TCCGTGGGAGGCTCCGCGCGGCCGTTCCACGTCCAGAGGAACCACGCGACCGCGGCGCCGATCGCGACCGTCACCAGCAGCCCGGGT

:CODBLACXX:7:1101:2731:2202 2:N:0:ATTAC

Bioinformatics Facility

JJIIIJJJGIJJDHGIIJJAHFHHHFFDDEDDDDDDBB<BDDDDDDDDDBD@BDDC?@BBBBDCDDDDDD"BBD######## :CODBLACXX:7:1101:2690:2223 2:N:0:ATCACG TCCACGTCAACACGCCGAGGAAGAAGCCGCCGATGAGCGCGGGCGAGCCGGTGGCGACCACGCCGAGGAAGAAGCCGCCGATGAGCGCGGGCGAGCCGGTGGCGACCACGC

ute-1-12 halorubrum_genomes_2012]\$ perl shuffleSequences_fastq.pl 7_Index-1.SP9-2_replacement_R1.fastq 7_Index-1.SP9-2_replacement_R2.fas

Center

Biotechnology/Bioservices

CODBLACXX:7:1101:1513:2226 1:N:0:ATCACG

JJJIJJJJJJJJIJIIIFGIIIIJJJIJJJJJIIIIEHHGFFFFDDDBDD@B?4@CCCD>@C@BBB?CC@?B@?BCDDB>?BDDDDB; CODBLACXX:7:1101:1899:2244 1:N:0:ATCACG

CGAGGACGGGCGGCTCCTGGCCGGCGTCGCGGTCGGTCCGTTCGAGGAGCGCGCCGTCGCGGCCGACGAGGTGGACGTCGACGT

Co-Heads : J.P. Gogarten, Paul Lewis Facility Scientist : Pascal Lapierre Hardware/Software Manager: Jeff Lary

JJJJIJJJJJJJJJJJJJJJHFHHFFFFFCCEDDDDDDDDDDDDDDDDDDDDDDDDDDBBBDECEDCC<@BDDBDDBDDBBBDBABBDBB>B :CODBLACXX:7:1101:2571:2193 1:N:0:ATCACG

Mandate of the Facility:

To provide computational power and technical support to both academia and industry. These services are available, free of charge, to faculty and students within the University system and are available at negotiated rates to other academic institutions

:CODBLACXX:7:1101:2174:2221 1:N:0:ATCACG

CGAGCGGTACGGGGTTCGGGCCGCCGAGGACGCTGACGGGACCGGAGTTCCGGGAGAGGGCGCGACGAGGGGCGACGCGGCGGCG

DDD> GEDDDA The usage of the cluster is for research



UIJJJJJJJJJJJJJJJGIJIJJJJJJ##-7CHIJHFFD#,;?BBBDDDDDDDDDDDDDD CODBLACXX:7:1101:2430:2247 1:N:0:ATCACG

\$AACTGGCCGCGGAACTGGAGCTGGATGACGAGGAGGTTCACGACCGGATCGCGTATCTCTCAACGTTCGACCGGGTCCGTCGGGATG

JJJJIJJJJJJJJJJJJJJJHFHHFFFFFCCEDDDDDDDDDDDDDDDDDDDDDDDBBBDECEDCC<@BDDBDDBDDBBBDBABBDBB>B :CODBLACXX:7:1101:2571:2193 1:N:0:ATCACG

GGTCGGCCTGATAGCCGGCTACGACNNCGGTCTCCCGCCCGATCGCGAGGCGGAACTCGCCCGCGACGCGGCGGCGGCCGTCCTCGTCGT

TCCGTGGGAGGCTCCGCGCGGCCGTTCCACGTCCAGAGGAACCACGCGACCGCGGCGCCGATCGCGACCGTCACCAGCAGCCCGGGT

ededdehg38F8F8';A;;AAH<<?@BDD;9:BD855??EB6heeB6>Fardyace:

TCCACGTCAACACGCCGAGGAAGAAGCCGCCGATGAGCGCGGGCGAGCCGGTGGCGACCACGAGGACCCCGGCGAGGTCGCTCCCGA

Dute-1-12 he or both 17 node Dell Linux cluster running Redhat EL5. Each accent_R2 fas accode Later running Redhat EL5. Each accent_R2 fas accenter running Redhat EL5. Each accenter running Redhat EL5. Each accenter running Redhat et acce

• 18-node, 36 processor Apple Xserve Cluster

JJJJJJJJJJJJJJIJIIIFGIIIIJJJIJJJJJIIIIEHHGFFFFDDDBDD@B?4@cccd>@c@BBB?cc@?B@?BcDDB>?BDDDDB> codblacxx:7:1101:1899:2244_1:N:0:ATCACG AGGACGGGcc•ctcctSmallcLinux=basedssatellitesclusterc

TCCGTGGGAGGCTCCGCGCGGCCGTTCCACGTCCAGAGGAACCACGCGACCGCGGCCGATCGCGACCGTCACCAGCAGCCCGGGT



@D@DDEHG3BFBFB';A;;AAH<<?@BDD;9:BDB55??BD8B8@C9?B-8BB>B@>955?#; ############### CODBLACXX:7:1101:2671:2213 2:N:0:ATCACG

Help can be found on the Uconn Bioinformatics Wiki page : http://137.99.47.91/wiki/index.php/Main_Page

ute-1-12 halorubrum_genomes_2012]\$ perl shuffleSequences_fastq.pl 7_Index-1.SP9-2_replacement_R1.fastq 7_Index-1.SP9-2_replacement_R2.fas

To log on the servers :

CODBLACXX:7:1101:2430:2247 1:N:0:ATCACG

bbcxsrv1.biotech.uconn.edu (Xserve cluster) bbcsrv3.biotech.uconn.edu (Dell server)



sername for accounts: actocggctacggccacttcgagt

Passwords are

When you login for the first time, create a new password by typing : passwd

Unix basic commands:

Is : List file of a directory

- cd : Change directory
- mkdir : make directory
- cp:copy
- - mv: move
 - more/less/cat: view content of a file
- man <commands>: manual of a command
 - pwd : display current path
- up arrow : cycles through previous commands tab : autocompletions of file names

Unix advance commands:

- Isload (Dell only) : See the current CPU loads
- Islogin (Dell) or rlogin (Xserver) : log to a sub-node
- qstat : Display the status of the queue
 - qsub : Submit a script to the queue (qsub perl run.pl)
 - bjobs: Display the status of your jobs (if any)
 - ssh compute-1-x (Dell) or ssh nodex (Xserver) : To

manually log on a sub-node ?BC ?BCDDB>?BDDDDB>

qdel : To terminate a job running on the queue ps ux: Display processes status



JJIIIJJJGIJJDHGIIJJJAHFHHHFFDDEDDDDDBB<B

VTCCACGTCAACACGCCGAGGAAGAAGCCGCCGATGAGCGCGGGCGAGCCGGTGGCGACCACGAGGACCCCGGCGAGGTCGCTCCCGA

:CODBLACXX:7:1101:1372:2240 1:N:0:ATCACG

ATCGACGGCGGCCGGCTCCACCGCGAGTTCGTCCTCCACAACAAGGCGCTCGTCGGCAGCGTCAACTCCGGCTACGGCCACTTCGAGT

JGEAGAHIFBHHBeasto?c?B@B@CDDDDDBDDI	
BLAST Suite	ассааааас HyPhy этссааааассаааааа
DBLACXX:7:1CLUSTALWACG	Mauve
DarkHorse	DBDD@B?4@CCCMOCHTCTCACGFFCFCGFGGFGAACGCC
DBLACXX:7:1EMBOSS 0:ATCACG	BAGGAGCGCGCCMrBayes
JJIIJJJIHDDD FastTree BDDDBDDDDDDDDD	DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD
BAGCCGTACGGCGCGCGCGCGCGCGCGCCCGGCGCCC	GAGTTCCCGGAC MUMCACGACGCCCACGCGGCCGCC
IDBLACXX: 7: IgsAssembler	PAML

nhPhyML Paup* Phycas PHYLIP Phylobayes PhyML RaxML

GTCGGCCTGATAGCCGGCTACGACNNCGGTCTCCCGCCCGATCGCGAGGCGGAACTCGCCCGCGACGCGGCGGCGGTCCTCGTCGT

8@D@DDEHG3BFBFBF';A;;AAH<<?@BDD;9:BDB5 :CODBLACXX:7:1101:2671:2213 2:N:0:ATC GAACCTCGTCGATATCGGGGTGCTCGCACAGGCGGCGCC



JJIIIJJJGIJJDHGIIJJJAHFHHHFFDDEDDDDDDBB<BDDBDDDDDDDDBD@BDDC?@BBBBDCDDDDDDBBD########## CODBLACXX:7:1101:2690:2223 2:N:0:ATCACG

• There are no official limits on the number of jobs you can run on the cluster but...please refrain from using all the nodes a the same time.

- NEVER run anything on the head node of a cluster (default login node).
- Keep track of what you are running and where so if something goes wrong, you can go back and kill the desired job.

 Each nodes have their own hard drive (scratch drive). It is advised when possible to run and write your output files on this drive (/scratch), then copy the file back to your home directory when done.

CODBLACXX:7:1101:2571:2193 1:N:0:ATCACG GTCGGCCTGATAGCCGGCTACGACNNCGGTCTCCCGCCCGATCGCGAGGCGGAACTCGCCGCGACGCGGCGGCGGCGGCCGTCCTCG

:CODBLACXX:7:1101:2731:2202 2:N:0:ATTACC

DEHGSBEFBE A RUNNING MPI JOBS ON THE DELL

oute-1-12 halorubrum_genomes_2012]\$ perl shuffleSequences_fastq.pl 7_Index-1.SP9-2_replacement_R1.fastq 7_Index-1.SP9-2_replacement_R2.fas .ed.fastq

B:CODBLACXX:7:1101:1372:2240 1:N:0:ATCACG

ATCGACGGCGGCCGG

GIIJGEAGAHIFBHH 8:CODBLACXX:7:1 CGCGCCGTCAAGAAG

HJJJJJJIIIJIJJJJ 3 : CODBLACXX : 7 : 1 FGCCCTCGTCTCGCC

HJJJJJJJJJJJJJJJJ 3 : CODBLACXX : 7 : 1 CGAGGACGGGCGGCT

HJJJJIIJJJIHDDD CODBLACXX:7:1 GGGAGCGGTACGGG

DEE:?E<?@:?DDD> CODBLACXX:7:1 CGTCGTCTCGCTGG

IJJIJJJJJJJJJJJJJ 3 : CODBLACXX : 7 : 1 GAACTGGCCGCGGAA

IJJJJIJJJJJJJJJJJ }:CODBLACXX:7:1 \GGTCGGCCTGATAG -Everything you need is on the Wiki page with example shell scripts.

-128 available CPUs, used at about 50% in the last few weeks.

- What is MPI?

Message Passing Interface (MPI) is a specification for an API that allows many computers to communicate with one another.

HJJJJHIJJJJJIJI 3 : CODBLACXX : 7 : 1 AGATCGATTGATACG

G8@D@DDEHG3BFBF B:CODBLACXX:7:1 GGAACCTCGTCGATA

HJJIIIJJJGIJJDE CODBLACXX:7:1 CCACGTCAACACG

UJJJJJJIIGGIGIJ oute-1-12 halor ed.fastq :CODBLACXX:7:1 UTCGACGGCGGCCGG

IIJGEAGAHIFBHH CODBLACXX:7:1 GCGCCGTCAAGAAG

HJJJJJJIIIJIJJJJ 3 : CODBLACXX : 7 : 1 FGCCCTCGTCTCGCC

HIJJJIJJJJJJJJIJ 3 : CODBLACXX : 7 : 1 CGAGGACGGGCGGCT

IJJJIIJJIHDDD 3 : CODBLACXX : 7 : 1 ACGGAGCGGTACGGG

DEE:?E<?@:?DDD> 3:CODBLACXX:7:1 CCGTCGTCTCGCTGG

HJJIJJJJJJJJJJJJJ 3 : CODBLACXX : 7 : 1 HAACTGGCCGCGGAA

HJJJJIJJJJJJJJJJJ 3 : CODBLACXX : 7 : 1 AGGTCGGCCTGATAG

AT THIS TIME THERE ARE A FEW MPI-ENABLED APPLICATIONS ON THE DELL CLUSTER (BBCSRV3):

- CLUSTALW-MPI (CLUSTALW-MPI VO.13, BASED ON CLUSTALW V1.82)

MB-MPI (MRBAYES V3.1.2)

PHYML-MPI (PHYML V3.0)

- R (2.9.2)

MPI-BLAST

cement R2 fas

TCCGTGGGAGGCTCCGCGCGGCCGTTCCACGTCCAGAGGAACCACGCGACCGCCGACCGCGACCGTCACCAGCAGCCCGGGT

:CODBLACXX:7:1101:2731:2202 2:N:0:ATTAC

BEFORE YOU RUN YOUR FIRST MPI JOB,

MAKE SURE THAT THE FOLLOWING CODE IS

CODBLACXX:7:1101:2690:2223 2:N:0:ATCA N YOUR BASHRC FILE:

module load mpi/openmpi-interconnects-gnu

ATCCACGTCAACACGCCGAGGAAGAAGCCGCCGATGAGCGCGGGCGAGCCGGTGGCGACCACGAGGACCCCGGCGAGGTCGCTCCCGA

Load saved modules

HJJJJJJIIGGIGIJJJJIIIJJHHHG oute-1-12 halorubrum genomes

led.fasto

B:CODBLACXX:7:1101:1372:2240

ndex-1.SP9-2_replacement_R2.fas

GIIJGEAGAHIFBHH9BBCEEDDDDDD0?C?B@B@CDDDDDDDDDDDDDDDBBDBBDD0;<58@B@ACDDDDBBCDDB@B@CDC@A

:CODBLACXX:7:1101:1513:2226 1:N:0:ATCACG

:CODBLACXX:7:1101:1595:2236 1:N:0:ATC/CG

-A	#!/bin/bash	
JI	#BSUB -q normal # submit the job to the normal queue, which is the default queue	
	#BSUB -o /home/ <yourusername>/clustalw-mpi-%J.o # name the output file; %J inserts the current job number</yourusername>	
	#BSUB -e /home/ <yourusername>/clustalw-mpi-%J.e # name the error file; %J inserts the current job number</yourusername>	
DD	#BSUB -J mpi-job # give the job a jobname, mpi-job	
11	#BSUB -n 4 # define the number of processors to use	
GG	#BSUB -a openmpi # define the type of MPI to use.	
>:	#	
11	cd /home/ <yourusername>/clustalw-mpi-0.13/</yourusername>	
GC	, # · · · · · · · · · · · · · · · · · ·	
ы	mpirun.lsfmca btl ^openib ./clustalw-mpi -infile=CFTR.input -newtree=CFTR.mytree	
1.1	# # themca btl ^openib part of the command line is telling mpirun	
AC	# to exclude using infiniband in the byte transfer layer (btl)	

JIJJJJJJJJJJJJJJHFHHFFFFFCCEDDDDDDDDDDDDDDDDDDDDDDDBBBDECEDCC<@BDDBDDBBBDBABBDBB>B

CODBLACXX:7:1101:2571:2193_1:N:0:ATCACG

GGTCGGCCTGATAGCCGG C'I ACG AC N IC 3GTCTCCCG ICC 3AT CHCGACGCGG AA I'I CHCCCCCG ACGCGGCGGCGGTCCTCGTCGT

TCCGTGGGAGGCTCCGCGGCCGTTCCACGTCCAGAGGAACCACGCGACCGCGGCCGATCGCGACCGTCACCAGCAGCCCGGGT

CODBLACXX:7:1101:2671:2213 (#!/] #BS #
JJIIIJJJGIJJDHGIIJJJAHFHHHFFI CODBLACXX:7:1101:2690:2223 FCCACGTCAACACGCCGAGGAAGAAGCC	cd / phy

#!/bin/bash

SUB -o /home/<USERNAME>/outputfile # Put output and errors in file outputfile

d /home/<USERNAME>/<working directory>/ hyml -i somefile.phy -d nt -b 20 -m JC69 -v 0 -c 4 -a e -s BEST -o tlr

pute-1-12 halorubrum_genomes 2	2012]\$ perl shuffleSequences fastq.pl 7 Index-1.SP9-2 replacement R1.fastq 7 Index-1.	SP9-2_replacement_R2.
B:CODBLACXX:7:1101:1372:2240 1	UB SCRIPTS TO RUN ON SCRATCH DRIV	- :
ATCGACGGCGGCCGGCTCCACCGCGAGTT(GIIJGEAGAHIFBHH9BBCEEDDDDDD0?(8:C0DBLACXX:7:1101:1513:2226 :	<pre>#! /bin/bash #BSUB -B # Send mail at beginning of job execution #BSUB -N # Send mail at end of job execution</pre>	
CGCGCCGTCAAGAAGGTGTGCGTCCGGAT(#BSUB -u Firstname.Lastname@uconn.edu # users email destination	
HJJJJJIIIJIJJJJIBGEHIIIJIIIJHI	#BSUB -J myjob # Give the job the name 'myjob'	
8:CODBLACXX:7:1101:1595:2236	#BSUB -o outputfile # Put output and errors in file outputfile	
ICCCTCOTCTCOCCATACCCOATCOAAC	if [!-d "/scratch/\$USER"]: then	
HIJJJIJJJJJJJJIIIFGIIIIJJJIJ	mkdir /scratch/\$USER	
CGAGGACGGGCGGCTCCTGGCCGGCGTCG	fi	
	#	
	if [! -d "/scratch/\$USER/subdirname"]; then	
ACGGAGCGGTACGGGGTTCGGGCCGCCGA(mkdir /scratch/\$USER/subdirname	
	fi	
DEE:?E @:?DDD :6@DDDA'554,3;	#	
B:CODBLACXX:7:1101:2262:2190 : CCGTCGTCTCGCTGCCGTACTACTTCNNC(cd /scratch/\$USER/subdirname # cd to the working directory in the scratch area	
	cp \$HOME/myprog.	
HJJIJJJJJJJJJJJJJJGIJIJJJJJ##	cp \$HOME/data/inputfile1.	
B:CODBLACXX:7:1101:2430:2247 1	./myprog < inputfile1 > outputfile1	
BAR TOOCCOCOGAACTOGAOCTOGATGA		
HJJJJIJJJJJJJJJJJJJHFHHFFFFFCCI	cp outputfile1 \$HOME/savedir # make sure that \$HOME/savedir exists!	
B:CODBLACXX:7:1101:2571:2193 :	#	
ADD I COOCCI IGA I AGCCOUCI ACGACINICI	rm -t / scratcn/ \$USEK/ subdirname/ "	
HIJIIJJJJJJJJJJJJHIJJJHHFEEFD##.5	32BDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	

CODBLACXX: 7:1101:2671:2213 2:N:0 ANALYSIS OVER MULTIPLE CPUS

IIIJJJGIJJDHGIIJJAHFHHHFFDDEDDDDDDBB<BDDDDDDDDDDDB0@BDDC?@BBBBDCDDDDDDBBD#########

#BSUB -J test2[1-575]%40 #Will cycle from 1 to 575 and start processes over a maximum of 40 processors are reached

ement R2.fas

#BSUB -o bootjob%J.log #Create log files named bootjob1.log to bootjob475.log

#BSUB \$LSB_JOBINDEX

perl ~/map_algor/bootstraps/boot_sphere.pl \$LSB_JOBINDEX