INFRASTRUCTURE

At the GHRC DAAC

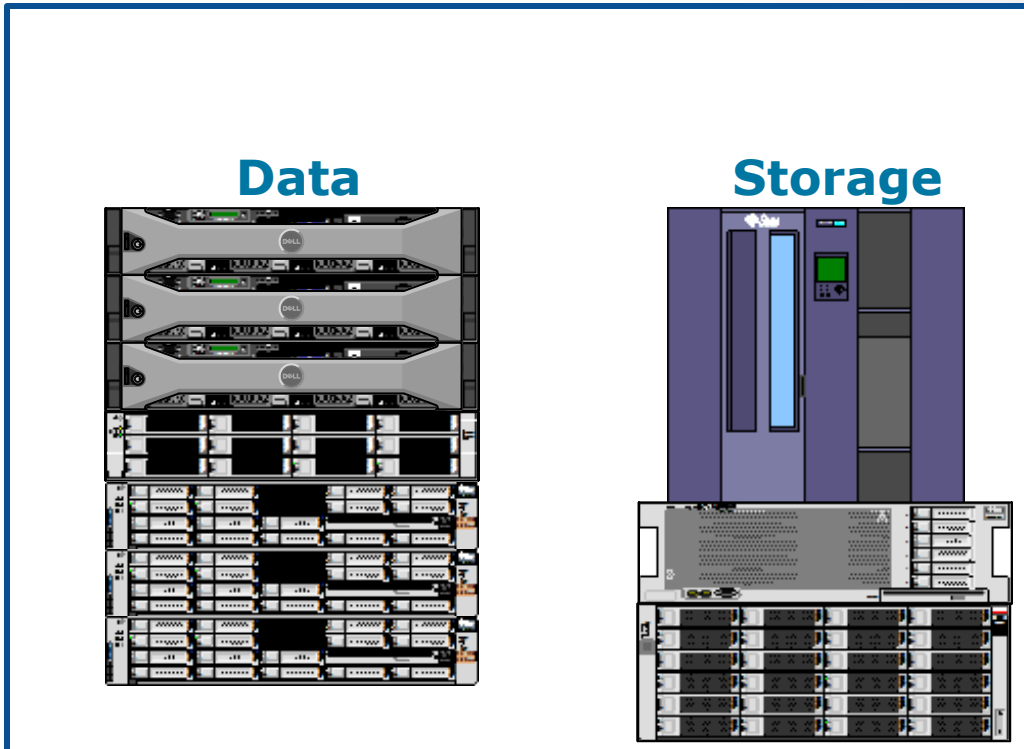
Will Ellett
IT Manager
sysadmin@itsc.uah.edu

Support: Michele Garrett, Michael McEniry, Jason Toone

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

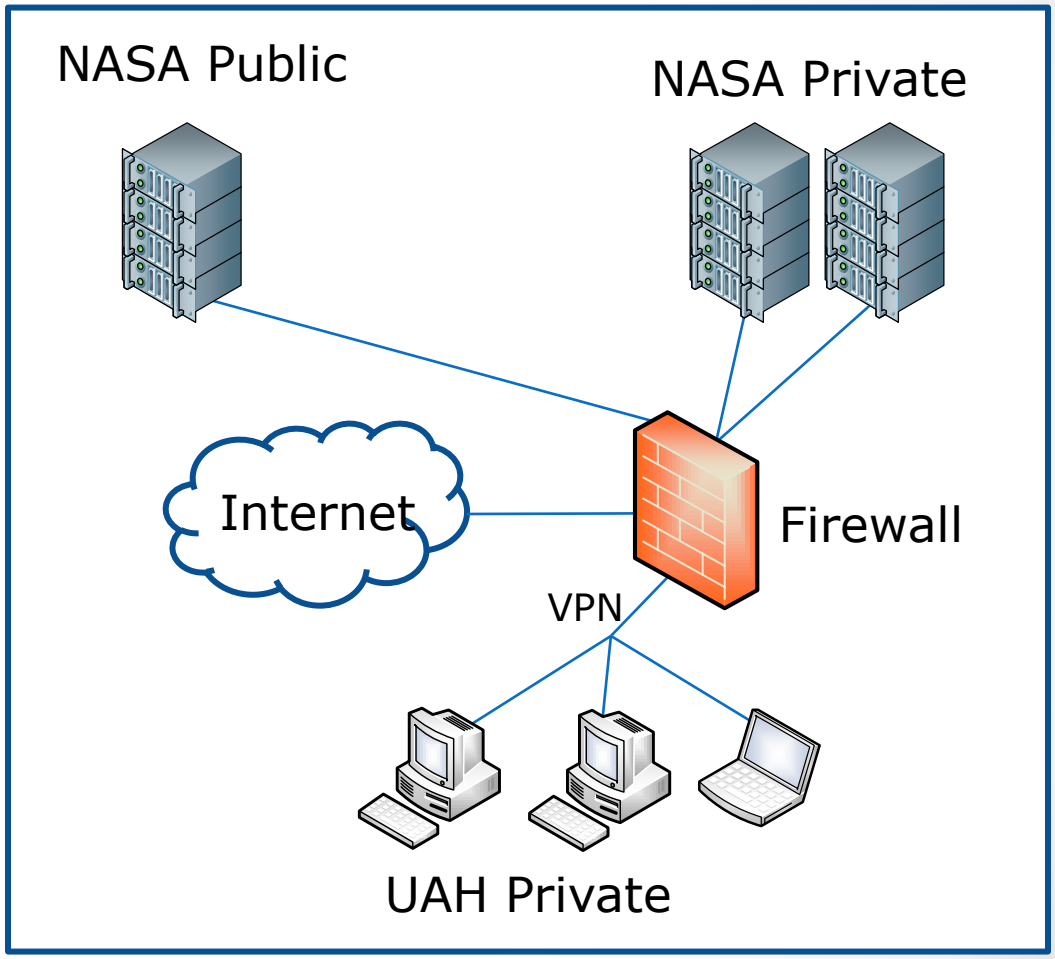
GHRC Overview

- **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.nsstc.nasa.gov
LIGHTNING.nsstc.nasa.gov
SCS3.nsstc.nasa.gov

Field Campaigns

Dell PowerEdge R510
AIRBORNESCIENCE.nsstc.nasa.gov
FCPORTAL.nsstc.nasa.gov
GPM.nsstc.nasa.gov

HS3 Project

Dell PowerEdge R510
HS3.nsstc.nasa.gov

LANCE Project

Dell PowerEdge R720
LANCE.nsstc.nasa.gov

RTMM Project

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

Private Network Systems

Ingest/Processing

Dell PowerEdge R510
gale

LANCE Processing

Dell PowerEdge R720
gwen1

LMA Processing

Dell Precision T7500
LMA Processing

Database

Sun Fire X4250
neptune

Backup/Logs

Dell PowerEdge R710
underdog

Storage

NetGear PowerNAS 4200
20TB NAS

AMSR Processing

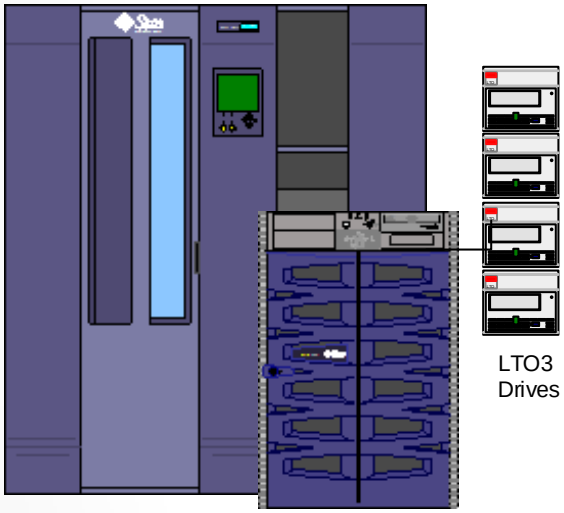
Sun Fire X4270
AMSR1-3

AMSR Storage

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

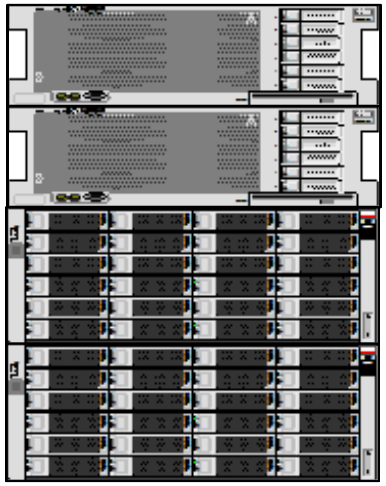
Private Network Systems

KELVIN



Sun V880/L700
90TB Tape Archive
75% full

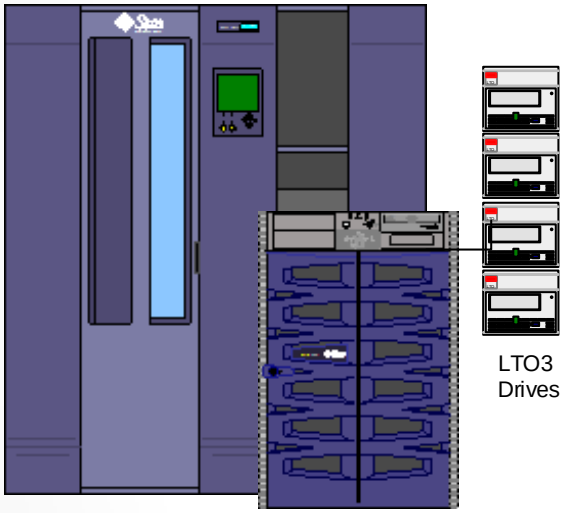
GHRCARC1-2



Sun ZFS Storage 7420
120TB Disk Archive
10% full

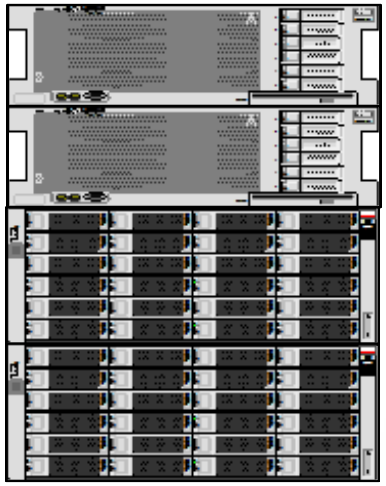
Archive Migration

Installed Sept 2002



Sun V880/L700
90TB usable
Scalable to 500TB

Installed June 2013

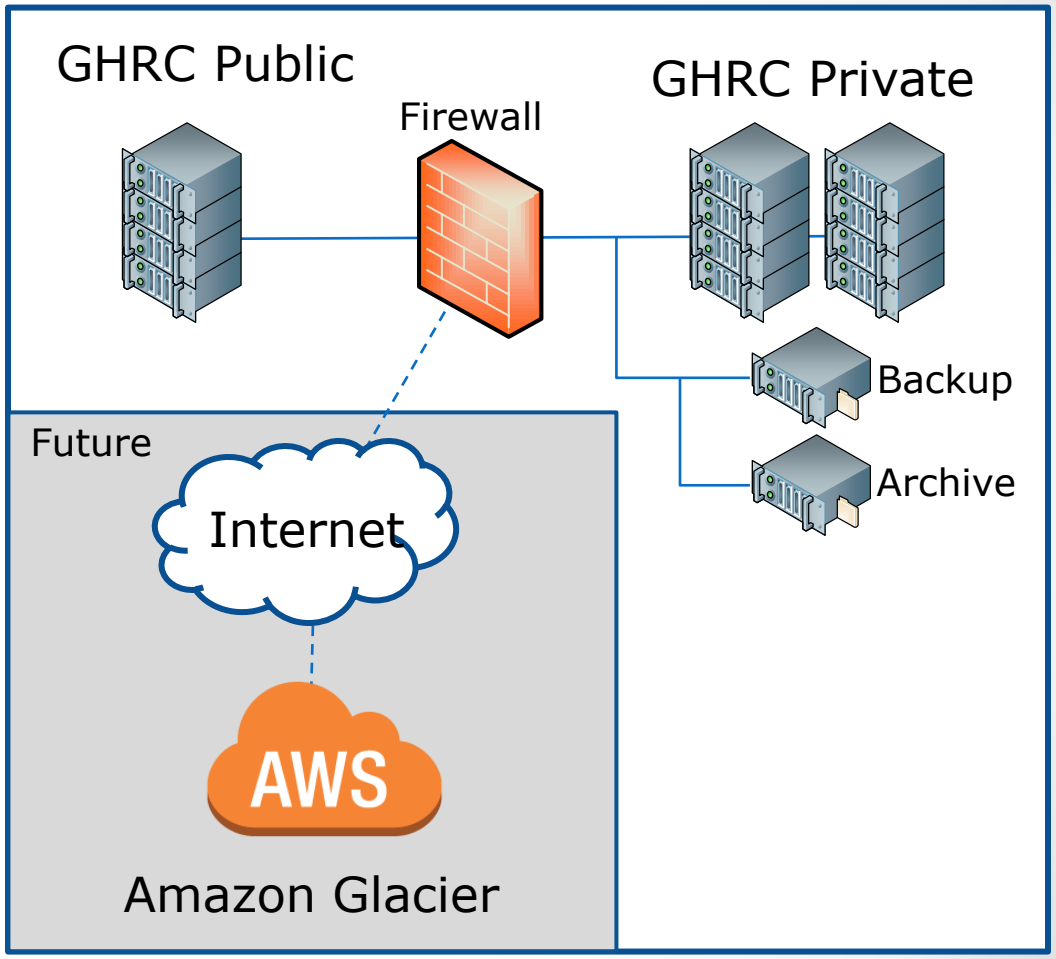


Sun ZFS Storage 7420
120TB usable
Scalable to 2PB

Replacing aging Tape Archive – to be competed by Summer 2015

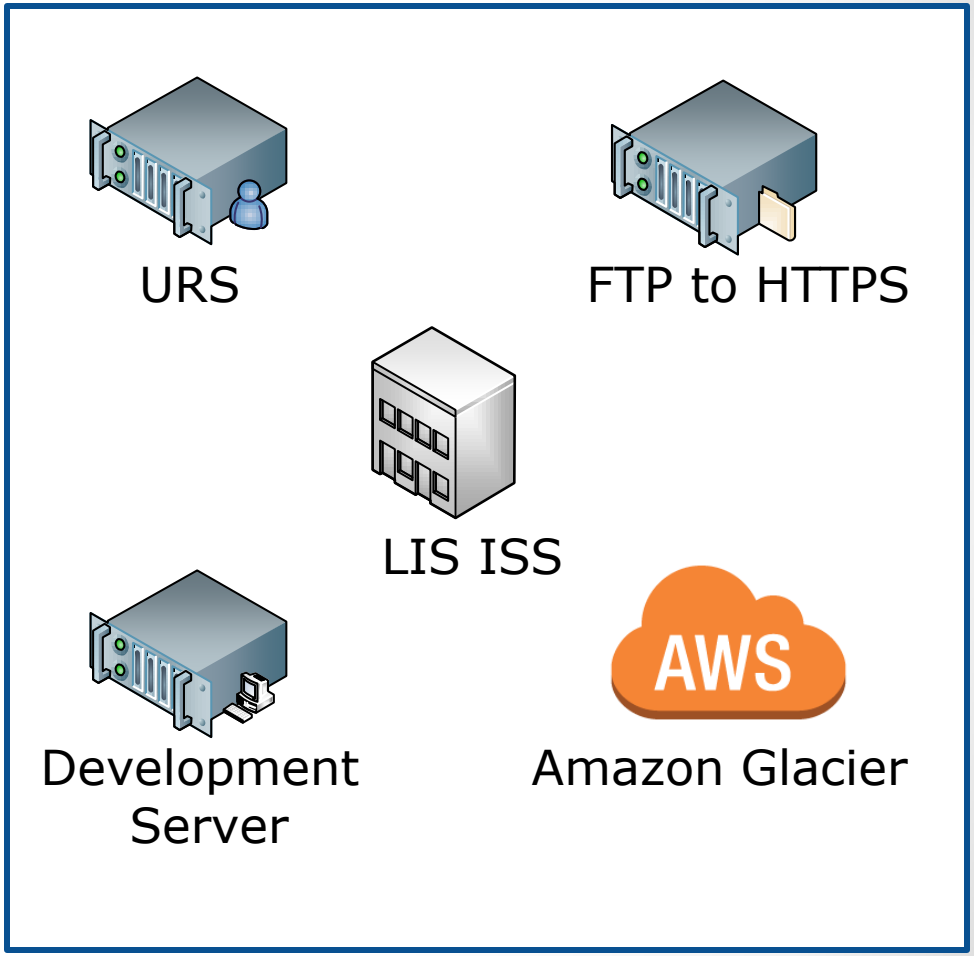
Data Backup

- **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

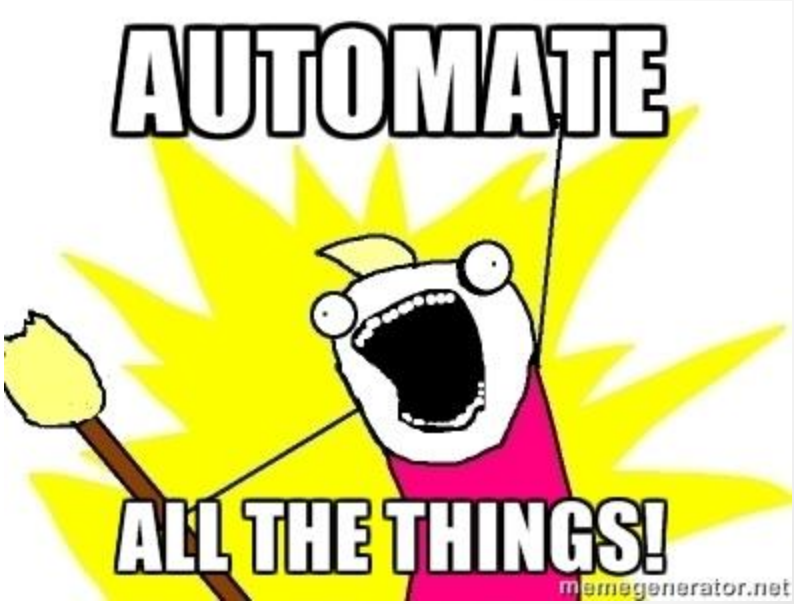
The Situation by the Numbers

- ~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

- 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

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

Data Distribution

- 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

NASA

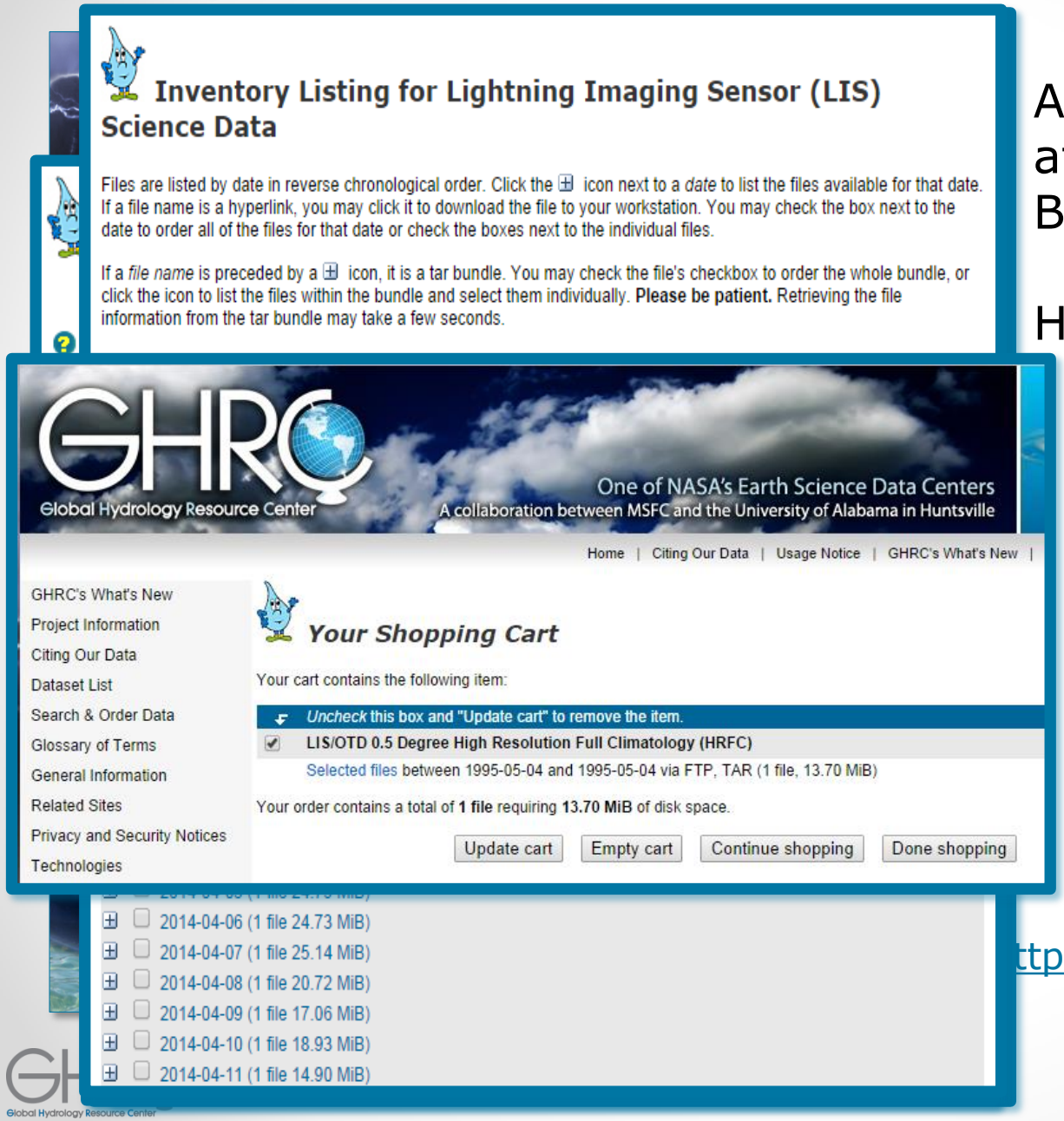
GHRC
Global Hydrology Resource Center

itsc
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

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

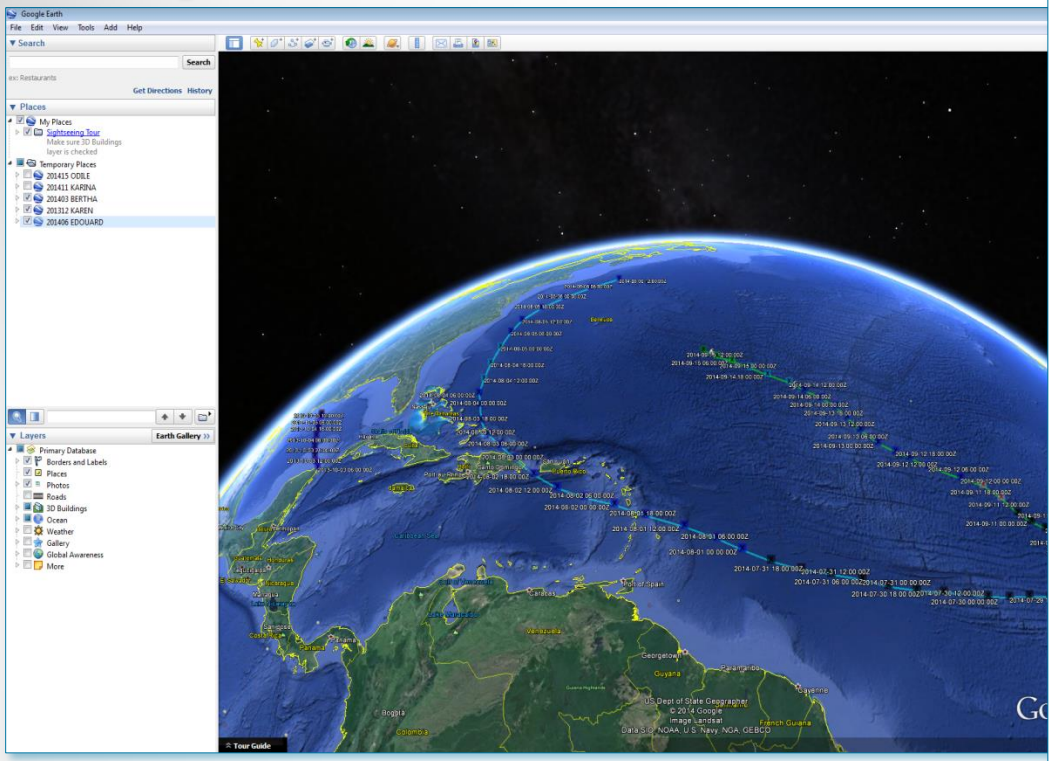
ttp://ghrc.nsstc.nasa.gov/hydro/

Data Search Tools

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

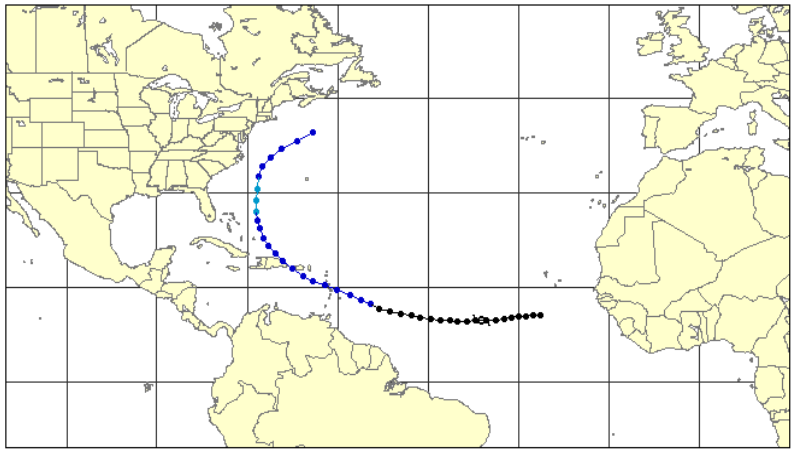
  http://ghrc.nss
  hydro/ghost.xr

Tropical Storm Tracks



STORM TRACK LISTING

• 2014 BERTHA •

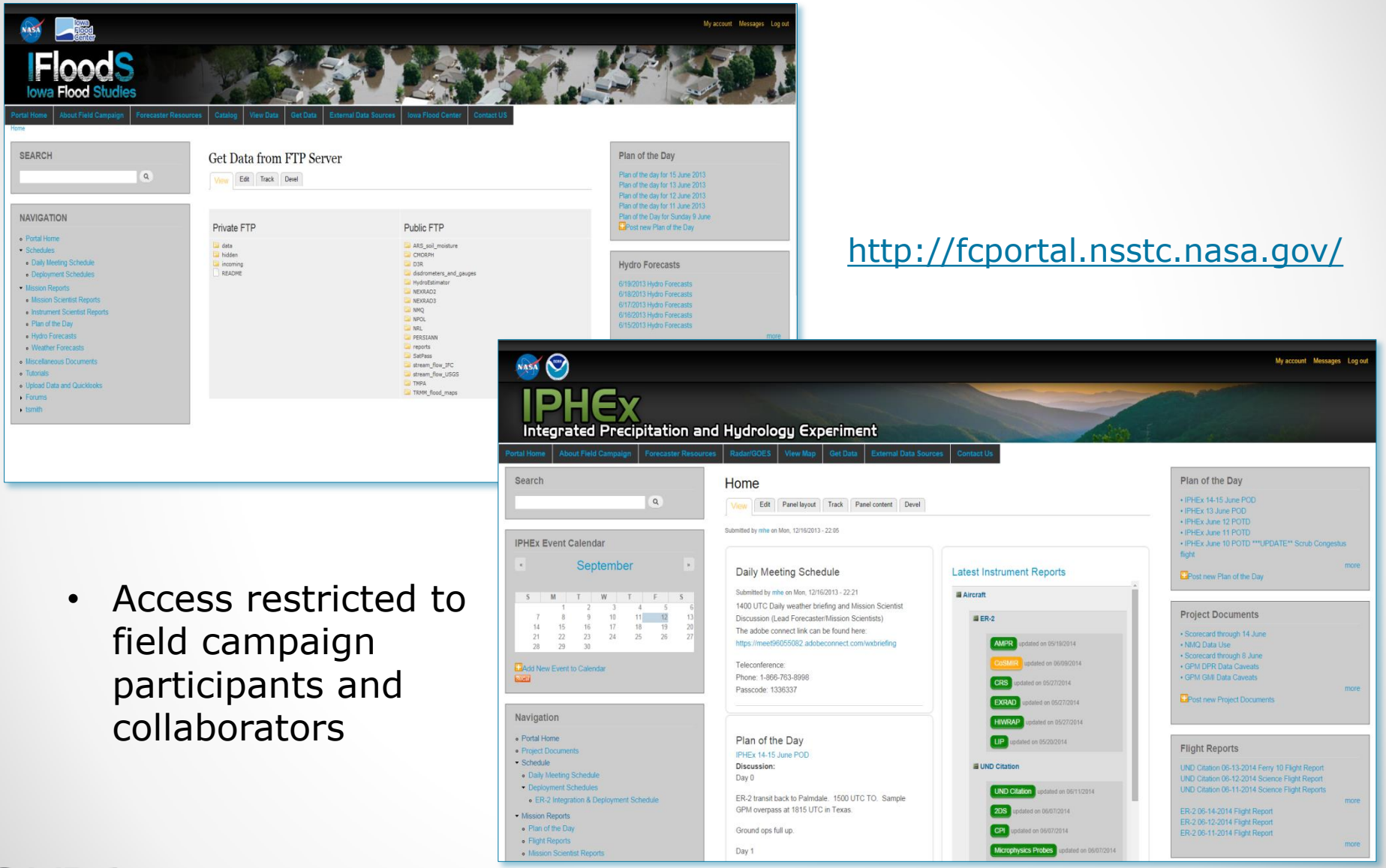


Date	Category	Latitude	Longitude	Wind speed	Pressure
2014-07-27 06:00:00Z	Disturbance	10.4°	-26.4°	20 kts	1014 mb
2014-07-27 12:00:00Z	Disturbance	10.4°	-27.6°	20 kts	1014 mb
2014-07-27 18:00:00Z	Disturbance	10.3°	-28.8°	20 kts	1014 mb
2014-07-28 00:00:00Z	Disturbance	10.2°	-30°	20 kts	1014 mb
2014-07-28 06:00:00Z	Disturbance	10.1°	-31.3°	20 kts	1013 mb
2014-07-28 12:00:00Z	Disturbance	9.9°	-32.5°	20 kts	1011 mb
2014-07-28 18:00:00Z	Disturbance	9.7°	-33.9°	20 kts	1011 mb
2014-07-29 00:00:00Z	Disturbance	9.6°	-35.5°	25 kts	1010 mb
2014-07-29 06:00:00Z	Disturbance	9.6°	-37.1°	30 kts	1010 mb
2014-07-29 12:00:00Z	Disturbance	9.5°	-38.6°	30 kts	1011 mb
2014-07-29 18:00:00Z	Disturbance	9.5°	-40.1°	30 kts	1012 mb
2014-07-30 00:00:00Z	Low pressure	9.6°	-41.5°	30 kts	1012 mb
2014-07-30 06:00:00Z	Low pressure	9.7°	-43°	30 kts	1012 mb
2014-07-30 12:00:00Z	Low pressure	9.8°	-44.7°	30 kts	1012 mb
2014-07-30 18:00:00Z	Low pressure	10°	-46.4°	30 kts	1011 mb
2014-07-31 00:00:00Z	Low pressure	10.4°	-47.9°	35 kts	1010 mb
2014-07-31 06:00:00Z	Low pressure	10.7°	-49.7°	35 kts	1009 mb
2014-07-31 12:00:00Z	Low pressure	11°	-51.4°	40 kts	1008 mb
2014-07-31 18:00:00Z	Low pressure	11.5°	-53.1°	40 kts	1008 mb
2014-08-01 00:00:00Z	Tropical storm	12.2°	-54.7°	40 kts	1008 mb
2014-08-01 06:00:00Z	Tropical storm	12.9°	-56.1°	40 kts	1008 mb

- 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



- Access restricted to field campaign participants and collaborators

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

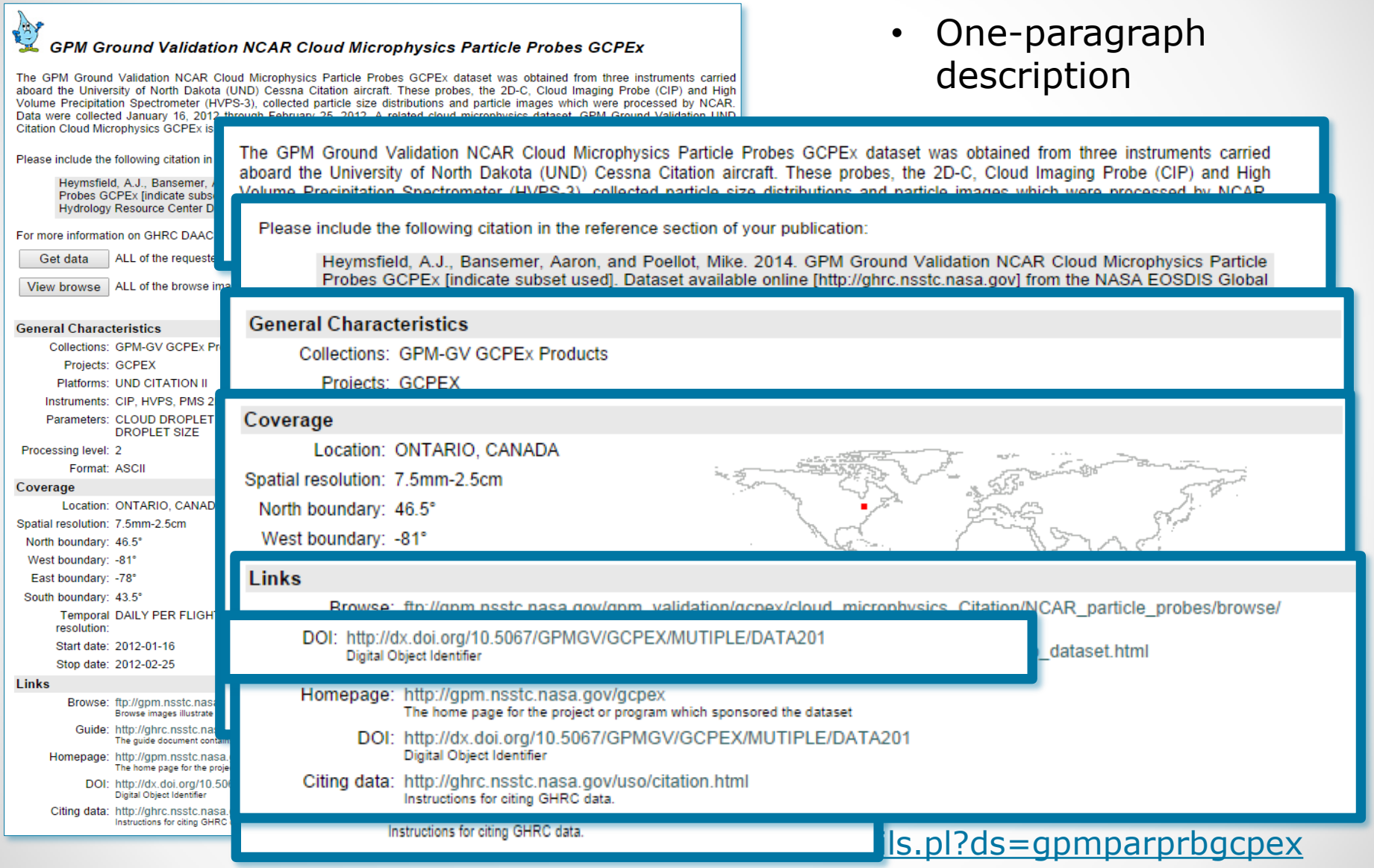
Digital Object Identifiers (DOIs)

- 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

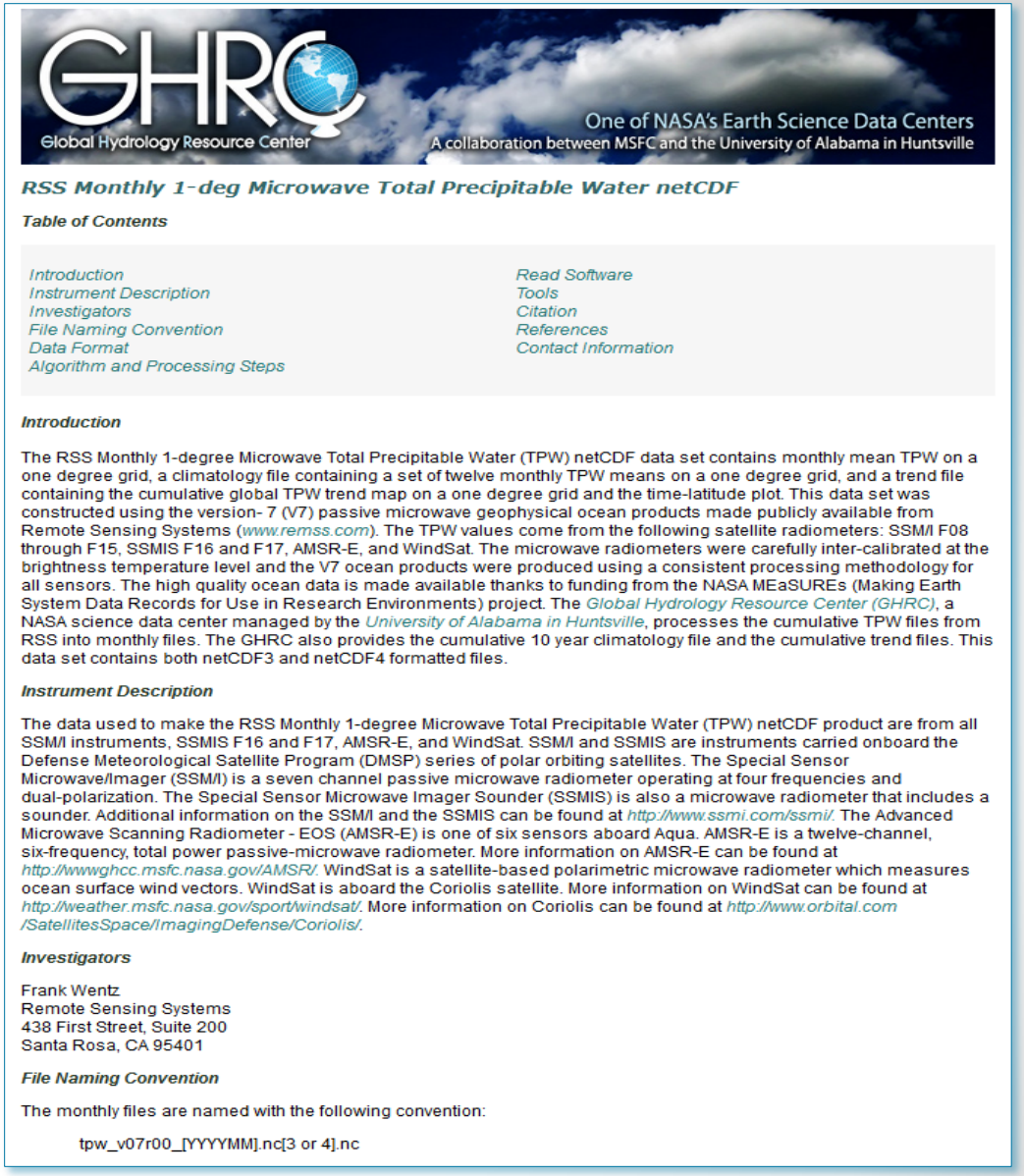
- One-paragraph description



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



GHRC Global Hydrology Resource Center — One of NASA's Earth Science Data Centers — A collaboration between MSFC and the University of Alabama in Huntsville

RSS Monthly 1-deg Microwave Total Precipitable Water netCDF

Table of Contents

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://wwwghcc.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

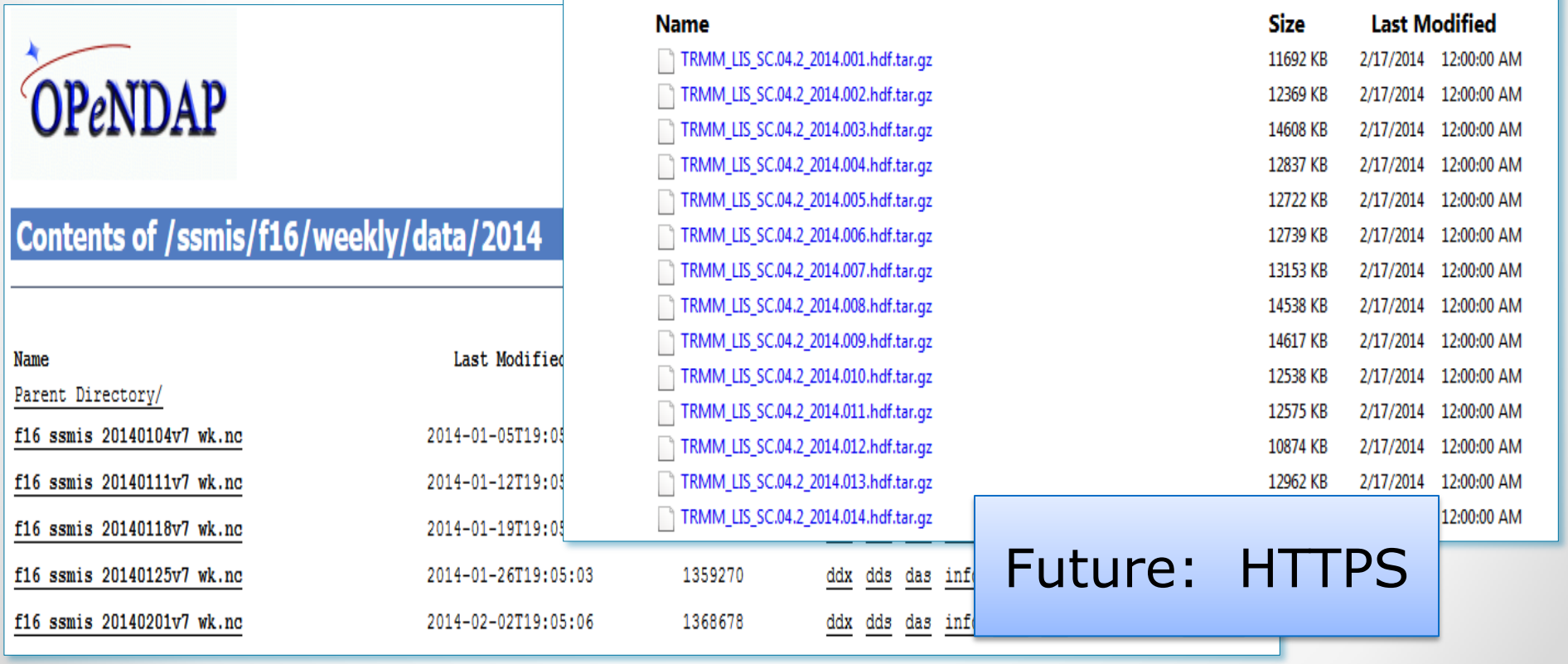
The monthly files are named with the following convention:

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

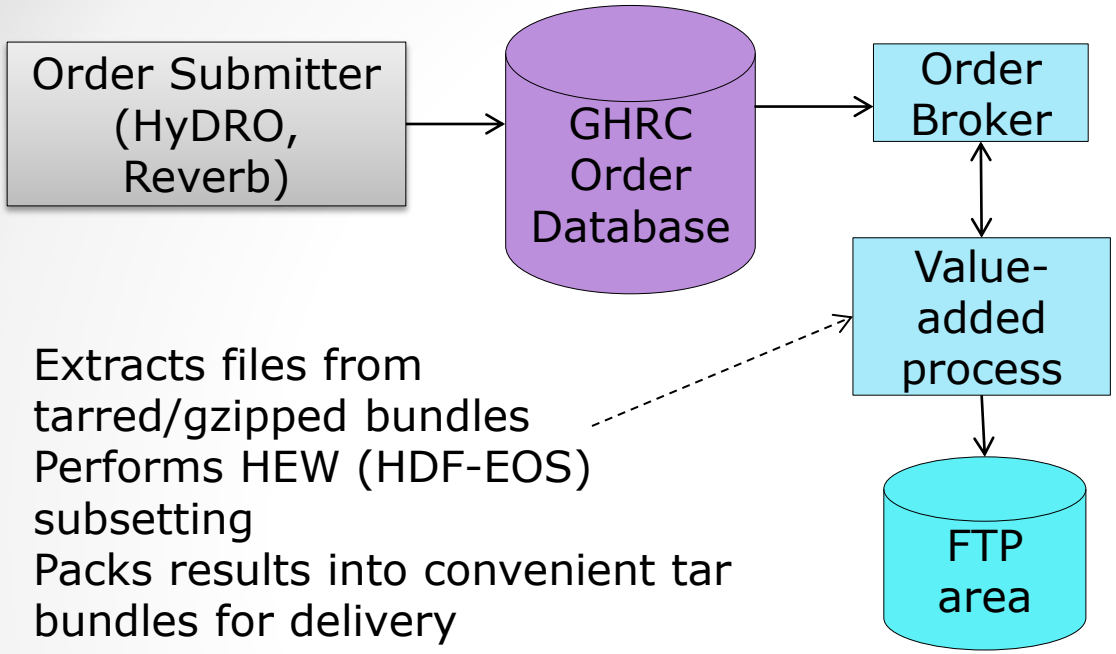
http://ghrc.nsstc.nasa.gov/uso/ds_docs/tpw/rssm1tpwn_dataset.html

Additional Access Methods

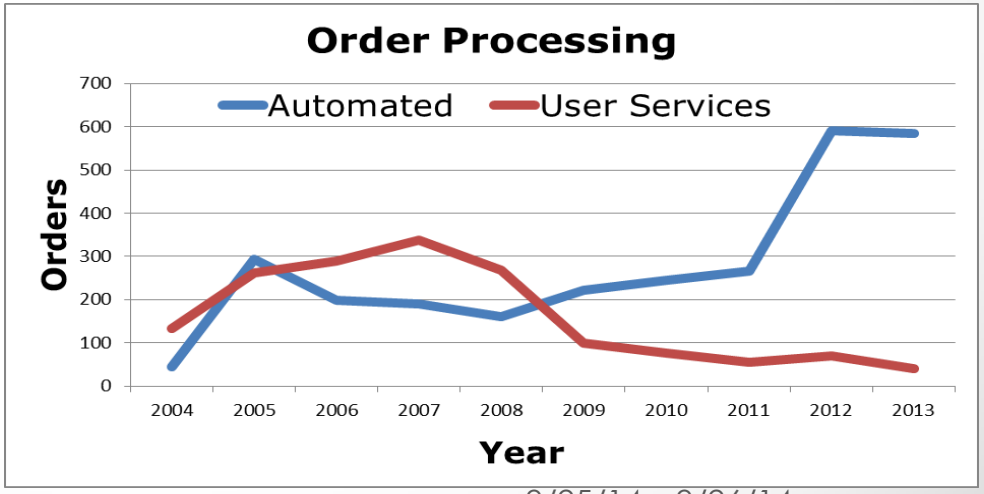
ftp://ghrc.nsstc.nasa.gov/
ftp://gpm.nsstc.nasa.gov/
http://ghrc.nsstc.nasa.gov/opendap/



Automated Order Processing

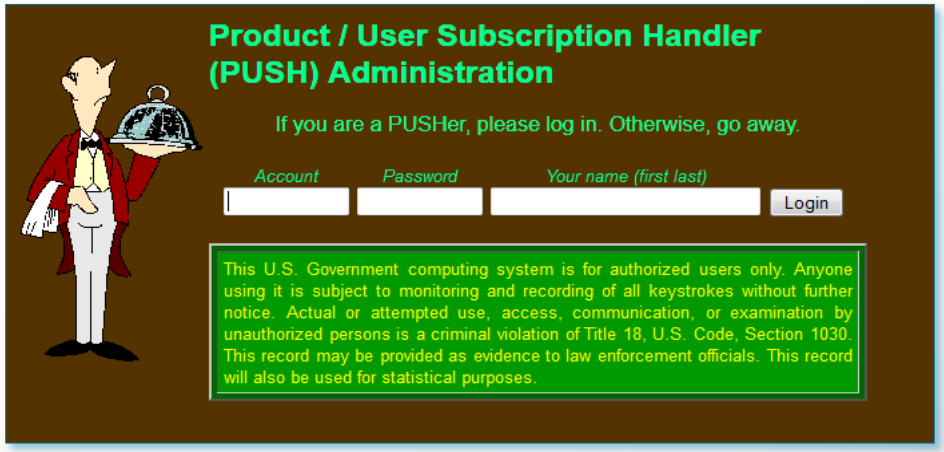


- Extracts files from tarred/gzipped bundles
- Performs HEW (HDF-EOS) subsetting
- Packs results into convenient tar bundles for delivery



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



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