Running Scientific Applications with DIRAC in Federated Clouds

Víctor Méndez Muñoz
PIC, UAB

ADVCOMP, 30 September 2013
Plan

• DIRAC overview
• Federated hybrid cloud computing model for scientific communities
• VMDIRAC: the Cloud extension of DIRAC
• VMDIRAC v0r9 main features
• Next milestones of VMDIRAC roadmap
• Current use cases of VMDIRAC
• Conclusions
DIRAC overview

User Communities

DIRAC Overlay System

Grid Infrastructure

Federated Cloud

Victor Mendez. ADVCOM 2013
User Communities

DIRAC Overlay System

Grid Infrastructure

Federated Cloud

Victor Mendez. ADVCOM 2013
DIRAC overview

User Communities

Dirac Overlay System

Grid Infrastructure

Federated Cloud

LHCb/WLCG Cloud
EGI Federated Cloud
FG Federated Cloud

Víctor Méndez. ADVCOM 2013
DIRAC overview

User Communities

Belle II

LHCb

BES

ILC

Clic

CREATIS

DIRAC Overlay System

Grid Infrastructure

Resources

LCG

Grid WN

Site Gatekeeper

Tier1 VO-box

Services

Job Monitor

DIRAC Job Submission Service

Bookkeeping

Configuration

Message

Agent

Agent

Agent

Federated Cloud

LHCb/WLCG Cloud

EGI Federated Cloud

FG Federated Cloud

Victor Mendez. ADVCOM 2013
• DIRAC has all the necessary components to build ad-hoc distributed infrastructures interconnecting computing resources of different types
• DIRAC has all the necessary components to build ad-hoc distributed infrastructures interconnecting computing resources of different types
Furthermore of grid and cloud resources, DIRAC also integrates:

- Standalone computing clusters:
  - LSF, BQS, SGE, PBS/Torque, Condor"
  - More to come: OAR, SLURM, LoadLeveler. Etc

- Volunteer grids based on BOINC + virtualization technology

- Third party Virtual Clusters (KVM) with vacuum resource allocation, from the grid site job queue instead of DIRAC job queue.
DIRAC Workload Management

User Job

Prod Job

Matcher Service

Task Queue

VO Policies

Sandbox repository

DIRAC WMS

Physicist User

Production Manager

EGI/WMCC Grid

NDG Grid

GISEL Grid

CREAM CE

Pilot Job

Pilot Job

Pilot Job

Pilot Job

EGEE Pilot Director

NDG Pilot Director

EELA Pilot Director

CREAM Pilot Director

Víctor Méndez. ADVCOM 2013
DIRAC Overview

DIRAC Workload Management

Job Priorities for different Groups and Activities

Víctor Méndez. ADVCOM 2013
DIRAC Workload Management

Job Priorities for different Groups and Activities

VO policies application in the central TQ dictates the use of Multiuser Pilot Agents

Víctor Méndez. ADVCOM 2013
Job Priorities for different Groups and Activities

VO policies application in the central TQ dictates the use of Multiuser Pilot Agents

Pilot pre-allocates the computing resources (Pilot auth)
DIRAC Workload Management

DIRAC overview

Víctor Méndez. ADVCOM 2013
DIRAC Workload Management

Job Priorities for different Groups and Activities

VO policies application in the central TQ dictates the use of Multiuser Pilot Agents

Job Brokering,
(User VOMS auth)

Pilot pre-allocates the computing resources
(Pilot auth)

Input/Output Sandbox and third party storage management
For DIRAC users, the use of Storage Element or File Catalog is transparent.
For DIRAC users, the use of Storage Element or File Catalog is transparent.
DIRAC Data Management

For DIRAC users, the use of Storage Element or File Catalog is transparent.

gLite/EGI SE: Standard SRM
Gridftp Globus

Storage Management By Site

LCG File Catalog

DIRAC File Catalog

DIRAC SE

Víctor Méndez. ADVCOM 2013
DIRAC overview

DIRAC Web Portal

Proxy Auth

n12 password:

while trying to make this process as secure as possible by using SSL to encrypt the n12 file with your credentials when it's sent to the server, for maximum security, we recommend that you manually convert and upload the proxy using DIRAC client commands:

dirac-cert-convert.sh YOUR_P12_FILE_NAME.p12
dirac-proxy-init -U -g GROUP_NAME

Víctor Méndez. ADVCOM 2013
DIRAC overview

DIRAC Web Portal

Pilot Management

Job Management

Data Management

Proxy Auth

Víctor Méndez. ADVCOM 2013
DIRAC overview

DIRAC Web Portal

Pilot Management

Job Management

Data Management

Proxy Auth

Job Launch Pad

Accounting Statistics

Víctor Méndez. ADVCOM 2013
Federated hybrid cloud computing model for scientific communities

- Federated Cloud requires SaaS:
  - Integrating different IaaS providers

- Hybrid Cloud requires:
  - Community Clouds IaaS providers
  - Commercial Clouds IaaS providers

- Scientific communities requires
  - A minimal set of functionalities for scientific computing communities:
    - Information, Monitoring, Auth, Image Metadata, Accounting
Federated hybrid cloud computing model for scientific communities

Rafhyc Overview: Resilient Architecture of Federated HYbrid Clouds

Science gateways, pilot factories, job factories
science workflows, scientific applications

Rafhyc VM Interface

Rafhyc Resilient Layer

Rafhyc Multi-Cloud Layer

Specific cloud manager REST APIs

Private Cloud

Public Cloud

Rafhyc Configuration Interface

Rafhyc Federated Hybrid Cloud Services Layer

Specific web services APIs

Federated Cloud Services

Victor Mendez. ADVCOM 2013
VMDIRAC: the Cloud extension of DIRAC

- Rafhyc adoption in DIRAC I
  - Resilient High-level Layer to provide service delivery and management (VMs):
    - Persistent Configuration: An extension of DIRAC Configuration System
    - VM Manager: Taking statistics of VM status
    - VM Scheduler: Integrated with DIRAC TQ job brokering, or with third party job brokers
    - VM Interface: An extension of the DIRAC Web portal
  - Multi-Cloud Low-level layer:
    - OpenNebula OCCI 0.8 and rOCCI 1.1, CloudStack 2, OpenStack nova 1.1 driver, EC2 Amazon
    - Image Context Manager
VMDIRAC: the Cloud extension of DIRAC

• Rafhyc adoption in DIRAC II

• Federated Hybrid Cloud Services Low-level layer:
  - V0r9: Federated static info managed by admins at DIRAC Configuration System
    • Image Metadata, IaaS Information
  - Third party VOMS authentication is integrated in DIRAC (OpenNebula, OpenStack)
  - DIRAC VM monitoring is provided for VOs
  - Third party ganglia monitoring system has been tested for IaaS monitoring of the VMs
  - Next: Accounting
VMDIRAC Architecture

VMDIRAC: the Cloud extension of DIRAC

VM Scheduler
- VM Launcher

VM Manager (VM Status Manager)

Image Context Manager
- OpenNebula Contextualized Image Driver
- OpenStack Contextualized Image Driver
- CloudStack Ad-hoc Image Driver

Configuration Server (VM part)
- RunningPods
- Cloud End-points
- DIRAC Image
- DIRAC Images

Configuration Interface

OCCI 0.8
- DIRAC Driver

Nova 1.1
- DIRAC Driver

CloudStack 2
- DIRAC Driver

OpenNebula PIC.es
OpenStack CC.IN2P3.fr
CloudStack USC.es
VMDIRAC: the Cloud extension of DIRAC

VMDIRAC Architecture

VM Scheduler
  VM Launcher

VM Manager
  (VM Status Manager)

Image Context Manager
  - OpenNebula Contextualized Image Driver
  - OpenStack Contextualized Image Driver
  - CloudStack Ad-hoc Image Driver

Configuration Server (VM part)
  - RunningPods
  - Cloud End-points
  - DIRAC Image

Configuration Interface

Running Pod: logical abstraction particular running conditions.

Víctor Méndez. ADVCOM 2013
VMDIRAC: the Cloud extension of DIRAC

VMDIRAC Architecture

- VM Interface
- Configuration Interface
- Configuration Server (VM part)
  - Running Pods
  - Cloud End-points
  - DIRAC Images
- VM Scheduler
  - VM Launcher
- VM Manager (VM Status Manager)
- Image Context Manager
  - OpenNebula Contextualized Image Driver
  - OpenStack Contextualized Image Driver
  - CloudStack Ad-hoc Image Driver
- OCCI 0.8 DIRAC Driver
- Nova 1.1 DIRAC Driver
- CloudStack 2 DIRAC Driver

Running Pod: logical abstraction particular running conditions.

DIRAC image: boot image and context

Víctor Méndez. ADVCOM 2013
VMDIRAC: the Cloud extension of DIRAC

VMDIRAC Architecture

- VM Interface
- Configuration Interface
- Configuration Server (VM part)
- Running Pods
- Cloud End-points
- DIRAC Images
- VM Scheduler
  - VM Launcher
- VM Manager (VM Status Manager)
- Image Context Manager
  - OpenNebula Contextualized Image Driver
  - OpenStack Contextualized Image Driver
  - CloudStack Ad-hoc Image Driver
- OCCI 0.8 DIRAC Driver
- Nova 1.1 DIRAC Driver
- CloudStack 2 DIRAC Driver
- OpenNebula PIC.es
- OpenStack CC.IN2P3.fr
- CloudStack USC.es

Running Pod: logical abstraction particular running conditions.

DIRAC image: boot image and Context

Cloud End-points: Specific IaaS API and Context

Víctor Méndez. ADVCOM 2013
VMDIRAC: the Cloud extension of DIRAC

VMDIRAC Web Portal
VMDIRAC: the Cloud extension of DIRAC

VMDIRAC Web Portal

Víctor Méndez. ADVCOM 2013
VMDIRAC: the Cloud extension of DIRAC

VMDIRAC Web Portal
VMDIRAC v0r9: main features

- X509 proxy authentication and authorization transparent for user by the DIRAC user/group system integrating third party VOMS

- DIRAC image setup to run VMs:
  - Ad-hoc image
  - Golden image and dynamic contextualization
    - HEPiX contextualization (OpenNebula, OpenStack)
    - SSH generic contextualization
VM horizontal auto-scaling setup I:

• VM allocation policy
  - Elastic: Depending in the work of the DIRAC TQ:
    • CPUPerInstance: VM Scheduler option to balance efficiency and wall time:
      - a) Zero to submit a new VM with no minimal CPU in the jobs of the tasks queue.
      - b) A longer value could be the average required CPU of the jobs as a compromise solution between VM efficiency and total wall time.
      - c) A very large value to maximize the efficiency in terms of VM creation overhead, for the cases where the production total wall time is not a constrain.
  - Static: A constant number of VMs defined by IaaS provider
• VM horizontal auto-scaling setup II:
  • VM stoppage policy:
    - Elastic: VM is automatically stopped by DIRAC if there are no more jobs running in the last VM halting margin time, which is an option to be setup.
    - Never: VMs is not automatically stopped by DIRAC
      • VM is requested to stop by the VO operator or by IaaS provider using DIRAC interface (Web or HEPiX machine features)
      • VM Monitor Agent manages to orderly shutdown
Next milestones of VMDIRAC roadmap

- v1r0: Cloud VM allocation by resource use accounting and top-down resource usage policy
- v1.r1: Integration of the Federated Cloud Accounting Service
- v1.rX: Integration of other Federated Cloud Services: Information System, Federated Market Place
- v1.rX: Cost/Price efficiency integration of commercial and community Clouds
Current use cases of VMDIRAC

Cloud Resources in the LHCb computing

- Running on Production at:
  - CERN (OpenStack)
  - PIC (OpenNebula)

- Jobs Types:
  - LHCb Monte Carlo Simulations
  - Data processing (LHCb offline)
Current use cases of VMDIRAC

Cloud Resources in the LHCb computing

- Integrating Cloud end-points as computing resources of DIRAC sites lcg.cern.ch and lcg.pic.es by HEPiX context

- VM multicore optimization on memory consumption of Nathalie Rauschmayr work with GaudiMP, the LHCb software framework. Integration with DIRAC in CHEP2013
Cloud Resources in the LHCb computing

Current use cases of VMDIRAC

Full Picture, courtesy of Mario Ubeda in CHEP2013
Víctor Méndez. ADVCOM 2013
Current use cases of VMDIRAC

Cloud Resources in the Belle II computing

- Testbed DIRAC WMS with VMDIRAC located in Krakow, but Configuration, Accounting and other systems are fetching from master Belle II DIRAC server located in KEK (Japan)
- Testbed is in simulation jobs, it is not using input data for jobs, but in future will be
Current use cases of VMDIRAC

Cloud Resources in the Belle II computing

- Output ROOT files are stored straightly on grid SEs and registered in LFC and Amga metadata catalog (in KEK)
- VMDIRAC is connected to CC1 cloud now by EC2 interface and soon to OpenStack cloud in Zurich
- Testbed successful, going to production in October 2013
Current use cases of VMDIRAC

Cloud Resources in the Belle II computing

Belle II Cloud scheme courtesy of Rafał Grzymkowski in CHEP2013
Current use cases of VMDIRAC

WeNMR-VMDIRAC EGI Fedcloud Use Case

- Validating and improving biomolecular NMR structures using VCing, a suite of ~25 programs
- Demonstrated using EGI Fedcloud infrastructure
- EGI Fedcloud computing model is moving beyond the frontier, integrating multiple IaaS providers, and deploying necessary services for eScience communities
WeNMR-VMDIRAC EGI Fedcloud Use Case

- **WeNMR with VMDIRAC** taking advantage of automatic VM management in VMDIRAC
- VMDIRAC portal as VM scheduler
- DIRAC broker is not involved, job payload for the VCing VM is provided by the ToPoS server
- A mix of *ad-hoc* VCing image and DIRAC ssh contextualization, which runs only the VM Monitor to manage VM status and stoppage
Current use cases of VMDIRAC

WeNMR-VMDIRAC EGI Fedcloud Use Case

Víctor Méndez. ADVCOM 2013
France Grilles Cloud and FG-DIRAC portal

- FG DIRAC portal is supporting different communities, mainly of life sciences
- FG Cloud aggregates multiple IaaS providers, using the FG-DIRAC portal with VMDIRAC server, successful in testbed and now is going to production
- Contextualization is ssh for generic image management, HEPiX with Cernvm can also be used
Conclusions

- VMDIRAC is a proved tool to aggregate IaaS providers in the level of NGIs supporting multiple VOs, and also in medium and big scientific communities.

- VMDIRAC provides solutions to SaaS deployment and management in a wide range of scientific communities.

- Small communities can be federated to face the DIRAC portal operations (EGI Fedcloud, FG Cloud).
Conclusions

- Medium and big communities may exploit their own DIRAC portal with Cloud resources (LHCb WLCG Clouds, Bell II private Cloud)

- VMDIRAC strategy is addressing sustainability through *industrial concentration* of SaaS management in Federated Hybrid Clouds, at the same time allowing *local development* by the aggregation of distributed IaaS resources

- Detailed instructions for potential IaaS providers and users in the “How to” paper in Proc. ADVCOMP 2013
News in
https://github.com/DIRACGrid/VMDIRAC/wiki

Thanks