P301 - MERIT and FITS reprocessing with new evt classifications

status: **Complete**
last update: 27 June 2015

This page is a record of the configuration and execution of the P301 reprocessing project, regeneration of MERIT and FITS files from the P300 MERIT using TMine.

Notice: there is now a JIRA for this task. Please check here for operational notes.

Pipeline tasks

- P301-MERIT - This task reads P300 MERIT and produces reprocessed MERIT. This is the first time TMineExt has been used to perform this function.
- P301-FITS - This task reads new P301 MERIT and produces fits files, EXTENDEDFT1 + EXTENDEDLS1 ** THIS TASK IS DEFUNCT, see P302-FITS **

Datafile names, versions and locations

Data file version numbers for this reprocessing will begin with v300.

XROOT location and file naming

Location template:

```
/glast/Data/Flight/Reprocess/<reprocessName>/<dataType>
```

Locations for P301:

```
/glast/Data/Flight/Reprocess/P301/merit

[soon]
/glast/Data/Flight/Reprocess/P301/ft1
/glast/Data/Flight/Reprocess/P301/extendedft1
/glast/Data/Flight/Reprocess/P301/electronft1
/glast/Data/Flight/Reprocess/P301/ls1
/glast/Data/Flight/Reprocess/P301/extendedls1
```

File naming:

Data Type	aka	Send to FSSC	Naming template
MERIT		No	r<run#>_<*version*>_<dataType>.root
FILTEREDMERIT		No	r<run#>_<*version*>_<dataType>.root
ELECTRONMERIT		No	r<run#>_<*version*>_<dataType>.root
ELECTRONFT1		No	gll_el_p<procVer>_*r<run#>*_<version>.fit
EXTENDEDFT1		No	gll_xp_p<procVer>_*r<run#>*_<version>.fit
FT1	LS-002	Yes	gll_ph_p<procVer>_*r<run#>*_<version>.fit
EXTENDEDLS1		No	gll_xe_p<procVer>_*r<run#>*_<version>.fit
LS1	LS-001	Yes	gll_ev_p<procVer>_*r<run#>*_<version>.fit

Note: 'procVer' is a field added to the file name (and the keyword "PROC_VER" in the primary header) added to the FFD 5/12/2010. Ref: http://fermi.gsfc.nasa.gov/ssc/dev/current_documents/Science_DP_FFD_RevA.pdf

Example:

```
/glast/Data/Flight/Reprocess/P301/recon/r0239557414_v301_merit.root
```

DataCatalog location and naming

Logical directory and group template:

```
Data/Flight/Reprocess/<reprocessName>:<dataType>
```

Note that the <dataType> field (following the colon) is a DataCatalog 'group' name, and file names are of the form r<run#>.

Naming examples:

```
Data/Flight/Reprocess/P301:MERIT r0239557414
```

Data Sample

The currently defined data sample (as of 4/18/2014) for P301 reprocessing includes:

First run	239557414 (MET), 2008-08-04 15:43:34 (UTC)
Last run	present

Bookkeeping

1. (This page): Define ingredients of reprocessing (processing code/configuration changes)
2. Processing History database: http://glast-ground.slac.stanford.edu/HistoryProcessing/HProcessingRuns.jsp?processingname=P301
    a. List of all reprocessings
    b. List of all data runs reprocessed
    c. Pointers to all input data files (-> dataCatalog)
    d. Pointers to associated task processes (-> Pipeline II status)
3. Data Catalog database: http://glast-ground.slac.stanford.edu/DataCatalog/folder.jsp
    a. Lists of and pointers to all output data files
    b. Meta data associated with each output data product

P301-MERIT

Status chronology

- 3/27/2014 - Begin setting up P301 task for Pass8 re-MERIT.
- 5/5/2014 - Please see the P301 JIRA for future chronology updates
- 6/9/2014 - Begin block 2, through end of April 2014 (see JIRA)
- 6/10/2014 - Block 2 complete
- 10/21/2014 - Begin block 3, through end of September 2014 (see JIRA)
- 10/22/2014 - Block 3 complete
- 02/17/2014 - Begin block 4, through end of February 2015
- 02/18/2015 - Block 4 complete
- 04/16/2015 - Begin block 5, through 7 Apr 2015
- 04/18/2015 - Block 5 complete
- 06/11/2015 - Block 6 through 3 Jun 2015 started and completed.
- 06/26/2015 - Block 7 - the final block - through 24 June 2015 started.
- 06/27/2015 - Finished with Engines.
- 08/05/2015 - Emergency reprocessing of 40 runs that suffered bad CAL calibrations. (See P300-ROOT for details.)

Configuration

Task Location	/nfs/farm/g/glast/u38/Reprocess-tasks/P301-MERIT
Task Status	http://glast-ground.slac.stanford.edu/Pipeline-II/index.jsp
TMineExt	3.3.8 (17 Apr 2014) as used from $GLAST_EXT/TMineExt
Classification worksheet	v0r16 from $GLAST_EXT/extFiles
Run Selection	All MERIT files produced by P300-ROOT task

s/c data	P202 FT2SECONDS which will eventually become a "standard" Public Release https://confluence.slac.stanford.edu/display/SCIGRPS/Official+LAT+Datasets
Input Run List	ftp://ftp-glast.slac.stanford.edu/glast.u38/Reprocess-tasks/P301-MERIT/config/runList.txt
Code Variants used	redhat5-i686-32bit-gcc41, redhat5-x86_64-64bit-gcc41, redhat6-x86_64-64bit-gcc44 (Optimized)
Output Data Products	MERIT

Timing and Scaling

job step	average CPU time
TMine	~0.5-1M evts/hour

P301-FITS

This task generates all desired FITS data products.

- A slightly updated task structure has been implemented for this task: the mergeClumps step is now a separate subTask which allows overriding the batch queue parameter for extra long runs. This restructuring required quite a lot of modification related to data passed between job steps (as now certain job steps are in the new subTask).
- Task stored in my github as P301-FITS

Status chronology

This task basically mirrors the P301-MERIT task, running after the aforementioned completes

Please see the associated JIRA for operational notes.

Configuration

Task Location	/nfs/farm/g/glast/u38/Reprocess-tasks/P301-FITS
Task Status	http://glast-ground.slac.stanford.edu/Pipeline-II/exp/Fermi/task.jsp?refreshRate=60&task=137594093
Input Data	MERIT (direct from P301-MERIT task)
spacecraft data	same as P301-MERIT
Input Run List	ftp://ftp-glast.slac.stanford.edu/glast.u38/Reprocess-tasks/P301-FITS/config/runList.txt
evtClassDefs	01-00-01
eventClassMap	EvtClassDefs_P8V1.xml
ScienceTools	09-33-01 (5/15/2014)
Code Variants	redhat5-x86_64-64bit-gcc41 & redhat6-x86_64-64bit-gcc44 (Optimized)
Diffuse Model	N/A
Diffuse Response	N/A
IRFs	N/A
Output Data Products	EXTENDEDFT1, EXTENDEDLS1
commonTools	00-02-01

Generation of output data products:

Data Product	destination	data content [1]	event selection [1]	makeFT1	gtselect	gtdiffrsp	gtmktime
EXTENDEDFT1	SLAC	pass8_FT1variables	pass8.1_Transient_cuts_L1	✓	✗	✗	✓

EXTENDEDLS1	SLAC	pass8_LS1variables	pass8.1_Transient_cuts_L1	✓	✗	✗	✓

[1] /afs/slac/g/glast/ground/releases/volume04/evtClassDefs/01-00-01/data

Note on 'Code Variant': The SLAC batch farm contains a mixture of architectures , both hardware (Intel/AMD 64-bit) and software (RHEL5-64, gcc v4.1, etc.). At this time, GlastRelease builds only on RHEL5-32 (RHEL6-64 are built but not yet validated), while ScienceTools builds for RHEL5-64, RHEL6-64.

Timing and Scaling

Approx 20 CPU minutes and 30 wall clock minutes per job (run).