Introduction to BioHPC Cloud

BioHPC Cloud Workshop

Jaroslaw Pillardy

Bioinformatics Facility
Institute of Biotechnology
Cornell University

http://biohpc.cornell.edu/lab/lab.aspx

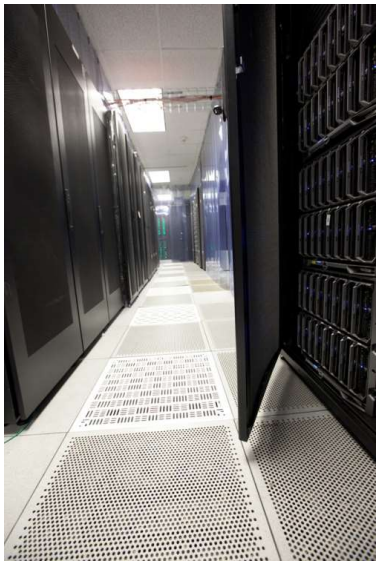
http://biohpc.cornell.edu/lab/doc/Introduction_to_BioHPC_Cloud_v8.pdf

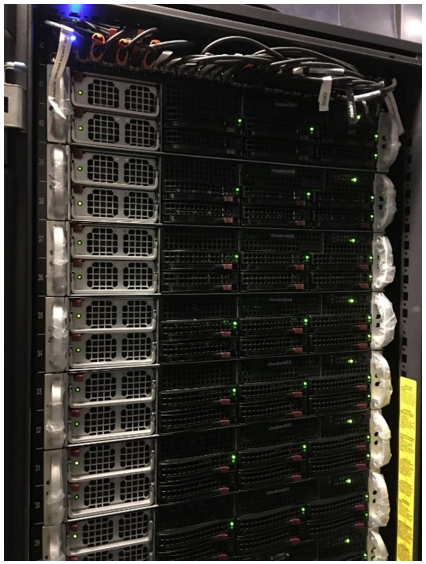
Part 1
Introduction to BioHPC Cloud

BioHPC Cloud

BioHPC Cloud 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 700 titles)
- Located in Rhodes Hall and accessible remotely and on-site



BioHPC Cloud

Renting computing resources

Hosting computing resources

Storage

Consulting and education

BioHPC Cloud: computing

Two ways of using BioHPC Cloud computing resources:

- Buy hours and run computations on BioHPC Cloud 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 better ask us to buy it for you) and host it in BioHPC Cloud

  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 BioHPC

BioHPC Cloud: hosting

Buy your own hardware and host it in BioHPC Cloud

- Fully customizable hardware – buy just what needed
- Access to all software and storage – same as rental
- Machine usage restricted to group: usage may be regulated internally (all have access) or by reservations (users need to reserve machine)
- No root access – installations that require full admin rights may be done by us or in a customized Docker container
- Use BioHPC scheduler or use SLURM
- Storage servers can also be hosted

BioHPC Cloud: storage

- 1,743 TB of networked storage available in two volumes: 1,510 TB Lustre volume and 233 TB Gluster volume.
- Very robust and scalable architecture: cluster storage based on RAID6/RAID7 servers connected by Lustre or Gluster, new servers can be added to expand storage.
- Very affordable! $95.01 per TB per year
- Users get free storage allocations with the accounts

BioHPC Cloud: Backup service

- Users are able to choose backup options using BioHPC website (what to backup, how many versions etc)

- Backup servers (currently 459TB) are located in different building (Weill Hall)

BioHPC Cloud: software

- BioHPC Cloud 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.
- 719 titles as of 1/31/2020
- Common genomic data is available locally in the Lab: sequence and annotation databases, preformatted for common programs

http://biohpc.cornell.edu/lab/labsoftware.aspx

BioHPC Cloud 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” servers:

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

“medium gen1” servers

1	16-core, 64GB RAM, 1TB HDD
16	24-core, 128GB RAM, 4TB HDD, 1TB SSD
	(cbsumm01-16)

“medium gen2” servers

12	40-core, 256GB RAM, 8TB HDD
	(cbsumm21-32)

“large gen1” servers

8	64-core, 512GB RAM, 12TB HDD
	(cbsulm02-05, cbsulm07-10)

BioHPC Cloud hardware infrastructure

"large gen2" servers

2	96-core, 512GB RAM, 12TB HDD, 1TB SSD (cbsulm12-13)
4	112-core, 512GB RAM, 12TB HDD, 1TB SSD (cbsulm14-17)
2	80-core, 512GB RAM, 10TB HDD, (cbsulm14-17)
3	88-core, 512GB RAM, 7TB NVMe SSD (cbsulm14-17)

"extra large" servers

1	64-core, 1024GB RAM, 9TB HDD, 1TB SSD (cbsuem01)
1	112-core, 1024GB RAM, 12TB HDD (cbsuem02)
1	88-core, 1024GB RAM, 7TB NVMe SSD (cbsuem02)

"gpu-equipped gen2" servers

2	32-core, 256GB RAM, 12TB HDD, 2x nVidia P100 (cbsugpu02-03)

BioHPC Cloud hardware infrastructure

“networked storage: total 1,743 TB available in two volumes

233 TB	Gluster cluster (4 servers)
1,510 TB	Lustre cluster (13 servers)

login machines (cbsulogin, cbsulogin2, cbsulogin3)

3	12-core, 64-128GB RAM

BioHPC servers – big picture

Infrastructure servers

Lustre, Gluster, management

No direct user access
No reservations
Restricted network access

Rental servers

general, medium, large, extra large, GPU

Users must make reservations
Any BioHPC user can make reservation
Local storage cleaned after reservation ends
Cornell campus network only

Hosted servers

Purchased by groups or departments

Users must be added to appropriate groups
Only users approved by PIs can be added
Local storage persistent – no cleaning
Cornell campus network only

Login servers

For data transfers and indirect access

Any BioHPC user can login
No local storage
No computing allowed
Open to the world network

Accessing BioHPC servers from outside Cornell campus

- Only login servers are accessible directly from outside network
  cbsulogin.biohpc.cornell.edu, cbsulogin2.biohpc.cornell.edu, cbsulogin3.biohpc.cornell.edu

- All other BioHPC servers are behind firewall, and NOT accessible from outside

- Preferred method of outside access is Cornell VPN, available for all Cornelians from CIT

- BioHPC users without Cornell Netid should use login servers for outside access.

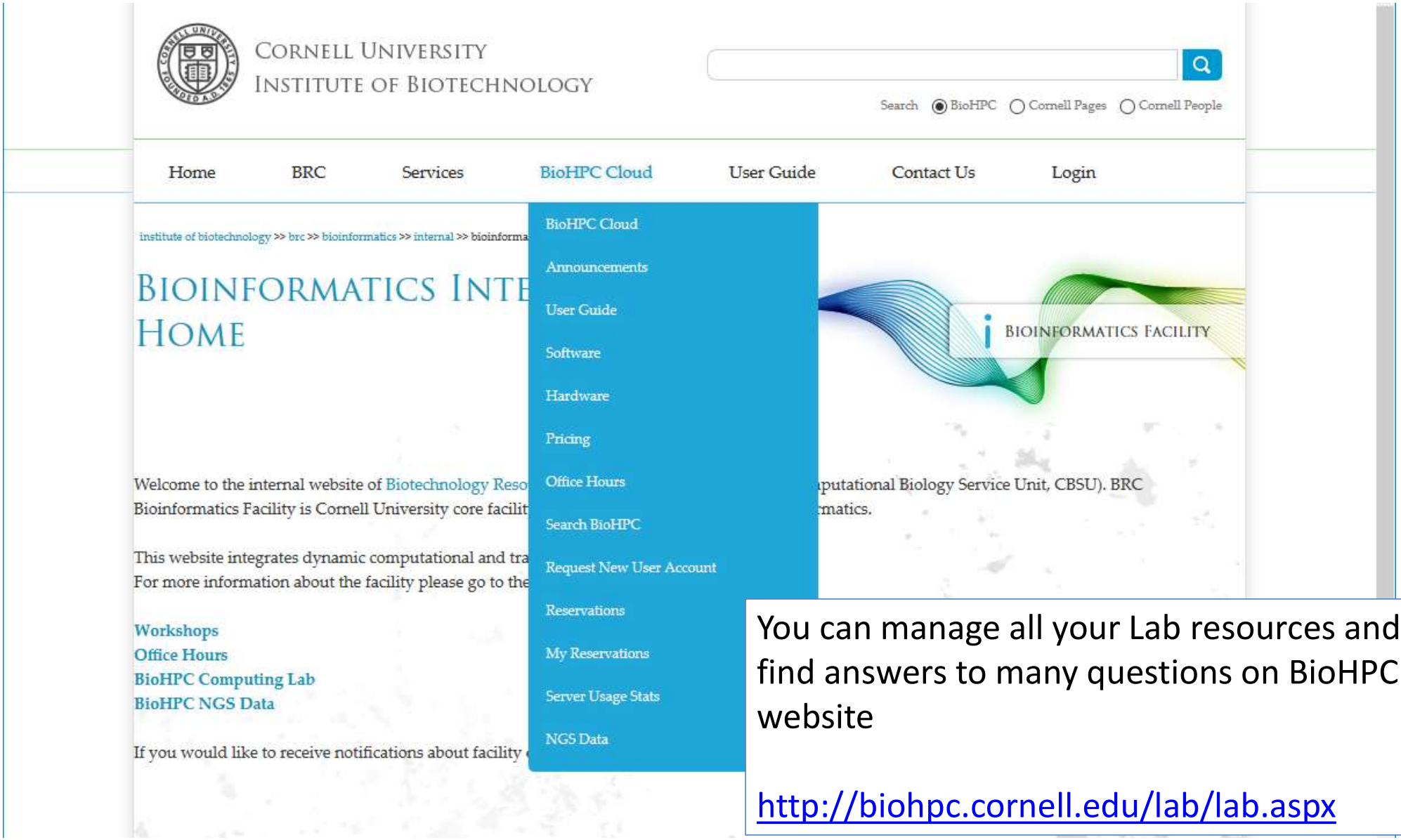
Setting up an account

- In order to get an account fill out account request form online https://biohpc.cornell.edu/NewUserRequest.aspx

- 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 BioHPC overall usage.


Cornell University
Institute of Biotechnology
Search
BioHPC
Cornell Pages
Cornell People
Home
BRC
Services
BioHPC Cloud
User Guide
Contact Us
Login
BioHPC Cloud
Announcements
User Guide
Software
Hardware
Pricing
Office Hours
Search BioHPC
Request New User Account
Reservations
My Reservations
Server Usage Stats
NGS Data
Bioinformatics Facility
Workshops
Office Hours
BioHPC Computing Lab
BioHPC NGS Data
You can manage all your Lab resources and find answers to many questions on BioHPC website
http://biohpc.cornell.edu/lab/lab.aspx

CORNELL UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY

SearchBioHPCCornell Pages ◉ ○ Cornell People ○

Home | BRC | Services | BioHPC Cloud | User Guide | Contact Us | Login

institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc cloud: user guide

BioHPC Cloud:
: User Guide

BIOINFORMATICS FACILITY

Guides | Overview | Quick Start Guide | Accounts | Access | Storage | Storage Space | Backups | Databases | Software | FAQ

Guides

BioHPC Cloud is a com arn Linux operating system, and do bioinfor ervers, storage and software. Below are link about BioHPC can be found in other tabs of

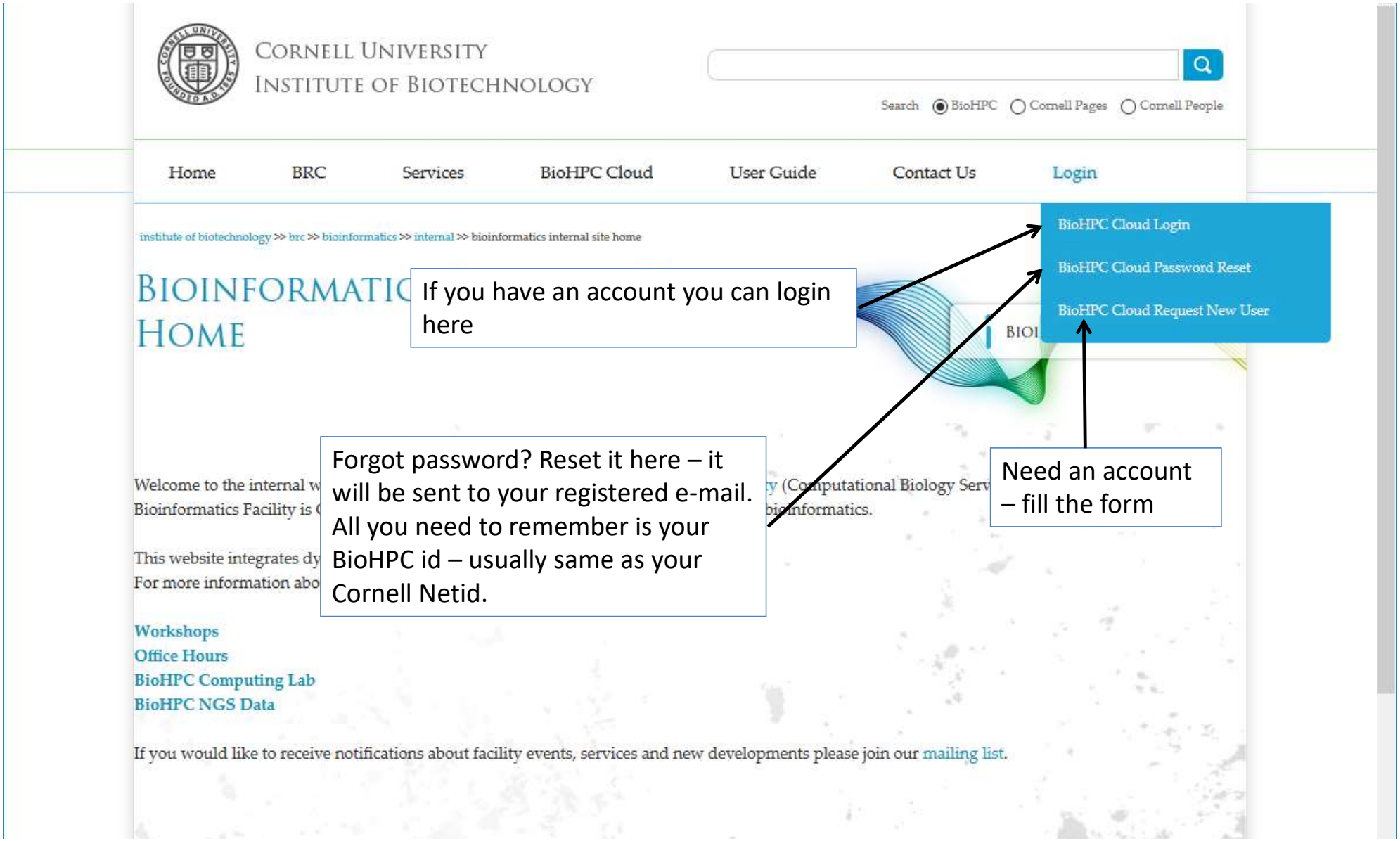
Getting Started with Bi

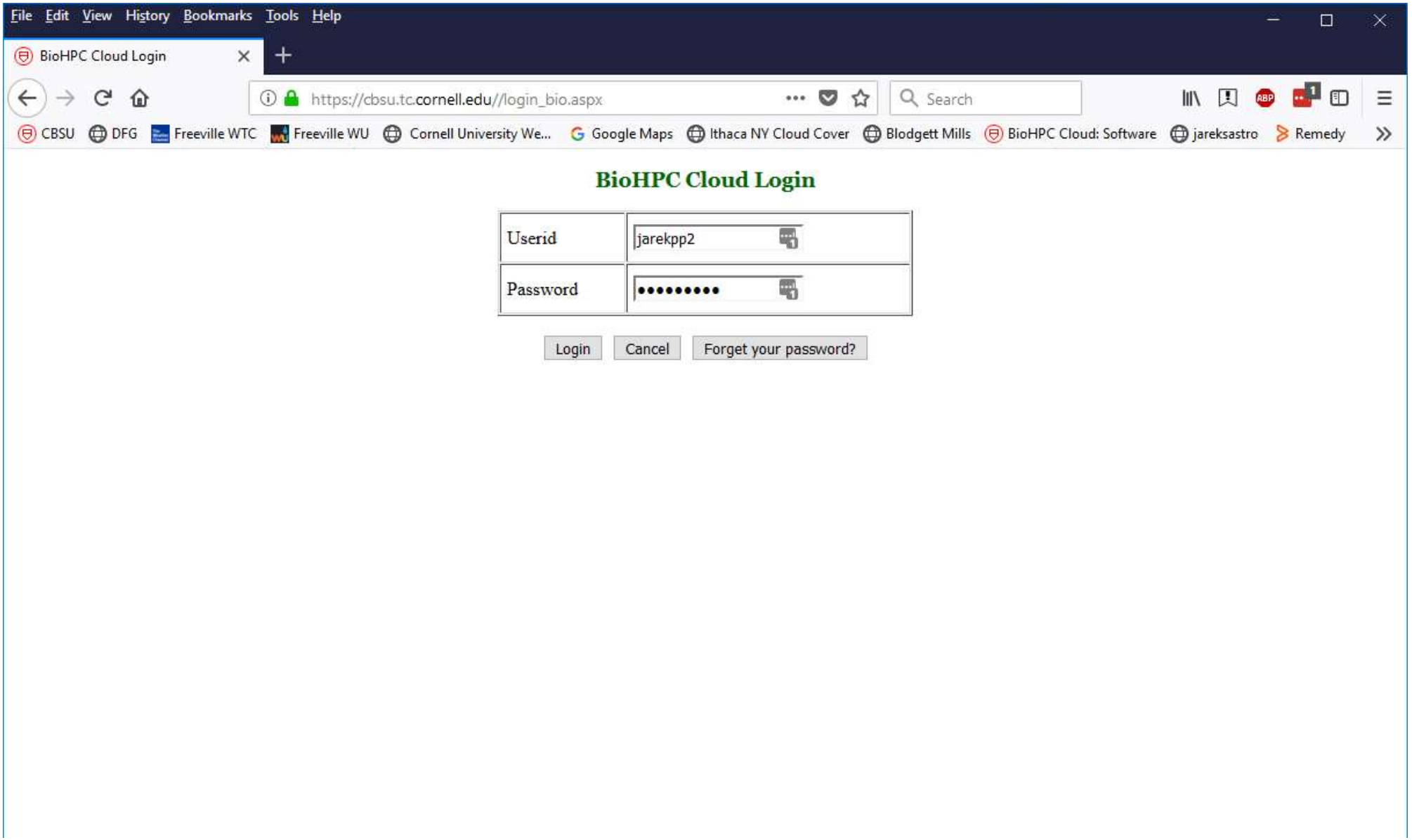
Hosted Servers at Bio

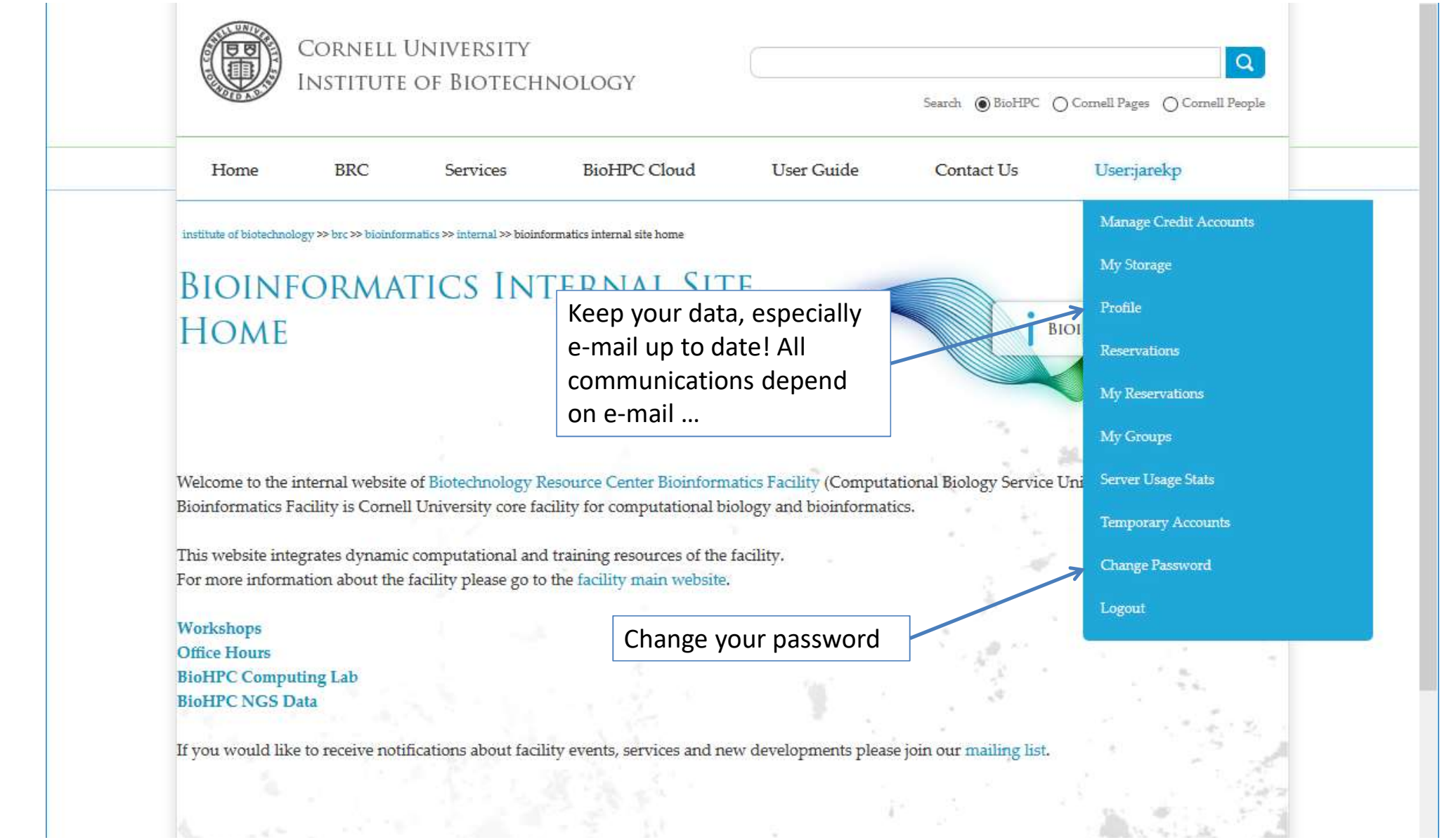
Using and Installing S

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

http://biohpc.cornell.edu/lab/userguide.aspx


Cornell University
Institute of Biotechnology
Search BioHPC Cornell Pages Cornell People
Home BRC Services BioHPC Cloud User Guide Contact Us Login
BioHPC Cloud Login
BioHPC Cloud Password Reset
BioHPC Cloud Request New User
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
Bioinformatic Home
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 BioHPC id – usually same as your Cornell Netid.
Need an account – fill the form
Welcome to the internal w
Bioinformatics Facility is
This website integrates dy
For more information abo
Workshops
Office Hours
BioHPC Computing Lab
BioHPC NGS Data
If you would like to receive notifications about facility events, services and new developments please join our mailing list.


File Edit View History Bookmarks Tools Help
BioHPC Cloud Login
https://cbsu.tc.cornell.edu//login_bio.aspx
Search
CBSU
DFG
Freeville WTC
Freeville WU
Cornell University We...
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
BioHPC Cloud: Software
jareksastro
Remedy
BioHPC Cloud Login
Userid
jarekpp2
Password
Login
Cancel
Forget your password?


Cornell University
Institute of Biotechnology
Search BioHPC Cornell Pages Cornell People
Home BRC Services BioHPC Cloud User Guide Contact Us User:jarekp
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
Bioinformatics Internal Site Home
Keep your data, especially e-mail up to date! All communications depend on e-mail …
Manage Credit Accounts
My Storage
Profile
Reservations
My Reservations
My Groups
Server Usage Stats
Temporary Accounts
Change Password
Logout
Change your password
Welcome to the internal website of Biotechnology Resource Center Bioinformatics Facility (Computational Biology Service Uni
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 NGS Data
If you would like to receive notifications about facility events, services and new developments please join our mailing list.

Getting started with a new account

- You need hours: create and fund your own Lab Credit Account or get added to one. Or get access to a hosted server(s).
- Get extra storage if needed – all users get free storage allocations, but it may not be enough
- Verify that your software is available and read instructions
- Transfer data to your Lab storage
- Make reservation(s)
- Connect to reserved workstations
- *Compute!*

Credit Accounts

In order to reserve workstations you need to have hours available in your BioHPC Credit Account, or you need to have access to hosted server(s).

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

If your group already has a 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 Credit Account and buy hours using Cornell Account or a credit card.

Lab Credit Accounts

Any server can be reserved using any Lab Credit Account.

The hours from Lab Credit Account, called “purchase hours” will be used at different speed depending on server type

http://biohpc.cornell.edu/Lab/Pricing.aspx

Credit Accounts

CLASS	EQUIVALENT TO
purchase hours	1.000
general	3.052
medium gen1	1.960
large gen1	1.078
gpu-equipped gen1	1.674
extra large	0.750
large gen2	0.833
medium gen2	1.691
gpu-equipped gen2	0.990

What is the best server type to use?

Depends very much on the task at hand!

You need large memory when doing assembly and similar tasks.

However, vast majority of tasks can be done using our **medium memory gen2** machines, at the same time they are quite affordable presenting the best price/performance ratio. Best overall choice!

Credit Accounts

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

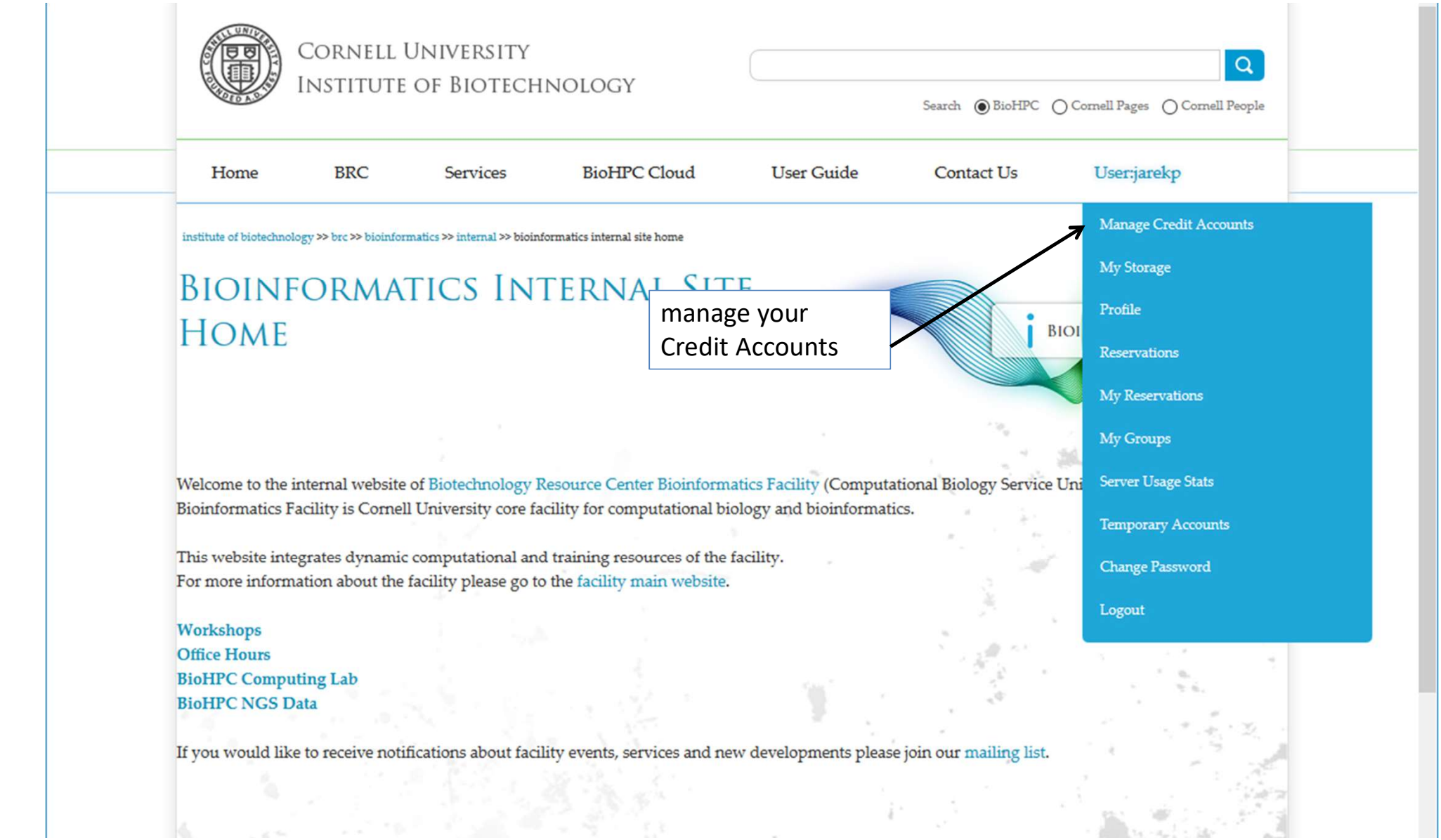
The more time you buy at a time the cheaper it gets.

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

medium gen2

This pricing applies to gen2 medium memory generally accessible workstations, i.e. cbsumm21-32 machines (40 cores; 256GB RAM; 8TB HDD).

purchase unit	current unit	hours	unit cost (Cornell)	cost per hour (Cornell)		unit cost (external)	cost per hour (external)	
				server	core		server	core
60 hours	101.5 hours	101.5 hours	$90.23	$0.89	$0.02 (40 cores)	$112.79	$1.11	$0.03 (40 cores)
200 hours	338.2 hours	338.2 hours	$300.78	$0.89	$0.02 (40 cores)	$375.97	$1.11	$0.03 (40 cores)
1 month	1.7 months	1,234.6 hours	$878.28	$0.71	$0.02 (40 cores)	$1,097.85	$0.89	$0.02 (40 cores)
6 months	10.1 months	7,407.5 hours	$4,215.75	$0.57	$0.01 (40 cores)	$5,269.69	$0.71	$0.02 (40 cores)
1 year	1.7 years	14,815.1 hours	$6,745.20	$0.46	$0.01 (40 cores)	$8,431.50	$0.57	$0.01 (40 cores)


CORNELL UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY
Search BioHPC Cornell Pages Cornell People
Home BRC Services BioHPC Cloud User Guide Contact Us User:jarekp
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
BIOINFORMATICS INTERNAL SITE HOME
manage your Credit Accounts
Manage Credit Accounts
My Storage
Profile
Reservations
My Reservations
My Groups
Server Usage Stats
Temporary Accounts
Change Password
Logout
Welcome to the internal website of Biotechnology Resource Center Bioinformatics Facility (Computational Biology Service Uni
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 NGS Data
If you would like to receive notifications about facility events, services and new developments please join our mailing list.

BioHPC Cloud:
: My Lab Credit Accounts

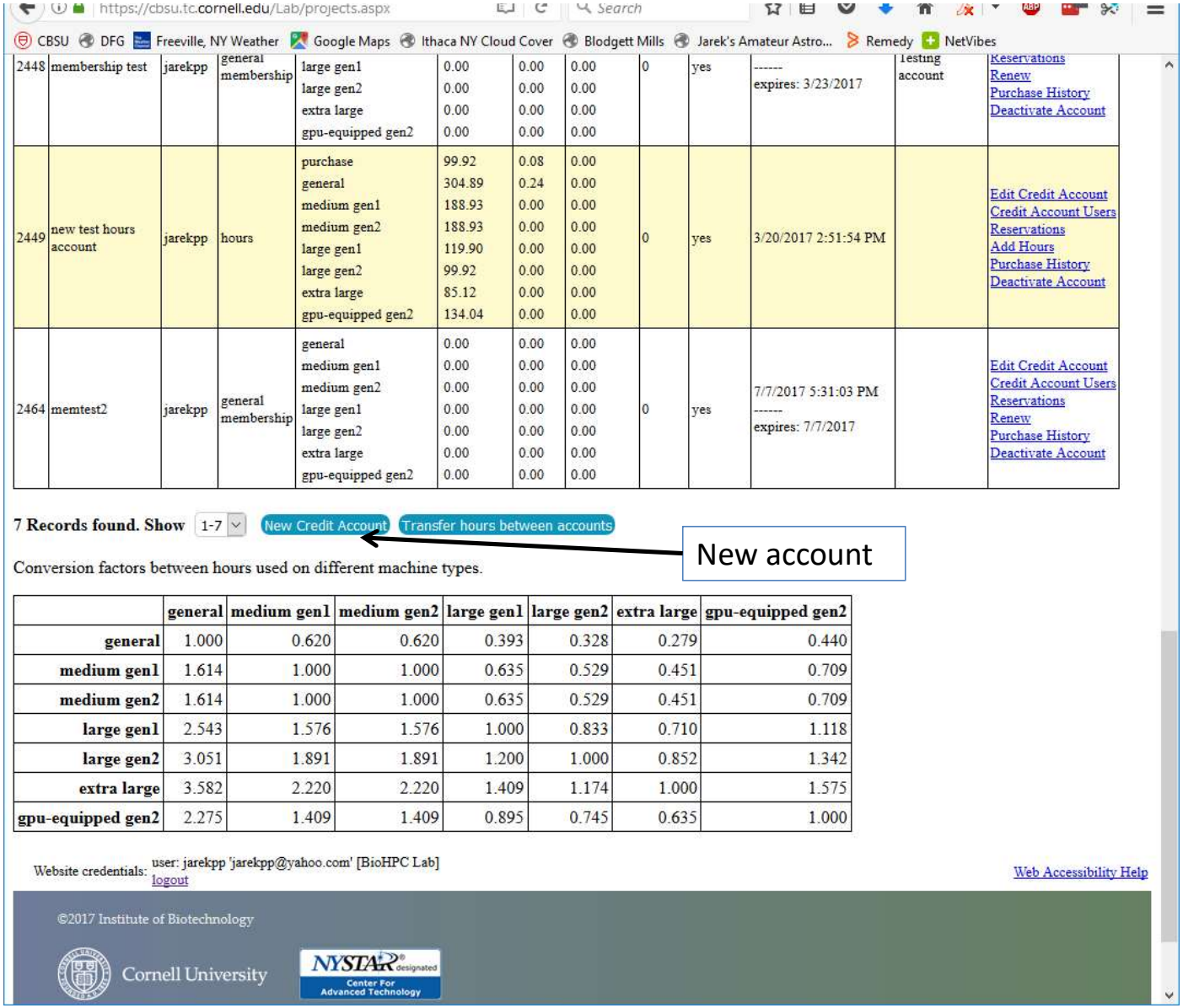
Lab Credit Accounts can be now used for all machines, there is no need to create an account for each type of machine. Conversion factors between hours used on different machine types are in the table at the bottom of the page. Each type of hours are listed for each Lab Credit Account.

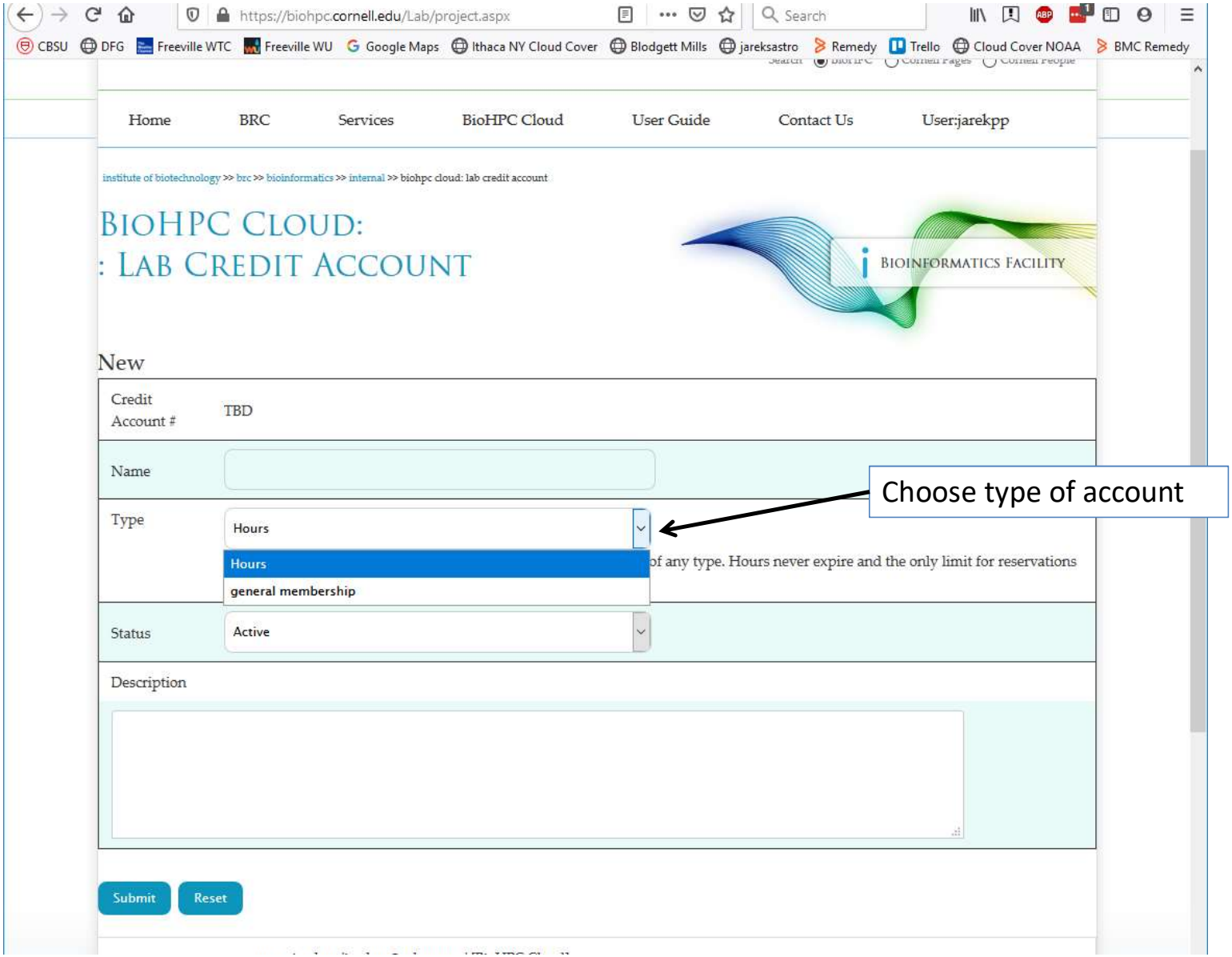
Filter by:

Name: * go Description: * go

Order by Created Ascending

#	Account Name	Owner	Type	Machine Class	Hours Available	Hours Used	Hours Reserved	Other Users	Active?	Created	Description	Action
5	CBSU	jarekp	hours	purchase general medium gen1 medium gen2 large gen1 large gen2 extra large gpu-equipped gen2	INF	192155.00 0.00 2278.50 0.00 0.00 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00	3	yes	11/28/2011 4:34:45 PM	CBSU staff unlimited access for restricted machines	Reservations Add Hours
7	jarekpp_general_	jarekpp	hours	purchase general medium gen1 medium gen2 large gen1 large gen2 extra large gpu-equipped gen2	440.61 1344.64 863.75 745.17 475.18 367.06 330.40 436.35	7930.88 12480.90 2150.97 554.98 1902.05 192.52 280.09 46.96	0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00	3	yes	11/29/2011 2:49:15 PM		Edit Credit Account Credit Account Users Reservations Add Hours Purchase History Deactivate Account


https://cbsu.tc.cornell.edu/Lab/projects.aspx
Search
CBSU
DFG
Freeville, NY Weather
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
Jarek's Amateur Astro...
Remedy
NetVibes
2448 membership test jarekpp general membership large gen1 large gen2 extra large gpu-equipped gen2 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0 yes ------ expires: 3/23/2017 Testing account Reservations Renew Purchase History Deactivate Account
2449 new test hours account jarekpp hours purchase general medium gen1 medium gen2 large gen1 large gen2 extra large gpu-equipped gen2 99.92 304.89 188.93 188.93 119.90 99.92 85.12 134.04 0.08 0.24 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0 yes 3/20/2017 2:51:54 PM Edit Credit Account Credit Account Users Reservations Add Hours Purchase History Deactivate Account
2464 memtest2 jarekpp general membership general medium gen1 medium gen2 large gen1 large gen2 extra large gpu-equipped gen2 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0 yes 7/7/2017 5:31:03 PM ------ expires: 7/7/2017 Edit Credit Account Credit Account Users Reservations Renew Purchase History Deactivate Account
7 Records found. Show 1-7
New Credit Account
Transfer hours between accounts
New account
Conversion factors between hours used on different machine types.
general medium gen1 medium gen2 large gen1 large gen2 extra large gpu-equipped gen2
general 1.000 0.620 0.620 0.393 0.328 0.279 0.440
medium gen1 1.614 1.000 1.000 0.635 0.529 0.451 0.709
medium gen2 1.614 1.000 1.000 0.635 0.529 0.451 0.709
large gen1 2.543 1.576 1.576 1.000 0.833 0.710 1.118
large gen2 3.051 1.891 1.891 1.200 1.000 0.852 1.342
extra large 3.582 2.220 2.220 1.409 1.174 1.000 1.575
gpu-equipped gen2 2.275 1.409 1.409 0.895 0.745 0.635 1.000
Website credentials: user: jarekpp 'jarekpp@yahoo.com' [BioHPC Lab]
logout
Web Accessibility Help
©2017 Institute of Biotechnology
Cornell University
NYSTAR designated Center For Advanced Technology

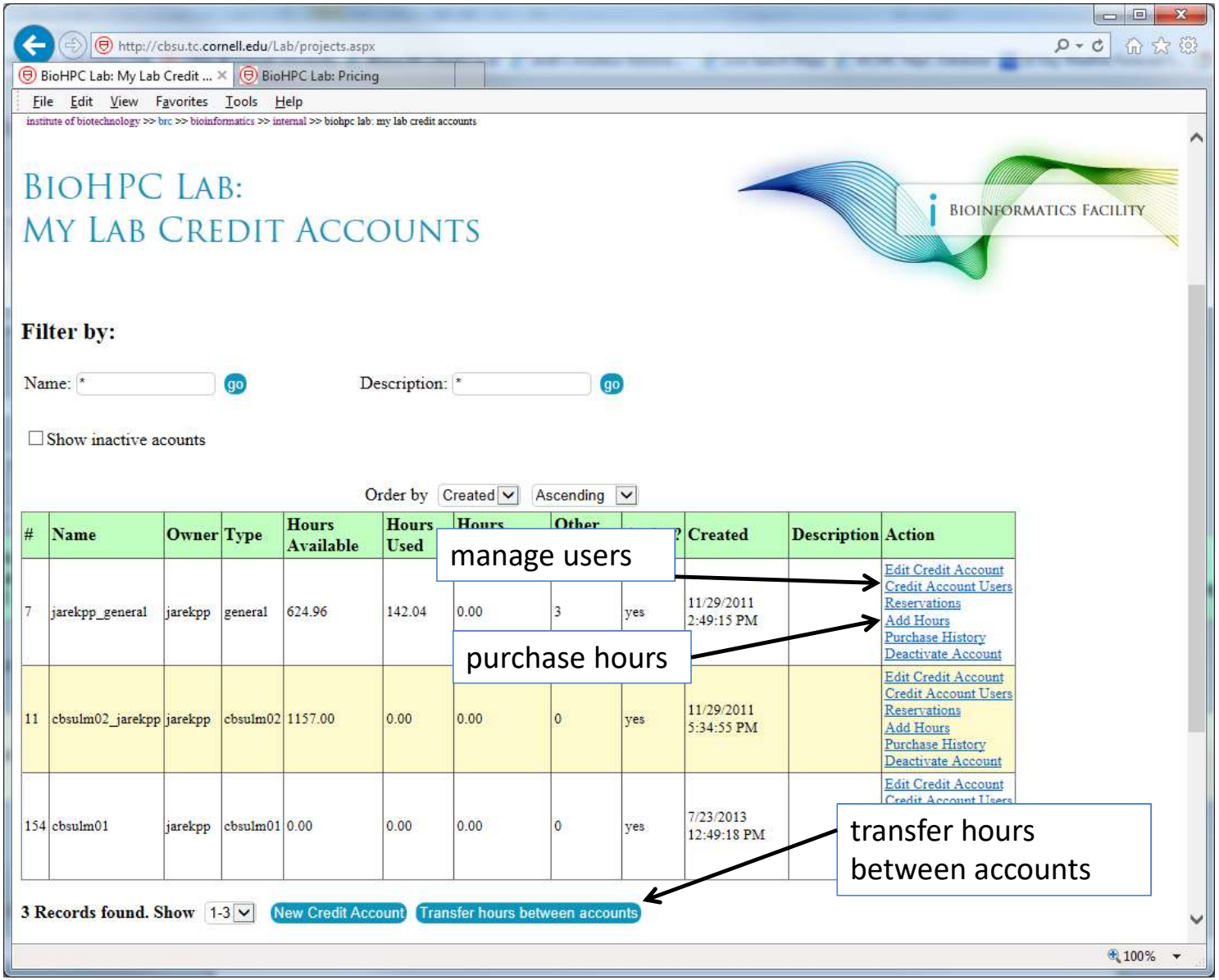

https://biohpc.cornell.edu/Lab/project.aspx
Search
CBSU
DFG
Freeville WTC
Freeville WU
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
jareksastro
Remedy
Trello
Cloud Cover NOAA
BMC Remedy
Home
BRC
Services
BioHPC Cloud
User Guide
Contact Us
User:jarekpp
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc cloud: lab credit account
BioHPC Cloud:
: Lab Credit Account
Bioinformatics Facility
New
Credit Account #
TBD
Name
Type
Hours
Hours
general membership
of any type. Hours never expire and the only limit for reservations
Status
Active
Description
Submit
Reset
Choose type of account

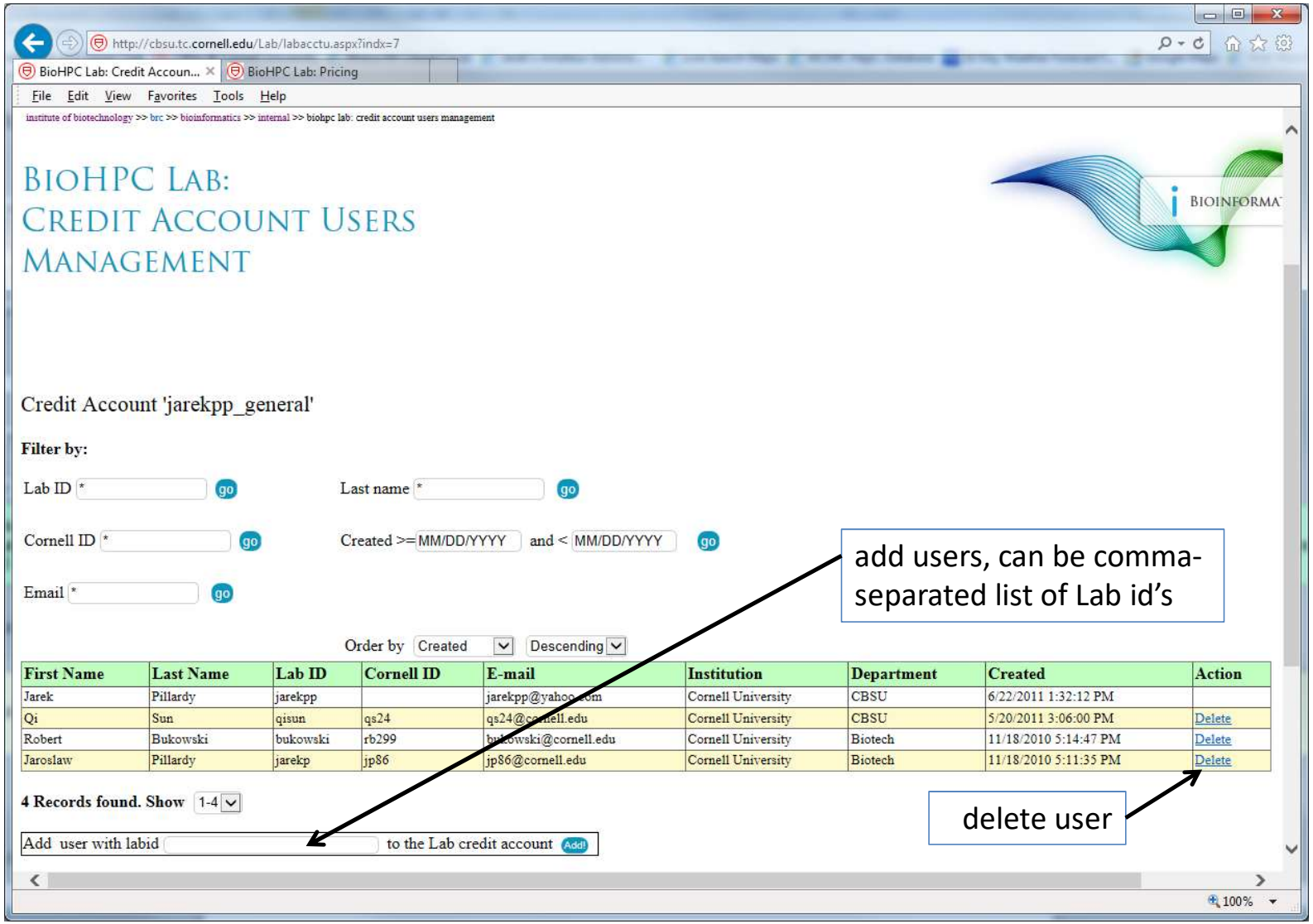
Credit Accounts

Membership: general membership

This membership allows you to use continuously the types of machines listed in the table below with some limits. The limits are 'rolling' which means that they have to be met an any given time, but get updated over time. For example if you make the maximum length reservation now, you will be able to extend it by one day tomorrow at the same time.

machine	maximum reservation length (rolling)		maximum combined reservation (rolling)		annual cost (Cornell)	annual cost (external)
	unit	hours	unit	hours		
general	1.0 weeks	168.0	1.0 weeks	168.0	$500.00	$600.00
medium gen1	1.0 days	24.0	1.0 days	24.0		


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
3 Records found. Show 1-3
New Credit Account
Transfer hours between accounts
manage users
purchase hours
transfer hours between accounts
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
BIOINFORMA
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%

CORNELL UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY

search ◉ Cornell Pages ○ Cornell People

Home | BRC | Services | BioHPC Lab | User Guide | 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

FROM:	'jarekpp_general_ 149.61 hrs left	100	hours
TO:	'jarekpp_general_ 149.61 hrs left	100.0 hours	

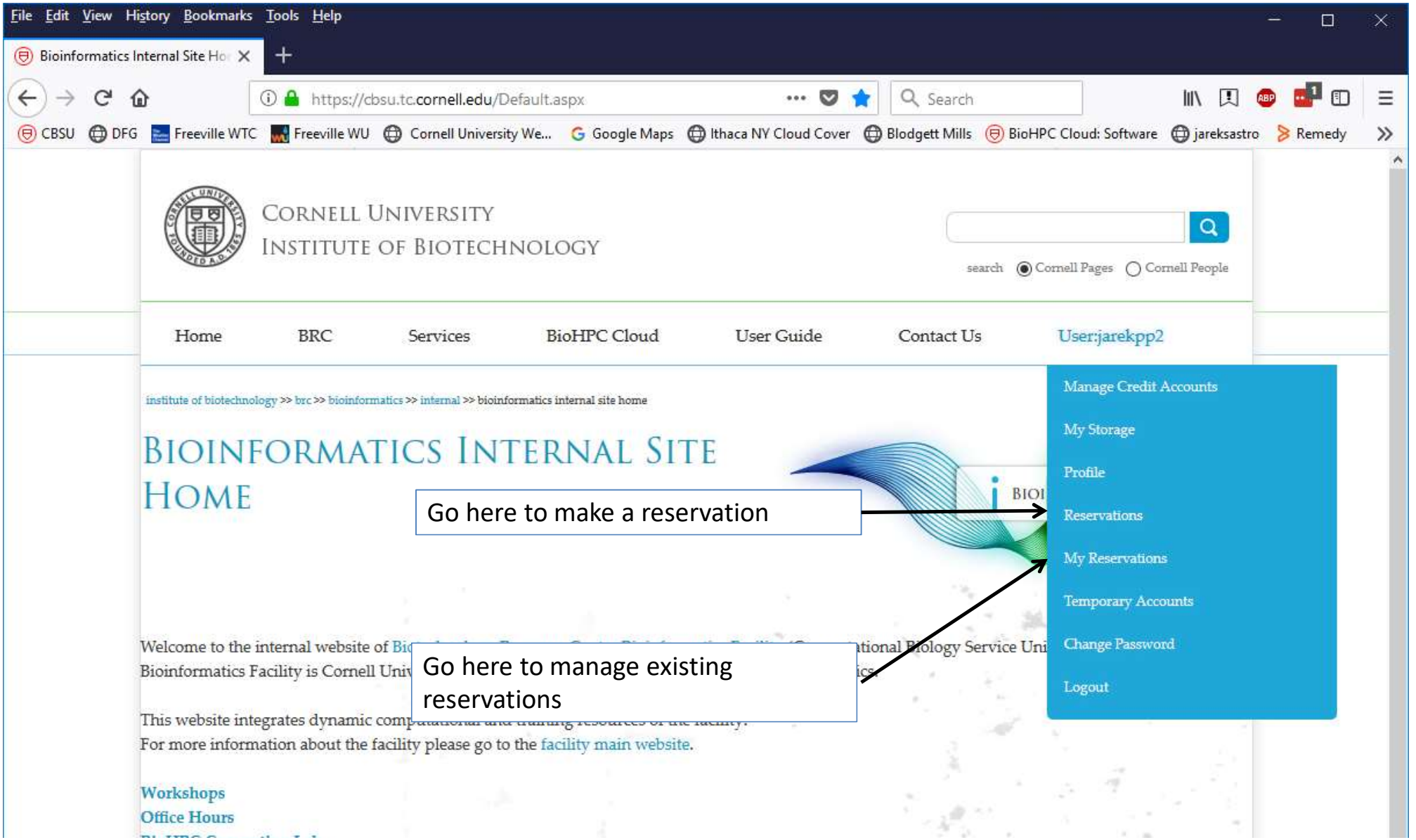
Transfer

Website credentials: user: jarekpp 'jarekpp@yahoo.com' [BioHPC Lab]
logout

Web Accessibility Help

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
- Verify that your software is available and read instructions
- Transfer data to your Lab storage
- Make reservation(s)
- Connect to reserved workstations
- *Compute!*


File Edit View History Bookmarks Tools Help
Bioinformatics Internal Site Ho
https://cbsu.tc.cornell.edu/Default.aspx
Search
CBSU DFG Freeville WTC Freeville WU Cornell University We... Google Maps Ithaca NY Cloud Cover Blodgett Mills BioHPC Cloud: Software jareksastro Remedy
CORNELL UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY
search Cornell Pages Cornell People
Home BRC Services BioHPC Cloud User Guide Contact Us User:jarekpp2
Manage Credit Accounts
My Storage
Profile
Reservations
My Reservations
Temporary Accounts
Change Password
Logout
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
BIOINFORMATICS INTERNAL SITE HOME
Go here to make a reservation
Go here to manage existing reservations
Welcome to the internal website of
Bioinformatics Facility is Cornell
This website integrates dynamic
For more information about the facility please go to the facility main website.
Workshops
Office Hours

CORNELL UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY

Home | BRC | Services | BioHPC Cloud | User Guide | Contact Us | User:jarekp

institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc cloud: reservations

BioHPC Cloud:
: Reservations

choose type of machines

My Reservations

medium gen1: 24 cores, 128GB RAM

These workstations can be only accessed remotely via an ssh client.
Total of 17 workstations available, 3 enhanced security

Display reservations from February 4, 2019 for 30 days Go!

New reservation from February 4, 2019 02 : 36 pm to February 4, 2019 02 : 36 pm for machine first available with Lab credit account 'BioMG7810' owner:jarekp 16.3hrs 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

	cbsulm01 Linux (CentOS 7.4) Dell PowerEdge R710 16 cores; 64GB RAM; 1TB HDD; VM supported AVX support: none	**cbsumm01** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported (**enhanced security**) AVX support: AVX	**cbsumm02** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX	**cbsumm03** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported (**enhanced security**) AVX support: AVX	**cbsumm04** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX	**cbsumm05** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX	**cbsumm06** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX
Mon Feb 04 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	yah6 11:58 AM - EOD	qx32 01:25 PM - 02:06 PM AVAILABLE	pds28 10:05 AM - 12:41 PM AVAILABLE	tad94 ALL DAY
Tue Feb 05 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	yah6 UNTIL 11:58 AM AVAILABLE	AVAILABLE	AVAILABLE	tad94 UNTIL 10:24 PM AVAILABLE
Wed Feb 06 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	AVAILABLE	AVAILABLE	AVAILABLE	AVAILABLE
Thu Feb 07 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	AVAILABLE	AVAILABLE	AVAILABLE	AVAILABLE
Fri Feb 08 2019	AVAILABLE	ncg37 ALL DAY	bukowski UNTIL 02:47 PM AVAILABLE	AVAILABLE	AVAILABLE	AVAILABLE	AVAILABLE
Sat Feb 09 2019	AVAILABLE	ncg37 ALL DAY	AVAILABLE	AVAILABLE	AVAILABLE	AVAILABLE	AVAILABLE

BioHPC Cloud:
: Reservations

My Reservations

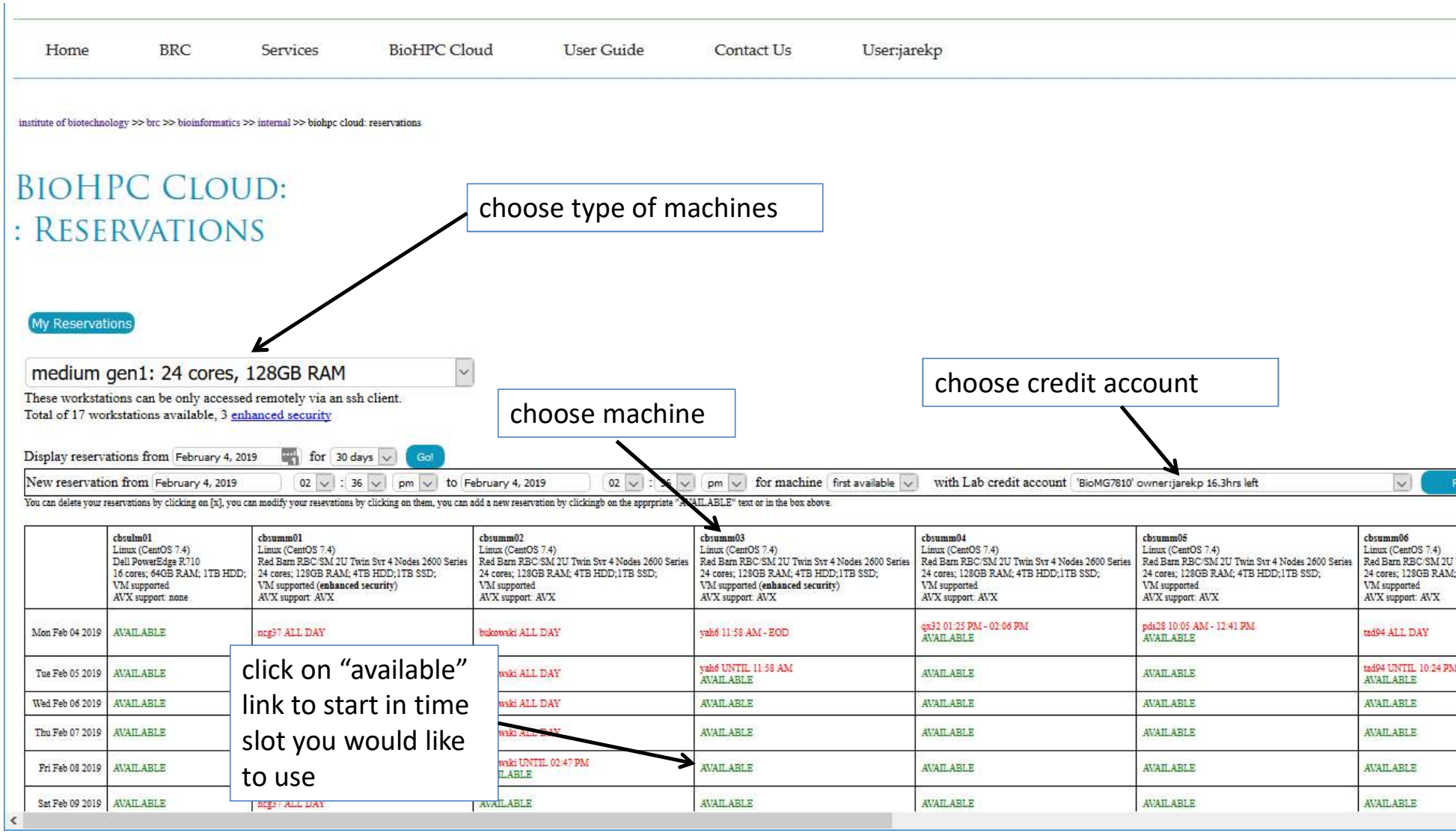
medium gen1: 24 cores, 128GB RAM

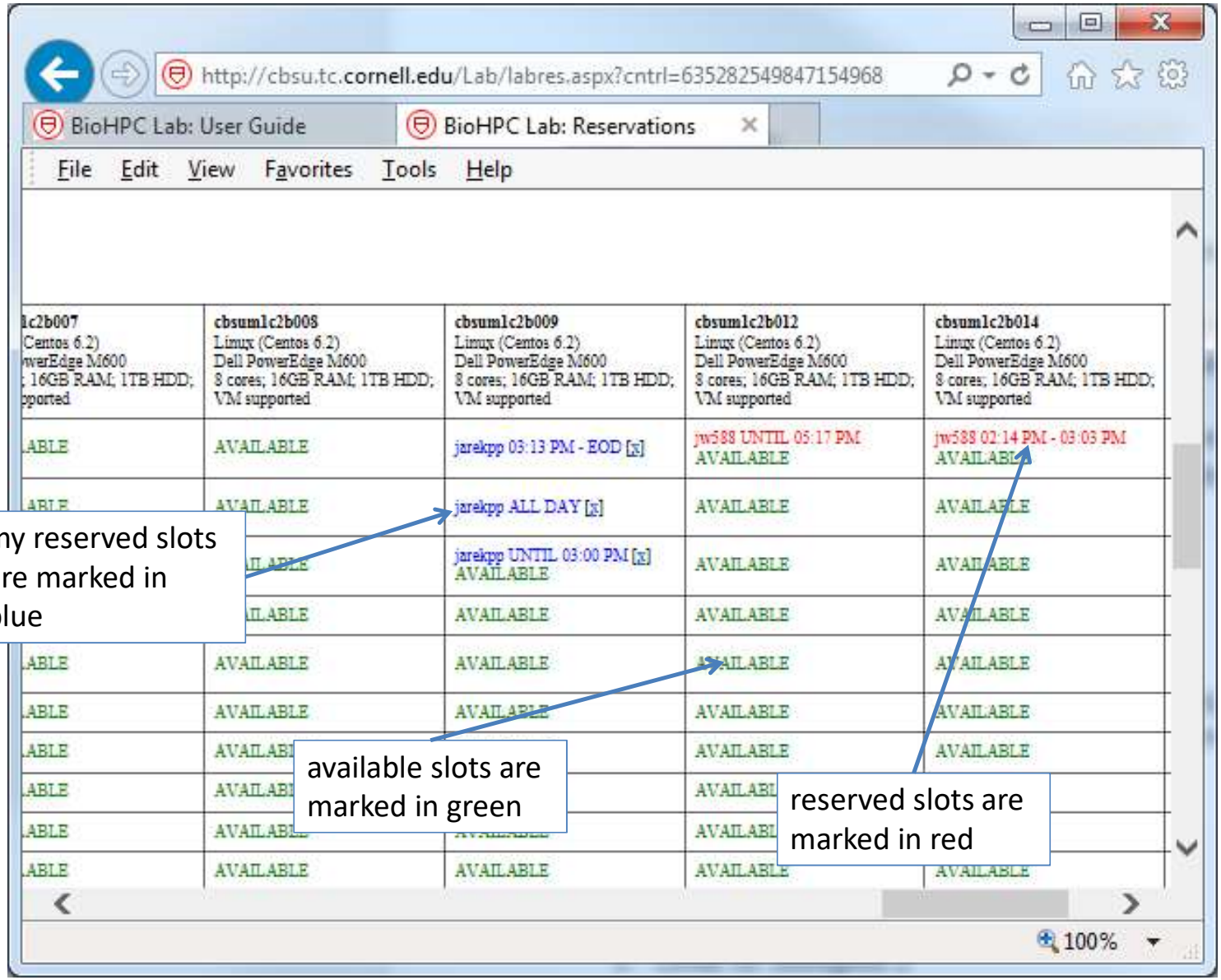
- interactive: 4 cores, 24GB RAM
- general: 8 cores, 16GB RAM
- medium gen1: 24 cores, 128GB RAM
- medium gen2: 40 cores, 256GB RAM
- large memory gen1: 64 cores, 512GB RAM
- large memory gen2: 96-112 cores, 512GB RAM
- extra large memory: 96-112 cores, 1024GB RAM
- gpu gen2: 32 cores, 256GB RAM, 2x NVidia P100
- database computing
- restricted

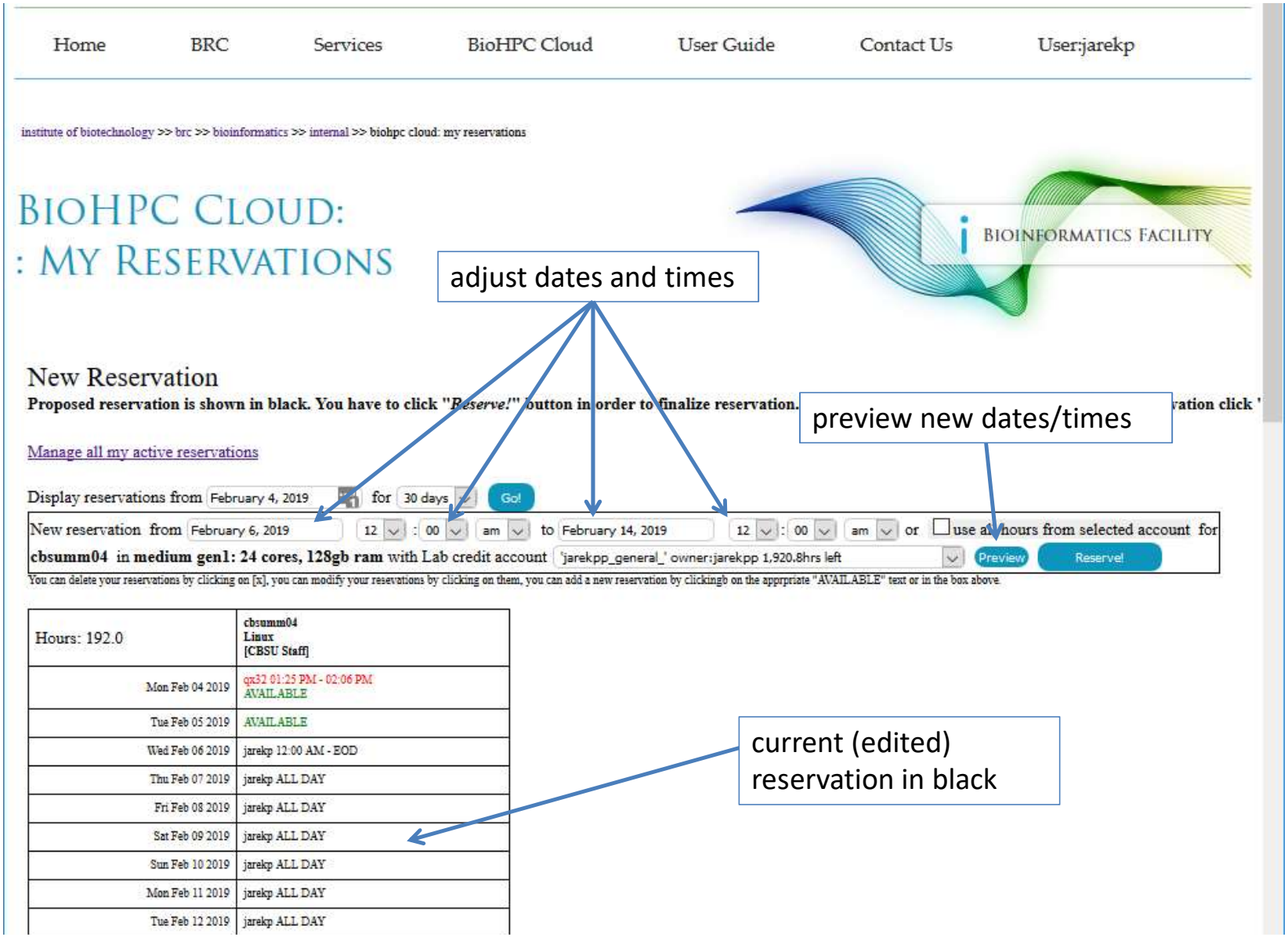
ebruary 4, 2019 | 02 : 36 | pm | for machine | first available | with Lab credit account | 'BioMG7810' owner:jarekp 16.3hrs left

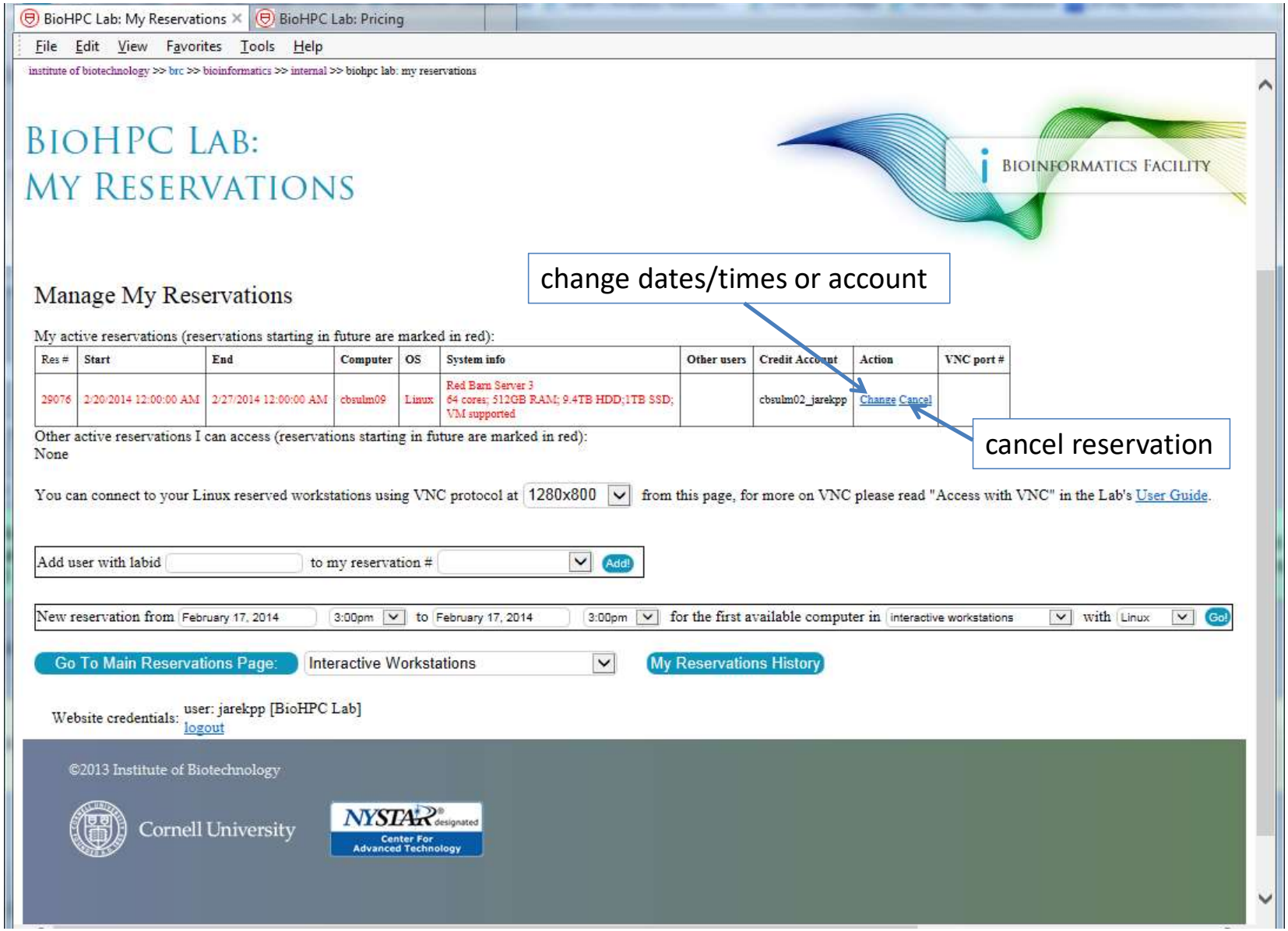
dd a new reservation by clickingb on the apprpriate "AVAILABLE" text or in the box above.

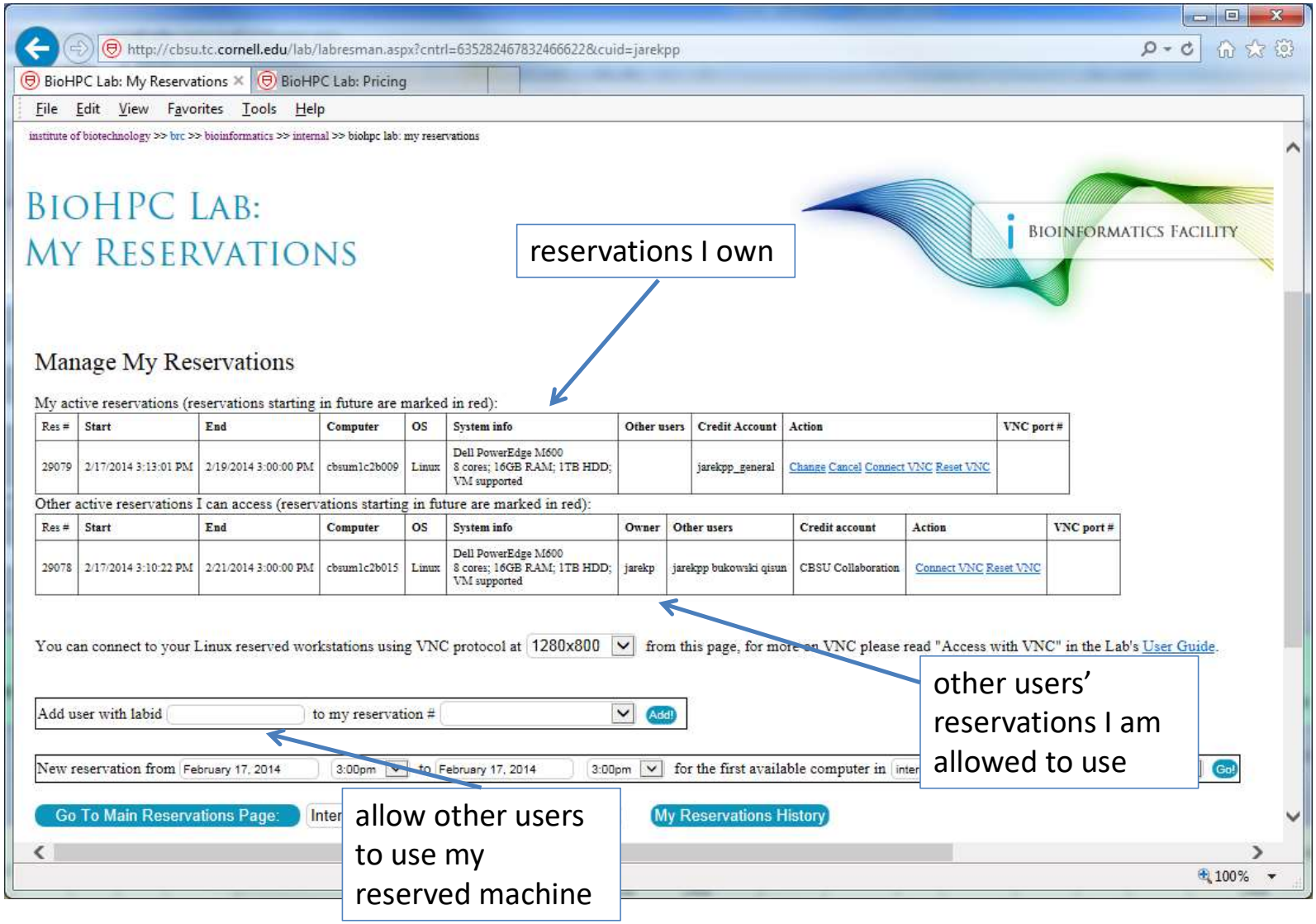
			cbsumm02 Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX	**cbsumm03** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported (**enhanced security**) AVX support: AVX	**cbsumm04** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX	**cbsumm05** Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX	**cbs** Lin Red 24 VM AV
Mon Feb 04 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	yah6 11:58 AM - EOD	qx32 01:25 PM - 02:06 PM AVAILABLE	pds28 10:05 AM - 12:41 PM AVAILABLE	tzd
Tue Feb 05 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	yah6 UNTIL 11:58 AM AVAILABLE	AVAILABLE	AVAILABLE	tzd AV
Wed Feb 06 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	AVAILABLE	AVAILABLE	AVAILABLE	AV
Thu Feb 07 2019	AVAILABLE	ncg37 ALL DAY	bukowski ALL DAY	AVAILABLE	AVAILABLE	AVAILABLE	AV
Fri Feb 08 2019	AVAILABLE	ncg37 ALL DAY	bukowski UNTIL 02:47 PM AVAILABLE	AVAILABLE	AVAILABLE	AVAILABLE	AV


Home
BRC
Services
BioHPC Cloud
User Guide
Contact Us
User:jarekp
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc cloud: reservations
BioHPC Cloud:
: Reservations
choose type of machines
My Reservations
medium gen1: 24 cores, 128GB RAM
These workstations can be only accessed remotely via an ssh client.
Total of 17 workstations available, 3 enhanced security
choose machine
choose credit account
Display reservations from February 4, 2019 for 30 days Go!
New reservation from February 4, 2019 02 : 36 pm to February 4, 2019 02 : 36 pm for machine first available with Lab credit account 'BioMG7810' owner:jarekp 16.3hrs left
You can delete your reservations by clicking on [x], you can modify your resevations by clicking on them, you can add a new reservation by clickingb on the apprpriate "AVAILABLE" text or in the box above.
cbsulm01 Linux (CentOS 7.4) Dell PowerEdge R710 16 cores; 64GB RAM; 1TB HDD; VM supported AVX support: none
cbsumm01 Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported (enhanced security) AVX support: AVX
cbsumm02 Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX
cbsumm03 Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported (enhanced security) AVX support: AVX
cbsumm04 Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX
cbsumm05 Linux (CentOS 7.4) Red Barn RBC/SM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM; 4TB HDD;1TB SSD; VM supported AVX support: AVX
cbsumm06 Linux (CentOS 7.4) Red Barn RBC/SM 2U 24 cores; 128GB RAM; VM supported AVX support: AVX
Mon Feb 04 2019 AVAILABLE ncg37 ALL DAY bukowski ALL DAY yah6 11:58 AM - EOD qx32 01:25 PM - 02:06 PM AVAILABLE pds28 10:05 AM - 12:41 PM AVAILABLE tad94 ALL DAY
Tue Feb 05 2019 AVAILABLE wski ALL DAY yah6 UNTIL 11:58 AM AVAILABLE AVAILABLE AVAILABLE tad94 UNTIL 10:24 PM AVAILABLE
Wed Feb 06 2019 AVAILABLE wski ALL DAY AVAILABLE AVAILABLE AVAILABLE AVAILABLE
Thu Feb 07 2019 AVAILABLE AVAILABLE AVAILABLE AVAILABLE AVAILABLE
Fri Feb 08 2019 AVAILABLE wski UNTIL 02:47 PM LABLE AVAILABLE AVAILABLE AVAILABLE AVAILABLE
Sat Feb 09 2019 AVAILABLE ncg37 ALL DAY AVAILABLE AVAILABLE AVAILABLE AVAILABLE AVAILABLE
click on “available” link to start in time slot you would like to use


http://cbsu.tc.cornell.edu/Lab/labres.aspx?cntrl=635282549847154968
BioHPC Lab: User Guide
BioHPC Lab: Reservations
File Edit View Favorites Tools Help
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
AVAILABLE
jarekpp 03:13 PM - EOD [x]
jarekpp ALL DAY [x]
jarekpp UNTIL 03:00 PM [x]
jw588 UNTIL 05:17 PM
jw588 02:14 PM - 03:03 PM
100%
my reserved slots are marked in blue
available slots are marked in green
reserved slots are marked in red


Home
BRC
Services
BioHPC Cloud
User Guide
Contact Us
User:jarekp
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc cloud: my reservations
BioHPC Cloud:
: My Reservations
BIOINFORMATICS FACILITY
adjust dates and times
preview new dates/times
New Reservation
Proposed reservation is shown in black. You have to click "Reserve!" button in order to finalize reservation.
Manage all my active reservations
Display reservations from February 4, 2019 for 30 days Go!
New reservation from February 6, 2019 12 : 00 am to February 14, 2019 12 : 00 am or use all hours from selected account for
cbsumm04 in medium gen1: 24 cores, 128gb ram with Lab credit account 'jarekpp_general_' owner:jarekpp 1,920.8hrs 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
Hours: 192.0
cbsumm04
Linux
[CBSU Staff]
Mon Feb 04 2019
qx32 01:25 PM - 02:06 PM
AVAILABLE
Tue Feb 05 2019
AVAILABLE
Wed Feb 06 2019
jarekp 12:00 AM - EOD
Thu Feb 07 2019
jarekp ALL DAY
Fri Feb 08 2019
jarekp ALL DAY
Sat Feb 09 2019
jarekp ALL DAY
Sun Feb 10 2019
jarekp ALL DAY
Mon Feb 11 2019
jarekp ALL DAY
Tue Feb 12 2019
jarekp ALL DAY
current (edited) reservation in black


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


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.

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 for 2 hours – 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. Same thing will happen after 2 hours past reservation

Reservations

What if I am running a program and need to end the reservation immediately when the program ends?

- Run the program form a *script* – more about scripting on "Linux for Biologists".
- After a line with program name add the following command */programs/bin/labutils/endres.pl*

```
#!/bin/bash
/home/myid/myprogram [options]
/programs/bin/labutils/endres.pl
```

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
- Verify that your software is available and read instructions
- Transfer data to your Lab storage
- ✖ Make reservation(s)
- Connect to reserved workstations
- *Compute!*

Storage

- **Networked storage**

  Very large storage (1,510TB+233TB), 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 (/).

Each file or directory has a unique **path** starting from root (/):

/home/jarekp/tmp/tmpfile

/programs/bin/labutils/endres.pl

Storage

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

/home/jarekp

/usr/local

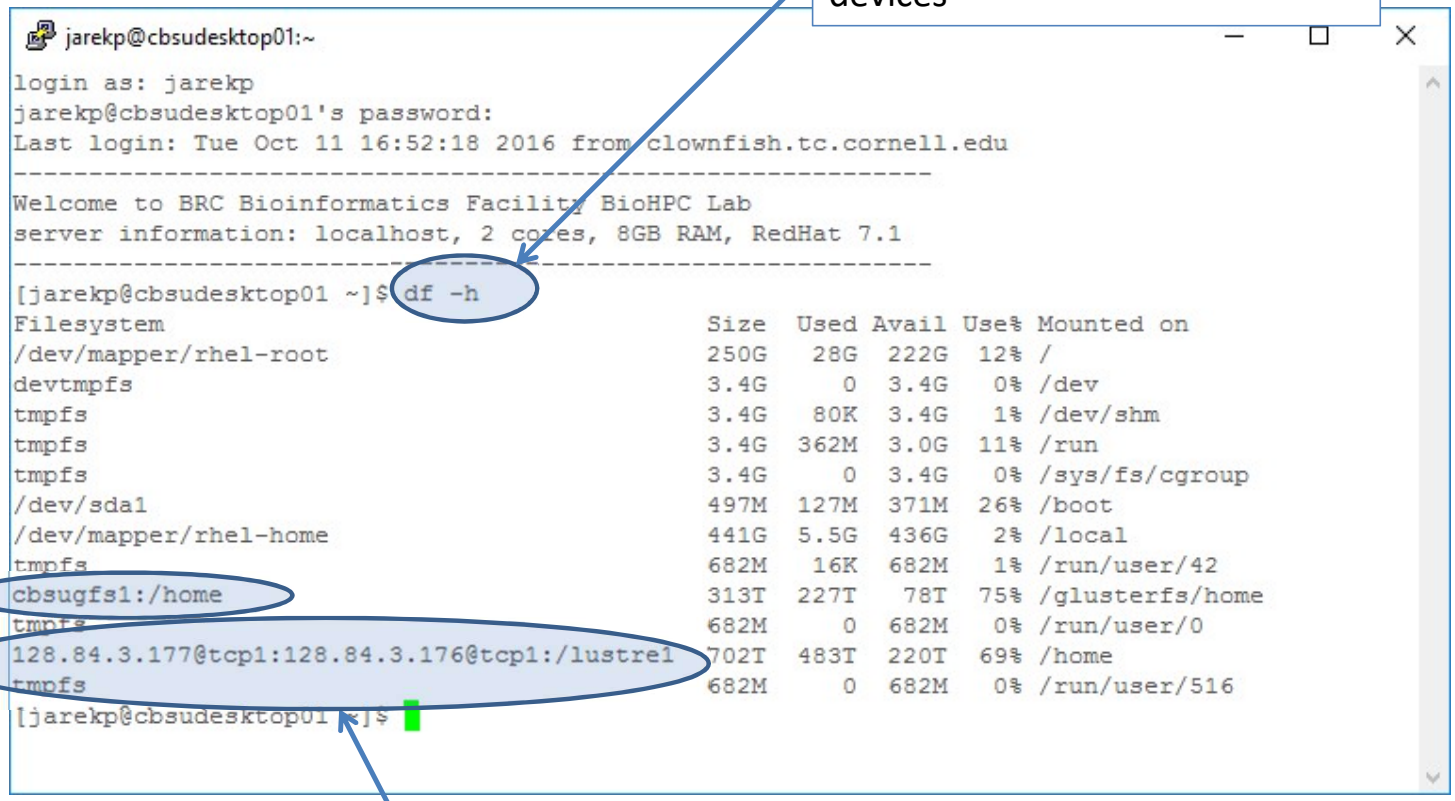
/workdir/jarekp

/shared_data

/local_data

Storage

If in doubt “df -h” lists all storage devices



These are network devices – starting with “computername:/”

Storage

- **Networked storage**
  /home
  /shared_data
  /programs

- **Local storage**
  /workdir
  /SSD
  /local

/workdir full

What if my local storage /workdir is full on my reserved server?

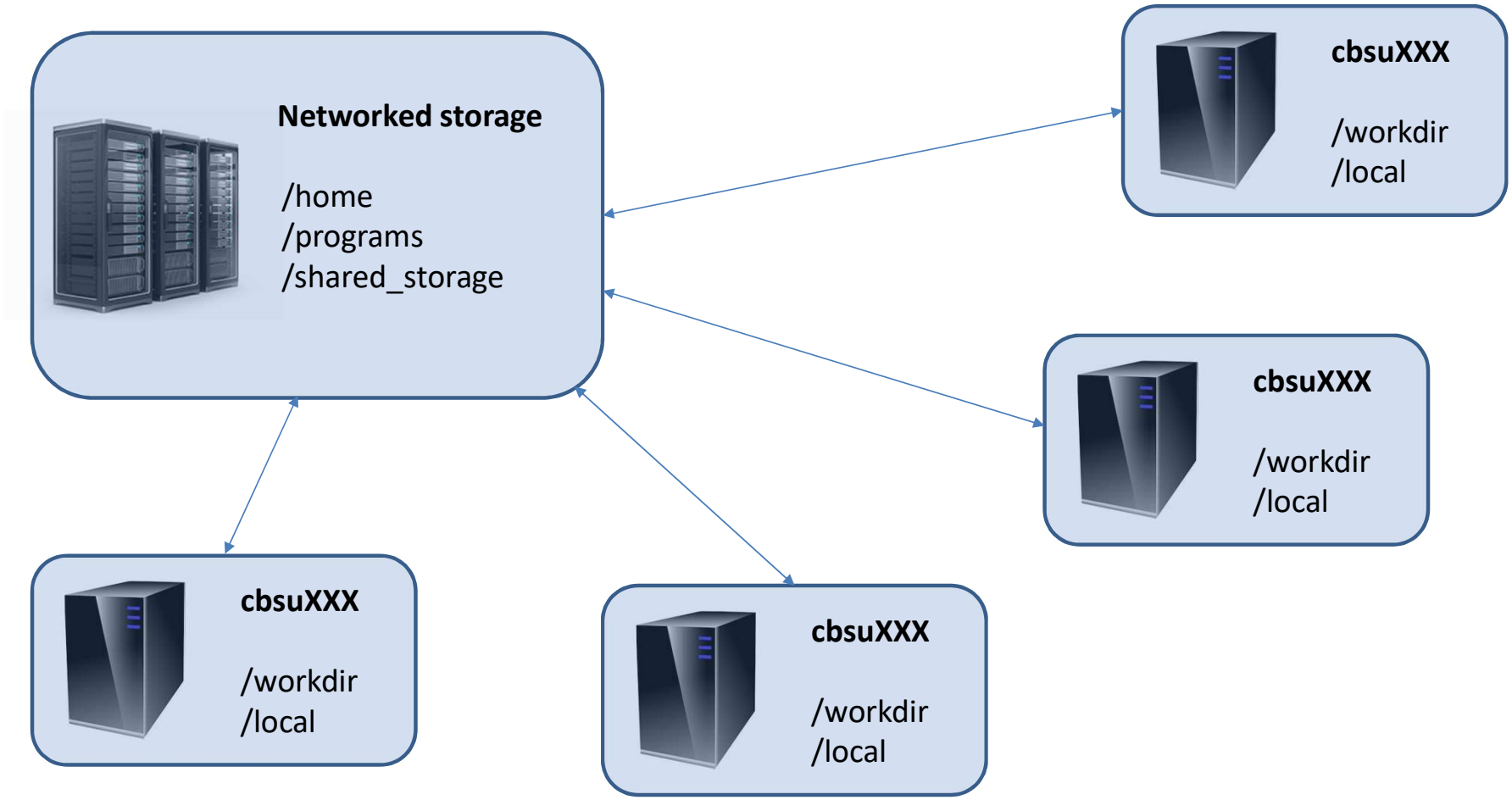
It should not happen, every night at 3am a script is trying to clean it, but if there is a lot of local storage available it will try to avoid deleting until after 5 days past reservation.

You can run this script yourself:

```
/programs/config/clean_workdir
```

This way 5 day rule will be ignored and all data from previous reservations will be deleted.

Storage – big picture



Home directories

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

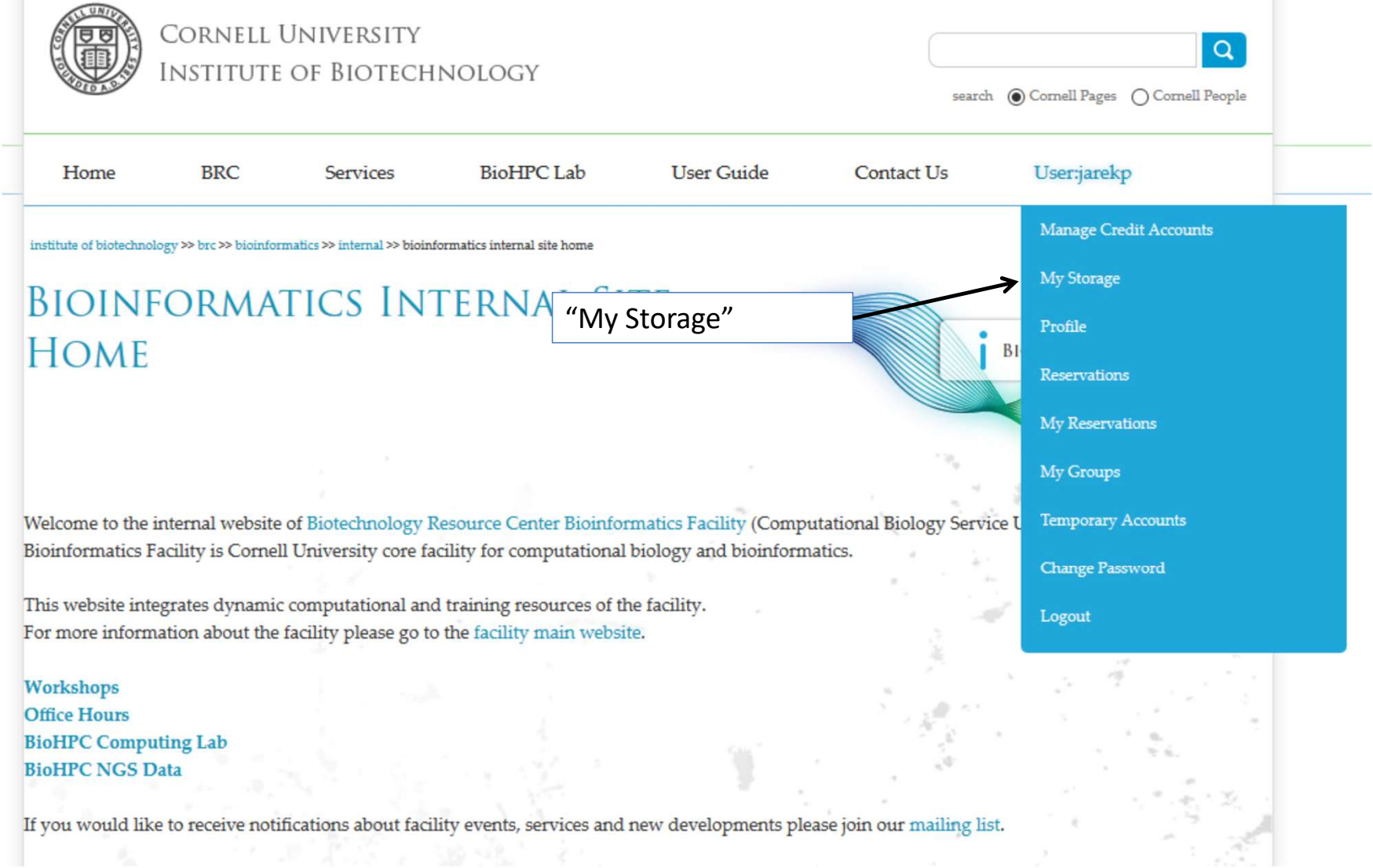
All home directories on BioHPC Cloud are networked and they are the same on all servers.

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 during the night.

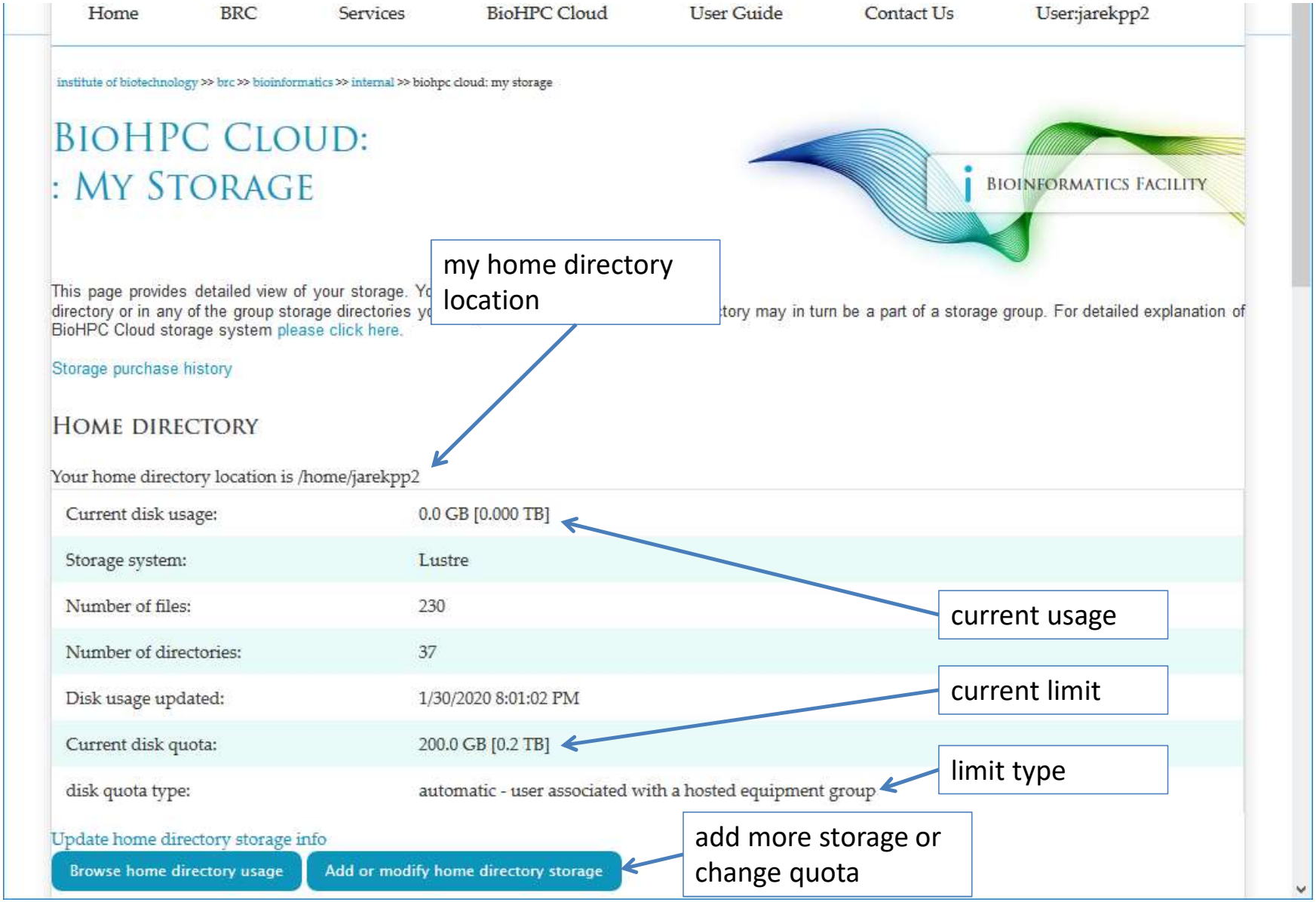

Cornell University
Institute of Biotechnology
search
Cornell Pages
Cornell People
Home
BRC
Services
BioHPC Lab
User Guide
Contact Us
User:jarekp
Manage Credit Accounts
My Storage
Profile
Reservations
My Reservations
My Groups
Temporary Accounts
Change Password
Logout
institute of biotechnology >> brc >> bioinformatics >> internal >> bioinformatics internal site home
Bioinformatics Internal Site Home
"My Storage"
Welcome to the internal website of Biotechnology Resource Center Bioinformatics Facility (Computational Biology Service U
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 NGS Data
If you would like to receive notifications about facility events, services and new developments please join our mailing list.

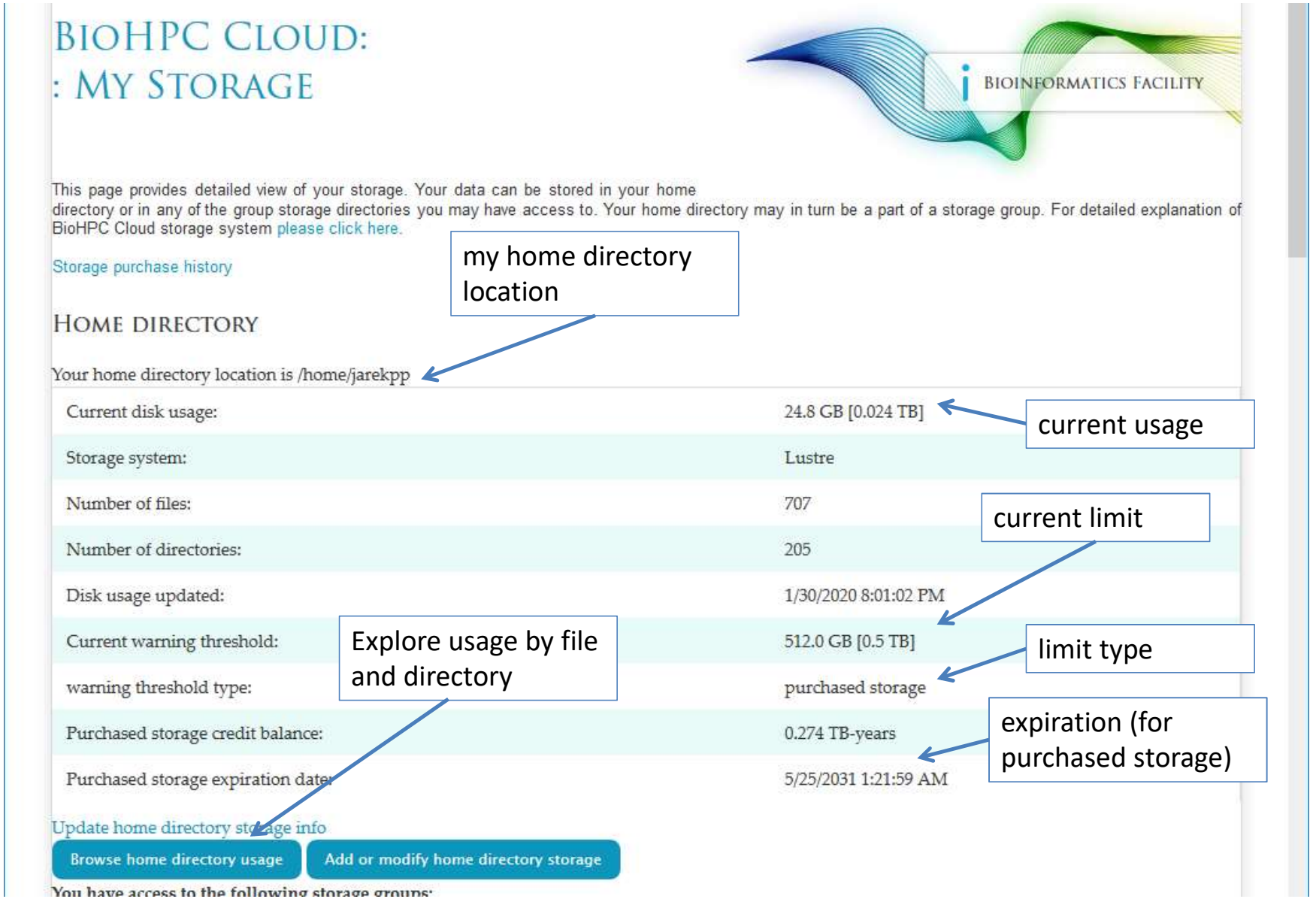
Home storage quotas

If 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.


Home
BRC
Services
BioHPC Cloud
User Guide
Contact Us
User:jarekpp2
institute of biotechnology >> brc >> bioinformatics >> internal >> biohpc cloud: my storage
BioHPC Cloud:
: My Storage
Bioinformatics Facility
my home directory location
This page provides detailed view of your storage. Yo
directory or in any of the group storage directories yo
tory may in turn be a part of a storage group. For detailed explanation of
BioHPC Cloud storage system please click here.
Storage purchase history
Home Directory
Your home directory location is /home/jarekpp2
Current disk usage: 0.0 GB [0.000 TB]
Storage system: Lustre
Number of files: 230
Number of directories: 37
Disk usage updated: 1/30/2020 8:01:02 PM
Current disk quota: 200.0 GB [0.2 TB]
disk quota type: automatic - user associated with a hosted equipment group
current usage
current limit
limit type
Update home directory storage info
Browse home directory usage
Add or modify home directory storage
add more storage or change quota


BioHPC Cloud:
: 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 Cloud storage system please click here.
Storage purchase history
my home directory location
Home directory
Your home directory location is /home/jarekpp
Current disk usage: 24.8 GB [0.024 TB]
current usage
Storage system: Lustre
Number of files: 707
current limit
Number of directories: 205
Disk usage updated: 1/30/2020 8:01:02 PM
Current warning threshold: 512.0 GB [0.5 TB]
Explore usage by file and directory
limit type
warning threshold type: purchased storage
Purchased storage credit balance: 0.274 TB-years
expiration (for purchased storage)
Purchased storage expiration date: 5/25/2031 1:21:59 AM
Update home directory storage info
Browse home directory usage
Add or modify home directory storage
You have access to the following storage groups:

BioHPC Cloud:
: Storage Usage Browser

/home/jarekp Storage

Show files/directories greater than 0 GB Filter by Size

```
492.7 Gb   /home/jarekp (44498 files / 3013 directories)
  183.2 Gb   scanlon                        (452 / 29)
   87.7 Gb   tmp                            (1781 / 191)
   66.6 Gb   Kerry                          (4 / 1)
     17.7 Gb   2691-ellison_QTL4E_Lp_R2.fastq.gz
     17.2 Gb   2691-ellison_QTL4E_Lp_R1.fastq.gz
     16.2 Gb   2690-Ellison_QTL4E_Lk_R2.fastq.gz
     15.5 Gb   2690-Ellison_QTL4E_Lk_R1.fastq.gz
   47.2 Gb   workdir.copy                   (73 / 3)
   24.7 Gb   perl_bio.2013                  (383 / 26)
   17.0 Gb   Wojtek                         (23 / 7)
   11.7 Gb   mouse                          (41 / 1)
    9.9 Gb   blastdb                        (206 / 7)
    5.2 Gb   wojtek_data                    (85 / 7)
    4.4 Gb   teresa                         (4 / 1)
```

Storage

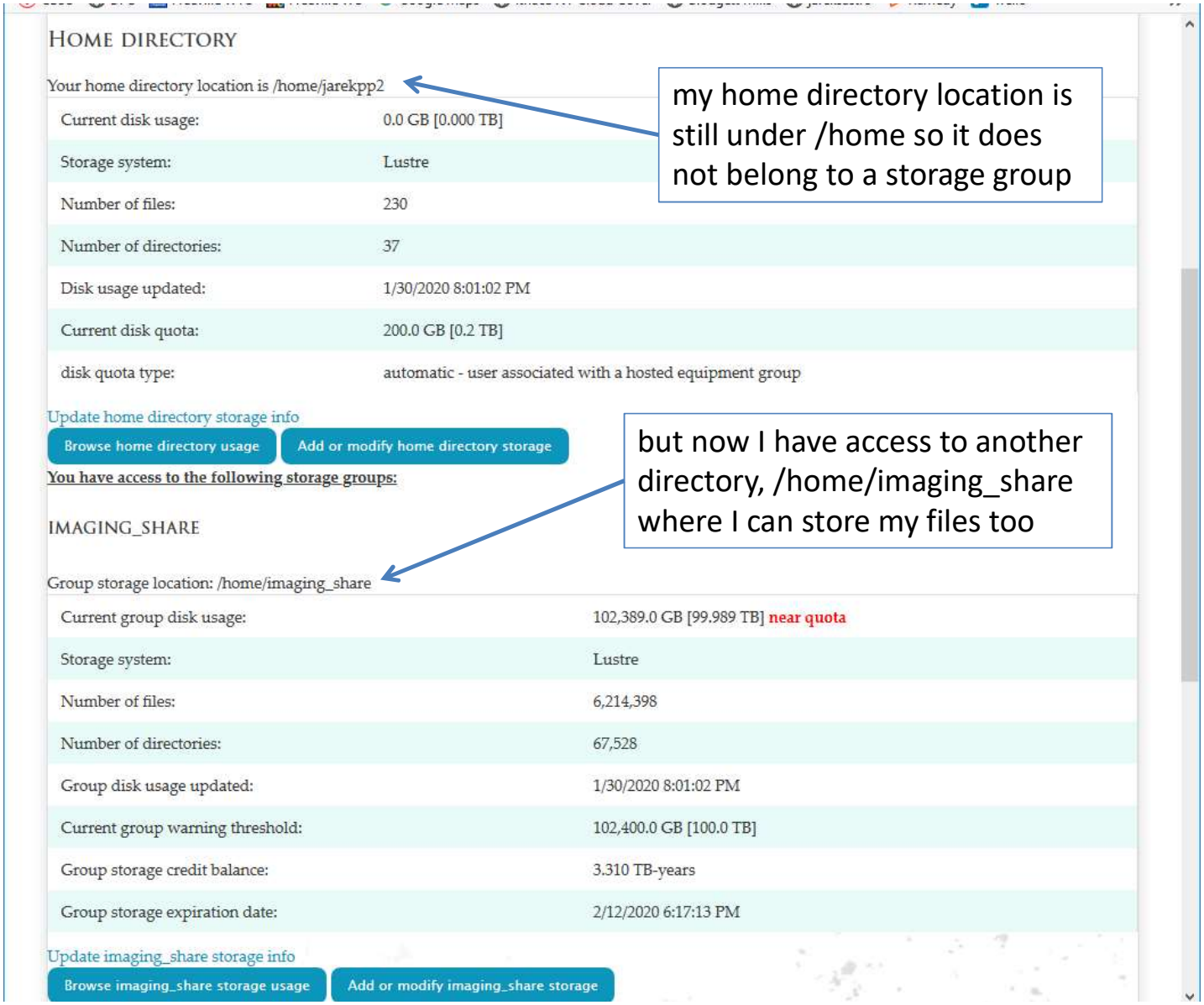
Extra storage can be purchased for $95.01 per TB per year.

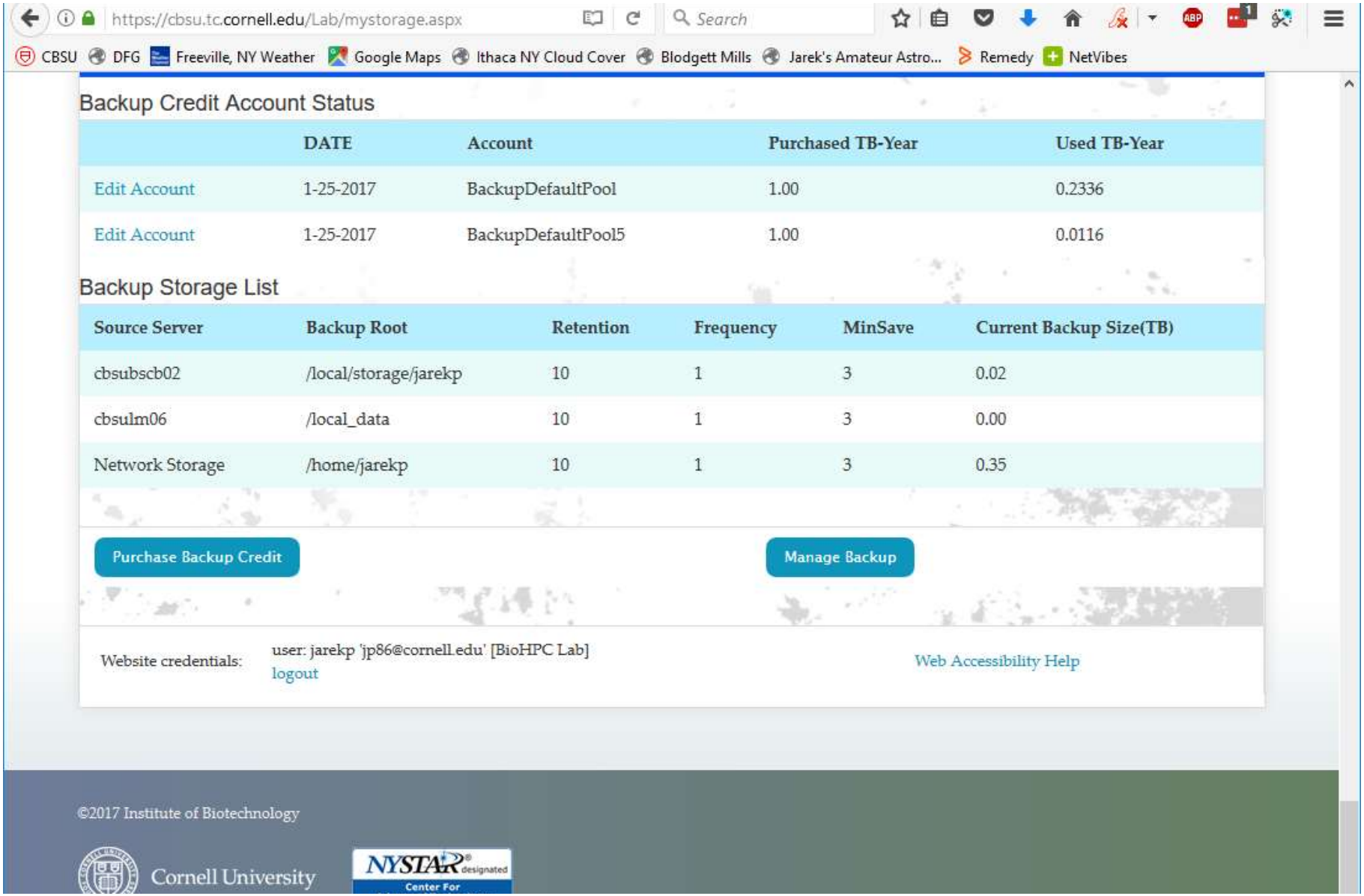
Extra storage can be applied to user's home directory.

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"


HOME DIRECTORY
Your home directory location is /home/jarekpp2
Current disk usage: 0.0 GB [0.000 TB]
Storage system: Lustre
Number of files: 230
Number of directories: 37
Disk usage updated: 1/30/2020 8:01:02 PM
Current disk quota: 200.0 GB [0.2 TB]
disk quota type: automatic - user associated with a hosted equipment group
Update home directory storage info
Browse home directory usage
Add or modify home directory storage
You have access to the following storage groups:
IMAGING_SHARE
Group storage location: /home/imaging_share
Current group disk usage: 102,389.0 GB [99.989 TB] near quota
Storage system: Lustre
Number of files: 6,214,398
Number of directories: 67,528
Group disk usage updated: 1/30/2020 8:01:02 PM
Current group warning threshold: 102,400.0 GB [100.0 TB]
Group storage credit balance: 3.310 TB-years
Group storage expiration date: 2/12/2020 6:17:13 PM
Update imaging_share storage info
Browse imaging_share storage usage
Add or modify imaging_share storage
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/imaging_share where I can store my files too


https://cbsu.tc.cornell.edu/Lab/mystorage.aspx
Search
CBSU
DFG
Freeville, NY Weather
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
Jarek's Amateur Astro...
Remedy
NetVibes
Backup Credit Account Status
DATE
Account
Purchased TB-Year
Used TB-Year
Edit Account
1-25-2017
BackupDefaultPool
1.00
0.2336
Edit Account
1-25-2017
BackupDefaultPool5
1.00
0.0116
Backup Storage List
Source Server
Backup Root
Retention
Frequency
MinSave
Current Backup Size(TB)
cbsubscb02
/local/storage/jarekp
10
1
3
0.02
cbsulm06
/local_data
10
1
3
0.00
Network Storage
/home/jarekp
10
1
3
0.35
Purchase Backup Credit
Manage Backup
Website credentials:
user: jarekp 'jp86@cornell.edu' [BioHPC Lab]
logout
Web Accessibility Help
©2017 Institute of Biotechnology
Cornell University
NYSTAR designated
Center For

BioHPC Cloud:
: Backup Storage

Server: Network Storage ? Backup Account Pool: BackupDefaultPool

Enter Backup Root: Add Directory to Backup ?

Source Server	Backup Root	Retention	Frequency	MinSave	Account			
Network Storage	/home/imaging_share /CURRENT	10	1	2	BackupDefaultPool	Edit	Stop Backup	Manage Excludes
Network Storage	/home/jarekp	14	1	14	BackupDefaultPool	Edit	Stop Backup	Manage Excludes

Website credentials: user: jarekp 'jp86@cornell.edu' [BioHPC Cloud]
logout

Web Accessibility Help

Storage

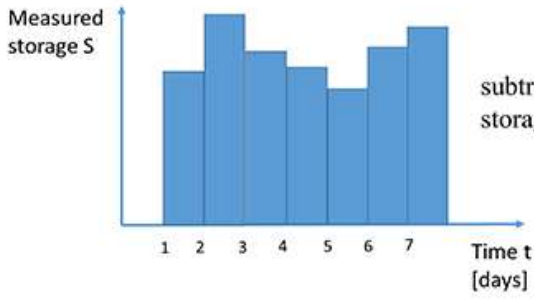
Storage can be only purchased in 1 TB-year chunks, it needs to be done up front, used storage will be subtracted every day, until your storage credit is gone. Expiration date is estimated based on current storage

PAYING FOR STORAGE

Storage is purchased in **1 TB-year increments**

example: 1 TB-year buys 1 TB of storage for a year, or 2 TB for ½ year, or 0.5 TB for 2 years, etc.
usage measured every day and subtracted from the purchased amount

Usage after N measurements (typically – days)



$$U_N = \sum_{i=1}^{N} S_{i-1} \times \Delta t_i$$

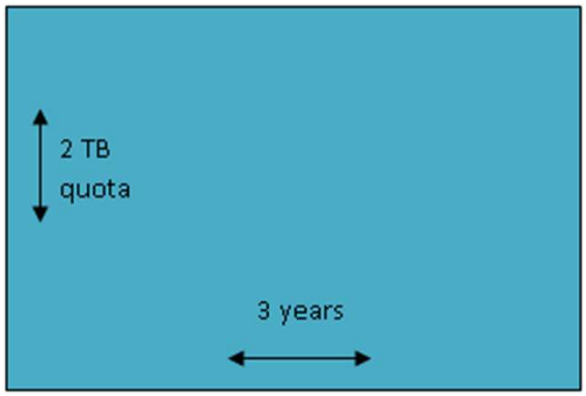
subtracted daily from the originally purchased storage amount, displayed on the website

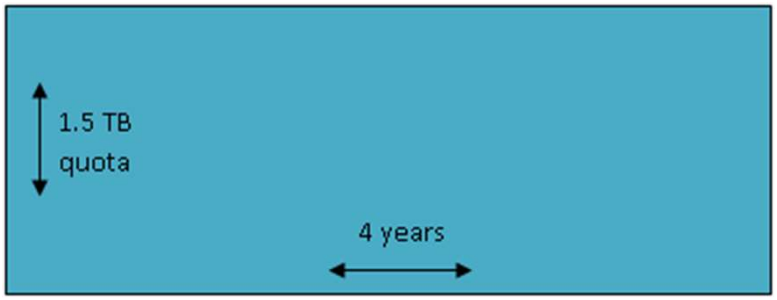
$\Delta t_i = 1$ day (typically)
$S_0 = 0$ (first day free!)

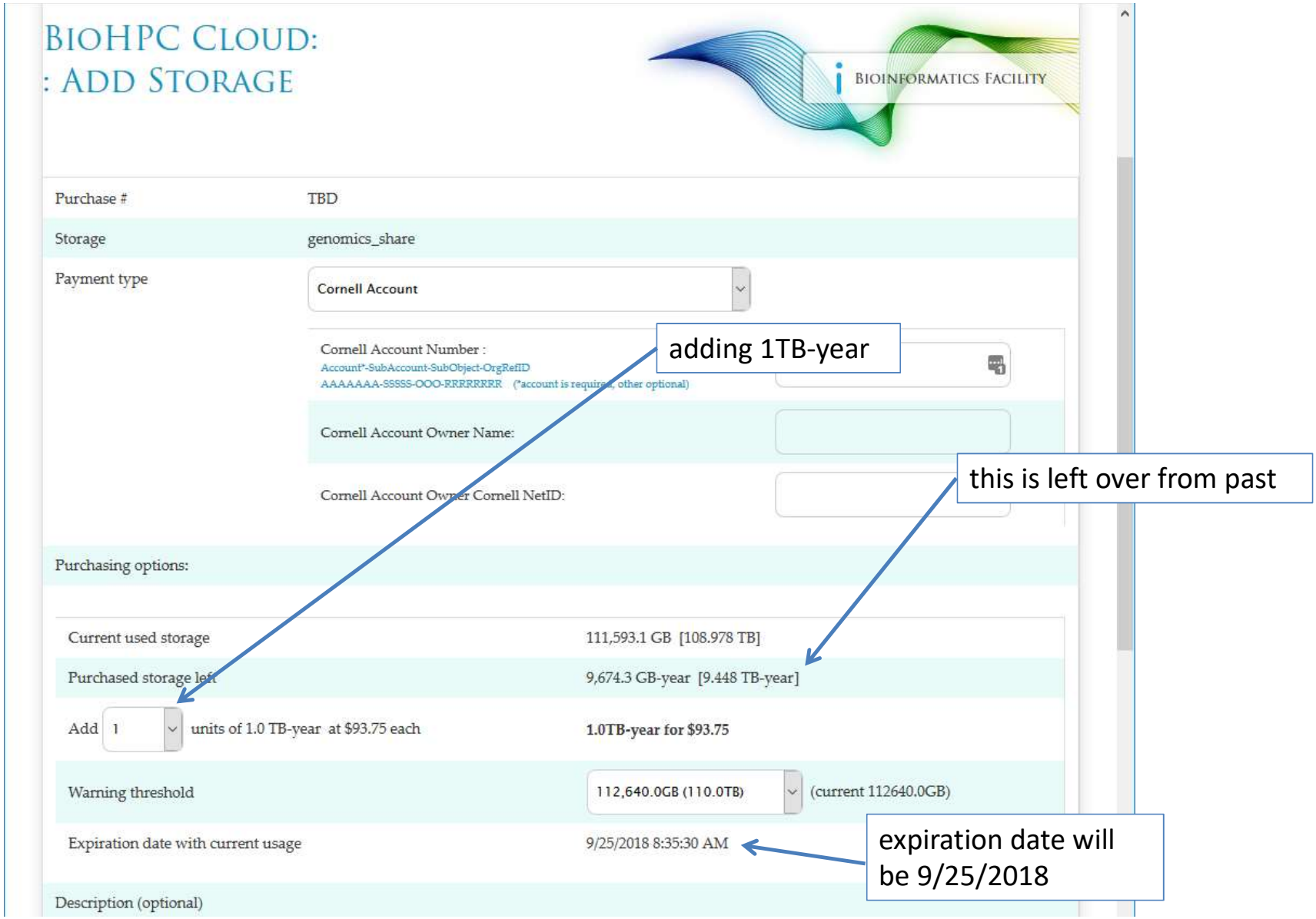
Storage: TB-years

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.






BioHPC Cloud:
: Add Storage
Bioinformatics Facility
Purchase #
TBD
Storage
genomics_share
Payment type
Cornell Account
Cornell Account Number :
Account*-SubAccount-SubObject-OrgRefID
AAAAAAA-SSSSS-OOO-RRRRRRRR (*account is required, other optional)
Cornell Account Owner Name:
Cornell Account Owner Cornell NetID:
Purchasing options:
Current used storage
111,593.1 GB [108.978 TB]
Purchased storage left
9,674.3 GB-year [9.448 TB-year]
Add 1 units of 1.0 TB-year at $93.75 each
1.0TB-year for $93.75
Warning threshold
112,640.0GB (110.0TB)
(current 112640.0GB)
Expiration date with current usage
9/25/2018 8:35:30 AM
Description (optional)
adding 1TB-year
this is left over from past
expiration date will be 9/25/2018

Local Storage Notes

Your local storage on a rental 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 may be deleted on the first night …

Part 2
Hands-on training

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
- Verify that your software is available and read instructions
- Transfer data to your Lab storage
- ✖ Make reservation(s)
- Connect to reserved workstations
- *Compute!*

Connecting to BioHPC Cloud 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.biohpc.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 (Windows) 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 or Linux terminal window. Type

    ssh labid@cbsuXXXX.biohpc.cornell.edu

    (replace the "**cbsuXXXX"** with the server that you just reserved, and "**labid**" 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 CBSU machines from outside of Cornell

Two ways to connect from outside:

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

- ❑ Log in to cbsulogin.biohpc.cornell.edu (or cbsulogin2 or cbsulogin3):

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

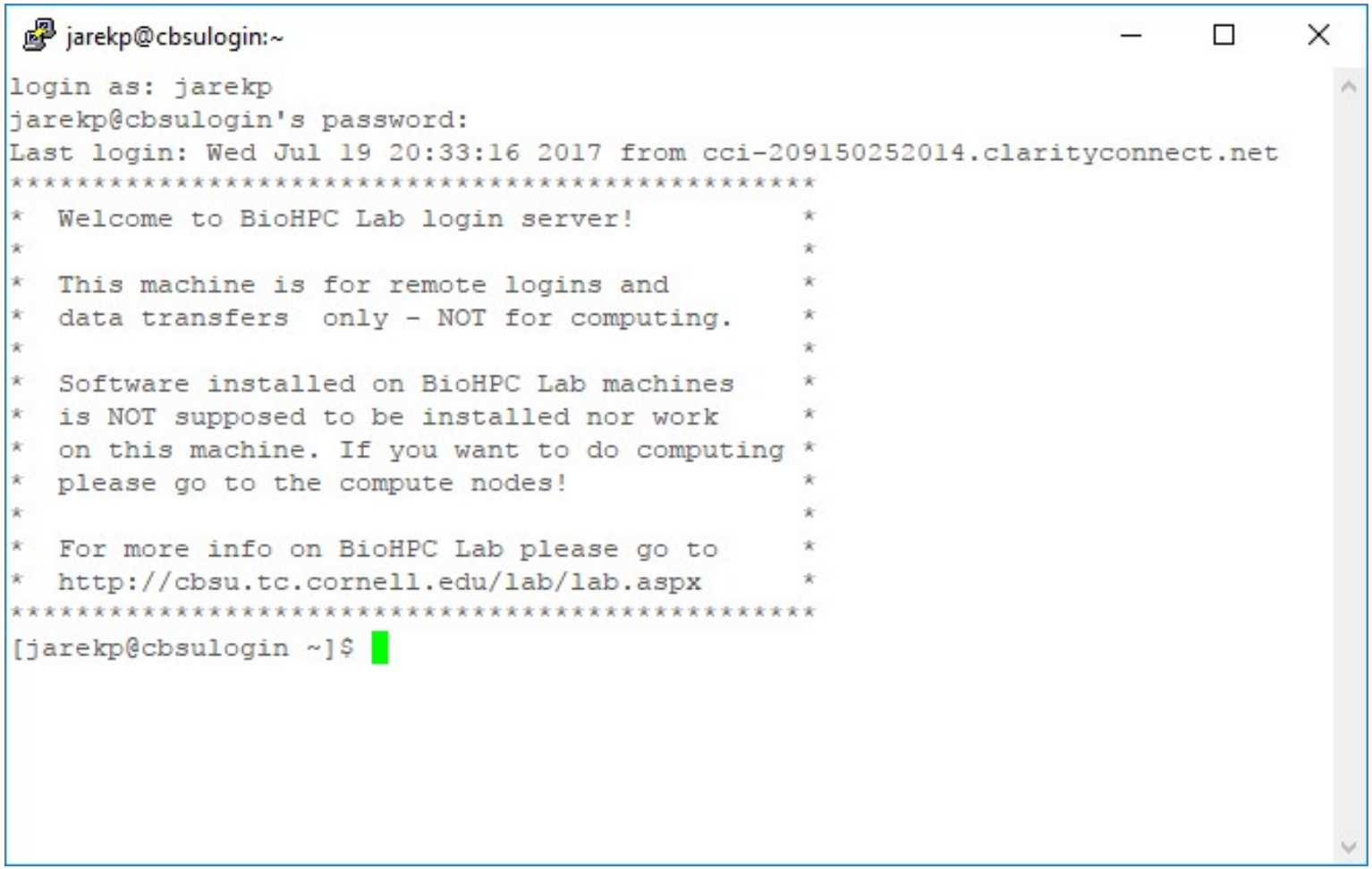
Once logged in to cbsulogin, ssh further to your reserved machine

ssh jarekp@cbsum1.biohpc.cornell.edu

Backup login machine is cbsulogin2.biohpc.cornell.edu and cbsulogin3.biohpc.cornell.edu

https://biohpc.cornell.edu/lab/doc/BioHPCLabexternal.pdf

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. Usually shell called **bash** is used (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 https://biohpc.cornell.edu/ww/machines.aspx?i=126

- 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@server.biohpc.cornell.edu”. Provide your user name and password when prompted.

- Once connected check storage configuration (`df -h`) and who else is there (`w`).

Connecting to BioHPC Cloud

GUI (graphical) connection: VNC

VNC is a remote desktop connection for Linux. It will open full graphical environment the same way one is created when logging in via a graphical console.

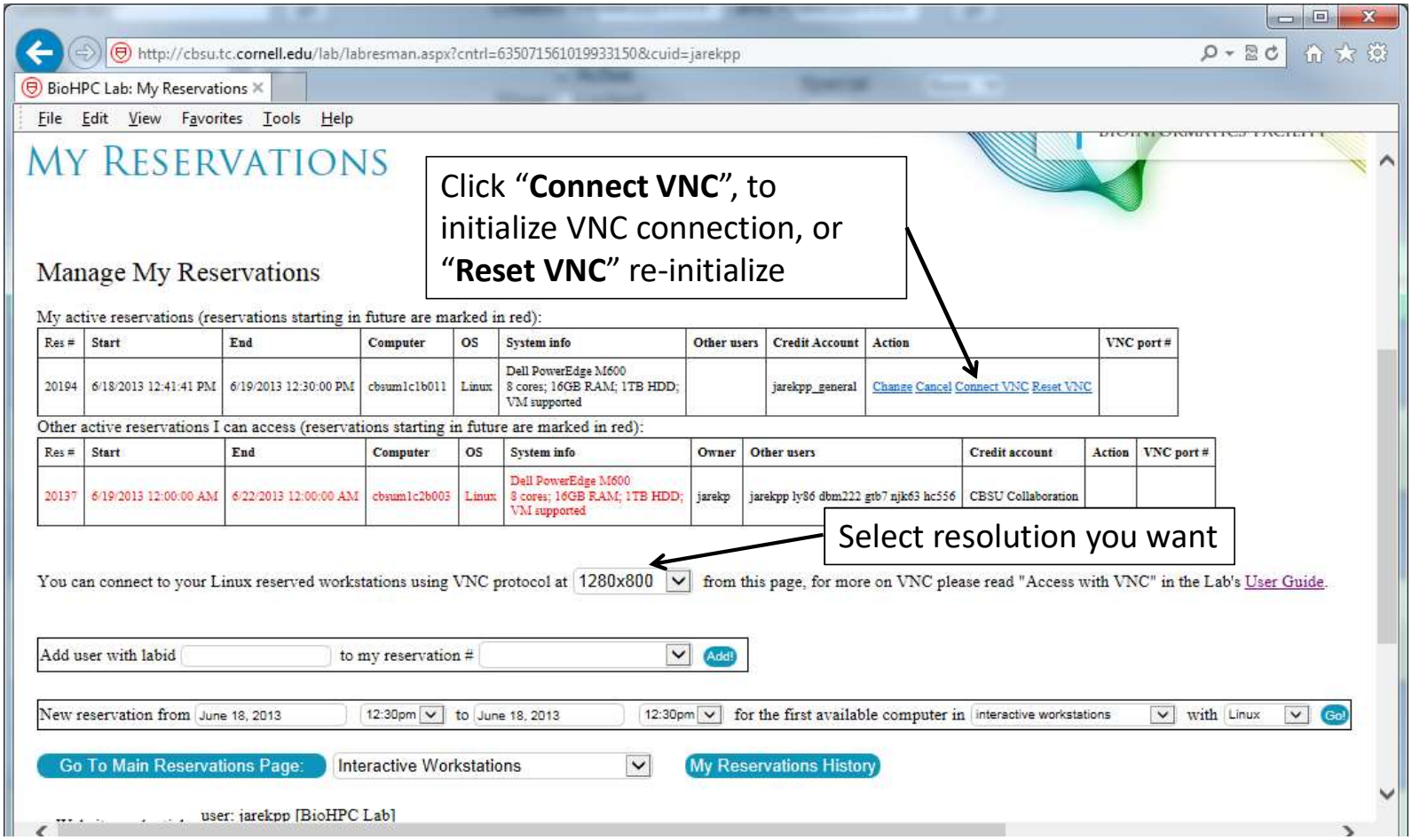
Please note that it is MUCH slower way to connect since lots of graphical data needs to be sent over the network. SSH is much faster, and should be used when no graphics is needed.

VNC stands for Virtual Network Computing

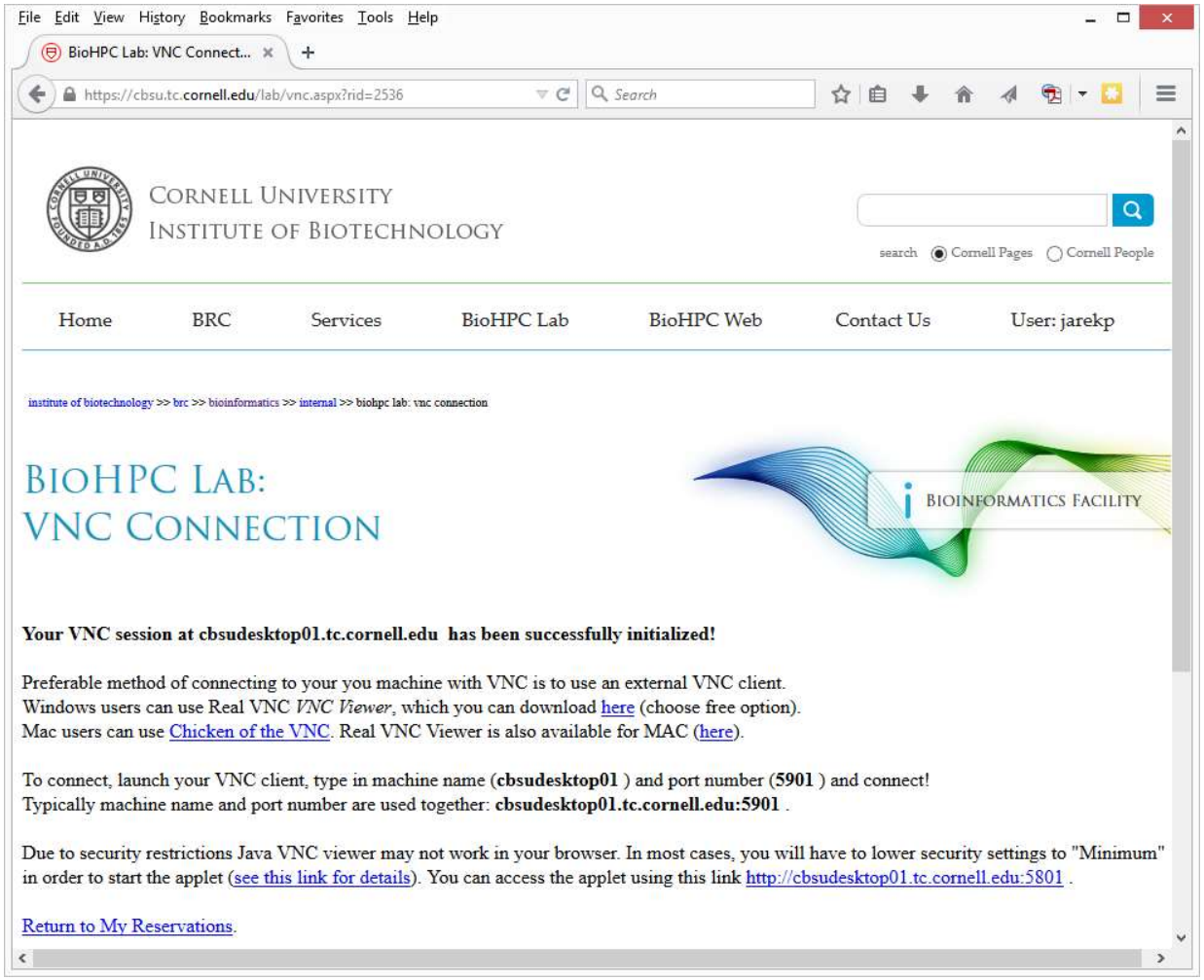
Logging in to a Linux workstation via VNC client

(GUI)

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



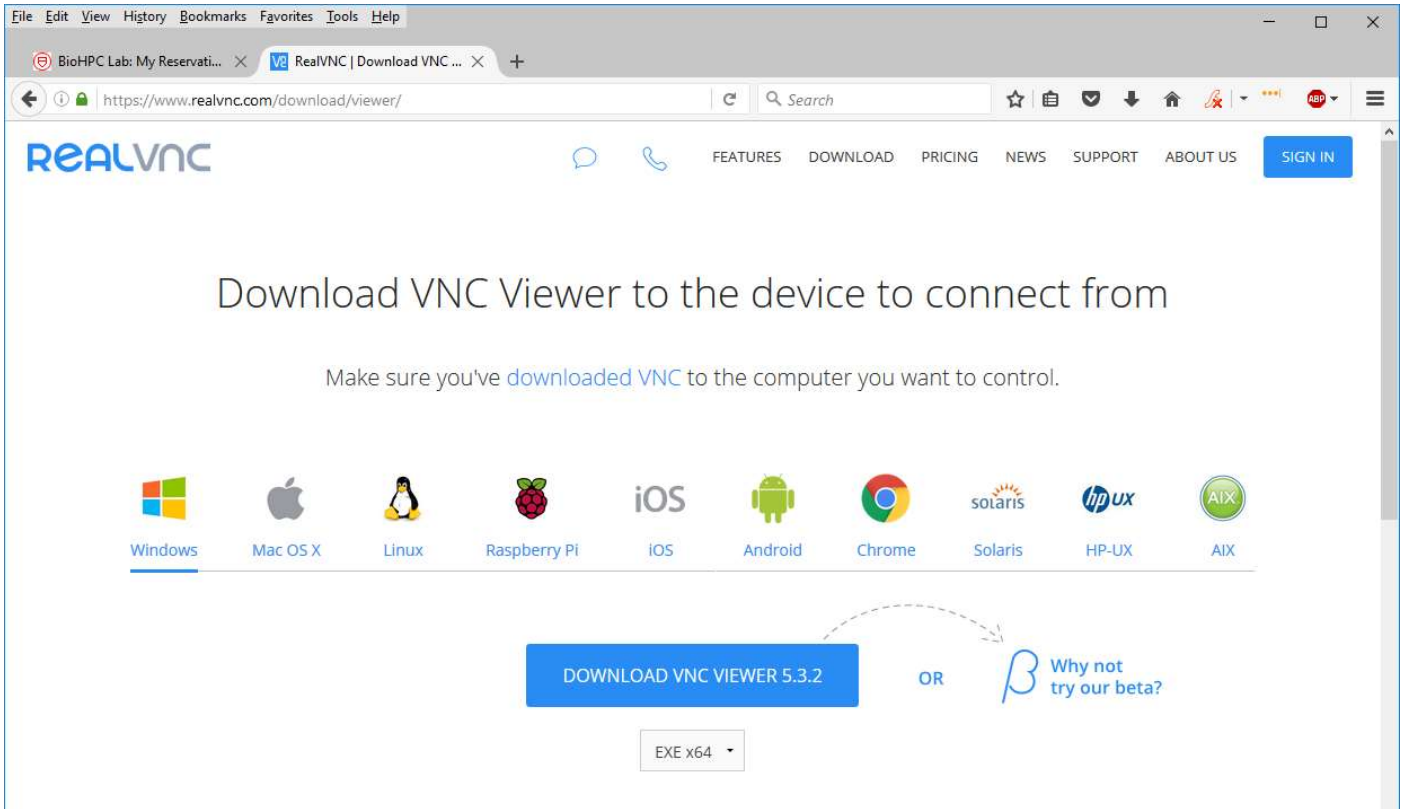
Logging in to a Linux workstation
(GUI)



Logging in to a Linux workstation
(GUI)

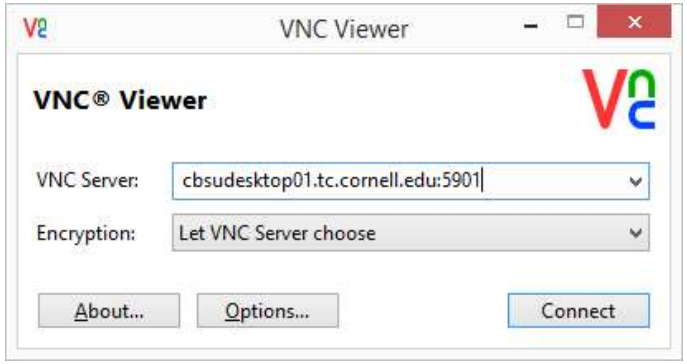
You need software client to connect to your machine via VNC.

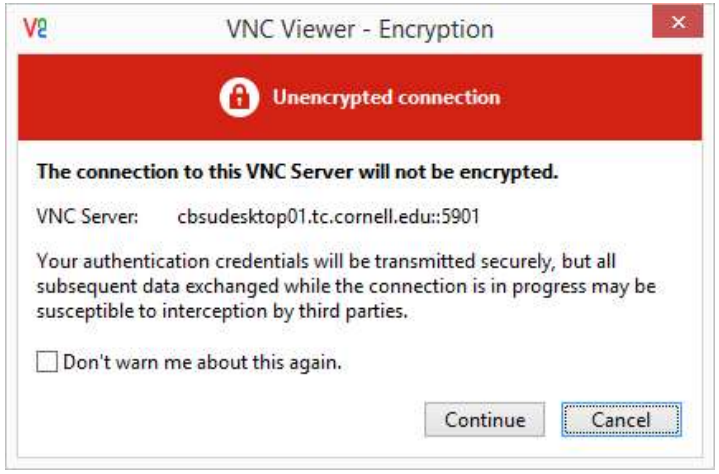
We recommend RealVNC VNC Viewer for all platforms.

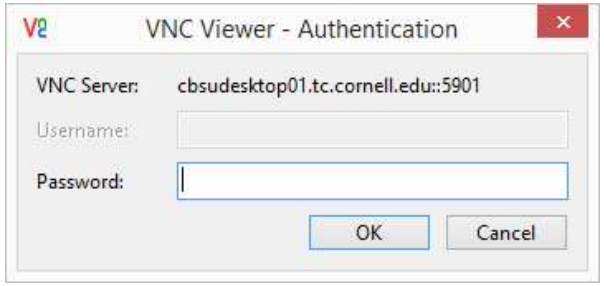


Logging in to a Linux workstation

(GUI)

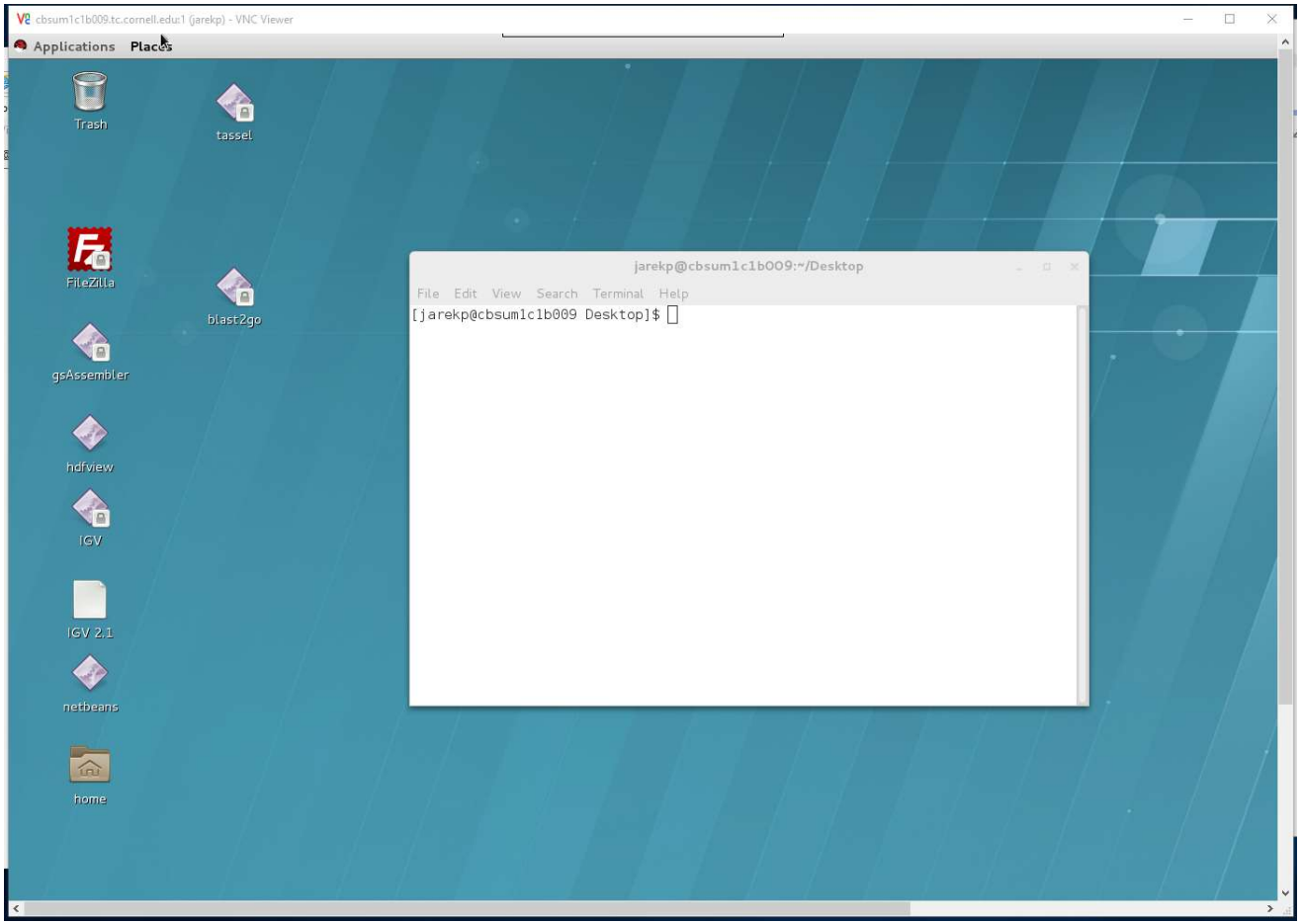






Logging in to a Linux workstation via VNC client
(GUI)

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



Connecting to BioHPC Cloud: VNC

VNC sessions are *persistent*.

VNC sessions run even when the client is disconnected.
Note: SSH can also do that with *screen* or *tmux*.

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

Equivalent to Windows Remote Desktop.

Exercise: connect to your assigned workstation using VNC

- Go to “My Reservations” page
  http://biohpc.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 Cloud

GUI (graphical) connection: X-Windows

Not persistent – programs will get killed when client disconnects, even when using screen or tmux.

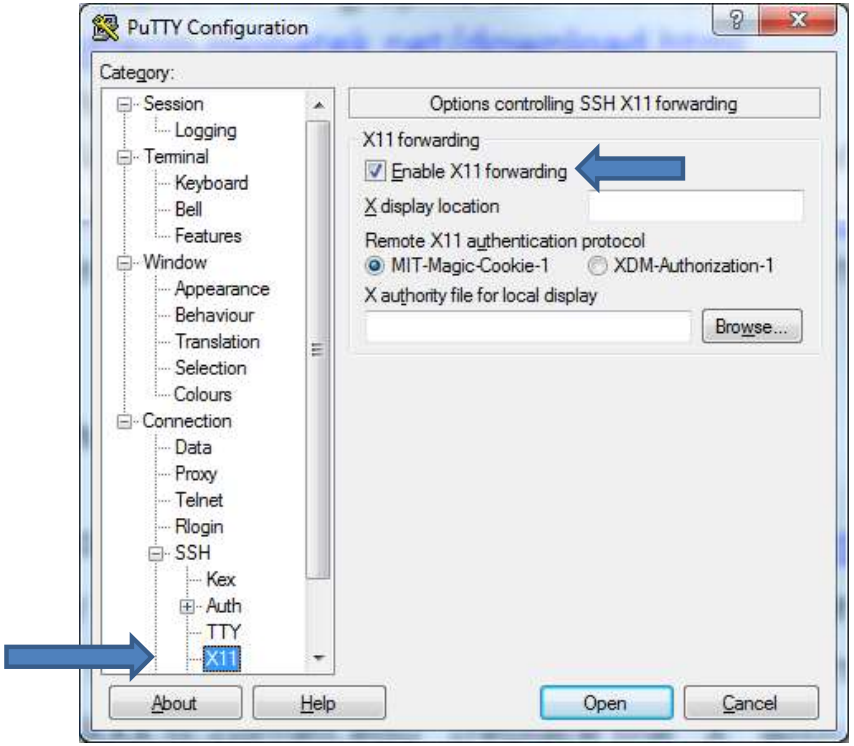
Often faster than VNC

Connecting to BioHPC Cloud: 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 Cloud machine using PuTTY. Make sure X11 forwarding is enabled. X11 is a synonym for X-Windows



Connecting to BioHPC Cloud: 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 Cloud machine, but it will display graphics on your local computer.

Connecting to BioHPC Cloud: X-Windows Linux or Mac

- Connect to BioHPC Cloud machine using ssh with X11 forwarding : ssh -X labid@workstation.biohpc.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 Cloud machine, but it will display graphics on your local computer.

- NOTE: Mac machines no longer have X11 as a part of their system, it must be installed by user

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.

```
eog /home/jarekp/picture.jpg
```

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
- Verify that your software is available and read instructions
- Transfer data to your Lab storage
- ✖ Make reservation(s)
- ✖ Connect to reserved workstations
- *Compute!*

Transferring data

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

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

- wget
  Transfer can be done to BioHPC machines only.

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.biohpc.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

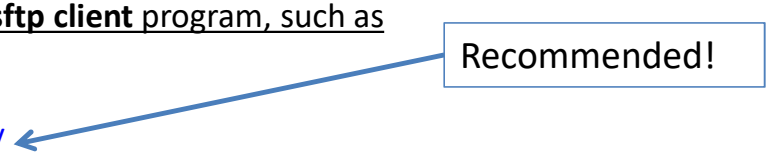
```
wget    ftp://ftp.ncbi.nih.gov/blast/matrices/BLOSUM100
```

File transfer: sftp

between PC or Mac and a BioHPC workstation

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

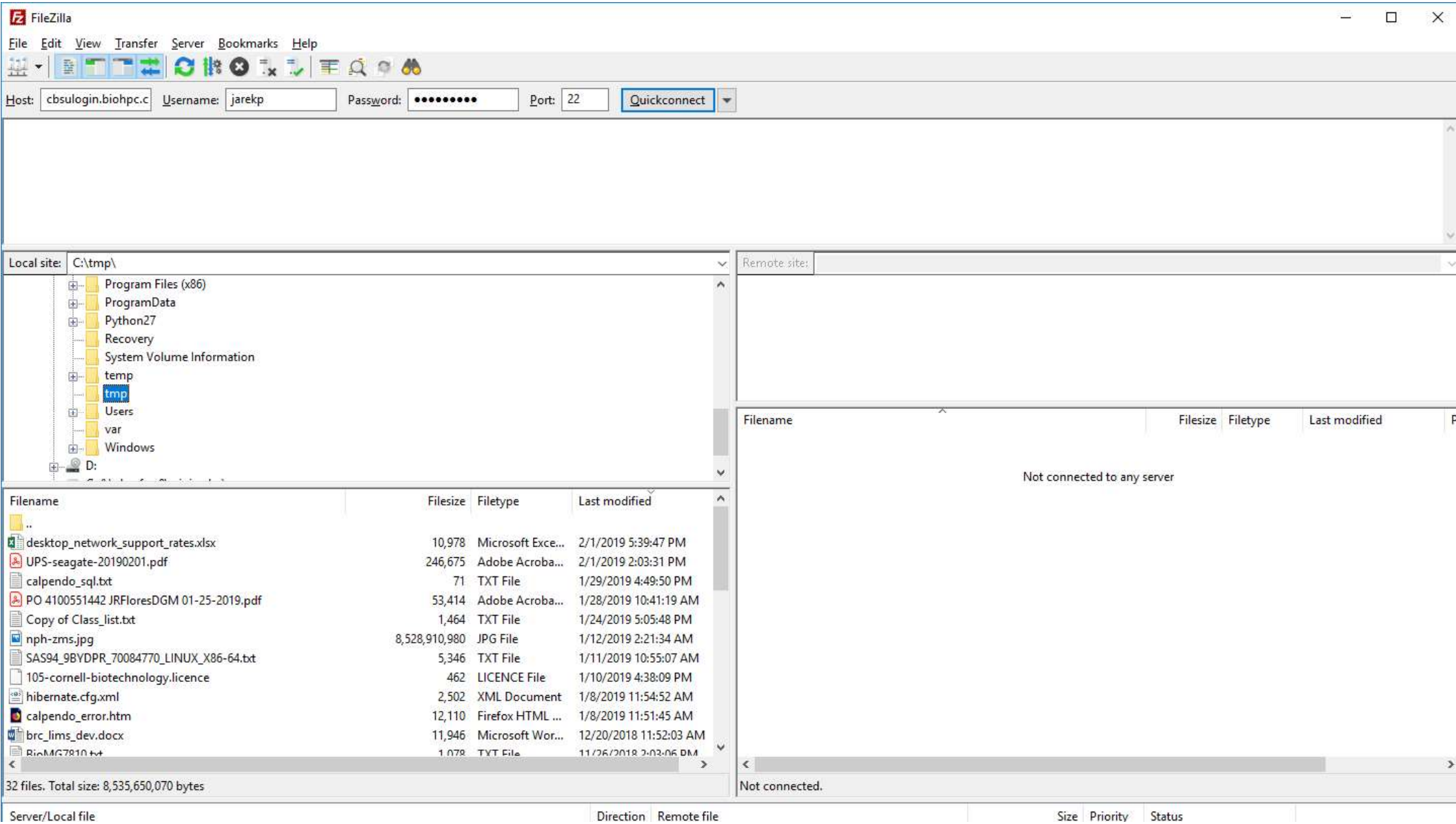
- **winscp**: http://winscp.net/eng/index.php
- **CoreFTP LE**: http://www.coreftp.com/
- **FileZilla** (client): http://filezilla-project.org/
- … others…
- When connecting to Lab workstations from a client, use the **sftp** protocol (port 22). You will be asked for your user name and password (the 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

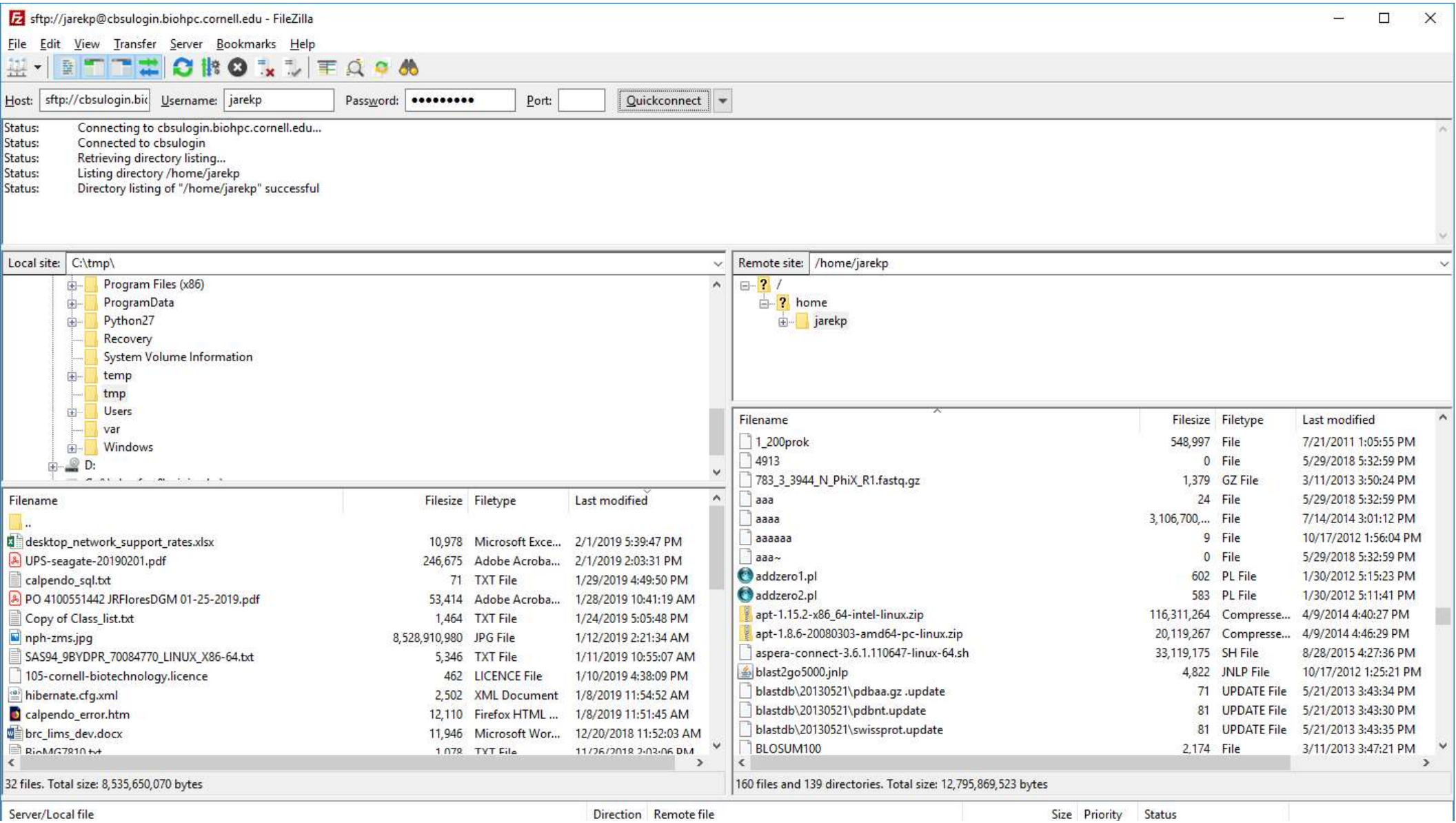


On a Mac: file transfer program is **fetch** (recommended by Cornell CIT)

- https://it.cornell.edu/managed-servers/transfer-files-using-fetch
- graphical user interface
- Drag-and-drop functionality

But **FileZilla** has now also Mac version which I like best …


FileZilla
File Edit View Transfer Server Bookmarks Help
Host: cbsulogin.biohpc.c Username: jarekp Password: ••••••••• Port: 22 Quickconnect
Local site: C:\tmp\
Program Files (x86)
ProgramData
Python27
Recovery
System Volume Information
temp
tmp
Users
var
Windows
D:
Filename Filesize Filetype Last modified
.. 10,978 Microsoft Exce... 2/1/2019 5:39:47 PM
desktop_network_support_rates.xlsx 246,675 Adobe Acroba... 2/1/2019 2:03:31 PM
UPS-seagate-20190201.pdf 71 TXT File 1/29/2019 4:49:50 PM
calpendo_sql.txt 53,414 Adobe Acroba... 1/28/2019 10:41:19 AM
PO 4100551442 JRFloresDGM 01-25-2019.pdf 1,464 TXT File 1/24/2019 5:05:48 PM
Copy of Class_list.txt 8,528,910,980 JPG File 1/12/2019 2:21:34 AM
nph-zms.jpg 5,346 TXT File 1/11/2019 10:55:07 AM
SAS94_9BYDPR_70084770_LINUX_X86-64.txt 462 LICENCE File 1/10/2019 4:38:09 PM
105-cornell-biotechnology.licence 2,502 XML Document 1/8/2019 11:54:52 AM
hibernate.cfg.xml 12,110 Firefox HTML ... 1/8/2019 11:51:45 AM
calpendo_error.htm 11,946 Microsoft Wor... 12/20/2018 11:52:03 AM
brc_lims_dev.docx
32 files. Total size: 8,535,650,070 bytes
Remote site:
Filename Filesize Filetype Last modified
Not connected to any server
Not connected.
Server/Local file Direction Remote file Size Priority Status


sftp://jarekp@cbsulogin.biohpc.cornell.edu - FileZilla
File Edit View Transfer Server Bookmarks Help
Host: sftp://cbsulogin.bic Username: jarekp Password: ••••••••• Port: Quickconnect
Status: Connecting to cbsulogin.biohpc.cornell.edu...
Status: Connected to cbsulogin
Status: Retrieving directory listing...
Status: Listing directory /home/jarekp
Status: Directory listing of "/home/jarekp" successful
Local site: C:\tmp\
Program Files (x86)
ProgramData
Python27
Recovery
System Volume Information
temp
tmp
Users
var
Windows
D:
Remote site: /home/jarekp
/
home
jarekp
Filename Filesize Filetype Last modified
desktop_network_support_rates.xlsx 10,978 Microsoft Exce... 2/1/2019 5:39:47 PM
UPS-seagate-20190201.pdf 246,675 Adobe Acroba... 2/1/2019 2:03:31 PM
calpendo_sql.txt 71 TXT File 1/29/2019 4:49:50 PM
PO 4100551442 JRFloresDGM 01-25-2019.pdf 53,414 Adobe Acroba... 1/28/2019 10:41:19 AM
Copy of Class_list.txt 1,464 TXT File 1/24/2019 5:05:48 PM
nph-zms.jpg 8,528,910,980 JPG File 1/12/2019 2:21:34 AM
SAS94_9BYDPR_70084770_LINUX_X86-64.txt 5,346 TXT File 1/11/2019 10:55:07 AM
105-cornell-biotechnology.licence 462 LICENCE File 1/10/2019 4:38:09 PM
hibernate.cfg.xml 2,502 XML Document 1/8/2019 11:54:52 AM
calpendo_error.htm 12,110 Firefox HTML ... 1/8/2019 11:51:45 AM
brc_lims_dev.docx 11,946 Microsoft Wor... 12/20/2018 11:52:03 AM
32 files. Total size: 8,535,650,070 bytes
Filename Filesize Filetype Last modified
1_200prok 548,997 File 7/21/2011 1:05:55 PM
4913 0 File 5/29/2018 5:32:59 PM
783_3_3944_N_PhiX_R1.fastq.gz 1,379 GZ File 3/11/2013 3:50:24 PM
aaa 24 File 5/29/2018 5:32:59 PM
aaaa 3,106,700,... File 7/14/2014 3:01:12 PM
aaaaaa 9 File 10/17/2012 1:56:04 PM
aaa~ 0 File 5/29/2018 5:32:59 PM
addzero1.pl 602 PL File 1/30/2012 5:15:23 PM
addzero2.pl 583 PL File 1/30/2012 5:11:41 PM
apt-1.15.2-x86_64-intel-linux.zip 116,311,264 Compresse... 4/9/2014 4:40:27 PM
apt-1.8.6-20080303-amd64-pc-linux.zip 20,119,267 Compresse... 4/9/2014 4:46:29 PM
aspera-connect-3.6.1.110647-linux-64.sh 33,119,175 SH File 8/28/2015 4:27:36 PM
blast2go5000.jnlp 4,822 JNLP File 10/17/2012 1:25:21 PM
blastdb\20130521\pdbaa.gz .update 71 UPDATE File 5/21/2013 3:43:34 PM
blastdb\20130521\pdbnt.update 81 UPDATE File 5/21/2013 3:43:30 PM
blastdb\20130521\swissprot.update 81 UPDATE File 5/21/2013 3:43:35 PM
BLOSUM100 2,174 File 3/11/2013 3:47:21 PM
160 files and 139 directories. Total size: 12,795,869,523 bytes
Server/Local file Direction Remote file Size Priority Status

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

- Connect to your workstation using sftp program (FileZilla)
- Download BLOSUM100 – you got it there before using wget

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 Cloud Globus endpoint servers are is **biohpc#cbsulogin**, **biohpc#cbsulogin2** and **biohpc#cbsulogin3**

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

http://biohpc.cornell.edu/lab/doc/Globus_at_BioHPC_Lab.pdf

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

globus

I Want To... Pricing Resources Support About Log In

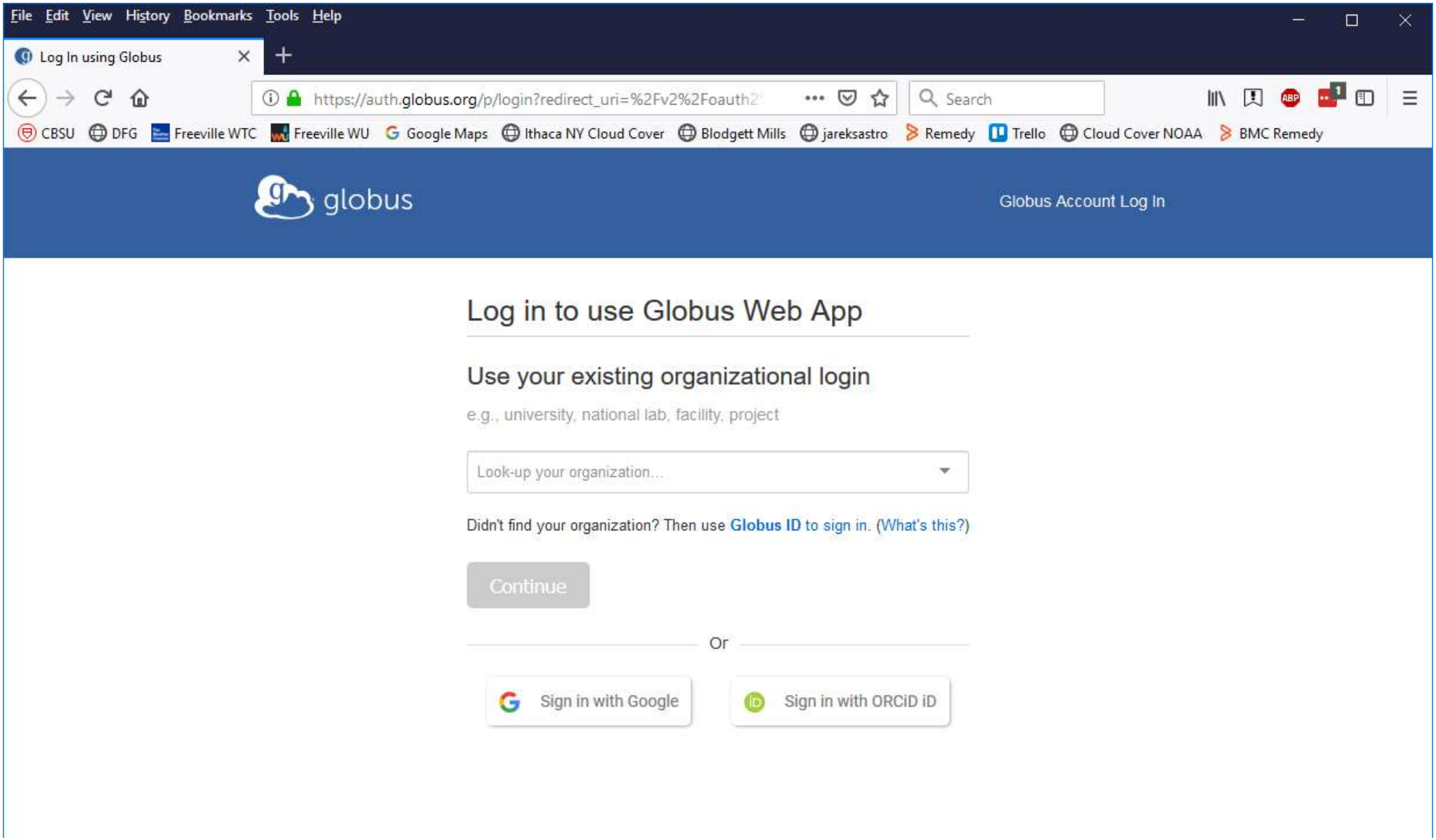
Protected Data Support

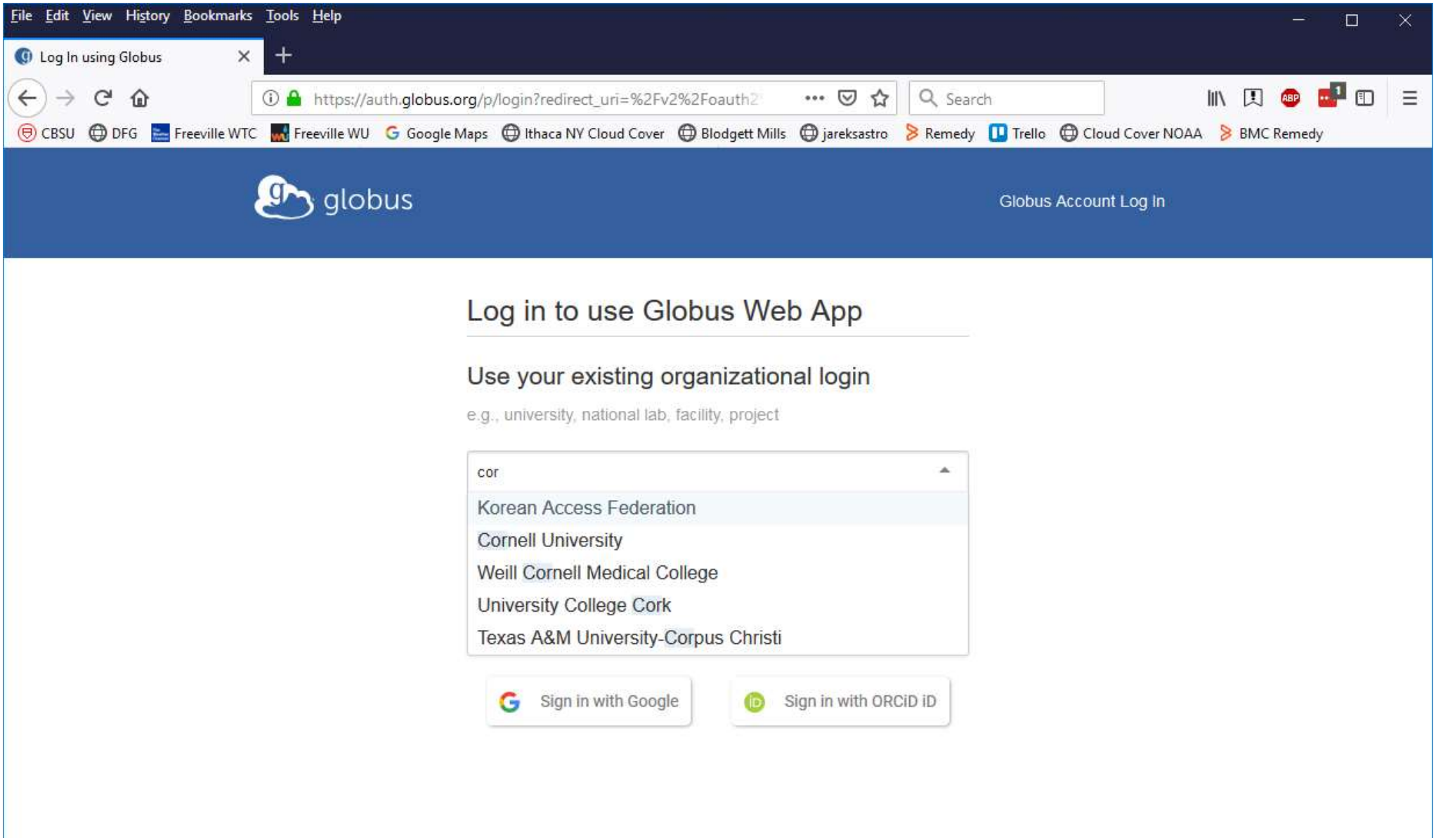
Working with PHI, PII, or CUI?
Need to manage HIPAA-regulated data?
Globus has you covered!

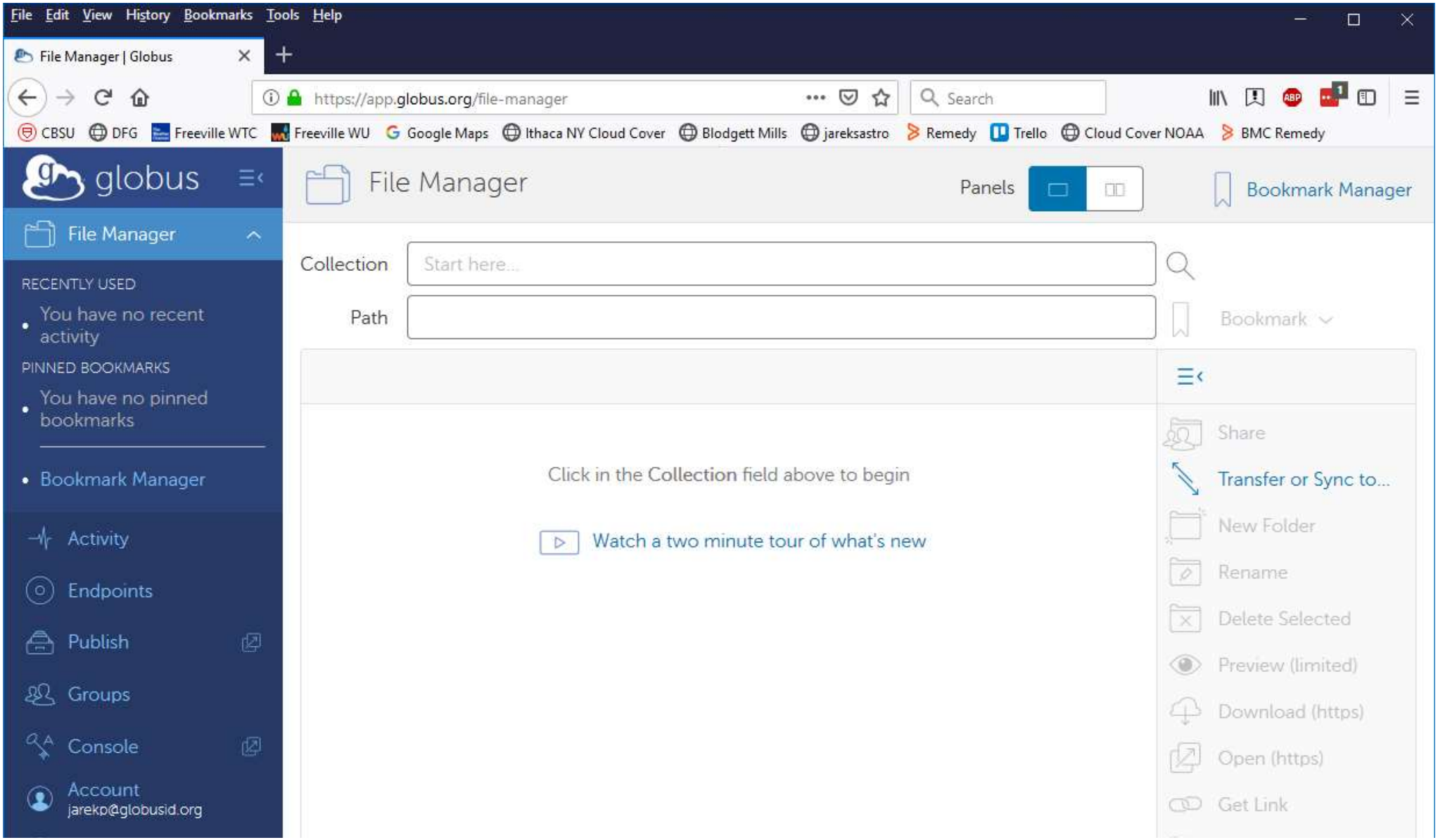
LEARN MORE REQUEST PRICING

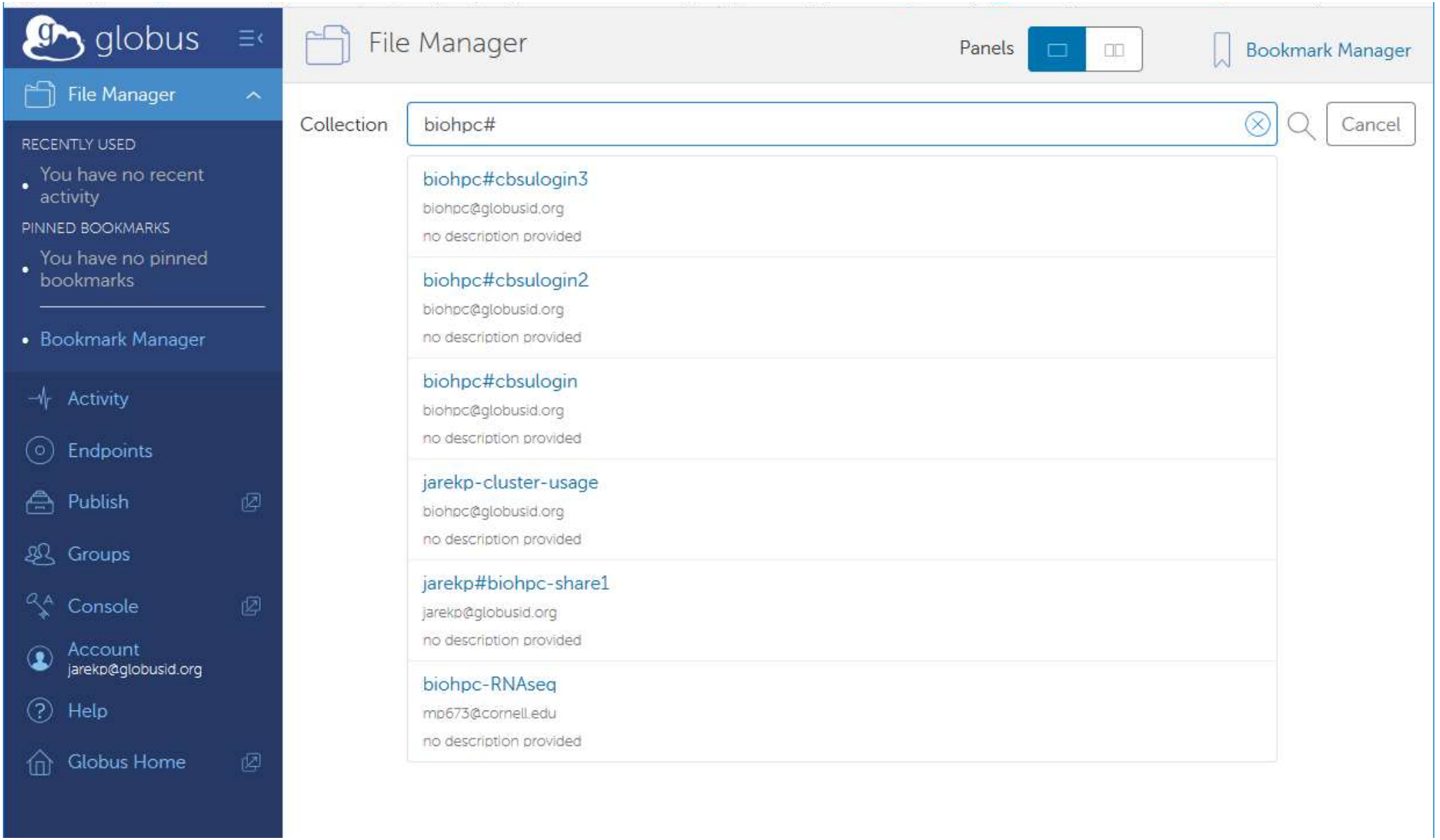
Research data management simplified.

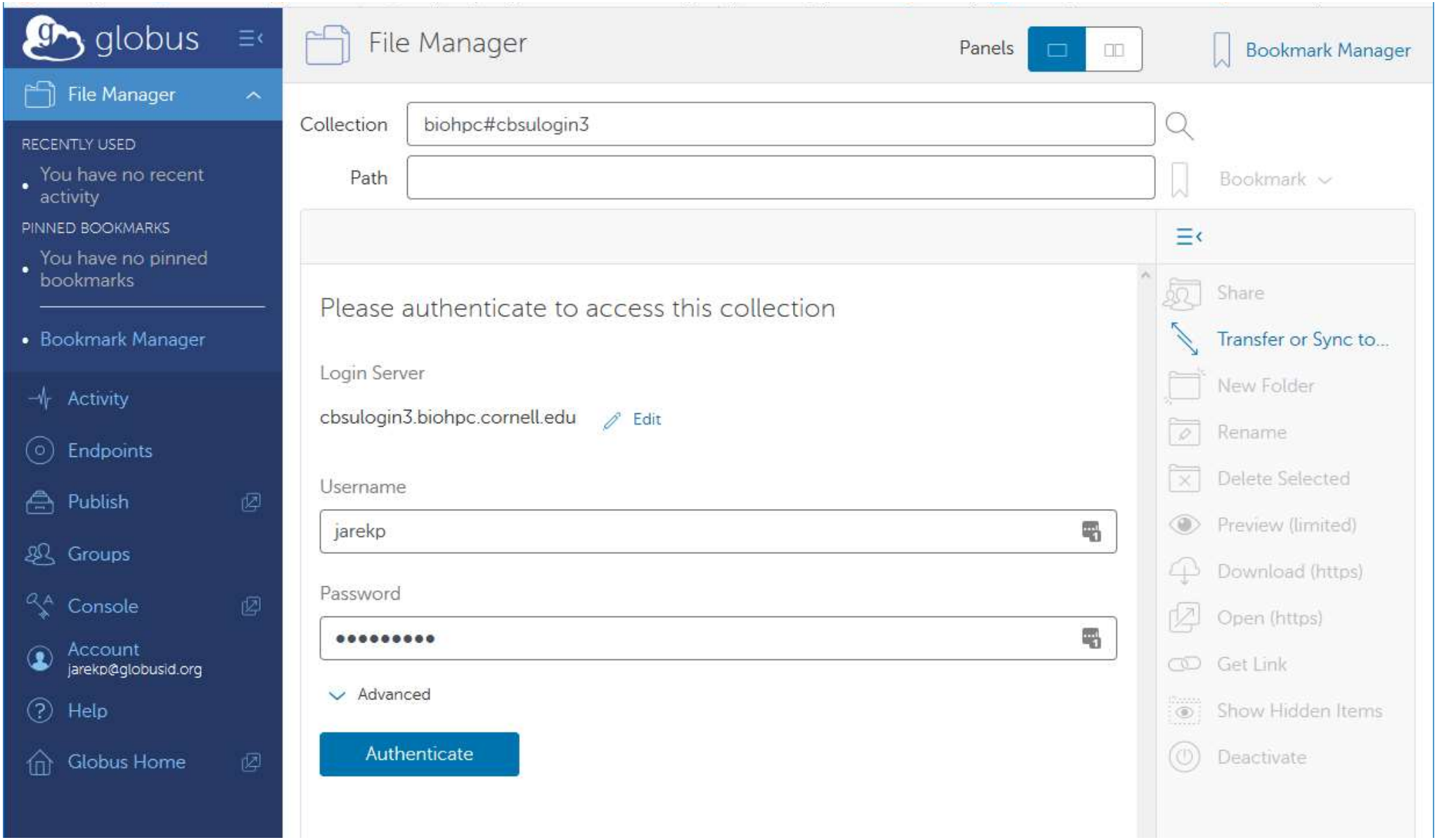
TRANSFER SHARE PUBLISH BUILD

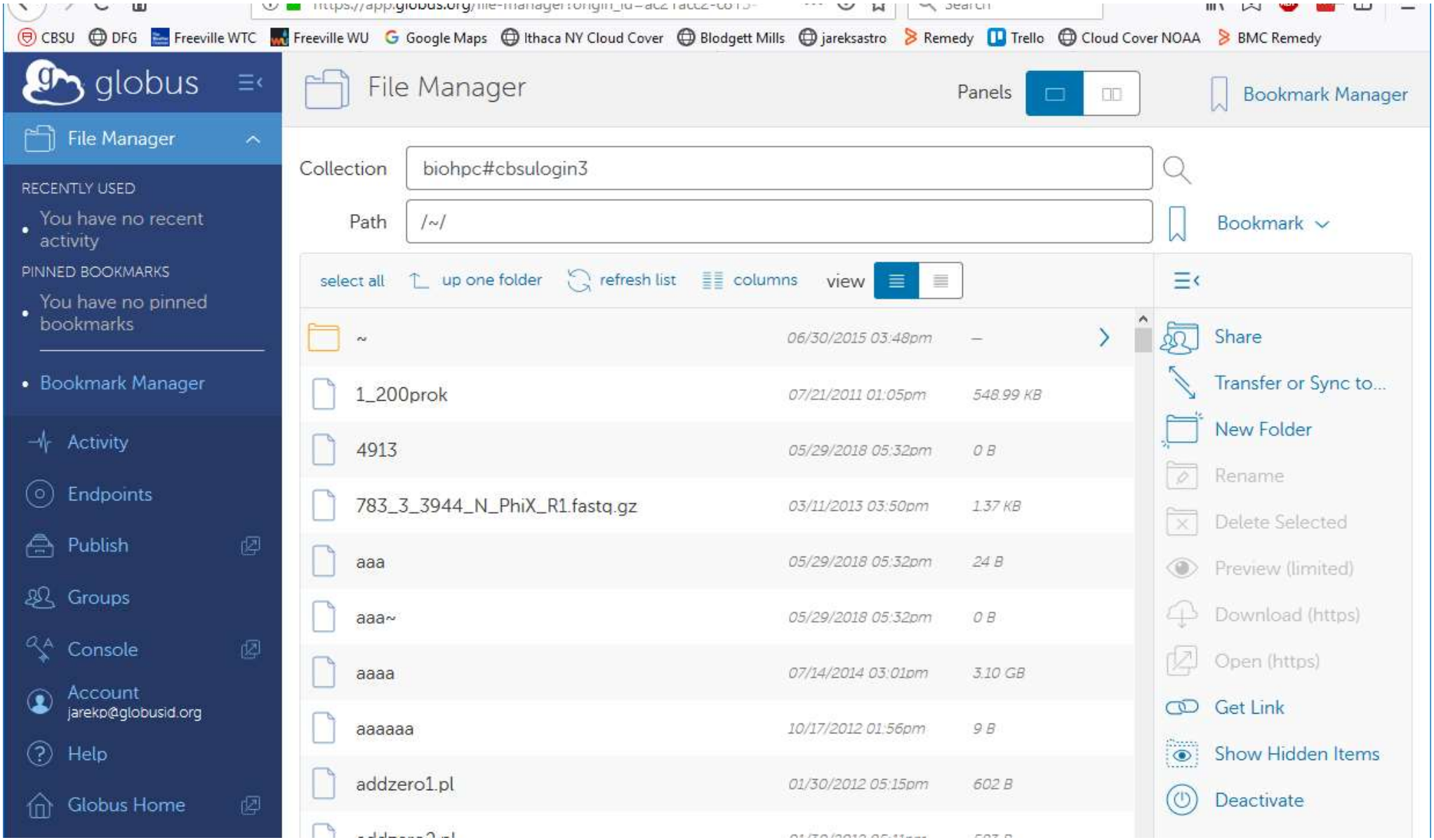

File Edit View History Bookmarks Tools Help
Log In using Globus
https://auth.globus.org/p/login?redirect_uri=%2Fv2%2Foauth2
Search
CBSU
DFG
Freeville WTC
Freeville WU
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
jareksastro
Remedy
Trello
Cloud Cover NOAA
BMC Remedy
globus
Globus Account Log In
Log in to use Globus Web App
Use your existing organizational login
e.g., university, national lab, facility, project
Look-up your organization...
Didn't find your organization? Then use Globus ID to sign in. (What's this?)
Continue
Or
Sign in with Google
Sign in with ORCiD iD

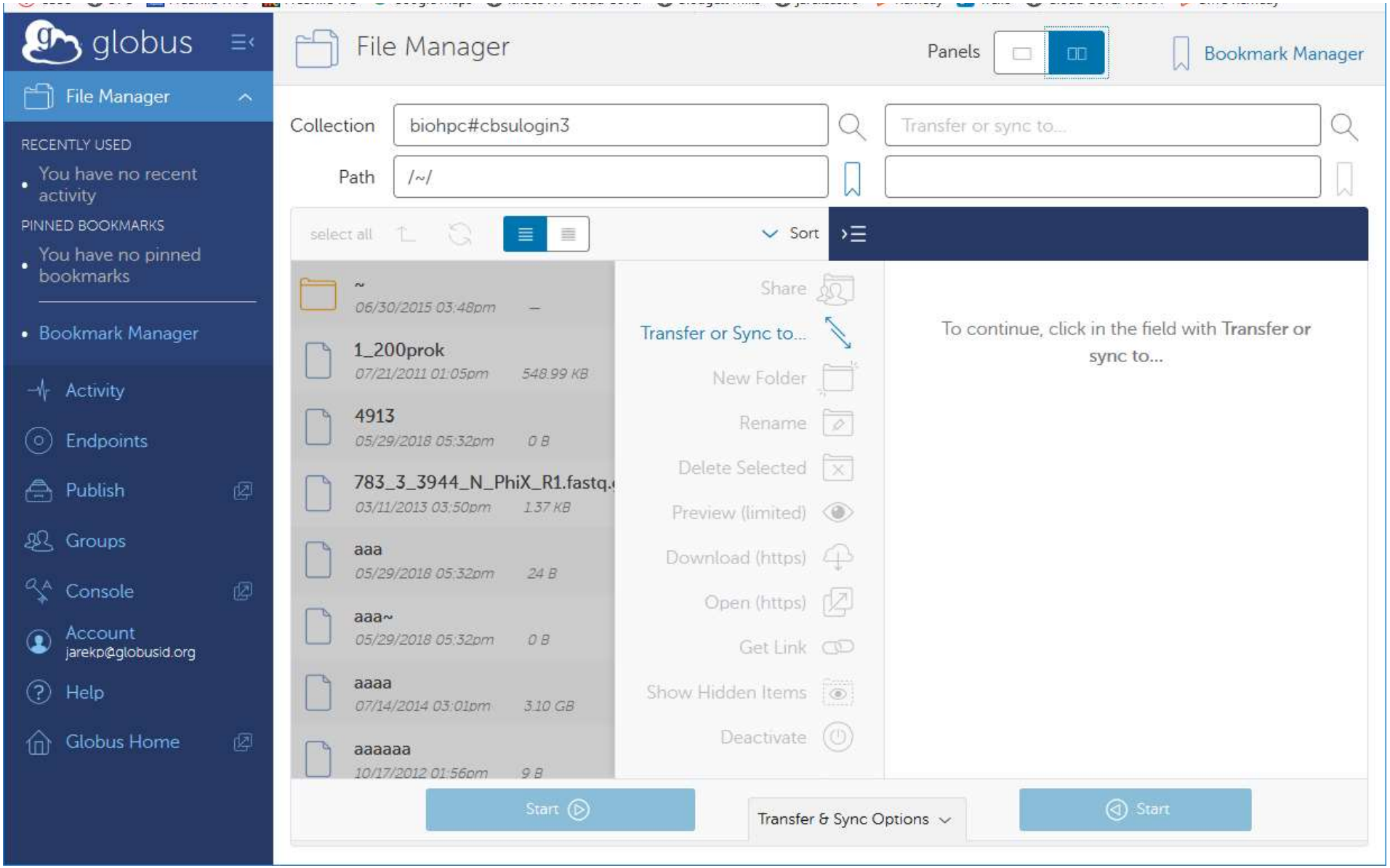

File Edit View History Bookmarks Tools Help
Log In using Globus
https://auth.globus.org/p/login?redirect_uri=%2Fv2%2Foauth2
Search
CBSU
DFG
Freeville WTC
Freeville WU
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
jareksastro
Remedy
Trello
Cloud Cover NOAA
BMC Remedy
globus
Globus Account Log In
Log in to use Globus Web App
Use your existing organizational login
e.g., university, national lab, facility, project
cor
Korean Access Federation
Cornell University
Weill Cornell Medical College
University College Cork
Texas A&M University-Corpus Christi
Sign in with Google
Sign in with ORCiD iD


File Edit View History Bookmarks Tools Help
File Manager | Globus
https://app.globus.org/file-manager
Search
CBSU
DFG
Freeville WTC
Freeville WU
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
jareksastro
Remedy
Trello
Cloud Cover NOAA
BMC Remedy
globus
File Manager
Panels
Bookmark Manager
File Manager
RECENTLY USED
You have no recent activity
PINNED BOOKMARKS
You have no pinned bookmarks
Bookmark Manager
Activity
Endpoints
Publish
Groups
Console
Account
jarekp@globusid.org
Collection
Start here...
Path
Bookmark
Click in the Collection field above to begin
Watch a two minute tour of what's new
Share
Transfer or Sync to...
New Folder
Rename
Delete Selected
Preview (limited)
Download (https)
Open (https)
Get Link


globus
File Manager
RECENTLY USED
You have no recent activity
PINNED BOOKMARKS
You have no pinned bookmarks
Bookmark Manager
Activity
Endpoints
Publish
Groups
Console
Account
jarekp@globusid.org
Help
Globus Home
File Manager
Panels
Bookmark Manager
Collection
biohpc#
Cancel
biohpc#cbsulogin3
biohpc@globusid.org
no description provided
biohpc#cbsulogin2
biohpc@globusid.org
no description provided
biohpc#cbsulogin
biohpc@globusid.org
no description provided
jarekp-cluster-usage
biohpc@globusid.org
no description provided
jarekp#biohpc-share1
jarekp@globusid.org
no description provided
biohpc-RNAseq
mp673@cornell.edu
no description provided


globus
File Manager
RECENTLY USED
You have no recent activity
PINNED BOOKMARKS
You have no pinned bookmarks
Bookmark Manager
Activity
Endpoints
Publish
Groups
Console
Account
jarekp@globusid.org
Help
Globus Home
File Manager
Panels
Bookmark Manager
Collection
biohpc#cbsulogin3
Path
Bookmark
Please authenticate to access this collection
Login Server
cbsulogin3.biohpc.cornell.edu
Edit
Username
jarekp
Password
Advanced
Authenticate
Share
Transfer or Sync to...
New Folder
Rename
Delete Selected
Preview (limited)
Download (https)
Open (https)
Get Link
Show Hidden Items
Deactivate

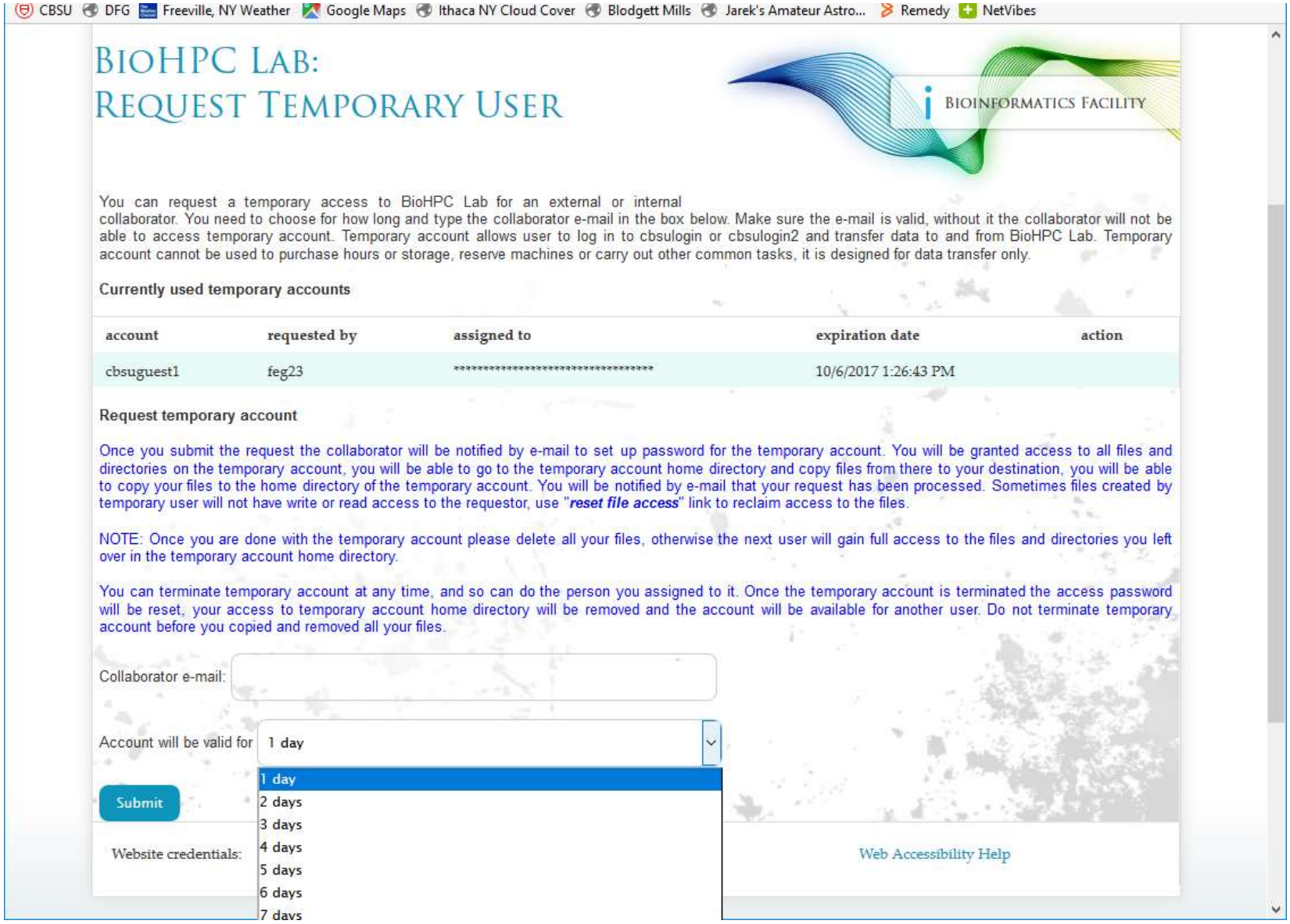

CBSU
DFG
Freeville WTC
Freeville WU
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
jareksastro
Remedy
Trello
Cloud Cover NOAA
BMC Remedy
globus
File Manager
RECENTLY USED
You have no recent activity
PINNED BOOKMARKS
You have no pinned bookmarks
Bookmark Manager
Activity
Endpoints
Publish
Groups
Console
Account
jarekp@globusid.org
Help
Globus Home
File Manager
Panels
Bookmark Manager
Collection
biohpc#cbsulogin3
Path
/~/
Bookmark
select all
up one folder
refresh list
columns
view
~ 06/30/2015 03:48pm –
1_200prok 07/21/2011 01:05pm 548.99 KB
4913 05/29/2018 05:32pm 0 B
783_3_3944_N_PhiX_R1.fastq.gz 03/11/2013 03:50pm 1.37 KB
aaa 05/29/2018 05:32pm 24 B
aaa~ 05/29/2018 05:32pm 0 B
aaaa 07/14/2014 03:01pm 3.10 GB
aaaaaa 10/17/2012 01:56pm 9 B
addzero1.pl 01/30/2012 05:15pm 602 B
Share
Transfer or Sync to...
New Folder
Rename
Delete Selected
Preview (limited)
Download (https)
Open (https)
Get Link
Show Hidden Items
Deactivate


globus
File Manager
RECENTLY USED
You have no recent activity
PINNED BOOKMARKS
You have no pinned bookmarks
Bookmark Manager
Activity
Endpoints
Publish
Groups
Console
Account
jarekp@globusid.org
Help
Globus Home
File Manager
Panels
Bookmark Manager
Collection
biohpc#cbsulogin3
Transfer or sync to...
Path
/~/
select all
Sort
~
06/30/2015 03:48pm
1_200prok
07/21/2011 01:05pm 548.99 KB
4913
05/29/2018 05:32pm 0 B
783_3_3944_N_PhiX_R1.fastq.
03/11/2013 03:50pm 1.37 KB
aaa
05/29/2018 05:32pm 24 B
aaa~
05/29/2018 05:32pm 0 B
aaaa
07/14/2014 03:01pm 3.10 GB
aaaaaa
10/17/2012 01:56pm 9 B
Share
Transfer or Sync to...
New Folder
Rename
Delete Selected
Preview (limited)
Download (https)
Open (https)
Get Link
Show Hidden Items
Deactivate
To continue, click in the field with Transfer or sync to...
Start
Transfer & Sync Options
Start

Exercise: Login to Globus using your BioHPC account

Sharing data with external collaborators

- Use Globus – you can share for reading any of your files or directories provided the other person has Globus account
- Request temporary BioHPC account for data sharing. This allows external collaborators both inbound and outbound transfer.
- Request permanent BioHPC account for external collaborator


CBSU
DFG
Freeville, NY Weather
Google Maps
Ithaca NY Cloud Cover
Blodgett Mills
Jarek's Amateur Astro...
Remedy
NetVibes
BioHPC Lab:
Request Temporary User
Bioinformatics Facility
You can request a temporary access to BioHPC Lab for an external or internal collaborator. You need to choose for how long and type the collaborator e-mail in the box below. Make sure the e-mail is valid, without it the collaborator will not be able to access temporary account. Temporary account allows user to log in to cbsulogin or cbsulogin2 and transfer data to and from BioHPC Lab. Temporary account cannot be used to purchase hours or storage, reserve machines or carry out other common tasks, it is designed for data transfer only.
Currently used temporary accounts
account
requested by
assigned to
expiration date
action
cbsuguest1
feg23
****
10/6/2017 1:26:43 PM
Request temporary account
Once you submit the request the collaborator will be notified by e-mail to set up password for the temporary account. You will be granted access to all files and directories on the temporary account, you will be able to go to the temporary account home directory and copy files from there to your destination, you will be able to copy your files to the home directory of the temporary account. You will be notified by e-mail that your request has been processed. Sometimes files created by temporary user will not have write or read access to the requestor, use "reset file access" link to reclaim access to the files.
NOTE: Once you are done with the temporary account please delete all your files, otherwise the next user will gain full access to the files and directories you left over in the temporary account home directory.
You can terminate temporary account at any time, and so can do the person you assigned to it. Once the temporary account is terminated the access password will be reset, your access to temporary account home directory will be removed and the account will be available for another user. Do not terminate temporary account before you copied and removed all your files.
Collaborator e-mail:
Account will be valid for
1 day
1 day
2 days
3 days
4 days
5 days
6 days
7 days
Submit
Website credentials:
Web Accessibility Help

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
- Verify that your software is available and read instructions
- ✖ Transfer data to your Lab storage
- ✖ Make reservation(s)
- ✖ Connect to reserved workstations
- *Compute!*

Software

https://biohpc.cornell.edu/lab/userguide.aspx?a=software
https://biohpc.cornell.edu/lab/labsoftware.aspx

Please read the information provided, and consult author's website if needed.

Some programs require special setup to run, some need to be run with full path, some are installed only on a subset of workstations.

Please check detailed information before running any program, it may contain important instruction on how to properly use the software in BioHPC Cloud.
Non-tabular Lab Software pagee

Filter by:

Name: * go About: * go

☑ Show Windows
☑ Show Linux

Order by Name Ascending

583 Records found. Show 1-200

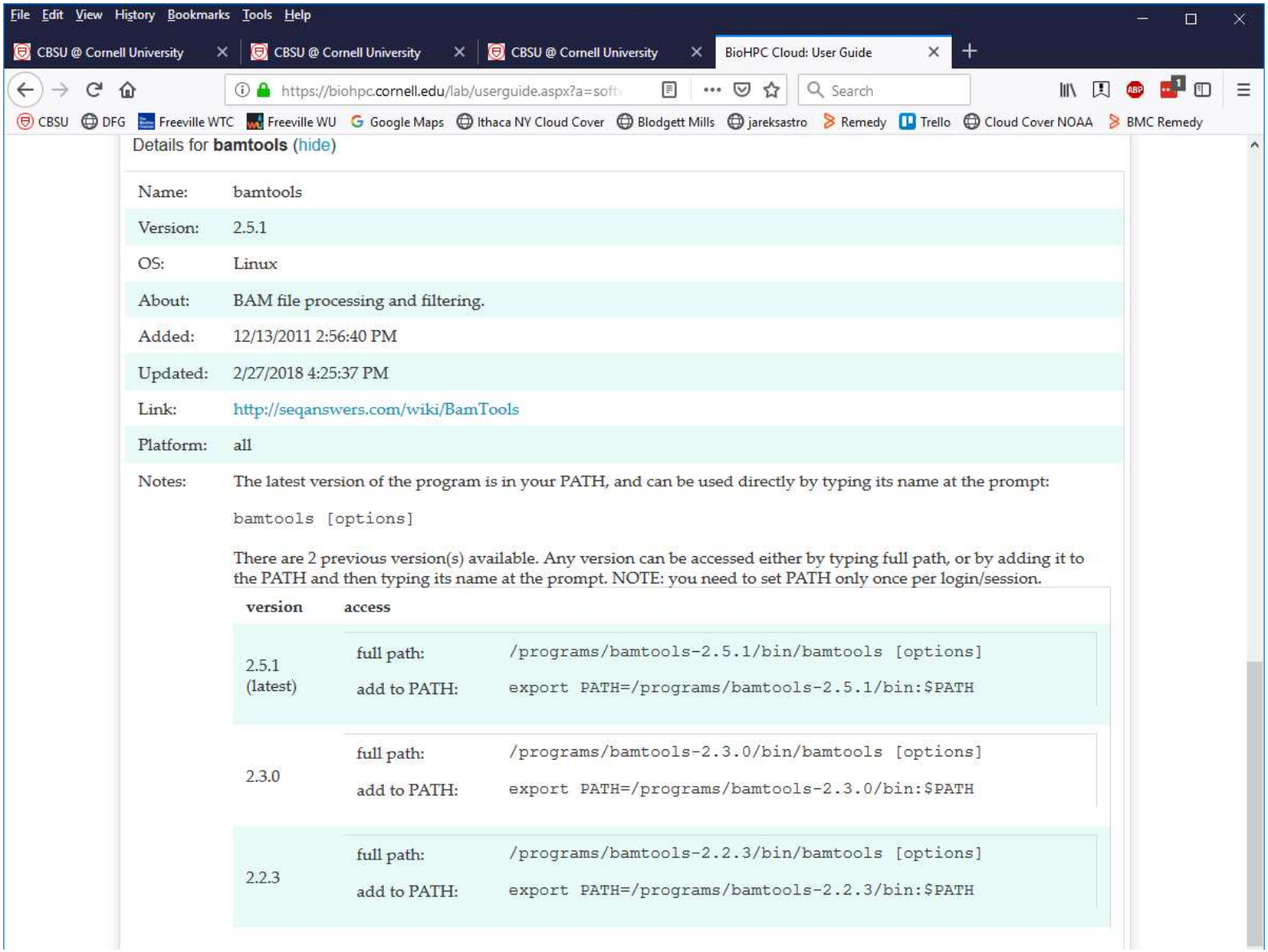
Name	Version	OS	About	Installed	Updated	More
454 gsAssembler or gsMapper	2.8	Linux	Assembly and alignment software for 454 or other long reads.	12/13/2011	3/24/2014	detailed information
a5	20150522	Linux	A genome assembly pipeline for bacteria and archaea	10/14/2013	6/3/2016	detailed information
ABRicate	v0.8	Linux	Mass screening of contigs for antimicrobial resistance or virulence genes.	6/11/2018		detailed information
ABruijn	20161221	Linux	ABruijn is a de novo assembler for PacBio and Oxford Nanopore Technologies reads.	12/21/2016		detailed information
ABySS	1.9.0	Linux	Illumina short reads assembly tool.	12/13/2011	5/9/2016	detailed information
AdapterRemoval	2.1.1	Linux	Remove adapters from sequences in either single end or paired end experiments	9/16/2015		detailed information
Admixtools	5.1	Linux	The ADMIXTOOLS package implements 5 methods described in Patterson et al (2012) Ancient Admixture in Human History.	11/6/2013	12/20/2018	detailed information
Admixture	1.23	Linux	Software tool for maximum likelihood estimation of individual ancestries from multilocus SNP genotype datasets.	2/19/2014	2/19/2014	detailed information
agrep	3.41.5	Linux	approximate GREP for fast fuzzy string searching.	7/12/2018		detailed information
albacore	2.3.4	Linux	Nanopore base caller.	6/2/2017	1/4/2019	detailed information
Alder	1.03	Linux	The ALDER software computes the weighted linkage disequilibrium (LD) statistic for making inference about population admixture.	11/6/2013	11/6/2013	detailed information
AlleleSeq	1.1	Linux	Detects SNVs from ChIP-seq or RNA-seq experiments.	4/2/2014		detailed information
ALLMAPS	20150710	Linux	ALLMAPS is capable of computing a scaffold ordering that maximizes the colinearity to a collection of maps, including genetic, physical or comparative maps into the final chromosome build.	7/10/2015		detailed information
ALLPATHS-LG	52415	Linux	Illumina short reads assembly tool.	12/14/2011	1/9/2018	detailed information
AMOS	3.1.0	Linux	AMOS is a collection of tools and class interfaces for the assembly of DNA reads.	1/12/2013	1/14/2013	detailed information
AMPHORA	2	Linux	AMPHORA is an Automated Phylogenomic Inference Pipeline for bacterial sequences	7/26/2017	7/27/2017	detailed information

BioHPC Cloud Software

There is 583 software titles installed in BioHPC Cloud. The sofware is available on all machines (unless stated otherwise in notes), complete list of programs is below, please click on a title to see details and instructions. Tabular list of software is available here

Please read details and instructions before running any program, it may contain important information on how to properly use the software in BioHPC Cloud.

454 gsAssembler or gsMapper, a5, ABRicate, ABruijn, ABySS, AdapterRemoval, Admixtools, Admixture, agrep, albacore, Alder, AlleleSeq, ALLMAPS, ALLPATHS-LG, AMOS, AMPHORA, analysis, ANGSD, Annovar, antiSMASH, apollo, Arlequin, aspera, atac-seq-pipeline, athena_meta, Atlas-Link, ATLAS_GapFill, ATSAS, Augustus, AWS command line interface, axe, bamtools, bamUtil, Basset, BayeScan, Bayescenv, BBmap, BCFtools, bcl2fastq, BCP, Beagle, Beast2, bedops, BEDtools, bfc, bgc, bgen, bigQF, bigWig, bioawk, biobambam, Bioconductor, biom-format, BioPerl, BioPython, Birdsuite, Bismark, blasr, BLAST, blast2go, BLAT, BMGE, bmtagger, Boost, Bowtie, Bowtie2, BPGA, BreedingSchemeLanguage, breseq, brocc, BSseeker2, BUSCO, BWA, bwa-meth, cactus, canu, CAP3, cBar, CBSU RNAseq, CCTpack, cd-hit, CEGMA, CellRanger, centrifuge, CFSAN SNP pipeline, CheckM, chromosomer, Circlator, Circos, Circuitscape, CLUMPP, Clustal Omega, CLUSTALW, Cluster, cmake, CNVnator, compat, CONCOCT, Conda, copyNumberDiff, cortex_var, CRISPRCasFinder, CRISPResso, CrossMap, CRT, cuda, Cufflinks, cutadapt, dadi, dadi-1.6.3_modif, dDocent, DeconSeq, deepTools, defusion, delly, destruct, DETONATE, diamond, diploSHIC, Discovar, Discovar de novo, distruct, DIYABC, Docker, dREG, dREG.HD, Drop-seq, dropEst, dropSeqPipe, dsk, ea-utils, ecopcr, ecoPrimers, ectyper, EDGE, edirect, eems, EIGENSOFT, EMBOSS, entropy, ephem, ermineJ, ete3, exabayes, exonerate, eXpress, FALCON, FALCON_unzip, Fast-GBS, fasta, fastcluster, FastME, FastML, fastp, fastq_species_detector, FastQC, fastsimcoal26, fastStructure, FastTree, FASTX, fineRADstructure, fineSTRUCTURE, FIt-SNE, flash, flash2, flexbar, Flexible Adapter Remover, Flye, FMAP, FragGeneScan, FragGeneScan, freebayes, FunGene Pipeline, GAEMR, Galaxy, GATK, gatk4, GBRS, gcc, GCTA, gdc-client, GEM library, GEMMA, geneid, GeneMark, GeneMarker, Genome STRiP, GenomeMapper, GenomeStudio (Illumina), GenomicConsensus, gensim, germline, gffread, giggle, GMAP/GSNAP, GNU Compilers, GNU parallel, gradle-4.4, graftM, graphviz, Grinder, GROMACS, GSEA, GTFtools, Gubbins, HapCompass, HAPCUT, HAPCUT2, hapflk, HaploMerger, Haplomerger2, HapSeq2, HarvestTools, HiC-Pro, HiCExplorer, HISAT2, HMMER, Homer, HOTSPOT, HTSeq, HUMAnN2, hyperopt, HyPhy, iAssembler, IBDLD, IDBA-UD, IDP-denovo, IgBLAST, IGoR, IGV, IMa2, IMa2p, IMAGE, ImageJ, Immcantation, impute2, IMSA-A, INDELseek, infernal, InStruct, InteMAP, InterProScan, ipyrad, IQ-TREE, iRep, jags, java, jbrowse, jellyfish, JoinMap, julia, jupyter, kallisto, Kent Utilities, keras, khmer, KmerFinder, kraken, kSNP, kWIP, LACHESIS, lammps, LAST, lcMLkin, LDAK, leeHom, Lep-MAP3, Lighter, LINKS, LocusZoom, longranger, LUCY, LUCY2, LUMPY, lyve-SET, MACS, MaCS simulator, MACS2, MAFFT, mafTools, Magic-BLAST, MAKER, MAQ, MASH, MaSuRCA, Mauve, MaxBin, mccortex, mcl, megahit, MeGAMerge, MEGAN, MELT, MEME Suite, MERLIN, MetaBAT, MetaCRAST, metaCRISPR, MetAMOS, MetaPathways, MetaPhlAn, MetaVelvet, MetaVelvet-SL, MGmapper, Migrate-n, mikado, Minimac4, minimap2, mira, miRDeep2, MISO (misopy), MITObim, MixMapper, MKTest, MMAP, MMSEQ, mosdepth, mothur, MrBayes, mrsFAST, msId, MSMC, msprime, MSR-CA Genome Assembler, msstats, MSTMap, mugsy,



Details for **bamtools** (hide)

Name:	bamtools
Version:	2.5.1
OS:	Linux
About:	BAM file processing and filtering.
Added:	12/13/2011 2:56:40 PM
Updated:	2/27/2018 4:25:37 PM
Link:	http://seqanswers.com/wiki/BamTools
Platform:	all
Notes:	The latest version of the program is in your PATH, and can be used directly by typing its name at the prompt: `bamtools [options]` There are 2 previous version(s) available. Any version can be accessed either by typing full path, or by adding it to the PATH and then typing its name at the prompt. NOTE: you need to set PATH only once per login/session.

version	access	
2.5.1 (latest)	full path:	`/programs/bamtools-2.5.1/bin/bamtools [options]`
	add to PATH:	`export PATH=/programs/bamtools-2.5.1/bin:$PATH`
2.3.0	full path:	`/programs/bamtools-2.3.0/bin/bamtools [options]`
	add to PATH:	`export PATH=/programs/bamtools-2.3.0/bin:$PATH`
2.2.3	full path:	`/programs/bamtools-2.2.3/bin/bamtools [options]`
	add to PATH:	`export PATH=/programs/bamtools-2.2.3/bin:$PATH`

Details for **Trinity** (hide)

Name:	Trinity
Version:	2.8.4
OS:	Linux
About:	De novo transcriptome assembly.
Added:	12/13/2011 3:22:56 PM
Updated:	10/25/2018 10:13:24 PM
Link:	http://trinityrnaseq.github.io/

Notes:	You need to use full path to the binaries:
	`/programs/trinityrnaseq-Trinity-v2.8.4/Trinity [options]`
	You can also add the program to your PATH:
	`export PATH=/programs/trinityrnaseq-Trinity-v2.8.4:$PATH`
	and then use it directly by typing program name at the prompt.
	
	It is recommended to run on medium memory nodes (e.g., cbsumm03). A large memory machines (e.g. cbsulm05) may be needed to process larger datasets.
	If using the latest version of Trinity (2.8.4), run the following before launching the Trinity script:
	`export PATH=/programs/jellyfish-2.2.3/bin:`/programs/salmon-0.11.3/bin**`:$PATH`**
	(this needs to be done once per session, and is not necessary for earlier versions of trinity).
	To lauch Trinity, use the full path to the executable, like so:
	`/programs/`trinityrnaseq-Trinity-v2.8.4**`/Trinity [options] >& trinity.log &`**
	If you wish to use eXpress software to estimake read abundance via Trinity's utility **`align_and_estimate_abundance.pl`**, add the location of the eXpress executable to the PATH:
	`export PATH=/programs/express:$PATH`
	TransDecoder is not part of Trinity release starting from 2.1.1.
	Here is the instruction of running TransDecoder on BioHPC Lab: https://cbsu.tc.cornell.edu/lab/userguide.aspx?a=software&i=209#c

Software

You can “lock in” the program version by using full path or prepending your version to the PATH.

The commands to do it are always listed on the Lab program page.

Want to install software yourself?

- Many software titles can be installed in user space, especially Python or Perl.
- You can use Conda to install software in a separate environamnt.
- You can download and use Docker image. You can install software inside Docker container.

⇒“Software installation on Linux”

⇒“Using Docker in BioHPC Cloud “

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
- ✖ Verify that your software is available and read instructions
- ✖ Transfer data to your Lab storage
- ✖ Make reservation(s)
- ✖ Connect to reserved workstations
- *Compute!*