Report from Middleware Working Group.

Kai Leffhalm, Hubert Simma

DESY, December 13, 2013

20th ILDG Workshop

HELMHOLTZ GEMEINSCHAFT

DESY

MWWG

New MWWG Members

- Massimo Di Pierro(USQCD)
- Oliver Witzel (USQCD)

Discussions and Activities

- ILDG requirements in Europe were raised in EoI for data services (by Simlab Cyprus-DESY-Jülich in reply to PRACE call)
- Porting of Itools to new OS'es (for LDG and others)
- Tackling generation of metadata generation (see also MDWG discussion)
- Policy and implementation of restricted access

Generation of XML for Configurations

. . . seems to be

- a difficult and painful work
- impossible without parsing of log files
- something that should be done by a computer

But the XML schema for configs has a total of just 18 mandatory elements:

- **4** of them can or need to be determined automatically when uploading the config
- **10** of them might be filled with dummy content, but in practice, something should really be known or acknowledged
- **1** of them must be known (precision of the simulation)
- **1** of them is typically empty (if algoritm parameters are not changed during the Markov chain, as they should not anyhow)
- the **remaining 2** have an obvious (default) value

A closer look to QCDmlConfig

The 18 mandatory elements in `QCDmlConfig1.3.1.xsd` are (in xpath notation):

(1) management / crcCheckSum
→ compute when uploading

(2–5) management / archiveHistory / *: who generated the config?
→ name and institution should be known / acknowledged
→ date could be approximate or dummy (1970-01-01 00:00:00)
→ fallback: who is uploading the config

(6–8) implementation / machine / *
→ something should be known (at least acknowledge host institution)
→ fallback: UNKNOWN

(9–11) implementation / code / *
→ something should be known (at least name of the code)
→ fallback: UNKNOWN (not very convincing)

A closer look to QCDmlConfig (cont.)

(12) algorithm / parameters
→ empty (if no changes during Markov chain)

(13) precision
→ should really be known

(14) markovStep / markovChainURI
→ determined when uploading

(15) markovStep / series
→ obvious default (e.g. 1 or “A”)

(16) markovStep / update
→ determined by configuration file

(17) markovStep / avePlaquette
→ re-compute when uploading

(18) markovStep / dataLFN

Generation of XML for Configurations

Possible solution:

- No or minimal changes to XML schema
  (leave all or most elements mandatory as they are!)
- Provide (or adjust existing) script which
  - generates config XML with sensible default values
  - takes update number as mandatory argument
  - reads minimal text input (or corresponding further arguments)
    to specify at least 3 and at most 17 elements,
    e.g. of the form:

```
markovChainURI = ...
dataLFN = ...
precision = single | double
generate/name = ...
generate/institution = ...
generate/date = ⟨at least a year⟩
```

ToDo for ltools

`ltools` are based on grid middleware and are pre-compiled for various OSes (e.g. RHEL54, ubuntu 10.04)

- Grid middle-ware development has been continued by EMI (successor of EGEE) till now
  - Funding in the future is unclear
  - Releases are available for SL6/5 and Debian 6
- Pre-compiling necessary as packages are not everywhere available eg: SLES
- Continuous updates of security features give problems (Even on SL6 latest middleware release had problems with proxies)
- Much of the middleware functionality is already in SL standard repository
  - voms-proxy-init, lcg commands for copying files
- Porting ltools to current OSes is necessary, however
  - Every new release of middleware creates new problems
  - Man power is limited

TODO

Plans and pressing issues:

- Porting of ltools to new OS'es
- Clarify with MDWG changes to XML schema or simple markup tools
- Clarify within ILDG requirements or policy for access restrictions
- Clean up and improve web pages