

VM5k and DVMS on Grid'5000

Deploying and Managing Thousands of Virtual Machines on Hundreds
of Nodes Distributed Geographically

Jonathan Pastor¹ Laurent Pouilloux²

¹Héméra Phd
ASCOLA - Mines Nantes / Inria

²Hemera Engineer
Inria / ENS Lyon

18-06-2014 / Grid'5000 School

Context

Cloud computing usage is becoming very popular.

- Ever-increasing demand \Rightarrow ever-increasing infrastructure size.
- Problems: scalability, reliability, network overhead, energy but also security, and jurisdiction
- Decentralise the computing ressources

concept of micro/nano datacenters [Greenberg2009] geographically spread.
 \Rightarrow nodes can be far from each other.

Discovery project

<http://beyondtheclouds.github.io/>

- And we want to maximise cooperation between close nodes/micro DCs.
- Example: The DVMS case.

Grid'5000 as a testbed

node
router
repeater
switch

- 10Gb interconnected network
- various hardware (cpu, memory size, disks, network bandwidth)
- KaVLAN: allow to have a single network over the sites
- full experiment stack control (hardware, OS, hypervisor)

Created by topo5k
2014-06-18 11:26:00+02:00
API commit 0b625b83fcdfe9cf2f8850cea0a875143388edc

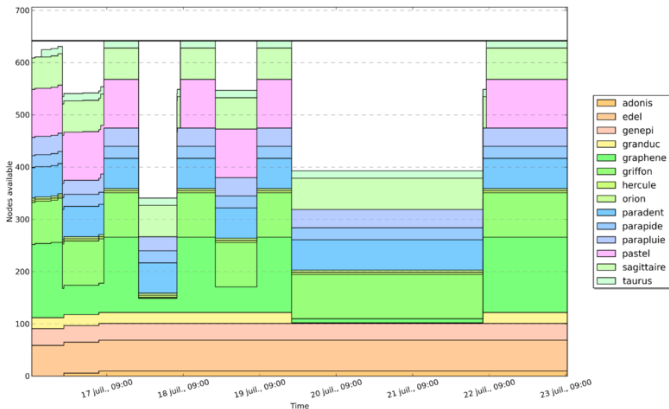
Experimental Workflow

- ① reserve many nodes on different sites, with a global-KaVLAN
- ② deploy thousands of Virtual Machines
- ③ initiate stress process on them
- ④ install DVMS
- ⑤ use vivaldi to compute hosts distances
- ⑥ generate random stress on the virtual machines
- ⑦ live experiment visualization
- ⑧ collect results

(F)ind yo(U)r (N)ode on g5(K)

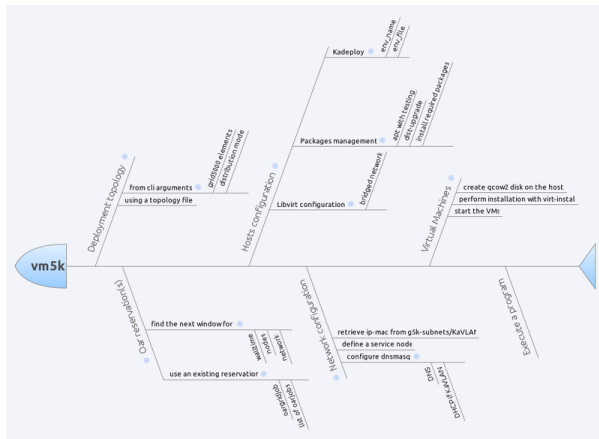
A advanced resources discovery tool for multisite reservation

```
funk -m free -r grid5000:200 -o "-t deploy" -w 12:00:00  
-b helios,sagittaire,nantes,reims,graphite -k -c
```



Automatic Virtual Machines deployment

Moving FLauncher (D. Balouek and F. Quesnel) to vm5k



Tested successfully up to 5 000 VMs on 300 nodes.

Stress initialization

On all running Virtual Machines and using execo

- upload the memtouch
- start the process
- set it's cpu usage to 1% using cpulimit

All VMs are ready to be stressed.

Live visualization

- infrastructure state (VM position and load)
- distance map from Vivaldi (used by DVMS to determine where to migrate the VMs)
- bonus: node live power usage

`http://localhost:9000`

Load events generation

We can tune:

- Gaussian stress distribution for the load
- events frequency

```
<?xml version="1.0" encoding="UTF-8"?>
<events>
  <event type="update_cpu_load" time="0.6322858606354911" target="vm-103" location="10.27.216.103" value="30"/>
  <event type="update_cpu_load" time="0.6501133097844356" target="vm-24" location="10.27.216.24" value="50"/>
  <event type="update_cpu_load" time="0.77156495657007" target="vm-59" location="10.27.216.59" value="100"/>
  <event type="update_cpu_load" time="1.001639188886804" target="vm-123" location="10.27.216.123" value="80"/>
  <event type="update_cpu_load" time="1.7390251934205876" target="vm-19" location="10.27.216.19" value="100"/>
  <event type="update_cpu_load" time="2.161792297887189" target="vm-44" location="10.27.216.44" value="80"/>
  <event type="update_cpu_load" time="3.1211590695284515" target="vm-41" location="10.27.216.41" value="100"/>
  <event type="update_cpu_load" time="3.525741904535117" target="vm-61" location="10.27.216.61" value="40"/>
  ...
</events>
```

Use an execo script to set the value of the load using `cpulimit`

Results Analysis

- Vivaldi map
- Migration statistics
- Bonus: fine-grained power consumption for some nodes

Conclusion

Large scale validation of DVMS taking into account node distance

- -almost- fully automatized experiment
- wide usage of Grid'5000 features (API, Kadeploy, KaVLAN, Kwapi)
- real execution up to 5000 Virtual Machines
- demo available on Challenge_DVMS_Live_-_School_2014

Jonathan Pastor, Marin Bertier, Frédéric Desprez, Adrien Lèbre, Flavien Quesnel, and Cédric Tedeschi. *Locality-aware Cooperation for VM Scheduling in Distributed Clouds*. In Euro-Par 2014, Porto, Portugal, August 2014.

Thank your for your attention. Questions ?