framework has to be adopted to OpenNebula/Cloud contextualization support
 - Automatic cloud site preparation



+ ROCED

IEKP Cluster Resources



Summary/Outlook

- Virtualization widely used for:
 - Legacy support
 - Test environments
 - Server consolidation (e.g. UNI-KARLSRUHE T3: BDII/DPM-SE VM)
- Virtualization in HTC Computing:
 - Usage of non-HEP resources (ViBatch/ROCED)
 - **ViBatch in KIT IEKP production usage for ~1 year**

Outlook:

- Further development of ViBatch+ROCED fusion.
- (Re-)evaluation of Xen vs. KVM as Xen made it into recent Linux kernel
- Testing of Interoperation of Grid and virtual Cloud resources (e.g. combine glideinWMS Grid farms and Cloud extended (ROCED) farms)