



JSOC Status Review



Purpose of Review



- **To confirm final hardware configuration prior to committing funds**
- **To examine the status of the software development tasks**
- **To review plans for completion**
- **Goal is to assess status and determine if there are any gaps, issues, etc. to being ready to support science analysis by February 2009.**
- **This is a forward looking status assessment so little time will be spent reviewing completed items.**
- **The review is Stanford-Centric, i.e. mostly dealing with the activities at Stanford**
- **Why now?**
 - Hardware purchases soon
 - Only 16 months until science data begins



Overview of HMI–AIA Joint Science Operations Center (JSOC) Science Data Processing (SDP)

Overview will be in reverse order than the detailed discussion that follows.

- Science Goals
- Science Products
- Software Infrastructure
- Hardware



HMI & AIA JSOC Scope



- **The HMI/AIA Joint SOC consists of two functional parts:**
 - Science Data Processing (**SDP**) – at Stanford and LMSAL
 - Instrument Operations Center (**IOC**) – at LMSAL
- **JSOC IOC includes:**
 - HMI and AIA Commanding and Health Monitoring
 - HMI and AIA Engineering support as needed
- **JSOC SDP includes:**
 - HMI and AIA Science Telemetry Data capture (from DDS) and archive
 - HMI and AIA Level-0 processing and archive
 - HMI processing through to level-2 with archiving of end products
 - AIA processing through level-1 with online archive at *Stanford*
 - AIA level-2 processing at *LMSAL*
 - Data export of the above and other HMI and AIA products as needed
- **JSOC does not include tasks such as:**
 - Science analysis beyond level-2 products
 - HMI and AIA EPO
 - HMI & AIA Co-I science support



JSOC – SDP Stanford Personnel



JSOC SDP Team

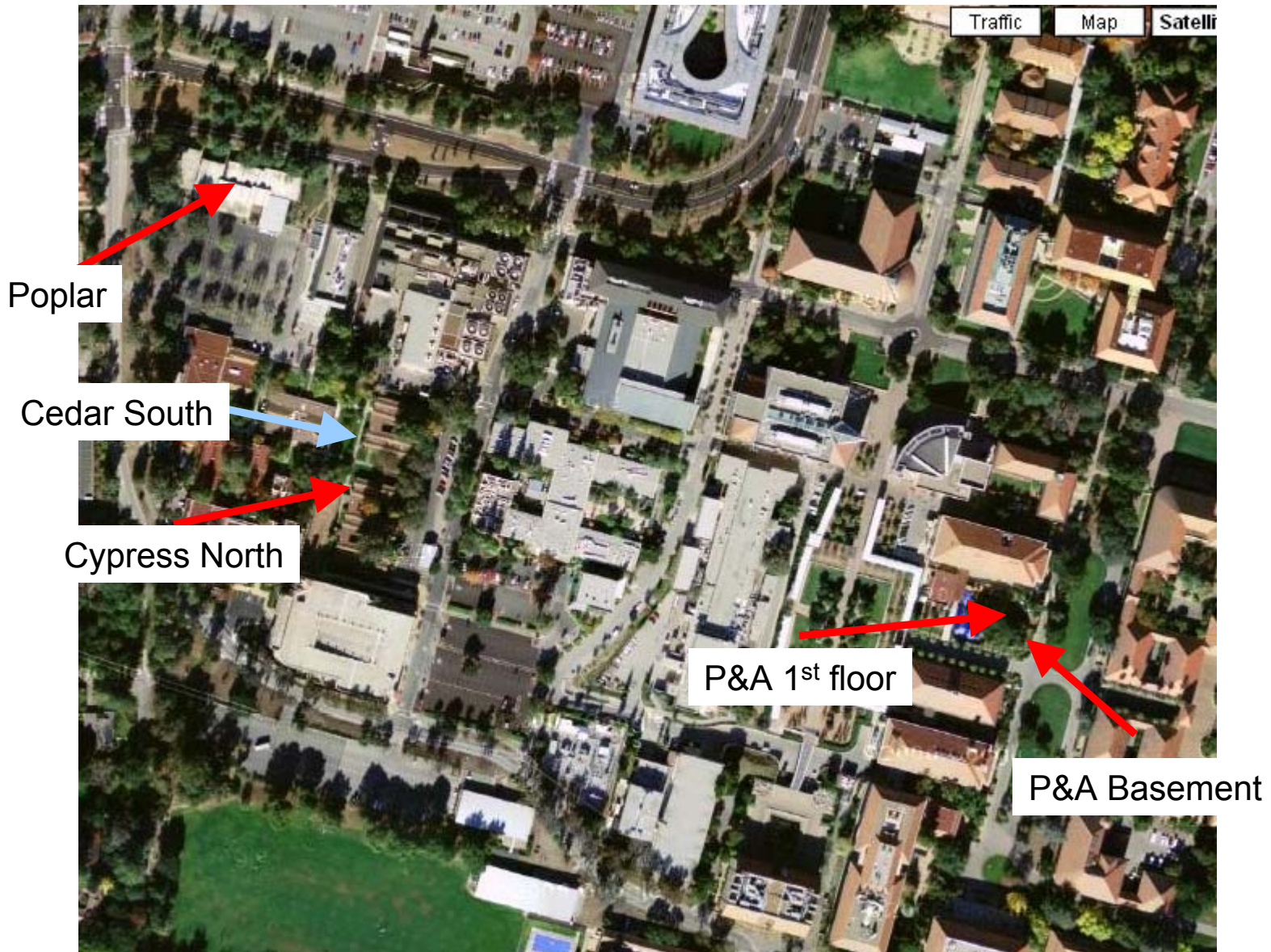
- Phil Scherrer, Acting SDP Lead
- SDP Software
 - Jim Aloise, SUMS, PUI
 - Karen Tian, DRMS, Export
 - Art Amezcua, CM, Level-2 Science Products
 - Jennifer Spencer, Database
 - Carl Cimilluca, Sys Engineer
- SDP Hardware
 - Keh-Cheng Chu, Lead
 - Brian Roberts, Sys Admin
- Data Operations
 - Jeneen Sommers
 - Hao Thai

Science Data Processing Teams

- Level-0 Team, Rock Bush, Lead
 - Jim Aloise
 - Keh-Cheng Chu
 - Jennifer Spencer
 - Carl Cimilluca
- Level-1 Team, Jesper Schou, Lead
 - Sebastien Couvidat
 - Cristina Rabello-Soares
 - Richard Wachter
 - Yang Liu
 - Steve Tomczyk, (HAO group lead)
- Level-2 Science Leads
 - Rick Bogart
 - Tom Duvall (GSFC)
 - J. Todd Hoeksema
 - Sasha Kosovichev
 - Jesper Schou



JSOC – SDP Locations at Stanford





HMI Primary Science Objectives

1. Convection-zone dynamics and solar dynamo
 - **Structure and dynamics of the tachocline**
 - **Variations in differential rotation.**
 - **Evolution of meridional circulation.**
 - **Dynamics in the near-surface shear layer.**
2. Origin and evolution of sunspots, active regions and complexes of activity
 - **Formation and deep structure of magnetic complexes.**
 - **Active region source and evolution.**
 - **Magnetic flux concentration in sunspots.**
 - **Sources and mechanisms of solar irradiance variations.**
3. Sources and drivers of solar activity and disturbances
 - **Origin and dynamics of magnetic sheared structures and delta-type sunspots.**
 - **Magnetic configuration and mechanisms of solar flares and CME.**
 - **Emergence of magnetic flux and solar transient events.**
 - **Evolution of small-scale structures and magnetic carpet.**
4. Links between the internal processes and dynamics of the corona and heliosphere
 - **Complexity and energetics of solar corona.**
 - **Large-scale coronal field estimates.**
 - **Coronal magnetic structure and solar wind**
5. Precursors of solar disturbances for space-weather forecasts
 - **Far-side imaging and activity index.**
 - **Predicting emergence of active regions by helioseismic imaging.**
 - **Determination of magnetic cloud Bs events.**



HMI Science Analysis Plan

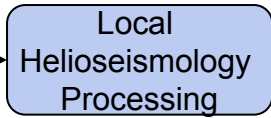
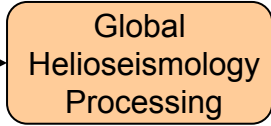
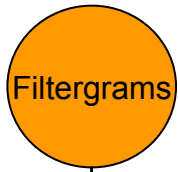


HMI Data

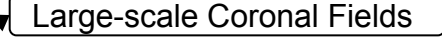
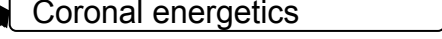
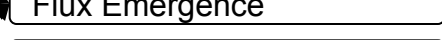
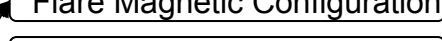
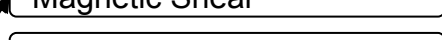
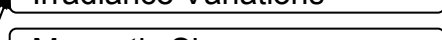
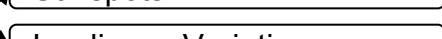
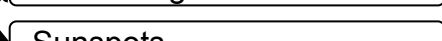
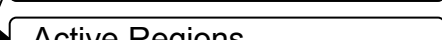
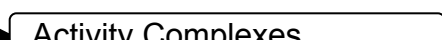
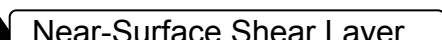
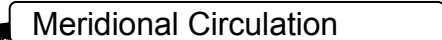
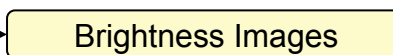
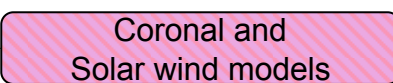
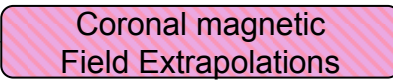
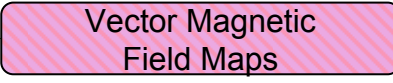
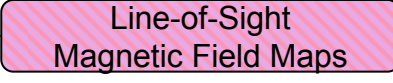
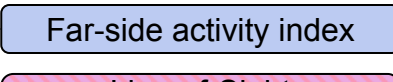
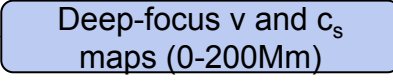
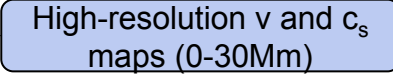
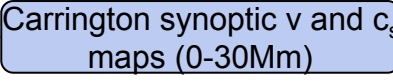
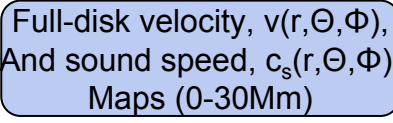
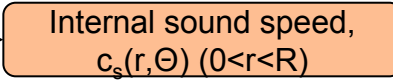
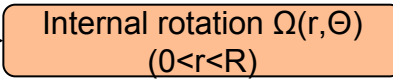
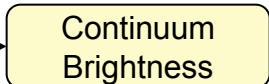
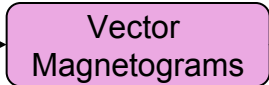
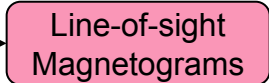
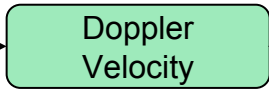
Processing

Data Product

Science Objective



Observables



Version 1.0w



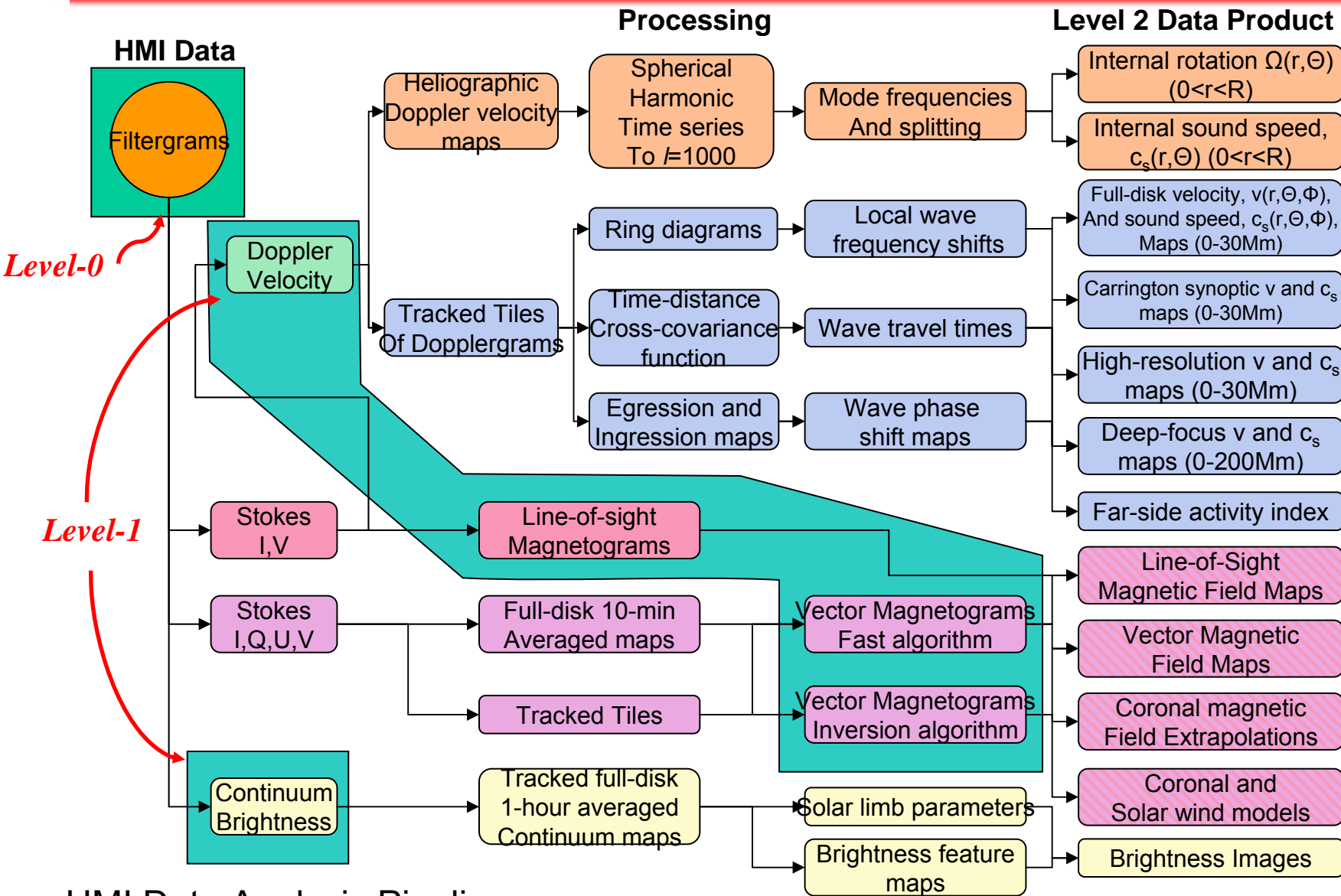
JSOC Processing “Levels”



- **Tlm is raw telemetry files as received from SDOGS**
- **Level-0 is images extracted from tlm with added meta-data, no change to pixels**
- **Level-1 is cleaned up and calibrated into physical units in standardized form**
- **Level-2 is science data products**
- **Level-3 is higher level products or user produced products and are not JSOC products but may be archived and distributed as desired by owner**



JSOC - HMI Pipeline



HMI Data Analysis Pipeline



Analysis modules: Co-I contributions and collaboration

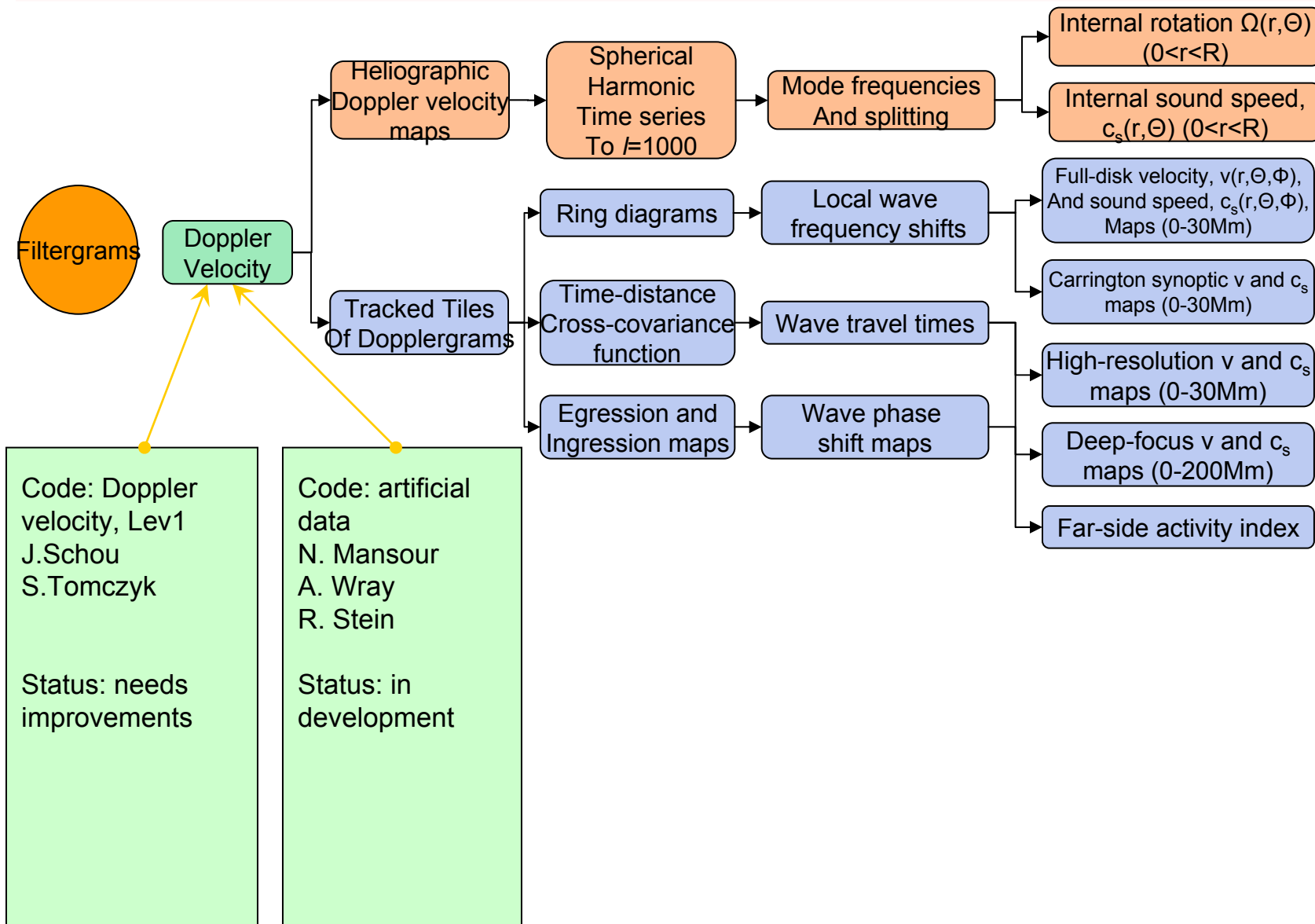


- **Contributions from co-I teams:**
 - Software for intermediate and high level analysis modules
 - Output data series definition
 - Keywords, links, data segments, size of storage units etc.
 - Documentation (detailed enough to understand the contributed code)
 - Test data and intended results for verification
 - Time
 - Explain algorithms and implementation
 - Help with verification
 - Collaborate on improvements if required (e.g. performance or maintainability)

- **Contributions from HMI team:**
 - Pipeline execution environment
 - Software & hardware resources (Development environment, libraries, tools)
 - Time
 - Help with defining data series
 - Help with porting code to JSOC API
 - If needed, collaborate on algorithmic improvements, tuning for JSOC hardware, parallelization
 - Verification

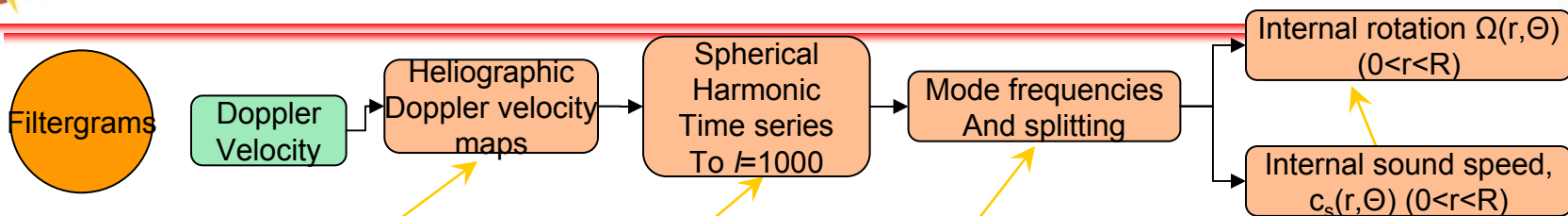


Example - Helioseismology





Example - Global helioseismology



Code: project
J. Schou

Status: ready to port

Code: qdotprod
J.Schou

Status: ready to port

Code: med-l peak bagging
J.Schou
High-l ridge fitting,
E. Rhodes

Status: needs improvements

Code: sound-speed inversions
A.Kosovichev

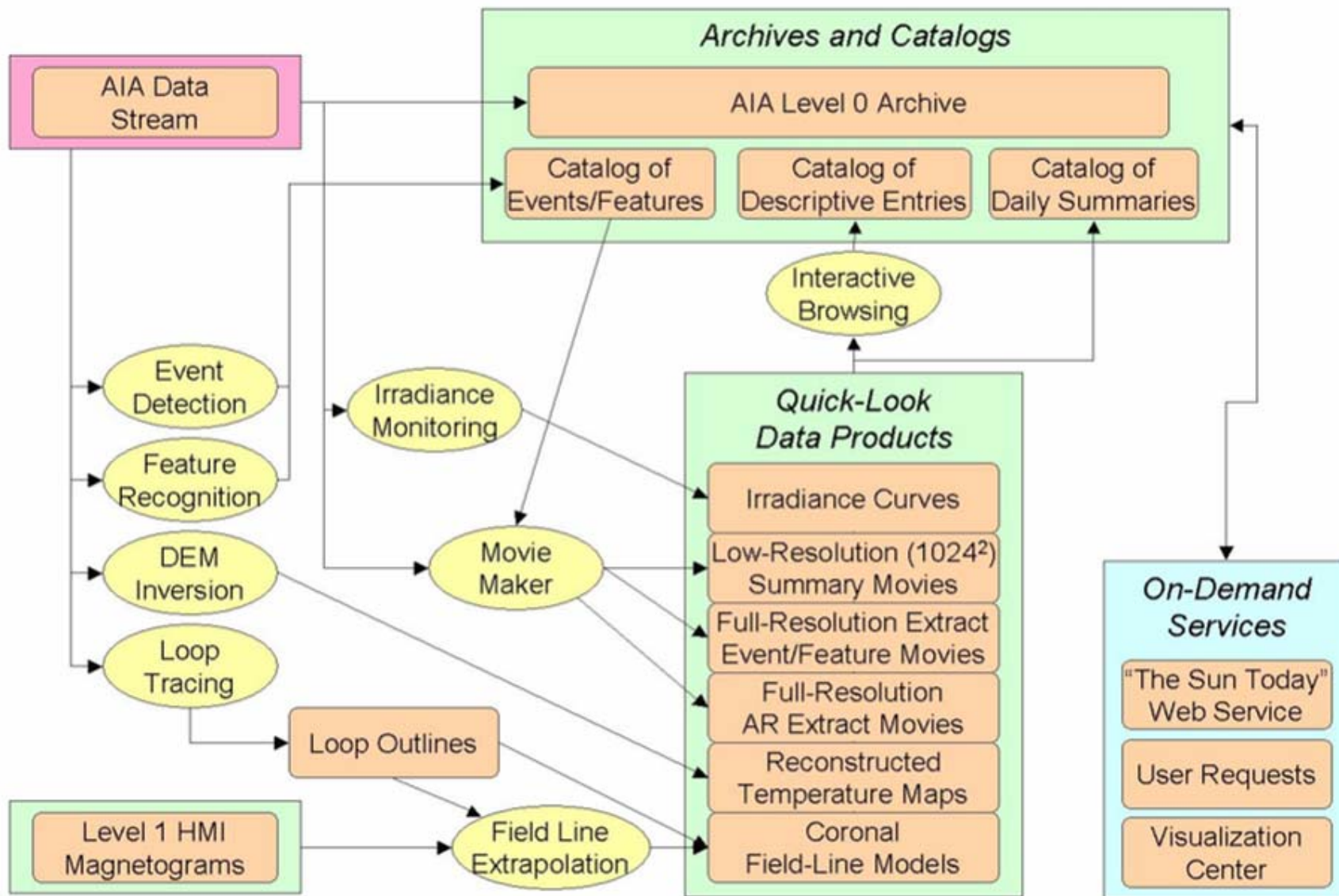
Status: ready to port

Code: rotation inversion
J.Schou
R. Howe

Status: ready to port



AIA Level-2





Level 1 HMI and AIA Basic Observable Quantities



- **HMI and AIA level-1 “levels” are similar but the details differ (a lot).**
- **HMI – from filtergrams to physical quantities**
 - 1.0 Flat field applied to enable limb fit and registration
 - 1.5 Final product types
 - 1.5q – Quicklook available in ~10 minutes, saved ~10 days
 - 1.5p – Provisional mix of 1.5q and 1.5 final
 - 1.5 – Final best possible product
 - Products
 - Continuum Intensity
 - Doppler Velocity
 - Line of Sight Field
 - Vector Field
- **AIA – Filtergrams are basic product**
 - Quicklook and Final both produced
 - Planning movies from quicklook
 - Full details in development (after AIA delivery)



HMI and AIA Level-0



- **Level-0 Processing is the same for HMI and AIA**

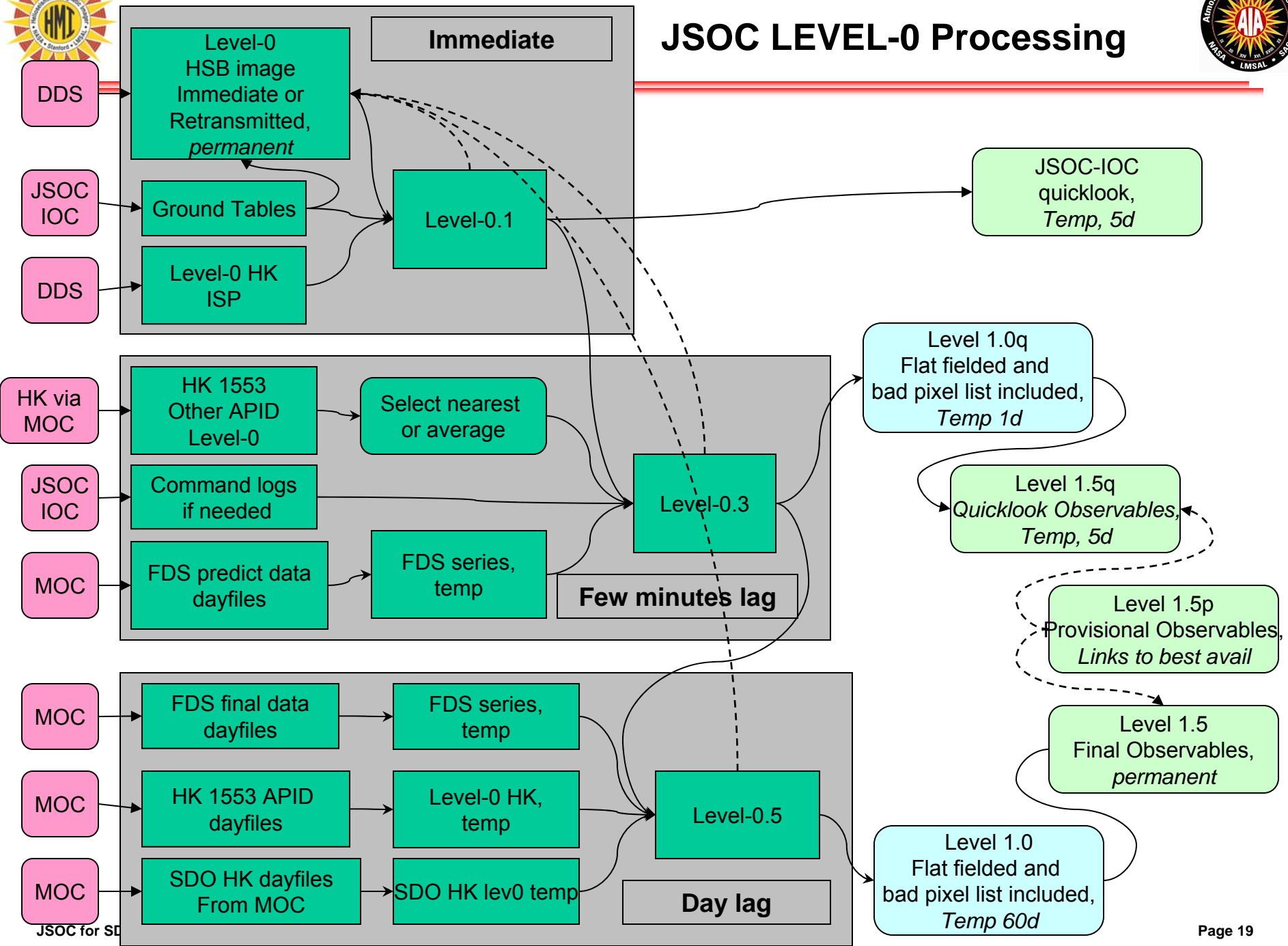
- **Level 0.1 – Immediate – Used for Ops quicklook**
 - Reformat images
 - Extract Image Header meta-data
 - Add “Image Status Packet” high-rate HK Packet (per image)
 - Export for JSOC IOC Quicklook

- **Level 0.3 – Few minute lag - Used for quicklook science data products**
 - Add other RT HK meta-data
 - Add FDS S/C info

- **Level 0.5 – Day or more lag – Used for final science data products**
 - Update FDS data
 - Add SDO roll info
 - Includes final images



JSOC LEVEL-0 Processing





- **ALL HMI and AIA data will be available for export at level-0 through standard products (level-1 for both and level-2 for HMI)**
- **It would be unwise to expect to export all of the data. It is simply not a reasonable thing to expect and would be a waste of resources.**
- **Our goal is to make all useful data easily accessible.**
- **This means “we” must develop browse and search tools to help generate efficient data export requests.**
- **Quicklook Products**
 - Quicklook to JSOC IOC
 - Quicklook Basic Products to Space Weather Users
- **Prime Science Users**
 - JSOC will support Virtual Solar Observatory (VSO) access
 - JSOC will also have a direct web access
 - There will be remote DRMS/SUMS systems at key Co-I institutions
 - JSOC In Situ Delivery
 - Special Processing at SDP as needed and practical
- **Public Access**
 - Web access for all data
 - E/PO
 - Solar Events



Configuration Management & Control



- **Capture System**
 - Managed by JSOC-SDP CCB after August freeze
 - Controlled in CVS
- **SUMS, DRMS, PUI, etc. Infrastructure**
 - Managed by JSOC-SDP CCB after launch
 - Controlled in CVS
- **PUI Processing Tables**
 - Managed by HMI and/or AIA Instrument Scientist
 - Controlled in CVS
- **Level 0,1 Pipeline Modules**
 - Managed by HMI and/or AIA Instrument Scientist
 - Controlled in CVS
- **Science Analysis Pipeline Modules**
 - Managed by program author
 - Controlled in CVS



JSOC DCS Science Telemetry Data Archive



- **Telemetry data is archived twice**
- **The Data Capture System (DCS) archives tlm files for offsite storage**
- **Archive tapes are shipped to the offsite location and verified for reading**
- **The Data Capture System copies tlm files to the Pipeline Processing System**
- **The Pipeline Processing System generates Level-0 images and archives both tlm and Level-0 data to tape**
- **Only when the DCS has received positive acks on both tlm archive copies does it inform the DDS, which is now free to remove the file from its tracking logic**



JSOC DRMS/SUMS Basic Concepts



- Each “image” is stored as a record in a data “series”.
- There will be many series: e.g. hmi_ground.lev0 is ground test data
- The image metadata is stored in a relational database – our Data Record Management System (DRMS)
- The image data is stored in SUMS (Storage Unit Management System) which itself has database tables to manage its millions of files.
- SUMS owns the disk and tape resources.
- Users interact with DRMS via a programming language, e.g. C, FORTRAN, IDL.
- The “name” of a dataset is actually a query in a simplified DRMS naming language that also allows general SQL clauses.
- Users are encouraged to use DRMS for efficient use of system resources
- Data may be exported from DRMS as FITS or other protocols for remote users.
- Several Remote DRMS (RDRMS) sites will be established which will “subscribe” to series of their choice. They will maintain RSUMS containing their local series and cached JSOC series.
- The JSOC may act as an RDRMS to access products made at remote sites.



JSOC data organization



- **Evolved from FITS-based MDI dataset concept to**
 - Fix known limitations/problems
 - Accommodate more complex data models required by higher-level processing
- **Main design features**
 - Lesson learned from MDI: Separate meta-data (keywords) and image data
 - No need to re-write large image files when only keywords change (lev1.8 problem)
 - No (fewer) out-of-date keyword values in FITS headers
 - Can bind to most recent values on export
 - Easy data access through query-like dataset names
 - All access in terms of sets of data records, which are the “atomic units” of a data series
 - A *dataset name* is a query specifying a set of data records (possibly from multiple data series):
 - Storage and tape management must be transparent to user
 - Chunking of data records into “storage units” and tape files done internally
 - Completely separate storage and catalog (i.e. series & record) databases: more modular design
 - Legacy MDI modules should run on top of new storage service
 - Store meta-data (keywords) in relational database (PostgreSQL)
 - Can use power of relational database to rapidly find data records
 - Easy and fast to create time series of any keyword value (for trending etc.)
 - Consequence: Data records for a given series must be well defined (i.e. have a fixed set of keywords)



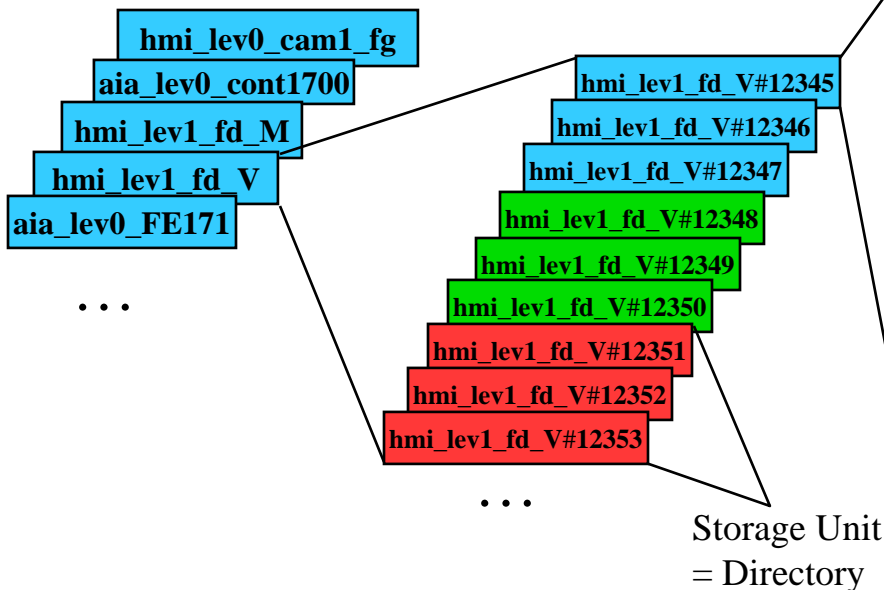
Logical Data Organization



JSOC Data Series

Data records for series hmi_fd_V

Single hmi_fd_V data record



Keywords:
 RECORDNUM = 12345 # Unique serial number
 SERIESNUM = 5531704 # Slots since epoch.
 T_OBS = '2009.01.05_23:22:40_TAI'
 DATAMIN = -2.537730543544E+03
 DATAMAX = 1.935749511719E+03
 ...
 P_ANGLE = LINK:ORBIT,KEYWORD:SOLAR_P
 ...

Links:
 ORBIT = hmi_lev0_orbit, SERIESNUM = 221268160
 CALTABLE = hmi_lev0_dopcal, RECORDNUM = 7
 L1 = hmi_lev0_cam1_fg, RECORDNUM = 42345232
 R1 = hmi_lev0_cam1_fg, RECORDNUM = 42345233
 ...

Data Segments:

Velocity =



JSOC Series Definition (JSD)



```

#===== Global series information =====
Seriesname:      "hmi_fd_v"
Description:     "HMI full-disk Doppler velocity. ..."
Author:         "Rasmus Munk Larsen"
Owners:         "production"
Unitsize:       90
Archive:        1
Retention:      40000
Tapegroup:      2
Primary Index:  T_Obs

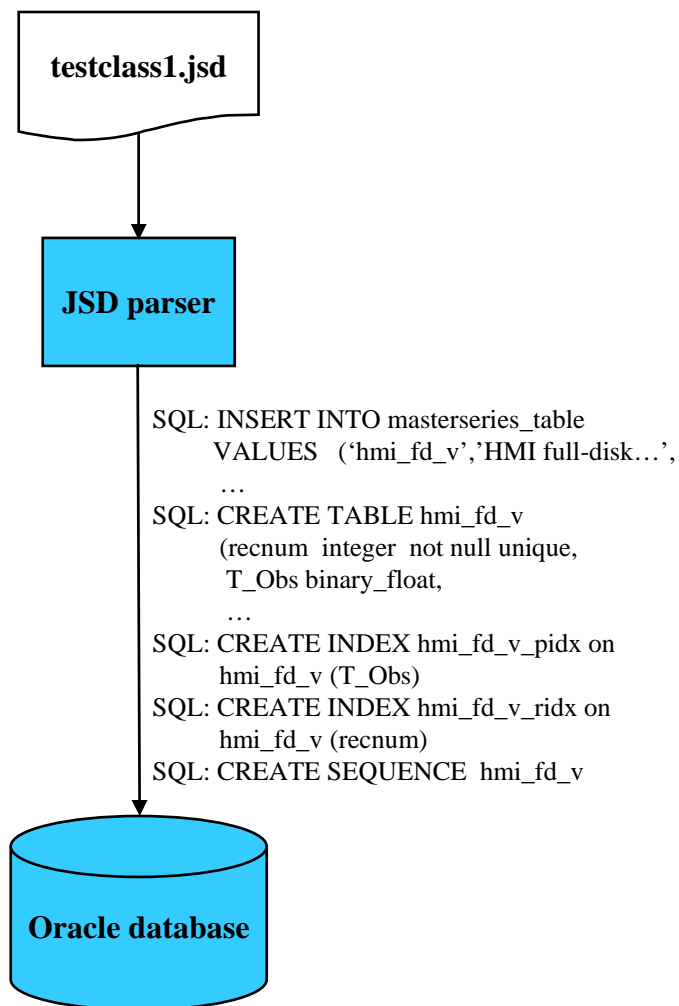
#===== Keywords =====
# Format:
# Keyword: <name>, <type>, <default value>, <format>, <unit>, <comment>
# or
# Keyword: <name>, link, <linkname>, <target keyword name>
#
Keyword: "T_Obs", time, "1970.01.01_00:00:00_TAI", "%F %T", "s",
        "Nominal observation time"
Keyword: "D_Mean", double, 0.0, "%lf", "m/s", "Data mean"
Keyword: "D_Max", double, 0.0, "%lf", "m/s", "Data maximum"
Keyword: "D_Min", double, 0.0, "%lf", "m/s", "Data minimum"
Keyword: ...
Keyword: "P_Angle", link, "Attitude", "P_Angle"

#===== Links =====
# Format:
# Link: <name>, <target series>, { static | dynamic }
#
Link: "L1", "hmi_lev0_fg", static
Link: "R1", "hmi_lev0_fg", static
Link: "L2", "hmi_lev0_fg", static
Link: "R2", "hmi_lev0_fg", static
Link: ...
Link: "Caltable", "hmi_dopcal", static
Link: "Attitude", "sdo_fds", dynamic

#===== Data segments =====
# Data: <name>, <type>, <naxis>, <axis dims>, <unit>, <protocol>
#
Data: "velocity", float, 2, 4096, 4096, "m/s", fitz

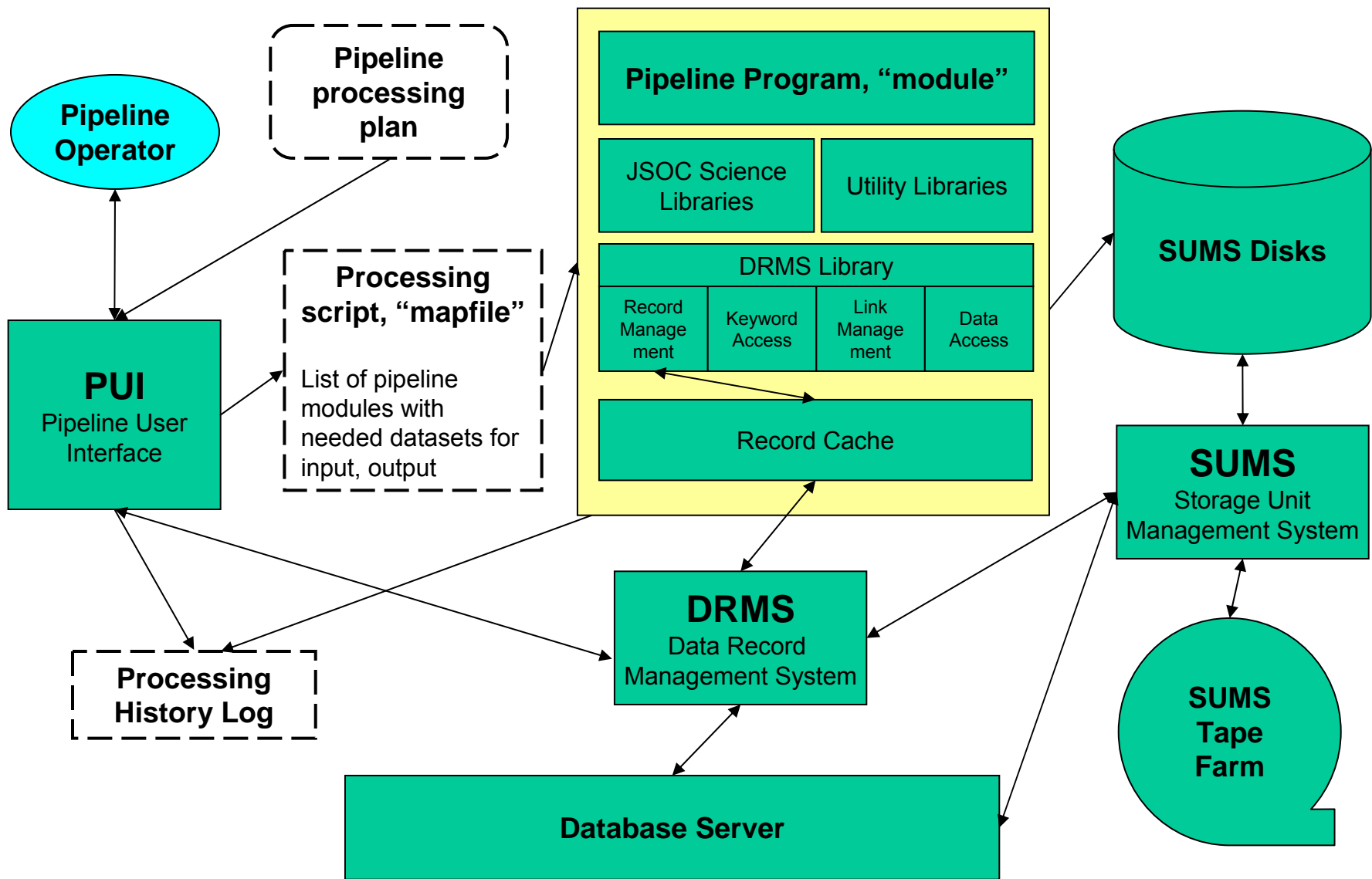
```

Creating a new Data Series:





JSOC Pipeline Processing System Components

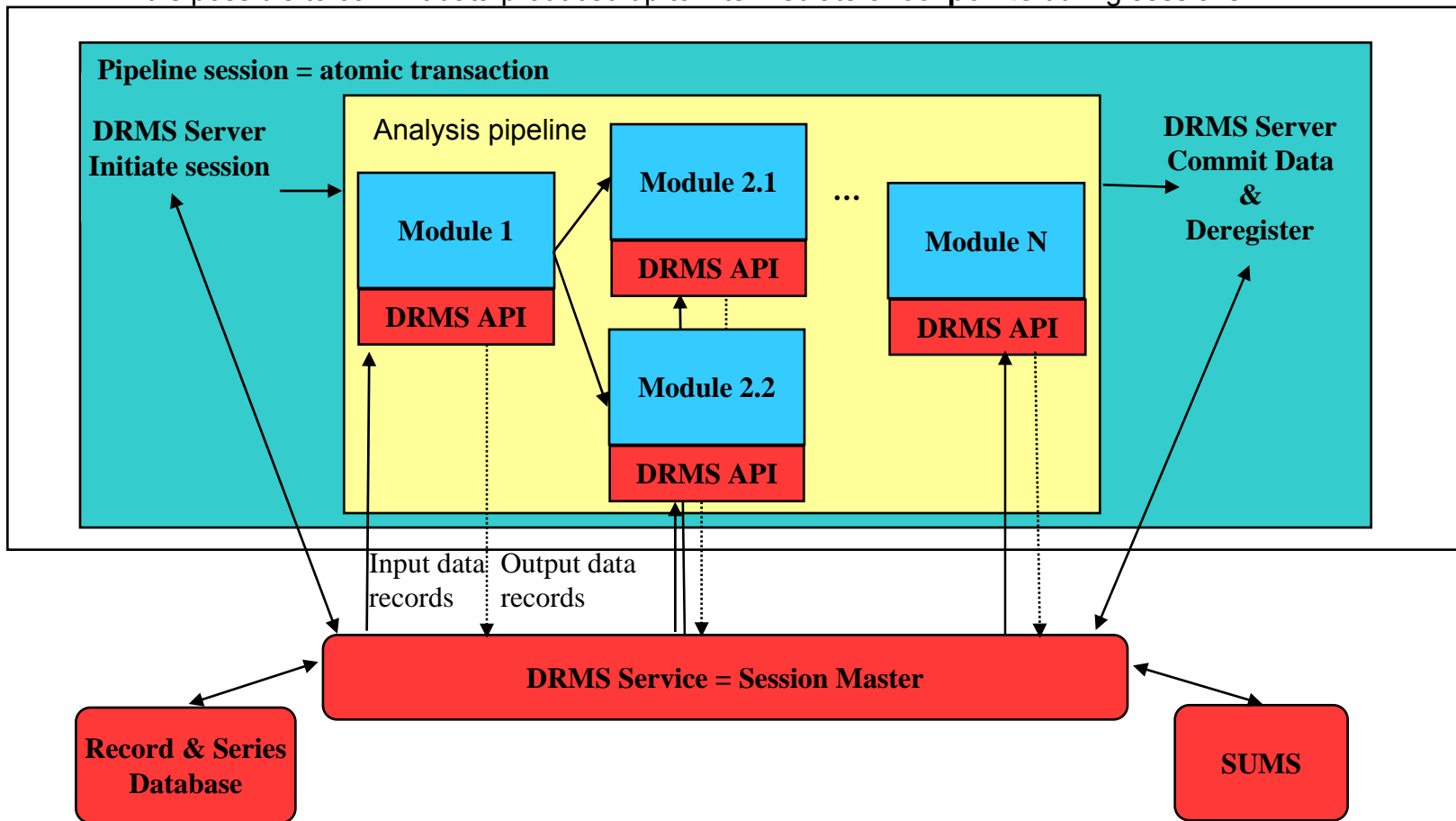




Pipeline batch processing



- A pipeline “session” is encapsulated in a single database transaction:
 - If no module fails all data records are **committed** and become visible to other clients of the JSOC catalog at the end of the session
 - If failure occurs all data records are deleted and the database **rolled back**
 - It is possible to commit data produced up to intermediate **checkpoints** during sessions





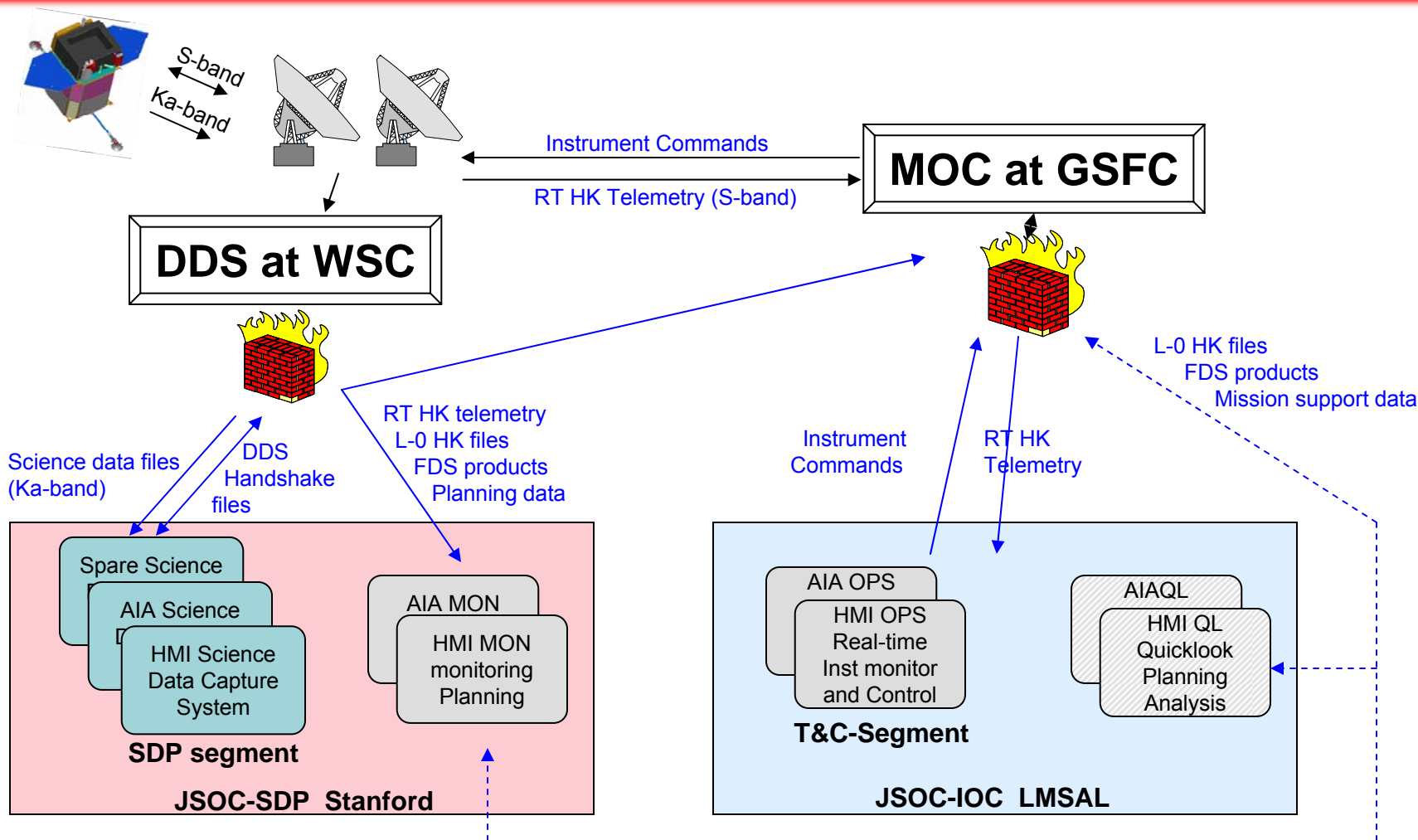
Software Development Status



- **DCS System complete, needs testing with DDS, 98% complete**
- **SUMS fully operational, needs tuning, 95% complete**
- **DRMS fully operational, some features not yet implemented, 90% complete**
- **PUI development expected to start shortly**
- **Level-0 Image Code being updated to reflect final flight software, 80% complete**
- **Level-0 HK, FDS and other metadata merge in development, 60% complete**
- **Level-1 (science observables) code will be started after instrument ship**
- **HMI Level-2 (science data products) work in development**
 - Local Helioseismology: Rings at risk due to funding, Time-Distance 40%, Holography 20%
 - Global Helioseismology: low-*l* will be MDI port to DRMS, work beginning, high-*l* needs funds
 - Magnetic Field standard products will be MDI ports, work beginning
 - Vector Field work at risk due to funding and loss of key Co-I
- **AIA Level-2**
 - Event catalog studies underway
 - DEM work beginning, will accelerate after AIA delivery
 - Merge with HMI fields under study, work accelerating
- **Summary: About where expected for L-14 and Phase-E–16 months**



JSOC Interfaces with SDO Ground System



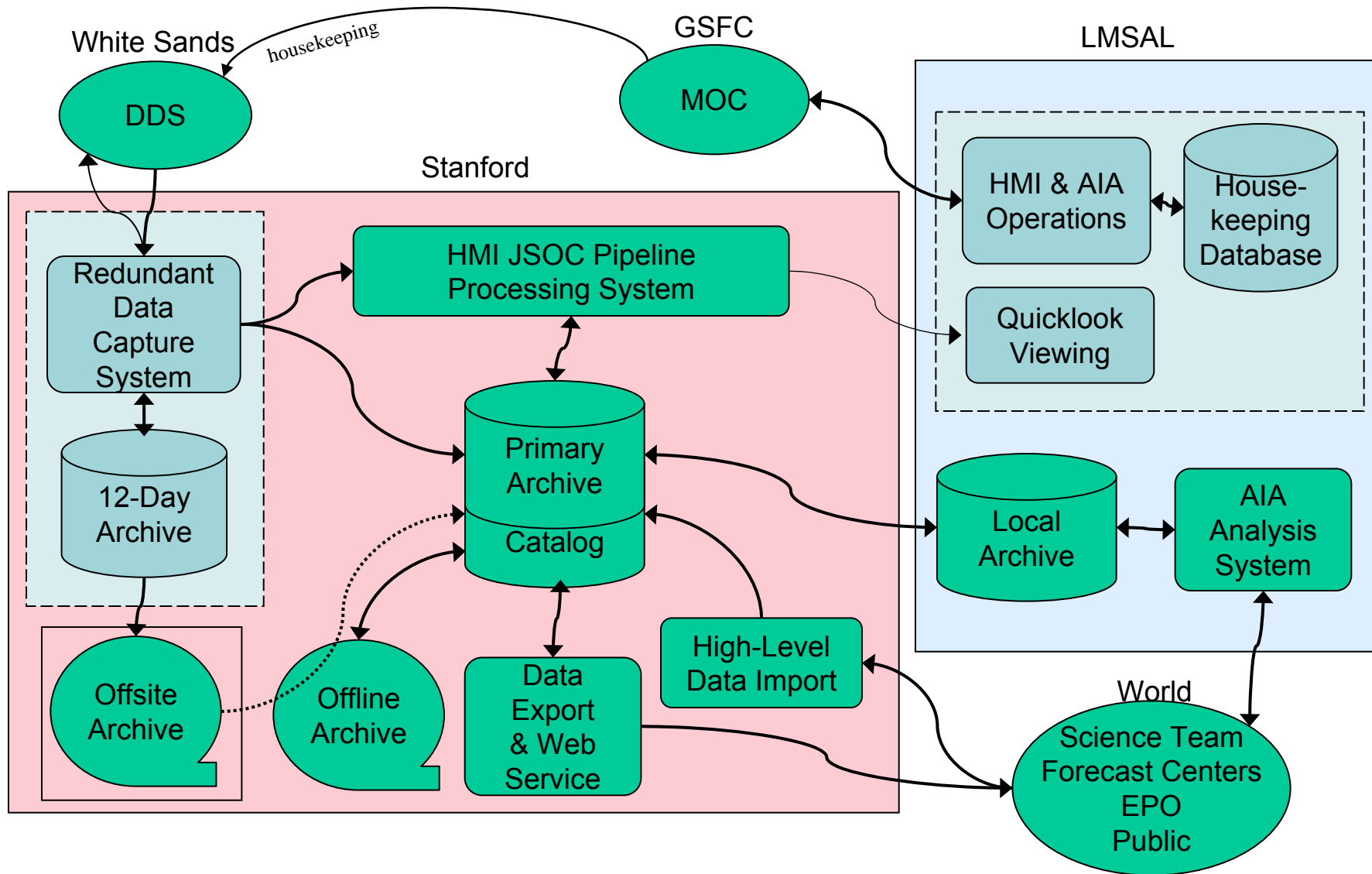
JSOC-SDP Primary responsibilities:
 Capture, archive and process science data

Additional: Instrument H&S monitoring

JSOC-IOC Primary responsibilities:
 Monitor instruments health and safety in real-time, 24/7
 Control instrument operations and generate commands
 Support science planning functions

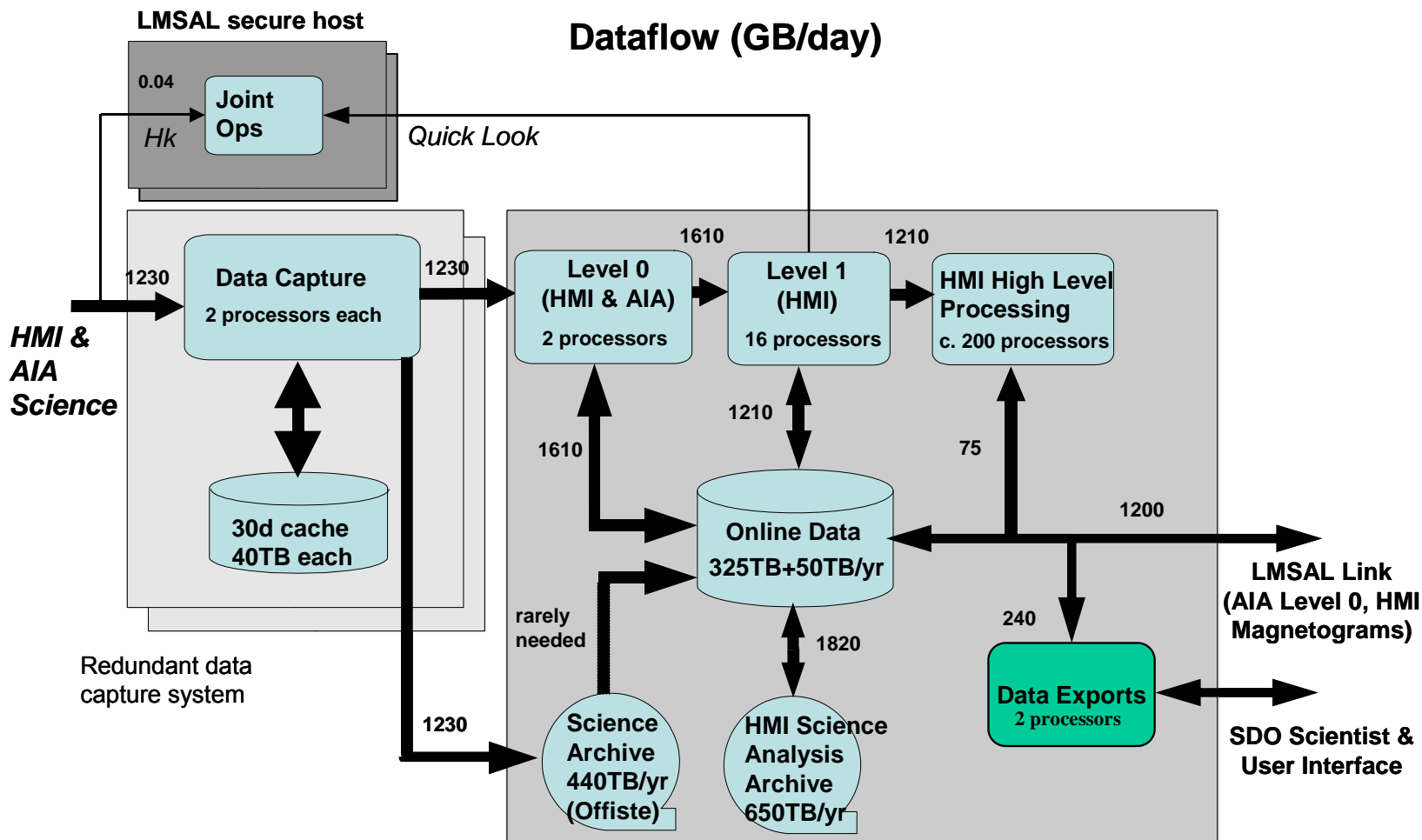


HMI & AIA JSOC Architecture



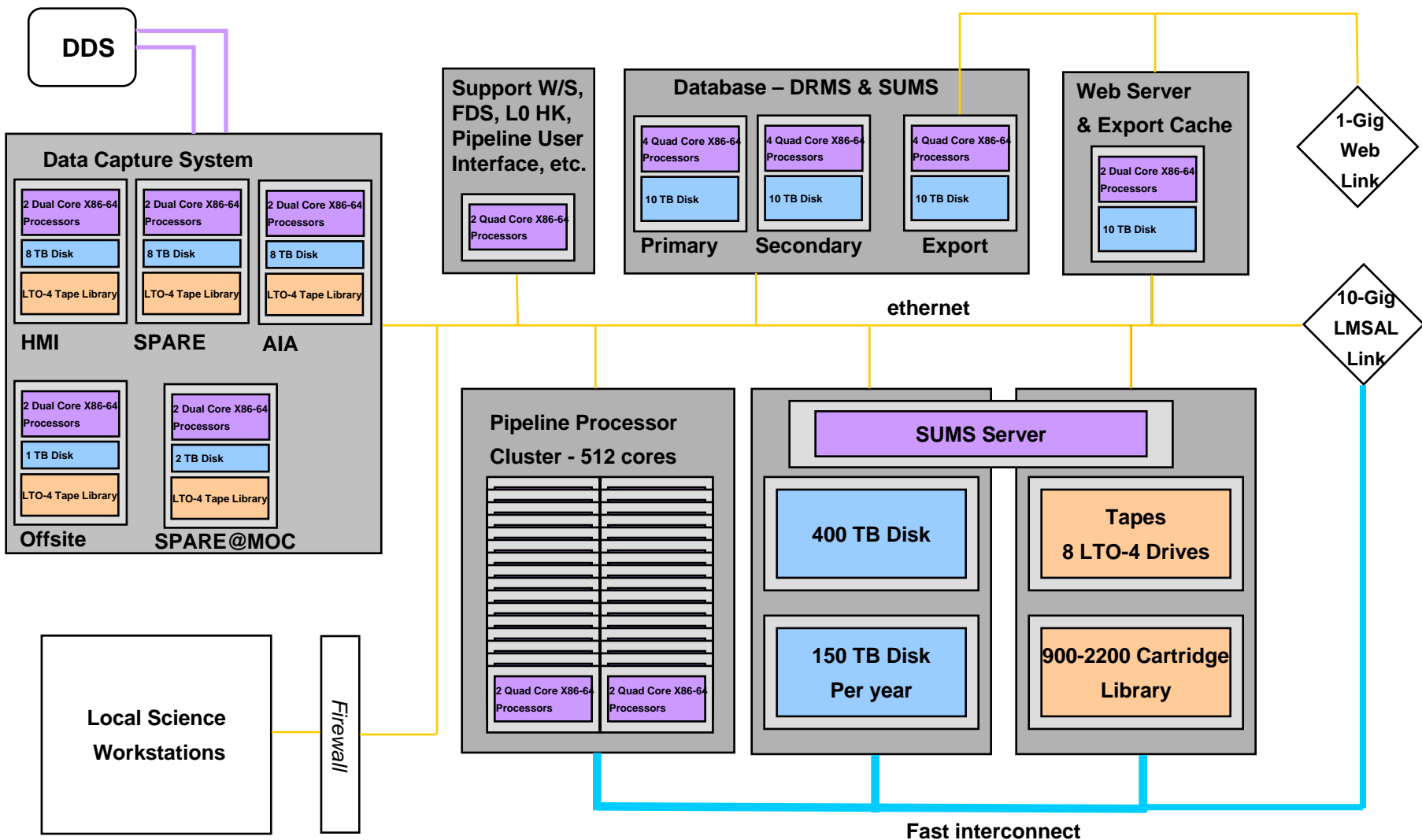


JSOC Dataflow Rates



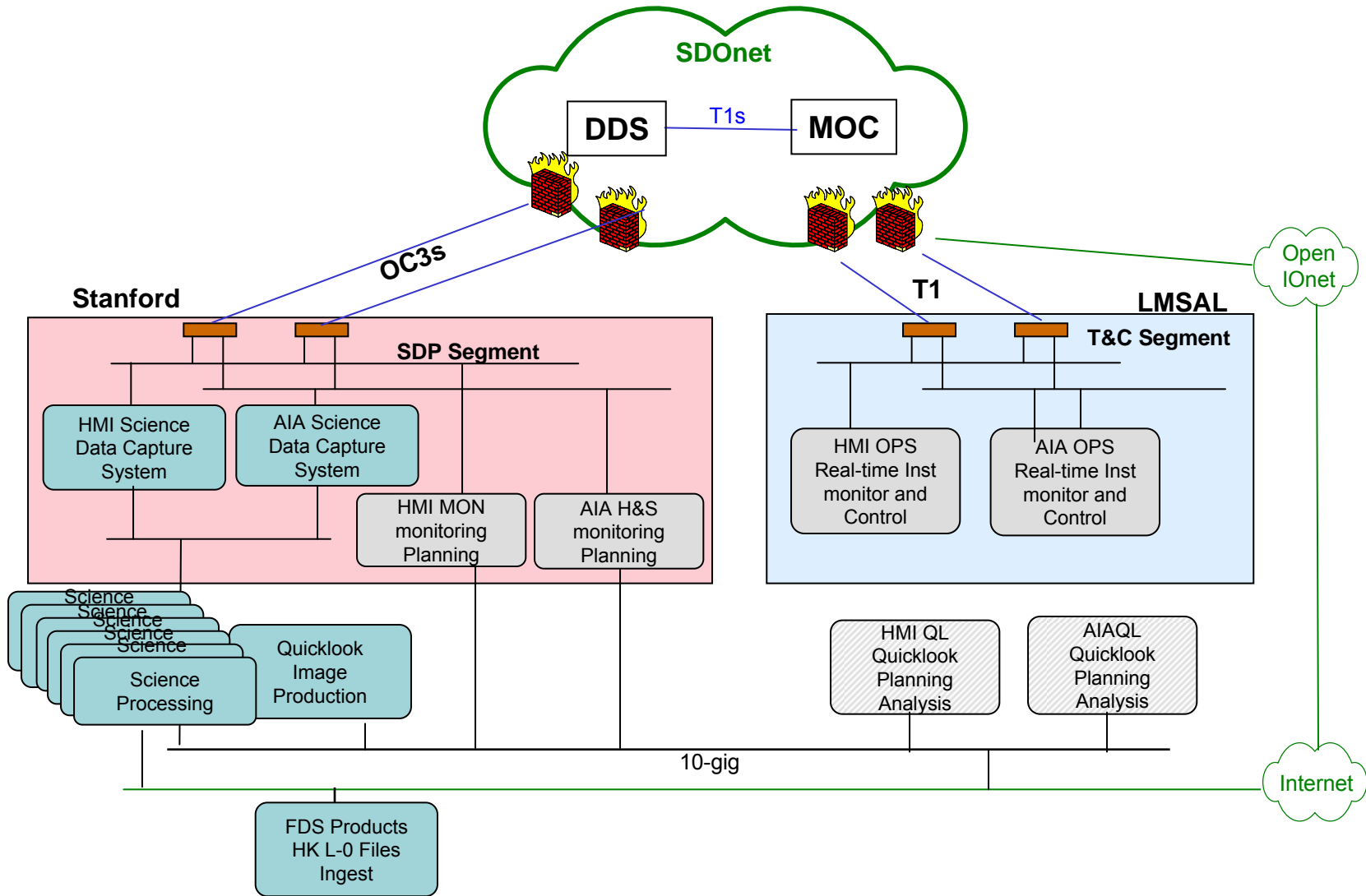


JSOC-SDP Major Components



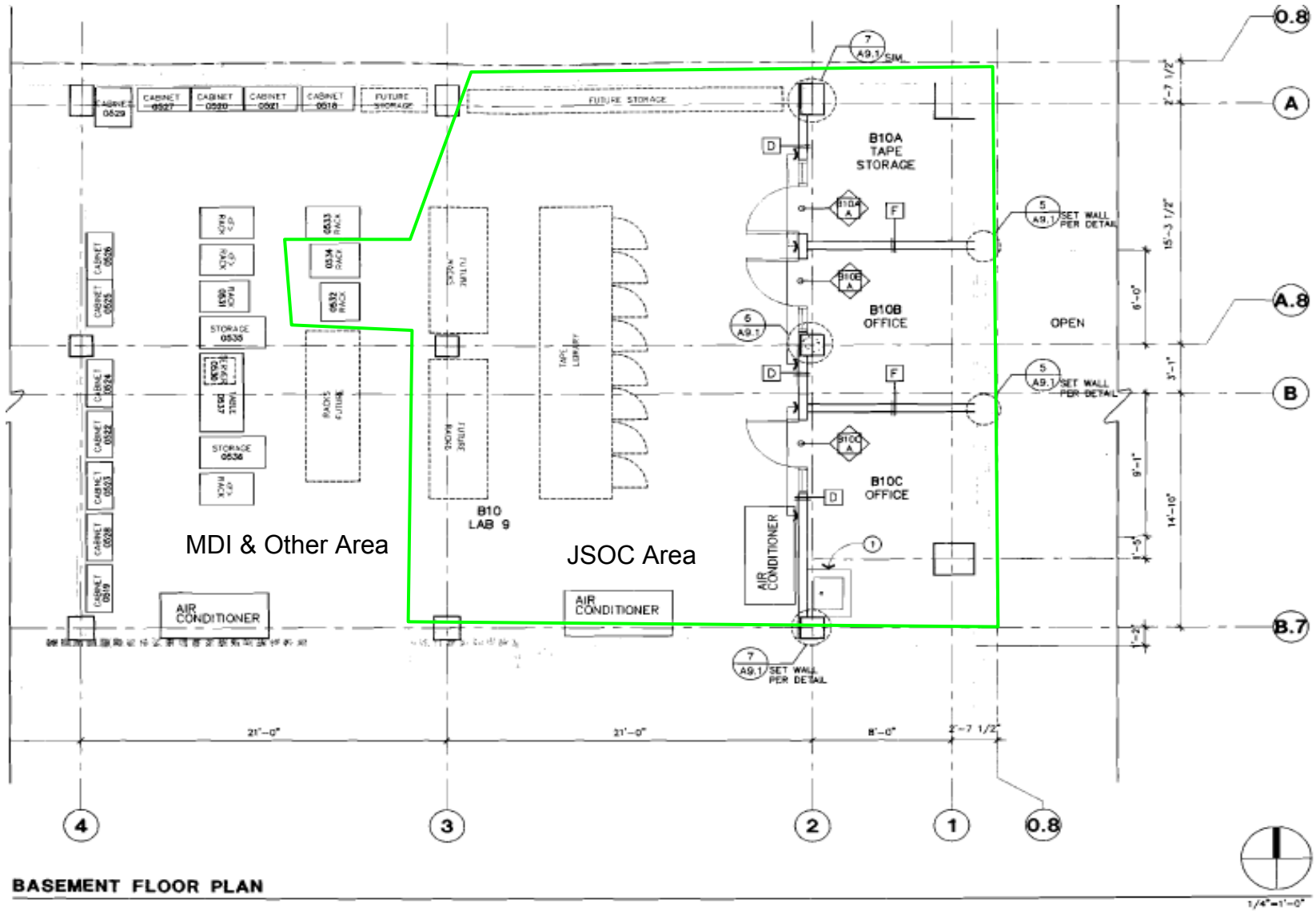


JSOC to SDOnet Network Connections





JSOC Computer Room Layout



BASEMENT FLOOR PLAN



Hardware Procurement Status



- **DCS system is in place and acceptance tested**
 - The 1st OC3 is in place
 - The DDS communications rack has been delivered and installed
 - Test data flows at expected rates and handshaking functions OK
- **Offsite and Catastrophe Systems ready to order now, tape systems in house**
- **Database, 16 nodes of Pipeline system, tape robotics and drives, disk server and initial 400TB procurement planned for December 2007**
- **48 nodes of Processor farm will be delayed until summer 2008**
- **Facility ready now**
 - Third chiller to be installed when needed
 - Door lock to be changed to SU ID card swipe