

Mise en place d'ownCloud au sein de l'université de Lorraine

Benjamin SECLIER

Direction du numérique - Sous-direction infrastructures
Université de Lorraine
24 rue Lionnois
54000 Nancy

Résumé

L'université de Lorraine a fait le constat que de nombreux enseignants-chercheurs utilisent des services de synchronisation et de partage de fichiers de type Dropbox ou Google Drive. Dans une optique de confidentialité et de conservation des données scientifiques et techniques, la direction du Numérique a souhaité mettre en place ce type de service au sein de notre université.

Après avoir étudié plusieurs solutions de synchronisation/partage de documents, la solution logicielle ownCloud a été retenue et déployée à titre expérimental dans un premier temps (fin 2014), puis ouvert à l'ensemble du personnel en septembre 2015.

Dans cette communication, je présenterai les différentes étapes de la mise en place de cet outil au sein de notre université :

- 1. la mise en place de l'infrastructure nécessaire pour ce service. Seront présentés, plus spécifiquement, les différents serveurs nécessaires ainsi que les choix de fonctionnement comme la gestion des sessions, centralisation des logs, etc*
- 2. la présentation des plugins owncloud (« apps ») utilisés dans notre établissement*
- 3. notre stratégie d'ouverture progressive du service à l'ensemble du personnel de l'université (7000 personnes)*
- 4. les outils/aides mis en place lors de l'ouverture du service tels qu'une documentation exhaustive, un screencast ou encore une enquête de satisfaction (je présenterai les résultats).*

Des chiffres tels que l'évolution des inscriptions ou encore l'espace moyen utilisé serviront d'illustration à la présentation.

Mots-clefs

Synchronisation, Stockage, Partage, Cloud, Drive, ownCloud

1 Contexte

L'université de Lorraine, créée au 1^{er} janvier 2012, est un grand établissement à caractère scientifique, culturel et professionnel, qui regroupe 8 collègiums (44 composantes d'enseignement) et 10 pôles de recherche (67 laboratoires). Composée de 3700 enseignants-chercheurs et enseignants, 3000 personnels administratifs et techniques et 52 000 étudiants, elle est répartie sur l'ensemble de la Lorraine.

Au cours du premier semestre 2014, le projet NumiLab a été initié afin de déterminer les besoins en services numériques pour les laboratoires de recherche. Au cours des entretiens menés avec chaque laboratoire, des besoins importants ont été identifiés dans le domaine du stockage et de la sauvegarde des données. Ils ont fait ressortir des pratiques dangereuses pour le stockage des données scientifiques (résultats de recherches, d'expérimentations, de calculs, etc.) :

- données stockées sur les postes de travail ;
- données sauvegardées sur des disques durs externes personnels ;
- données stockées sur des services en ligne de type Dropbox (problème de propriété des données) ;
- données stockées par certains laboratoires sur des infrastructures de stockage (serveurs, baies) avec des niveaux de fiabilité plus ou moins satisfaisants.

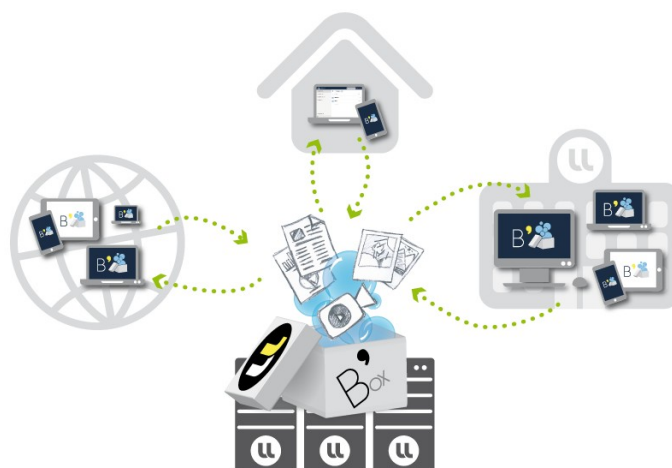
La façon de gérer les données étant très variable d'un laboratoire et d'un individu à l'autre, nous avons constaté qu'il y a souvent confusion entre stockage des données et sauvegarde du poste de travail. Il convient donc d'appréhender cette problématique dans son ensemble et de différencier les besoins (administratifs, personnels, données scientifiques, partages ...) afin d'apporter des réponses ciblées à chaque cas d'utilisation.

L'université disposant déjà de services pour le stockage des données administratives, ainsi que pour la sauvegarde des postes de travail, il a été décidé de mettre en œuvre un service de type « drive » afin de permettre aux chercheurs :

- de stocker leurs fichiers de travail sur des serveurs de l'université, et ainsi éviter l'utilisation de service en ligne du type DropBox ;
- de partager leurs données de travail, avec d'autres des personnels de l'université ou d'autres chercheurs dans le monde.

Le besoin d'un service de type « drive » n'étant pas spécifique à la recherche, nous avons dès le début de ce projet eu comme objectif de fournir un service aux laboratoires mais également à l'ensemble du personnel de l'université de Lorraine pour septembre 2015.

La mise en place de ce projet a été possible grâce à plusieurs personnes au sein de différentes sous-directions de la direction du numérique. Les décisions « fonctionnelles » (choix des fonctionnalités à retenir ou à exclure, délais de mise en production, etc.) ont été validées par mon chef de service et mon sous-directeur. Les images, design et vidéo de promotion ont été réalisés par l'équipe « audiovisuel et multimédia ». J'étais pour ma part chargé des tests, de la mise en place technique du projet (décrite dans cet article) ainsi que de la rédaction de la documentation.



2 Évaluation et choix d'une solution

J'ai commencé le recensement et l'évaluation des outils de type « Drive » en fin d'année 2013. Notre choix se dirigeait plutôt vers des outils open-source, ce que nous avons l'habitude de privilégier aux solutions propriétaires. J'ai retenu et testé quatre solutions libres : Seafile, Pydio, ownCloud et SparkleShare.

J'ai ensuite créé une grille d'évaluation recensant les caractéristiques indispensables (ou fortement souhaitables) que nous devrions retrouver dans la solution que nous retiendrions. Voici ces caractéristiques :

- Authentification LDAP (au minimum), authentification CAS (au mieux)
- Interface Web (indispensable)
- Accès hors-ligne à ses documents (indispensable)
- Mise en cluster possible (indispensable)
- Langage de programmation facilement compréhensible et modifiable (fortement souhaité)

Voici le résultat de ma grille d'évaluation :

	Seafile	Pydio	ownCloud	SparkleShare
<i>Authentification LDAP</i>	Non	Oui (plugin)	Oui (de base)	Non
<i>Authentification CAS</i>	Non	Oui (plugin)	Oui (plugin)	Non
<i>Interface web</i>	Oui	Oui	Oui	Non
<i>Accès hors-ligne</i>	Oui	Non	Oui	Oui
<i>Passage à l'échelle</i>	Non	Oui	Oui	Difficilement
<i>Langage</i>	Python	PHP	PHP	Mono

SparkleShare a rapidement été écarté. Son langage mais surtout l'absence d'interface web (et donc d'accès à ses fichiers ailleurs que sur l'ordinateur où est installé le client de synchronisation) a été rédhibitoire.

Pydio ne disposait pas (en 2013) d'outil de synchronisation des données, qui n'étaient donc accessibles qu'en ligne depuis une interface web. Pas d'accès à ses documents sans connexion à Internet. Notons qu'un client de synchronisation (PydioSync) est maintenant disponible depuis fin 2014.

Seafile était un projet fonctionnel (la synchronisation multi-OS fonctionnait correctement) mais sa jeunesse se ressentait dans l'absence de système d'authentification facilement adaptable à notre SI, rendant son intégration difficile (en particulier en visant une authentification CAS par la suite).

OwnCloud, quant à lui, regroupait les fonctionnalités de base attendues d'un outil de « Drive » à intégrer au sein de notre établissement. Par ailleurs, nous avons déjà quelques retours d'expériences dans le monde de l'enseignement supérieur et de la recherche où le déploiement d'ownCloud avait été un succès (projet ownCore au CNRS, mise en place dans les universités Lille 3, Aix-Marseille, Toulouse 3, Lyon, etc). De plus, l'application étant écrite en PHP, la modification du code pour l'adapter à nos besoins serait grandement facilitée.

Pour ces différentes raisons, nous avons choisi de mettre en place ownCloud au sein de l'université de Lorraine.

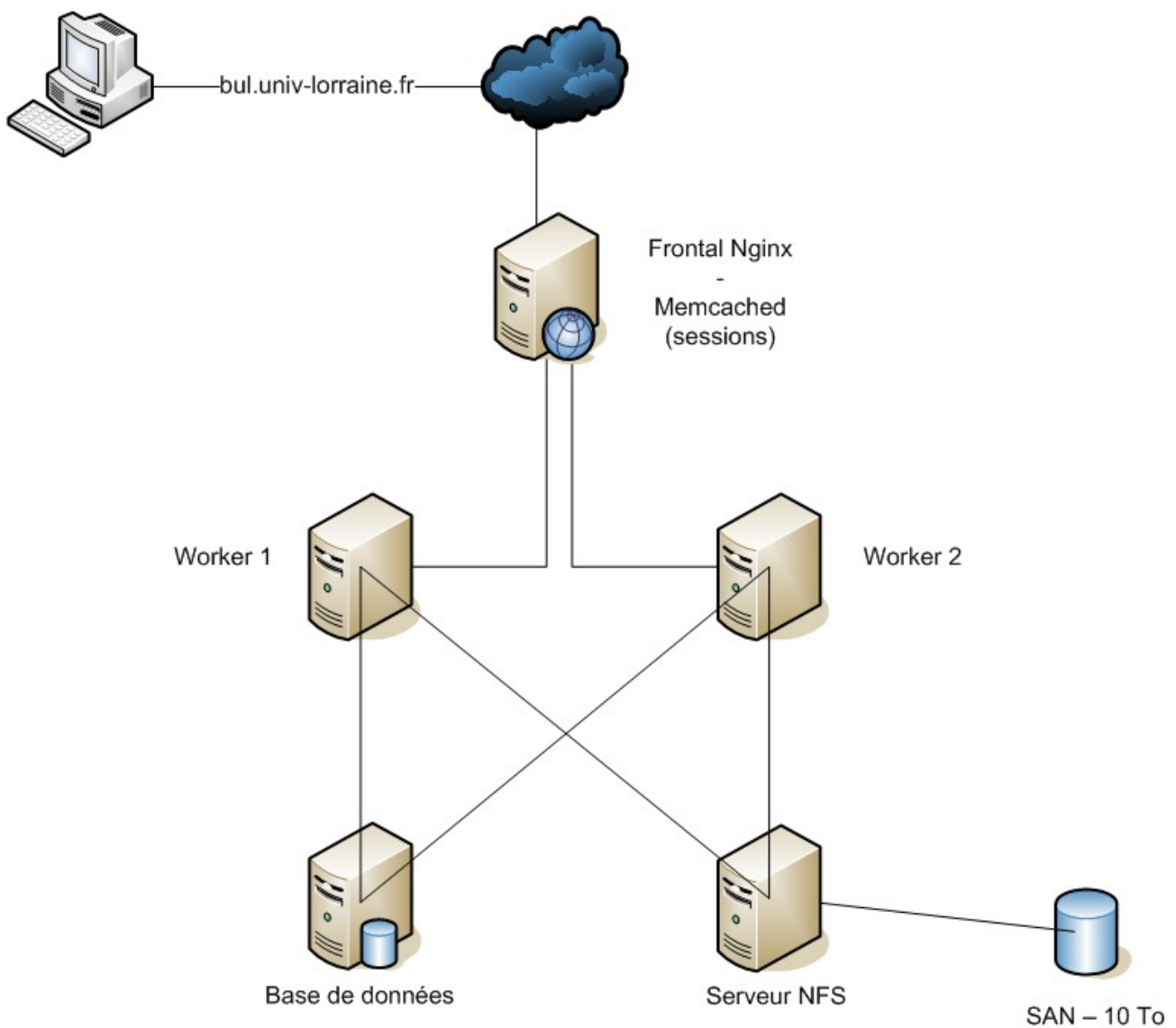
3 Mise en place de l'infrastructure

3.1 Vue d'ensemble

J'ai commencé la mise en place de l'infrastructure ownCloud lorsque le choix de cette solution a été validé en juillet 2014. J'avais déjà une assez bonne expertise des infrastructures d'hébergement d'application web permettant de faire de la tolérance de panne ou de la répartition de charge. Nous utilisons ce schéma pour nos infrastructures d'hébergement de l'ENT ou de notre système de cours en ligne Moodle. Le choix des applications/services pour l'infrastructure d'ownCloud et leur mise en place se sont donc faits assez rapidement.

J'ai mis en place le schéma suivant : un frontal web dans un VLAN public, deux machines hébergeant le code PHP d'ownCloud (workers), un serveur de base de données et un serveur de stockage de fichiers (NFS) dans un VLAN privé. Toutes les machines sont des machines virtuelles hébergées dans notre cluster VMWare.

Voici le schéma de l'infrastructure actuelle :



3.2 Frontal Web

Nginx est un logiciel de serveur Web et de **reverse proxy** très robuste. Les études réalisées s'accordent à dire qu'il est bien plus performant (consommation mémoire et requêtes par seconde traitées¹) que son principal concurrent Apache, en particulier quand il s'agit de faire du reverse proxy. C'est pour cette dernière raison que j'ai décidé d'utiliser Nginx dans notre infrastructure ownCloud. J'ai choisi le mode de fonctionnement standard de reverse proxy, c'est à dire que chaque requête est envoyée alternativement sur un worker puis sur l'autre, sans considération de charge des workers sous-jacents ni de sessions (sticky sessions).

Je me suis également servi de cette machine pour y stocker les **sessions des utilisateurs** grâce à memcached. Ainsi la session d'un client est stockée de manière centralisée sur le frontal et ne dépend donc pas du worker sur lequel il arrive.

Enfin, les **logs** sont centralisés sur cette machine. Qu'il s'agisse des logs Nginx, Apache (workers) ou applicatifs (ownCloud), toutes sont stockés dans un même répertoire, ce qui facilite grandement leur consultation.

Détail matériel et logiciel : 2 CPU, 2 Go de RAM ; CentOS 6.7, Nginx 1.8

3.3 Workers

Cette infrastructure est composée de deux machines (workers). Ces workers accueillent le code source de l'application ownCloud (une instance de test et une instance de production). Ces répertoires sont synchronisés automatiquement toutes les nuits et un script a été développé pour forcer une synchronisation en journée (en cas de modification du code/mise à jour la journée). Un même volume NFS est monté sur chacun de ces deux workers (voir Serveur NFS) afin que chacun ait la même vision des répertoires stockant les données brutes.

Détail matériel et logiciel : 4 CPU, 4 Go de RAM ; Apache 2.2, PHP 5.3

3.4 Serveur de base de données

Le serveur de base de données est déporté afin de permettre l'accès aux bases depuis les deux workers. Ainsi, si un worker s'arrête ou n'est plus disponible, l'accès aux bases est encore possible pour l'autre worker.

Détail matériel et logiciel : 4 CPU, 4 Go de RAM ; MariaDB 5.5

3.5 Serveur NFS

Le serveur NFS est branché sur un montage SAN de 10 To. Ce montage est exposé sur les deux workers, de sorte que les opérations effectuées sur les fichiers par un worker sont immédiatement visibles sur l'autre.

Détail matériel et logiciel : 2 CPU, 2 Go de RAM

¹ http://wiki.dreamhost.com/Web_Server_Performance_Comparison

4 Adaptation du service aux besoins de l'université

4.1 Désactivation de certaines fonctionnalités

OwnCloud est fourni avec de nombreux plugins (« apps ») intégrés. Quelques plugins étaient redondants avec certains services rendus par d'autres applications à l'université tandis que d'autres nous ont semblé peu intéressants/utiles. Voici la liste des fonctionnalités que nous avons décidé de désactiver :

- **Calendrier** : La messagerie de l'université fonctionne avec la suite Zimbra qui, dans son interface web, propose un calendrier qui est très utilisé. Nous n'avons pas souhaité faire de doublons avec ownCloud, c'est pourquoi cette fonctionnalité a été désactivée.
- **Carnet d'adresses** : Cette option a été désactivée pour les mêmes raisons que le calendrier. En effet, Zimbra propose également un outil de gestion de contact assez élaboré (et en général déjà synchronisé avec les périphériques mobiles du personnel de l'université).
- **Documents** : La fonctionnalité d'édition de documents bureautiques de type Word, Excel ou LibreOffice n'est pas encore complètement opérationnelle. Nous avons rencontré un certain nombre de problèmes de compatibilité/affichage lors de l'édition de ces documents. Par précaution, nous avons décidé de la désactiver. Il y a cependant toujours la fonctionnalité d'édition en ligne de fichiers texte simples (.txt).
- **Activity** : ownCloud permet de lister toute l'activité de tous les utilisateurs en base de données afin de leur proposer un tableau de bord (dans l'interface web) de toutes leurs actions. Cela nous semblait consommateur dans la base de données pour un gain de fonctionnalité mineur, c'est pourquoi nous avons décidé de désactiver cette fonctionnalité.
- **Versionning** : Par défaut, ownCloud propose un versionning des documents présents. Autrement dit, pour chaque modification de chaque document, une nouvelle version de ce dernier est créée. Nous redoutions une perte d'espace de stockage effectif pour nos utilisateurs, due aux multiples versions de nombreux documents. Nous avons donc décidé de désactiver cette fonctionnalité en nous basant sur le fait que si un document est supprimé, il peut être restauré de la corbeille. En revanche, cela retire la possibilité de revenir sur une version antérieure dès qu'un document est modifié.

Les fonctionnalités essentielles d'ownCloud ont été conservées. Le service permet de déposer ou télécharger des documents depuis une interface web, de faire de la synchronisation de fichiers sur différents ordinateurs/périphériques, de partager des documents avec du personnel de l'université ou des intervenants extérieurs, etc.

4.2 Mise en place de plugins spécifiques

J'ai ajouté deux plugins à ownCloud afin de l'adapter à notre utilisation :

- **Superlog²** : Une fonctionnalité de log des différentes actions effectuées dans ownCloud était indispensable pour nous. En effet, en cas de réclamation/problème, nous souhaitions avoir un historique de ces actions. Ce plugin Superlog permet de les afficher dans l'interface web. Deux problèmes sont apparus : seul un administrateur de la plateforme ownCloud pouvait voir cela (alors qu'on ne souhaite pas avoir un tel rôle dans l'interface web) et l'inscription en base de toutes ces actions était trop consommateur. J'ai donc modifié ce plugin afin qu'il n'écrive plus les actions en base de données mais directement dans un fichier de log sur le frontal web. Ainsi, la consultation/recherche/historisation de ces derniers par des membres de l'équipe en charge du projet a été grandement facilitée. Voici un exemple de rendu de log applicative :

```
10-06-2015 08:31 oc1 [*** via webdav] get /DP/tableau.xlsx
10-06-2015 08:33 oc2 [*** via webdav] delete /MOOC/NouveauDossier
10-06-2015 08:33 oc1 [*** via web] rename /20150925.docx => /20150926.docx
10-06-2015 08:35 oc2 [*** via web] share /W8
```

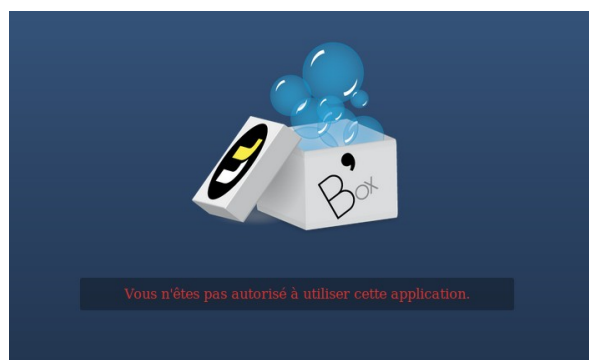
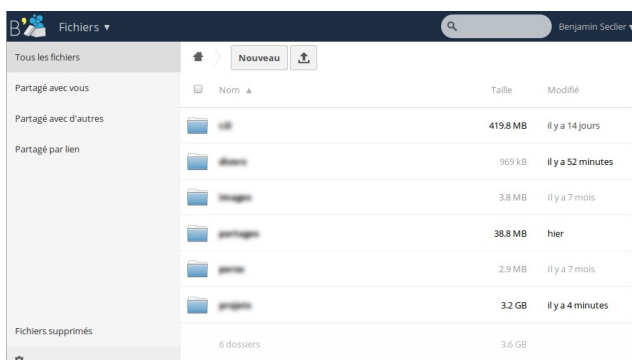
² <https://apps.owncloud.com/content/show.php?content=159124>

- **CAS user and group backend**³ : Ce plugin est développé par Pascal AVALLE de l'université Aix-Marseille. Il permet une authentification CAS ainsi qu'un filtrage des utilisateurs en se basant sur leur groupe Grouper⁴. Je l'ai modifié afin de l'adapter à notre mode de fonctionnement (utiliser un attribut LDAP contenant le groupe de l'utilisateur et non le format par défaut de Grouper). Il nous permet donc de gérer l'authentification ainsi que l'autorisation d'accès au service. En effet, nous n'avons autorisé l'utilisation du service qu'au personnel de l'université.

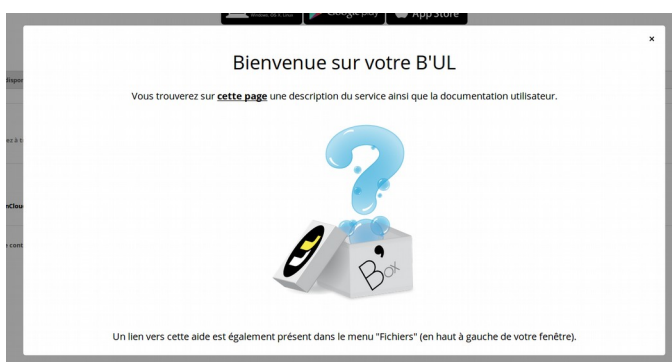
4.3 Adaptation graphique d'ownCloud aux couleurs de l'université de Lorraine

Nous avons souhaité modifier le thème graphique au minimum. Voici les pages que nous avons personnalisées/créées :

- La page de présentation des fichiers et la page de redirection en cas d'accès non autorisé au service.



- La popup s'affichant lors de la toute première connexion au service et quelques éléments réutilisables pour nos communications autour du service.



³ <https://apps.owncloud.com/content/show.php/CAS+user+and+group+backend?content=165973>

⁴ <http://www.internet2.edu/products-services/trust-identity-middleware/grouper/>

5 Ouverture progressive du service

N'ayant pas de retour d'expérience sur l'utilisation d'ownCloud au sein d'une grande population, nous avons décidé d'ouvrir le service progressivement. En septembre 2014, nous avons mis le service en production pour les membres de la Direction du numérique, soit 250 personnes. L'objectif était d'une part de démarrer sur une petite population et d'autre part, que les informaticiens de proximité prennent en main cet outil avant leurs utilisateurs.

L'ouverture s'est ensuite faite de façon progressive, au fur et à mesure des demandes. Le principe était le suivant : lorsqu'une personne de l'université remontait un besoin d'un service de ce type, je l'ouvrais à toute sa composante.

Voici l'évolution des ouvertures du service :

	Nombre d'inscrits au service
Septembre 2014	250
Décembre 2014	550
Janvier 2014	650
Février 2015	1000
Mars 2015	1700
Avril 2015	1900
Mai 2015	2350
Juin 2015	3000
Juillet 2015	Ouverture à tout le personnel

D'un point de vue technique, à chaque demande, je faisais un export LDAP de tous les membres de la composante concernée puis j'ajoutais ces personnes au groupe Grouper autorisé à accéder au service. Le plugin d'authentification CAS permettait de rediriger les personnes non-autorisées vers une page d'erreur (voir page précédente).

Un mail était ensuite envoyé à tous les nouveaux personnels inscrits afin de les informer de l'ouverture du service B'UL :

Nous venons de mettre en place la B'UL, un nouveau service de partage/synchronisation/stockage de vos données. À la différence de Dropbox ou Google Drive, les données que vous déposez/synchronisez sur la B'UL sont stockées sur les serveurs de l'université. La B'UL vous permet :

- d'accéder à vos données de plusieurs manières (interface web, votre système d'exploitation, votre téléphone mobile) ;
- partager des données/répertoires avec plusieurs personnes, qu'ils aient accès à la B'UL ou pas, qu'ils soient des personnels de l'UL ou pas ;
- synchroniser vos données entre votre(vos) ordinateur(s) et le serveur.

*Vous trouverez une documentation complète (présentation de l'interface web, installation de l'application de synchronisation, FAQ, etc.) à l'adresse suivante : ****

*Si vous rencontrez des problèmes, que vous avez des questions (qui ne sont pas dans la documentation), vous pouvez contacter vos informaticiens de site ou l'équipe en charge de la B'UL à l'adresse ***.*

Aujourd'hui, cette procédure d'ouverture progressive est terminée, le service est désormais ouvert à l'ensemble du personnel présent dans notre annuaire LDAP. Il n'y a plus aucune distinction de groupe. Tout compte personnel dans notre LDAP a également un compte ownCloud.

6 Outils mis en place

6.1 Fiche service et documentation

Une fiche a été mise en place⁵ sur le site web de présentation des services fournis par la Direction du numérique. Sur cette fiche sont présents les éléments importants tels qu'une présentation du service, l'URL d'accès au service, une adresse de contact ainsi qu'un lien vers la documentation réalisée.

L'objectif était d'avoir une documentation la plus exhaustive possible afin de rediriger les gens se posant des questions vers celle-ci. Elle est accessible depuis plusieurs endroits : la fiche service, le wiki de documentation de l'université mais également directement depuis l'interface Web d'ownCloud (dans une iframe).

Cette documentation contient des tutoriels illustrés sur les différentes fonctionnalités du service comme, par exemple, la présentation de l'interface principale, le partage de documents ou encore l'installation d'un client de synchronisation. En plus de cela, une foire aux questions est disponible, permettant de répondre aux questions les plus fréquemment posées par les utilisateurs (quota, gestion des doublons, etc.)

6.2 Vidéos

Deux vidéos ont également été réalisées par la Direction du numérique :

1. Une vidéo de présentation générale du service⁶ qui explique le concept et l'utilisation de cet outil. Nous y évoquons la possibilité d'accéder aux documents depuis n'importe quelle connexion Internet via le navigateur, de synchroniser des documents sur n'importe quel ordinateur (via le client de synchronisation) ainsi que de partager des documents avec du personnel de l'université ou des intervenants extérieurs.
2. Une vidéo « screencast » tutorielle d'installation d'un client de synchronisation (les étapes du téléchargement, de la configuration, etc.) ainsi que du partage des documents (pour les partages internes ou externes à l'université).

6.3 Enquête de satisfaction

Au mois d'avril, nous avons souhaité avoir l'avis des premiers utilisateurs du service. Nous avons donc mis en place une enquête afin de recueillir leurs opinions/besoins/souhaits. Voici les principales questions posées et leurs réponses (les détails ne sont pas indiqués dans cet article) :

	Oui	Non	Pas totalement
Utilisez-vous votre B'UL (sinon, pourquoi) ?	61 %	39 %	
Êtes-vous satisfait du service (sinon, pourquoi) ?	79 %	1 %	20 %
Trouvez-vous que 10Go d'espace soit suffisant ?	62 %	38 %	
Avez-vous trouvé la documentation claire et complète ?	93 %	7 %	
Quelle(s) fonctionnalité(s) souhaiteriez-vous voir apparaître dans la B'UL ?	QCM, réponses diverses		
Utilisez-vous le client de synchronisation sur votre ordinateur ?	79 %	21 %	
Utilisez-vous la fonctionnalité de partage de documents ou de répertoires ?	87 %	13 %	
Avez-vous des commentaires / suggestions à faire sur le service B'UL ?	Question ouverte, réponses diverses		

⁵ <http://numerique.univ-lorraine.fr/outils-collaboratifs/bul-owncloud>

⁶ <http://videos.univ-lorraine.fr/index.php?act=view&id=2573>

7 Quelques chiffres

Voici quelques chiffres à la date du **6 décembre 2015**. L'annonce officielle d'ouverture du service a été faite la semaine dernière au sein de l'ensemble du personnel de l'université.

7.1 Consommation de l'espace disque

- Espace utilisé : 602 Go
- Espace restant disponible : 7,9 To

7.2 Utilisateurs et quotas

- Utilisateurs connectés au moins une fois : 1009
- Espace moyen utilisé par personne : 1,1 Go

<i>Espace utilisé</i>	Pourcentage d'utilisateurs
<i>Entre 0 et 1 Go</i>	76 %
<i>Entre 1 et 3 Go</i>	11 %
<i>Entre 3 et 5 Go</i>	5 %
<i>Entre 5 et 7 Go</i>	4 %
<i>Entre 7 et 10 Go</i>	4 %

8 Évolutions à venir

Voici les différents projets d'évolution du service pour l'avenir :

- **Mettre à jour ownCloud** en version 8 (nous sommes actuellement en version 7). Deux problèmes principaux empêchent actuellement cette mise à jour. Le premier est l'absence de compatibilité du plugin Superlog qu'il va donc falloir modifier pour l'adapter. Le second est qu'ownCloud 8 nécessite une version de PHP non présente dans les paquets de base de notre distribution actuelle (Centos 6.6). Je dois déterminer s'il est préférable d'ajouter des dépôts ou s'il vaut mieux mettre à jour la distribution pour passer en Centos 7.
- **Changer notre support de stockage**. Nous stockons actuellement nos données sur un espace SAN exposé via notre Serveur NFS. Un groupe de travail de notre sous-direction est en train d'étudier et de mettre en place une solution de stockage distribué (basé sur la solution Ceph). Cela nous permettrait de fournir davantage d'espace de stockage à nos utilisateurs car le coût de stockage est nettement inférieur sur une solution de ce type que sur une baie SAN.
- Continuer à faire vivre le **wiki**⁷ et la **liste de diffusion**⁸ mis en place chez Renater.
- Mettre en place une procédure de **purge automatique** des anciens comptes.

⁷ <https://groupes.renater.fr/wiki/owncloud-fr/index>

⁸ owncloud-fr@groupes.renater.fr