Alice plans for EDG-2 and LCG-1

Stefano Bagnasco, INFN Torino
Piergiorgio Cerello, INFN Torino

Running Productions

- **Goal 1:** simulate and reconstuct 5000 Pb-Pb central events
  - Output size: about 1.5 GB/event (8 files), so **7.5 TB**
  - Job duration: 12 to 24 hrs

- **Goal 2:** analyse these events
  - 1 single job to analyse all the events

- **Status:**
  - about 300 central Pb-Pb events simulated
  - output stored on the EDG SE's and registered on the EDG RC
  - Output is being replicated on CASTOR at CERN and CNAF and registered in the AliEn Data Catalogue
  - About 100 events "natively" on MSS somewhere
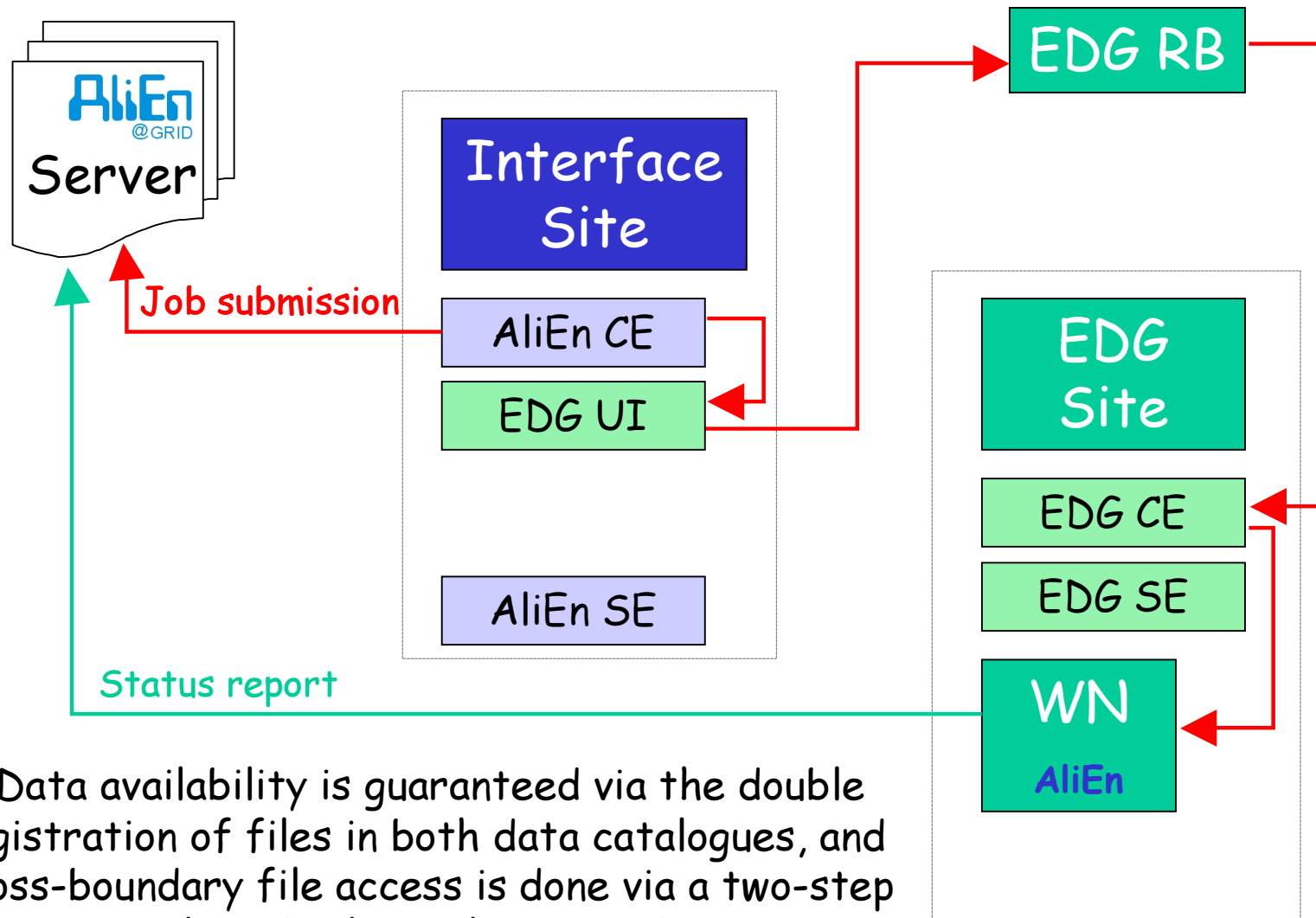
Running Productions

- **Comments**

  - Application Testbed unstable
  - AliEn interface functionality available but unreliable
  - Many jobs (at least twice the completed sample) failed for services failures
  - It takes a long time to track down the errors and clean up the RC by hand when needed
  - We have access to only one CE at Lyon

    ➲ See next slide

Running Productions

💣 **The "Lyon problem"**

- WN's publish the total available memory in the IS
- The JDL memory requirement is compared to the published values
- When more than a job is allowed on the WN, the memory is shared.
- Aliroot jobs break because they need more memory than the actually available amount
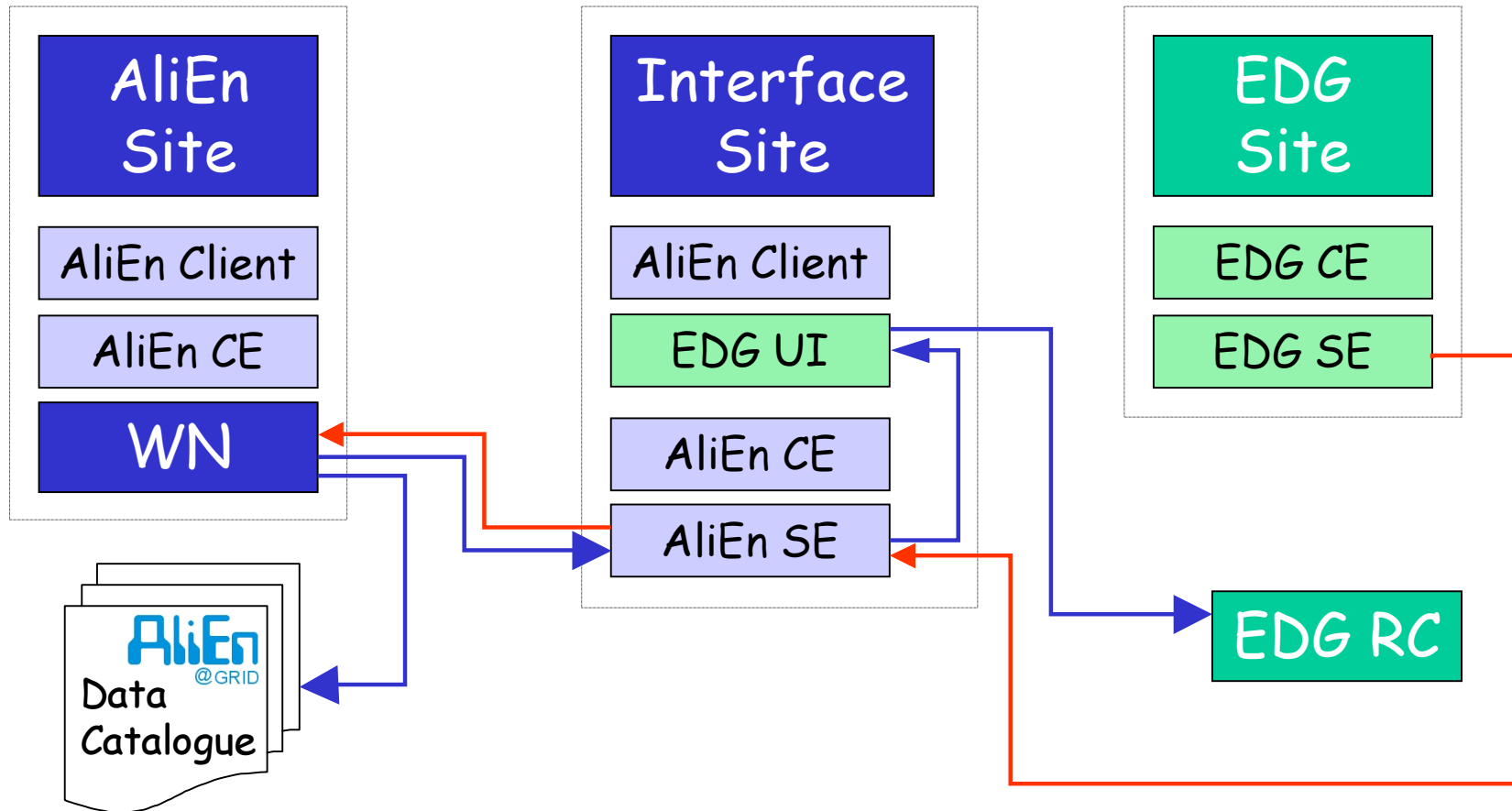- Workaround by F. Hernandez, but on 1 CE only - why?

AliEn - EDG Interface

- Data availability is guaranteed via the double registration of files in both data catalogues, and cross-boundary file access is done via a two-step process involving both catalogues and an "intermediate" logical file name.

AliEn - EDG Interface

- **Complex cross-boundary data access: file replication**

AliEn - EDG Interface

- **Goal:** manage productions on the EDG testbed from the AliEn portal in parallel with AliEn "native" productions

  - ✓ Job submission through AliEn: ok (but not used for production)

  - ✓ Write Data to the AliEn Catalogue from EDG: ok
  - ✓ Read Data in the AliEn Catalogue from EDG: ok
  - ✓ Read Data in the EDG Catalogue from AliEn: ok
    BUT performance problem (investigating)
  - ✗ Data requirements passing to the RB: coming soon.

  - ✗ Error management and recovery: difficult.
    Badly need more streamlined API's...

Plans and issues - Short Term

- **Open issues and short-term plans:**

  ✗ Testbed Stability the major problem.

  ✓ Require OUTBOUND connectivity on WN's - ok, but for how long?

  ✗ Testbed Resources: **storage** - LCG-1

    ✗ Complete the test production - not yet

    ✓ Complete the AliEn-EDG interface - ok (or nearly so)

    ✗ Run a high level (analysis) use cases - no

    – Use ALICE resources + LCG testbed under AliEn control - functionality available, in place on LCG-1

Plans - 2003/2004

- **Goal 0:** complete the ongoing production - at the present rate, it will take 18 months...

- Full functionality for data analysis is now missing: data access.

- Evaluate EDG-2 functionality with Metadata

- Try to develop an analysis use case on EDG-2

- If not possible, use AliEn for data analysis:
  - ROOT API to AliEn available, integration of software with Data Catalogue may be easy
  - Pseudo-interactive analysis already possible with distributed input

Data GRID

EU DataGrid Project Conference
Barcelona may 13, 2003

Alice plans for EDG-2/LCG-1 9/8
Stefano Bagnasco, INFN Torino

INFN Istituto Nazionale di Fisica Nucleare