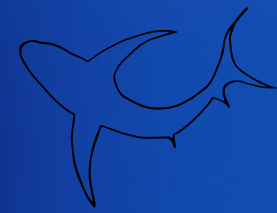




LEIDS UNIVERSITAIR MEDISCH CENTRUM



Shark Cluster Overview

51 Execution Nodes

1 Head Node (shark)

2 Graphical login nodes

800 Cores = slots

714 TB Storage RAW



```

      .'.
0MMMMMMWKxc.      'LOWMMNc
0MMMMMMMMMMMMW0d:.;xWMMMMWl
lWMMMMMMMMMMMMMMMMMMMMMM'
.xWMMMMMMMMMMMMMMMMMMMMMW
.onMMMMMMMMMMMMMMMMMMMo
cXMMMMMMMMMMMMMMMMMMO;.      .;xNMMWWMX
cMMMMMMMMMMMMMMMMMMMMN0o;'.....';lknMMMMWx. .;l
cMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMNo.
dMMMMMMWXK00XWMMMMMMMMMMMMMMMMMMMMMMWk,
xMMmk'      .,cdOKNMMMMMMMMMMXd,
dMMN'      .....k
.WN.
;;

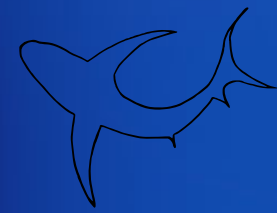
```

Welcome to the Shark cluster @ the research network
This cluster is still under development.
Please report problems to M.P.Villierius.





LEIDS UNIVERSITAIR MEDISCH CENTRUM



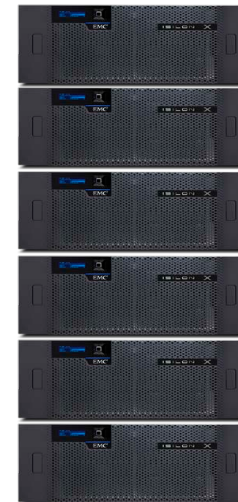
Introduction

What is a High Performance Compute (HPC) cluster?

- A HPC cluster is a group of servers all connected together with a network.
 - The group can work as one large computer or as single computers called nodes.
 - This cluster is controlled by a **resource manager** Open Grid Scheduler (OGS)

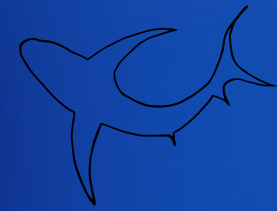
There are different type of nodes

- Head/master/login node
 - Controls/manage the complete cluster
- Compute node
 - A single blade server which does one thing compute





LEIDS UNIVERSITAIR MEDISCH CENTRUM



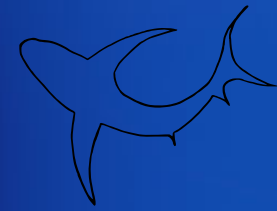
Open Grid Scheduler

Resource manager

- OGS is a scheduler which divides the available resources(CPU,mem,etc) into queues
- OGS places the compute tasks (jobs) in a large waiting list.
- OGS searches the right queue (all.q) for the requested resources in a job and starts that job.
- Jobs with a higher priority are placed higher in the list
- The priority is based on the priority resource and the fair-share policy



LEIDS UNIVERSITAIR MEDISCH CENTRUM



Connecting to Shark head node

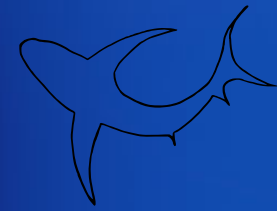
- Connect to Shark with a ssh client
 - Within LUMC
 - **ssh** RESEARCHLUMCusername@145.88.65.151
 - Outside the LUMC use ssh jump server
 - **ssh** RESEARCHLUMCusername@145.88.35.10
 - **ssh** 145.88.65.151

Start an interactive job **qlogin**

Start submitting batch jobs **qsub**



LEIDS UNIVERSITAIR MEDISCH CENTRUM



Connecting to Shark GUI node

Only from the LUMC:

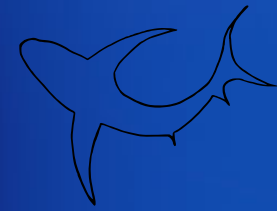
Open your Remote Desktop Connection Client

fill in for computer name : 145.88.65.200
145.88.66.103

Use your {researchlumc username} and password to login.

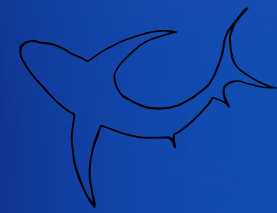


LEIDS UNIVERSITAIR MEDISCH CENTRUM



Resources

- Storage / disk space
- CPU / cores / slots
- Memory
- Runtime



Storage

Directory	Description	Size
/exports	Mount point for HPC Isilon storage, every department has bought their own export that is mounted here. HPC isilon storage has no backup!!!!	714TB
/exports/archive	Long Term Storage	~1P
/home/	Home dir, can hold: scripts, R library, compiled code	2TB
/bam-export	Everyone can save BAM files here that can be reached from the UCSC genome browser.External address : https://barmsijs.lumc.nl/.....	2TB

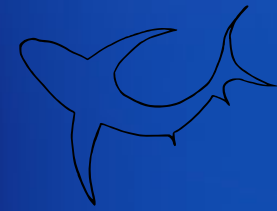
Location of important directories

- **/usr/local** *Where programs are located*
- **/usr/local/Genomes** *Where the reference genomes are located*

Command to check storage limits : `df -kh`



LEIDS UNIVERSITAIR MEDISCH CENTRUM

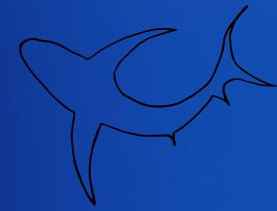


CPU / cores / slots

- **CPU** = Central Processing Unit
- **Multi-core-processor** = single chip containing two or more CPUs called CORES
- **Slot** = unit from Open Grid Scheduler that defines one core
- On Shark : 1 slot = 1 core = 1 cpu
- Number of slots are bound by execution node
- Command to find out number of slots per node = **qhost**
- Submit with more then 1 slot use -pe BWA {nr. Of slots}
 - `qsub -pe BWA 4 /home/noname/test.sh`



LEIDS UNIVERSITAIR MEDISCH CENTRUM

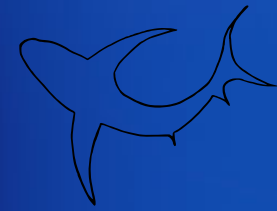


Memory

- The maximum amount of memory you can use is execution host related
 - Check the amount of memory per node with the command **qghost**
- The amount of memory you specify is per slot
 - `qsub -pe BWA 2 -l h_vmem=10G` (you ask for 20G mem in total)
- The default amount of memory per job is 3G
- If you need more memory you have to specify the amount
- If you exceed the requested amount of memory your job will be killed



LEIDS UNIVERSITAIR MEDISCH CENTRUM



Environment Modules

- The Environment Modules package sets your environment to use a specific software/pipeline

#module list

Prints Currently Loaded Modulefiles

#module avail

Lists available modules

#module load R/3.0.2

Load the R 3.0.2 module

#module unload R/3.0.2

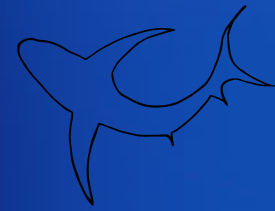
Unload the R 3.0.2 module

#module switch R/3.1.1

Switch to R 3.1.1

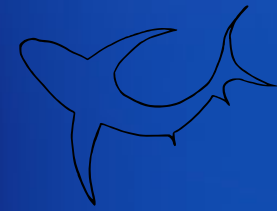


LEIDS UNIVERSITAIR MEDISCH CENTRUM



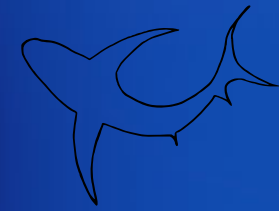
Job submission

- Interactive job submission
 - Command : **qlogin**
- Batch mode job submission
 - Command : **qsub**
- To request resources you need to set command options



Scheduler submit options

- Some useful submit options.
 - -V pass on all current environment variables to the job
 - -N <jobname> give your job a name.
 - -cwd run in current directory
 - -o <path/output_filename> path to and name for output file
 - -e <path/error_filename> path to and name for error file
 - -m <b e a s n> begin,end,abort,suspended,no mail is send
 - -M <email address>
 - -w <e,w> validate the options, error and reject, warning
- Resource options
 - -l h_vmem=<size> specify the maximum amount of memory to use per slot
 - -pe <name> <nr._slots> specify the amount of slots to use
 - Print the parallel environment names : **qconf -spl**
 - -l h_rt=<hh:mm:ss> specify the maximum run time



Interactive use of OGS

username@[shark](#):~\$ **qlogin**

Your job 822125 ("QLOGIN") has been submitted

waiting for interactive job to be scheduled ...

Your interactive job 822125 has been successfully scheduled.

Establishing /usr/local/Scripts/qlogin_wrapper session to host [angelshark.researchlumc.nl](#) ...

username@[angelshark](#):~\$ **qstat**

Job-ID	prior	name	user	state	submit/start at	queue	slots	ja-task-ID
822125	0.50075	QLOGIN	username	r	06/04/2014 13:17:25	qlogin.q@ angelshark.researchlumc.nl	1	

username@[shark](#):~\$ **qlogin -pe BWA 4**

Your job 110997 ("QLOGIN") has been submitted

waiting for interactive job to be scheduled ...

Your interactive job 110997 has been successfully scheduled.

Establishing built in session to host [angelshark.researchlumc.nl](#) ...

vill@[angelshark](#):~\$

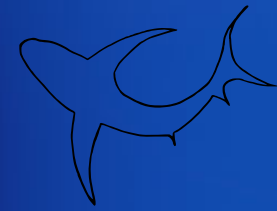
username@[angelshark](#):~\$ **qstat**

job-ID	prior	name	user	state	submit/start at	queue	slots	ja-task-ID
110997	0.53227	QLOGIN	vill	r	08/03/2011 03:13:59	qlogin.q@ angelshark.researchlumc.nl	4	

Important Note : default memory of every job is set to
3GB / slot



LEIDS UNIVERSITAIR MEDISCH CENTRUM



Qsub your job

Only a script can be submitted with qsub !!!!!!!

Syntax qsub command:

qsub [options] [scriptfile | -- [script args]]

Binary files can not be submitted directly to SGE.

Example submit "date" command to OGS (with script):

#!/bin/bash

/bin/date

Submit this script with:

\$ qsub sge-date

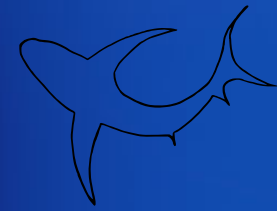
OGS will then run the program, and place two files in your home directory:

sge-date.e sge-date.o

Important Note : default memory of every job is set to 3GB / slot
-l h_vmem=3G



LEIDS UNIVERSITAIR MEDISCH CENTRUM



Qsub your job

Qsub options can be placed inside your script.
username@shark:\$ cat run_my_first_job.sh

```
#!/bin/bash
#$ -S /bin/bash
#$ -q all.q <= use with care
#$ -N my_first_job
#$ -pe BWA 4 <==== the number 4 is only if you need 4 CPU's
#$ -l h_vmem=4g <== maximum memory that can be used.
#$ -cwd
#$ -j Y
#$ -V
#$ -m be
#$ -M email@address.lumc
```

```
echo Start time : `date`
/bin/hostname
echo End time : `date`
```

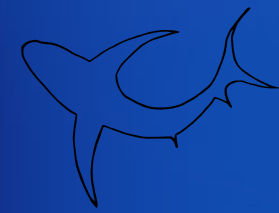
username@shark:\$ **qsub run_my_first_job.sh**

Your **job 220282** ("my_first_job") has been submitted

username@shark:~/shark/trunk\$ **qstat**

job-ID	prior	name	user	state	submit/start at	queue	slots	ja-task-ID
220282	0.60500	my_first_j	username	r	09/07/2011 09:26:55	all.q@kitefinshark.cluster.loc	4	

Important Note : default memory of every job is set to 3GB
-l h_vmem=4G



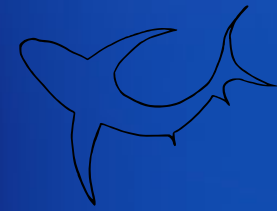
username@shark:~/shark/trunk\$ **qacct -j 220282**

```
=====
qname      all.q
hostname   kitefinshark.cluster.loc
group      NexGenSeq
owner      vill
project     NONE
department defaultdepartment
jobname     my_first_job
jobnumber   220282
taskid      undefined
account     sge
priority    20
qsub_time  Wed Sep  7 09:26:53 2011
start_time Wed Sep  7 09:26:54 2011
end_time   Wed Sep  7 09:27:15 2011
granted_pe  BWA
slots       4
failed     0
exit_status 0
ru_wallclock 21
ru_etime     0.040
ru_stime     0.010
ru_maxrss    1500
ru_ixrss     0
ru_ismrss    0
ru_idrss     0
ru_isrss     0
ru_minflt    2332
ru_majflt    0
ru_nswap     0
ru_inblock   16
ru_oublock   192
ru_msgsnd    0
ru_msgrcv    0
ru_nsignals  0
ru_nvcsw     148
ru_nivcsw     5
cpu        0.050
mem        0.000
io         0.000
iow        0.000
maxvmem    30.523M  <= note the maximum memory that has been used
arid         undefined
```

Important Note : default memory of every job is set to 3GB

Job has been submitted with 4GB times 4 slots = 16GB can be submitted with 15MB

-l h_vmem=15m



Summary

- **qsub** submit a job to the batch scheduler
- **qlogin** submit an interactive login session to Grid Engine
- **qstat** examine the job queue
- **qdel** delete a job from the queue
- **qacct** report and account for Grid Engine usage
- **unique-user-slots**
- /home Home dir do not store DATA here
- /data/... for storage of your data
- /usr/local/Genomes/ dir where ref. Files are located
- /bam-export dir for temp. bam files and share files.
- <https://barmsijs.lumc.nl/>