Networking

Pandora: la piattaforma di "storage in the cloud" dell'Area della Ricerca RM1 di Montelibretti.†

Giuseppe Nantista,[a] Andrea Lora,[a] Augusto Pifferi,[a]

In questo rapporto tecnico descriveremo il setup di una piattaforma di cloud storage accessibile dagli utenti dell'Area della Ricerca RM1 che hanno già un account email attivo sul server di posta elettronica locale, Mercurio. Dopo la fase iniziale di setup abbiamo collegato una share NFS ospitata su uno storage esterno su cui memorizzare i dati degli utenti, quindi configurato il bridge LDAP per l'autenticazione single sign on mediante le credenziali di posta elettronica.
Keywords: OwnCloud, cloud storage, single sign on, collaboration tools

1 Introduzione

La collaborazione fra membri di un gruppo di lavoro ha bisogno di opportuni strumenti, in questo rapporto tecnico discutiamo di uno strumento software che da la possibilità di accedere da remoto ai dati condivisi del gruppo di lavoro, senza precluderne l'accesso da terminali mobili come tablet o smartphone.

Esistono servizi come Dropbox o Google Drive, tanto per citare i più diffusi, che mettono a disposizione questa possibilità anche in maniera gratuita con spazi allocati di modesta entità. Tuttavia la necessità di allocare spazio in quantità elevata, unita alla preferenza per ospitare dati sensibili della ricerca scientifica su piattaforme a gestione interna, ci ha spinti a sviluppare un sistema analogo basato su software open source reperibili in rete e opportunamente supportati dalla comunità.

Il progetto Pandora si basa su software di cui il core è costituito da ownCloud e permette la memorizzazione di file sui server dell'Area di Ricerca RM1 di Montelibretti, per poi accedervi dovunque si abbia una connessione ad internet.

Lo scopo di Pandora è quello di offrire agli utenti un modo facile, veloce ed affidabile di conservare e condividere file, per ulteriori informazioni sul software, le guide sull'utilizzo e per scaricare i client di sincronizzazione è possibile far riferimento al sito del prodotto: http://owncloud.org/

Cos'è in pratica Pandora? Un sistema di storage remotizzato, che permette di mantenere i propri file memorizzati in sicurezza su una cartella ospitata presso un server remoto, accessibile via web da qualunque parte del mondo. Non solo, l'accesso tramite Sync Client offre la possibilità di mantenere una copia di tutti i file sui propri PC, replicando le modifiche su tutte le copie attive contemporaneamente, così da poter modificare i propri documenti da più postazioni senza dover salvare su supporti non affidabili (le classiche chiavette USB) e consentendo anche il lavoro off line nei momenti in cui l'accesso a internet non è disponibile.

2 Setup

Abbiamo approntato un server con distribuzione Debian Wheezy su cui abbiamo installato i package di base richiesti da owncloud, quindi apache2, php5, mysql server, curl e altre librerie richieste.

Data la riservatezza dei dati in transito e delle credenziali di login degli utenti è fondamentale che tutte le comunicazioni siano cifrate. Per questo abbiamo installato dei certificati SSL ottenuti gratuitamente tramite l'accordo con il GARR e Terena Certification Authority.

L'installazione del pacchetto owncloud si effettua semplicemente estraendo dentro la cartella /var/www il package scaricato dal sito http://owncloud.org, lo step successivo è stato quello di collegare la cartella /var/www/-data, dove verranno memorizzati i dati degli utenti, allo storage esterno tramite share NFS.

Per consentire l'accesso tramite la login della propria casella e-mail è stato quindi necessario abilitare il plugin LDAP e configurarlo opportunamente. È necessario configurare non solo l'accesso in lettura all'albero LDAP relativo agli account utente, ma anche distinguere un utente abilitato ad avere il suo spazio su Pandora da un utente

[a] *Istituto di Cristallografia, C.N.R. via Salaria km 29.300, 00015 Monterotondo, Italy*

† Rapporto tecnico 2014/21 con protocollo CNR-IC 2213 del 11/12/2014

che non ne ha diritto.

Questa operazione si ottiene filtrando gli utenti in base al servizio attivo "owncloud".

User Filter:

```
(&
  (objectclass=mailUser)
  (enabledService=owncloud)
)
```

Login Filter:

```
(&
 (&
  (objectclass=mailUser)
  (enabledService=owncloud)
 )
 (
   (mailPrimaryAddress=%uid)
   (mail=%uid)
 )
)
```

Poiché la base dati usata dal nostro sistema di posta elettronica è di tipo LDAP è stato necessario aggiungere un "enabledService" a tutti gli utenti abilitati al servizio e, in fase di login, effettuare una verifica sulla presenza dello stesso flag nella porzione di albero relativa all'utente che si sta loggando.

3 Configurazioni Avanzate

Di default a un utente che effettua il login tramite credenziali LDAP[1] verrà assegnato un uid (user identifier) esadecimale e i suoi file verranno memorizzati nella directory /var/www/data/uid. Per ragioni di semplicità nel management abbiamo fatto sì che il percorso dati coincidesse con la login dell'utente, modificato il campo "Attributo nome utente interno" affinché coincidesse con l'indirizzo email.

Un'altra personalizzazione è relativa al campo email associato a ogni utente. Nonostante esso sia di default assegnato al valore "mail" della struttura LDAP, fintanto che l'utente non effettua il primo login tale valore è vuoto. Come conseguenza ogni qual volta un utente condivide un file della propria area con un collega che non ha mai effettuato il login la mail di notifica non viene consegnata. Con uno script ad hoc abbiamo popolato tale tabella, inoltre lo script è stato inserito in crontab per essere eseguito ogni giorno. Si veda in appendice il listato dello script.

I file gestiti dalla piattaforma sono memorizzati su una porzione di disco offerta da una SAN Solaris montata via nfs4 con le opzioni **hard** e **initrd**. La quota è stata configurata a 2 TB.

È necessario evitare che il demone apache parta prima del nfs, questo è evitabile aggiungendo l'opzione **netdev** al file **/etc/fstab**. Inoltre, per evitare qualsiasi problema, è bene modificare il file **/etc/init.d/apache** affinché esso impedisca l'avvio del servizio qualora la share NFS non fosse montata. A tal proposito è sufficiente creare un file sulla share NFS che contiene i dati (nel nostro caso **/var/www/data**), mediante il comando **touch we_are_on_nfs** ed eseguire prima dell'avvio del servizio una verifica sull'esistenza del file. Il seguente snippet di codice fornisce una possibile implementazione

```
if [ ! -f /var/www/data/we_are_on_nfs ]
then
    echo Missing nfs mount
    exit 1
fi
```

Di default i log vengono salvati nella directory **/var/www/data**. Ciò può portare diversi problemi di performance a causa dei numerosi accessi I/O. Una possibile soluzione è quella di spostare il file di log in altra posizione. Questa posizione è configurabile attraverso il file di configurazione con il parametro logfile.

Esempio:

$$\text{"}\mathbf{logfile}\text{"} => \text{"}/\mathbf{var}/\mathbf{log}/\mathbf{owncloud}.\mathbf{log}\text{"}$$

4 Online Editor

La piattaforma da anche la possibilità di condividere documenti di testo in formato odt (open document text) con l'aggiunta di un sistema online di editing, che permette quindi di accedere in modalità lettura/scrittura sullo stesso file contemporaneamente senza imbattere in conflitti di accesso e con in più la possibilità di vedere in diretta le modifiche che stanno facendo i partecipanti alla sessione di editing. Le modifiche del singolo editore vengono evidenziate con colori differenti.

5 Monitoraggio della Piattaforma

Tutti i server gestiti dal nostro gruppo viene tenuto sotto controllo tramite una piattaforma di monitoring basata sul software open source Zabbix[2], che si occupa di tenere sotto controllo lo stato del sistema e dei processi tramite l'uso di un agent.

La procedura di installazione dell'agent prevede 3 passi:

```
# wget http://repo.zabbix.com/zabbix/2.0/
debian/pool/main/z/zabbix-release/
zabbix-release_2.0-1wheezy_all.deb
# dpkg -i zabbix-release_2.0-1wheezy_all.deb
# apt-get update
# apt-get install zabbix-agent
```

La configurazione del template linux è sufficiente a monitorare una gran quantità di parametri. Resta da configurare il controllo sullo spazio disco della share NFS. Avendo assegnato una quota in fase di configurazione basterà interrogare il sistema operativo tramite df, il cui output è riportato in appendice.

Quindi un check con parametro

$$[\mathbf{vfs.fs.size}[/\mathbf{var}/\mathbf{www}/\mathbf{data}, \mathbf{pfree}]]$$

fornirà le informazioni richieste in zabbix.

6 Appendice

Script di popolazione della tabella mysql oc_preferences

```
#!/bin/bash
tmp=$(mktemp)
if [ -f $tmp ];
then
    echo "Il file con le istruzioni sql e' $tmp"
else
    echo "$tmp non e' stato creato"
    exit
fi
lista_utenti=$(ldapsearch -h *ip_ldap* -p 389 -D "cn=Manager,dc=mlib,dc=cnr,dc=it" -w *password* \
-b "o=domains,dc=mlib,dc=cnr,dc=it" enabledService=owncloud mail | grep "mail: " | cut -d " " -f 2) \
for utente in $lista_utenti
do
echo
"INSERT
IGNORE
into
oc_preferences
(userid,appid,configkey,configvalue)
VALUES
(\"$utente\", \ \"settings\",\"email\",\"$utente\");" >> $tmp
done
echo File $tmp contiene gli statement SQL
```

Output del comando df

```
# df -h /var/www/data
File system
Dim. Usati Dispon. Uso% Montato su
server_nfs:/volumes/tank/data 2,0T 47G 2,0T 3% /var/www/data
```

Riferimenti

1 G. Nantista, G. Righini, L. Ianniello, A. Lora, A. Pifferi, Servizi DNS e DHCP con Backed LDAP in Business Continuity, SMART eLAB 1 (2013) 1–8. doi:10.30441/smart-elab.v1i0.15.

2 A. Pifferi, G. Nantista, L. Ianniello, A. Lora, M. Simonetti, Analisi e implementazione di sistemi per il monitoraggio della rete wireless relativa al progetto add (anti digital divide) e delle infrastrutture di campus adr rm1., SMART eLAB 2 (2013) 1–9. doi:10.30441/smart-elab.v2i0.46.