

SYNNEFO + GANETI + CEPH.
VANGELIS KOUKIS, TECHNICAL LEAD, SYNNEFO

Running a public cloud: ~okeanos

Ceph Day London
vkoukis@grnet.gr

History

- Design started late 2010
- Production since July 2011

Numbers

- Users: > 3500
- VMs: > 5500 currently active
- More than 160k VMs spawned so far, more than 44k networks

Running a public cloud: ~oceanos

Ceph Day London
vkoukis@grnet.gr

Our choices

- Build own AWS-like service (Compute, Network, Storage)
- Persistent VMs
- Everything open source
- Production-quality IaaS
- Super-simple UI

How?

Running a public cloud: ~oceanos

Ceph Day London
vkoukis@grnet.gr

The tough stuff

- Stability
- Persistent VMs: VMs are not cattle, they are pets
- Commodity hardware
- Scalability
- Manageability: Gradual rollout of upgrades and new features

Running a public cloud: ~okeanos

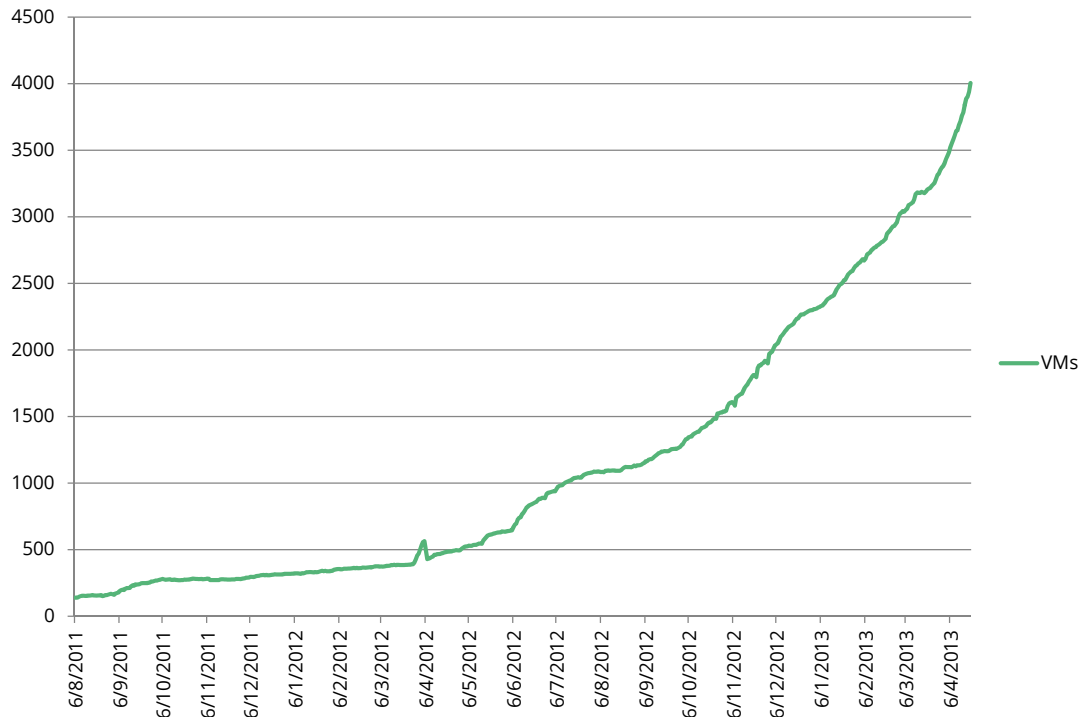
Ceph Day London
vkoukis@grnet.gr

Our approach

- Synnefo
- Google Ganeti
- DRBD
- Ceph
- OpenStack APIs

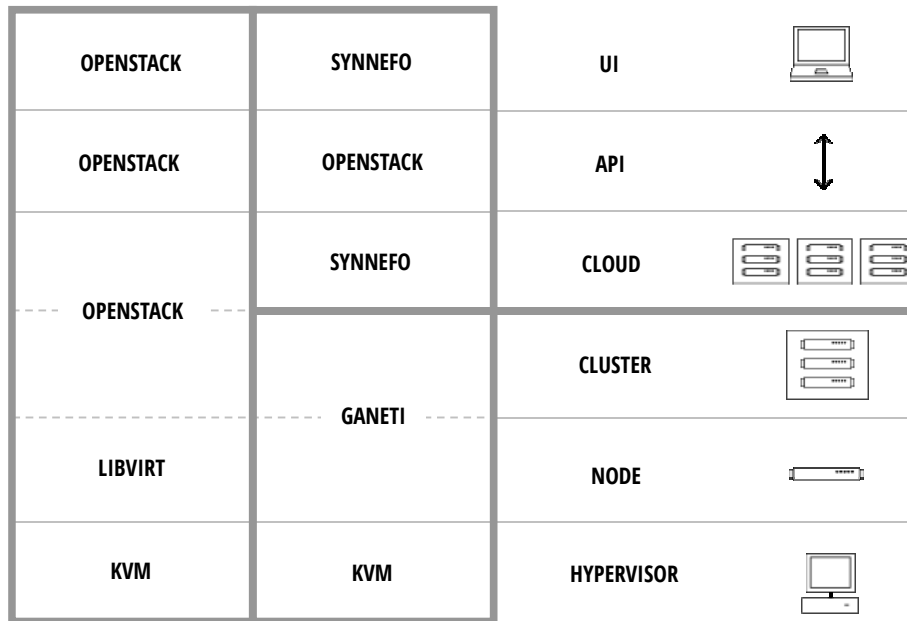
~okeanos VMs

Ceph Day London
vkoukis@grnet.gr



Cluster vs Cloud

Ceph Day London
vkoukis@grnet.gr



Google Ganeti

Mature, production-ready VM cluster management

- used for Google's corporate infrastructure

Multiple storage backends out of the box

- LVM, DRBD
- Files on local or shared directory
- RBD (Ceph/RADOS)

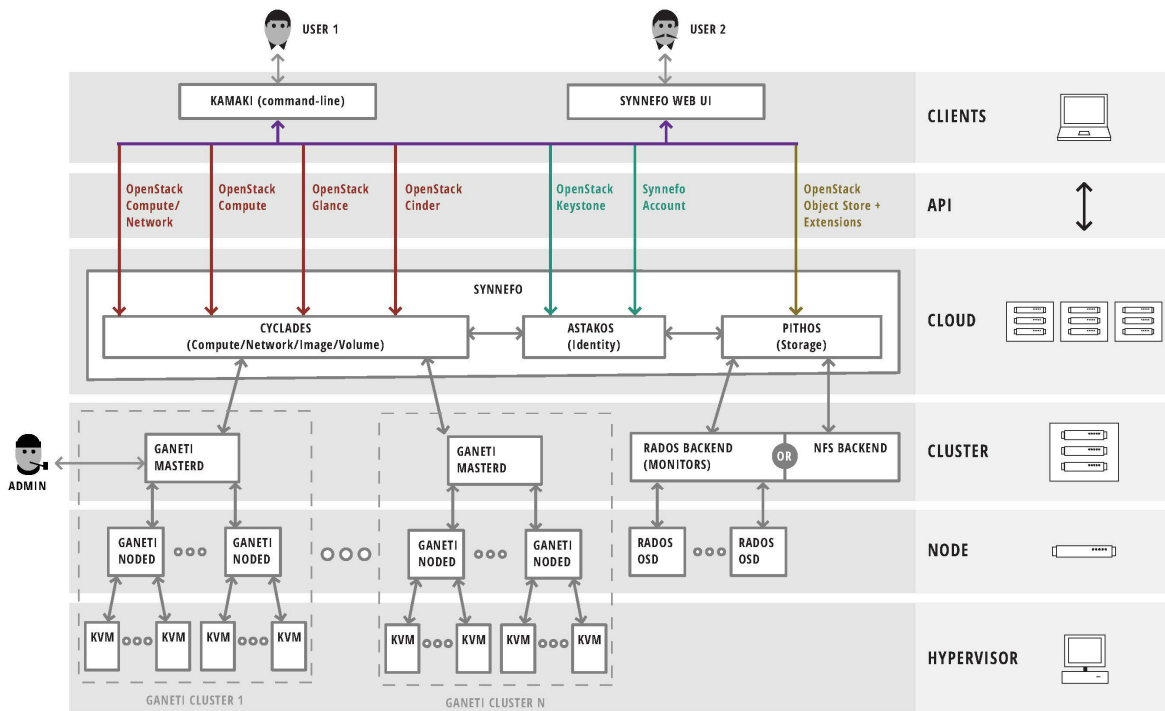
External Storage Interface for SAN/NAS support

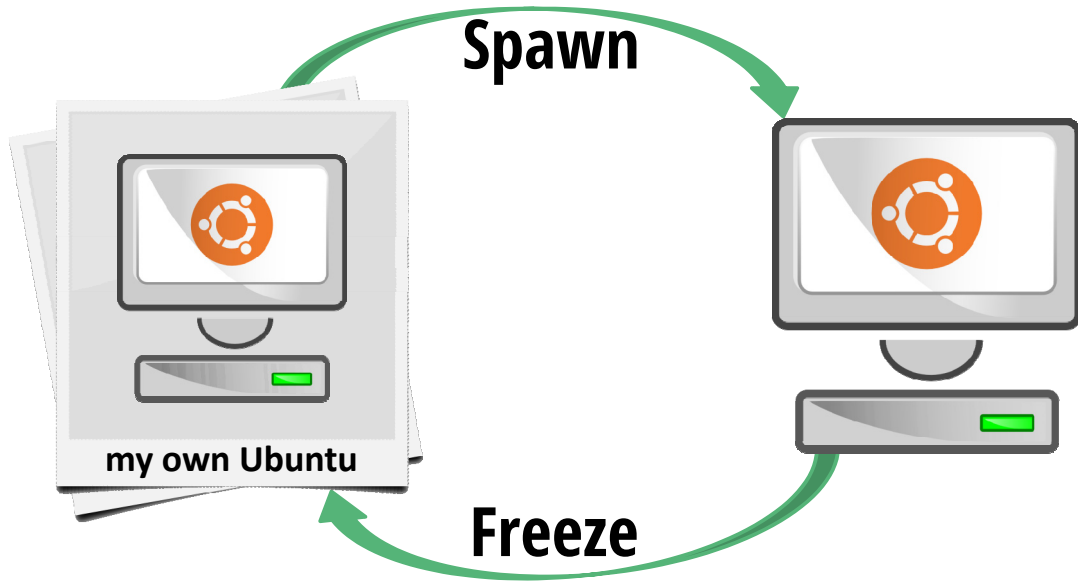
Ganeti cluster = *masterd* on master, *noded* on nodes

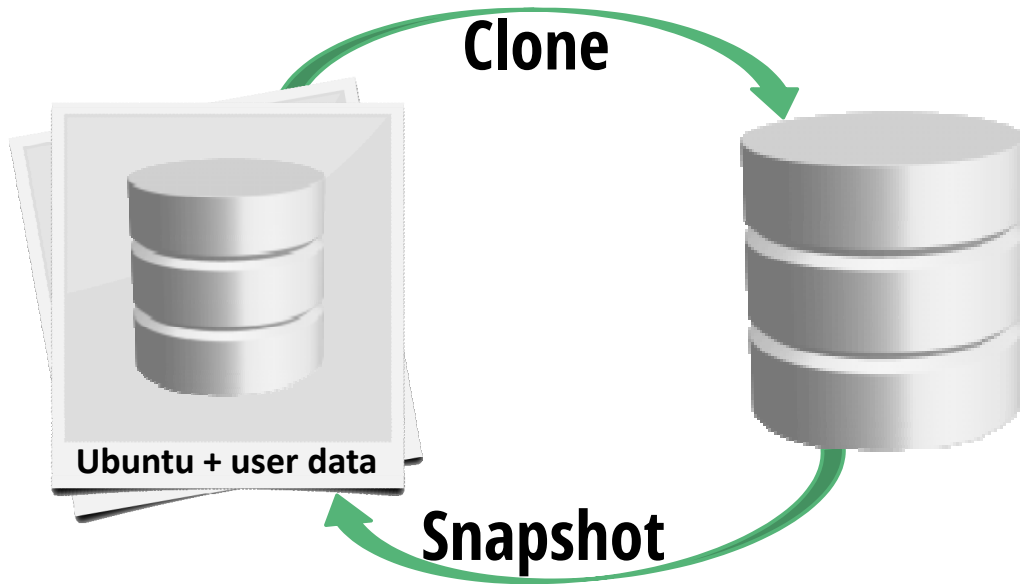
Easy to integrate into existing infrastructure

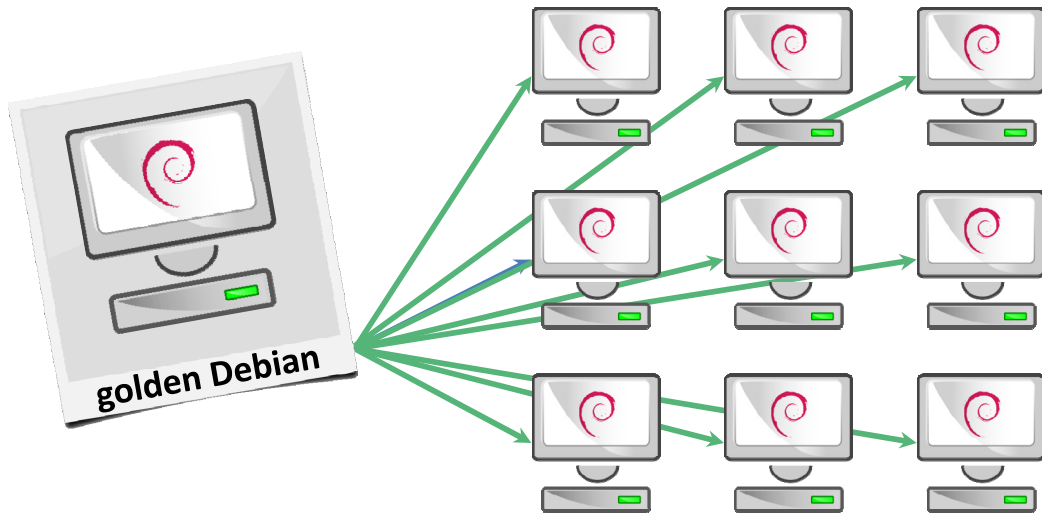
- Remote API over HTTP, pre/post hooks for every action!

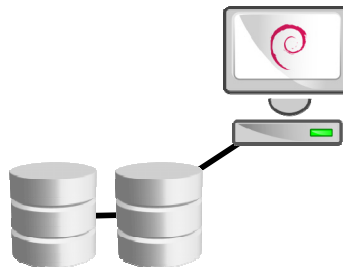
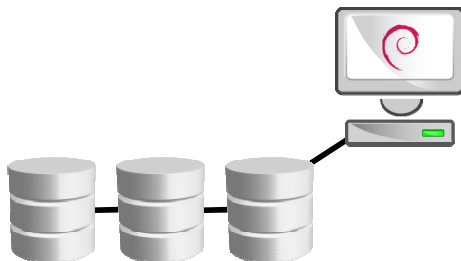
Architecture

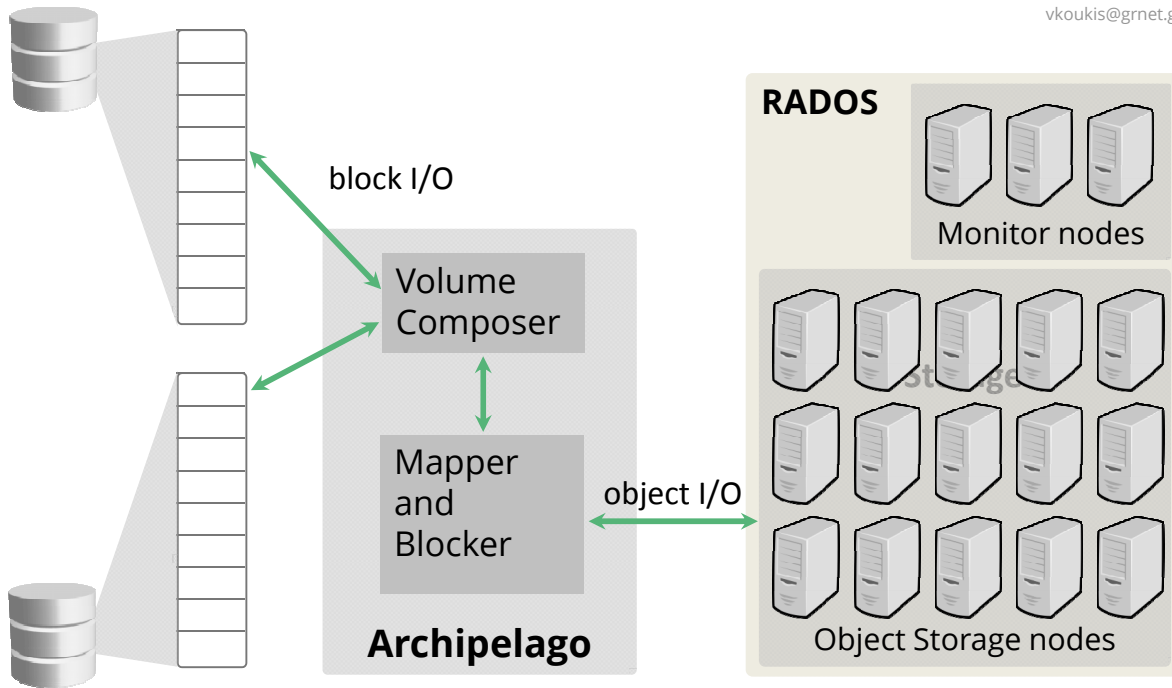




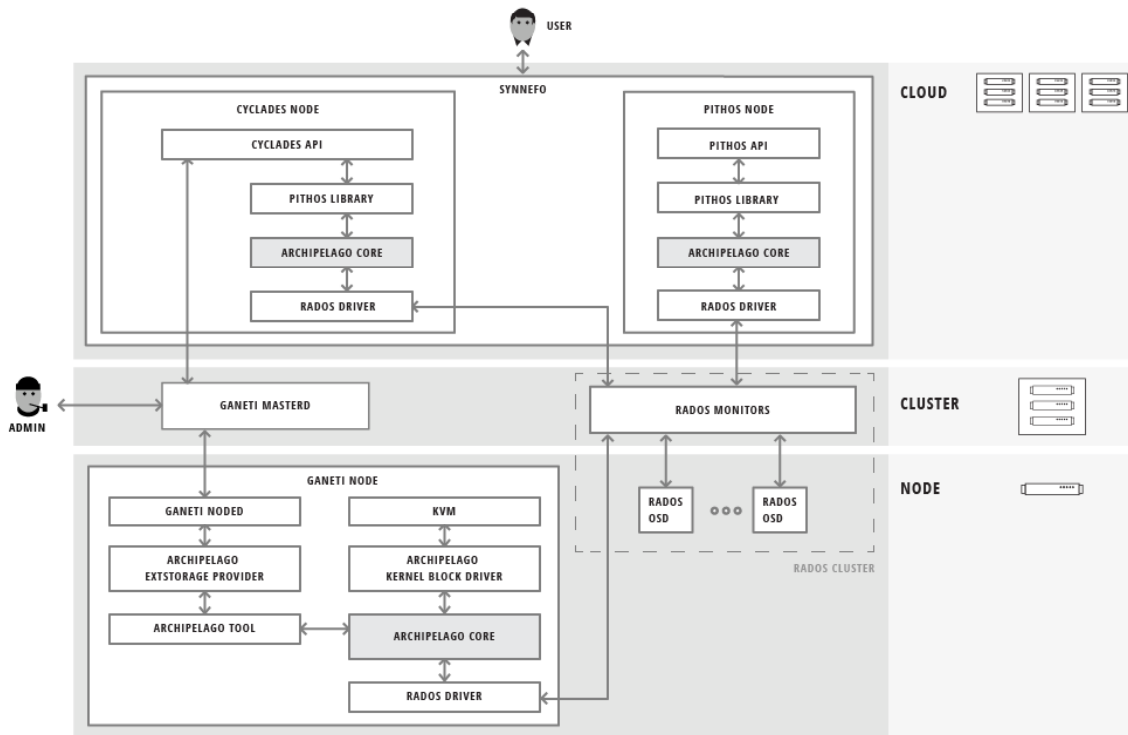








Volumes: Archipelago



Volumes: Archipelago

Ceph Day London
vkoukis@grnet.gr

Unified storage for Files, Images ↔ Volumes

Thin layer over the actual storage cluster

Storage backend agnostic

- NFS, RADOS, ...

Efficient syncing / sharing of Images as files on the storage service

Zero-copy cloning of volumes from Images

Experience

Ceph Day London
vkoukis@grnet.gr

Operations

- Rolling hardware and software upgrades
 - kernel, Ganeti, RADOS, Synnefo
 - with no VM downtime
- Node evacuations with live VM migrations
- Cross-datacenter move, Intel → AMD, no VM downtime
- On-the-fly migration from NFS-backed storage to RADOS
- IP renumbering of all VMs

Experience

Ceph Day London
vkoukis@grnet.gr

Scalability

- From few physical hosts to multiple racks
- dynamic addition of Ganeti clusters

Diverse workloads

- Different network and storage backends
- Choice exposed to the user

synnefo



Try it out!

Ceph Day London
vkoukis@grnet.gr

<http://www.synnefo.org>

synnefo

Thank you!



Ceph Day London
vkoukis@grnet.gr