Software Workshop 15-18 July, SLAC



#### **Level 1 Processing Pipeline**

#### **Daniel Flath**

For

#### Daniel Flath (Stanford), Alex Schlessinger (SLAC)

**Daniel Flath & Alex Schlessinger** 

**Data Processing Pipeline** 



#### Overview

- Requirements
  - Processing
  - Cataloguing
- How it works
- OPUS
  - Introduction & Desirable features
  - LAT-Specific Additions
- Implementation Status
  - OPUS
  - Add-Ons
- Goals
  - DC1
  - **EM**



### Requirements

- Processing
  - Conversion of L0 (near-raw) to L1 data upon receipt from MOC
    - XFer of HSK data to OF
    - Reconstruction & Digitization
    - Must be complete before next downlink
    - XFer of diagnostics to OF
  - Reprocessing of L1 data as needed
  - Production of simulated data
- Data Cataloguing
  - Classification and storage of L1 data after processing
    - Summary information on each dataset
  - Tracking status and schedule of reprocessing



#### **How it Works**

- Pushes data through a sequence of processing steps
- Monitors status of a dataset as it proceeds through processing
- Provides notification upon failure of a processing stage
- Catalogues the result of processing for each dataset



#### **OPUS – Introduction & Features**

- Developed by AURA for the Space Telescope
- In use by several NASA experiments
- Does almost everything we need
  - Will run any \*program\* that can be "wrapped" by a shell script
  - Provides hooks for trapping processing errors
  - Distributes processing over a network of machines
  - Will run multiple processing sequences simultaneously
- Supports Extension of Functionality
  - Provides a C++ API to develop "OPUS-Aware" applications that have access to OPUS state information
- Displays live processing status (see next slide)
  - User can modify or override individual jobs & statuses

Software Workshop 15-18 July, SLAC



#### **OPUS in Action (PMG)**

2 Pr	1G: OF	PUS Proces	ss Mana	ger				
<u>F</u> ile	Edit	<u>M</u> anage	⊻iew	<u>T</u> ools	<u>H</u> elp			
	PUS		Pending	ending		whopper	vorion	
🖗 🗖	Proc	Delete	Pendin	g	pid	process	proc_stat	t year
	⊳⊡। ⊳⊡।	Susperd Selected			00014	dc_acs	pending	2001
		Resum Deguast comp			5493h		idle	2001
ାତ	ן ⊑ם ≺		Reque		r to susp	end select	ed tasks	
ę		Reinit	Selecte	a				
	ዋ (	Termir	nate Sel	ected				
		Delete	Absent					
		Cala	acs 	1000000	_			

(Image Courtesy of OPUS manual)

**GLAST LAT Project** 

.000

GLAST Software

Software Workshop 15-18 July, SLAC

#### **OPUS in Action (OMG)**

CMG: OPUS Observation Manager: bab5												
File Edit Manage View Tools Help												
OSFs	Modify											
datas	Modify (Unsafe!)	np	data_id	dcf_num	DP	DE	SS	DV	WC	GC	CA	HT
o4r103wbc	Suspend	:52:14	sti	061	C	n	C	C	C	C	C	w
o4r103w7c	13w7d		sti	060	С	n	C	C	C	C	C	w
o4r10301d	Resume	:53:01	sas	000	_	_	_	_	_	C	C	w
o4r101vnq	2000 09/20 13	3:51:14	sti	054	C	n	C	C	C	C	_	_
o4r101vkq	2000 09/20 13	3:51:14	sti	054	C	n	C	C	C	C	_	_
o4r101vhq	2000 09/20 13	8:51:12	sti	054	С	n	C	C	C	C	C	W
o4r101vdq	2000 09/20 13	3:51:02	sti	053	С	n	C	С	C	С	C	W
o4r101010	2000 09/20 13	8:51:59	sas	000	_	_	_	_	_	C	C	W
o4rj05b3q	2000 09/20 13	3:49:37	sti	005	С	n	C	С	C	С	C	W
o4rh05dyq	2000 09/20 13	3:49:44	sti	018	С	n	C	С	C	C	C	w
o41m52fdq	2000 09/20 13	3:50:46	sti	026	С	n	C	С	C	C	_	_
o41m52f6q	2000 09/20 13	3:50:37	sti	025	С	n	C	С	C	C	_	_
41 5061	0000 00 000 10			00.4	-		-	-	-	-		

#### (Image Courtesy of OPUS manual)

**Daniel Flath & Alex Schlessinger** 

**Data Processing Pipeline** 



# **LAT-Specific Additions**

- Support for the SLAC LSF Batch Processing Farm (Alex)
  - OPUS supports processing over multiple nodes using R/SSH
  - Alex is developing an interface layer that will extend this to the Batch Farm at SLAC
- Support for the ORACLE processing DB
  - OPUS saves log files containing the status of each processing stage
  - Dan is developing a set of scripts to wrap the processing database
  - At completion of processing for a particular dataset, these will catalog the status, location, and a summary of the dataset



## **Implementation Status**

- OPUS
  - Sample pipeline has been developed
    - Runs GLEAM (Sim/Digi/Recon)
    - Runs ROOT macros to verify the output
  - Support garnered from development team
- Database & LSF
  - Database routines about 50% complete
  - LSF layer is researched and ready to be developed
  - Awaiting OPUS source code & documentation
    - Upon receipt, can develop the LSF layer and the Cataloguing program (that will use the DB routines)



GOALS

- DC1
  - Stress-test Pipeline components (OPUS, LSF-Layer, DB-Layer) by running GLEAM on many machines
    - Full implementation contingent on getting source code from STSC
    - Kludge implementation possible failing this in short
      order
- EM
  - Test a toy version of what will be the L1P using data as it comes off the instrument