

INFRASTRUCTURE

At the GHRC DAAC

Will Ellett

IT Manager

sysadmin@itsc.uah.edu

Support: Michele Garrett, Michael McEniry, Jason Toone

Presented at the GHRC User Working Group Meeting
September 25-26, 2014



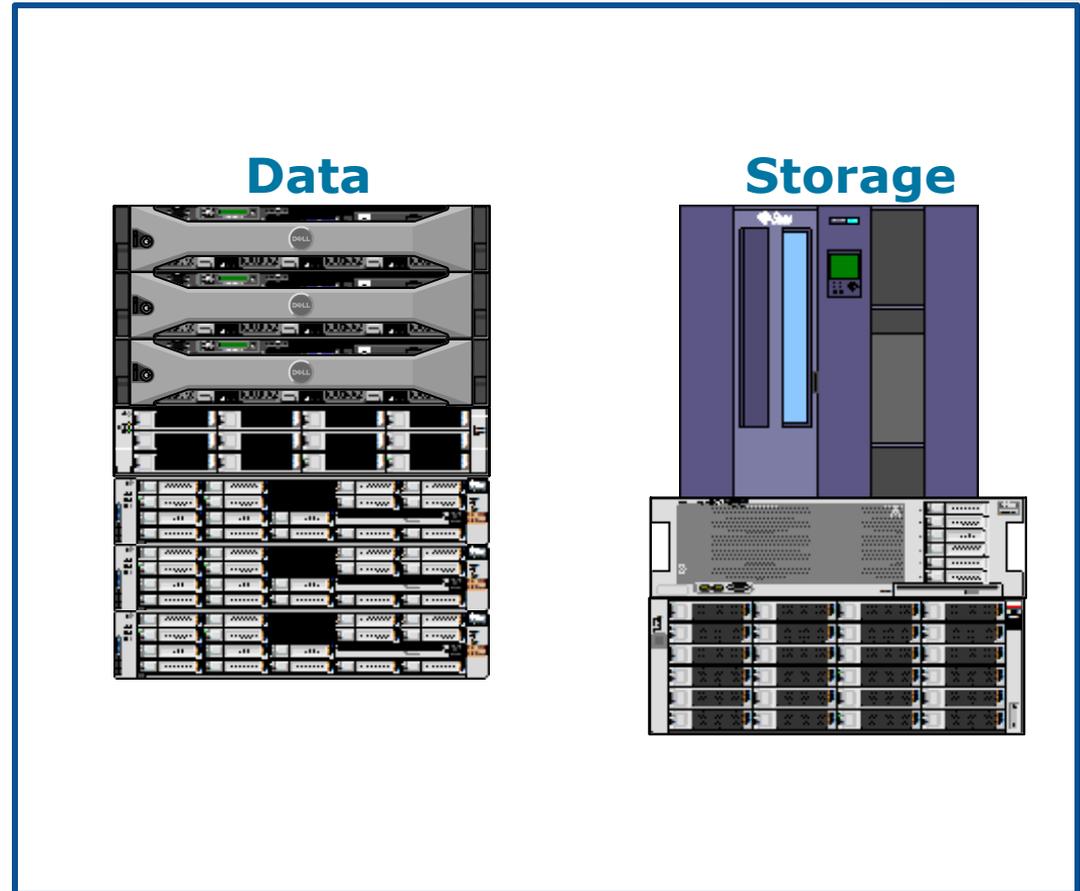
GHRC Overview

- **Data Systems**

- Ingest & Processing
- Public (Web, FTP)
- Database

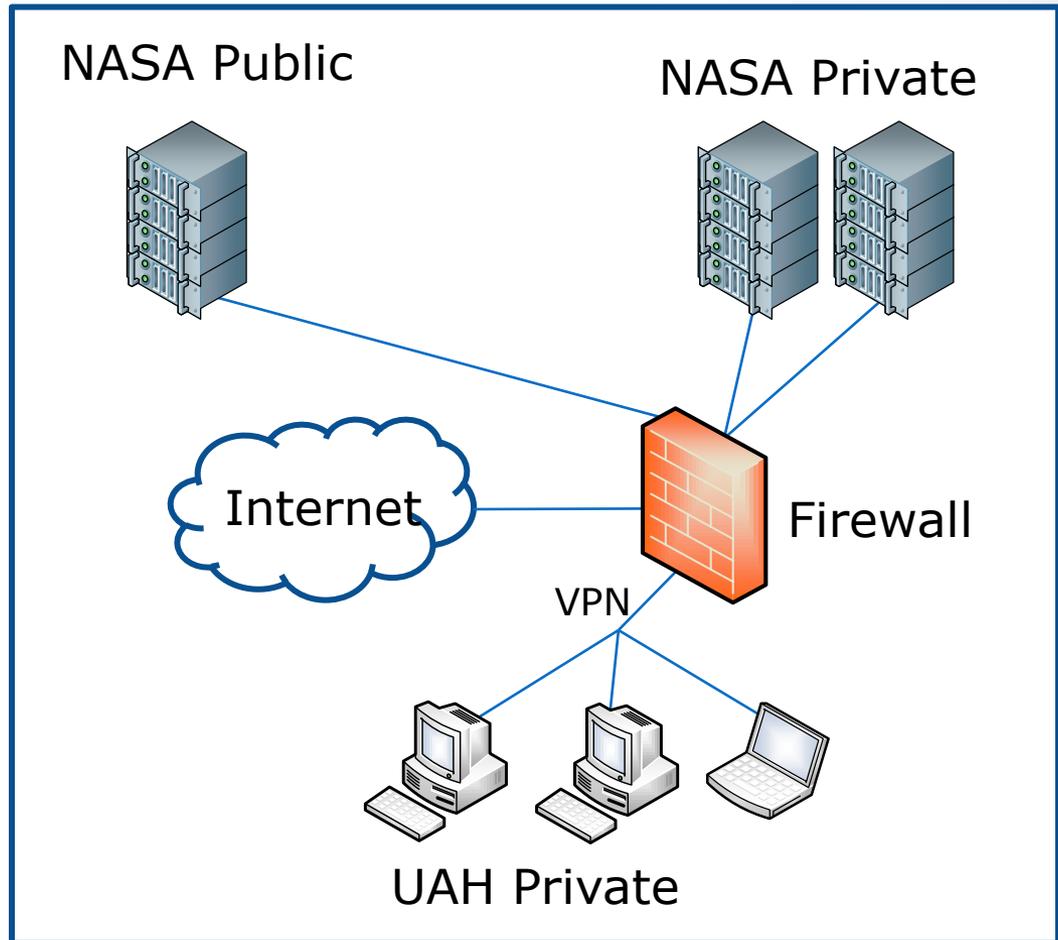
- **Storage Systems**

- Tape-based Archive
- Disk-based Archive
- Backup
- NAS



GHRC Network

- **NASA Public**
 - Web
 - FTP
- **NASA Private**
 - Ingest
 - Processing
 - Archive
 - NAS
- **UAH Private**
 - User Workstations



Public Network Systems

web/ftp

Production Sites



Dell PowerEdge R510
GHRC.nssc.nasa.gov
LIGHTNING.nssc.nasa.gov
SCS3.nssc.nasa.gov

Field Campaigns



Dell PowerEdge R510
AIRBORNE SCIENCE.nssc.nasa.gov
FCPORTAL.nssc.nasa.gov
GPM.nssc.nasa.gov

HS3 Project



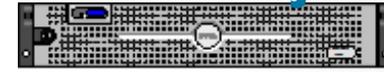
Dell PowerEdge R510
HS3.nssc.nasa.gov

LANCE Project



Dell PowerEdge R720
LANCE.nssc.nasa.gov

RTMM Project



Dell PowerEdge 2950
RTMM2.nssc.nasa.gov
(retiring soon)

Private Network Systems

Ingest/Processing



Dell PowerEdge R510
gale

Database



Sun Fire X4250
neptune

AMSR Processing



Sun Fire X4270
AMSR1-3

LANCE Processing



Dell PowerEdge R720
gwen1

Backup/Logs



Dell PowerEdge R710
underdog

LMA Processing



Dell Precision T7500
LMA Processing

Storage



NetGear PowerNAS 4200
20TB NAS

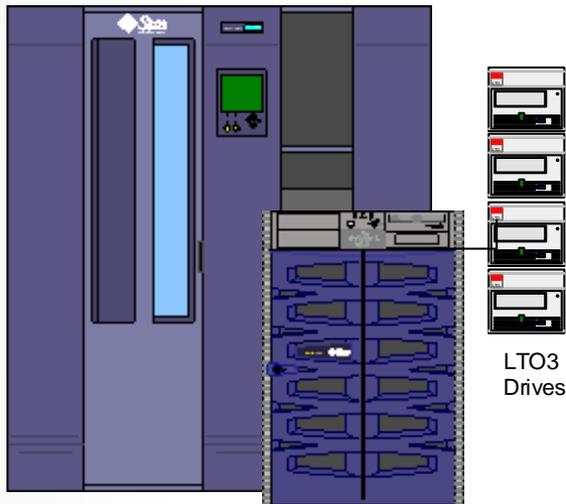
AMSR Storage



Sun Storage
amsrnas1: 16TB NAS
amsrnas2: 20TB NAS
(scaleable to 60TB)

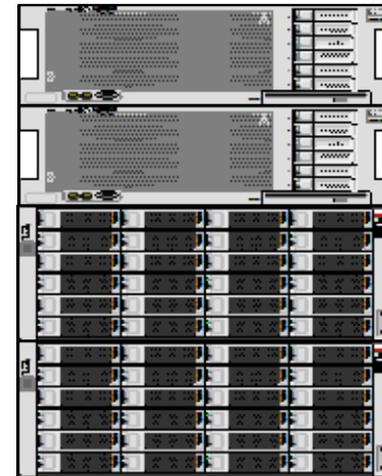
Private Network Systems

KELVIN



Sun V880/L700
90TB Tape Archive
75% full

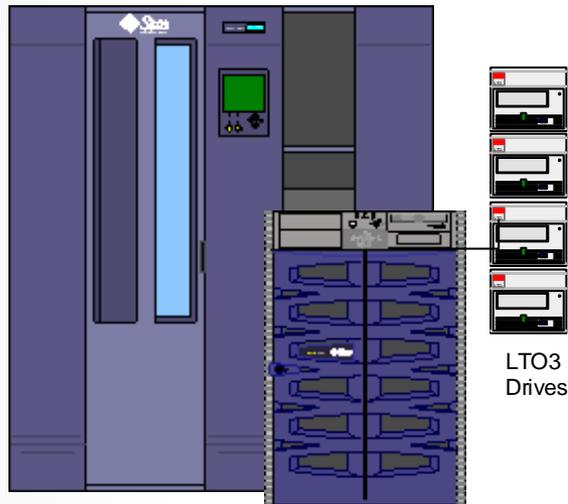
GHRCARC1-2



Sun ZFS Storage 7420
120TB Disk Archive
10% full

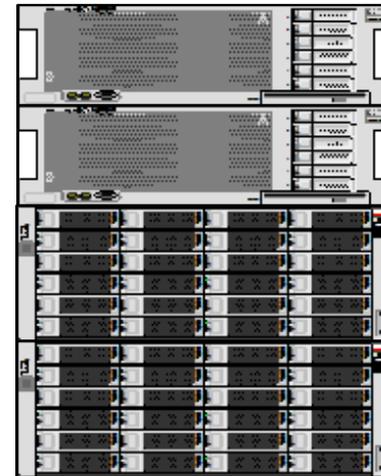
Archive Migration

Installed Sept 2002



Sun V880/L700
90TB usable
Scalable to 500TB

Installed June 2013

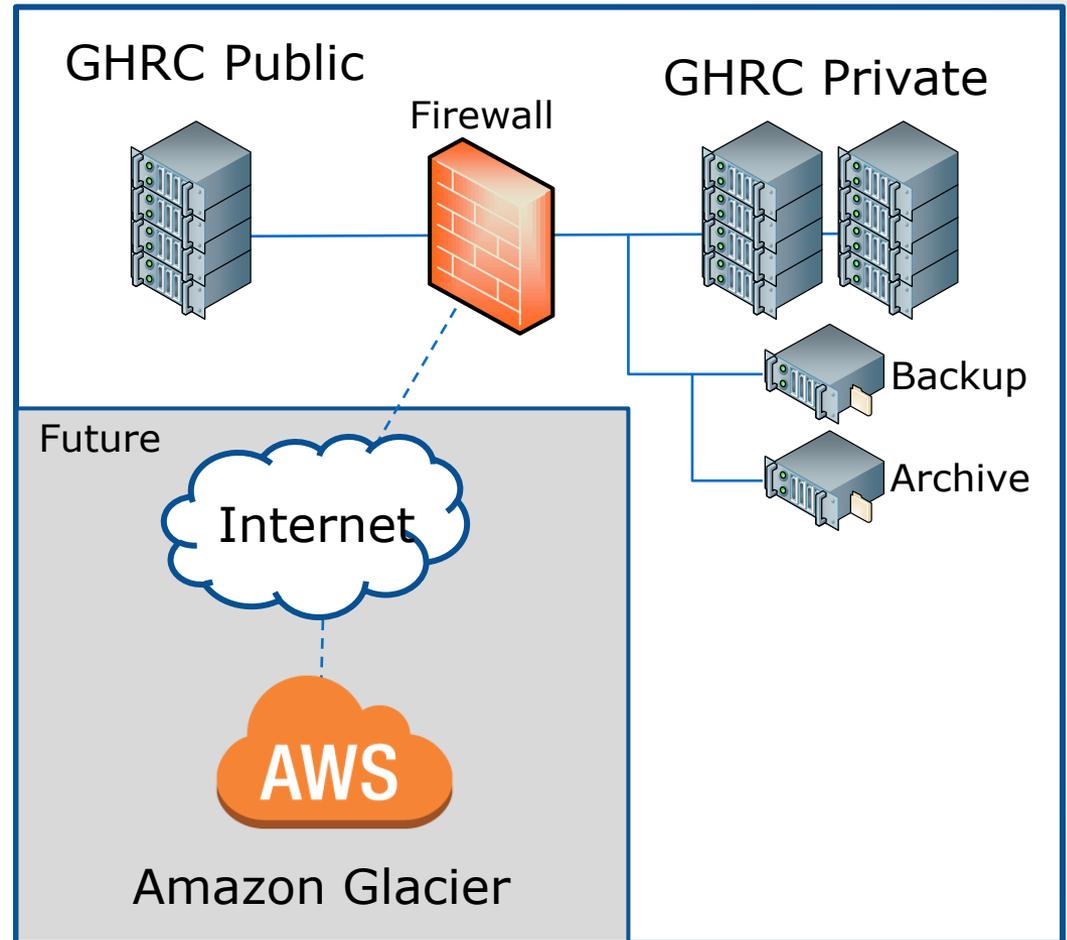


Sun ZFS Storage 7420
120TB usable
Scalable to 2PB

Replacing aging Tape Archive – to be completed by Summer 2015

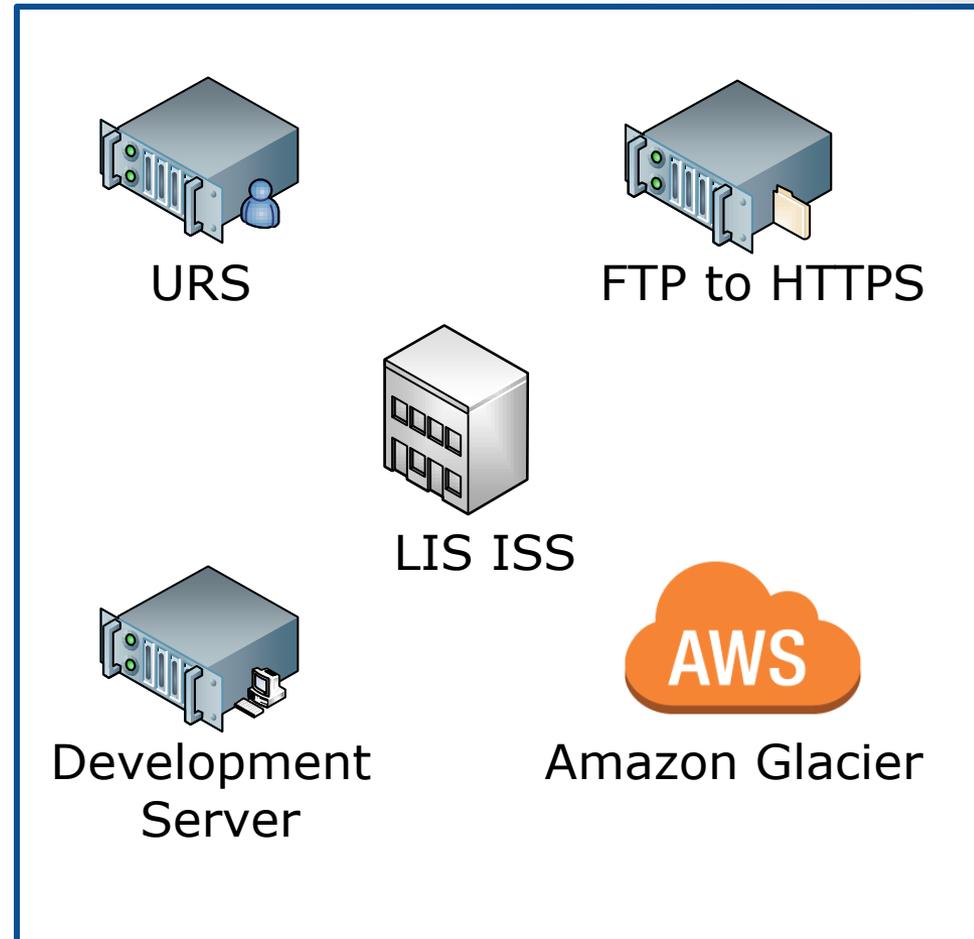
Data Backup

- **Tape Backup**
 - System files
 - Source code
 - Critical data
- **Tape/Disk Archive**
 - Datasets (multiple copies)
- **Researching Off-Site Archive**
 - Datasets



Future Projects

- **User Registration System (URS)**
 - Require registration for data access
- **FTP to HTTPS**
 - Evaluate Impact on Users
- **LIS Space Station**
 - Setup new Operations Center
- **Development Server**
 - Help reduce load on **gale**
 - Additional Storage
- **Off-Site Archive**
 - Amazon Glacier



GHRC DATA PROCESSING

Lamar Hawkins

Operations Manager

dhawkins@itsc.uah.edu

Bruce Beaumont

Lead Software Engineer

beaumont@itsc.uah.edu

Presented at the GHRC User Working Group Meeting
September 25-26, 2014



The Situation by the Numbers

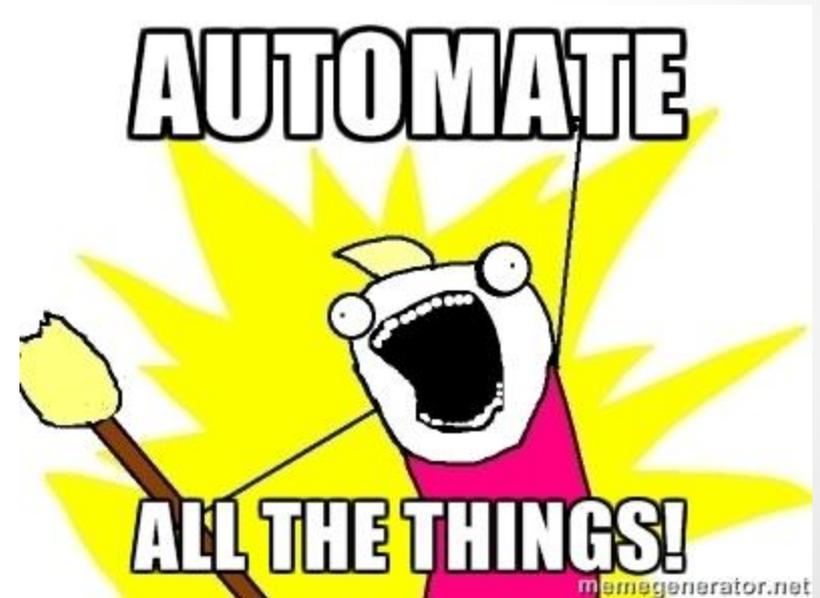
- ~300 cataloged datasets
- ~30 ongoing datasets
- Frequent field campaigns
 - ~25 real time data ingests (each)
- 1-1/2 Operations staff



Goals

Automate everything!

- Standardize data processing
- Simplify data flow
- Reduce duplicated code
- Increase maintainability
- Document everything
- Automated watchdogs



Environments

- DEV (development)
 - Writable by all developers
 - Basic (unit) testing done here
- TEST (integration & test)
 - Writable by Operations staff only
 - Acceptance testing done here
- OPS (production)
 - Writable by SysAdmin only
 - Certain directories are writable by Ops staff
 - Operational processing done here



Overall Data flow



Data Ingest

Ingest

- PUSH method
 - Remote site delivers data to us periodically
 - Standard SW discovers new data
- PULL method
 - We poll a remote site for new data
 - Standard SW handles new data
- Other method
 - Data delivered on media
 - Other PUSH method (socket, LDAP)
- Ingest metrics are generated for most streams

Processing

Process

- Science processing for some data
- May include reformatting, renaming, etc.
- Processing is *not* required
- Modules are stream-specific

Data Distribution

Distribute

- Data distribution is handled by a common module
- Distribution may include
 - Copying files to public or private FTP areas
 - Putting files on the archive (in OPS only!)
 - Staging files for delivery to external users via PUSH
- File-level metadata are generated for most streams

GHRC DATA SEARCH, ACCESS AND ORDER

Sherry Harrison

User Services and Data Management Team Member

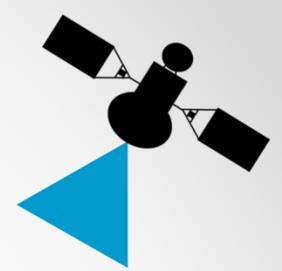
sharrison@itsc.uah.edu

Mary Nair

DBA and Data Management Team Member

mnair@itsc.uah.edu

Presented at the GHRC User Working Group Meeting
September 25-26, 2014



Overview

- Search
 - HyDRO
 - Reverb
 - GCMD
 - Data Set List
 - OpenSearch
 - Tropical Storm Tracks
- Access
 - Field Campaign Portals
 - DOIs
 - Data Set Landing Pages
 - Guides
 - OPeNDAP
 - Ftp
 - Future: https
- Order
 - Automated Order Processing
 - Data Subscriptions: PUSH & GDX

Hydrologic Data Search, Retrieval, and Order System (HyDRO)

Application developed at the GHRC by Bruce Beaumont

Highlights

- Quick Search
- Advanced Search
- Data Sets by Collection
- Data Set Information
- Download Data
- Order Data



Inventory Listing for Lightning Imaging Sensor (LIS) Science Data

Files are listed by date in reverse chronological order. Click the  icon next to a date to list the files available for that date. If a file name is a hyperlink, you may click it to download the file to your workstation. You may check the box next to the date to order all of the files for that date or check the boxes next to the individual files.

If a file name is preceded by a  icon, it is a tar bundle. You may check the file's checkbox to order the whole bundle, or click the icon to list the files within the bundle and select them individually. **Please be patient.** Retrieving the file information from the tar bundle may take a few seconds.



One of NASA's Earth Science Data Centers
A collaboration between MSFC and the University of Alabama in Huntsville

[Home](#) | [Citing Our Data](#) | [Usage Notice](#) | [GHRC's What's New](#) |



Your Shopping Cart

Your cart contains the following item:

- Uncheck this box and "Update cart" to remove the item.**
- LIS/OTD 0.5 Degree High Resolution Full Climatology (HRFC)
Selected files between 1995-05-04 and 1995-05-04 via FTP, TAR (1 file, 13.70 MiB)

Your order contains a total of 1 file requiring 13.70 MiB of disk space.

- | | | |
|---|--------------------------|-------------------------------|
|  | <input type="checkbox"/> | 2014-04-06 (1 file 24.73 MiB) |
|  | <input type="checkbox"/> | 2014-04-07 (1 file 25.14 MiB) |
|  | <input type="checkbox"/> | 2014-04-08 (1 file 20.72 MiB) |
|  | <input type="checkbox"/> | 2014-04-09 (1 file 17.06 MiB) |
|  | <input type="checkbox"/> | 2014-04-10 (1 file 18.93 MiB) |
|  | <input type="checkbox"/> | 2014-04-11 (1 file 14.90 MiB) |

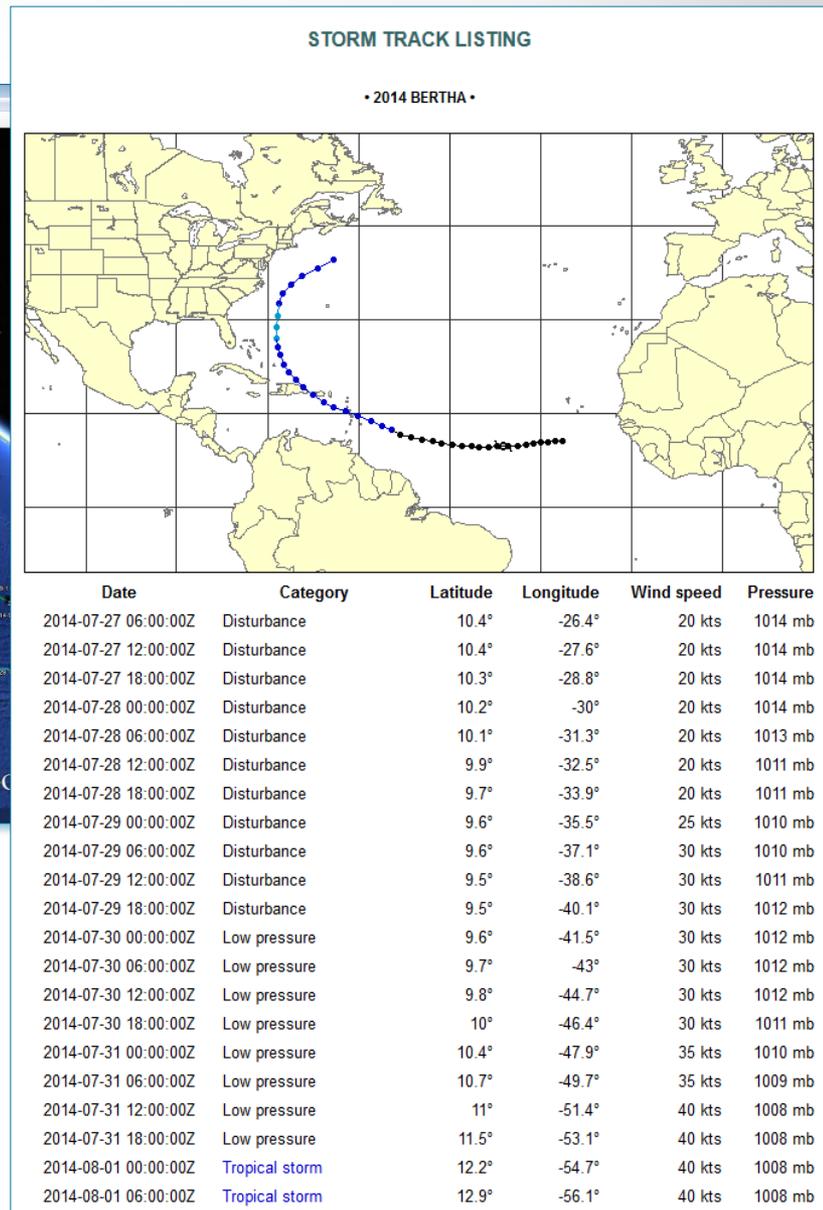
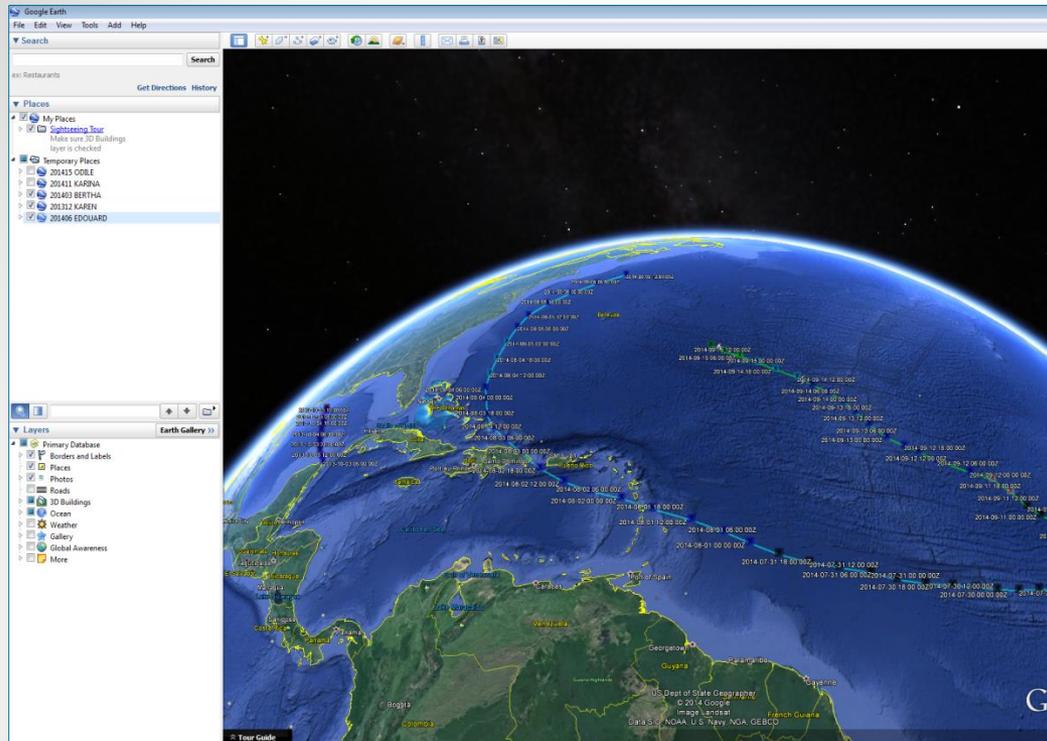
<http://ghrc.nsstc.nasa.gov/hydro/>

Data Search Tools

- Reverb
<http://reverb.e>
- Global Change
Directory (GC
<http://gcmd.gs>
- Data Set List
<http://ghrc.nss>
hydro/search.p
- OpenSearch
 - Provides a
API for sea
GHRC cata
<http://ghrc.nss>
hydro/ghost.xr



Tropical Storm Tracks



- Application developed at the GHRC
- Storm data from the National Hurricane Center
- ~ 6 hour interval updates during active storms

<http://ghrc.nsstc.nasa.gov/storms/>

Field Campaign Portals

<http://fcportal.nsstc.nasa.gov/>

- Access restricted to field campaign participants and collaborators

9/25/14 – 9/26/14

23

Digital Object Identifiers (DOIs)

- What is a DOI?
 - Unique alphanumeric string used to identify a digital object
 - Provides persistent identification with a permanent online link
 - Enables easier access to research data
 - Assigned and regulated by The International DOI Foundation (IDF)
 - Often used in online publications in citations
- DOIs at the GHRC
 - DOIs have been defined for most of the approximately 300 datasets in the GHRC catalog, with about 65% of these registered through ESDIS.
 - Dataset Landing Pages are already provided for all GHRC datasets, whether or not a DOI is in place.
 - DOI example: <http://dx.doi.org/10.5067/MEASURES/DMSP-F17/SSMIS/DATA302>

Data Set Landing Pages



GPM Ground Validation NCAR Cloud Microphysics Particle Probes GCPEX

The GPM Ground Validation NCAR Cloud Microphysics Particle Probes GCPEX dataset was obtained from three instruments carried aboard the University of North Dakota (UND) Cessna Citation aircraft. These probes, the 2D-C, Cloud Imaging Probe (CIP) and High Volume Precipitation Spectrometer (HVPS-3), collected particle size distributions and particle images which were processed by NCAR. Data were collected January 16, 2012 through February 25, 2012. A related cloud microphysics dataset, GPM Ground Validation UND Citation Cloud Microphysics GCPEX is also available.

Please include the following citation in the reference section of your publication:

Heymsfield, A.J., Bansemer, Aaron, and Poellot, Mike. 2014. GPM Ground Validation NCAR Cloud Microphysics Particle Probes GCPEX [indicate subset used]. Dataset available online [http://ghrc.nsstc.nasa.gov] from the NASA EOSDIS Global Hydrology Resource Center Data Access Viewer.

For more information on GHRC DAAC datasets, visit the GHRC Data Access Viewer.

[Get data](#) ALL of the requested data

[View browse](#) ALL of the browse information

General Characteristics

Collections: GPM-GV GCPEX Products
Projects: GCPEX
Platforms: UND CITATION II
Instruments: CIP, HVPS, PMS 2
Parameters: CLOUD DROPLET SIZE, DROPLET SIZE

Processing level: 2

Format: ASCII

Coverage

Location: ONTARIO, CANADA
Spatial resolution: 7.5mm-2.5cm
North boundary: 46.5°
West boundary: -81°
East boundary: -78°
South boundary: 43.5°
Temporal DAILY PER FLIGHT resolution:
Start date: 2012-01-16
Stop date: 2012-02-25

Links

[Browse](#): ftp://gpm.nsstc.nasa.gov/validation/gcpeex/cloud_microphysics_Citation/NCAR_particle_probes/browse/ Browse images illustrate the data.
[Guide](#): http://ghrc.nsstc.nasa.gov/guide/gcpeex/ The guide document contains information on the data.
[Homepage](#): http://gpm.nsstc.nasa.gov/gcpeex/ The home page for the project.
[DOI](#): http://dx.doi.org/10.5067/GPMGV/GCPEX/MUTIPLE/DATA201 Digital Object Identifier
[Citing data](#): http://ghrc.nsstc.nasa.gov/uso/citation.html Instructions for citing GHRC data.

The GPM Ground Validation NCAR Cloud Microphysics Particle Probes GCPEX dataset was obtained from three instruments carried aboard the University of North Dakota (UND) Cessna Citation aircraft. These probes, the 2D-C, Cloud Imaging Probe (CIP) and High Volume Precipitation Spectrometer (HVPS-3), collected particle size distributions and particle images which were processed by NCAR. Data were collected January 16, 2012 through February 25, 2012. A related cloud microphysics dataset, GPM Ground Validation UND Citation Cloud Microphysics GCPEX is also available.

Please include the following citation in the reference section of your publication:

Heymsfield, A.J., Bansemer, Aaron, and Poellot, Mike. 2014. GPM Ground Validation NCAR Cloud Microphysics Particle Probes GCPEX [indicate subset used]. Dataset available online [http://ghrc.nsstc.nasa.gov] from the NASA EOSDIS Global Hydrology Resource Center Data Access Viewer.

General Characteristics

Collections: GPM-GV GCPEX Products
Projects: GCPEX

Coverage

Location: ONTARIO, CANADA
Spatial resolution: 7.5mm-2.5cm
North boundary: 46.5°
West boundary: -81°



Links

[Browse](#): ftp://gpm.nsstc.nasa.gov/validation/gcpeex/cloud_microphysics_Citation/NCAR_particle_probes/browse/

[DOI](#): http://dx.doi.org/10.5067/GPMGV/GCPEX/MUTIPLE/DATA201
Digital Object Identifier

[Homepage](#): http://gpm.nsstc.nasa.gov/gcpeex/ The home page for the project or program which sponsored the dataset

[DOI](#): http://dx.doi.org/10.5067/GPMGV/GCPEX/MUTIPLE/DATA201
Digital Object Identifier

[Citing data](#): http://ghrc.nsstc.nasa.gov/uso/citation.html
Instructions for citing GHRC data.

Instructions for citing GHRC data.

- One-paragraph description

<http://ghrc.nsstc.nasa.gov/uso/citation.html>

Guides

- Data set overview document composed by the GHRC from PI provided information
- Features
 - Instrument Overview
 - Data Format and File Naming Convention
 - Investigator Information
 - Algorithm Details
 - PI Documentation and Software Information and Links
 - Citations and References



RSS Monthly 1-deg Microwave Total Precipitable Water netCDF

Table of Contents

<i>Introduction</i>	<i>Read Software</i>
<i>Instrument Description</i>	<i>Tools</i>
<i>Investigators</i>	<i>Citation</i>
<i>File Naming Convention</i>	<i>References</i>
<i>Data Format</i>	<i>Contact Information</i>
<i>Algorithm and Processing Steps</i>	

Introduction

The RSS Monthly 1-degree Microwave Total Precipitable Water (TPW) netCDF data set contains monthly mean TPW on a one degree grid, a climatology file containing a set of twelve monthly TPW means on a one degree grid, and a trend file containing the cumulative global TPW trend map on a one degree grid and the time-latitude plot. This data set was constructed using the version- 7 (V7) passive microwave geophysical ocean products made publicly available from Remote Sensing Systems (www.remss.com). The TPW values come from the following satellite radiometers: SSM/I F08 through F15, SSMIS F16 and F17, AMSR-E, and WindSat. The microwave radiometers were carefully inter-calibrated at the brightness temperature level and the V7 ocean products were produced using a consistent processing methodology for all sensors. The high quality ocean data is made available thanks to funding from the NASA MEaSUREs (Making Earth System Data Records for Use in Research Environments) project. The *Global Hydrology Resource Center (GHRC)*, a NASA science data center managed by the *University of Alabama in Huntsville*, processes the cumulative TPW files from RSS into monthly files. The GHRC also provides the cumulative 10 year climatology file and the cumulative trend files. This data set contains both netCDF3 and netCDF4 formatted files.

Instrument Description

The data used to make the RSS Monthly 1-degree Microwave Total Precipitable Water (TPW) netCDF product are from all SSM/I instruments, SSMIS F16 and F17, AMSR-E, and WindSat. SSM/I and SSMIS are instruments carried onboard the Defense Meteorological Satellite Program (DMSP) series of polar orbiting satellites. The Special Sensor Microwave/Imager (SSM/I) is a seven channel passive microwave radiometer operating at four frequencies and dual-polarization. The Special Sensor Microwave Imager Sounder (SSMIS) is also a microwave radiometer that includes a sounder. Additional information on the SSM/I and the SSMIS can be found at <http://www.ssmi.com/ssmi/>. The Advanced Microwave Scanning Radiometer - EOS (AMSR-E) is one of six sensors aboard Aqua. AMSR-E is a twelve-channel, six-frequency, total power passive-microwave radiometer. More information on AMSR-E can be found at <http://www.ghcc.msfc.nasa.gov/AMSR/>. WindSat is a satellite-based polarimetric microwave radiometer which measures ocean surface wind vectors. WindSat is aboard the Coriolis satellite. More information on WindSat can be found at <http://weather.msfc.nasa.gov/sport/windsat/>. More information on Coriolis can be found at <http://www.orbital.com/SatellitesSpace/ImagingDefense/Coriolis/>.

Investigators

Frank Wentz
Remote Sensing Systems
438 First Street, Suite 200
Santa Rosa, CA 95401

File Naming Convention

The monthly files are named with the following convention:

tpw_v07r00_[YYYYMM].nc[3 or 4].nc

http://ghrc.nsstc.nasa.gov/uso/ds_docs/tpw/rssm1tpwn_dataset.html

Additional Access Methods

<ftp://ghrc.nsstc.nasa.gov/>

<ftp://gpm.nsstc.nasa.gov/>

<http://ghrc.nsstc.nasa.gov/opendap/>



Contents of /ssmis/f16/weekly/data/2014

Name	Last Modified
Parent Directory/	
f16 ssmis 20140104v7 wk.nc	2014-01-05T19:03
f16 ssmis 20140111v7 wk.nc	2014-01-12T19:03
f16 ssmis 20140118v7 wk.nc	2014-01-19T19:03
f16 ssmis 20140125v7 wk.nc	2014-01-26T19:05:03
f16 ssmis 20140201v7 wk.nc	2014-02-02T19:05:06

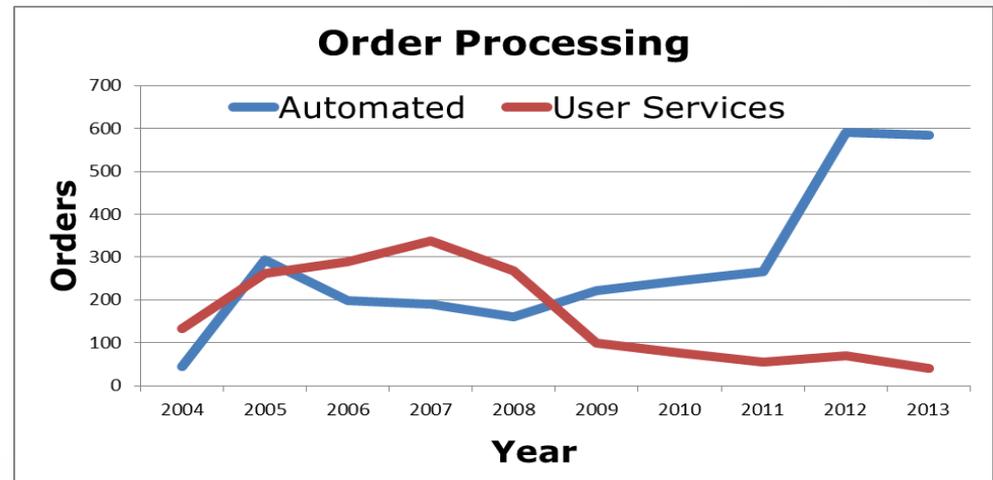
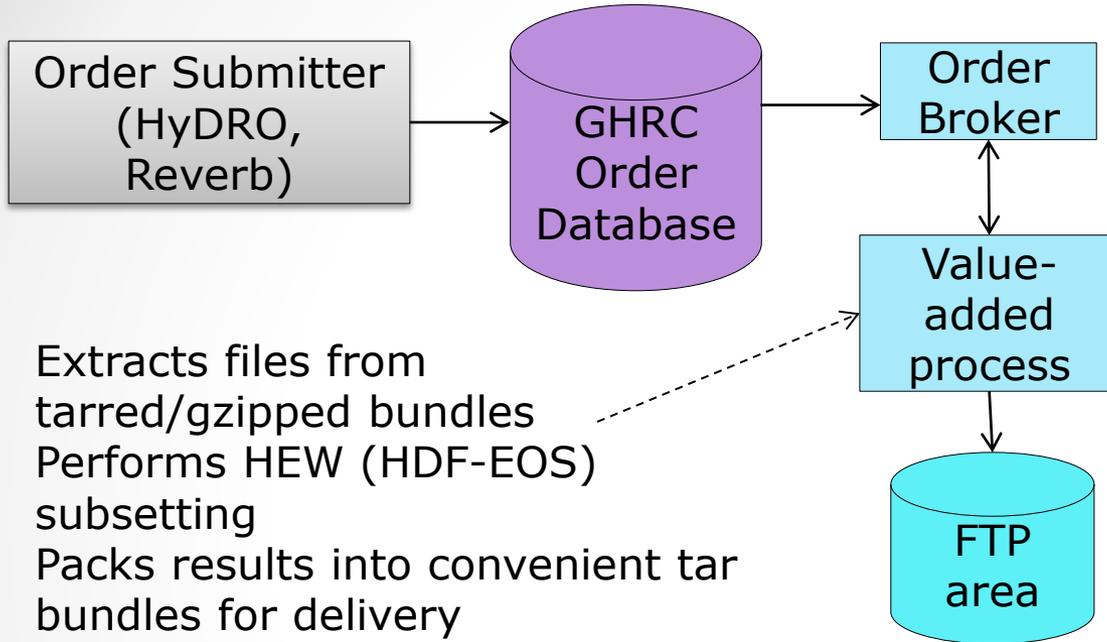
Index of ftp://ghrc.nsstc.nasa.gov/pub/data/lis/science/2014/

[Up to higher level directory](#)

Name	Size	Last Modified
TRMM_LIS_SC.04.2_2014.001.hdf.tar.gz	11692 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.002.hdf.tar.gz	12369 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.003.hdf.tar.gz	14608 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.004.hdf.tar.gz	12837 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.005.hdf.tar.gz	12722 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.006.hdf.tar.gz	12739 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.007.hdf.tar.gz	13153 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.008.hdf.tar.gz	14538 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.009.hdf.tar.gz	14617 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.010.hdf.tar.gz	12538 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.011.hdf.tar.gz	12575 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.012.hdf.tar.gz	10874 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.013.hdf.tar.gz	12962 KB	2/17/2014 12:00:00 AM
TRMM_LIS_SC.04.2_2014.014.hdf.tar.gz		12:00:00 AM

Future: HTTPS

Automated Order Processing



Data Subscriptions

- Data subscription
 - Scheduled delivery of data on a near-real-time basis to individual subscribers
 - Delivery via applications developed at the GHRC (PUSH, GDX)
 - Access to subscription applications is limited to GHRC operations staff
- Product / User Subscription Handler (PUSH)
 - Primary Data Subscription Service
 - Configurable for the dataset and the transfer interval
- GPM Data Interchange (GDX)
 - Command line mechanism for data transfer which includes handshaking
 - Near-real-time LIS provided to PPS (Erich Stocker)
 - Configurable to transfer various data sets



Product / User Subscription Handler (PUSH) Administration

If you are a PUSHer, please log in. Otherwise, go away.

Account Password Your name (first last)

This U.S. Government computing system is for authorized users only. Anyone using it is subject to monitoring and recording of all keystrokes without further notice. Actual or attempted use, access, communication, or examination by unauthorized persons is a criminal violation of Title 18, U.S. Code, Section 1030. This record may be provided as evidence to law enforcement officials. This record will also be used for statistical purposes.

Discussion

THANK YOU

for your attention!

- Please cite your data.
 - When the DOI is available, please use it in your data citation.
 - When your publication cites our data, please notify us.
- What data formats do you prefer?
- What metadata is most useful to you?
- Do you find the user guide documents useful?
- Are there additional data access methods to consider?

If you have not already done so, please respond to the ESDIS survey for the GHRC DAAC.

Please contact **GHRC User Services** for any help or questions
ghrcdaac@itsc.uah.edu

