RESPONSE TO MARS RECONNAISSANCE ORBITER'S END-TO-END DATA ACCOUNTABILITY CHALLENGES[1]

Y. Lee

Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, California, 91109, U.S.A., E-mail: Young.H.Lee@jpl.nasa.gov

ABSTRACT

NASA launched the Mars Reconnaissance Orbiter (MRO) on August 12, 2005. It carries six science instruments and three engineering payloads. Because MRO will produce an unprecedented number of science products, it will transmit a much higher data volume via high data rate than any other deep space mission to date. Keeping track of MRO products as well as relay products would be a daunting, expensive task without a well-planned data-product tracking strategy. To respond to this challenge, the MRO project developed the End-to-End Data Accountability System by utilizing existing information available from both ground and flight elements. Therefore, a capability to perform first-order problem diagnosis is essential in order for MRO to answer the questions, *where is my data?* and *when will my data be available?* This paper details the approaches taken, design and implementation of the tools, procedures and teams that track data products from the time they are predicted until they arrive in the hands of the end users.

1.0 INTRODUCTION

All spacecraft carrying multiple instruments must face end-to-end data accountability issues during flight, for example, lost commands, spacecraft anomalies, and problems in spacecraft-to-ground communications can contribute to unexpected data loss and degradation of data quality. To identify and address these data loss problems in the past, arduous workarounds and last-minute tools have been developed during the mission operations phase, resulting in significant cost impacts to the overall project. Projects can achieve significant cost benefits by confronting this problem in a systematic manner during *the development phase* of the mission. Mars Reconnaissance Orbiter (MRO) was the first NASA mission to proactively address end-to-end data accountability during the development phase.

Over the course of its mission lifetime, MRO is expected to return more than 34 terabits of data—over three times as much data as two previous and three current NASA missions combined.[2]

The MRO spacecraft carries six science instruments and three engineering payloads. In addition to providing a platform for several science instruments, MRO will be the telecom relay asset for some upcoming Mars missions. Since MRO will produce an unprecedented number of science products it will transmit in a much higher data volume than any other deep space mission to date. Keeping track of MRO products as well as relay products would be a daunting, expensive task without a well-planned data-product tracking strategy.

To respond to this challenge, the MRO project developed the End-to-End Data Accountability System by utilizing pre-existing information available from both ground and flight elements. For example, large data products require multiple days to deliver (for instance, the High Resolution Imaging Science Experiment [HiRISE] is expected to create images of anywhere from a few gigabits to 28 gigabits), passing through complex flight and ground data flow processes, which utilize many new flight and ground components. The capability to perform first-order problem diagnosis is essential in order for MRO to answer the questions, *where is my data?* and *when will my data be available?*

The main goals of the End-to-End Data Accountability System are to (1) provide automated end-to-end data product status information, from science and relay data product planning through their entire life cycle until delivered to end-users and (2) assist the team in performing data analysis and data problem resolution in a timely manner.

2.0 MRO END-TO-END DATA ACCOUNTABILITY SYSTEM CHALLENGES

Current JPL projects have experienced significant problems with data accountability issues because end-to-end data accountability related issues were not addressed in a systematic manner during the projects'

[1] The work described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

[2] DS1 produced 15 Gb of data, Mars Odyssey 1,012 Gb, Mars Global Surveyor 1,759 Gb, Cassini 2,550 Gb, and Magellan produced 3,740 Gb of data whereas MRO is expected to produce 34 Tb.

development phase, resulting in arduous and costly workarounds last-minute tools-development during the operations phase. For MRO, these problems would have been exacerbated because of the unprecedented data volume MRO will produce during the primary science phase (around November 2006), where at times, large science data products will require delivery times of multiple days. Thus, there is a significant need for a capability to track science and relay products from product planning through data product delivery to science and instrument teams to answer the important end-user question, *where is my data?* Moreover, a capability to perform first-order data problem diagnosis is essential because MRO has lots of new flight/ground components and will use complex flight/ground data flow processes.

3.0 APPROACHES TAKEN

To address MRO data accountability challenges proactively, the project decided to create an end-to-end data accountability system during the mission-development phase. Since the development of this accountability system was initiated six months prior to the Critical Design Review (CDR), an approach was taken to utilize existing design elements as much as possible while minimizing any critical change requests to the current mission design so that both cost and schedule impact could be avoided. As a result, the following key system elements that were already in work were heavily relied on in the design and implementation of this accountability system:

- On-board science data management capabilities
- Advance Orbiter System (AOS) frame processing and generation of AOS frame accountability
- Product telemetry processing including relay product telemetry
- Product telemetry files generation and management
- Defined uplink/downlink processes

MRO's End-to-End Data Accountability System is composed of four major components:

1. Data management software,
2. Mission data management team processes,
3. end-to-end Data Accountability Software (eeDAS), and
4. end-to-end Data Accountability Team (eeDAT)

The following sections of this paper will only focus on the eeDAS components and eeDAT roles and its processes since they are the new elements in MRO (data management software and mission data management team processes are JPL multi-mission based that adapted for MRO).

4.0 END-TO-END DATA ACCOUNTABILITY SOFTWARE

The eeDAS (Figure 1) was created with the objectives to provide:

- Automated end-to-end data product status information from science and relay data-product planning at any point in the delivery cycle, where the software displays the status of data products and helps locate the first problem occurrence of product delivery/quality (real/suspected)
- Tools to support the eeDAT ability to diagnose and coordinate resolution of suspected data product-flow problems and perform data analysis and data problem resolution

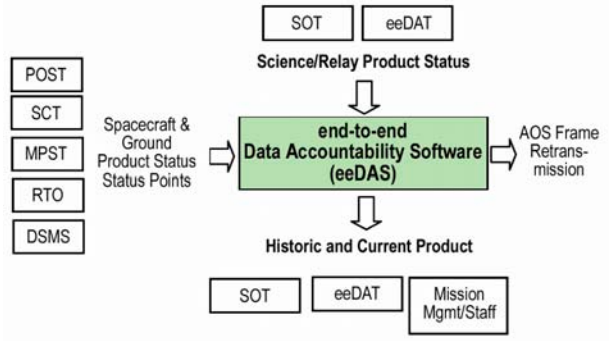

Fig. 1. eeDAS Context Diagram[3]

Product Predictor is a software tool that captures science and relay product identification provided by instrument and relay teams during the science planning cycle. Then, the automated web-based tracking system, Tailorable Report Utility for Science product Tracking (TRUST)—a window to the product status for science/ instrument teams—will use this Product ID to track product status information as this product is traveling throughout entire MRO data delivery path. TRUST facilitates science end users, Mission Operation System (MOS) users, and Ground Data System (GDS) users in monitoring science or relay product delivery status via defined tacking locations in uplink, spacecraft, and downlink processes.

When data dropout is detected on the ground based the AOS frame accountability reports, the Retransmission

[3] POST: Payload Operation Science Team, SCT: Spacecraft Team, MPST: Mission Planning and Sequencing Team, RTO: Real-time Operations Team, DSMS: Deep Space Mission System Team, SOT: Science Operations Team

Manager will generate command sets based on the gaps, spacecraft commanding timing, and a set of user-specified criteria, and monitor the submission of retransmission commands to fill the data gaps.

eeDAS is composed of three independent tools:

- Product Predictor: generates product predict list (PPL) file to validate MRO science and relay product files based on the final product of the science planning processing
- Tailorable Report Utility for Science product Tracking (TRUST): provides capability to detect product status in several locations throughout the system such as uplink, spacecraft, and downlink processes
- Retransmission Manager: monitors execution of retransmission commands generation based on the frame accountability reports to handle data dropouts

4.1 Product Predictor

Product Predictor takes the science planning results as inputs, updates them throughout the execution cycle, and provides the science teams and eeDAT with up-to-date expected product lists for each instrument. Product Predictor integrates the information from the integrated payload target file (IPTF), which is used to generate integrated target lists (ITLs), with the non-interactive payload express commanding (NIPCEC) and non-interactive file load (NIFL) submissions from the various science teams that support the observations during the primary science phase (PSP) and the relay phase (RP) of the MRO mission.

The major purpose of the Product Predictor is to provide TRUST with the product predict list, which is the time-ordered, expected science products list that MRO will generate as a result of executing the ITLs and the science submitted NIPCEC/NIFLs (Figure 2). From PSP science observations on, Product Predictor will provide TRUST with observation IDs, product IDs, estimated raw size of predicted products, uplink status of the predicted products, observation support files, uplink product origin, predicted observation start time, and so on. In addition, the program generates a human-readable PPL to assist eeDAT to perform data analysis and data problem resolution.

```
# Product Predict List:0000
# FileName: PPL_IPTF
# Creation Time: 2004-352T14:00:48
# Producer: PP
# Creation Reason: POST_IPTF
# Expectation Period: 2007-339T14:00:48 2007-339T20:37:18
# EOH
0000,2007-339T13:58:06,324.00,4A-030000006400,UNK,45.311,1.48,X,0,PLANNED,1,UNK,1,100
0000,2007-339T13:58:06,324.00,4A-020000006400,UNK,203.752,1.48,X,0,PLANNED,1,UNK,2,100
0000,2007-339T14:06:36,173.89,4A-07000003E800,UNK,838860800.000,1,X,0,PLANNED,1,UNK,3,d:/shr/SH_SVT_MAP1_1_1000.OST
0000,2007-339T14:06:36,173.89,4A-07000003E900,UNK,838860800.000,1,X,0,PLANNED,1,UNK,4,d:/shr/SH_SVT_MAP1_1_1000.OST
0000,2007-339T14:06:36,173.89,4A-07000003EA00,UNK,838860800.000,1,X,0,PLANNED,1,UNK,5,d:/shr/SH_SVT_MAP1_1_1000.OST
0000,2007-339T14:06:36,48.32,4A-07000003EB00,UNK,233119416.320,1,X,0,PLANNED,1,UNK,6,d:/shr/SH_SVT_MAP1_1_1000.OST
0000,2007-339T14:25:23,11.25,4A-051UNKNOWN00,UNK,703.125,1.00,X,0,PLANNED,1,UNK,7,
0000,2007-339T14:26:10,15.40,4A-04UNKNOWN00,UNK,316.000,1.00,X,2,PLANNED,1,UNK,9,d:/ctx/ctx_itl1.nifl
0000,2007-339T14:26:15,5.00,4A-01F80148BC00,UNK,3222.000,DISABLE,X,2,PLANNED,1,UNK,8,d:/hir/r004160833724.mod
0000,2007-339T14:33:07,98.18,4A-04UNKNOWN00,UNK,2014.500,1.70,X,0,PLANNED,1,UNK,10,
0000,2007-339T14:40:51,332.00,4A-030000006500,UNK,314.481,1.48,X,2,PLANNED,1,UNK,11,101
0000,2007-339T14:40:51,332.00,4A-020000006500,UNK,1328.722,1.48,X,2,PLANNED,1,UNK,12,101
0000,2007-339T14:43:29,15.40,4A-04UNKNOWN00,UNK,316.000,1.00,X,2,PLANNED,1,UNK,14,d:/ctx/ctx_itl1.nifl
0000,2007-339T14:43:34,5.00,4A-01F801492C00,UNK,1617.000,DISABLE,K,4,PLANNED,1,UNK,13,d:/hir/r004160833836.mod
0000,2007-339T14:48:17,98.18,4A-04UNKNOWN00,UNK,2014.500,1.00,X,0,PLANNED,1,UNK,15,
```

Fig. 2. PPL Contents as prepared for PSP Sequence Verification Tests

4.2 TRUST

TRUST is used by the eeDAT to monitor the movement of science and relay product telemetry, from project planning to spacecraft processing, through data downlink. TRUST is a product status tracker, automated web-based tracking system, similar to the FedEx package tracking system. TRUST will compare received telemetry and tracking accountability data against predicted product information to note irregularities anywhere along the data delivery path as a first step in performance monitoring.

After the entire mission data flow was examined to identify the most effective and efficient product status checking points, considering cost and implementation schedule, the following tracking points were selected for TRUST to implement:

- Product Predicts: Provided by Product Predictor to TRUST with sufficient information for creating an "expected product" record. This is the beginning of the product tracking process.
- Product sequenced and commanded status
- The following spacecraft onboard accountability status points, which were part of onboard science data management capabilities, were selected:
  - Sequence/Command: To ensure that products are not rejected or lost in sequence processing
  - Onboard Command Status: To verify that the spacecraft has loaded ITL into the Target Determination Tool (located on the spacecraft) and executed each line correctly
  - Instrument Downlink Started: To verify that the ground is receiving the data for a specified instrument
  - Onboard Execution Status: To verify the last read product ID
  - Onboard Downlink Status: To verify that the spacecraft has begun sending the product to the ground
- Product Construction Status: To obtain detailed product status information as ground processing occurs such as product size, product status (i.e., partial or full), transaction and timer status and number of gaps detected and so on.
- Product Availability Status to Users: To notify users when the product is available for retrieval.

In order to start tracking science data products, TRUST receives the PPL from Product Predictor (Figures 3 and 4). Using Product ID, the Onboard Product Status Monitor (OPSM) monitors products as they are assembled on the spacecraft and captures all onboard status values for ITL, Sequences, Solid State Recorder (SSR), Raw Buffer, Frame Buffer through the Real-time Telemetry Translator module which translates telemetry information per product ID to TRUST. OPSM also captures spacecraft event messages through its Event Reports (EVR) Telemetry Translator so that eeDAT can use the information for understanding spacecraft behaviors in relation to data accountability.

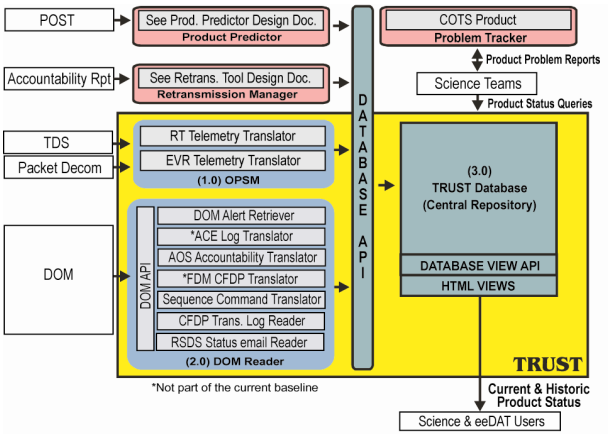

Fig. 3. TRUST Context Diagram[4]

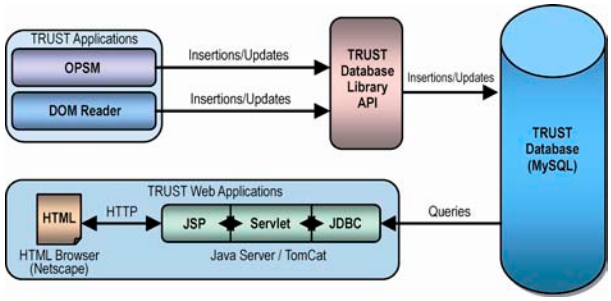

Fig. 4. TRUST Data Base Design

The Distributed Object Manager (DOM) Reader polls various ancillary files for TRUST to detect events and display them as they occur.

The TRUST has an encrypted web server using standard secure socket layer (SSL) and Lightweight Directory Access Protocol (LDAP) to ensure secure communication with authorized mission personnel. Open-source packages were used extensively by TRUST: for example, Apache and TomCat were used for web-server functionality; Openldap, openssl were used for encryption and access control to the web server; and MySQL was used as the database. The web interface was written in Java servlets and Java Server Pages (JSP) using Java Database Connectivity (JDBC) as the interface to the MySQL database (Figure 5).

Fig. 5. TRUST Product Selection Window

As a user selects either a product ID, an instrument type, or a predicted planned time range along with other options of interest, TRUST will display its high-level product summary page (Figure 6) so that the user can get a quick overview of their status.

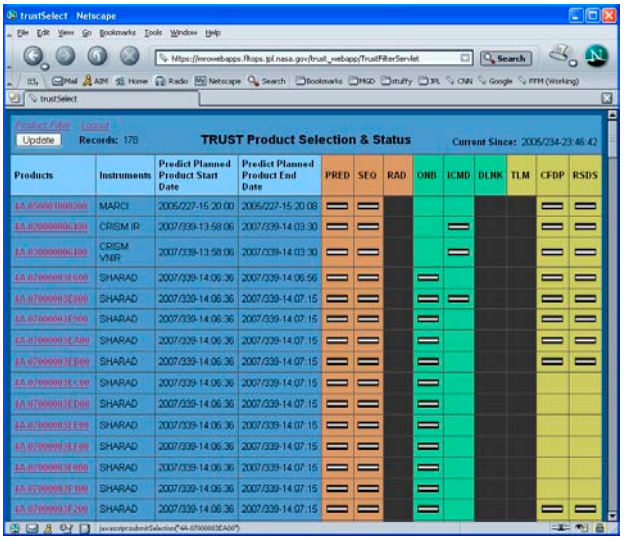

Fig. 6. TRUST Product Summary List Window

[4] API: application program interface, CDFP: Consultative Committee for Data Protocol, COTS: commercial off-the-shelf, DOM: Distributed Object Manager, EVR: Events Report, FDM: File Delivery Manager, RSDS: Remote Science Data Server, RT: real time, TDS: Telemetry Delivery System

After a user selects a particular product, TRUST will display the detailed product status information (Figure 7). It will have uplink, spacecraft onboard, and downlink-related product status information along with important spacecraft events, system errors, and anomaly information that are relevant to product accountability.

Fig. 7. TRUST Detailed Product Info Window

TRUST was operational in the configuration shown in Figure 8 during instrument checkout periods in August 2005 although its primary usage will be during the PSP in November 2006.

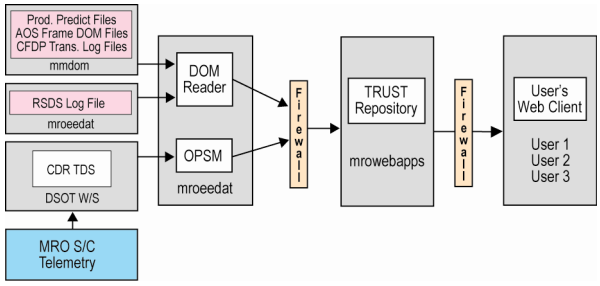

Fig. 8. TRUST System Configuration

4.3 Retransmission Manager

Retransmission Manager is used by the eeDAT to enhance the overall science and relay product completeness when a data gap is observed on the ground. Retransmission is scheduled to be automated during PSP background sequence execution.

Retransmission Manager consists of two software modules. The first, Retransmission Scheduler, determines how and when the Retransmission Command Generator (RCG) responds, based on the Retransmission Configuration File (RCF), which is updated by the Payload Operation Science Team (POST) for each 28-day science planning cycle during both the PSP and the RP. The purpose of the RCF is to set thresholds and parameter values that influence the retransmission program execution and determination of how the automated retransmission process is to be operated by Retransmission Manager.

The second module, RCG, generates the sequence to retransmit missing data during the time interval specified in the Retransmission Schedule File (RSF), which is generated by the POST, based on the POST-maintained RCF. The main purpose of RSF is to supply timing threshold information to the RCG, which executes within the maximum threshold values specified from the schedule file and configuration file.

5.0 END-TO-END DATA ACCOUNTABILITY SYSTEM PROCESS

The MRO Mission Data Management Process produces and distributes raw science data products and ancillary data. It also provides accountability for data production.

Raw science data products are the result of instrument observations and experiments as recovered from the telemetry stream or the results of tracking activities. Ancillary mission data generally concerns the circumstances under which an observation or experiment occurred. Such information is typically required in order to interpret the data. Processing levels for ancillary data collected on MRO vary.

Production and distribution of MRO raw science and ancillary data vary according to the manner in which the data is generated and transported. Consequently, the Mission Data Management Process can be viewed as the result of various combinations of the following seven MRO subprocesses:

- Monitor Mission Data Management Performance
- Produce and Distribute Raw Science Products (File Telemetry)
- Produce and Distribute Instrument Engineering Products (Packet/Channel Telemetry)
- Distribute Raw Science Products (Tracking Data)
- Distribute Ancillary Data (Ops Files)
- Prepare and Archive Level 0 Data
- Distribute Ancillary Data (SPICE Kernels)

Most of these subprocesses are developed and operated by the JPL Deep Space Mission System (DSMS) in the form of multimission operations services. The exception is Monitor Mission Data Management Performance (MDMP), the implementation and operation of which is specific to MRO and developed and operated by eeDAT.

The MDMP receives input from both external processes as well as internal MDMP. From the external processes, the MDMP subprocess receives planning, stored sequence generation, relay coordination, and command radiation information. From the three MDM sub-

processes that produce and distribute raw science products (i.e., File Telemetry, Packet/Channel Telemetry and Tracking Data), the MDMP subprocess receives accountability information. Output data is used to provide problem notifications to the Anomaly Response Process, MDM Performance Reports to the Status and Coordination Process, and Corrective Action Requests to the Configuration and Management Process.

6.0 END-TO-END DATA ACCOUNTABILITY TEAM

MRO decided to form a team that could serve as the point of contact for support to science operations teams. Its purpose is to track and troubleshoot science data-product delivery issues. Established during the MRO mission-development phase, the eeDAT was formed to:

- Define, implement, and operate an MRO end-to-end product accountability process, where "end-to-end" means from end-user planning and scheduling to delivery to end-user, of science- and relay-related products and "accountability" means tracking product flow and perform first-order problem diagnosis
- Participate in the MRO MOS, GDS and flight software designs, to ensure that adequate capabilities exist to support the end-to-end product accountability process
- Design overall MDMP
- Be cognizant of specific allocations of DSMS-supplied data-management responsibilities and interface
- Plan, design, and coordinate integration and test of process implementation
- Assess and certify process verification
- Monitor and assess end-to-end performance of MDMP
- Lead diagnosis and/or troubleshooting where multiple contributors are involved

However, during the mission operations phase, especially during the Primary Science Phase and the Relay Phase, eeDAT is responsible for:

- Science and Relay Product Telemetry Accountability Process: Tracking Science/Relay Data Product Status from planning to delivery
- TRUST Operations
- Monitoring and assessment of end-to-end data flow performance:
  - Notify science users of product gaps
  - Provide systematic feedback for tuning production parameters
  - Diagnose and coordinate resolution of suspected data product flow problems
    - If in-depth analysis is required to resolve an anomaly, contact appropriate team upstream of the problem location
  - Report performance assessment to MOS and science users
- Retransmission Process
- Product Predict List File Generation Process

In addition, eeDAT coordinates DSMS service for:

- Raw science product management, generation, distribution, and archival
- Engineering data management, generation, distribution, and archival
- MOS and science user support
- Frame and packet accountability report generation

Based on system-monitoring and performance assessments, during nominal operations, eeDAT will notify science users of product gaps. However, during off-nominal operations, eeDAT will execute four steps for data accountability for problem analysis and resolution:

- Step 1: eeDAT uses the web-based tracking tool to view routine product status or respond to science/instrument requests.
- Step 2: eeDAT identifies the team that the problems belong to, such as Science Operations Team, Payload Operation Science Team, and Spacecraft Team.
- Step 3: Cognizant team performs detailed analysis and eeDAT coordinates dissemination of the findings.
- Step 4: eeDAT collects lessons learned and enhances accountability processes for continual improvement.

One typical operations scenario that used the above-mentioned eeDAT approach follows.

A project scientist calls eeDAT to say Product A has not yet been received. The eeDAT evaluates accountability and tracking data. Two possible outcomes of the analyses are:

- eeDAT discovers the data was received and is being processed correctly. For this outcome, eeDAT will notify the scientist that Product A is somewhere in the science data product delivery cycle, that the scientist will receive the data later, and when to expect the data.

Or

- eeDAT discovers the scientist should have already received the product. Using the available monitoring data via TRUST, the eeDAT analyzes what happened and tries to answer the following types of questions:
  - Was project science planning correct?
  - Did the spacecraft perform the data capture activity correctly?
  - Did the spacecraft process the data (e.g., packetize it) correctly?
  - Did the ground system receive the data?
  - Did the ground system process the data correctly?

Then, the scientist will be informed via the Problem Notification process and the findings will also be passed to the Initiate Corrective Actions process for appropriate action.

Furthermore, eeDAT will gather and analyze metrics about the product to further understand the problem that needs to be resolved and to create performance report statistics and problem summaries for management review. Then, the metrics and other analysis input will be used to answer the following kinds of questions:

- Do the metrics show a repeating problem pattern?
- Do the metrics help identify where in the product cycle the problem occurs?
- Is the problem the result of something the scientist consistently forgets to plan?
- Is the problem caused by something that the spacecraft overrides consistently in the SSR?
- Is the Ground Data System consistently loosing the data?

These analysis results will facilitate continual improvement of the MOS/GDS processes throughout the long mission operations phases.

7.0 STATUS OF THE END-TO-END DATA ACCOUNTABILITY SYSTEM

Although the End-to-End Data Accountability System was heavily used and participated in PSP sequences validation tests during MRO's development phase, its prime usage will be during mission's PSP phase in November 2006. However, numerous system demonstrations have been conducted for MOS/GDS and science teams. Recently, the system was completely tested for full functionality with live MRO flight data during the L+18 instrument calibration checkout on August 30, 2005. Since then, the system has been used to support major project tests. MRO has agreed to accept a few bug fixes, rearranged screen layouts, and performance-related updates as a part of GDS version 6.0, due to be delivered by December 1, 2005 to be fully ready for its "prime time" activities.

8. DESIRED ENHANCEMENTS

Throughout system demonstrations and participations in project milestones, the following items were identified as strong candidates for system enhancements:

- While the End-to-End Data Accountability System was under development, the HiRISE instrument team submitted a change request that the HiRISE image data product be redefined. The revised product would be the channel image data products, which would be created by the Remote Science Data Server (RSDS) breaking the image product into channel image products prior to transmittal to HiRISE science team. Thus, it would be very beneficial for the HiRISE Instrument Team to have status information about channel image products from the RSDS.
- Currently, the TRUST OPSM module that collects instrument status information onboard from telemetry data cannot obtain spacecraft-related information such as onboard command status and instrument space event time because the Context Imager (CTX) instrument software cannot interface with the spacecraft flight software. It would be beneficial for the CTX team and eeDAT to have spacecraft-related information via TRUST because it would provide insight into the CTX instrument software interaction with the spacecraft flight software.
- Electra instrument products are "double CFDP-wrapped" (first wrapped by the relay asset and second wrapped by the MRO spacecraft) since they are relay data products. At present, TRUST only tracks status information for the first unwrapped Electra products on the ground. Hence, status information about the second unwrapped Electra-extracted relay products, which would be useful to have, is unavailable.
- Since the TRUST database collects onboard Mars Climate Sounder (MCS) product generation related information, it would be beneficial to the MCS team if MCS-specific displays could be designed and implemented via TRUST.

9. BIOGRAPHY

Young H. Lee is the Mars Reconnaissance Orbiter end-to-end Data Accountability Team chief, responsible for development of the MRO End-to-End Data Accountability System, operations team formulation and implementation. Since joining JPL in 1985, she has gained a wealth of flight project development and operations experience, such as oversight, hands-on development, and operational support for 14 NASA missions, including MRO, Mars Exploration Rovers, and Cassini. She has an M.S. in management of information systems from the Claremont Graduate University in California and is pursuing her Ph.D. at the Claremont Graduate University.

10. ACKNOWLEDGEMENT

Many thanks to JPL colleagues, the end-to-end Data Accountability System design and implementation team members: Bryan Allen, William C. Eggemeyer, Lamar Gilliam, Tim Reeve, Ted Specht, Robert Wendlandt, and Randall Reed from MissionSpace Corporation.

Additional special thanks for strong support from the MRO management: Jim Graf, Ben Jai, B. Gentry Lee, and Richard Zurek.