CLIC Production Workflow

André Sailer

CERN-EP-LCD

DIRAC User Workshop
May 22, 2018

Production Chain

- Chain of Generation, Split, Simulation, Reconstruction, Replication transformations
- Script sets workflow module parameters (Steering file, detector model, etc.)
- Reads 'conf' file, creates chain of given *ProdTypes*, for each *process* or *prodID*

```
[Production Parameters]
detectorModel = CLIC_o3_v14
version = 2018-04-10
softwareVersion = ILCSoft-%(version)s_gcc62
clicConfig = ILCSoft-%(version)s
productionLogLevel = VERBOSE
outputSE = CERN-DST-EOS
finalOutputSE = CERN-SRM
# ttbar at 380
prodGroup = %(detectorModel)s_%(version)s
ProdTypes = Sim, Rec, RecOver # Gen, Split
prodids =  4539, 4542, 4545, 4548, 4551, ...
eventsInSplitFiles =
MoveTypes = SIM, REC #GEN
move = True
additionalName = CT_%(detectorModel)s_%(version)s
cliReco = --Config.Tracking=Conformal
overlayEvents = 350GeV
```

Main Production Script

- Metadata used to define input data (ProdID) and metadata for folders
- Productions implicitely coupled. ProdID part of input data query for next step
- https://gitlab.cern.ch/CLICdp/iLCDirac/ILCDIRAC/blob/Rel-v28r0/ILCTransformationSystem/scripts/dirac-clic-make-productions.py

```
#[...]
if self._flags.sim:
  simMeta = self.createSimulationProduction(metaInput, prodName, parameterDict)
  self._updateMeta(metaInput, simMeta, eventsPerJob)

if self._flags.rec:
  recMeta = self.createReconstructionProduction(metaInput, prodName, parameterDict, over=False)

if self._flags.over:
  overMeta = self.createReconstructionProduction(metaInput, prodName, parameterDict, over=True)

if genMeta:
  self.createMovingTransformation(genMeta, 'MCGeneration')

if splitMeta:
  self.createMovingTransformation(splitMeta, 'MCGeneration')
#[...]
```

Transformation Metadata Example

- Metadata set at creation of transformation

```
- Registered metadata:
    /ilc/prod/clic/3tev/ = {'Energy': '3000'}
    /ilc/prod/clic/3tev/tt/ = {'EvtType': 'tt'}
    /ilc/prod/clic/3tev/tt/CLIC_o3_v14 = {'DetectorType': 'CLIC_o3_v14'}
    /ilc/prod/clic/3tev/tt/CLIC_o3_v14/DST = {'Datatype': 'DST'}
    /ilc/prod/clic/3tev/tt/CLIC_o3_v14/DST/00010229 = {'NumberOfEvents': 25, 'ProdID': 10229L}
    /ilc/prod/clic/3tev/tt/CLIC_o3_v14/REC = {'Datatype': 'REC'}
    /ilc/prod/clic/3tev/tt/CLIC_o3_v14/REC/00010229 = {'NumberOfEvents': 25, 'ProdID': 10229L}
- Registered non searchable metadata:
    /ilc/prod/clic/3tev/tt/CLIC_o3_v14/DST/00010229 = {'BeamParticle2': 'E1',
                                                       'SWPackages': 'marlin.ILCSoft-2018-05-18_gcc62',
                                                              'BeamParticle1': 'e1', 'EPA_B1': 'F', 'EPA_B2': 'F'}
    /ilc/prod/clic/3tev/tt/CLIC_o3_v14/REC/00010229 = {'BeamParticle2': 'E1',
                                                              'SWPackages': 'marlin.ILCSoft-2018-05-18_gcc62',
                                                              'BeamParticle1': 'e1', 'EPA_B1': 'F', 'EPA_B2': 'F'}
```

Data Operation Transformations

- scripts to define replication/moving transformations
- Set ProdID, SourceSE, TargetSE, Datatype, optional GroupSize
- http://lcd-data.web.cern.ch/lcd-data/doc/ilcdiracdoc/DOC/ILCDIRAC/ILCTransformationSystem/scripts/dirac-ilc-replication-transformation.html
- http://lcd-data.web.cern.ch/lcd-data/doc/ilcdiracdoc/DOC/ILCDIRAC/ILCTransformationSystem/scripts/dirac-ilc-moving-transformation.html

DataRecovery-, FileStatusTransformationAgent

- DataRecoveryAgent: Checks all production jobs for consistency, treat inconsistent states
  - All output files present for successful jobs, remove output files that shouldn't exist, input file exists, one input file has only one set of descendents,. . .
- FileStatusTransformationAgent: checks consistency of *DataOperation* transformations, reset *Requests*
  - output files exist at destinations