SYNNEFO + GANETI + CEPH.
VANGELIS KOUKIS, TECHNICAL LEAD, SYNNEFO
Running a public cloud: ~okeanos

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: ~okeanos

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: ~okeanos

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

Our approach

- Synnefo
- Google Ganeti
- DRBD
- Ceph
- OpenStack APIs
~okeanos VMs

Ceph Day London
vkoukis@grnet.gr
## Cluster vs Cloud

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OPENSTACK</td>
<td>SYNNEFO</td>
<td>UI</td>
<td></td>
</tr>
<tr>
<td>OPENSTACK</td>
<td>OPENSTACK</td>
<td>API</td>
<td></td>
</tr>
<tr>
<td>OPENSTACK</td>
<td>SYNNEFO</td>
<td>CLOUD</td>
<td></td>
</tr>
<tr>
<td>LIBVIRT</td>
<td>GANETI</td>
<td>CLUSTER</td>
<td></td>
</tr>
<tr>
<td>KVM</td>
<td>KVM</td>
<td>NODE</td>
<td></td>
</tr>
<tr>
<td>KVM</td>
<td>KVM</td>
<td>HYPERVISOR</td>
<td></td>
</tr>
</tbody>
</table>
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!
my own Ubuntu

Spawn

Freeze
Clone

Ubuntu + user data

Snapshot
golden Debian
Archipelago

RADOS

Monitor nodes

Object Storage nodes

block I/O

object I/O

Volume Composer

Mapper and Blocker
Volumes: Archipelago

Ceph Day London
vkoukis@grnet.gr

Volumes: Archipelago

SYNNEFO

CLUSTER

NODE

ARCHIPELAGO MODE
ARCHIPELAGO API
ARCHIPELAGO CORE
ARCHIPELAGO LIBRARY
ARCHIPELAGO DRIVER

GASTE/ NODE
GASTE/ NODE
ARCHIPELAGO EXISTENCE PROVIDER
ARCHIPELAGO CORE
ARCHIPELAGO DRIVER

GASTE/ NODE
GASTE/ NODE
ARCHIPELAGO EXISTENCE PROVIDER
ARCHIPELAGO CORE
ARCHIPELAGO DRIVER

Cloud

SIDER

ARCHIPELAGO MODE
ARCHIPELAGO API
ARCHIPELAGO CORE
ARCHIPELAGO LIBRARY
ARCHIPELAGO DRIVER

GASTE/ NODE
GASTE/ NODE
ARCHIPELAGO EXISTENCE PROVIDER
ARCHIPELAGO CORE
ARCHIPELAGO DRIVER

GASTE/ NODE
GASTE/ NODE
ARCHIPELAGO EXISTENCE PROVIDER
ARCHIPELAGO CORE
ARCHIPELAGO DRIVER

GASTE/ NODE
GASTE/ NODE
ARCHIPELAGO EXISTENCE PROVIDER
ARCHIPELAGO CORE
ARCHIPELAGO DRIVER
Volumes: Archipelago

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

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

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
Try it out!

http://www.synnefo.org
Thank you!