Introduction to BioHPC Lab

BioHPC Lab Workshop

Jaroslaw Pillardy

Bioinformatics Facility
Institute of Biotechnology
Cornell University

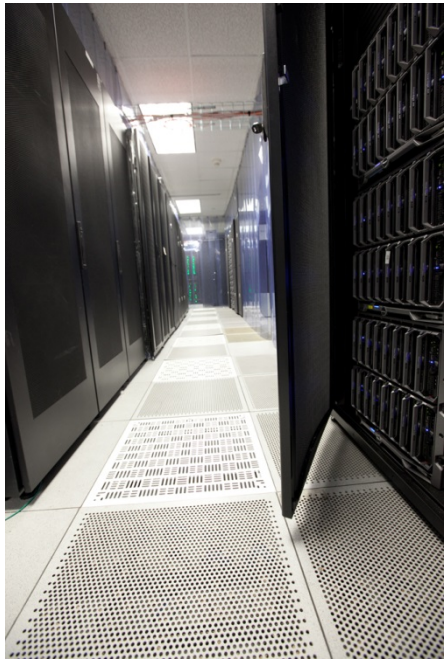
http://cbsu.tc.cornell.edu/lab/lab.aspx

http://cbsu.tc.cornell.edu/lab/doc/Introduction_to_BioHPC_Lab.pdf

BioHPC Laboratory

BioHPC Laboratory is a computational resource configured and optimized for the needs of computational biology and bioinformatics

- Available to Cornell and external users
- Provides access to computing, storage and software (over 116 titles)
- Located in Rhodes Hall and accessible remotely and on-site



BioHPC Lab: computing

Two ways of using BioHPC Lab computing resources:

- Buy hours and run computations on BioHPC Lab workstations/servers

  Only pay for the hours you need, avoid problems with maintaining your own hardware, software and resource allocation.

- Buy your own hardware (or ask us to buy it for you) and host it in BioHPC Lab

  If you need a special hardware, or you expect to use it 24/7, you can still avoid problems related to maintaining your own hardware, software and resource allocation by hosting the machine with the Lab

BioHPC Lab: storage

- 256 TB of networked storage available in one volume
- Very robust and scalable architecture: cluster storage based on RAID6 servers connected by Gluster, new servers can be added to expand storage.
- Very affordable! $80 per TB per year is lower than even Amazon archival storage (Amazon Glacier: $120 per TB per year + data transfer fees)
- Users get free storage allocations with the accounts

BioHPC Lab: software

- BioHPC Lab is pre-configured for bioinformatics with software and related software infrastructure (libraries, development tools etc.).
- All software installed as a response to our or our users computing needs.
  If you need a program that is not installed, ask us. We may install it, it depends on possible usage level and time investment required to deploy.
- 116 titles as of 2/13/2014
- Common genomic data is available locally in the Lab: sequence and annotation databases, preformatted for common programs

http://cbsu.tc.cornell.edu/lab/labsoftware.aspx

Setting up an account

- In order to get an account send an e-mail to cbsu@cornell.edu with your name, Cornell Netid and affiliation

- All Cornell employees, students and alumni are eligible

- Any collaborators of Cornell employees, students and alumni are eligible

- External users requests are considered on case-by-case basis, depending on Lab overall usage.

Getting started with a new account

- You need hours: create and fund your own Lab Credit Account or get added to one
- Get extra storage if needed – all users get free storage allocations, but it may not be enough
- Transfer data to your Lab storage
- Make reservation(s)
- Connect to reserved workstations
- *Compute!*

Getting started with a new account

In order to execute last item of the workflow (*Compute!*) some additional knowledge may be needed

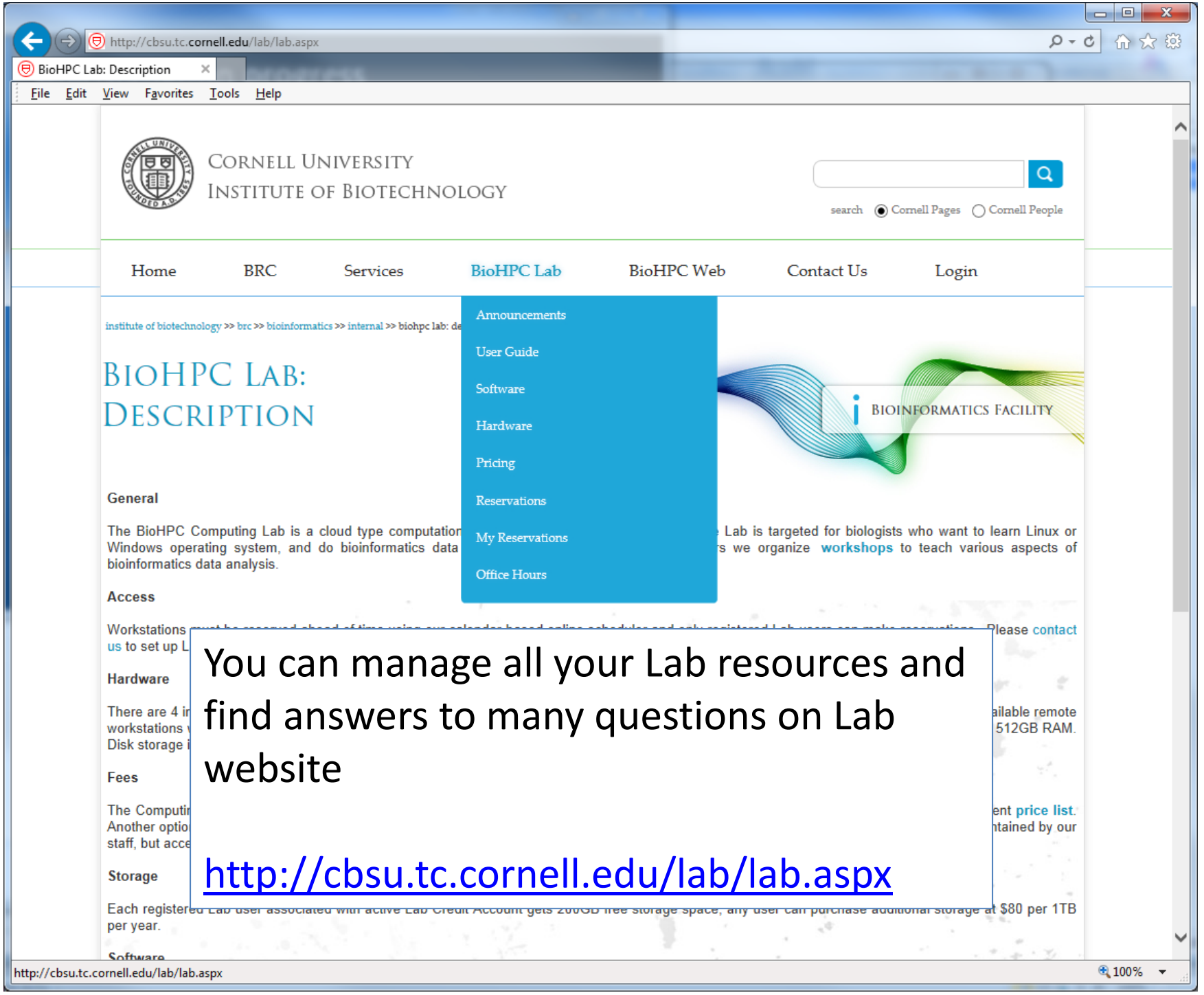
- how to work in Linux environment?
  => see our "***Linux for Biologists***" workshop
  http://cbsu.tc.cornell.edu/ww/1/Default.aspx?wid=45

- what programs to use for my bioinformatics problems? How?
  => use our ***facility office hours*** to discuss problems and solutions
  http://cbsu.tc.cornell.edu/lab/office1.aspx

- sometimes custom data processing is needed that requires a little bit of tinkering with software and scripts

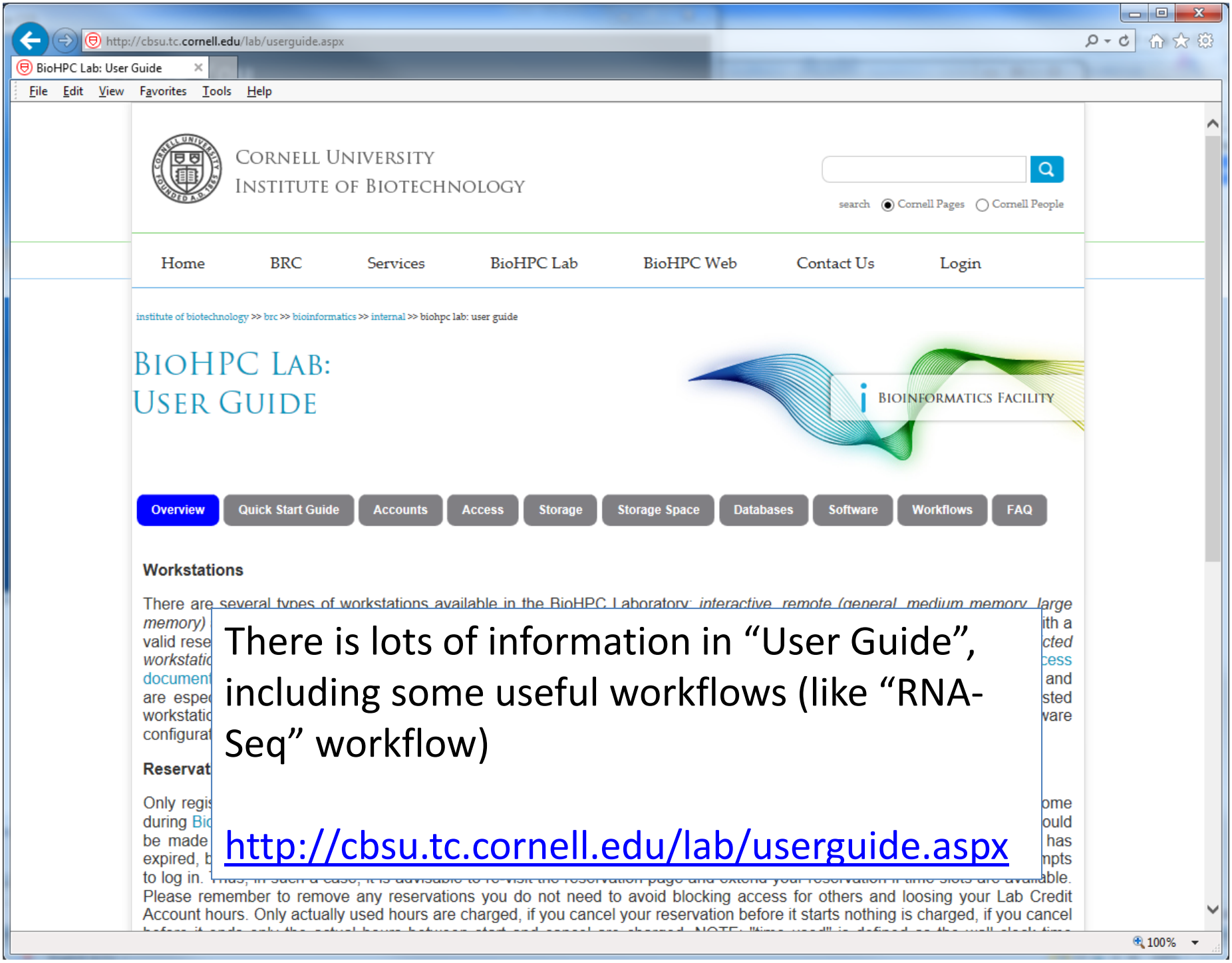
  => see our "***Perl for Biologists***" workshop
  http://cbsu.tc.cornell.edu/ww/1/Default.aspx?wid=46

  => use our ***facility office hours*** to discuss problems and solutions
  http://cbsu.tc.cornell.edu/lab/office1.aspx



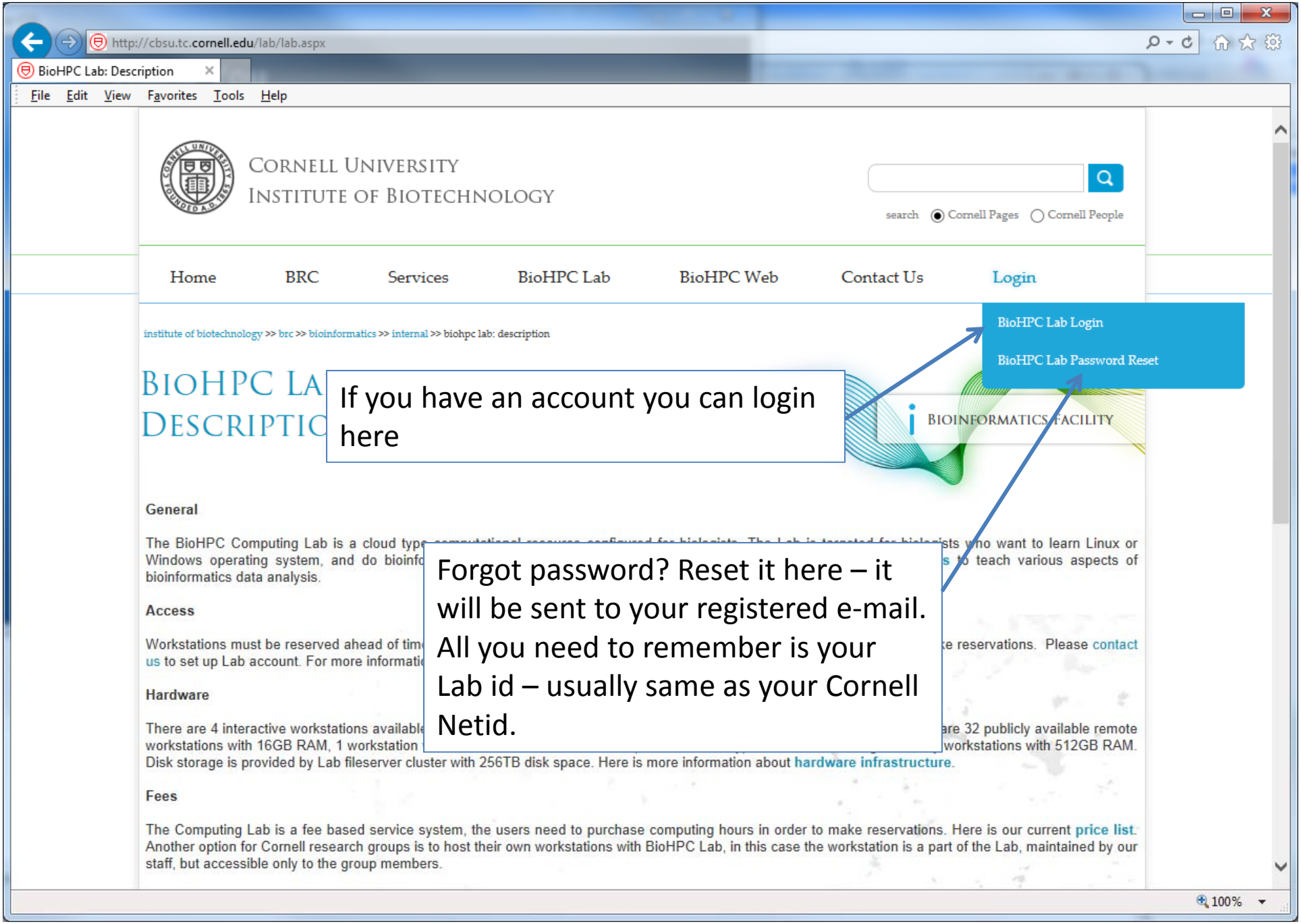
You can manage all your Lab resources and find answers to many questions on Lab website

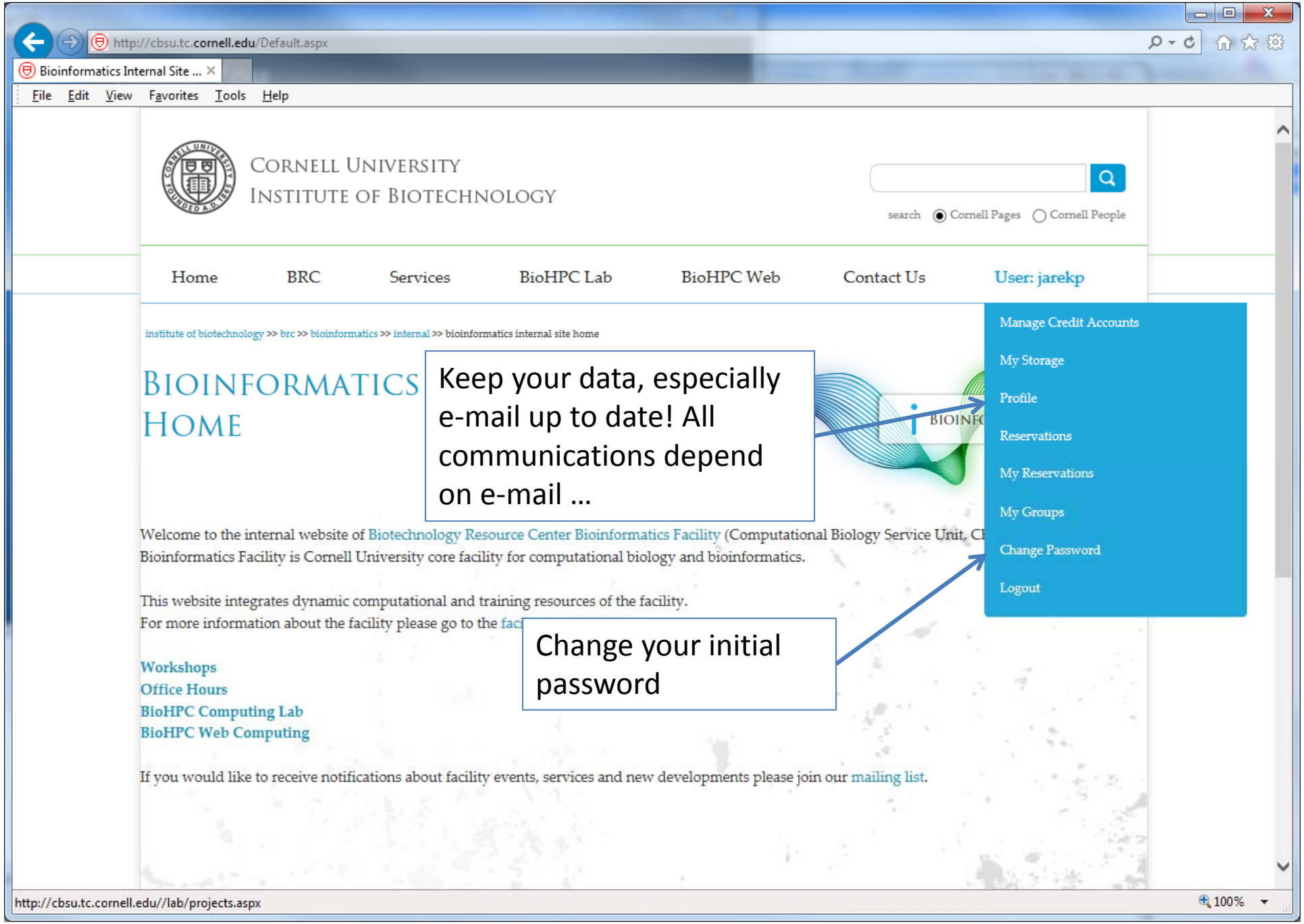
http://cbsu.tc.cornell.edu/lab/lab.aspx



There is lots of information in “User Guide”, including some useful workflows (like “RNA-Seq” workflow)

http://cbsu.tc.cornell.edu/lab/userguide.aspx


http://cbsu.tc.cornell.edu/lab/lab.aspx
BioHPC Lab: Description
File Edit View Favorites Tools Help
CORNELL UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY
search Cornell Pages Cornell People
Home BRC Services BioHPC Lab BioHPC Web Contact Us Login
BioHPC Lab Login
BioHPC Lab Password Reset
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: description
BIOINFORMATICS FACILITY
If you have an account you can login here
Forgot password? Reset it here – it will be sent to your registered e-mail. All you need to remember is your Lab id – usually same as your Cornell Netid.
General
Access
Hardware
Fees
hardware infrastructure
price list
100%


http://cbsu.tc.cornell.edu/Default.aspx
Bioinformatics Internal Site ...
File Edit View Favorites Tools Help
Cornell University
Institute of Biotechnology
search Cornell Pages Cornell People
Home BRC Services BioHPC Lab BioHPC Web Contact Us User: jarekp
Manage Credit Accounts
My Storage
Profile
Reservations
My Reservations
My Groups
Change Password
Logout
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
Bioinformatics Home
Keep your data, especially e-mail up to date! All communications depend on e-mail ...
Change your initial password
Welcome to the internal website of Biotechnology Resource Center Bioinformatics Facility (Computational Biology Service Unit, CI
Bioinformatics Facility is Cornell University core facility for computational biology and bioinformatics.
This website integrates dynamic computational and training resources of the facility.
For more information about the facility please go to the fac
Workshops
Office Hours
BioHPC Computing Lab
BioHPC Web Computing
If you would like to receive notifications about facility events, services and new developments please join our mailing list.
http://cbsu.tc.cornell.edu//lab/projects.aspx
100%

Getting started with a new account

- You need hours: create and fund your own Lab Credit Account or get added to one

- Get extra storage if needed – all users get free storage allocations, but it may not be enough

- Transfer data to your Lab storage

- Make reservation(s)

- Connect to reserved workstations

- *Compute!*

Lab Credit Accounts

In order to reserve workstations you need to have hours available in your Lab Credit Account.

Multiple users can be assigned to the same Lab Credit Account, but only one person (owner) can manage it: add hours, add/remove users, view usage etc.

If your group already has a Lab Credit Account you can ask the owner to be added to it, and get instant access to its hours.

You can set up your own Lab Credit Account and buy hours using Cornell Account or a credit card.

Lab Credit Accounts

There are **3 types of workstations** linked to **3 types of hours**:

- **general**
  cbsum1c1*, cbsum1c2*, cbsuwrkst*

- medium memory (**cbsulm01**)
  cbsulm01, cbsumm*

- large memory (**cbsulm02**)
  cbsulm*

http://cbsu.tc.cornell.edu/Lab/Pricing.aspx

BioHPC Lab hardware infrastructure

interactive workstations with nice consoles ("general"):

4	4-core, 24GB RAM, 4TB HDD
	(cbsuwrkst1,2,3,4 – can be used directly in 625 Rhodes)

"general" remote workstations:

32	8-core, 16GB RAM, 1TB HDD
	(cbsum1c1b0NN, cbsum1c2b0NN)

"medium memory" remote workstations

1	16-core, 64GB RAM, 1TB HDD
16	12-core, 128GB RAM, 4TB HDD, 1TB SSD
	(cbsummNN)

"large memory" remote workstations

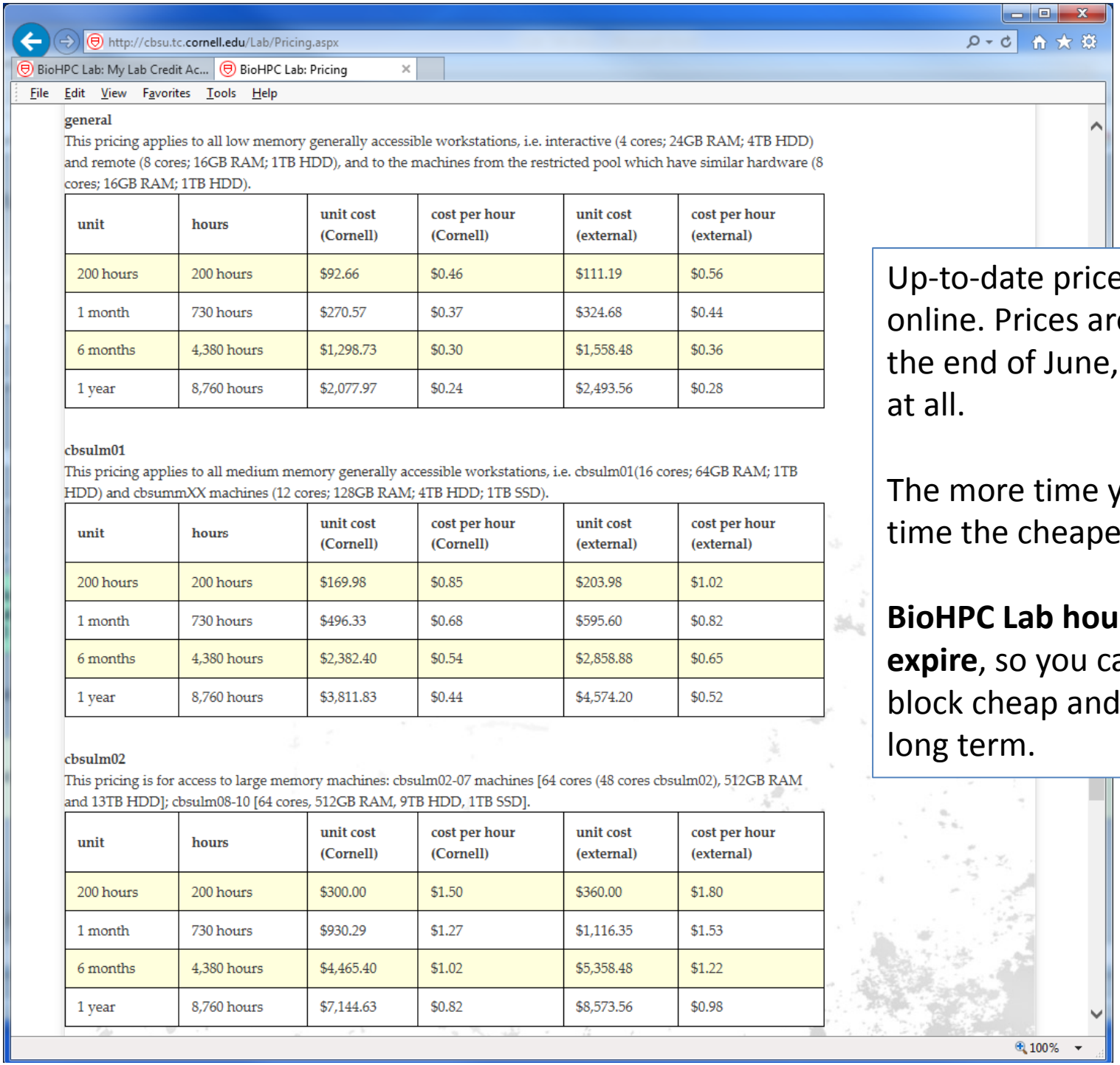
1	48-core, 512GB RAM, 12TB HDD
5	64-core, 512GB RAM, 12TB HDD
3	64-core, 512GB RAM, 9TB HDD, 1TB SSD
	(cbsulmNN)

networked storage: total 256TB available in one volume

2	40TB HDD server
2	88TB HDD server

login machine (cbsulogin)

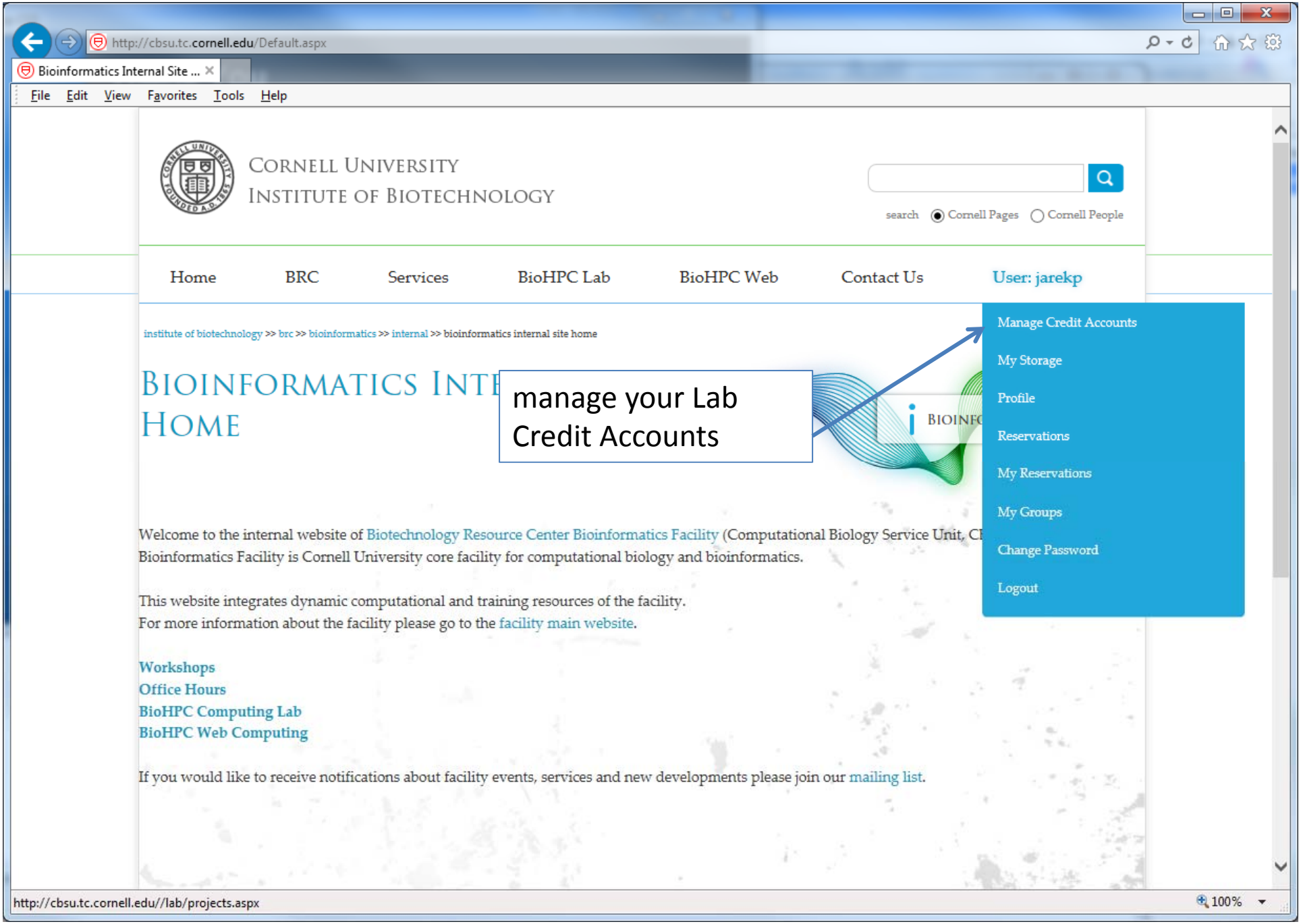
1	12-core, 64GB RAM, 1.5TB HDD

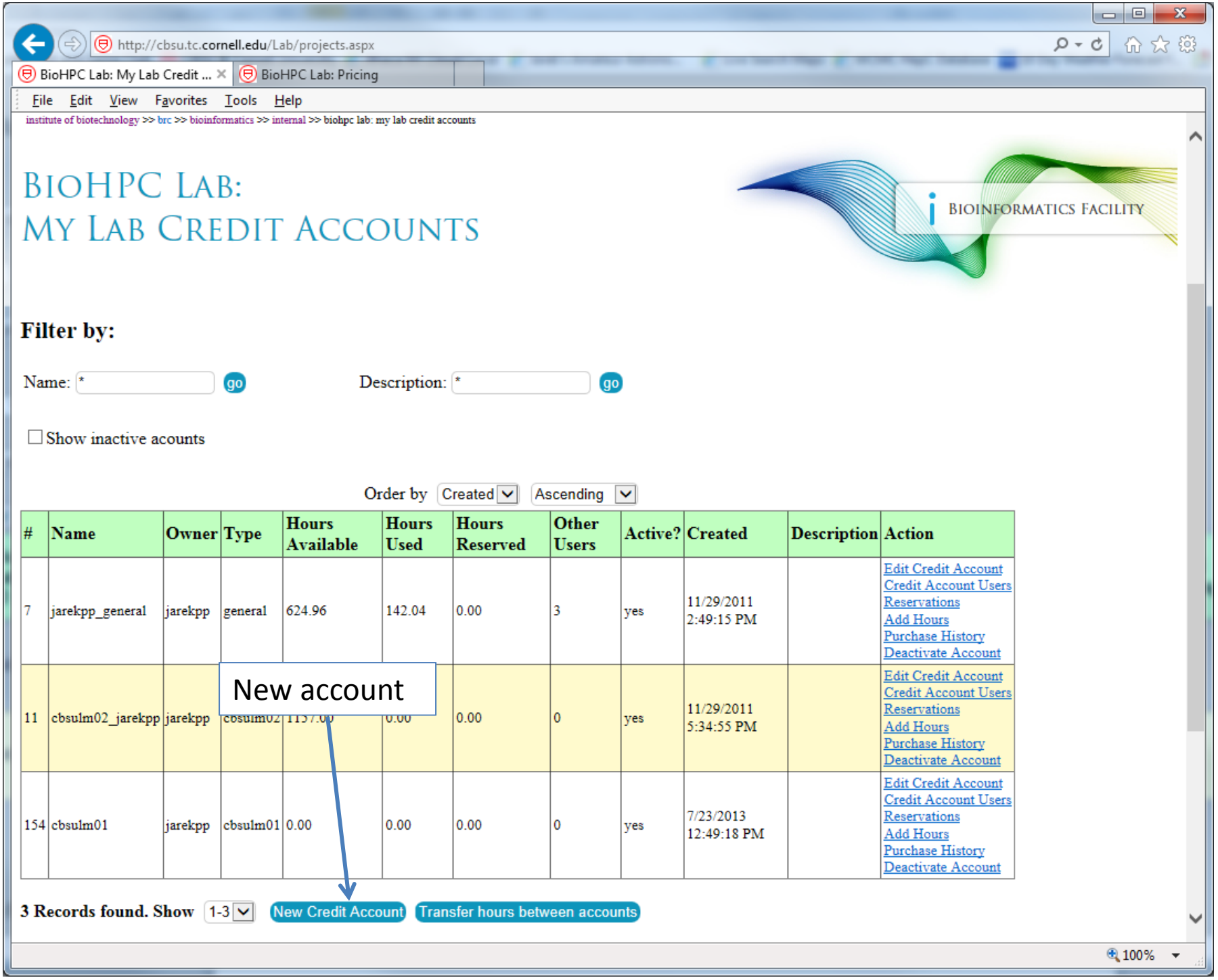


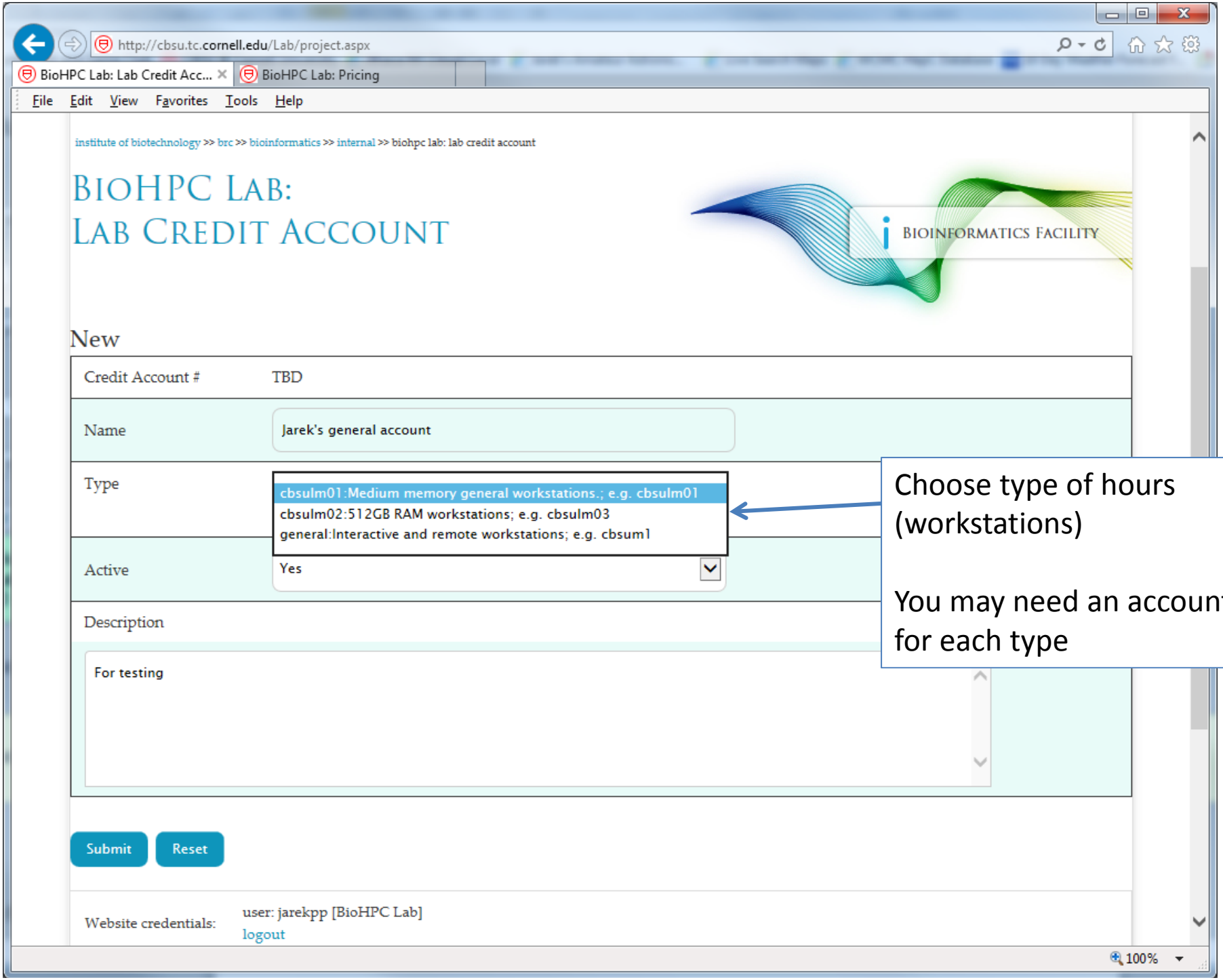
Up-to-date price list is always online. Prices are updated at the end of June, if they change at all.

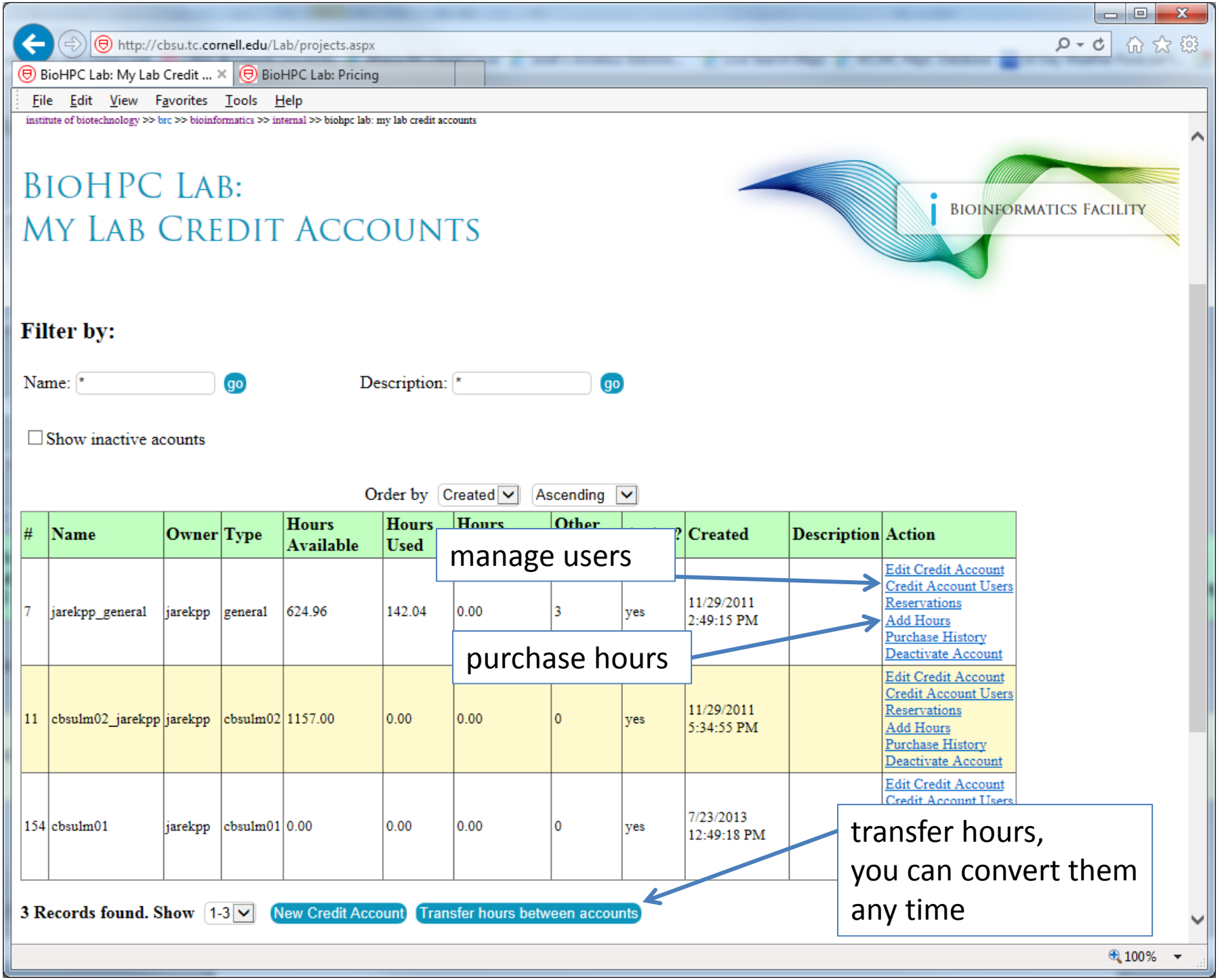
The more time you buy at a time the cheaper it gets.

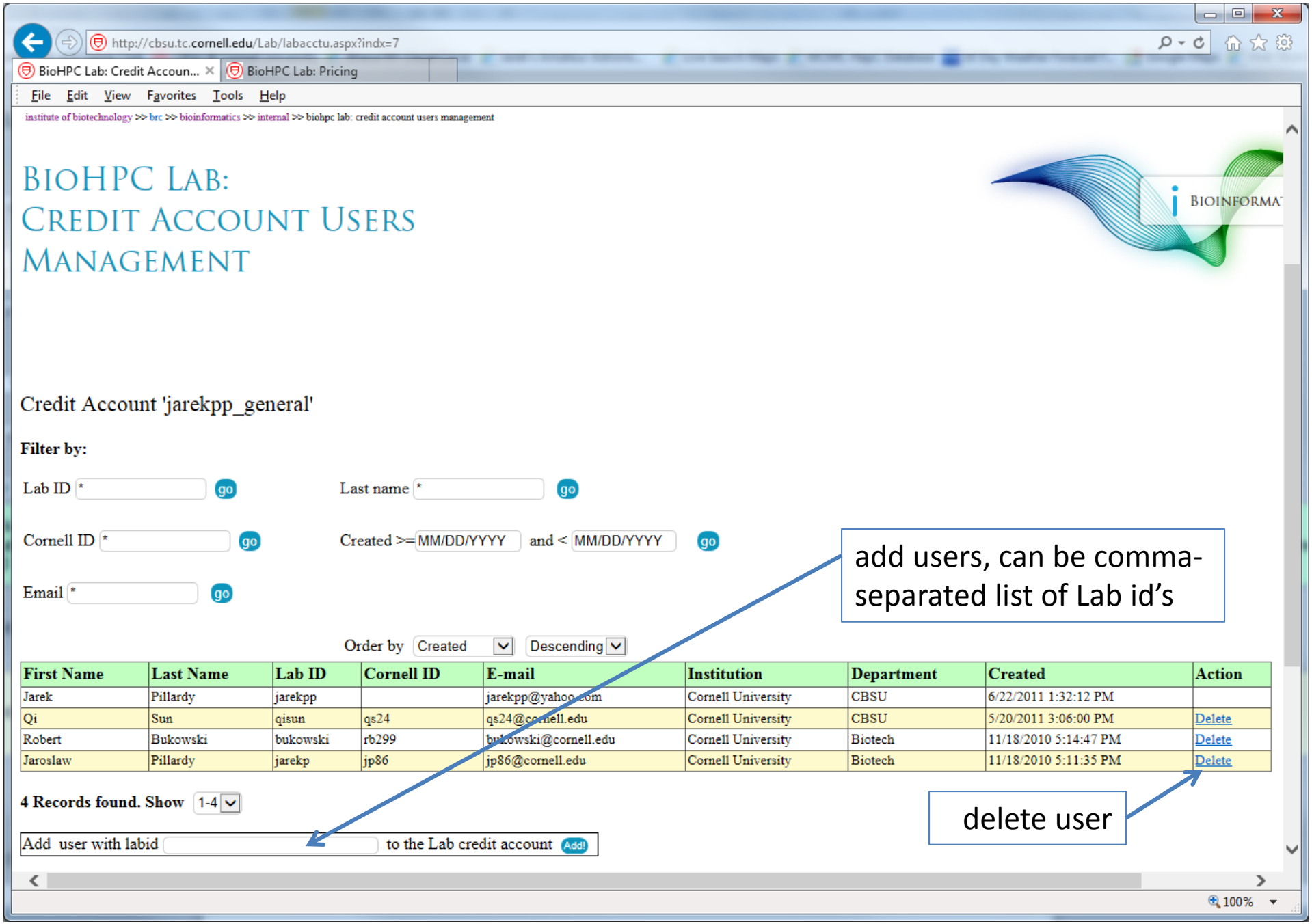
BioHPC Lab hours NEVER expire, so you can buy a large block cheap and use them in a long term.

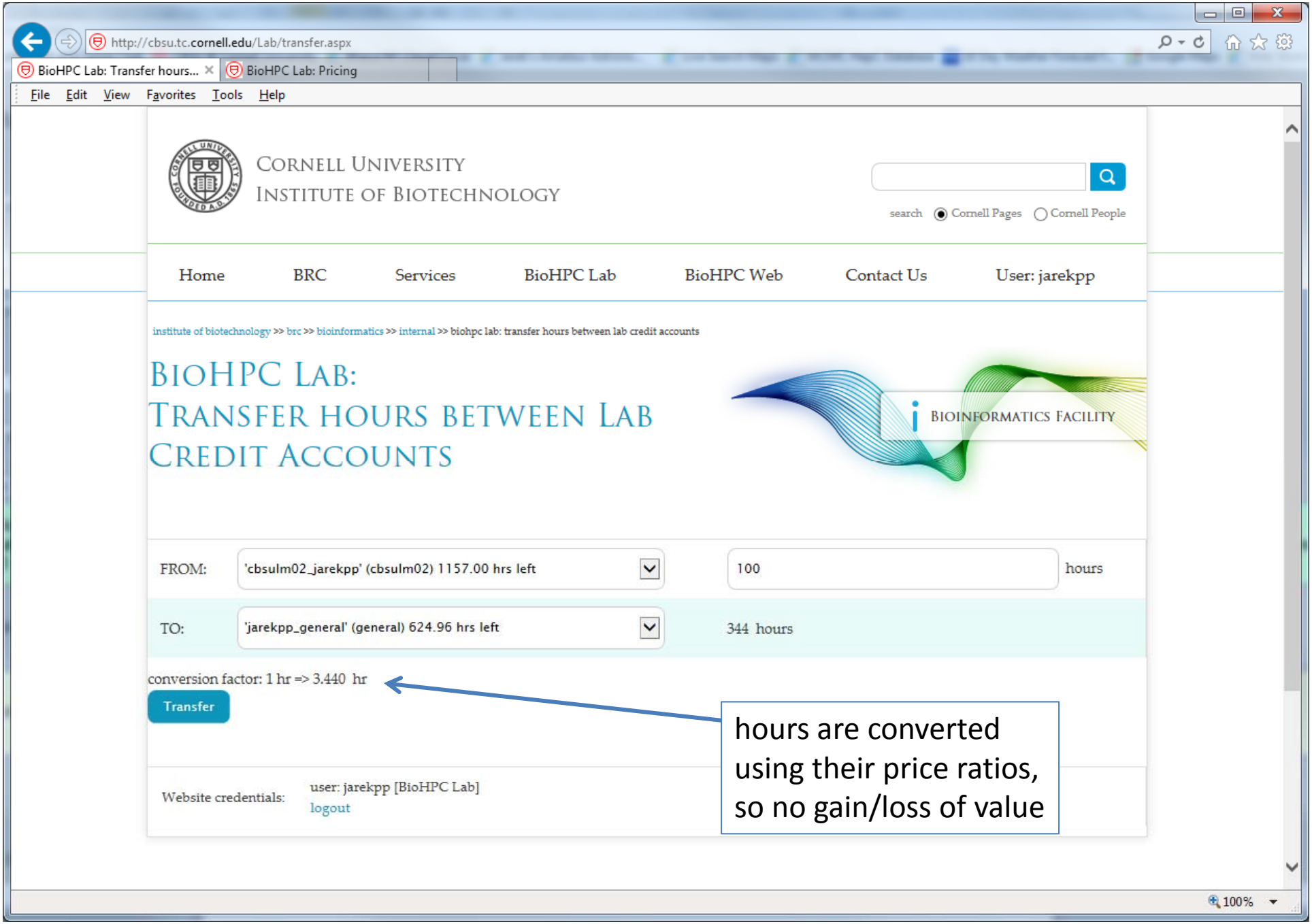

http://cbsu.tc.cornell.edu/Default.aspx
Bioinformatics Internal Site ...
File Edit View Favorites Tools Help
Cornell University
Institute of Biotechnology
search Cornell Pages Cornell People
Home BRC Services BioHPC Lab BioHPC Web Contact Us User: jarekp
Manage Credit Accounts
My Storage
Profile
Reservations
My Reservations
My Groups
Change Password
Logout
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
Bioinformatics Inte
Home
manage your Lab
Credit Accounts
Welcome to the internal website of Biotechnology Resource Center Bioinformatics Facility (Computational Biology Service Unit, C
Bioinformatics Facility is Cornell University core facility for computational biology and bioinformatics.
This website integrates dynamic computational and training resources of the facility.
For more information about the facility please go to the facility main website.
Workshops
Office Hours
BioHPC Computing Lab
BioHPC Web Computing
If you would like to receive notifications about facility events, services and new developments please join our mailing list.
http://cbsu.tc.cornell.edu//lab/projects.aspx
100%


http://cbsu.tc.cornell.edu/Lab/projects.aspx
BioHPC Lab: My Lab Credit ...
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: my lab credit accounts
BioHPC Lab:
My Lab Credit Accounts
BIOINFORMATICS FACILITY
Filter by:
Name: * go
Description: * go
Show inactive acounts
Order by Created Ascending
Name Owner Type Hours Available Hours Used Hours Reserved Other Users Active? Created Description Action
7 jarekpp_general jarekpp general 624.96 142.04 0.00 3 yes 11/29/2011 2:49:15 PM Edit Credit Account Credit Account Users Reservations Add Hours Purchase History Deactivate Account
11 cbsulm02_jarekpp jarekpp 0.00 0 yes 11/29/2011 5:34:55 PM Edit Credit Account Credit Account Users Reservations Add Hours Purchase History Deactivate Account
154 cbsulm01 jarekpp cbsulm01 0.00 0.00 0.00 0 yes 7/23/2013 12:49:18 PM Edit Credit Account Credit Account Users Reservations Add Hours Purchase History Deactivate Account
New account
3 Records found. Show 1-3
New Credit Account
Transfer hours between accounts
100%


http://cbsu.tc.cornell.edu/Lab/project.aspx
BioHPC Lab: Lab Credit Acc...
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: lab credit account
BioHPC Lab:
Lab Credit Account
Bioinformatics Facility
New
Credit Account # TBD
Name Jarek's general account
Type
cbsulm01:Medium memory general workstations.; e.g. cbsulm01
cbsulm02:512GB RAM workstations; e.g. cbsulm03
general:Interactive and remote workstations; e.g. cbsum1
Active Yes
Description
For testing
Submit
Reset
Website credentials: user: jarekpp [BioHPC Lab]
logout
100%
Choose type of hours (workstations)
You may need an account for each type


http://cbsu.tc.cornell.edu/Lab/projects.aspx
BioHPC Lab: My Lab Credit ...
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: my lab credit accounts
BioHPC Lab:
My Lab Credit Accounts
BIOINFORMATICS FACILITY
Filter by:
Name: * go
Description: * go
Show inactive acounts
Order by Created Ascending
Name Owner Type Hours Available Hours Used Hours Other ? Created Description Action
7 jarekpp_general jarekpp general 624.96 142.04 0.00 3 yes 11/29/2011 2:49:15 PM
Edit Credit Account
Credit Account Users
Reservations
Add Hours
Purchase History
Deactivate Account
11 cbsulm02_jarekpp jarekpp cbsulm02 1157.00 0.00 0.00 0 yes 11/29/2011 5:34:55 PM
Edit Credit Account
Credit Account Users
Reservations
Add Hours
Purchase History
Deactivate Account
154 cbsulm01 jarekpp cbsulm01 0.00 0.00 0.00 0 yes 7/23/2013 12:49:18 PM
Edit Credit Account
Credit Account Users
3 Records found. Show 1-3
New Credit Account
Transfer hours between accounts
manage users
purchase hours
transfer hours,
you can convert them
any time
100%


http://cbsu.tc.cornell.edu/Lab/labacctu.aspx?indx=7
BioHPC Lab: Credit Accoun...
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: credit account users management
BioHPC Lab:
Credit Account Users
Management
Credit Account 'jarekpp_general'
Filter by:
Lab ID * go
Last name * go
Cornell ID * go
Created >= MM/DD/YYYY and < MM/DD/YYYY go
Email * go
Order by Created Descending
First Name | Last Name | Lab ID | Cornell ID | E-mail | Institution | Department | Created | Action
Jarek | Pillardy | jarekpp | | jarekpp@yahoo.com | Cornell University | CBSU | 6/22/2011 1:32:12 PM |
Qi | Sun | qisun | qs24 | qs24@cornell.edu | Cornell University | CBSU | 5/20/2011 3:06:00 PM | Delete
Robert | Bukowski | bukowski | rb299 | bukowski@cornell.edu | Cornell University | Biotech | 11/18/2010 5:14:47 PM | Delete
Jaroslaw | Pillardy | jarekp | jp86 | jp86@cornell.edu | Cornell University | Biotech | 11/18/2010 5:11:35 PM | Delete
4 Records found. Show 1-4
Add user with labid to the Lab credit account Add!
add users, can be comma-separated list of Lab id's
delete user
100%


http://cbsu.tc.cornell.edu/Lab/transfer.aspx
BioHPC Lab: Transfer hours...
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
Cornell University
Institute of Biotechnology
search Cornell Pages Cornell People
Home BRC Services BioHPC Lab BioHPC Web Contact Us User: jarekpp
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: transfer hours between lab credit accounts
BioHPC Lab:
Transfer hours between Lab Credit Accounts
Bioinformatics Facility
FROM: 'cbsulm02_jarekpp' (cbsulm02) 1157.00 hrs left
100 hours
TO: 'jarekpp_general' (general) 624.96 hrs left
344 hours
conversion factor: 1 hr => 3.440 hr
Transfer
hours are converted using their price ratios, so no gain/loss of value
Website credentials: user: jarekpp [BioHPC Lab]
logout
100%

Getting started with a new account

- You need hours: create and fund your own Lab Credit Account or get added to one

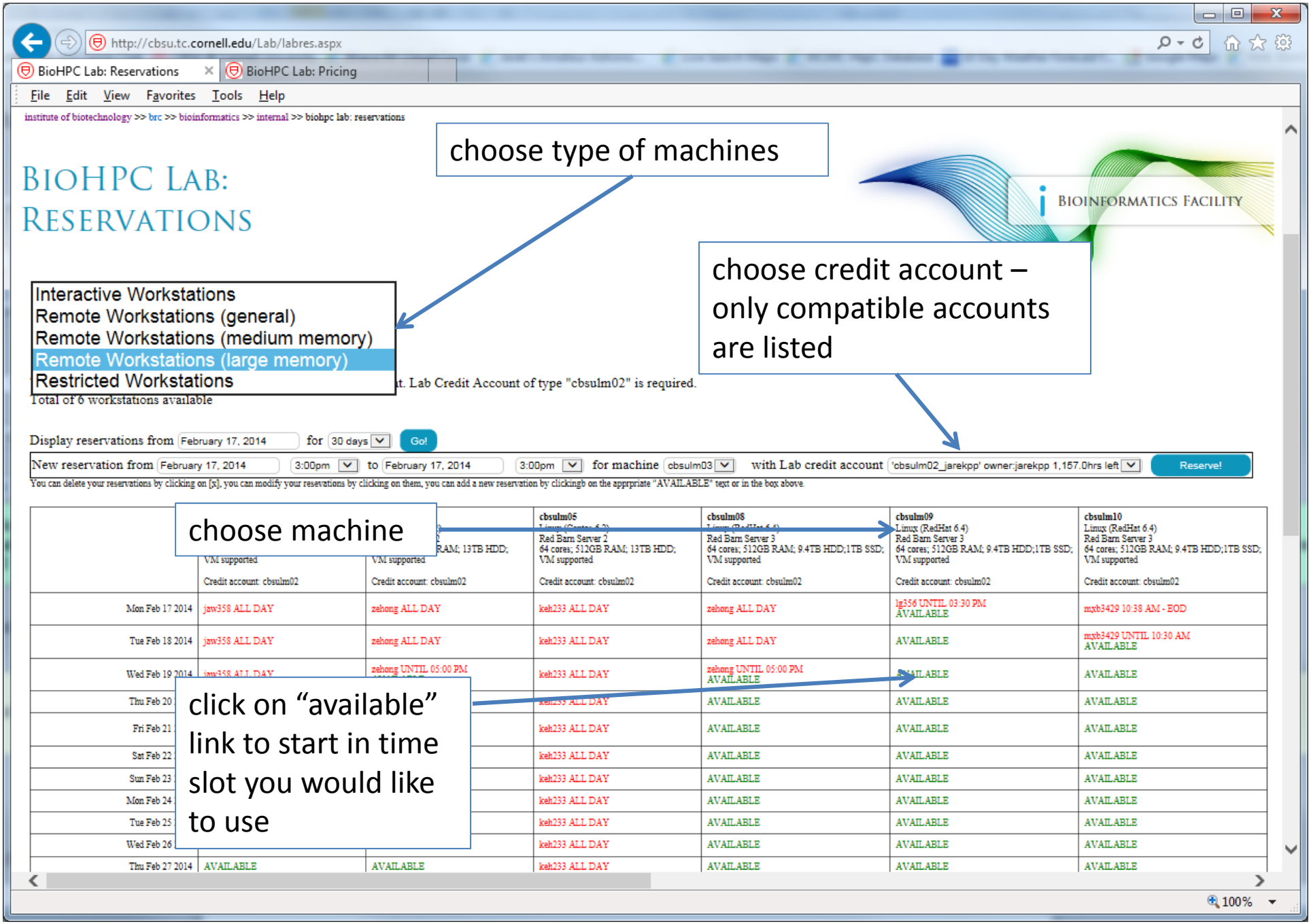
- Get extra storage if needed – all users get free storage allocations, but it may not be enough

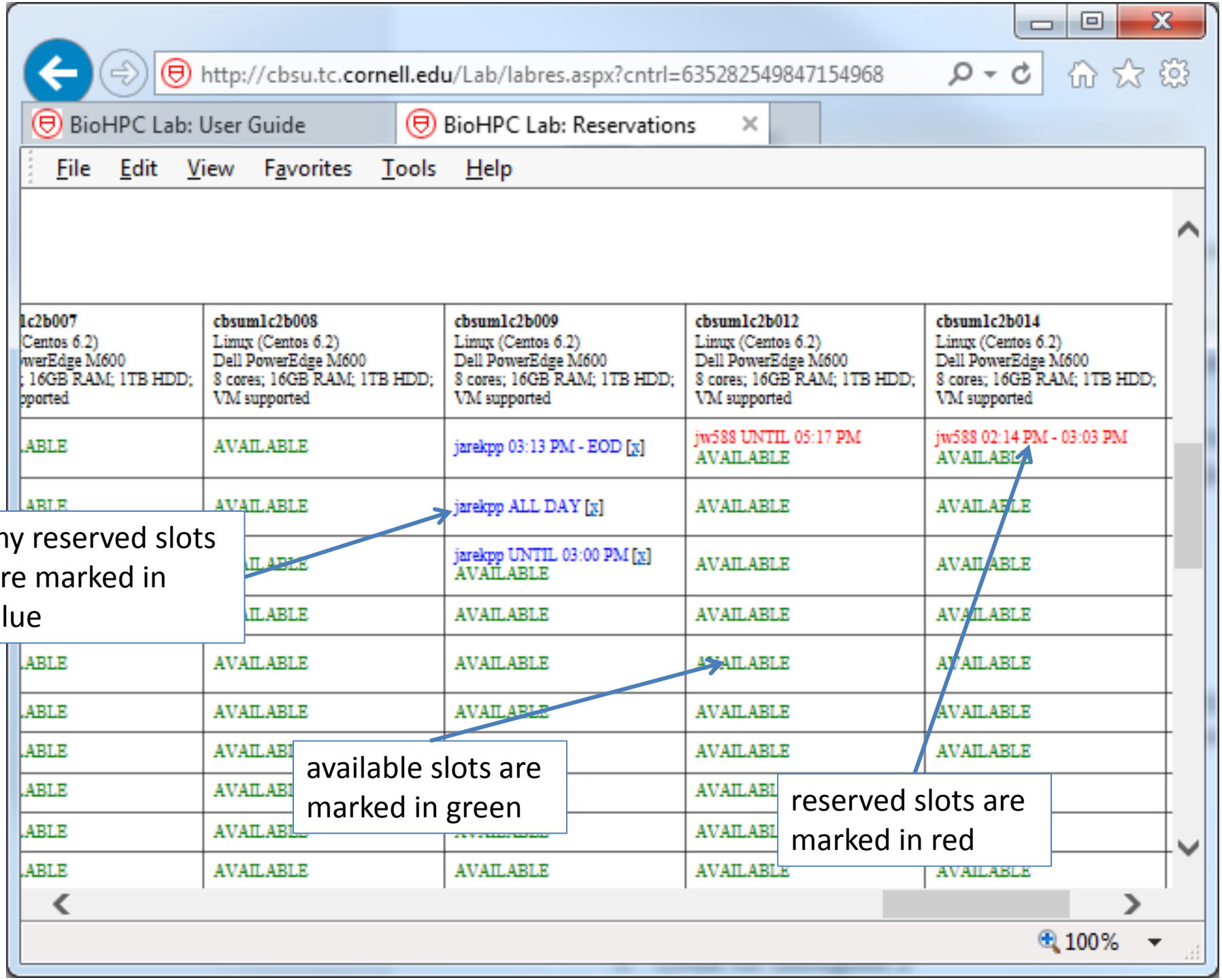
- Transfer data to your Lab storage

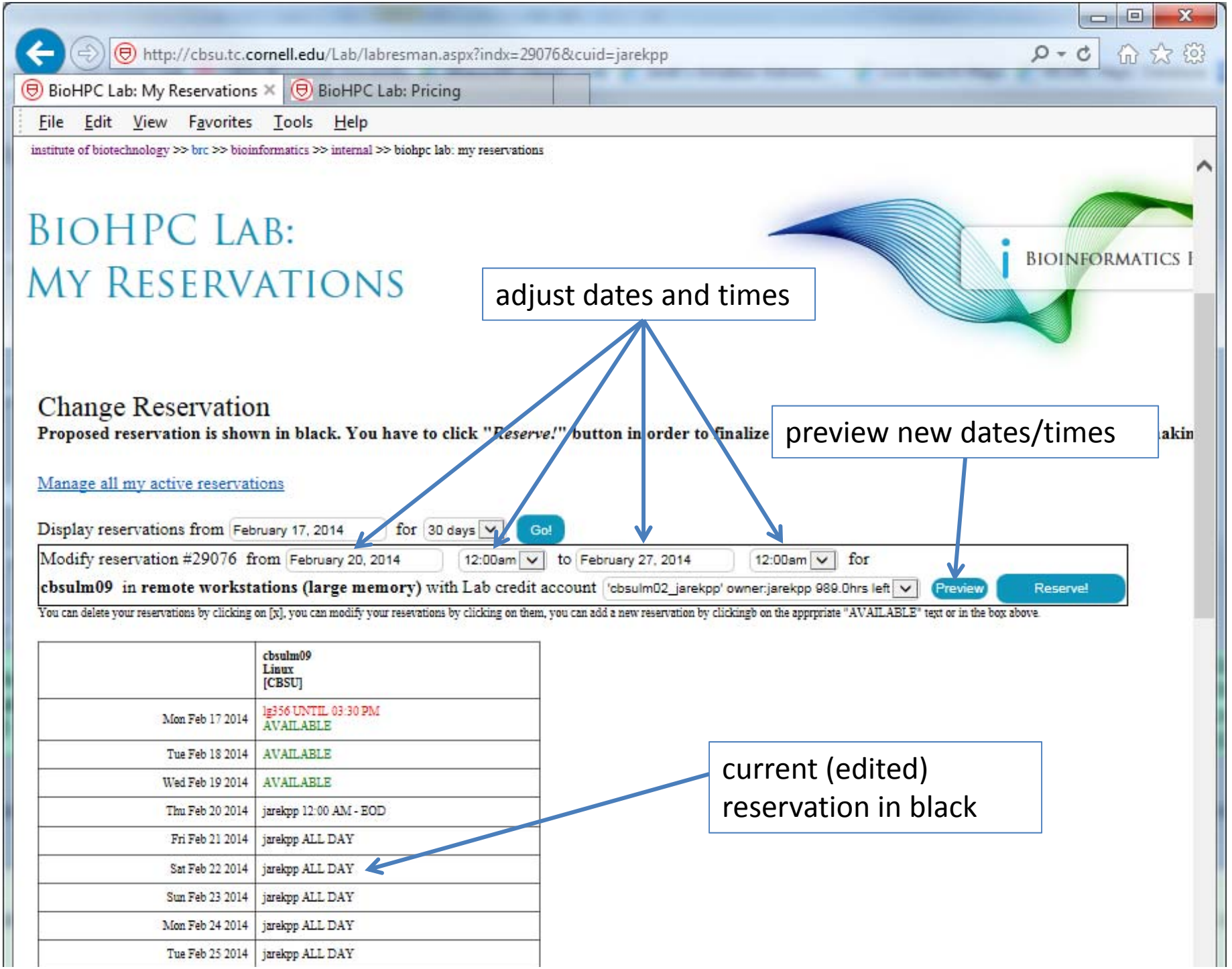
- Make reservation(s)

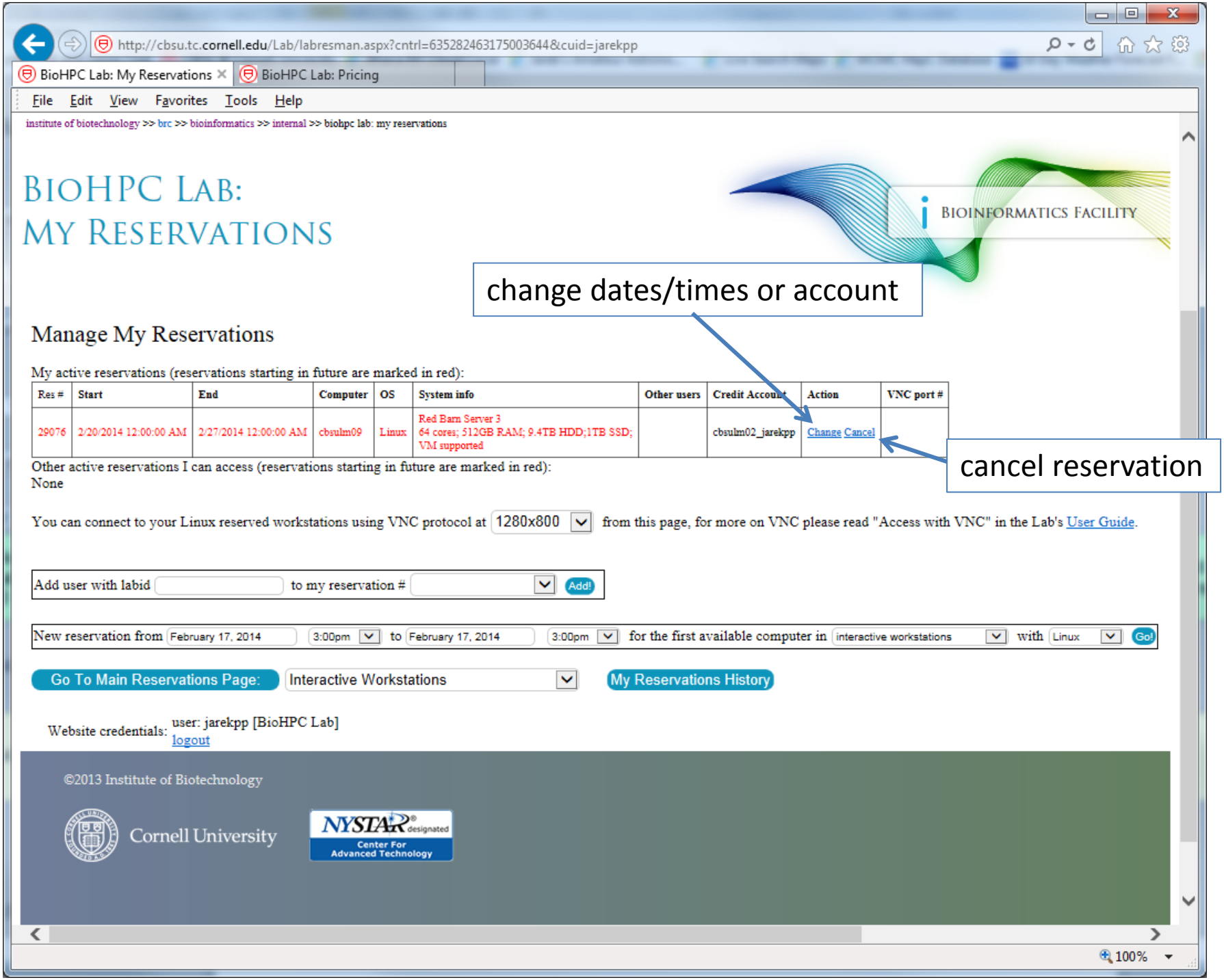
- Connect to reserved workstations

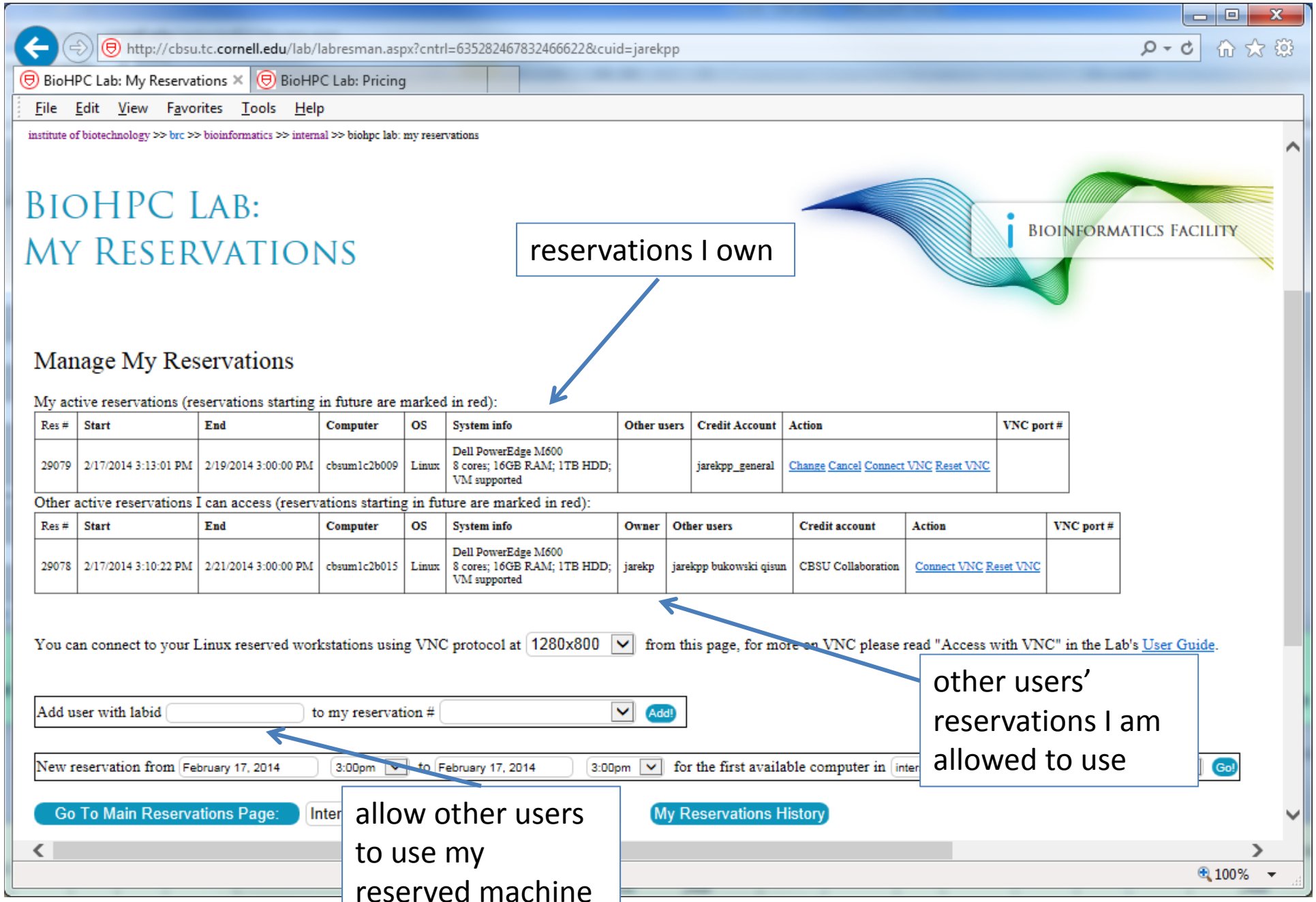
- *Compute!*


http://cbsu.tc.cornell.edu/Lab/labres.aspx
BioHPC Lab: Reservations
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: reservations
BioHPC Lab: Reservations
choose type of machines
Bioinformatics Facility
choose credit account – only compatible accounts are listed
Interactive Workstations
Remote Workstations (general)
Remote Workstations (medium memory)
Remote Workstations (large memory)
Restricted Workstations
t. Lab Credit Account of type "cbsulm02" is required.
Total of 6 workstations available
Display reservations from February 17, 2014 for 30 days Go!
New reservation from February 17, 2014 3:00pm to February 17, 2014 3:00pm for machine cbsulm03 with Lab credit account 'cbsulm02_jarekpp' owner:jarekpp 1,157.0hrs left Reserve!
You can delete your reservations by clicking on [x], you can modify your reservations by clicking on them, you can add a new reservation by clickingb on the apprpriate "AVAILABLE" text or in the box above.
choose machine
cbsulm05 Red Barn Server 2 64 cores; 512GB RAM; 13TB HDD; VM supported Credit account: cbsulm02
cbsulm08 Red Barn Server 3 64 cores; 512GB RAM; 9.4TB HDD;1TB SSD; VM supported Credit account: cbsulm02
cbsulm09 Linux (RedHat 6.4) Red Barn Server 3 64 cores; 512GB RAM; 9.4TB HDD;1TB SSD; VM supported Credit account: cbsulm02
cbsulm10 Linux (RedHat 6.4) Red Barn Server 3 64 cores; 512GB RAM; 9.4TB HDD;1TB SSD; VM supported Credit account: cbsulm02
Mon Feb 17 2014
jaw358 ALL DAY
zehong ALL DAY
keh233 ALL DAY
zehong ALL DAY
lg356 UNTIL 03:30 PM AVAILABLE
mxb3429 10:38 AM - EOD
Tue Feb 18 2014
jaw358 ALL DAY
zehong ALL DAY
keh233 ALL DAY
zehong ALL DAY
AVAILABLE
mxb3429 UNTIL 10:30 AM AVAILABLE
Wed Feb 19 2014
zehong UNTIL 05:00 PM
keh233 ALL DAY
zehong UNTIL 05:00 PM AVAILABLE
AVAILABLE
AVAILABLE
Thu Feb 20
AVAILABLE
AVAILABLE
AVAILABLE
Fri Feb 21
keh233 ALL DAY
AVAILABLE
AVAILABLE
AVAILABLE
Sat Feb 22
keh233 ALL DAY
AVAILABLE
AVAILABLE
AVAILABLE
Sun Feb 23
keh233 ALL DAY
AVAILABLE
AVAILABLE
AVAILABLE
Mon Feb 24
keh233 ALL DAY
AVAILABLE
AVAILABLE
AVAILABLE
Tue Feb 25
keh233 ALL DAY
AVAILABLE
AVAILABLE
AVAILABLE
Wed Feb 26
keh233 ALL DAY
AVAILABLE
AVAILABLE
AVAILABLE
Thu Feb 27 2014
AVAILABLE
AVAILABLE
keh233 ALL DAY
AVAILABLE
AVAILABLE
AVAILABLE
click on “available” link to start in time slot you would like to use
100%


http://cbsu.tc.cornell.edu/Lab/labres.aspx?cntrl=635282549847154968
BioHPC Lab: User Guide
BioHPC Lab: Reservations
File Edit View Favorites Tools Help
lc2b007 (Centos 6.2) werEdge M600 ; 16GB RAM; 1TB HDD; pported
cbsum1c2b008 Linux (Centos 6.2) Dell PowerEdge M600 8 cores; 16GB RAM; 1TB HDD; VM supported
cbsum1c2b009 Linux (Centos 6.2) Dell PowerEdge M600 8 cores; 16GB RAM; 1TB HDD; VM supported
cbsum1c2b012 Linux (Centos 6.2) Dell PowerEdge M600 8 cores; 16GB RAM; 1TB HDD; VM supported
cbsum1c2b014 Linux (Centos 6.2) Dell PowerEdge M600 8 cores; 16GB RAM; 1TB HDD; VM supported
ABLE
AVAILABLE
jarekpp 03:13 PM - EOD [x]
jw588 UNTIL 05:17 PM
jw588 02:14 PM - 03:03 PM
jarekpp ALL DAY [x]
jarekpp UNTIL 03:00 PM [x]
100%
my reserved slots are marked in blue
available slots are marked in green
reserved slots are marked in red


http://cbsu.tc.cornell.edu/Lab/labresman.aspx?indx=29076&cuid=jarekpp
BioHPC Lab: My Reservations
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: my reservations
BioHPC Lab:
My Reservations
BIOINFORMATICS I
adjust dates and times
preview new dates/times
Change Reservation
Proposed reservation is shown in black. You have to click "Reserve!" button in order to finalize
Manage all my active reservations
Display reservations from February 17, 2014 for 30 days Go!
Modify reservation #29076 from February 20, 2014 12:00am to February 27, 2014 12:00am for
cbsulm09 in remote workstations (large memory) with Lab credit account 'cbsulm02_jarekpp' owner:jarekpp 989.0hrs left Preview Reserve!
You can delete your reservations by clicking on [x], you can modify your reservations by clicking on them, you can add a new reservation by clickingb on the apprpriate "AVAILABLE" text or in the box above.
cbsulm09 Linux [CBSU]
Mon Feb 17 2014 lg356 UNTIL 03:30 PM AVAILABLE
Tue Feb 18 2014 AVAILABLE
Wed Feb 19 2014 AVAILABLE
Thu Feb 20 2014 jarekpp 12:00 AM - EOD
Fri Feb 21 2014 jarekpp ALL DAY
Sat Feb 22 2014 jarekpp ALL DAY
Sun Feb 23 2014 jarekpp ALL DAY
Mon Feb 24 2014 jarekpp ALL DAY
Tue Feb 25 2014 jarekpp ALL DAY
current (edited) reservation in black


http://cbsu.tc.cornell.edu/Lab/labresman.aspx?cntrl=635282463175003644&cuid=jarekpp
BioHPC Lab: My Reservations
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: my reservations
BioHPC Lab:
My Reservations
Bioinformatics Facility
change dates/times or account
Manage My Reservations
My active reservations (reservations starting in future are marked in red):
Res # | Start | End | Computer | OS | System info | Other users | Credit Account | Action | VNC port #
29076 | 2/20/2014 12:00:00 AM | 2/27/2014 12:00:00 AM | cbsulm09 | Linux | Red Barn Server 3 64 cores; 512GB RAM; 9.4TB HDD;1TB SSD; VM supported | | cbsulm02_jarekpp | Change Cancel |
cancel reservation
Other active reservations I can access (reservations starting in future are marked in red):
None
You can connect to your Linux reserved workstations using VNC protocol at 1280x800 from this page, for more on VNC please read "Access with VNC" in the Lab's User Guide.
Add user with labid to my reservation # Add!
New reservation from February 17, 2014 3:00pm to February 17, 2014 3:00pm for the first available computer in interactive workstations with Linux Go!
Go To Main Reservations Page: Interactive Workstations
My Reservations History
Website credentials: user: jarekpp [BioHPC Lab]
logout
©2013 Institute of Biotechnology
Cornell University
NYSTAR designated Center For Advanced Technology
100%


http://cbsu.tc.cornell.edu/lab/labresman.aspx?cntrl=635282467832466622&cuid=jarekpp
BioHPC Lab: My Reservations
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: my reservations
BioHPC Lab:
My Reservations
BIOINFORMATICS FACILITY
reservations I own
Manage My Reservations
My active reservations (reservations starting in future are marked in red):
Res # | Start | End | Computer | OS | System info | Other users | Credit Account | Action | VNC port #
29079 | 2/17/2014 3:13:01 PM | 2/19/2014 3:00:00 PM | cbsum1c2b009 | Linux | Dell PowerEdge M600 8 cores; 16GB RAM; 1TB HDD; VM supported | | jarekpp_general | Change Cancel Connect VNC Reset VNC |
Other active reservations I can access (reservations starting in future are marked in red):
Res # | Start | End | Computer | OS | System info | Owner | Other users | Credit account | Action | VNC port #
29078 | 2/17/2014 3:10:22 PM | 2/21/2014 3:00:00 PM | cbsum1c2b015 | Linux | Dell PowerEdge M600 8 cores; 16GB RAM; 1TB HDD; VM supported | jarekp | jarekpp bukowski qisun | CBSU Collaboration | Connect VNC Reset VNC |
You can connect to your Linux reserved workstations using VNC protocol at 1280x800 from this page, for more on VNC please read "Access with VNC" in the Lab's User Guide.
other users' reservations I am allowed to use
Add user with labid to my reservation # Add!
New reservation from February 17, 2014 3:00pm to February 17, 2014 3:00pm for the first available computer in Go!
allow other users to use my reserved machine
Go To Main Reservations Page:
My Reservations History
100%

Reservations

You can only make reservation if you have enough hours of appropriate type.

Your hours will be “reserved” when you make the reservation so you cannot use them twice – but not immediately deducted.

Only actually used hours are charged, if you cancel your reservation before it starts nothing is charged, if you cancel before it ends only the actual hours between start and cancel are charged.

NOTE: “Time used" is defined as the wall clock time elapsed when your reservation is active - it has nothing to do with how much you actually used the CPU/RAM during this time. "Time used" reflects the span for which you had the workstation reserved.

Reservations

What happens when a reservation ends, but you are still working?

- You will not be able to login to the machine anymore.
- If you are logged in you will stay logged in until the user from the next reservation logs in.
- Your programs will continue to run, as long as the machine is not used – this is to give you time to extend reservation if you need it.
- When a person that has the next reservation (now current) logs in all your programs and processes will be killed and you will be logged out.

Getting started with a new account

- ✖ You need hours: create and fund your own Lab Credit Account or get added to one

- Get extra storage if needed – all users get free storage allocations, but it may not be enough

- Transfer data to your Lab storage

- ✖ Make reservation(s)

- Connect to reserved workstations

- *Compute!*

Storage

- **Networked storage**

  Very large storage (256TB), access limited by network speed, much slower than local storage. *Unsuitable for direct computing*. Very good for storing data long-term or sharing data between workstations – networked storage is the same on all machines.

- **Local storage**

  Fast storage, especially on large memory and medium memory workstations. Small – between 1TB (general) and 12TB (large memory). *Designed to be used for computations*.

Storage

Linux directory structure is *continuous*, i.e. regardless of the physical location of storage it all seems to be part of one directory tree starting from root (/).

Not easy to tell which storage is local and which global just by a name.

/home/jarekp

/usr/local

/workdir/jarekp

/shared_data

/local_data

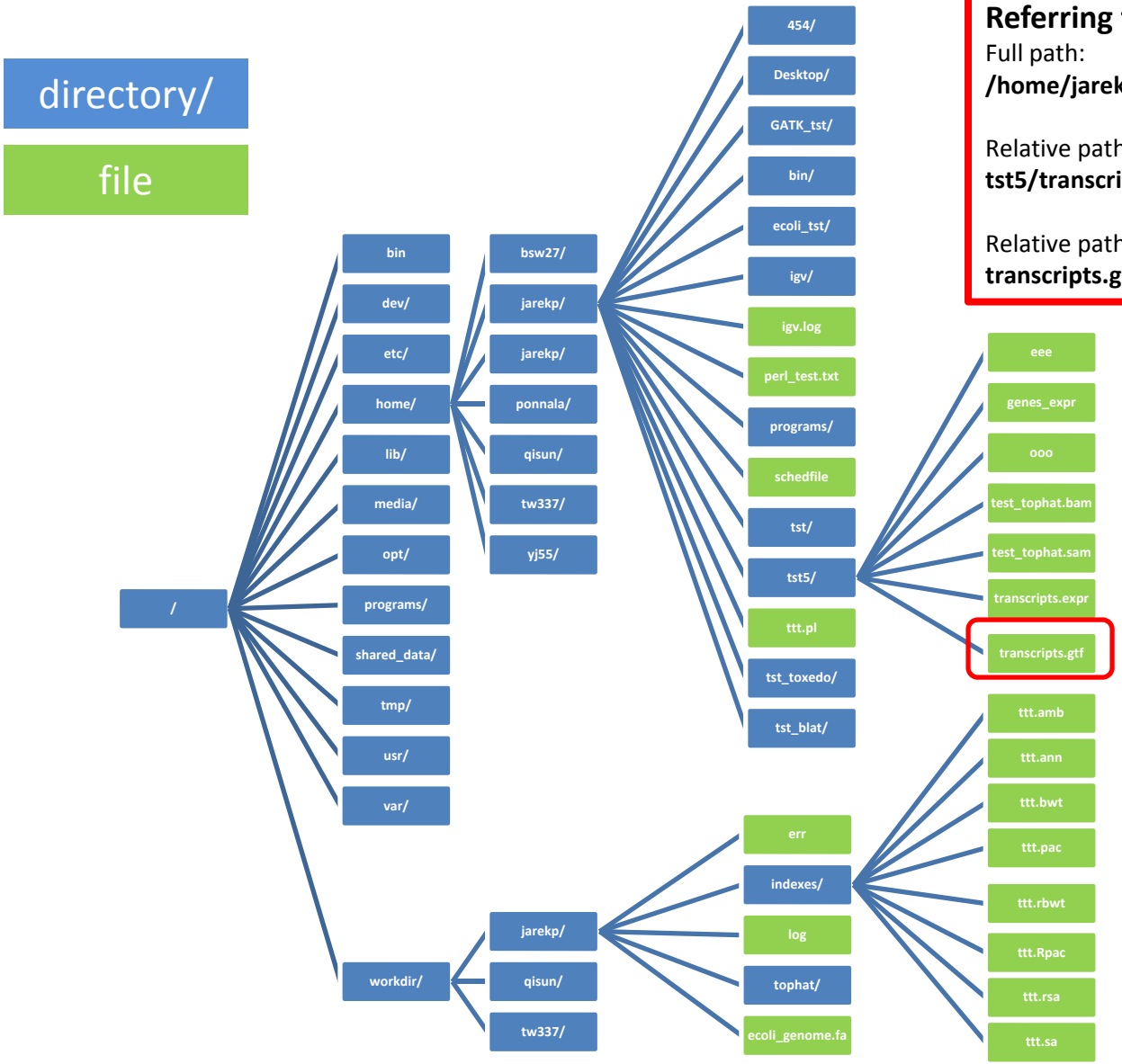
Storage

Linux directory structure is *continuous*, i.e. regardless of the physical location of storage it all seems to be part of one directory tree starting from root (/).

Not easy to tell which storage is local and which global just by a name.

/home/jarekp	networked
/usr/local	local
/workdir/jarekp	local
/shared_data	networked
/local_data	local

Example of directory structure



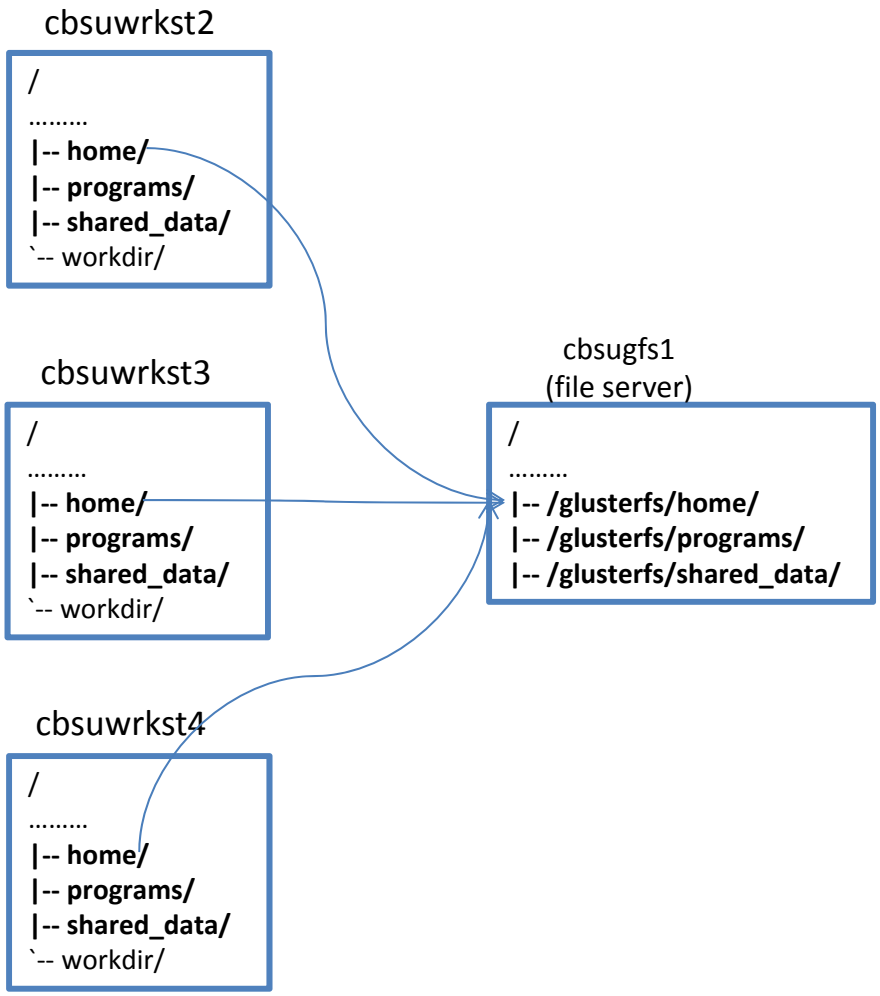
Referring to files:

Full path:
/home/jarekp/tst5/transcripts.gtf

Relative path (i.e., relative to /home/jarekp)
tst5/transcripts.gtf

Relative path (i.e., relative to /home/jarekp/tst5)
transcripts.gtf

Local vs. network directories



Network directories

/home, /programs, /shared_data (with all subdirectories)

- Physically located on the file server
- Visible from all workstations
- SLOW access – DO NOT run any calculations there

Local directories

/workdir (with all subdirectories), **/local_data**

/SSD (if present)

- Physically attached to “its own” workstation
- Not visible from other workstations
- Fast access – all calculations should be run in **/workdir** or **/SSD**

Home directories

Each Unix (Linux) user has a personal storage space called home directory usually referred as /home/userid.

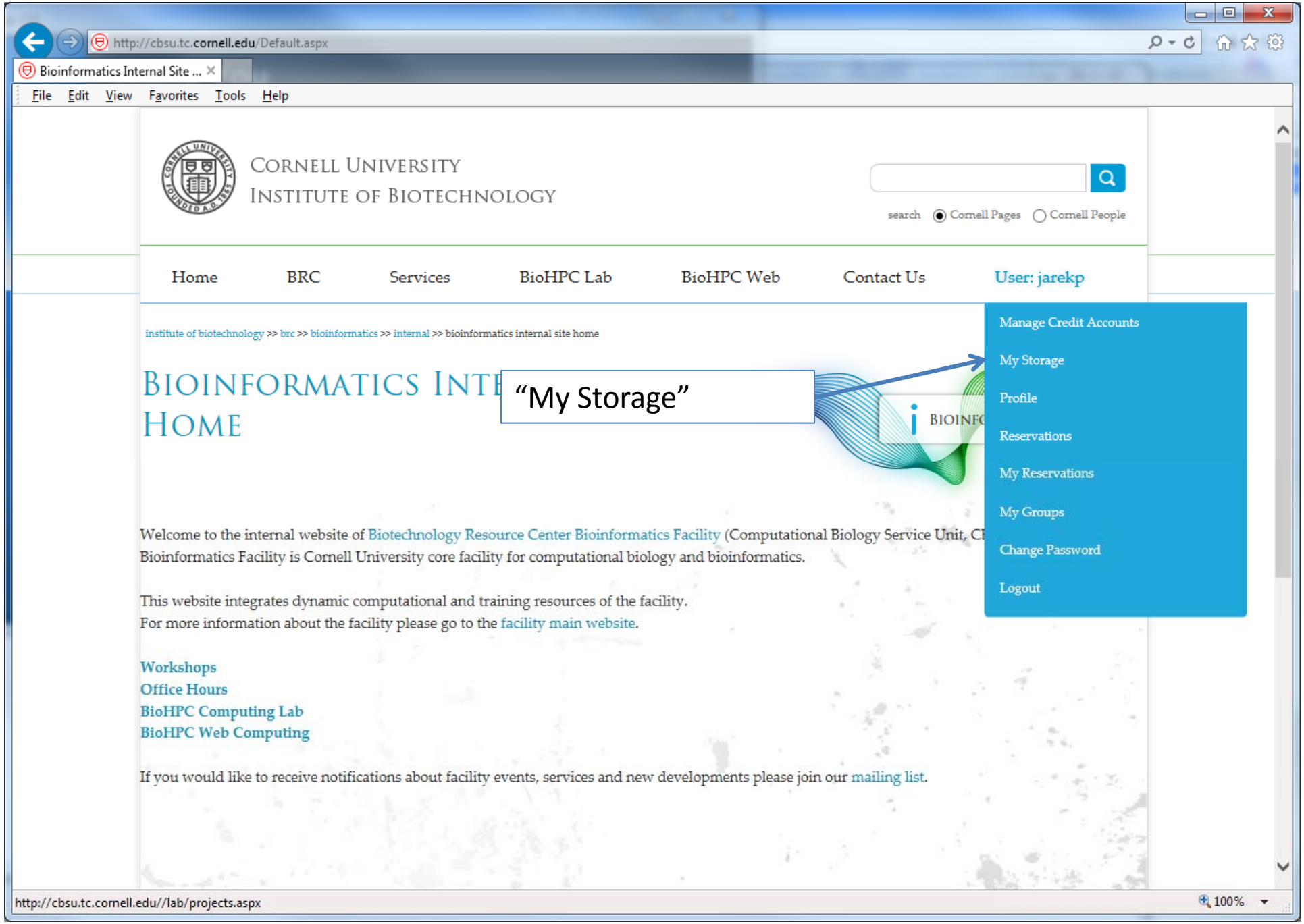
All home directories on BioHPC Lab are networked.

DO NOT RUN ANY COMPUTATIONS IN YOUR HOME DIRECTORY!

Copy your files to /workdir/mylabid first and run computations there!

Space available for each user in home directory is limited by a quota, which depends on type of the user and his resources.

You can always see your current limits and storage under “My Storage” menu. The storage info is updated daily at 5 am.

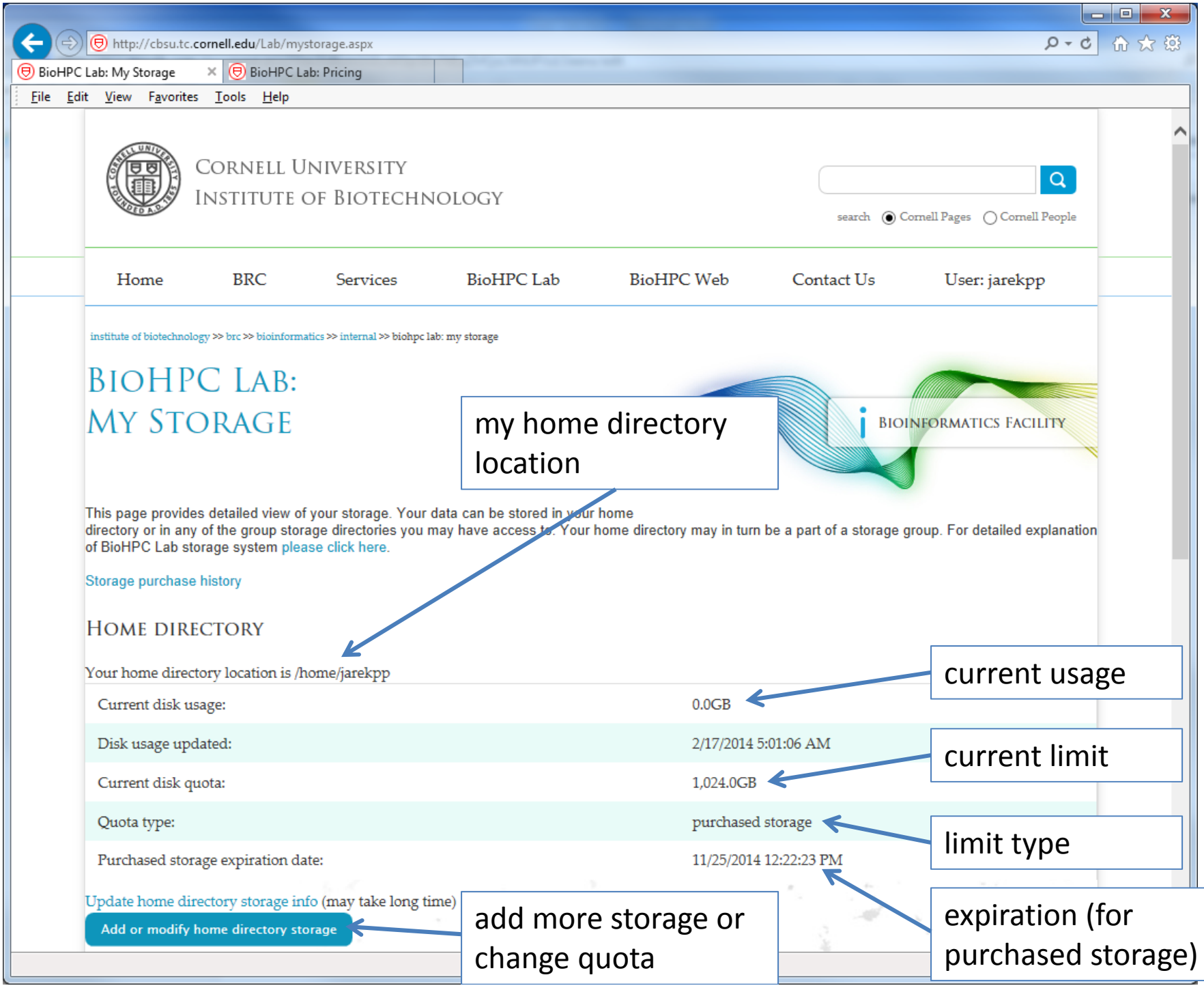

http://cbsu.tc.cornell.edu/Default.aspx
Bioinformatics Internal Site ...
File Edit View Favorites Tools Help
Cornell University
Institute of Biotechnology
search Cornell Pages Cornell People
Home BRC Services BioHPC Lab BioHPC Web Contact Us User: jarekp
Manage Credit Accounts
My Storage
Profile
Reservations
My Reservations
My Groups
Change Password
Logout
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
Bioinformatics Inte
Home
"My Storage"
Welcome to the internal website of Biotechnology Resource Center Bioinformatics Facility (Computational Biology Service Unit, C
Bioinformatics Facility is Cornell University core facility for computational biology and bioinformatics.
This website integrates dynamic computational and training resources of the facility.
For more information about the facility please go to the facility main website.
Workshops
Office Hours
BioHPC Computing Lab
BioHPC Web Computing
If you would like to receive notifications about facility events, services and new developments please join our mailing list.
http://cbsu.tc.cornell.edu//lab/projects.aspx
100%

Home storage quotas

For user DOES NOT have access to paid storage

- User is associated with an active Lab Credit Account. Home directory storage limit is 200 GB.
- User is associated with an active hosted hardware resource. Home directory storage limit is 200 GB.
- User is NOT associated with an active Lab Credit Account or hosted hardware. Home directory storage limit is 20 GB.

Free storage quotas cannot be combined, added to purchased storage or used for multiple accounts. They are just to make sure users can carry out common computations without purchasing extra storage.


http://cbsu.tc.cornell.edu/Lab/mystorage.aspx
BioHPC Lab: My Storage
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
Cornell University
Institute of Biotechnology
search Cornell Pages Cornell People
Home BRC Services BioHPC Lab BioHPC Web Contact Us User: jarekpp
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc lab: my storage
BioHPC Lab:
My Storage
Bioinformatics Facility
This page provides detailed view of your storage. Your data can be stored in your home directory or in any of the group storage directories you may have access to. Your home directory may in turn be a part of a storage group. For detailed explanation of BioHPC Lab storage system please click here.
Storage purchase history
Home Directory
Your home directory location is /home/jarekpp
Current disk usage: 0.0GB
Disk usage updated: 2/17/2014 5:01:06 AM
Current disk quota: 1,024.0GB
Quota type: purchased storage
Purchased storage expiration date: 11/25/2014 12:22:23 PM
Update home directory storage info (may take long time)
Add or modify home directory storage
my home directory location
current usage
current limit
limit type
expiration (for purchased storage)
add more storage or change quota

Storage

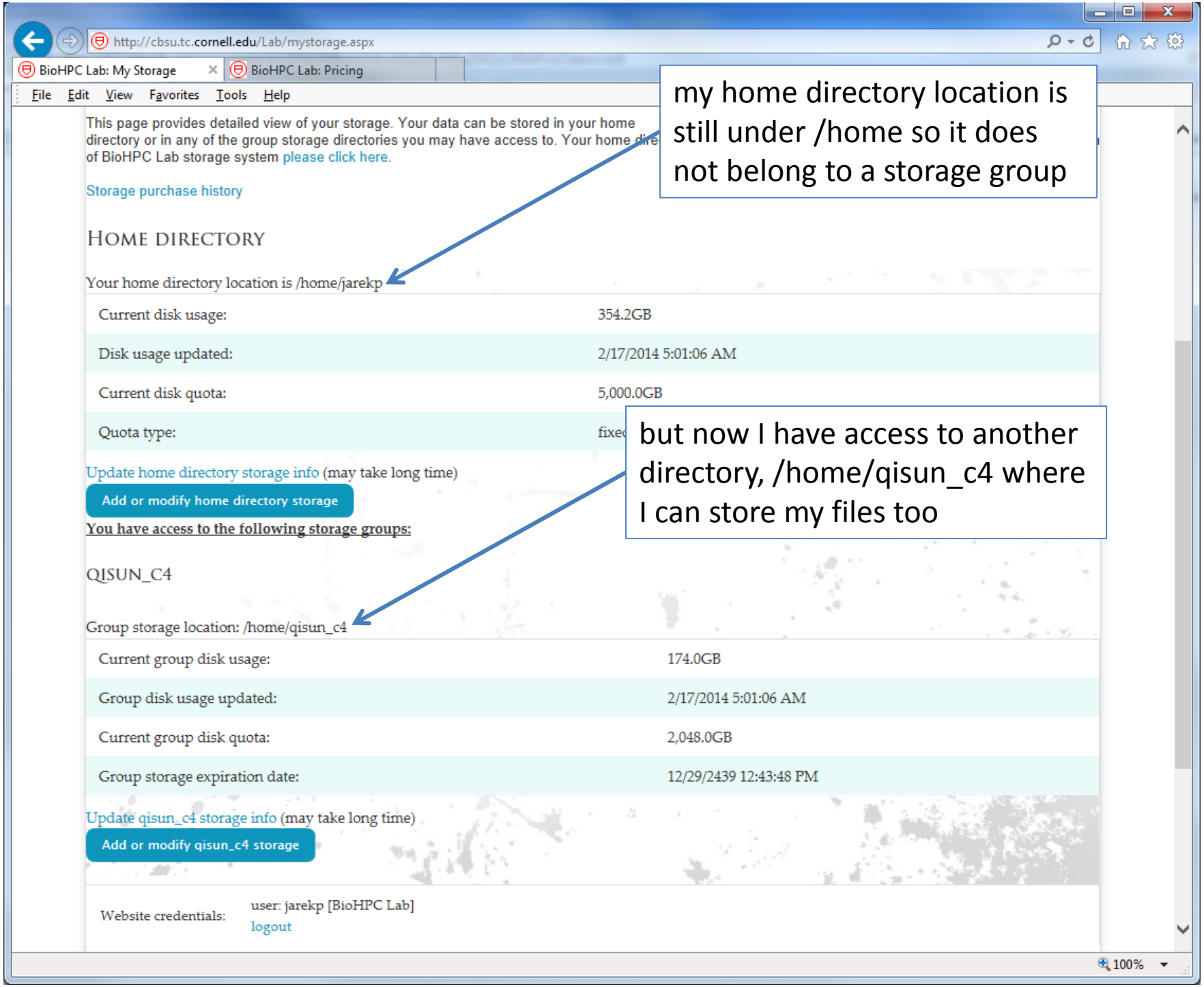
Extra storage can be purchased for $80 per TB per year.

Extra storage can be applied to user's home directory.

Extra storage can be shared among group of users in a **storage group**

Users can have their home directories placed in a storage group, their combined storage limit is then equal to the limit of the storage group.

Please contact us to create a storage group, once created it can be managed online under "My Groups" and "My Storage"


http://cbsu.tc.cornell.edu/Lab/mystorage.aspx
BioHPC Lab: My Storage
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
This page provides detailed view of your storage. Your data can be stored in your home directory or in any of the group storage directories you may have access to. Your home dire
of BioHPC Lab storage system please click here.
Storage purchase history
HOME DIRECTORY
Your home directory location is /home/jarekp
Current disk usage: 354.2GB
Disk usage updated: 2/17/2014 5:01:06 AM
Current disk quota: 5,000.0GB
Quota type: fixe
Update home directory storage info (may take long time)
Add or modify home directory storage
You have access to the following storage groups:
QISUN_C4
Group storage location: /home/qisun_c4
Current group disk usage: 174.0GB
Group disk usage updated: 2/17/2014 5:01:06 AM
Current group disk quota: 2,048.0GB
Group storage expiration date: 12/29/2439 12:43:48 PM
Update qisun_c4 storage info (may take long time)
Add or modify qisun_c4 storage
Website credentials: user: jarekp [BioHPC Lab]
logout
100%
my home directory location is still under /home so it does not belong to a storage group
but now I have access to another directory, /home/qisun_c4 where I can store my files too

Storage

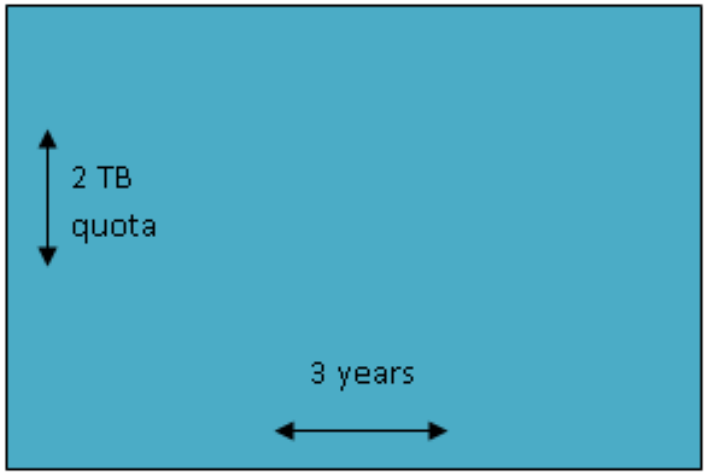
- The storage can be only purchased in 1 TB-year chunks, it needs to be done up front, and you can set your quota to an appropriate size, which in turn will decide the expiration date.

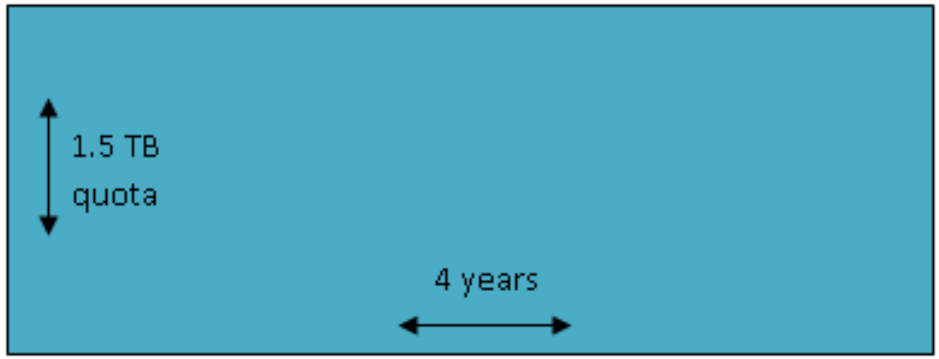
- You can buy as many of the 1TB-year chunks as you want and then set the quota at the level you want, the expiration date will be computed as the result.

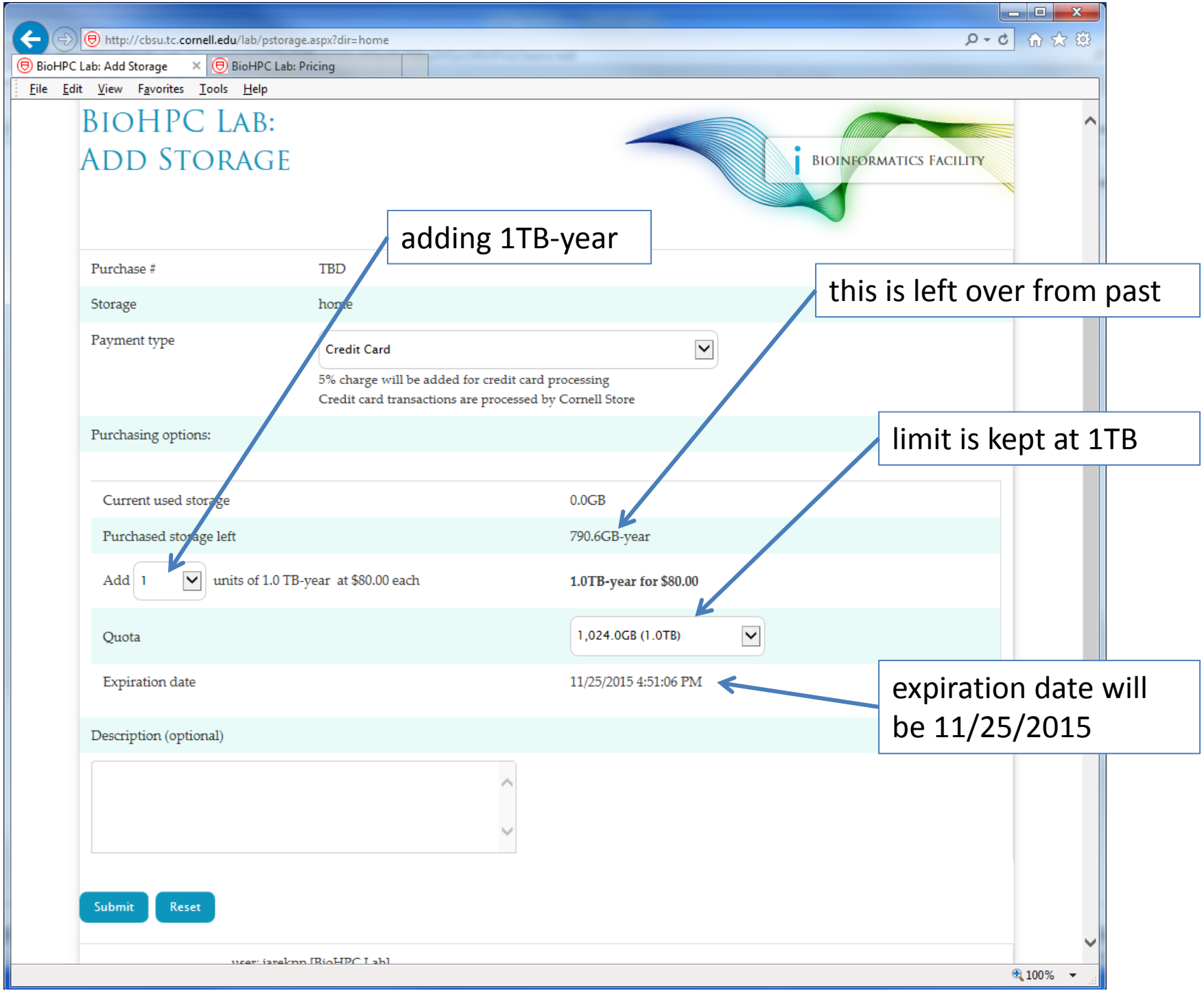
Storage: TB-years

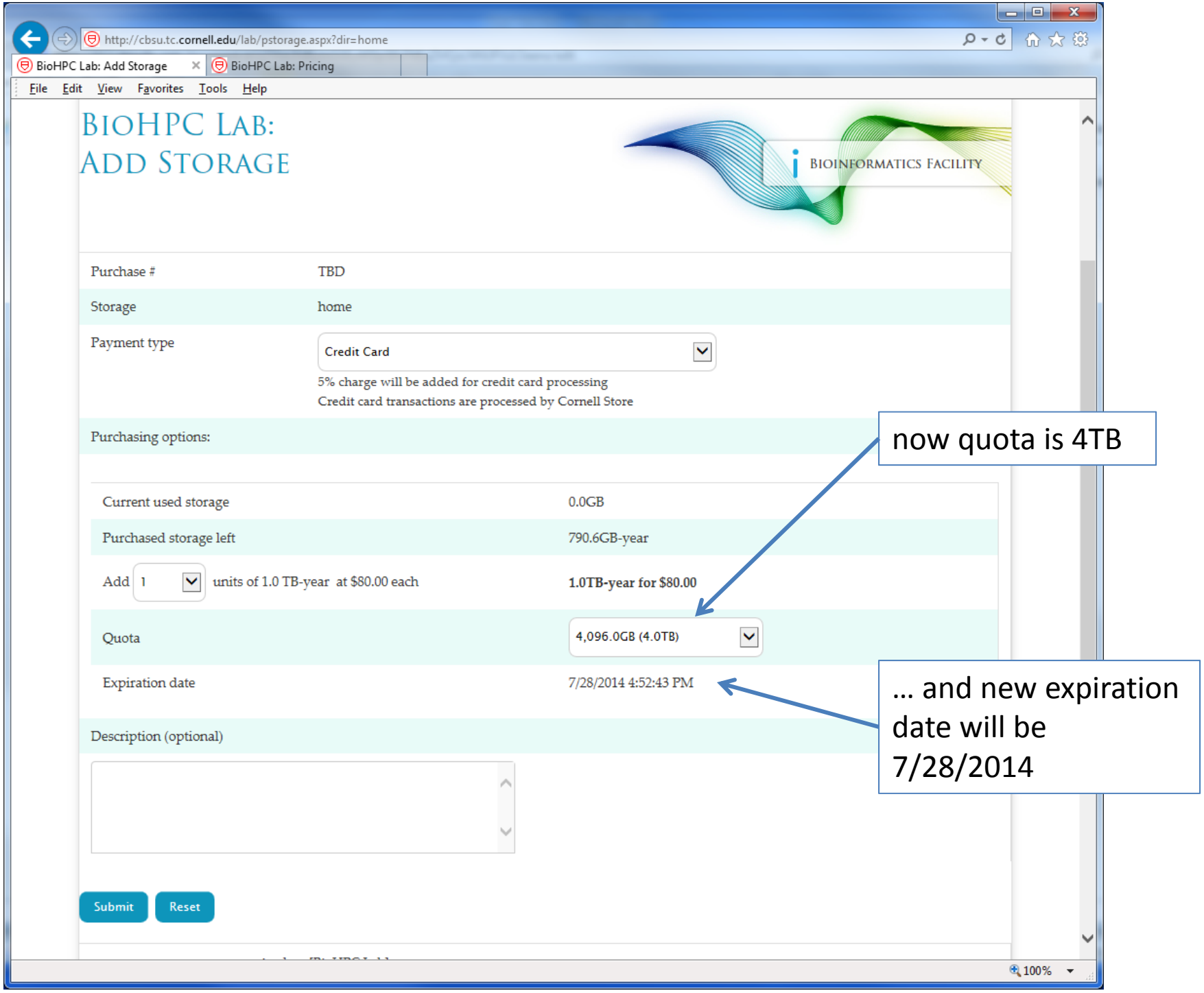
Storage in TB-years represents an *area* and it is always preserved, but either of the rectangle sides can be changed resulting in the other one adapting: lower quota extends expiration time, higher quota shortens time span.

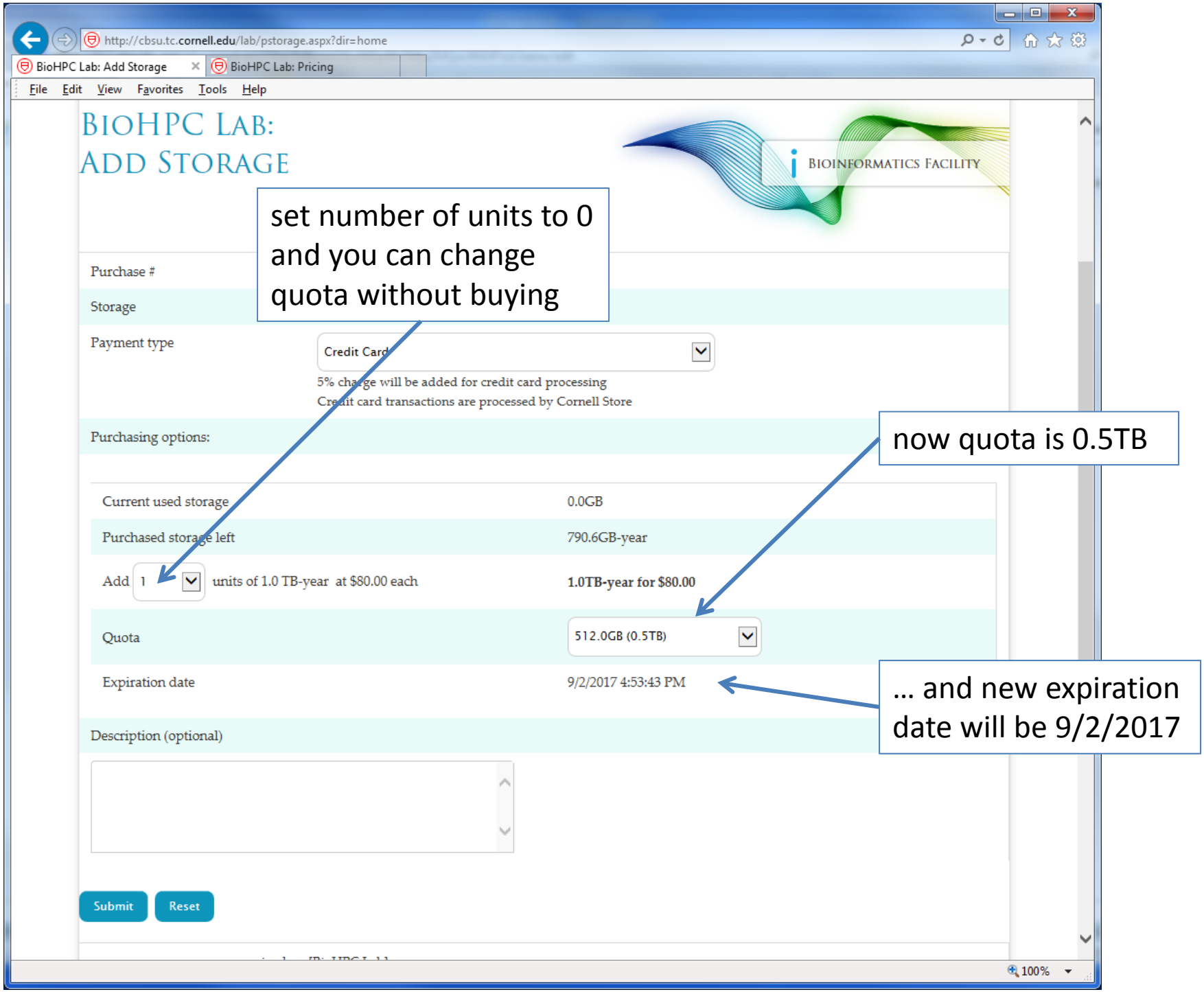
Both rectangles represent the same purchase: 6 TB-years.






http://cbsu.tc.cornell.edu/lab/pstorage.aspx?dir=home
BioHPC Lab: Add Storage
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
BIOHPC LAB:
ADD STORAGE
BIOINFORMATICS FACILITY
adding 1TB-year
this is left over from past
limit is kept at 1TB
expiration date will be 11/25/2015
Purchase # TBD
Storage home
Payment type
Credit Card
5% charge will be added for credit card processing
Credit card transactions are processed by Cornell Store
Purchasing options:
Current used storage 0.0GB
Purchased storage left 790.6GB-year
Add 1 units of 1.0 TB-year at $80.00 each
1.0TB-year for $80.00
Quota
1,024.0GB (1.0TB)
Expiration date 11/25/2015 4:51:06 PM
Description (optional)
Submit
Reset
100%


http://cbsu.tc.cornell.edu/lab/pstorage.aspx?dir=home
BioHPC Lab: Add Storage
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
BioHPC Lab: Add Storage
BIOINFORMATICS FACILITY
Purchase # TBD
Storage home
Payment type
Credit Card
5% charge will be added for credit card processing
Credit card transactions are processed by Cornell Store
Purchasing options:
Current used storage 0.0GB
Purchased storage left 790.6GB-year
Add 1 units of 1.0 TB-year at $80.00 each
1.0TB-year for $80.00
Quota
4,096.0GB (4.0TB)
Expiration date 7/28/2014 4:52:43 PM
Description (optional)
Submit
Reset
100%
now quota is 4TB
… and new expiration date will be 7/28/2014


http://cbsu.tc.cornell.edu/lab/pstorage.aspx?dir=home
BioHPC Lab: Add Storage
BioHPC Lab: Pricing
File Edit View Favorites Tools Help
BioHPC Lab: Add Storage
Bioinformatics Facility
set number of units to 0 and you can change quota without buying
Purchase #
Storage
Payment type
Credit Card
5% charge will be added for credit card processing
Credit card transactions are processed by Cornell Store
Purchasing options:
Current used storage
0.0GB
Purchased storage left
790.6GB-year
Add 1 units of 1.0 TB-year at $80.00 each
1.0TB-year for $80.00
Quota
512.0GB (0.5TB)
now quota is 0.5TB
Expiration date
9/2/2017 4:53:43 PM
… and new expiration date will be 9/2/2017
Description (optional)
Submit
Reset
100%

Storage

Similarly as with Lab Credit Accounts computing hours you are charged for *reservation* of storage, i.e. your TB-year storage purchased is used and subtracted based on your *quota*, NOT the amount of actually stored data.

Local genomic data

Available on the local drives (/local_data – you can use directly)

- Arabidopsis_thaliana_tair10
- Caenorhabditis_elegans_ce10
- Drosophila_melanogaster_dm3
- Homo_sapiens_UCSC_hg19
- Mus_musculus_UCSC_mm10
- Saccharomyces_cerevisiae_sacCer3
- Zea_mays_agpv3

Local genomic data

Available on the network file server (/shared_data/genome_db – must be copied to local directory before using)

- NCBI BLAST database (nt, nr and others)
- interproscan
- Arabidopsis_thaliana
- Caenorhabditis_elegans
- Drosophila_melanogaster
- Homo_sapiens
- Mus_musculus
- Saccharomyces_cerevisiae
- Zea_mays
- apple
- grape

Local Storage Notes

Your local storage on a machine (in /workdir or /SSD) is NOT persistent!

It will be deleted some time after your reservation ends, so you must copy all the files back to your home directory if you want to keep them.

Local directories are cleaned at 3am every night.

If the machine is not used the files may stay up to 5 days.

If the machine is used by other users, the files will be deleted on the first night …

Getting started with a new account

- ✖ You need hours: create and fund your own Lab Credit Account or get added to one
- ✖ Get extra storage if needed – all users get free storage allocations, but it may not be enough
- Transfer data to your Lab storage
- ✖ Make reservation(s)
- Connect to reserved workstations
- *Compute!*

Connecting to BioHPC Lab machines

Text-based connection: ssh (Secure SHell)

GUI (graphical) connection: X-Windows or VNC

Logging in to a Linux machine

- On any Linux machine, you need
  - **network name** of the machine (e.g. cbsumm10.tc.cornell.edu)
  - an account, i.e., **user ID** and **password**
  - on your local computer: remote access software (typically: **ssh client**)

- Linux is a multiple-access system: multiple users may be logged in and operate on one machine at the same time

Logging in to a Linux machine

- Remotely from a PC via **ssh client**
  - Install and configure remote access software (**PuTTy**).
  - Use **PuTTy** to open a terminal window on the reserved workstation using **ssh** protocol;
  - You may open several terminal windows, if needed.

Logging in to a Linux machine

- Remotely from other Linux machine or Mac via native **ssh client**

  - Launch the Mac's terminal window. Type

    ssh jarekp@cbsuwrkstX.tc.cornell.edu

    (replace the "**cbsuwrkstX"** with the workstation that you just reserved, and "**jarekp**" with your own user ID). Enter the lab password when prompted.

  - You may open several terminal windows, if needed, and log in to the workstation from each of them.

Logging in to a Linux machine

- Directly from the console (consoles of cbsuwrkst2, 3, and 4 are in 625 Rhodes)
  - Turn on the monitor, then log in using the login screen
  - Right-click anywhere on the desktop, then select “Open terminal” – a terminal window will open
  - You can open additional terminals window the same way

Logging in to CBSU machines from outside of Cornell

Two ways to connect from outside:

- Install and run the CIT-recommended the VPN software (http://www.it.cornell.edu/services/vpn) to join the Cornell network, then proceed as usual

- Log in to cbsulogin.tc.cornell.edu:

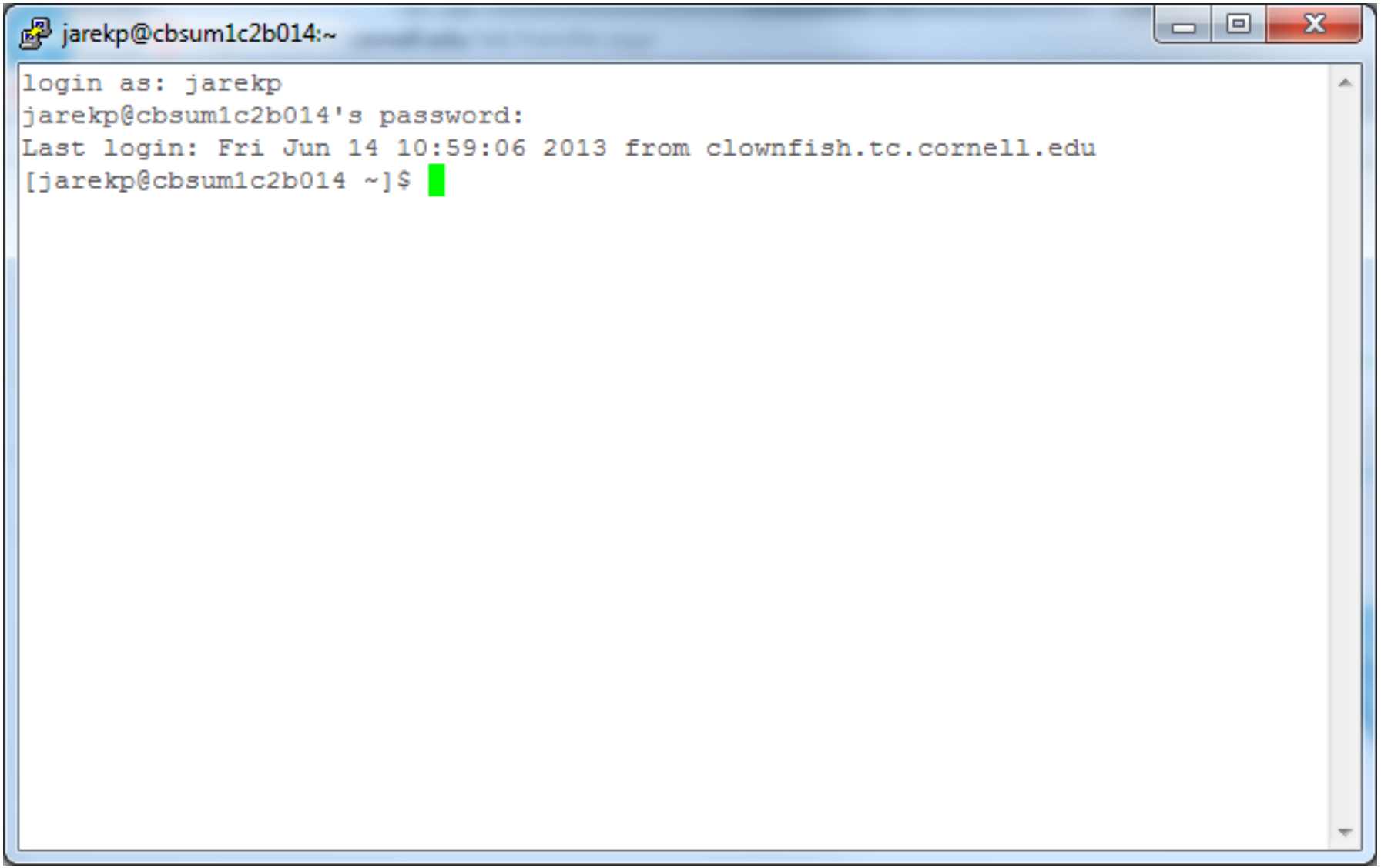
  ssh jarekp@cbsulogin.tc.cornell.edu (using PuTTy or other ssh client program)

Once logged in to cbsulogin, ssh further to your reserved machine

ssh jarekp@cbsuwrkst3.tc.cornell.edu

Backup login machine is cbsuss02.tc.cornell.edu

Terminal window



Terminal window

- User communicates with the machine via **commands** typed in the terminal window
  - Commands are interpreted by a program referred to as **shell** – an interface between Linux and the user. We will be using the shell called **bash** (another popular shell is **tcsh**).
  - Typically, each command is typed in one line and "**entered**" by hitting the **Enter** key on the keyboard.
  - Commands deal with **files** and **processes**, e.g.,
    - request information (e.g., list user's files)
    - launch a simple task (e.g., rename a file)
    - start an application (e.g., Firefox web browser, BWA aligner, IGV viewer, …)
    - stop an application

Logging out of a Linux machine

- While in terminal window, type **exit** or **Ctrl-D** - this will close the current terminal window

Exercise: connect to your assigned workstations using ssh

- Find your assigned machine on the list on workshop page http://cbsu.tc.cornell.edu/ww/1/Default.aspx?wid=44

- Windows: open PuTTY program, enter your workstation name and connect. Provide your user name and password when prompted.

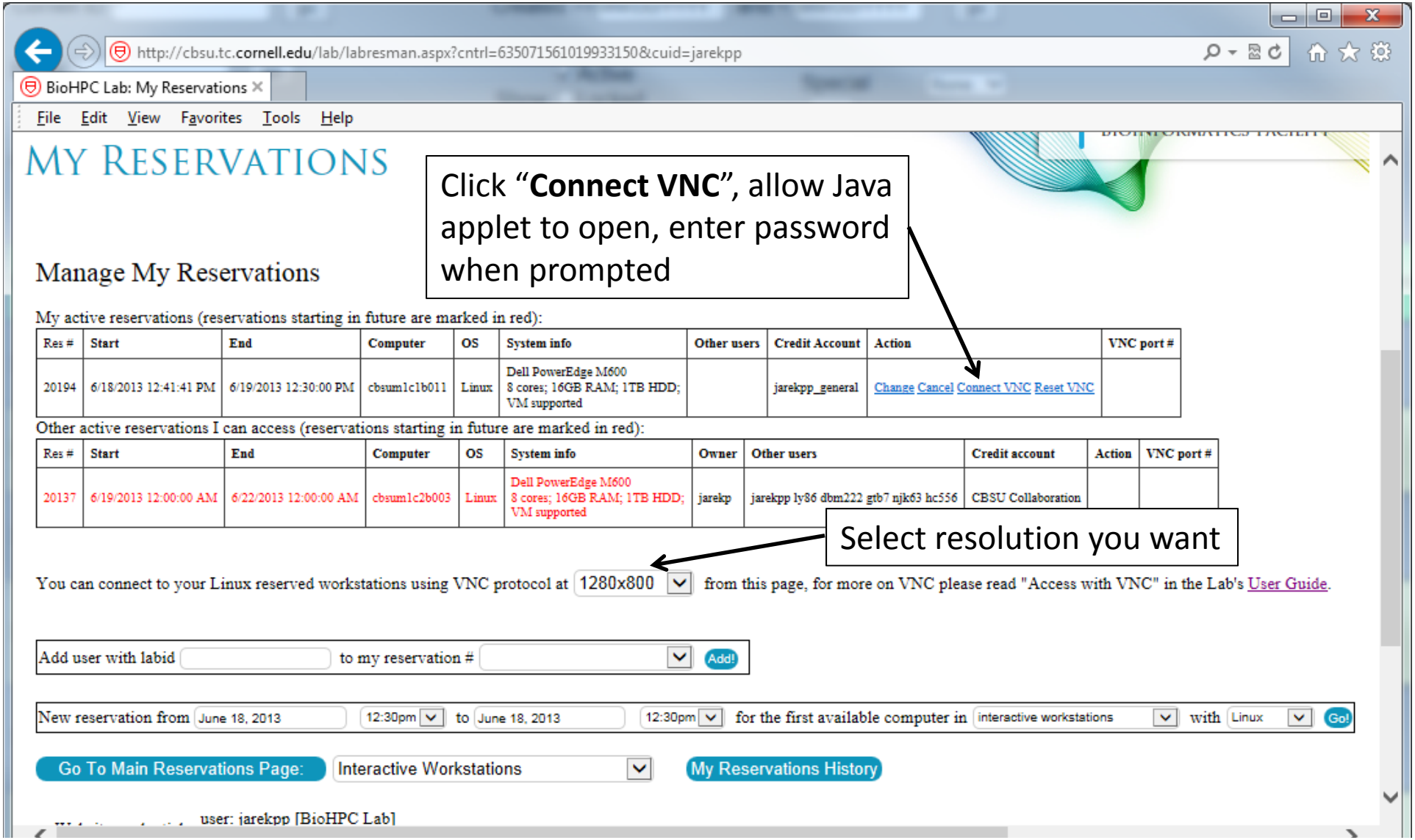
- Linux or Mac: Open terminal window and type ssh command “ssh labid@workstation.tc.cornell.edu”. Provide your user name and password when prompted.

Connecting to BioHPC Lab

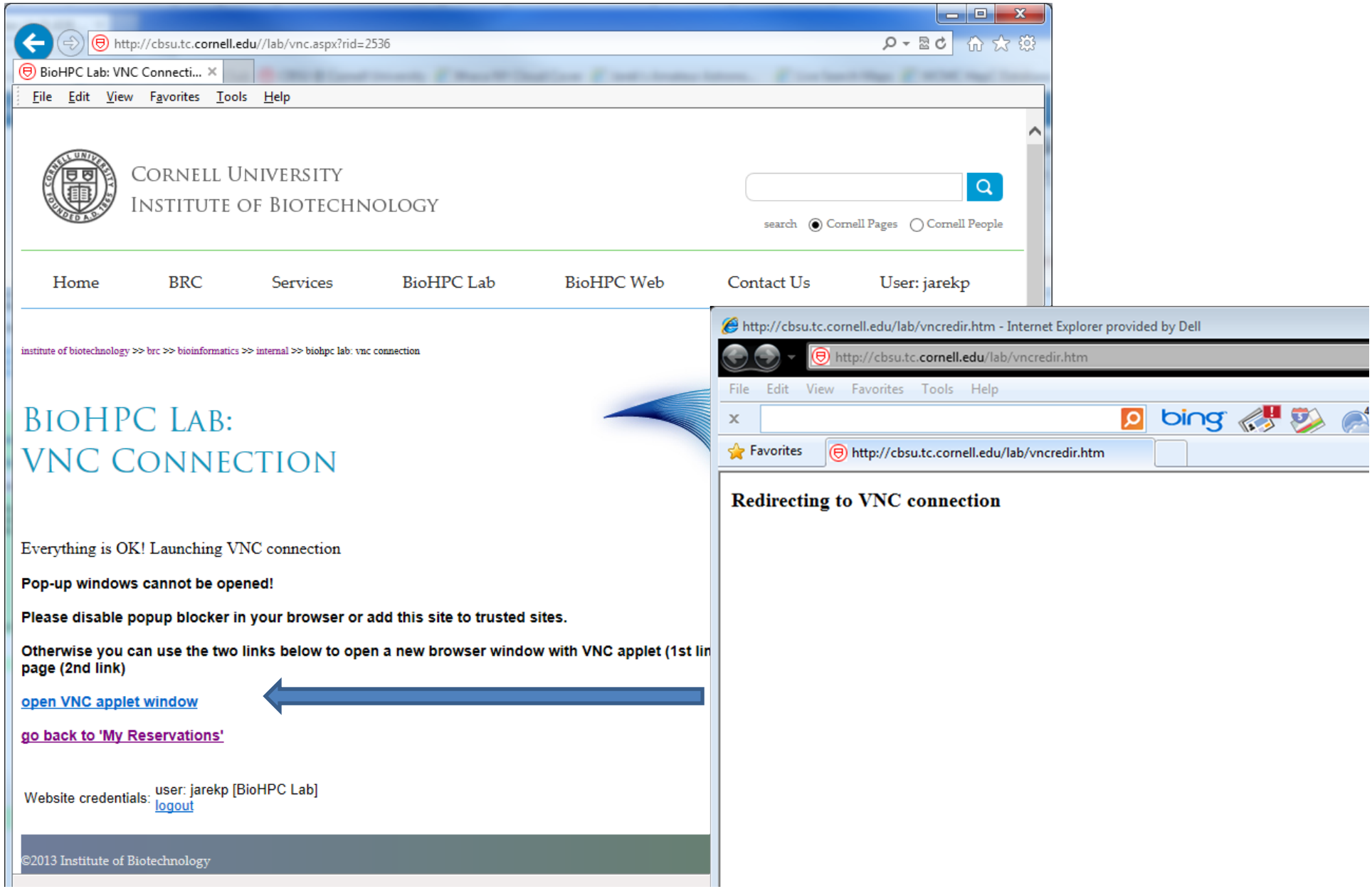
GUI (graphical) connection: VNC

Logging in to a Linux workstation via web browser VNC client (GUI)

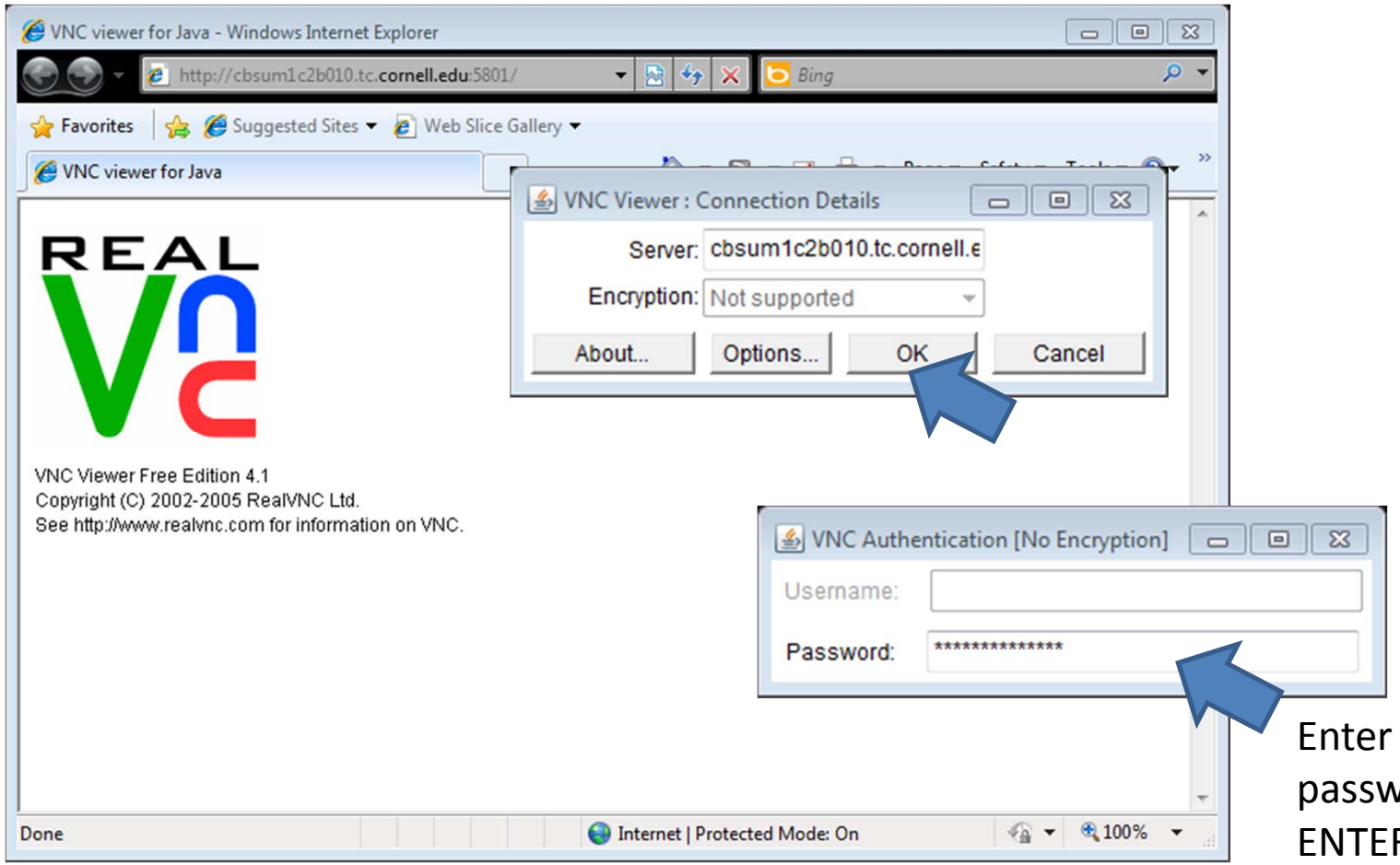
In web browser, navigate to http://cbsu.tc.cornell.edu/, log in (if not yet logged in), click on **User:your_id**, select tab **My Reservations**



Logging in to a Linux workstation
(GUI)



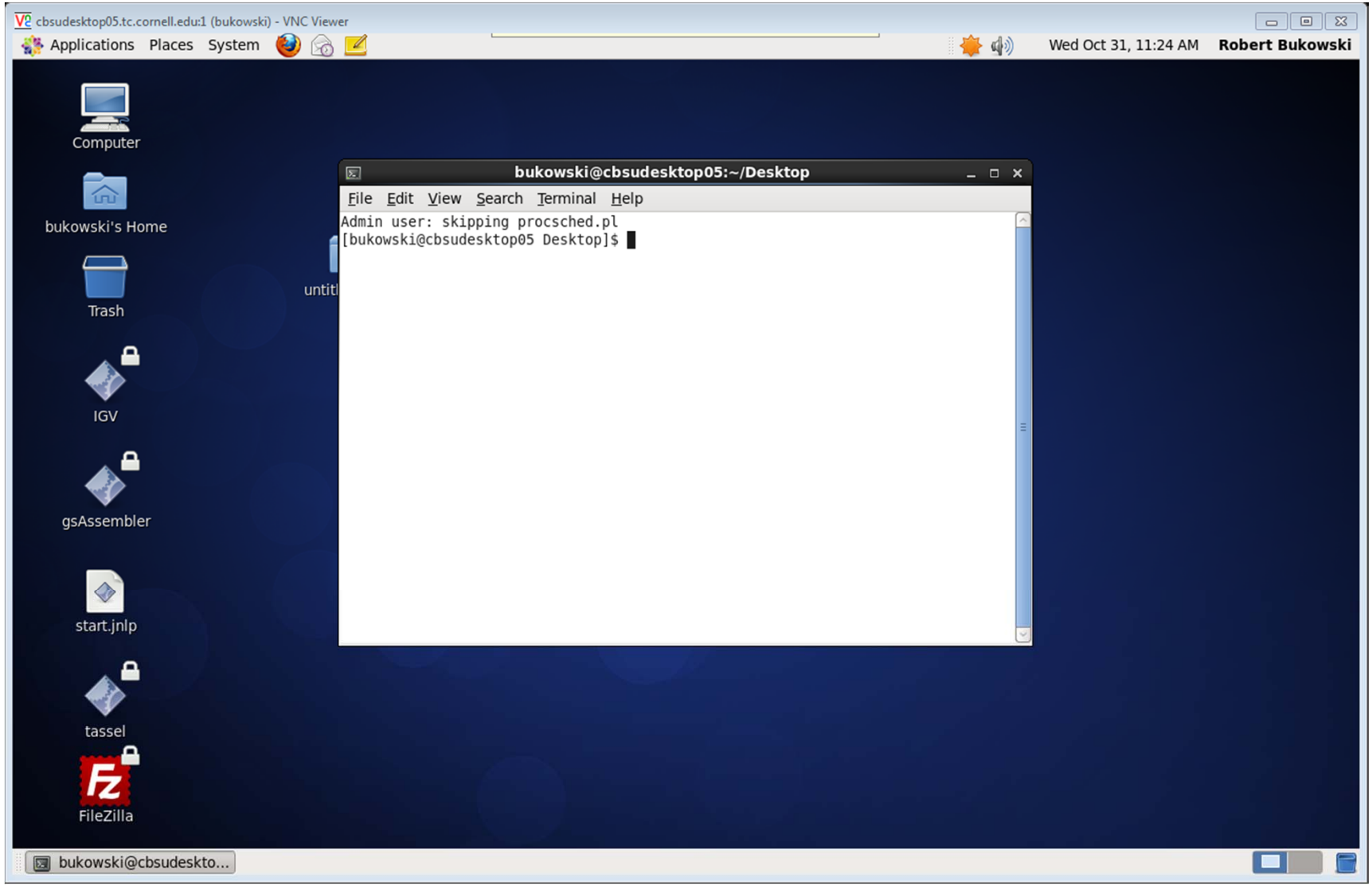
Keep this browser window open



Enter your Lab password, hit ENTER

Logging in to a Linux workstation via web browser VNC client
(GUI)

Right-click anywhere within blue desktop, select **Open Terminal** …. or
…. click **Applications -> Accessories -> Terminal**



Exercise: connect to your assigned workstation using VNC

- Go to “My Reservations” page
  http://cbsu.tc.cornell.edu/lab/lab.aspx , log in, click on “My Reservations” menu link
- Choose resolution (depends on your monitor)
- Click on “Connect VNC”
- Follow prompts
- Open terminal window in the VNC desktop by right-click on the desktop background and choosing “Open Terminal”.
- Disconnect (close browser windows) and then reconnect. Is the session still alive?

Connecting to BioHPC Lab: VNC

VNC sessions are *persistent*.

They run even when the client is disconnected.

If you need to reset the session you need to use "Reset VNC" link.

Equivalent to Windows Remote Desktop.

Connecting to BioHPC Lab

GUI (graphical) connection: X-Windows

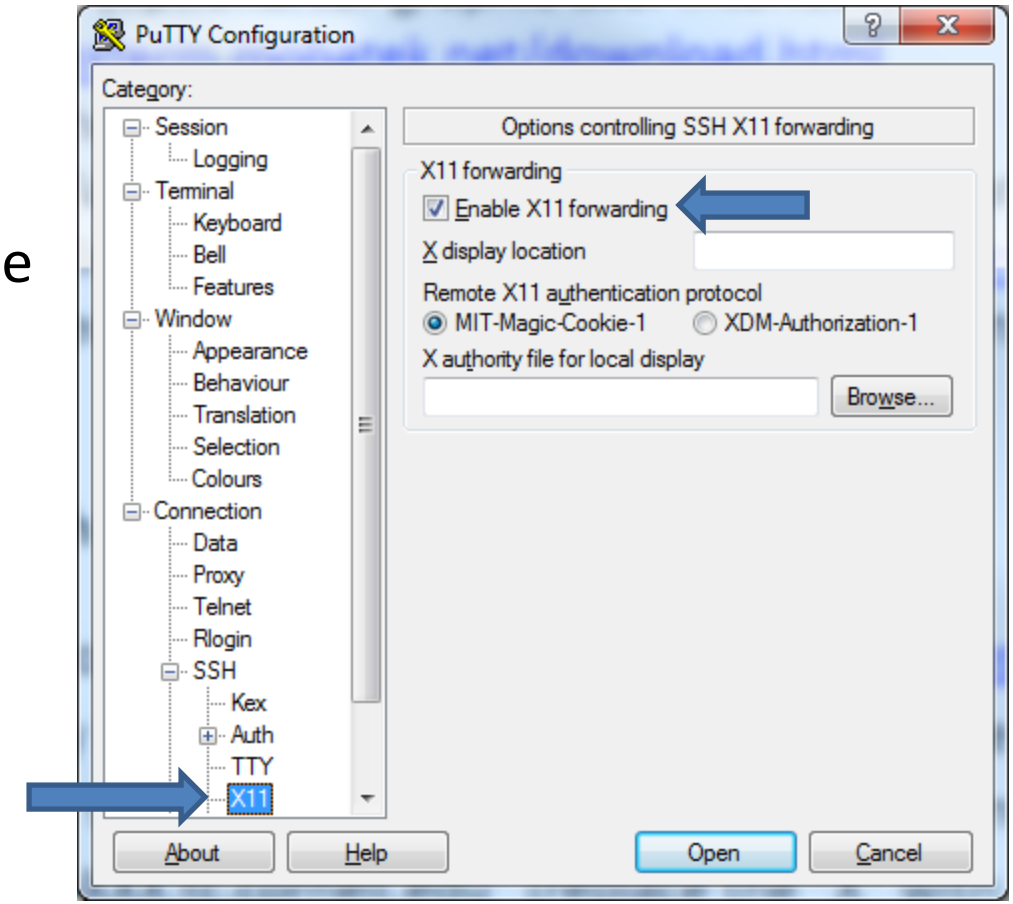
Not persistent – programs will get killed when client disconnects.

Connecting to BioHPC Lab: X-Windows Windows

- Install X-Windows software on your computer. We recommend MobaXterm free (http://mobaxterm.mobatek.net/download.html)

- Start MobaXterm

- Connect to BioHPC Lab machine using PuTTY. Make sure X11 forwarding is enabled. X11 is a synonym for X-Windows



Connecting to BioHPC Lab: X-Windows Windows

- Start your GUI (graphical) software in terminal window. For testing you may try eog (Linux image viewer) or firefox.

- You may need to allow MobaXterm connections through Windows firewall.

- New window with your GUI program will appear. The program will physically run on the BioHPC Lab machine, but it will display graphics on your local computer.

Connecting to BioHPC Lab: X-Windows Linux or Mac

- Connect to BioHPC Lab machine using ssh with X11 forwarding :
  ssh -X labid@workstation.tc.cornell.edu

- Start your GUI program in remote terminal window. For testing you may try eog (Linux image viewer) or firefox.

- New window with your GUI program will appear. The program will physically run on the BioHPC Lab machine, but it will display graphics on your local computer.

Exercise: connect to your assigned workstation with X-Windows

- Connect to your workstations using ssh with X11 forwarding enabled
- Windows: start MobaXterm
- Start eog in remote terminal.

Getting started with a new account

- ✖ You need hours: create and fund your own Lab Credit Account or get added to one
- ✖ Get extra storage if needed – all users get free storage allocations, but it may not be enough
- Transfer data to your Lab storage
- ✖ Make reservation(s)
- ✖ Connect to reserved workstations
- *Compute!*

Transferring data

- sftp (secure file transfer protocol)
  Transfer can be done to and from Lab machines.

- Globus
  Transfer can be done to and from Lab machines.

- wget
  Transfer can be done to Lab machines only.

File Transfer: summary

Another **Linux** or **Mac** machine (call it **cbsuss04)**

web

Initiate **sftp** connection from terminal of either machine:

sftp jarekp@cbsumm10.tc.cornell.edu

OR

sftp myid_remote@cbsuss04.tc.cornell.edu

Linux workstation e.g., **cbsumm10**

Download file using **firefox** web browser

OR

Use the URL directly with **wget** command, e.g.,

wget ftp://ftp.ncbi.nih.gov/blast/matrices/BLOSUM100

SFTP: secure file transfer protocol

Mac

Use **fetch** – a graphical file transfer program for MACs to **sftp** to Linux machine

sftp to Linux machine using any graphical **sftp client program**, such as:

FileZilla

WinScp

CoreFTP

…..

Windows PC

File transfer: wget

from web- and ftp sites to lab workstations

Option 1: run **wget** command on the workstation (if you know the URL of the file)

- **Examples:**

wget ftp://ftp.ncbi.nih.gov/blast/matrices/BLOSUM100

(will download the file BLOSUM100 from the NCBI FTP site and deposit it in the current directory under the name BLOSUM100)

- **another Example (the following should be typed on one line):**

wget -O e_coli_1000_1.fq
"http://cbsuapps.tc.cornell.edu/Sequencing/showseqfile.aspx?cntrl=646698859&laneid=487&mode=http&file=e_coli_1000_1.fq"

(the command above can be used to download files given by complicated URLs; note the **""** *marks around the link and the* **–O** *option which specifies the name you want to give the downloaded file)*

Exercise: download BLOSUM100 with wget

- Connect to your workstation using ssh
- Type wget command

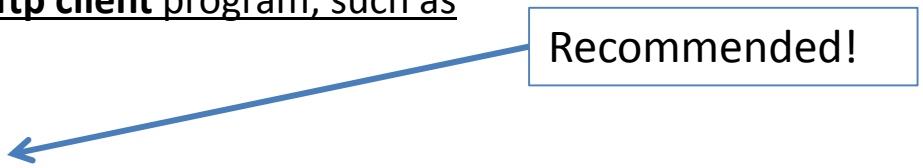
  wget ftp://ftp.ncbi.nih.gov/blast/matrices/BLOSUM100

File transfer: sftp

between PC or Mac and a Lab workstation

On Windows PC: install and use your favorite **sftp client** program, such as

- **winscp**: http://winscp.net/eng/index.php
- **CoreFTP LE**: http://www.coreftp.com/
- **FileZilla** (client): http://filezilla-project.org/
- … others…
- When connecting to Lab workstations from a client, use the **sftp** protocol. You will be asked for your user name and password (the same you use to log in to the lab workstations).
- Transfer text file in text mode, binary files in binary mode (the "default" not always right).
- All clients feature
  - File explorer-like graphical interface to files on both the PC and on the Linux machine
  - Drag-and-drop functionality



On a Mac: file transfer program is **fetch** (recommended by Cornell CIT)

- http://www2.cit.cornell.edu/services/systems_support/filefetch.html#fetchinst
- graphical user interface
- Drag-and-drop functionality

File transfer: sftp

between a lab workstation and another Linux machine

Suppose we want to transfer a file from **cbsuss04.tc.cornell.edu** (another Linux machine; substitute "your" Linux machine here) and **cbsuwrkst2** lab workstation.

Option 1: when logged in to **cbsuwrkst2**, sftp to **cbsuss04** by running the following commands:

cd /workdir/jarekp	*(this is where we want the file to be on cbsuwrkst2)*
sftp jarekp@cbsuss04.tc.cornell.edu	*(instead of "jarekp", use your own user name on cbsuss04; you will be asked for password)*
cd /data/jarekp/blat	*(on cbsuss04, go to the directory where the file is)*
get e_coli_1000_1.fq	*(transfer, or "get" the file from cbsuss04)*
quit	*(exit sftp client and disconnect from cbsuss04 – we are back on cbsuwrkst2)*

Option 2: when logged in to **cbsuss04**, sftp to **cbsuwrkst2** by running the following commands:

cd /data/jarekp/blat	*(this is where the file is on cbsuss04)*
sftp jarekp@cbsuwrkst2.tc.cornell.edu	*(instead of "jarekp", use your own user name on cbsuss04; you will be asked your lab password)*
cd /workdir/jarekp	*(on cbsuwrkst2, go to the directory where the file is supposed to be stored)*
put e_coli_1000_1.fq	*(transfer, or "put" the file on cbsuwrkst2)*
quit	*(exit sftp client and disconnect from cbsuwrkst2– we are back on cbsuss04)*

Option 3: **Filezilla** is installed on Lab workstations. Connect to Lab workstations with GUI support (VNC or X-Windows), type filezilla and connect to the other servers using sftp protocol.

Exercise: download BLOSUM100 file from Lab machine to your local computer

- Connect to your workstation using sftp program
- Download BLOSUM100

Transferring data: Globus

Globus Online is an online system based on GridFTP engine that supports fast and reliable data transfer and can be scheduled and controlled online in a web browser.

Data can be transferred between endpoints, which can be data servers or Globus Connect clients.

BioHPC Lab Globus endpoint server is **biohpc#cbsulogin**.

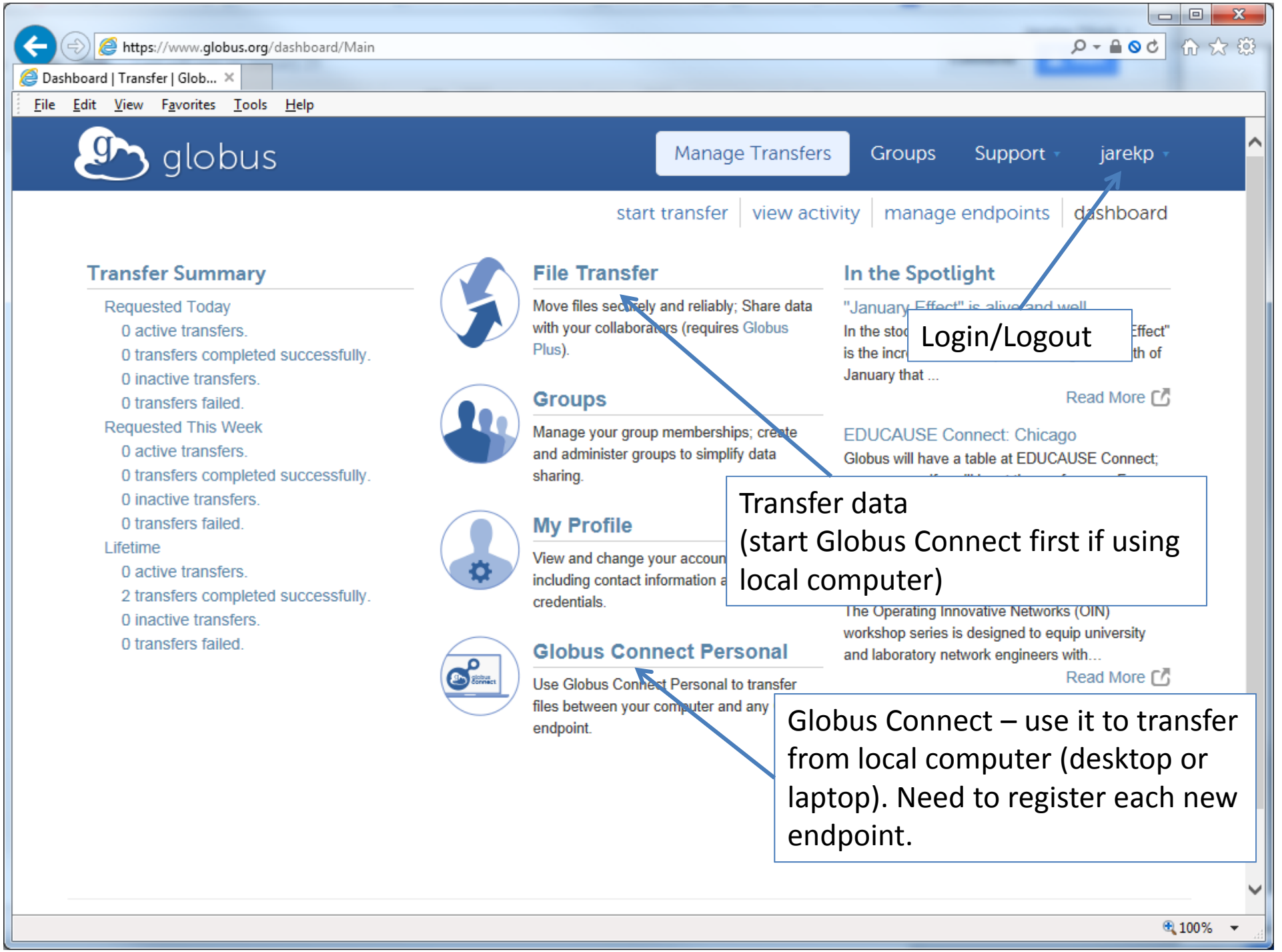
You can transfer data to another server or to your computer (using Globus Connect client).

http://cbsu.tc.cornell.edu/lab/doc/Globus_at_BioHPC_Lab.pdf

https://www.globusonline.org/quickstart/

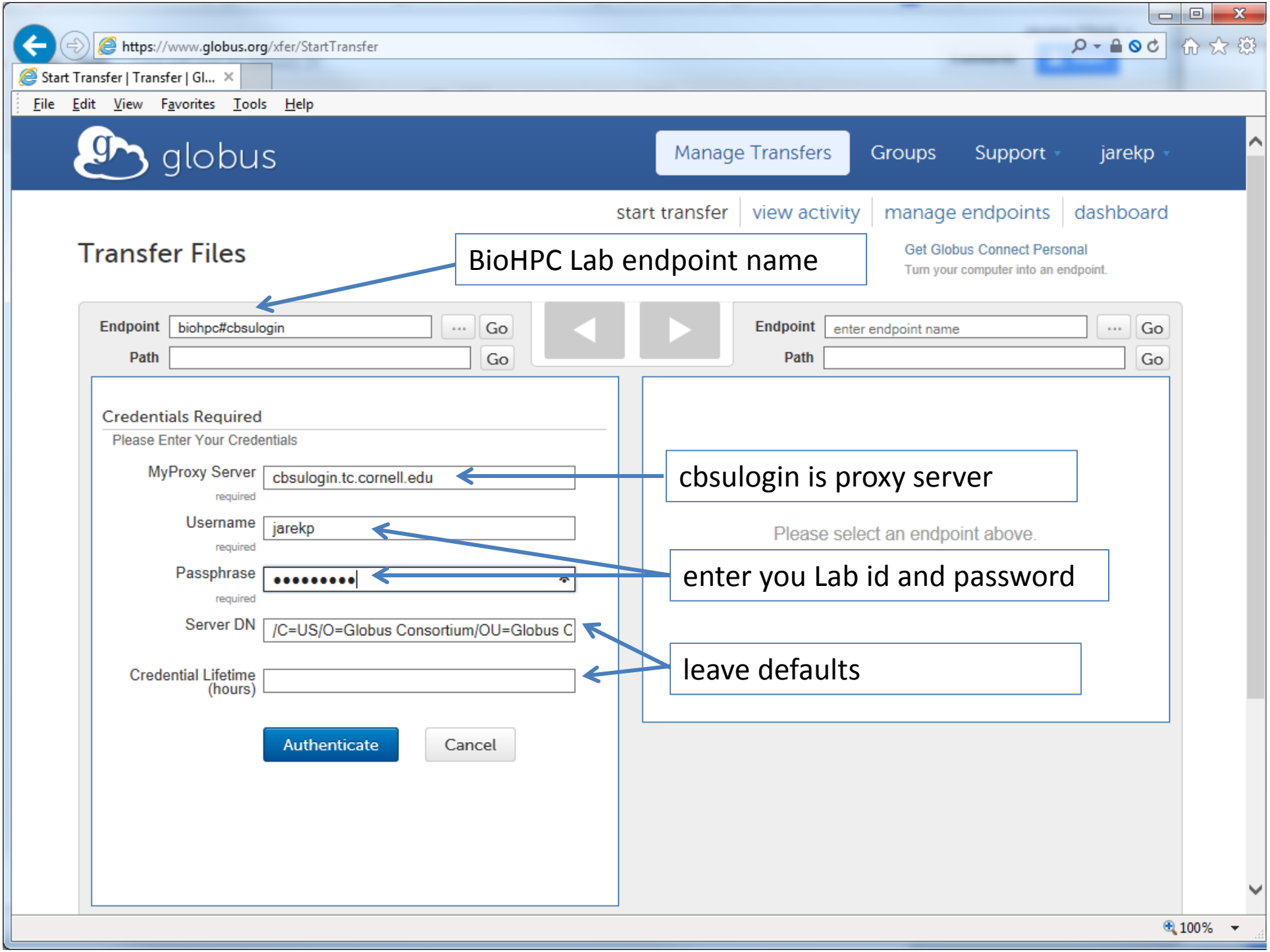
Transferring data: Globus

- Sign up for Globus online account if you don't have one (use the link above).
- Sign in to Globus online using your id and password
- Go to File Transfer, in the left panel "Endpoint" field type **biohpc#cbsulogin** and click go. Login window will pop up, you need to use your BioHPC Lab user id and password to connect to the endpoint, click Authenticate
- When authentication is successful files from your home directory will show up in the left panel.
- If you would like to connect to another **server** endpoint you can just type it in the right side "Endpoint" field and connect.
- If you want to transfer to the local laptop (which is not a server) you need to install Globus Connect (if it is already installed, just start it).


https://www.globus.org/dashboard/Main
Dashboard | Transfer | Glob...
File Edit View Favorites Tools Help
globus
Manage Transfers
Groups
Support
jarekp
start transfer
view activity
manage endpoints
dashboard
Transfer Summary
Requested Today
0 active transfers.
0 transfers completed successfully.
0 inactive transfers.
0 transfers failed.
Requested This Week
0 active transfers.
0 transfers completed successfully.
0 inactive transfers.
0 transfers failed.
Lifetime
0 active transfers.
2 transfers completed successfully.
0 inactive transfers.
0 transfers failed.
File Transfer
Move files securely and reliably; Share data with your collaborators (requires Globus Plus).
Groups
Manage your group memberships; create and administer groups to simplify data sharing.
My Profile
View and change your accoun
including contact information a
credentials.
Globus Connect Personal
Use Globus Connect Personal to transfer files between your computer and any endpoint.
In the Spotlight
"January Effect" is alive and well
In the sto
Effect"
is the incr
th of
January that ...
Read More
EDUCAUSE Connect: Chicago
Globus will have a table at EDUCAUSE Connect;
The Operating Innovative Networks (OIN) workshop series is designed to equip university and laboratory network engineers with...
Read More
Login/Logout
Transfer data
(start Globus Connect first if using local computer)
Globus Connect – use it to transfer from local computer (desktop or laptop). Need to register each new endpoint.
100%

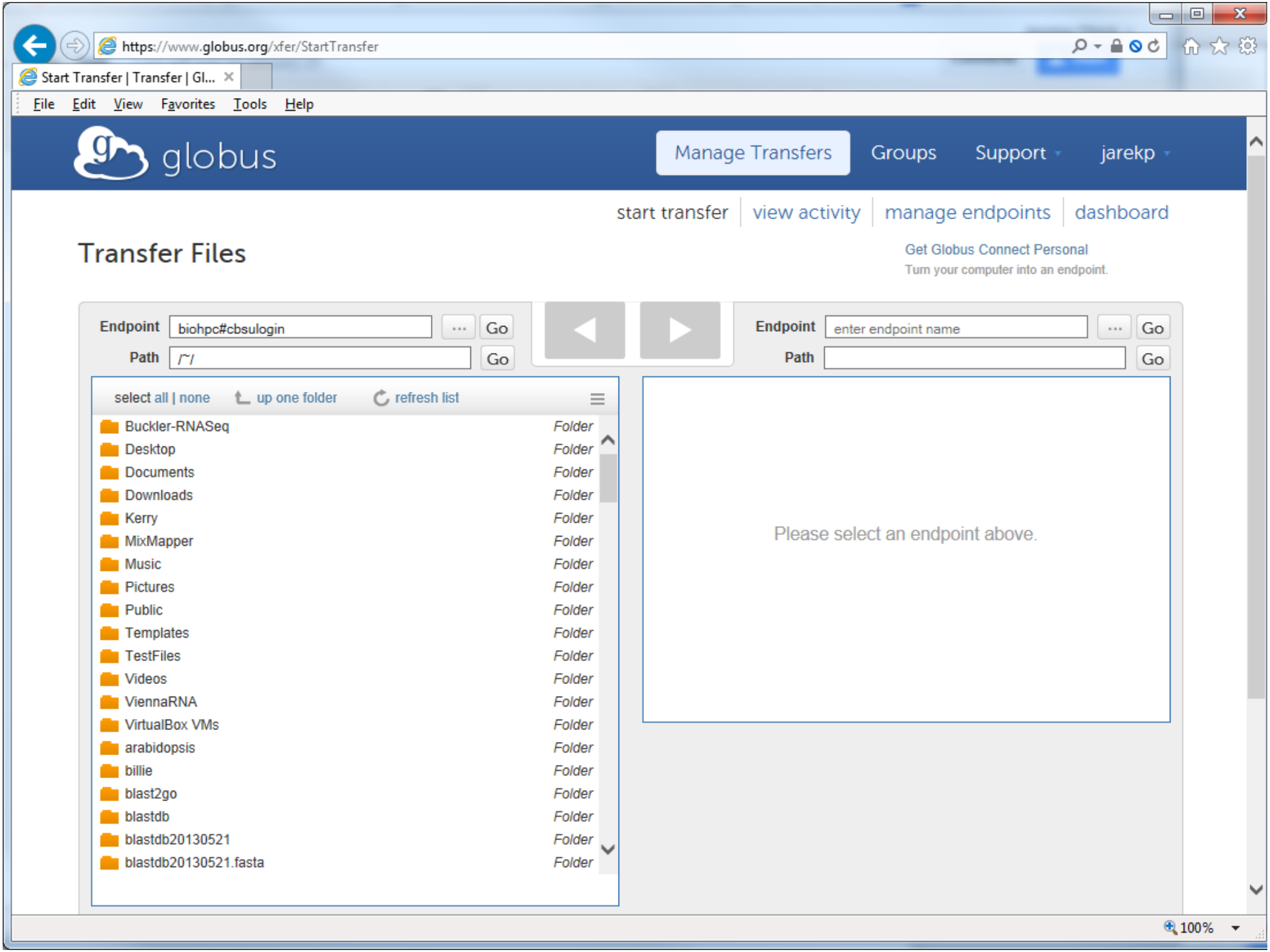
Transferring data: Globus

- Sign up for Globus online account if you don't have one (use the link above).
- Sign in to Globus online using your id and password
- Go to File Transfer, in the left panel "Endpoint" field type **biohpc#cbsulogin** and click go. Login window will pop up, you need to use your BioHPC Lab user id and password to connect to the endpoint, click Authenticate
- When authentication is successful files from your home directory will show up in the left panel.
- If you would like to connect to another **server** endpoint you can just type it in the right side "Endpoint" field and connect.
- If you want to transfer to the local laptop (which is not a server) you need to install Globus Connect (if it is already installed, just start it).


https://www.globus.org/xfer/StartTransfer
Start Transfer | Transfer | Gl...
File Edit View Favorites Tools Help
globus
Manage Transfers
Groups
Support
jarekp
start transfer
view activity
manage endpoints
dashboard
Transfer Files
BioHPC Lab endpoint name
Get Globus Connect Personal
Turn your computer into an endpoint.
Endpoint
biohpc#cbsulogin
Go
Path
Go
Endpoint
enter endpoint name
Go
Path
Go
Credentials Required
Please Enter Your Credentials
MyProxy Server
cbsulogin.tc.cornell.edu
required
Username
jarekp
required
Passphrase
required
Server DN
/C=US/O=Globus Consortium/OU=Globus C
Credential Lifetime (hours)
Authenticate
Cancel
cbsulogin is proxy server
Please select an endpoint above.
enter you Lab id and password
leave defaults
100%

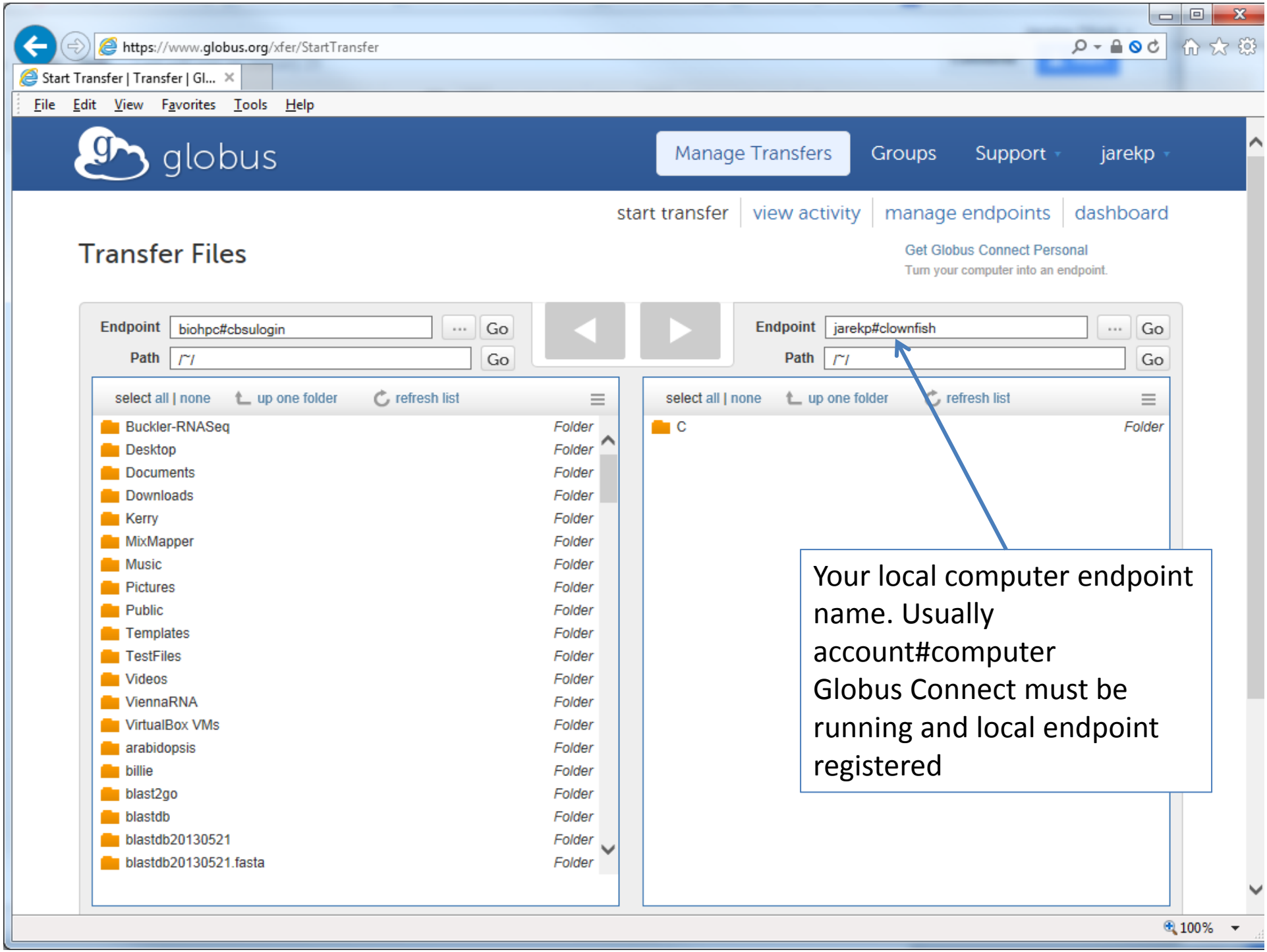
Transferring data: Globus

- Sign up for Globus online account if you don't have one (use the link above).
- Sign in to Globus online using your id and password
- Go to File Transfer, in the left panel "Endpoint" field type **biohpc#cbsulogin** and click go. Login window will pop up, you need to use your BioHPC Lab user id and password to connect to the endpoint, click Authenticate
- When authentication is successful files from your home directory will show up in the left panel.
- If you would like to connect to another **server** endpoint you can just type it in the right side "Endpoint" field and connect.
- If you want to transfer to the local laptop (which is not a server) you need to install Globus Connect (if it is already installed, just start it).


https://www.globus.org/xfer/StartTransfer
Start Transfer | Transfer | Gl...
File Edit View Favorites Tools Help
globus
Manage Transfers
Groups
Support
jarekp
start transfer
view activity
manage endpoints
dashboard
Transfer Files
Get Globus Connect Personal
Turn your computer into an endpoint.
Endpoint
biohpc#cbsulogin
Go
Path
/~/
Go
Endpoint
enter endpoint name
Go
Path
Go
select all | none
up one folder
refresh list
Buckler-RNASeq Folder
Desktop Folder
Documents Folder
Downloads Folder
Kerry Folder
MixMapper Folder
Music Folder
Pictures Folder
Public Folder
Templates Folder
TestFiles Folder
Videos Folder
ViennaRNA Folder
VirtualBox VMs Folder
arabidopsis Folder
billie Folder
blast2go Folder
blastdb Folder
blastdb20130521 Folder
blastdb20130521.fasta Folder
Please select an endpoint above.
100%

Transferring data: Globus

- Sign up for Globus online account if you don't have one (use the link above).
- Sign in to Globus online using your id and password
- Go to File Transfer, in the left panel "Endpoint" field type **biohpc#cbsulogin** and click go. Login window will pop up, you need to use your BioHPC Lab user id and password to connect to the endpoint, click Authenticate
- When authentication is successful files from your home directory will show up in the left panel.
- If you would like to connect to another **server** endpoint you can just type it in the right side "Endpoint" field and connect.
- If you want to transfer to the local laptop (which is not a server) you need to install Globus Connect (if it is already installed, just start it).


https://www.globus.org/xfer/StartTransfer
Start Transfer | Transfer | Gl...
File Edit View Favorites Tools Help
globus
Manage Transfers
Groups
Support
jarekp
start transfer
view activity
manage endpoints
dashboard
Transfer Files
Get Globus Connect Personal
Turn your computer into an endpoint.
Endpoint
biohpc#cbsulogin
Go
Path
/~/
Go
select all | none
up one folder
refresh list
Buckler-RNASeq Folder
Desktop Folder
Documents Folder
Downloads Folder
Kerry Folder
MixMapper Folder
Music Folder
Pictures Folder
Public Folder
Templates Folder
TestFiles Folder
Videos Folder
ViennaRNA Folder
VirtualBox VMs Folder
arabidopsis Folder
billie Folder
blast2go Folder
blastdb Folder
blastdb20130521 Folder
blastdb20130521.fasta Folder
Endpoint
jarekp#clownfish
Go
Path
/~/
Go
select all | none
up one folder
refresh list
C Folder
Your local computer endpoint name. Usually account#computer
Globus Connect must be running and local endpoint registered
100%

Exercise: transfer BLOSUM100 file to your laptop using Globus

Getting started with a new account

- You need hours: create and fund your own Lab Credit Account or get added to one
- Get extra storage if needed – all users get free storage allocations, but it may not be enough
- Transfer data to your Lab storage
- Make reservation(s)
- Connect to reserved workstations

- *Compute!*

Getting started with a new account

In order to execute last item of the workflow (*Compute!*) some additional knowledge may be needed

- how to work in Linux environment?
  => see our "***Linux for Biologists***" workshop
  http://cbsu.tc.cornell.edu/ww/1/Default.aspx?wid=45

- what programs to use for my bioinformatics problems? How?
  => use our ***facility office hours*** to discuss problems and solutions
  http://cbsu.tc.cornell.edu/lab/office1.aspx

- sometimes custom data processing is needed that requires a little bit of tinkering with software and scripts

  => see our "***Perl for Biologists***" workshop
  http://cbsu.tc.cornell.edu/ww/1/Default.aspx?wid=46

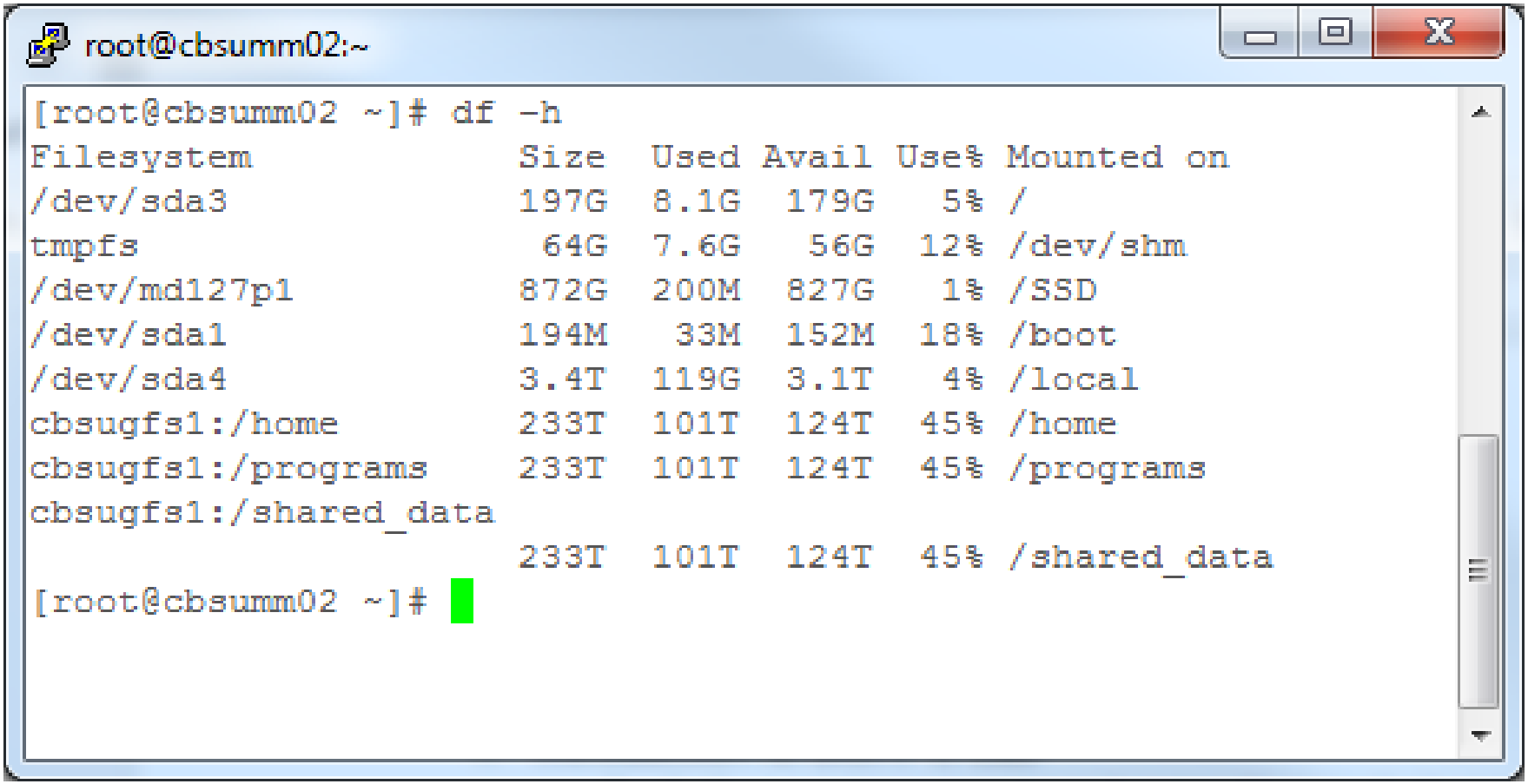
  => use our ***facility office hours*** to discuss problems and solutions
  http://cbsu.tc.cornell.edu/lab/office1.aspx

APPENDIX
Storage

Which storage is local?

When in doubt, use “df -h”.

Entries with name starting with “/” are local, entries in the form of *name:/path* are networked (name being the server name and /path exported directory or share)



Storage

For example you can buy 30 x 1TB-year chunks and set the quota for 30TB, and it will last for 1 year, at which point you will need to buy storage again.

You can buy 60 x 1 TB-year chunks and set the quota for 30TB, this will last 2 years.

You can change the quota at any time, the remaining TB-years (not rounded) will be used to compute new expiration date.

You can add TB-year units at any time (and change or not change quota as you like), you can lower your quota at any time (and push back your expiration date as a result), but you cannot get a refund (i.e. convert the TB-year units left back to $$).

Storage

If you need extra storage for a short time, you can raise your quota temporarily, and then lower it back when not needed, it will use more of your TB-year units, but only so many as needed (and usage is computed based on quota and time, where TB-year are counted with floating point numbers).

For example an additional 3 TB quota increase for 6 months will cost you 1.5 TB-year units. Your usage of your TB-year units solely depends on quota you set, essentially you pay for reserving certain amount of storage.