Accessing the UNAVCO Dataworks for GNSS

May 1, 2016

Contents

Installing Dataworks for GNSS from the UNAVCO **Gitlab**.

Installing Dataworks for GNSS on **AWS**
using the Dataworks for GNSS **AWS Image**.

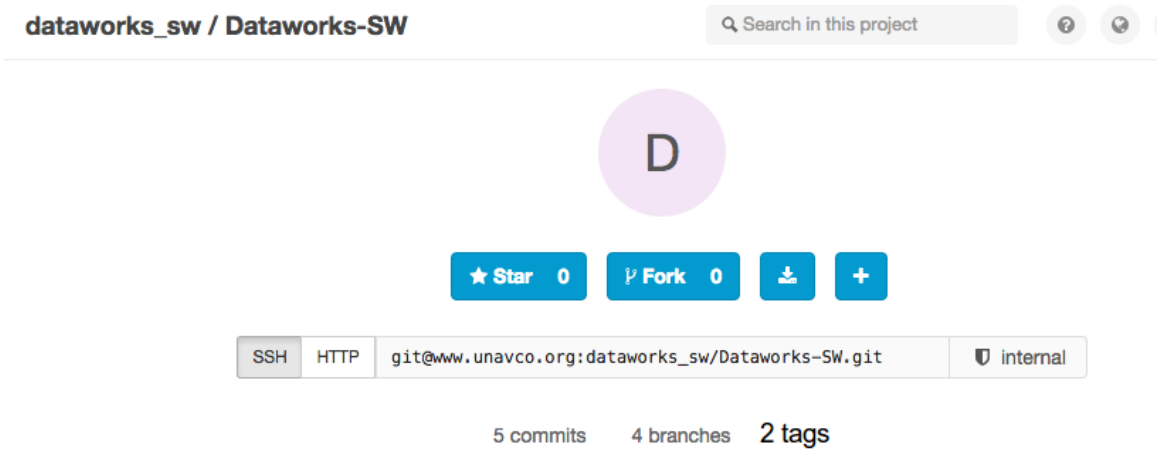
Installing Dataworks for GNSS from the UNAVCO **Gitlab**.

To install Dataworks for GNSS with source code from Gitlab you will need a server able to run online web applications (typically with Tomcat), with Linux, with "svn" and "ant", and where you can install a MySQL database and build Java 7 code. See all the system requirements for Dataworks for GNSS in the Dataworks for GNSS manual at the UNAVCO Dataworks for GNSS website,
http://www.unavco.org/software/data-management/dataworks/lib/docs/Dataworks_for_GNSS.pdf.

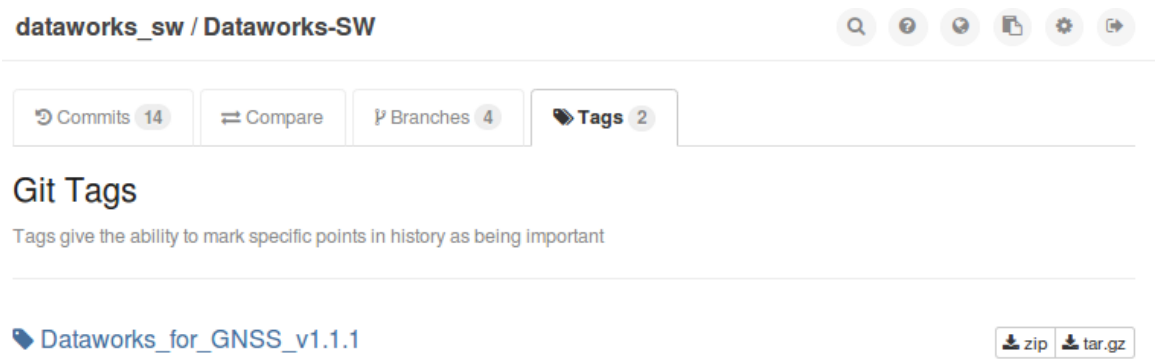
System setup instructions are in:
http://www.unavco.org/software/data-management/dataworks/lib/docs/Dataworks_System_Services_Guide.pdf.

To download Dataworks software from Gitlab, use a Chrome browser. Go the UNAVCO GitLab web site at http://www.unavco.org/gitlab/users/sign_in.

Ask for a new account. There may be a delay, waiting to get approval. Then go to the project page for "Dataworks-SW":



Click on "**2 tags**" to see the page of tags:



Click on one of the choices to download the Dataworks for GNSS package of source code. This file is 110 MB in size. Save the file, with a name like Dataworks-SW-0e76527022ff9e52c681ab6aa77365a5ba8.tar.gz,
to a directory where you will build Dataworks.

In that directory unpack the Dataworks for GNSS file. Do command cd Dataworks-SW.git (Dataworks-SW.git is a directory, not a file), where you will see these directories and a README file:

Click on the *most recent* tag, in this example, Dataworks_for_GNSS_v1.1.1.

Note that this tag, Dataworks_for_GNSS_v1.1.1, is for the COCONet Community Workshop, May 2016.

In future UNAVCO will provide newer GitLab production tags for Dataworks for GNSS. Find them by consulting the installation guide in the UNAVCO Dataworks for GNSS web site.

Use the "**zip**" or "**tar.gz**" buttons to download a Dataworks package of source code. Click on either one of the choices to download. Save the file, with a name like Dataworks-SW-0e76527022ff9e52c681ab6aa77365a5ba8.tar.gz,
to a directory where you will build Dataworks.

In that directory unpack the Dataworks file. Then do command

cd Dataworks-SW.git/Dataworks-SW/

in which folder you will see these directories and a README file:

```
-rw-r—r    README_Dataworks_Installation_Operation
drwxr-xr-x database
```

```
drwxr-xr-x dataworks-gsac
drwxr-xr-x documentation
drwxr-xr-x downloader
drwxr-xr-x mirror
```

To build operating software (GSAC) and to operate Dataworks for GNSS, begin with the instructions in

- the top level "README" file, `README_Dataworks_Installation_Operation`
- the Dataworks for GNSS manual at http://www.unavco.org/software/data-management/dataworks/lib/docs/Dataworks_for_GNSS.pdf

There are several README files in the Dataworks for GNSS package, one in each of the above-listed directories. Read all of them.

To customize the appearance of the GSAC web site for your agency, see the file README_Dataworks_Installation_Operation.

Installing Dataworks for GNSS on **AWS**
using the UNAVCO Dataworks for GNSS **AWS Image**

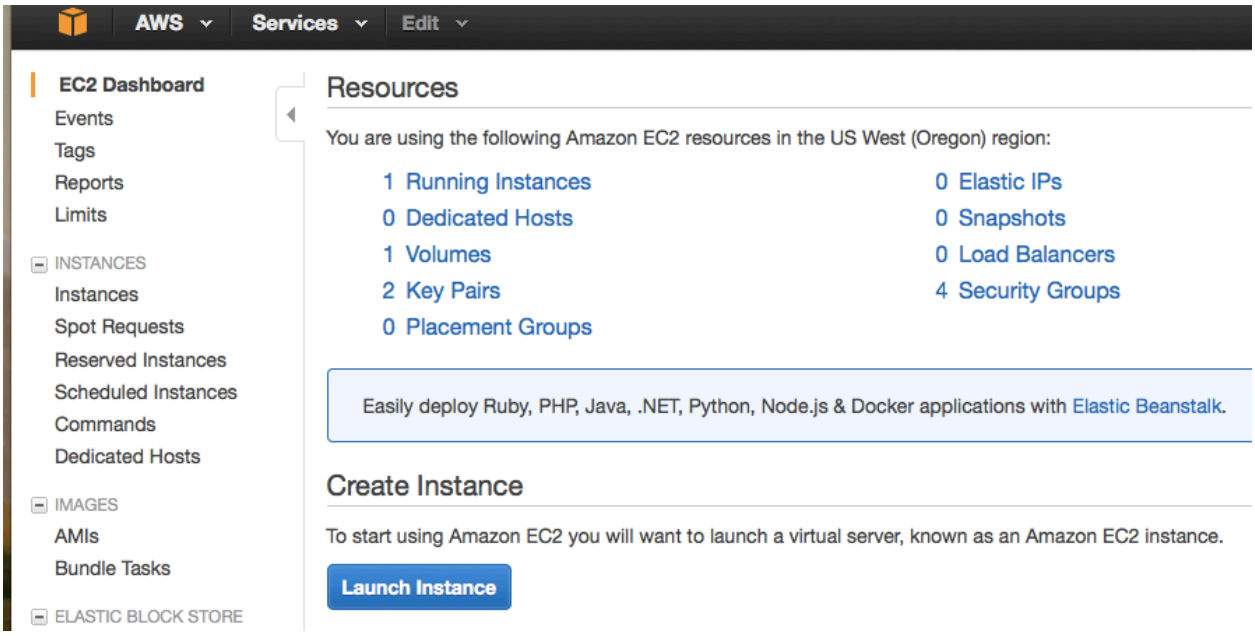
For a small cost, you can run Dataworks for GNSS on Amazon Web Services. This is, in effect, a complete computer system including all the software you need to run Dataworks for GNSS including GSAC, the database, FTP downloads, and a connection to the internet.

Advantages of using an "AWS instance" of Dataworks for GNSS include having already built and operating software, and having a fast and reliable internet connection. You do not need to own and operate a server.

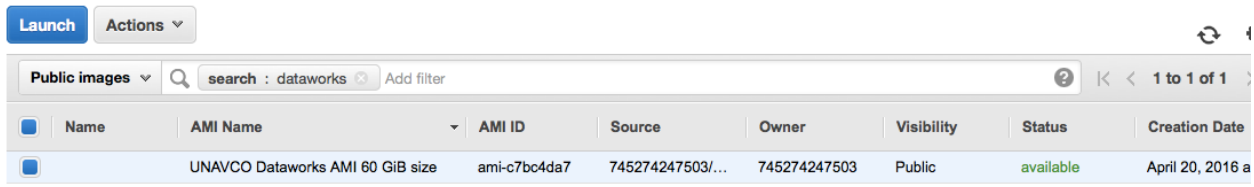
You can estimate your monthly cost by using the AWS Simple Monthly Calculator. We recommend using a Chrome browser. Go to http://calculator.s3.amazonaws.com/index.html. The Dataworks for GNSS installation outlined here should cost approximately $25 US / month.

First make a new AWS account. Go to https://aws.amazon.com/ and create a new account, but not a "free tier" account. You will need a credit card, and a telephone number, which Amazon will call from Seattle Washington USA when creating your account. You will need to answer the phone call, and follow one instruction in English about a security number.

Go to your account on the AWS web site; and go to the "EC2" dashboard:

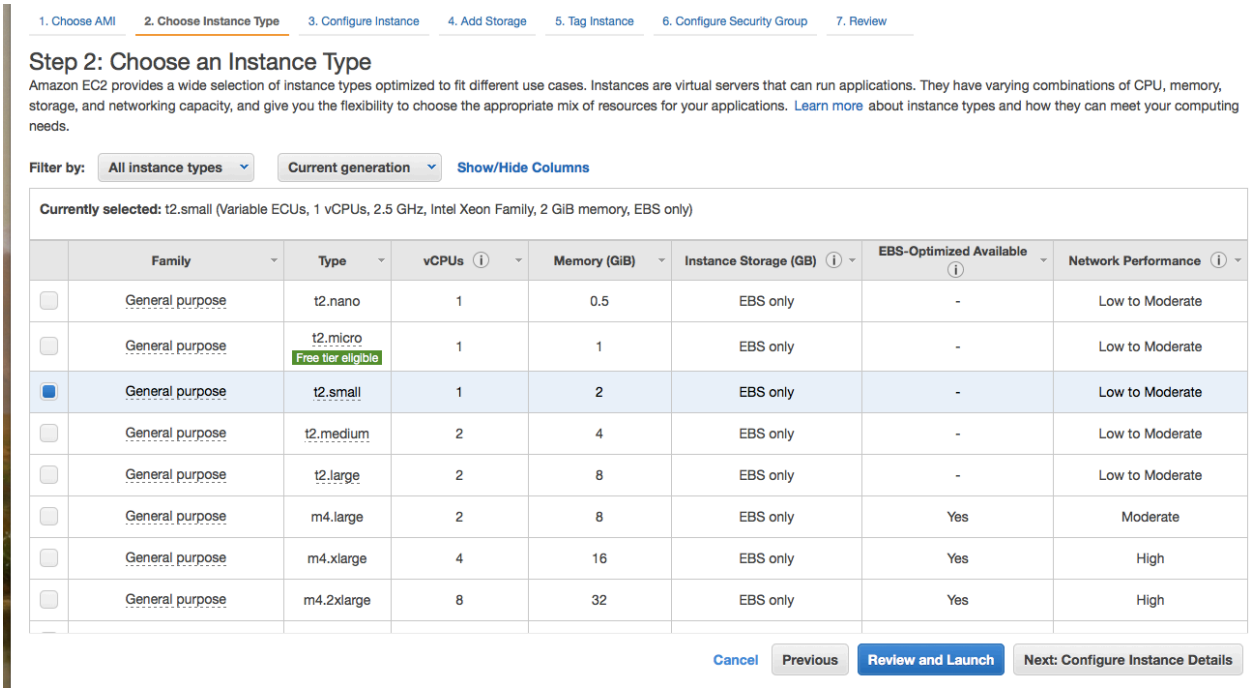


To "launch" a Dataworks for GNSS as a new "AWS instance" to make a running Dataworks) in your account, next click on AMIs in the left menu, to see this page. (*do not click* on the blue button "Launch Instance" or on menu item "Instances"!) You will choose which AMI for the source.



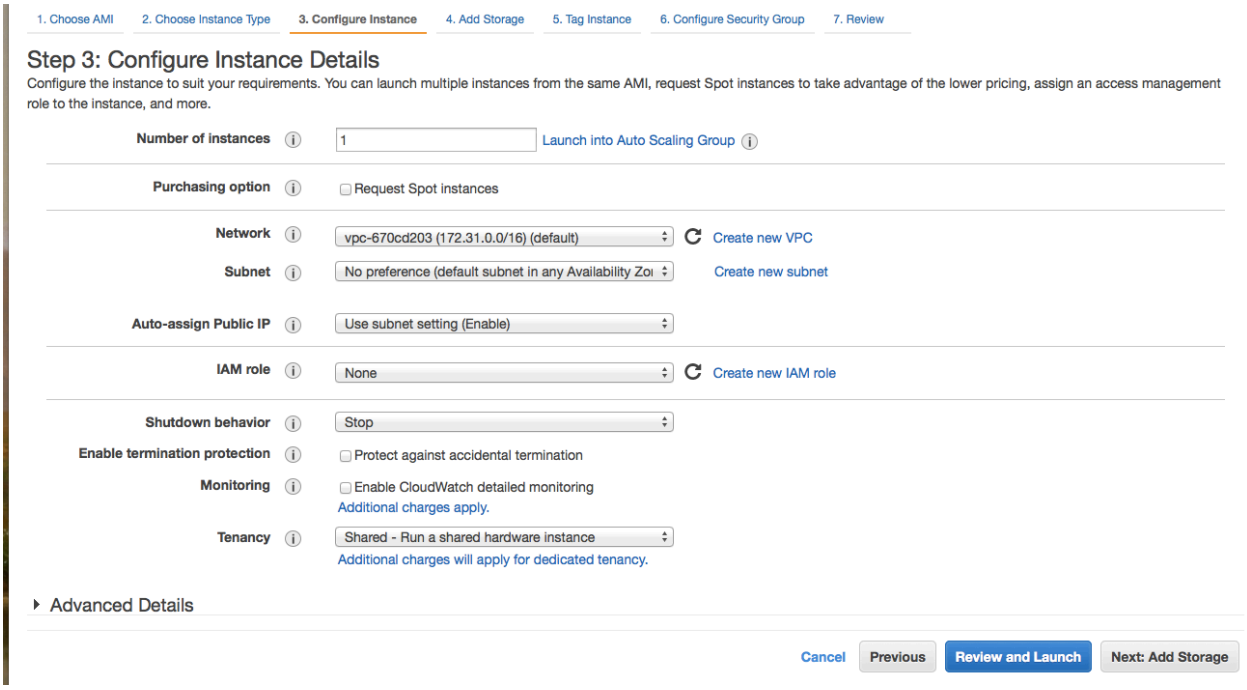
Find and choose the AMI from UNAVCO to build your AWS instance. (An AMI is an Amazon Machine Image, with all the code to make a running "instance." An 'image' is not a running virtual server.) Change the choice **Owned by me** to **Public images**. In the search box, enter *dataworks* and click enter. The item named "UNAVCO Dataworks AMI 60 GiB size" will be listed under AMI Name. This is the name of the *AMI* or *image* from which your *instance* will be made. An *instance* is a running virtual machine; an *image* is stored files only.

Click on the blue button "Launch" to see:

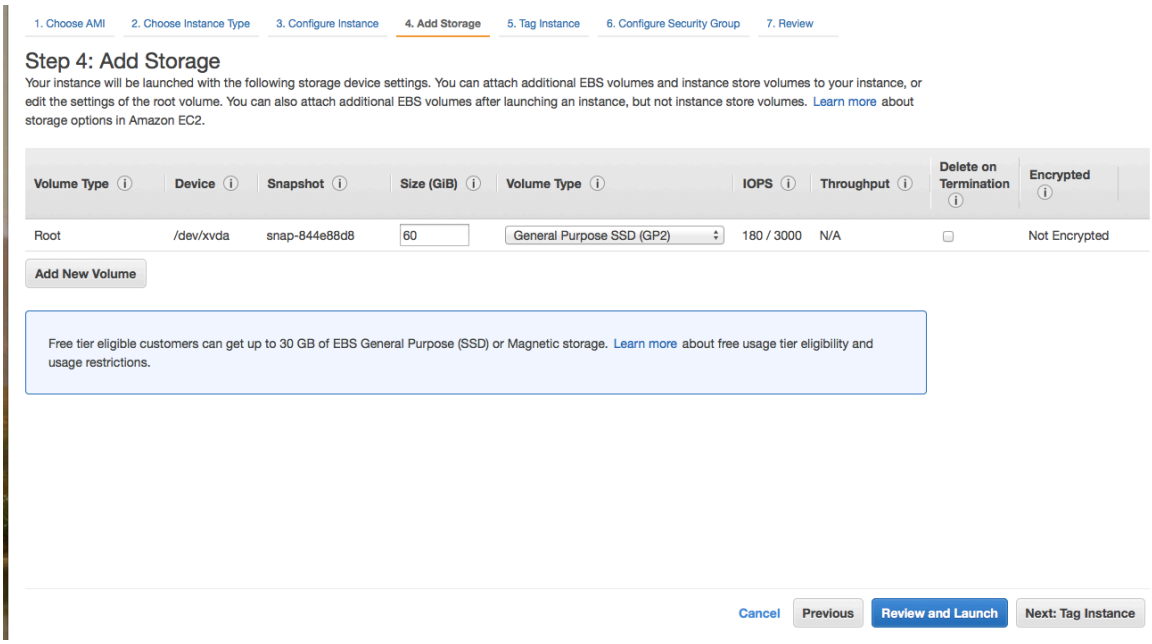


Check the left box in the row with Type of "t2.small" to choose a cpu with 2 GB RAM.

Click on Next: Configure Instance Details to a new page, where no changes are needed:



Then click on Next Add Storage to see:



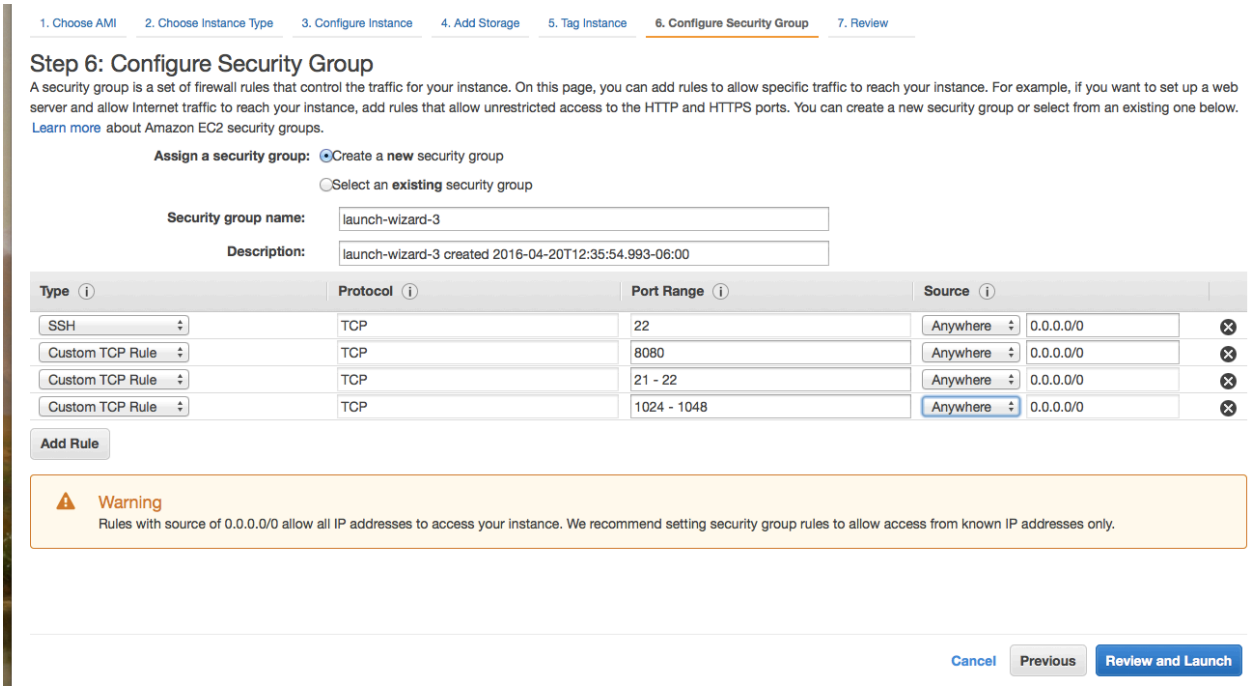
Enter 60 in the Size (GiB) box. This is the disk or "storage" size available size for GPS data files. 60 GiB will hold about 75,000 to 100,000 RINEX obs files and also their related nav and net files.

You may wish to use a larger disk size. Adding each additional 10 GiB above 60 increases your monthly cost about $1. Each 10 Gib storage will hold about 13,000 to 17,000 RINEX obs files (15 or 30 sec sample interval) and their related nav and net files. Three current

Dataworks for GNSS in COCONet or TLALOCNet use from 37 to 109 GiB for GPS data file storage. UNAVCO has not tested launching a Dataworks for GNSS image with more than 60 Size (GiB), and cannot be sure results. We do know using less than 60 probably will not be permitted by AWS since the UNAVCO AMU has 60 GiB. You should be able to increase the disk storage size for your AWS instance later, after you have it running.

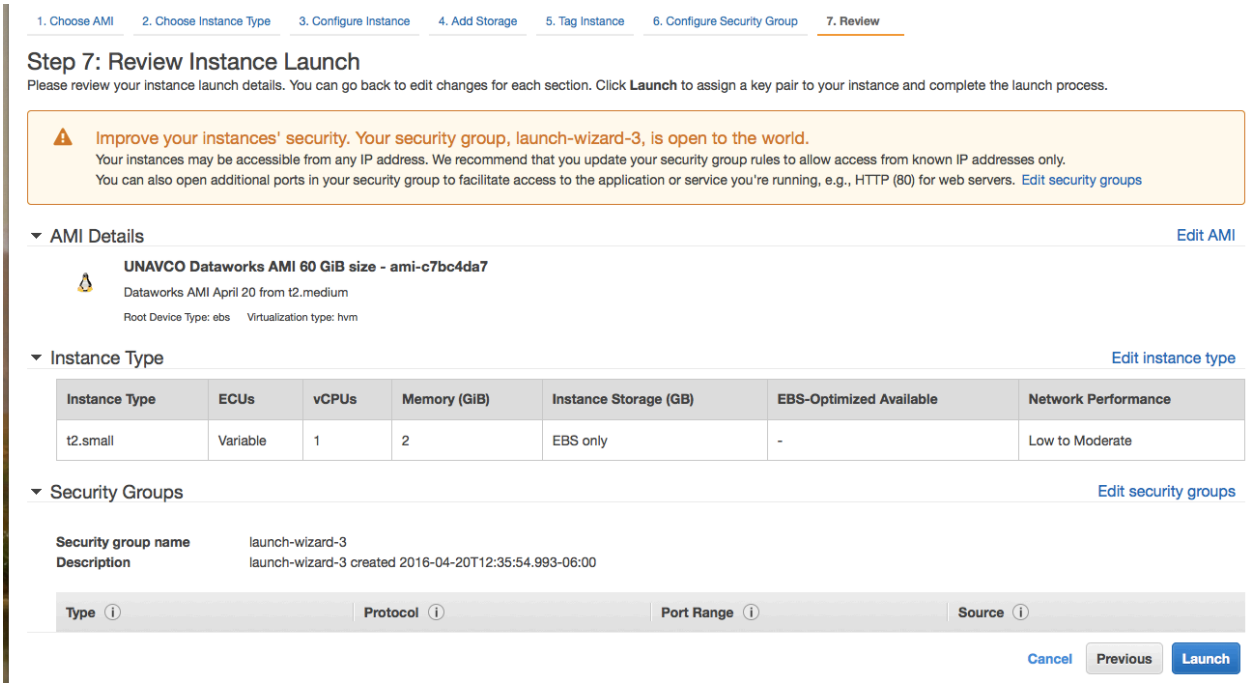
After you enter your value for Size (GiB), click on Next: Tag Instance, and in that next page click on Next: Configure Security Group. (A AWS "tag" is not a Gitlab "tag"!) In the Configure Security Group page you need to make some important additions, firewall rules.

Click on **Add Rule** three times, and each time enter values to make the three new rules (rows) shown in the next picture. For each line you need to set Port Range numbers, and set Source to Anywhere.



The port numbers shown will be enabled for outside users. Port 8080 is needed for the GSAC web site. Port numbers 21 and 1024 to 1048 are for the FTP server. By default port 22 is available as shown; it is used for logins with ssh.

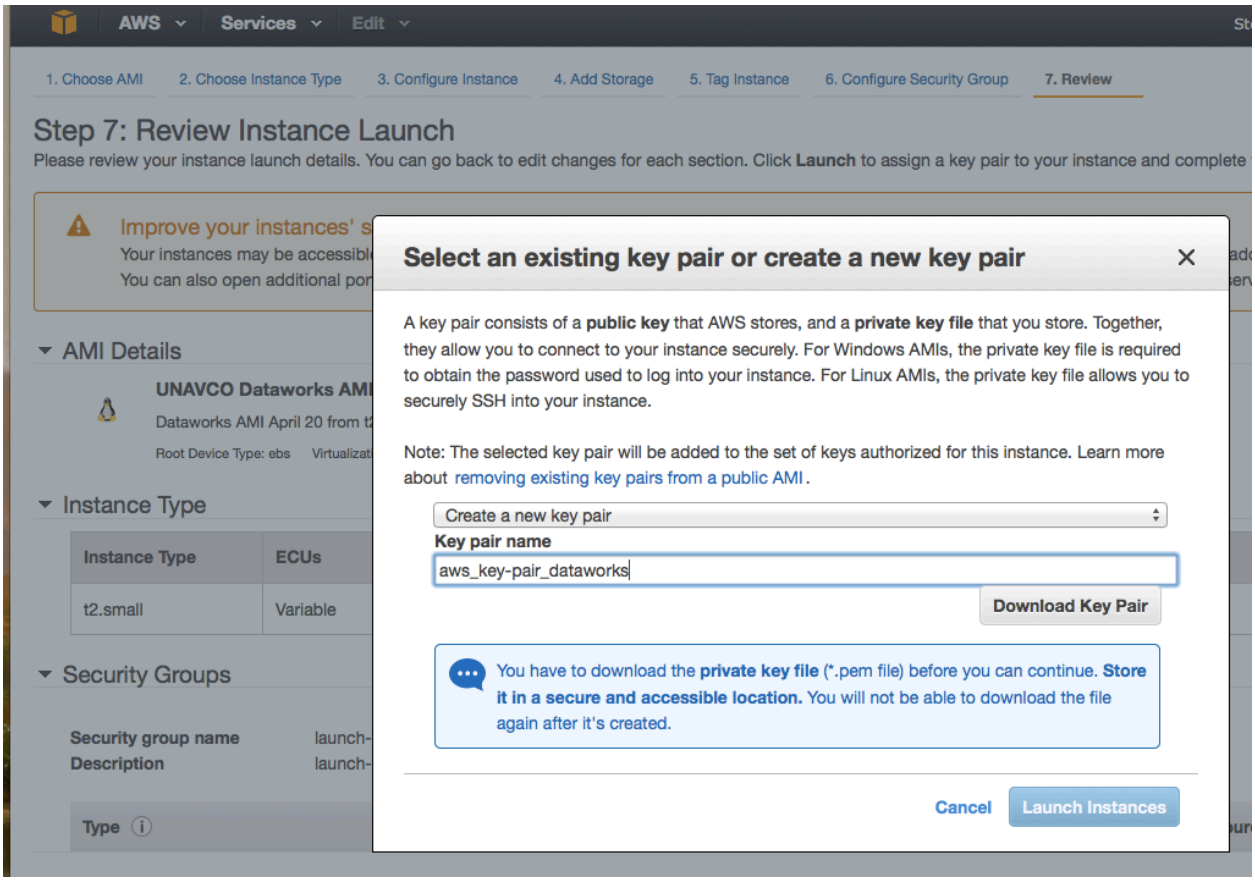
Next click on Review and Launch.



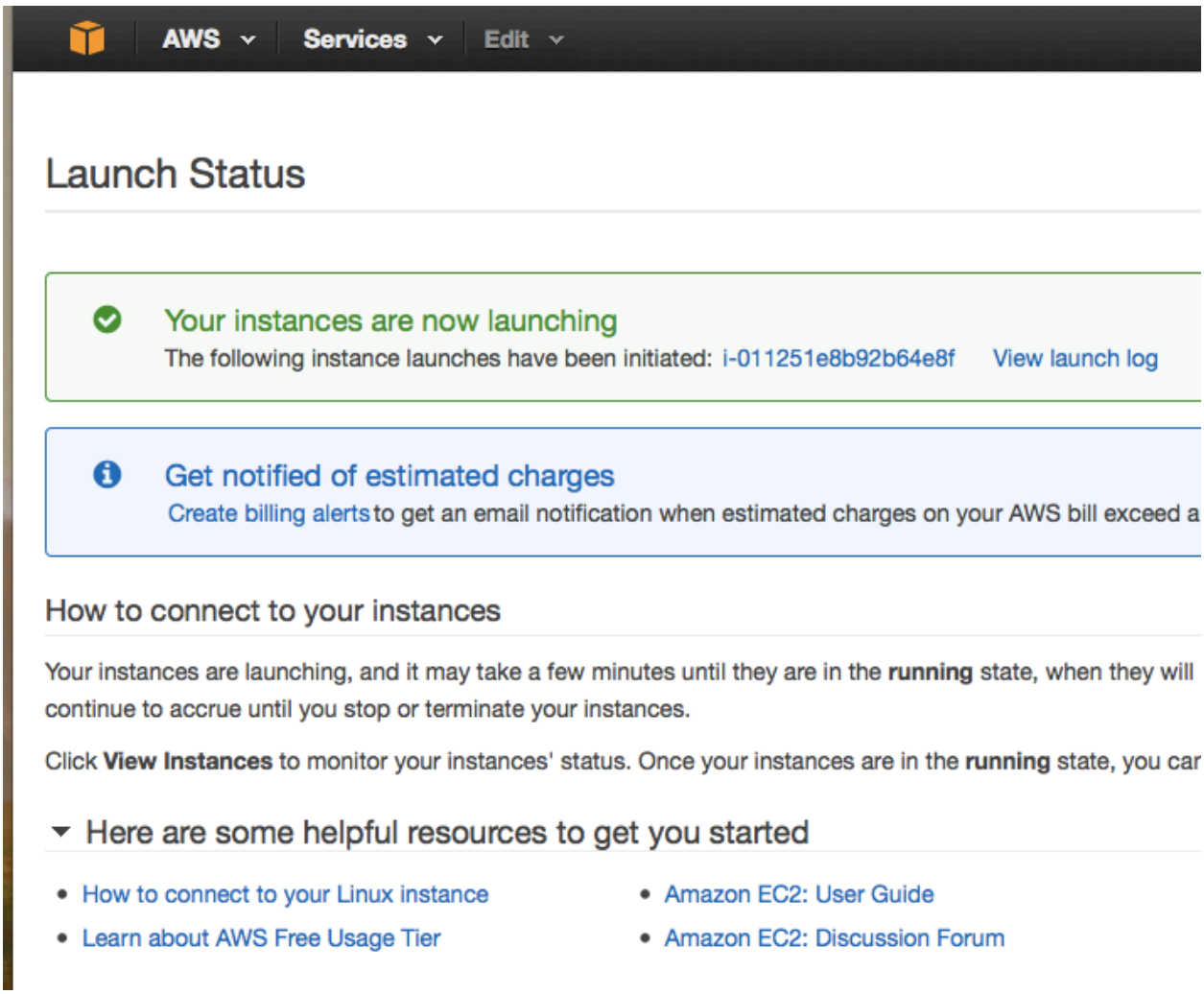
Everything should be OK.

"Memory (GiB) 2 " is correct – this is 2GB RAM for the cpu, not disk storage capacity.

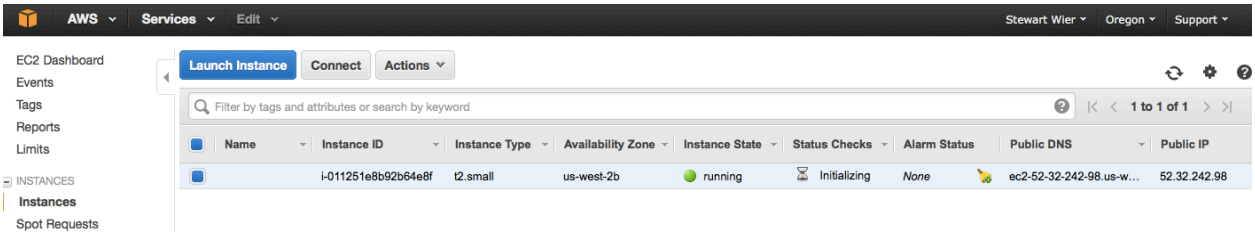
To allow you to log in to your new virtual machine, you need a "key pair," that is actually a file on your computer with the file name ending in ".pem". Create the 'key pair' with this box, choosing "Create a new key pair" and where you enter a suitable name for part of your new .pem file name:



Click Launch Instance to start your new Dataworks for GNSS "virtual machine" running. You will see this page:

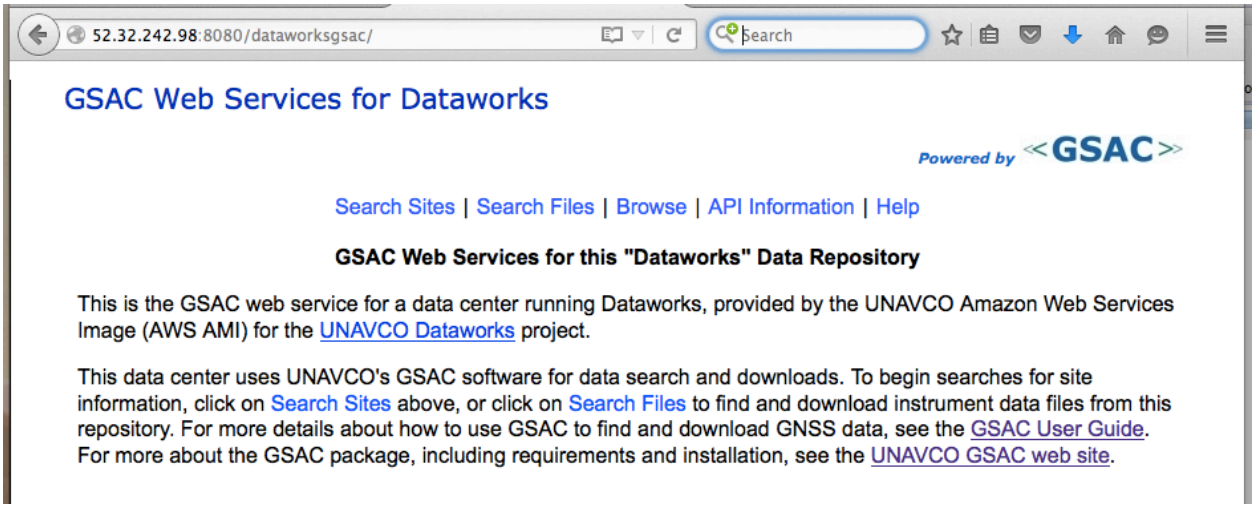


This starting process will take about 10 minutes. You can click anytime on the main left column menu to see "Instances:"



Your new AWS Dataworks for GNSS "instance" is installed and running and available online.

Note the "Public IP" number. This number is used as the domain in the URL to see your Dataworks for GNSS web site (GSAC) in a browser. In this case the IP is 52.32.242.98, and the complete URL is http://52.32.242.98:8080/dataworksgsac/.
Here is what you can see:



This GSAC is fully operational. A little demo data is in the Dataworks for GNSS database to allow you to test GSAC.

To log in to your AWS instance, use a Linux command like

```
ssh -v -i awsdevKeyPair.pem ec2-user@54.123.45.67
```

using your public IP number in place of 54.123.45.67 and with your ".pem" filename. This command is done in the directory with the .pem file.

Operating Dataworks for GNSS in the AWS instance

For anonymous ftp do the following:

Edit the ftp config file at /etc/vsftpd/vsftpd.conf to set the correct IP for your server. For exampla if your IP is 56.221.100.9

```
sudo vi vsftpd.conf


Insert (or change) the line:
pasv_address=56.221.100.9
```

Save and quit.

Then do linux commands

```
sudo usermod -d /data/pub ftp

sudo chmod a-w  /data/pub

sudo /sbin/service vsftpd restart
```

See instructions in

- the top level “README” file, `/home/ec2-user/README_Dataworks_Installation_Operation`
- the Dataworks for GNSS manual at http://www.unavco.org/software/data-management/dataworks/lib/docs/Dataworks_for_GNSS.pdf

There are several README files in the Dataworks for GNSS package, one in each of the above-listed directories. Read all of them.

To customize the appearance of the GSACweb site for your agency, log in to AWS. See the file /home/ec2-user/Dataworks-SW/README_Dataworks_AWS_Operation. You will do some Linux file operations to build a new GSAC, and deploy it to Tomcat.

To allow a URL like dw.myserver.org to work in a browser to show your GSAC from your AWS, it's necessary to create an "A" record on your DNS server, defining that number. For example, dw.myserver.org points to an Amazon instance: the “host” Linux command shows for example:
> host dw.myserver.org
> dw.myserver.org has address 54.187.10.298