The SIM Science Data Center Prototype: An Integral Element of the SIM End-to-End Testbed

Raymond J. Bambery

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109
email: Raymond.J.Bambery@jpl.nasa.gov

Abstract

The Space Interferometry Mission (SIM) will be the first flight of an optical interferometer where the scientific return is co-equal with technology development objectives. The SIM Science Data Center (SDC) will be designed to monitor, process and archive the instrument engineering and science data collected over the five year operational life of the mission. Furthermore, it will host the postprocessing to be carried out for three years after the end of mission operations. Since the instrument is a complex technological and operations challenge, and due to SIM's ambitious astrometry and imaging goals, a prototype Science Data Center will be built and operated in parallel with the development of the instrument. In particular, this prototype system will be tested with SIM System Testbed-3 (STB-3) to command and monitor that instrument. After launch the SIM Science Data Center will continue to have access to STB-3 for testing of sequences and instrument operations modes.

1. Introduction

The Space Interferometry Mission[1] (SIM) will launch a 10-meter baseline optical interferometer into an Earth-trailing, heliocentric orbit on a Boeing Delta-3 launch vehicle in June, 2005. SIM will demonstrate important new technologies for future missions but, more importantly, will address significant science objectives during its five year operational lifetime. As part of NASA's Origins Program, the mission will have as one of its primary goals, the indirect detection of planets orbiting other stars.[2]

SIM's interferometer will observe the universe in two operational modes: astrometry and imaging. Over the life of the mission the astrometry measurements will provide a catalog of about 50,000 objects with position, proper motion and parallax to a precision of ±4 microarseconds. In imaging mode, SIM will perform both high resolution (10 mas per pixel) narrow angle multispectral imaging of a broad range of astrophysical phenomena as well as starlight nulling imaging of planetary dust disks. Imaging will be performed on some 500-1000 sources.

The astrometry mode provides the best chance of detecting planets in the range of 10 Earth masses to 10 Jupiter masses at 1-5 AU distances from the central star for distances up to 100 parsecs. It accomplishes this goal by the following scenario. Before launch, some 3000-4000 "stable" candidate grid stars, the **Astrometric Reference Grid**, which consists of stars randomly placed about the 4-π steradian celestial sphere, will be selected. Once in space, the interferometer will perform periodic campaigns to precisely measure their positions to ±4 microarseconds over the life of the mission. Other observations of stars, which are likely to have planets, will also be made periodically during the mission. Planets orbiting about these targets can be detected by measuring

"astrometric jitter", i. e., motions that are not accountable by proper motion and parallax, when compared to nearby Astrometric Reference Grid stars.

SIM has ambitious science goals for both astrometry and imaging. The mission's *defining characteristic* is that the spacecraft, the instrument and the observing scenario interact strongly in producing scientifically valid results. As a consequence, there needs to be a single point where all of the engineering and science data come together and where the investigator has a wide range of tools at his or her disposal to produce quality science. Since it is the first of many interferometer missions it will also be a training ground for future missions. The remainder of this paper will focus on the development of this center.

2. Science Data Center Development

Since SIM doesn't launch for seven more years, why begin the development of a Prototype SIM Science Data Center *now*? SIM was funded as a new start in October, 1997. In the two preceding years, though, a vigorous science program was initiated outlining the case for this relatively expensive interferometry mission in the era of the "faster, better, cheaper NASA". Technologically, astronomical optical and infrared interferometers are in their infancy in spite of the fact that they have been in use for over a century.[3] Many ground-based optical and infrared interferometers are under development or in the planning stages. The large number of papers submitted to this proceedings are a testament to that fact. The Hubble Space Telescope's Fine Guidance Sensors are an example of space-based interferometers for which some astrometry has been performed.[4] Most of these interferometers, though, have spent the lion's share of their early resources on overcoming difficult technology hurdles and so only recently has there been publications of any significant science results.

Each of these ground-based systems have developed analysis software based on their unique needs or attributes. The emphasis has been on real time control of the interferometer subsystems. The processing of science data has been left largely to professors and students of the collaborating universities usually back at their home institutions. SIM, on the other hand, is a flight project. The science results are the ultimate validation of the flight instrument performance. Timely feedback of science processing from a local data center to the instrument operations team will thus be critical to the success of SIM.

A prototype Science Data Center (SDC) must be available in 2001. In addition to the objectives in monitoring the instrument as part of the SIM System Test Bed 3 (STB-3) discussed in Section 8 below, it will be the focal point for the analysis of interferometry data from ground-based observatories as well. Before launch, the Science Data Center will be used for science target selection and planning. For example, candidate grid stars have to be selected for the initial Astrometric Reference Grid. Selection criteria have to be generated and results of imaging, spectroscopic and interferometric ground-based observing programs will be analyzed in order to produce the 3000-4000 initial candidates. Already the full statistical analysis algorithms to perform this "Astrometric Grid Reduction" processing in a systematic fashion are underway and an automated processing pipeline will be optimized at the SDC before launch.

After launch SIM undergoes a 30 day checkout period. During this period the instrument will be exercised in all of its modes, an in-flight calibration of all instrument components will be performed and the in-flight performance model will be computed. In addition, the first astrometric reference grid campaign will be initiated as soon as the instrument checkout is complete, All ground software, thus, will be exercised in every conceivable mode to its ensure its robustness for use during the five years of mission operations. Based on experiences in development of ground-based interferometers, the technical challenges of this flight instrument, and in view of the fact that the science is the ultimate validation of the instrument operation, then, all software development for the standard products at the Science Data Center must be robust at launch. Development must begin at the first opportunity to receive data from the instrument testbed. Currently, that first opportunity is in April, 2001 on the STB-3 instrument.

3. Science Data Center Design Goals

Briefly, the design goals are encapsulated in the keywords autonomy, automation, and simplicity.

Autonomy. The development of the SIM Science Data Center itself will begin in 2003. In the current vision this facility will host resident as well as visiting scientists. The computing resources of the center are being designed to function as an autonomous facility and it will appear to an astronomer to be merely an extension of his or her workstation. It will have a minimum in non-science ground support staff such as system administrators and maintenance personnel. In order to achieve this, the facility will have maximum hardware redundancy, automated hardware failure recovery schemes, minimal, but automated, backup of data sets, automated load balancing between servers and clients, and maximal interconnectivity to all elements of the project and the scientific community **(Figure 1)**. In brief, the computing facility will be so designed that during mission operations, maintenance will be a scheduled activity.

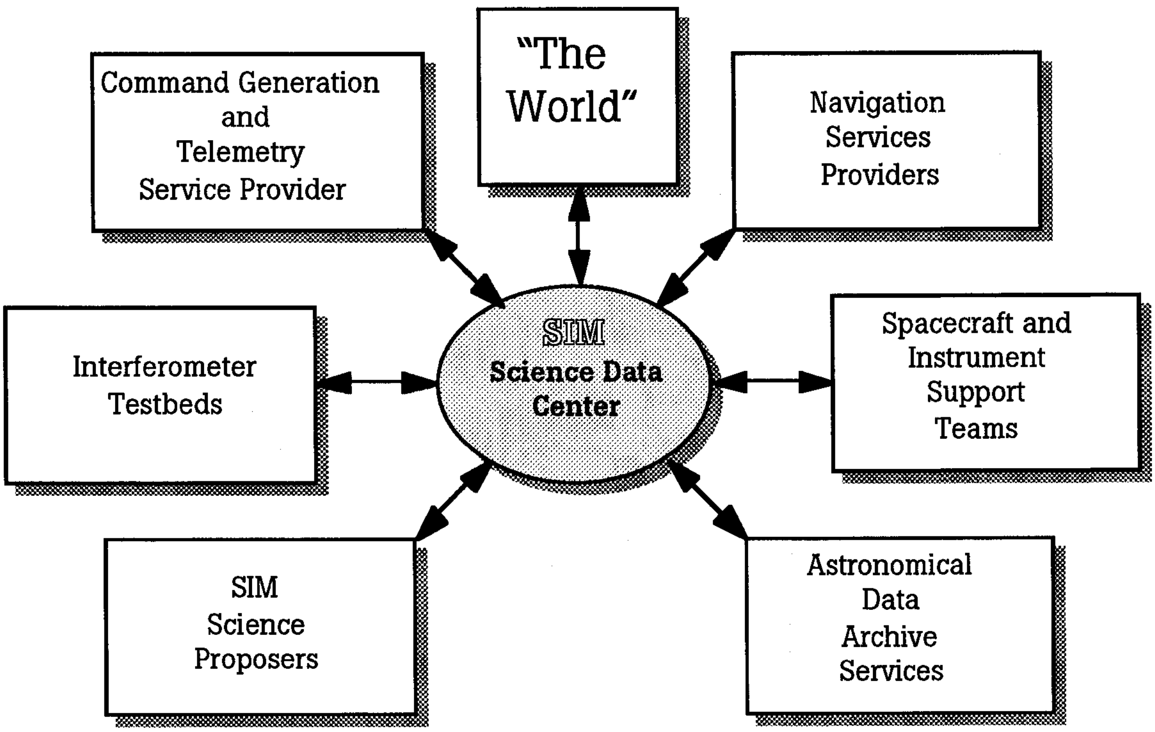

Figure 1. SIM Science Data Center Connectivity

Automation. All the primary (standard) data products of the system will be produced by systematic pipeline processing driven by a preset set of parameters. Some of these pipelines are discussed in Section 6 below. SIM needs to have significant optimization in software developments for automated pipeline processing and analysis algorithms to lessen the cost of product creation. This includes data archiving and distribution. It is envisioned that the raw data and standard products will reside on DVD or similar archive medium which will always be available on line. In the current design, data distribution during the mission would not be significant in quantity and may be done electronically.

Simplicity. The principal interface to the Science Data Center's computing resources will be through a WEB browser with pages specifically tailored to viewing SIM data in a variety of schemes. This will aid novices and allow control of proprietary rights, for example. Remote login windows will be available for the computing sophisticates. In addition to their uses in pipeline processing, many commercial off the shelf tools and software development packages will be available to users. These tools should satisfy all data visualization, statistics, mathematics, format interconversions and other processing needs. In order to facilitate science analysis many catalogs and tools to analyze other types of observational data will be readily accessible either at the center or via the internet.

4. Science Data Center Functions

The SIM Science Data Center will perform all six of the following functions for the science team during mission operations. and the last four for the three years of post-operations data analysis:

1. Science Observation Planning,
2. Instrument Performance Monitoring,
3. Science Processing and Analysis,
4. Data Product Production,
5. Data Storage and Management, and
6. Data Archiving and Distribution.

Science Observation Planning. Due to the complicated nature of the instrument operations it will be a planning challenge for the flight team. The SDC will provide tools for the scientists to plan their observations. The output of these tools will be placed into a queue which will be merged with all other types of observing requests, i.e., instrument calibrations, repeat requests, spacecraft maneuvering, etc. The merged sequences will be processed, validated and uplinked by other tools of the flight team by means of the Telemetry Service Provider. The uploaded sequence schedule will be archived at the SDC so that the scientists can monitor the progress of their observations and later verify that they were accomplished.

Instrument Performance Monitoring. The SDC will allow the user to monitor instrument performance from the raw telemetry. Instrument stability is a significant technology challenge and remote monitoring of this instrument will be a significant operations challenge. The instrument will operate in a number of submodes for both astrometry and imaging. The system will be so designed that the user can replay the telemetry under a variety of conditions depending on the instrument mode and the nature of data downlinked,

Science Processing and Analysis. The SDC will have a repository of tools for science analysis of the interferometer data. The users will also have the option of developing software on the SDC for analysis or perform processing at their home institutions.

Data Product Production. The SDC will perform Committee on Data Management, Archiving and Computing (CODMAC) definition levels[5] 1 through 4 instrument processing routinely via automated pipelines. The pipeline parameters will normally be fixed by the instrument scientists and standard output data will be archived in catalogs or in data files pointed to by the catalogs. Users, however, will be able to review the processing or resubmit selected data files to the processing pipeline with different parameters for their own purposes.

Data Storage and Management. Raw instrument and processed data at the SDC will be archived on CD-ROM, DVD or other media and managed via a data base manager as specified by the SIM Science Team.

Data Archiving and Distribution. The SDC is the repository for all SIM data, both engineering and science. It will serve as the distribution point for this data and will ultimately distributed to the science community and a NASA long term archive.

5. Typical SIM Data Products

The SDC will archive both raw and processed data of the mission. Archived standard data products run the gamut of CODMAC processing levels 0 through 4. The majority of processed products listed below fall in the CODMAC level 4 category, i. e., data products which are from model output or results from analyses of lower level data, e.g., variables derived from multiple measurements. Such products for SIM are:

1. Position coordinates, proper motions and parallaxes of all stars in the astrometric reference grid,
2. Position coordinates for all science targets,
3. Proper motions and parallaxes of many, if not most, science targets,
4. Other astrometric data such as binary star orbital elements,
5. Rotational synthesis images (u,v amplitude & phases),
6. Interferometric nulling images (u,v amplitudes only),
7. Instrument telemetry (engineering and science), and
8. Instrument performance model.

6. Typical Science Data Center Processing

Note the data flow into the Science Data Center from the spacecraft in **Figure 2**. Raw instrument data are the telemetry packets. After depacketizing the data is rearranged into time-tagged telemetry raw science data records and associated engineering information such as metrology, collector mirror offset data, etc. Other ancillary information regarding the spacecraft such as spacecraft velocity, heliocentric position, pointing data, etc. will also be archived. The basic first order astrometric product is a unit called a **scan;** a one-dimensional measurement of visibility and delay line position. A **tile** is composed of several scans for multiple objects at a particular inertial pointing of the spacecraft within the collector mirror's **field of regard**; typically one or more science targets along with three or more grid stars. A target's celestial coordinates are generated from two, or more, nearly-orthogonal tiles **(see Figure 3).** The delay line position, visibility phasor (complex) data from two or more **visits** are computed against some reference grid solution to produce Right Ascension, Declination and magnitudes for all objects within a tile.

A major processing effort will be expended the **Grid Campaign Sky Coverage** task. This grid of ~3000 stars distributed over the 4-π celestial sphere will serve as the basis for the global astrometry task and instrument astrometric calibration. After 5 years these positions should be accurate within ±4 microarcseconds.

Two types of image products (synthesis imaging and interferometric nulling) will be generated and two types of data flow will created. The measurement process for synthesis imaging is shown in **Figure 4**. This diagram shows that for imaging of resolved sources a bright, unresolved source will have to be nearby (within 0.5°) of the resolved source to act as a fiducial. This is required because the Guide Interferometers will have to reacquire different guide (not grid) stars as the baseline rotates through a 180° imaging observation. The Science Interferometer may have to collect as many as

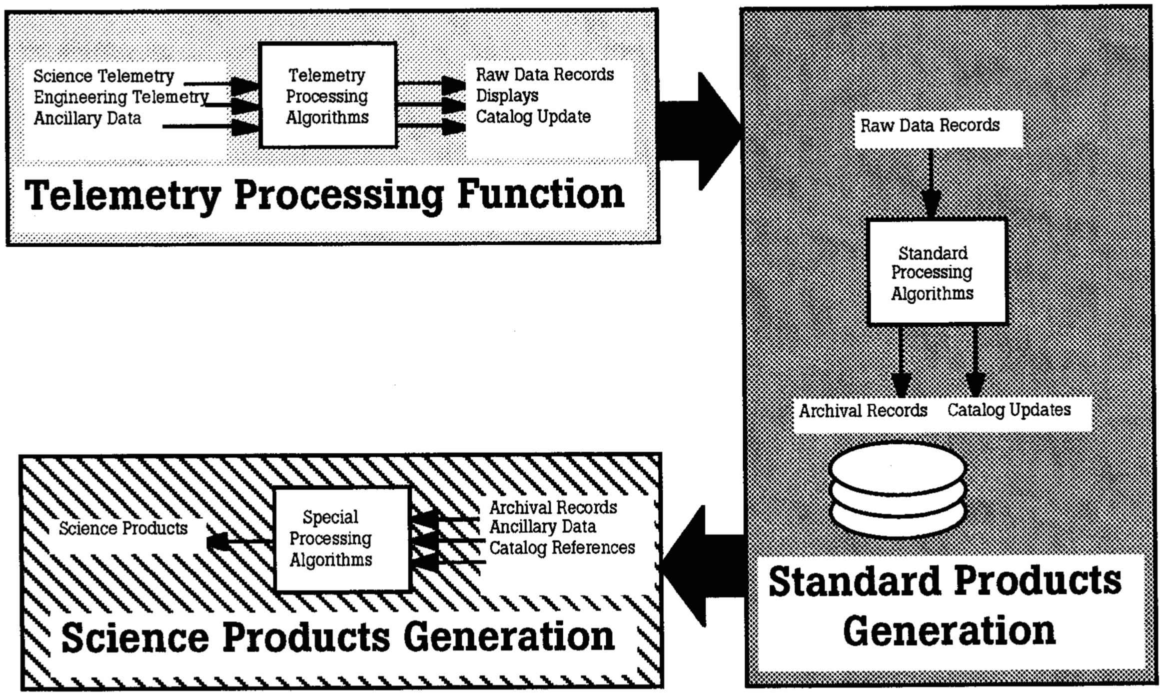

Figure 2. SIM Science Data Center Data Flow

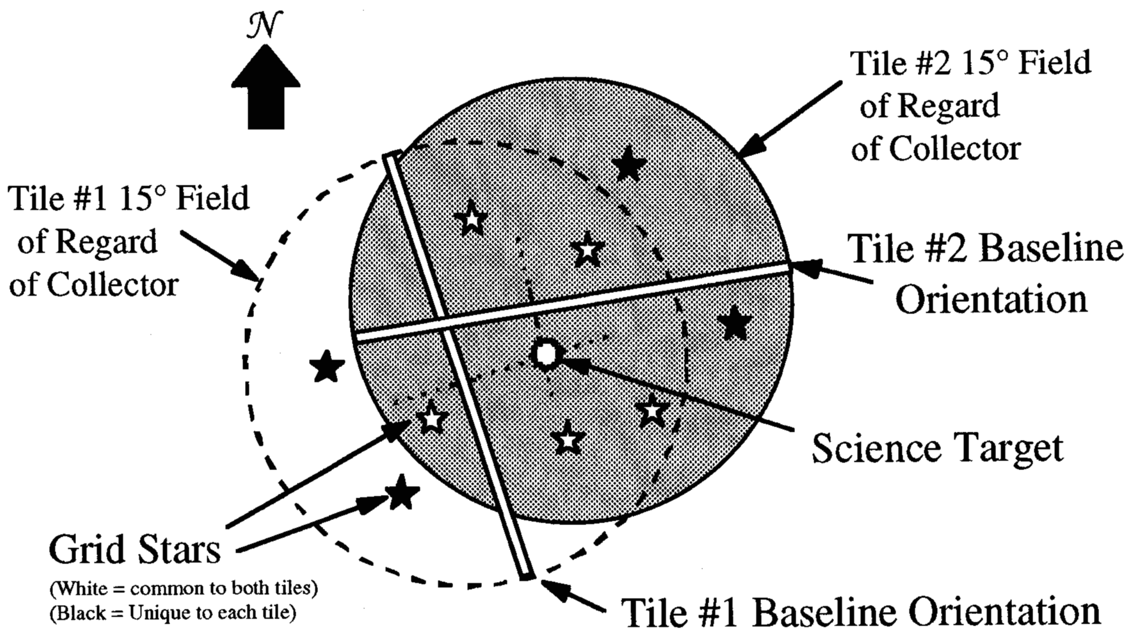

Figure 3. Astrometry with SIM

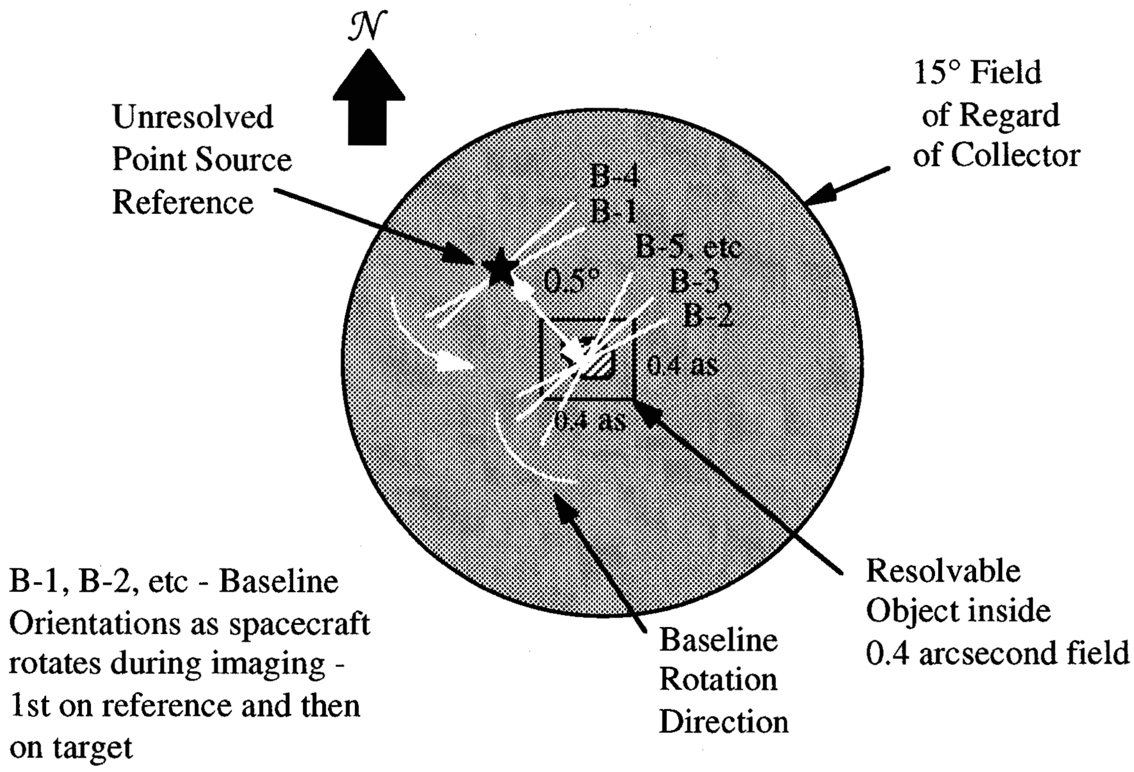

Figure 4. Synthesis Imaging with SIM

two delay line and visibility measurements for the point source and the resolved source for each baseline orientation.

Another major processing task will be calibration of the instrument. A number of routine processing tasks will be performed in this regard. For this, before launch, a ground instrument performance model will be created and then serve as a basis for an in-flight performance model. This in-flight performance model will constantly be updated during mission operations and aid in instrument health analyses and the Astrometric Reference Grid solutions.

Science analysis software tools for computation of binary orbits, stellar diameters, star cluster dynamics, galactic rotations, etc. with full statistical analyses will also reside on the Science Data Center. It is expected that a number of hypothesis testing algorithms will be developed and tested during mission operations.

7. SIM System Testbed 3

The purpose of the SIM System Testbed-3 (STB-3) is twofold: to reduce technology risk to SIM and to support SIM design, integration and test, and flight operations test objectives.[6] In the latter category two SIM data handling schemes are to be validated: (1) the on-board data handling, and (2) the ground system data processing. As part of the ground system a prototype Science Data Center will be integrated and begin testing in April, 2001. Before launch, STB-3, JPL's Flight System Testbed, the Ground System Testbed and the Science Data Center prototype constitute the SIM End-to-End Testbed **(Figure 5).** This testbed will operate throughout the project lifecycle.

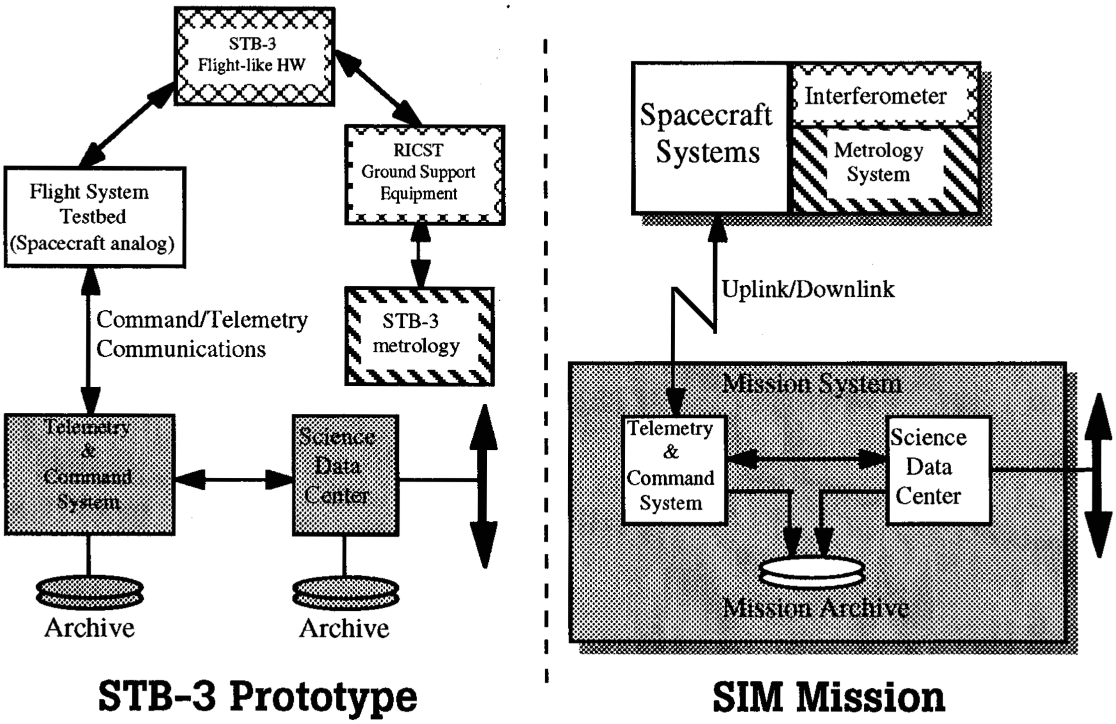

Figure 5. Architecture for STB-3 and SIM

STB-3 will consist of a suspended hardback structure having a combiner pod, two collector pods and a fixed delay compensator in the Spacecraft Assembly Facility high-bay at JPL. It will support the testing of three interferometers (two guides and one science) and all interferometer control electronics with real time control software. Of particular interest to the prototype science data center is the use of pseudo stars as test targets for tracking and instrument jitter to emulate on-orbit performance. These test data will be used in testing the prototype telemetry and the Instrument Health Display.

8. The SIM Science Data Center Prototype

The SIM Science Data Center Prototype is the proving ground for flight operations. Initially the prototype will consist of a workstation located at JPL near the STB-3 structure. As the prototype it will perform three useful functions for SIM in astrometric measurements. First, in concert with the ground data system prototype, it will allow for receipt and processing of the prototype telemetry data. This in turn, will provide a prototype Instrument Health Display. At the same time ancillary data such as instrument, spacecraft and support sensor data necessary for the proper cataloging of the raw data and initialization of the pipeline parameters will be prototyped. Second, it will allow prototyping of the pipeline processing which produces astrometric grid data and generation of the astrometric catalog. The initial goal will be to perform a "SIM tile reduction"; the minimal processing needed to produce a tile data set. Another activity in this category is the designing the processing for the instrument performance model for use with the Instrument Health Display and generation of the astrometric catalog. Third, it will provide a test platform for repository of useful software libraries for science algorithm prototyping.

At this time it is not known how much software development can be done with the STB-3 instrument data to prototype image products. Certainly, the two types of imaging telemetry packets for synthesis imaging and nulling can be created and processed. It is likely that some form of simulated science data can be created to prototype the image reconstruction processing for both image synthesis and nulling. However, details in this area will require further study.

The SDC and Ground Systems prototypes will be integrated with STB-3 in early 2001 and be ready for testing with the interferometer by April, 2001. When the Science Data Center itself is started in 2002-3, the prototype repository and pipeline processing functions will be transferred there where they will be refined until launch. The telemetry development will continue on the SDC prototype until launch. After that telemetry processing software will be made available at a variety of locations including the SDC.

9. Conclusion

In summary, the SIM Science Data Center is central to the success of SIM. It is the location where all of the mission data comes together to validate the design and operation of the instrument. Development will begin at the first opportunity to obtain instrument testbed data in April 2001. This paper addresses the design goals and the development of that Science Data Center. The data processing requirements of the center are now being developed with the cooperative efforts of the instrument designers, the SIM project science personnel and data system engineering professionals.

To keep up with the latest information on the dynamic SIM project please monitor the SIM home page at http://sim.jpl.nasa.gov/sim.

Definitions of Terms[7]

- **Astrometric Jitter** - Complex object motions remaining after removal of parallax and proper motions from the multi-year cumulative astrometric positions of the science targets with respect to the Astrometric Reference Grid.
- **Astrometric Reference Grid** - A grid of stars covering the whole sky, whose positions (parameterized to include proper motions, etc.) are known to high precision. The reference grid itself is not a science objective but provides the astrometric calibration for the SIM instrument and is fundamental for both wide-angle and narrow-angle astrometric measurements.
- **Field of Regard** - Defined as the conical angle of the sky area available around the nominal look direction of the collector mirror where the stars within the cone are observable by allowing the collector mirror to move throughout its maximum range of motion while the spacecraft is inertially pointed. This angle is nominally 15°.
- **Field of View** - Defined as the conical angle of sky area available around the nominal look direction of the collector mirror where the stars within the cone are observable without changing the orientation of the collector mirror while the spacecraft is inertially pointed. This angle is nominally 1/3 of arcsecond.
- **Grid Campaign Sky Coverage** - A systematic interlaced brickwork-like pattern of sky coverage with the spacecraft oriented in discrete pointings known as tiles. Adjacent tiles overlap one another to establish relative object positions and geometric continuity. As the interferometer is necessarily a one-dimensional device, the available sky is traversed twice in two quasi-orthogonal baseline orientations to provide the Right Ascension and Declination positions and isotropic measurement errors. Each campaign takes about 10-15 days and is repeated approximately four-and-half times per year.
- **Scan** - A continuous observation of a single target, made with the instrument non-rotating in inertial space. The duration of a scan is used to calculate instrument sensitivity, and therefore, excludes time taken after the instrument has stopped moving but is not ready to acquire data.
- **Tile** - A set of astrometric observations which is performed while the SIM spacecraft is intertially pointed. A tile is the second-lowest unit of astrometric data for SIM, the lowest being a delay line measurement of a single science star. Loosely, a "tile"

is all the observations within a given radius of the optics center of field of view, typically 7.5°.

- **Visit** - A set of scans of a target to define its stellar coordinates within a particular window of time, i.e., days. Multiple visits over longer periods of time, i.e., months or years, allow computations of parallax and proper motion.

Acknowledgments

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The author thanks Stephen Unwin for his thoughtful comments and suggestions in preparation of this manuscript.

References

1. M. Shao, "The Space Interferometry Mission", in *The Proceedings of the SPIE Conference #3350,* Astronomical Telescopes and Instrumentation, Kona, HI, 20-24 March 1998.

2. S. Unwin, "Science Goals of the Space Interferometry Mission", in *The Proceedings of the SPIE Conference #3350,* Astronomical Telescopes and Instrumentation, Kona, HI, 20-24 March 1998.

3. Peter R. Lawson (editor), "Selected Papers on Long Baseline Stellar Interferometry", *SPIE Milestone Series Volume MS-139,* SPIE Optical Engineering Press, Bellingham, WA, 1997. Note especially the first seven papers published between 1868 and 1896.

4. G. F. Benedict, et. al., "Working with a space-based interferometer: Hubble Space Telescope find guidance sensor 3/small field astrometry", Poster paper [3350-91], in *The Proceedings of the SPIE Conference #3350,* Astronomical Telescopes and Instrumentation, Kona, HI, 20-24 March 1998.

5. For the CODMAC instrument data processing levels definitions please see WEB site: http://eospso.gsfc.nasa.gov/eos_reference/TOC.html.

6. R. Laskin, "SIM System Testbed-3 (STB-3)" , in *SIM Project Preliminary Instrument System Requirements Review,* Jet Propulsion Laboratory, Pasadena CA, 17-18 March 1998.

7. S. C. Unwin, M. Shao, D. M. Peterson, and C. Jones, "SIM Science Requirements Document (DRAFT)", JPL Internal Document, March 16, 1998.