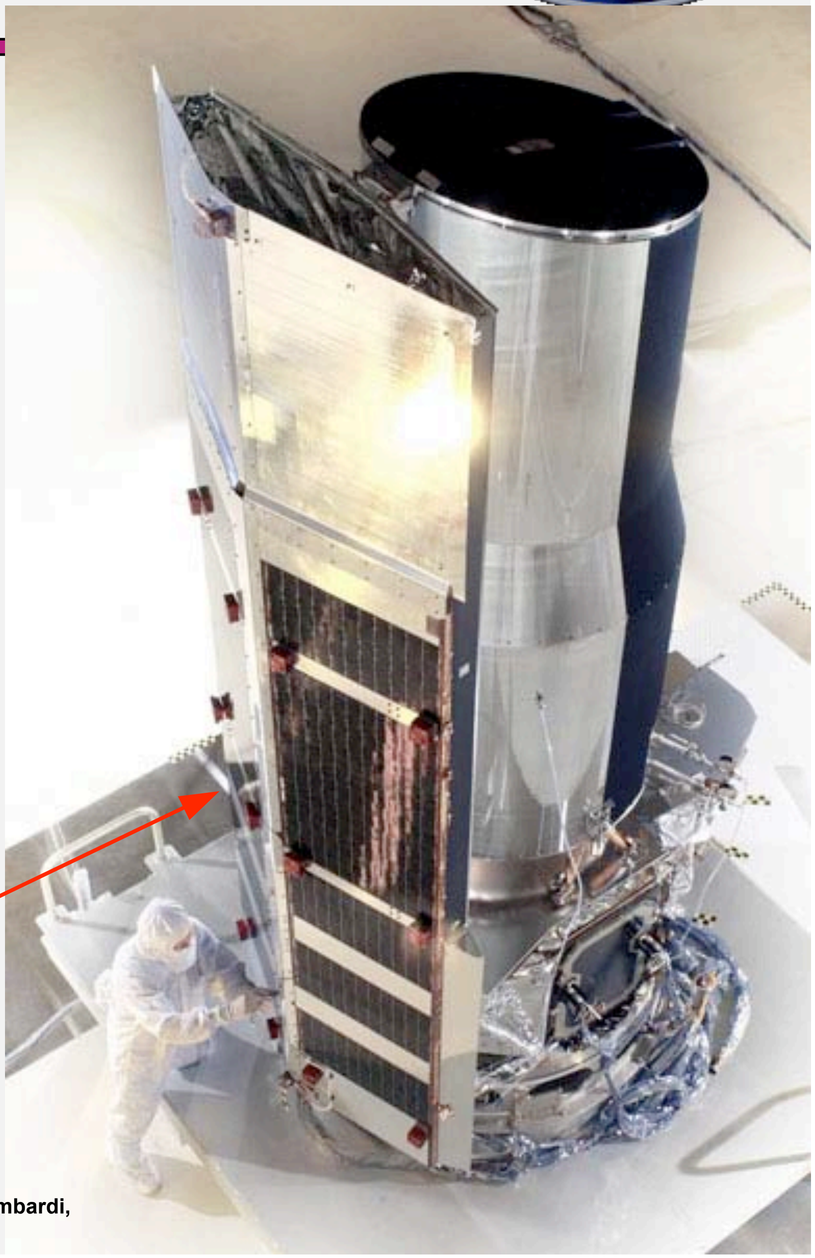
Spitzer Science Center within an Enterprise Architecture

Tom Handley, Tung Nguyen, Steven Lo, Rick Ebert, Laraine Amy, Jack Lampley, Jeff Jacobson, Steve Schurr, Carl Smay, Ed Jackson, Jonathan Kakumasu, Alex Hui, Jake Llamas, John White, Tracy Wang, Dennis Wittman, Mary Wittman, Daniel Moreno, Rosanne Scholey, Louisa Fung, Johnnie Higgins, Ron Beck

Infrared Processing and Analysis Center (IPAC) and Spitzer Science Center (SSC),
California Institute of Technology,
Pasadena, California, 91125

The Spitzer Space Telescope*

- Multi-purpose observatory cooled passively and with liquid-helium for astronomical observations over six octaves of wavelength in the infrared
- Liquid helium to last ~5 years
- 1-2 orders of magnitude improvement in sensitivity and performance
- Warm launch, solar orbit
- Completes NASA's Great Observatories
- Provides cornerstone science for NASA's Origins Theme, especially JWST



Assembled SIRTF Observatory at Lockheed-Martin, Sunnyvale.
<u>Key Characteristics:</u>
Aperture – 85 cm
Wavelength Range - 3-to-180um
Telescope Temperature – 5.5K
Mass – 870kg
Height – 4m

*Background slides "Spitzer Science Center: Science and Operations", Lisa Storrie-Lombardi, OMB Briefing, September 13, 2005

Spitzer Focal Plane Instruments

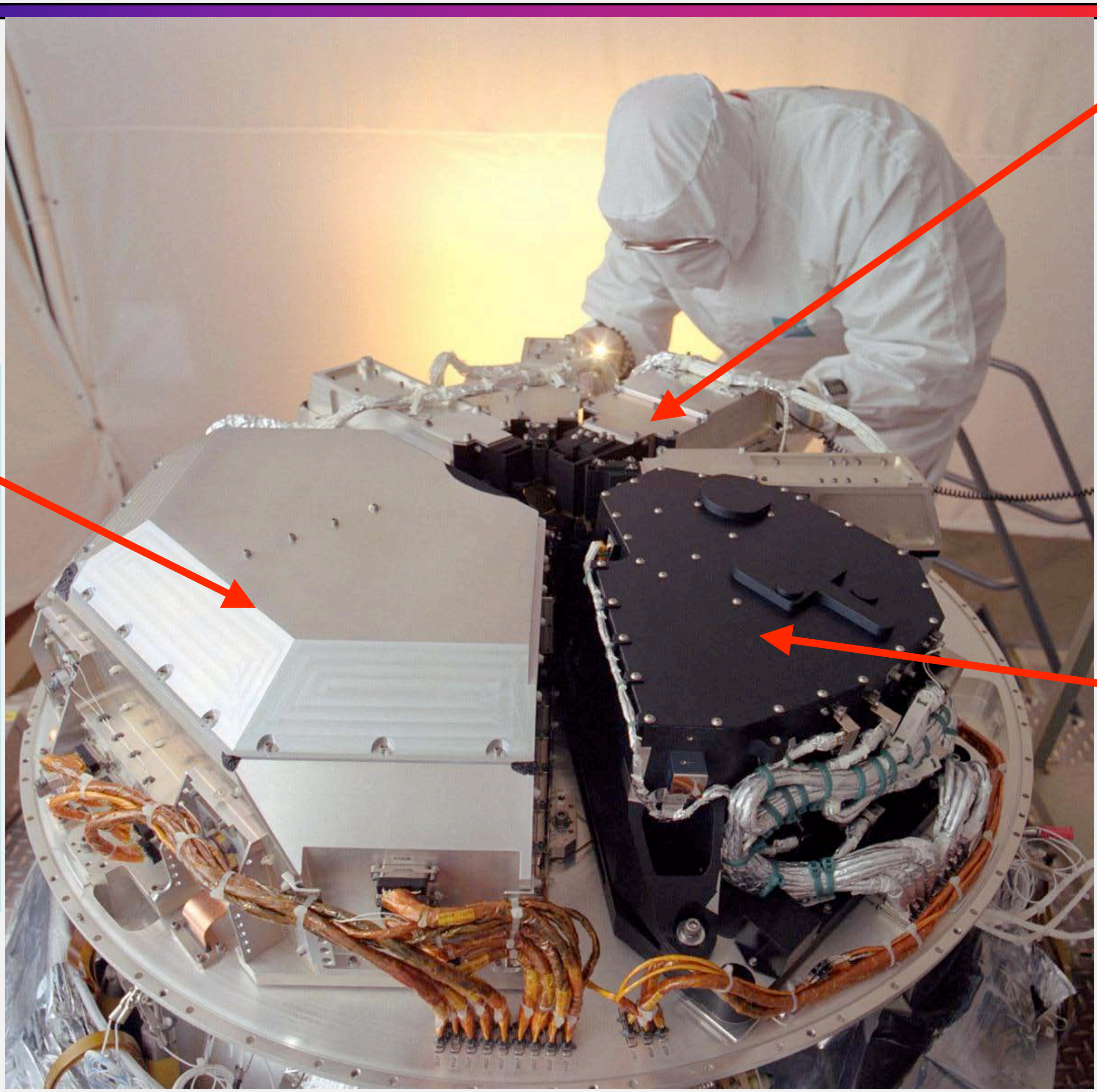
Integrated at Ball Aerospace, Boulder, May, 2001

MIPS

Imaging 5'x5'
24μm: ~0.1 mJy
70 μm:~2 mJy
160μm: ~4 mJy
[1σ in 100 sec]

Spectroscopy
50-100 μm R~20

G.Rieke,
U Arizona/BATC



IRS

Spectroscopy
5 to 40μm R~100
~0.1 to 0.3 mJy

10 to 40μm R~600
~1 to 3 mJy
[1σ in 500 sec]

J.R. Houck
Cornell/BATC

IRAC

Imaging 5'x5'
3.5 μm: ~2 μJy
4.6 μm : ~2 μJy
5.8 μm : ~10 μJy
8 μm : ~10 μJy
[1σ in 100 sec]

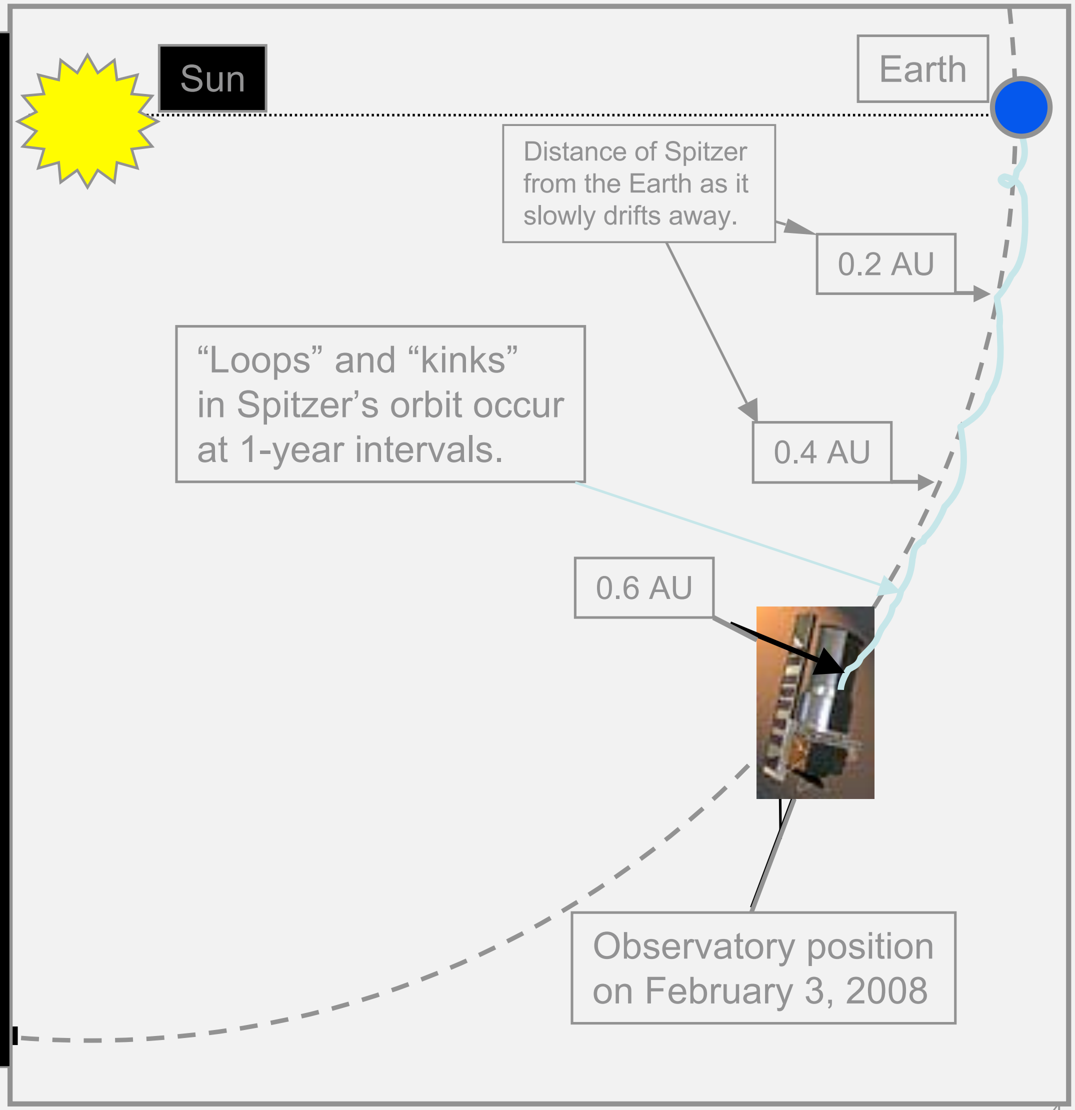
G. Fazio
SAO/GSFC

Spitzer's Solar Orbit

Why a Better Choice?

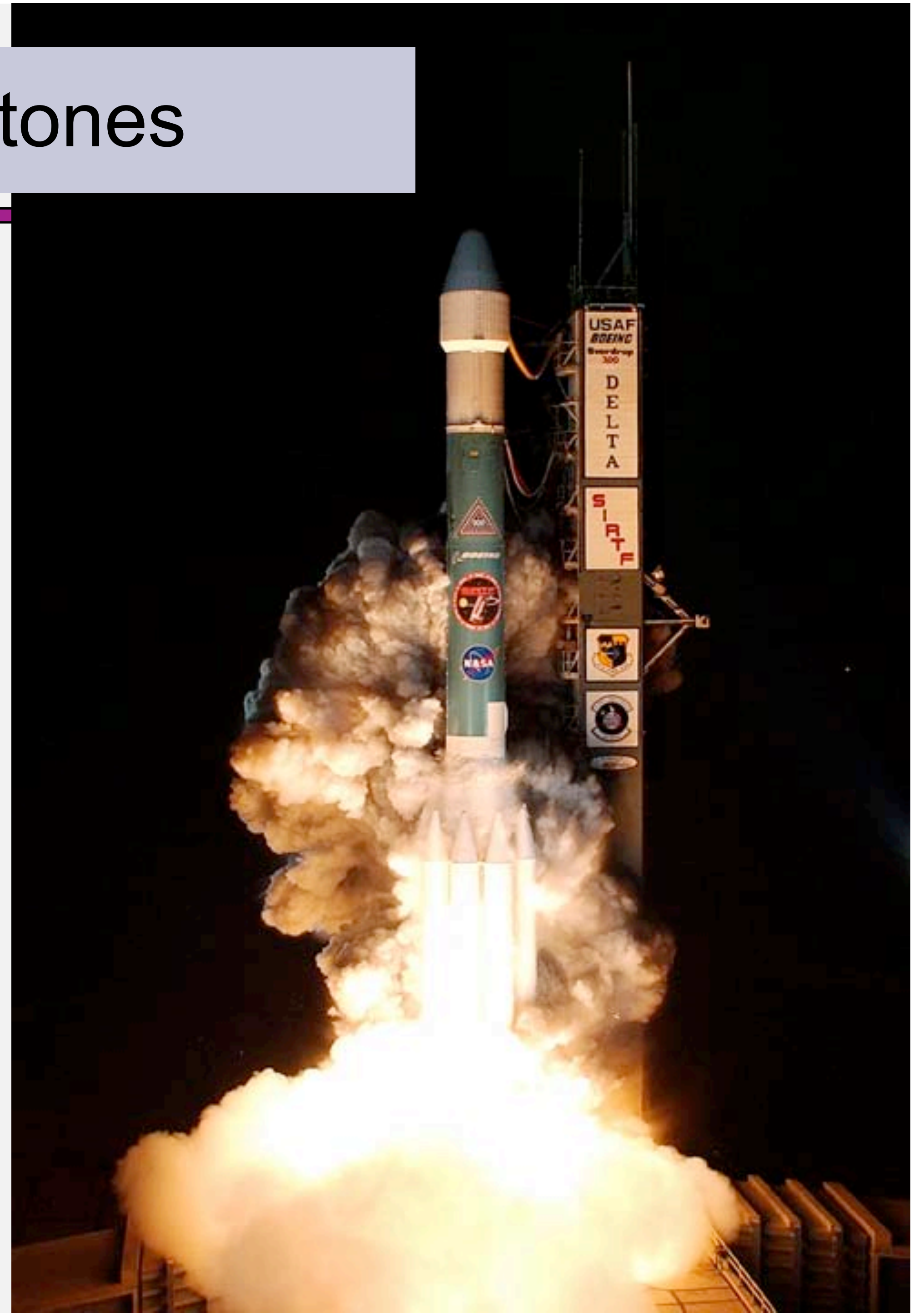
- **Stable Thermal Environment**
  (allows passive cooling)
- **No Earth Radiation Belt**
  (no damage to detectors or electronics)
- **No Need for Earth-Moon Avoidance**
  (Maximizes observing time, simplifies scheduling)

We are currently scheduling 6500-7000 hours of science per year!



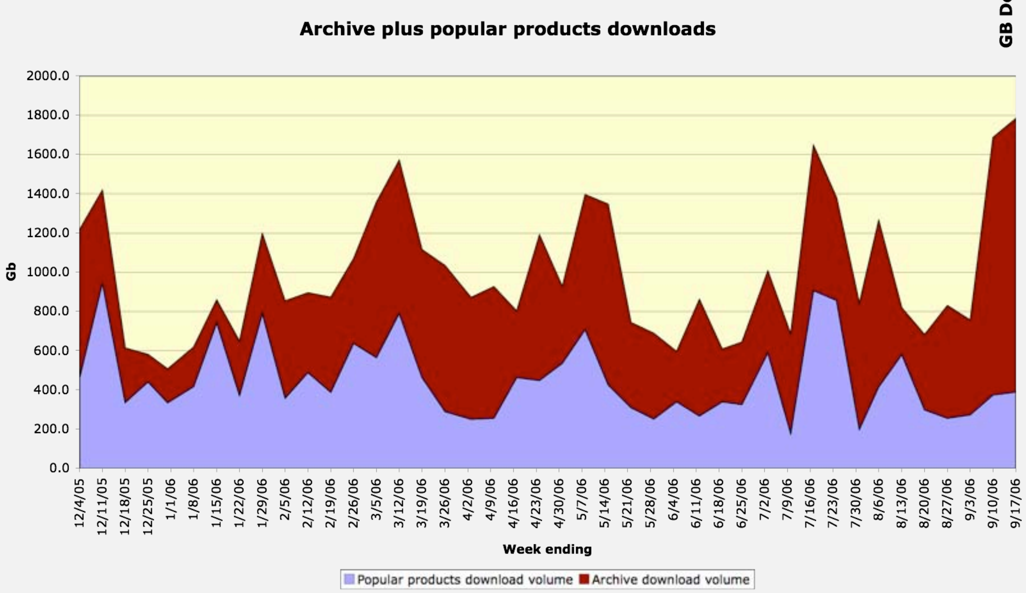
Milestones

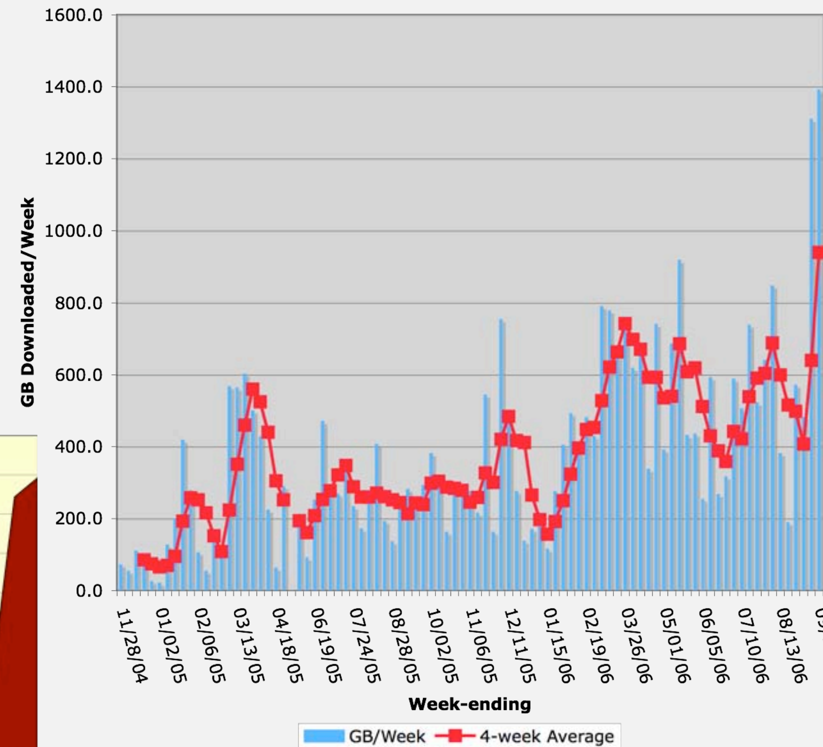
- Launch + 3 years
  - August 25, 2003 launch
  - May 11, 2004 – opened the Science Data Archive
  - May, 2006 – completed the Third General Observer Proposal Call
  - November,2006 – Issue Cycle-4 Call for Proposals
  - Observatory is operating nominally (very well behaved)
  - Expected helium lifetime is 5.3 years.
  - Executed our 17,830th hour of **science observations** October 11, 2006.



Archive Downloads*

Public downloads from the SSC archive peaked to 1400 GB/week a few weeks ago.





To get Leopard: http://ssc.spitzer.caltech.edu/tools/
To get popular products: http://data.spitzer.caltech.edu/popular/

*Cumulative statistics as of September 17, 2006 courtesy of Mark Lacy, SSC.

Process: SSC's approach to architecture, design, & development

- SSC chose an evolutionary approach*-
  - Allowed SSC to deal with uncertainty and evolution in requirements and in technology.
    - Especially important for the SSC : a long development phase (4-5 years) and a long system life (8-11 years)
  - Requirements could be more abstract and therefore subject to interpretation at any point in time.
    - As level 1 & 2 requirements were virtually immutable, level 3 detailed requirements remained more abstract to give SSC flexibility.
  - Alternative solutions could be explored and could be pursued further as new technology options became available.
  - Intermediate designs could be saved for future use.
  - Intermediate designs could be implemented as prototypes but never operationally implemented.
  - A functional architecture was used to partition and allocate requirements to different subsystems and operational modes.

* SE Tutorial Sys Architectures, Kasse Initiatives, LLC

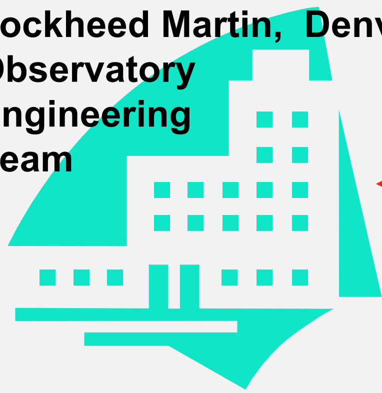
Development: SSC's approach to architecture, design, & development

- SSC chose a "freeze, build, & deliver" approach
  - When there is a need to build a system, the available solution that best meets the current requirements is selected and implemented using the traditional engineering approach: incremental system deliveries ("S" denotes "System")
    - S1.0 December , 1998 (SOS/FOS interfaces)
    - S2.0 October, 1999

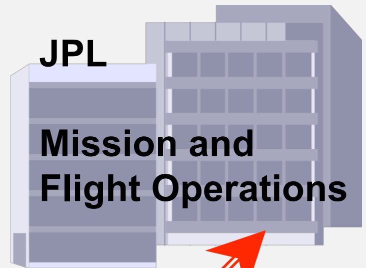
    ….
    - S15.0 November 1,2006.

Where is the Spitzer Science Center (SSC) within the Spitzer Enterprise?

Lockheed Martin, Denver – Observatory Engineering Team



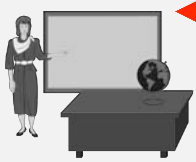
JPL



Mission and Flight Operations

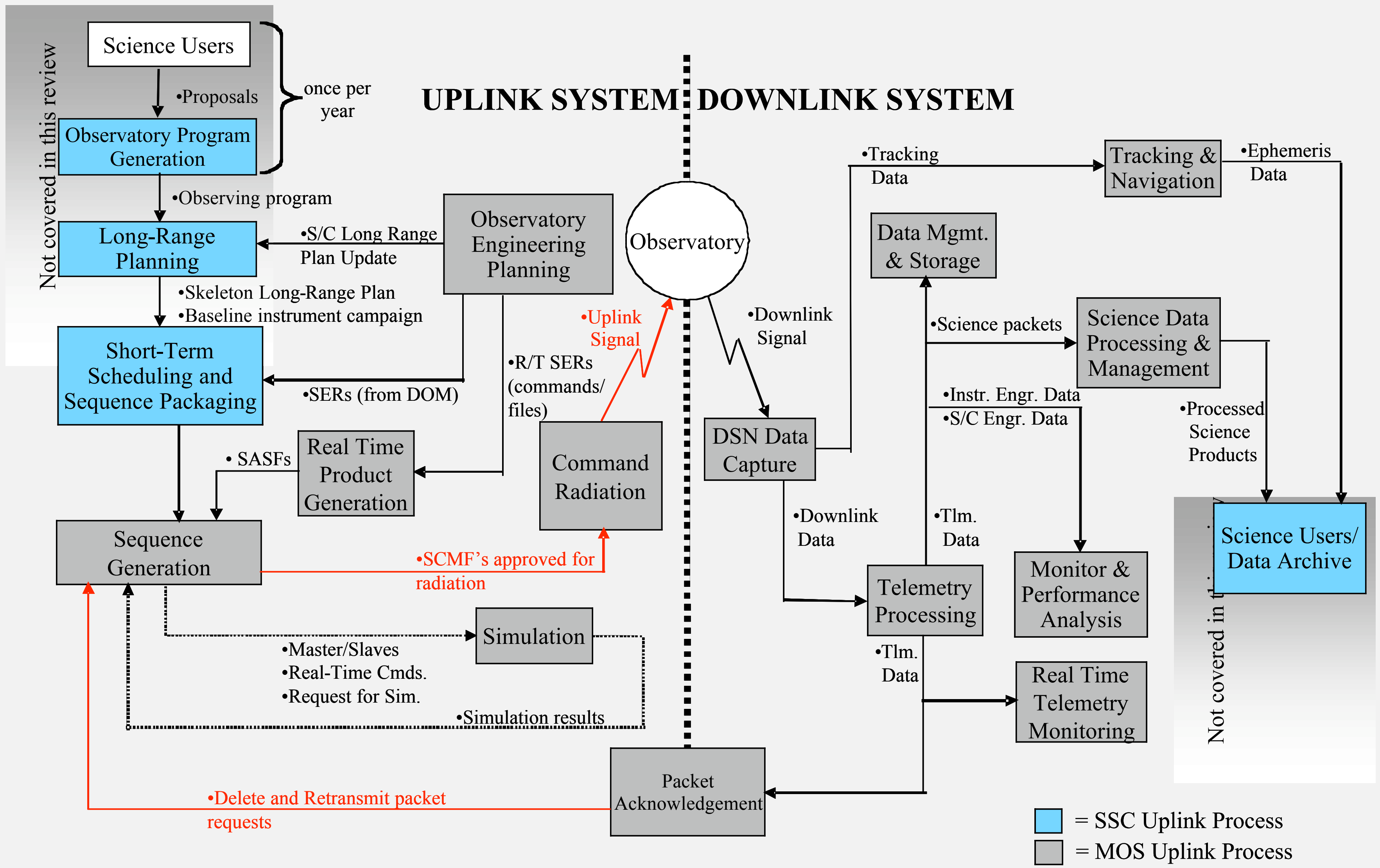
Spitzer Science Center, Caltech

Science User Community



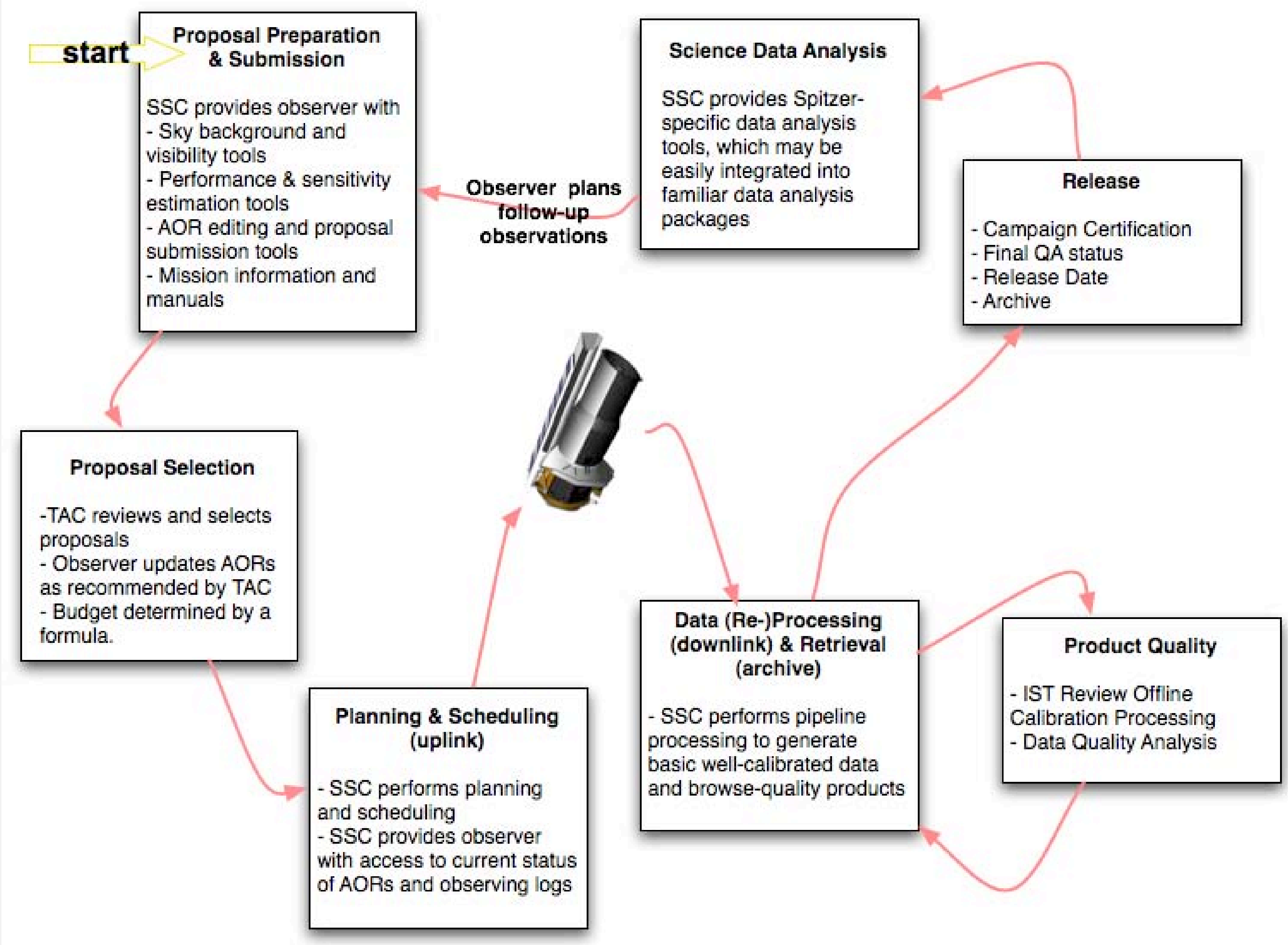
Education & Public Outreach

Spitzer Ground Data System

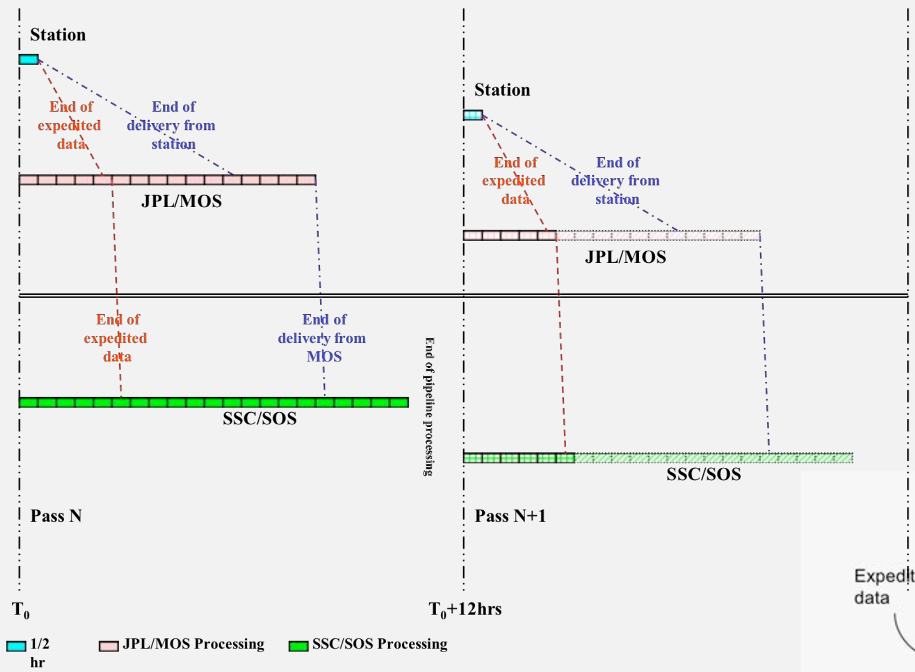


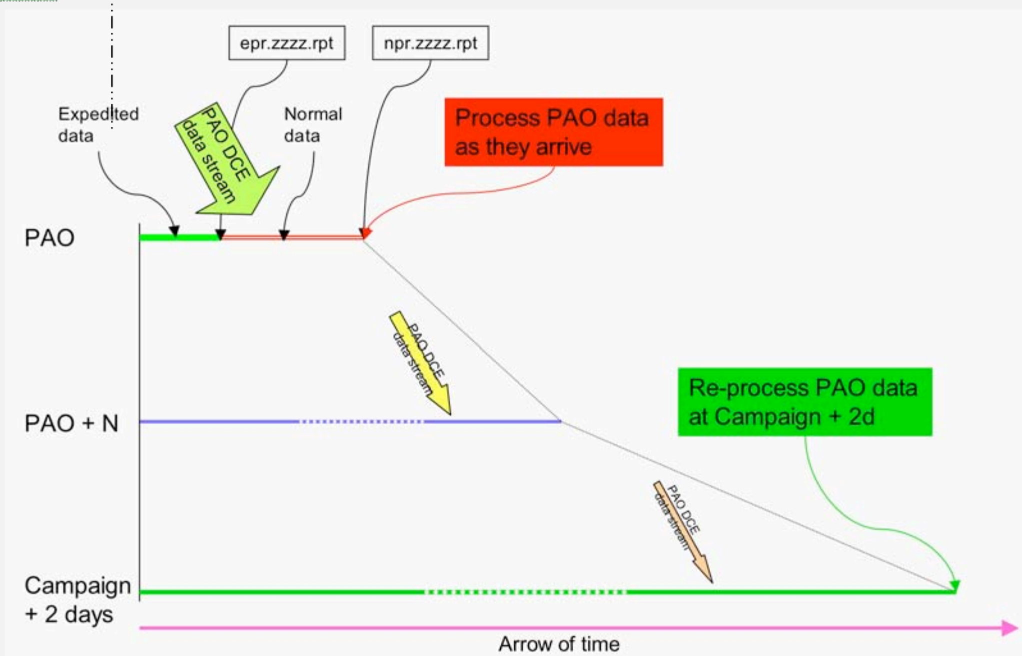
"Spitzer observatory operations — increasing efficiency in mission operations", Charles P. Scott, et al, SPIE 2006

SSC Operations Flow



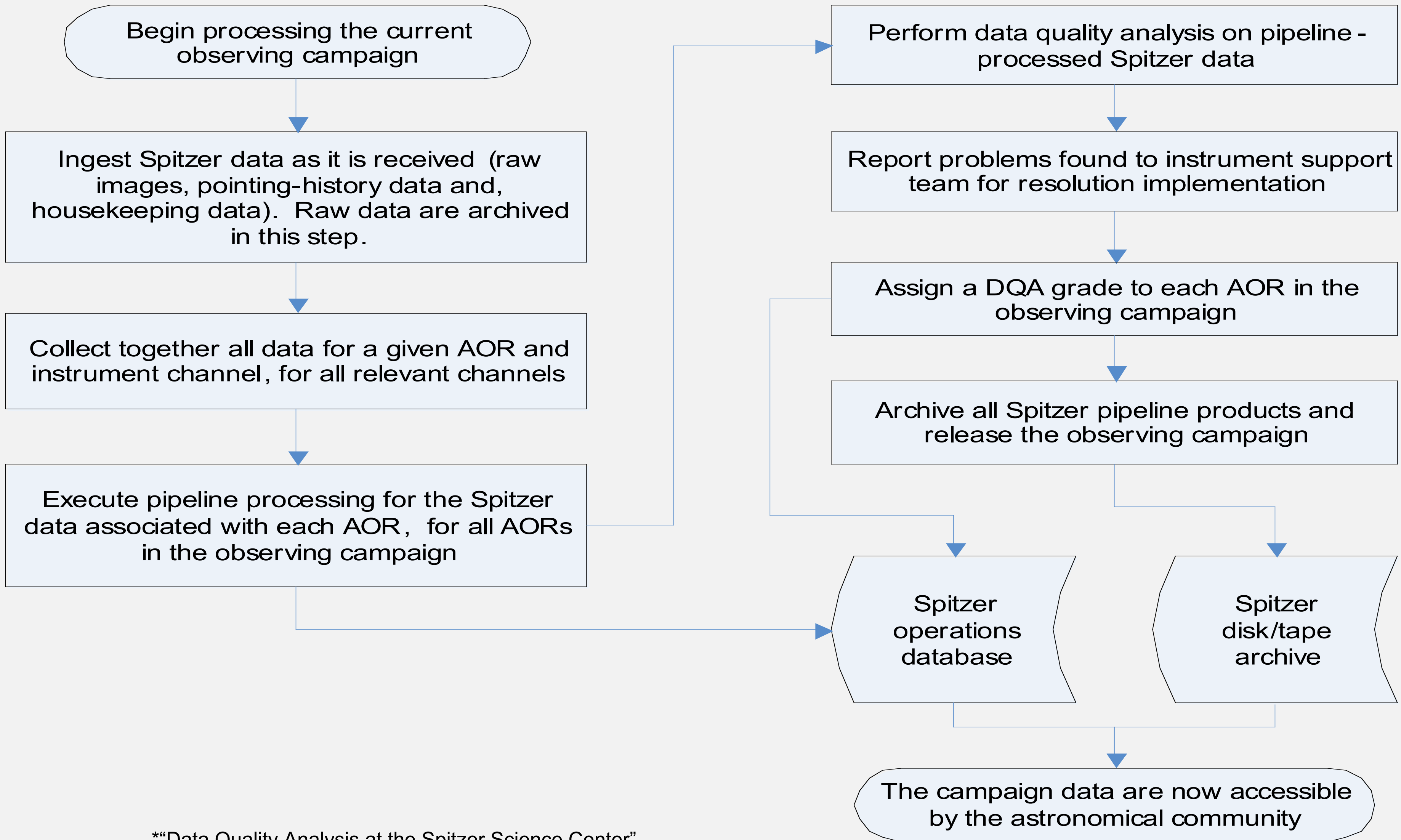
Data Transfer and PAOs*





*"Downlink Infrastructure and Pipelines", Russ Laher June, 2002

Data (re-)processing and DQA Activities*



*"Data Quality Analysis at the Spitzer Science Center", V. Mannings, et al, SPIE, 2006

Archived/Released/Public Data

- An AOR can be considered to consist of:
  - Metadata from the SODB, e.g.:
    - AOR Parameters
    - Scheduling history
    - Observer and Program Information
    - Data processing and Quality history
  - FITS and ASCII files from the file system
- An AOR has been "archived" when:
  - Metadata has been replicated to the Archive Database
    - This happens continuously as changes are made to the SODB
    - Users can track the status of their AOR from the SSC Archive or Planning Tools
  - ProductArchiver: Moves a campaign of files from the sandbox to the archive file system
    - FITS files + ASCII files (data, masks, logs, source lists, etc.)
- An AOR is "released" (and visible to its owner) when:
  - The campaign in which the AOR was taken is archived has a CERTIFIED date set in the database
  - The AOR is not embargoed
  - The final AOR STATUS has been set to a RELEASABLE type in the database
  - The RELEASED date has been set in the database
- An AOR is "public" when:
  - It has been released
  - Release date + QA Interval = Date prior to search date

"Overview of Archive Components, Design and Schedule", Lisa Storrie-Lombardi, June 20, 2003

Spitzer Science Operations System (SOS) at Launch

- **Downlink**
  - **Data ingestion**
  - **Pipeline processing**
  - **Data quality analysis**
- **Uplink**
  - **Planning/scheduling**
  - **Resource planning**
  - **Science user tools**
- **Archive**
  - **Leopard**
- **Science Data Management**
  - **SSC dbms (sodb)**
  - **Public dbms (archive -replicated sodb)**
  - **Proposal dbms (staging, proposals)**
  - **File management system**
- **Support and Operations**
  - **Science user support**
  - **Community support**
  - **Educational Outreach**
  - **Public Affairs**

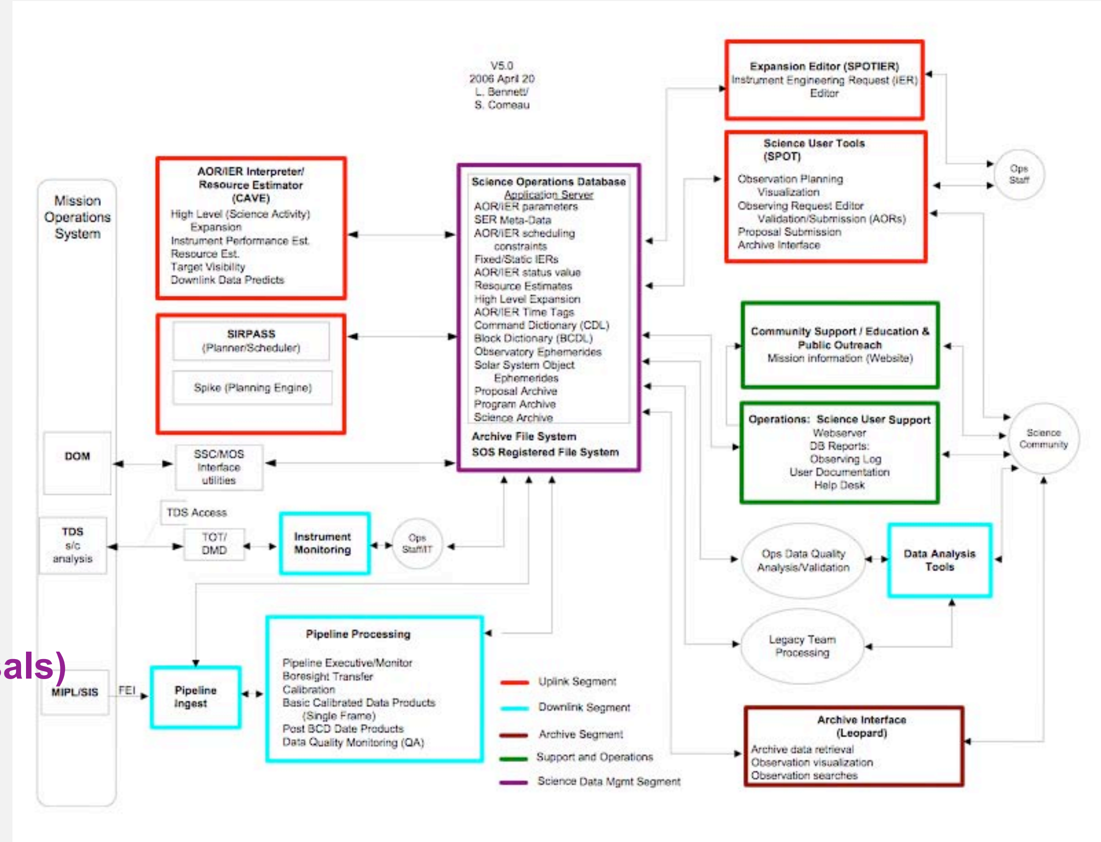
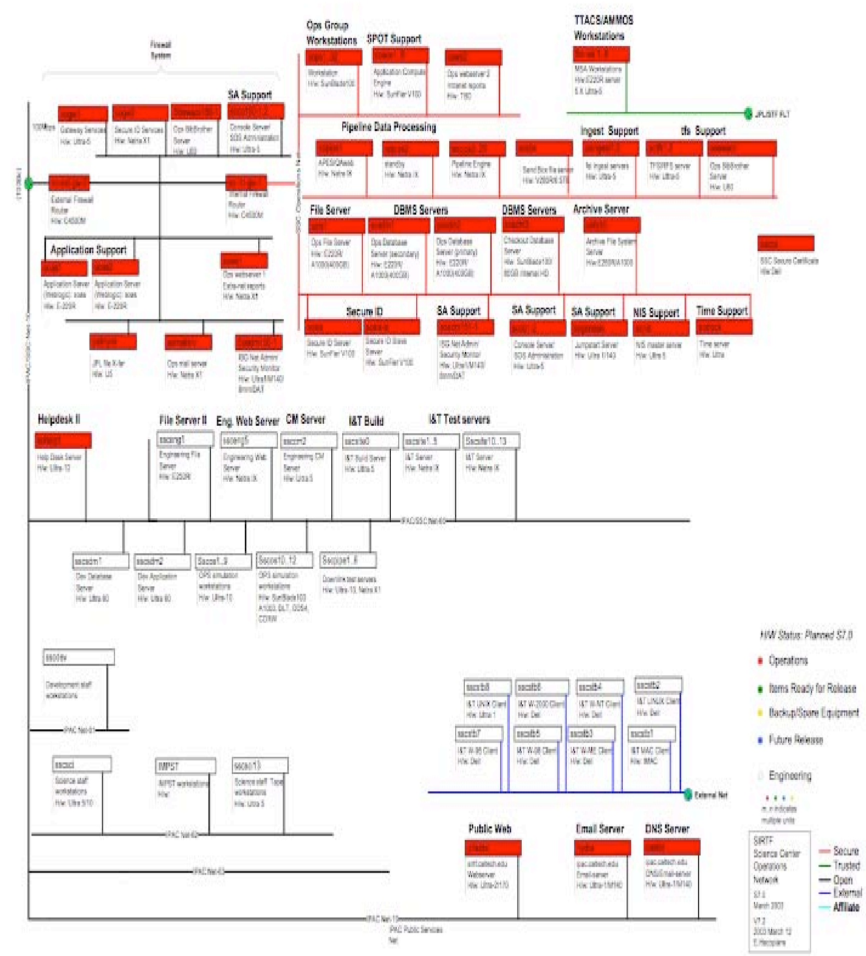


Diagram from "The Spitzer Science Center System Description and Lessons learned Due to Two Years of Operations", Lee Bennett, et al, SPIE, 2006

Launch SSC Operations System

- Core comprised of
  - common/interchangeable components
  - Single vendor Sun with Platinum maintenance for launch and IOC.
- Servers - Sun E220R, E250R, & E450
- Storage - Sun A1000 & T3
- Processing - Sun Netra
- Workstations - Sunblade 100
- Archival storage - HSM
  - E450, L700
  - SAM-FS
- 100 Mbs Network



Changes in the first 1100 days (1 of 4)

- Storage
  - From direct-attached storage to network-attached storage with "2Gb fiber channel"
  - Sandboxes
    - Requirements went from 0.6TB to 48.0TB.
    - File system from UFS to QFS to support 750M to 1000M files files systems
  - On-line data - raw, other, proc0, proc1, proc2, proc3
    - Server (V240) + Nexsan SataBeast each
  - Archive
    - HSM: Sun V440, Qualstar tape library, 5 (and growing) Nexsan Satabeasts.

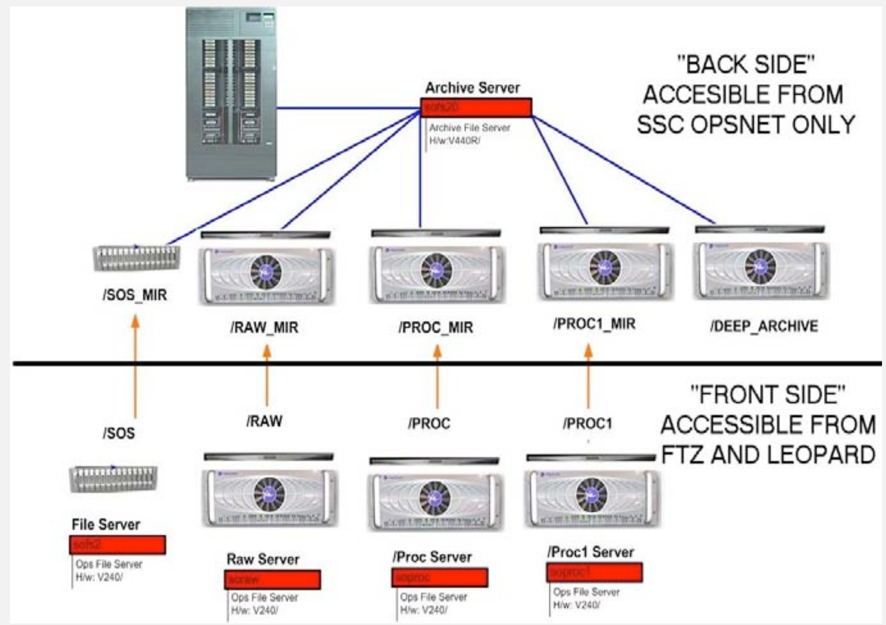


Diagram from “Deep Archive Proposal”, Eugean Hacopians, June 20, 2006

Changes in the first 1100 days (2 of 4)

- Pipeline processing
  - Initially 30 Sun Netra X1
  - Added a Sun E450
  - Replaced all with 18 SunFire V240
- Network
  - SSC operations network from 100Mb/s to1Gb/s
  - Consolidated network connections to new core switch with 720Gbps capable backplane.
    - Dedicated Cisco 6509.
    - Topology changed from numerous local switches to “direct run” to new core switch.
  - Improved packet forwarding in and out of SSCON
  - New core switch delivers application and service continuity; and redundant system components provide hardware-level resiliency.
  - Expanded network capability for storage management

Changes in the first 1100 days (3 of 4)

- DBMS
  - Informix 9.4 to Informix 10.0 to enhance replication.
  - Server: Sun E220R to Sun V880 to Sun V890
  - Storage: Sun A1000 to Sun T3 to Sun 3510.
  - Off-load obsolete dbms and file system content.
    - Reprocessing
      - Causes dbms and on-line, active file systems to grow unboundedly.
      - Older data become obsolete upon reprocessing
      - Need to retain older versions until end-of-mission.
    - Solution
      - Deploy new deepArchive subsystem
        - DBMS server
        - File system
      - Periodically move obsolete data to the deepArchive
    - Operational advantages
      - Manageable database maintenance
      - Improve database backup and recoveries
      - Faster database snapshots
      - Faster SODB to ARCHIVE replication and synchronization validation

Changes in the first 1100 days (4 of 4)

- Public archive interface (leopard) developed post-launch.
  - Client application (Leopard + Subscriber) : client software, query and manage downloads
  - Web interface: query and manage downloads
  - Middleware Storage
    - Weblogic Middleware lives on 2 servers (SOAS1, SOAS2)
    - 1.5 TB for each SOAS server
    - Marshalling, packaging, staging before sending to users
  - FTP Server
    - Public FTP server is used for staging public data
    - Popular Products: Web site where popular data is staged for download (http://data.spitzer.caltech.edu/popular/)

What are today's issues?

- Extended mission and an additional technology refresh.
- Change dbms or not: Oracle dbms from Informix?
- SAM-FS/QFS to Solaris 10 ZFS file system conversion.
- Explosive archive growth
- Managing file systems where the number of files is 3B-5B.
- Managing currency of data products
  - dbms (sodb/archive) and file systems
  - deepArchive.
- Storage network from 2Gb to 4Gb
- Support 10Gb network throughput
- Reprocessing and publication of final standard products to the community via IPAC's Infrared Science Archive (IRSA).

What did we get really…
+ Right? - Wrong?

+ System kept operational in spite of growth of data by more than two orders of magnitude from original estimates.
+ Hardware and software partitioning and scalability enable focused improvements, changes and upgrades.
+ As built pipelines - processing times (from end-of-campaign to release of data)
  + IRAC ~8-10 days
  + MIPS ~ 9-11 days
  + IRS ~ > 16 day days
  + Include calibration file selection, reprocessing and archiving.

- Inaccurate data products' definition and forecast.
  - Unplanned dbms and file system growth (SSC explicitly chose to record all relevant information,e.g.,everything)
- Limited I&T resources (environment and staffing)
  - Risk-based (re-)testing

What would we do differently next time?

- Increase the formality of the engineering lifecycle and its processes.
  - Improve accuracy and fidelity for intermediate- and final- data products.
  - Improve knowledge of as-built through better configuration management.
- An all-disk solution for on-line storage, relegating SAM-FS/QFS to archive only.
- Service level agreements with Spitzer project and instrument teams.
- Improved network architecture
  - Perimeter network for DMZ services
  - Improved internal network providing services to the operations “dark network”