

GRNET Cloud Services

Offering Cloud Services to the Greek R&E Community

Yannis Mitsos¹ ymitsos@noc.grnet.gr, Panos Louridas¹ louridas@grnet.gr

¹Greek Research and Technology Network

September 2010

Enabling the Future Service-Oriented Internet

- ▶ A centralised services provisioning system
- ▶ Virtual working environment services
- ▶ Services targeted to the academic, scientific, research community...
- ▶ ... Provide support to production / administration services as well

Current Services

- ▶ GRNET Virtual Machine (ViMa) Service
- ▶ GRNET AAI Federation
- ▶ HellasGrid infrastructure
- ▶ Pithos online storage service

Coming Services

Based on our experience with IaaS we will offer additional SaaS facilities:

- ▶ Scientific Software as a Service
- ▶ Labs / Science 2.0 Environments as a Service
- ▶ Administration Software as a Service
- ▶ Database as a Service
- ▶ Storage as a Service
- ▶ Grid as a Service (cf. StratusLab)
- ▶ Private cloud as a Service
- ▶ Hybrid cloud as a Service
- ▶ Disaster recovery aaS
- ▶ Data archiving / curation as a Service
- ▶ Backup as a Service

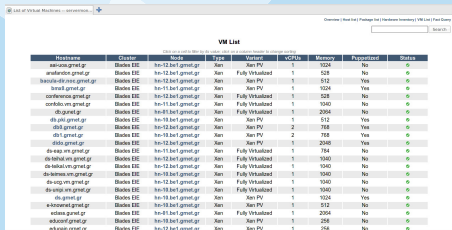
Scope

- ▶ Core services hosting
- ▶ Virtual Private Server (VPS) hosting for our clients

ViMa Service: Initial Deployment

What do we offer currently:

- ▶ Debian 5.0/Xen 3.2
- ▶ Manual administration of the VMs
- ▶ VM hardware node initially assigned in ldap.noc.grnet.gr
- ▶ Control dashboard with Django + libvirt + MySQL



VM List

Click on a VM to view its details, click on a column header to change sort order

Hostname	Cluster	Node	Type	Variant	vCPUs	Memory	Powered	Status
anetson.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Xen PV	1	1024	No	0
anetson.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Fully Virtualized	1	512	No	0
bavaria-dr.noc.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Xen PV	1	512	Yes	0
bavaria.grnet.gr	Blades EDC	he-11.be1.grnet.gr	Xen	Xen PV	1	1024	Yes	0
conference.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Fully Virtualized	1	512	No	0
confobs.vms.grnet.gr	Blades EDC	he-11.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	0
dbi.grnet.gr	Blades EDC	he-11.be1.grnet.gr	Xen	Fully Virtualized	1	2056	No	0
dbi.pilot.grnet.gr	Blades EDC	he-18.be1.grnet.gr	Xen	Xen PV	1	512	Yes	0
dbi1.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Xen PV	2	768	Yes	0
dbi1.grnet.gr	Blades EDC	he-18.be1.grnet.gr	Xen	Xen PV	2	768	Yes	0
dbi2.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Xen PV	1	2048	Yes	0
ds-eap.vms.grnet.gr	Blades EDC	he-11.be1.grnet.gr	Xen	Fully Virtualized	1	784	No	0
ds-holal.vms.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	0
ds-holal.vms.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	0
ds-holal.vms.grnet.gr	Blades EDC	he-18.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	0
ds-vop.vms.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	0
ds-vop.vms.grnet.gr	Blades EDC	he-18.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	0
dsi.grnet.gr	Blades EDC	he-18.be1.grnet.gr	Xen	Xen PV	1	1024	Yes	0
e-kosovnet.grnet.gr	Blades EDC	he-18.be1.grnet.gr	Xen	Xen PV	1	512	No	0
edison.grnet.gr	Blades EDC	he-11.be1.grnet.gr	Xen	Fully Virtualized	1	2056	No	0
edison.grnet.gr	Blades EDC	he-18.be1.grnet.gr	Xen	Xen PV	1	256	No	0
edison.grnet.gr	Blades EDC	he-12.be1.grnet.gr	Xen	Xen PV	1	256	No	0

ViMa Service: Initial Deployment

List of Virtual Machines -- servermon...

Overview | Host list | Package list | Hardware Inventory | VM List | Fact Query

Search

VM List

Click on a cell to filter by its value; click on a column header to change sorting

Hostname	Cluster	Node	Type	Variant	vCPUs	Memory	Puppetized	Status
aai-uoa.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Xen PV	1	1024	No	✓
anafandon.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Fully Virtualized	1	528	No	✓
bacula-dir.noc.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Xen PV	1	512	Yes	✓
bmx0.grnet.gr	Blades EIE	hn-11.be1.grnet.gr	Xen	Xen PV	1	1024	Yes	✓
conference.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Fully Virtualized	1	528	No	✓
confolio.vm.grnet.gr	Blades EIE	hn-11.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	✓
db.gunet.gr	Blades EIE	hn-01.be1.grnet.gr	Xen	Fully Virtualized	1	2064	No	✓
db.pki.grnet.gr	Blades EIE	hn-10.be1.grnet.gr	Xen	Xen PV	1	512	Yes	✓
db0.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Xen PV	2	768	Yes	✓
db1.grnet.gr	Blades EIE	hn-10.be1.grnet.gr	Xen	Xen PV	2	768	Yes	✓
dido.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Xen PV	1	2048	Yes	✓
ds-eap.vm.grnet.gr	Blades EIE	hn-11.be1.grnet.gr	Xen	Fully Virtualized	1	784	No	✓
ds-teihal.vm.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	✓
ds-teikal.vm.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	✓
ds-teimes.vm.grnet.gr	Blades EIE	hn-10.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	✓
ds-ucg.vm.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	✓
ds-unipi.vm.grnet.gr	Blades EIE	hn-10.be1.grnet.gr	Xen	Fully Virtualized	1	1040	No	✓
ds.grnet.gr	Blades EIE	hn-10.be1.grnet.gr	Xen	Xen PV	1	1024	Yes	✓
e-knownet.grnet.gr	Blades EIE	hn-10.be1.grnet.gr	Xen	Xen PV	1	512	No	✓
eclass.gunet.gr	Blades EIE	hn-01.be1.grnet.gr	Xen	Fully Virtualized	1	2064	No	✓
educonf.grnet.gr	Blades EIE	hn-10.be1.grnet.gr	Xen	Xen PV	1	256	No	✓
edugain.grnet.gr	Blades EIE	hn-12.be1.grnet.gr	Xen	Xen PV	1	256	No	✓

Problems

1. Xen instabilities when CPU in high load → kernel panics (fixed)
2. Xen scheduler races when CPU in high load + I/O loads: dom0 starvation
3. CPU starvation under high load when using SMP dom0 and high vCPU/CPU ratio
4. Unacceptable I/O performance!
5. Unsupported graceful shutdown of HVM guests
6. Unstable during live migrations (no-go) and problems with VM clocks (large drifts)
7. Lock-in in one kernel flavor

Minor improvements Xen 3.2 → 3.4

KVM seems to be the only solution

ViMa Service Redux

Redesigned ViMa based on our initial experiences to meet the needs of a production-level Infrastructure as a Service cloud.

Tests with KVM

Exhaustive tests with KVM over the last 8 months.

- ▶ Noticable improvements
- ▶ Major problems were tackled

Problem

Filesystem corruption during consecutive migrations between two hosts

Solution

Caused by page-cache incoherency between the hardware nodes. KVM switch `cache=none` bypasses the host's page-cache and solves the problem.

Problem

Instabilities in live migration of paravirtualized Linux instances with > 1 vCPUs

Solution

Caused by kvm-clock (paravirt) problems, fix exists in kernel 2.6.33+.
Workaround with LD_PRELOAD wrapper for the `ioctl()` that suppresses the kvm-clock capability advertisement.

ViMa Service: Second Development

Prerequisites:

- ▶ Actual cluster management
- ▶ A firm hypervisor
- ▶ Web frontend for OOB access / power control
- ▶ Open architecture :)

Deployment:

- ▶ KVM
- ▶ Google's Ganeti project with GRNET NOC contributions
- ▶ Custom web frontend in Django (under development)

Results

- ▶ 10-node cluster in GRNET's main Data Center in production
- ▶ 150 VMs in KVM running core services (DNS, mail, radius, IdP, ldap etc.)
- ▶ Kernel 2.6.32 + qemu-kvm 0.11.1 with KSM support → ≈ 1.2 GB RAM de-duplicated
- ▶ Shared block storage over FC
- ▶ Working live-migrations for all O/Ses (Linux, Windows, FreeBSD)
- ▶ Working paravirtualized drivers for Linux, Windows
- ▶ No stability/performance problems have been identified

Ganeti: Shared-nothing Cluster

- ▶ No SPoF in the cluster
- ▶ Master node: cluster control, master IP address
- ▶ Master candidate: cluster configuration copy as master candidate
- ▶ Simple node: only as node controller
- ▶ If master fails: master failover to a master candidate

Ganeti: Storage Support

Four types of storage (“disk templates”):

- ▶ diskless: only used for special purpose operating systems or for testing
- ▶ file: the instance uses plain files as backend for its disks
- ▶ plain: the instance uses LVM devices as backend for its disks
- ▶ DRBD: 1 DRBD-device per guest disk on an LVM. Failover / migration capabilities

GRNET Extensions

- ▶ Shared block storage (SAN) support (in production)
- ▶ Shared file storage (NAS) support (planning)

Ganeti GRNET NOC Contribution

Ganeti modifications, submitted (Google)

- ▶ Shared block-storage support (working migrations and failover)
- ▶ Changes in the KVM backend:
 - ▶ Boot from HTTP + minor bugfixes
 - ▶ VNC forwarding
- ▶ Minor changes in the API

Cooperative upstream, keen in extending Ganeti's functionality

Conclusions

There are many supporting currents in our view of cloud computing:

- ▶ Economies of scale (favourable economic / market environment)
- ▶ Lack of sysadmin talent
- ▶ Quick to market (production)
- ▶ Long-term experience with providing quality services to the academia
- ▶ Excellent cooperation with the public sector / government, possible to exploit synergies (e.g., disaster recovery).