

P302 - Pass8 FITS data product generation

status: **Complete**

last update: 21 July 2015

This page is a record of the configuration and execution of the P302 reprocessing project, generation FITS files from the P301 MERIT files.

Notice: there is now a [JIRA](#) for this task. Please check here for operational notes.

Pipeline tasks

- P302-FITS - This [task](#) reads new P301 MERIT and produces fits files, FT1 + LS1 + EXTENDEDFT1 + EXTENDEDLS1

Datafile names, versions and locations

Data file version numbers for this reprocessing will begin with v302.

XROOT location and file naming

Location template:

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

Locations for P302:

```
/glast/Data/Flight/Reprocess/P302/ft1
/glast/Data/Flight/Reprocess/P302/extendedft1
/glast/Data/Flight/Reprocess/P302/ls1
/glast/Data/Flight/Reprocess/P302/extendedls1
```

File naming:

Data Type	aka	Send to FSSC	Naming template
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/P302/recon/r0239557414_v302_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/P302:MERIT r0239557414
```

Data Sample

The currently defined data sample (as of 1/16/2015) for P302 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=P302>
 - 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
-

P302-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 P302-FITS

Status chronology

This task basically mirrors the P301-FITS task but updating certain packages and expanding the generated data products to include FT1 and LS1.

Please see the associated JIRA for operational notes.

- 1/27/2015 - Completed block 1, through 30 Sep 2014

Name	Files	Events	Size	Created (UTC)
EXTENDEDFT1	34154	13,822,731,339	1.2 TB	21-Jan-2015 20:32:16
EXTENDEDLS1	34154	13,822,731,339	2.6 TB	21-Jan-2015 20:32:16
FT1	34154	691,710,015	64.1 GB	21-Jan-2015 20:32:16
LS1	34154	1,835,112,272	353.2 GB	21-Jan-2015 20:32:16

- 2/17/2015 - Rerun stream 29158, Run 405333211, after Warren revises the FT2 file – which adds back the final 10 minutes of this run. (Jean Ballet found this problem.)
- 2/18/2015 - Completed block 2, through 31 Jan 2015

Name	Files	Events	Size	Created (UTC)
EXTENDEDFT1	36020	14,613,698,010	1.3 TB	21-Jan-2015 20:32:16
EXTENDEDLS1	36020	14,613,698,010	2.7 TB	21-Jan-2015 20:32:16
FT1	36020	740,190,525	68.6 GB	21-Jan-2015 20:32:16
LS1	36020	1,956,141,539	376.5 GB	21-Jan-2015 20:32:16

- 3/9/2015 - Create P302-FITS v 2.0 task. The only difference is changing evtClassDefs from 01--01--01 to 01--01--04 which updates event selection to eliminate sweep and periodic trigger events. (This was discovered by Don Horner at FSSC as duplicate (sweep) events in files transferred there a couple weeks ago.)
Begin processing entire mission once again...

- 3/11/2015 Reprocessing complete. New Datacatalog summary:

Name	Files	Events	Size	Created (UTC)
EXTENDEDFT1	36020	14,613,584,327	1.3 TB	21-Jan-2015 20:32:16
EXTENDEDLS1	36020	14,613,504,070	2.7 TB	21-Jan-2015 20:32:16
FT1	36020	740,187,922	68.6 GB	21-Jan-2015 20:32:16
LS1	36020	1,956,117,915	376.5 GB	21-Jan-2015 20:32:16

Problem with xFT1 and xLS1 files: #evts should be identical. After studying logs, decide there must be a silent failure mode to gtmtime. Rollback and now the statistics look like this:

Name	Files	Events	Size	Created (UTC)
EXTENDEDFT1	36020	14,613,589,234	1.3 TB	21-Jan-2015 20:32:16
EXTENDEDLS1	36020	14,613,589,234	2.7 TB	21-Jan-2015 20:32:16
FT1	36020	740,188,083	68.6 GB	21-Jan-2015 20:32:16
LS1	36020	1,956,127,293	376.5 GB	21-Jan-2015 20:32:16

- 04/18/2015 - Begin block 2, through 7 Apr 2015
- 04/19/2015 - Block 2 complete. Summary from dataCatalog:

Name	Files	Events	Size	Created (UTC)
EXTENDEDFT1	37023	15,032,071,005	1.3 TB	21-Jan-2015 20:32:16
EXTENDEDLS1	37023	15,032,071,005	2.8 TB	21-Jan-2015 20:32:16
FT1	37023	763,368,036	70.7 GB	21-Jan-2015 20:32:16
LS1	37023	2,015,775,627	388.0 GB	21-Jan-2015 20:32:16

- 06/11/2015 - Block 3, 7-Apr through 3-Jun-2015 complete (MERIT statistics from P301-MERIT task)

Name	Files	Events	Size	Created (UTC)
MERIT	37888	82,148,825,606	139.7 TB	28-Mar-2014 22:40:39
EXTENDEDFT1	37888	15,381,147,951	1.4 TB	21-Jan-2015 20:32:16
EXTENDEDLS1	37888	15,381,147,951	2.9 TB	21-Jan-2015 20:32:16
FT1	37888	779,586,312	72.2 GB	21-Jan-2015 20:32:16
LS1	37888	2,059,797,732	396.4 GB	21-Jan-2015 20:32:16

- 6/27/2015 - Block 4, through 24 June 2015 **FINAL** block for Pass8!!
- 8/5/2015 - Block 5, an emergency reprocess of 40 runs in the period 7/27-7/30/2015 which suffered bad CAL calibrations.

Configuration

Task Location	/nfs/farm/g/glast/u38/Reprocess-tasks/P302-FITS
Task Status	complete
Input Data	MERIT (direct from P301-MERIT task)
spacecraft data	P203 (released 3 June 2014)
Input Run List	ftp://ftp-glast.slac.stanford.edu/glast.u38/Reprocess-tasks/P302-FITS/config/runList.txt
evtClassDefs	01-01-01 (version 1.1), 01-01-04 (version 2.0)
eventClassMap	EvtClassDefs_P8R2.xml
ScienceTools	09-35-02 (1/26/2015)
Code Variants	redhat6-x86_64-64bit-gcc44 (Optimized)
Diffuse Model	N/A
Diffuse Response	N/A
IRFs	N/A
Output Data Products	FT1 , LS1 , EXTENDEDFT1 , EXTENDEDLS1

commonTools	00-02-01
-------------	----------

Generation of output data products:

Data Product	destination	data content [1]	event selection [1]	makeFT1	gtselect	gtdiffrsp	gtmtime	Notes
EXTENDEDFT1	SLAC	pass8.2_Extended_cuts_L1	pass8.2_Transient_cuts_L1	✓	✗	✗	✓	
FT1	FSSC	pass8_FT1variables	evclass=128	✗	✓	✗	✓	produced from extendedFT1
EXTENDEDLS1	SLAC	pass8.2_Extended_cuts_L1	pass8.2_Transient_cuts_L1	✓	✗	✗	✓	
LS1	FSSC	pass8_LS1variables	evclass=65544	✗	✓	✗	✓	produced from extendedLS1

[1] /afs/slac/g/glast/ground/releases/volume04/evtClassDefs/01-01-04/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).