The NOAO Data Management System:
Commissioning the NOAO Data Management System

Howard H. Lanning, Rob Seaman, Chris Smith
National Optical Astronomy Observatory (NOAO)
NOAO Data Products Program

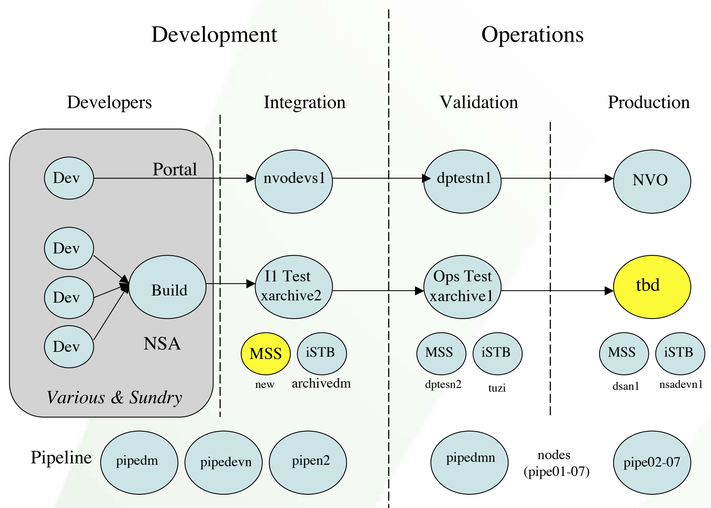
The NOAO Data Management System is comprised of several large subsystems. Its Data Transport System annually conveys terabytes of data between six remote intercontinental sites. The NOAO Science Archive has been safeguarding key NOAO data products for almost five years - NSA release 3.0 has dramatically increased data holdings as well as updated the entire suite of technologies. The NOAO High-Performance Pipeline System addresses the need for scientifically verified pipeline processed data products from major NOAO instrumentation. The NOAO Virtual Observatory Portal is the observatory's keystone VO project.

This integrated, yet highly distributed, system is the result of a large software project known as the "NOAO End-to-End System." E2E involved the development of numerous interfaces and tools requiring careful and thorough review and testing. Extensive test plans were developed to assure that the science and functional requirements of the entire E2E system were met. Integration tests were run by the developers before the individual subsystems were delivered to the Data Products Program Operations Group. Acceptance tests were then run by the Operations staff to ensure the delivered system was ready for commissioning and deployment. Performance tests and scientific verification were done concurrently to assure the resulting data quality of the processed data met their science requirements. Testing of infrastructure and user interfaces was invaluable not only in ensuring that functional requirements were met for the current version, but in developing new requirements for future versions. In short, commissioning is an ongoing process, not a milestone.

INTRODUCTION

- Commissioning the NOAO Archive
  - Large system: 6 distributed systems (Tucson, La Serena, KPNO, CTIO, Cerro Pachon, NCSA)
  - OPS configuration and functional testing (automated scripts and manual verification of fits, etc.)
  - TEST personnel test all subsystems as well as the integrated (E2E) system using test plans, manual testing, automated testing, iterations with Development Team and Acceptance testing to verify product
  - Customer Team user and science evaluation
    - the Operations and Scientific personnel charged with determining features that need to be built and verified

E2E System Deployment



TESTING BENEFITS

- Independent view to software functionality
  - effective testing is careful analysis of the product as well as creating tests/procedures
- Results in improved software quality
- Value-added software testing
  - Customer input required
  - improved user interface
  - not just finding problems; making system more productive for the user/customer

ITERATIVE RELEASE TESTING

- Testers involved in iterative delivery/testing with Development Teams
- Bugs, improvement suggestions, new features, clarifications, etc. filed using JIRA bug tracking system
  http://www.atlassian.com/software/jira/
- Issues addressed/fixed for next delivery/test
- Iterative testing has proved invaluable to Portal quality in preparation for final release
- Similar process being used for NOAO Science Archive (w/ automated FitNesse tests)
  - FitNesse: http://fitnesse.org/

ACCEPTANCE TESTING

- Science and Functional requirements identified (E2E, NSA, NVO Portal, Pipeline)
- Detailed test plans prepared and executed
  - regression tests/procedures
  - science requirement test plans
  - functional requirements test plans
  - performance test plans
- Customer Team Science Verification
- Detailed test reports filed at completion

SCIENCE VERIFICATION AND THE CUSTOMER TEAM

- Customer Team science verification goals
  - testing astrometric and photometric accuracy to assure E2E Specifications are met
  - comment on data quality issues
  - identify nature of problems and ways to improve results
- Customer Team tests
  - inspection of pipeline review pages (PNG graphics, etc.)
  - visual and quantitative inspection of processed FITS images
  - quantitative assessment of observed characteristics

SCIENCE VERIFICATION: PIPELINE PROCESSING OF OBSERVATIONS

- Mosaic pipeline processing to remove instrumental and telescope signatures
  - astrometric & photometric characterization
  - measurement of data quality parameters
- Goals:
  - relative astrometric calibration of 0.5 arcsec (RMS) 90% of time
  - absolute accuracy of 0.5 arcsec 90% of time
  - relative photometric accuracy of 5% (RMS) 90% of time
  - 20% absolute accuracy for BVRI filters

SCIENCE VERIFICATION: QUALITY AND CHARACTERISTICS EVALUATION

- Artifacts (Pupil ghosts, Fringes, Bad pixels, etc.)
- Examination of resampled/reprojected images
- Handling of extended objects, crowded fields, poor observing conditions, etc.
- World Coordinate Systems (astrometric accuracy, internal residuals, 'reference reductions')
- Relative/Absolute photometry evaluation (accuracy, zeropoint, photometric depth)
- Image noise
- PSF FWHM

SCIENTIFIC VERIFICATION: NVO PORTAL

- Verify scientific capabilities and goals are met
  - queries of NOAO Science Archive
  - queries of external archives
  - extraction, download of datafiles, image display
  - VO Plotting, WESIX, ?XMatch
  - ease of use of Portal

- Customer Team evaluation critical to Acceptance of Portal and NOAO Data Management System
  - valuable resource for verification and future improvements to the user interface

SUMMARY

Commissioning a system as extensive as the distributed NOAO Data Management System is critically dependent upon good testing practices and Customer Team input. The process to date has demonstrated the value of such practices in isolating problem issues, improving scientific usefulness and functional capabilities of the product, providing recommendations for enhancement of tools and addition of new features and ultimately providing a more user friendly interface. While testing is never 100% foolproof, serious problems and general issues are addressed and resolved early resulting in a higher quality and more efficient end-product.

III. Test Specifications (*Requirements Testing, Interface Testing, Data Quality Verification*)

{E2E Build Delivered to TEST x/xx/06}

Requirements Doc	No. of Reqs.	Assigned To	PASSED	FAILED	LIENS	Effective Date
• E2E Science Requirements	39 [*8, Portal Only*]	*See Reqs*	8			09/28/06 [*Portal V1.0-beta Only*]
System Software Requirements						
• NVO Portal Functional Requirements	26 [*8 V1.0-beta; 18 Deferred to V1.1 or N/A*]	*See Reqs*	8			09/28/06 [*Portal V1.0-beta Only*]
• Science Archive Functional Requirements	9 (*+5 non-func*)	*See Reqs*				
• Pipeline Functional Requirements	16	*See Reqs*				
Web/GUI Interface Test Plans						
• NVO Portal	...	Lanning			Docs; INT link	09/28/06
• NSA Portal Form	...	Lanning				
Pipeline Test Plan/Science Verification Test Plan						
• Pipeline Processing	...	F. Valdes				
• Science Verification(Customer Team) [*E2E Req #20*]	...	M. Dickinson				
Regression End-to-End Testing						

E2E V1.0 Science Requirements

[*Version: 20 July 2006*]

(+Testing Status)

BLUE = In Test
GREEN = PASSED
RED = FAILED or LIEN

- *Click on Req. # to view detailed test plan*
- *Test Reports will be linked to the Assignee's name*

Req. #	Assigned To	STATUS	Science Requirement Description
1	S. Lowry	TEST	The E2E system will ingest all raw observations (FITS files) generated by NOAO observers using CTIO and KPNO (including partner) telescopes and facility instruments. "Ingest" includes populating the archive with the data and the database tables with the needed information. Goal: visitor instruments.
2	R. Seaman	TEST	The E2E system will provide access to the survey data products that currently reside in NSA R2 and will provide access to similar NOAO survey data products as acquired in the future.
3	F. Valdes	TEST	The E2E system will be delivered with the raw data in DCI storage that have accumulated up to that time.
4	T. Cline	TEST	The E2E system will provide raw data and metadata to the NOAO pipelines and ingest data products generated by NOAO pipelines.
5	Lanning	TEST	The E2E system will be able to identify data objects within its holdings that match queries delivered to the archive. The availability of such metadata will be subject to approved policies.
6	N. Zarate	TEST	The E2E system will allow for users to download, via the internet, selected data identified through such queries. The availability of such data will be subject to approved policies.
7	C. Miller	TEST	External queries and external data transport will be accomplished through VO-approved standard protocols, where they exist.
			The E2E system will support queries of and return values for the following parameters:

IV. Procedure: Step-by-step flow through Portal Operations

Throughout, the tester may directly access the test database to view content (tables/relations/fields) verifying the datasets to be retrieved, etc.) The tester may also view files/configuration on the specified host noted above. Currently, the temporary cache is identified as the anonymous ftp area.

A. Public Access Features

- **SKY**
  - **PAN:**
    - Click and hold mouse to pan around sky image (in the SKY pane)
  - **ZOOM:**
    - Select zoom level(s); verify field scale change
    - *Currently takes ~15sec to display new zoom level (7/26/06) - network between Chile and Tucson*
  - **Data Available:**
    - Double-click on wire-frame image to view list of data available at that position of sky. **NOTE:** zoom level must be set to '**5**' or higher in most cases

  [NOTE: ALL COORDS MUST BE IN DECIMAL DEG]
  - Direct Entry
    - Select '***Add a Marker to NOAOSky***' in the window pane to left of SKY pane to mark on sky image for requested source position
    - Input RA/DEC (decimal degrees; comma separated format)
    - Click '**Add**'
    - Verify red '**X**' marker at desired location on SKY pane
  - File Upload
    - Create file (*in your local area*) containing list of coordinates (1 or more)
    - Select '***Browse***'
    - Select coordinate list file
    - Click on '**U**' to Upload the file
    - Coordinate(s) will appear in the entry field below
    - Select desired coordinate (click on coord)
    - Select '**Add**' to add a Marker for that (or those) coordinate(s) selected
  - **List of Flags (RA, Dec, [TIME])**
    - Note entry in '**List of Flags (RA, Dec, [TIME])**' after **Add**ing the Marker
    - Verify Marker placed in desired location on SKY pane
    - **Remove** marker
      - Click on coord flag in list to be removed
      - Select '**Remove**'

The National Optical Astronomy Observatory (NOAO) is operated by the Association of Universities for Research in Astronomy (AURA), Inc. under cooperative agreement with the National Science Foundation.