Link Samba to Cloud Storage

May 2011

Fabrizio Manfredi Furuholmen

ZEROPIU
Difference as value

Link Cloud Storage

Goals

Solution

Cloud Overview

Architecture

Components

Examples

Limitations

Samba Management Console

Update

Design

New Functions

Demo

* One Session two talks

5/11/11

ZEROPIU SpA, founded in 1994 by a group of professionals from leading multinational Information Technology companies.

We operate at international level and our Headquarters are in Italy. We have subsidiaries in Scandinavia called ZEROPIU Nordic, and in Libya with ZEROPIU MEA.

- Competence Center, able to support customers in meeting the standard regulation requirements (D.Lgs. 196/2003), Basilea II, Sarbanes Oxley Act, BS7799/ Iso and 17799/Iso 27001.

Consulting
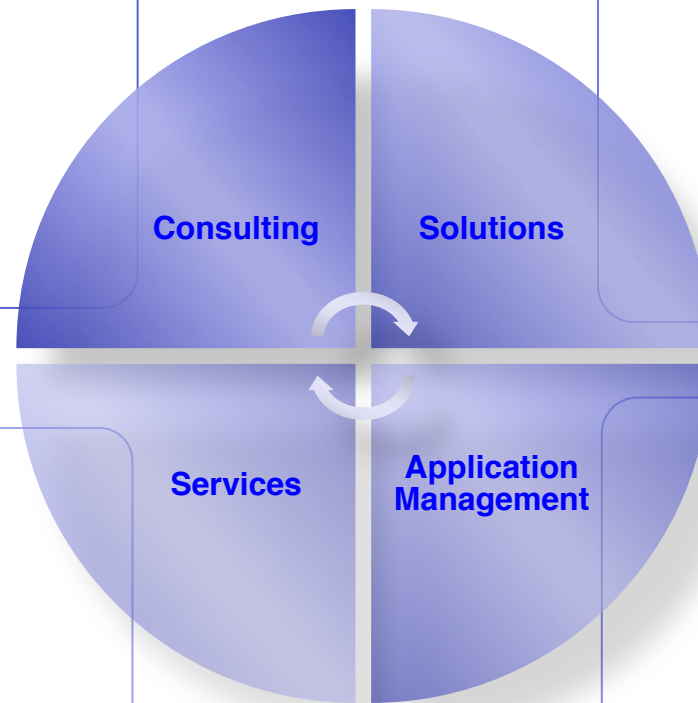
Solutions

- Identity Access Management (IAM), Authentication and Authorization, Single Sign-On and Provisioning
- Video on Demand solutions
- Intranet, Extranet and Enterprise Portal integrated vertical solutions and applications
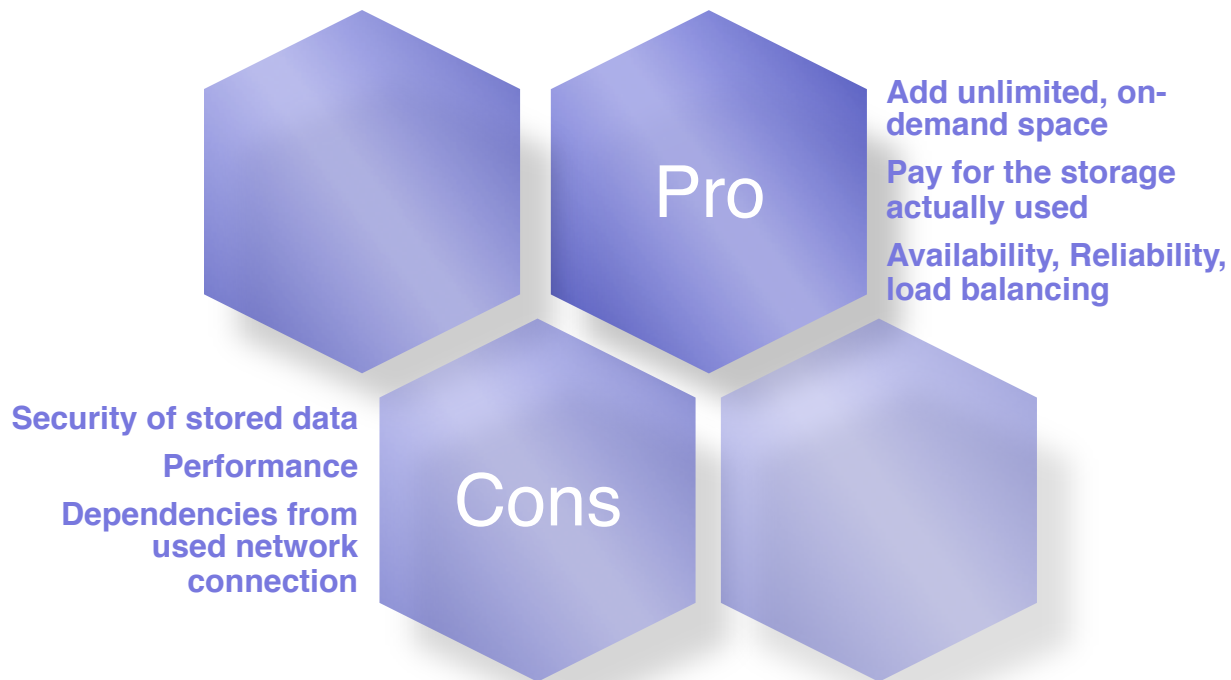
Services

Application Management

- Application and Facility management
- Service Desk and technical Call Center.

- Software infrastructure design, management and maintenance
- Application upgrade and evolution
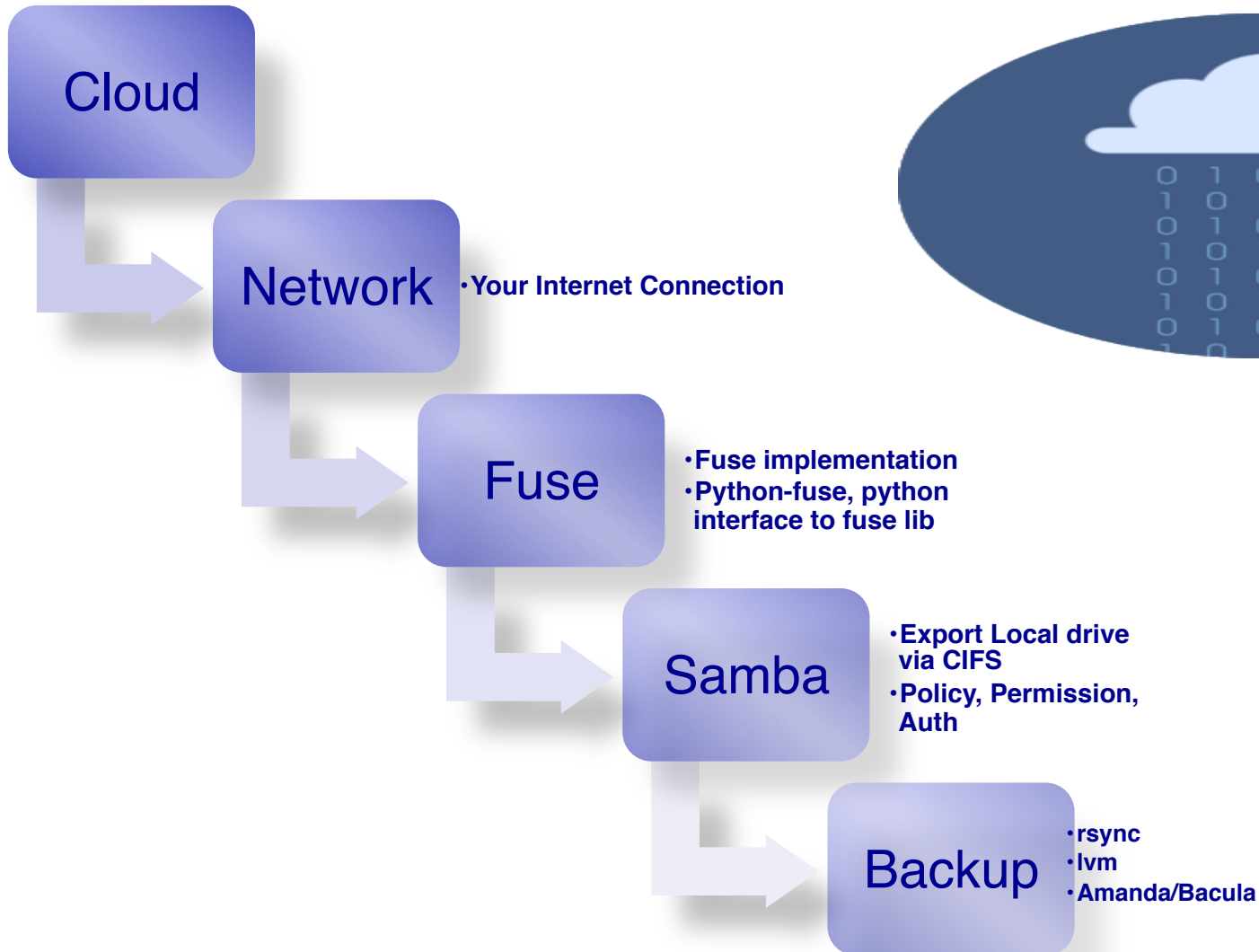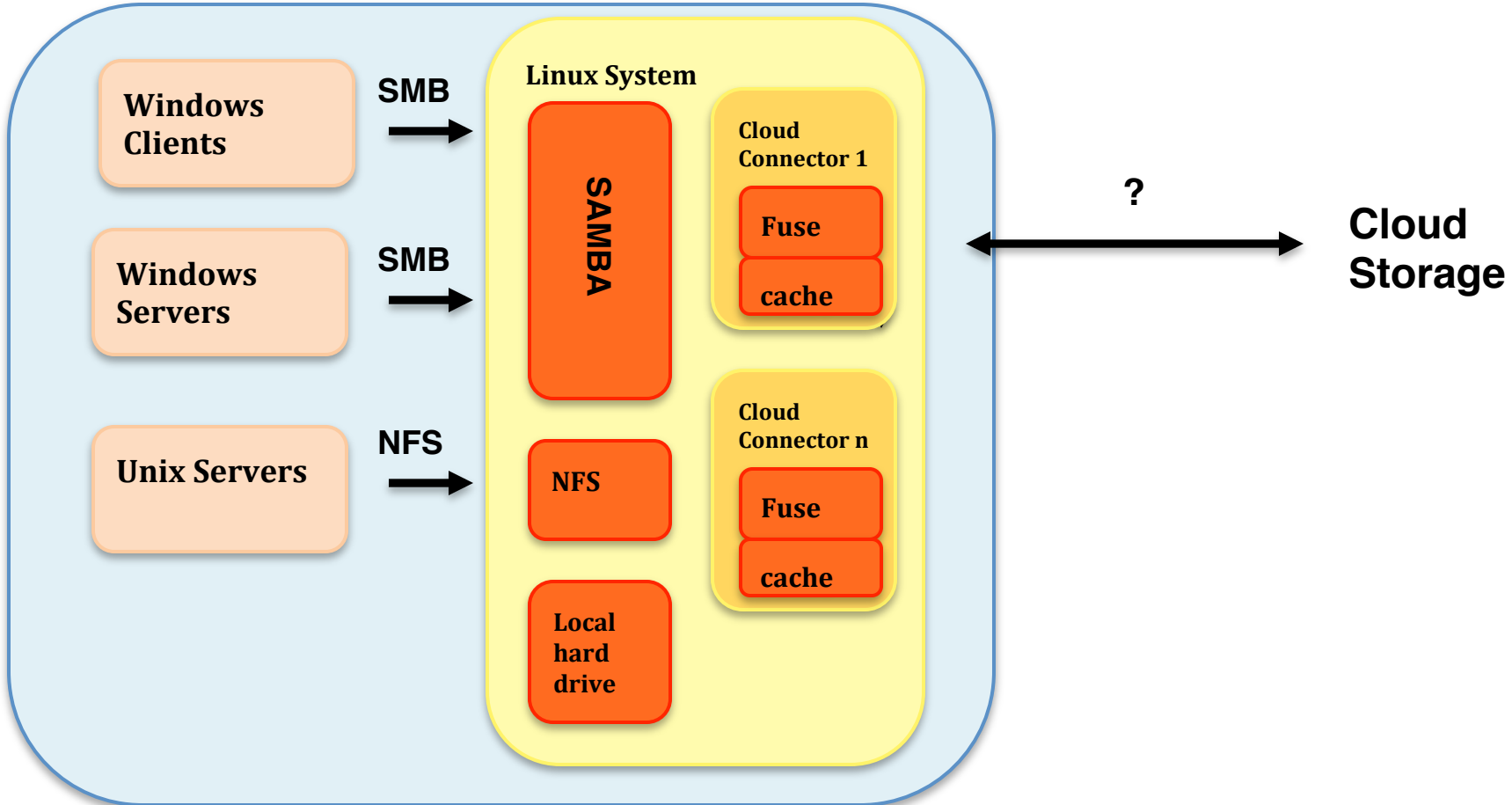- Testing and Quality Assurance

5/11/11

5/11/11
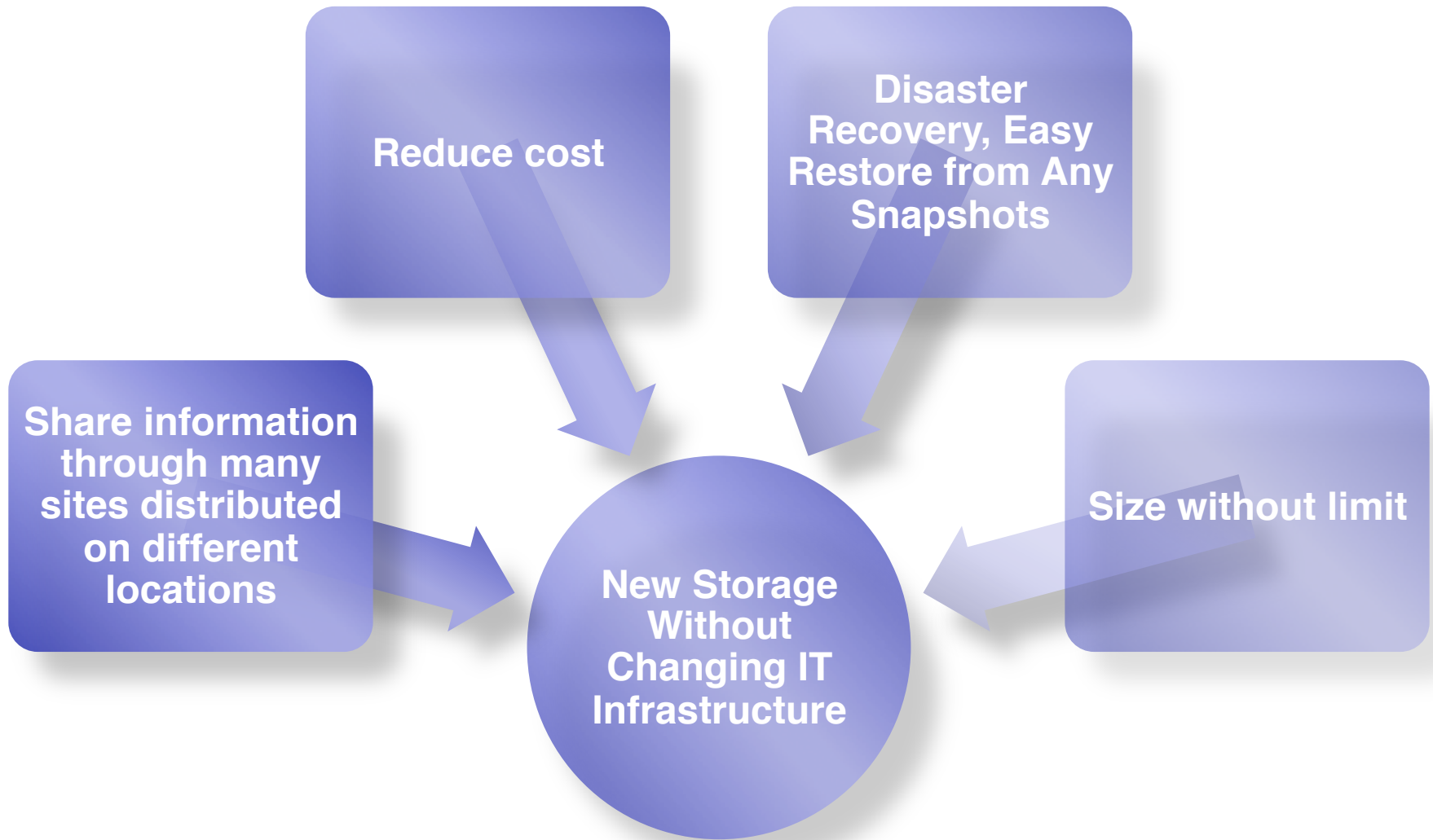
"Cloud storage is a model of networked online storage where data are stored on multiple virtual servers, generally hosted by third parties, rather than being hosted on dedicated servers…" (wikipedia)

CLOUD STORAGE

Pro

Add unlimited, on-demand space

Pay for the storage actually used

Availability, Reliability, load balancing

Cons

Security of stored data

Performance

Dependencies from used network connection

5

Cloud

Network
- Your Internet Connection

Fuse
- Fuse implementation
- Python-fuse, python interface to fuse lib

Samba
- Export Local drive via CIFS
- Policy, Permission, Auth

Backup
- rsync
- lvm
- Amanda/Bacula
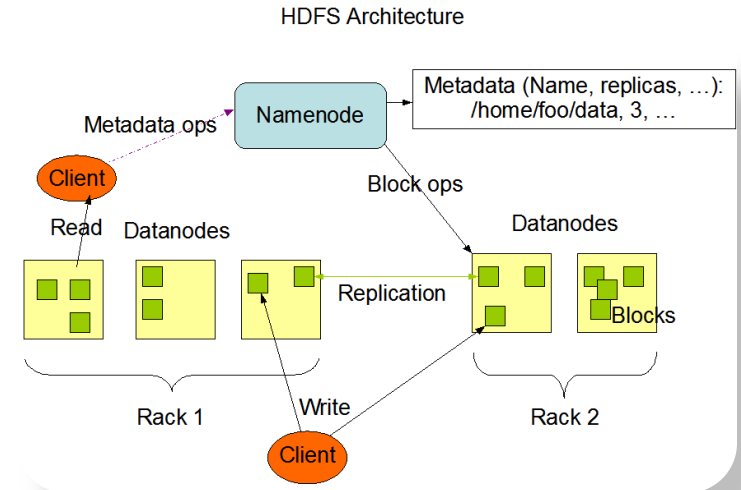
6

5/11/11

Cloud storage gateways

Namenode

- An HDFS cluster consists of a single NameNode
- It is a master server that manages the file system namespace and regulates access to files by clients.

Datanodes

- Datanode manage storage attached to the system it run on
- Applay the map rule of MapReduce

Blocks

- File is split into one or more blocks and these blocks are stored in a set of DataNodes

HDFS Architecture

Metadata ops

Namenode

Metadata (Name, replicas, …):
/home/foo/data, 3, …

Client

Block ops

Read   Datanodes

Datanodes

Replication

Blocks

Rack 1   Write   Rack 2

Client

Block Replication

Namenode (Filename, numReplicas, block-ids, …)
/users/sameerp/data/part-0, r:2, {1,3}, …
/users/sameerp/data/part-1, r:3, {2,4,5}, …

Datanodes
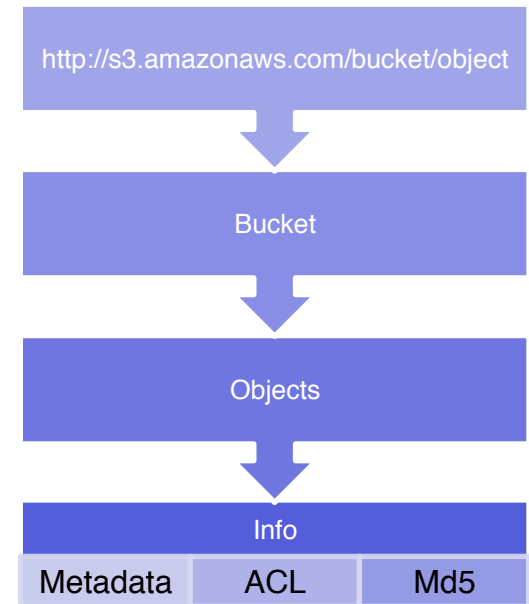
Amazon Simple Storage Service
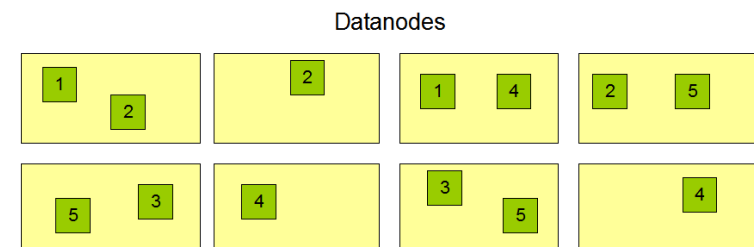
Amazon S3 provides a simple web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the web. It gives any developer access to the same highly scalable, reliable, secure, fast, inexpensive infrastructure…

Designed to provide 99.999999999% durability and 99.99% availability of objects over a given year.
**(Amazon disclaimer)**

❑ **S3 stores arbitrary objects up to 5 terabytes in size**
❑ **Objects are organized into buckets, and identified unique name**
❑ **Buckets and objects can be created, listed, and retrieved using either a REST-style HTTP interface or a SOAP interface.**
❑ **Objects can be downloaded using the HTTP GET interface and the BitTorrent protocol.**
❑ **Objects Versioning**

http://s3.amazonaws.com/bucket/object
Bucket
Objects
Info

| Metadata	ACL	Md5

**S3 storage service now hosts more than 100 billion objects !**

FUSE , filesystem in user space

❑ Simple library API
❑ Simple installation (no need to patch or recompile the kernel)
❑ Secure implementation
❑ Userspace – kernel interface is very efficient
❑ Usable by non privileged users
❑ Runs on Linux kernels 2.4.X and 2.6.X
❑ Lot of Language bindings
❑ Lot of file systems supported (> 50)
❑ Mount as a local filesystem

5/11/11

Cloud storage
- more application programming interface
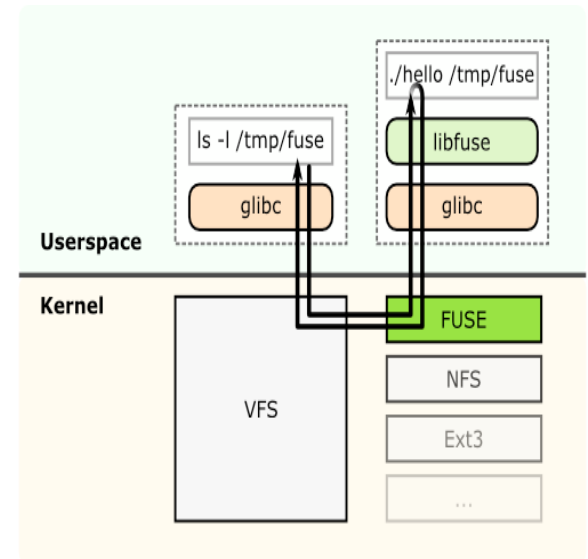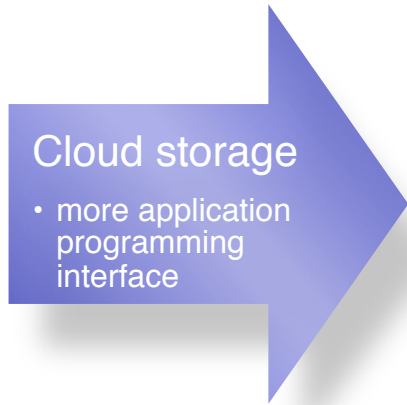
Cloud	Unix
Bucket	Mount point
Objects	File
Objects empty	Directory
Amz-meta-custom fields	File Attributes
Md5 comparison	Dirty cache
Remote copy	Rename/move

Storage
- based on block- or file-based

5/11/11
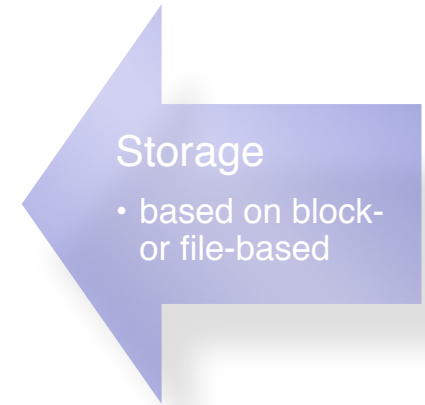
CIFS Translator

Permission Control

Samba

Locks

Authentication

5/11/11

5/11/11

S3fs
- use_cache
- retries
- connect_timeout
- readwrite_timeout
- max_stat_cache_size
- prefix

Fuse
- allow_other
- kernel_cache

Samba
- It is simple share
- Locks handle by samba

5/11/11

□ The buckets sharing is not safe due to the Amazons S3 consistency data model.

□ Only module supported by Fuse

□ Extended attributes and locks not supported

□ On change, the entire file is transferred
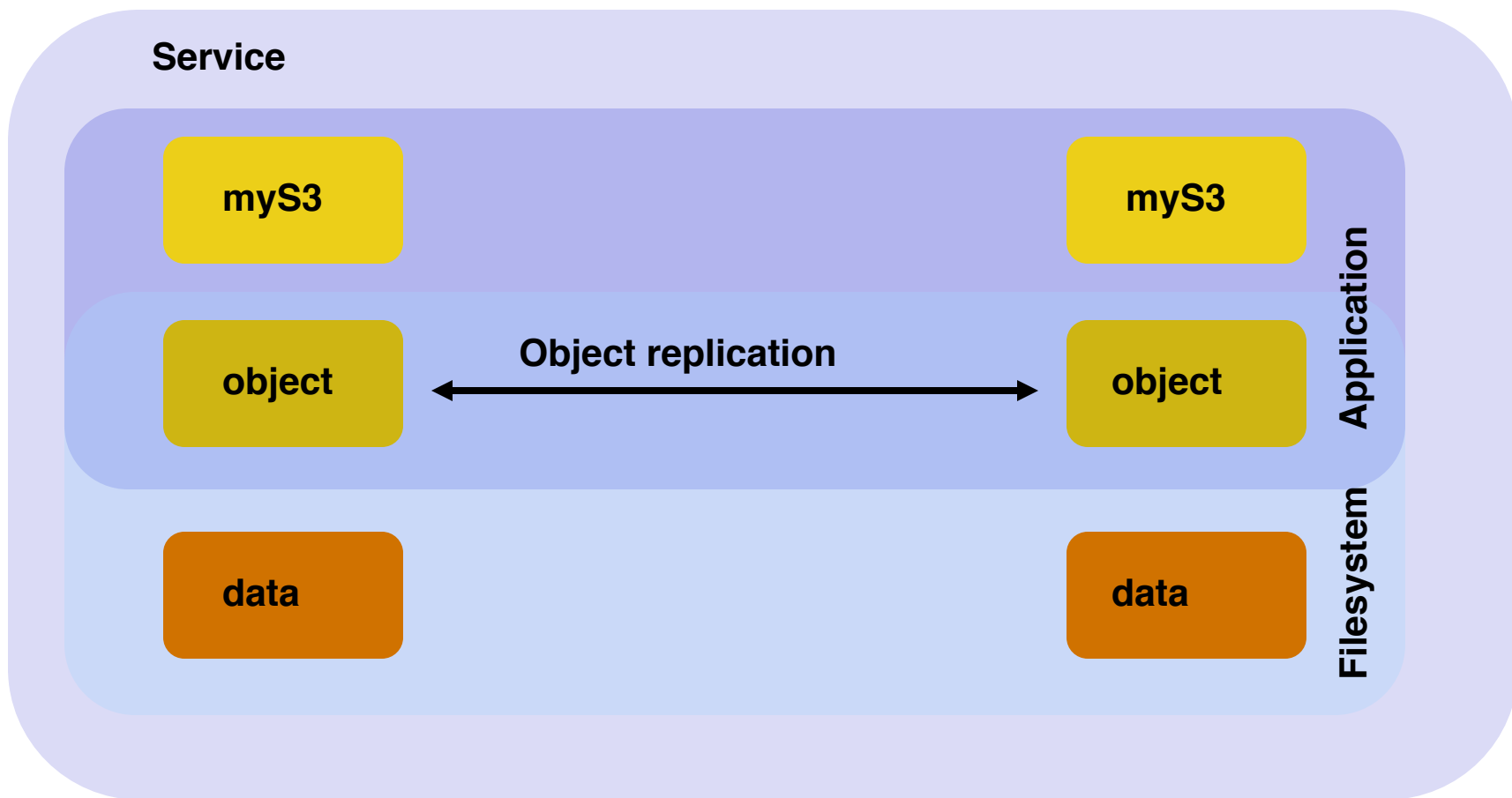
□ Network bandwidth

Short-term

- Better user and acl handling
- More cache triggers
- Encryption
- Compression

Middle-term

- Move s3fs to samba VFS plugin
- Locks
- Write only differences
- Split file in block (rsync model)

In the future

- Mixed local disk with cloud in a more intelligent way
- Software appliance
- Distributed across different cloud storages

5/11/11

Short-term

• Better user and acl handling
• More cache triggers
• Encryption
• Compression

Middle-term

• Move s3fs to samba VFS plugin
• Locks
• Write only differences
• Split file in block (rsync model)
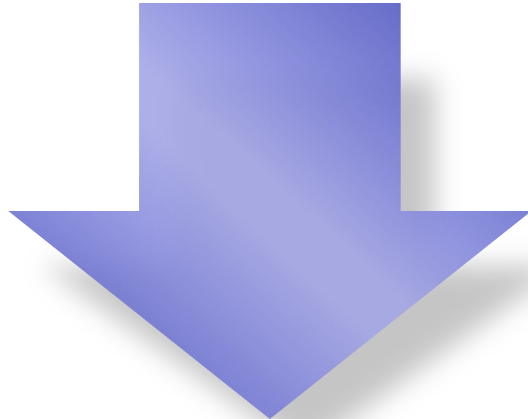
In the future

• Mixed local disk with cloud in a more intelligent way
• Software appliance
• Distributed across different cloud storages

5/11/11

5/11/11

Move from local server to cloud storage

No Change
- CIFS/NFS

Share
- WEB/Internet Disk

Disaster Recover
- Backup

5/11/11
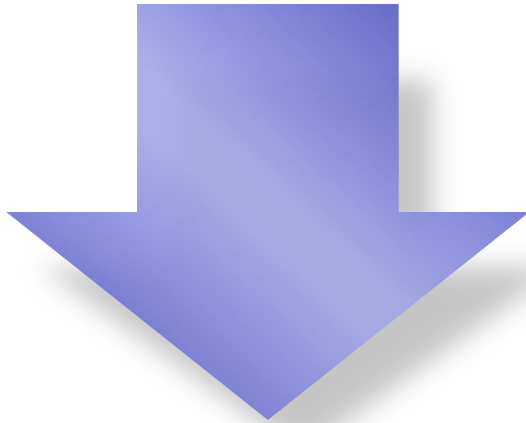
ZEROPIU
Difference as value

Advantage

- Pay for the storage used.
- No maintenance contracts, administration ..
- Very high levels of availability.
- Very good for Backups and Disaster recover
- You can test for Free
- Good for share (web site)
- Many reads, low write rate

Disadvantage

- Edit large file
- Network performance
- Concurrence
- Changing one bit is the same of all file
- Not Posix
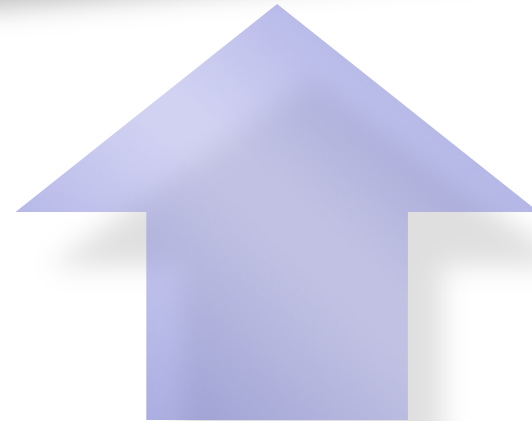
5/11/11

ZEROPIU
Difference as value
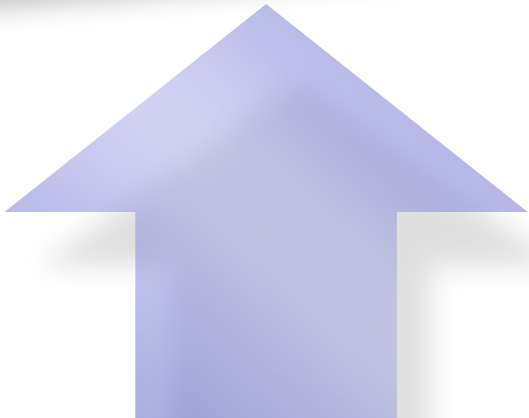
Advantage

- Low hw cost
- Very high levels of availability.
- You can test for Free
- Many reads, low write rate
- Good performance in read

Disadvantage

- Small file
- Not Posix
- Administration

5/11/11

Thank you

manfred@freemails.ch
manfred@zeropiu.it