

A large version of the okeanos logo, with the stylized wave icon and the word "okeanos" in a large, blue, lowercase sans-serif font.

Delivering IaaS for the Greek
Academic and Research Community



Vangelis Koukis
vkoukis@grnet.gr
Technical Coordinator, ~okeanos Project

Outline

- ◆ ~oceanos ?
- ◆ Rationale
- ◆ Design
- ◆ Platform
- ◆ Features
- ◆ Internals
- ◆ Opensource
- ◆ Upcoming



What is ~oceanos?



What is ~oceanos?

'oceanos' is Greek for 'ocean'.



What is ~oceanos?

'oceanos' is Greek for 'ocean'.

Oceans capture, store and deliver energy, oxygen and life around the planet.



Simplicity







Compute



Network



Storage



Security



Virtual Machines



Virtual Ethernets



Virtual Disks

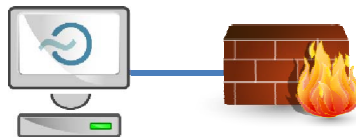


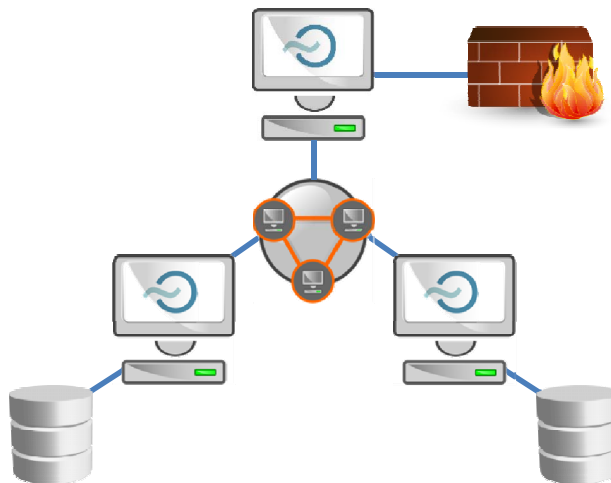
Virtual Firewalls

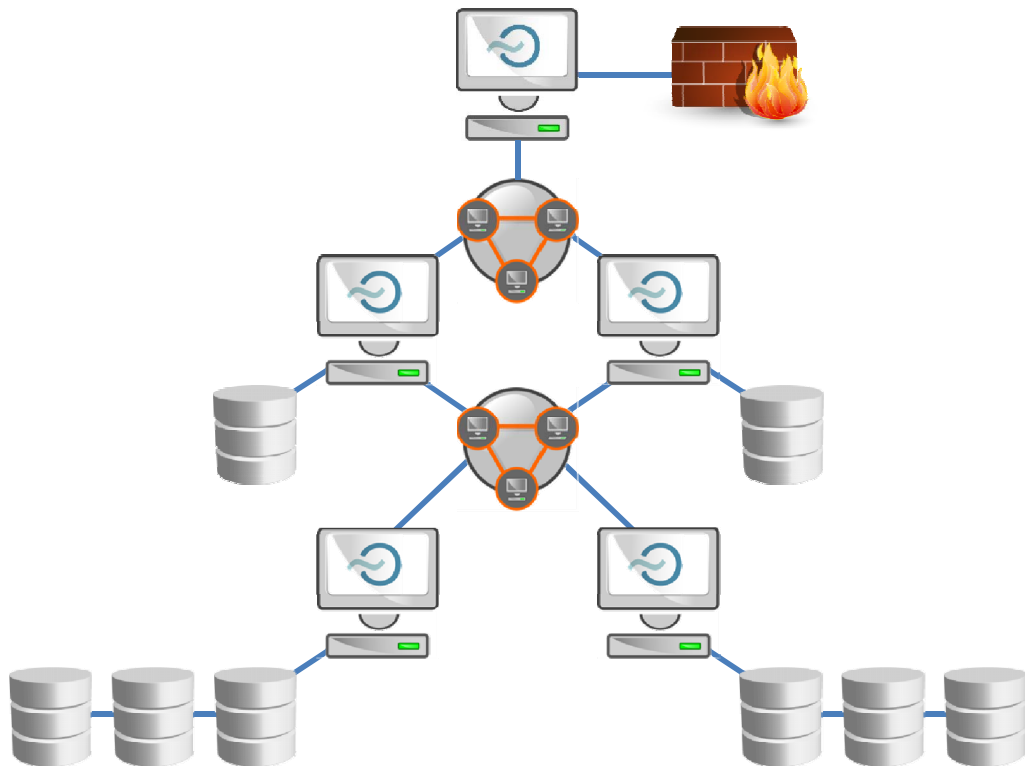
Flexibility







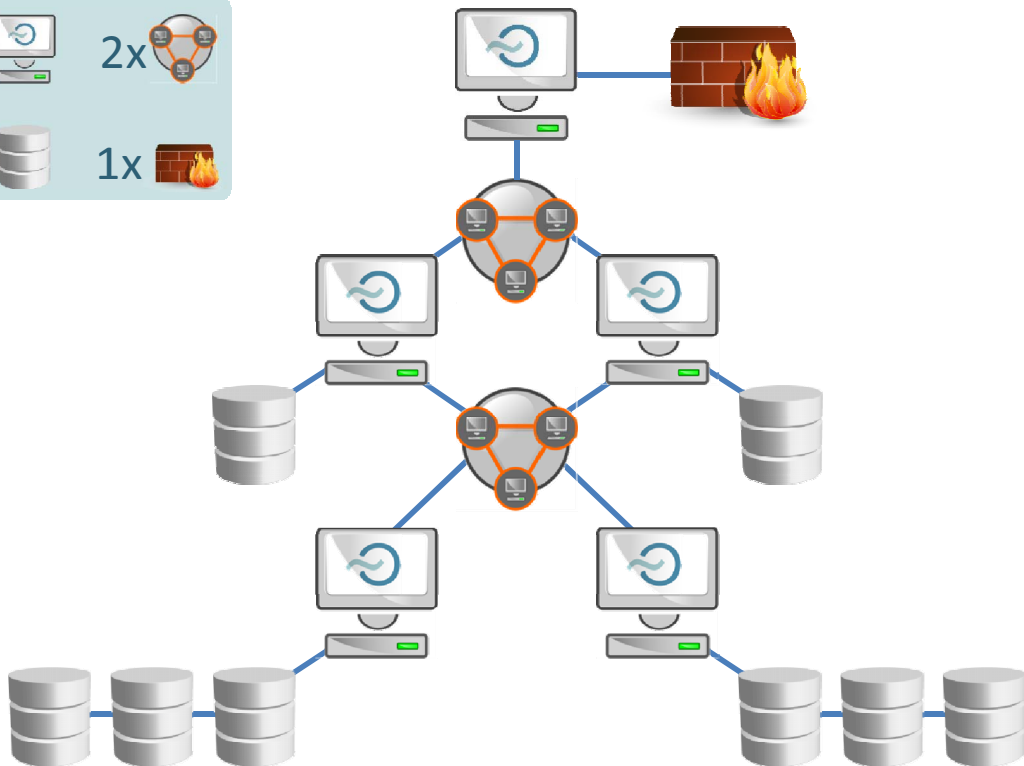








5x		2x	
8x		1x	



~okeanos service

- ◆ **Goal:** Production-quality IaaS
- ◆ **Working Alpha:** from July 2011 / 350 VMs / 200 alpha users
- ◆ **Target group:** GRNET's customers
 - ➔ direct: IT depts of connected institutions
 - ➔ indirect: university students, researchers in academia
- ◆ **Users manage resources over**
 - ➔ a simple, elegant UI, or
 - ➔ a REST API, for full programmatic control



~oceanos service

- ◆ Goal: Production quality IaaS
 - working Alpha from July 2011 / 350 VMs / 200 alpha users
- ◆ A jigsaw puzzle of many pieces
 - UI, API, Image registry, VM management, Networking, Storage, Monitoring, identity management, crediting, issue handling, helpdesk
- ◆ Synnefo
 - custom cloud management software to power ~oceanos
 - Google Ganeti backend



Rationale

How it all started



How it all started

- ◆ Need for easy, secure access to GRNET's datacenters
 - User friendliness, simplicity
- ◆ Scalable to the thousands
 - #VMs, TBs, users (Pithos: ~10k)
- ◆ running within GRNET's AAI Federation
- ◆ Resell or build your own?



Build on commercial IaaS?

◆ Commercial IaaS

- ➔ Amazon EC2 not an end-user service
- ➔ Need to develop custom UI, AAI layers
- ➔ Vendor lock-in
- ➔ Unsuitable for IT depts
 - persistent, long-term servers, custom networking requirements

◆ Gain know-how, build on own IaaS → reuse for own services



What about opensource?

- ◆ Eucalyptus, OpenNebula, OpenStack
- ◆ Need a mature opensource core to *build* around
- ◆ Maturity, production-readiness?
 - proven in production environments, predictable
- ◆ Extensibility?
- ◆ Flexibility?
- ◆ Upgradeability, maintainability?

Design

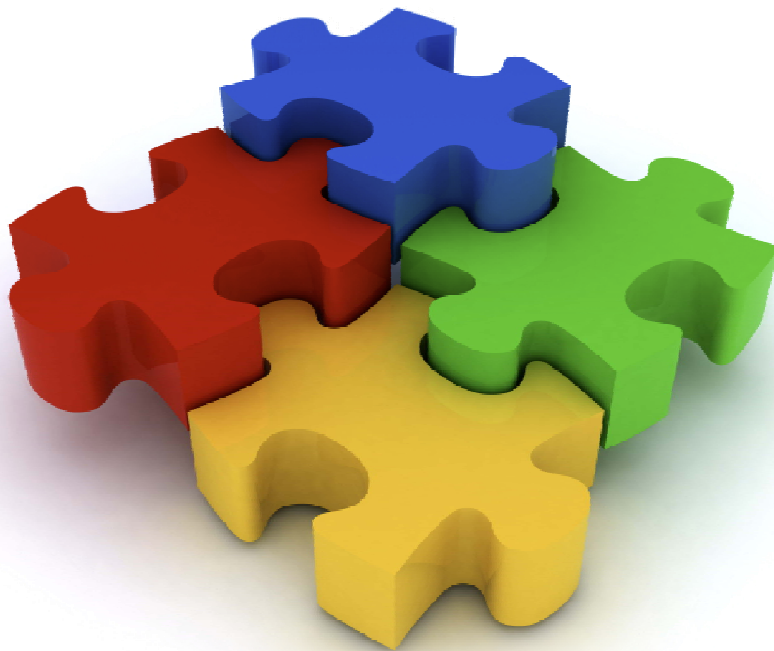
~okeanos design decisions

- ◆ Reuse existing components
- ◆ Build on Google Ganeti
- ◆ target commodity hardware
- ◆ release to the community as opensource



~okeanos design principles

- ◆ No need to make the world
- ◆ No need to support *everything*
 - ➔ Service developed and maintained by ~10-15 people
- ◆ Start from the architecture...
 - ➔ ...then discover, combine, reuse the right components
- ◆ And for everything that's not already available
 - ➔ Do it yourself!



Jigsaw puzzle

- ◆ Synnefo
 - custom cloud management software to power ~oceanos
- ◆ Google Ganeti backend
 - VM cluster management: physical nodes, VMs, migrations
- ◆ OpenStack Compute API v1.1
 - with custom extensions whenever necessary
- ◆ Then everything comes together
 - UI, Networking, Images, Storage, Monitoring, Identity management, Accounting, Billing, Clients, Helpdesk

Why Ganeti?

- ◆ No need to reinvent the wheel
- ◆ Scalable, proven software infrastructure
 - ➔ Built with reliability and redundancy in mind
 - ➔ Combines open components (KVM, LVM, DRBD)
 - ➔ Well-maintained, readable code
- ◆ VM cluster management in production is serious business
 - ➔ reliable VM control, VM migrations, resource allocation
 - ➔ handling node downtime, software upgrades



Why Ganeti?

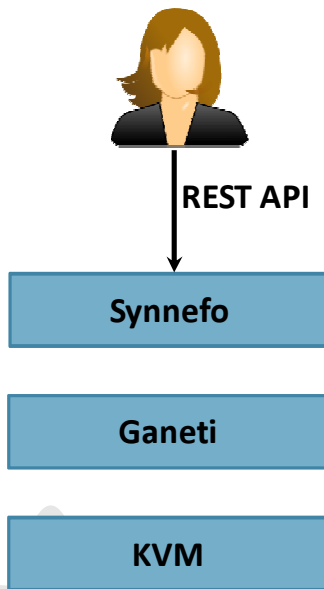
- ◆ GRNET already had long experience with Ganeti
 - ➔ provides ~280 VMs to NOCs through the ViMa service
 - ➔ involved in development, contributing patches upstream

- ◆ Build on existing know-how for ~oceanos
 - ➔ Common backend, common fixes
 - ➔ reuse of experience and operational procedures
 - ➔ simplified, less error-prone deployment

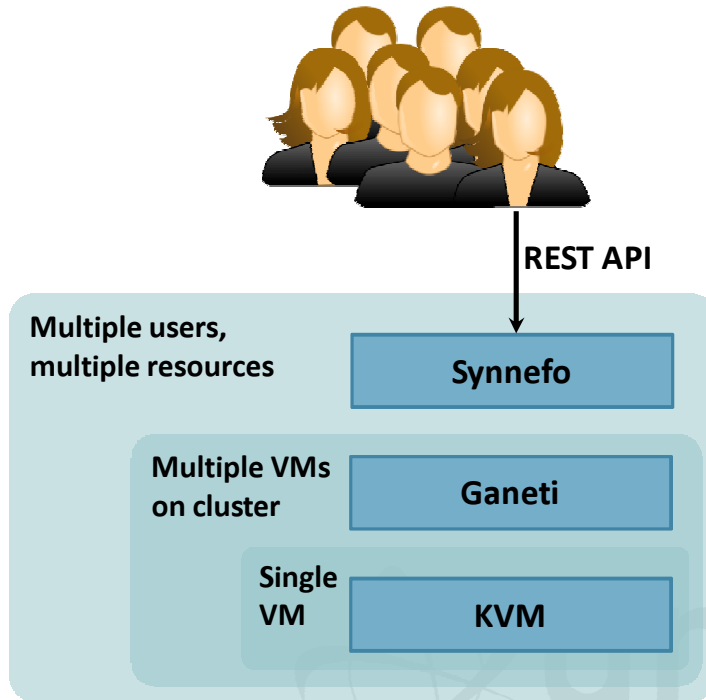


Platform

Software Stack



Software Stack



Platform Design

user@home

admin@home

GRNET
datacenter



Virtual
Hardware



Platform Design

user@home

Web Client

CLI Client

Web Client 2

admin@home

GRNET
datacenter



Synnefo cloud management software

Google Ganeti

KVM

Debian

Virtual
Hardware



Platform Design

user@home

Web Client

CLI Client

Web Client 2

admin@home

GRNET
datacenter



Synnefo cloud management software

Google Ganeti

KVM

Debian

Virtual
Hardware



Platform Design

user@home

Web Client

CLI Client

Web Client 2

admin@home

 GRNET
datacenter


Synnefo cloud management software

Google Ganeti

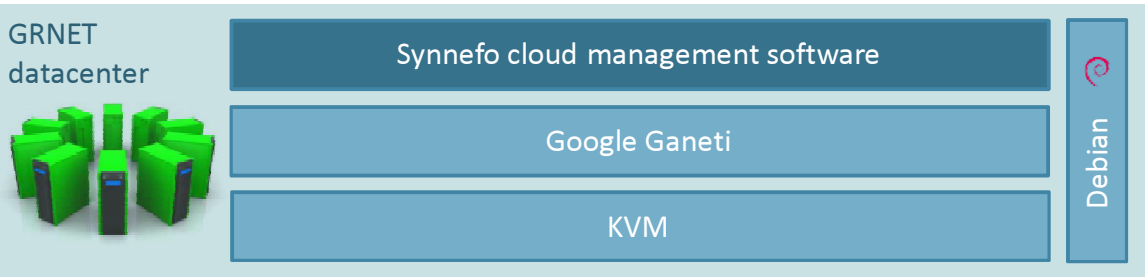
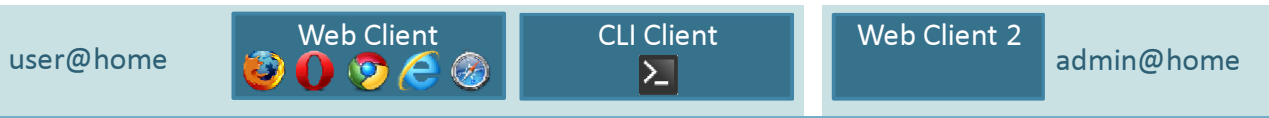
KVM

Debian

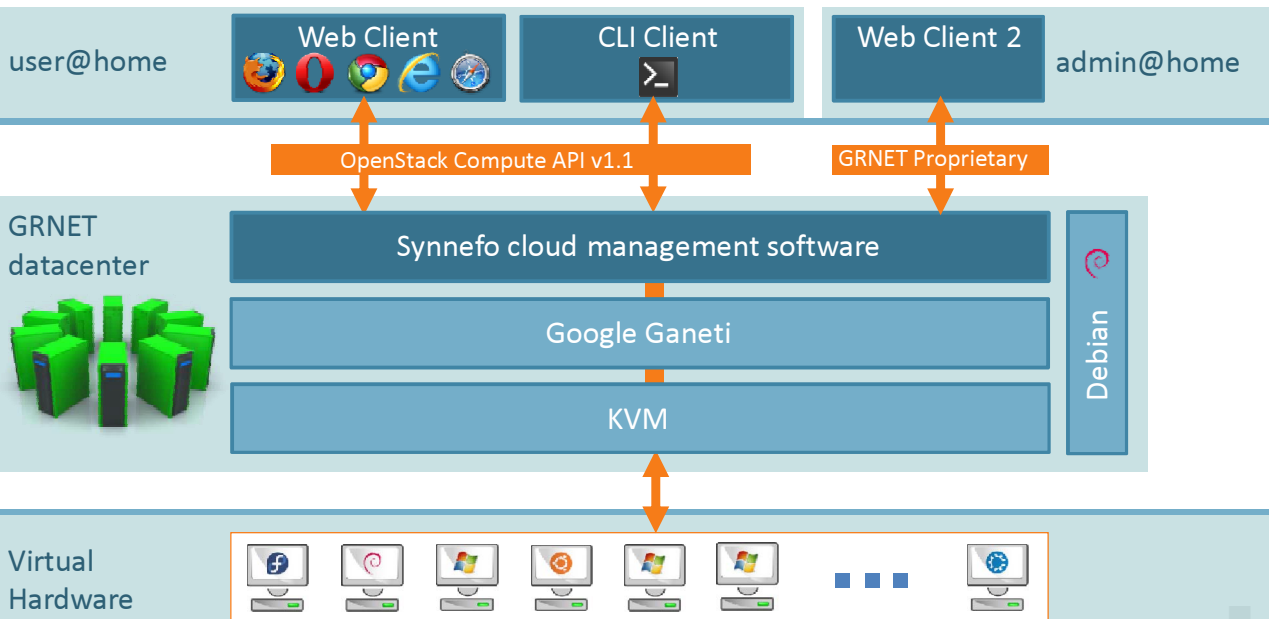
 Virtual
Hardware



Platform Design



Platform Design



Features

Virtual Machine Actions



My_Windows_desktop

Virtual Machine Actions



My_Windows_desktop



Start



Reboot



Shutdown

Virtual Machine Actions



My_Windows_desktop



Start



Console



Reboot



Shutdown



Destroy



IaaS – Compute (1)

◆ Virtual Machines

- ➔ powered by KVM
 - Linux and Windows guests, on Debian hosts
- ➔ Google Ganeti for VM cluster management
- ➔ accessible by the end-user over the Web or programmatically (OpenStack Compute v1.1)

IaaS – Compute (2)

◆ User has full control over own VMs

➔ Create

- Select # CPUs, RAM, System Disk
- OS selection from pre-defined Images
- popular Linux distros (Fedora, Debian, Ubuntu)
- Windows Server 2008 R2

➔ Start, Shutdown, Reboot, Destroy

➔ Out-of-Band console over VNC for troubleshooting



IaaS – Compute (3)

- ◆ REST API for VM management
 - ➔ OpenStack Compute v1.1 compatible
 - ➔ 3rd party tools and client libraries
 - ➔ custom extensions for yet-unsupported functionality
 - ➔ Python & Django implementation
- ◆ Full-featured UI in JS/jQuery
 - ➔ UI is just another API client
 - ➔ All UI operations happen over the API

IaaS – Network (Virtual Ethernets)



Internet



Private Network 1

IaaS – Network (Virtual Ethernets)



Internet



Private Network 1

IaaS – Network (Virtual Ethernets)



Internet



Private Network 1



IaaS – Network (Virtual Ethernets)



Internet



Private Network 1



IaaS – Network (Virtual Ethernets)



Internet



Private Network 1



Private Network 2



Private Network 3

IaaS – Network (Virtual Ethernets)



Internet



Private Network 1



Private Network 2



Private Network 3



IaaS – Network - Functionality

- ◆ Dual IPv4/IPv6 connectivity for each VM
- ◆ Easy, platform-provided firewalling
 - ➔ Array of pre-configured firewall profiles
 - ➔ Or roll-your-own firewall inside VM
- ◆ Multiple private, virtual L2 networks
- ◆ Construct arbitrary network topologies
 - ➔ e.g., deploy VMs in multi-tier configurations
- ◆ Exported all the way to the API and the UI

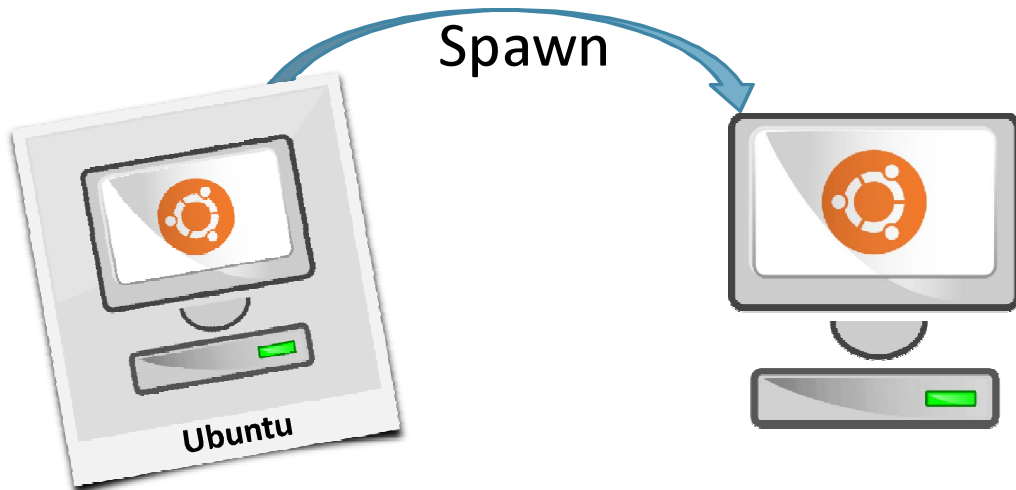


Unity

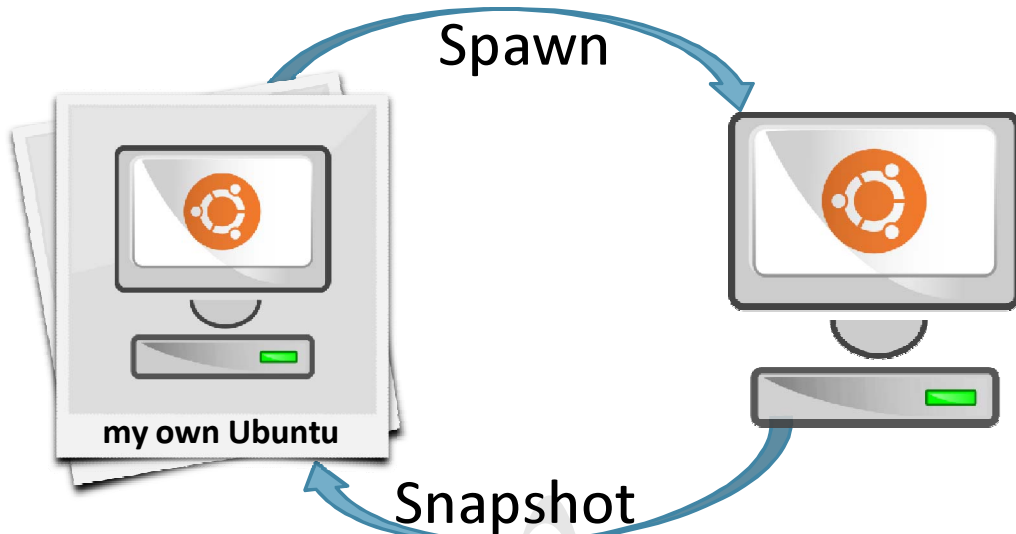
Images



Images



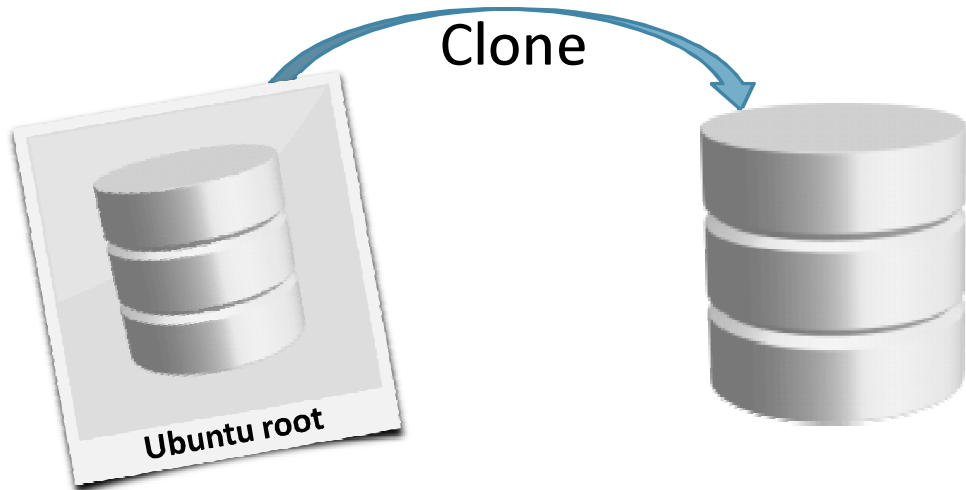
Images



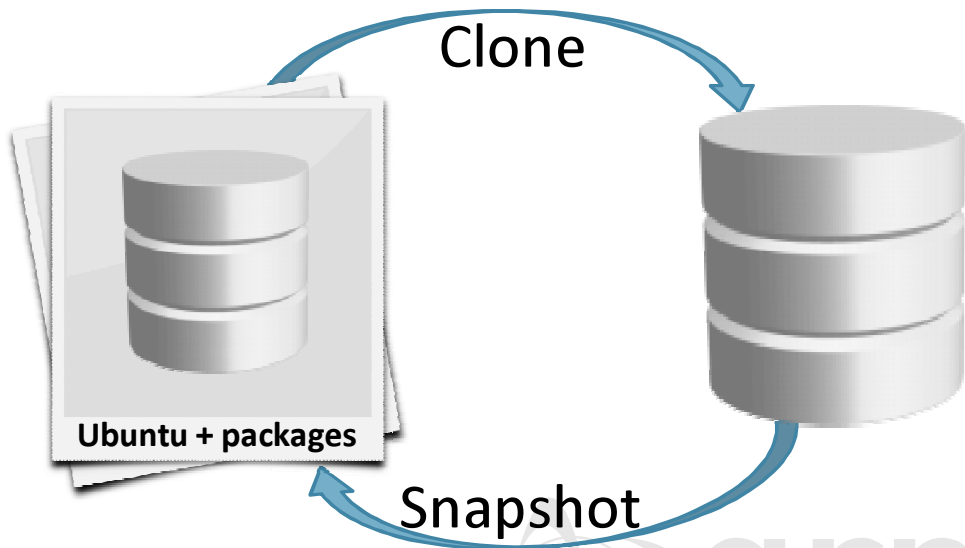
Images ↔ Storage



Images ↔ Storage



Images ↔ Storage



Images – Golden Image



Images – Golden Image



IaaS – Storage



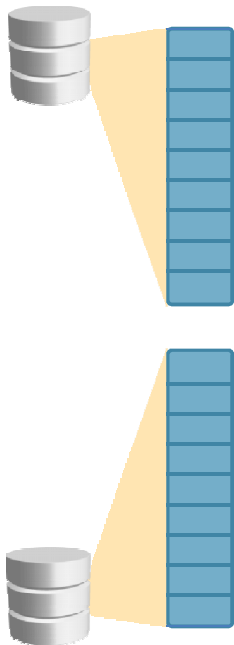
IaaS – Storage



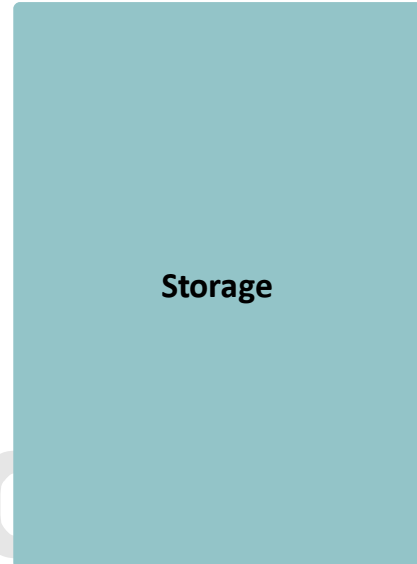
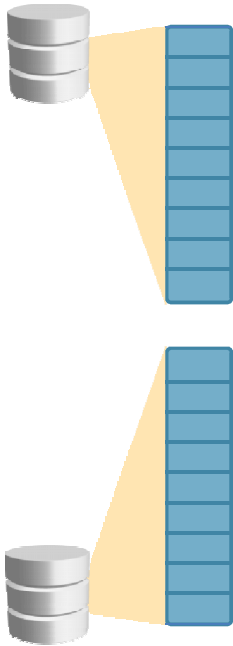
IaaS – Storage



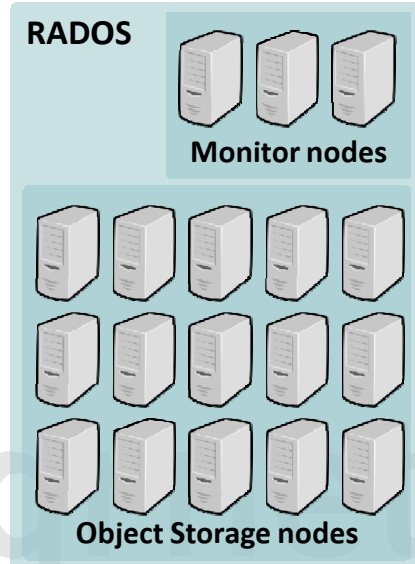
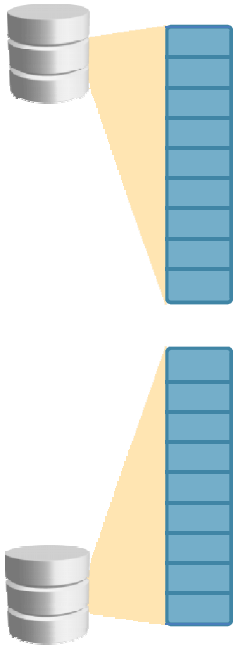
IaaS – Storage



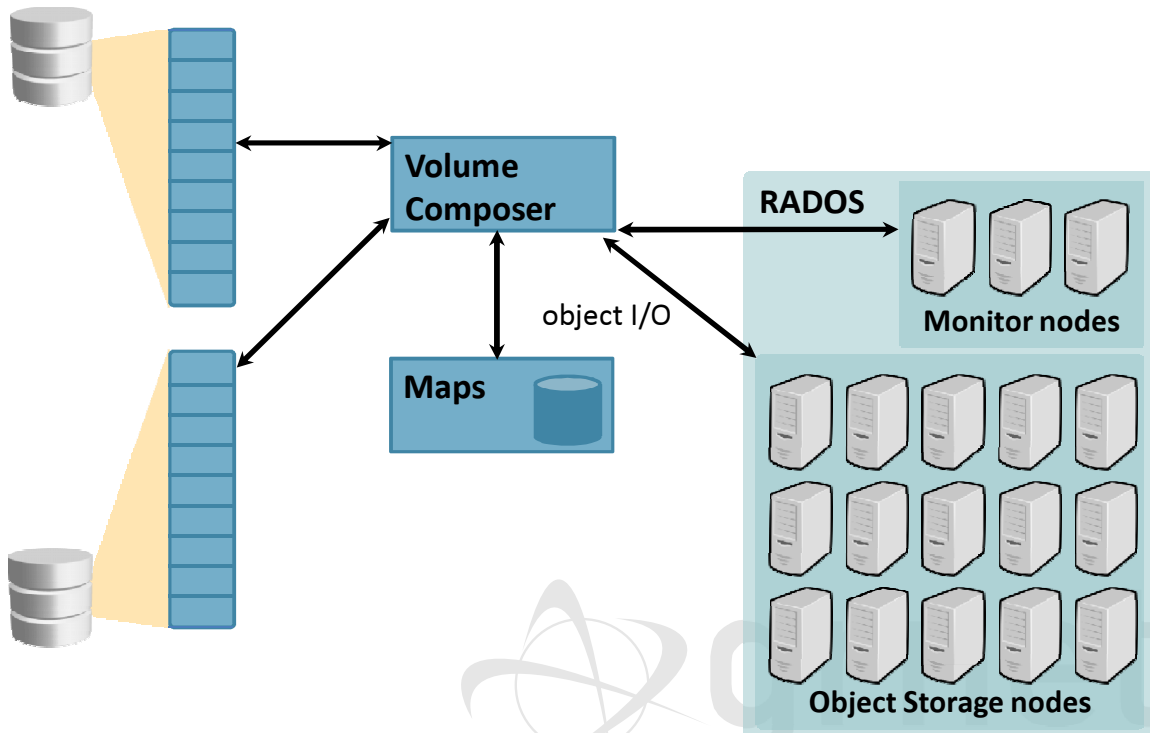
IaaS – Storage



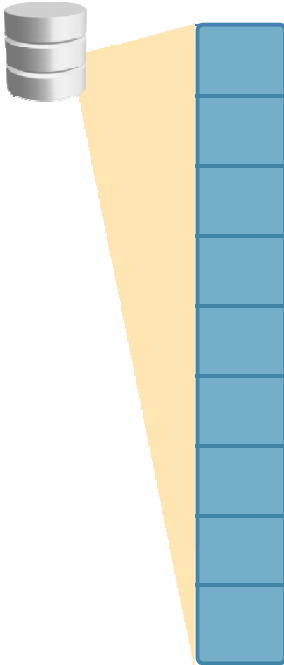
IaaS – Storage



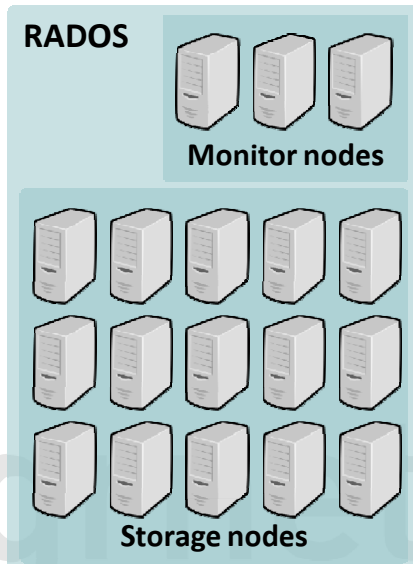
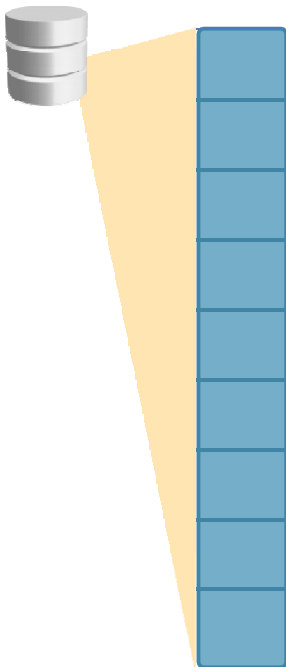
IaaS – Storage



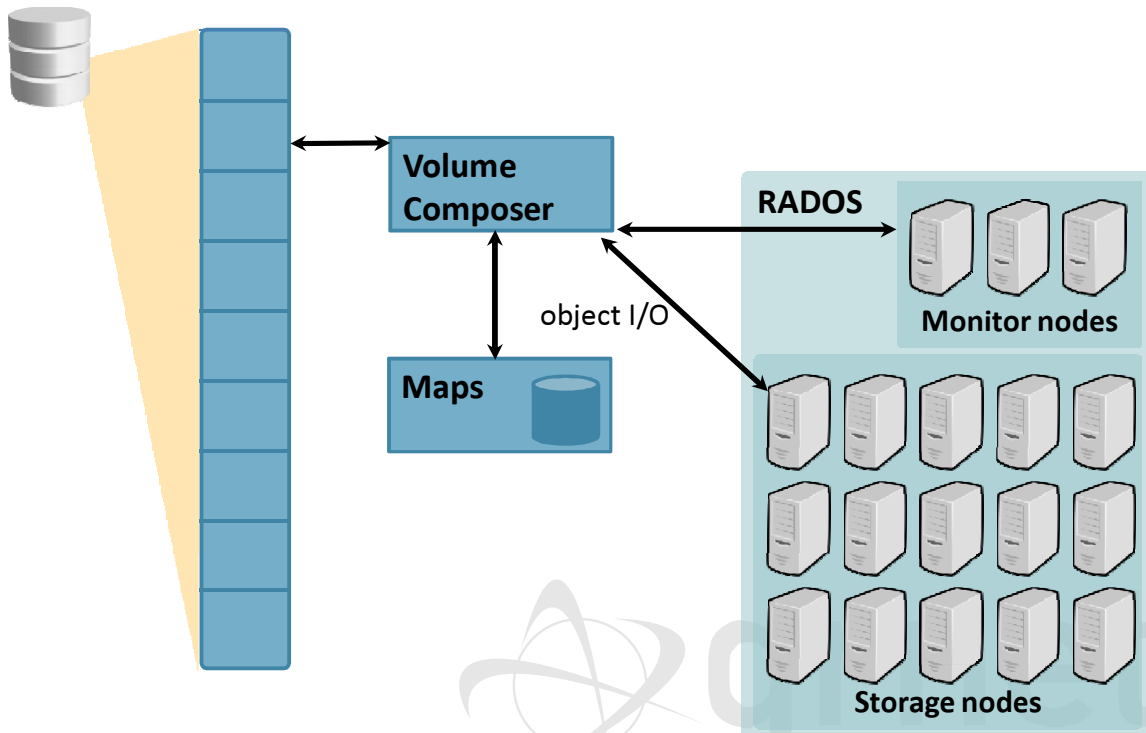
IaaS – Storage



IaaS – Storage



IaaS – Storage



IaaS – Storage (1)

◆ First-phase deployment

- ➔ Pre-defined images of common OSs
- ➔ Redundant storage based on DRBD
- ➔ VMs survive node downtime or failure

◆ Currently under testing

- ➔ Reliable distributed storage over RADOS
- ➔ Combined with custom software for snapshotting, cloning
- ➔ Dynamic virtual storage volumes

IaaS – Storage (2)

- ◆ Multi-tier storage architecture
 - ➔ Dedicated Storage Nodes (SSD, SAS, and SATA storage)
 - ➔ OSDs for RADOS
- ◆ Custom storage layer
 - ➔ manages snapshots, creates clones over RADOS
 - ➔ OS Images held as snapshots
- ◆ VMs created as clones of snapshots

Custom Images: snf-image

◆ *Untrusted* images

- ➔ Host cannot touch user-provided data
- ➔ Resize fs, change hostname, change passwords, inject files

◆ Split design

- ➔ snf-image-host
- ➔ snf-image-helper

◆ All customization in helper VM

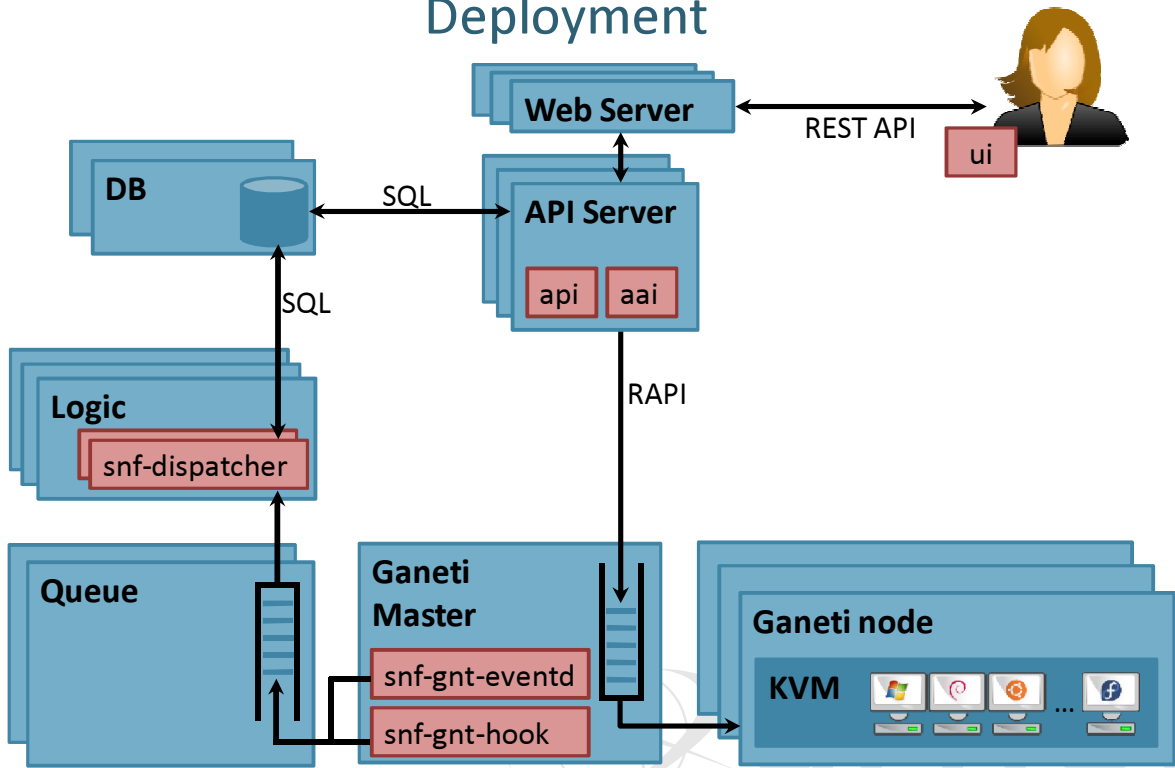


The logo for Pithos+ features a stylized icon of stacked horizontal bars on the left, followed by the text "pithos+" in a lowercase, sans-serif font.

- ◆ OpenStack Object Storage API
- ◆ Block storage
- ◆ Content-based addressing for blocks
- ◆ Every file is a collection of blocks
- ◆ Web-based, command-line, and native clients
- ◆ Synchronization, deduplication
- ◆ Goal
 - ➔ Image registry on Pithos+, common RADOS-based backend

Internals

Deployment



Context

./kamaki

```
$ ipython
```

```
In [1]: from kamaki.client import Client
In [2]: c = Client('http://localhost:8000/api/v1.1', "1234527db2...")
In [3]: c.list_flavors()
...
In [4]: i = c.list_images()
In [5]: i[5]
{u'created': u'2011-06-09T00:00:00+00:00',
 u'id': 7,
 u'metadata': {u'values': {u'OS': u'windows',
                           u'size': u'11000'}}},
 u'name': u'Windows',
 u'progress': 100,
 u'status': u'ACTIVE',
 u'updated': u'2011-09-12T14:47:12+00:00'}
In [6]: c.create_server('mywin1', 3, 5)
```

./kamaki

```
$ ./kamaki
```

```
Usage: kamaki <group> <command> [options]
```

```
...
```

```
--api=API      API can be either openstack or synnefo  
--url=URL      API URL  
--token=TOKEN  use token TOKEN
```

```
...
```

```
Commands:
```

```
flavor info      get flavor details  
flavor list      list flavors
```

```
...
```

```
image create     create image  
image delete     delete image
```

```
$ ./kamaki server shutdown 101 --url=http://localhost:8000/api/v1.1  
--token=1234527db2...
```



More interaction

◆ Identity Management

- ➔ Provides the user base for ~oceanos
- ➔ Once authenticated, the user retrieves an auth token for programmatic access

◆ Pithos+ storage service

- ➔ Aim is to provide the Image service for ~oceanos
- ➔ Sharing a common storage backend on RADOS

◆ Aquarium

- ➔ Common crediting and billing infrastructure



Upcoming

Upcoming goals

- ◆ short-term: Synnefo v0.8
 - ➔ user-visible storage backends
 - ➔ VMs with RADOS-based storage
 - ➔ Custom Image registry, custom Image deployment
- ◆ medium-term
 - ➔ Dynamic disks, clonable / snapshottable / attachable disks
 - ➔ integration with Pithos+, single user base
- ◆ Beta when datacenter is fully populated

Upcoming goals

- ◆ Credit-based resource allocation
- ◆ Abstract away the Ganeti backend, replace with backend connector behind the MQ
 - ➔ Release to community as reference implementation of OpenStack Compute v1.1
- ◆ Support live modification of VMs in Ganeti
- ◆ Snapshots, clones in storage layer
 - ➔ Dramatic decrease in VM initialization time
 - ➔ Support workloads with 100s of ephemeral VMs
 - e.g. for scientific computation, MPI jobs

Sights

[New Machine +](#)

Welcome to ~okeanos !

From this panel you will be able to manage your Virtual Machines (VMs).

The panel is currently empty, because you don't have any VMs yet. Start by clicking the orange button on the top left. The wizard will guide you through the whole process.

For more information or help, click [here](#).

close

Create new machine

1

Image

Select an OS
Choose your preferred image

2

3

4

Image type

System

Custom

Categories

no categories available

Available Images

Debian Base	473.00 MB
Debian Squeeze Base System	
Debian Desktop	3.18 GB
Debian Squeeze Desktop	
Ubuntu	2.23 GB
Ubuntu 11.04	
Kubuntu	2.22 GB
Kubuntu 11.04	
Fedora Desktop	2.40 GB
Fedora 15 Desktop Edition	
CentOS	990.00 MB
CentOS 5.6	
Windows	10.74 GB
Windows 2008 R2, Aero Desktop Ex...	

Ubuntu

Description
Ubuntu 11.04

OS
Ubuntu

Size
2.23 GB

GUI
Unity/GNOME 2.32.1

Kernel
2.6.38

cancel

next

Create new machine

close

1

2

Flavor

Select CPUs, RAM and Disk Size

Available options are filtered based on the selected image

3

4

Predefined

Medium

Small

Large

CPUs

Choose number of CPU cores

1 x

2 x

4 x

Memory size

Choose memory size

1024 MB

2048 MB

4096 MB

Disk size

Choose disk size

20 GB

30 GB

40 GB

Storage

Select storage type

DRBD

DRBD storage.

previous

next

Create new machine

close

1

2

3


Personalize

Virtual machine custom options

Virtual machine custom options

4

Machine name

 My Ubuntu server

Public SSH keys

[manage keys](#)

Select ssh keys

No ssh keys in your account. [Create/import a new key now.](#)

Suggested tags

You may change machine tags later from the machines view.

Role

Database server

File server

Mail server

Web server

Proxy

previous

next

Create new machine

close

1 2 3 **4 Confirm**

Confirm your settings

Confirm that the options you have selected are correct

Machine name

 **My Ubuntu server**

Image

Ubuntu

Ubuntu 11.04

OS **Ubuntu**

Size **2.23 GB**

GUI **Unity/GNOME
2.32.1**

Kernel **2.6.38**

Flavor

CPUs **4x**

Memory **4096 MB**

Disk **20.00 GB**

Storage
type **DRBD**

Machine Tags

Role **Web server**

SSH Keys

public key

previous

create machine

Machine password

close

Your new machine is now buidling... (this might take a few minutes)

Write down your password now:

UqPt4Sqk



You will need this later to connect to your machine.

After closing this window you will *NOT* be able to retrieve it again

[view machine](#)



My Debian Base server

Initializing...

info

Building...





machines

New Machine +



icon



list



single



My Ubuntu server

Initializing...

info

Building...



CPU: 4
RAM: 4096MB
System Disk: 20GB

Image: Ubuntu
Image Size: 2.23 GB

CPU



Net



[Full report](#)

OS : ubuntu

[Manage Tags](#)



machines

New Machine +



icon



list



single


My Debian Base server

IPv4 62.217.112.41 IPv6 ...a80c:eaff:fe2a:d59a

info

Running



Reboot

Shutdown

Console

Destroy

Confirm


My Debian Base server 1

IPv4 62.217.112.220 IPv6 ...a80c:eaff:fe6a:a62e

info

Running





machines

New Machine +



icon



list



single



My Debian Base server

IPv4 62.217.112.41 IPv6 ...a80c:eaff:fe2a:d59a

info

Running



Reboot

Shutdown

Console

Destroy

Confirm



My Debian Base server 1

IPv4 62.217.112.220 IPv6 ...a80c:eaff:fe6a:a62e

info

Running



Reboot

Shutdown

Console

Destroy

Confirm



Your actions will affect 2 machines

Cancel All

Confirm All



machines

New Machine +



icon



list



single



My Debian Base server

IPv4 62.217.112.41 IPv6 ...a80c:ea:ff:fe2a:d59a

info

Shutting down...



000



My Debian Base server 1

IPv4 62.217.112.220 IPv6 ...a80c:ea:ff:fe6a:a62e

info

Rebooting...



000



machines

New Machine +



icon



list



single


My Debian Base server 1

Running

IPv4 62.217.112.220 IPv6 ...a80c:eaaff:fe6a:a62e

info


My Debian Base server

Stopped

IPv4 62.217.112.41 IPv6 ...a80c:eaaff:fe2a:d59a

info





machines

New Machine +



icon



list



single

Search:

<input type="checkbox"/>	OS	Name	Flavor	Status	
<input type="checkbox"/>		My Debian Base server 1	4 CPU, 4096MB, 20GB	Running	Start Reboot Shutdown
<input type="checkbox"/>		My Debian Base server	4 CPU, 4096MB, 20GB	Stopped	Destroy



Public IPv6:

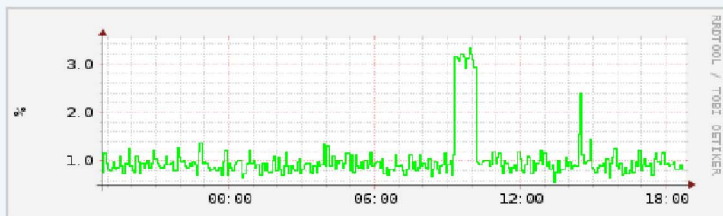
2001:648:2ff:1112:a00c:eaff:fe6a:a62e

tags

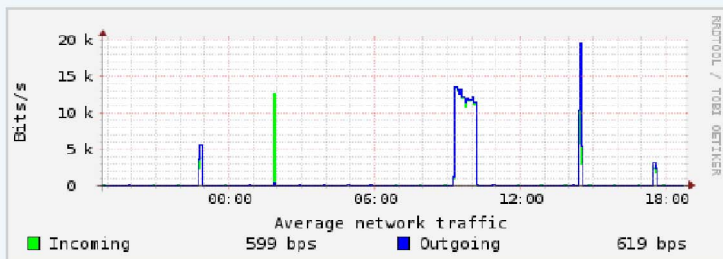
debian web se...

mail server

CPU Utilization



Network Utilization





machines

New Machine +



icon



list



single


mail server

Running

IPv4 62.217.112.220 IPv6 ...a80c:eaff:fe6a:a62e



info

 CPUs: 4
 RAM:4096MB
 System Disk: 20GB

 Image: Debian Base
 Image Size: 501.00 MB

CPU

CPU: 0.0%

Net

TX/RX: 0.00/0.00 Mbps

[Full report](#)

OS : debian

[Manage Tags](#)

debian web server

Stopped

IPv4 62.217.112.41 IPv6 ...a80c:eaff:fe2a:d59a



info

mail server close

Manage tags

Key: Value:

Owner	Role	
<div style="display: flex; align-items: center;"> <div style="background-color: #f44336; color: white; padding: 2px 5px; margin-right: 5px;">OS</div> <div style="border: 1px solid #ccc; padding: 2px;">debian</div> </div>	<div style="display: flex; align-items: center;"> <div style="background-color: #f44336; color: white; padding: 2px 5px; margin-right: 5px;">Software</div> <div style="border: 1px solid #ccc; padding: 2px;">Apache, Post...</div> </div>	<div style="background-color: #8bc34a; color: white; padding: 5px; display: inline-block;"> Add new tag + </div>

CPU: 4
 RAM: 4096MB
 System Disk: 20GB
 Image: Debian Base
 Image Size: 501.00 MB

Net
 Tx/RX: 0.00/0.00 Mbps
[Full report](#)

OS: debian
 Soft... : Apache, Po...
[Manage Tags](#)



debian web server
 IPv4 62.217.112.41 IPv6 ...a80c:eaff:fe2a:d59a
[info](#)

Stopped



New Network +



Internet

machines (2) ▾

Public network





New Network +



Internet

Public network



machines (2)


debian web server


Firewall (Off)

IPv4: 62.217.112.41

IPv6: 2001:648:2ffc:1112:a80c:eaff:fe2a:d59a


mail server


Firewall (Off)

IPv4: 62.217.112.220

IPv6: 2001:648:2ffc:1112:a80c:eaff:fe6a:a62e



New Network +



Internet

Public network



machines (2) ▲



debian web server

IPv4: 62.217.112.41



Firewall (off) ▲

IPv6: 2001:648:2ff:c:1112:a80c:eaaff:fe2a:d59a

- Unprotected mode (Firewall off)
- Fully protected mode (Firewall on)
- Basically protected mode (Firewall on)

Apply



mail server

IPv4: 62.217.112.220



Firewall (Off) ▼

IPv6: 2001:648:2ff:c:1112:a80c:eaaff:fe6a:a62e

Networks close

Create new private network

Network name: create network



Internet

machines (2)



debian web server

Firewall (off)

IPv6: 2001:648:2ffc:1112:a80c:eaff:fe2a:d59a

IPv4: 62.217.112.41

- Unprotected mode (Firewall off)
- Fully protected mode (Firewall on)
- Basically protected mode (Firewall on)

Apply



mail server

Firewall (off)

IPv6: 2001:648:2ffc:1112:a80c:eaff:fe6a:a62e

IPv4: 62.217.112.220

Public network





New Network +



Internet

machines (2) ▾

Public network



my private network

machines (0) ▾

Private network



my private network
close

Connect machine

Select machines to add

debian web server
✖

debian

mail server
+

debian

connect machines



my private network

machines (0) ▼

Private network





New Network +



Internet

machines (2)

Public network



my private network

machines (1)

Private network



Add Machine

Destroy


debian web server

[Connect](#) to manage private IPs




machines

New Machine +



icon



list



single



mail server

IPv4 62.217.112.220 IPv6 ...a80c:caff:fe6a:a62e

info

Rebooting...



000



debian web server

IPv4 62.217.112.41 IPv6 ...a80c:caff:fe2a:d59a

info

Rebooting...



000

SSH keys

Manage your ssh keys

close

SSH public keys list

generate new 

create/import new +

You can use SSH keys to establish a secure connection between your computer and the virtual machines.

rsa public key 1

fingerprint: 29:8f:92:15:c9:ba:42:2d:9d:b4:88:3e:2d:7f:e1:c6

rsa public key

hide key -

edit 

remove 

fingerprint: 40:c7:5f:6d:ec:64:0f:60:e4:8d:d0:06:c3:f1:4b:26

```
ssh-rsa
AAAAB3NzaC1ye2EAAAADAQABAAQDRKDLtabp+MFe0kwhWJ1AYjqNylDnq+EAdCRk
Cay5McycleRo6p67AF5rXqNx21itLYJMYeOed8ft1b4olu+XtRWBRSDx2SBYTKFScorC
xgs9QgMBbRAAWZ3WPmLd8ZHPaf2pQpJeQ1UlwNtOoJhrQj8xqoRtsmVy5ID5eGFyppR
/+iXXio7XXqIsPSX0vpj/t0KRM9WyMCLERdEL5SOD/nh1/Km5Ao9gnqaO2YgCeRDN7N
5WdUfaMcY2V.TeGA23Ua7zEaP+TnsLFed72khr+gPRhY+HEhr5ZPPwnNdCILVTKY5+d
```

Open

Opensource



Opensource

- ◆ Synnefo
 - <https://code.grnet.gr/projects/synnefo>
- ◆ Pithos+
 - <https://code.grnet.gr/projects/pithos>
- ◆ Ganeti
 - <https://code.google.com/p/ganeti>
- ◆ snf-image
 - <https://code.grnet.gr/projects/snf-image>
- ◆ kamaki
 - <https://code.grnet.gr/projects/kamaki>
- ◆ vncauthproxy
 - <https://code.grnet.gr/projects/vncauthproxy>

Opensource

- ◆ Synnefo
 - <https://code.grnet.gr/projects/synnefo>
- ◆ Pithos+
 - <https://code.grnet.gr/projects/pithos>
- ◆ Ganeti
 - <https://code.google.com/p/ganeti>
- ◆ snf-image
 - <https://code.grnet.gr/projects/snf-image>
- ◆ kamaki
 - <https://code.grnet.gr/projects/kamaki>
- ◆ vncauthproxy
 - <https://code.grnet.gr/projects/vncauthproxy>



The logo for Okeanos features a stylized blue wave icon to the left of the word "okeanos" in a lowercase, blue, sans-serif font.

okeanos

<https://okeanos.grnet.gr>