Ligo-VirgoCC data transfer

A. Colla, C. Palomba

DAC call April 19, 2013

Current schema

Virgo

LIGO (AEI) ←LDR→ Cascina ←LCG→ IN2P3

Cascina ←iRODS→ CNAF
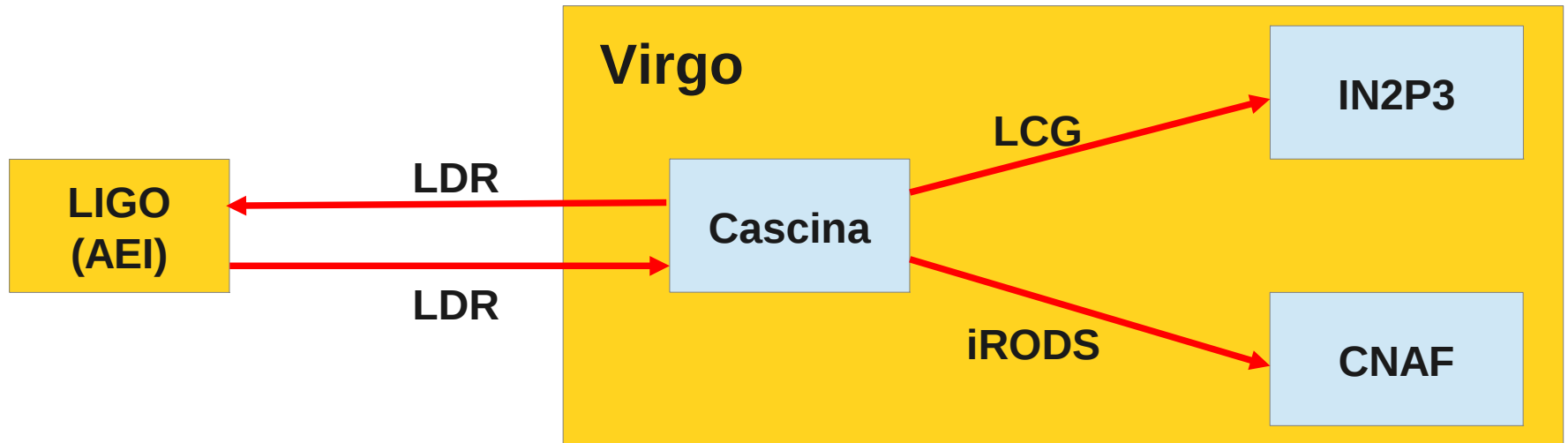
Current implementation: Ligo RDS data are transferred to Cascina with LDR, then copied to Virgo CCs with the same framework as the Virgo bulk DT.

Drawbacks:
- ✗ Ligo data are not used in Cascina! They are needed at the CCs for offline analysis
- ✗ DT occupies bandwidth, disk buffers at Cascina.
- ✗ One more possible point of failure to monitor

Proposed schema
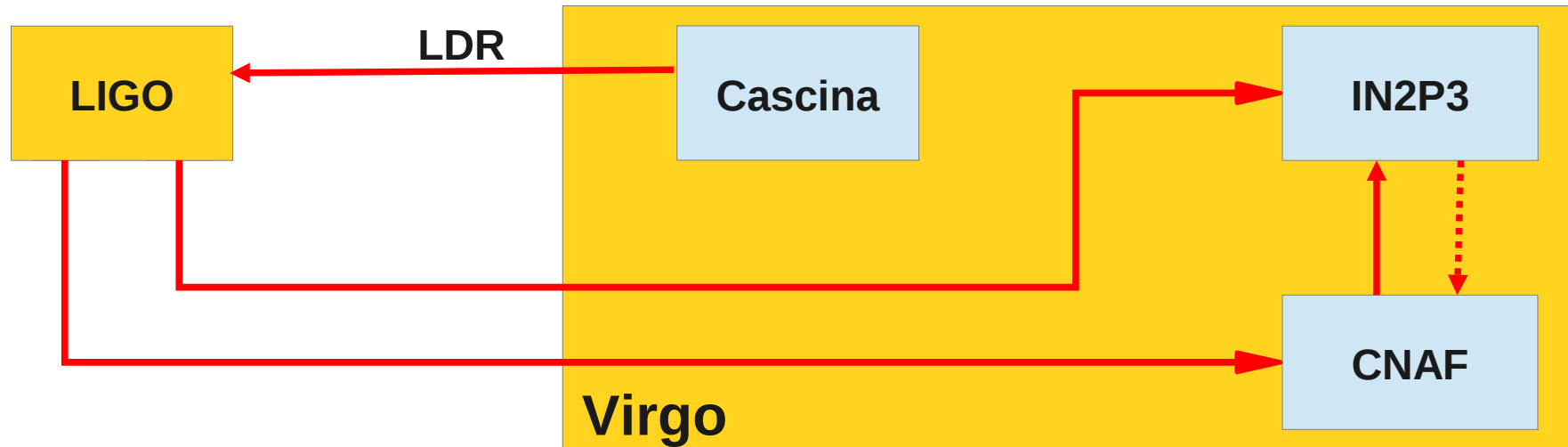
Copy Ligo RDS data directly to Virgo Ccs, either in a "star" or "daisy chain" configuration
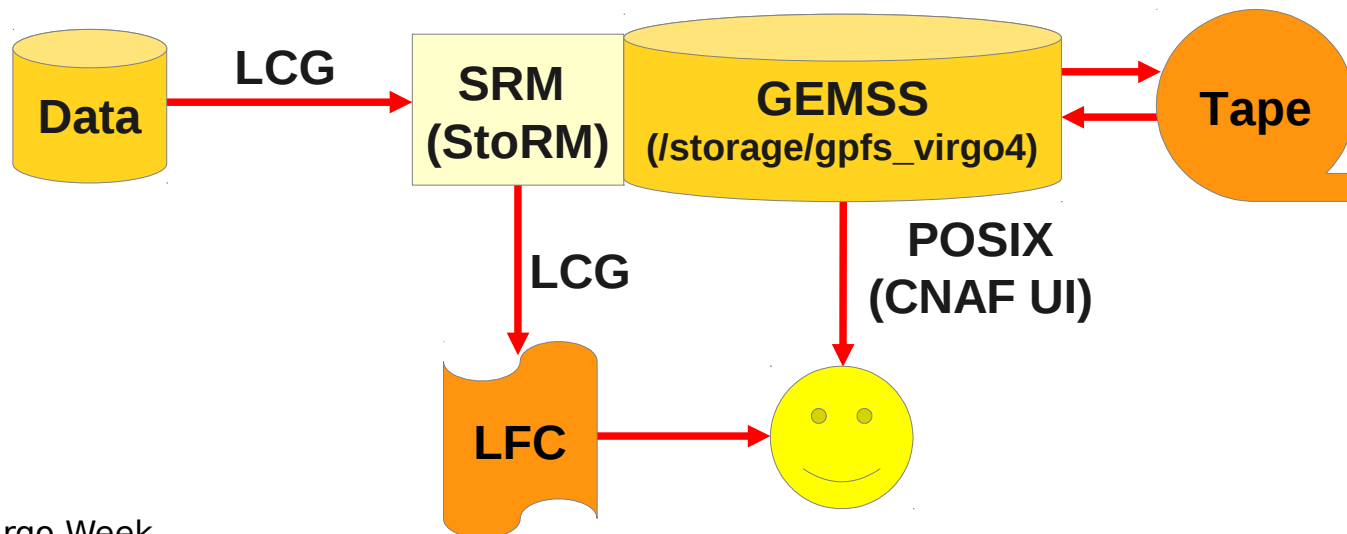
Expected RDS data flow for ADE: O(10 TB) per year

Hot issues

LIGO RDS data transfer should be done with the same tools used for the Virgo bulk DT
- ✔ **LCG** for **CNAF, iRODS** for **Lyon**

In this way management, bookeeeping, access of Virgo and LIGO data would be coherent

e.g. at CNAF:

Investigations

We started investigating the possibility of use **LCG tools** to transfer data from LIGO CC to CNAF

We should ask to install a "Grid User Interface" on a LIGO machine which "sees" the data

- ✔ A few RPMs developed and maintained by the EGI community
- ✔ Possibly at Caltech, which uses SL6, fully compatible with EGI middleware
  - ✗ AEI uses Debian for which EGI middleware is not yet ready

Investigations

If we have the Grid UI we could install and test the "Grid DT" prototype developed for the bulk data transfer

✔ already implements many useful features: error recovery, data integrity checks, web monitoring

✔ Installation and setup would be straightforward

  ✔ we could plan an "official" test already for **ER4**!

Data transfer to IN2P3 could then be done from CNAF, installing iRODS on a user interface (to be verified)

✔ Of course also this project would highly benefit of using the same technology for the data transfer at CNAF and IN2P3 (The Computing Model encourages to go in this direction)

We need to negotiate...

LIGO colleagues do not like very much the idea of installing external libraries on their computers :(

They propose to use either:

✔ **LDR**
  - ✗ to be installed at CCs, which we don't think is feasible

✔ **plain gridFTP**
  - ✗ Again, not sure that Virgo CCs allow that for a production environment
  - ✗ LCG tools do use gridFTP embedded in their specific functionalities

In any case, then we should do extra work to integrate data in the same environment as Virgo data (e.g. LCG) !!

We are organizing a call with LIGO experts to discuss these issues in the next days!