~okeanos

Unveiling ~okeanos: A public cloud IaaS service coming from the depths of the GRNET's DataCenter facilities

Nectarios Koziris, GRNET

National Technical University of Athens
CSLab

grnet
Networking Research and Education

What is Okeanos?

- ‘Okeanos’ is the Greek word for ‘ocean’

What is Okeanos?

- ‘Okeanos’ is the Greek word for ‘ocean’

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

Outline

- ~okeanos IaaS
- Compute
- Synnefo architecture
- Network
- Storage
- Infrastructure

Motivation

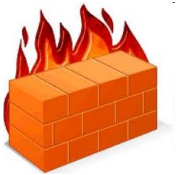
- Deliver IaaS to GRNET's customers
  - direct: IT depts of connected institutions
  - indirect: university students, researchers in academia
- Other IaaS efforts (Amazon, Rackspace, Nimbula, Nebula, GoGrid etc)
  - Amazon EC2 not an end-user service
  - Need to develop custom UI, AAI layers
  - Vendor lock-in
  - Unsuitable for IT depts
    - persistent, long-term servers, custom networking requirements
- Nimbula Director offers similar functionality
  - Gain know-how, build on own IaaS → new services

~okeanos IaaS

Compute

Network

Storage



firewall

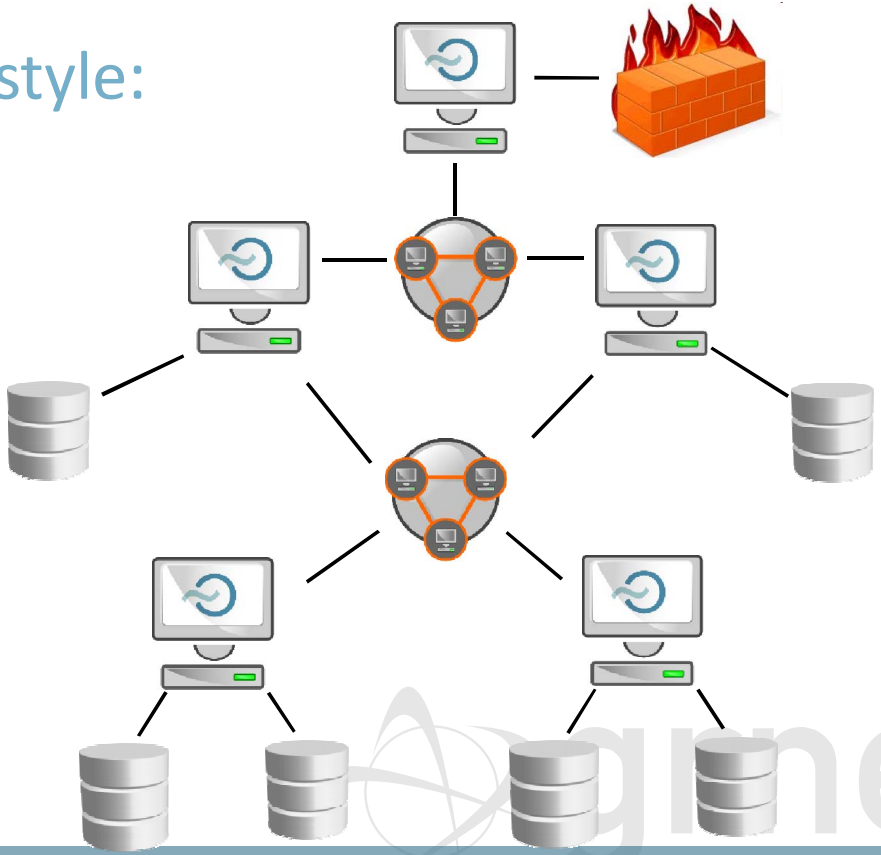
simplicity

LEGO
4+
5576
375
pcs/Stck./pzs/db

flexibility

HOTEL
HOTEL

Lego style:



~okeanos service:

- Goal: Production quality IaaS
  - working Alpha from July 2011/ 350 VMs / 200 alpha users
- A jigsaw puzzle of many pieces
- Synnefo sw
  - custom cloud management software to power ~okeanos
  - Google Ganeti backend

Outline

- ~okeanos IaaS
- Compute
- Network
- Storage
- Infrastructure

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
  - Remote Desktop
  - Out-of-Band console over VNC for troubleshooting

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

Shutdown

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

Why Google 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

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

user@home

Web Client

CLI Client

Web Client 2

admin@home

GRNET

datacenter

synnefo Cloud Management Software

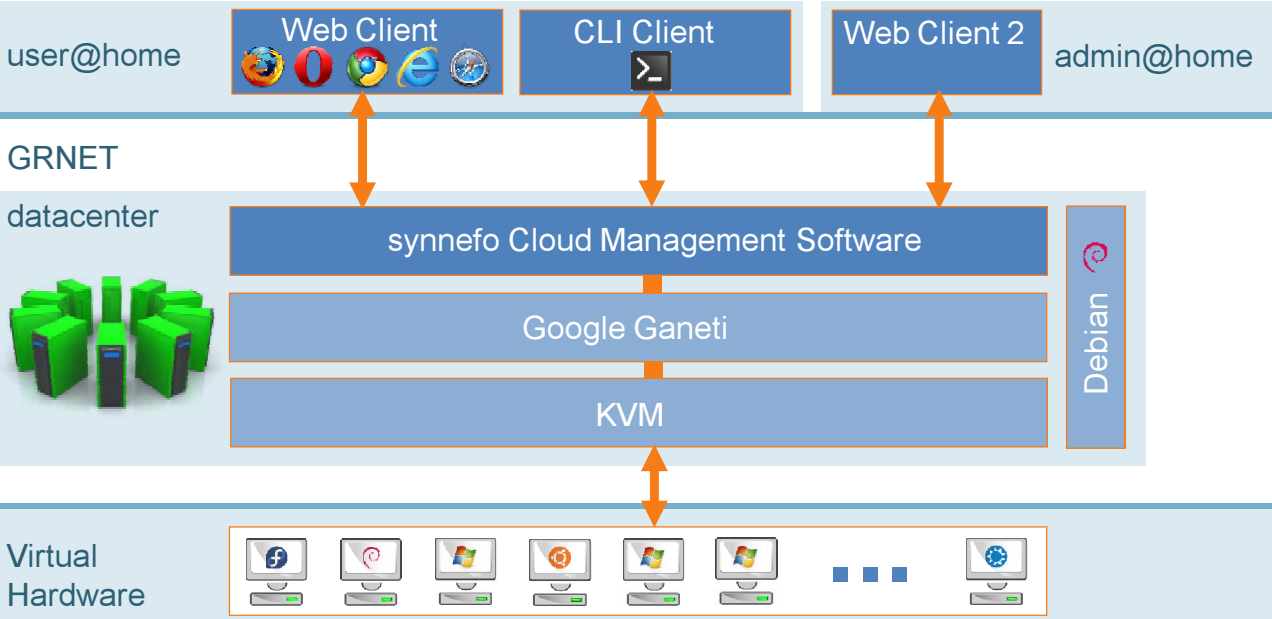
Google Ganeti

KVM

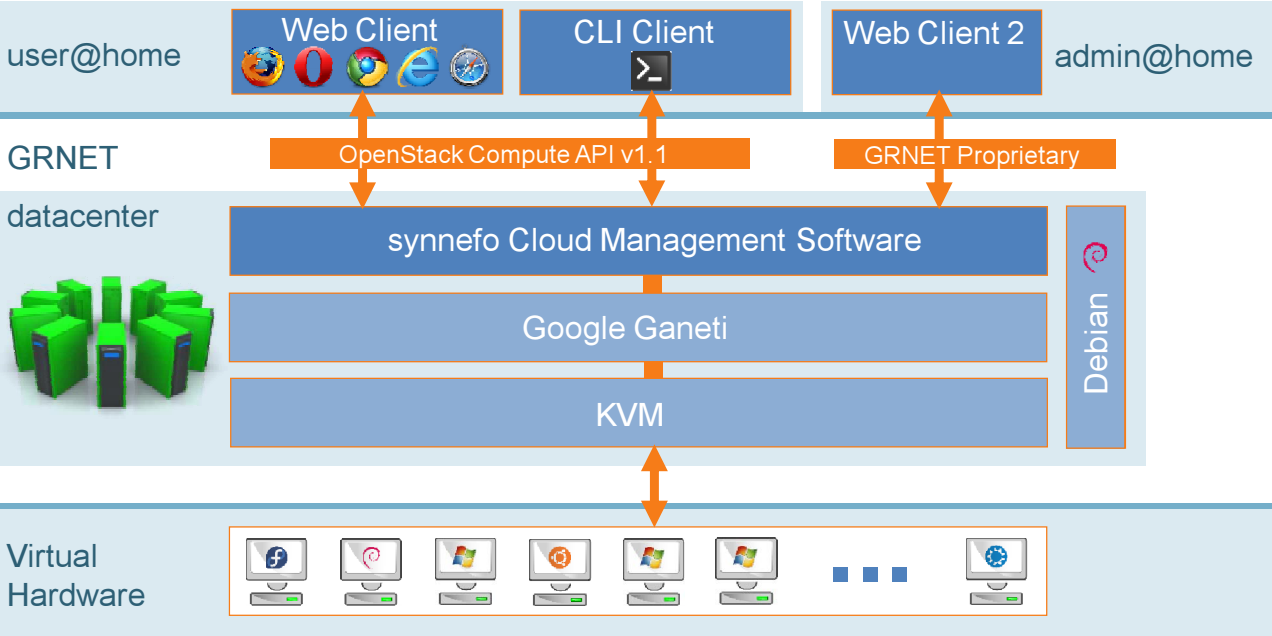
Debian

Virtual Hardware

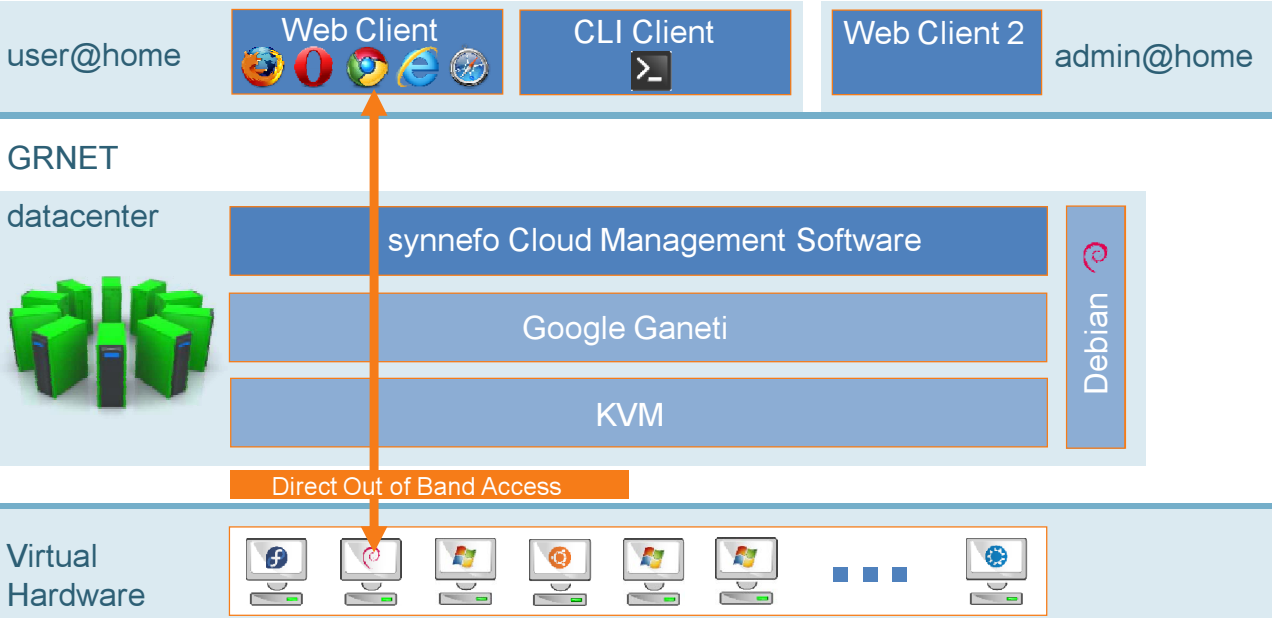
Platform Design



Platform Design



Platform Design



Outline

- ~okeanos IaaS
- Compute
- Network
- Storage
- Infrastructure

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

Virtual Networks

Internet

Private Network 1

Virtual Networks

Internet

Private Network 1

Virtual Networks

Internet

Private Network 1

Virtual Networks

Internet

Private Network 1

Virtual Networks

Internet

Private Network 1

Private Network 2

Virtual Networks

Internet

Private Network 1

Private Network 2

Private Network 3

Virtual Networks

Internet

Private Network 1

Private Network 2

Private Network 3

Outline

- ~okeanos IaaS
- Compute
- Synnefo architecture
- Network
- Storage
- Infrastructure

IaaS – Storage (1)

- First-phase deployment
  - Ability to customize VM contents based on predefined images of common OSs
  - Redundant storage based on DRBD, VMs survive scheduled node downtime
- Soon:
  - Reliable distributed storage over RADOS
  - Combined with custom software for snapshotting, cloning to provide 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

Interaction with other GRNET services

- GRNET AAI Federation
  - Provides the user base for ~okeanos
  - Once authenticated, the user retrieves a Synnefo-specific auth token for programmatic access
- Pithos File service
  - Currently being overhauled
  - Aim is to provide the Image service for ~okeanos
  - Sharing a common storage backend

Outline

- ~okeanos IaaS
- Compute
- Synnefo architecture
- Network
- Storage
- Infrastructure

- main Datacenter (40 racks)
  - green high-density architecture (Lamda Hellix)
  - PUE <1.8 / 1600KVA
  - In-row cooling
  - Fully virtualized
- 'High' Density (but not too dense…)

  ~30.000 VMs (20 racks, 1500 VM/rack)
  - VMs with 1.5GB RAM
  - QoS in disk/core overcommitment
  - Block storage provisioning/elastic volumes per user
  - Persistent Storage
  - 100GB/user – 6K users already in Pithos storage (S3 like online storage service)

πίθος
www.pithos.gr

Design Specs for a Cloud Provider

Common denominator:

#VM instances (in various VM ‘flavors’)

Guaranteed ratios:
VMs/Core, RAM/Core, disks(spindle)/core

€/VM/h?

‘Going Green’ necessity

- Operation Costs (power) are huge!
- Space is always not enough – density against green

We need both green & dense infrastructure solutions...

while ’elegant’ ones in terms of services!

Design Considerations for Public Cloud Service Budgeting

- High density in terms of watts: Maximise Watt/rack?
- Power budget per rack (i.e. 16KW spec)

  maximize # of Cores/1U,TB/1U, RAM/1U, disks/spindles/1U ?

  or

  find the perfect balance for your services (QoS driven)?

  Plain PUE is everything?
  - If you are a facility…maybe yes
  - if you offer managed services …not enough, need more

Conclusion: Consider an 'efficiency' metric

- e.g. If VM is your service:
  - ➡ the '*moment of truth*' green spec is: **#VMs/Watt**.
    - It's a function of all previous!

And of course (@ the end of the day): **€/VM !!**

So: get the 'best' out of every watt consumed

- Add to our budget envelope: Watt per service, may alter design specs

Service driven power/euros budgeting

Cloud Facilities (cont.)

- Disaster Recovery DC (@RFI stage)
  - Container based solution
  - Up to 1MW
  - Low PUE (<1.6)
  - Outdoor installation (close to hydroelectric plant facility)
  - Water cooling per rack / Freecooling

See more on: http://okeanos.grnet.gr

okeanos

test@synnefo.gr | el ~ en

machines | | |

Create New +

Welcome to ~okeanos !

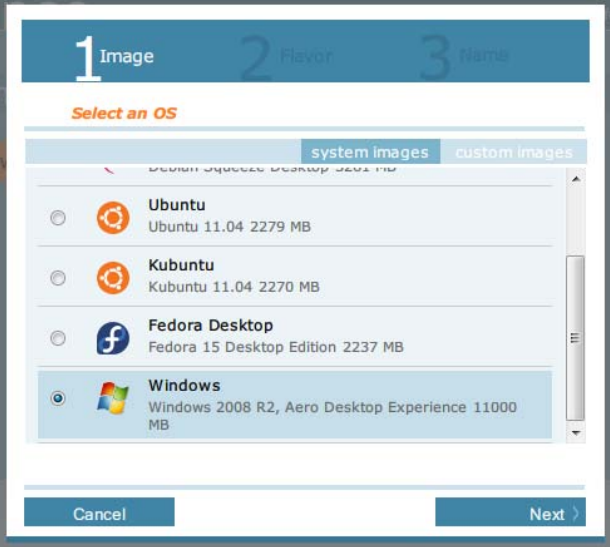
From this panel you will be able to manage your Virtual Machines (VMs). If you don't know what a VM is: take the tour.

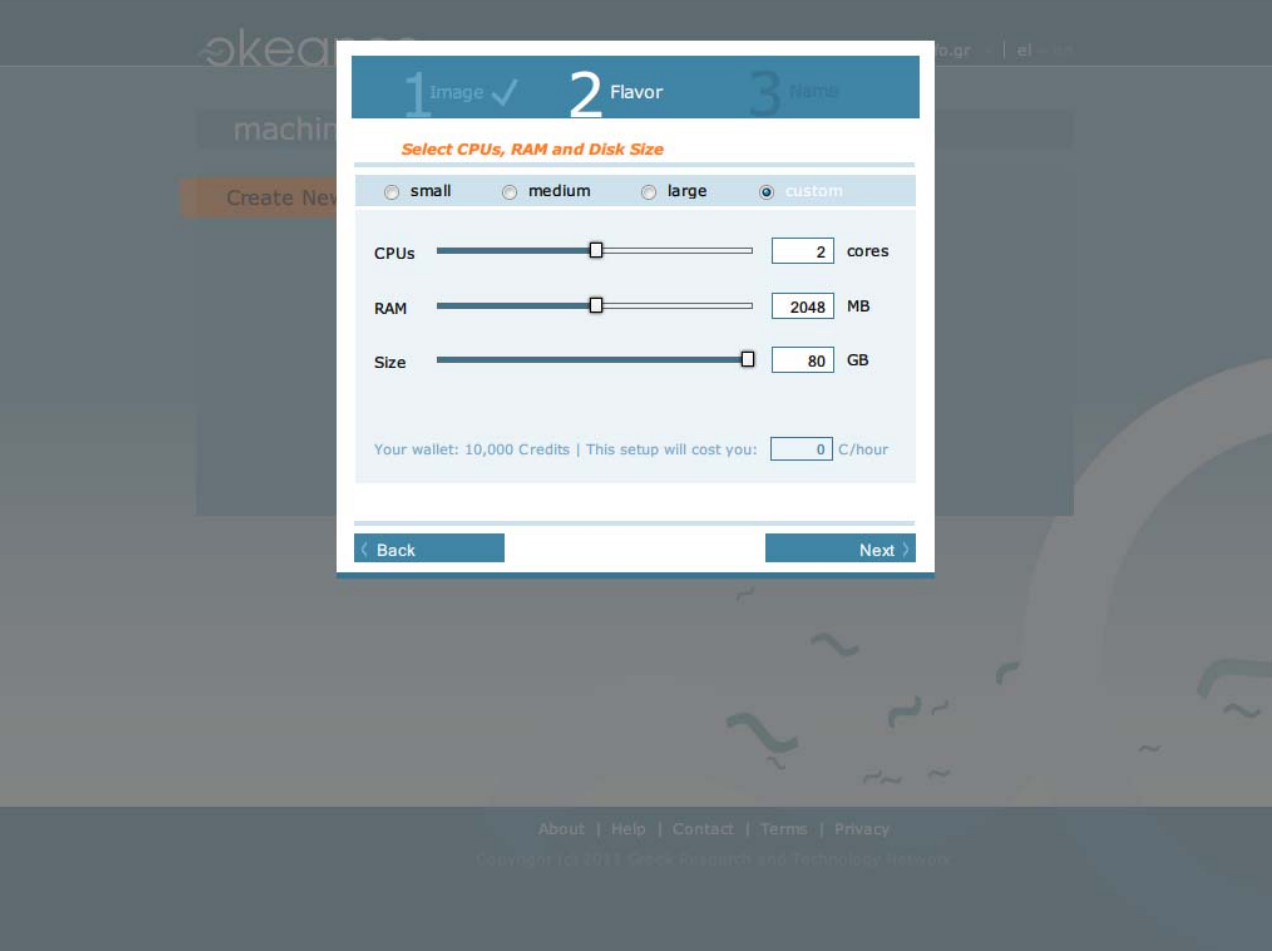
The panel is currently empty, because you don't have any VMs yet. Start by clicking the orange button on the top left. The wizard will guide you through the whole process.

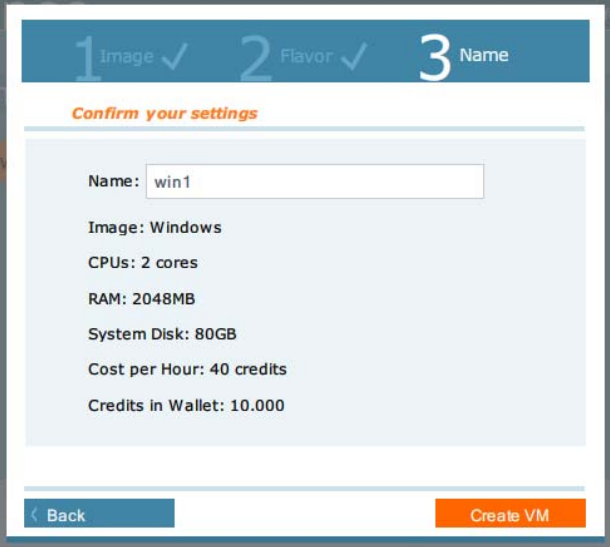
For more information or help, click here.

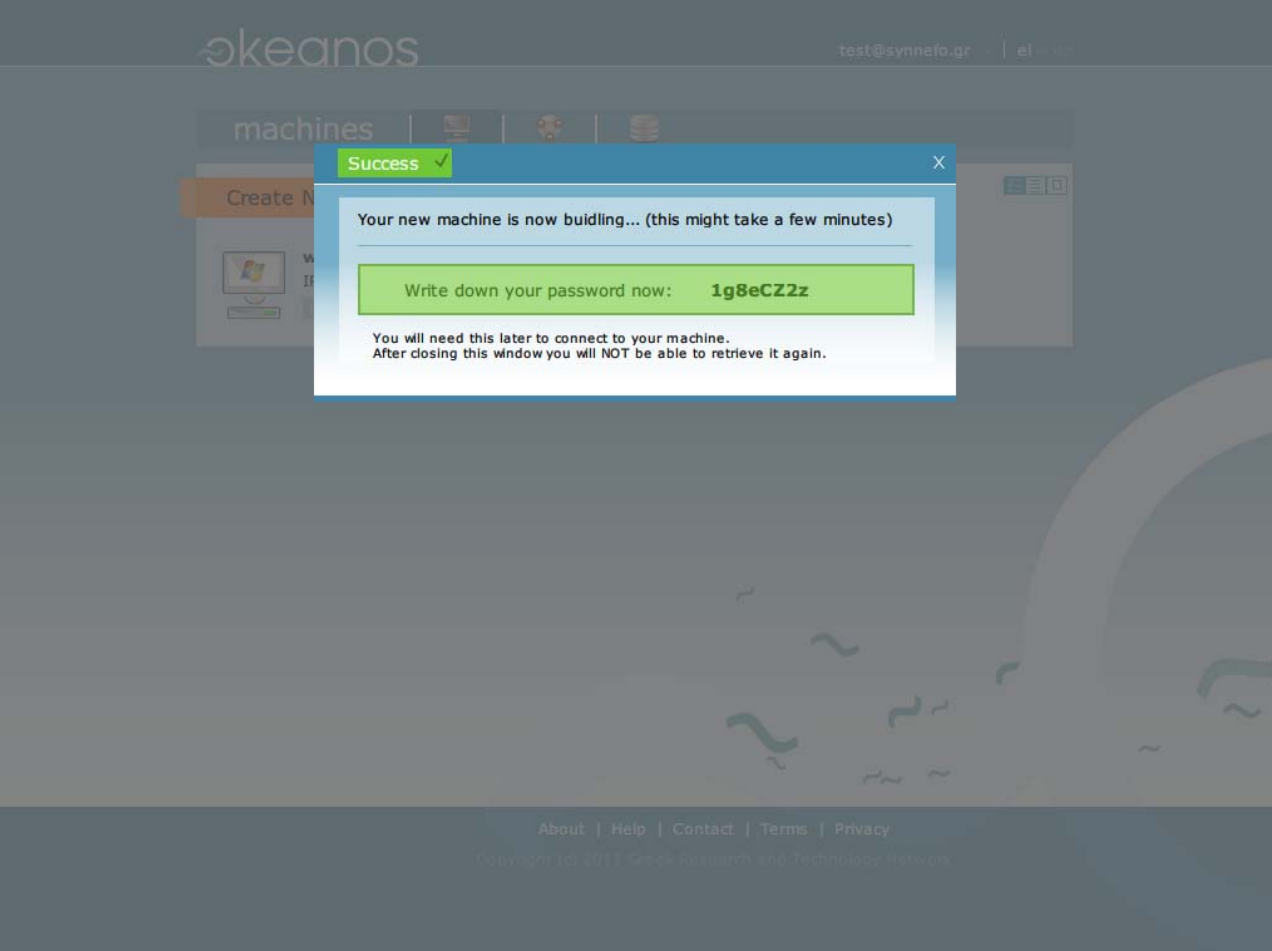
About | Help | Contact | Terms | Privacy

Copyright (c) 2011 Greek Research and Technology Network


1 Image
2 Flavor
3 Name
Select an OS
system images
custom images
Ubuntu
Ubuntu 11.04 2279 MB
Kubuntu
Kubuntu 11.04 2270 MB
Fedora Desktop
Fedora 15 Desktop Edition 2237 MB
Windows
Windows 2008 R2, Aero Desktop Experience 11000 MB
Cancel
Next


1 Image
2 Flavor
3 Name
Select CPUs, RAM and Disk Size
small
medium
large
custom
CPUs
2
cores
RAM
2048
MB
Size
80
GB
Your wallet: 10,000 Credits | This setup will cost you:
0
C/hour
Back
Next
About | Help | Contact | Terms | Privacy


1 Image ✓
2 Flavor ✓
3 Name
Confirm your settings
Name: win1
Image: Windows
CPUs: 2 cores
RAM: 2048MB
System Disk: 80GB
Cost per Hour: 40 credits
Credits in Wallet: 10.000
Back
Create VM


okeanos
test@synnefo.gr
machines
Create N
Success
X
Your new machine is now buidling... (this might take a few minutes)
Write down your password now: 1g8eCZ2z
You will need this later to connect to your machine.
After closing this window you will NOT be able to retrieve it again.
About | Help | Contact | Terms | Privacy

okeanos

test@synnefo.gr | el ~ en

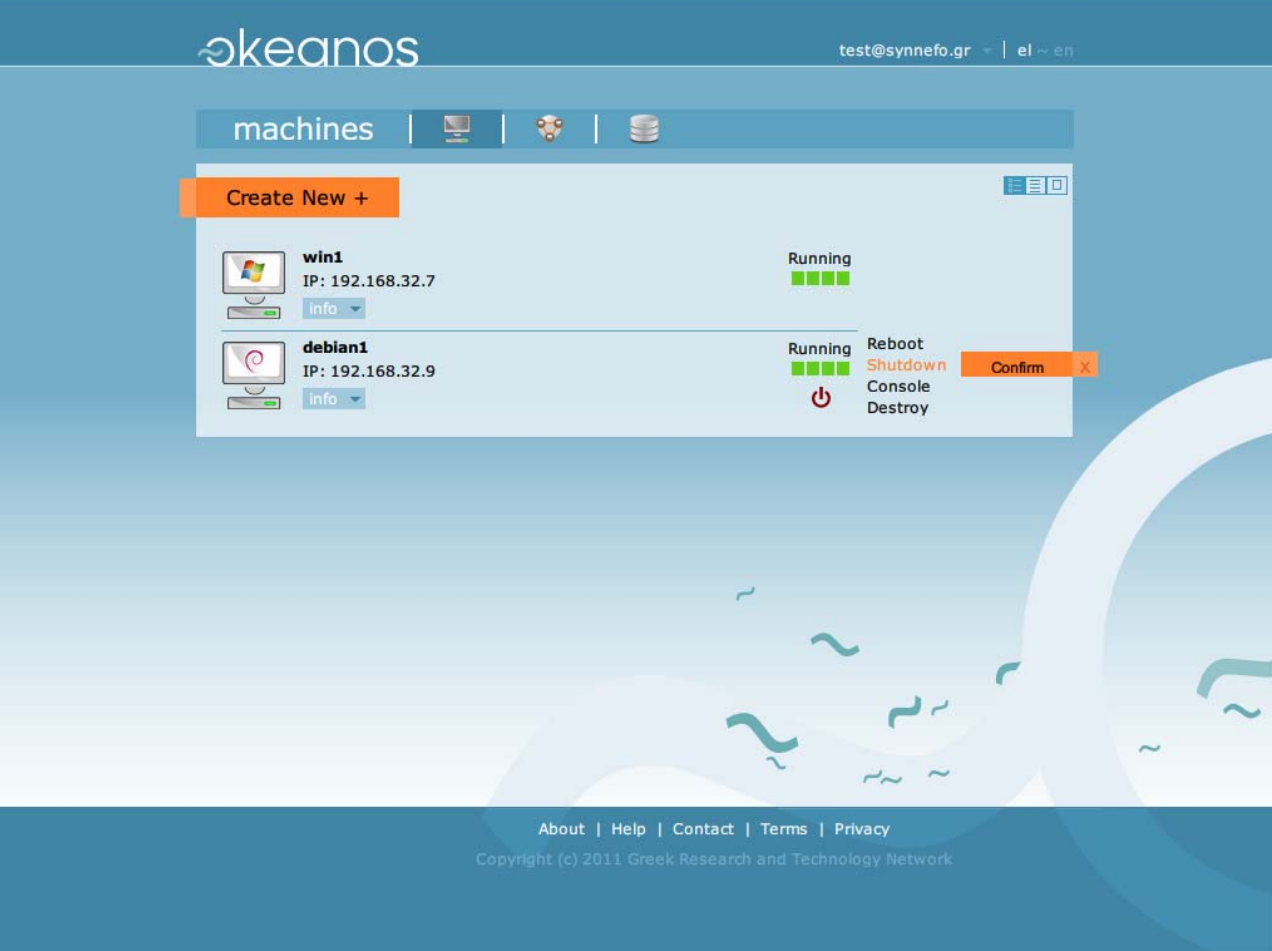
machines

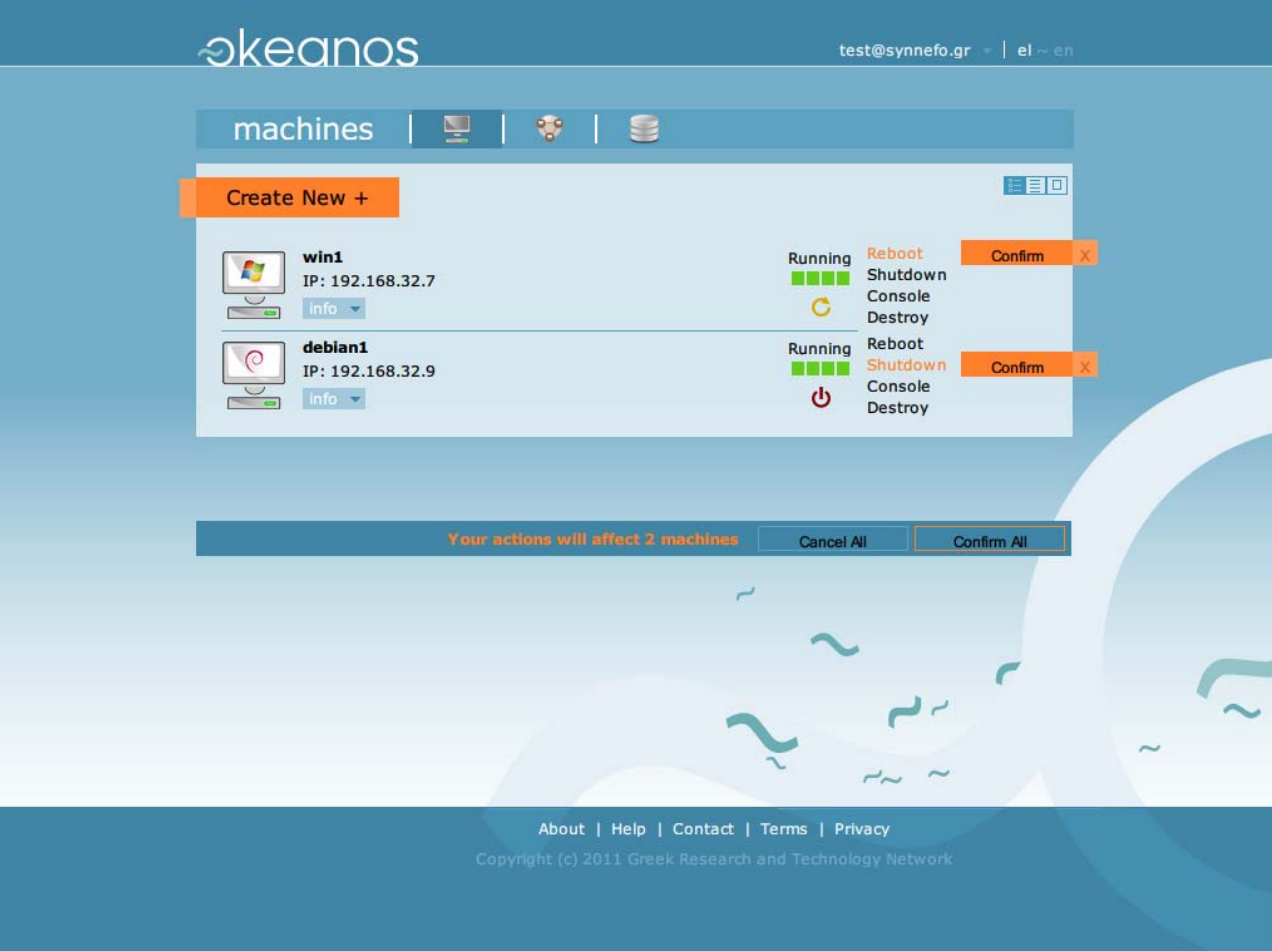
Create New +

win1
IP: undefined
info

Building

Copyright (c) 2011 Greek Research and Technology Network


okeanos
test@synnefo.gr | el ~ en
machines
Create New +
win1
IP: 192.168.32.7
info
Running
debian1
IP: 192.168.32.9
info
Running
Reboot
Shutdown
Confirm
Console
Destroy
About | Help | Contact | Terms | Privacy
Copyright (c) 2011 Greek Research and Technology Network


okeanos
test@synnefo.gr | el ~ en
machines
Create New +
win1
IP: 192.168.32.7
info
Running
Reboot
Shutdown
Console
Destroy
Confirm
debian1
IP: 192.168.32.9
info
Running
Reboot
Shutdown
Console
Destroy
Confirm
Your actions will affect 2 machines
Cancel All
Confirm All
About | Help | Contact | Terms | Privacy
Copyright (c) 2011 Greek Research and Technology Network

okeanos

test@synnefo.gr | el ~ en

machines

Create New +

win1
IP: 192.168.32.7
info

Shutting down

debian1
IP: 192.168.32.9
info

Running

Copyright (c) 2011 Greek Research and Technology Network

okeanos

test@synnefo.gr | el ~ en

machines

Create New +

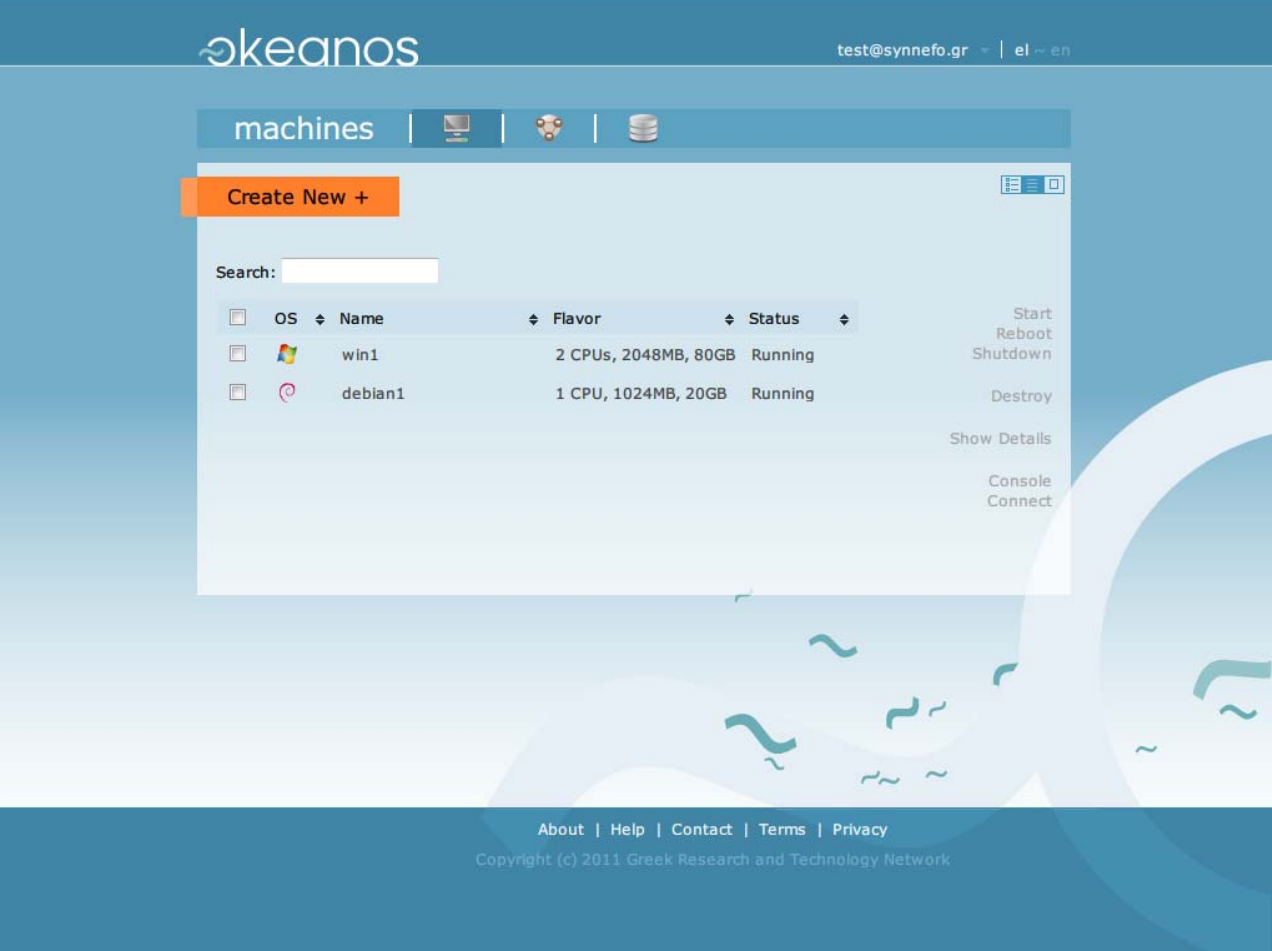
debian1
IP: 192.168.32.9
info

Running

win1
IP: 192.168.32.7
info

Stopped

Copyright (c) 2011 Greek Research and Technology Network


okeanos
test@synnefo.gr | el ~ en
machines
Create New +
Search:
OS
Name
Flavor
Status
win1
2 CPUs, 2048MB, 80GB
Running
debian1
1 CPU, 1024MB, 20GB
Running
Start
Reboot
Shutdown
Destroy
Show Details
Console
Connect
About | Help | Contact | Terms | Privacy
Copyright (c) 2011 Greek Research and Technology Network

okeanos

test@synnefo.gr | el ~ en

machines

Create New +

Running

Name:	win1
CPUs:	2
RAM (MB):	2048
System Disk (GB):	80
Image Name:	Windows
Image Size (GB):	11000
Public IPv4:	192.168.32.7
Public IPv6:	2001:db8::a800:ff:fe7c:3d80

Tags

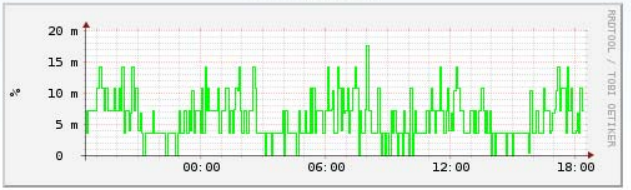
Reboot
Shutdown
Console
Destroy

previous next

win1

debian1

CPU Utilization



Network Utilization

Running

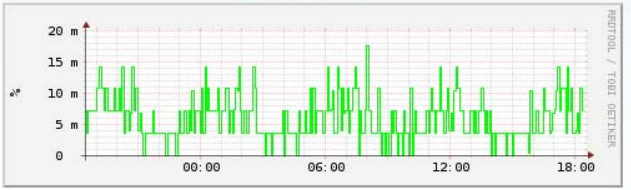
Image Name:	Windows
Image Size (GB):	11000
Public IPv4:	192.168.32.7
Public IPv6:	2001:db8::a800:ff:fe7c:3d80

Tags

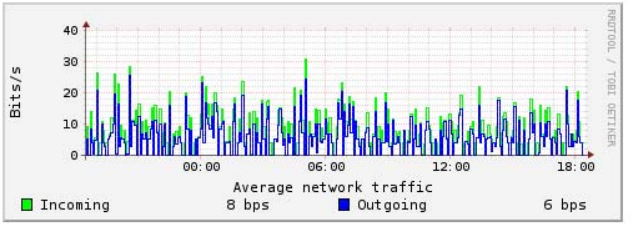
win1

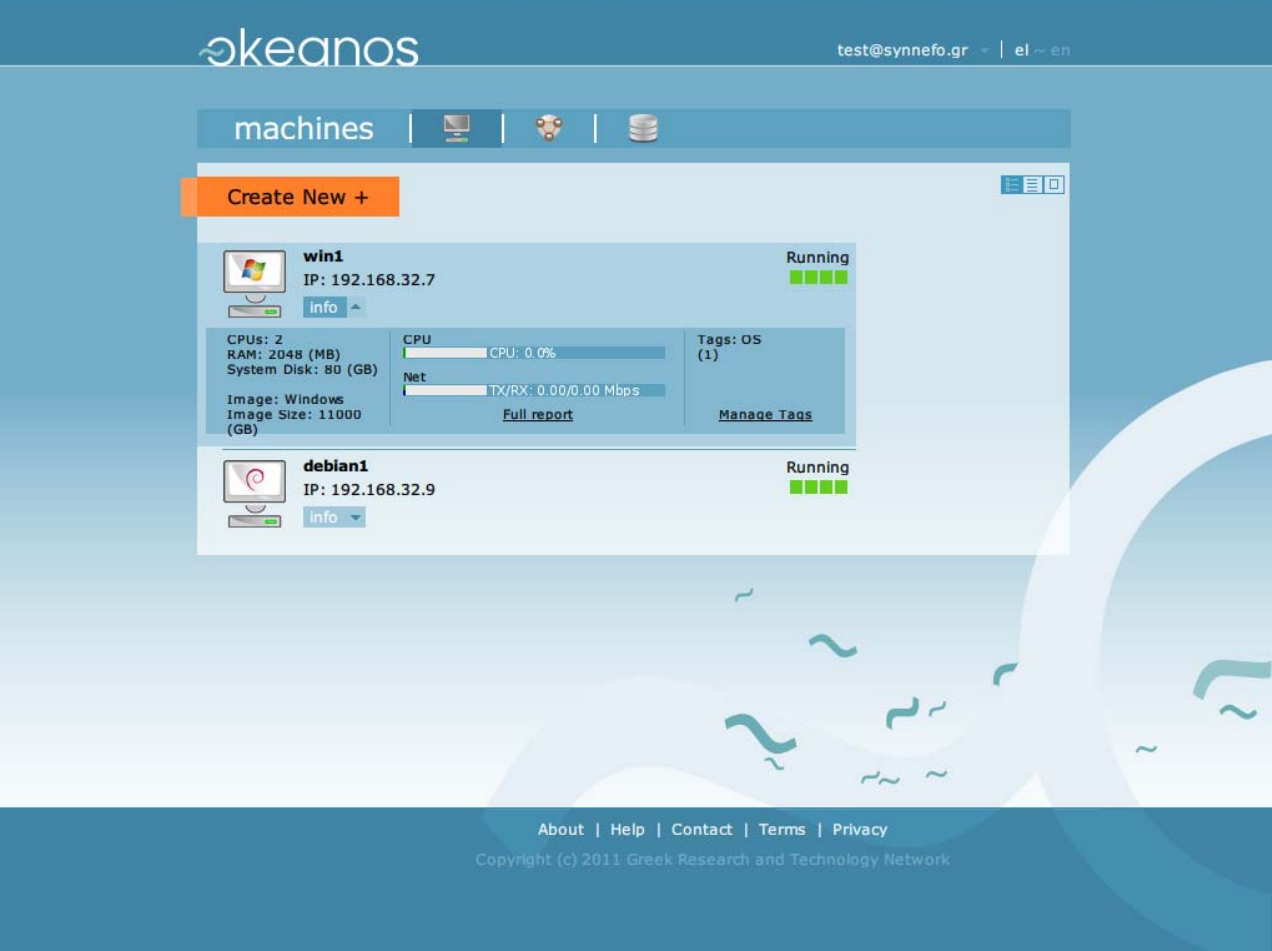
debian1

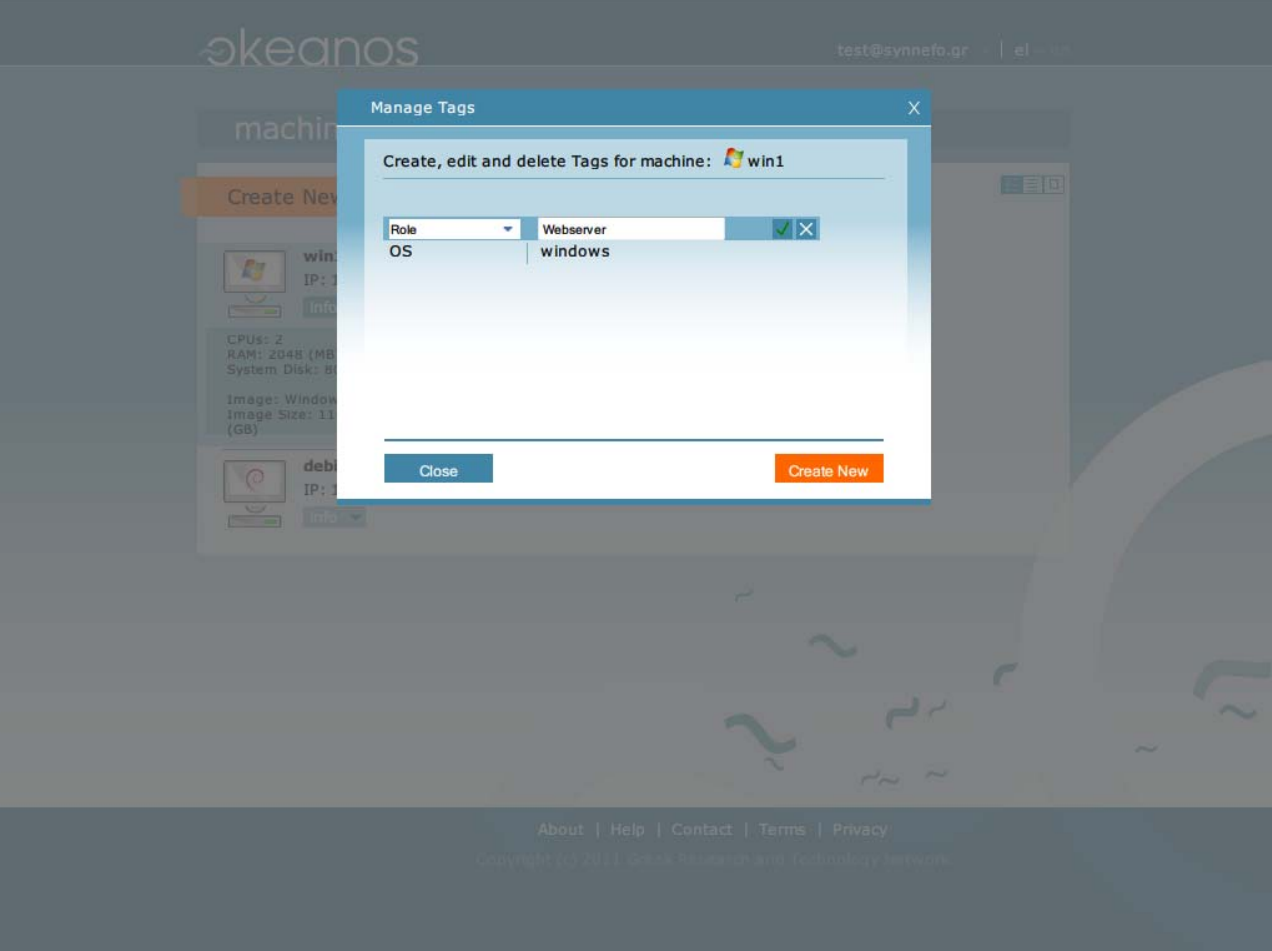
CPU Utilization



Network Utilization




okeanos
test@synnefo.gr | el ~ en
machines
Create New +
win1
IP: 192.168.32.7
Running
info
CPUs: 2
RAM: 2048 (MB)
System Disk: 80 (GB)
Image: Windows
Image Size: 11000 (GB)
CPU
CPU: 0.0%
Net
TX/RX: 0.00/0.00 Mbps
Full report
Tags: OS (1)
Manage Tags
debian1
IP: 192.168.32.9
Running
info
About | Help | Contact | Terms | Privacy
Copyright (c) 2011 Greek Research and Technology Network


okeanos
test@synnefo.gr
Manage Tags
Create, edit and delete Tags for machine: win1
Role
Webserver
OS
windows
Close
Create New
About | Help | Contact | Terms | Privacy

okeanos

test@synnefo.gr | el ~ en

networks

Create New +

Internet

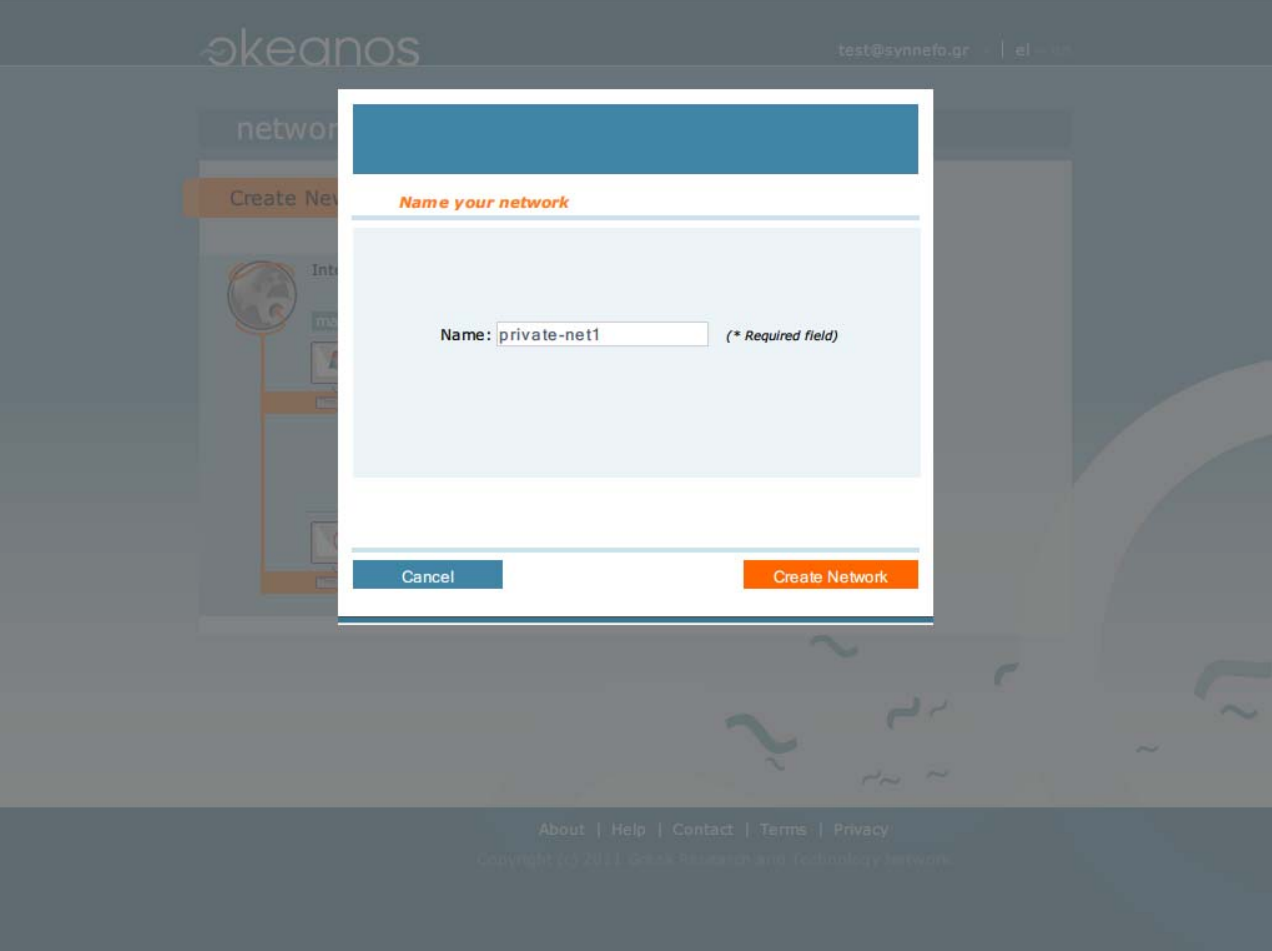
machines (2)

Public Network

Copyright (c) 2011 Greek Research and Technology Network

okeanos
test@synnefo.gr | el ~ en
networks
Create New +
Internet
Public Network
machines (2)
win1
IPv4: 192.168.32.7
IPv6: 2001:db8::a800:ff:fe7c:3d80
Firewall (Off)
debian1
IPv4: 192.168.32.9
IPv6: 2001:db8::a800:ff:fe81:cd6a
Firewall (Off)
About | Help | Contact | Terms | Privacy
Copyright (c) 2011 Greek Research and Technology Network

okeanos
test@synnefo.gr
el ~ en
networks
Create New +
Internet
Public Network
machines (2)
win1
IPv4: 192.168.32.7
IPv6: 2001:db8::a800:ff:fe7c:3d80
Firewall (Off)
Unprotected mode (Firewall off)
Fully protected mode (Firewall on)
Basically protected mode (Firewall on)
Apply
debian1
IPv4: 192.168.32.9
IPv6: 2001:db8::a800:ff:fe81:cd6a
Firewall (Off)
About | Help | Contact | Terms | Privacy
Copyright (c) 2011 Greek Research and Technology Network


okeanos
test@synnefo.gr
Create Ne
Name your network
Name: private-net1
(* Required field)
Cancel
Create Network
About | Help | Contact | Terms | Privacy

okeanos

test@synnefo.gr | el ~ en

networks

Create New +

Internet

machines (2)

Public Network

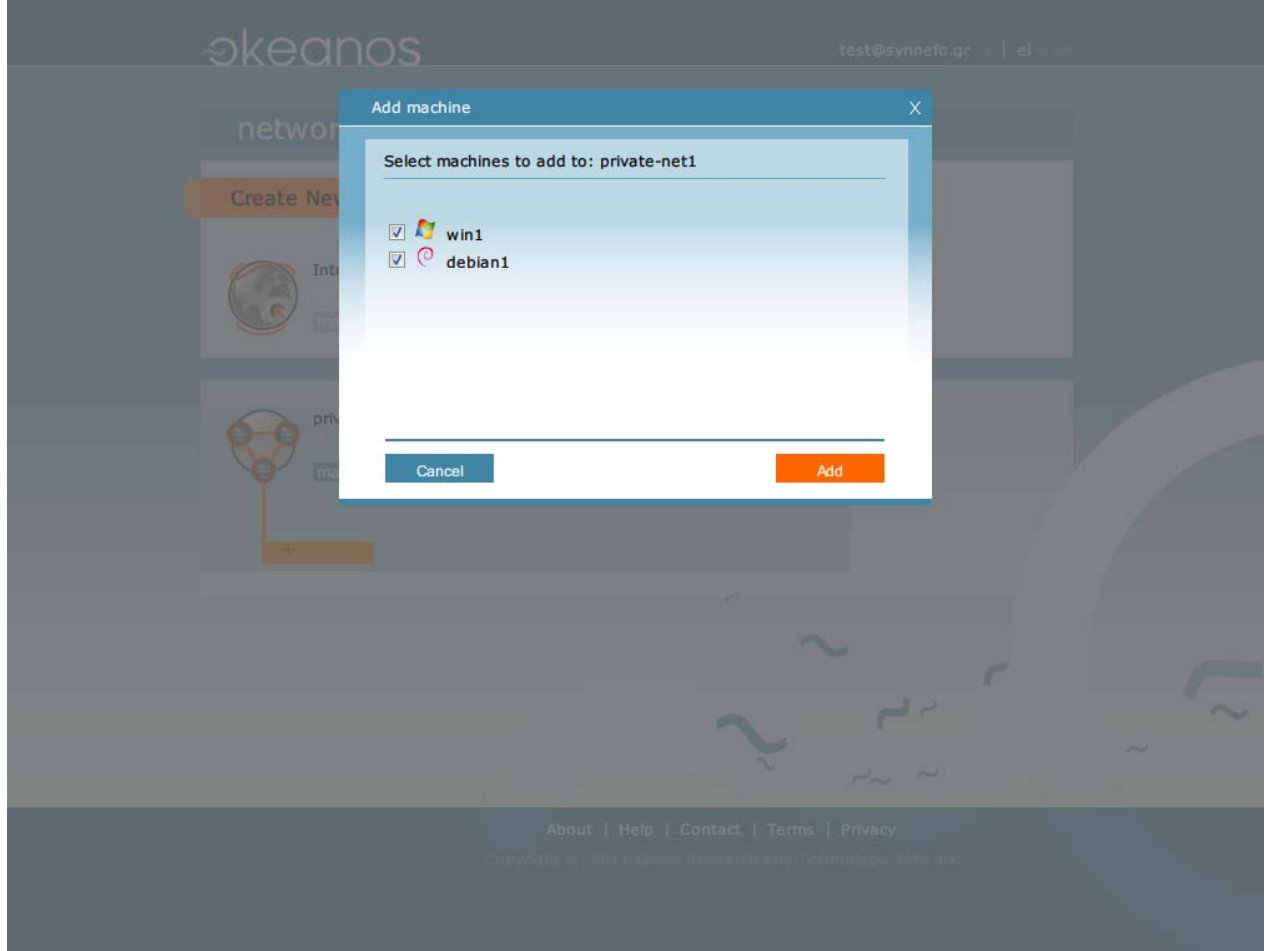
private-net1

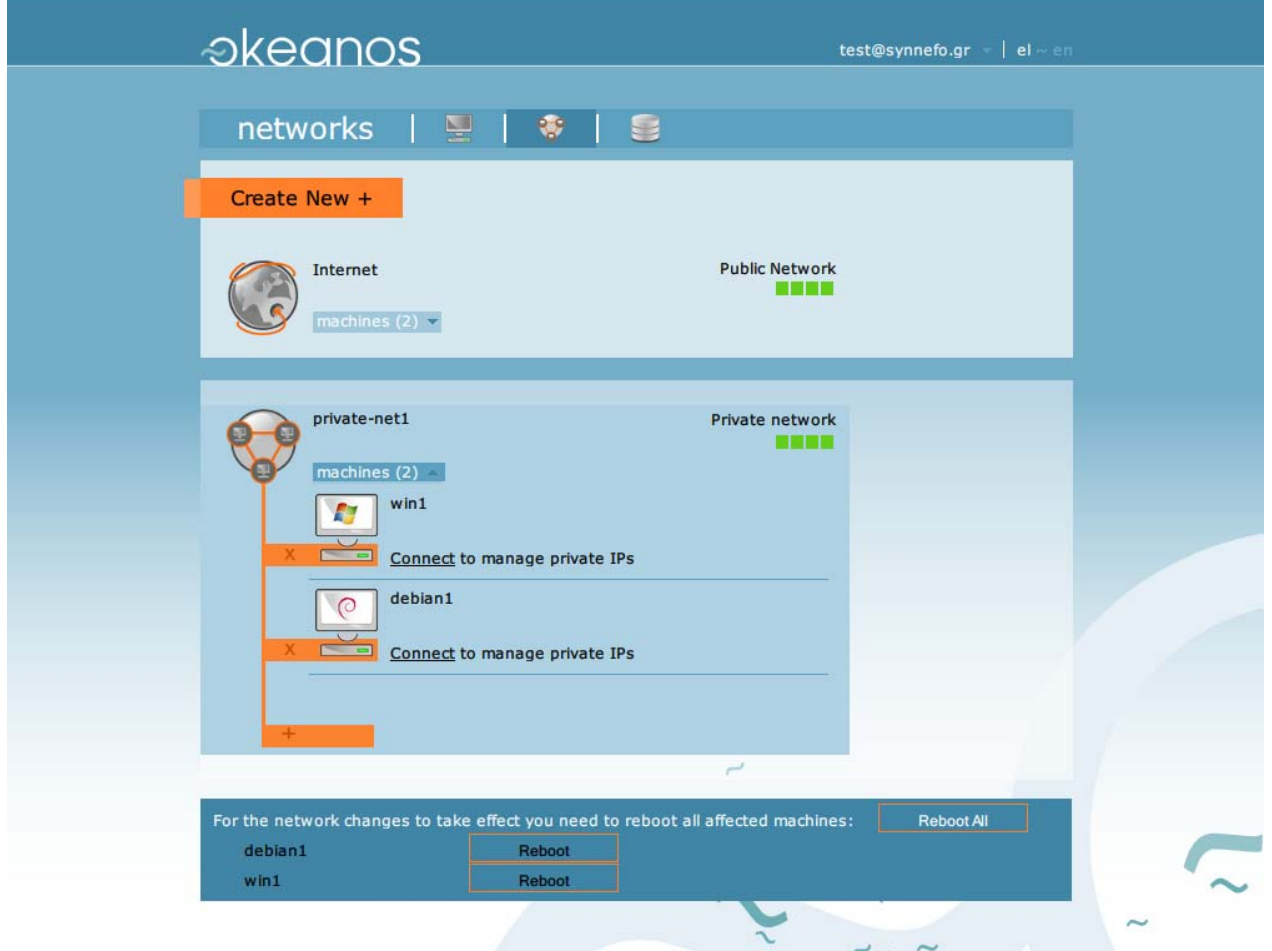
machines (0)

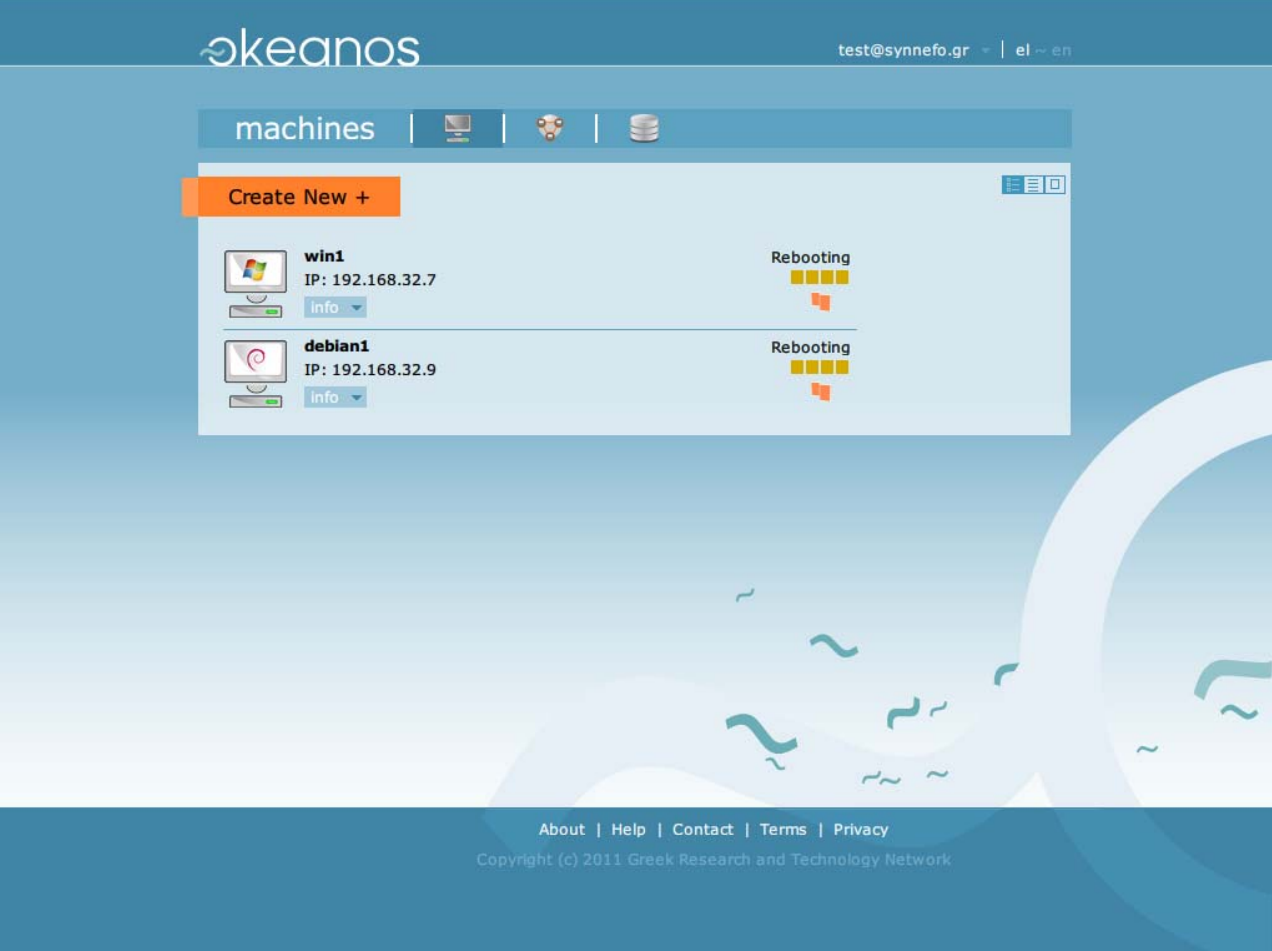
Private Network

About | Help | Contact | Terms | Privacy

Copyright (c) 2011 Greek Research and Technology Network


okeanos
test@synnefo.gr
Add machine
Select machines to add to: private-net1
win1
debian1
Cancel
Add
About | Help | Contact | Terms | Privacy


okeanos
test@synnefo.gr | el ~ en
networks
Create New +
Internet
Public Network
machines (2)
private-net1
Private network
machines (2)
win1
Connect to manage private IPs
debian1
Connect to manage private IPs
For the network changes to take effect you need to reboot all affected machines:
Reboot All
debian1
Reboot
win1
Reboot


okeanos
test@synnefo.gr | el ~ en
machines
Create New +
win1
IP: 192.168.32.7
info
Rebooting
debian1
IP: 192.168.32.9
info
Rebooting
About | Help | Contact | Terms | Privacy
Copyright (c) 2011 Greek Research and Technology Network