

BioHPC: A High Performance Computing System for Life Sciences

Bioinformatics Facility Biotechnology Resource Center Cornell University

The BioHPC Cloud has 4 types of computers

- **1.** Hosted Servers: Computers owned by individual labs but managed by the BioHPC team;
- 2. Rental Servers: Computers available for hourly rent by BioHPC users;
- 3. BioHPC Network Storage: Central storage accessible from each BioHPC server;
- **4.** Login Nodes: Computers used for accessing BioHPC data storage without reservation, and for BSCB users to access their cluster. The login nodes can also be used for accessing compute servers from outside campus without VPN. No computing is allowed on these machines.



On each server,

- Your home directory /home/<userID> is located on the BioHPC network storage. You can access the same home directory from any BioHPC computer. However, it is slower to read/write files in the home directory than the local storage. Do not compute directly on files in home directory. You also have /home2/<userID> network directory which is fast – you can compute directly on files in home2, you can access the same home2 directory from any BioHPC computer.
- Local storage: There is a /workdir on each server which is mounted on the local hard drives inside each server. It is much faster and much more reliable to access data files in the /workdir. Make sure to copy data files from your home directory to the /workdir before computing.

cbsumm01



Limit on storage size

- Your home and home2 directory: Up to 200 GB free storage for each account. You can purchase more storage if needed. The data files in your home directory are kept as long as your account is active.
- The working directory /workdir: There is no restriction on size, up to the local storage capacity of each server. On <u>rental machines</u>, the /workdir is cleared after your reservation ends. On <u>hosted</u> <u>machines</u>, files on /workdir are kept permanently.

Data backup:

- By default, **there is no second copy or backup** for files in your home directory. As the BioHPC storage system uses RAID, the risk of data loss due to hard drive failure is very low. However, if files are deleted due to human errors, or your account gets compromised, they cannot be recovered.
- For important data files, we recommend that you keep a second copy. BioHPC does provide optional data backup service on user-specified directories. More information on the backup service can be found at https://biohpc.cornell.edu/lab/userguide.aspx?a=backupguide .

The Firewall

 The three login nodes (cbsulogin, cbsulogin2, cbsulogin3) are open to the external network. You can access these three machines from anywhere in the world.

 The rest of BioHPC servers are behind the Cornell firewall. On the Cornell campus, they can be accessed directly. From outside Cornell, you can either connect through Cornell VPN (<u>https://it.cornell.edu/cuvpn</u>), or "ssh" via the login nodes.

Use the BioHPC Website to Manage User Access

https://biohpc.cornell.edu

Reserve a Rental Machine

	C Closul: valions					THE REAL PROPERTY
	gent. Sticzew, 1	and the second state		a) (n. 1) ann bhail (an 1920)	Concernances	100° - 1000 (17 Juni
	dand DECKAR	Annel 1970 - Maria Maria I., and an an 1970 - Maria Maria I., and an 1970 - Maria Maria I., and an 1970 - Maria Maria I., and an 1970 - Andrea Maria I., and an an an 1970 - Andrea Maria I., and an an an an an an 1970 - Andrea Maria I., and an	(and	Anna Bartana Anna Anna Anna Anna Anna Anna Anna	d and 1990 - State State and 1990 - State State State State 1990 - State State State State 1990 - State State State State 1990 - State State State State State 1990 - State State State State State State 1990 - State State State State State State State 1990 - State State State State State State State State 1990 - State State State State State State State State 1990 - State State State State State State State State 1990 - State State State State State State State State State 1990 - State State State State State State State State State State 1990 - State S	ine di Constanti ang
	and the second second		An annual state of the second state of the sec	And the second second		
		20.00 M		100.00.0	0.000	A
NO. 7 15	441411	441410	80.00	C2.61	10100	800 - 707
			200 - 20 F	100 M 1	1.01.000	807-001
1940 A 144			Annual Contraction		1000 C	And a local distance of the local distance o
			And the second s	10000	and the second s	And the second s
540 S 15		101000	100-000	CRA1	10100	100 C 100
teres a ter			Augusta .	*******		And a fee
			Paral I	100.004	10000	Autority .
			Contractor 1	10.00.00.0	Logo and States	And the second s
	21010	21010	50 - SV	P2041	27008	80.00
	10.0181	2010/00 2010/00	899-997	CRA1	10100	87.011
		21080 44100 41100	800-00 800-00	CEAL COLOR		807-007 807-007
	2010/01 4414/01 4414/01		100-00 100-00			
			100-00 100-00			

Change Password

BioHPC Clo	uu	I BIOINFORMATICS FACILITY
Password		BOINFORMATICS FACILITY
hange BioHPC Cloud Us	er password. Applies to all workstations and web pages.	
User:	qisun	
Old password:	=	
New Password:	•	
New Password:	•	
sually is in the path starti ubdirectory of /workdir o	rked storage cannot be used directly in computing. Networked storage includ ng with /home. The data residing on networked storage must be copied first to (/SD or /local/storage) and then computing must be carried out on that local n storage access for everyone and therefore may result in your programs being	o a local storage space on a given server (e.g. a storage. Excessive use of networked storage fo
	ser: aisun 'as24@cornell.edu' (BioHPC Cloud)	

Make Payment

				1	Trend Backers Factor
Furthers I		TED			
Lei-Coste A	ALC F	7			
Lei Cristit A	porant Nates	postpp_provid_			
Type .		Hears			
Paymail 172		Corrall Account			
		Catal Account Na	2014) Alban-Sylvill Albanisylvill	and stargest	
		Carnet Account O	ther Name:		
		Control: Account Of	viter Comell X4521		
Dechastre o office types o decos	etterne Horars dag (machiras et the) unit			Sarga" machines, Pierces are com	veneur, table to separadent hours for total
1994 (1794) a	(machines at the	leyed any "standard" has oftens of the page.	era aquad tu busuna su		
dana	(machines at the) sent	leyd yn "randed" hes oton of fie ysge hearr	era aquad to Jacasa en anti ecot	exercises of write	total
dana a	and the series of the series o	leyd wy 'vanded' ley ofner ef fie yegt lewe dithean	era aqual to Joces en anti ecol 800.23	examines of wests	1042 19122
datar a 0	Emaddenes at the 1 settl 40 hours 200 hours	laysi see 'standard' lace often of the year. Inters All locars 200 broas	anti otat 980.23 9302.78	number of units	1014 19523 5580.76

Mange VNC

	HPC y Re		aud: ation	5				-		i
1	-				inne er se de la			0.0		
	***	~	Register.	•	passes.	***	Paran .	-		r
	1000	100	1000		Conception of the local division of the loca		propagation for the set of the se	49	Const. The Associate	
-	100	12.2		•	and to ratify a first frame data ratify rational	10	program School 1, 8	-13	Const. The Associate	
-	1223	12.2			AND THE OWNER OF THE OWNER OWNER OF THE OWNER OWNER OF THE OWNER OWNE		quality (and the first product and a second		Const. The Asso Cold	
	1253	12.2	dealers for		AND THE OWNER OF THE OWNER OWNER OF THE OWNER OWNER OF THE OWNER OW	10.	Adamah (pada galama (1990) and Pala) ang Pangalan ang Pangalan ang Pangalan		Constant Coll.	*
ar.	1000	12.2			AND TAXABLE PARTY CARDINAL PARTY PARTY AND TAXABLE PARTY		hiland (an Ingen		Const The Area City	×
-	1000	100	deside the fit	-	and scratcher Last Last Lotter Last Lotter		program Theorem Stary B	-1	Const The local City	
	100	100	-		No. Instak		bland dagan gilim in ad bil að lýt að i gar íslan di gili gil længa	-1	Const The local City	*
**	123	22	dender -	line .	Include La March 19 Process Process Royal La Ha	жа.	land an e-failm staff latter to Kanada and	~	Construction (HC	-
w14	322	22	denter 1		Section 1988 Long Sections		provide all angle of the standard sector of the sector of	43.	Constant of Constant	
6.4	-	010			Constant Con	~	ten en discussement na Loci du Levi municis nul 16 si sul 16 man este municipare a constructivante statistica del se a Estate a suder na una estate estate para estate estate estate a constructiva de constructiva de la constructiva de la construct na una estate estate para estate estate estate a constructiva de constructiva de la constructiva de la construct	~	Construction of the	
	1000	22			A Tradition A Tradition And A Strategy		Propagation of the constrained start, which for the start of the constraint of the start of t		Care The Later	

Storage & Backup

Backup Shi	angr				N.	-	
		Talay (south				-	
In this Im		- All cost	e e adar				
least service	Rober Sect	trains.	Deserve	Malari	Annual		
lang.	\$1000		:	3	Section.	14 Sept. 14	King Talaw
Service and Services	Service years	24	1	2	we pull day	14 Sept. Sep	Kang Talawa
H Smith	See your		1	2	profession (M. Helder	Karang Kalon n
·		C-Labor					
the state of the s		$(-\infty)^{-1}$				fee a fa a fa	

Book Office Hours

Office H	lours		-	Report Section In
NUTE Office have Finane net up an ay limit. Due to high	festand for consultations we can C Cloud account please login to		e, bat you will have in confirm. I	1
Office	Binholematics Stall	Data	Right 3	Net 1
Thoday 413 Diselect21	Qt Ten Jed Glaskes	Transfey February 28 2019	10.00956-82.00956 80/11	ELEPH - ELEP
Thesias #15 Thesias #21	Qr. han Self Clauting	Monday March & 2018	03.00754-32.00754 BUTS	CLARNA - BLOOP
11 famil	Qillion Self Gloubite	Tuesday Marcin 1 200	01/0254-86/0254 #C251	04.00Phi - 00.00P
Rhudes #18 Rhodes #21	Qa Sun Judi Claubere	Thursday March 7 2019	03-00954 - 82-00954 90/51	RE SERVE - AN OWN
Rhodau Ald. Rhodau 421	Qi fun Sull Gaulete	Monder March 11 2019	es conse - es conse BUTU	CO DIPLA - ELLER BUDY
11 Breach	Q: than Jed Glaukete	Transfer March 12 2019	10-10754 - SA-00754 AVALABLE	14.00PM - 01-00P
Straige (1) Stades (2)	Q: has 3et Claules	Transfer March 14 2019	01.00954-82.00954 #VALABLE	CONTRACTOR
Dodas (1) Diedas (2)	Qt fran Self Glavitette	Monday March 18 2014	AVALABLE	12-00PM - 89-00P 20142-4912
33 Barris	Q: Non Self Claubra	Torotay March 18 2018	03-00754 - D4/00754 41/4/5-476.5	0400954-05009 20102-0015
Zitodes #18 Titodes #21	Qi fras Jeff Glaubre	Thursday Macils 21 2019	IN BRITH - B2-ROPEL AVAILABLE	42.0074-03.007 47.42.4813
Zhadse filt Biodas 621	Q: San Jef Claubra	Number March 25 2019	O.SOTA-SZOTA AVALARIS	ALLER ALLER
10. Borech	Q: Tran 3elf Gauline	Torontey March 28 2019	CO CONTRA - Description and the second	(MARTIN - DECOR
Ennio (1) Data (1)	Q1 from	Thursday March 28 2218	01.00754 - 82.00754	CLASPIC - 01-007

Quick start Step 1. Transfer data files to BioHPC

You can use one of the login nodes or any other BioHPC machines to transfer data files to your home directory. Any sftp client software will work, e.g. FileZilla (<u>https://filezilla-project.org/download.php?show_all=1</u>)

🔁 cbsum1c2b010 - sftp://qisun@	cbsum1c2b010.biohp	c.cornell.edu - Filez	Zilla							-		×
File Edit View Transfer Serv												
₩ • 8 • • • • • • •	1 🗱 🗓 표	TE 🔍 🍠 👌	6									
Host: sftp://cbsuvitisgen/ Usern	ame: qisun	Password:	••••••	Port: C	uickconnect	-						
Status: Disconnected from se Status: Creating directory '/h Status: Retrieving directory lis Status: Listing directory /hon Status: Directory listing of "/hon	ome/qisun/hapmap/fa sting of "/home/qisun/ ne/qisun/hapmap/fast	/hapmap/fast1" 1										~
Local site: c:\tmp\mydata\				· · · ·	Remote site:	/home/qisun/hapmap/fast1						~
	ing			~								^ ~
Filename	Filesize	Filetype	Last modified		Filename	V		Filesize F	Filetype	Last modified	Permis	ssions
fasta-data		File folder	3/2/2019 9:48:00 AM				Empty directo	ory listing				
1 directory					< Empty director	prv.						>
-	Direction Remote file	:	Size Prio	ority Status	1							
Queued files Failed transfer	rs (1) Successful tra	nsfers (1)										
									? 0	Queue: empty		•

Quick start Step 2. Reserve a rental machine

(skip this step if you are using a hosted machine)

- 1. Login to the BioHPC website (<u>https://biohpc.cornell.edu</u>) and click "Reservations";
- Reserve a computer from one of the categories. Each category is priced differently (<u>https://biohpc.cornell.edu/lab/Pricing.aspx</u>);
- 3. You can only access a machine while your reservation is active.
- 4. Reserve time longer than what is needed to run your job. You can adjust or cancel a reservation at any time through the "My Reservations" page.

BioHPC Cloud:

: Reservations

My Reservation

iny reserve	luona					
medium	gen1: 24 cores, 1	28GB RAM	7			
interactiv	e: 4 cores, 24GB	RAM	1			
general:	8 cores, 16GB RA	M				
medium g	gen1: 24 cores, 1	28GB RAM				
medium g	gen2: 40 cores, 2	56GB RAM	arch 2, 2019 09 ¥ : 55	▼) (am ▼) for machine (first available	with Lab credit account (jarekpp	general 'owner:jarekpp
	mory gen1: 64 co		a new reservation by clickingh on the apprpriate "AV/	AILABLE" text or in the box above.		
extra larg gpu gen2	je memory: 96-11	2 cores, 512GB RAM 2 cores, 1024GB RAM B RAM, 2x NVidia P100	bsumm02 inux (CentOS 7.4) ted Bare REC/SM 2U Twin Svr 4 Nodes 2600 Series 4 cores; 128GB RAM; 4TB HDD;1TB SSD; /// support: AVX	cbrumm03 Linux (CentOS 7.4) Red Barn RPGCSM 2U Twin Svr 4 Nodes 2600 Series 24 cores; 128GB RAM, 47B HDD;1TB SSD; VM support cleanacce security) AVX support: AVX	cbsumm04 Linux (CentOS 7.4) Red Barn RPGCSM 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 5 24 cores; 128GB RAM; 4TB 1 VM supported AVX support: AVX
restricted			WAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
San Mar 03 2019	db247 UNTIL 11:23 PM AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Mon Mar 04 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Tue Mar 05 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Wed Mar 06 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Thu Mar 07 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Fri Mar 08 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Sat Mar 09 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Sun Mar 10 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Mon Mar 11 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 ALL DAY	AVAILABLE	AVAILABLE
Tue Mar 12 2019	AVAILABLE	AVAILABLE	AVAILABLE	tad94 UNTIL 01:13 PM AVAILABLE	AVAILABLE	AVAILABLE

Quick start Step 3. Login to the computer and run software

- 1. Login to the compute server through ssh (e.g., Putty on Windows, Terminal on Mac);
- Create your working directory under /workdir and copy your data files there or use /home2/userid directory;
- You can find all pre-installed software on this page: <u>https://biohpc.cornell.edu/lab/labsoftware.aspx</u> (Use "*" as wildcard when searching for a software name). Follow instructions on each software detail page;
- If you do not find the software you want to use, email <u>brc bioinformatics@cornell.edu</u> or install the software by yourself in your home directory;
- 5. If you are using a rental machine, make sure to copy results back to home directory after the work is done. You can always access your home directory through the login nodes.

Miscellaneous information

 When using software like Python, R, Java, GCC, Docker, Tensorflow, etc. for the first time, please consult software details page under <u>https://biohpc.cornell.edu/lab/labsoftware.aspx</u> (use the search box to find the entry)

 Please read our user guide to find more information about data file sharing, using GUI software, etc.: <u>https://biohpc.cornell.edu/lab/userguide.aspx</u>

Miscellaneous information

• Use "screen" when running software;

- **Parallelization** running jobs on multiple cores:
 - perl_fork_univ.pl: perl_fork_univ.pl jobList 5
 - GNU parallel
 - Local scheduler (SLURM)

How to get help ?

Where to find more information?

- User guide https://biohpc.cornell.edu/lab/userguide.aspx
- Workshops https://biohpc.cornell.edu/workshops.aspx

Prefer to talk to a real person?

Book an office hour https://biohpc.cornell.edu/lab/office1.aspx

More questions?

Email: support@biohpc.ornell.edu

Happy Computing !

The BioHPC Team support@biohpc.cornell.edu