WISE Science Data Center Hardware Peer Review

Roc Cutri - IPAC
WISE Science Data Manager
Tim Conrow - IPAC
WSDC Architect and Lead Engineer
Heidi Brandenburg – IPAC
WSDC Hardware and QA Cogniscent Engineer

Review Panel

- Wendy Burt (IPAC)
- Steve Groom (IPAC)
- Sharon Brunett (CACR)

Charge to Panel

The panel is asked to comment on the following:

- Is the WSDC processing system capable of handling our expected data volume and throughput?
- Are there any fragilities that might impede our ability to smoothly manage the ops system for the mission duration?
- Are the backup plans sufficient to recover from hardware outages?
- Please share any cluster or high-volume disk server lessons learned.

What Is WISE?

- A NASA Medium Explorer (MIDEX) Mission launching Nov. '09
- The Wide-field Infrared Survey Explorer (WISE)
  - An all-sky survey at 3.3, 4.7, 12 & 23 μm with 3 to 6 orders of magnitude more sensitivity than previous surveys
  - A cold 40 cm telescope in a sun-synchronous low Earth orbit
- WISE will deliver to the scientific community
  - An Image Atlas containing ~20,000 calibrated, rectified images covering the whole sky in 4 infrared bands
  - A Source Catalog of $\approx 5 \times 10^8$ objects seen in these 4 IR bands
- WISE will find
  - The most luminous galaxies in the Universe.
  - The closest stars to the Sun.
  - Most Main Belt asteroids larger than 3 km.

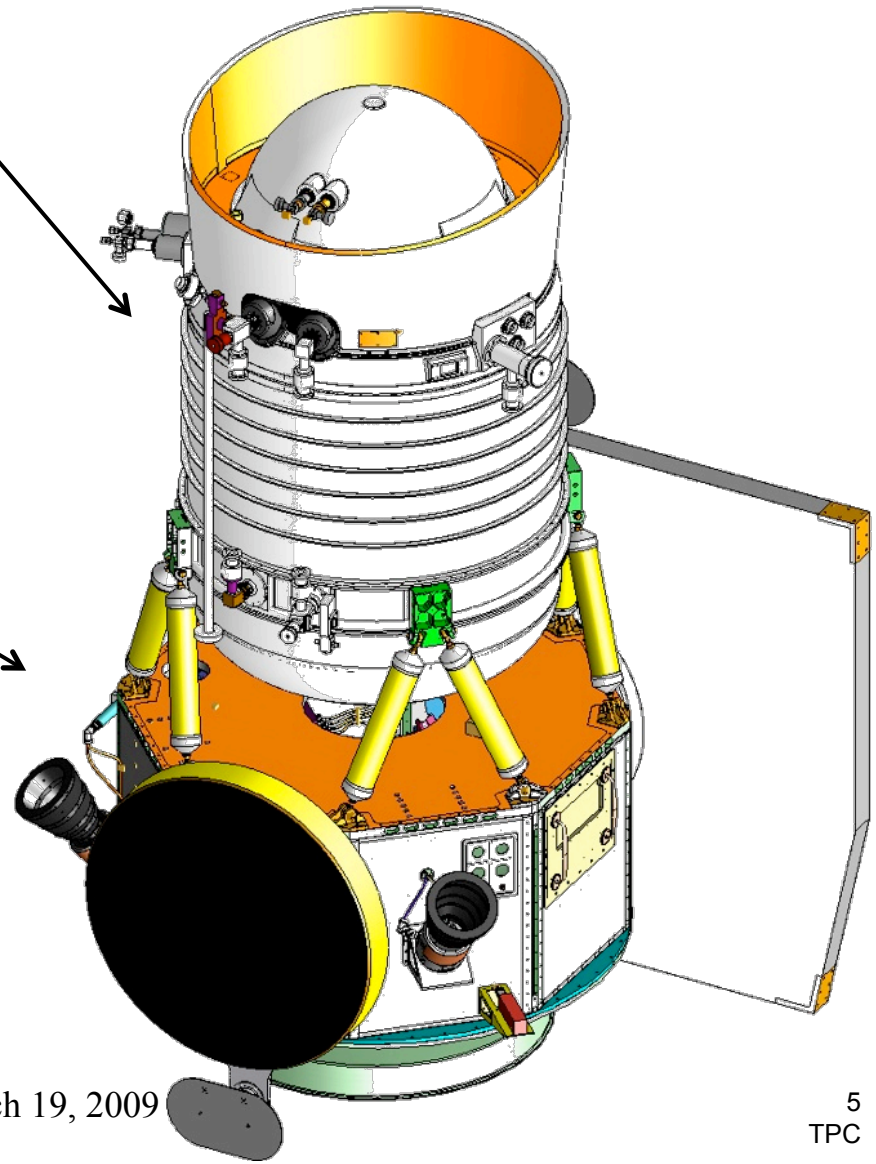
Flight System

Payload (Space Dynamics Lab)

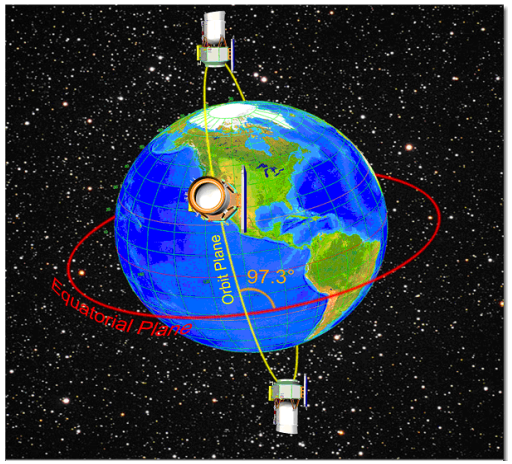
- 2-Stage Solid H_2 cryostat
  - **13.5 months life time (7 required)**
- All aluminum reflective optics: <17K
  - 40-cm telescope
- Dichroic beamsplitters separate wavelengths onto four 1024^2 pixel arrays (2.75”/pix)
- 2 HgCdTe detectors: 3.3, 4.7 microns (32K)
- 2 Si:As detectors: 12, 23 microns (7.8K)
- 3 electronics boxes (mounted in spacecraft)

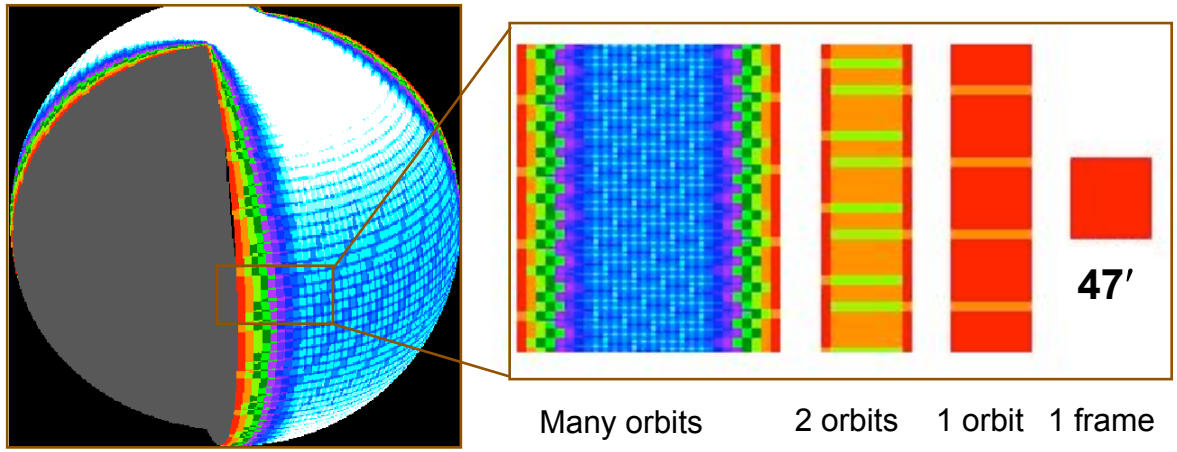
Spacecraft (Ball Aerospace)

- Orbital Express architecture
- Augmented single string
- No mechanisms, no deployables, no propulsion
- 3-axis stabilized
- Pointing stability/accuracy: ~ 1”/ ~1’
- Ku band science data link: 100Mbps
- 3.5 days (96 GB) of science data storage



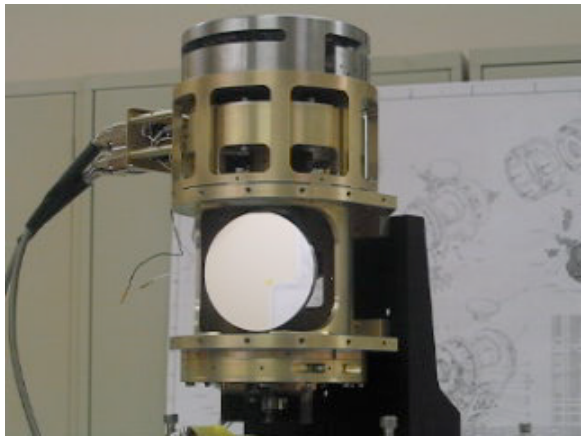
Simple Mission Design





- Delta 7320 launch – WTR
- 523 km, circular, polar sun -synchronous orbit
  - Nodal crossing time 6:00 PM
  - One month of checkout
  - **6 months of survey operations**
  - Extended 13 month mission proposed
- One simple observing mode – half orbit scan

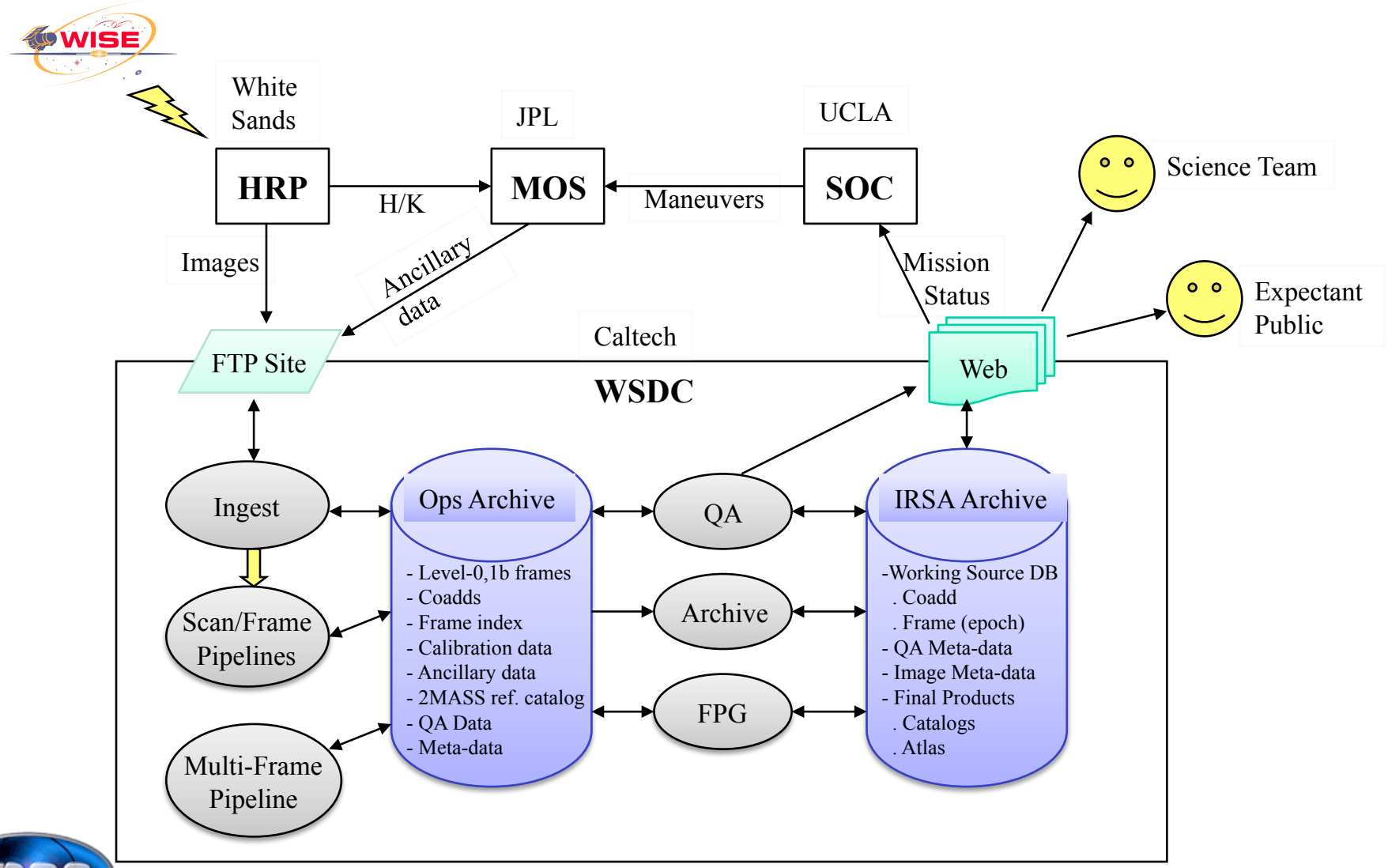
- Scan mirror "freezes" orbital motion enabling efficient mapping
  - 8.8-s exposure/11-s duty cycle
  - 10% frame to frame overlap
  - 90% orbit to orbit overlap
- Expect to achieve at least 8 exposures/position after losses to Moon and SAA
- Uplink, downlink, calibrations at poles
  - 4 TDRSS contacts per day



WSDC's Role

- Critical Mission Operations Functions
  - Receive, decompress, archive raw telemetry
  - Support In Orbit Checkout to tune and commission instrument
  - Perform "quicklook" QA to confirm proper instrument function
- Science Processing Functions
  - Produce level-0 data archive
  - Process level-0 data to produce level-1 calibrated frame products
  - Perform multi-frame processing to produce final, level-3 products
- Data Access Functions
  - Provide data access to engineers and Science Team during ops
  - Provide public data access to data products
    - Preliminary: EOO+**6 months**
    - Final: EOO+17 months

WSDC Functional Block Diagram



Requirements and Limitations

Tim Conrow - IPAC
WSDC Lead Engineer

Data Volumes

- Ops archive
  - Raw telemetry: 25 GB/day (over 4 downlinks)
  - L0: 13 MB/FS * 7500 FS/day = ~100 GB/day
  - L1: 75 MB/FS * 7500 FS/day = ~570 GB/day (300 GB images)
  - L3: <100 GB/day (removable?)
  - Total: < **~800 GB/day** uncompressed
    - 7 mo. Mission total: **~150 TB**
  - Total: < **~650 GB/day** compressed (33% image compression)
    - 7 mo. Mission total: **~120 TB**
- Node-local temporary work space (mostly L2 images)
  - Max: 600 MB/FS * 7500 FS/day / 35 nodes = **130 GB/day/node**
  - Min: 350 MB/FS * 7500 FS/day / 35 nodes = **75 GB/day/node**

Requirements and Limitations

- **Top-level Throughput Requirements**
  - Quicklook results within 24 hours (L4WSDC-32).
  - Scan pipeline results within 3 days (L4WSDC-39). A different requirement may be driven by NEO-WISE turn-around requirements, but these have not been clarified.
  - Preliminary public data release EOO+**6 mo**.s (L4WSDC-4).
  - Final public data release EOO+17 mo.s (L4WSDC-8).

The preliminary data release might impose a more demanding requirement than the latency requirements since under some FPG scenarios they imply a need to do final reruns while we are still operating.

Requirements and Limitations

- **Derived Throughput Requirements**
  - 24 hours of data ingested in 2 hours
  - 24 hours of data processed by scan/frame pipelines in 6 hours
  - Daily ops coadds produced in 4 hours

    Note there is no top-level requirement imposed on ops coadds. These are for QA /trending and care and feeding of the Science Team but can be squeezed if needed
  - Daily ops allocation
    - 2 hours ingest
    - 6 hours scan/frame primary processing
    - 6 hours scan/frame re-processing
    - 4 hours ops coadds
    - 6 hours backups, msc. Maintenance and downtime
  - A **goal** is to fit 24 hours of scan/frame processing in 3 hours

Requirements and Limitations

- **Ops Archive size and network limitations**
  - Ops Archive size limited by
    - Budget
    - Backup Time (disaster recovery)
    - Network capacity
  - New ops archive processed data generated per day: ~800GB (uncompressed)
  - Estimated backup time for 1 TB: ~6 hours
  - Network usage assuming 1 TB data read/written
    - 6 hours: 1TB/6 hours = ~50 MB/s
    - 3 hours: ~100 MB/s
    - Assuming the sustained max. loads will be about twice average, the max. sustained network load is **~200 MB/s**
  - Network usage by coadds still largely unknown

Requirements and Limitations

- **Node-local temporary work space usage**
  - Local storage limited by
    - Budget
    - Disk slots and rack space
    - Disk interface capacity
    - Clean-up latency
  - Read/write speed required
    - 130 GB / 6 hours = 6 MB/s average => ~12 MB/s sustained peak
    - 130 GB / 3 hours = 12/ MB/s average => ~25 MB/s sustained peak
  - Clean-up latency: how long can we wait before removing temp. files, i.e. how long do QA analysts have access to the data
    - 490 GB/node / 130 GB/day/node = 3.8 days max.
    - 490 GB/node / 75 GB/day/node = 6.5 days max.

Requirements and Limitations

- **RAM usage**
  - RAM usage limited by
    - Budget
    - Dell 1950 max. of 64 GB/node
  - More memory means fewer nodes at a trade rate of 8 GB added per node costs ~1 node
    - Could be cheaper using 3rd party RAM