Data Transfer

Ritu Arora

For any questions, comments, or concerns, please contact me at:
rauta@tacc.utexas.edu

October 13, 2015

Protocols for Data Transfer

- Different protocols exist for data transfer between remote sites, or a local computer and a remote system *e.g.*,
  1. Linux command-line utilities **`scp`** & **`rsync`**
     1. **These will work on both Linux and Mac computers**
  2. **`WinSCP`**
     1. **This is used on Windows**
  3. Globus Connect
  4. Globus' globus-url-copy command-line utility
- We will only cover `scp`, `rsync`, and `WinSCP` in these slides

Data Transfer Using **scp**

- If your local computer is a Mac or a Linux laptop, you can use the `scp` commands to transfer data to and from a remote resource like Stampede

```
localhost% scp filename
username@stampede.tacc.utexas.edu:/path/to/project/
directory
```

- If you are using a Windows computer, you can download and use the WinSCP application (GUI-based) or download and use Cygwin (command-line based, can run the aforementioned commands)

Data Transfer Using `rsync` (1)

- The `rsync` command is another way to transfer data and to keep the data at the source and destination in sync
- If transferring the data for the first time to a remote resource, `rsync` and `scp` might show similar performance except when the connection drops
  - If a connection drops, upon restart of the data transfer, `rsync` will automatically transfer only the remaining files to the destination, it will skip the already transferred files
- `rsync` transfers only the actual changed parts of a file (instead of transferring an entire file)
  - this selective method of data transfer can be much more efficient than `scp` because it reduces the amount of data sent over the network

Data Transfer Using `rsync` (2)

- The following example demonstrates the usage of the `rsync` command for transferring a file named `myfile.c` from the current location on Stampede to Lonestar's `$WORK` directory
  - Lonestar is another supercomputer a TACC
  - For running this command, you should be connected to Stampede

```
login1$ rsync myfile.c
username@lonestar.tacc.utexas.edu:/work/01698/username/
data
```

Data Transfer Using **rsync** (3)

- Transferring an entire directory from Stampede to Lonestar
  - To preserve the modification times use the `-t` option
  - To preserve symbolic links, devices, attributes, permissions, ownerships, etc. transfer in the archive mode using the `-a` option
  - To increase the amount of information displayed during transfer use the `-v` option (verbose mode)
  - To compress the data for transfer, use the `-z` option
  - The following example demonstrates the usage of the `-avtz` options for transferring a directory named `gauss` from the present working directory on Stampede to a directory named `data` in the `$WORK` file system on Lonestar

```
login1$ rsync -avtz ./gauss
username@lonestar.tacc.utexas.edu:/work/01698/
username/data
```

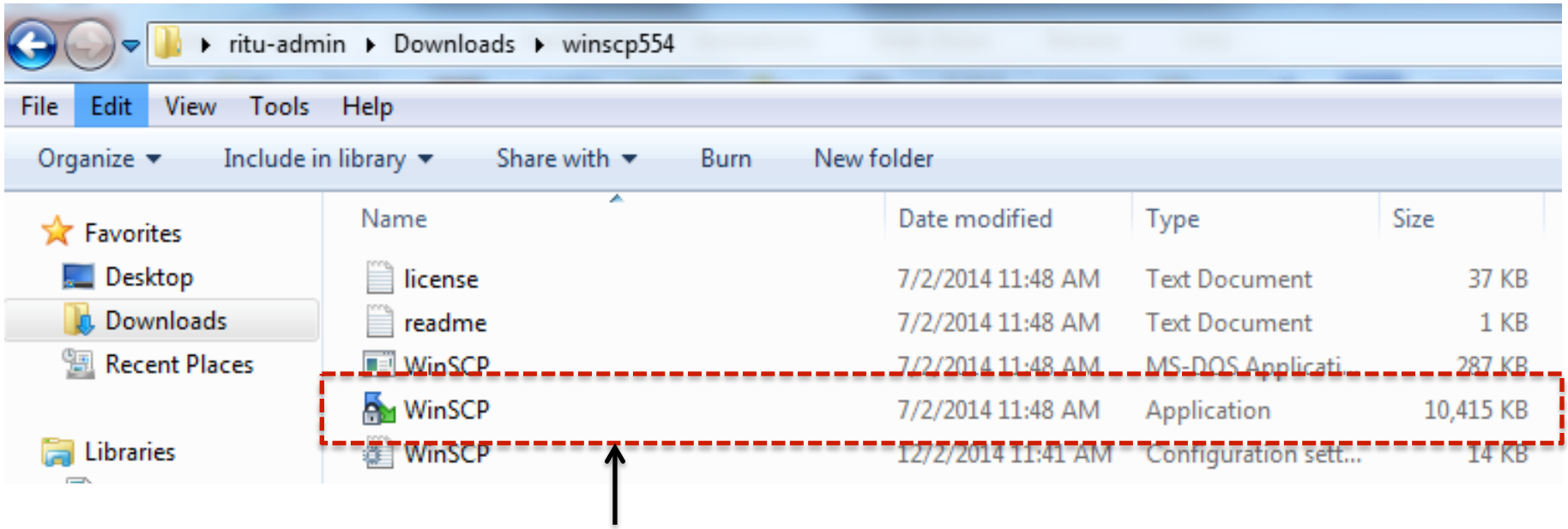
WinSCP

- WinSCP is a free software that can be used for transferring files between a Windows computer and a remote Linux system

- Link to download and install WinSCP is as follows:

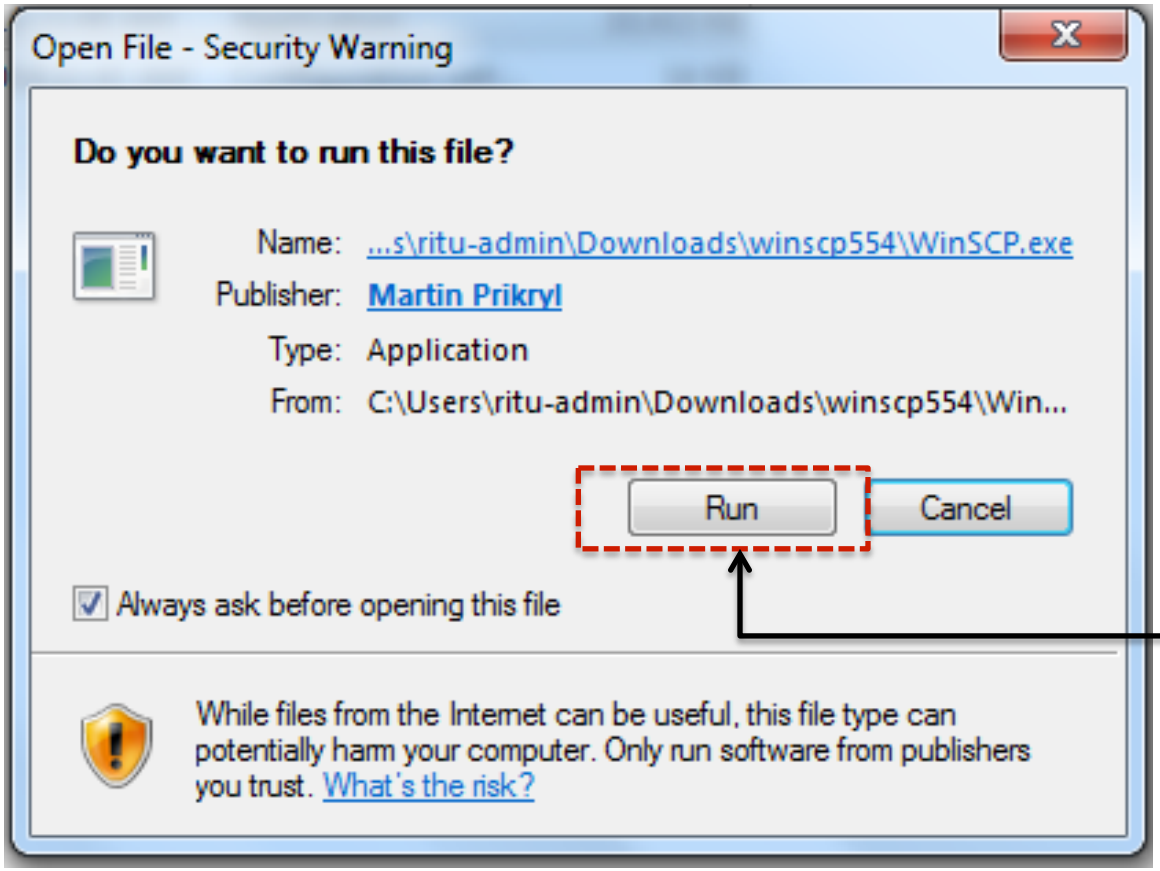
  http://winscp.net/eng/download.php

Launching WinSCP: Step 1



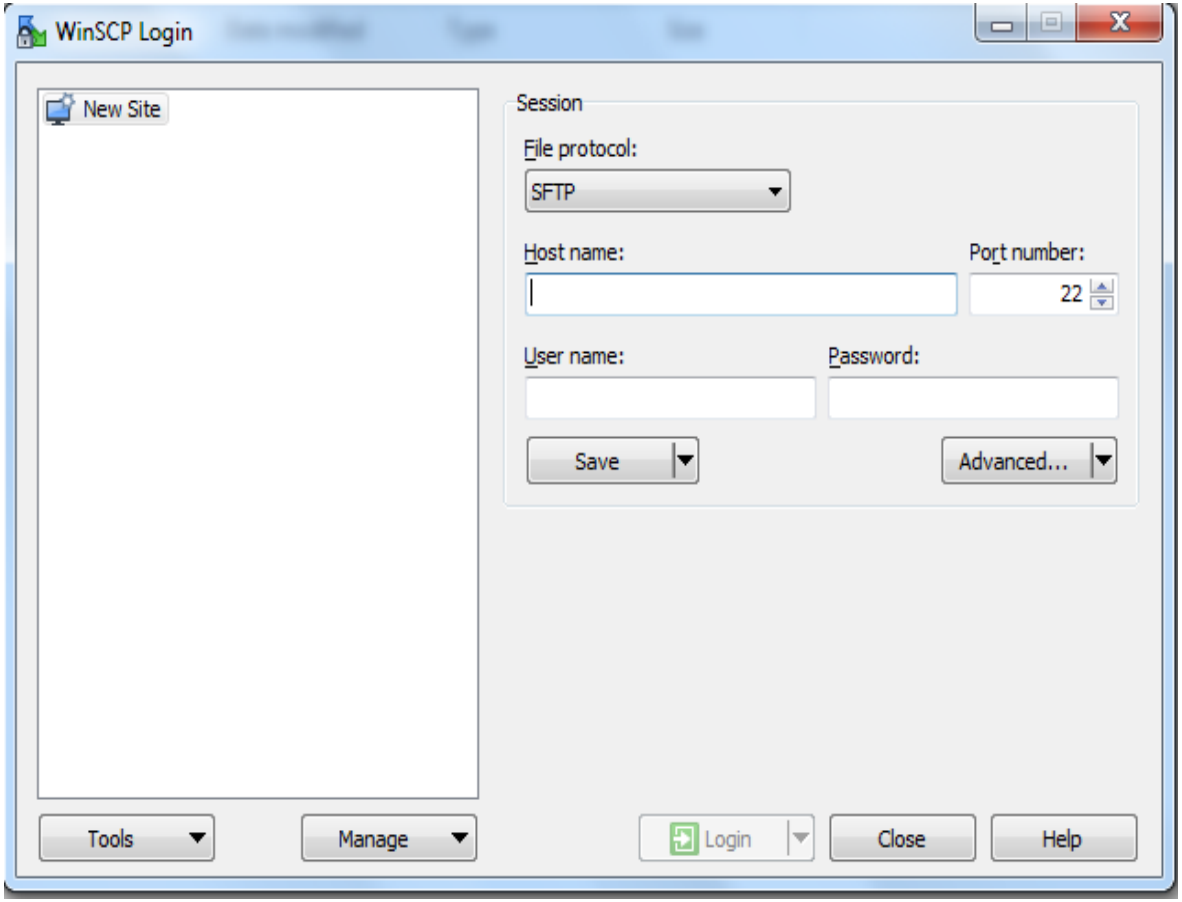
- After installing WinSCP on your Windows computer, go to the folder in which you have downloaded it, and click on the icon for WinSCP application to launch it

Launching WinSCP: Step 2



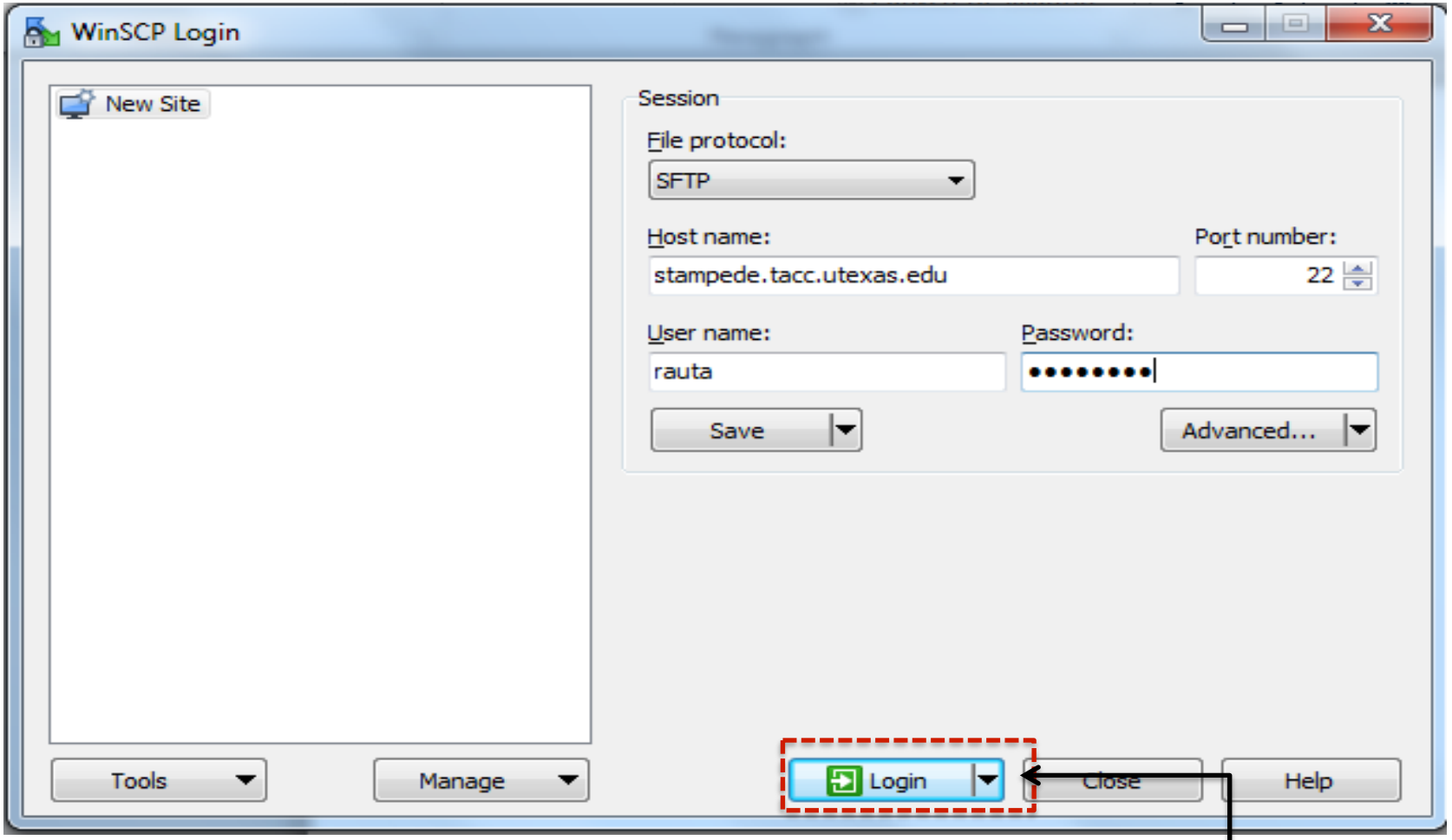
- Upon clicking the WinSCP application icon, a window like the one shown on the slide may pop-up
- Click on “Run”

Launching WinSCP: Step 3 A



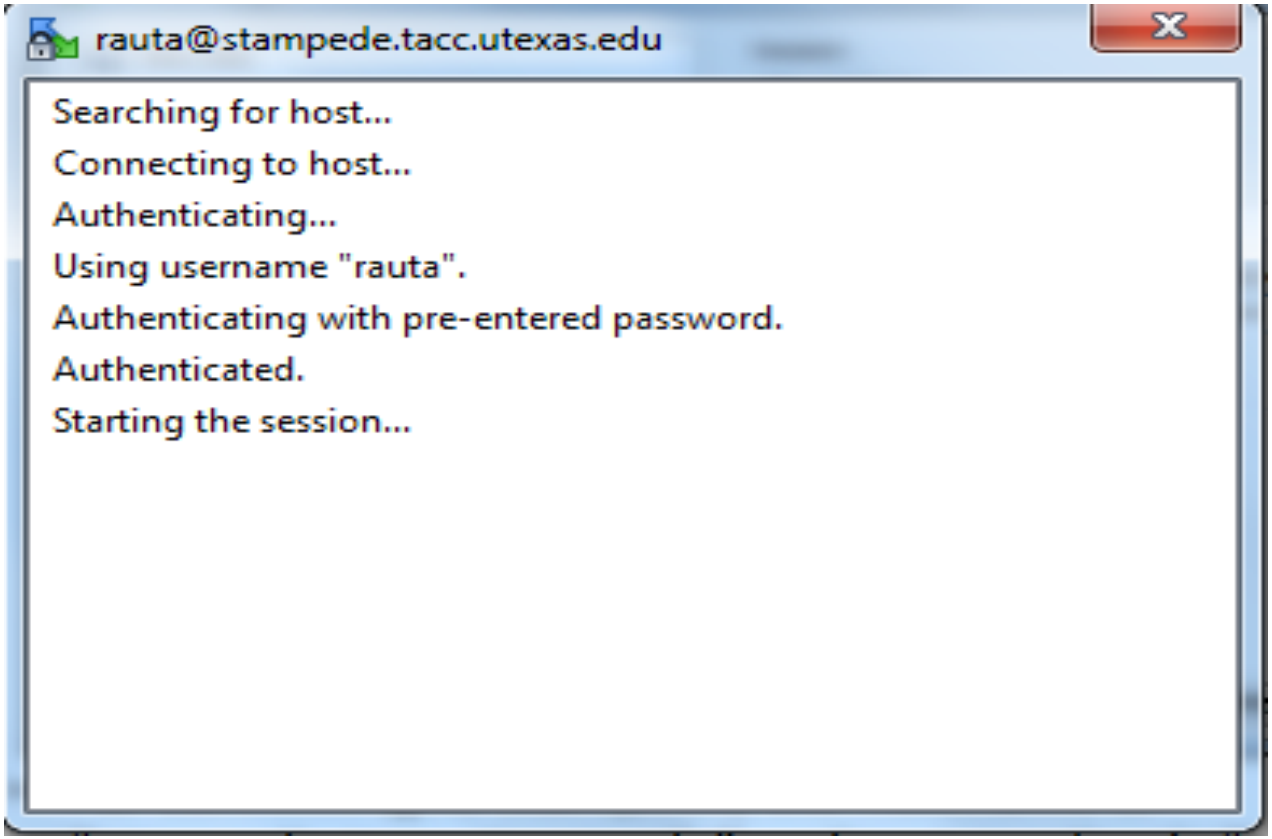
- After clicking on "Run" in Step 2, a window as shown on this slide will pop-up
- Type the Host-name, user-name, and password in this window and click on "Login" – see next slide for this step

Launching WinSCP: Step 3 B



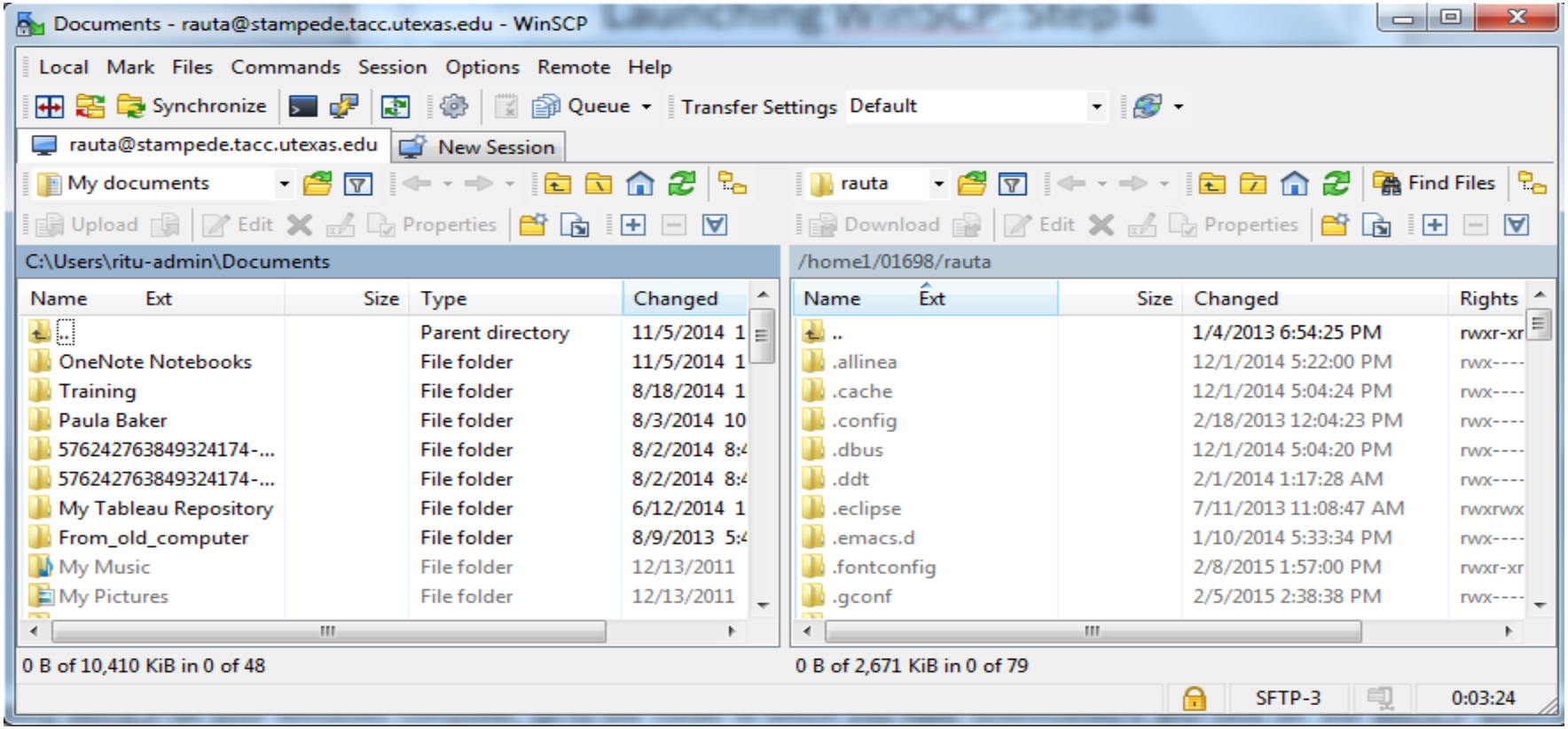
- For connecting to Stampede, the Hostname is "stampede.tacc.utexas.edu", and port number is "22"
- Type your Stampede user-name and password and click on "Login"

Launching WinSCP: Step 3 C



- After you click on “Login” in Step 3 B, you will see a window like the one shown on the slide – you do not need to do anything at this stage

Launching WinSCP: Step 4



- After Step 3 C (assuming your credentials were accepted on Stampede), you will see a window for file transfer as shown on the slide
- Left-hand-side of the window shows the local file-system (from your Windows computer), and the right-hand-side shows the $HOME file-system on the remote Linux system, which is Stampede here

Drag and Drop Files to Transfer

- Once you have reached at Step 4, you can select the directories for file transfer – a source directory on the left-hand-side window and a target directory on the right-hand-side window

- Please note: Stampede is connected to three file systems, each having its own quota and storage policies
  - Some information on these file systems in on the next slide
  - More information can be found in the Stampede user-guide at:
    https://portal.tacc.utexas.edu/user-guides/stampede