

Deploying OpenStack Object Storage



Juan J. Martínez <jjm@usebox.net>
<http://www.usebox.net/jjm/>



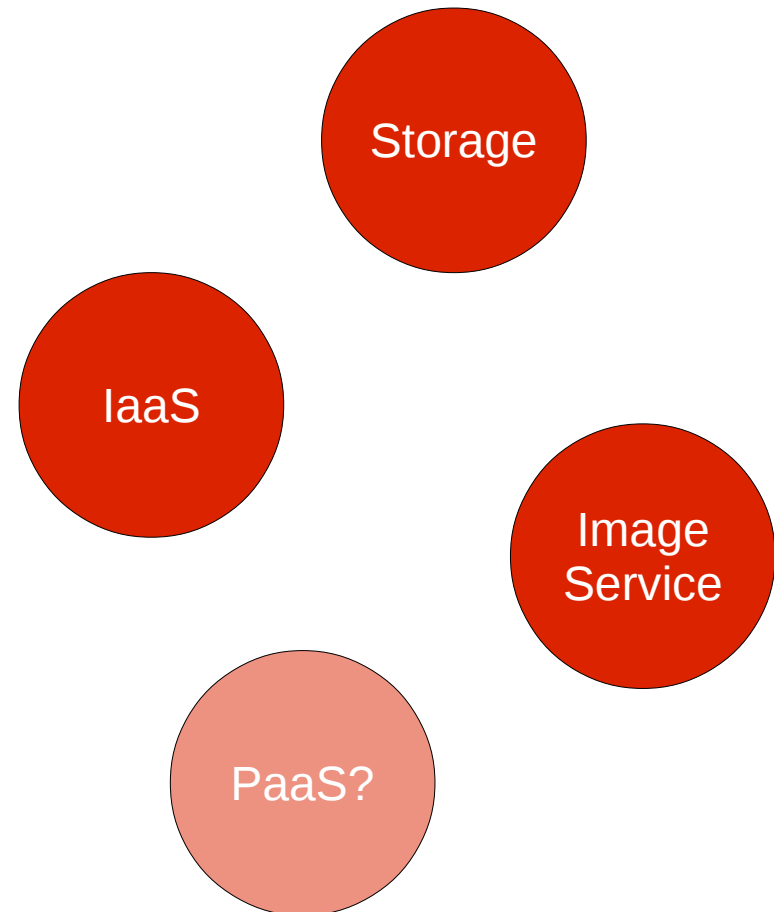
10-11 September 2011
University Club, Oxford

What is OpenStack?



Open Source software for “building clouds”

- OpenStack Object Storage (Swift)
- OpenStack Compute (Nova)
- OpenStack Image Service (Glance)
- Incubating projects:
Keystone, Dashboard, Crowbar, ...



What is Swift?



Open Source software for creating redundant, scalable **object storage** using clusters of standardized servers to store petabytes of accessible data.

Object Storage

- Not a real file system
- RESTful HTTP API
- Blob and metadata (name, content type, date-time, size, MD5)
- Support for pseudo-hierarchical directories

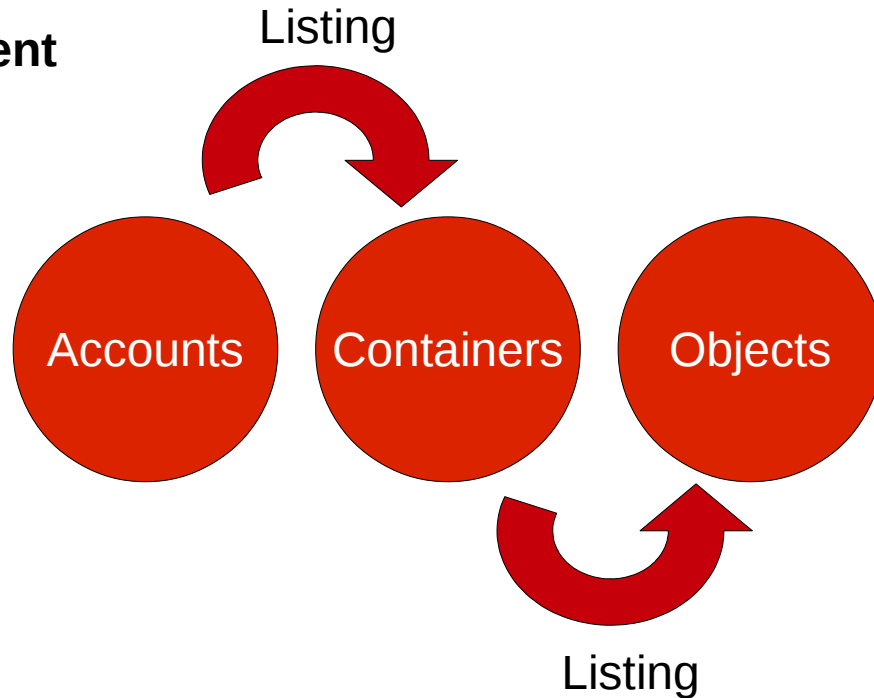
Real File System

- open, read, write, seek, close
- chdir, mkdir, rmdir
- remove, rename
- chmod, chown, access
- etc

Swift Architecture

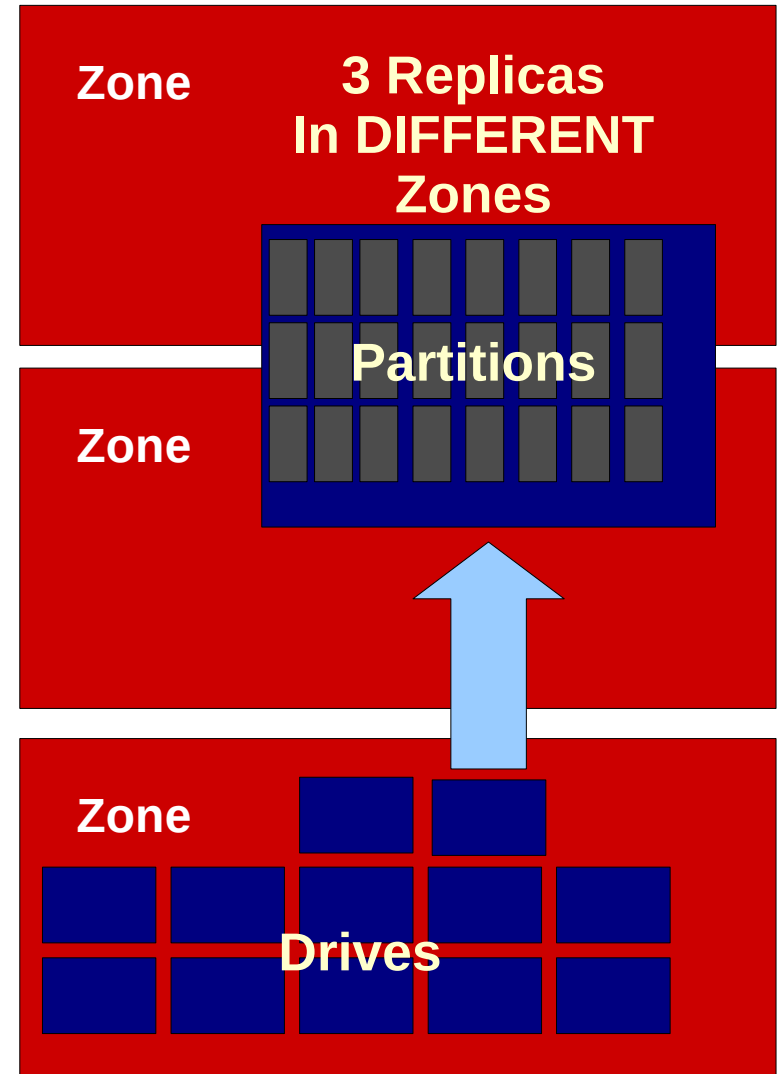


3 Different entities:



The Ring: location of an entity in the cluster

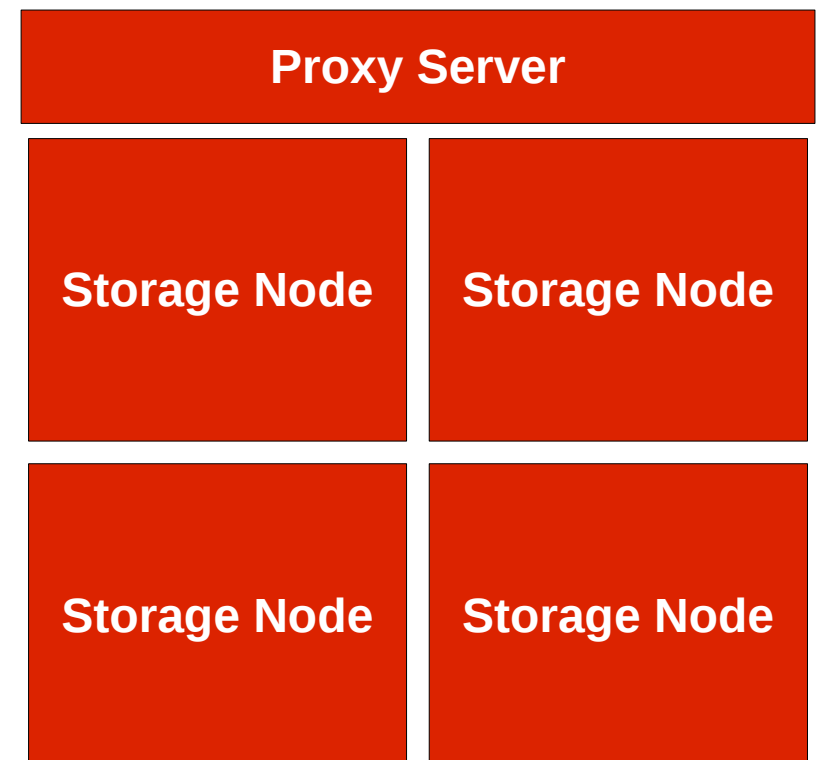
- Three rings (were intended for the Elves)
- Weights can be used to balance the distribution
- Three different logical levels



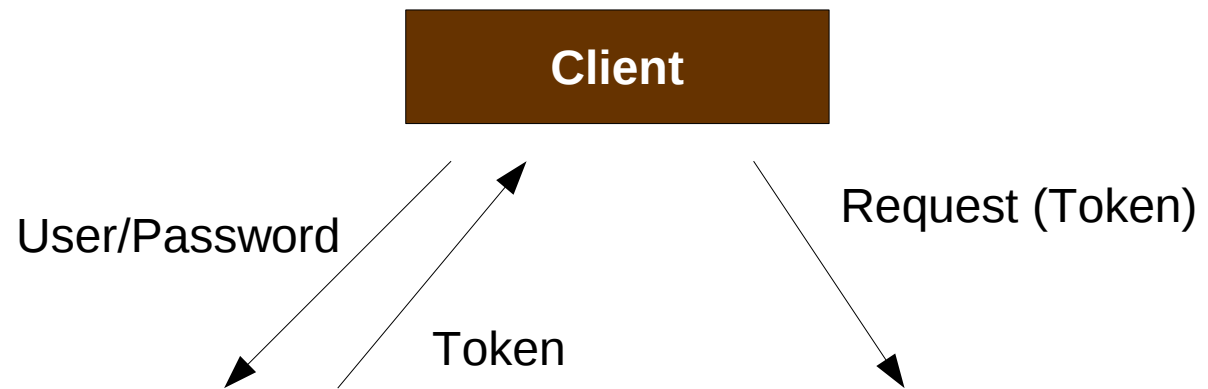
Swift Services



- proxy server
- account server
- container server
- object server
- replicators (rsync based)
- updaters (container, object)
- auditors
- reappers (account, container)

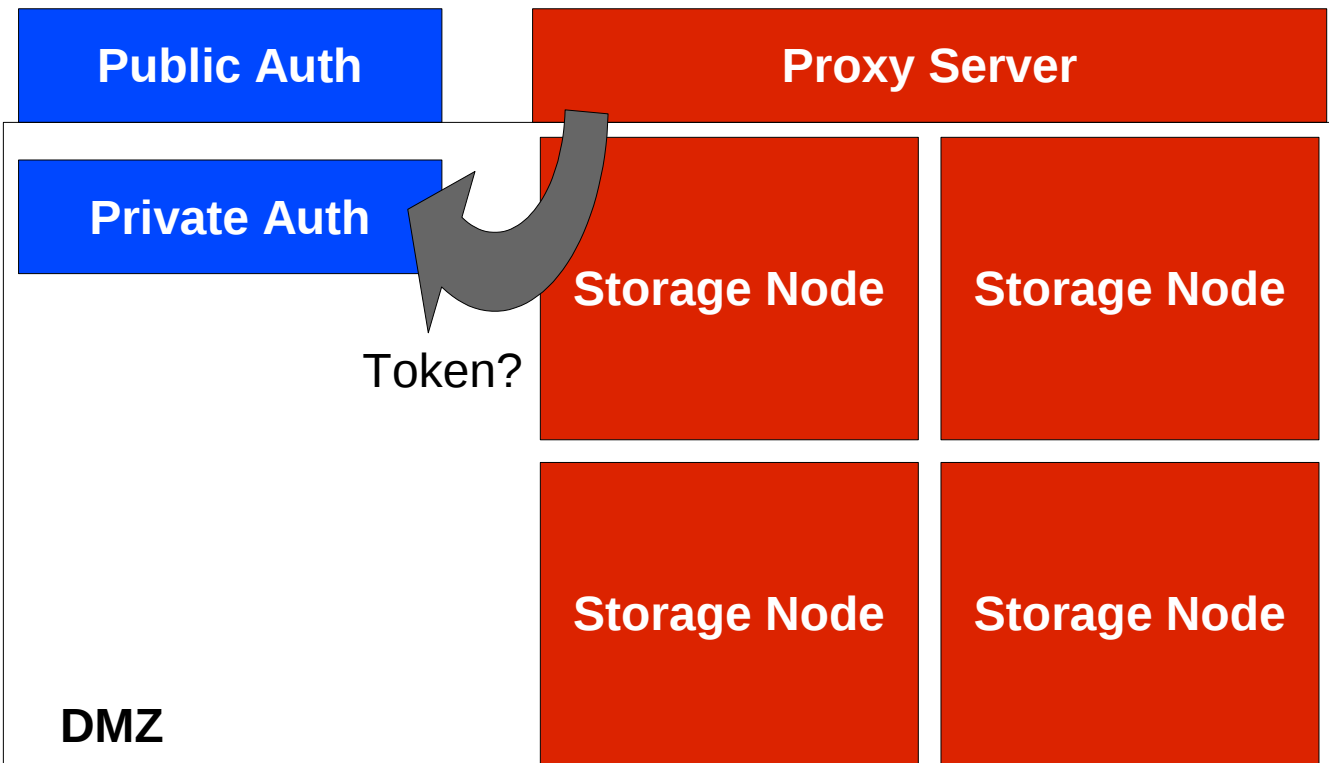


Swift Authentication (pluggable)

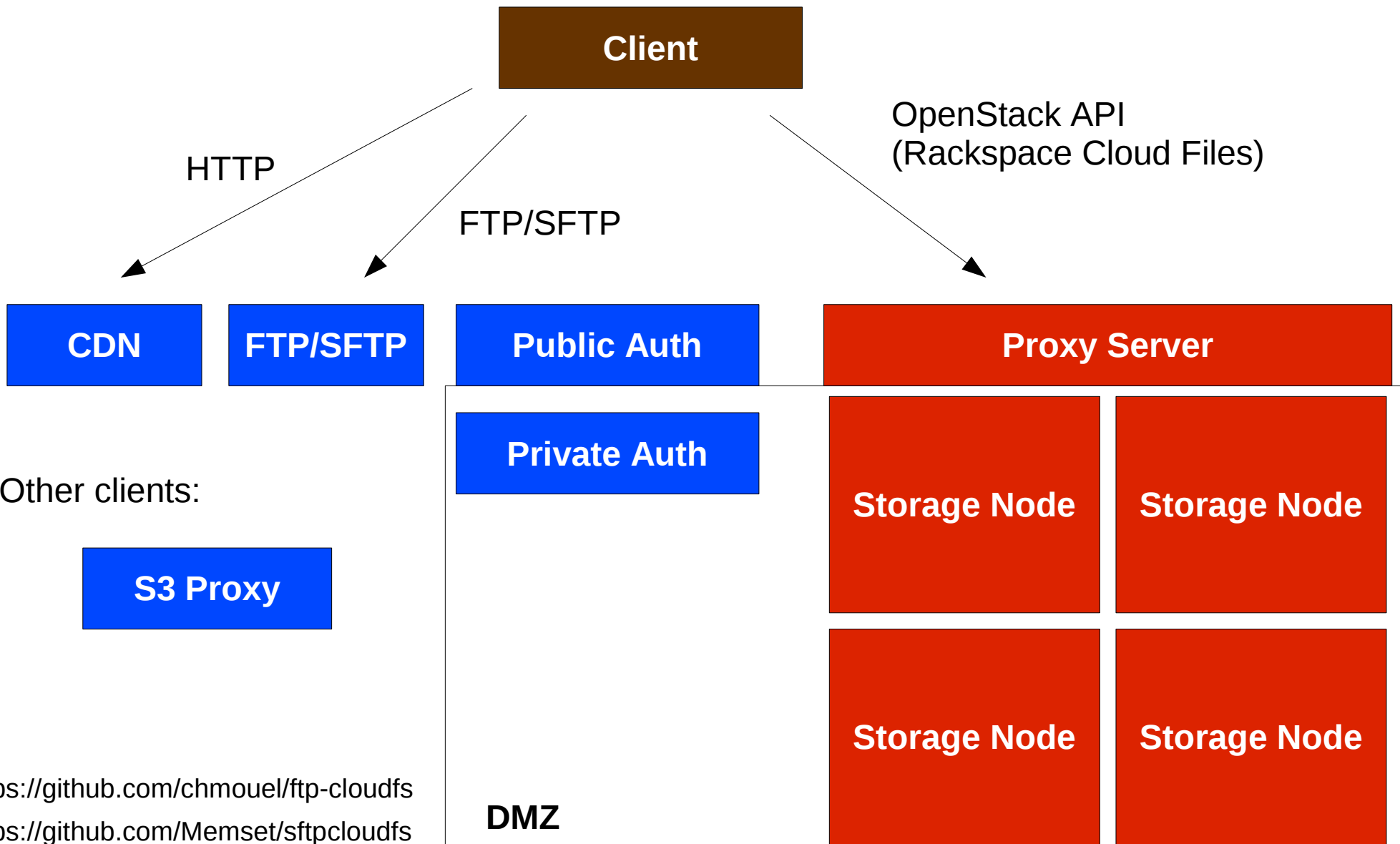


Different solutions:

- devauth (obsolete)
- swauth (external project)
- tempauth (testing)
- keystone (the future!)



Swift Interfaces



<https://github.com/chmouel/ftp-cloudfs>
<https://github.com/Memset/sftpcloudfs>

Swift Operations



- Managing the rings (adding/removing devices, zones, search for devices, rebalance the ring)
- Upgrading services (one zone at a time)
- Handling driver failure (unmount; optionally remove it from the ring, mount a new EMPTY drive)
- Zone failure (temporal: nothing!)
- Detecting failing disks (device audit)
- Object auditor (manually after a system crash)

What could possibly go wrong?

Any questions? Thanks for your attention!



Juan J. Martínez <jjm@usebox.net>
<http://www.usebox.net/jjm/>

