Overview of recent ATF activities

Leanne.Guy@cern.ch http://cern.ch/leanne

Overview

- ATF reformed in early 2002
- Meets for 2 days every 1-2 months
- Began by looking at the interactions between the different components from each WP
- Moved then to a component based approach as opposed to WP based.
- Application use case documents were basis for the architectural analysis

Role of the ATF

- **ATF does not intend to define the EDG architecture**
  - focus is on the merits on one architecture vs another
- **Reconciliation of the designs from the various work packages**
  - WPs began to design components in isolation
  - Interdependencies are complex
  - Need to ensure one interoperable & efficient grid

ATF – Aims & Goals

- **Definition of TB2 architecture**
  - Lessons and feedback from TB1 was used to analyze the architecture,
  - Many problems and misunderstandings between WPs were found.
- **Assist middleware WPs to understand the interaction between their components**
- **Applications can understand how to build application layers on top of the middleware**
- **Formal documentation of EDG architecture**

Methodology (1)

- **Based on the Class-Responsibilities-Collaborators technique [1]**
- **Each middleware package defines one or more classes**
  - classes represent components the WP designs
- **An example use case is taken from the application areas**
- **ATF members act out the interactions between their classes to execute the use case**

Methodology (2)

- **Think about the integrated system**
  - not the individual components only.
- **Decomposition by WP was broken down**
- **Many misunderstandings between WPs became apparent**
  - better now than at integration time !
- **Results captured in UML**
  - sequence diagrams to capture message flow between components over time
  - class diagrams show static relationship between objects

Use cases covered (1)

- **Create grid resources**
  - create computing element
  - create VO
  - assign CE to VO
- **Grid login**
  - obtain grid certificate
  - obtain VO authorization
  - grid login

Use cases covered (2)

- **Job submission**
  - simple job submission
    - no data management services
  - job submission with static file selection
    - all LFN required are known at job submission time
  - job submission with dynamic file selection
    - data required are determined at execution time
- **Output data access**
  - fetch output data
    - data returned to user with no registration
  - register job output
    - output data stored on an SE and registered in the RLS & RMC

Job with static file selection (1)

1. **Job submitted at UI**
   - LFN needed are specified in the JDL
2. **The WMS generates a list of candidate CEs**
   - interaction with the Information system to CEs that accept this VO's jobs
   - CEs filtered against job requirements specified in the JDL
3. **ERM calculates costs for each LFN per CE**
   - WMS calls getAccessCost(lfn[],ce[])
   - ERM returns:
     - best PFN/LFN and the data cost associated for each CE
       - combination of network and SE costs (WP7 & WP5)
     - size of files that need to be copied to the close SE

Job with static file selection (2)

4. **WMS ranks CEs again**
   - CE ranking is combined with ERM ranking to produce new CE ranking
5. **Space reservation on SE**
   - if there is not enough space available, WMS will query the closeSE of the next ranked CE. (currently not guaranteed)
   - job is scheduled when space is reserved
6. **Replication of files triggered by WMS**
   - WMS calls replicateFile(LFN,SE) to transfer data between SEs
   - WMS requests SE to pin files for duration of job
7. **Job runs**
   - PFN on closeSE obtained from brokerinfo file by job

Job with static file selection (3)

Questions

- the WM uses replicate file to start replication. At some earlier point we used initFilePrefetch. Now the WMS calls getBestFile. What should be used

- if initFilePrefetch is used, how do we ensure that the file transfers are finished before the job starts

Job with dynamic file selection (1)

- **Differences to static file selection example**
  - all LFNs required by job not known apriori
- **In TB 2, job speaks POSIX only**
  - transfer files to closeSE
  - make a temporary copy to the WN scratch space
- **Job copies file to local scratch space**
  - issues MD query to extract LFNs
  - ask ERM for be PFN w.r.t CE
  - job copies PFN to local scratch space of WN

Job with dynamic file selection (2)

- **Job requests replication to closeSE**
  - calls getBestFile for LFN
  - PFN returned
  - ERM initiates replication to closeSE

Registration of job output

- **Interaction between ERM – SE- Job wrapper**
- **Job wrapper asks ERM to store file on SE with LFN**
- **ERM generates a GUID for the file**
  - assocation between GUID and LFN inserted into RMC
- **ERM requests SE to return URL as destination for file**
  - ERM is expected to specify PFN for file & attributes
- **ERM sends file to SE (URL) via gridFTP**
  - tells SE to commit file
- **GUID – PFN mapping stored in RLS**
- **RMC updated if additional information exists**

Architectural assumptions

Description	Comments	Priority
Sandboxes fill up on RB	In future sandboxes will be stored on the SE not the RB	Low
How does the UI know the VOMS server name when voms-proxy-init is issued	Default is that it is stored in the home directory, needs further discussion	Low
Outbound connectivity from worker nodes	Needs further discussion	Deferred
Service authentication and authorization (RB & ERM). RB uses delegation, ERM has server certificate	Still under discussion. Need to ensure traceability, some sites prefer a proxy if access is per user and not VO (denied user list)	Medium
Submit job with AdHoc files. Application should provide an estimate of the storage space but RB will not match on SE attributes	Storage requirements of 10s MB not a big problem	Low

References

[1] Kent Beck Ward Cunningham. A laboratory for teaching object oriented thinking. Proc. OOPSLA 1989

[2] EU DataGrid: Software Architecture Models. WP12-NOTE-368971. EU DataGrid Architecture Task Force. Editor: Lee Momtahan