



# HDF Group Support for NPP/JPSS

Mike Folk, Elena Pourmal, Larry Knox, Albert Cheng  
The HDF Group  
DEWG Meeting  
June 19, 2012



# Goal

---

Provide HDF5 risk-reduction support for the distribution of NPP/JPSS VIIRS, OMPS, and other sensor and environmental data products



# PROJECT INFORMATION



# Project Information

- Project Web site  
<http://www.hdfgroup.org/projects/npoess/>
- Project background
  - Started in 2009
  - NASA POCs are Richard Ullman and Alfreda Hall
  - Areas of primary needs
    - Easy and intuitive NPP/JPSS data access and use by diverse communities
    - Allow producers and consumers to view content, manage metadata, and convert data to other formats.
    - High-quality, rapid-response for NPP/JPSS users at all levels.



# Project Information

---

- 2011-2012 Priorities
  - Maintain HDF software and software developed for NPP/JPSS on platforms critical to NPP/JPSS project
  - Provide rapid and high priority support for data producers and users
  - Develop software to facilitate access and management of NPP/JPSS products



# **HDF5 SOFTWARE MAINTENANCE**



# HDF5 Software Maintenance

---

- HDF5 library and command-line utilities
- <http://www.hdfgroup.org/HDF5/>
  - C, C++ and Fortran90 APIs
  - Command-line utilities to view, compare and copy data stored in HDF5 files
    - h5dump, h5ls, h5diff, h5copy



## HDF5 Software Maintenance

---

- Latest release HDF5 1.8.9 on May 15, 2012
- For platforms and compilers supported for each release see <http://www.hdfgroup.org/HDF5/release/obtain5.html>
- Tested daily with NASA HDF-EOS5 and NPP/JPSS files from GRAVITE system





# HDF5 Software Maintenance

---

- Let us know your needs! We will help with:
  - Porting to new platforms and compilers
  - HDF software features
  - Performance tuning
- Report defects and suggest improvements to HDF products
- How can you reach us?



# PRIORITY USER SUPPORT



## Support for NPP/JPSS Users

- Priority support to NPP/JPSS users
  - Send email to [help@hdfgroup.org](mailto:help@hdfgroup.org)
  - Responding to “NPOESS”, “NPP”, or “JPSS” in subject field
    - Subject: *[NPOESS] Where is the plug-in?*
    - Subject: *[NPP] Why my program doesn't work?*
    - Subject: *[JPSS] Need a tool to edit files with references*
    - **This has worked well, though occasionally we have to guess**



## Support for NPP/JPSS Users

---

- Join us at the HDF and HDF-EOS Workshops
  - Tutorials
  - Consultations
  - Discussions of future directions
  - <http://www.hdfeos.net/workshops/index.php>



Software developed at The HDF Group for NPP/JPSS

# **SPECIAL PROJECTS**



# Software Developed for JPSS

The screenshot shows the HDF Group website. The top navigation bar includes links for Home, Why HDF?, Products, Services, About Us, News, and Contact Us, along with a search bar. The left sidebar contains a 'Quick Links' dropdown and a 'JPSS' section with links for Team Members, Software, Documentation, and Download. The 'Software' link is circled in red, and a red arrow points from it to the main content area. The main content area is titled 'IPO/JPSS SUPPORT (FORMERLY NPOESS)' and contains text about the project's aim, the JPSS system, and the HDF Group's support. The text is as follows:

[HOME](#) > [PROJECTS](#) > [NPOESS](#)

## IPO/JPSS SUPPORT (FORMERLY NPOESS)

The aim of this project is to develop and adapt HDF5 technologies and provide HDF5 support for the [Joint Polar Satellite System \(JPSS\)](#) and [NPOESS preparatory project \(NPP\)](#). JPSS is part of the restructured National Polar-orbiting Operational Environmental Satellite System (NPOESS). This project is funded through [NOAA's National Environmental Satellite, Data & Information Service \(NESDIS\)](#).

JPSS is a new generation of low earth orbiting satellites that monitor environmental conditions and provide data for long-range weather and climate forecasts. The NPOESS project was established by a 1994 Presidential Decision Directive to converge the polar-orbiting satellite systems operated by the Department of Commerce and by the Department of Defense into a single, cost-efficient integrated system. The new system would also take advantage of the advanced remote sensing and spacecraft technologies provided by NASA through its Earth Observing System.

The JPSS will circle the Earth approximately once every 100 minutes, delivering up to eight terabytes of data concerning the Earth's weather, atmosphere, oceans, land, and near-space environment. The data will be processed on the ground and distributed within 30 minutes of observation to its central users, and to the world within 24 hours.

The volume and complexity of the JPSS data, and the enormously high speed at which it must be processed, requires powerful technology and expertise in using this technology. JPSS data will be archived and distributed in HDF5.

The HDF Group provides support for JPSS/NPP data producers and users, as it has done for more than 15 years for the [NASA ESDIS program](#) and its estimated 1.6 million users of petabytes of earth science data.

---

The HDF Group JPSS project currently focuses on the following areas:

- **Improving accessibility and usability of JPSS/NPP data**

JPSS data will be distributed in HDF5 files containing raw data and indexing metadata that allows fast access to the raw data. The



# **“JPSS FRIENDLY” HDF5 TOOLS**



# HDFView

- HDF Java Browser
- <http://www.hdfgroup.org/hdf-java-html/hdfview/>
- Version 2.8 released in December 2011
- Platforms supported
  - Linux, Mac OS X, Windows, Solaris
- JPSS specific features
  - Displays values of quality flags
  - Navigates through object and region references





# JPSSSS Quality Flags in HDFView

Dataset with quality flags

Flags packed in 8 bits

Extracted bits (3, 4)

Extracted bits

The screenshot shows the HDFView application interface. On the left, a tree view displays the dataset hierarchy: SVI-NPP.h5 > All\_Data > VIIRS-I1-SDR\_All > QF\_VIIRS11SDR\_Array. The main window displays two TableView windows showing the data as a table. The top TableView shows the raw data, and the bottom TableView shows the data after bit extraction. A red circle highlights the 'Extract bits (3, 4)' option in the TableView menu. A red arrow points to the 'Extracted bits' section of the bottom TableView. A red circle highlights the '4' and '3' bit positions in the Bitmask section of the Dataset Selection dialog. The Dataset Selection dialog is open, showing the 'Display As' section with 'Spreadsheet' selected, the 'Index Base' section with '0-based' selected, and the 'Bitmask' section with 'Show Value of Selected Bits' checked. The 'Dimension and Subset Selection' section shows the dimensions of the dataset: Height (dim 0) 0 to 511, Width (dim 1) 0 to 6399, and Depth (dim 2) 0 to 0.

	0	1	2	3	4	5	6	7
0	0	0	0	1	0	2	0	3
1	40	87	43	147	47	109	47	2
2	0	0	0	0	0	0	0	0
3	38	50	41	184	48	48	50	1
4	0	0	0	0	0	0	0	0
5	35	32	39	59	41	236	47	1
6	0	0	0	0	0	0	0	0
7	42	254	48	48	49	230	55	8
8	49	180	49	189	49	189	49	1

QF\_VIIRS11SDR\_Array (3800, 4)  
8-bit unsigned character, 512 x 6400 x 1  
Number of attributes = 0



# HDF5 Object and Region References in HDFView

Dataset with HDF5 object references

The screenshot shows the HDFView application interface. On the left, a tree view displays the dataset structure for 'SVI-NPP.h5'. The 'radiance\_Array' dataset is highlighted. The main pane shows a table view of the selected region, with a red circle highlighting the value '2928.292...'. A red arrow points from this value to the 'radiance\_Array' dataset in the tree view. Another red arrow points from the 'radiance\_Array' dataset to the 'radiance\_Array' dataset in the tree view. The bottom pane shows the metadata for the selected dataset, including the object reference, number of attributes, and aggregate beginning date.

Table view showing data pointed by region reference:

	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	12614	12623	12623	12623	12614	12623	12623
5	12050	12050	12059	12050	12068	12041	12059
6	12764	12801	12810	12792	12783	12801	12783
7	12642	12642	12641	12632	12651	12651	12641
8	12869	12896	12878	12887	12887	12878	12878
9	13215	13224	13224	13224	13224	13206	13233
10	12764	12746	12764	12746	12783	12755	12737
11	13060	13033	13060	13051	13051	13051	13060

Table view showing data of the selected region:

	0
0	2928.292...
1	3528.252...
2	3800.380...
3	4072.407...
4	4344.434...
5	4616.461...
6	4888.488...
7	5160.516...
8	5432.543...
9	6032.603...
10	6304.630...

Table view showing data of the selected region:

	0
0	2928
1	3528
2	3800
3	4072
4	4344
5	4616
6	4888
7	5160
8	5432
9	6032
10	6304

Metadata for VIIRS-I1-SDR\_Aggr (111503304, 4):

- Object reference, 11
- Number of attributes = 9
- AggregateBeginningDate = 2003125
- AggregateBeginningTime = 111503304.000000

Data of the selected region



# h5dump

- h5dump is a command line tool to display the content of an HDF5 file
- Example: dump metadata information (no data displayed)

```
h5dump -H SVI-NPP.h5
```

- Example: display datasets pointed by object references

```
h5dump -d /Data_Products/VIIRS-I1-SDR/VIIRS-I1-SDR_Aggr -R SVI-NPP.
```

- Example: display quality flags

```
h5dump -d /All_Data/VIIRS-I1-SDR_All/QF_VIIRSI1SDR_Array -M 3,1 SVI-NPP.h5
```



# h5edit

- h5edit is a command line tool to create and edit attributes
  - The tool is under development; more functionality is coming
- Example: add “scale\_factor” attribute

```
h5edit -c "CREATE /Radiance scale_factor  
{H5T_IEEE_F32LE SIMPLE(1) DATA{2.8339462E-4}};"  
file.h5
```

- Example: add “units” attribute:

```
h5edit -c "CREATE /Longitude units {H5T_STRING  
{ STRSIZE 12 } DATA {'degrees_east'}};" file.h5
```



## High-Level Library

- Functions for conveniently handling HDF5 object and region references and packed bits (quality flags)
  - C and Fortran API to support NPP/JPSS data
  - Examples:
    - Create a region reference to a rectilinear region
    - Create and write a datasets with object references
    - Read quality flags
  - Fortran functions require HDF5 1.8.5 and later
  - Tested daily on Linux and AIX



# JPSS SOFTWARE



## JPSS Software

---

- Software developed at The HDF Group to support the JPSS project
  - High-level library for handling HDF5 object and region references
  - Augmentation tool h5augjpss
  - Aggregation tool nagg

[http://www.hdfgroup.org/projects/npoess/jpss\\_software.html](http://www.hdfgroup.org/projects/npoess/jpss_software.html)



## Augmentation tool – h5augjpss

---

- h5augjpss is a command line tool to change an HDF5 JPSS file to make it accessible by netCDF-based applications
  - Hide HDF5 elements not supported by netCDF applications
  - Add metadata or data needed by netCDF or netCDF applications
- Attention!
  - h5augjpss modifies the file! Make a copy if you need to preserve original data





# Clarification

- **netCDF-3 files**
  - Based on netCDF **classic** data model
- **netCDF-4 files**
  - Based on netCDF **enhanced** model
  - Uses HDF5 as a storage layer
  - Group hierarchy, user-defined data types, etc.
  - But can be restricted netCDF classic
- **NPP files**
  - HDF5 file with primary data
    - Incompatible with netCDF, unless modified
  - XML metadata file
    - Important information, including dimensions
  - Geo data in separate file, or group in primary file



# Example: ncdump of an augmented file

```
netcdf SVM07_ter_augmented-step1-step2 {  
  // global attributes:
```

```
  .....
```

```
  group: All_Data {
```

```
    group: VIIRS-M7-SDR_All {
```

```
      dimensions:
```

```
        AlongTrack = 768 ;
```

```
        CrossTrack = 3200 ;
```

```
        ...
```

```
      variables:
```

```
        int AlongTrack(AlongTrack) ;
```

```
        ...
```

```
        float Radiance(AlongTrack, CrossTrack) ;
```

```
          Radiance:DatumOffset = 0 ;
```

```
          Radiance:Scaled = 1 ;
```

```
          string Radiance:MeasurementUnits = "W/(m^2 μm sr)"
```

```
          ...
```

```
    } // group VIIRS-M7-SDR_All
```

```
  } // group All_Data
```

```
}
```

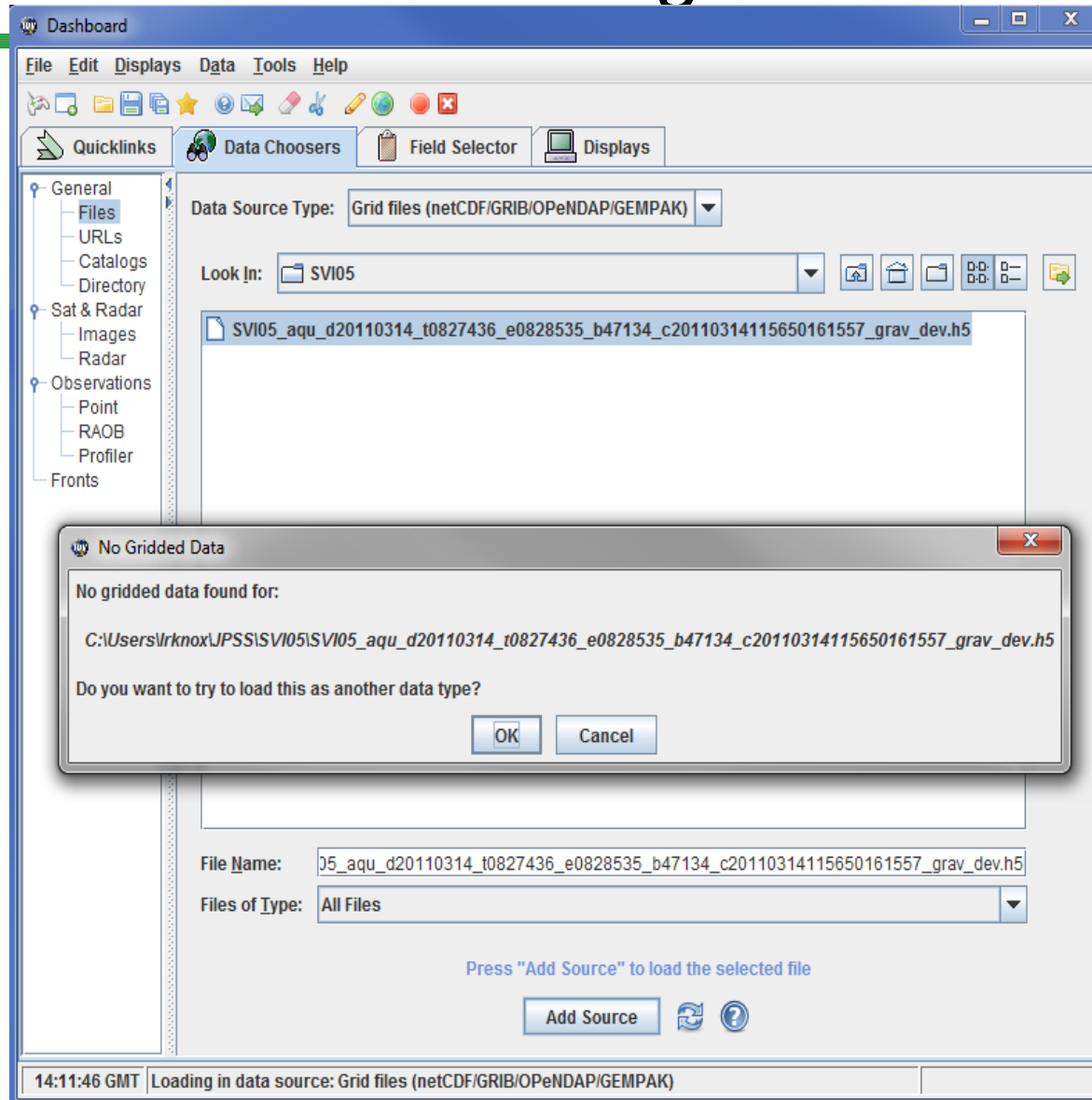
**Meaningful dimensions**

**Coordinate variables**

**Product attributes**



# IDV before augmentation





## Two problems

- Has to be netCDF-3 conformant

```
h5augjpss -o4 SVI05.h5
```

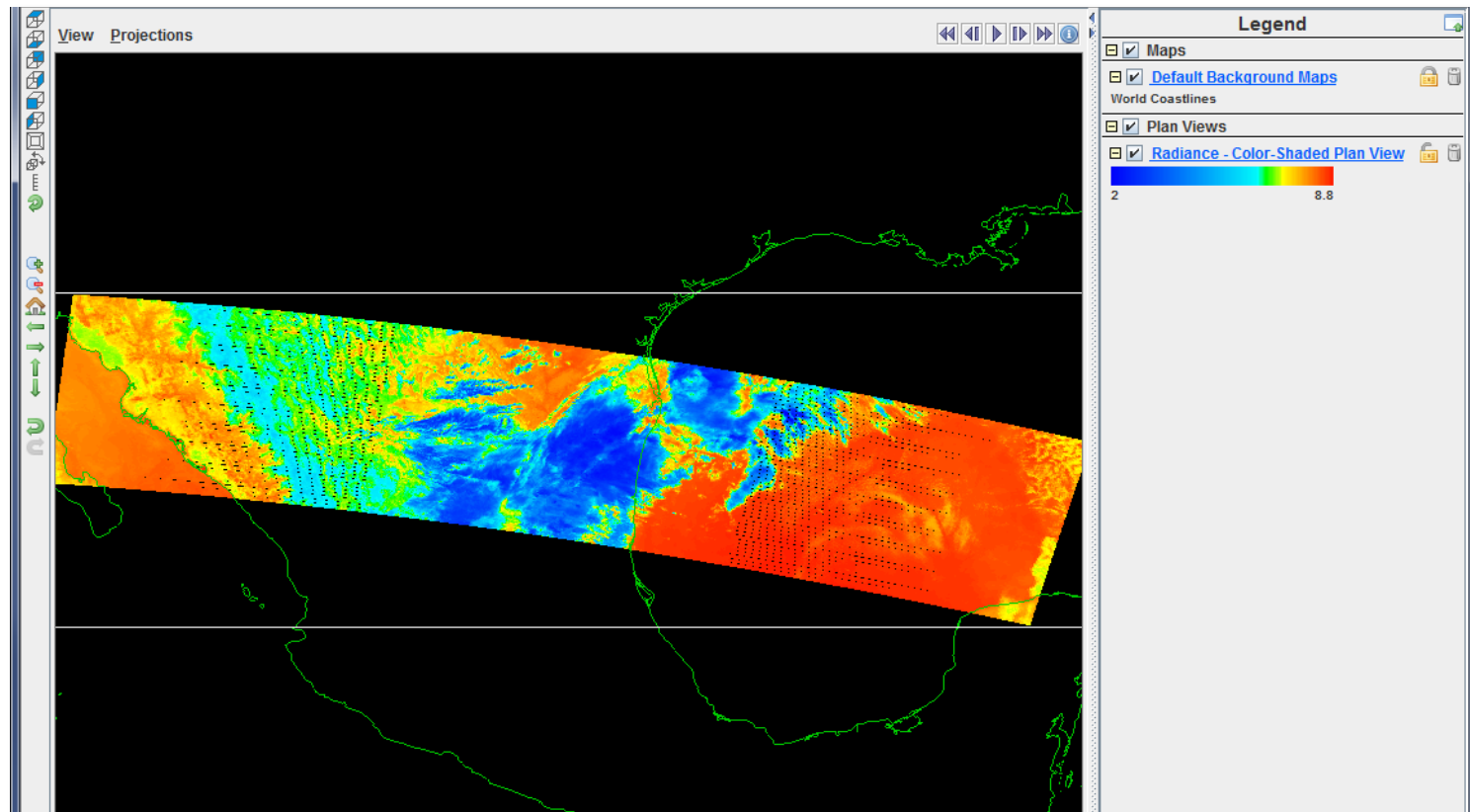
- Missing CF metadata

```
h5edit -c "CREATE /Radiance scale_factor  
{H5T_IEEE_F32LE SIMPLE(1) DATA  
{2.8339462E-4}};" SV105.h5
```

```
h5edit -c "CREATE /Longitude units {H5T_STRING  
{ STRSIZE 12 } DATA {'degrees_east'}};"  
SV105.h5
```



# IDV after augmentation



IDV with augmented JPSS file



## Aggregation tool – nagg

---

- nagg is a command line tool for aggregating JPSS data granules from existing files into new files with
  - Different number of granules per file or
  - Different combinations of compatible products than in the original files



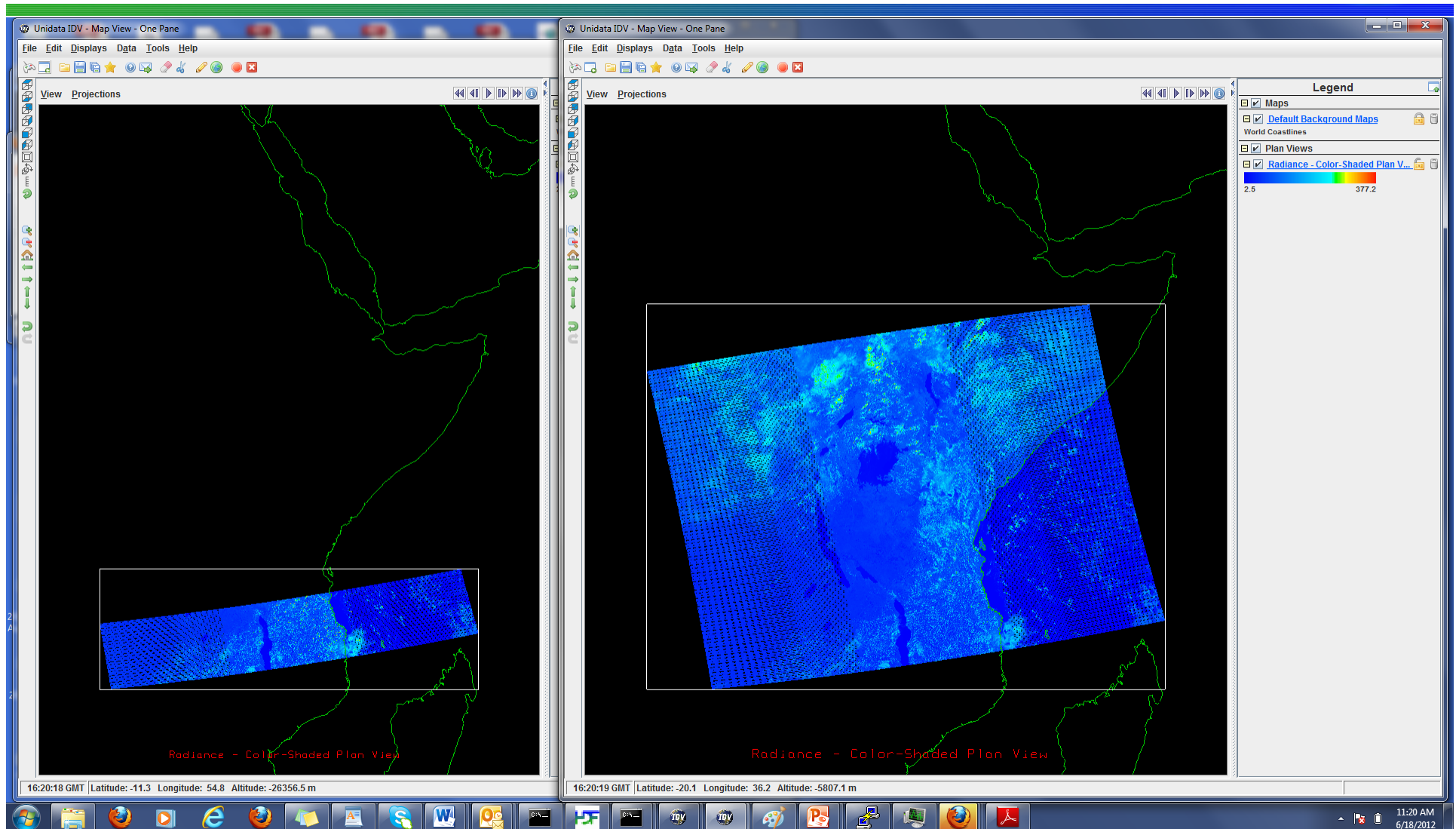
## Why nagg?

- NPP data products organized as “granules.”
- Granules are relatively small.
- Several granules may be packaged per file.
- Several products may be packaged per file.
- For convenience of a particular application, we may want to re-package them.
- May also want only a subset of them.

(Thanks to Richard Ullman)



# Aggregated file in IDV

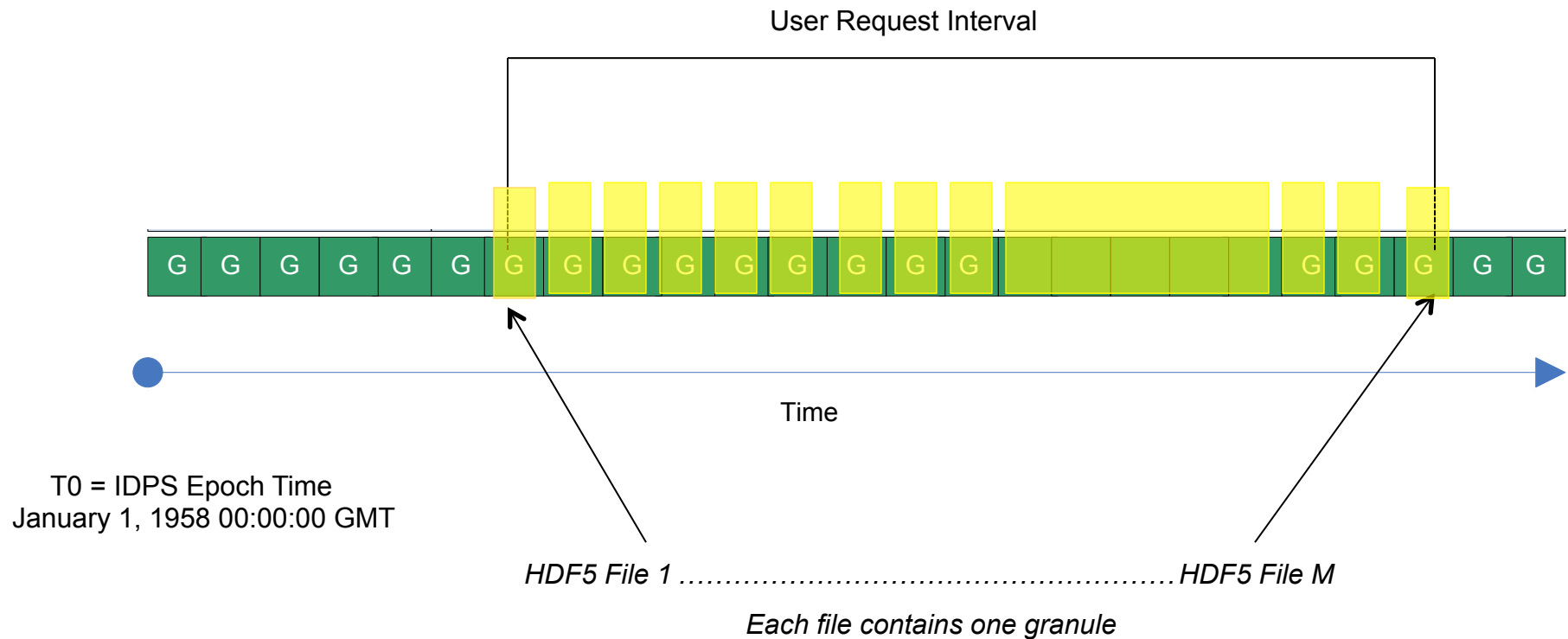


One granule is displayed on the left; four aggregated granules are displayed on the right





# Aggregation Example

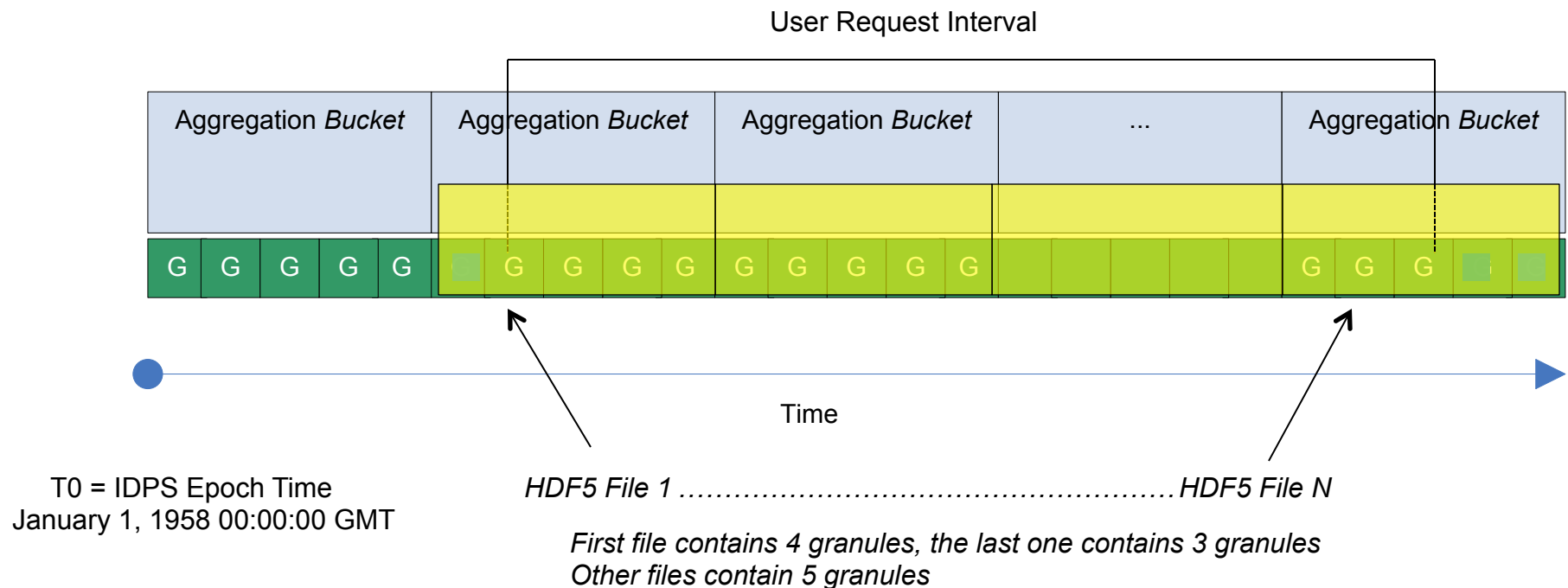


- User requests data from the IDPS system for a specific time interval
- Granules and products are packaged in the HDF5 files according to the request
- This example shows one granule per file for one product



# Aggregation Example

Example: `nagg -n 5 -t SATMS SATMS_npp_d2012040*.h5`

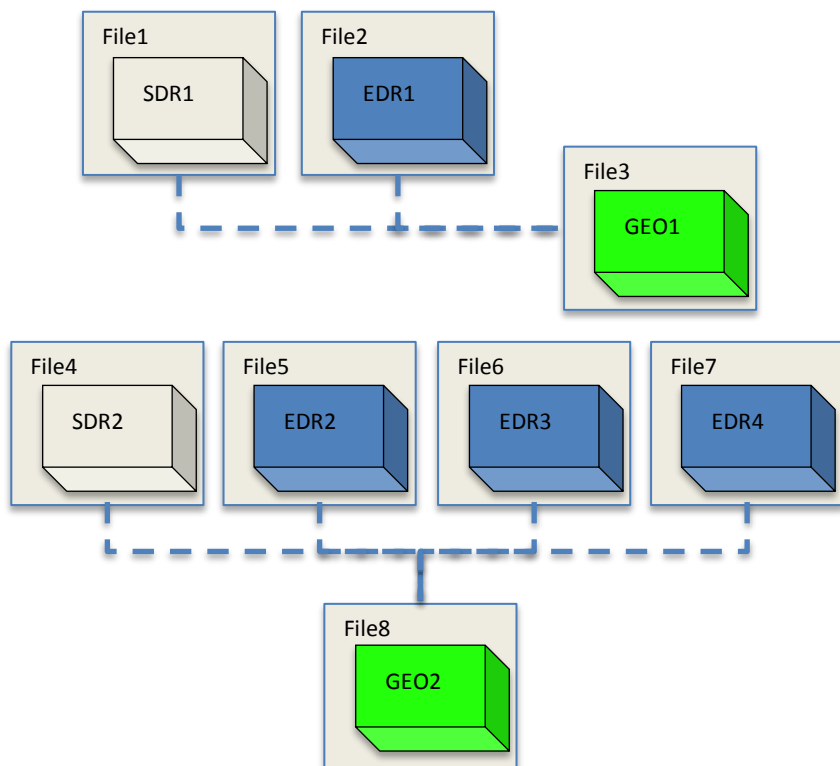


- Produced files co-align with the aggregation bucket start
- HDF5 files are 'full' aggregations (full, relative to the aggregation period)
- Geolocation granules are aggregated and packaged; see `-g` option for more control



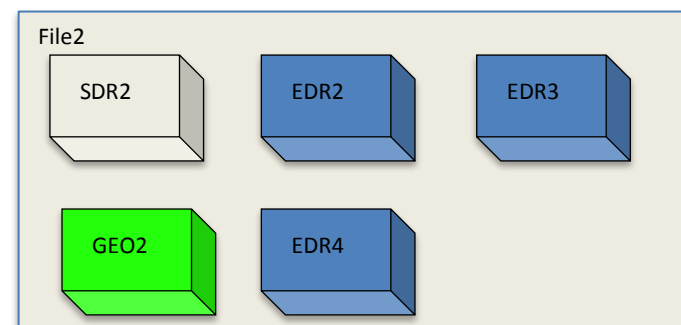
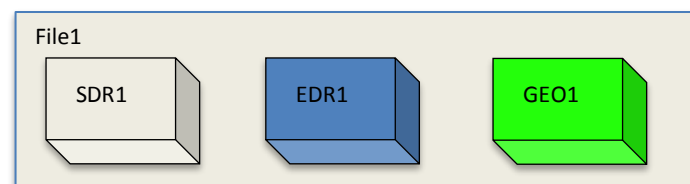
# IDPS Packaging Baseline

## CDFCB-X Volume I



Packaging Off

Packaging only applies to products with geolocation data



Packaging On

Always makes a new copy.  
Doesn't destroy the original  
file.



# OTHER SOFTWARE OF INTEREST



# H5Py

- HDF5 Python APIs
- <http://code.google.com/p/h5py>
- Written and supported by Andrew Collette (not The HDF Group)
- Contains
  - High-level Python methods for HDF5
  - Low-level (follows C) Python APIs
- Easy to learn and use



## H5Py Example

- Traverse the file, print object name and names and values of its attributes

```
import h5py
def print_info(name, obj):
    print name
    for name, value in obj.attrs.iteritems():
        print "    "+name+":", value

f = h5py.File('GATMO-SATMS-npp.h5', 'r+')
f.visititems(print_info)
f.close()
```



## H5Py Example Output

.....

```
Data_Products/ATMS-SDR/ATMS-SDR_Gran_0
  Ascending/Descending_Indicator: [[1]]
  Beginning_Time: [['040230.416997Z']]
  N_LEOA_Flag: [['On']]
  Band_ID: [['N/A']]
  Beginning_Date: [['20030126']]
  East_Bounding_Coordinate: [[-19.96759987]]
  Ending_Date: [['20030126']]
  Ending_Time: [['040301.825163Z']]
  G-Ring_Latitude: [[ 25.41589928]]
```

..... . .



# **FUTURE DEVELOPMENT EFFORTS**





## Future Efforts

---

- Finish nagg development
  - Add missing command line flags
  - Enhance test suite
  - Improve documentation
- Continue development of h5edit
- Continue with software maintenance and priority user support



- Questions?
- Requests?



The HDF Group



# Thank You!