DISASTER PLAN, A DRAFT

Riccardo Del Gratta

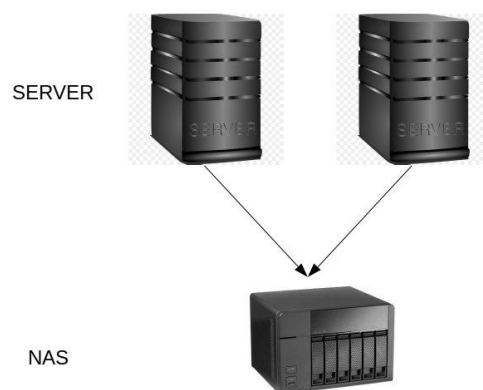dspace-clarin-it-ilc-help@ilc.cnr.it

1 Executive Summary

This document describes the draft disaster plan at the ILC4CLARIN[1] repository[2] of the CLARIN-IT[3] Consortium.

2 REPOSITORY AT ILC4CLARIN: CURRENT ARCHITECTURE

The current architecture of the repository at ILC4CLARIN consists of two Dell PowerEdge R630 Rack Servers[4] in the active/passive configuration.

The servers are configured in a High Availability (HA) arrangement[5], with the addition of data replication using the Distributed Replicated Storage System DRBD[6].

The local infrastructure is completed with a Synology NAS[7] hosted at a different CNR institute in a separate building.

---

[1] https://ilc4clarin.ilc.cnr.it/en/
[2] https://dspace-clarin-it.ilc.cnr.it/repository/xmlui/
[3] https://clarin-it.it/
[4] http://www.dell.com/en-us/work/shop/productdetails/poweredge-r630
[5] http://corosync.github.io/corosync/
[6] https://linbit.com/drbd/
[7] https://www.synology.com/en-global/products/RS819

3 CONNECTED SOFTWARE REPOSITORIES

3.1 DSPACE GITHUB

Main software repository
https://github.com/ufal/clarin-dspace

3.2 OVERLAYS GITHUB

ILC4CLARIN GIT https://github.com/cnr-ilc/ilc4clarin-overlays
This repository contains customization made by CNR-ILC. See
https://github.com/ufal/lindat-dspace/wiki/Overlays

3.3 HEADER/FOOTER

ILC4CLARIN GIT https://github.com/cnr-ilc/ilc4clarin-common
This repository contains the common theme (from https://github.com/ufal/lindat-common.git)
of the ILC4CLARIN repository.

3.4 SSO/AAI

ILC4CLARIN GIT https://github.com/cnr-ilc/ilc4clarin-aai
This repository contains customization made by CNR-ILC for aai and discojuice.

4 PERIODIC BACKUPS OF DSPACE

The strategy adopted at ILC4CLARIN is to backup databases and assetstore every week to
maintain consistency between databases and data (assetstore)

The NAS is mounted on /mnt/sinology-nas
Copies of the databases and assetstore are backed up on /mnt/sinology-nas

4.1 Weekly use of dspace backups commands

backup_databases to backup main and utils databases.
Every WED at 3 am two backups are created over a default folder. Then they are copied
under /mn/sinology-nas, at 4 am.

```
0 3 * * 4 cd /opt/git/ilc4clarin-dspace/utilities/project_helpers/scripts;  sudo make backup_databases
0 4 * * 4 cd /opt/git/ilc4clarin-dspace/utilities/project_helpers/scripts;  sudo copy_backup_db.sh
```

4.2 Weekly backup of DSPACE assetstore

Copy the entire assetstore over /mnt/sinology-nas

Every WED at 4:30 am the assetstore is zipped and saved under /mn/sinology-nas.

```
30 4 * * 4 cd /opt/git/ilc4clarin-dspace/utilities/project_helpers/scripts;  sudo copy_andzip_assetstore.sh
```

5 Service Level Agreement

IL4CLARIN guarantees a weekly backup of data. This is a good trade-off between storage usage and ILC4CLARIN data volatility.

6 SCRIPTS

Copy database backups (copy_backup_db.sh)

```
#! /bin/bash
# copy_backup_db.sh
today=$(date +"%Y-%m-%d")
NASDIR=/mnt/sinology-nas
DBDIR=/opt/java/ilc4clarin-dspace/ilc4clarin/database_backup
LOGDIR=${NASDIR}/logs
LOGNAME=db_logs
DBDIR=database_backup
DBNAME=ilcforclarin-db
DBUTILNAME=ilcforclarin-db-utils
DB1=${DBDIR}/${DBNAME}
DB2=${DBDIR}/${DBUTILNAME}

LOGNAME=${LOGDIR}/${LOGNAME}-${today}.log
CMD1="cp -ra ${DB1} ${NASDIR}/${DBDIR}/${DB1}-${today}.sql >> ${LOGNAME}"
CMD2="cp -ra ${DB2} ${NASDIR}/${DBDIR}/${DB2}-${today}.sql >> ${LOGNAME}"
printf "Copying backup file ${DB1} to the NAS server'%s'\n"
"${NASDIR}/${DB1}-${today}.sql"
`${CMD1}`
printf "Copying backup file ${DB2} to the NAS server'%s'\n"
"${NASDIR}/${DB2}-${today}.sql"
`${CMD2}`
```

```
#! /bin/bash
# copy_andzip_assetstore.sh
today=$(date +"%Y-%m-%d")
```

```
NASDIR=/mnt/sinology-nas
ASSETDIR=/opt/java/ilc4clarin-dspace/ilc4clarin/installations/assetstore
LOGDIR=${NASDIR}/logs
LOGNAME=asset_zip
TARNAME=assetstore-${today}.tar.gz
LOGNAME=${LOGDIR}/${LOGNAME}-${today}.log
CMD1="tar cfz  ${NASDIR}/${TARNAME} ${ASSETDIR} >> ${LOGNAME}"
printf "Copying asset file ${ASSETDIR} to the NAS server'%s'\n"
"${NASDIR}/${TARNAME}"
`${CMD1}`
```