SciServer How To's

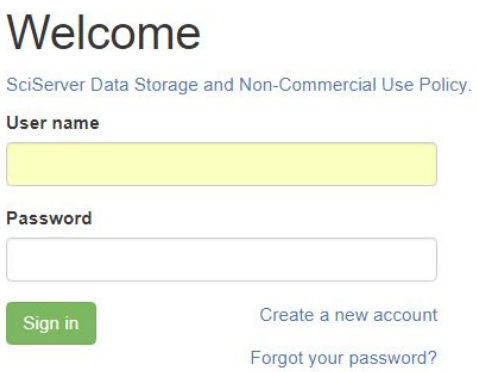
07/25/2018

Contents

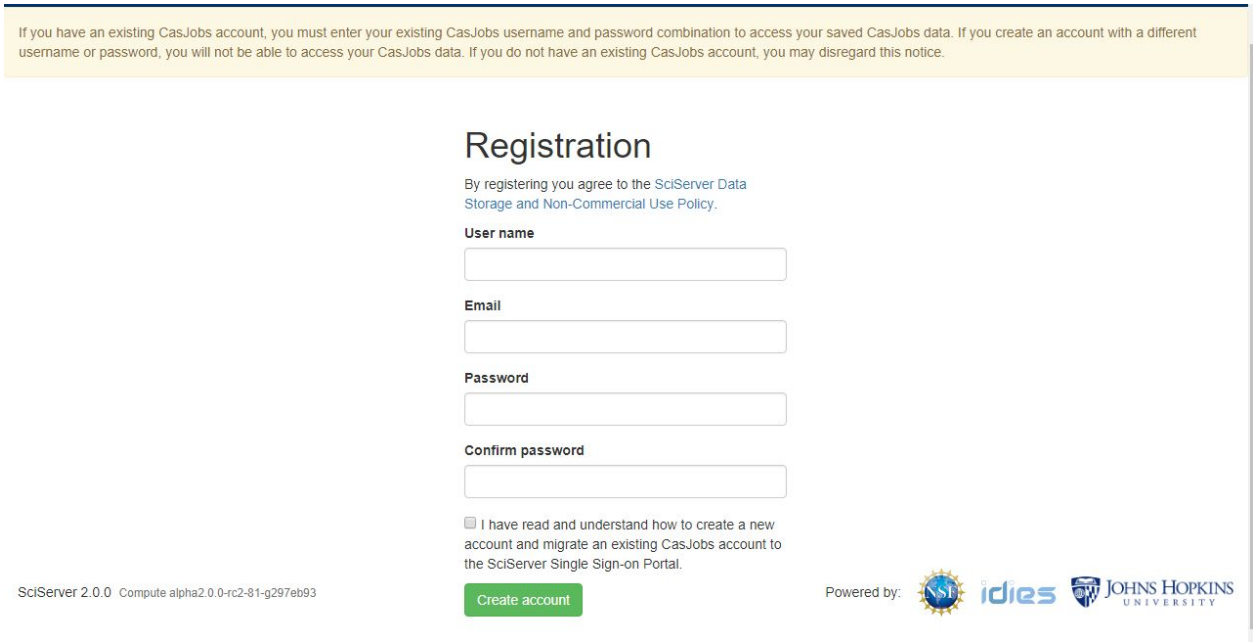
1. Account Management

1.1. “How do I register a new account?”

- Goto the following page https://apps.sciserver.org/login-portal

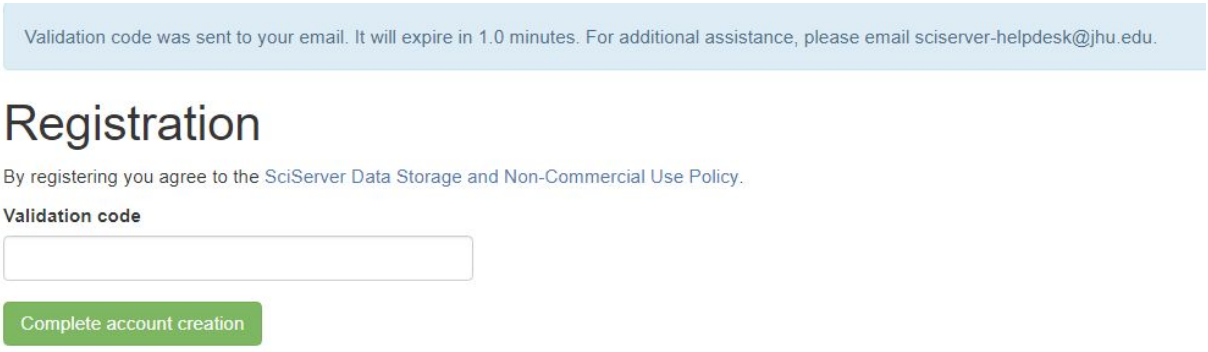


- Click the “Create a new account” link



- At the Registration page, enter the following:
  - User name
  - Email
  - Password (twice)
  - Check terms of service

- Then click “Create Account” which will send you an email and take you to a Validation page, where you enter your verification code:

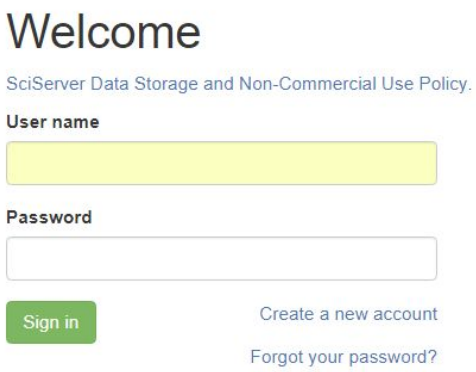


- Click “Complete account creation”
- Then you can go back to the Login Page to access your new account!

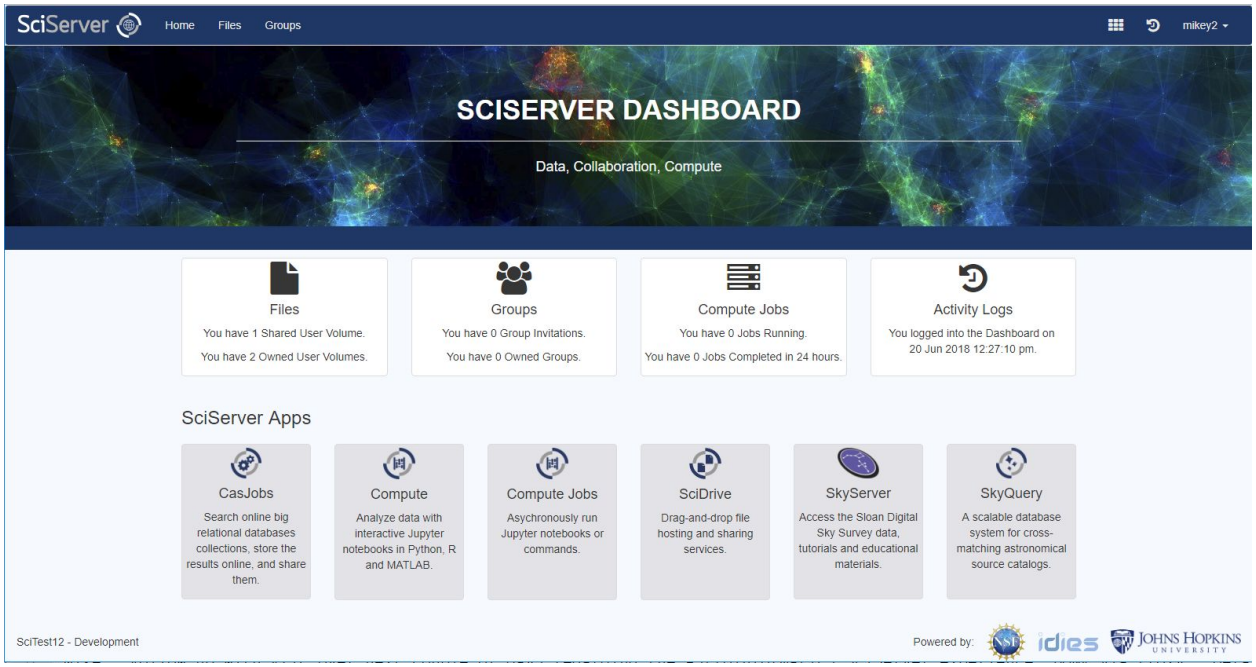
1.2. “How do I Login?”

- Go to the following page https://apps.sciserver.org/login-portal

- At the Login page, enter the following:
  - User name
  - Password

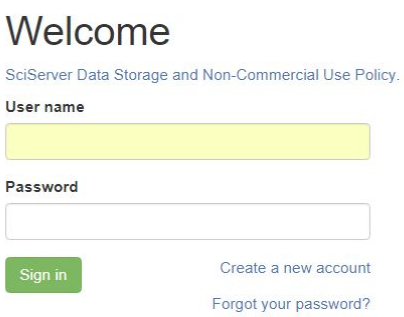


- Press the “sign in” button to Access your account!
- You will be taken to the SciServer Dashboard:

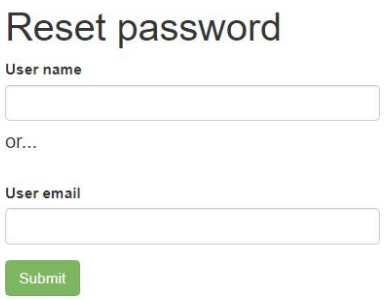


1.3. “I have forgotten my password, how do I reset it?

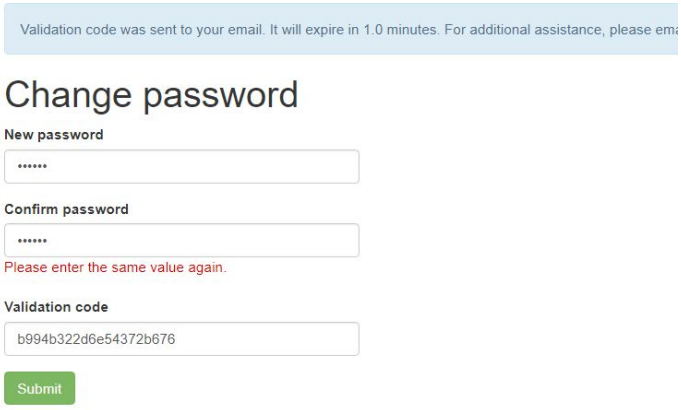
- Go to the Login Page https://apps.sciserver.org/login-portal



- Click the link “Forgot your password?”
- This will take you to the Reset Password page:



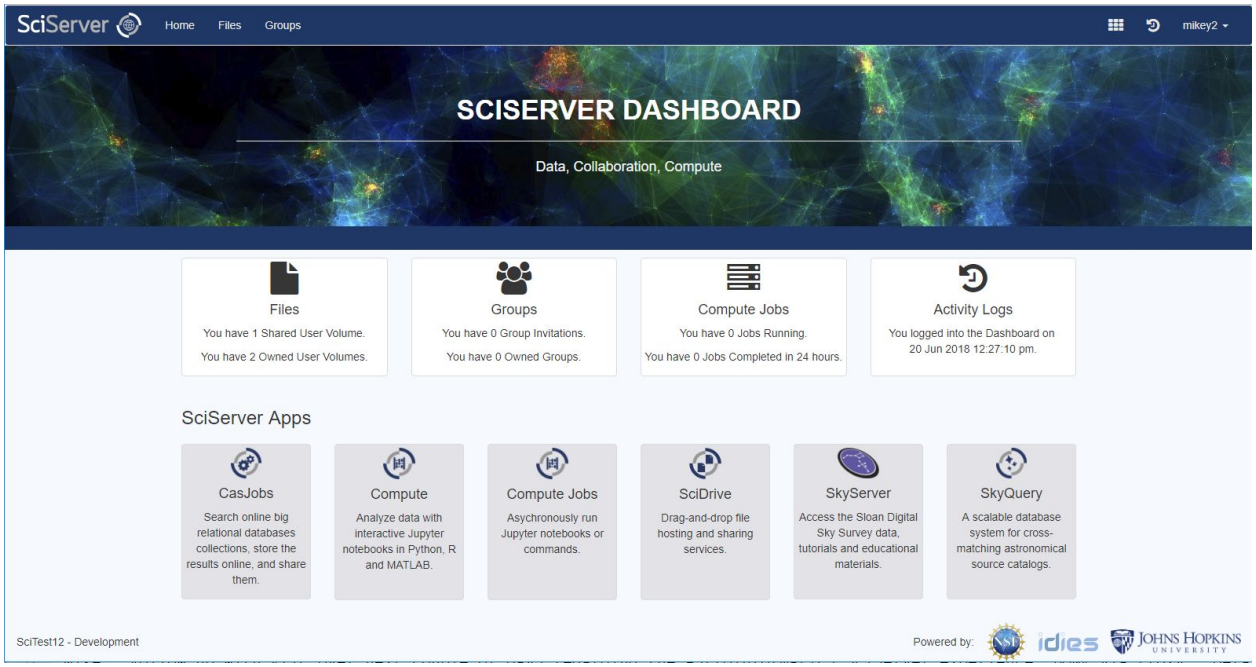
- Enter either your Username or your Email, and press the Submit button.
- This will take you to a page to reset your password, which you type in twice:



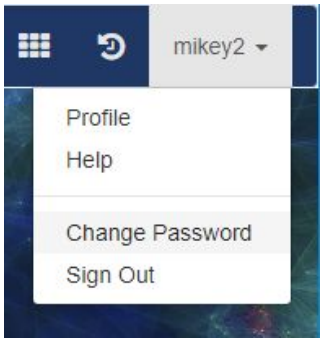
- You will also be sent a verification code to the registered email address that you also need to enter.
- Press the “Submit” button and your password will be reset!

1.4. “How do I Change my Password”

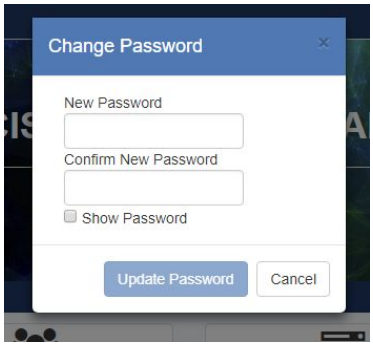
- Got to the SciServer Dashboard:



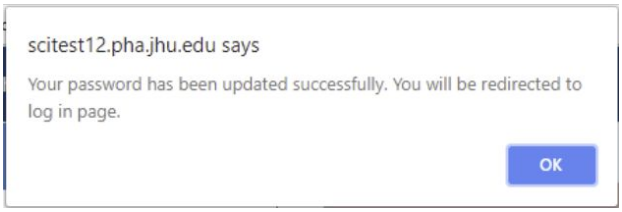
- Access the Menu in the top right hand side corner:



- Select “Change Password” which will display the “Change Password” dialog box:



- Enter your new password (twice) and press the “Update Password” button
- This will change your password, and then display a message that you will be redirected to the Login Page to re-login

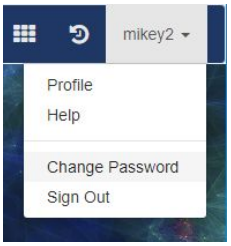


- You will then need to re-login using your new password to access your account again.

2. SciServer Help

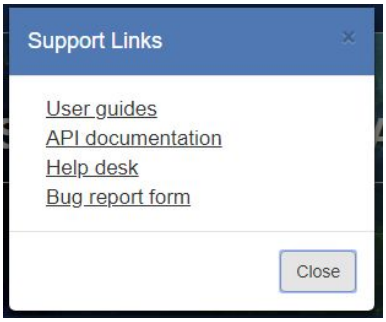
2.1. “Where can I find Help Documentation?”

- Access the Menu in the top RHS corner:



v

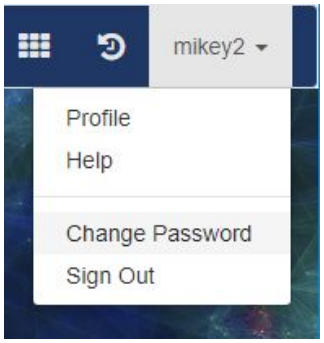
- Select “Help” to display a small dialog with a number of links:



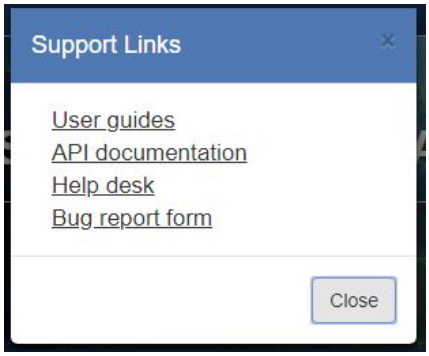
- Here you can find “User Guides”, “API Documentation” and “Help Desk”

2.2. “How do I report a Bug or Issue?”

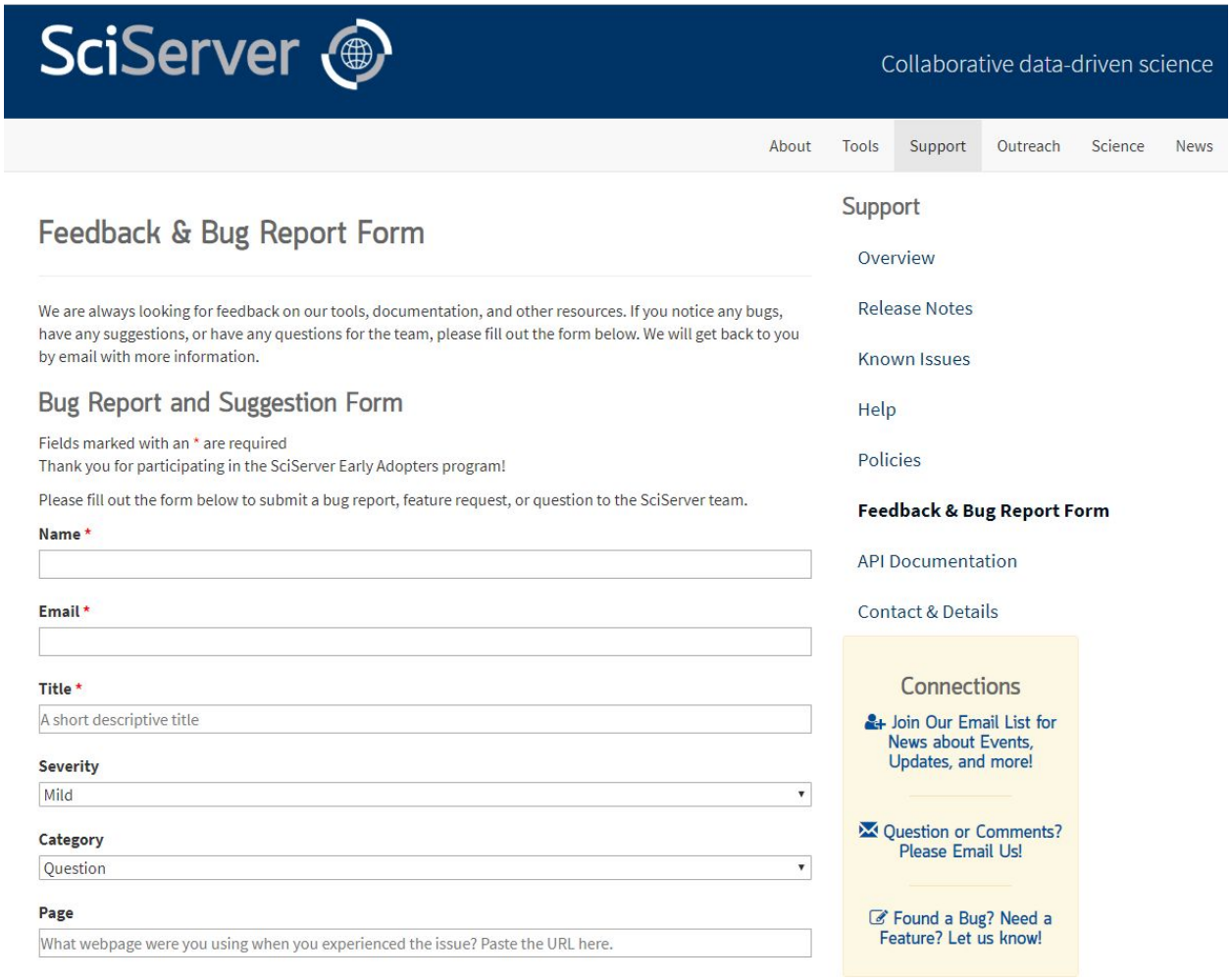
- Access the Menu in the top RHS corner:



- Select “Help” to display a small dialog with a number of links:



- Here you can find “Bug Report Form”
- Select this and it will open a new Web Page on the sciserver.org website:



2.3. "Contextual Clues"

3. Little "?" icons are on most of the UI app pages with links to specific section of the help pages "SciServer How-Tos".

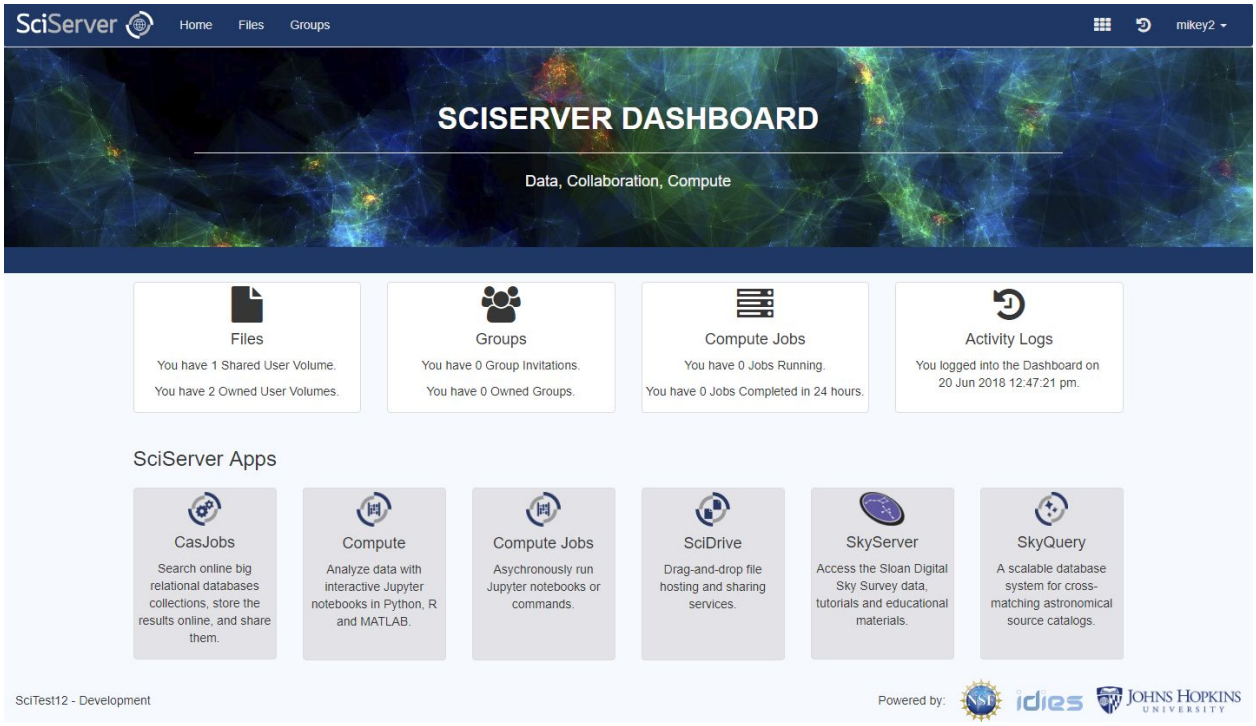
4. Dashboard

4.1. "What is the SciServer Dashboard?"

The SciServer Dashboard is the application that allows users to manage the work they are doing with SciServer and the other applications available. A user will be sent to the Dashboard Home Page after logging in from the Login Portal, and can navigate back to the Dashboard from most other SciServer Applications via shortcuts.

Home Page

The Homepage for the Dashboard looks like this:



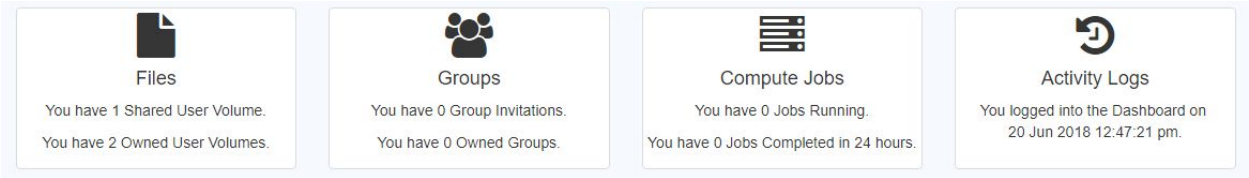
Menu and Functions

The Dashboard has a main “Menu Bar” across the top which is always visible in all views of the Dashboard, providing access to the following core features:

- Home Brings the user back to the Home Page
- Files Navigates the User to a Tab to manage Files and Folders
- Groups Navigates the User to a Tab to manage user defined Groups and sharing
- Displays a drop-down menu allowing the User to launch any of the supporting SciServer Applications (Home, Compute, Compute Jobs, CasJobs, SciDrive, SkyServer, SkyQuery)
- Navigates the User to a tab displaying a tabular summary of the history of their activity with SciServer functions and applications
- mikey2 A Menu dropdown that displays additional options for the user (access Profile, Access Help, Change Password, Sign Out)

Function Shortcuts

A set of shortcuts are displayed to access the many functions in the Dashboard application itself:



In particular the User can access:

- Files
- Groups
- Compute Jobs
- Activity Logs

Each of these shortcuts also displays information about recent activity on, or notifications about, each function.

Application Shortcuts

A second row of shortcuts allows the User to launch

Each of these shortcuts will launch a separate web app in a separate Browser tab.
Note that “Compute Jobs” is in both shortcut lists.

5. User Activities

5.1. “How do I View My Recent Activities?

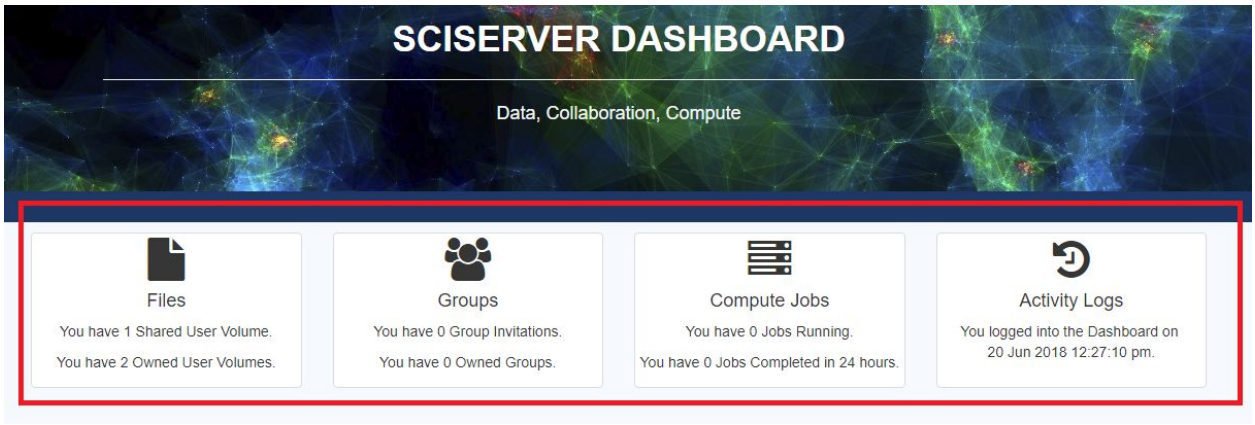
SciServer provides two ways to view your recent activities:

- The Dashboard displays the most recent activity related to File Management, Group Management, and Job Execution
- A separate “Activity View” accessible from the main Menu in the RHS corner provides a detailed table of all logged activities within SciServer, with filtering and sorting.

Dashboard Activity Summary

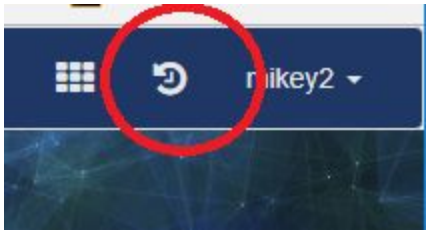
- On the main dashboard, a summary of Jobs, Groups and Files is provided as shown below

- In particular, Invites to New Group are shown, that the User can select to take them to the Groups tab in SciServer

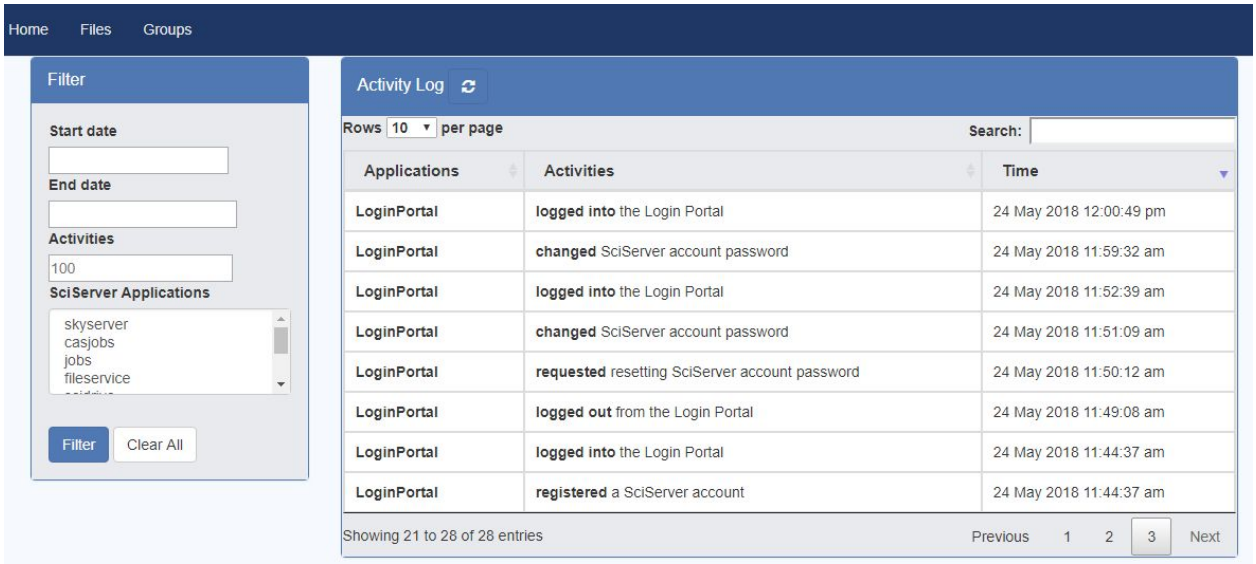


Activities Logs

- Access the top RHS menu bar and click the  icon:



- This will display the Activities table:



6. Group Management

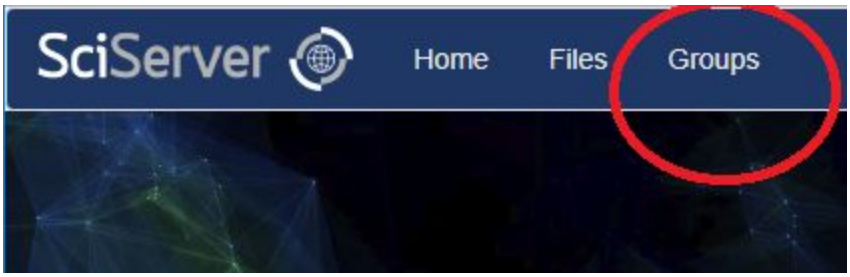
6.1. "What can I do with the 'Groups' feature in SciServer?"

SciServer provides a feature called 'Groups' for users to share their resources with their collaborators privately. The resources that users can share are file folders, databases, volume containers, and Docker images. Also SciServer provides a Group view which lists all the shared resources among group members.

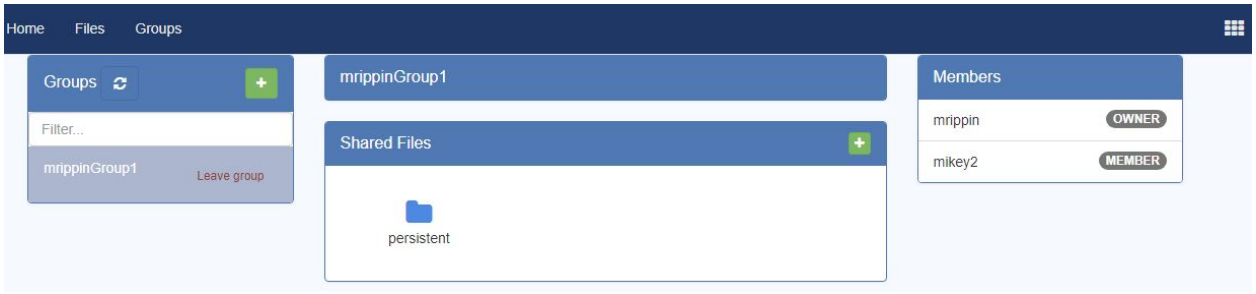
SciServer Groups allow you to create lists of Users and to share resources, such as file Folders, to all of them at once. You can manage a team of people in a project by creating a group with the relevant users in, and allowing everyone to work from the same shared folder. Importantly, you can also make sure that no one who is not in the group can access the shared folder, so you can keep things private.

6.2. "How do I create a new Group?"

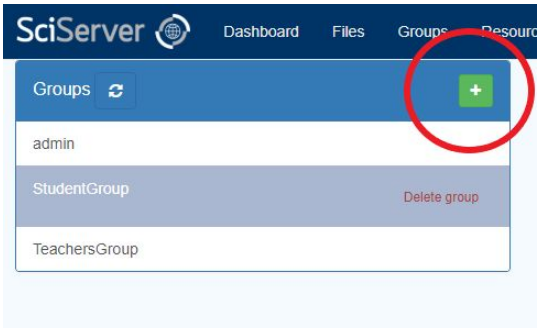
- Select the Groups tab in SciServer:



- This will show the Groups View:

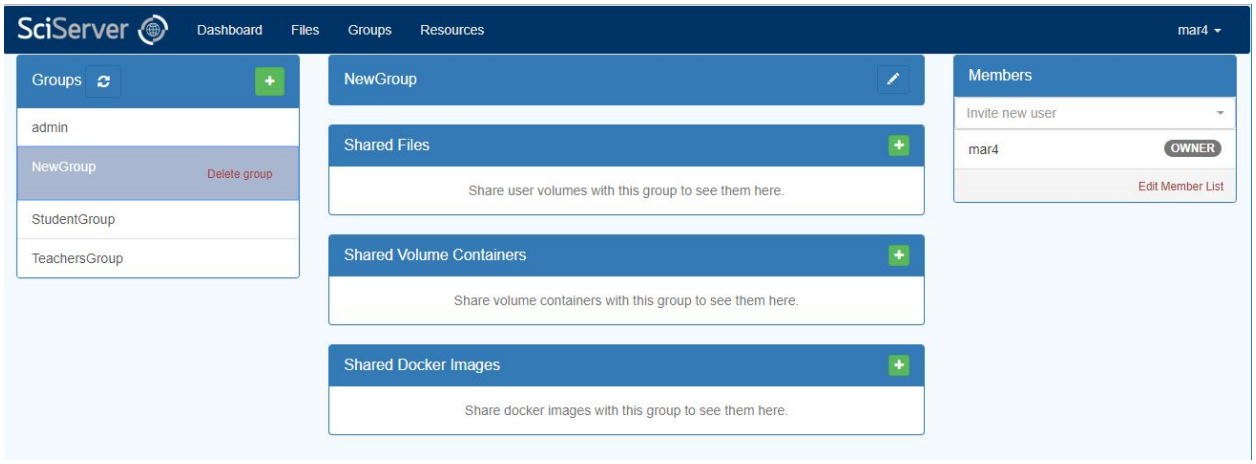


- The left hand box will show you all the groups that you are a member of. As you select each one, the middle set of boxes will show you which resources have been shared with that group, and the right hand box will show you all the members in the group.
- To Create a New Group, click the "+" button on the LHS "Groups" box:



- This will open the “New Group” dialog box:

- Enter a name for the Group and an optional Description, then press the “Create” button. This will show your new Group in the “Groups” box, and will show you as the only member in the RHS box:

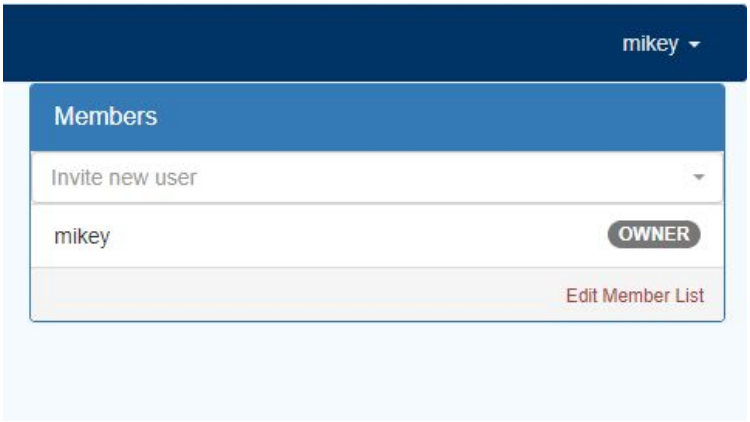


- This is your new group!

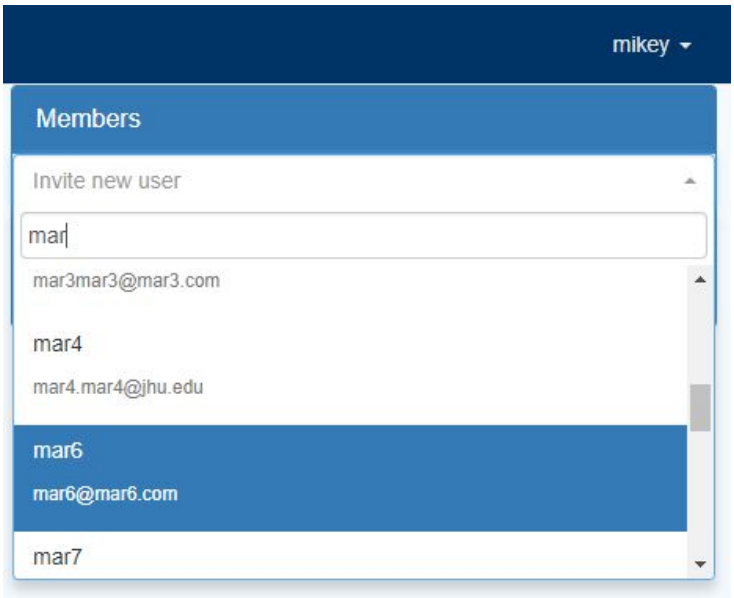
6.3. “How Do I Invite Users to a Group?”

- You can only **invite** users to a group that you own, or that someone else has given you the “grant” privilege on.

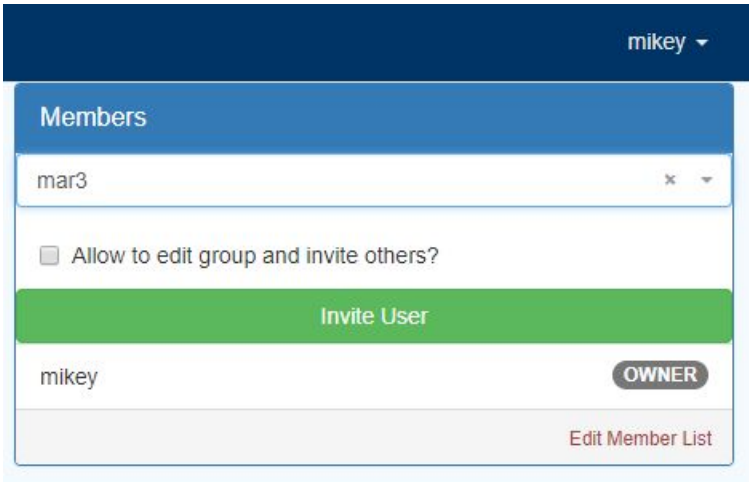
- In the groups view, go to the “members” box on the right hand side:



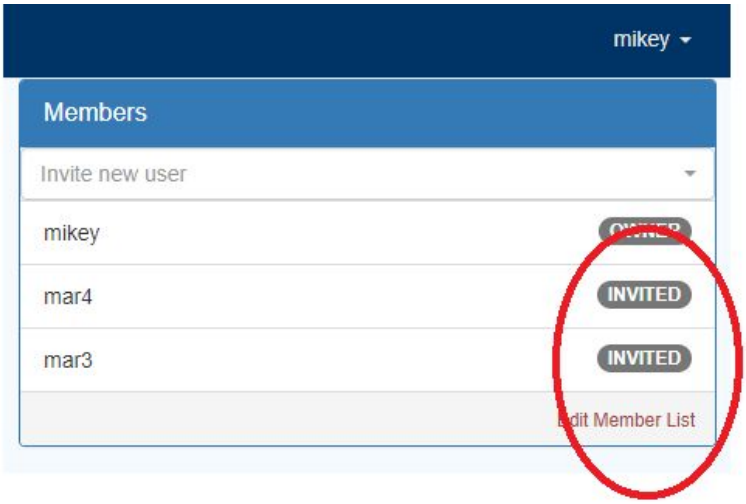
- In the box “Invite New User”, start to type in the names of users to invite. This will create a drop down filtered by what you enter, so pick the name that is correct.



- You will be given the option to Grant this user the ability to edit the group and invite other people:

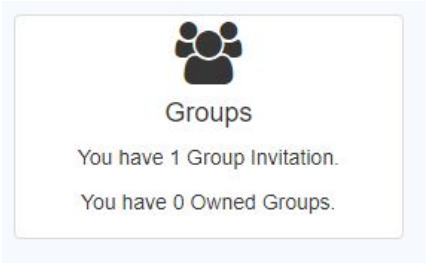


- Do this for all users you want to add.
- NOTE: This will **invite** the users. They will need to accept the invitation before being a member of the group:

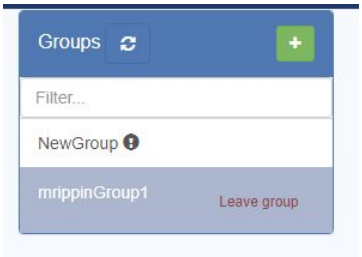


6.4. "How do I Accept a Group Invite?"

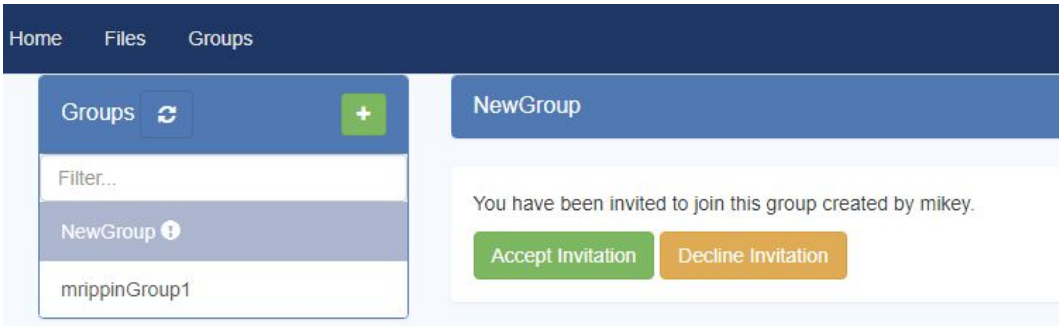
- On the Dashboard tab, you will see an Invite:



- Select the Groups tab, and you will see the Group listed with an icon next to it:



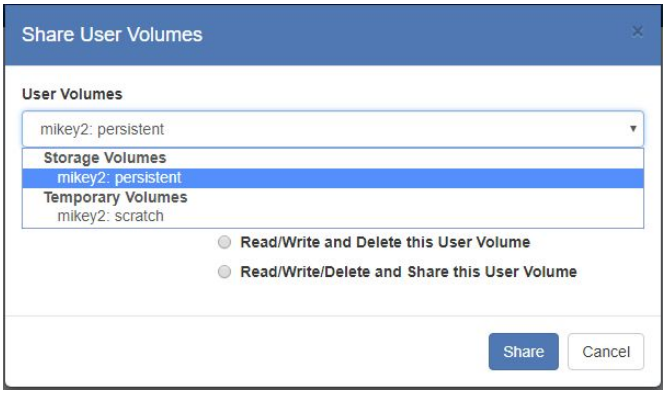
- Select the group and you will see an option to accept the invitation:



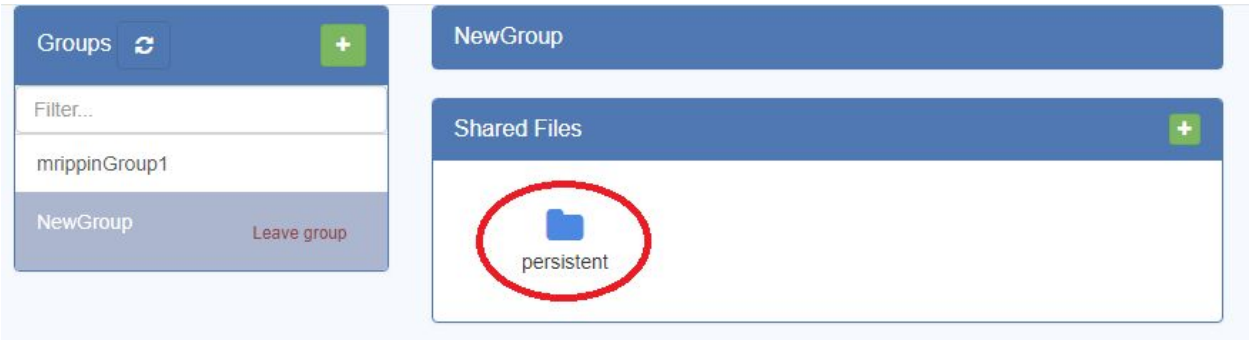
- Select “Accept Invitation”, and you will join the group and be able to access it.

6.5. “How do I Share Resources with a Group?”

- In the Groups Tab, selecting any of the groups on the left hand panel will show the resources already shared with that group. Resources that can be shared are: (1) File Folders (2) Compute Docker Images (3) Compute Volume Containers (4) Databases 9in some circumstances). It will only show each box if there are any resources that you are able to share.
- Each of these resources is shared in a separate box, and each has a  button.
- Pressing this will display a dialog box where you can select resources that you have access to:

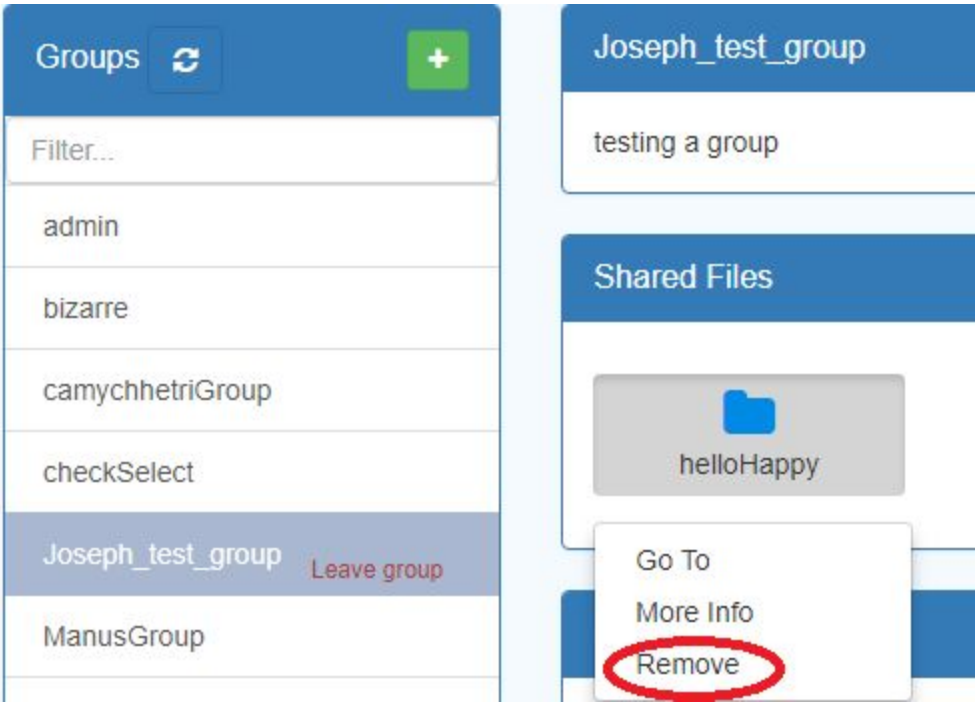


- Select the resource and permission level, then press “Share” and the resource will appear in the group:



6.6. "How do I Remove Resources from a Group?"

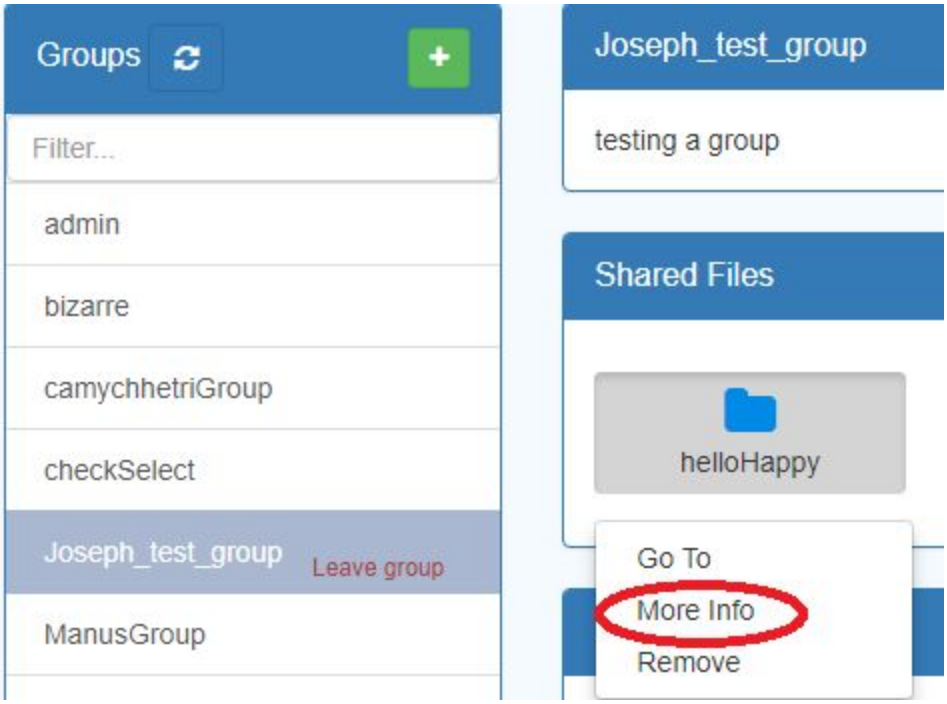
- In the Groups Tab, selecting any of the groups on the left hand panel will show the resources already shared with that group.
- Click on the shared resource
- Click "Remove"



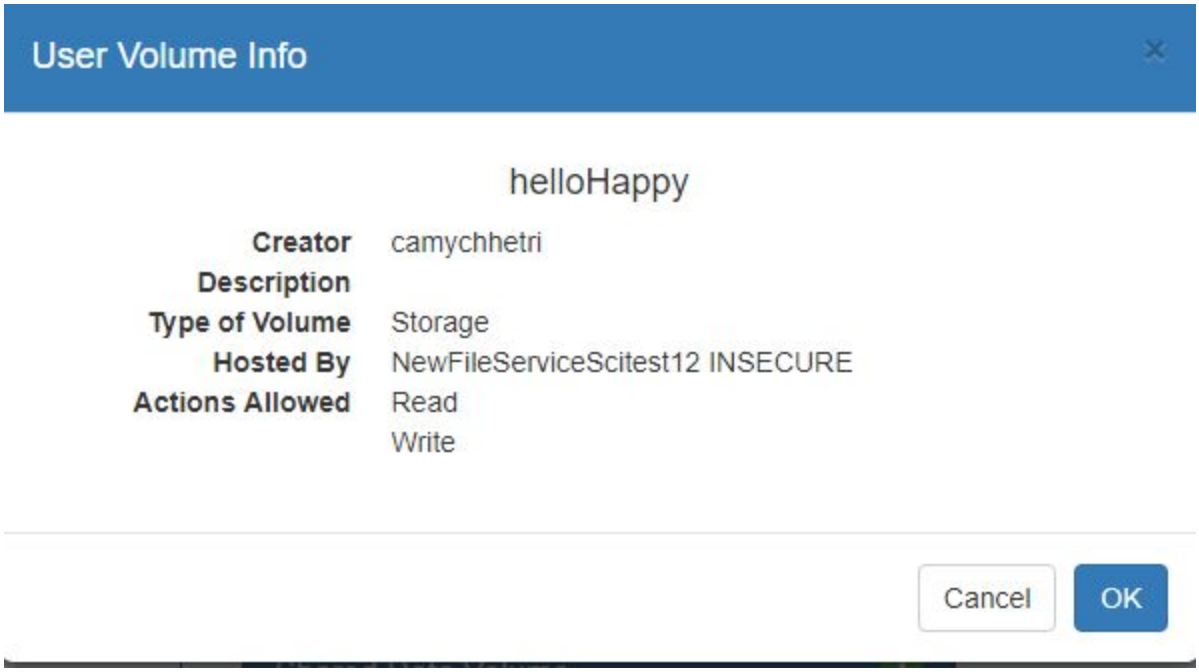
- Delete confirmation alert will appear
- Click ok to proceed

6.7. "How do I view Shared Resources Information?"

- In the Groups Tab, selecting any of the groups on the left hand panel will show the resources already shared with that group.
- Click on the shared resource
- Click "More Info"



- Pressing this will display a dialog box where you will find more information:



7. File Management

7.1. "What File Management Capabilities does SciServer provide?"

SciServer provides a number of features for managing files and compute notebooks, all of which will be made available in both the Sciserver dashboard application and in SciServer Compute containers from within Jupyter Notebooks. All the common capabilities of File management are

provided, such as renaming, moving, copying etc, but also sharing of “User Volumes” for collaborative work. NOTE: the File Management features provided in the SciServer UI are different to the services provided by SciDrive, which is an older storage system that is still supported for legacy applications.

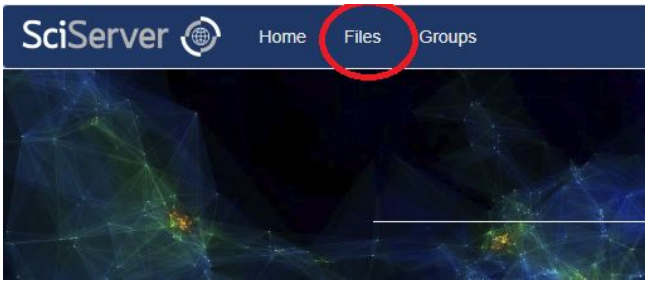
Permanent or Temporary Storage?

Users can create a number of top level “User Volumes” under which new folders and files may be added. User Volumes can be created in one of two different storage pools: a permanent pool called “Storage” and a short-lived pool called “Temporary”. Folders and files in User Volumes under “Storage” will be backed up and permanent, but there is a quota limit of 10GB. Folders and files in User Volumes unde “Temporary” are not backed up, and will be deleted after a particular time period, but there is no imposed limit or quota on how much data can be stored (because it will be deleted).

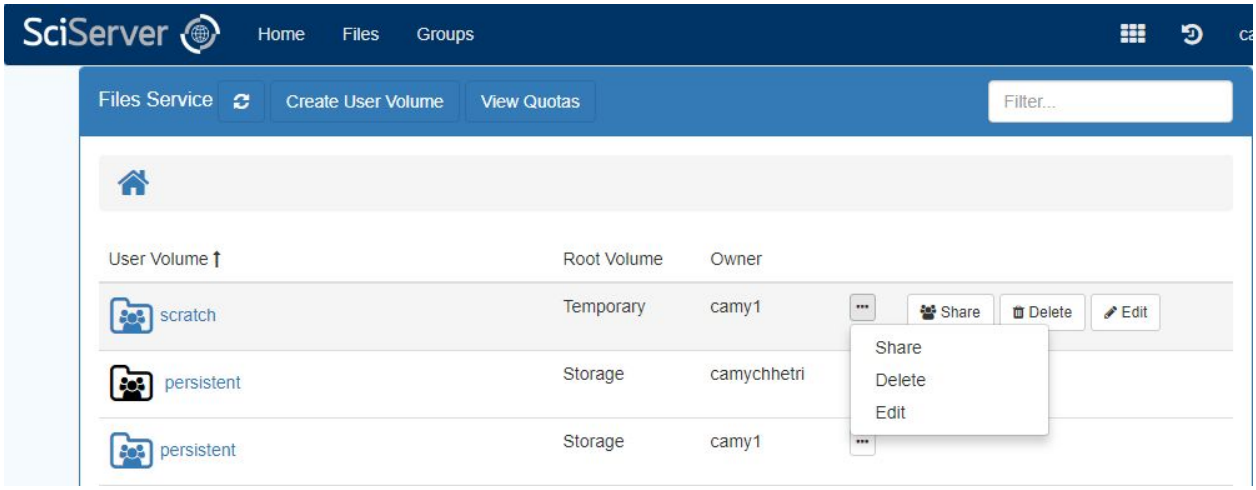
Most users should store their data and files in a “Storage” User Volume. The “Temporary” User Volumes are meant to be used as intermediate storage for SciServer Compute calculations.

SciServer Dashboard View

- Login to the SciServer dashboard and select the “Files” tab:



- This will show the Files View at the top level of “User Volumes”:

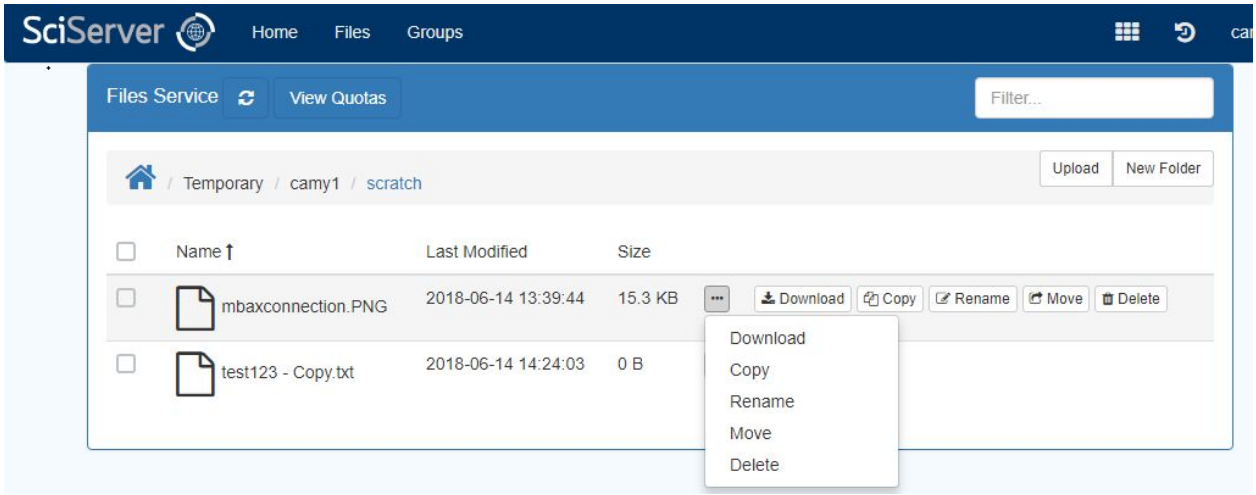


- There are a number of User Volume operations available in this view

- Hover over each row to view available operations
- Click on the ellipsis button to view available operations
    - *Share*: If the user has the appropriate permissions, they can share a User Volume with other users, or groups of users.A User Volume created by the user is always shareable by them,
    - *Delete*: If the user has the appropriate permissions, they can delete a User Volume. A user can always delete a User Volume they create.
    - *Edit*: The User Volume name can be changed, and a description provided for it.

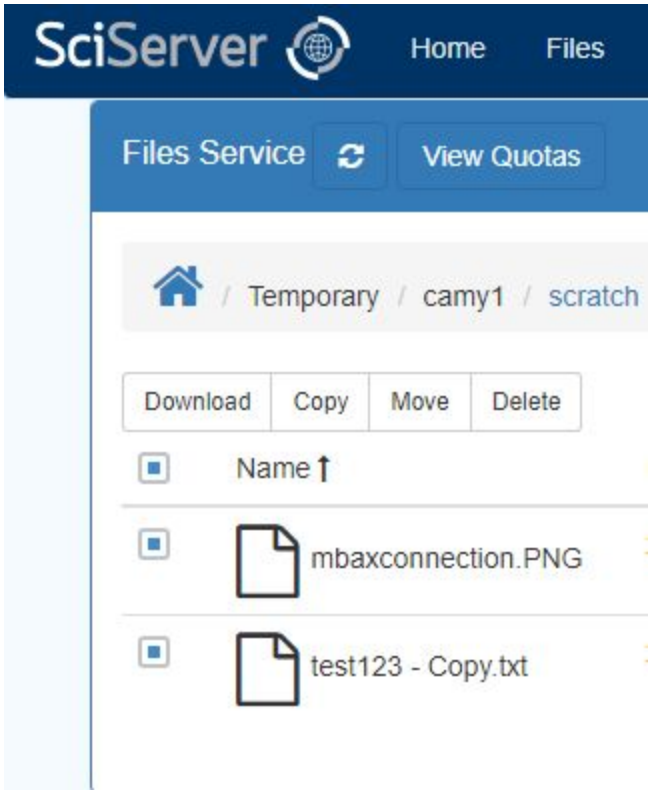
- Different icons refer to different levels of sharing:
    -  This User Volume is owned by the user, and has not been shared
    -  This User Volume is owned by the user, but **has** been shared with another user or group.
    -  This User Volume is not owned by the user, but has been shared with the user by another user or group.

- Selecting a Folder in the LHS column will open that folder up, one level at a time. *It does not provide a tree view.*



- There are a number of file and folder operations available in this view
- Hover on each to view available operations
- Click on the ellipsis button to view available operations
    - Copy
    - Rename
    - Move

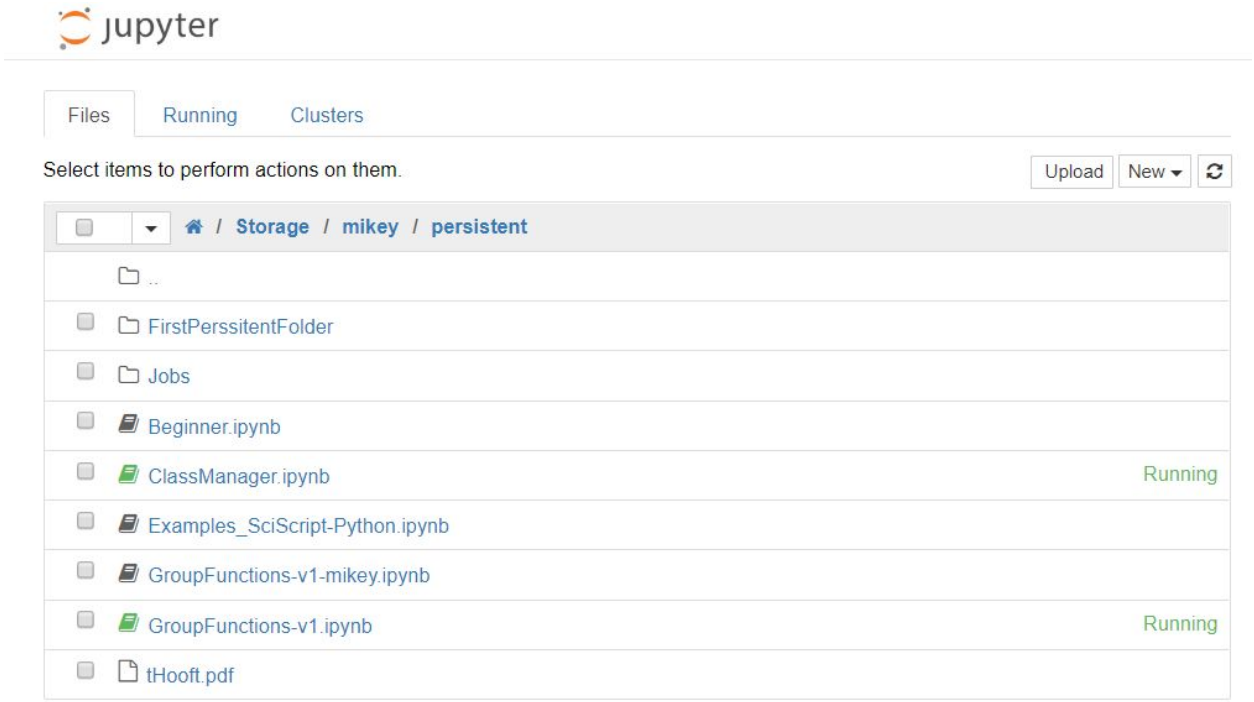
- Delete
- Download: A file can be downloaded (but a Folder cannot)



- Perform multiple operations on files or folders
    - Check one or multiple checkbox
    - Menu gets displayed with available operations on top
    - Click on any menu item to perform an operation

Jupyter Notebook View

- When running SciServer Compute, the same filesystem is presented by the Jupyter application, but is presented in a more traditional hierarchical manner with a full path access that supports working with files in a Linux Console:



- The operations on files and folders available in this view are provided by Jupyter.

7.2. "What is a User Volume? How is that different to a Folder?"

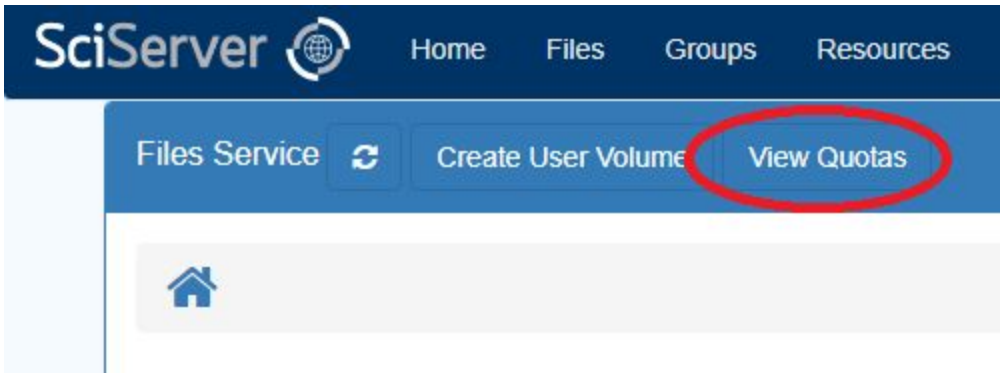
A User Volume is a top level folder that is created and owned by a SciServer User. It is different to a 'normal' Folder in a few ways:

- There is only the one top level of User Volumes, all subsequent lower levels are normal Folders
- A User Volume can be shared with others users and groups, but a normal folder cannot. This is very important.

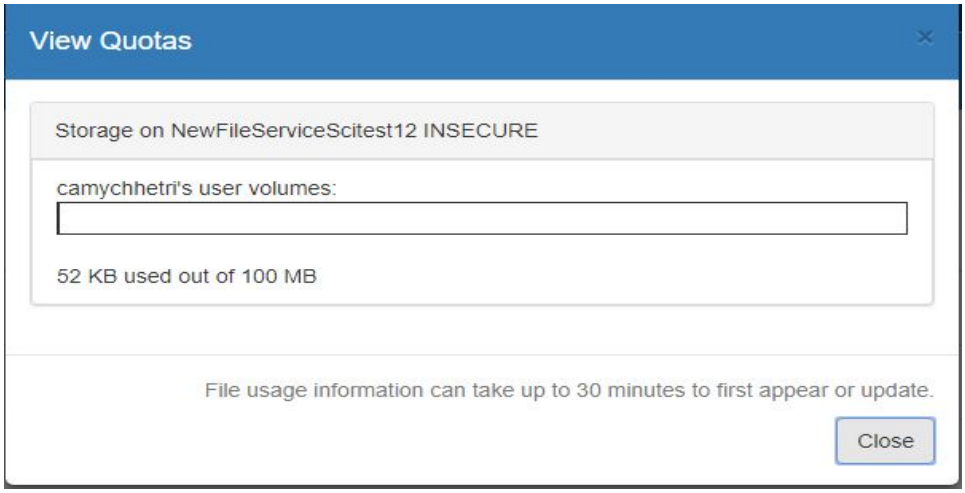
- A User Volume can be selectively “mounted” in a Compute Container, and made accessible to a Jupyter Notebook. Folders at lower levels under a Volume Container cannot be.
- A User Volume cannot be moved or copied.

7.3. “How to check Storage usage information?”

- On the SciServer “Files” tab, click the “View Quotas” button:

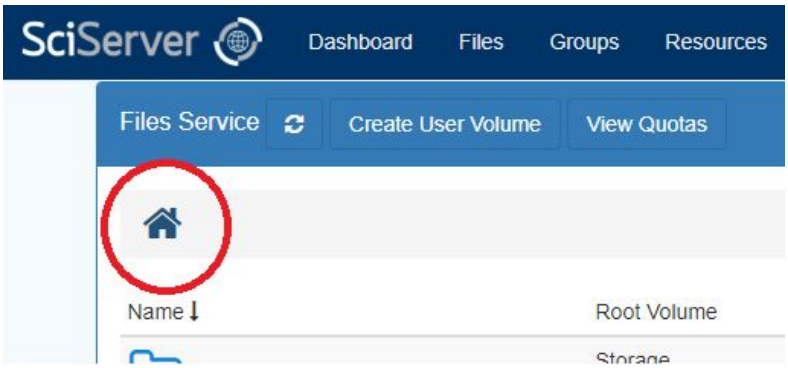


- “View Quotas” dialog pops up.

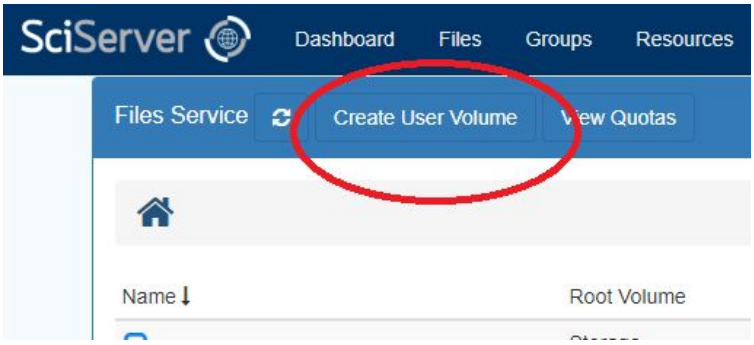


7.4. “How do I create a new User Volume?”

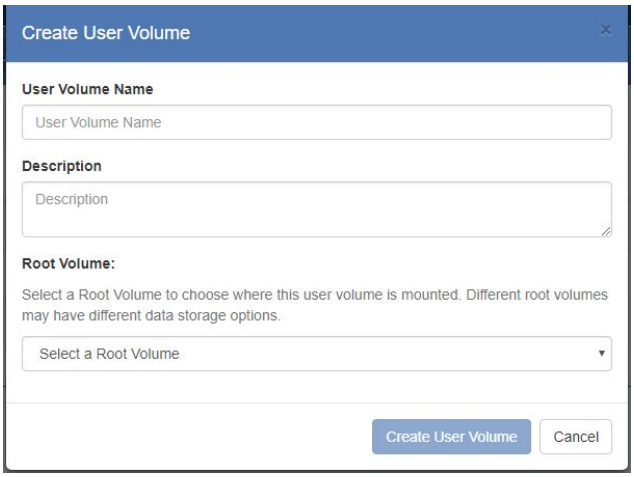
- On the SciServer “Files” tab, click the “Home” button to get back to the User Volume view:



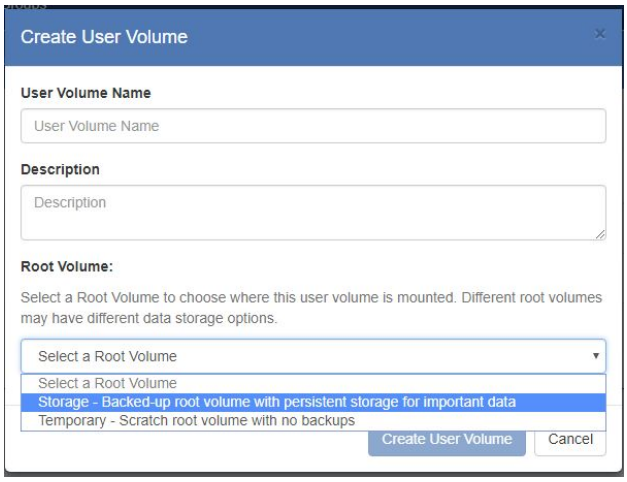
- Press the “Create User Volume” button:



- In the dialog that pops up, enter a name and optional description:



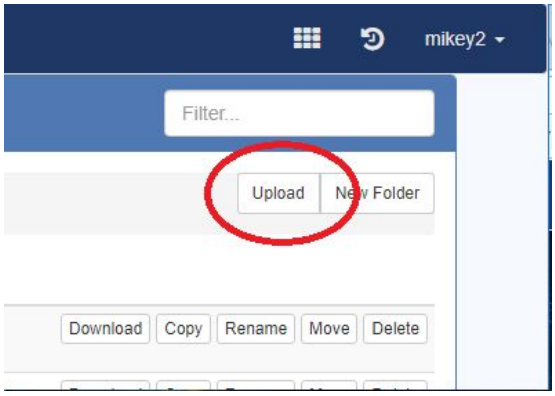
- Select a “Root Volume”, which means decide whether to create the new User Volume in a permanent and backed-up storage pool, or to create it in a temporary storage pool, knowing that it will be deleted after a certain period of time (defined by the SciServer Data Storage Policy Rules).



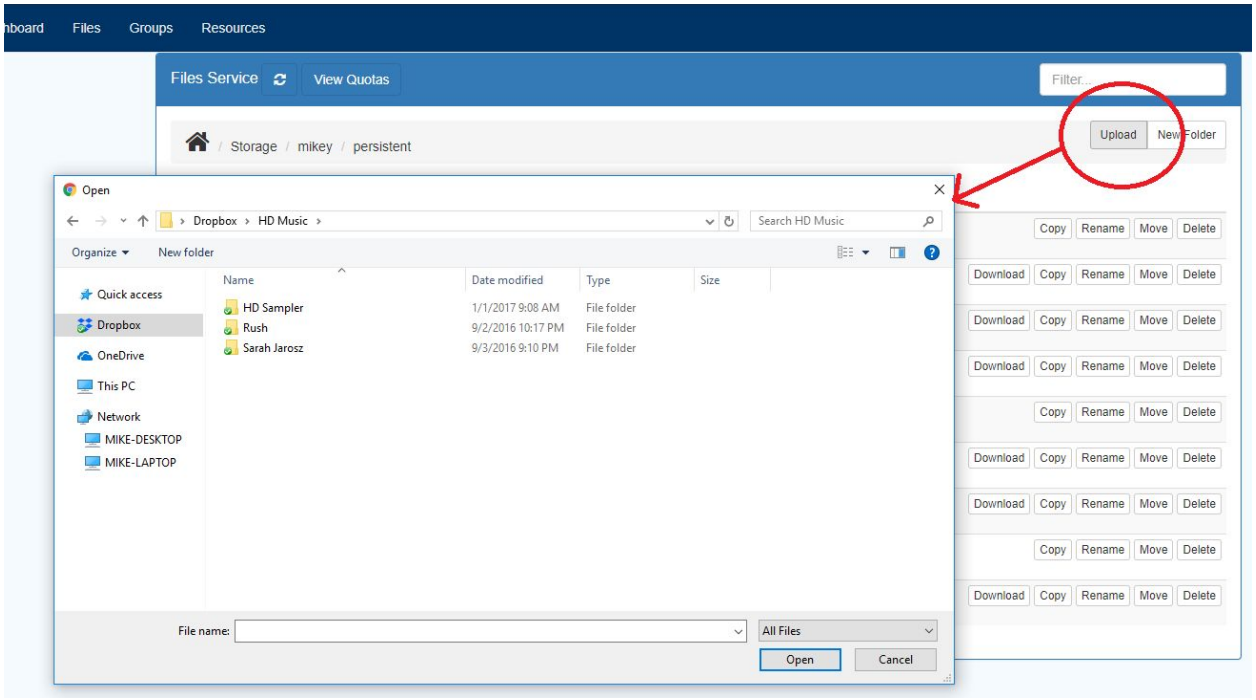
- Press the “Create User Volume” button and the new User Volume will be created.

7.5. “How do I Upload Files?”

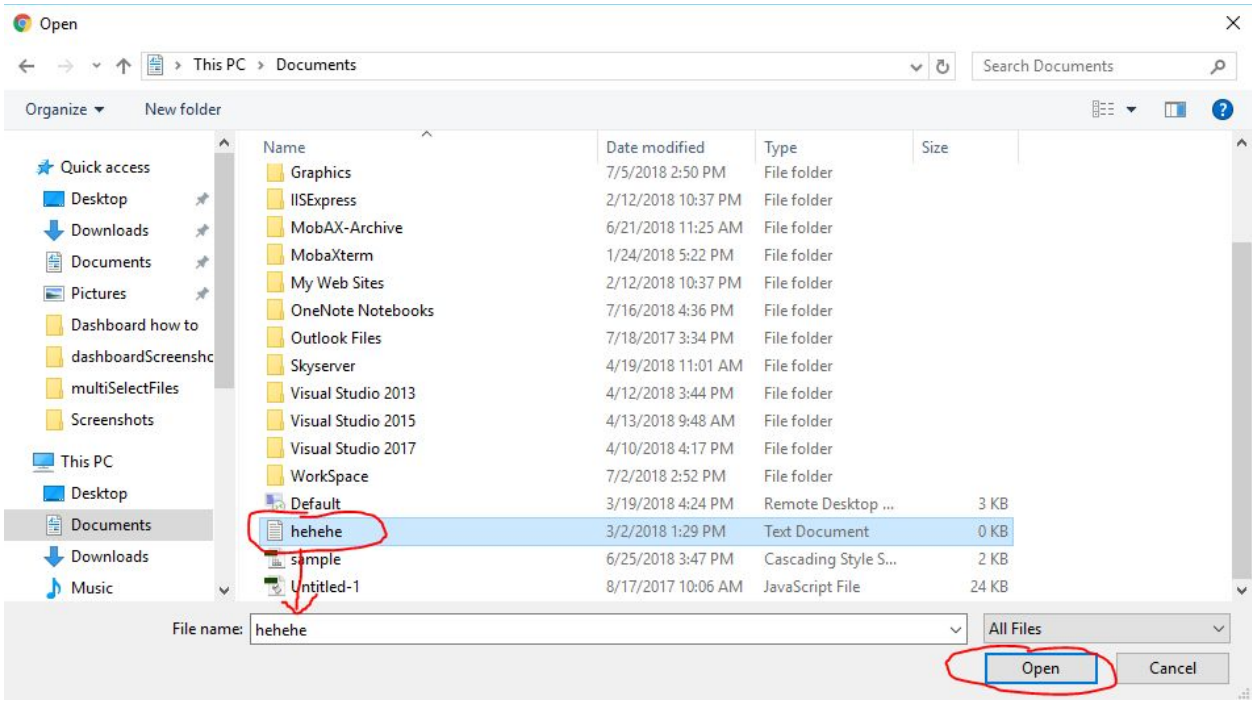
- Files cannot be uploaded directly to the top level User Volume level.
- Files can be uploaded to any Folder under a User Volume, including directly under the User Volume
- Navigate into a User Volume and you will see an “Upload” button:



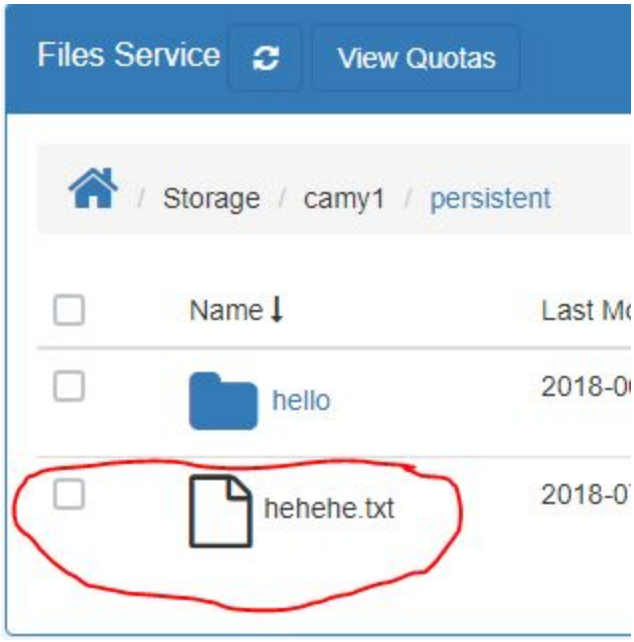
- Press this to display a dialog box:



- You must navigate to a *File*, select it then press the “Open” button”



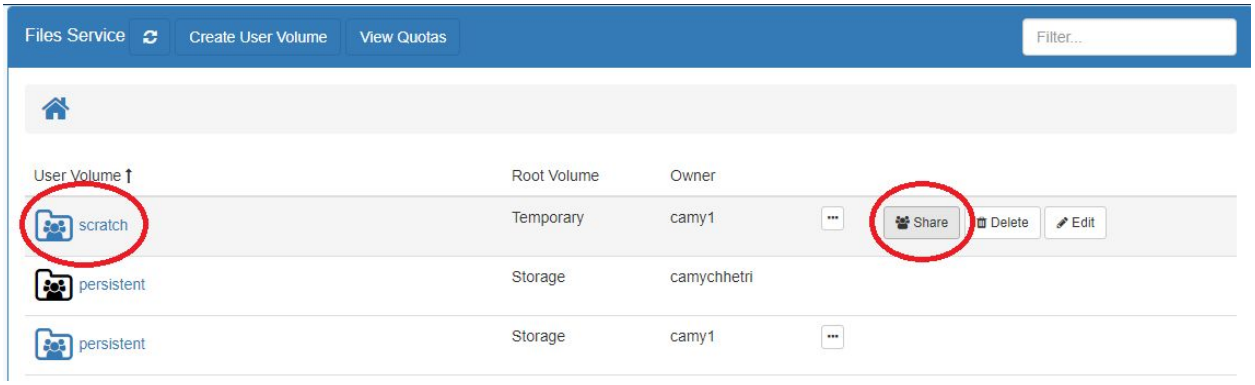
- This will then show the file you just uploaded:



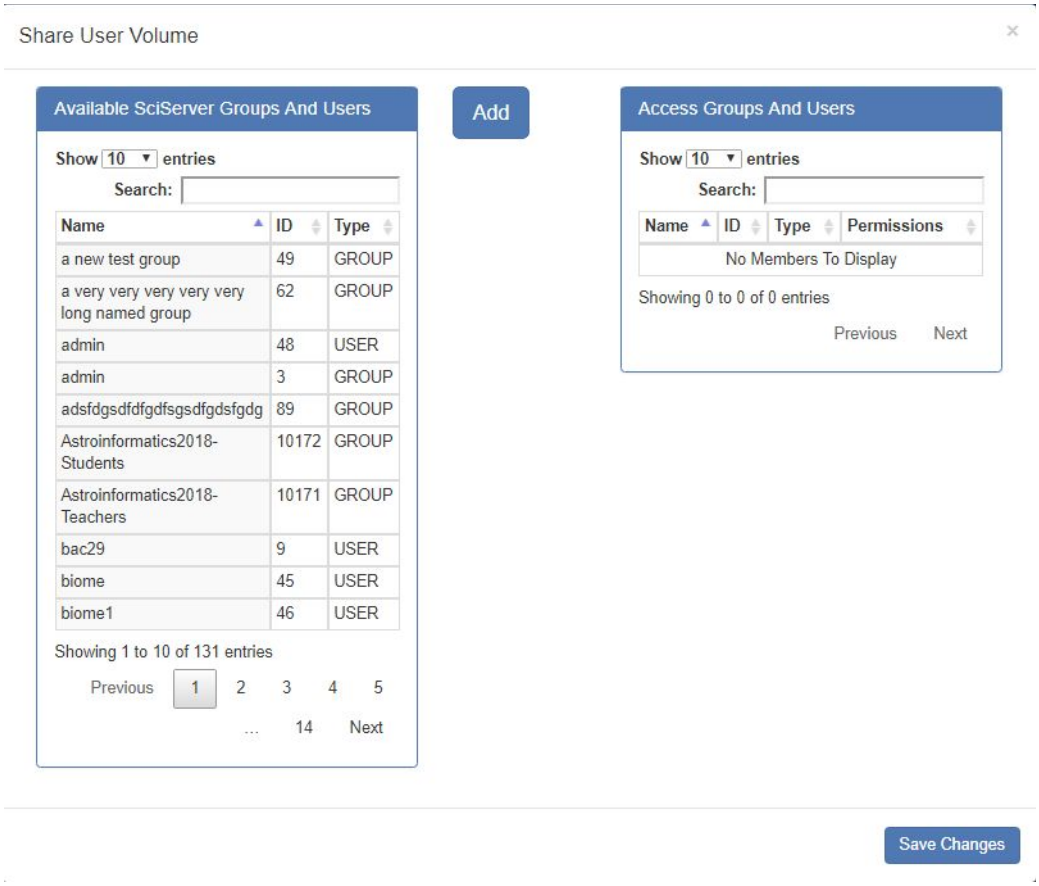
7.6. “How do I share User Volume?”

- You can share any User Volume that you own, and you can share others User Volumes if they assigned you that permission when they shared it with you

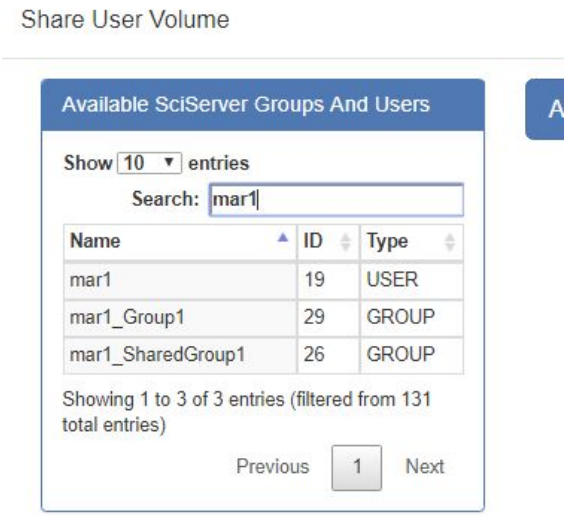
- At the Files tab, click the “home” button  to get to the top level User Volume view, and you will see the list of User Volumes. All User Volumes that you are allowed to share will have a “Share” button next to them:



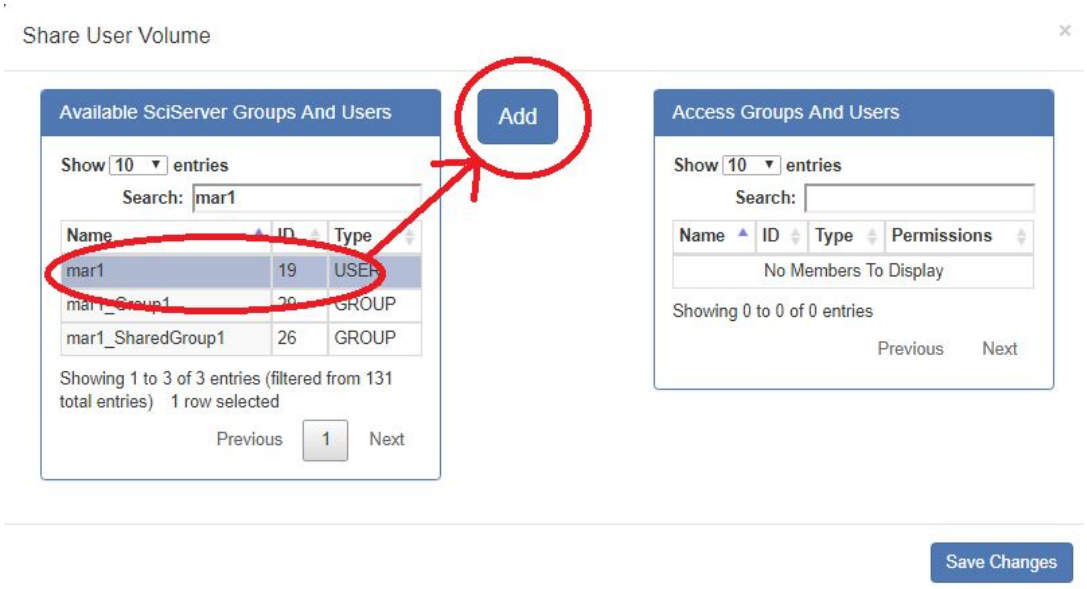
- Click the “Share” button and it will display a pop up dialog box:



- The left hand panel is where you choose, one at a time, the users, or groups of users, that you wish to share with. To find users more easily, type something in the “Search:” box and the list will be filtered:

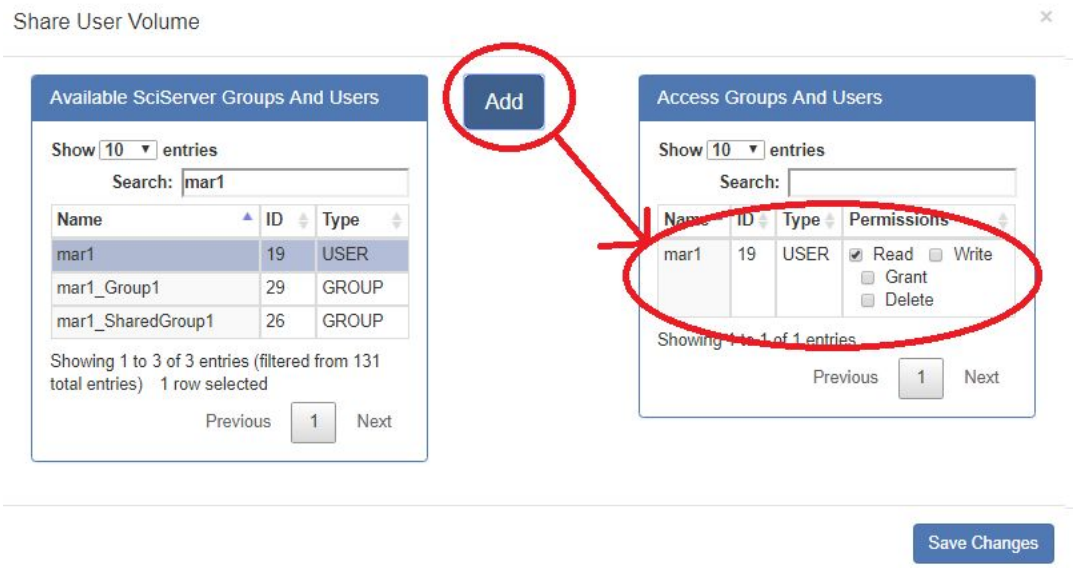


- Click on the name you want to add to the group to highlight it, then press the “Add” button:



- This will add that name to the right hand panel, where you choose the Permissions that you want the user/group to have on your User Volume. These permissions are as follows:
    - *Read*: The user is able to view and read the User Volume and its contents, but not change anything
    - *Write*: The user is able to write new files and folders to the User Volume
    - *Delete*: The user is able to delete the User Volume (not recommended!)

- *Grant*: The user has the ability to also further Share this User Volume with other users.



- Select the Permissions you want to assign by checking the checkboxes.
- You can keep adding more people/groups.
- Press “Save Changes” when you are done.
- The User Volume you have just shared will have its Icon changed to indicate that it has been changed:

7.7. “What do the different Folder Icons mean?”

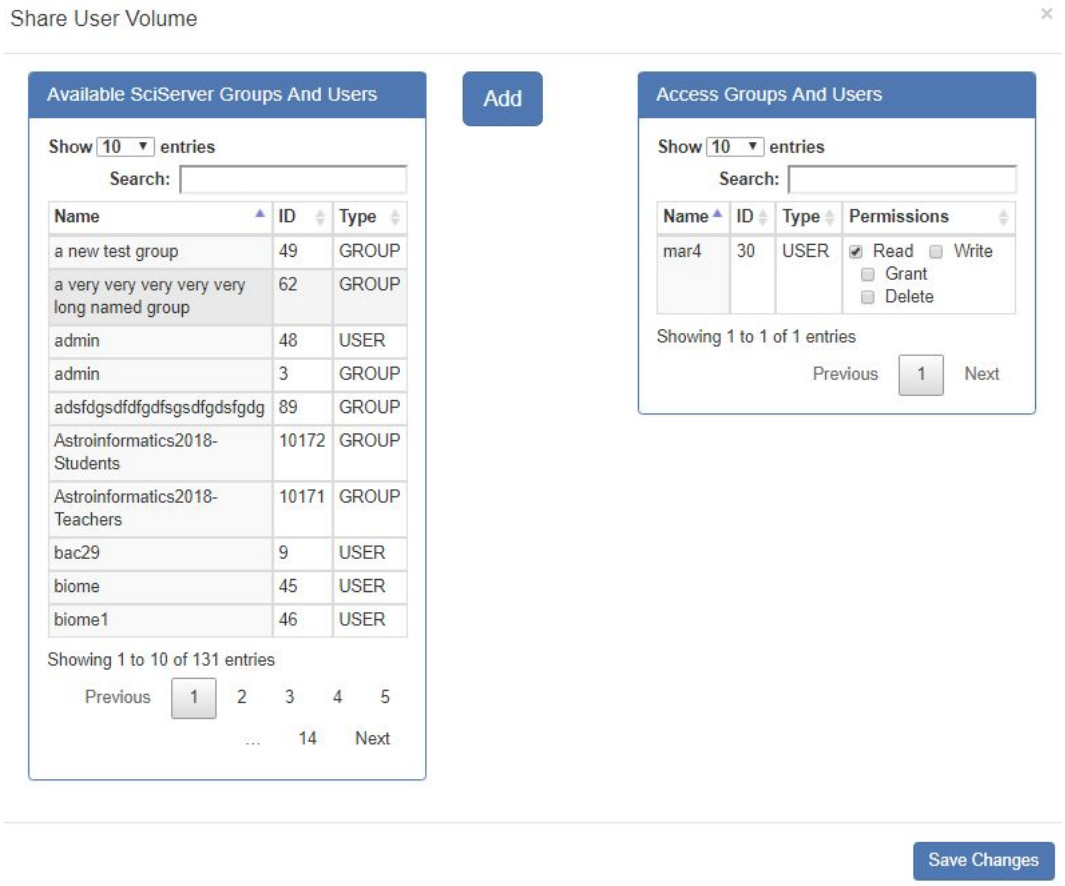
- Regular Folder
- User Volume that has NOT been shared

- User Volume that has been shared BY the User

- User Volume that has been shared WITH the User

7.8. “How do I UNSHARE a User Volume?”

- At the Files tab, click the “home” button  to get to the top level User Volume view, and you will see the list of User Volumes. Click Share button
- Display the Sharing Dialog box for a User Volume that you want to unshare



- Left hand panel ‘Access Groups And Users’ displays the list of users and groups with different permission.
- For each User or Group that you want to unshare, UNCHECK ALL THE CHECKBOXES in the right hand panel.
- Press Save Changes, and the User Volume will be displayed with the regular blue Icon ()indicating that it is no longer shared.

8. Interactive Compute

8.1. "What is SciServer Compute?"

SciServer Compute is an application that allows users to easily create and run Jupyter Notebooks containing code and instructions to analyse and process SciServer hosted data sets. SciServer provides a rich API to access all aspects of SciServers resources, including databases, Filesystems, user and group management, and even Compute Jobs.

There are a couple of steps involved, which SciServer makes easy:

- Create a new "Container" to run the Jupyter Notebook in. A container defines the "environment" for the user, and is configurable
- Open the container, which will start Jupyter, and create, save and execute Notebooks.
- The full capabilities of Jupyter (which is a third party application) are available to the user and will not be covered here.

An important step in setting up a Container environment is specifying what external file systems will be accessible to the Jupyter environment.

8.2. "What is a 'Container'?"

A 'Container' in SciServer Compute is a defined environment within which Jupyter Notebooks can be run. Its is technically a "Docker Container" (Docker is the technology used), and provides a way to isolate the user and their code from the rest of the SciServer system, and other users.

A Container in SciServer Compute is a "long-lived' resource, and as such there are some resource management issues you need to know about

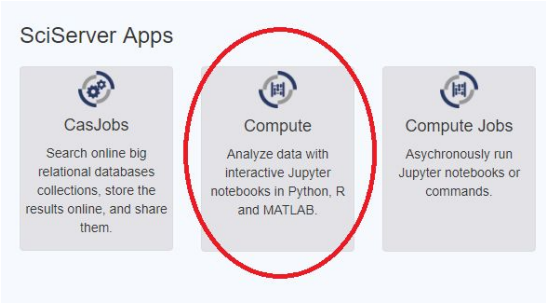
- Each User can create up to 3 containers at any given time. If you need another one, you need to delete one first.
- Containers have a lifecycle, and can be "running" or "stopped". SciServer keeps Containers running for a certain period of time, even if the User is not actively working with it, to ensure that when a user comes back to it, it starts up nice and fast without delays.
- Running Containers consume system resources like memory etc, so after 24hrs (TBC) the Container will be stopped. This has two effects:
    - If code was running, it will be terminated
    - It will take a bit longer to start up the Container next time it is accessed
- SciServer could, and sometimes does, delete containers that have not been used for a long time. This frees up storage resources.
- Whereas data (files, folders) can be stored "in" the container, ***you should never do this for any data that you need to keep.*** **Always store data files in the "Storage" or "Temporary" storage pools, which are external to the container.** Data in these storage pools are accessible when all containers are closed or deleted, and the same

data is accessible across any containers that includes those Volume Containers in their environment.

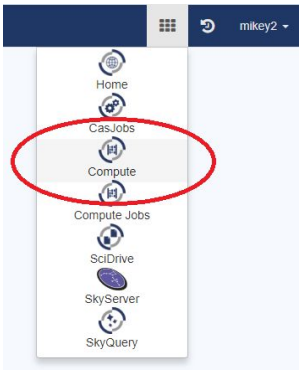
8.3. "How do I create a new Container to run my Jupyter Notebooks"?

Creating a new Compute Container is easy! However there a few parameters required to define the compute environment that you need.

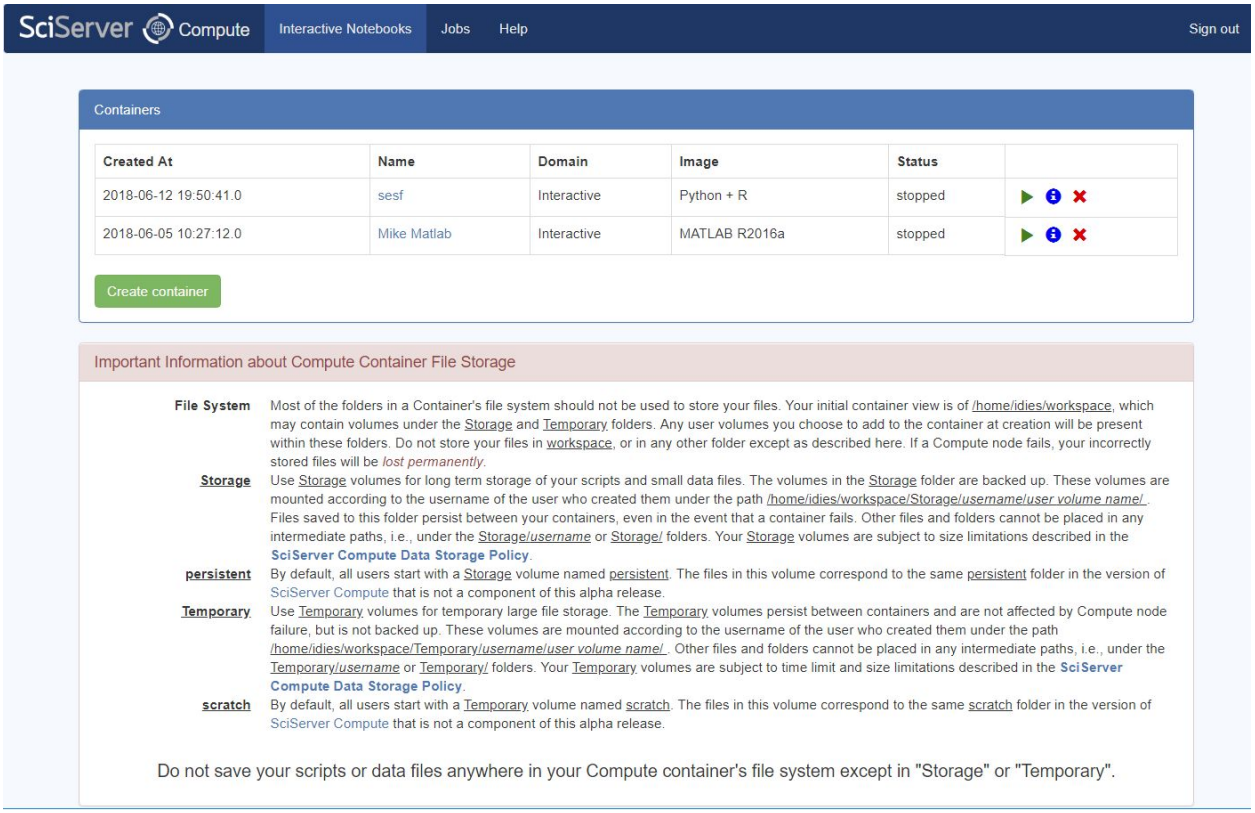
- Access the Compute Application by clicking on the "Compute" icon on the SciServer Dashboard:



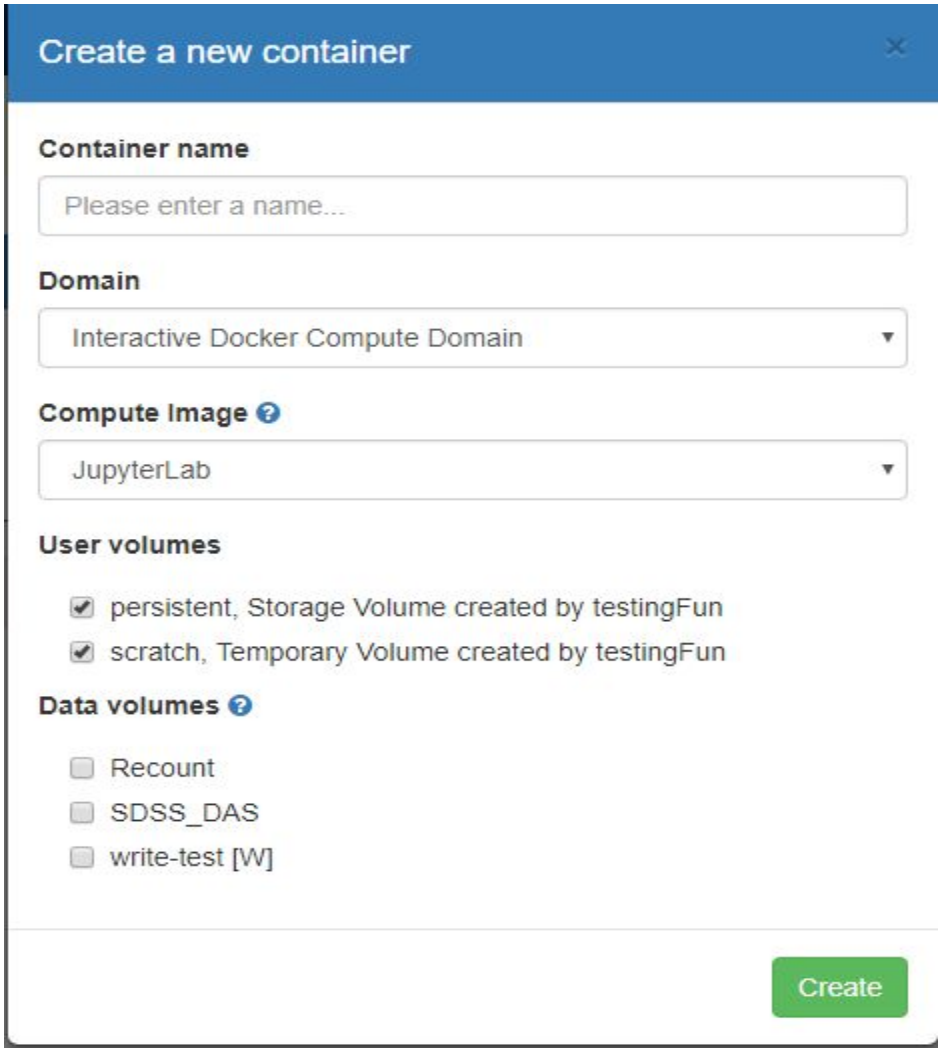
- OR from the "Apps Menu"  :



- This will take you to the "Compute dashboard"



- You will see a large text box explaining how to use the Storage capabilities to ensure that you do not lose data if you accidentally store it in the Container itself
- Press the “Create Container” button, and this will pop a dialog box:



- The following needs to be entered:
  - ***Container Name***: you choose this
  - ***Domain***: This is a drop down from which you should most always leave at the default value “Interactive Docker Compute Domain”.
  - ***Image***: This define a “software environment” for the Jupyter notebooks that you want to run. The images contain libraries tailored to different needs. For the most part you will choose an image that supports the language you are interested in (python, R, Matlab etc), but there are “specialty” science domain specific images that you may have access to if the creator of those images has shared it with you. Additional information on “Images” can be found at http://www.sciserver.org/support/compute-images/
  - ***User Volumes***: This is a list of all User Volumes that you have access to, either which you own, or which have been shared with you. When you select some of these, on container creation these folders will be “mounted” and will be accessible as if they were local files. This makes file access and management

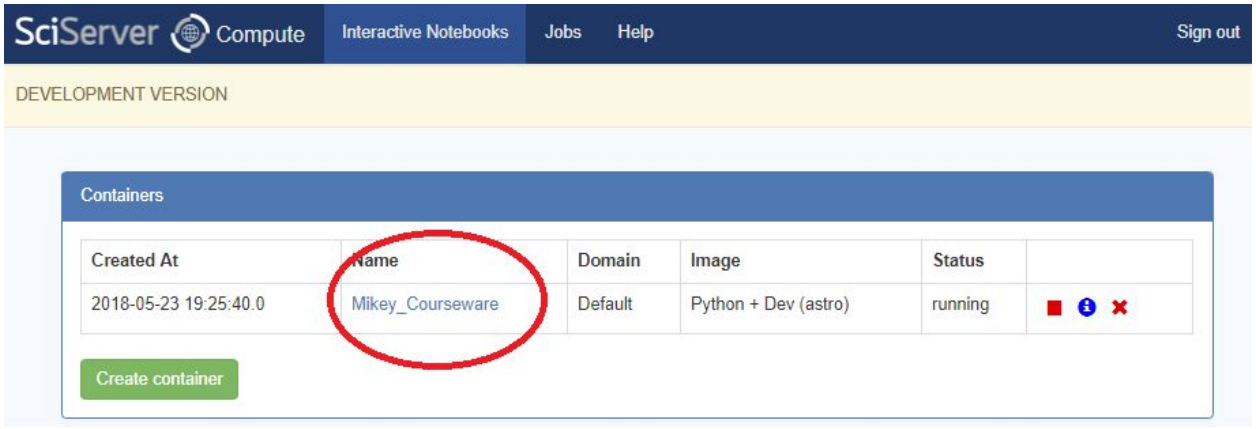
much simpler. NOTE: they will be mounted with the same access controls as you would have in the SciServer File UI i.e. “readonly” or “readwrite”.

- ***Data Volumes***: These are a series of special Data Volumes that are either shared publicly with all users, or for which you have been given special access privileges to see. Again, selecting these Volumes will mount them, and make them appear “local” in the Container. These will always be mounted “readonly”. Additional information on “Data Volumes” can be found at http://www.sciserver.org/datasets/

- Once created, it is not possible to add additional User Volumes or Data Volumes, so you should be sure to get this right.
- Pressing the “Create” button will create a new Container and show it in the table with the name you provided.
- Clicking on the link in the column titled “name” will launch Jupyter

8.4. “How do I create a new Jupyter Notebook?”

- Start one of your Compute Containers to Launch Jupyter:



9. Compute Jobs

9.1. “What is Compute Jobs?”

Compute Jobs allows a user to run a Jupyter Notebook or a standard script in offline batch mode. The same exact capabilities are provided as for Interactive Compute:

- Compute Images and software environment
- Mounting external volume folders

Executing Job will put it in a queue, and it will be run when there are resources available on the server cluster.

9.2. “Why would I create a Job?”

You should create a Job for the following reasons:

- Executing your notebook may take a long time and you want to set it running and do something else without worrying about browser sessions timing out etc
- You may develop your code interactively to make sure the algorithm works, using a small amount of data to test it out. But you really want to run your code against a full dataset which will require massive resources for memory and CPU, as well as execution time.
- You are provided with far more resources (CPU and memory) to execute a Job than you are in an Interactive Session.

9.3. “What are the two types of Jobs I can create?”

SciServer allows you to define two “types” of job:

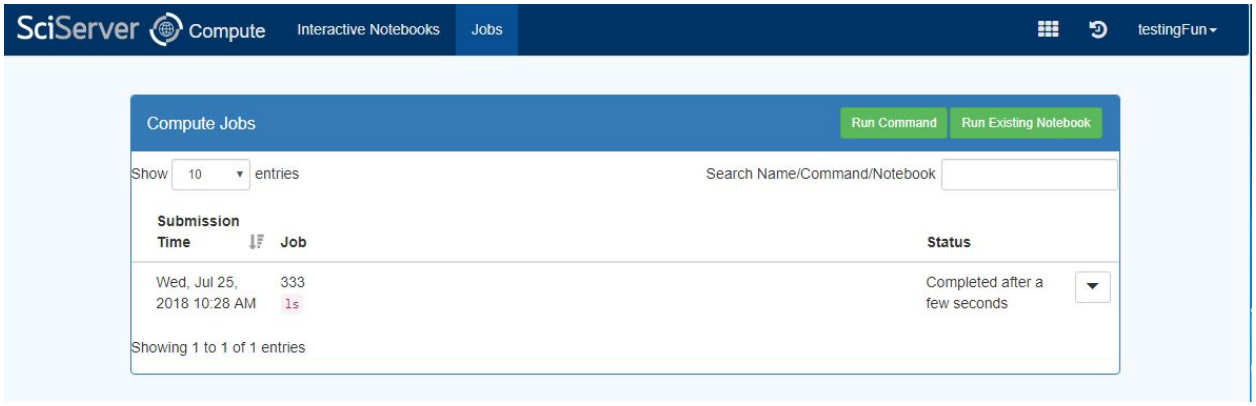
- Specify a script to execute, or a command line command
- Specify an existing Jupyter Notebook that you have previously developed

The second of these is the most useful in that you can develop your Jupyter Notebook interactively then “submit” the exact same notebook as a Job.

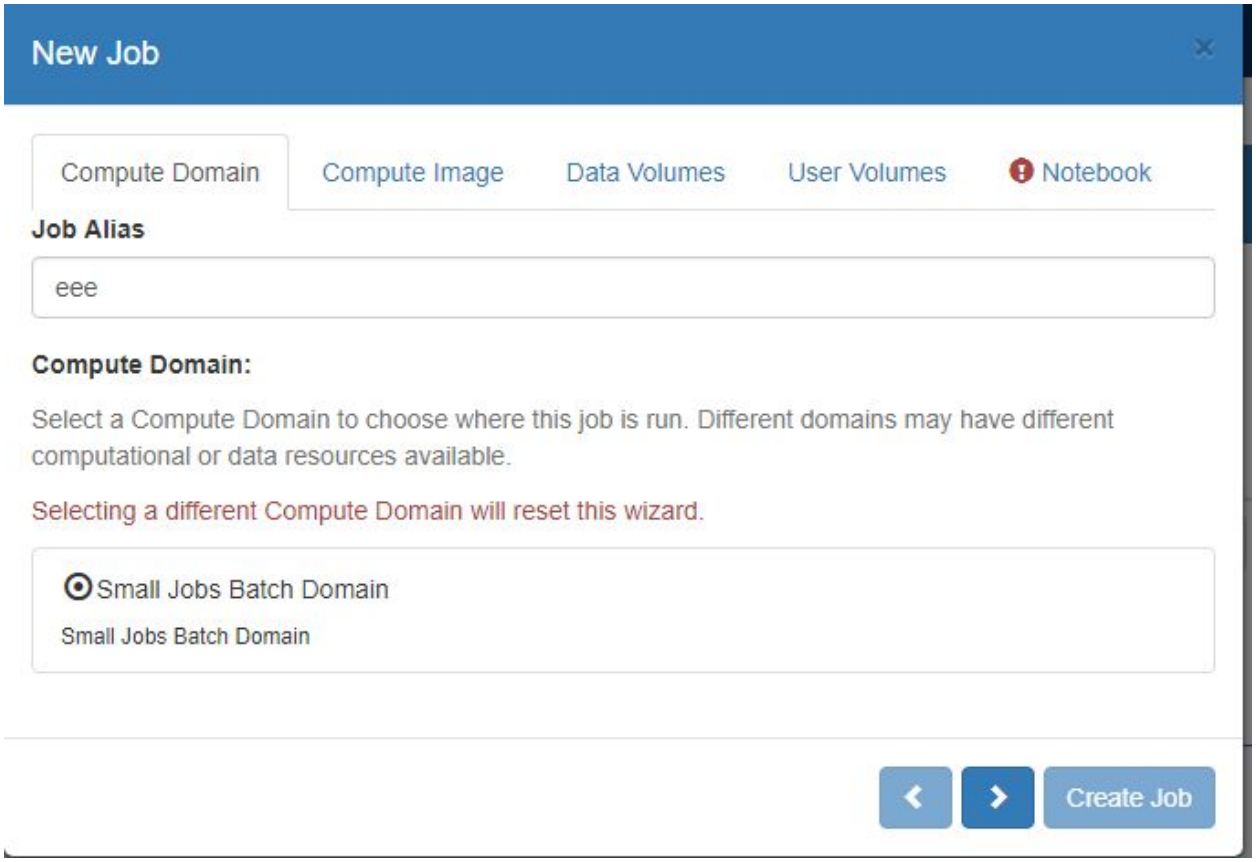
9.4. “How do I create and run a Job?”

Creating a new Job is easy! We explain how to create a notebook based job, but creating a script based job is very similar.

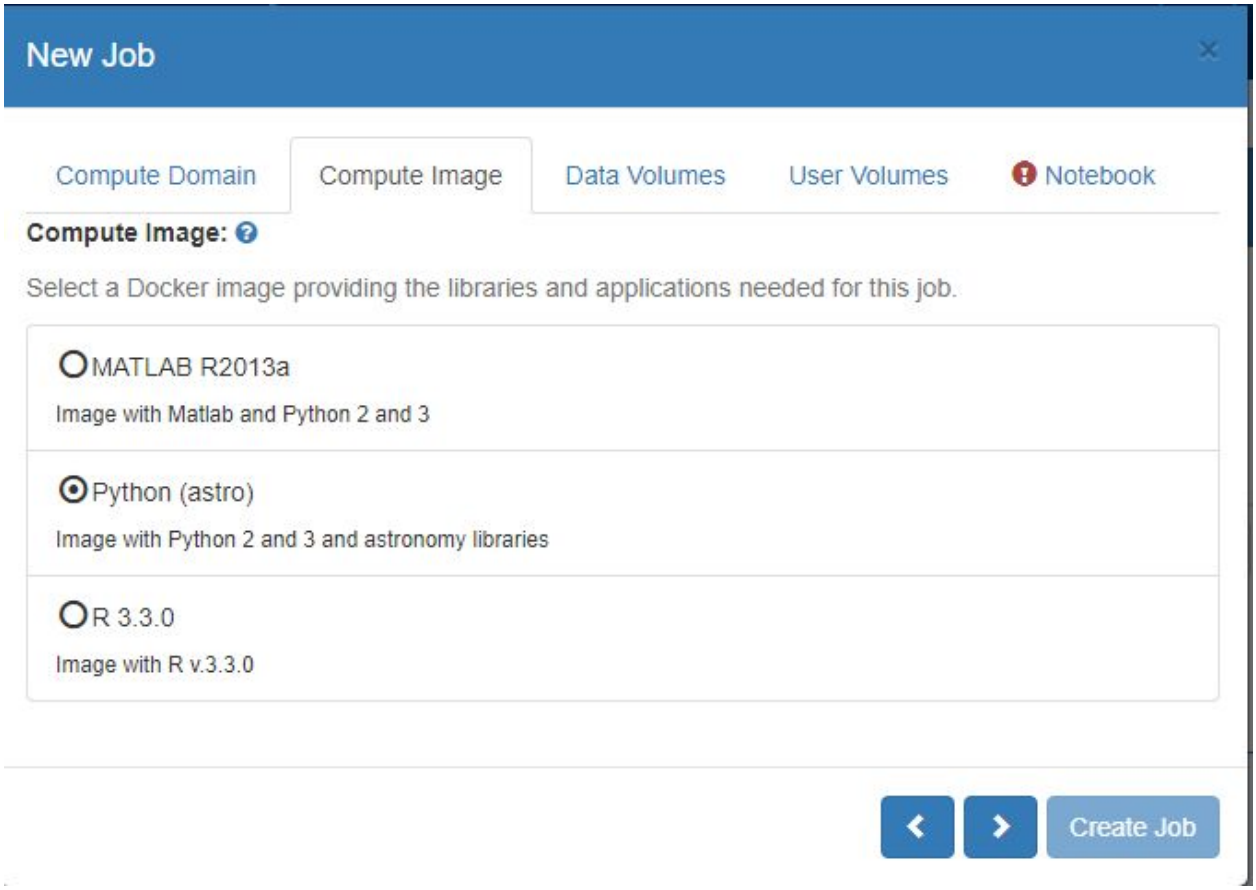
- Go to the Compute Jobs Page:



- Click “Run Existing Notebook”

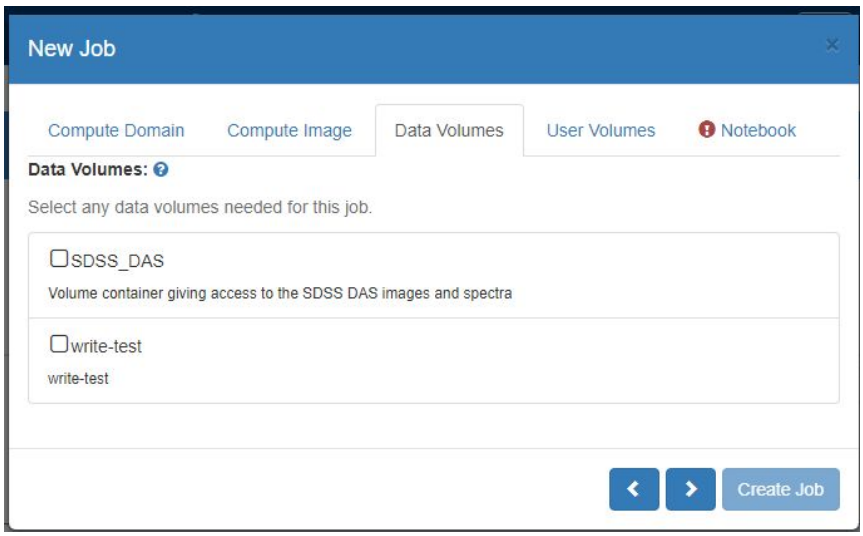


- On the '**Compute Domain'** Tab:
    - Choose the Compute Domain, for which in most cases currently there will only be one option
    - Optionally enter a "Job Alias' to easily identify your Job later
- On the '**Compute Image**' Tab:



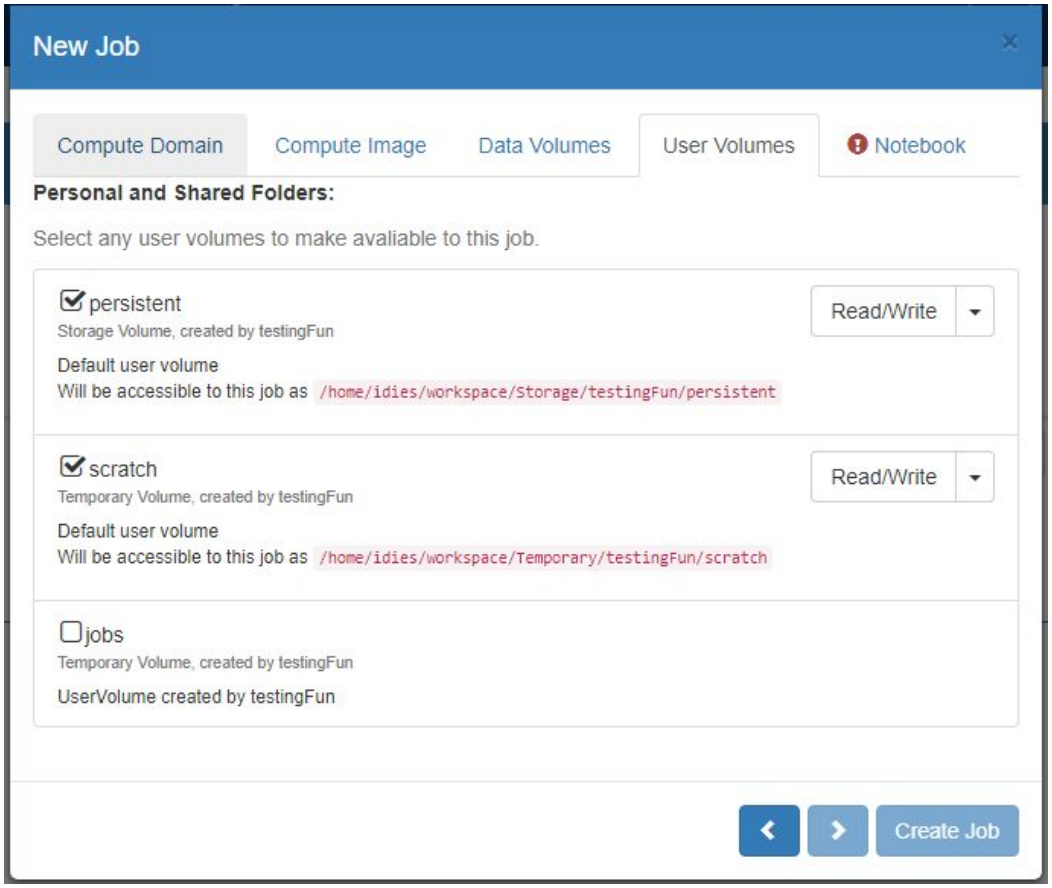
- Pick the 'Image' you need to use. Each image contains different tools and programming language support.
- (Compute Images are described in more detail at http://www.sciserver.org/support/compute-images/
-

- On the **'Data Volumes**' tab:



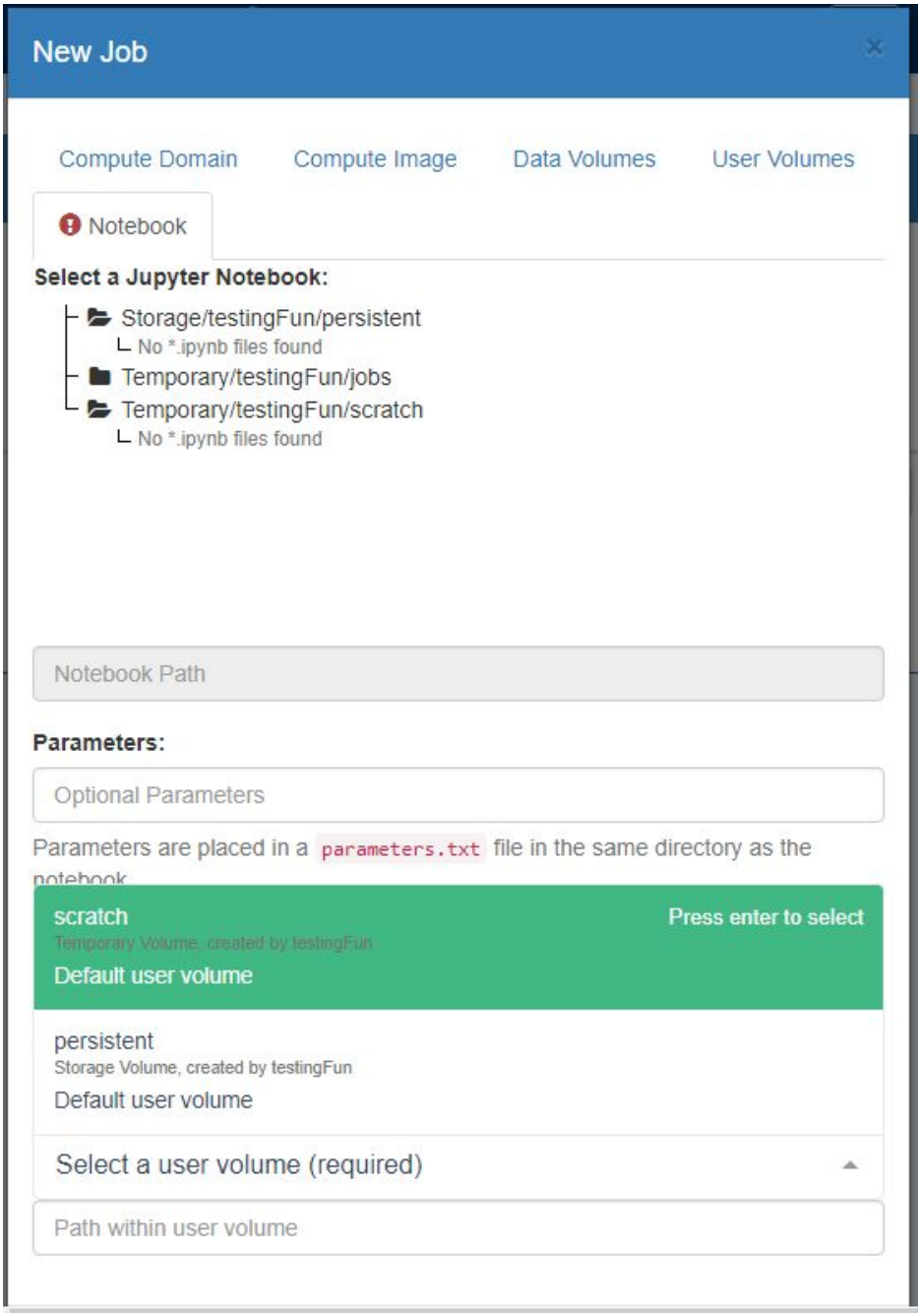
- Select all the data volumes with appropriate permissions needed for this job.

- On the '**User Volumes'** Tab:



  - Selectall the Folder systems that you would like to be made accessible to your Compute notebook
  - For Folders that you own, or that have been shared with you and you were given the appropriate permissions, you can select whether a given folder is read only or writable. Folders that you do not own will be readonly by default.

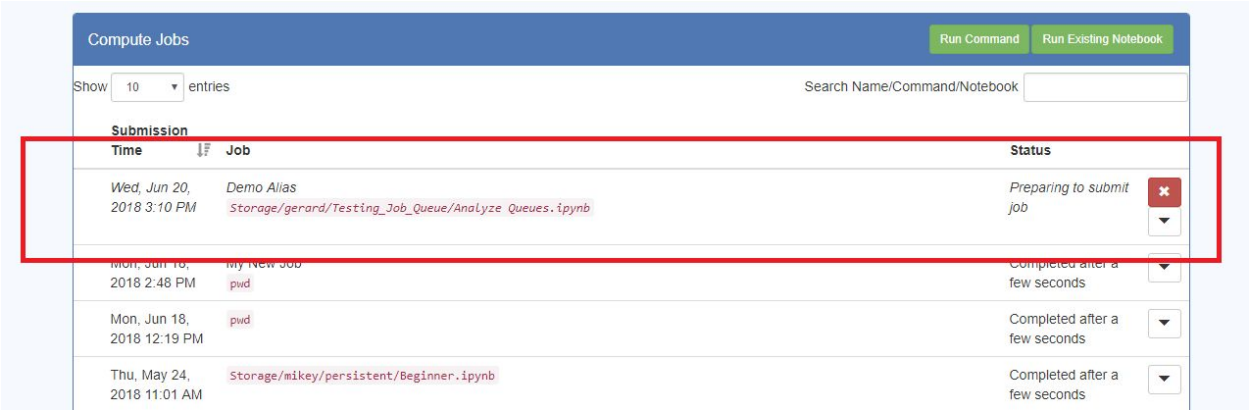
- On the **'Notebook'** tab:



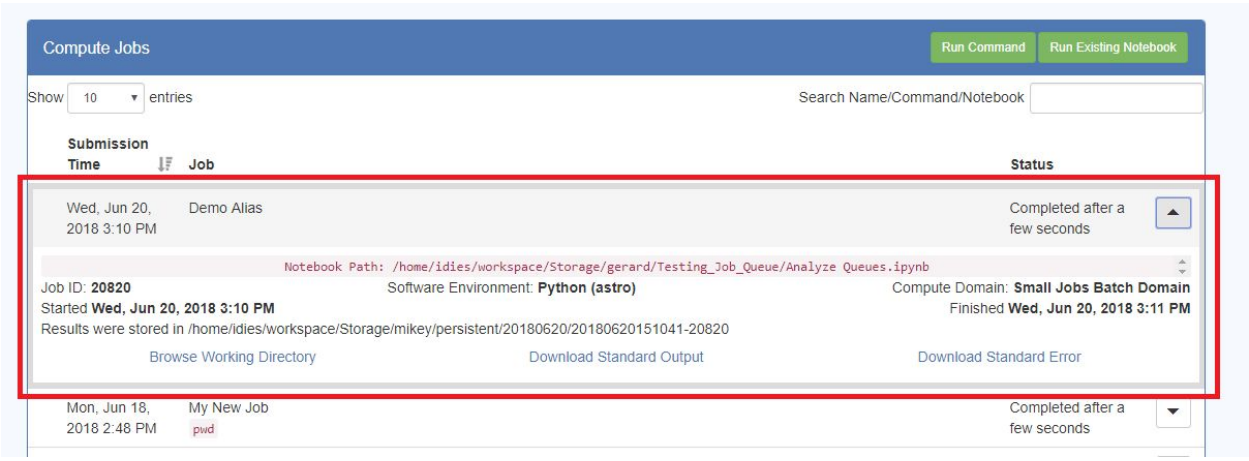
- Navigate to the Notebook you wish to use as the basis for your Job, and select it
- Enter any additional parameters that the Notebook can read in to affect how the code is executed
- Choose a directory where the output results will go.
    - By default these will go to jobs within which subdirectories will be created and your results written to.

- Alternatively you can choose a specific directory to output results. The directory you choose will be a 'root' within which subdirectories will be created and your results written to.

- When everything has been entered you can press 'Create Job', and the Job will be submitted, and displayed in a Jobs Table view:



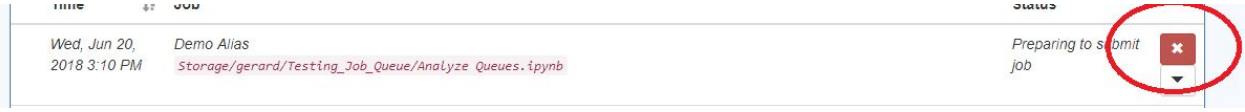
- The Table will be refreshed every several seconds, telling you the status of the Job.
- While the Job is still running there will be a red "X" button, and pressing this will Cancel the job.
- Pressing the down triangle on the RHS will expand the view and show more information about the Job. This is what you see for a completed Job:



- This give status information about the Job, the path to the location of the results, and links to the results output

9.5. "How do I Cancel a Job?"

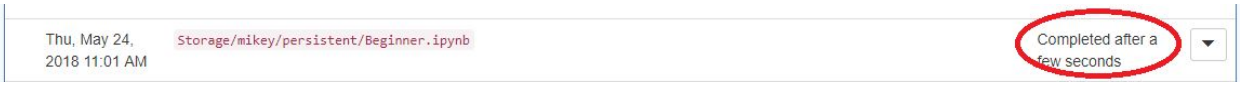
- After submitting a Job, and before it has completed execution, it can be cancelled.
- Display the Jobs table, and your running Jobs will be identified by having a red "X" next to them:



- Press the “X” button to cancel the job.

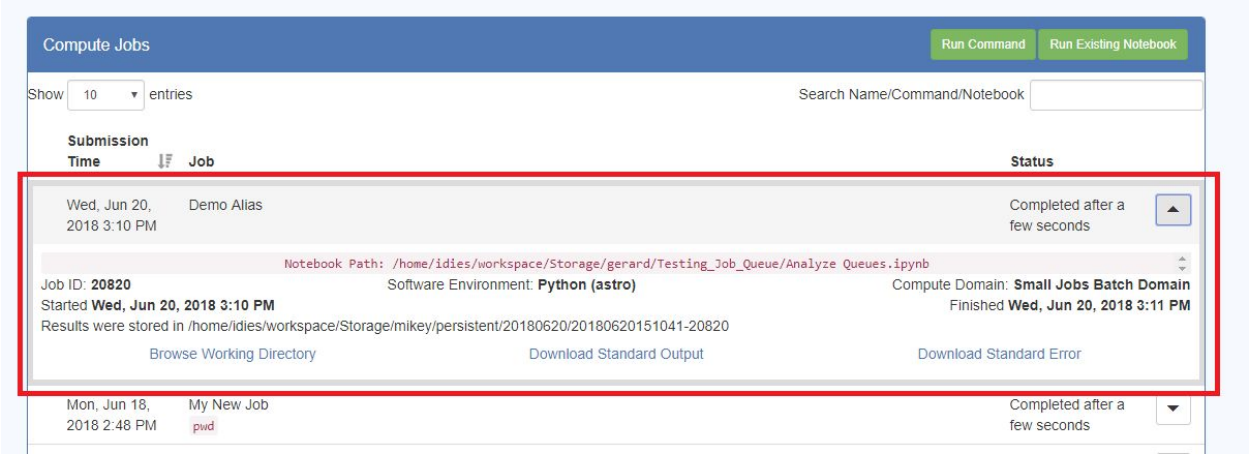
9.6. “How do I know when my Job is complete?”

- When your job is complete the Jobs Table will display the status:



9.7. “How do I access the output from my Jobs?”

- The results output will go to the location specified in the Job Definition.
- In the Jobs Table expand the job of interest:



- This give status information about the Job, the path to the location of the results, and three hyperlinks:
  - ***‘Browse Working Directory’*** will take you to the Dashboard Files tab and show you the output files as well as the original Python Notebooks.
  - ***‘Download Standard Output’ and ‘Download Standard Error’*** will allow you to download these two text files as appropriate. They will be downloaded according to your Browser settings.