



## My CoRe: ownCloud at CNRS

David Rousse  
CNRS - DSI

## Content

- 1 Background and context
- 2 Service summary
- 3 Functional choices
- 4 Technical choices
- 5 Project feedbacks and roadmap
- 6 Appendixes

## Content

### 1 Background and context

### 2 Service summary

### 3 Functional choices

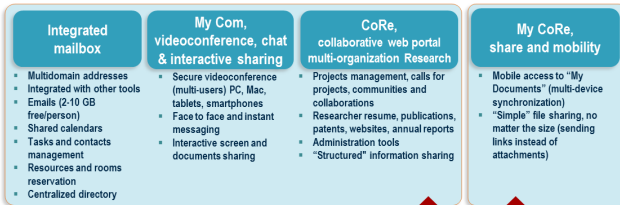
### 4 Technical choices

### 5 Project feedbacks and roadmap

### 6 Appendixes

# Global context

**Users “ecosystem”** : native integration of tools ➡ easy to use ...



“Structured” or “simple”  
file sharing, “DROPBOX like”

**Base infrastructure** : Secure and private Cloud for mutualisation ➡ source of savings...



**Helpdesk**

IT Teams  
in local  
offices



Local  
helpdesk  
(advice and IT  
support to the  
implementation of  
services)

\* Laboratories, administrative units and functional departments can benefit from these services

More information (in French) at <http://ods.cnrs.fr>

## P. 5 2 different business needs <sup>(identified mainly through end users polls)</sup>

- Synchronization and sharing service to provide a secure alternative to Dropbox for CNRS users
- Backup service for CNRS users' documents

## A unique choice

- ownCloud because it had the required functionality, it was already used in some local CNRS units with a good end users feedback and it was (and still is) open source (so easy to customize)
- Service deployed in CNRS' IN2P3 Computing Center, in order to keep the data safe and to be able to use the existing network bandwidth <sup>(10GB/s link)</sup> and also the local backup and restore service <sup>(based on Tivoli)</sup>
- Service operated by an IT service provider, to be able to provide it as a 24-7 service

### Schedule and deployment steps

- January to September 2013: market survey
- October 2013 to May 2014: ownCloud functional and technical evaluation <sup>(in collaboration with Linagora)</sup>
- May to December 2014: implementation
- January to September 2015: beta service for 2.000 end users (5GB per user)
- Since October 2015: service opened to all CNRS laboratories (20GB per user)

## Service summary

<b>Status:</b>	Production
<b>Number of users (current/target):</b>	4500 <sup>(at 2016/01/11)</sup> / 15.000 <sup>(for end 2016)</sup>
<b>Default quota:</b>	20GB
<b>Linux/Mac/Win user ratio:</b>	20/40/40 <sup>(estimated)</sup>
<b>Desktop/mobile/web clients access ratio:</b>	Unknown yet
<b>Technology:</b>	ownCloud <sup>(web)</sup> with MariaDB-Galera <sup>(DB)</sup> and Scality <sup>(storage)</sup>
<b>Target communities:</b>	CNRS members
<b>Integration in current environment:</b>	None <sup>(except our existing Shibboleth SSO backend)</sup>
<b>Risk factors:</b>	Load on DB
<b>Most important functionality:</b>	Files synchronization and sharing, central files backup
<b>Missing functionality:</b>	Versioning, accounts for external users, sharing files temporary and then delete theses files (once downloaded)

## Content

1 Background and context

2 Service summary

**3 Functional choices**

4 Technical choices

5 Project feedbacks and roadmap

6 Appendixes



### ownCloud

- ☐ ownCloud core, community edition<sup>(version 7)</sup>
- ☐ Antivirus app
- ☐ Without Versions app<sup>(too much load generated on DB)</sup>

### And some apps developed/forked by CNRS

- ☐ "Dashboard" app for metrics on service usage<sup>(current metrics in appendix 3)</sup>
- ☐ "Lotsofgroups" app for managing a lot of groups
- ☐ "Group custom" app for end users group management
- ☐ "Password Policy" app for password policy enforcement
- ☐ "GTU" app for GTU online agreement

*Theses developments have been made on the ownCloud server part, nothing has been yet developed on the clients side  
Source code is available on Github at <https://github.com/CNRS-DSI-Dev/>*

### And some apps developed/forked by CNRS <sup>(continued)</sup>

- ☐ "Gatekeeper" app for filtering access depending on end user groups
- ☐ "User Servvars 2" app for end users authenticate and account provisionning <sup>(in relation with Shibboleth SSO backend)</sup>
- ☐ "User account actions" developed for performing actions when users are created/deleted
- ☐ "User files migrate" app for transferring files between two accounts of a unique user
- ☐ "User files restore" app and shell scripts for managing backup and restore of end users files <sup>a</sup>
- ☐ Shell script "Mycore sympa" to integrate ownCloud users list to an SYMPA mailing list manager
- ☐ A specific theme
- ☐ Various housekeeping shell scripts

*These developments have been made on the ownCloud server part, nothing has been yet developed on the clients side  
Source code is available on Github at <https://github.com/CNRS-DSI-Dev/>*

---

*a. See detail on appendix 2 of this document*

## Content

1 Background and context

2 Service summary

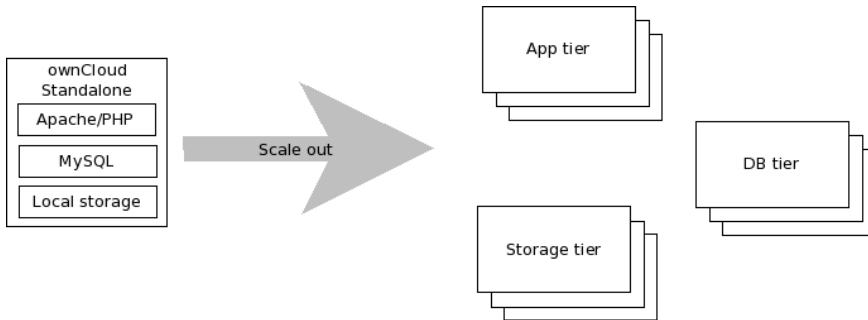
3 Functional choices

**4 Technical choices**

5 Project feedbacks and roadmap

6 Appendixes

## Initial key points



## 2 key points

- ☐ Is ownCloud scalable and if so how (how much web, DB nodes do we need, ...)?
- ☐ Which components to choose in order to implement the service?

## Method

- ☐ Functional hypothesis on the service usage
- ☐ Deployment of a dedicated infrastructure for theses load tests studies
- ☐ Technical hypothesis on the web, DB and storage behaviours

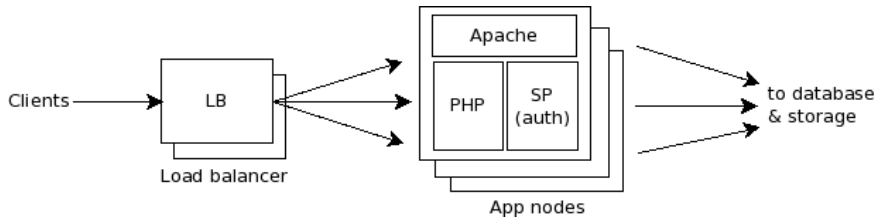
See in detail at [https://github.com/CNRS-DSI-Dev/mycore\\_press/raw/master/CERN-CNRS-meeting-20141117.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/raw/master/CERN-CNRS-meeting-20141117.pdf), appendix 1

## Results

- ☐ "Go" to deploy such a service, with an accurate visibility on the architecture evolution  
(number of nodes, number of servers, ... depending on number of users, number of devices per user, ...)
- ☐ Main concern: cost on the long term for such a service
- ☐ Functional hypothesis can have a huge impact on sizing

See in detail at [https://github.com/CNRS-DSI-Dev/mycore\\_press/raw/master/CERN-CNRS-meeting-20141117.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/raw/master/CERN-CNRS-meeting-20141117.pdf), appendix 2

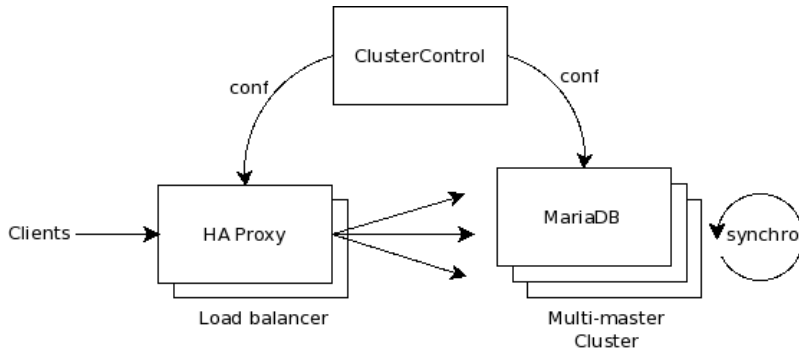
## App tier choice



### ownCloud, PHP, Shibboleth's SP and Apache choices

ownCloud for business reasons, Apache because of internal teams knowledge, Shibboleth's SP on the web nodes in order to have the same web nodes

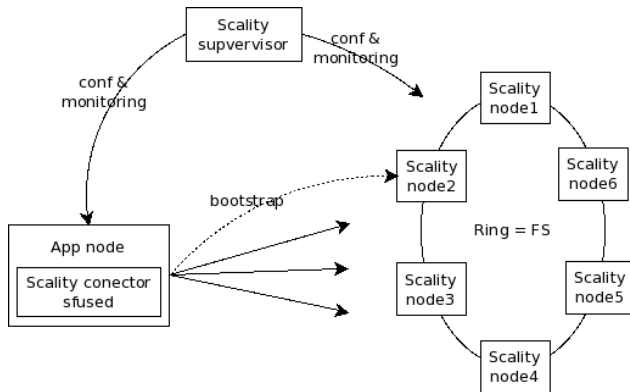
## DB tier choice



### MariaDB-Galera cluster choice

open source<sup>(except ClusterControl tool, detail at <http://severalnines.com/product/clustercontrol>)</sup>, 32 nodes supported, already known internally and easy to install and administrate through ClusterControl tool

## Storage tier choice



## Scality software defined storage choice

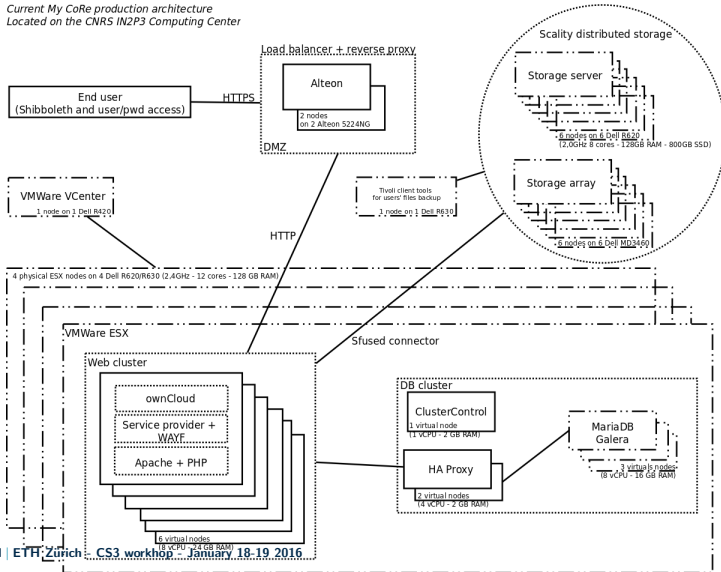
Reliable file system access, hardware agnostic solution and scalable by design

See in detail at [https://github.com/CNRS-DSI-Dev/mycore\\_press/raw/master/CERN-CNRS-meeting-20141117.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/raw/master/CERN-CNRS-meeting-20141117.pdf), appendix 3



## Current architecture

Current My CoRe production architecture  
Located on the CNRS IN2P3 Computing Center



## Content

1 Background and context

2 Service summary

3 Functional choices

4 Technical choices

**5 Project feedbacks and roadmap**

6 Appendixes

### Positive points

- ☐ Useful project, so real motivation to make it a success
- ☐ Very positive end user feedbacks
- ☐ Meaningful exchanges between project team members and with other partners (especially CNRS' IN2P3 Computing Center)

### Negative points

- ☐ Hard to conciliate the needs for a "Dropbox like service" and for a pure files backup and restore service
- ☐ Project length too long (over 2 years!), mainly due to difficulties encountered during implementation
- ☐ An agile process instead of a classical V-model approach would have been a better choice to reduce the project length
- ☐ Deploying a dedicated backend (web, DB, storage) just for My CoRe was time consuming : re-using an IaaS platform and building the My CoRe service on top of that would have been more appropriate

## Functional roadmap

- ☐ App for setting local password just after the first login <sup>a</sup>
- ☐ ownCloud 8.2 migration
- ☐ "Share manager" app for an easy management of the files shares <sup>b</sup>
- ☐ Guest accounts <sup>c</sup>
- ☐ App Versions activated
- ☐ Enhanced end users administration backend
- ☐ Users UID migration, because of new CNRS referential "Réséda"
- ☐ Specific mobile and PC clients <sup>d</sup>

<sup>a</sup>. See detail at [https://github.com/CNRS-DSI-Dev/user\\_set\\_password](https://github.com/CNRS-DSI-Dev/user_set_password)

<sup>b</sup>. See detail at <https://github.com/LydSC/sharewatcher>

<sup>c</sup>. Accounts for non-CNRS members, with a 0GB quota, in order to ease collaborations between CNRS users and external CNRS users

<sup>d</sup>. For instance to have a unique access through our Shibboleth SSO, whatever the client is

## Technical roadmap

- Architecture evolution
  - ▶ basic option with more web and/or DB nodes, see detail at = [https://github.com/CNRS-DSI-Dev/mycore\\_press/raw/master/architecture/schema\\_2016.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/raw/master/architecture/schema_2016.pdf)
  - ▶ alternative option with DB migration from virtual to physical nodes = [https://github.com/CNRS-DSI-Dev/mycore\\_press/raw/master/architecture/schema\\_2016\\_option.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/raw/master/architecture/schema_2016_option.pdf)
- Access storage backend through an object interface instead of a file system access
- Try to investigate the "Server to server sharing" ownCloud function
- End-to-end testing tool, to test the ownCloud core functionalities <sup>a</sup>
- File encryption (on server side)

---

<sup>a</sup>. For instance with the CERN' smashbox tool

## Questions?



### Projet team (previous and current members)

David Bercot (IT department manager) Jonathan Bouchiquet (end users support), Marc Dexet (developer), Philippe Dubrulle (communication), Gilian Gambini (system administrator), Eric Gervasoni (end users committee manager), Jérôme Jacques (system administrator), Nadine Marouzé (previous IT department manager), Paulo Mora de Freitas (end users committee manager), Olivier Lenormand (IT coordinator), Jean-Yves Lopez (previous IT department manager), Patrick Paysant (developer), David Rousse (previous project leader), Lyderic Saint-Criq (developer) et Alexandre Salvat (project leader<sup>(since 2015/11)</sup>)

### Partners involved in the project

DSI' CNRS IT department, DSI' CNRS graphist, CNRS' IN2P3 Computing Center and the following external partners Atos, Dell, Linagora, Scalify and SeveralNines

## Content

- 1 Background and context
- 2 Service summary
- 3 Functional choices
- 4 Technical choices
- 5 Project feedbacks and roadmap
- 6 Appendixes**

### P. 2 URLs in relation with My CoRe

- Previous My CoRe presentation during 1<sup>st</sup> CERN workshop =  
[https://github.com/CNRS-DSI-Dev/mycore\\_press/raw/master/CERN-CNRS-meeting-20141117.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/raw/master/CERN-CNRS-meeting-20141117.pdf)
- My CoRe presentation at JRES conference on Dec 2015 (in French) =  
[https://github.com/CNRS-DSI-Dev/mycore\\_press/tree/master/JRES2015](https://github.com/CNRS-DSI-Dev/mycore_press/tree/master/JRES2015)
- ownCloud load test in detail = [https://github.com/CNRS-DSI-Dev/mycore\\_press/blob/master/CERN-CNRS-meeting-20140513.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/blob/master/CERN-CNRS-meeting-20140513.pdf)
- Comparison between theoretical load tests and real load after beta period service =  
[https://github.com/CNRS-DSI-Dev/mycore\\_press/raw/master/myCore\\_comparison\\_estimate-load-real-load\\_on\\_ownCloud.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/raw/master/myCore_comparison_estimate-load-real-load_on_ownCloud.pdf)
- JoSy conference (in French), Strasbourg 2014 May =  
[https://github.com/CNRS-DSI-Dev/mycore\\_press/blob/master/CNRS-JoSy-20140519.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/blob/master/CNRS-JoSy-20140519.pdf)
- Scalify in detail, press made for the CNES (in French) = [https://github.com/CNRS-DSI-Dev/mycore\\_press/blob/master/CNES-CNRS-Scalify-20140619.pdf](https://github.com/CNRS-DSI-Dev/mycore_press/blob/master/CNES-CNRS-Scalify-20140619.pdf)
- My CoRe, how is it built? = [https://github.com/CNRS-DSI-Dev/mycore\\_build](https://github.com/CNRS-DSI-Dev/mycore_build)
- Other My CoRe press to done/to come = [https://github.com/CNRS-DSI-Dev/mycore\\_press](https://github.com/CNRS-DSI-Dev/mycore_press)



## Appendix 2: details on backup/restore function <sup>(1/2)</sup>

### End users files backup in My CoRe service

- ☐ The backup is made through a shell script that interfaces My CoRe infrastructure and Tivoli backup service = [https://github.com/CNRS-DSI-Dev/mycore\\_backup\\_restore\\_user\\_files/blob/master/backup\\_mycore.sh](https://github.com/CNRS-DSI-Dev/mycore_backup_restore_user_files/blob/master/backup_mycore.sh)
- ☐ The backup is scheduled in background, without end users interaction: files are backuped in background and stored by the CNRS' IN2P3 Computing Center Tivoli service

### End users files restore in My CoRe service

- ☐ The restore is also made through a shell script that interfaces My CoRe infrastructure and Tivoli backup service = [https://github.com/CNRS-DSI-Dev/mycore\\_backup\\_restore\\_user\\_files/blob/master/mycore\\_restore.sh](https://github.com/CNRS-DSI-Dev/mycore_backup_restore_user_files/blob/master/mycore_restore.sh)
- ☐ An ownCloud app allows end users to ask for files restore = [https://github.com/CNRS-DSI-Dev/user\\_files\\_restore](https://github.com/CNRS-DSI-Dev/user_files_restore)

### Global principle of a restore

- ☐ An end user can ask for a file/folder restore, at day-1, day-6 and day-15. Once restored, the restored file/folder overwrites the existing one (if exists)
- ☐ Restore requests are not processed in real time: instead, each request is written in a DB table and a scheduled shell script reads this table and execute the restore in background
- ☐ A user can cancel a pending restore request

## Appendix 2: detail on backup/restore function <sup>(2/2)</sup>

Files

All files

Shared with you

Shared with others

Shared by link

New

Upload

schema\_odu\_global.png

1 file

Download

Restore

Share

140 KB

seconds ago

1 day ago

Restore

140 KB

6 days ago

Restore

15 days ago

Restore

Restore

Restoration requests

Informations préalables

- If you want to restore a mistakenly deleted files, have a look on trashbin, it's simpler and quicker.
- You also can request a restoration on the application default page.
- Restoration requests are processed as background jobs. You can find on this page the state and results of your requests.
- Important:** restored files will override (ie delete and replace) existing files with same path and name.
- If you have been targeted by a crypto-virus, please ask your CSSI / RSSI before creating any restoration request.

Request a complete restoration

1 day

Validate

To be processed

/ (dir ; d - 1)

Cancel

Running

(dir) /test

(dir) /test

Done

(dir) /rec15mtt

## Appendix 3: detail on service usage (screenshot of Dashboard app on 2016/01/11)

CS3 CORE Dashboard ▾

### Dashboard



**Disk space**

**3.7 TB**

User data dir: /scality/owncloud/data  
Default quota per user: 20 GB



**Users**

**4571**

Size / user: 852.6 MB  
Files / user: 1,005.72 (standard deviation: 7,381.44)  
Folders / user: 100.73 (standard deviation: 715.05)



**Folders**

**460418**

Size / folder: 8.5 MB  
Files / folder: 9.98



**Files**

**4597130**

Size / file: 868 kB



**Shares**

**4319**

Shares / user: 0.94  
Standard deviation: 7,381.44

nbUsers ▾

Last month ▾

none ▾

