

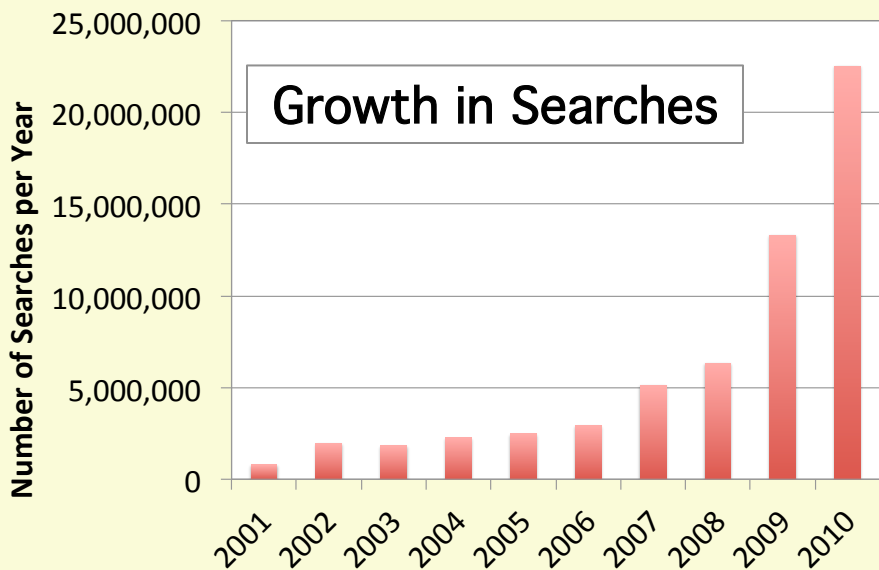
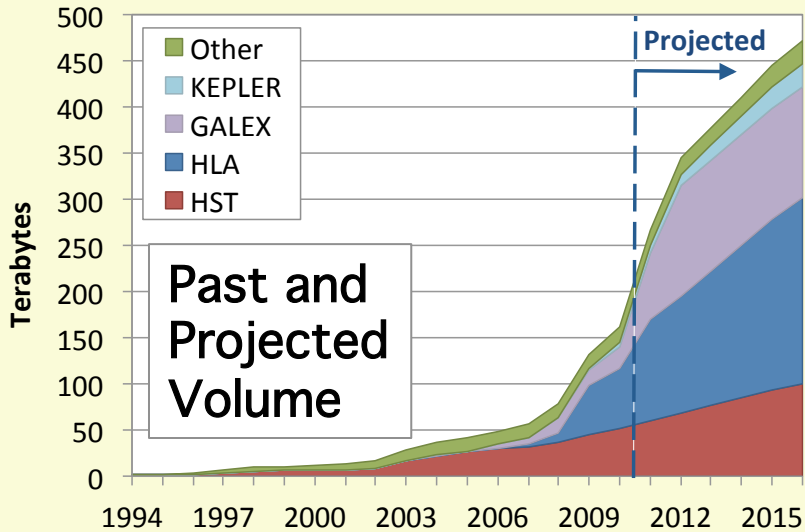


Highlights of MAST Activities

Rick White & Karen Levay

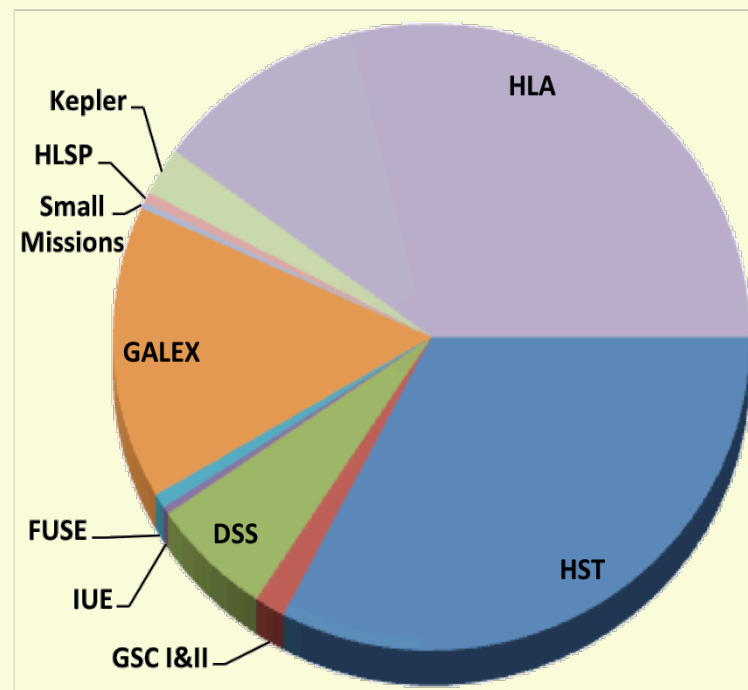
MAST Users Group

2011 September 23



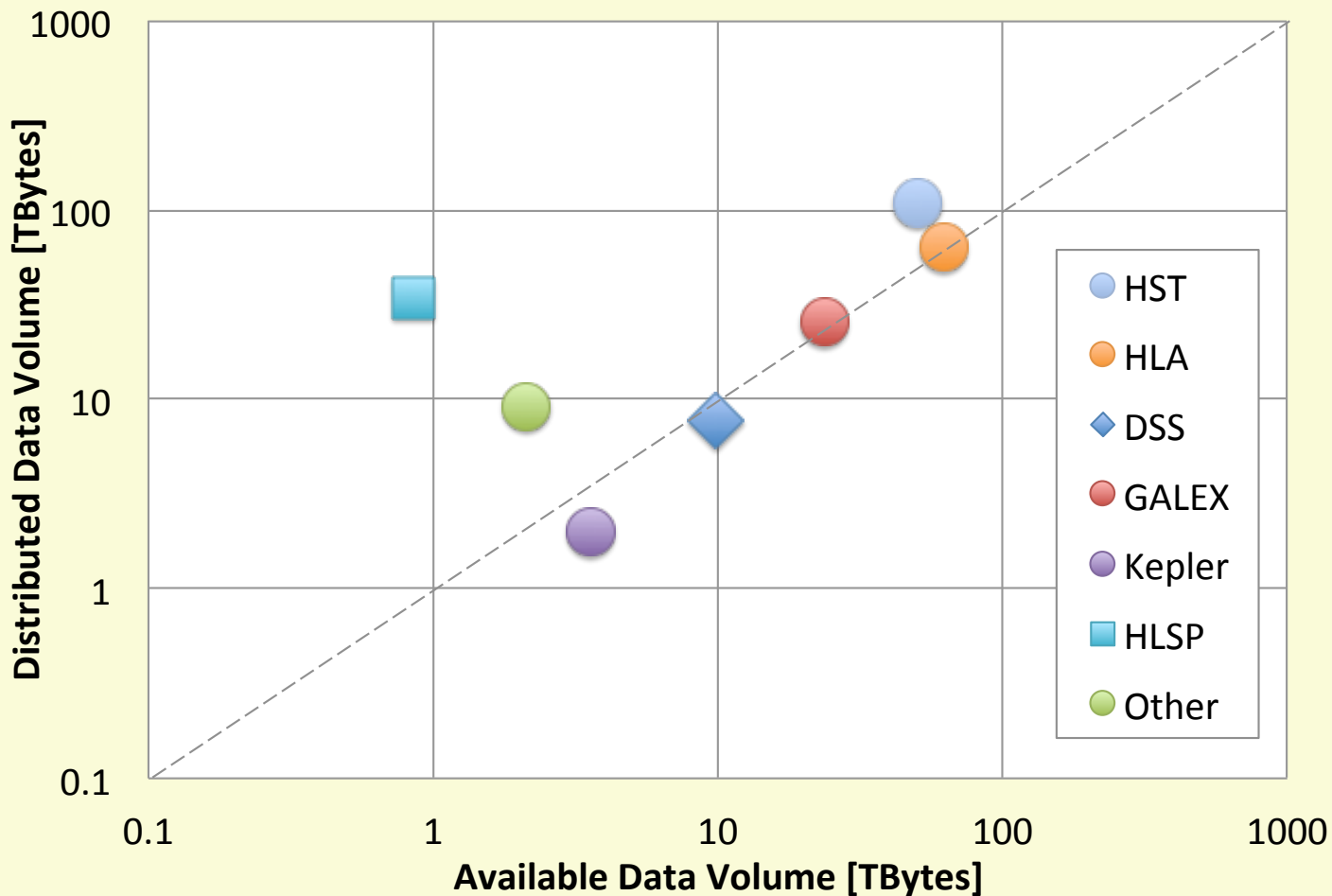
MAST Data & Usage Growth

Current Data Holdings



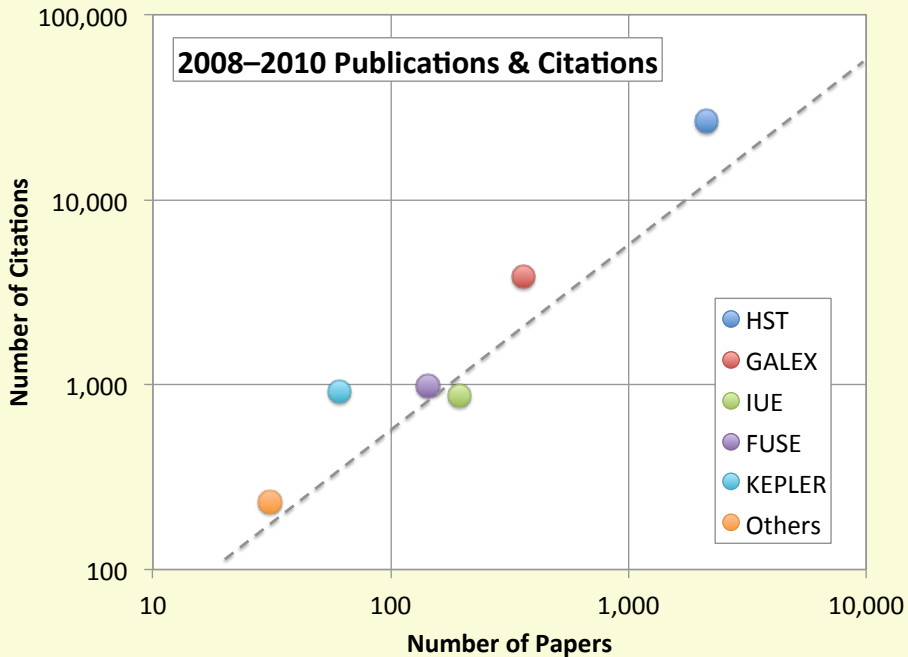


Distribution (2008–2010) vs. Data Holdings



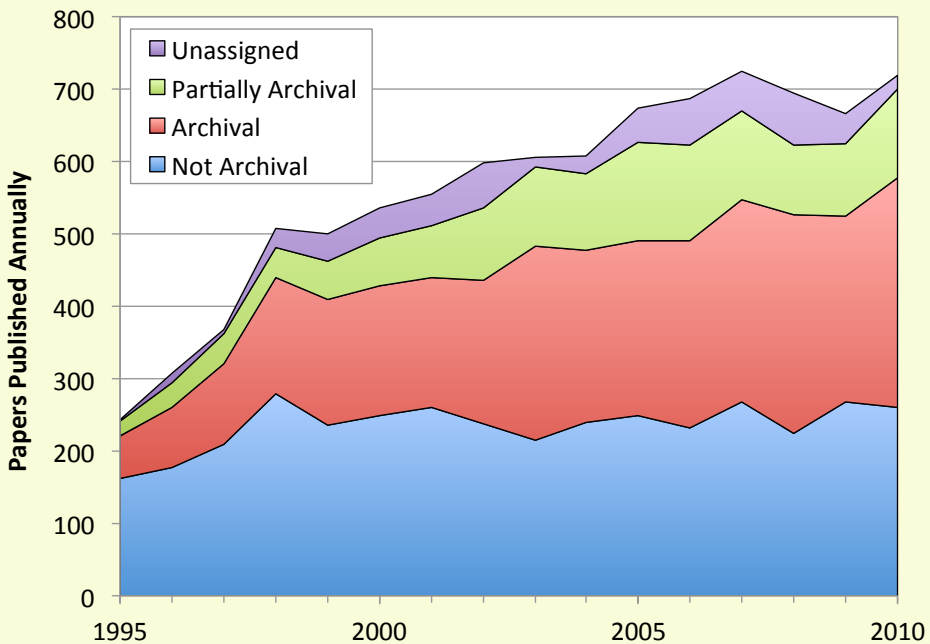


2008–2010 Publications & Citations



MAST Publications & Citations

HST Archival Publication Growth





Archive Center

- “**Archive center**” organization was established to help unify various STScI archive projects (Dec 2008)
 - **Carl Johnson** (project manager), **Gretchen Greene** (chief engineer), **Rick White** (archive center scientist) are triumvirate coordinating group activities
 - Includes:
 - Archive Sciences Branch (**K. Levay**)
 - Data Systems Branch (**M. Kyprianou**)
 - Operations (**F. Abney**)
 - Science Software Branch (**P. Greenfield**)
 - IT staff



MAST/Archive Science Branch Recent Staffing Changes

- Additions:
 - **Tom Donaldson** (software engineer, September 2010)
 - **Steve Handy** (software engineer, January 2011)
 - **Geoff Wallace** (software engineer, September 2011)
- Departures:
 - **Quyín Fan** (software engineer, arrived November 2010, departed July 2011)



MAST/Archive Center Staff

ASB Branch

Alberto Conti	Maria Nieto-Santisteban
Tom Donaldson	Tony Rogers
Theresa Dower	Bernie Shiao
Steve Handy	Myron Smith
Tim Kimball	Randy Thompson
Charlie Loomis	Shui-Ay Tseng
Brian McLean	Geoff Wallace
Karen Levay (lead)	

Other MAST/HLA Staff

Stefano Casertano	Mark Kyprianou
Dorothy Fraquelli	Steve Lubow
Niall Gaffney	Lee Quick
Gretchen Greene	Rick White
Anton Koekemoer	Michael Wolfe

Also collaborators outside STScI, principally at CADC and ST-European Coordinating Facility

Archive staff list (including ops & dev):

<http://archive.stsci.edu/staff.html>



2010–2011 Hubble Highlights

- HST [*M.Kyprianou, F.Abney*]
 - Change to Linux improved processing/delivery speeds
 - Hubble Multi-Cycle Treasury programs underway
 - CANDELS: Cosmic Assembly Near-IR Deep Extragalactic Legacy Survey
 - CLASH: Cluster Lensing and Supernova Survey with Hubble
 - PHAT: Pan-chromatic Hubble Andromeda Survey



Space Telescope-European Coordinating Facility (ST-ECF) closed in December 2010

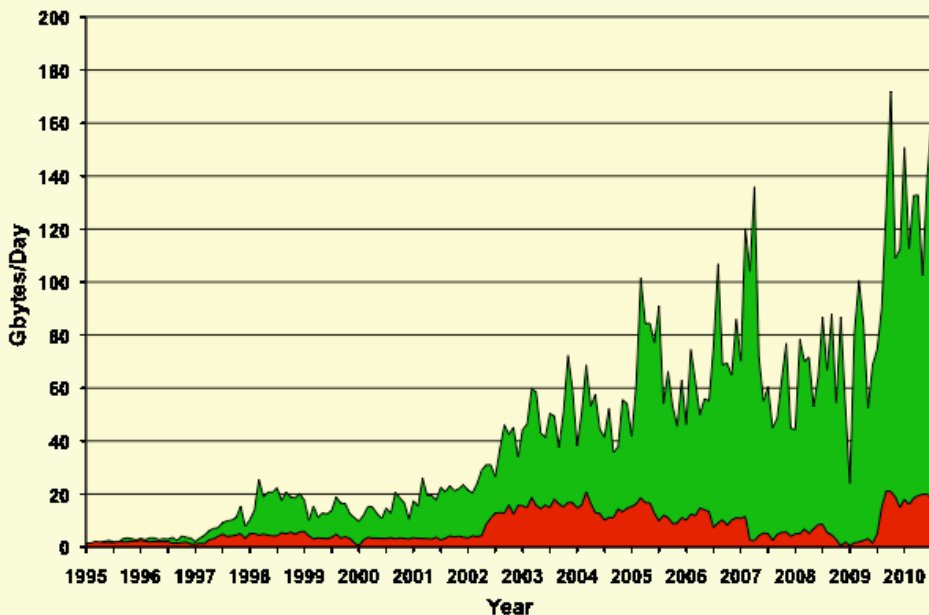
- European Space Astronomy Center (ESAC, Spain) will operate European mirror of Hubble archive, but will not have scientific expertise of ST-ECF



HST Data Ingest and Retrievals

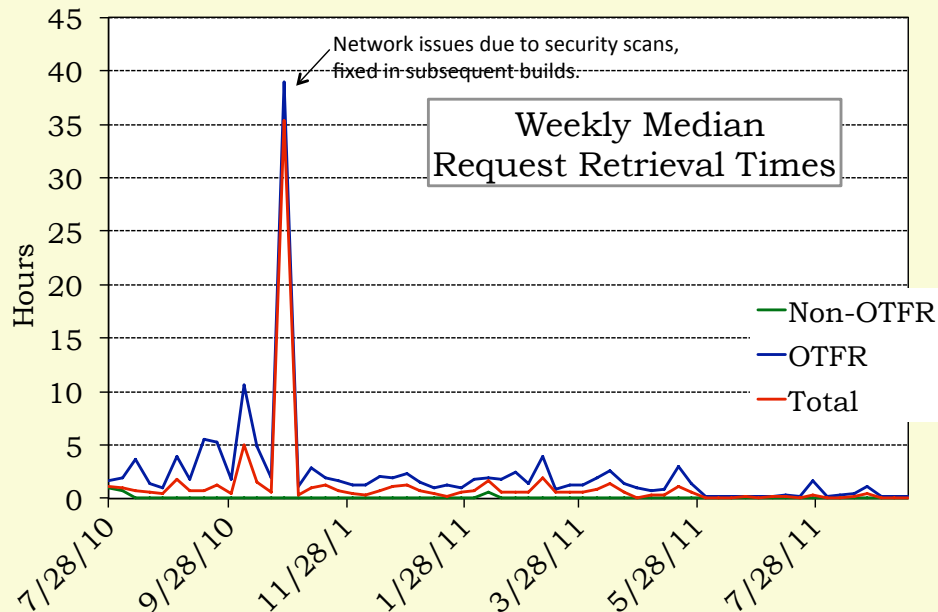
HST & FUSE Data Archive

Retrievals Ingest



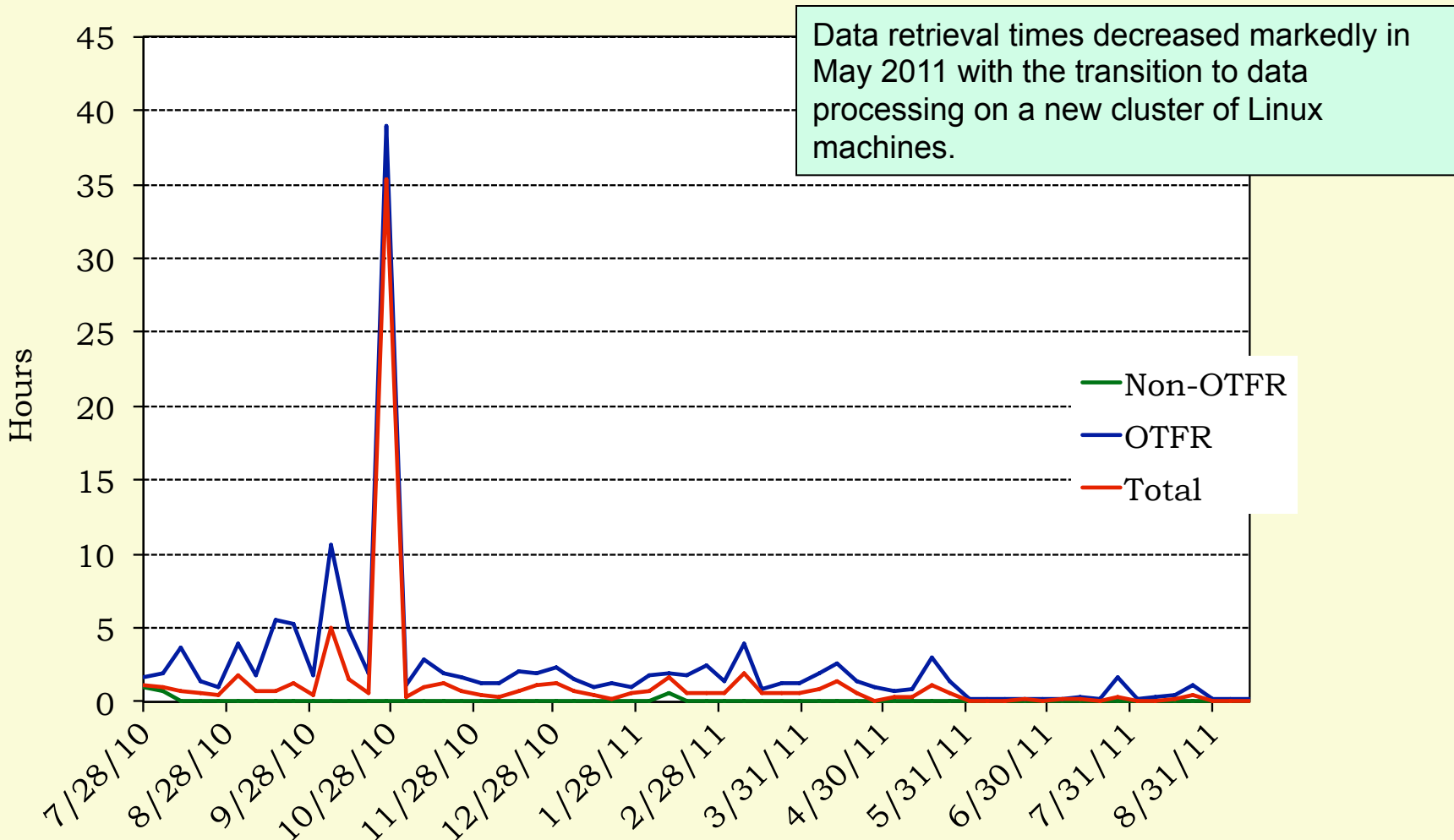
Both HST data volume into archive (red) and volume distributed (green) continue to grow.

Average time needed to fulfill requests for OTFR much improved.



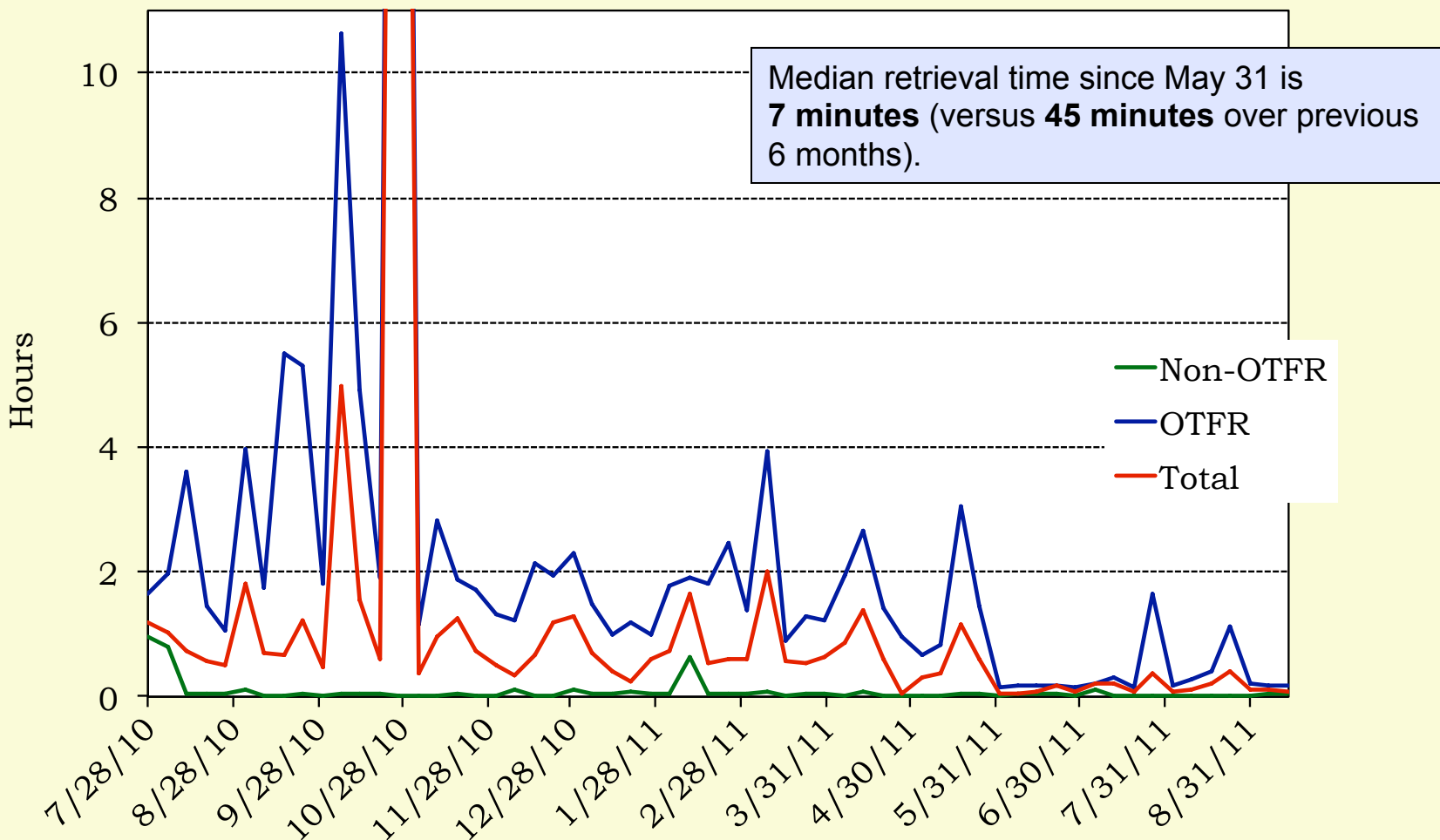


HST Weekly Median Retrieval Times (1 year)





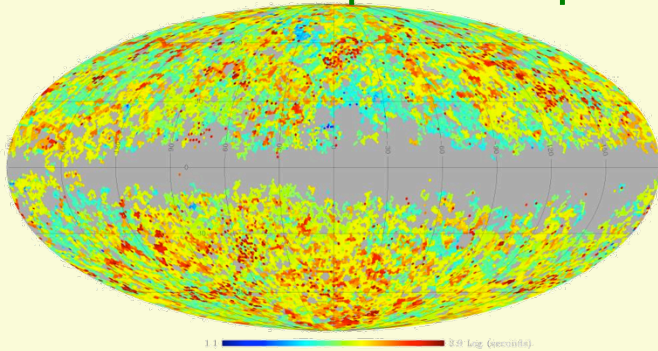
HST Weekly Median Retrieval Times (1 year)





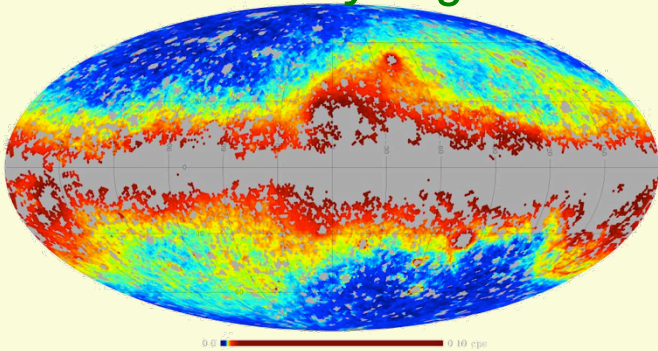
2010–2011 GALEX Highlights

GCAT NUV Exposure Map



- Satellite operations to end Sept. 30, 2011
[M.Smith]
 - Limited GALEX team support in FY12
 - MAST & GALEX teams working to preserve as much mission closeout information as possible for final archive
 - MAST will acquire GALEX photon list data
- GR6 data release completed in Nov. 2010
- SDSS DR7 to GALEX GR6 cross-match
 - Available through CasJobs

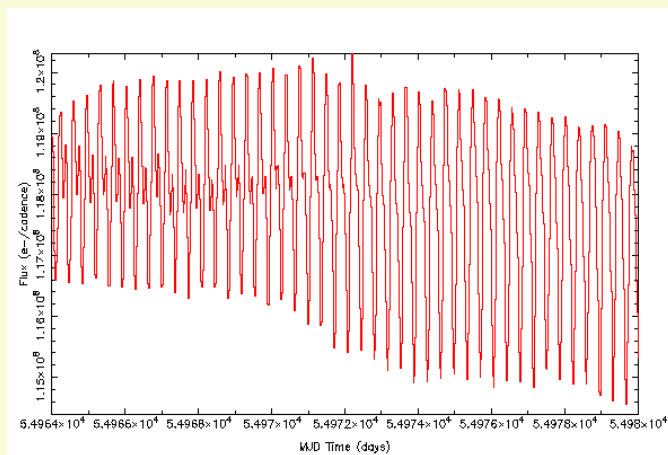
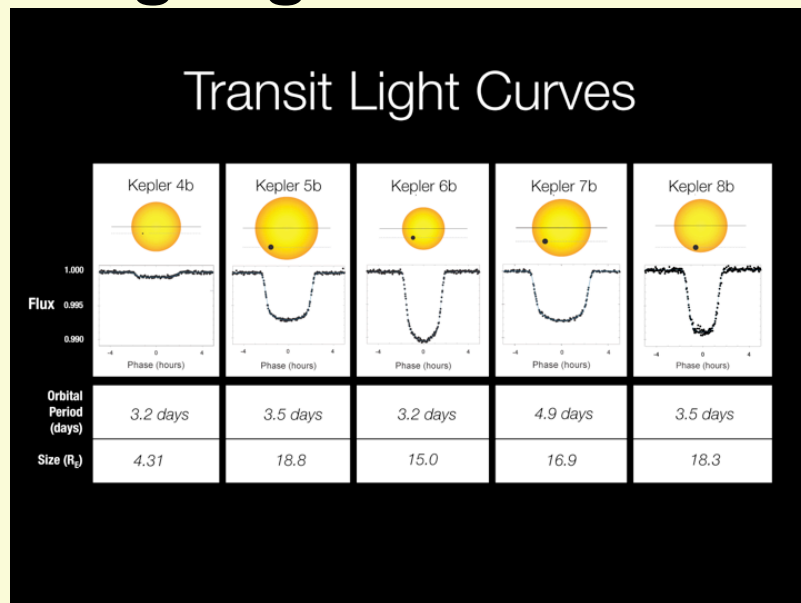
GCAT NUV Sky Brightness





2010–2011 Kepler Highlights

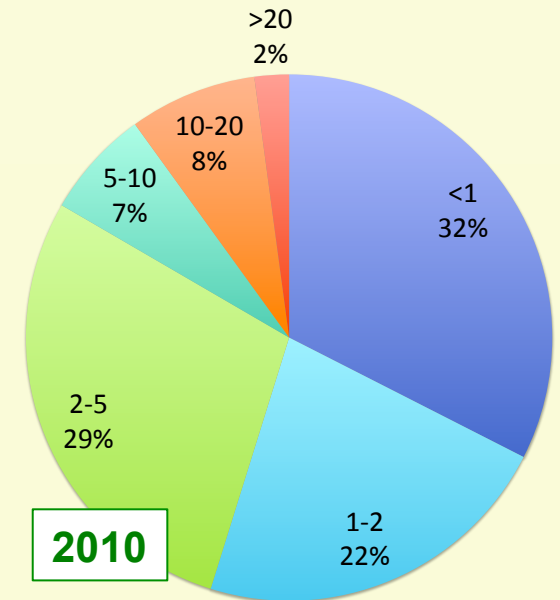
- Kepler is more than halfway through its nominal 3.5 year lifetime
[D.Fraquelli & M.Smith]
 - Launched March 6, 2009
- 549,049 public light curves!
 - Available for bulk downloads
- 645,668 searches & 16.7 TB distributed to date!
- Many new data products including target pixel files, Kepler/GALEX cross-match, Kepler/USNO cross-match



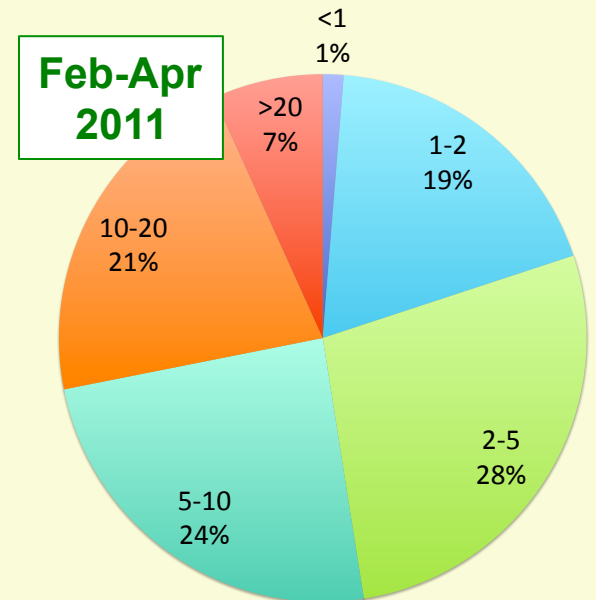


Internet Upgrade

- STScI Internet connection was upgraded from 100 Mbit/s to 500 Mbit/s in February
 - Extensible to 1 Gbit/s (for \$\$)
 - Includes Internet-2 access
- Much improved performance
 - Reduced loading so there is capacity to respond to requests
 - Factor of >2 improvement in average bandwidth for large transfers
 - Low transfer rates are much less common
 - Peak bandwidth also doubled to >100 Mbit/s



Download Rates (Mbit/s)





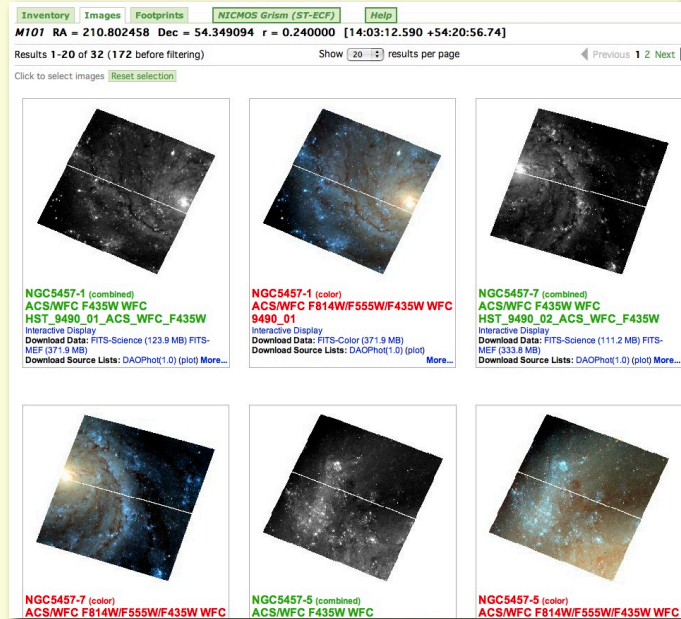
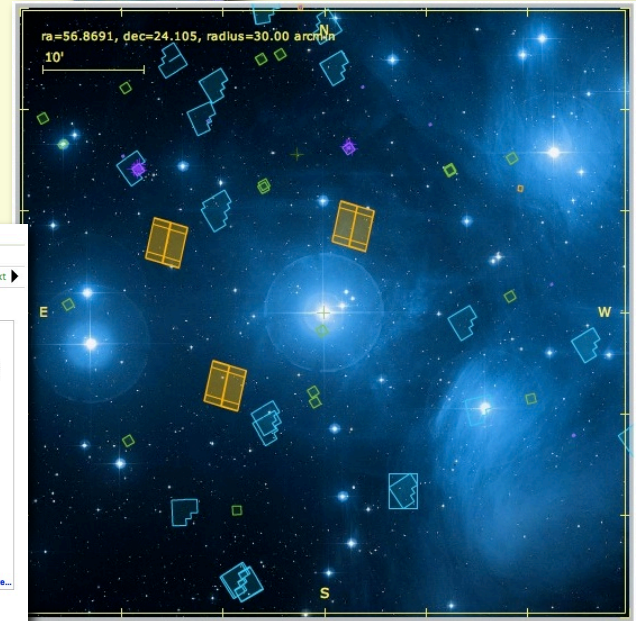
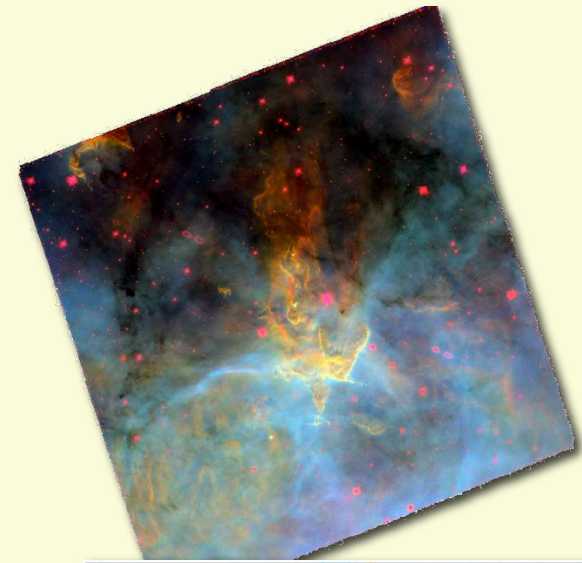
More 2010–2011 Highlights

- Working to incorporate Swift UVOT data in MAST
 - Multi-filter UV/optical data, similar to XMM-OM
 - Collaboration with HEASARC and Swift UVOT team (Penn State)
- New MAST portal in development
[T.Donaldson & T.Rogers]
 - Will provide new unified user interface for all missions
 - Sharing infrastructure with VAO portal
- Much progress on Common Archive Observation Model
 - Unifies MAST mission data
 - Jointly developed with CADC



Hubble Legacy Archive

- Hubble Legacy Archive (HLA) DR5 in March 2011 [*S. Casertano*]
 - New footprint interface
 - Popup info for catalog sources in interactive display
 - New ACS, WFPC2; WFC3 in progress
 - <http://hla.stsci.edu>





2010–2011 HLSP Highlights

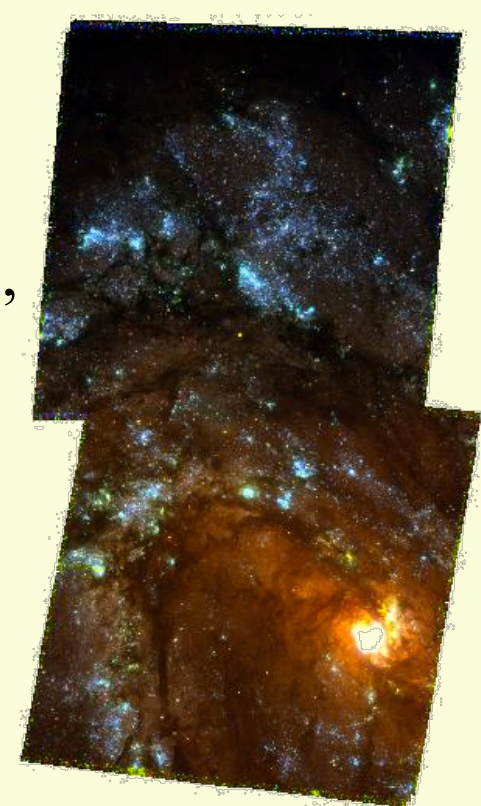
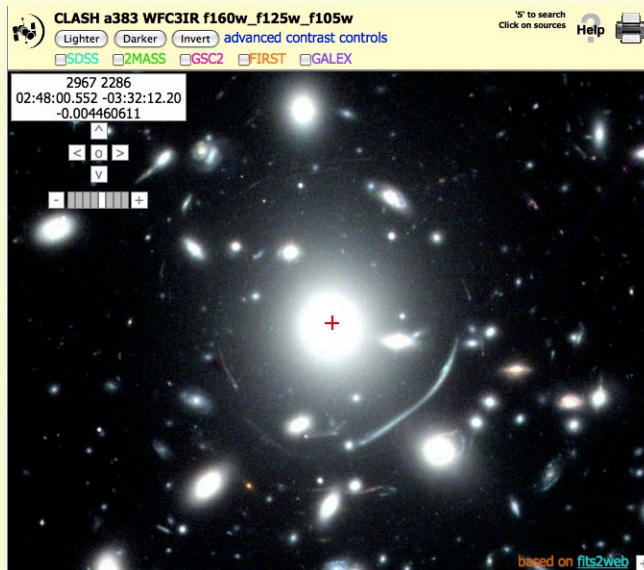
New High Level Science Products

- Coma Cluster, CANDELS (7 releases!), CLASH (5 clusters), WFC3 ERS, GALEX Hot-Star Candidates, ACS Globular Clusters, HIPPIES, HUDF09

Globular Cluster M13
ACS/WFPC2



CLASH Abell 383
WFC3 IR F160W/F125W/F105W



M83 WFC3
F438W/F336W/F225W



Some Current/Planned Activities

- Cross-mission
 - Storage broker for integrated file storage
 - MAST Portal (ongoing)
 - Common Archive Observation Model
 - Includes integrating all MAST missions into footprint database (that work is nearly complete)
 - Extend CasJobs database access (previously GALEX-only) to other MAST missions
 - Moving Target Database/Interface
 - Starting to think about interface; may use the VO SkyBoT tool to identify observations
 - Input from Max Mutchler/Alberto Conti HST archival project and from ST-ECF legacy database work



Some Current/Planned Activities

- HST/DADS [*M.Kyprianou, F.Abney*]
 - “Future of OPUS” project to replace pipeline software (selected Condor as workflow manager)
 - Move on-the-fly-reprocessing (OTFR) to the background so retrievals are much faster
 - ACS CTE correction being incorporated into CALACS pipeline
 - WFC3/IR files corrected for image persistence are being generated and will initially be distributed through the HLSP system



Some Current/Planned Activities

- Kepler
 - New tools for light curve browsing and manipulation
- GALEX
 - Closeout activities and photon lists
 - Note activity is heavily concentrated in the coming months
- JWST [*G. Greene*]
 - Early efforts are focused on design, integration & test data



2010 MUG Report Followup (1)

- *Could astroview be used for HLA-style viewing in the portal? Need to be sure there is not duplication of effort here.*
 - The technologies to be used for the portal image display tool are still under discussion, but astroview is one of the contenders. It is possible that an HTML5/Javascript version of the interface could both have good performance and be portable to more platforms than the Flash-based astroview. A GALEX interface using astroview is currently in development.
- *It would be useful to have some kind of a "depth" map that shows not just how many exposures there are at a point, but how deep it goes.*
 - The newest HLA image pipeline (for WFC3) includes information that should make computing such a depth map straightforward. An interesting question is how to store this so that it can be used in database searches (rather than simply being another image that is available). An enhanced footprint database might be able to incorporate depth information. We are still thinking about this issue and agree it would be very helpful to users.



2010 MUG Report Followup (2)

- *It will be important to keep the low-level (single-image) data available in the new (portal) interface.*
 - We agree that this is a good idea. The data model (and the current HLA database) both include the required links between exposures and composite data products. So the information is available, but we don't currently have an interface designed to display them in an easy-to-use form. We do plan to include this in the portal interface, but exactly how it will work from the user's point of view is TBD. Suggestions from the MUG on possible designs for this interface are welcome.
- *Provide links to associated high-level science products in searches for MAST data.*
 - We are already doing this in the standard MAST interface (but not in the HLA interface). We will plan to incorporate these links in the data model and in the portal interface. Note that this is coupled with the question above about single-image data; from the user's point of view, a HLSP combined dataset is just another way to use the exposure level data and should be linked in the same way.



2010 MUG Report Followup (3)

- *Is MAST prepared to meet the JWST ramp-up in terms of monetary and people resources?*
 - In the near-term, the costs of JWST archive development should be fully covered by the JWST contract. MAST will work to ensure that the JWST development proceeds in a way consistent with our interfaces and missions. As an example, the JWST mission is already working on mapping its data into the MAST data model (CAOM). That effort is paid for by JWST and in fact is fed back into the MAST data model development to make sure the model is able to handle the data produced by the more complex JWST instruments.
- *Is MAST/HST planning on a point-source catalog of objects detected in archive images?*
 - Yes, this is part of the planned work for the HLA over the next couple of years. We intend to develop an "all HST-sky" object catalog that links the sources detected in different images and filters and that is searchable to find objects of interest in HST images. This is a challenging project but will indeed provide a great benefit to users.



2010 MUG Report Followup (4)

- *Documentation and working examples for CasJobs should be provided for novices.*
 - We plan to do this going forward as part of our continuing CasJobs enhancements. The Senior Review proposal also includes a novice interface to CasJobs that works by providing automatically generated, editable CasJobs queries that are built using the standard MAST search forms. That will allow users to quickly generate a query that does roughly what is needed as a starting point for CasJobs.
 - We also plan to provide small database subsets that give fast turn around while debugging queries. Once the queries are working they can then be submitted as longer jobs on the entire DB.
- *Should MAST have its own dedicated internet line to the outside that does not share or compete with the rest of the Institute?*
 - We upgraded the STScI internet connection this year, and the archive has greatly benefited from the improved bandwidth. A substantial fraction of all the data moving over our external network goes to & from the archive. But we will keep this suggestion in mind as we consider future improvements to the archive bandwidth.



2010 MUG Report Followup (5)

- *The MUG suggested some ways to increase the responses to the MAST user survey.*
 - The response rate was vastly better this year, as we had the largest number of responses ever. This is attributed both to a broader email list and to late August & early September being a better time than June for surveys.
- *We suggest a MAST specific documentation page that links to the latest and greatest documents.*
 - Unfortunately documentation improvements have been neglected this year due to a lack of resources. We do agree this would be helpful to users.
- *Additional public presence for MAST would be helpful in terms of advertisement and as a way to get direct user feedback. Send MAST members out to the community to give how-to colloquia. The MUG encourages an archive workshop.*
 - MAST and the HLA had an increased presence at the STScI booth at both the January and June 2011 AAS meetings and plan to continue this. We are also doing a demo at the November 2011 ADASS meeting.
 - We are in active discussions with the other NASA archive centers about a joint archive workshop that would include all the NASA centers. MAST is taking the lead on this discussion, but it is not yet clear if/when the workshop will take place.



MAST 2011 Senior Review

- NASA Senior Review
 - MAST and other NASA archive centers are reviewed by NASA HQ every 3 years to set funding
 - Used to be every 4 years; this is the first time for a review after only 3 years
 - Proposal includes a 5-year plan, but the next review will be in 3 years (2014)
 - Proposal submitted April 8, presentation to the review committee May 18
 - All contributed to proposal
 - White, Casertano, Christian, McLean at presentation





Highlights of Senior Review Proposal

- Support active missions (HST, GALEX, Kepler, XMM-OM) and projects (HLA, Portal, CAOM)
- Cross-mission products & tools:
 - Search and explore time-domain data (Kepler, HST, GALEX)
 - Portal including HLA capabilities and more
 - Science keyword index for archive
 - ST-ECF “One-box” (<http://archive.eso.org/archive/hst/search/>)
 - Simpler interface for large database queries
 - Advanced plotting and display tools in browser



Highlights of Senior Review Proposal (cont.)

- Mission-specific projects:
 - GALEX:
 - Final archive closeout
 - Incorporate photon lists (*requires additional funding*)
 - Kepler:
 - Store variability info in database
 - Develop new tools for Kepler time-series data
 - HLA:
 - New imaging and spectral data products
 - All-HST sky object catalog
 - Incremental data releases with improved quality control
 - Combined images for moving targets
 - High-resolution images (*requires additional funding*)
 - Distribute HLA tools to the community



Results of Senior Review

- Senior review panel report was released on June 27
 - https://science.nasa.gov/media/medialibrary/2011/08/10/ApArchSR_2011report_final.pdf
- Results for MAST were outstanding: Tied for first (with ADS, HEASARC) with Excellent ranking

Archive Center	Ranking
ADS	Excellent
HEASARC	Excellent
MAST	Excellent
IRSA	Very Good
NED	Very Good
NStED	Fair



Quotes from Senior Review Report

“The panel found the first three centers (ADS, HEASARC, MAST) to be at essentially the same, excellent level of performance and merit, so ranked them identically. These were very strong proposals that fully responded to the Senior Review Call, and contained no major weaknesses.”

“MAST is a lean operation that delivers substantial scientific returns for a relatively small investment.”

“The archive has been extremely responsive to input from its user community, improving access to the data through an increase in network bandwidth and by developing a common interface to the MAST data.”



(Minor) Weaknesses Noted in Report

“The proposal did not address the question of how the institutional knowledge associated with the NASA missions that are, or will be, in close-out can be preserved by the MAST team. [...] In particular, the team needs to consider long term planning for Kepler information and instrument expertise.”

“[It] was felt by the panel that the proposal would have been improved by a discussion of a long-term strategic plan for the archive including the long-term preservation and distribution of data and the expected evolution in technological resources.”



Overall Recommendations

“The panel recommends continued support for MAST through its in-guide budget. We consider MAST and its services an indispensable part of the NASA data archives. While we did not feel that MAST made a strong scientific case for its over-guide requests, the panel believes that the potential loss of the GALEX photon data is of sufficient importance that the request for over-guide funds to support preservation of the photon data was warranted. We do, however, consider that the over-guide request for support for high-resolution [HLA] images to be of low priority.”

Bottom line: The MAST funding prospects are healthy for the next 3+ years, which is very good news.