



National Aeronautics and Space  
Administration  
Jet Propulsion Laboratory  
California Institute of Technology



---

---

# WISE Data Processing Overview

R. Cutri and the WSDC Team @ IPAC



WSDC Data Processing Overview – 3/21/08



# WISE Will Release Three Primary Science Data Products



- ***Image Atlas*** - Calibrated images covering the sky in the four WISE bands. Photometric and astrometric calibration to be tied to the WISE Source Catalog.
  - $\sim 3.0e5$  1.375"/pix FITS format images (number depends on footprint adopted)
  - Combination of all available exposures covering each field
  - Depth-of-coverage maps for each image
  - Combined image metadata
- ***Source Catalog*** - Accurate positions and fluxes in four WISE bands
  - $\sim 3e8$  objects detected on combined images
  - Detection and measurement quality flags
- **Explanatory Supplement**
  - Mission and data product descriptions
  - User's guide (e.g. data formats, access modes)
  - Cautionary notes describing limitations of data and known idiosyncrasies





# Data Product Requirements



<b>WISE Source Catalog Requirements</b>		
Reliability	>99.9%	SNR>20; unconfused
Completeness	>95%	SNR>20; unconfused
Photometric Accuracy	<7%/band	SNR>100; unconfused
Photometric Sensitivity	3.3 $\mu$ m: 0.12mJy 4.7 $\mu$ m: 0.16mJy 12 $\mu$ m: 0.65mJy 23 $\mu$ m: 2.6mJy	SNR=5
Astrometric Accuracy	0.5" RMS wrt 2MASS PSC	

<b>WISE Image Atlas Requirements</b>	
Coverage	Constructed by combining all available images in each band
Band Registration	Four bands resampled to common pixel grid and scale
Calibration	Tied to Source Catalog
Format	Standard FITS





# Data Release Schedule



- **Preliminary Image Atlas and Source Catalog**
  - Derived from first 50% of sky surveyed
  - Conservative SNR limits
  - Release 6 months after end of on-orbit data collection (12/2010)
- **Final Image Atlas and Source Catalog**
  - Derived from all survey data
  - Aggressive SNR limits
  - Release 17 months after end of on-orbit operations (11/2011)
- **Project team, science community and general public will access WISE image and tabular data via the on-line services of NASA/IPAC Infrared Science Archive (IRSA)**
  - IRSA is NASA's designated archive center for infrared and sub-mm astronomy missions
  - IRSA will host Level 1,2 and final product archive
  - Project will maintain option to provide bulk distribution of WISE Source Catalog (e.g. via DVD-Rom)





# IPAC is the WISE Science Data Center (WSDC)



- **Data Ingest**
  - Receive raw science image packets from White Sands and H/K telemetry and ancillary files from MOS
  - Decompress and assemble image data into Level 0 FITS format, and mate with engineering data
- **Science Data Processing**
  - Convert raw imaging and engineering data into a photometrically and astrometrically calibrated Image Atlas and extracted Source Catalog
  - Compile Explanatory Supplement to the WISE Data Products
  - Generate ancillary data products to support mission requirements
- **Science Data Quality Assurance**
  - Quicklook QA feedback for on-orbit performance (i.e. scan mirror synch)
  - Overall science data QA for survey planning, data product generation and characterization
- **Science Data Archiving and Distribution**
  - Raw data (Level 0) archive during mission
  - Long-term “living” archive for primary and intermediate data products
  - Serve WISE science data products to project team, astronomical community and general public along with user’s guide documentation and descriptive analysis





# WSDC Design Philosophy - 1



- Leverage experience and intellectual heritage from past IPAC missions
  - 2MASS - Similar data acquisition, functional requirements, scope, archive size
  - Spitzer - Similar detectors, wavelength ranges
  - GALEX - Automated ingest
- Utilize a highly automated “pipeline” processing system to keep up with high data rate
  - Little opportunity for human intervention
- Plan a modular system design (subsystems) for efficient development, testing and for processing parallelization
- Enforce strict CM and interface management
- Make us of extensive, science-based automated quality assessment to deal with very large data sets and ambitious release schedule





## WSDC Design Philosophy - 2



- Adopt “two-pass” processing and release strategy
  - “Can’t get it right the first time”
  - First pass is best-effort during survey operations. Enables early data release
  - Second pass, post-ops, uses best instrument, calibration, sky and algorithm knowledge for optimal product development
- Focus on production of highly useable data products that enables science by a broad community
  - Extensive analysis and characterization (with Science Team support)
  - Well-documented (on-line Explanatory Supplement)
  - Ease of data access for Project Team and public (using proven IRSA interfaces)
- Build archive within IRSA architecture
  - Minimizes new development cost by taking advantage of existing expertise and re-use of infrastructure
  - Implicitly delivers products to NASA’s designated archive center for IR and sub-mm astronomy missions
  - Insures VO interoperability





# WISE Data Will Be Processed Two Complete Times



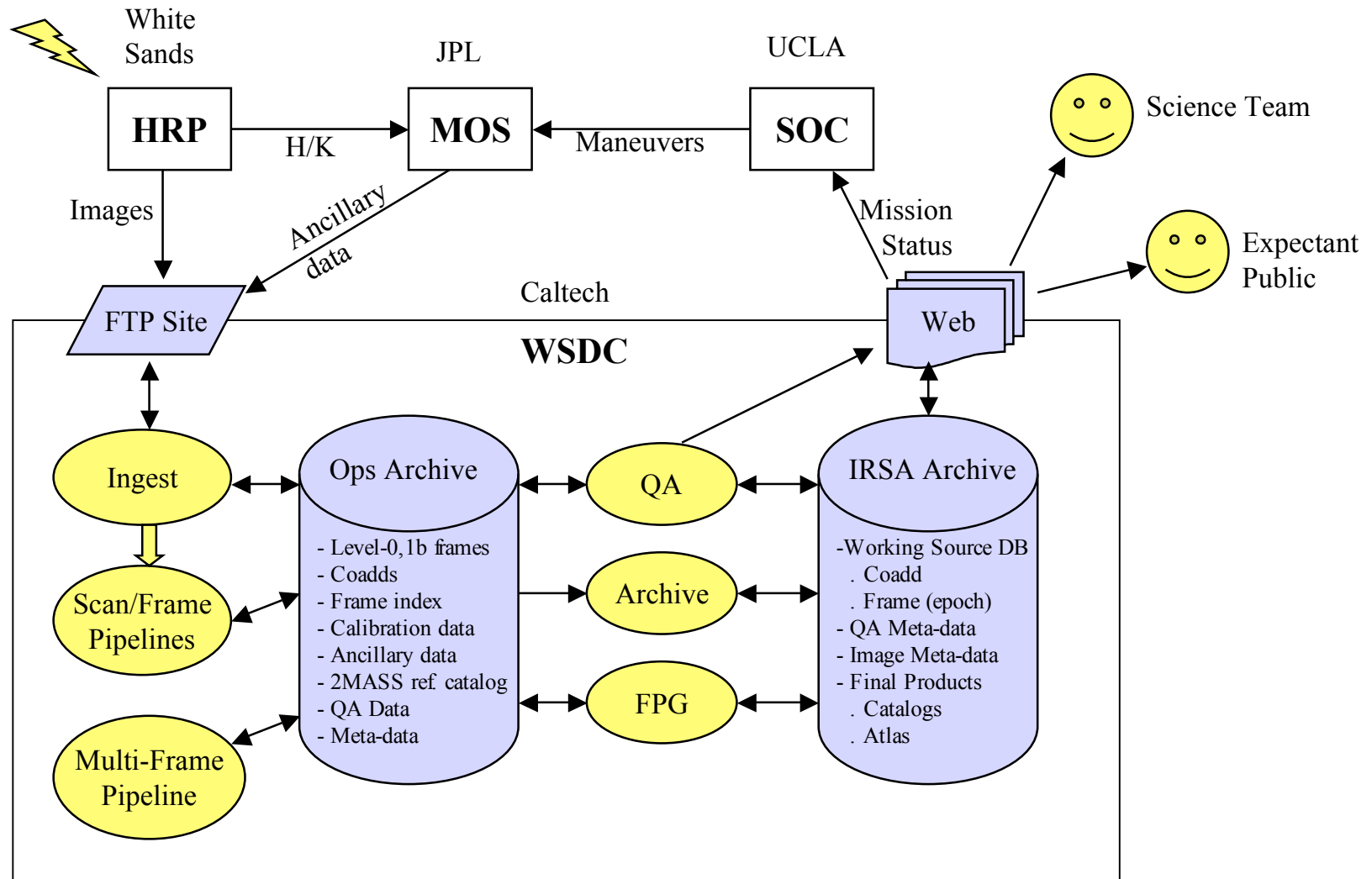
- **First Pass Processing**
  - Starts ~1 month after normal survey operations begin, finishes ~1 month after end of on-orbit data collection. Supports Quicklook QA, nominal survey operations
  - Use best available calibration data from ground testing and IOC; generate dynamic on-orbit calibrations and optimize windows/parameters (desired accuracy, anneals, SAA passages, stability, transients, bright source avoidance etc..)
  - Preliminary reduction algorithms based on pre-launch simulation and IOC tuning
  - Preliminary Image Atlas and Source Catalog will be drawn from results of first-pass processing
- **Second Pass Processing**
  - Starts ~4 months after on-orbit data collection, duration ~4 months.
  - Incorporates best knowledge of actual flight system performance, properties of the infrared sky, data processing algorithms derived from analysis of first-pass processing
  - Best available calibration using full data-set time history, and optimal “interval matching” to survey data
  - Final Image Atlas and Source Catalog will be drawn from results of the second-pass processing







# WISE Science Data System (WSDS) Executes Functions of the WSDC





# WSDS Consists of Six Subsystems

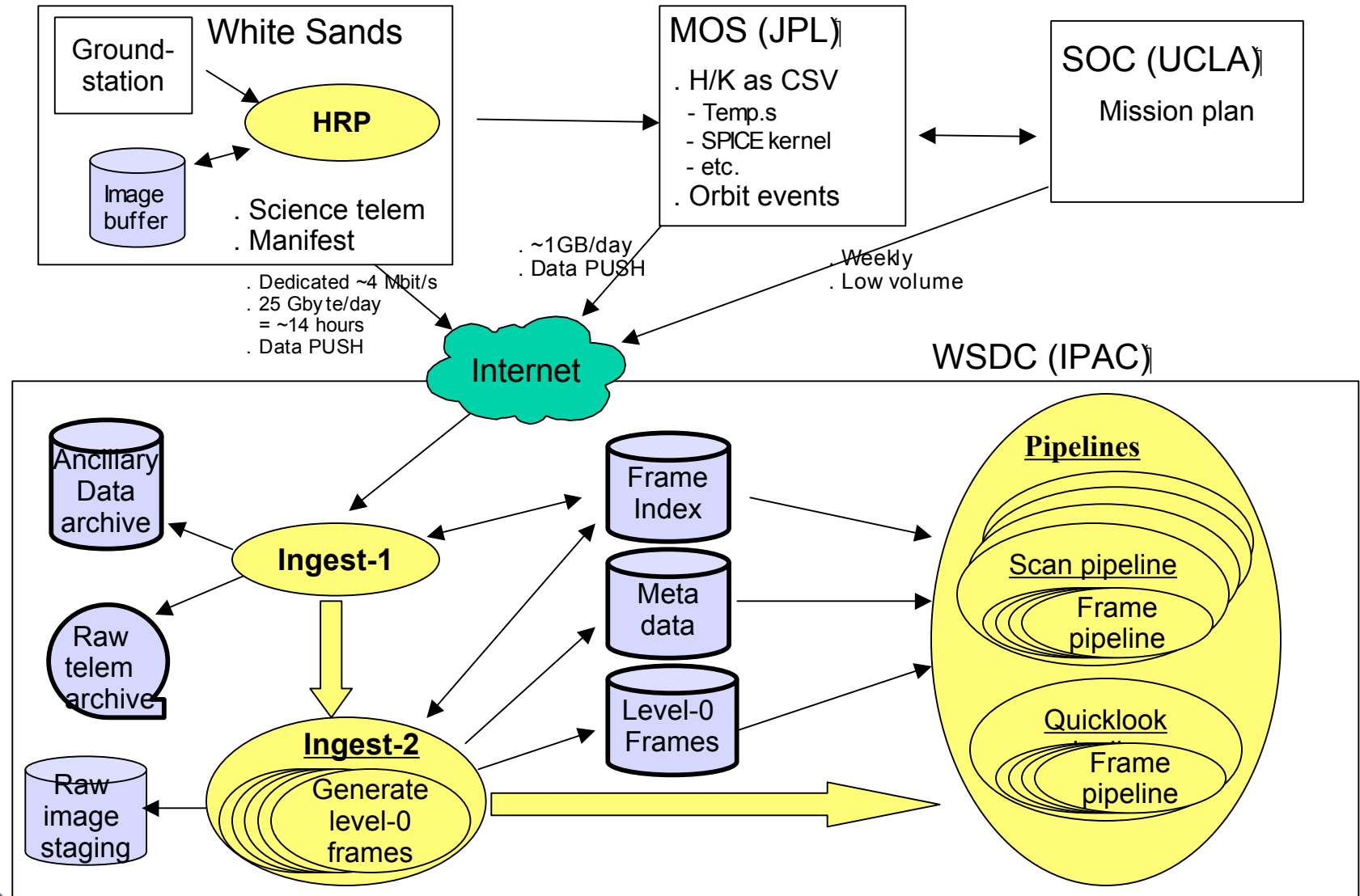


- INGEST
  - Receives science data packets, engineering telemetry and NAV products from MOS and assembles Level 0 FITS-format files. Stages Level 0 images and metadata for pipeline processing.
- Data Reduction PIPELINES
  - Converts Level 0 imaging data into calibrated images and extracted source *Working Databases*
    - **Frame/Scan pipeline** operates on individual frames within one “scan” (=1/2 orbit)
    - **Multiframe pipeline** operates on data from multiple orbits
- Quality Assurance (QA)
  - Generates concise reports summarizing science data quality using summary outputs from other subsystems. Web-based report, with capability to drill-down to detailed image, graphical and tabular data
  - Reports reviewed by QA scientists at WSDC. Final quality assignment approved by PI or designee
- EXEC
  - Provides interface-related services to wrappers and pipelines. Mediates between external callers and applications, providing a uniform interface, binding execution units (modules) together into a unified pipeline
- ARCHIVE/Distribution System
  - Archives raw and processed mission data and metadata. Serves Image Atlas, Source Catalog and mission metadata to WISE project team and astronomical community. Integrated into Infrared Science Archive (IRSA) at IPAC.
- Final Product Generator (FPG)
  - Constructs WISE Preliminary and Final Image Atlases and Source Catalogs from *combined* image and source *Working Databases*. Includes validation, characterization and documentation.





# WSDS Subsystems: Ingest





# WSDS Subsystems: Ingest



---

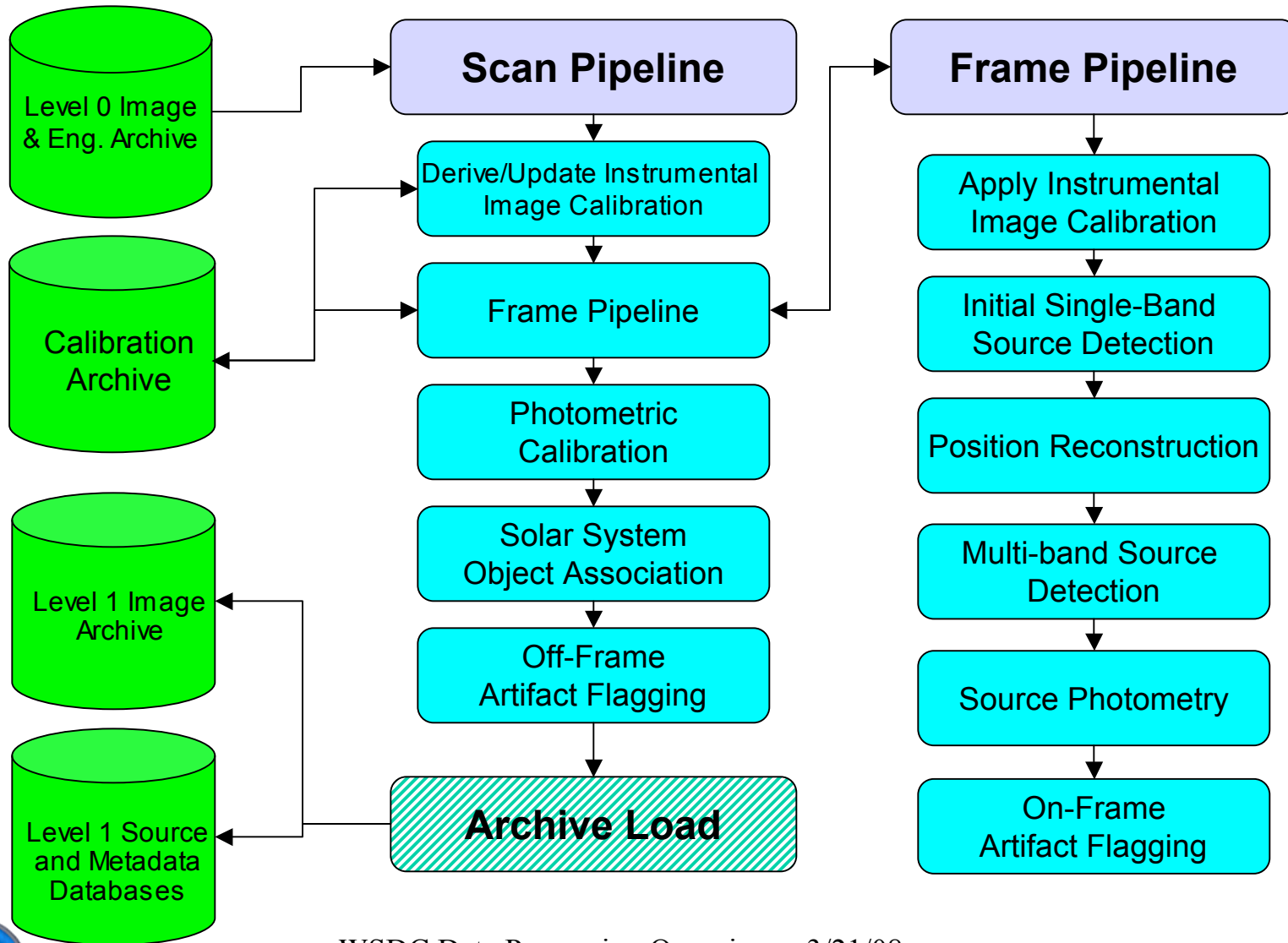
*Receive and verify high-rate science telemetry from White Sands. Decompress and assemble complete frames with ancillary data from MOS. Write level-0 image archive. Start scan pipeline.*

- Receive manifest at start of transfer from White Sands
- Detect completed transfers from HRDP and verify against manifest
- Archive raw frame telemetry and ancillary data from MOS
- Decompress and assemble all images and update frame index
  - Images may come out of order from multiple deliveries, so a scan may need to be stitched together using an index of data in the image archive
- Correlate data to ancillary data and generate frame meta-data
  - WCS info, time, other (e.g. temperatures, voltages, correlation to science plan?)
  - Write to image headers and stand-alone meta-data file
- Write compressed FITS files with meta-data to level-0 frame archive
- Identify frames in complete scans (pole crossing to pole crossing) and start the scan pipeline
  - Locate frames for each complete scan in the delivery. For incomplete scans, process when re-transmissions complete the scan or 72 hours (TBD) elapses
  - The quicklook pipeline is launched after each delivery without regard for missing frames





# Scan/Frame Pipeline Reduces Data From Each Half-Orbit





# WSDS Subsystems: Scan/Frame Pipeline



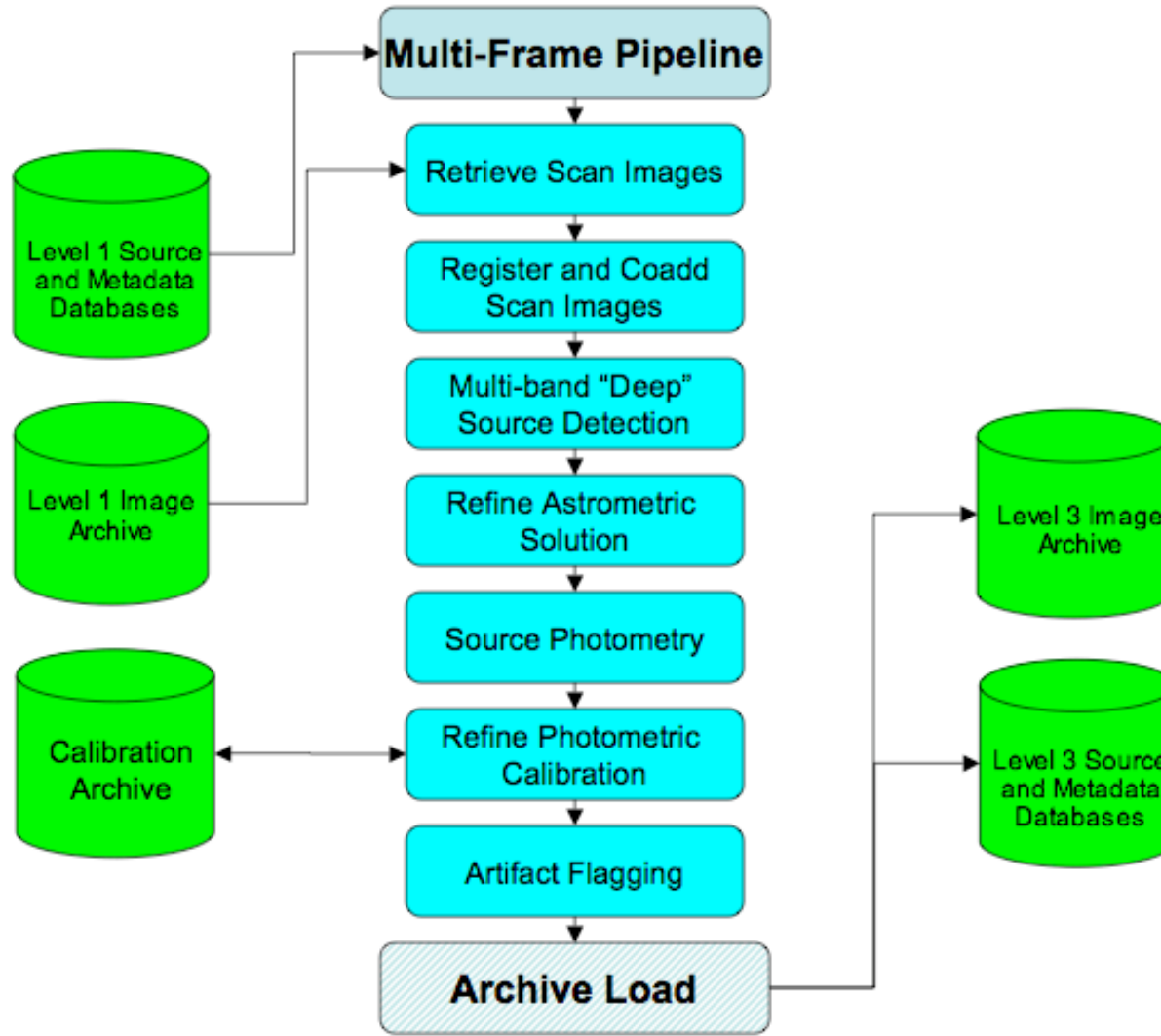
*For a complete **scan** (=1/2 orbit), calibrate frames, produce FITS Level 1 frame images, archive extracted sources to Level 1 Working Database, save meta-data and QA data.*

- Start and monitor **Frame Pipeline** for designated frames in scan
- Calibrate frames
  - Flag and mask hot pixels, artifacts, saturation, cosmic ray hits, etc.
  - Offset correction from reference pixels (TBD)
  - Subtract darks, apply pixel response, linearity, and droop corrections
  - Sky offset correction from dynamic illumination profile
- Extract and characterize sources from Level 1a frames
- Refine frame position/orientation/scale from comparison with 2MASS Point Source Catalog matched to high SNR extraction list
- Apply refinements to frame meta-data and source positions
- Apply photometric offsets to frame meta-data, source fluxes
  - Compute offsets from matched cal stars in most recent polar fields
- Search for associations with known solar system objects in extracted source lists
- Write source list and meta-data to frame Level 1 Working Database
- Compute and write QA data





# Multi-Frame Pipeline Combines Data From Multiple Scans/Frames





# WSDS Subsystems: Multi-Frame Pipeline



*Combine frame data covering the coadd geometry, extract sources, identify artifacts, Update calibration. Output coadd images and sources to the Level 2 Working Database.*

- Create coadds as often as once per delivery or as rarely as once per week
- Select pre-defined or manual coadd geometry (RA, Dec, size, epoch)
- Gain- and background-correct Level 1 frames, Resample to produce coadd pixels
- Construct intensity, coverage and uncertainty images
  - Examine frame flag images and meta-data
    - Reject frames failed by QA
    - Apply flags (artifacts, saturation, etc.), flag (possibly mask) out-of-bed frame pixels (cosmic ray hits, reflections, other outliers)
  - Record depth-of-coverage in supplementary images
- Extract and characterize sources, update astrometric and photometric solutions
- Write coadds to Level 2 Image Archive
- Write sources to Level 2 Working Database and meta-data to archive
- Write QA data







# WSDS Subsystems: Quality Assurance



*Generate concise, web-based reports summarizing science data quality. QA analysts generate a QA score for each frame. Key results are archived.*

- Draws from:
  - Multi-band scan pipeline output, including specialized QA results
  - IRSA searches: 2MASS, extractions from overlaps, etc.
  - Ancillary-data: Ephemeris (SAA or moon proximity), temp.s, etc.
- Manual examination of trend plots with data from every frame
- Specialized tools allow interactive QA analysis
- Evaluate, scan mirror synchronization, PSF trend, efficacy of artifact removal, flat performance, scan coverage, astronomical properties (logN/logS, color-color plots, astrometry evaluation, etc.)
- Multi-frame analysis adds N/M results (C&R, photometric repeatability), artifact trending, coadd image quality
- Write QA report and key results to QA archive
  - Compare QA metrics to predetermined thresholds, generate QA score
  - Generate human (HTML) and machine-readable reports
  - Final QA disposition approved by ST designee



# WSDS Subsystems: 2MASS QA Example



2MASS I  
 QA Web  
 001020s

2MASS QA Report  
 Focus Summary  
 Observation Date: 001020s



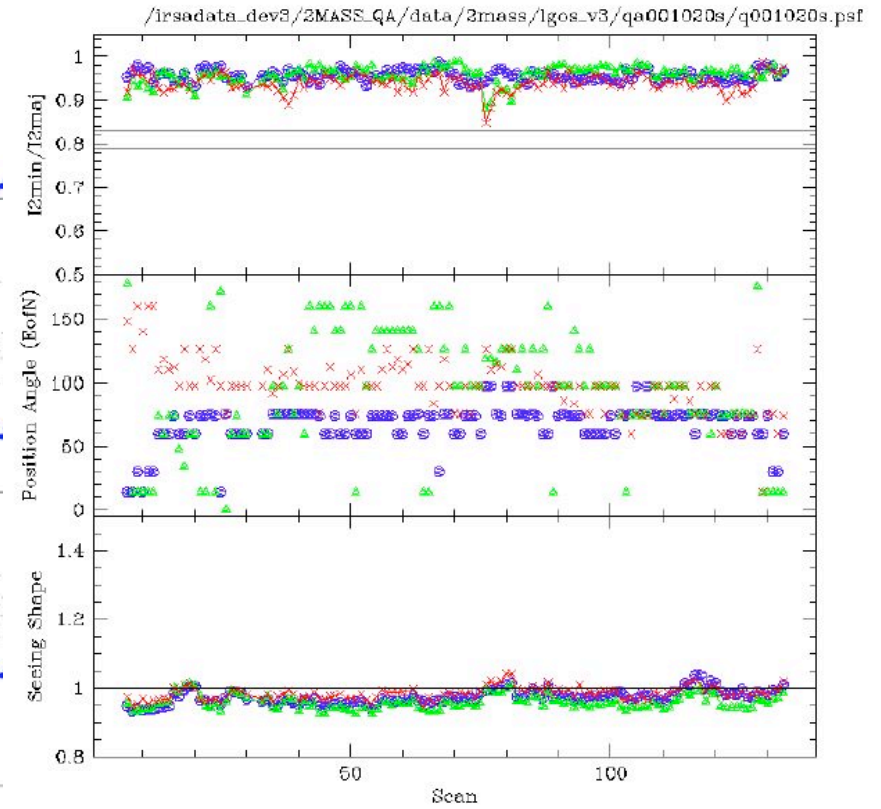
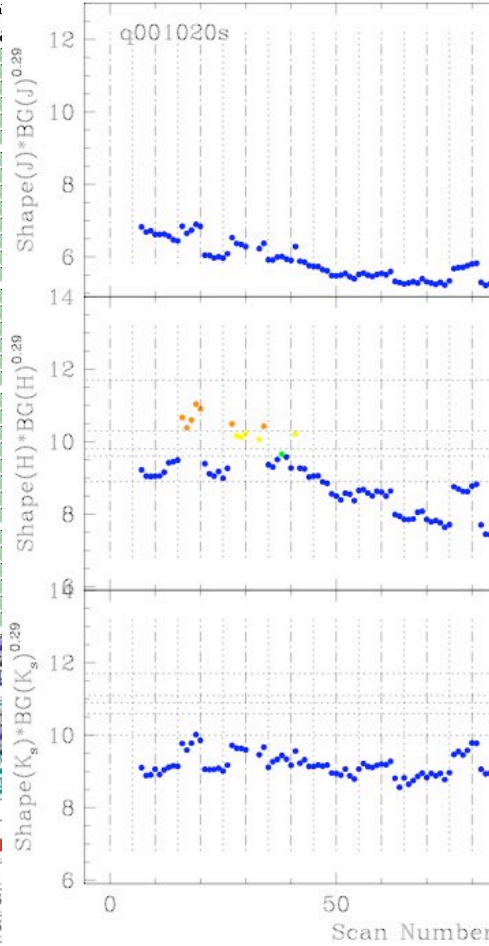
- Scan C
- Photon
- Bkg/See
- Gal

Scan#	Type	Average of I2min/I2maj			Seeing tracker score			Untracked area		
		J	H	Ks	J	H	Ks	J	H	Ks
007	CAL	0.948	0.909	0.923	-0.221	-0.200	-0.271	0	0	0
008	CAL	0.957	0.939	0.977	-0.141	0.100	-0.017	0	0	0
009	CAL	0.977	0.923	0.940	0.000	-0.229	-0.135	0	0	0
010	CAL	0.967	0.946	0.963	-0.020	-0.316	-0.208	0	0	0
011	CAL	0.976	0.930	0.951	-0.160	-0.160	-0.032	0	0	0
012	CAL	0.938	0.917	0.964	-0.100	1.058	-0.229	0	900	0
013	SUR	0.958	0.959	0.926	-0.041	-0.054	0.892	0	0	0
014	SUR	0.959	0.967	0.919	-0.062	-0.068	-0.009	0	0	0
015	SUR	0.960	0.961	0.926	0.045	-0.027	-0.040	0	0	0
016	SUR	0.947	0.951	0.924	-0.023	-0.097	0.152	0	0	0
017	SUR	0.944	0.937	0.936	0.211	-0.043	-0.009	0	0	0
018	SUR	0.961	0.962	0.936	0.125	-0.024	-0.068	0	0	0
019	SUR	0.937	0.929	0.924	-0.043	-0.042	0.040	0	0	0
020	SUR	0.933	0.909	0.938	0.252	-0.157	0.224	0	0	0
021	CAL	0.968	0.959	0.960	-0.260	-0.220	0.015	0	0	0
022	CAL	0.968	0.948	0.948	0.130	0.060	0.320	0	0	0
023	CAL	0.972	0.952	0.948	0.015	-0.061	0.039	0	0	0
024	CAL	0.972	0.976	0.960	-0.035	0.120	0.017	0	0	0
025	CAL	0.979	0.963	0.956	-0.014	0.380	-0.040	0	0	0
026	CAL	0.959	0.961	0.962	-0.085	-0.020	0.053	0	0	0
027	SUR	0.954	0.940	0.929	-0.037	-0.058	0.132	0	0	0
028	SUR	0.956	0.943	0.931	0.026	-0.060	0.104	0	0	0
029	SUR	0.958	0.948	0.934	-0.023	-0.086	-0.024	0	0	0
030	SUR	0.931	0.915	0.927	-0.019	-0.079	-0.083	0	0	0
031	SUR	0.944	0.950	0.940	0.080	0.190	0.695	0	0	0
032	SUR	0.967	0.969	0.939	0.486	0.626	0.065	0	0	0
033	SUR	0.956	0.941	0.936	0.001	-0.140	-0.035	0	0	0
034	SUR	0.957	0.957	0.931	-0.048	0.023	-0.029	0	0	0
035	CAL	0.959	0.958	0.934	-0.060	0.460	-0.157	0	0	0
036	CAL	0.955	0.938	0.931	-0.125	-0.230	0.017	0	0	0
037	CAL	0.945	0.924	0.909	0.018	-0.246	0.069	0	0	0
038	CAL	0.974	0.955	0.897	-0.140	-0.388	-0.058	0	0	0
039	CAL	0.955	0.964	0.911	-0.100	0.000	-0.052	0	0	0
040	CAL	0.961	0.961	0.944	-0.382	-0.109	-0.316	0	0	0

- Zero Point Default
- Overlays Summary
- Sensitivity
- Meteor Blanking
- RI:R2 Diff Monitor
- Distortion Monitor
- PSF/AT
- RI:R2 Ph

Error summary

filename	line #	error type
input.c	2169	scan number not in list data in updates
input.c	2169	scan number not in list data in updates





# WSDS Subsystems: Final Product Generator



---

*Human analysis and database-intensive activity to generate release Source Catalog and Image Atlas from L3 source Working Database and Coadded Image archive (not a discrete software subsystem per se)*

- Source Catalog tasks:
  - Select subset of entries in WDB that satisfies reliability criteria
  - Eliminate duplicate entries resulting from Atlas Image tile overlaps
  - Define final Catalog schema and derive new columns, if necessary
  - Validate Catalog quality relative to L1/1.5 requirements
  - Document
- Image Atlas tasks:
  - Define final FITS headers, image metadata and derive new parameters, if necessary
  - Validate Image quality relative to L1/1.5 requirements
  - Document
- Tailored to characteristics of data and actual system performance
  - Extensive use of processing QA output and off-line analysis





# WSDS Subsystems: Archive



*Supply permanently saved data to both internal (pipelines and QA) and external (Sci. Team, public) users.*

- Multi-tiered approach to data archiving and access
- Raw mission telemetry archive (from INGEST)
  - Copies to tape. On- and off-site storage
  - Live disk access
- Level-0 frame data in distributed, multi-node filesystem
  - Optimized for parallel access
- Mission ancillary data from MOS
- Level 1 and 2 Image Archives in file systems
- Working databases, catalogs, meta-data, and QA results in DBMS integrated into IRSA infrastructure
  - Catalog products are implicitly delivered to NASA-designated archive
  - Development minimized
  - IRSA/WISE web-based tools are primary Project interface to processed data
  - Public Interfaces are a subset of Project interfaces





# WSDS Subsystems: Executive Functions



---

*Provide a uniform interface for execution and control of routinely-executed WSDC applications and utilities and interfaces aiding automation and resource management.*

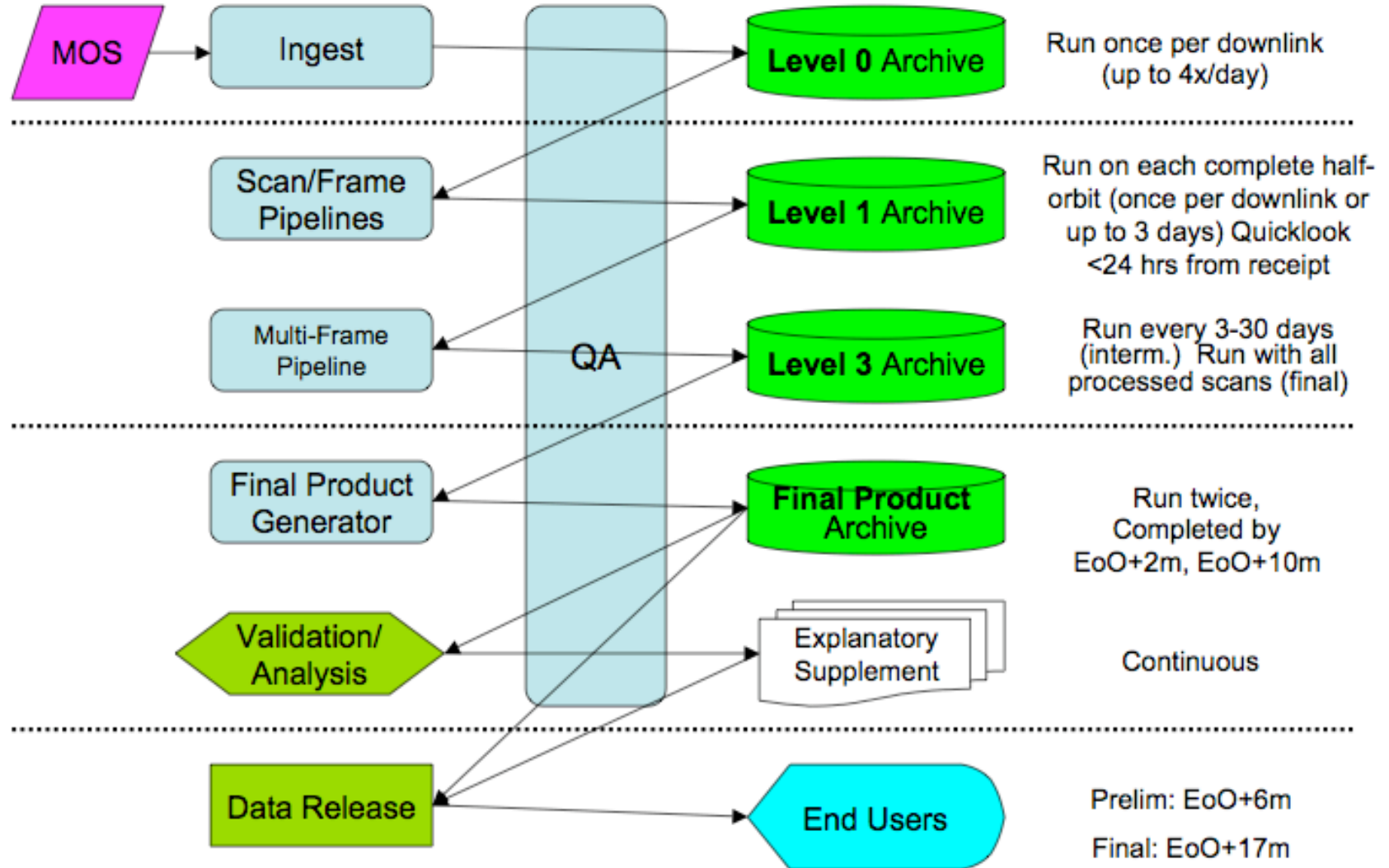
- **Application wrappers**
  - Standard parameter interface
  - Data dependency setup (moving, renaming, or massaging files, etc.)
  - Textual error, warning and informatory output management
  - Process status handling and error notification
  - Internal sub-process initiation and monitoring
- **Pipeline initiation and management**
  - Dependency-driven automatic start-up, e.g. ingest launches scan pipeline
  - Manual parameter-controlled CLI start-up, e.g. some coadd generation, special runs
  - Frame pipeline **concurrent** execution
- **Execution monitoring**
  - Web-accessible centralized process display
  - Controller notification of completion and failures
- **Resource monitoring**
  - Centralized monitoring of disk space, CPU and network load







# Internal Data Flow and Operational Cycle





# Data Processing Levels



Name	Description
Telemetry	CCSDS source packets
Raw	FITS, integer 2.75" pixels, band+time meta-data
Level-0	Raw + real pixels, add much meta-data
Level-1	Level-0 + instrumental, astrometric/photometric calibration
Level-1a	Level-0 + instrumental calibration applied to pixels
Level-1b	Level-1a + astrometric/photometric calibration in header
Level-2	Level-1 + upsampled and undistorted (rubber-sheeted)
Level-2a	Level-1a + upsampled and undistorted
Level-2b	Level-1b + upsampled and undistorted
Level-3	Multiple frame coadds, upsampled, undistorted
Atlas Images	Selected FPG Level-3 products

