Improving experimentations on Grid'5000 HEMERA IPL Evaluation

D. Balouek-Thomert T. Buchert S. Delamare M. Imbert L. Lefèvre M. Liroz L. Nussbaum J. Pastor <u>L. Pouilloux</u> F. Quesnel F. Rossigneux J. Rouzaud-Cornabas C. Ruiz

ASCOLA - ALGORILLE - AVALON - MYRIADS - ZENITH

17/12/2014

1 Grid'5000 overview

- 2 Experiment automation
- 3 Virtualization
- 4 Storage
- 5 Simulation

6 Energy

Grid'5000 overview

The Grid'5000 infrastructure



- 10 independent sites across France and Luxembourg
- 10 Gb dedicated network backbone
- 23 clusters with various hardware
- more than 7000 cores

A powerful infrastructure for computer science experimentations

Grid'5000 overview

A large research applicability



Full control to the experimental conditions

HEMERA IPL

Improving experimentations on Grid'5000

Platform tools

OAR: versatile batch scheduler

Users can reserve any kind of ressources (nodes, network, storage)

KADEPLOY: scalable, efficient and reliable deployment system

Use the nodes as bare-metal machines with their operating system

$GANGLIA: \ distributed \ monitoring \ system$

Users can retrieve metrics on the nodes (cpu, memory, network, ...)

Grid'5000 REST API

Hardware and network information, jobs, deployment and metrics

Limits of the platform

Long learning curve, especially for large scale experiments

- a lot of source of failures
- multi-sites experiments difficult to be automated
- No virtualization support
- Small storage capacity
- 4 Lack of information for simulation
- Embryonic energy monitoring support

Hemera develops several tools to push away these limits!

Easing experimentations on distributed systems



EXECO

Python library to rapidly develop reproducible experiments on Grid'5000

Rapid experiment development on Grid'5000 $_{\rm Execo}$

Originally developed by M. Imbert (SED-Inria Grenoble)

- manage thousands of process on remote hosts
- interact easily with the Grid'5000 infrastructure
- fully automated experimental engine

Hemera contribution

- contribute to API_UTILS, KADEPLOY, OAR, OARGRID modules
- add PLANNING module to simplify large scale reservation
- \bullet add ${\rm API_CACHE}$ module to manage ${\rm REST}$ ${\rm API}$ caching
- add TOPOLOGY module to obtain network information

Helping users to plan their large scale experiments FUNK: Find your Nodes on $g5\kappa$

A command line tool, based on **PLANNING** that helps you to find:

- how many nodes are available at a given date
- find the maximum number of nodes you can have
- find a time slot where a combination of resources is available

Key facts

- allow to automate multi-sites experiments
- used by more than 20 users
- multi-sites: 20% of jobs, 60 hosts/job (5 hosts/job for other)
- Grid'5000 admin use it for maintenance jobs on the whole platform

Virtualization

Adding virtualization capacities to Grid'5000

A new requirements for cloud computing experiments

Evolution of the network infrastructure

Grid'5000 team

- change the configuration of the DHCP servers
- create a new tool to manage subnets reservations

Integration in the production environment

The production environment now ships with KVM hypervisor

Hemera leads these evolution and develops 2 tools for virtualization: $$\rm VM5\kappa$$ and $\rm KAMELEON$

Virtualization

Developing tools for large scale VM management $_{\rm VM5K}$

A Python package based on ExeCO that provides:

- command line tool to deploy virtual machines (up to 10000)
- automated experimental engine
- API to create customized tools/experiments



More details during A. Lèbre talk at 14:05

17/12/2014 11 / 19

Virtualization

A tool for generating customized appliances KAMELEON

General purpose

- create recipe that describes how to create step by step your distribution
- generate custom KVM, VIRTUALBOX, ISO, KADEPLOY ... images

Known users

- OAR team
- Grid'5000 team
- Stimergy

17/12/2014 12 / 19

Evolution of the storage infrastructure

New tool from the Technical Team

 $_{\rm STORAGE5K:}$ allows to reserve large volume ($\sim 1~\text{TB})$ inside Grid'5000 for a month

Long term large storage on Grid'5000

a new storage policy has been setup and allows teams to have 20 TB for several months

Big data experiments are now much easier!

Hadoop cluster management

What is Hadoop?

The Apache Hadoop software library is a framework for distributed storage and distributed processing of Big Data on clusters.

$HADOOP_G5K$

A Python package based on EXECO that provides:

- command line tool to deploy the hadoop on the platform
- automated experimental engine
- API to create customized tools/experiments

Setup a hadoop cluster on 200 hosts in less than 10 minutes

Experimental campaigns to obtain hardware performance

Measuring flops of clusters, by M. Imbert

An EXECO_ENGINE that use classical benchmarks (ATLAS, HPL) to gather the core and host flops of every cluster. Values have been added in the Grid'5000 API

Measuring disks performance, with P. Veyre (CC-IN2P3)

An EXECO_ENGINE that use FIO to gather the disk performance of every cluster.

Values will be added in the Grid'5000 API

Experimental values can be used in simulator (*e.g.* SimGrid)

Simulation

Accessing the network topology of Grid'5000 $_{TOPO5\kappa}$

A tool based on Execo that create a graph from the Grid'5000 network API.

- visualize the topology of site, a cluster or a job
- exported as SimGrid platform XML file



We can compare experiments with simulation!

Energy

Development of a large monitoring framework KWAPI

- Originally developed for OpenStack in XLCloud project (AVALON)
- Adapted and extended by Hemera and Grid'5000 team
- Provides for energy and network traffic
 - experiment live visualization
 - 2 real-time access to the metric
 - ost-experiment long term storage



Developments is currently being pushed to OpenStack repository

What can we do by combining some of the Hemera tools?

1	create custom environments for PMs and VMs	KAMELEON
2	reserve 400 nodes on the platform	Funk
3	deploy 5000 virtual machines on these nodes	VM5K
4	setup hadoop cluster on these VMs	нд5к
5	follow energy consumption and network traffic	Kwapi
6	compare with SimGrid simulation	TOPO5K + API

Summary

Tool	Domain	Status	Maintenance
ехро	experiments	research prototype	?
xpflow	experiments	research prototype	ADT Cosette
execo	experiments	production	M. Imbert
funk	planification	production	Grid'5000 team
vm5k	virtualization	production	?
kameleon	virtualization	production	ADT OAR-Blossom
kwapi	energy	production	Grid'5000 team
topo5k	simulation	beta	?
hg5k	big data	beta	M. Liroz