



Data Management for Digital Archivists

*Jan Just Keijser
Nikhef / BiG Grid
10 April 2013*

J.J. Keijser
Nikhef
Amsterdam
Grid Group

Data Archiving and Networked Services
DANS *Door data gedreven*



BiG Grid
the dutch e-science grid

What is DANS?

“DANS promotes sustained access to digital research data. For this purpose, DANS encourages researchers to archive and reuse data in a sustained manner, e.g. through the online archiving system EASY. DANS also provides access, via NARCIS.nl, to thousands of scientific datasets, e-publications and other research information in the Netherlands.”

(from: <http://www.dans.knaw.nl/en/content/about-dans>)

Typical DANS Datasets

- > 23,000 data sets
- Every data set consists of 1+ data files, sometimes 1000+
- Most data sets are small (98% < 1 Gb)
- For example, the entire population census of 1960 (>11 million records) fits on one CD-ROM (< 700 Mb)
- Total number of files >1 million
- Total storage volume ca. 70 Tb
- Long processing times with large numbers of datasets and files

Usecase: Fixity

- Preserve a copy of (a part of) the DANS archives on the grid
- Long-term archival: > 10 years
- Continually check the integrity of the archives *as an end-user (archivist)*

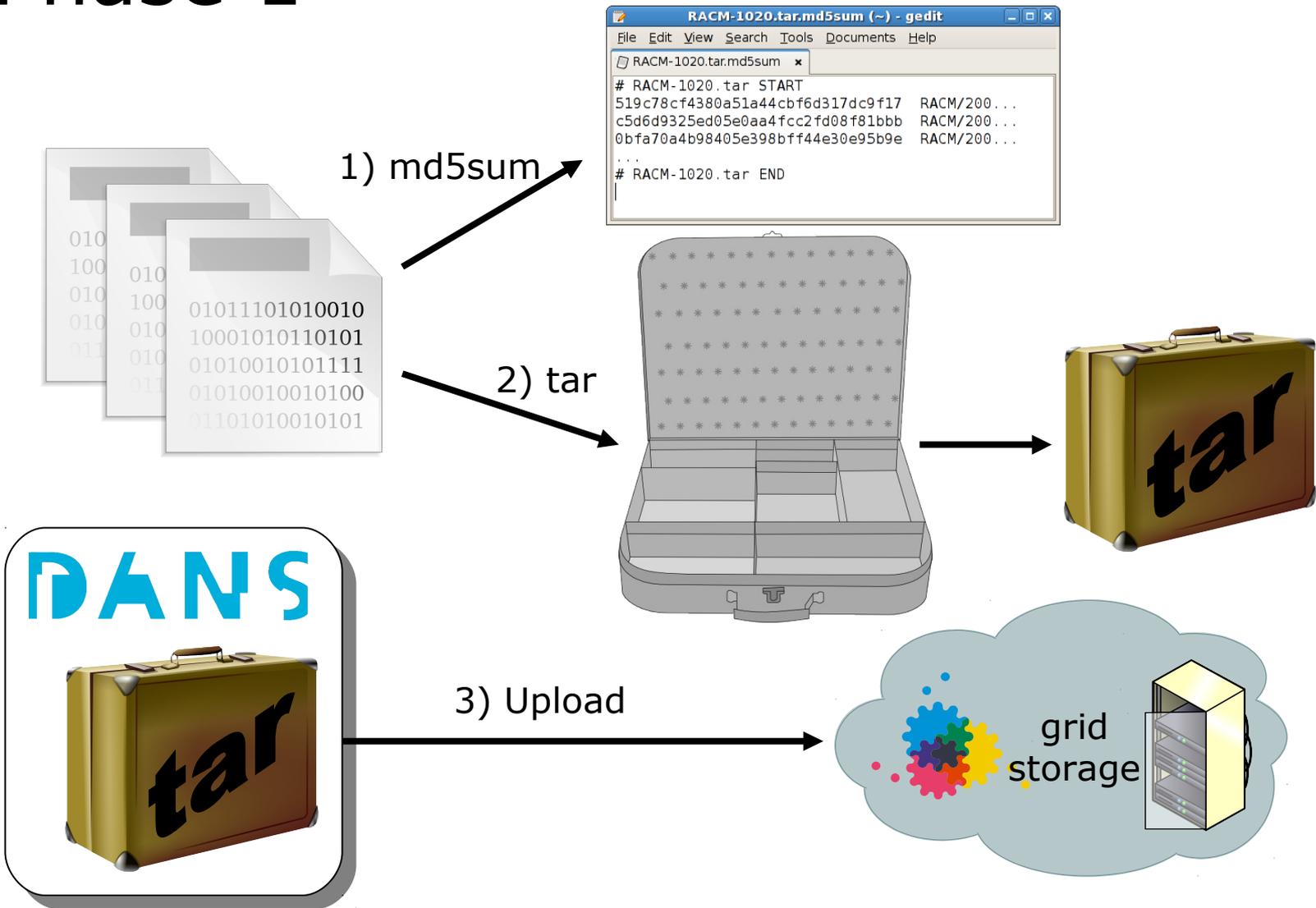
Excellent grid usecase:

Checksumming large amounts of data is quite compute intensive

Design

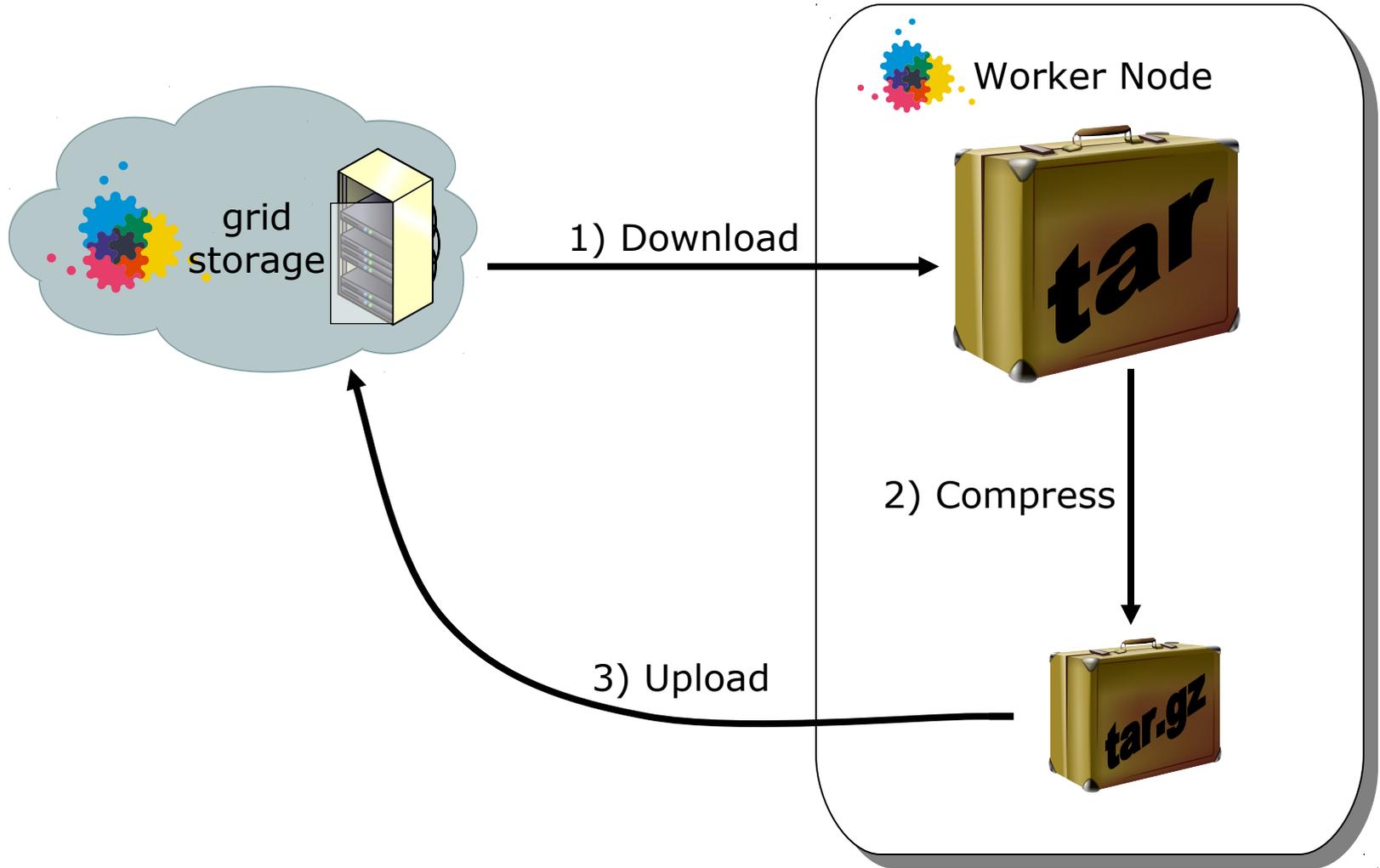
- 3 phase workflow
- Upload, compress and verify archives
- Store two copies of each archive on tape: one in Amsterdam, one in Almere
- Fully end-user controlled
- Hide complexity of grid commands as much as possible
- KISS

Phase 1



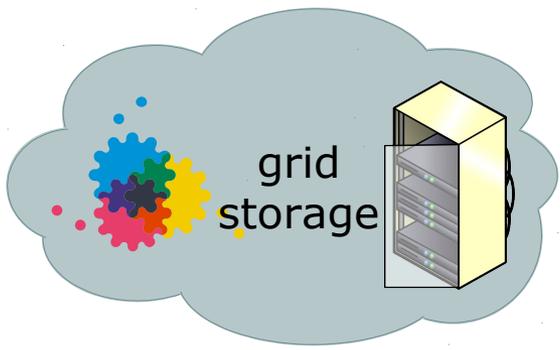
J.J. Keijser
Nikhef
Amsterdam
Grid Group

Phase 2



J.J. Keijser
Nikhef
Amsterdam
Grid Group

Phase 3

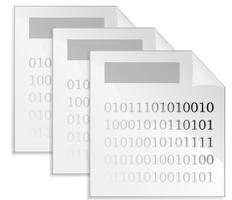


1) Download

Worker Node



2) Unpack



3) md5sum

```
*RACM-1020.tar.md5sum (~) - gedit
File Edit View Search Tools Documents Help
*RACM-1020.tar.md5sum x
# RACM-1020.tar.gz START
519c78cf4380a51a44cbf6d317dc9f17 RACM/200...
c5d6d9325ed05e0aa4fcc2fd08f81bbb RACM/200...
0bfa70a4b98405e398bfff44e30e95b9e RACM/200...
# RACM-1020.tar.gz END
```

4) Compare

DANS

```
RACM-1020.tar.md5sum (~) - gedit
File Edit View Search Tools Documents Help
RACM-1020.tar.md5sum x
# RACM-1020.tar.gz START
519c78cf4380a51a44cbf6d317dc9f17 RACM/200...
c5d6d9325ed05e0aa4fcc2fd08f81bbb RACM/200...
0bfa70a4b98405e398bfff44e30e95b9e RACM/200...
# RACM-1020.tar.gz END
```

J.J. Keijser
Nikhef
Amsterdam
Grid Group

Results

- Generic set of scripts for all workflow phases
- Scripts written in bash/awk/sed, plus glite commands
- Verification checksums calculated using 'md5sum' and 'md5deep'
- Most grid functionality hidden from end-users
- Time to checksum all files is drastically reduced
- 5 archives
- 290,341 "source" files
- 1,695 "tar.gz" files of ~ 5 GB each
- Total ~ 8.5 TB

Problems & workarounds

- First attempt: store files directly
 - Problem: SRM tools do not like spaces in filenames
- Second attempt: use ZIP format
 - Problem: zip files cannot be larger than 4 GB

Solution: Compressed tarballs (.tar.gz files)

- Can be larger than 4 GB
- Works around spaces in filenames
- Adds extra checksums
- Saves space

Remaining issues

- Validating a single file takes long in this setup
- Set of scripts only: no web interface
- Creating an archive is a bit tricky
- Grid credentials are required for the digital grid archivist, although a robot proxy could be used
- No way to validate both tape copies (grid tools don't allow for it)

Non-issue

- During one 'verify' run a tarball seemed corrupt. Turned out the disk on the worker node that was verifying the tarball was bad.

Recommendation

- Ensure that the paths are the same everywhere:
 - local archive: `$ARCHIVE_PATH`
 - LFC : `$LFC_HOST:$LFC_HOME/$ARCHIVE_PATH`
 - SRM : `$SRM_HOST:$SRM_HOME/$ARCHIVE_PATH`

The LFC/GFAL commands do not do this by default!

Conclusion & Future work

- A “fixity” service on the grid has been set up
- The service ensures that data stored on the grid remains valid (fixed)
- DANS is now actively using the grid as a backup resource

Future work

- Open up the data stored on the grid to large-scale computing
- Web interface to validate a single file
- Integration with EASY



J.J. Keijser
Nikhef
Amsterdam
Grid Group