



# Σχεδίαση και Υλοποίηση Μηχανισμού Κρυφής Μνήμης για Κατανεμημένο Σύστημα Αποθήκευσης σε Περιβάλλον Υπολογιστικού Νέφους

1. Καλημέρα σας, ονομάζομαι Αλέξιος Πυργιώτης  
Θα σας παρουσιάσω τη διπλωματική μου με τίτλο:..
2. Ακούγεται κάπως περίεργο στα ελληνικά...  
αυτό που πραγματεύεται είναι την δημιουργία ενός caching  
μηχανισμού για το Archipelago, ένα distributed, storage layer
3. Συγκεκριμένα, στην παρουσίαση αυτή θα μιλήσουμε για τον cached,  
δηλαδή τον caching μηχανισμό μας, αλλά και για το synapsed, ένα  
συμπληρωματικό εργαλείο που στόχος του είναι να δώσει στον  
cached δικτυακές δυνατότητες



Αλέξιος Πυργιώτης

Εθνικό Μετσόβιο Πολυτεχνείο

October 15, 2013

# Contents

---

Introduction

Request handling

Caching

Cachedi design

## 1. Ο κορμός της παρουσίασης είναι ο εξής:

- Αρχικά, παρουσιάζουμε κάποια εισαγωγικά που αφορούν το background της εργασίας μας. Αναφέρουμε τι είναι το Synnefo, τι είναι η υπηρεσία okeanos και τι είναι το Αρχιπέλαγο
- Έπειτα, δείχνουμε τον τρόπο με τον οποίο η υποδομή μας χειρίζεται αιτήματα δεδομένων από ένα VM.



# Table of Contents

Introduction

Request handling

Caching

Cached design



Ας ξεκινήσουμε με την παρούσα κατάσταση.  
Το software που τα ξεκίνησε όλα είναι το Synnefo

..by GRNET -> Και φυσικά τα παιδιά που βλέπετε εδώ

- Compute service, είναι η υπηρεσία η οποία προμηθεύει τους χρήστες με VMs και επιτρέπει το χειρισμό τους
- Network service, είναι η υπηρεσία η οποία δίνει τη δυνατότητα στους χρήστες να δημιουργήσουν ιδιωτικά δίκτυα και να συνδέσουν τα VMs τους σε αυτά.
- Storage service, που κοινώς αποθηκεύει τα αρχεία των χρηστών. Στην περίπτωση του Synnefo όμως, έχουμε ένα κοινό σημείο για τα πάντα: είτε είναι αρχεία, είτε δίσκοι των VMS, είτε images
- Image Service, υπεύθυνο για το deployment ενός VM από ένα image. Επίσης, κάνει και παραμετροποιήσεις (παράδειγμα ssh κλειδιά)

## Synnefo



Open source, production-ready, cloud software.  
Designed since 2010 by GRNET.

Synnefo, as most cloud software, has the following services:

- Compute Service
- Network Service
- Storage Service
- Image Service
- Identity Service



- IaaS είναι πρακτικά η παροχή εικονικής υποδομής σε χρήστες (δηλαδή πάρε υπολογιστή (VM), δίκτυα, αρχεία κτλ)
- Δωρεάν για τους Ακαδημαϊκούς σκοπούς, ήδη γίνονται εργαστήρια στο EMP και απ' αυτό το εξάμηνο σε άλλες σχολές

## okeanos



- IaaS service
- Targeted at the Greek Academic and Research Community
- Designed by GRNET
- In production since 2011



## okeanos



- IaaS είναι πρακτικά η παροχή εικονικής υποδομής σε χρήστες (δηλαδή πάρε υπολογιστή (VM), δίκτυα, αρχεία κτλ)
- Δωρεάν για τους Ακαδημαϊκούς σκοπούς, ήδη γίνονται εργαστήρια στο EMP και απ' αυτό το εξάμηνο σε άλλες σχολές

- IaaS service
- Targeted at the Greek Academic and Research Community
- Designed by GRNET
- In production since 2011
- ...and of course powered by Synnefo.



# Table of Contents

Introduction

Request handling

Caching

Cached design



## What is request handling?

Τι είναι η διαχείριση των αιτημάτων ενός VM?

Είναι η εφαρμογή πολιτικών και επεξεργασία των αιτημάτων σε όλη την πορεία τους μέχρι το να φτάσουν στο storage.

Δηλαδή έχουμε ένα εικονικό μηχάνημα <κλικ>

... το storage μας <κλικ>

και πρέπει με κάποιο τρόπο τα δεδομένα του μηχανήματος να φτάσουν σε εμάς <κλικ>

Ένας απλός τρόπος θα ήταν να τα συνδέσουμε. Άλλωστε όταν τρέχει VM, ο hypervisor κοιτάει block device. Θα μπορούσε να ήταν κομμάτι του storage Είναι αυτό αρκετό; <κλικ>

Όχι, χρειαζόμαστε επίσης **FIXME:**







## What is request handling?

Τι είναι η διαχείριση των αιτημάτων ενός VM?

Είναι η εφαρμογή πολιτικών και επεξεργασία των αιτημάτων σε όλη την πορεία τους μέχρι το να φτάσουν στο storage.

Δηλαδή έχουμε ένα εικονικό μηχάνημα <κλικ>

... το storage μας <κλικ>

και πρέπει με κάποιο τρόπο τα δεδομένα του μηχανήματος να φτάσουν σε εμάς <κλικ>

Ένας απλός τρόπος θα ήταν να τα συνδέσουμε. Άλλωστε όταν τρέχει VM, ο hypervisor κοιτάει block device. Θα μπορούσε να ήταν κομμάτι του storage Είναι αυτό αρκετό; <κλικ>

Όχι, χρειαζόμαστε επίσης **FIXME:**





## What is request handling?

Τι είναι η διαχείριση των αιτημάτων ενός VM?

Είναι η εφαρμογή πολιτικών και επεξεργασία των αιτημάτων σε όλη την πορεία τους μέχρι το να φτάσουν στο storage.

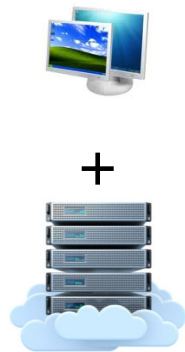
Δηλαδή έχουμε ένα εικονικό μηχάνημα <κλικ>

... το storage μας <κλικ>

και πρέπει με κάποιο τρόπο τα δεδομένα του μηχανήματος να φτάσουν σε εμάς <κλικ>

Ένας απλός τρόπος θα ήταν να τα συνδέσουμε. Άλλωστε όταν τρέχει VM, ο hypervisor κοιτάει block device. Θα μπορούσε να ήταν κομμάτι του storage Είναι αυτό αρκετό; <κλικ>

Όχι, χρειαζόμαστε επίσης **FIXME:**





Τι είναι η διαχείριση των αιτημάτων ενός VM?

Είναι η εφαρμογή πολιτικών και επεξεργασία των αιτημάτων σε όλη την πορεία τους μέχρι το να φτάσουν στο storage.

Δηλαδή έχουμε ένα εικονικό μηχάνημα <κλικ>

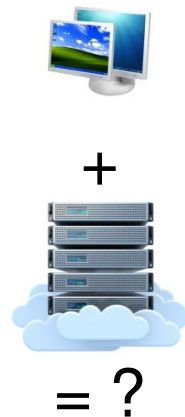
... το storage μας <κλικ>

και πρέπει με κάποιο τρόπο τα δεδομένα του μηχανήματος να φτάσουν σε εμάς <κλικ>

Ένας απλός τρόπος θα ήταν να τα συνδέσουμε. Άλλωστε όταν τρέχει VM, ο hypervisor κοιτάει block device. Θα μπορούσε να ήταν κομμάτι του storage Είναι αυτό αρκετό; <κλικ>

Όχι, χρειαζόμαστε επίσης **FIXME:**

## What is request handling?





= ?

- Policy enforcement?
- Storage agnosticity?

## What is request handling?

Τι είναι η διαχείριση των αιτημάτων ενός VM?

Είναι η εφαρμογή πολιτικών και επεξεργασία των αιτημάτων σε όλη την πορεία τους μέχρι το να φτάσουν στο storage.

Δηλαδή έχουμε ένα εικονικό μηχάνημα <κλικ>

... το storage μας <κλικ>

και πρέπει με κάποιο τρόπο τα δεδομένα του μηχανήματος να φτάσουν σε εμάς <κλικ>

Ένας απλός τρόπος θα ήταν να τα συνδέσουμε. Άλλωστε όταν τρέχει VM, ο hypervisor κοιτάει block device. Θα μπορούσε να ήταν κομμάτι του storage Είναι αυτό αρκετό; <κλικ>

Όχι, χρειαζόμαστε επίσης **FIXME:**



+



= ?

- Policy enforcement?
- Storage agnosticity?

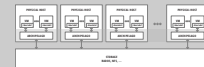


Request handling

Our solution

Our solution

Archipelago



Key features: 1) Software-defined 2) Distributed 3) Modular  
Copy-On-Write Storage agnostic

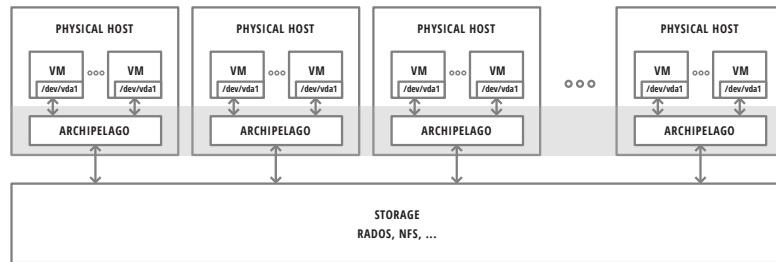
Η λύση που χρησιμοποιήσαμε είναι το Archipelago

- Software-defined: αν και είναι ένα όρος μαρκετινγκ, εμείς κανονικά. Σημαίνει με το software ΟΡΙΖΕΙΣ το storage (εφαρμογή policy, αλλαγή πορείας του request)
- τρέχει σε πολλούς κόμβους
- αποτελείται από διακριτά κομμάτια
- κάνει CoW (εξήγησε ότι τα images είναι λίγα, τα VMs πολλά, όπως όταν ένα process κάνει fork)
- μπορούμε χρησιμοποιήσουμε ότι θέλουμε

## Request handling

### Our solution

### Archipelago



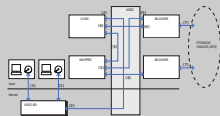
Key features: 1) Software-defined 2) Distributed 3) Modular  
Copy-On-Write Storage agnostic



Request handling

Archipelago Architecture

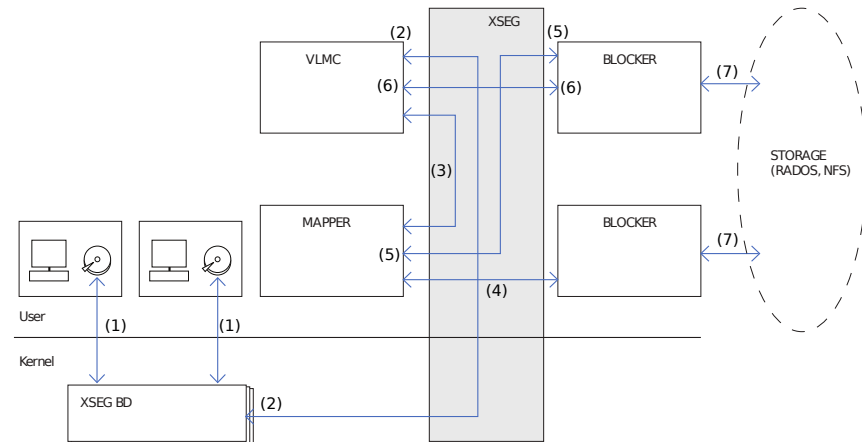
Archipelago Architecture



- Το VM στέλνει αίτημα στο δίσκο του, ο δίσκος είναι εικονικός, θα το δει ο hypervisor (εξήγησε τι είναι ο hypervisor) και θα το στείλει στον δίσκο που το έχουμε πει. (xsegbd)
- 2) **FIXME:**

## Request handling

### Archipelago Architecture



└ Request handling

└ RADOS

RADOS

Self replicating self managing  
TODO

Request handling

## RADOS

Το RADOS έχει τα εξής πολύ θετικά:

- Self-replication

Self replicating self managing

**TODO:**



└ Request handling

└ RADOS

#### RADOS

Self replicating self managing  
**TODO:**  
Speed issues.

Page-cache: > 90MB/s, < 1ms  
Rados: < 7MB/s, 10ms

## Request handling

# RADOS

Το RADOS έχει τα εξής πολύ θετικά:

- Self-replication

Self replicating self managing

**TODO:**

Speed issues.

Page-cache: > 90MB/s, < 1ms

Rados: < 7MB/s, 10ms





└ Request handling

└ RADOS

#### RADOS

Self replicating self managing  
**TODO:**  
Speed issues.

Page-cache: > 90MB/s, < 1ms  
Rados: < 7MB/s, 10ms  
Thesis goal, make this faster.

## Request handling

# RADOS

Το RADOS έχει τα εξής πολύ θετικά:

- Self-replication

Self replicating self managing

**TODO:**

Speed issues.

Page-cache: > 90MB/s, < 1ms

Rados: < 7MB/s, 10ms

Thesis goal, make this faster.



# Table of Contents

---

Introduction

Request handling

Caching

Cached design



No doubt we must do caching Caching is ... similar to the memory  
hierarchy model, ...

## Caching

No doubt we must do caching Caching is ... similar to the memory  
hierarchy model, ...



## Block solutions

- Bcache
- ...

Pros: ...

Cons: ...



## Object solutions

- Memcached
- Couchbase

Pros: ...

Cons: ...



## Page-cache

What if we used the page-cache?

Pros: ...

Cons: ...



- Cannot use those reliably with with Archipelago, may use them later.
- Have to create our own solution

## Conclusion

- Cannot use those reliably with with Archipelago, may use them later.
- Have to create our own solution



# Table of Contents

[Introduction](#)[Request handling](#)[Caching](#)[Cachedi design](#)



└─ Cachedi design

└─ Our decision

Our decision

- create something close understands Archipelago logic
- measure the best possible performance we can get

## Our decision

- create something close understands Archipelago logic
- measure the best possible performance we can get



- Nativity
- Pluggability
- In-memory
- Low indexing overhead

## Requirements

For our implementation we pose the following requirements:

- Nativity
- Pluggability
- In-memory
- Low indexing overhead



└─ Cachedi design

└─ Cached design

Cached design

operations image  
component image

Cachedi design

## Cached design

operations image  
component image



## Xcache design

show with red the xcache design  
go to xcache picture  
Show what xcache does



## Xworkq design

show the xworkq design  
what xworkq does



└─ Cachedi design

└─ Xwaitq design

Xwaitq design

show xwaitq design  
what xwaitq does

Cachedi design

## Xwaitq design

show xwaitq design  
what xwaitq does



## Bucket pool

Buckets have been designed to do this  
list

