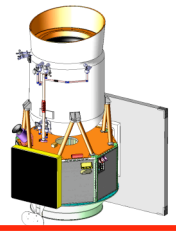
# WISE Science Data Center CDR

## WSDC Overview, Requirements and Implementation

R. Cutri - IPAC
WISE Science Data Manager

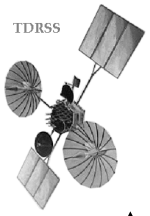
# MOS Architecture





Commands/Telemetry
S-Band

Science/Stored Telemetry
100Mbs Ku-Band




Science/Stored Telemetry
100Mbs Ku-Band

Commands/Telemetry
S-Band

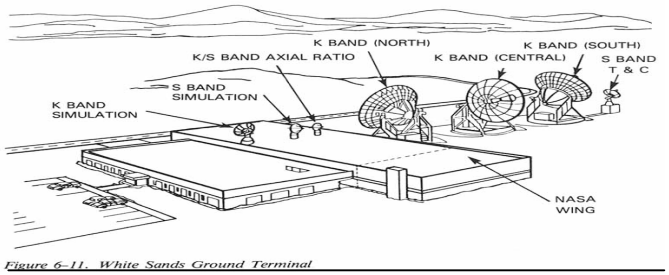



*Figure 6-11. White Sands Ground Terminal*

Science Survey Planning (UCLA)

Survey Plan Parameters

Engineering Operations System (JPL)

- Scheduling and Navigation (TDRSS and WISE ephemeris)
- Sequencing and Command Generation
- Real-Time Operations Heath and Safety Monitoring
- Data Archiving Engineering Data Processing
  - S/C Engineering Support (BATC)
  - Instrument Engineering Support (SDL)

WISE Science Data Center (IPAC)

- Science/Engineering Data Ingest
- Pipeline Processing And Data Archive
- Final Product Generation Image Atlas/Catalog
  - Science Community
  - Science Data And Survey Quality Assurance

White Sands Ground Terminal

- High Rate Science Data Processor
- TDRSS Uplink
- Real-time and Stored Telemetry

# WSDC Responsibilities



- Data Ingest
  - Receive raw science image packets from White Sands and engineering telemetry from MOS
  - Decompress and assemble image data into Level 0 FITS format, and mate with engineering telemetry
- 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
  - Provide rapid QA feedback for selected on-orbit performance (i.e. scan mirror synch)
  - Overall science data QA for survey planning and data product generation
- Science Data Archiving and Distribution
  - Archive raw data (Level 0) during mission
  - Provide an operational archive to store and serve intermediate data products and metadata to project team during mission
  - Provide a long-term "living" archive that stores and serves WISE science data products to the astronomical community and general public along with user's guide documentation and descriptive analysis

# Implementation Features



- The WSDC is an independent *task* at IPAC
  - Analogous to 2MASS, NHSC, IRSA, NED
  - IPAC/WISE Task Lead reports to IPAC Executive Director and works with IPAC Manager to organize staff and resources to perform task
  - Administrative, facilities and common infrastructure cost shared with all IPAC projects
  - WSDC will use dedicated hardware that is integrated into IPAC network structure
  - Resources and expertise shared with other IPAC tasks along with SSC, MSC which reside under a common administrative umbrella at IPAC
- Results in a substantial savings and improved quality because of access to expertise and synergy with other IPAC projects, including the Spitzer and Michelson Science Centers
- Science Data Processing/Archiving/Distribution integrated into WISE planning and design from the outset
  - WSDC task lead participates in WISE management meetings, telecons
  - Task lead, lead engineer and lead QA scientist participate in MOS, MSET, V&V meetings, telecons
  - WSDC personnel participate in Calibration WG, IOC planning, etc.

# Implementation - 2



- On-going interaction with WISE Science Team during all phases of the project
  - Five WISE co-Investigators are IPAC staff members (Cutri, Jarrett, Kirkpatrick, Lonsdale, Padgett) that provide strong scientific oversight for WSDC activities
  - "Cognizant" Science Team member will be assigned to each data processing pipeline subsystem and will work closely with WSDC cognizant engineer to develop and validate algorithms, and analyze on-orbit data (adopted from 2MASS)
  - Science Team will be heavily involved with development of data reduction algorithms, analysis of on-orbit data, validation of preliminary and final WISE data products, and in writing explanatory documentation
- WSDC software developers remain on staff throughout mission
  - Necessary for two-stage processing strategy - final processing takes place after the end of on-orbit operations
  - Retains personnel with key expertise in WSDC software systems, algorithms as well as WISE on-orbit data

# Key Implementation Assumptions



- The WISE project will deliver to the WSDC all pre-launch calibration products necessary to process WISE data
  - IPAC does not have pre-flight instrument characterization staff
  - IPAC will participate in updates for calibration products post-launch
- No requirements specified on intermediate data products such as single frames, solar system object associations other than they must be sufficient to support production of final product deliverable that meet requirements
- Software development at IPAC follows proven IPAC management methods and standards

# WSDC Status



- WSDC staffing and activity was kept at minimal level until start of Phase C/D (April 2007)
- Phase B work focused on cost and schedule refinement, requirements definition, high-level system design and key data and operational interface definition (MOS-GDS/WSDC)
- Phase C/D WSDC staff ramp-up in full swing
  - 7 of projected 11 development staff on-board. 4 to be added in FY08.
- Mature designs completed for many subsystems
  - Peer reviews complete for subsystems with significant new features
  - Selected additional peer reviews to follow CDR
- Work on some subsystems intentionally scheduled for start later this year (e.g. ARCHIVE, FPG, artifact ID, photometric calibration)

# Key Activities Leading up to WSDC CDR



- Successful Mission CDR (6/07), MOS CDR (7/07)
- Finalized WSDC Functional Requirements (FRD v2)
- Advanced versions of all major design documents (PDMP, FDD, SMP, Ops Plan)
- Completed top-level design of key INGEST and PIPELINE components
- Participated in Project-level Calibration Peer Review (9/07)
  - Peer reviews of ICAL, PCAL concepts
- Conducted peer reviews of PIPELINE subsystems with significantly new algorithms components
  - AWAIC (11/15/07), MDET (12/7/07), PREX (12/13/07)
- Demonstrated HRP science data INGEST in support of Spacecraft Mission Unique Board (MUB) testing
- WSDS v0 delivered
  - Internal data flow, operational data structure demonstration
  - Selected subsystem prototype functionality (detection, position reconstruction)
- Data product formats (image specifications/headers, catalog columns) discussion started with Science Team
- Supported detector testing to assist with flight FPA selection

# Action Item Status



- PDR
- Mission CDR
- MOS CDR

# WSDC Requirements are Mature



- Listed in WSDC Functional Requirements Document (WSDC D-R001)
- Tracing to Level 3 MOS requirements, where appropriate
- Self-derived requirements identified
- No TBDs

# Driving Requirements - 1



| Higher Level Requirement | Level 4 Requirement | Compliance | Verification Method |
|---|---|---|---|
| **L3MOS-366 (L1PP-8)** | **Image Atlas:** The WSDC shall produce a digital Image Atlas that combines multiple survey exposures at each position on the sky. | By design | Demonstration |
| **L3MOS-374 (L1PP-9)** | **Source Catalog:** The WSDC shall produce a Source Catalog derived from the WISE digital Image Atlas. | By design | Demonstration |
| **L3MOS-366,374 (L1PP-34)** | **Data Release:** The WSDC shall release to the public an image atlas and source catalog covering the full survey area within 17 months after the end of on-orbit data collection. | By design | Demonstration |
| **L3MOS-355, 363 (L1.5SRD-50)** | **Preliminary Data Release:** The WSDC shall release to the public a preliminary image atlas and source catalog covering at least 50% of the surveyed area within 6 months after the end of on-orbit data collection | By design | Demonstration |
| **L3MOS-417 (L1PP-010)** | **Catalog Reliability:** The final WISE Source Catalog shall have greater than 99.9% reliability for sources detected in at least one band with SNR>20, where the noise includes flux errors due to zodiacal foreground emission, instrumental effects, source photon statistics, and neighboring sources. This requirement shall not apply to sources that are superimposed on an identified artifact. | By design | Test: Comparison with external "truth tables"; detection confirmation |
| **L3MOS-363, 418 (L1PP-011)** | **Catalog Completeness**: The final WISE Source Catalog shall be at 95% complete for sources detected with SNR>20 in at least one band, where the noise includes flux errors due to zodiacal foreground emission, instrumental effects, source photon statistics, and neighboring sources. This requirement shall not apply to sources that are superimposed on an identified artifact. | By design | Test: Detection repeatability in deep coverage areas (e.g. the ecliptic poles) |

# Driving Requirements - 2



| Higher Level Requirement | Level 4 Requirement | Compliance | Verification Method |
|---|---|---|---|
| **L3MOS-376 (L1PP-012)** | **Photometric Accuracy:** The root mean square error in relative photometric accuracy in the WISE Source Catalog shall be better than 7% in each band for unsaturated point sources with SNR>100, where the noise flux errors due to zodiacal foreground emission, instrumental effects, source photon statistics, and neighboring sources. This requirement shall not apply to sources that superimposed on an identified artifact. | By design | Test: Photometric repeatability; stellar color stability over sky |
| **L3MOS-370 (L1PP-013)** | **Astrometric Accuracy:** The root mean square ($1\sigma$) error in WISE catalog positions with respect to 2MASS All-Sky Point Source Catalog positions shall be less than 0.5" on each axis | By design | Test: Comparison with 2MASS PSC and other astrometric catalogs (e.g UCAC) |
| **L3MOS-420 (L1PP-4)** | **Photometric Sensitivity:** Flux measurements in the WISE Source Catalog shall have a SNR of five or more for point sources with fluxes of 0.12, 0.16, 0.65 and 2.6 mJy at 3.3, 4.7, 12 and 23 micrometers, respectively, assuming 8 independent exposures and where the noise flux errors due to zodiacal foreground emission, instrumental effects, source photon statistics, and neighboring sources. | By design | Test: Photometric repeatability; Colors of normal stars, galaxies |
| **L3MOS-272** | **Quicklook Quality Assurance:** Approximately 3% of the science data from each orbit shall be processed and data quality summary reports posted to the WISE internal web site within 24 hours of receipt at the WSDC. | By design | Demonstration |

# Driving Requirements - 3



| Higher Level Requirement | Level 4 Requirement | Compliance | Verification Metho d |
|---|---|---|---|
| **L3MOS-286** | **Single Orbit Data Processing Latency:** Within 3 days of receipt of a given data set, the WSDC shall process all single orbit data through the first stage of pipeline processing . | By design | Demonstration |
| **L3MOS-387 (L1PP-34)** | **Archive Longevity:** The WSDC shall make the Image Atlas and Catalog products accessible to the astronomical community in collaboration with the NASA/IPAC science archive infrared to ensure long-term availability beyond the end WISE missions operations and data processing phase, and insure interoperability with other NASA mission archives.with other NASA mission archives. | By design | Demonstration |

# Changes Since MOS PDR



- Launch date moved to November 2009
- Extended Phase B completed
- Phase C/D started Feb. 2007
- Mission CDR June 2007
- MOS CDR July 2007
- Added prototype ingest system development to support MUB testing and in response to MOS PDR report recommendation
- Ground test data archive deployed
- Profile-fit photometry introduced into baseline design

# WSDC Mission Deliverables - 1



- WISE Image Atlas
  - FITS format, all exposures registered and combined, 4 bands registered. Footprint approx. Metadata table describing each image
  - Internal product footprint 1°x1°, pixel scale 1.375”/pix
  - Current size estimated to be ~300,000 2048x2048 images covering entire sky in four WISE bands

- WISE Source Catalog
  - Basic attributes of each object detected on combined images
    - Position on the sky (J2000) and associated uncertainties
    - Photometry in four WISE bands and associated uncertainties
    - Flux upper limits in non-detected bands
  - Source detection and measurement quality flags and parameters (*e.g.* detection statistics, reliability estimate, photometric quality, confusion and contamination)
  - A unique identifier (*i.e.* source name)
  - Additional information to enhance usability (*e.g.* association with 2MASS)
  - Current estimate of size is ~300 million entries (*Working Database* will contain ~$10^9$ entries)

# WSDC Mission Deliverables - 2



- WISE Explanatory Supplement
  - Mission and data product description,
  - User's guide (*e.g.* data formats, access modes)
  - Cautionary notes

- Ancillary Products

  *Not specified in higher level requirements. To be developed as resources and schedule allow*
  - Atlas Image coverage and noise maps
  - Solar system object association list (derived from single-epoch images)

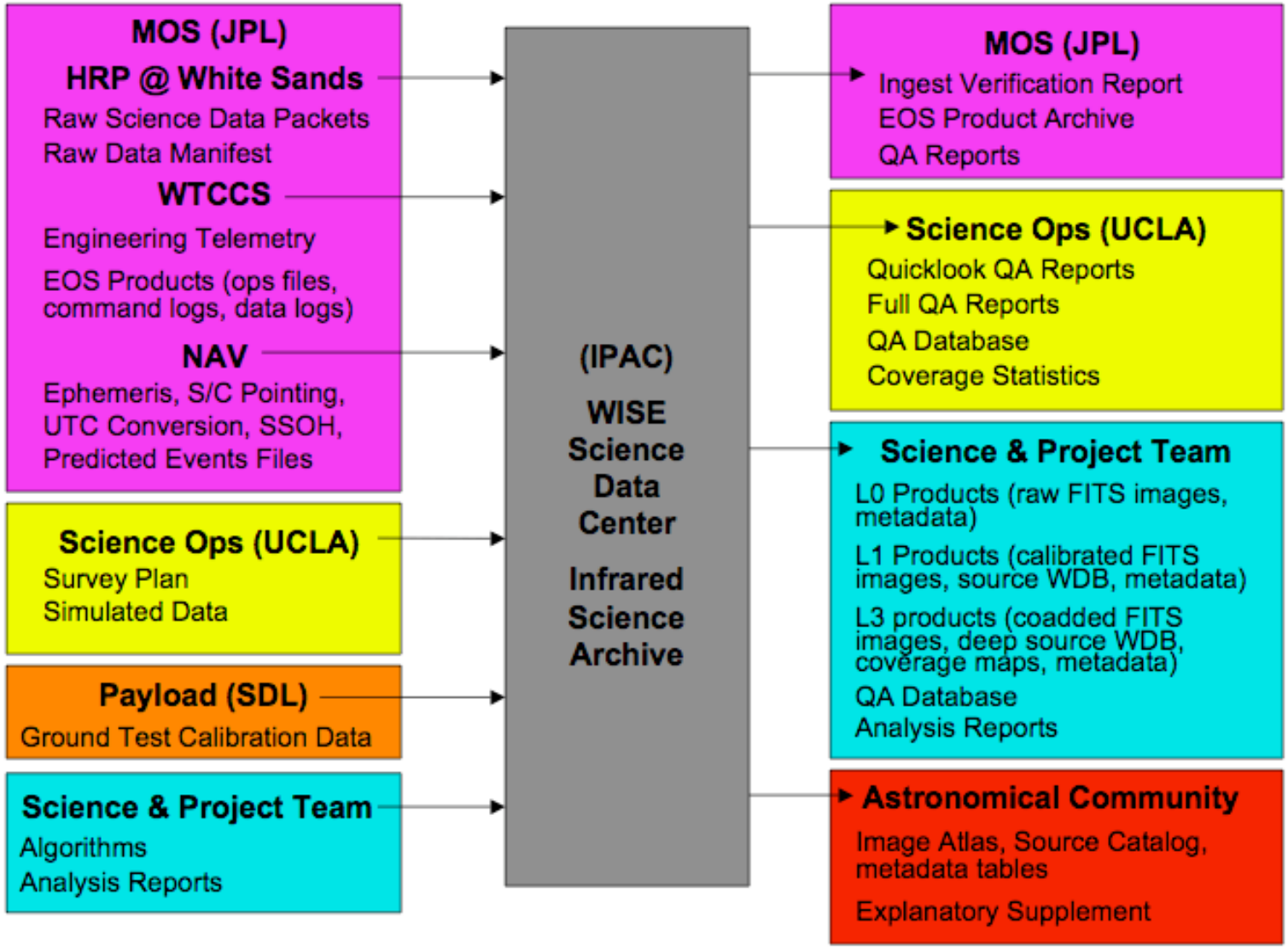
- Single-epoch (Level 1) images and extracted source database maintained during mission operational period for Project/Science Team access, *but not baselined for public release*

# Data Product Delivery



- Two-stage Data Release
  - Preliminary Image Atlas and Source Catalog
    - Derived from first 50% of sky surveyed
    - Release 6 months after end of on-orbit operations (12/2010)
  - Final Image Atlas and Source Catalog
    - Derived from all survey data
    - Release 17 months after end of on-orbit operations (11/2011)
- Data Access Mode (project team, science community and general public)
  - On-line services of NASA/IPAC Infrared Science Archive (IRSA)
  - IRSA will host Level 1, 3 (for project team access) and final product archive (for public access)
  - Option to provide bulk distribution of Source Catalog
- Level 0 Data Archive
  - At IPAC and off-site during mission
  - Permanent Level 0 Data Archive at NSSDC (Letter of agreement to be arranged in FY08)

# External Interfaces - 1



# External Interfaces - 2



- MOS/WTCCS & NAV (receive and acknowledge)
  - Receive Spacecraft ephemeris, pointing, VT-UTC conversion telemetry files
  - Receive Selected stored-state-of-health files
  - Receive Predicted Orbital Events files
- MOS/HRDP@White Sands (receive and acknowledge)
  - Receive High-rate science data
  - Receive Science data manifest
- Science Operations (receive and send)
  - Receive Science observation plan
  - Send Science data processing QA report to close loop on survey planning (on-line)
- Project Team (send and receive)
  - Send Intermediate processed data products and working databases
  - Send Science data processing QA reports
  - Receive data processing algorithms
  - Receive science data analysis reports and contributions for Explanatory Supplement
- Astronomical Community (send)
  - Image Atlas, Source Catalogs, ancillary products
  - Data product documentation (on-line Explanatory Supplement)

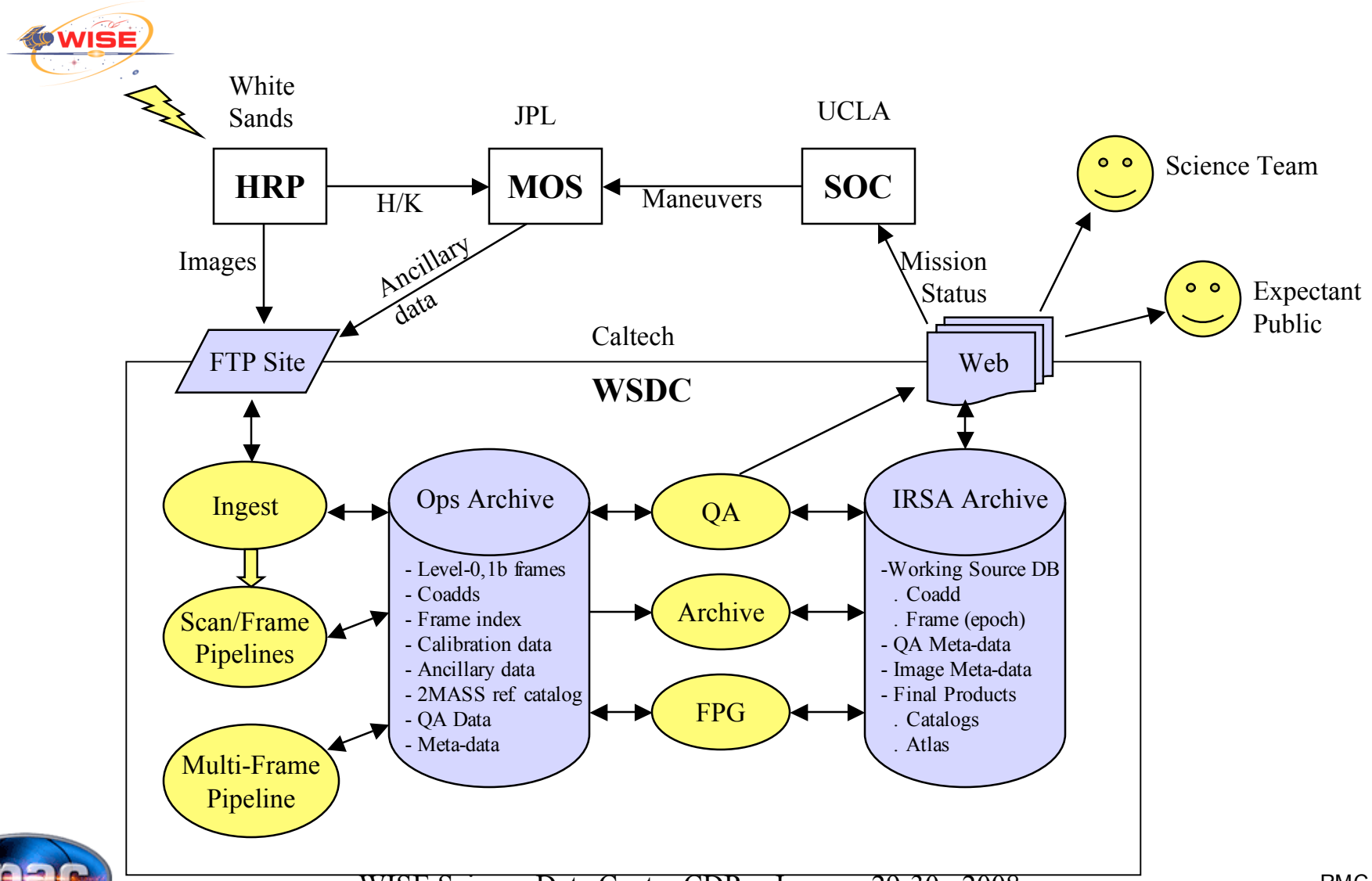
# WISE Science Data System (WSDS) Design Approach



- Software and operations system to execute WSDC functions
- Consists of six primary subsystems
  - EXEC, INGEST, PIPELINES, FPG, ARCHIVE, QA
- Based closely on systems used for 2MASS and other IPAC projects
  - Highly automated, "industrial strength" data processing software system designed for high-throughput, reliable operation
  - Extensive use of automated QA reporting
  - Modular system to facilitate parallel development, unit-testing
  - Extensive use of parameter control files to allow "tuning" for actual on-orbit performance
- Planned two-stage data processing and data release
  - "Can't get it right the first time"
  - Allows incorporating best knowledge of actual instrument performance, calibration and sky
  - Gets data out as rapidly as possible, and uses community as "beta-testers"
- Design from outset with product development and distribution in-mind
  - Source and image databases/archives developed within IRSA
  - Empower Science and Project Team by enabling easy access to intermediate and final data products

# WSDS Functional Block Diagram





# WSDS Subsystems

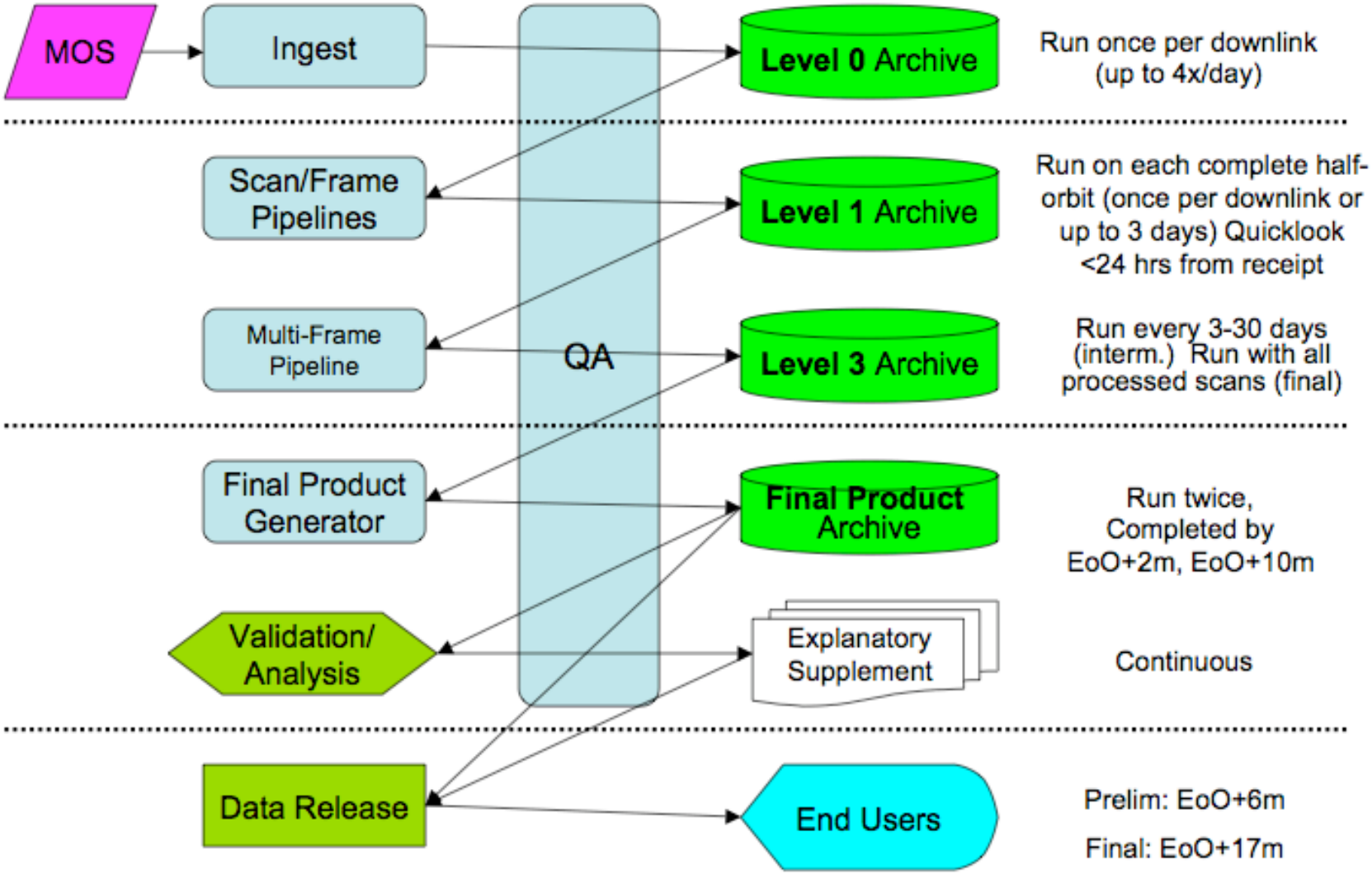


- INGEST
  - Receives science data packets and engineering telemetry from MOS and assembles Level 0 FITS-format files. Stages Level 0 images and metadata for pipeline processing.
- Data Reduction PIPELINES
  - Converts Level 0 imaging data into calibrated images and extracted source *Working Databases*
    - **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.

# Data Processing Levels

# Internal Data Flow and Operational Cycle





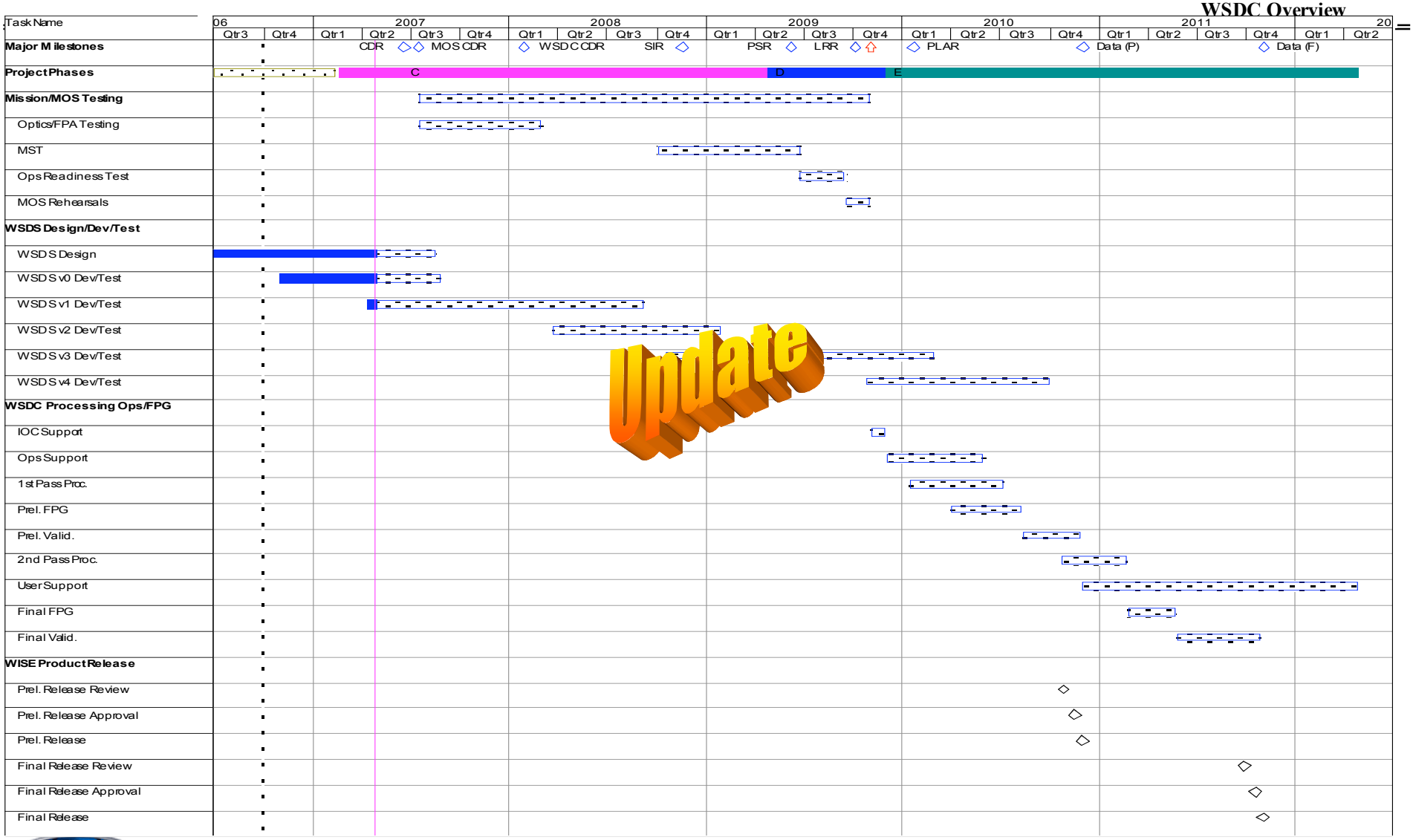
# Key Schedule Drivers



- MUB testing (mid-2007)
- Mission Scenario Testing (MST)
- Operations Readiness Testing (ORT)
- Launch/IOC (Nov. 2009 + 1 month)
- Short (6 month) operations lifetime
- Data releases
  - Preliminary (end of on-orbit data acquisition + 6 months)
  - Final (end of on-orbit data acquisition + 17 months)

# WSDC Dev/Ops Schedule





# WSDC WBS and Task Plan

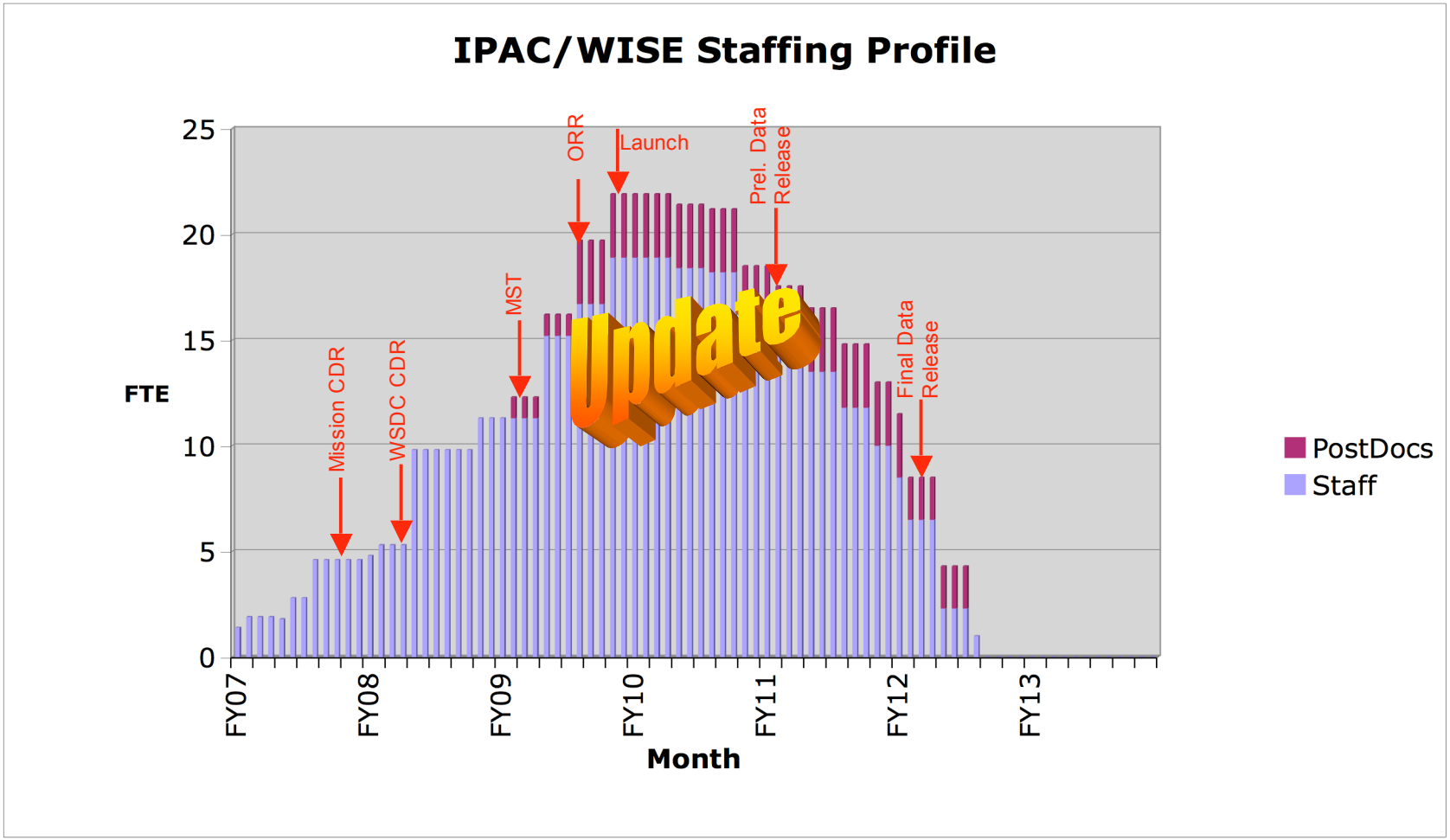


| Fiscal Year | FY07 | FY08 | FY09 | FY10 | FY11 | FY12 | FY13 | Total |
|---|---|---|---|---|---|---|---|---|
| Month | | | | | | | | |
| Phase | B/C | C | C/D | D/E | E | E | E | |
| ***Project Lead*** | | | | | | | | |
| Lead Scientist | 0.60 | 0.73 | 0.80 | 0.80 | 0.80 | 0.50 | 0.00 | 4.79 |
| Lead Engineer | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.17 | 0.00 | 3.08 |
| Lead QA/Analysis Scientist | 0.30 | 0.50 | 0.50 | 0.50 | 0.50 | 0.31 | 0.00 | 2.91 |
| ***Subystem Development Engineers*** | | | | | | | | |
| Pipeline Executive/Config. Mgmt./I&T | 0.13 | 0.81 | 1.00 | 0.92 | 0.29 | 0.00 | 0.00 | 3.15 |
| Data Ingest System | 0.18 | 0.47 | 1.00 | 0.58 | 0.00 | 0.00 | 0.00 | 2.23 |
| Instrumental Calibration | 0.00 | 0.50 | 0.50 | 0.50 | 0.22 | 0.00 | 0.00 | 1.72 |
| Image Construction/Mosaicing | 0.21 | 0.83 | 1.00 | 1.00 | 0.54 | 0.17 | 0.00 | 3.75 |
| Source Detection | 0.21 | 0.50 | 0.50 | 0.50 | 0.29 | 0.00 | 0.00 | 2.00 |
| Source Photometry | 0 | 0 | 0 | 1.50 | 1.08 | 0.17 | 0.00 | 5.63 |
| Position Reconstruction | 0 | 0 | 0 | 0.92 | 0.50 | 0.17 | 0.00 | 4.00 |
| Bandmerge | 0 | 0 | 0.50 | 0.42 | 0.00 | 0.00 | 0.00 | 1.25 |
| Artifact Identification | 0 | 0 | 0.83 | 1.00 | 0.54 | 0.10 | 0.00 | 3.19 |
| Photometric Calibration | 0 | 0.33 | 0.83 | 1.00 | 0.88 | 0.08 | 0.00 | 3.13 |
| Known Solar System Object ID | 0.00 | 0.00 | 0.27 | 0.28 | 0.00 | 0.00 | 0.00 | 0.55 |
| Quality Assurance System Development | 0.00 | 0.33 | 0.71 | 0.92 | 0.46 | 0.02 | 0.00 | 2.44 |
| ***Operations, QA, Analysis, Support*** | | | | | | | | |
| Pipeline Operations | 0.00 | 0.08 | 1.00 | 2.00 | 1.58 | 0.04 | 0.00 | 4.71 |
| Instrument Characterization Scientists | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Quality Assurance Scientists | 0.00 | 0.00 | 1.08 | 2.00 | 2.00 | 0.33 | 0.00 | 5.42 |
| Documentation and User Support | 0.00 | 0.00 | 0.03 | 0.20 | 0.48 | 0.31 | 0.00 | 1.02 |
| ***Archive*** | | | | | | | | |
| Archive Interface/Load | 0.00 | 0.04 | 0.42 | 0.50 | 0.26 | 0.07 | 0.00 | 1.29 |
| Archive Support (DBA,Metadata Mgmt,Tool Dev) | 0.00 | 0.04 | 0.92 | 2.00 | 2.00 | 0.56 | 0.00 | 5.52 |
| **Total Staff (FTE)** | 3.13 | 8.51 | 14.89 | 18.03 | 12.92 | 2.99 | 0.00 | 61.75 |
| | | | | | | | | |
| **PostDoctoral Analysts (FTE)** | 0.00 | 0.00 | 1.75 | 3.00 | 3.00 | 1.25 | 0.00 | 9.00 |
| | | | | | | | | |
| **Total Workforce by Month (FTE)** | **3.13** | **8.51** | **16.64** | **21.03** | **15.92** | **4.24** | **0.00** | **70.75** |

Update

# WSDC Staffing Profile





# What will not be covered in this CDR



- No presentations for subsystems for which design work has not yet been started, by design
  - Artifact identification/mitigation (ARTID)
  - Final Product Generation Subsystem (FPG)

  *Both have considerable design heritage from 2MASS, Spitzer*

- Conceptual design presentations only for subsystems which have only recently started work, by design
  - Multi-frame position reconstruction (MFPREX)
  - Source photometry (WPHOT)
  - Archive Subsystem

# WSDC Concerns



- Sound ground test plan in place, but responsibility for generation and delivery of pre-flight calibration products to WSDC is still being negotiated

  *Mitigation*: Product generation and delivery plan in active discussion at project level.

- Surprises in actual on-orbit instrument performance

  *Mitigation*: Modular S/W system with extensive parameter file control. Close involvement with payload ground test design discussion. Take advantage of lessons learned from Spitzer (Si:As). Lien for additional instrument characterization scientists.

- Aggressive data release schedules allows very little time for validation

  *Mitigation*: Relaxed requirements for Preliminary Release. Design for automated QA concurrent with processing, tied to science metrics linked to key science requirements (2MASS heritage).

- Handling deliveries of fragmented, out-of-order, and/or incomplete data frames

  *Mitigation*: Design for expected fragmentation. Working with MOS to insure that sufficient information to reassemble fragmented files will be provided.

# Summary

- WISE data processing, archiving and distribution tasks are challenging, but we know how to do them
- The WSDC processing and archiving system design is based closely on proven systems
  - Scope and functionality of WISE data processing, quality assurance, final product generation and archiving tasks are similar to those of 2MASS.
  - 2MASS photometric/astrometric accuracy, completeness and reliability requirements were more stringent than those for WISE, and all were met or surpassed.
  - WSDC Level 0 data ingest system design draws upon operational GALEX system (not strictly an IPAC institutional activity, but supported by IPAC staff detailed to GALEX)
  - Leveraging existing infrastructure of the NASA/IPAC Infrared Science Archive
- Staff ramp-up, design and development proceeding on schedule and on-budget
- Key challenges are identified