GRNET NOC
Providing IaaS to Greek Academic Users

Datacenter IaaS workshop 2014

George Kargiotakis (kargig@noc.grnet.gr)

whoami

Systems & Services Engineer @ GRNET

Messing with Linux, Security, Privacy

Provisioning Cloud Services

- 1st step
  - From hardware documentation to OS installation
- 2nd step
  - From OS configuration to Virtualization service deployment

Provisioning hardware

- Documentation before anything else!
- Hardware database → **Servermon**
- Servermon:
  - Inhouse project
  - Facilitates server management + monitoring through puppet
  - Documentation source
  - Django (feel free to use/hack)

Servermon

- HWDOC Database with:
  - DC→Row→Rack→Server
  - Server → Rack, Unit, Vendor, Model, Project
    - → Serial, IPMI, def. Admin password, IPMI MAC
- After provisioning a server with puppet, combine puppet database/facts with hwdoc
  - CPU, RAM, nic MAC, OS, IP addresses, firmware, etc
- Fact Query
- Package updates

Servermon

st1-03.okeanos.grnet.gr: Host information

System information

Architecture	amd64
BIOS Date	12/08/2012
BIOS Version	A18
Machine Type	physical
Memory	189.27 GB (182.14 GB free)
Model	ProLiant DL385 G7
Operating System	Debian 7.6
Processor type	AMD Opteron(tm) Processor 6172
Processors	24
Serial number	CZ220106L9
System Vendor	HP
Uptime	2 days
Last updated	2 weeks, 3 days

Location information

Datacenter	ΥΠΕΠΘ
IPMI Hostname	ILOCZ220106L9.serv-mgmt.grnet.gr
IPMI MAC	e4:11:5b:b2:93:d6
IPMI Method	ilo3
Rack	R10
Rack Row	R01-R20
Rack Unit	19

Network information

Toggle

Interface name	IPv4 Address	IPv6 Address	MAC Address
bond0	62.217.118.13/24	2001:648:2ffc:500::13	e4:11:5b:b2:93:ce

Servermon

Rack Info

Name:	R10
Model:	APC NetShelter SX
Mounted Depth:	80
Rack Row:	R01-R20
Position in Rack Row:	11
In Row AC:	False

Equipment Info

Serial	Model	Rack	Unit	Front	Interior	Back	IPMI Hostname	Project	Tickets	Hostname
			42							
			41							
CZ21520115	HP DL380 G7	R10	40 39	✓	✓	✓	ILOCZ21520115.serv-mgmt.grnet.gr	~okeanos soc	—	staging-rd0-01.okeanos.grnet.gr
5C7132P16H	HP DS2600	R10	38 37	✓	✓	✓		~okeanos soc	—	—
5C7135P363	HP DS2600	R10	36 35	✓	✓	✓		~okeanos soc	—	—
CZ2152010C	HP DL380 G7	R10	34 33	✓	✓	✓	ILOCZ2152010C.serv-mgmt.grnet.gr	~okeanos soc	—	staging-rd0-00.okeanos.grnet.gr
5C7143P2AX	HP DS2600	R10	32 31	✓	✓	✓		~okeanos soc	—	—
5C7135P358	HP DS2600	R10	30 29	✓	✓	✓		~okeanos soc	—	—
			28							
			27							
CZ220106KM	HP DL385 G7	R10	26 25	✓	✓	✓	ILOCZ220106KM.serv-mgmt.grnet.gr	~okeanos soc	—	st0-02.okeanos.grnet.gr
CZ220106KR	HP DL385 G7	R10	24 23	✓	✓	✓	ILOCZ220106KR.serv-mgmt.grnet.gr	~okeanos soc	—	st0-03.okeanos.grnet.gr
CZ220106KZ	HP DL385 G7	R10	22 21	✓	✓	✓	ILOCZ220106KZ.serv-mgmt.grnet.gr	~okeanos soc	—	st1-02.okeanos.grnet.gr
CZ220106L9	HP DL385 G7	R10	20 19	✓	✓	✓	ILOCZ220106L9.serv-mgmt.grnet.gr	~okeanos soc	—	st1-03.okeanos.grnet.gr
CZ220102PZ	HP DL385 G7	R10	18 17	✓	✓	✓	ILOCZ220102PZ.serv-mgmt.grnet.gr	~okeanos soc	—	st0-01.okeanos.grnet.gr
CZ220102Q3	HP DL385 G7	R10	16 15	✓	✓	✓	ILOCZ220102Q3.serv-mgmt.grnet.gr	~okeanos soc	—	—
CZ220102PD	HP DL385 G7	R10	14 13	✓	✓	✓	ILOCZ220102PD.serv-mgmt.grnet.gr	~okeanos soc	—	—
CZ220102NS !	HP DL385 G7	R10	12 11	✓	✓	✓	ILOCZ220102NS.serv-mgmt.grnet.gr	~okeanos soc	—	—
CZ220102P1 !	HP DL385 G7	R10	10 09	✓	✓	✓	ILOCZ220102P1.serv-mgmt.grnet.gr	hp-support-lab	—	—
CZ220102NZ	HP DL385 G7	R10	08 07	✓	✓	✓	ILOCZ220102NZ.serv-mgmt.grnet.gr	~okeanos soc	—	gnt5-04.gnt.grnet.gr
CZ220102PV	HP DL385 G7	R10	06 05	✓	✓	✓	ILOCZ220102PV.serv-mgmt.grnet.gr	~okeanos soc	—	gnt5-03.gnt.grnet.gr
CZ220102NW	HP DL385 G7	R10	04 03	✓	✓	✓	ILOCZ220102NW.serv-mgmt.grnet.gr	~okeanos soc	—	gnt5-02.gnt.grnet.gr

re-usable info

- Servermon::hwdoc → LDAP
- LDAP
  - cn==hostname + NIC MACs + puppet classes/variables
- LDAP → DHCP + DNS
  - Extract info from LDAP, publish into DHCP+DNS
  - Servers get static IPs & hostnames based on nic MAC

OS installation

- FAI (Fully Automated Installation)
  - DHCP + PXE
  - Reusable classes per hw type / sw service
  - DHCP IP (hostname) defines class → Install/Wipe/Rescue
    - fai-chboot -c VIMADRBD hostname

    After Installation (or automatically):
    - fai-chboot -d hostname
  - Machines always do PXE boot
    - Fallback to booting from Hard Disk

Configuration Management

- Puppet
  - LDAP Terminus
  - Puppet database

Provision

- users
- packages
- services
- backup

Automatic monitoring

- Icinga
- Munin
- Ganglia
- Logstash/graylog

Configuration management

- Availability (multiple workers)
- Scalability (multiple workers)
- More puppet classes == fewer human errors
- Re-usable components by different services
- Massive changes in < 30' across datacenters
- Git + puppet = complete history of changes
  - Accountability (who did what and when)

Automate!

- Automation == less effort, fewer time spent
- From an empty node to hosting VMs:
  - Documentation: <1' (usually already done at DC install)
  - Servermon→LDAP: 1-2'
  - LDAP→DHCP: 1'
  - FAI: 15'
  - Puppet: 10'
  - Total: <30'
- Can be fully parallelized

Virtualization Solution

- Ganeti (+ KVM)
  - Cluster management by Google
    - GRNET #2 committer
  - Scalable
  - Multiple storage backends
  - Remote API
  - KISS principle
- Multiple clusters
  - Different HW nodes (CPU, RAM, Disks)
  - Different Storage Backends (NFS, Shared block/NetApp, DRBD, RADOS)
  - Different networking (Bridged, Routed)
  - Different users/quotas/resources given

IaaS Platforms

- ViMa
  - Software project name: ganetimgr
  - Stable VPS service
  - Apply → **Approve** / **Install** → Run/Re-install
  - Long-running VMs
  - Geared towards Power users/Administrators
  - Controlled resource usage
  - Monitoring of clusters/nodes/jobs
  - Stateless architecture
  - FAST
  - (Very) easy to setup
  - No API (yet)

IaaS Platforms

- ~okeanos
  - Software project name: synnefo
  - Operates on ganeti clusters
  - Exposes OpenStack APIs (Nova, Neutron, Glance, Cinder) on top of Ganeti
  - Services:
    - Identity
      - incl. Shibboleth authentication
    - Object Storage
    - Compute
      - Quotas per user/project
    - Network
      - Users can create their own virtual networks (mac-filter based private networks)
      - Floating IPs
      - NIC hotplugging

IaaS Platforms

- ~okeanos services (cont):
  - Image
    - User-created custom images
  - Volume
    - VM's disks, snapshots
- Archipelago
  - Unified cloud storage resources
  - Decouples storage resources from storage backends (NFS, RADOS, GlusterFS, etc)
- Very simple UI
  - No administrative interface (yet)

~okeanos

- Cyclades
  - Compute UI ← VMs
  - Create/Expand/Destroy VMqs/Networks
- Pithos+
  - Storage UI ← Storage
  - Sync Clients for Windows/Mac/iOS
- Kamaki
  - cli tool to work with API

Clients/Users

- Who uses our IaaS platforms
  - Students (~okeanos)
  - Teachers/Classes/Labs (~okeanos)
  - Science (~okeanos)
  - NOCs (ViMa)
  - Libraries (ViMa)
  - Research institutions (ViMa)
  - Ministries/Government (ViMa)

Stats

- ViMa:
  - ~1200 Active VMs
  - 125 Users
  - 7 Clusters
- ~okeanos
  - ~7000 Active VMs (>380k spawned)
  - ~3500 Users with VMs (>10k total)
  - 13 clusters

THE problem

How to regulate resource usage ?

- Academic users don't pay for resources.
- If you allow X,Y,Z (with Z as max) for resources they always ask for Z even if they don't need it.

- Possible Solutions:
  - VPS: Co-design solutions, train admins, publish solutions, “rewards”
  - Cloud: Strict quotas per user/project (+exceptions...)

Security

- Deal with thousands of abuse requests
- Dedicated security-aware helpdesk
- Training!

Resources

- Servermon: https://github.com/servermon/
- FAI: http://fai-project.org
- ganetimgr: https://github.com/grnet/ganetimgr/
- synnefo: https://github.com/grnet/synnefo/

Thank you

Questions ?

@grnetnoc

/noc.grnet.gr