Metadata Working Group Report

◆Update of Ensemble XML
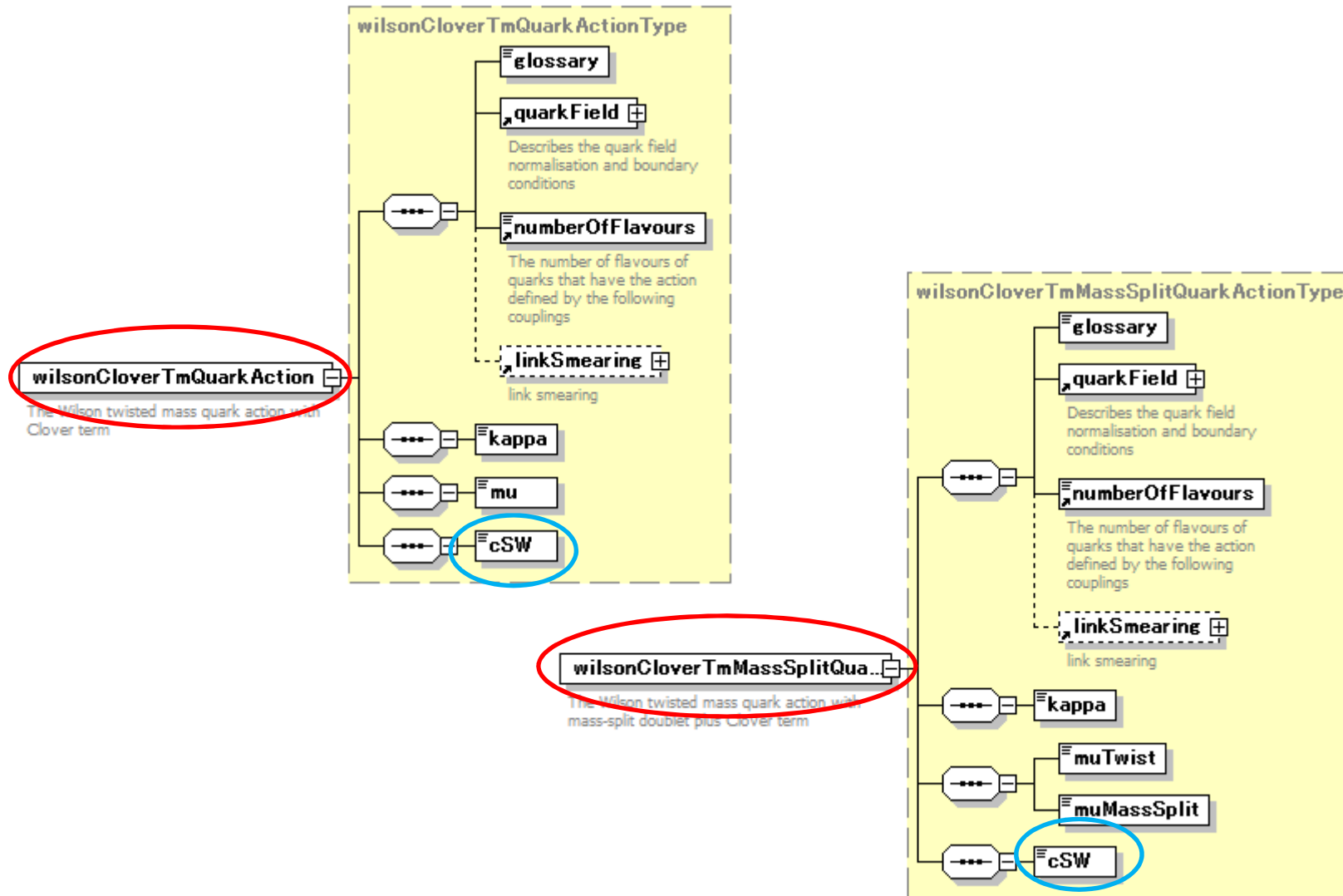◆Simplification of Config. XML
　(still under discussion)

T. Yoshie
CCS, Univ. of Tsukuba
@ILDG21, Dec 13th 2013

Update of Ensemble XML

- Twisted mass quark actions with clover term
  - request from ETMC
  - twisted mass actions were marked up in a previous version
  - realize them as extensions to twisted mass actions
  - simply add a coupling "cSW"
  - released as QCDmlEnsemble1.4.8 on June 8th, 2013

diagrams

Simplification of Config. XML

- Human costs for markup
  - Many collaborations feel that marking up their ensembles and configurations is time consuming
  - started discussions to simplify QCDml (on going)
- Ensemble and/or Configuration
  - cost for marking up ensemble is not high
    - copy an existing sample and edit it
    - has to be done only once for each ensemble
  - marking up configurations requires high human costs
    - method to gather metadata depends on collaborations
    - origin of high costs is that metadata comes from various sources

- Origin of config. metadata (JLDG case)
  - configuration itself for check-sum and plaquette value
  - log file for generation date (in archiveHistory)
  - supplemental file for machine/code information (in implementation)
  - prepare a shell-script to gather metadata and generate a config XML (and convert and pack the configuration)
  - the shell script is different from project to project, thus has to be written and tested for each physics project
- Note
  - Metadata Capture project by Edinburgh-Tsukuba collab.
  - feel not simple enough
- Possible solution
  - make metadata from configuration (and tags for configuration) mandatory
  - make metadata from other sources optional

Minimal Config. XML sample

```
<gaugeConfiguration>
<management>
   <crcCheckSum>347437280</crcCheckSum>
</management>
<precision>double</precision>
<markovStep>
   <markovChainURI>
    mc://JLDG/CP-PACS+JLQCD/RCNF2+1/RC28x56_B2050Kud013560
   </markovChainURI>
   <series>2</series>
   <update>002200</update>
   <avePlaquette>6.04326596981212e-01</avePlaquette>
   <dataLFN>
    lfn://JLDG/CP-PACS+JLQCD/RCNF2+1/RC28x56_B2050Kud013560-2-002200
   </dataLFN>
</markovStep>
</gaugeConfiguration>
```

metadata from configuration

other elements are tags for configuration

- with this simplification
  - it will be much easier to develop tools for ILDG
  - and possibly to incorporate tools to production code

- How to realize this idea (discussion within WG)
  - modify schemata
  - keep schemata unchanged, but allow dummy values
    (see report from Middleware working group)

To do soon

- joint discussion with MWWG on configuration XML
- DOI tag in ensemble XML