National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AIRS Product Development

The AIRS Software Development System

Steven Friedman
AIRS Science Processing

April 26, 2011

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration.

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Topics

- **AIRS Software Development Activities: V6 Specific**
  - CM Processes
  - Development and Testing
- **AIRS Computing Facility Status**
- **Impact of System Intrusion on AIRS Activities**
- **V6 Completion Phasing**

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AIRS Software Development Activities - CM Processes

- **AIRS uses the same successful software development process it started with:**
  - Science Team develops code
  - JPL integrates, builds and tests code
  - JPL delivers "production code" to GES DISC
- **Code packages are configuration managed - Harvest CM System**
  - One change per package
  - Packages can be backed-out from the configuration
  - Several packages typically collected into a build
  - Each build is tested/evaluated against standard test dataset
  - Significant builds are "verified" or "validated" with extended test sets and correlative data

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AIRS Software Development Activities - Development Status

- **Version 6 development began on 2007-07-11 w/ V5.1.0.0**
  - Since then more than 45 builds
  - Several significant tests conducted to assess our progress
  - Current V6 development version is V5.7.5.0

    - V5.1.7.0 – 2007-10-02 - MW RTA an tuning
    - V5.3.0.0 – 2008-02-27 – AMSU-A Ch4 quick-fix (still using)
    - V5.3.2.0 – 2008-09-20 – Surface emissivity hinge-points
    - V5.4.0.0 – 2009-01-26 – Improved CO2, surface retrieval
    - V5.4.5.0 – 2009-05-20 – ECMWF as climatology startup opt.
    - V5.4.14.0 – 2009-12-14 – Remove AMSU-A Ch 4,5 from MW and final retrievals
    - V5.5.1.0 – 2010-02-25 – Incorporate SCCN as startup option
    - V5.6.2.0 – 2010-07-19 – Stability parameters
    - V5.6.3.0 – 2010-07-17 – New Regression
    - V5.7.0.0 – 2010-10-07 – MODIS Emissivity, New var. freq. RTA, cloud phase
      ALL MAJOR V6 COMPONENTS!

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AIRS Software Development Activities - Testing Status

- **A few more builds, getting us to another benchmark test:**
  - Testing Start-up options…
  - Cloudy-Regression (tried and true)
  - Climatology (AIRS Climatology)
  - SCCN
    - *ECMWF as startup previously dismissed*
- **Large test dataset produced**
  - Focus Days (24-day cycle) for full mission
  - RaObs Match-ups every 5-days for most of mission
  - Special topical Match-ups: Desert and GPS
- ***Start-up of these efforts affected by system intrusion***
  - Most products produced by April 1

AIRS Computing Facility Status

- **The AIRS Team Leader Science Computing Facility (TLSCF) is aging and reaching "*End of Service Life*"**
  - Maintaining the TLSCF beyond this point is not viable without an upgrade… too expensive!
  - Previous system replacement took place more than 5 years ago
  - It is time to rebuild the TLSCF for the next 5-years.
  - TLSCF System Replacement Plan formulated
    - Leverages technology improvements since last upgrade
    - More compute power and storage
    - No additional energy footprint (electricity and cooling)
  - For hardware, this is a natural evolutionary process!

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AIRS Computing Facility Status - Historical Perspective

- **Pre-launch configuration:**
  - Monolithic SGI processors (2) - approx. 2.2x capability
  - Data archive 500 GB disk archive upgraded to 1.5 TB all with tape backup for data not on disk

- **Current configuration:**
  - TLSCF "cluster" 20 dual processors running Linux (5-10x)
  - Supported and appended to by many quad and eight-CPU systems, enhancement to "science computers"
  - Data archive is now 120 TB and soon to be 200+ TB still with tape backup for data not on disk but… not all data can be housed in the tape enclosure

AIRS Computing Facility Status, cont'd.

- **Future Configuration**
  - Multiple 16+ processors per server unit
  - Data archive consisting of 200 – 500 TB disk
    plus next-generation tape archive system
    inevitably, not all tapes will fit in the tape enclosure
- **First system components have been replaced this year**
  - Other hardware is in-house but not installed due to V6 activity
  - Additional replacements in key systems planned soon
- **Goal is to replace necessary systems without:**
  - Service interruptions
  - Extending the AIRS version improvement schedule
  - Incurring excessive costs overall and in any given year

AIRS/PEATE System Intrusion

- **Hackers broke into AIRS email server on December 12, 2010 - gaining access via recently identified vulnerability in server**
- **2010-12-13 - Intrusion detected by AIRS System Administrator**
  - Affected systems blocked from internet – *good and bad aspects*
    - *blocks hacker access*
    - also cripples desired external network traffic
    - allows local processing to continue
      - 25 servers/workstations affected
      - JPL Computer Security and NASA Inspector General's Office begin investigation
  - 2010-12-14 to 2011-01-03 - forensic analysis
    - Systems pulled 1 by 1 by JPL Computer Security and NASA

AIRS/PEATE System Intrusion (cont'd.)

2010-12-27 to 2011-01-04	Systems are released after forensics (one at a time)
2010-01-07	All affected systems taken offline
2011-01-07 to 2011-01-12	Restoration of critical systems
2011-02-16	Last affected system restored to operation
2011-02-22 to 2011-03-16	Restoration of data archive
2011-02-22	Recommenced data processing operations

AIRS/PEATE System Intrusion

- **Impacted both AIRS and NPP PEATE Projects**
  - 6-week schedule slip for PEATE
    - AIRS schedule impact TBD, but at least 6 weeks
  - Individual work efforts hampered
  - More than 900 work hours expended to rebuild the system!
  - Impact assessment for NASA IG under way
  - Net impact of intrusion estimated to be more than $585,000
- **In the end, since intrusion was almost immediately detected**
  - No loss/corruption of data or source code!
  - Hackers left back-door entry points, SPAMBots, loggers
    - all removed during restoration
  - The original point of access was removed with upgrade patch
  - *Fast action saves the day!*

Back to V6

- We've been working on V6 for quite some time…

- We are close to making some major decisions

and

We are close to moving beyond V6

- There is still much to be accomplished!

Remaining V6 Goals

- **We may have decided which start-up state to use…**

  but other V6 decisions still needed to be made.

  - Further improvement of boundary layer retrievals (?)
  - Refinement of QC and other status information
  - Incorporation of back-end features (cloud properties, trace gases, …)
  - Level 3, possibly new Level 3 climate products

V6 Milestones

- **AIRS Project Goals**
  - V6 must meet AIRS Project's and Science Team's goals
  - Deliver V6 to GES DISC before end of FY11
- **To accomplish this task…**
  - Concur on which start-up state to use
  - Determine whether Key V6 goals have been attained
    - agree to fix what we can
    - possibly defer some goals to V7
- **Milestones…**
  - June 30 - L2 must be completed
  - August 1 - Testing ends
  - August 23 – Documentation and code to GES DISC

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

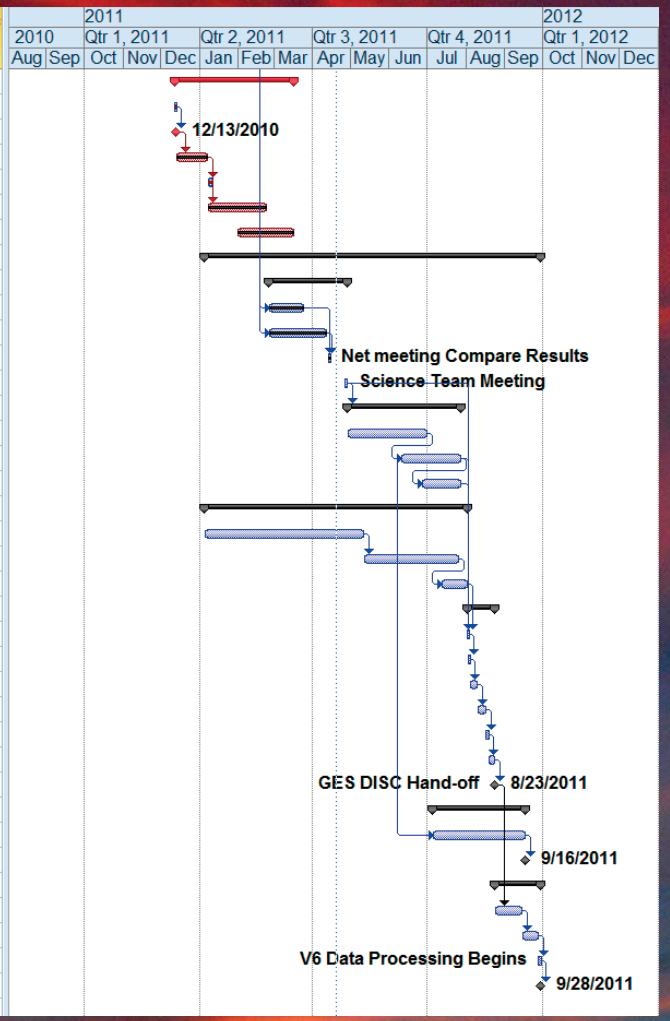
Backup Materials

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The Latest V6 Schedule

	Task Name	Start	Finish	Duration	% Complete	Predecess	R
14	**System Intrusion (Unplanned)**	**Sun 12/12/10**	**Wed 3/16/11**	**64 days**	**100%**		
15	Intrusion	Sun 12/12/10	Sun 12/12/10	1 day	100%		
16	Identification	Mon 12/13/10	Mon 12/13/10	0 days	100%	15	
17	Forensics	Mon 12/13/10	Thu 1/6/11	16 days	100%	16	
18	Systems Unavailable	Fri 1/7/11	Tue 1/11/11	3 days	100%	17	
19	Rebuild Systems	Fri 1/7/11	Tue 2/22/11	31 days	100%	17	
20	Rebuild Archive	Mon 1/31/11	Wed 3/16/11	32 days	100%		
21	**Final V6 Development and Checkout**	**Wed 1/5/11**	**Wed 9/28/11**	**185 days**	**13%**		
22	**L2 PGE Testing - Startup States**	**Fri 2/25/11**	**Thu 4/28/11**	**45 days**	**95%**		
23	Match-Up and Cal Subset	Fri 2/25/11	Thu 3/24/11	20 days	100%	12FS-10 days	
24	Generate L2 Products (extended period)	Fri 2/25/11	Mon 4/11/11	32 days	100%	12FS-10 days	
25	Science Team Net meeting	Thu 4/14/11	Thu 4/14/11	1 day	100%	23,24	
26	Science Team Meeting	Tue 4/26/11	Thu 4/28/11	3 days	0%		
27	**Post Science Team Meeting Activities**	**Fri 4/29/11**	**Wed 7/27/11**	**61 days**	**0%**	**26**	
28	Final L2 Mods	Fri 4/29/11	Thu 6/30/11	44 days	0%		
29	JPL Testing	Fri 6/10/11	Wed 7/27/11	32 days	0%	28FS-15 days	
30	Science Team V6 Checkout	Mon 6/27/11	Wed 7/27/11	21 days	0%	29FS-21 days	
31	**Level 3 Algorithm Development**	**Wed 1/5/11**	**Mon 8/1/11**	**144 days**	**0%**		
32	L3 Concept Development	Wed 1/5/11	Wed 5/11/11	89 days	0%		
33	L3 Coding	Thu 5/12/11	Mon 7/25/11	50 days	0%	32	
34	L3 Testing	Tue 7/12/11	Mon 8/1/11	15 days	0%	33FS-10 days	
35	**Package and Deliver**	**Tue 8/2/11**	**Tue 8/23/11**	**16 days**	**0%**		
36	Delivery Readiness Review	Tue 8/2/11	Tue 8/2/11	1 day	0%	34,26,29,30	
37	V6 Delivery Build	Wed 8/3/11	Wed 8/3/11	1 day	0%	36	
38	Final Checkout	Thu 8/4/11	Tue 8/9/11	4 days	0%	37	
39	V6 Packaging	Wed 8/10/11	Tue 8/16/11	5 days	0%	38	
40	V6 Code Delivery	Wed 8/17/11	Thu 8/18/11	2 days	0%	39	
41	JPL Unpack and Checkout	Fri 8/19/11	Tue 8/23/11	3 days	0%	40	
42	Hand-off to GES DISC	Tue 8/23/11	Tue 8/23/11	0 days	0%	41	
43	**Documentation**	**Tue 7/5/11**	**Fri 9/16/11**	**53 days**	**0%**		
44	Documentation - all testing	Tue 7/5/11	Fri 9/16/11	53 days	0%	29SS+15 days	
45	Deliver Documentation	Fri 9/16/11	Fri 9/16/11	0 days	0%	44	
46	**GES DISC I&T**	**Wed 8/24/11**	**Wed 9/28/11**	**25 days**	**0%**		
47	V6 Integration and Test (GES DISC)	Wed 8/24/11	Wed 9/14/11	15 days	0%	42	
48	V6 End-to End Flow Testing	Thu 9/15/11	Tue 9/27/11	9 days	0%	47	
49	V6 Data Processing Begins	Wed 9/28/11	Wed 9/28/11	1 day	0%	48	
50	V6 Product Support Begins	Wed 9/28/11	Wed 9/28/11	0 days	0%	49	



Upcoming V6 Milestones

- V6 Concluding Timeline
  - JUN 30 - Final "final" L2 V6 coding/mods incorporated
  - JUL 25 - Final L3 coding mods
  - Mid-JUN - Testing and checkout begins (incl. SciTeam)
  - AUG 1 - Testing/checkout ends
  - AUG 9 - Final Build and checkout
  - AUG 23 - Hand-off to GES DISC, code and documents
  - SEP 28 - Public Release V6 Data Products
- *This schedule allows for some minimal adjustments to be made during the testing period. Schedule risk is involved!*