WP10 macro use cases :
common and specific mini use cases

V. Breton, N. Jacq, J. Montagnat

AWG, Barcelona

2003/05/14

WP10 macro use cases

◆ Bioinformatics : Data analysis with a regularly updated database

- exponential growth of the data => exponential growth of jobs computing time

- need of updated version of public and private databases

- need of metadata to manage data (versioning, property, dependencies…)

◆ Medical Imaging : handling medical images in a grid environment

- Automatic indexation of new images

- Image processing using a registered algorithm

- Complex query

- Pipeline processing

- Interactive modeling

Steps of Bioinformatics use case

◆ Update of public or private databases on the Grid

  ▪ archive of the old version, storage of the new version on a master SE, replication

◆ Distributed job submission

  ▪ splitting of the job input to reduce computing time

◆ sharing of data on the Grid

  ▪ storage of results and their metadata
  ▪ access to data with metadata

Steps of medical use case

- A new image is produced and registered in the medical data system

- A patient queries the medical data system to retrieve one of its images

- A physician sends an image to a colleague for his expert diagnosis

- A physician or a researcher retrieves a patient image, performs a processing and stores the output
  - Pipeline processing, parallel processing, high priority processing

- A researcher queries the medical system to retrieve all images with given parameters

- A researcher retrieves all images obtained from an original image

- A physician retrieves the original image used to obtain a processed image

- A researcher deletes an image previously registered

Security on the medical data system

- ◆ Access rights must be defined at the file level

  - Individuals have free read access to their data. They can query the data management system to retrieve their own data

  - Physicians have free read access to their patient's data. They may delegate their full access right to another physician or limited access right to a researcher. A physician may query the medical information system to find patient data. The patient or another physician may grant him the access.

  - Researchers may access to blanked images if authorized by a physician who has access right to these images.

  - Every other grid user, including not accredited local system administrators) should not be able to read nominative data

- ◆ Data "anonymization" is a service provided

  - The information concerning the patient is separated from the image itself. This way, the image can be used for research purposes while preserving patient anonymity.

- ◆ Abnormal data usage must be detected

Common and specific mini use cases

◆ All HEP mini use cases are used by WP10 applications

◆ Specific WP10 mini use cases

- Time constrained job submission (medical imaging)
  - Dataset access cost evaluation, resource reservation, resource estimation related to time constrained job submission

- Parallel job submission (medical imaging, bio-informatics)
  - MPI needs

- Metadata anonymization (medical imaging)

- Transfer of access rights to a file from one grid user to the other

What is not clear in HEPCAL document

◆ Access rights to files : who owns data on the grid ?

◆ For mini use case "Physical data set instance deletion", the user needs deletion rights on the data set