T2K on the GRID

Ben Still
Queen Mary, University of London

What is the GRID?

- We use the LHC Computing Grid (LCG)
  - Developed and supported by CERN
  - Plus additions E.g. t2ksrm.nd280.org
- Distributed CPU and Storage accessible locally.
  - Not supercomputers just ordinary CPU
  - Large scale disk/tape storage at many locations
  - Part of the general research infrastructure in some countries.

Why Use It?

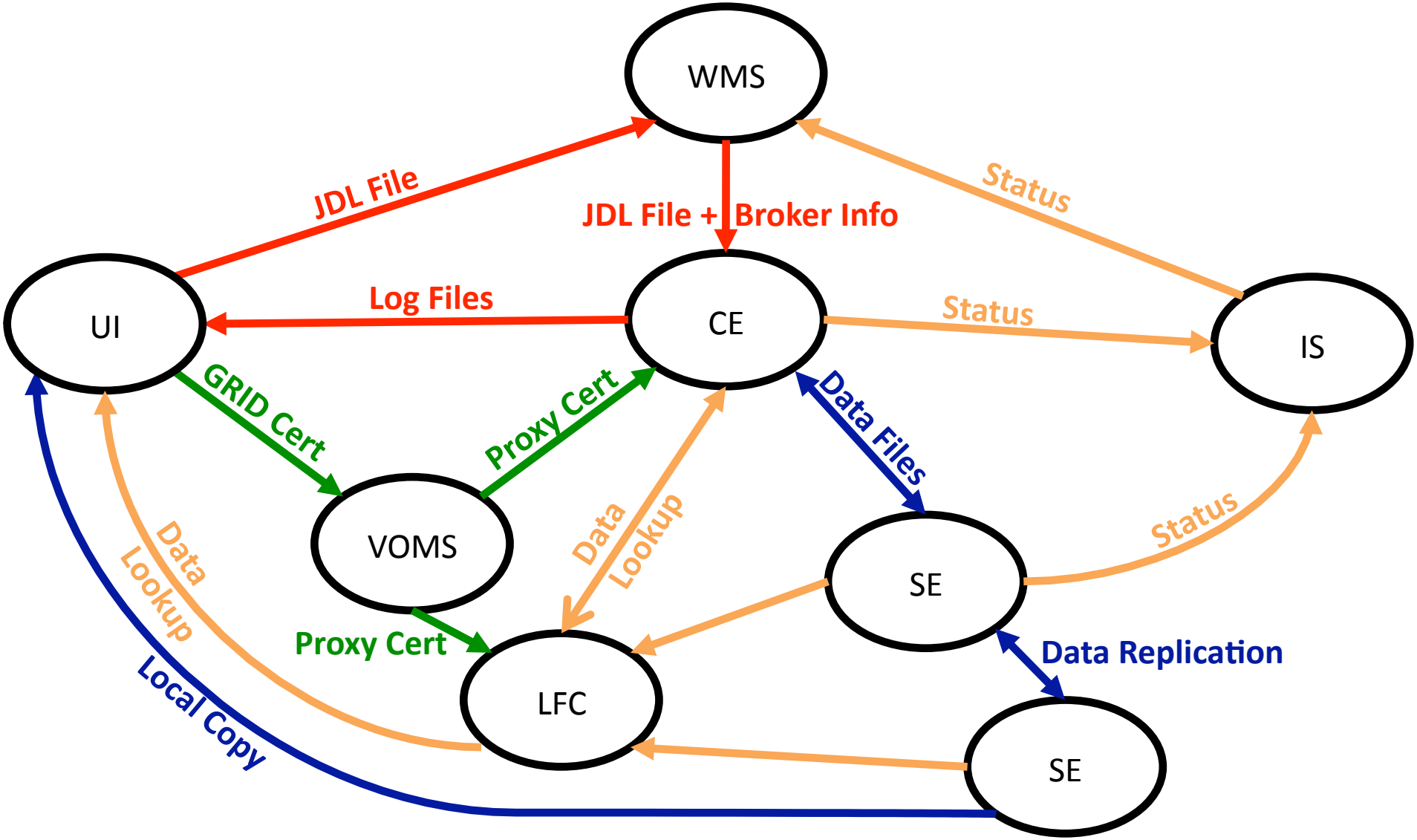
- Global maturing infrastructure with support
  - CPU and disk all ready for us to use in a number of locations
  - Computing experts develop tools and offer support
- Cheap CPU and storage
  - Part of the general research infrastructure in some countries.
- Secure authorisation for data and resource access.

T2K on the GRID at…

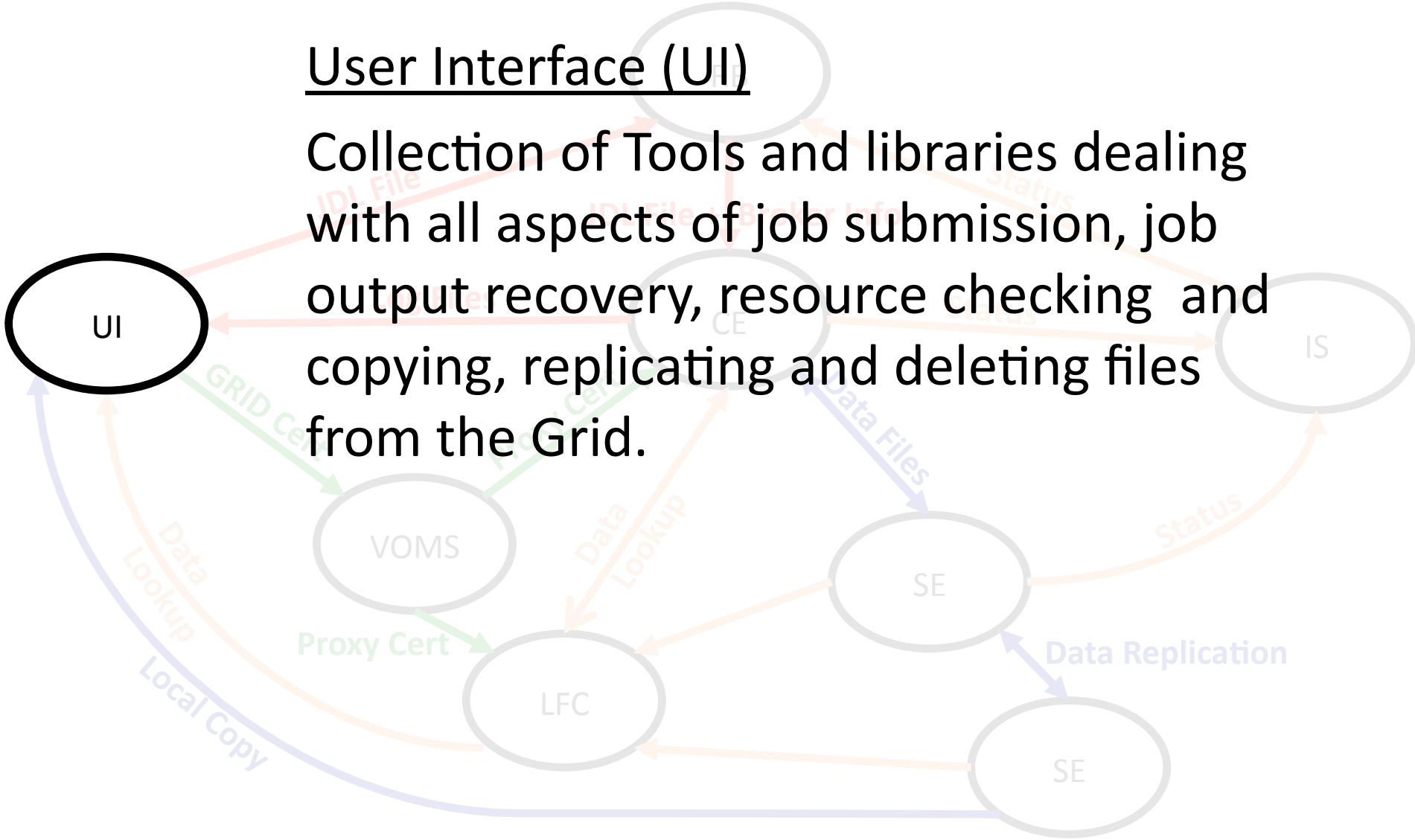
- VOMS
  - voms.gridpp.ac.uk
- Logical/LCG File Catalogue
  - lfc.gridpp.rl.ac.uk
- Storage elements
  - RAL-LCG2
  - UKI-LT2-IC-HEP
  - UKI-LT2-QMUL
  - UKI-NORTHGRID-LIV-HEP
  - UKI-NORTHGRID-MAN-HEP
  - UKI-NORTHGRID-SHEF-HEP
  - UKI-SOUTHGRID-OX-HEP
  - VICTORIA-LCG2
- Computing Elements
  - IN2P3-CC
  - RAL-LCG2
  - UKI-LT2-IC-HEP
  - UKI-LT2-QMUL
  - UKI-LT2-UCL-HEP
  - UKI-NORTHGRID-LANCS-HEP
  - UKI-NORTHGRID-LIV-HEP
  - UKI-NORTHGRID-MAN-HEP
  - UKI-NORTHGRID-SHEF-HEP
  - UKI-SOUTHGRID-OX-HEP
  - UKI-SOUTHGRID-RALPP

cf: http://gstat-prod.cern.ch/gstat/vo/t2k.org/

GRID Overview



Authorisation



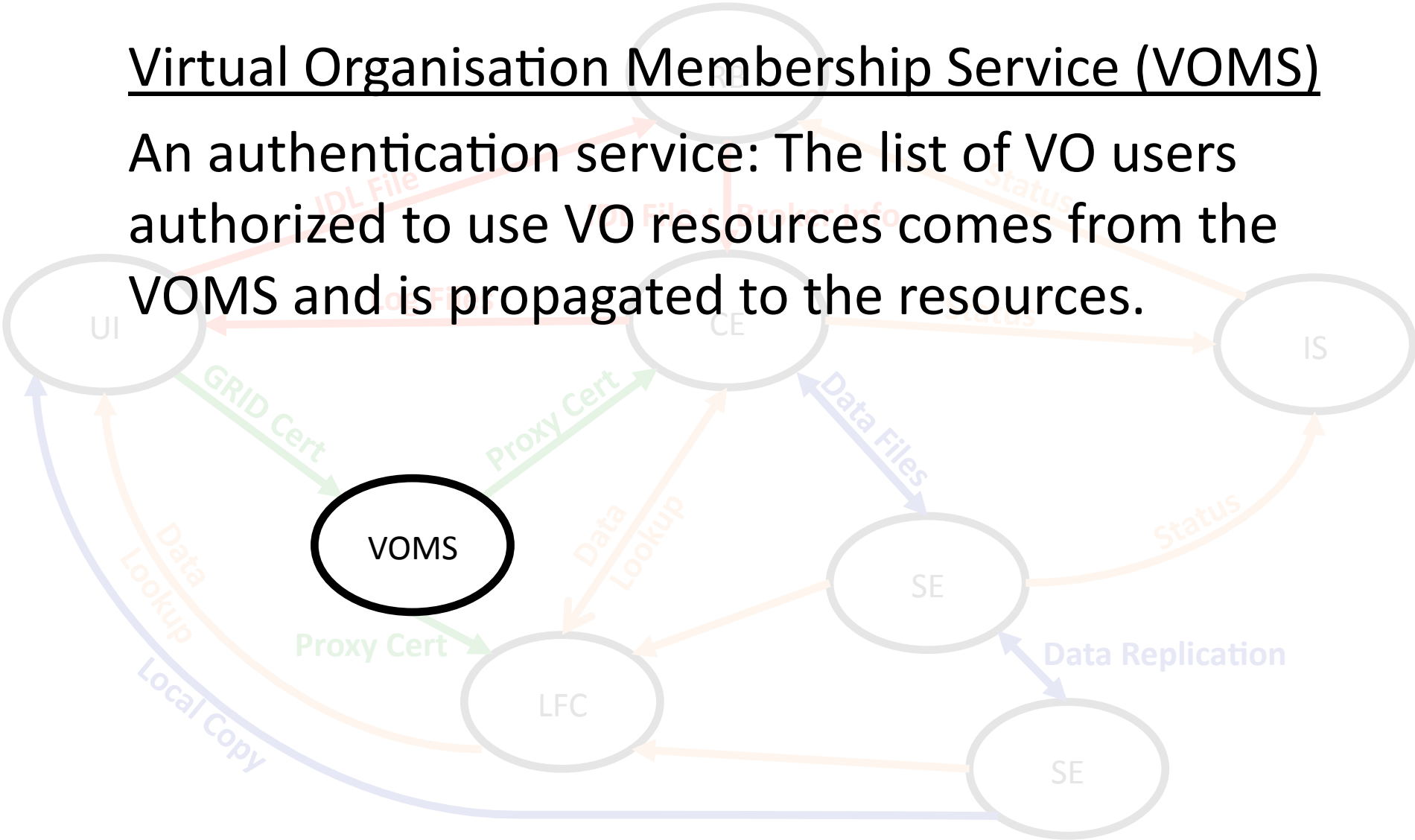
User Interface (UI)

Collection of Tools and libraries dealing with all aspects of job submission, job output recovery, resource checking and copying, replicating and deleting files from the Grid.

Authorisation

Virtual Organisation Membership Service (VOMS)

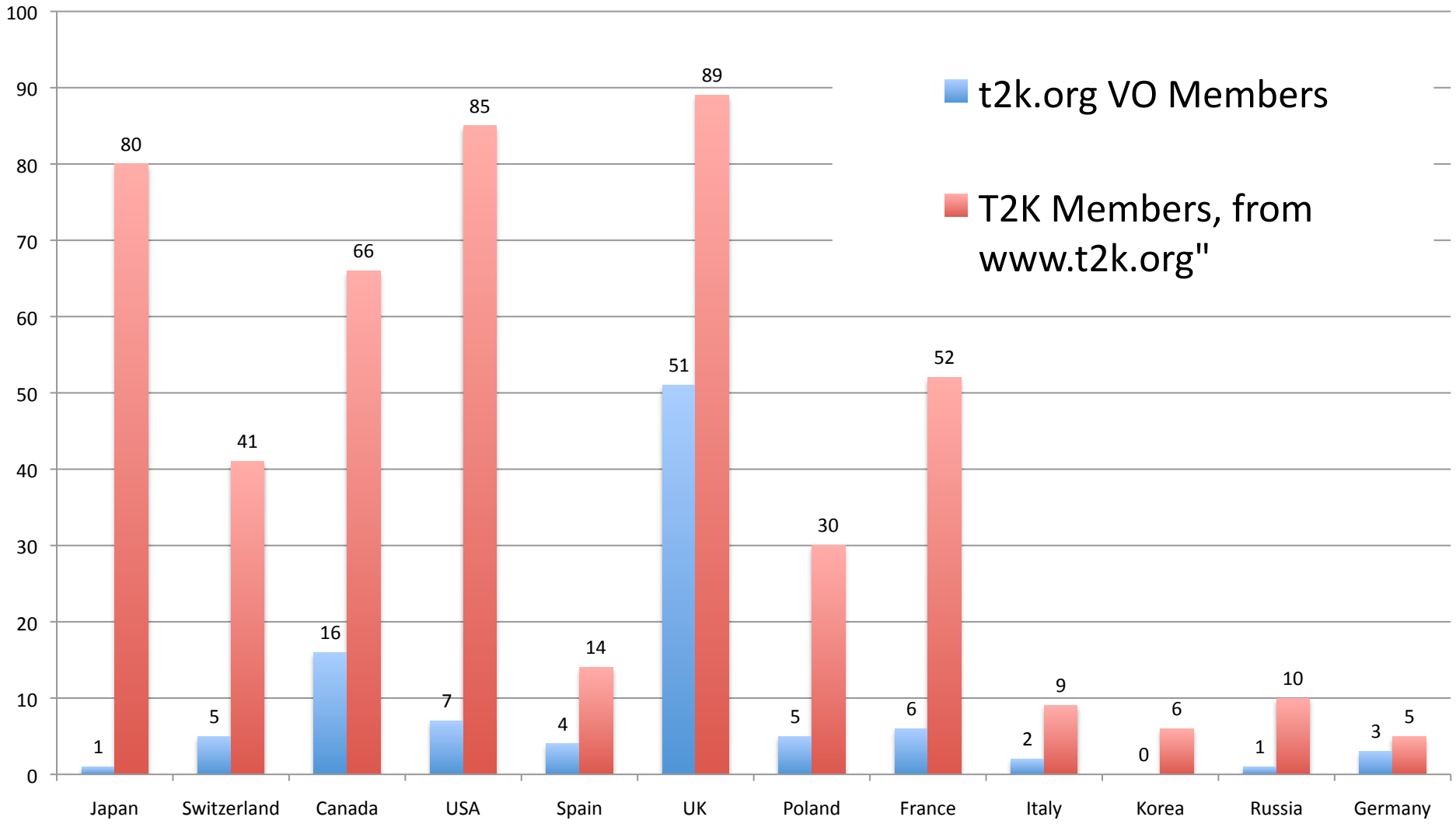
An authentication service: The list of VO users authorized to use VO resources comes from the VOMS and is propagated to the resources.



Authorisation: t2k.org Virtual Organisation (VO)

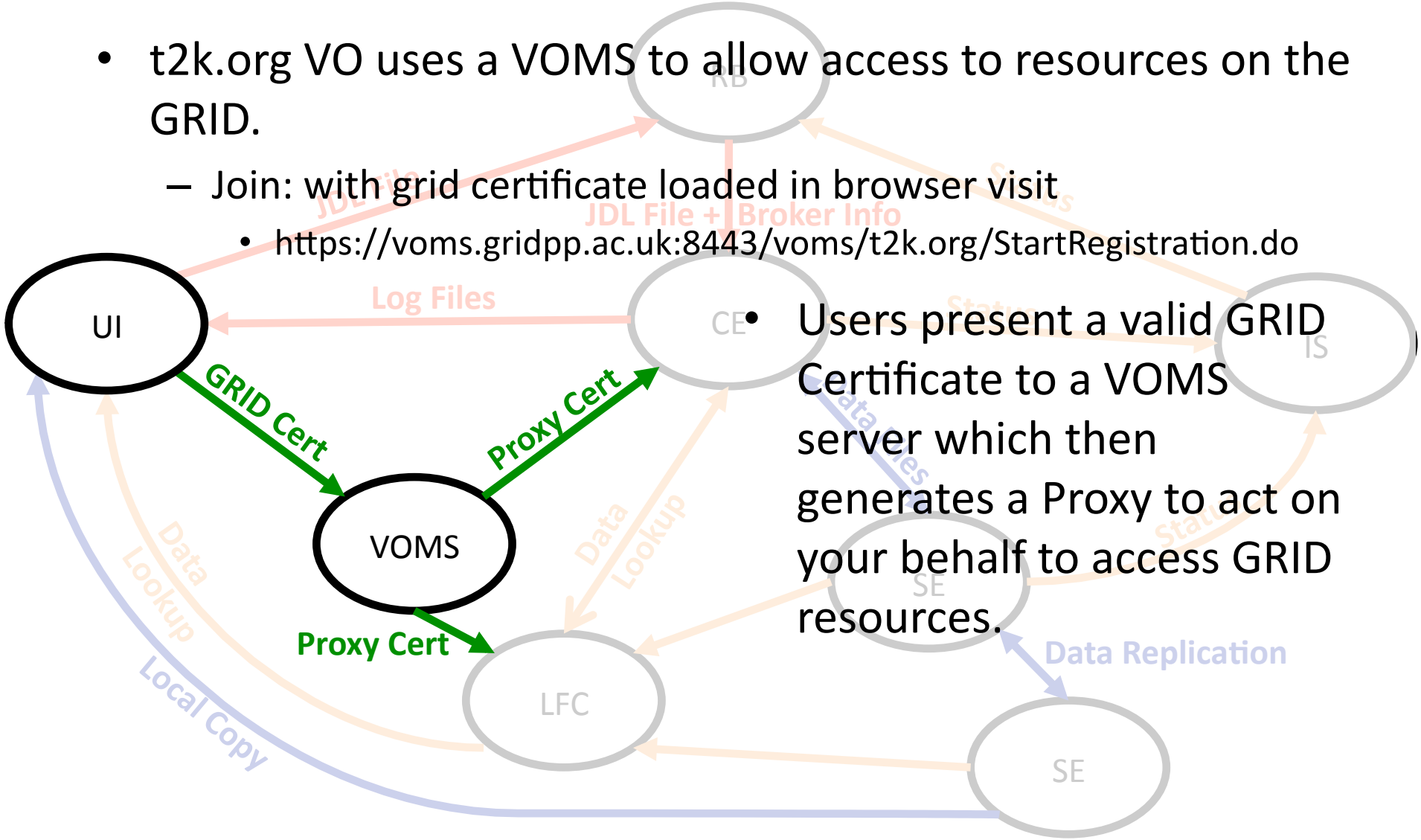
- Hosted by voms.gridpp.ac.uk VOMS server
- 100+ members across all collaborating countries
- t2k.org VO members may have privileges assigned to them via a 'Role':
  - VO-Admin: Manage VO and deal with security.
  - lcgadmin: Install/Maintain Software
  - production: Run production jobs

VO Membership



Authorisation

- t2k.org VO uses a VOMS to allow access to resources on the GRID.
  - Join: with grid certificate loaded in browser visit
    - https://voms.gridpp.ac.uk:8443/voms/t2k.org/StartRegistration.do



- Users present a valid GRID Certificate to a VOMS server which then generates a Proxy to act on your behalf to access GRID resources.

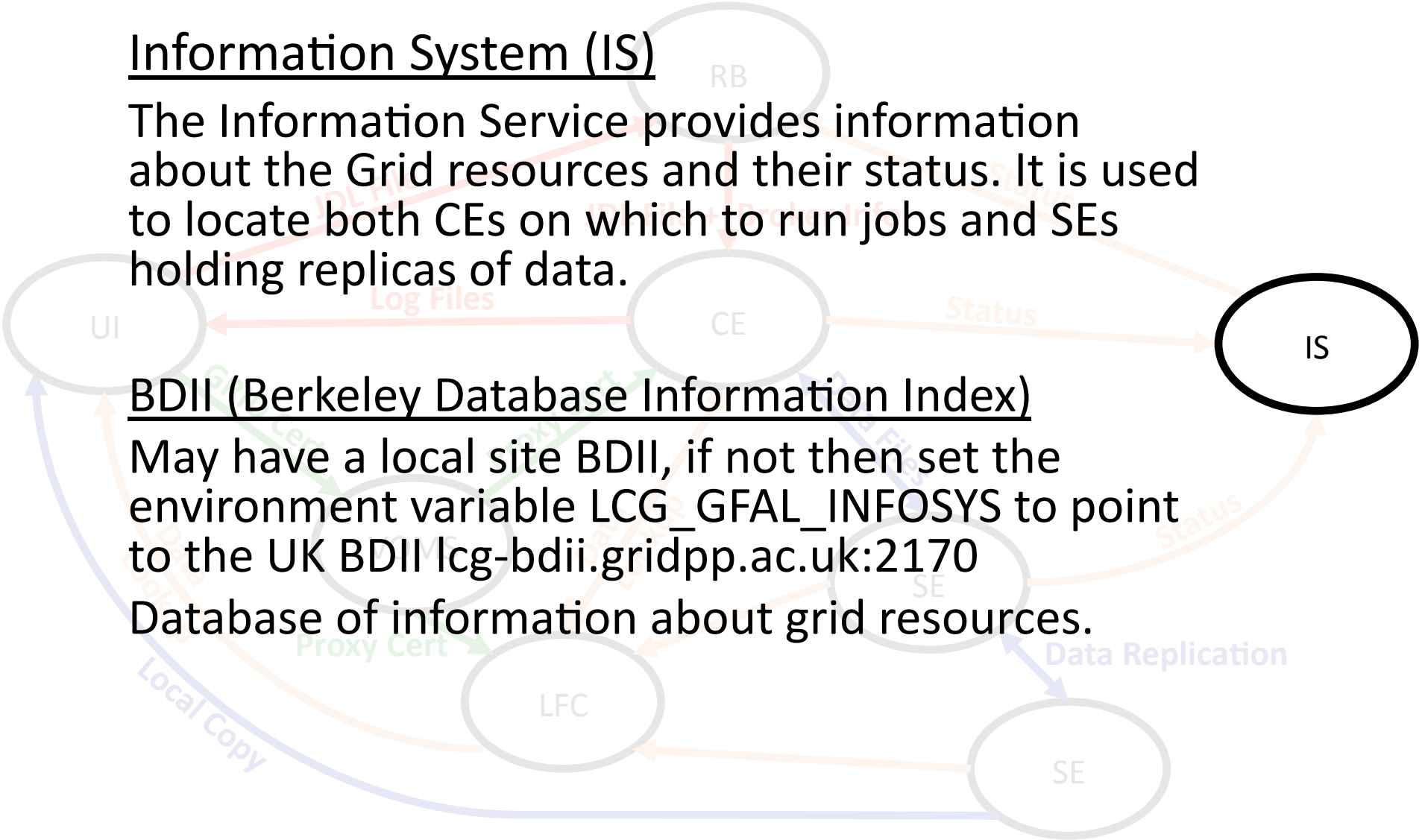
Task 1: Generating your first proxy

1. Get and prepare GRID Certificate
   - http://www.t2k.org/nd280/datacomp/howtoaccessdata/RegionalGrid
2. Join the t2k.org VO
   - With the certificate loaded in the browser visit
     - https://voms.gridpp.ac.uk:8443/voms/t2k.org/StartRegistration.do
3. Generate the voms proxy
   - $ voms-proxy-init -voms t2k.org

Information

Information System (IS)

The Information Service provides information about the Grid resources and their status. It is used to locate both CEs on which to run jobs and SEs holding replicas of data.



BDII (Berkeley Database Information Index)

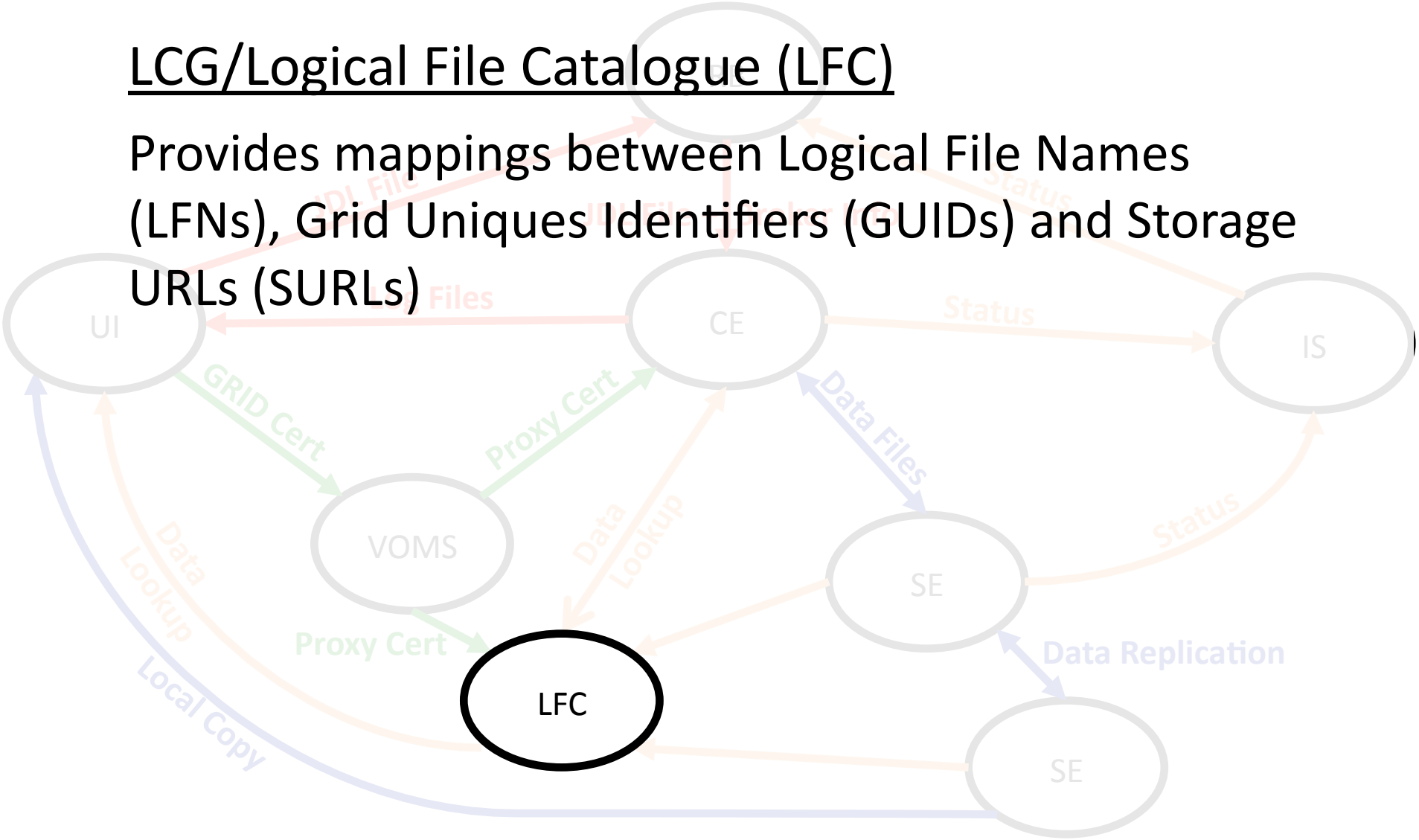
May have a local site BDII, if not then set the environment variable LCG_GFAL_INFOSYS to point to the UK BDII lcg-bdii.gridpp.ac.uk:2170

Database of information about grid resources.

Data Storage, Retrieval and Replication

LCG/Logical File Catalogue (LFC)

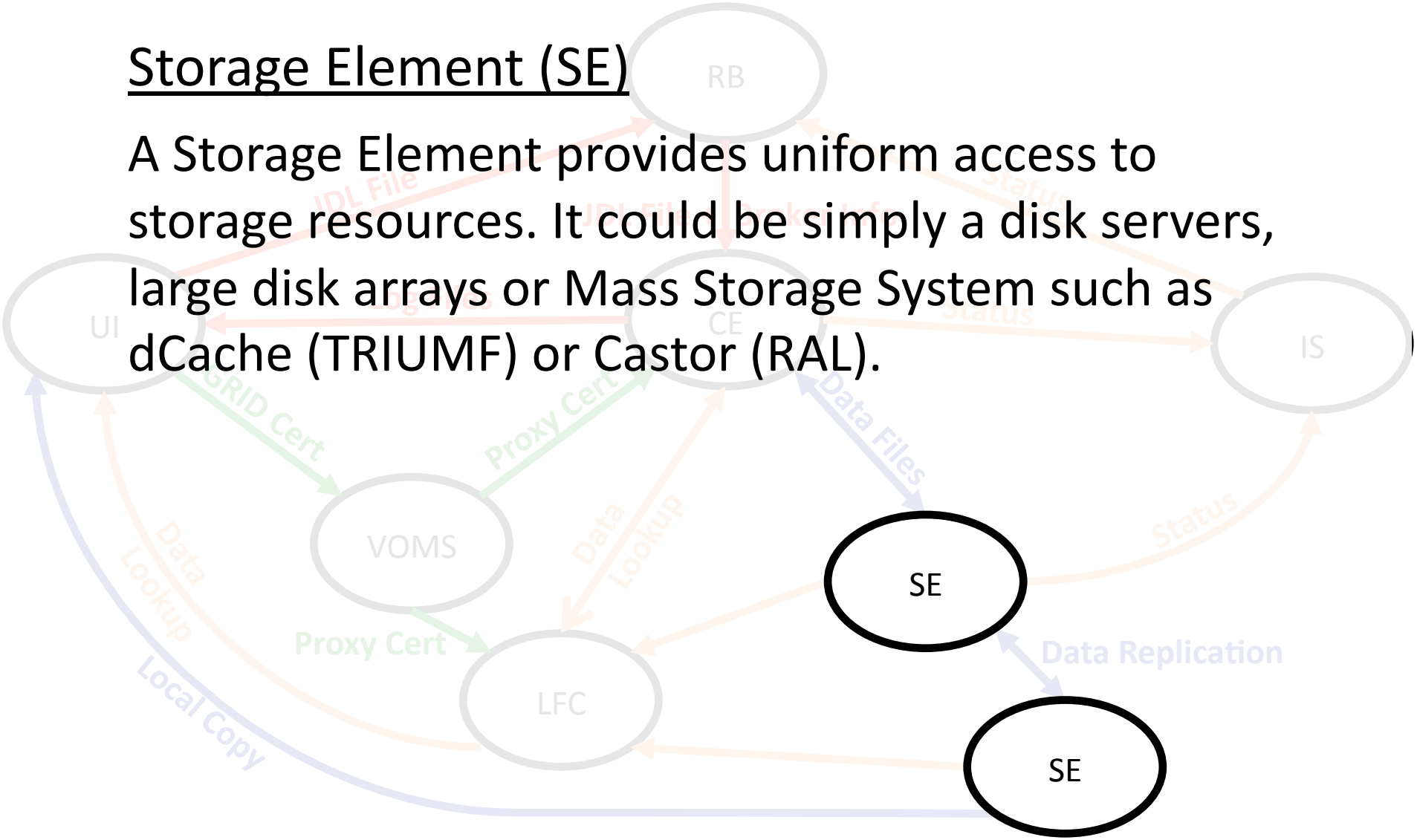
Provides mappings between Logical File Names (LFNs), Grid Uniques Identifiers (GUIDs) and Storage URLs (SURLs)



Data Storage, Retrieval and Replication

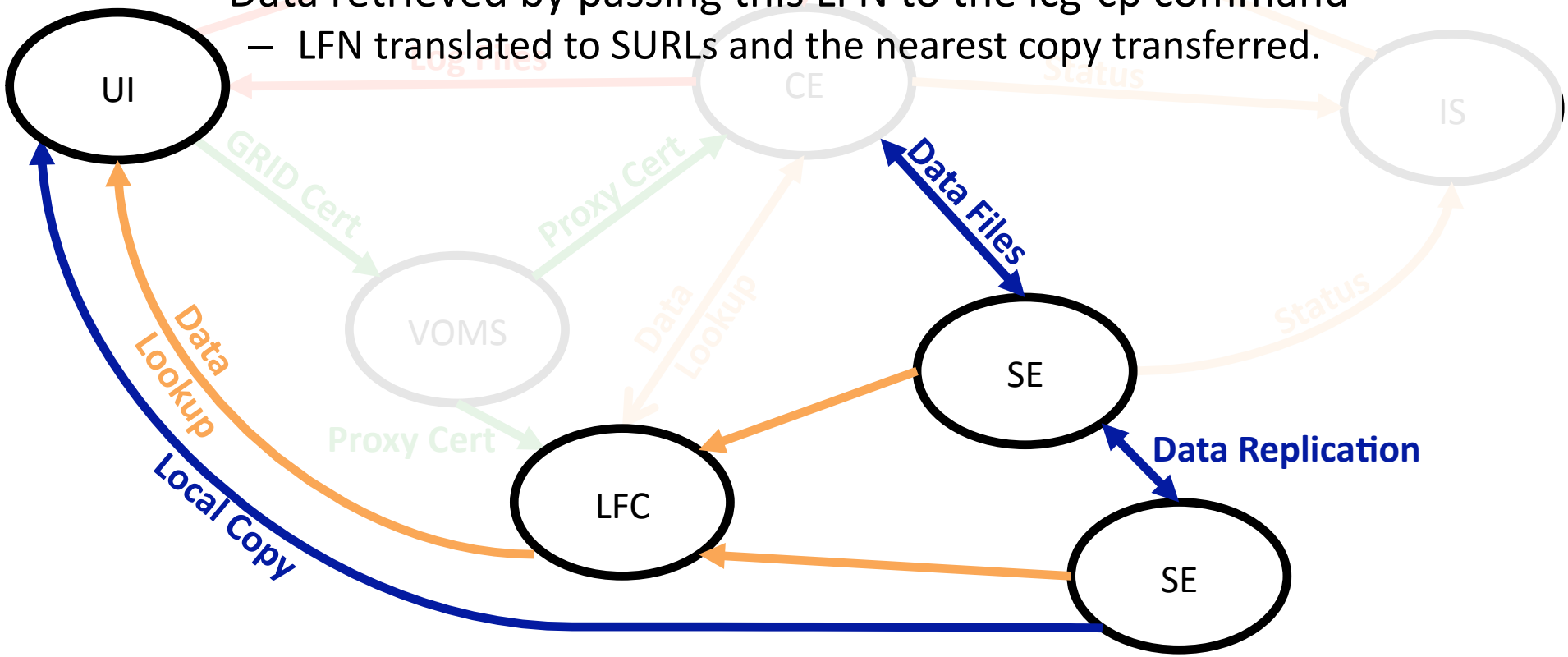
Storage Element (SE)

A Storage Element provides uniform access to storage resources. It could be simply a disk servers, large disk arrays or Mass Storage System such as dCache (TRIUMF) or Castor (RAL).



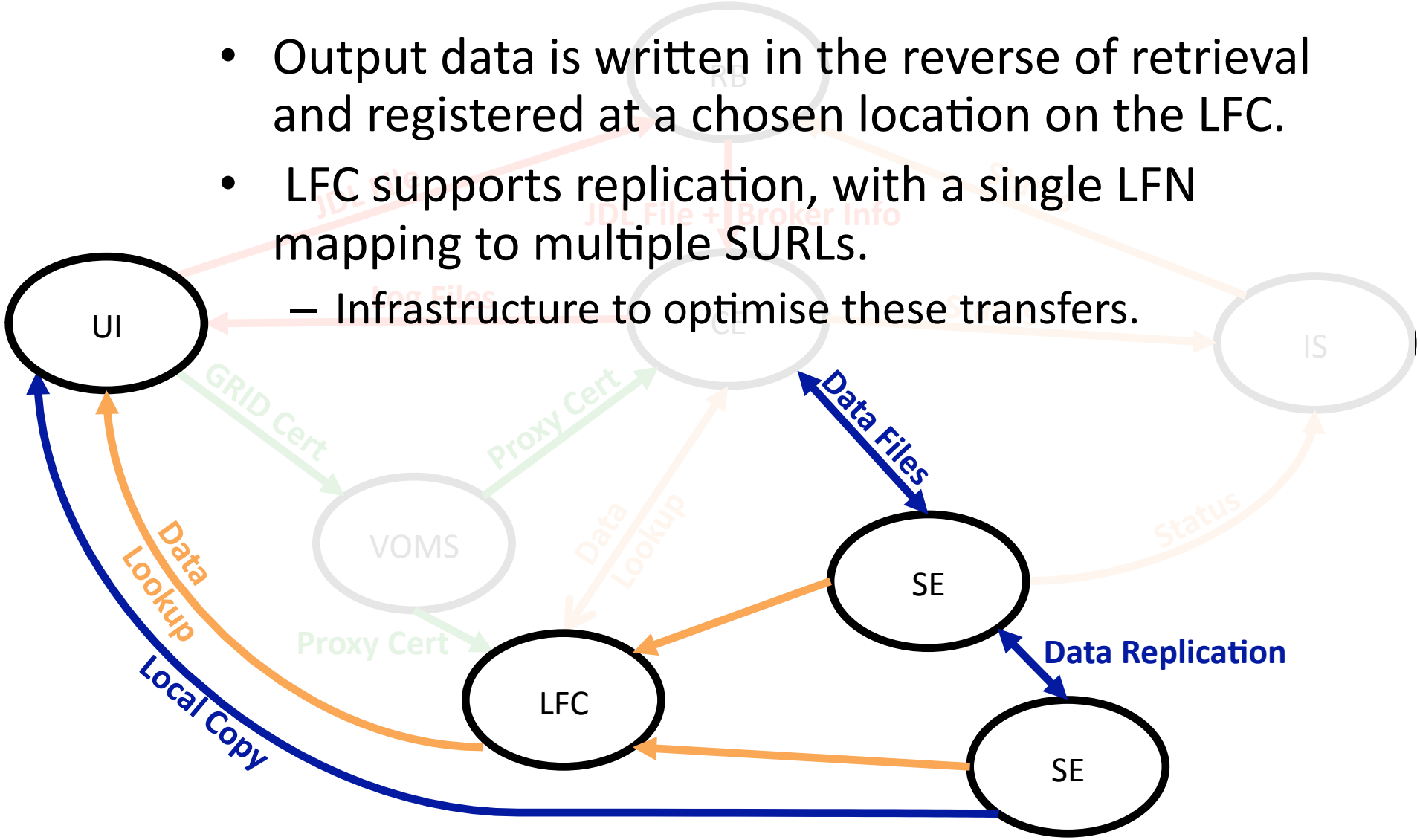
Data Storage, Retrieval and Replication

- Files are refered to symbolically using LFNs (Logical File Names)
  - LFNs passed to LFC server which translates them into SURLs (Storage URLs) which are physical locations on an SE.
- Data is listed and browsed in the LFC by LFN's
- Data retrieved by passing this LFN to the lcg-cp command
  - LFN translated to SURLs and the nearest copy transferred.



Data Storage, Retrieval and Replication

- Output data is written in the reverse of retrieval and registered at a chosen location on the LFC.
- LFC supports replication, with a single LFN mapping to multiple SURLs.
  - Infrastructure to optimise these transfers.



Task 2: Getting Data

1. Let the UI know where the LFC is
   - Set the environment variable LFC_HOST to lfc.gridpp.rl.ac.uk
2. Browse the file system to find the file
   - $ lfc-ls /grid/t2k.org/nd280/raw/ND280/ND280/00003000_00003999/nd280_00003999_0000.daq.mid.gz
3. Copy the file locally
   - $lcg-cp lfn:/grid/t2k.org/nd280/raw/ND280/ND280/00003000_00003999/nd280_00003999_0000.daq.mid.gz nd280_00003999_0000.daq.mid.gz

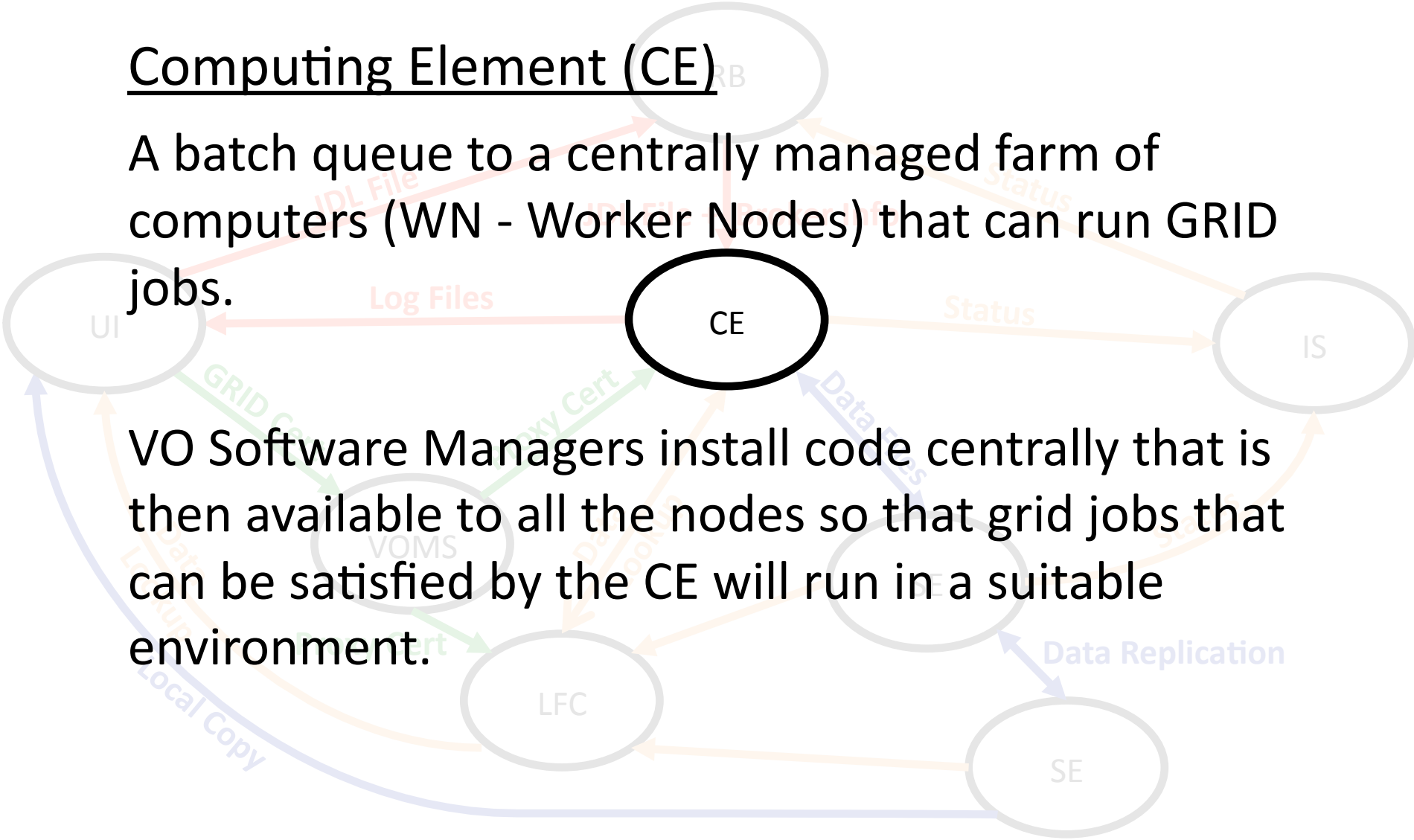
T2K Directory Structure

- Raw
  - /grid/t2k.org/nd280/raw/ND280/ND280/[00000000_00000999,00001000_00001999,0...,00004000_00004999,...]/
- Processed
  - /grid/t2k.org/nd280/[v7r19p1,v7r...]/[anal,cali,reco,unpk]/ND280/ND280/[00000000_00000999,00001000_00001999,0...,00004000_00004999,...]/
- MCP1
  - /grid/t2k.org/nd280/mcp1/[genie,neut]/[geobaseline E.g. 2010-02-water]/[magnet,basket]/[beam,cosmic, ccpizero,ncpizero]/[anal,cali,reco,unpk]/
    - The latter ncpizero,ccpizero is only available in the basket.
    - The geobaseline is the geometry baseline chosen plus water or nowater.
    - Currently only water in P0D has been generated.
- Here lists in [ ] parentheses are options

Job Submission

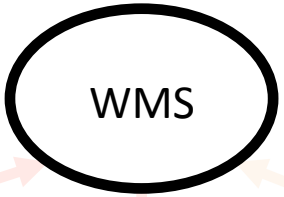
Computing Element (CE)

A batch queue to a centrally managed farm of computers (WN - Worker Nodes) that can run GRID jobs.



VO Software Managers install code centrally that is then available to all the nodes so that grid jobs that can be satisfied by the CE will run in a suitable environment.

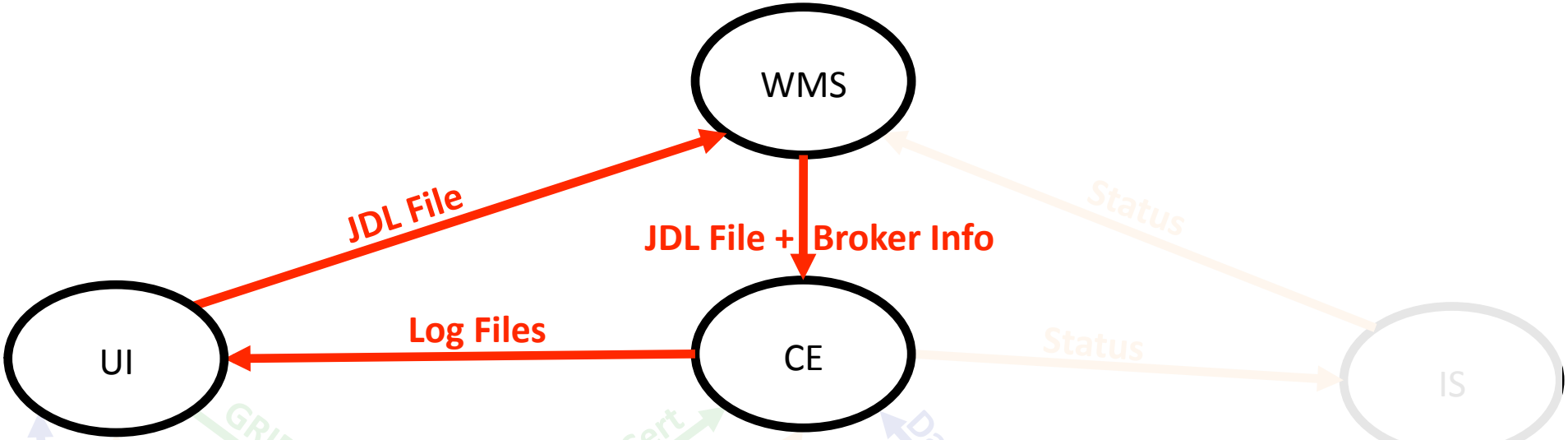
Job Submission



Workload Management System (WMS)

The WMS examines Requirements and Rank expressions (and also any data requirements) in the JDL. All CE are filtered against the Requirements, and the Rank is calculated for all the CEs which match and on the basis of Rank determines which CE to submit the job to. For that purpose, the WMS must retrieve information from the LFC

Job Submission



- Job submission via a JDL (Job Definition Language) file
  - Sent from a UI to a RB which examines the JDL and determines, with the help of the IS (Information Service), the best CE to run the job.
- You can send small user files, typically the script file to run, along with the JDL in what is called the JDL Input SandBox.
- Once the job is complete the the standard and error output, along with any small output files, e.g. log files, as specified in the JDL Output SandBox are returned to you.

Job Description Language (JDL) Files

- JDL is the language used to specify the resources that a job requires.
- To submit a job to the GRID a JDL file is created and passed to a WMS that examines the JDL and determines, with the help of the IS the best CE on which to run the Job.

Task 3: Hello World

Located at http://pprc.qmul.ac.uk/~still/GRID/helloworld.jdl

```
###############Hello World###################
Executable = "/bin/echo";
Arguments = "Hello welcome to the Grid ";
StdOutput = "hello.out";
StdError = "hello.err";
OutputSandbox = {"hello.out","hello.err"};
VirtualOrganisation = "t2k.org";
Requirements=Member("VO-t2k.org-ND280-v7r21p1",other.GlueHostApplicationSoftwareRunTimeEnvironment);
##########################################
```

Task 3a: List Software

Located at http://pprc.qmul.ac.uk/~still/GRID/listSoft.jdl

```
############### List Software Dir ###################
Executable = "/bin/ls";
Arguments = " -R $VO_T2K_ORG_SW_DIR/ ";
StdOutput = "list.out";
StdError = "list.err";
OutputSandbox = {"list.out","list.err"};
VirtualOrganisation = "t2k.org";
Requirements=Member("VO-t2k.org-ND280-v7r21p1",other.GlueHostApplicationSoftwareRunTimeEnvironment);
#############################################
```

More on JDL Files

- Comments using ##
- Executable: resident or uploaded file
- Arguments: Arguments to pass to the executable
- StdErr/StdOut: The name of file to store the std::err and std::out of the job
- InputSandbox: Any ***small*** files to be used by the job E.g. executable scripts or configuration files.
- OutputSandbox: Any ***small*** files to be sent back when the job output is collected
- VirtualOrganisation: the VO associated with this job.
- Requirements: Can specify information for the IS to select certain sites for processing.
  - Software tags are the only thing we currently plan to use in t2k.
    - Of the form VO-t2k.org-*SOFTWARE-VERSION,* E.g. VO-t2k.org-ND280-v7r21p1

Task 3: Hello World

a. Create the JDL
b. Get the config file from
   - http://pprc.qmul.ac.uk/~still/GRID/autowms.conf
c. Submit the job:
   - $ glite-wms-job-submit -a -c autowms.conf –o helloworld.jid helloworld.jdl
     - Config containing defaults
     - jid file: stores the job is and is used later to check status and get OutputSandbox of job.
     - The jdl file to send to the WMS.
d. Check the status
   - $ glite-wms-job-status –I helloworld.jid
e. Get the output
   - $ glite-wms-job-output --dir /some/dir/ -I helloworld.jid

T2K Software On the GRID

- VO Software are defined by environment variable
  - VO_T2K_ORG_SW_DIR
- Only lcg-admin users can install/modify software area
- Current software:
  - ND280 (at some sites)
    - $VO_T2K_ORG_SW_DIR/nd280v#r#p#
  - GENIE (at QMUL and RAL)
    - $VO_T2K_ORG_SW_DIR/GENIE
  - There is a setup.sh in each directory which sets up the environment

T2K Future On the GRID

- nd280Computing package to be a central resource for processing, data management and software installation tools and scripts.
- nd280Computing will be modified to work with Ganga, a job submission package developed by LHCb and Atlas.
  - Ease book keeping
  - Track failures
  - Automate processing

nd280Computing (*IN DEVELOPMENT*)

- New Python package in T2KRepository
- Many features to run locally and on a GRID system
  - Tools: Handy classes with GRID and local features, used by scripts
  - Installation Scripts: Scripts to install ND280 and GENIE plus more to follow.
  - Processing Scripts: Scripts to process data and MC
  - Data Scripts: Scripts to duplicate data, synchronise directories and check integrity.

nd280Computing (*IN DEVELOPMENT*)

- Script is the executable to be run in JDL file.
- Upload all tools as well in InputSandbox using the wild card tools/*.py
- May decide to upload to software areas in future depending on how large the package becomes.
- I will work on integration with Ganga.

Volunteers

- Software Installation and verification (2 people)
  - Run install jobs for frozen releases
  - Clean up software areas of old releases
  - Maintain software tags
- Processing data and MC (n people)
  - Responsible for a subset of processing and associated bookeeping
- Data integrity (2-3 people)
  - check the existence of data on the GRID
  - Oversee site-site duplication of data

Further Info

- Nick West has written some great informstion from his years looking after GRID in the UK for the MINOS experiment
  - https://www-numi.fnal.gov/offline_software/srt_public_context/GridTools/docs/index.html
- The CERN twiki is also another good source of information, E.g. the Atlas workbook
  - https://twiki.cern.ch/twiki/bin/view/Atlas/WorkBook
- I will write, and encourage others to write, t2k.org specific information in a similar format.