



MAST
Users
Group
Meeting

Dec 2,
2014

The HST Online Cache: Changes to the User Experience

Karen Levay

Faith Abney

Mark Kyprianou



Current Process for HST data

- User submits request for data
 - On the Fly Reprocessing for active instruments
 - “Static” archive for legacy instruments
- All requests submitted are batch requests via “DADS”
 - Requests for legacy instrument data is quick as the request is just copying data to stage or to user’s platform
 - All other data is completely reprocessed even if only flt files requested and delivered. All local copies of the processed data are deleted.

The screenshot displays the 'HST Search Form' interface. At the top, it includes 'Archive Status', 'HST Search Form', and 'Help Field Descriptions'. Below this are tabs for 'Standard Form' and 'File Upload Form'. The main search area contains fields for 'Target Name', 'Right Ascension', 'Declination', 'Resolver', 'Radius (arcmin)', and 'Equinox'. There are also sections for 'Imagers', 'Spectrographs', and 'Other' with various instrument filters. The bottom section, 'Retrieval Options', includes 'User Information' (anonymous or archive user), 'Delivery Options' (staging area, DVD, or FTP), and 'File Options' (file types and extensions). A 'Send retrieval request to ST-DADS' button is at the bottom.



MAST
Users
Group
Meeting

Dec 2,
2014

Future HST Online Cache

- All data will be stored online.
- Public data will be url accessible.
- Proprietary data will be protected. Initially not available via browser/ftp or through the portal.
 - Later – maybe next year, as the archive becomes comfortable with SSO and data protections, the proprietary data may become available through the portal and via proxy with proper authentication.
- Data for active instruments will be reprocessed as new calibrations / software are introduced.



MAST
Users
Group
Meeting

Dec 2,
2014

Classic HST search and DADS retrieval behavior remain the same

- Users will still be able to use the familiar HST search and batch retrieval interfaces, but should have faster response time for active instruments, as the data will be copied from the online-cache.
- The batch request will appear to be the same but retrieve all data from the “online cache” and deliver to stage or to the users platform as designated. Hard media (DVDs, CDs) will remain an option for HST retrievals.



... but there will be new options

- Data that are in the HST public online cache will be directly available via the MAST Discovery Portal and the download basket as the Common Archive Observation Model (CAOM) database is updated.

The screenshot shows the MAST Discovery Portal interface. The search results table is as follows:

	Actions	Preview	Mission	Instrument	Filters	Waveband	Target Name	Observation ID
9			HST	NICMOS/NIC2	detection		3C265	HST_7454_01_NIC_NIC
10			HST	NICMOS/NIC2	F160W	INFRARED	3C265	N48301020
11			HST	STIS/NUV-M...	MIRNUV	UV	3C265	O6GK08020
12			HST	STIS/NUV-M...	MIRNUV	UV	3C265	O6GK0806Q

The interface also includes a left sidebar with filters for Product Type, Mission, Instrument, and Waveband. The right sidebar shows the AstroView panel with a star field and observation coordinates: RA DEC 11:45:04.806 +31:22:18.96.



... but there will be new options

- Download basket options include:
 - Downloading a tar or zip files of data through the browser (must be < 5 GB)
 - Creation of cURL and wget scripts so users can download to platform of choice in the background
- If a user knows the specific dataset of interest, the data are easy to find in the public online cache. Users will be able to easily build programmatic scripts to download public data from the cache.

Search: 3C265

Examples: M80, 13:29:56.47:13:50 r=1m, More Examples..., Random Search

30 files

Download Basket

13.63 MB Selected / 5120 MB Max

24:02.89
33:49.43

Displaying 10 of 30 Total Files from 2 Datasets

Mission	Product Type	Observation ID	Description	Project
1 HST	image	O6GK08DTQ	DADS FLT file - Calibrated exposure AC...	DADS
2 HST	image	O6GK08DTQ	HSTonline fit FITS compressed - Calibr...	HSTonline
		O6GK08DTQ	DADS JIF file - histogram of jitter data	DADS
		O6GK08DTQ	DADS JIT file - table of jitter data	DADS
		O6GK08DTQ	DADS TRL file - Processing log	DADS
		SGK08E6Q	DADS FLT file - Calibrated exposure AC...	DADS
		SGK08E6Q	HSTonline fit FITS compressed - Calibr...	HSTonline
		SGK08E6Q	DADS JIF file - histogram of jitter data	DADS
		SGK08E6Q	DADS JIT file - table of jitter data	DADS
		SGK08E6Q	DADS TRL file - Processing log	DADS

Download Files As...

File Name*: MAST_2014-11-18T1629

Format: tar.gz

Remove completed files: tar.gz zip cURL wget

*Note: The filename MAST_2014-11-18T1629 will be used to where you want to save. Your browser will prompt you for the path when the download begins.

Product Group ID

Order Values by Count

2000919781 (0 of 4)
2000919782 (0 of 1)
(0 of 2)
(0 of 8)
(1 of 2)
(0 of 8)
(0 of 2)

Observation ID

Order Values by Count

O6GK08DTQ (0 of 2)
O6GK08E6Q (0 of 43)
(0 of 28)
(2 of 2)

Description

Order Values by Count

DADS FLT file - Calibrated exposure ACS,WFC3,STIS,COS (2 of 2)
DADS JIF file - histogram of jitter data (2 of 2)
DADS JIT file - table of jitter data (2 of 2)
DADS RAW file - Raw exposure (0 of 2)



MAST
Users
Group
Meeting

Dec 2,
2014

Over the next 2-3 years some new features may be introduced.

- Some of these features are primarily being developed for future JWST use and the HST/MAST data and interfaces will be used as a test bed. This is the reason for the longer development time-line.
- Logged-in users may be able to use a set of “subscription” services, e.g. notification when data become public, notification when data are reprocessed.... .



MAST
Users
Group
Meeting

Dec 2,
2014

Over the next 2-3 years some new features will be introduced.

- Users will be told when datasets of interest are in the queue for reprocessing.
- Users who are logged in and authorized to view specific proprietary datasets, will be able to see previews of proprietary data and to request them through the download basket



Timeline (1/3)

- Dec 2014
 - Population of the online cache begins with static, legacy instrument data
 - CAOM updates required for access to the online cache through the MAST Discovery portal
 - Implementation of the Owl/Condor workflow replacing OPUS pipelines for several functions



Timeline (2/3)

- Jan – May 2015
 - Processing to fill cache for active instruments begins.
 - CAOM population and preview population done via operational tools.
 - As CAOM is updated, the newly processed data will become available through the MAST discovery portal.
 - New previews will be created. HLA preview software being used for image data. New spectral plotting software for spectral data.
 - MAST discovery portal search options expand beyond spatial searches.
 - OTFR remains in place for data not yet in the online-cache.



Timeline (3/3)

- March 2015
 - STScI Single Sign-on (SSO) implemented for archive users. Initially will affect the standard retrievals. Planning very extensive public announcement and mailings about how the changes will affect users.
- June 2015
 - Replacing the OPUS “pre-archive” pipelines with Condor/Owl workflows.
 - CAOM updates and preview creation part of the Data Processing workflow.