GRAVITE ICF User Guide

By: David Tagatac - david.tagatac@nasa.gov
Updated: 7/14/2011

The GRAVITE Investigator Computing Facility (ICF) consists of several Red Hat Enterprise Linux servers for use by NPP Calibration/Validation teams and their colleagues. In addition to a home area, and some storage space, there exist on the GRAVITE ICF several tools to facilitate common interactions with the GRAVITE file catalog system, dubbed Data Storage Distribution (or DSD). This document walks through the steps to gain access to the GRAVITE ICF, to prepare your account for working with our tools, and to work with those tools.

Gaining Access to the GRAVITE ICF

Access to the GRAVITE ICF requires the submission of five forms:

1. GRAVITE Access Request Form – general access request
2. GRAVITE LAN Remote User Security Agreement Form – details about the computer you will use to access the ICF
3. GRAVITE Remote Access Token Request Form – request for an RSA token
4. I-9 – to verify permission to access ITAR data
5. NDA – to protect confidential program information

Contact your team lead or Janet Fennema (janet.t.fennema@nasa.gov, 240-684-0798) to obtain these forms and for instructions on where to send them.

Login to GRAVITE

The GRAVITE lab is located in room S625 in the GreenTech-IV building.

- Login to a machine with your user name and password.
- ssh to any of the procs from 12 – 17. (for example proc12)

Instructions for accessing via VPN are in the document "GRAVITE VPN-RSA Instructions_v4.pdf" which is distributed with this document.

Setting Up The GRAVITE Environment Variable

The first step to using these tools is making sure that the environment variable $GRAVITE is set, and the script $GRAVITE/cfg/gravite_cshrc has been sourced. To automate this, open your ~/.cshrc file for editing and add the following lines in the following order:

- *setenv GRAVITE /apps/gravite_ops*
- *if (-e $GRAVITE) source $GRAVITE/cfg/gravite_cshrc*

Your user area is linked as /home/<your user name> on all of the servers in the GRAVITE ICF network. Once you have added the lines above to your .cshrc file, you will have access to the tools discussed in this document upon your next login to the GRAVITE ICF. You can verify that this is the case after login by running the following command:

- *echo $DSD*

which should return */apps/gravite_ops/dsd*.

Which Server Should You Use?

The answer to this question depends on what you plan to do. For most users and most use cases, the answer will be one of the proc servers. In fact, any of the procs 4-17 will do. In order to login to a proc server type, for example `*ssh –X proc17*`. Here are some things to be aware of:

- Prior to launch, procs 1-12 are being used to generate proxy data. Therefore, prior to launch, we recommend using one of the newer procs 13-17.
- After launch, procs 1-3 will be used for pulling external data and internal testing. Therefore, after launch, we recommend using one of the procs 4-17.
- Our SA team has recommended keeping user areas under 100GB each. If you find that you need more space than this, there is a possibility to store data in a user area on our raid arrays. Create a folder with your name under /daq3_raid3/user/ (which is linked to the folder /data3 on server daq3). Keep in mind that, although we do not as yet put hard limits on how much each user can store in this area, space is not unlimited. **Please take care to clean up data that you no longer have a need to access on GRAVITE in the near future**.
- In addition to the proc servers, the ICF includes several dev servers (e.g. `*ssh –X dev1*`). These servers are a bit slower than the proc servers, and as yet they do not provide access to many of the tools described in this document. However, in certain cases, they host software that is not in the proc baseline. For example, IDL 8.0 is only installed on dev4 at the moment.

Listing, Downloading, and Obtaining Metadata about DSD Files

From Outside the ICF Network

When outside the ICF network, the best way to list, download, and obtain metadata about DSD files is using the GTP (GRAVITE Transfer Protocol) client. The download link for GTP is https://casanosa.noaa.gov/svn/gtp/gtp-client/trunk/GTPClient/release/. See this blog post for information on how to get started with GTP: Getting Started with GTP https://casanosa.noaa.gov/weblog/weblogger.php?entry_id=3771&weblog_id=190. Michael Iwunze <michael.iwunze@nasa.gov> maintains GTP. Let him know if you have any issues with GTP specifically.

Other possibly useful posts:

- **GTP Basics**
  https://casanosa.noaa.gov/weblog/weblogger.php?entry_id=3772&weblog_id=190
- **GTP Examples**
  https://casanosa.noaa.gov/weblog/weblogger.php?entry_id=3773&weblog_id=190
- **GTP Conditional Queries and the Virtual Directory Structure**
  https://casanosa.noaa.gov/weblog/weblogger.php?entry_id=3774&weblog_id=190
- **GRAVITE Data Type Descriptions**
  https://casanosa.noaa.gov/weblog/weblogger.php?entry_id=3813&weblog_id=190
- **Change Your GTP Password**
  https://casanosa.noaa.gov/weblog/weblogger.php?entry_id=3840&weblog_id=190

From Inside the ICF Network

When logged into the ICF network, the preferred method for listing, downloading, and obtaining metadata about DSD files is an internal tool called "gdata". After sourcing gravite_cshrc (see above), gdata will be on your path. This tool provides a variety of functions for interacting with the GRAVITE Data Storage Distribution. The most common functions are:

- gdata types – lists the data types present in DSD
- gdata list – lists a given number of files of a given data type
- gdata get – downloads a DSD file to the working directory
- gdata tree – lists the files that were generated from a given file (e.g. if a MODIS L0 file is specified, the granulated L0 files, L1B files, VIIRS SDRs, and VIIRS custom processed files will all be listed)
- gdata origin – lists the file most immediately upstream of the given file in GRAVITE processing
- gdata insert – inserts a given file into DSD from the working directory

For more details about these functions, as well as other functions that the tool performs, simply run gdata from any proc server with no arguments.

Setting Up a User Area for Development

(from https://casanosa.noaa.gov/docman/wiki_doc.php?id=2597&group_id=342)

GRAVITE is configured to use some key environment variables (e.g. $GRAVITE, $DSD, $GPROC, $NPROX). These environment variables define absolute paths to the various packages that make up GRAVITE, and allow the scripts to be abstracted from these paths. For development purposes, the scripts also check for user environment variables (e.g. $UGRAVITE, $UDSD, $UGPROC, $UNPROX), and if the user environment variables exist, they are used instead of the production environment variables. This guide walks through the recommended steps for setting up a user area for development and testing.

1. Log into GRAVITE.
2. Check out the package(s) you are interested in developing.
   - Open a terminal
   - *$ cd ~*
   - *$ mkdir guser*

- *$ cd ~/guser*
- Run `*$ svn ls https://casanosa.noaa.gov/svn/gravite/packages*` for a listing of the packages
- Run `*$ svn co https://casanosa.noaa.gov/svn/gravite/packages/cfg/trunk cfg*`
- For each other <package> you are interested in, run `*$ svn co https://casanosa.noaa.gov/svn/gravite/packages/<package>/trunk <package>*` **WARNING: Be sure to only check out the trunk. If you check out the package, you will get a copy of every tag and branch - many times the size of the trunk.** For more details on SVN, run `*$ svn help*`. For even more details, please see the online manual at http://svnbook.red-bean.com/en/1.4/index.html.
- For some packages (e.g. nprox), there are further instructions in a top-level readme file explaining how to install low-level dependencies (e.g. SeaDAS and ATMS_proxy_gen) if desired.

3. Edit your executable scripts to point to your user area.
    - *$ cd ~/guser/<package>/bin* (or in nprox: `*$ cd ~/guser/<package>/<sensor>/bin*`)
    - Edit the executable scripts that you wish to run from your user area instead of from the production area.
        - This almost always consists of changing the environment variables in those scripts from production environment variables to user environment variables (e.g. from $NPROX to $UNPROX).
        - There is usually only one environment variable to change. The binary scripts are usually no more than 5 lines.
4. Edit your ugravite_cshrc file.
    - *$ cd ~/guser/cfg*
    - *$ $EDITOR ugravite_cshrc*
    - Uncomment the source line for all installed packages you wish to develop.
    - Save and quit.
5. Set your ugravite_cshrc file to be automatically sourced at login.
    - *$ cd ~*
    - *$ $EDITOR .cshrc*
    - *You should already have the following lines in your .cshrc file to run the production version of GRAVITE. If not, add them to the end of the file now:
        - *setenv GRAVITE /apps/gravite_ops*
        - *if (-e $GRAVITE) source $GRAVITE/cfg/gravite_cshrc*
    - Add the following two lines immediately before the lines to source the production paths:
        - *setenv UGRAVITE ~/guser*
        - *source $UGRAVITE/cfg/ugravite_cshrc*
    - Note that in order to go back to the production GRAVITE at any time, you need only comment out those two lines which source the user area paths.
    - Save and quit.
6. Relog into GRAVITE to refresh your paths and environment variables.
    - If you receive an error upon opening a terminal, there is likely a problem with your .cshrc file.
    - Check your user environment variables (e.g. `*$ echo $UGRAVITE*`), and your path (e.g. `*$ echo $PATH*`, `*$ which vprox*`) to make sure that they are what you expect.
7. You are now set to run your own versions of the packages you have specified in your ugravite_cshrc file. These packages are located in $UGRAVITE (~/guser). Modify them as you need to for your development and testing purposes. Your checkouts are working copies of the production trunk (and can be updated via `*$ svn up*`), so **do not commit your changes unless you really intend them to be included in production GRAVITE immediately.**

If you do not know what this means, please do not commit anything to the server until a GRAVITE team member has a chance to explain it to you. Also, never commit anything to the tags folders other than a new tag copied from a branch or the trunk.

Some of the future enhancements to this document:

- A comprehensive "GRAVITE User's Guide" that includes info on GTP, NSIPS Web Portal, Common CM, Findings Tool, ADA, and CasaNOSA access.
- Include gproc instructions.
- Include instructions for checking code in and out of the CasaNOSA Subversion repository (this is NOT a user's guide for Subversion, just enough essential information for the user to be able to interact with the CasaNOSA Subversion repository).
- Instructions for uploading/ingesting data on GRAVITE, and sharing user-produced datasets.
- Include information on who to contact with suggestions, questions, and for support.