# How do I figure out how many resources I need for my job?

Cat/monkey meme: https://i.pinimg.com/originals/11/cc/fa/11ccfaaa09c179e322f1e61529d038a7.jpg
HPC animal: https://hpcc.usc.edu/about/

# Do I even need to figure out how many resources I need for my job?

One-off jobs will not be worth the time to figure out exactly how many resources they need


https://www.tate.org.uk/art/artworks/creed-work-no-890-dont-worry-ar01149

# Do I even need to figure out how many resources I need for my job?

One-off jobs will not be worth the time to figure out exactly how many resources they need

If your job ‘checkpoints’ (e.g. writes out files as it goes, that it can restart from if stopped), then if it hits the walltime, not such a biggie.


https://www.tate.org.uk/art/artworks/creed-work-no-890-dont-worry-ar01149

# Why should I figure out how many resources I need for my jobs?

**Selfish benefits:**

- If you request exactly what you need, your job will run as quick as it can (if you request too many resources, your job will sit in the queue for longer than it would if you requested less)

# Why should I figure out how many resources I need for my jobs?

**Selfish benefits:**

- If you request exactly what you need, your job will run as quick as it can (if you request too many resources, your job will sit in the queue for longer than it would if you requested less)
- If you request too few resources, your job will not go to completion

# Why should I figure out how many resources I need for my jobs?

**Selfish benefits:**

- If you request exactly what you need, your job will run as quick as it can (if you request too many resources, your job will sit in the queue for longer than it would if you requested less)
- If you request too few resources, your job will not go to completion
- If you request too many resources, you will burn through your allocation more quickly and your jobs will lose priority

# Why should I figure out how many resources I need for my jobs?

**Selfish benefits:**

- If you request exactly what you need, your job will run as quick as it can (if you request too many resources, your job will sit in the queue for longer than it would if you requested less)
- If you request too few resources, your job will not go to completion
- If you request too many resources, you will burn through your allocation more quickly and your jobs will lose priority

**Unselfish benefits:**

- Allows more efficient use of the cluster resources

# How do I figure out how many resources I need for my jobs?

For your first job in your set of jobs:

1) Request too few resources, have your job time out/hit the memory limit (OOM) and not complete, and then request more resources next time

2) Request too many resources, profile your job, reduce the amount of resources you need for next time

# How do I figure out how many resources I need for my jobs?

For your first job in your set of jobs:

1) Request too few resources, have your job time out/hit the memory limit (OOM) and not complete, and then request more resources next time

2) Request too many resources, profile your job, reduce the amount of resources you need for next time

# Example: filtering a genome assembly for contigs above a certain length

South Island robin

**Why do inbred males fire blanks? Unravelling the relationship between inbreeding and infertility**

MARSDEN FUND

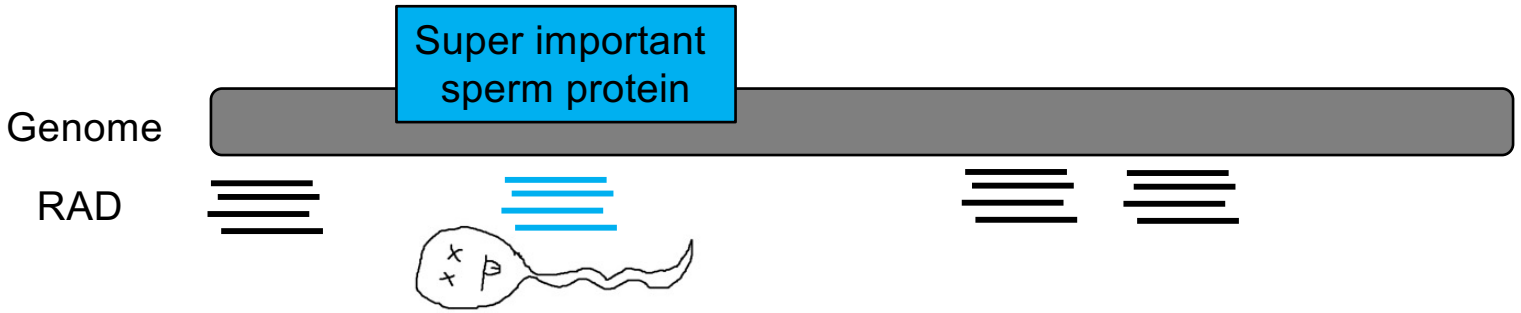
TE PŪTEA RANGAHAU A MARSDEN

https://commons.wikimedia.org/wiki/File:South_Island_robin,_Lewis_Pass.jpg

# Example: filtering a genome assembly for contigs above a certain length

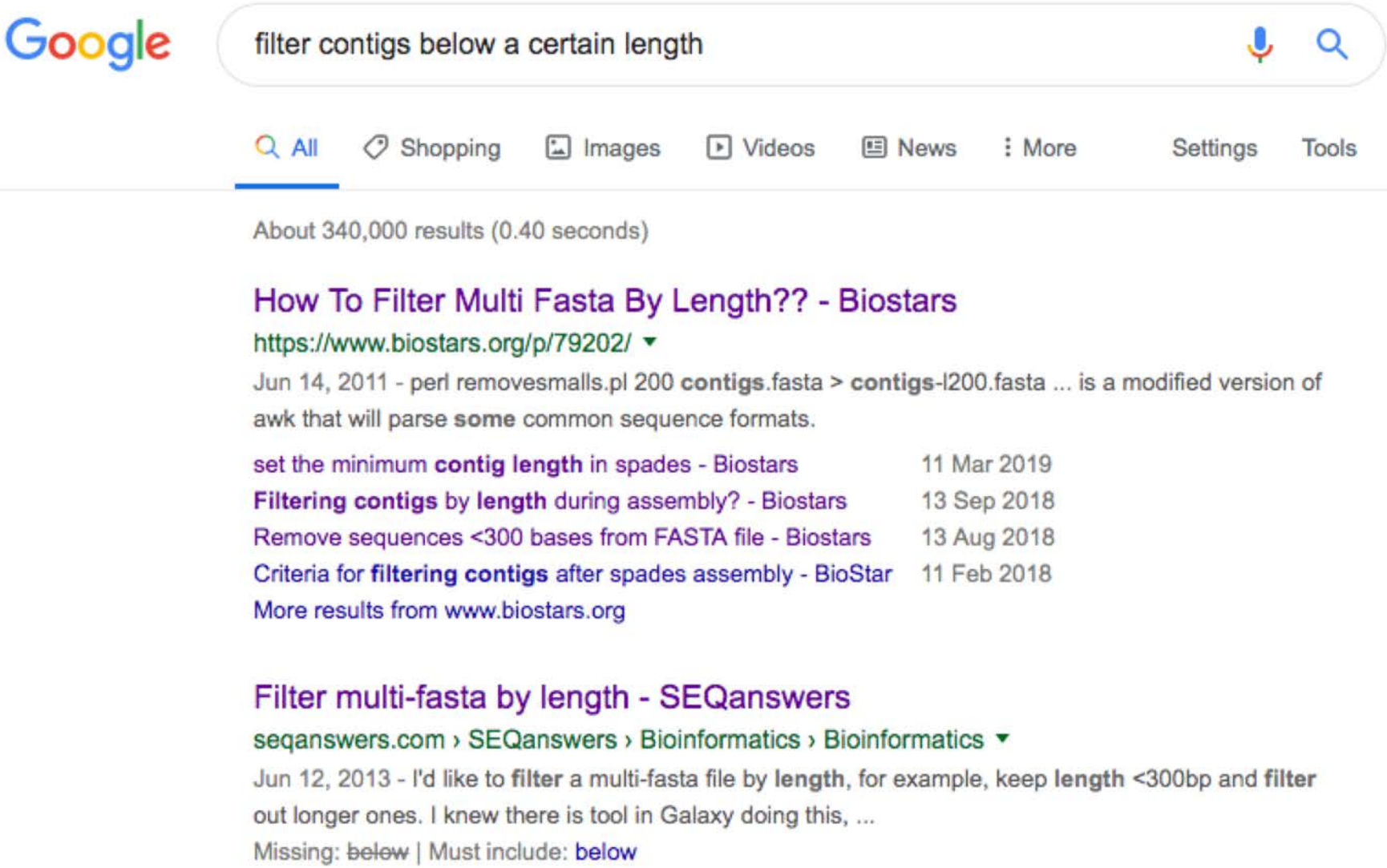
1) Provides a “Rosetta stone” to match up microsats and GBS targets


2) Provides a target for annotation




https://www.doc.govt.nz/news/media-releases/2017/birds-breed-at-bluff-for-first-time-in-more-than-100-years/

# How are we going to run this job?



# How are we going to run this job?

LATEST OPEN RNA-SEQ CHIP-SEQ SNP ASSEMBLY TUTORIALS TOOLS JOBS FORUM PLANET ALL »

Biostars — BIOINFORMATICS EXPLAINED

Welcome to Biostar! about • faq • rss

Community Log In Sign Up Add New Post

**Question: set the minimum contig length in spades**

0

Dear Sir/Madam:

I have the following script for spades, quast, and prokka. It is working perfect. However, there are several contigs with less than 200 in length. I need the minimum contig length 200. Spades manual does not say how to do it. Can you please help me with this?

3 months ago by babdalhamid • 0

```
#!/bin/bash
#SBATCH --time=7:00:00
#SBATCH --mem=200GB
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=4
#SBATCH --workdir=/lustre/work/bahaabd/baha000/genomes_untrimmed_fastq/2018/group15
#SBATCH --job-name=group15

module load spades quast prokka python/2.7

combined_data=combined.quast.results.txt


while read r1 r2 output; do

spades.py -1 $r1 -2 $r2 -o $output
spades.py -1 $r1 -2 $r2 -o $output -t 4 --only-assembler

cd $output
quast.py contigs.fasta
prokka --genus Escherichia -usegenus contigs.fasta

cd ..
echo -e "$output\t$(tail -n 1 $output/quast_results/latest/transposed_report.tsv)" >> $com

done < group15.txt
```

Thanks a lot. Baha

assembly • 273 views

ADD COMMENT • link • Not following • modified 3 months ago by genomax ♦ 68k • written 3 months ago by babdalhamid • 0

You can easily post-filter the contigs file to eliminate contigs smaller than 200bp.

ADD REPLY • link written 3 months ago by genomax ♦ 68k

I am a beginner in this field. I really do not know how to postfilter contigs. Would you please tell me how? Highly appreciate it. Baha

ADD REPLY • link written 3 months ago by babdalhamid • 0

You can use a separate program. One recommendation `reformat.sh` from BBMap suite. You would do something like

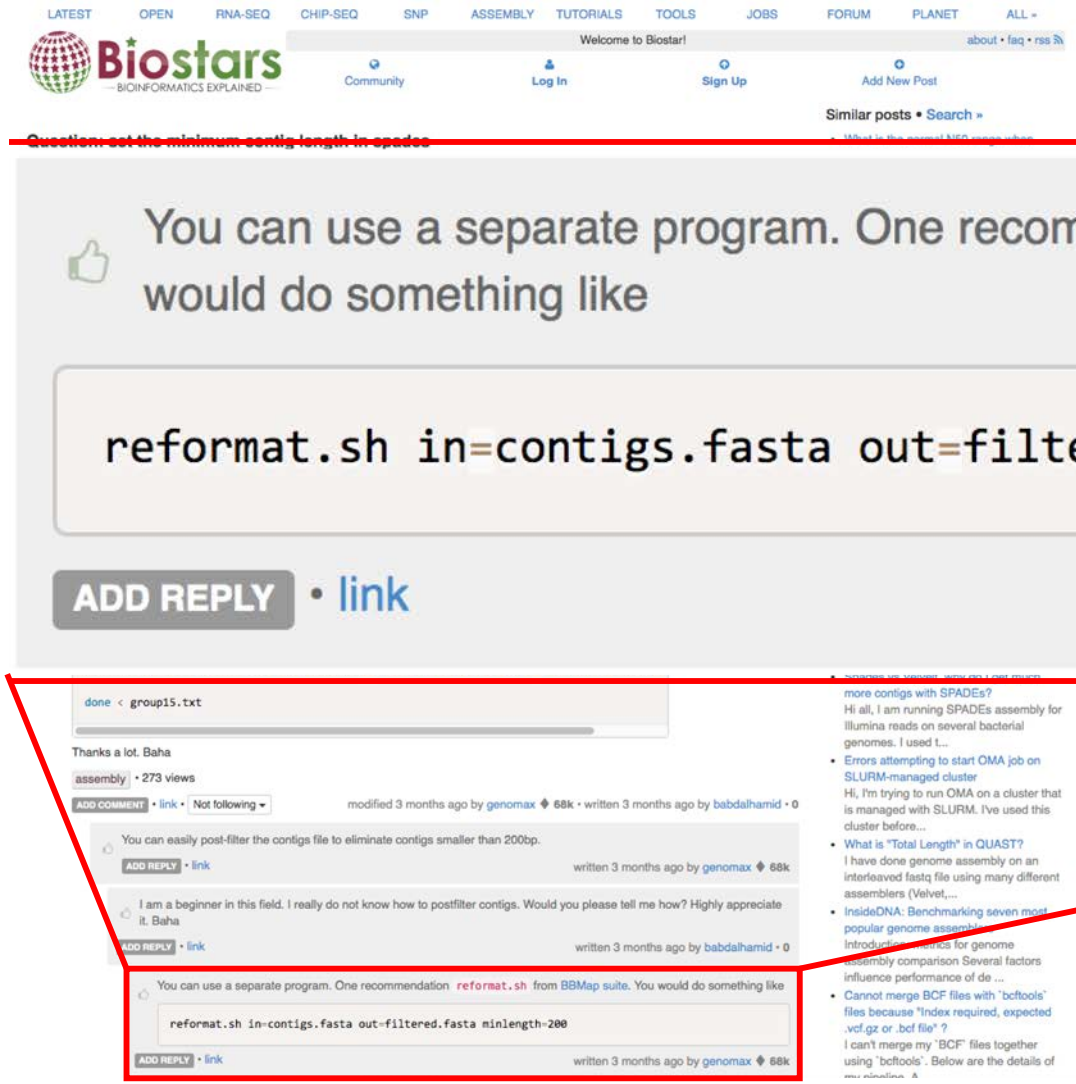
```
reformat.sh in=contigs.fasta out=filtered.fasta minlength=200
```

ADD REPLY • link written 3 months ago by genomax ♦ 68k

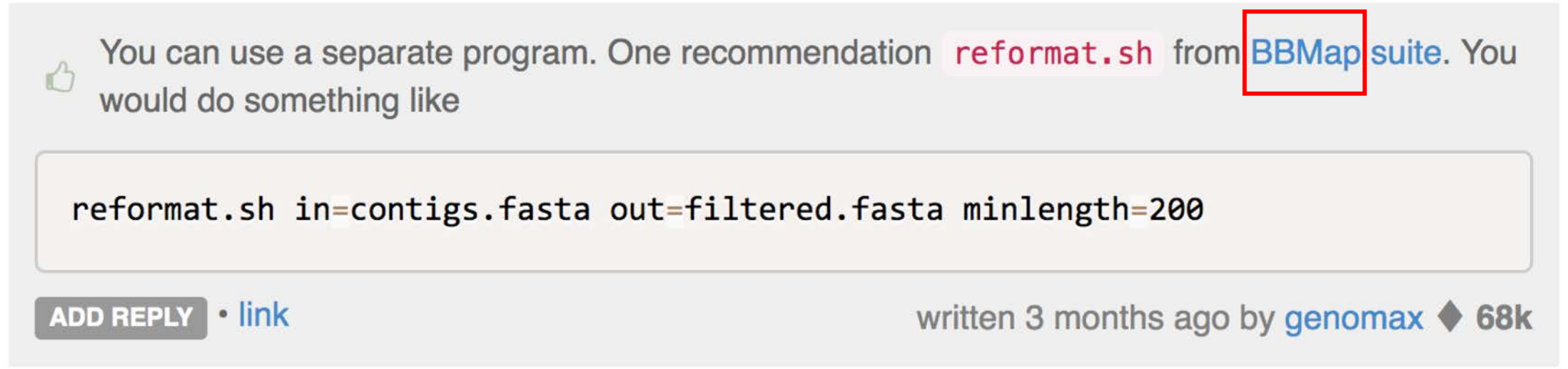
Similar posts • Search »

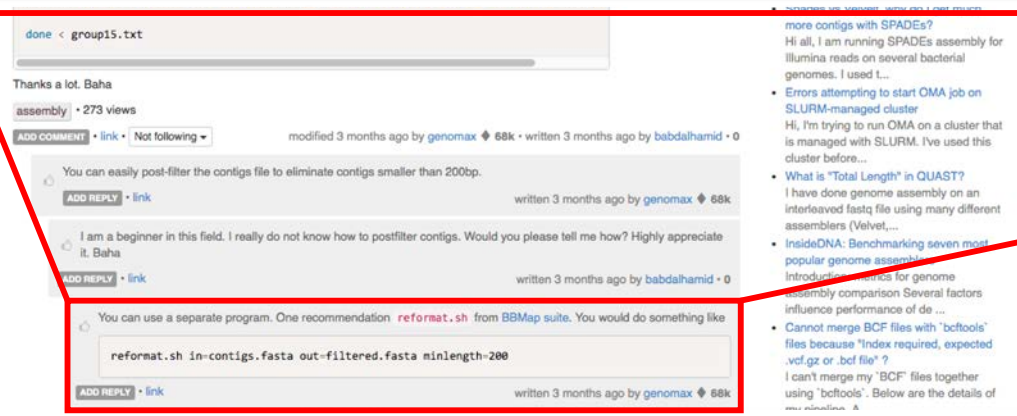
- What is the normal N50 range when assembly the virus genome.
  Dear All, I just finished my very first virus assembly by SPAdes. The following is the result fro...
- Assembly finished but output contig file is empty
  Hi everyone, I'm a beginner using spades 3.11.0 since March2019 , trying to practice doing some ...
- How to call a R variable in a loop with Slurm?
  Hello, I have a R (RHO_COR.R) script and I would like to create a loop in order to split the job...
- De-novo assembly using SPADES
  Before, I used SOAPdenovo2 to assemble my target reads. In order to use multiple k-mer and have b...
- BWA batch script does not execute
  I'm running an alignment with BWA on multiple exomes files on our server. For some reason it does...
- How to run Spades For Nextseq data
  Hi guys, I was having a try to run Nextseq data set with Spades. I have 4 fr reads and 4 re rea...
- How to extract bitmap images from .slide file generated by a CytoVision Platform
  I need help to extract images from a file ( .slide ) generated by a cytovision platform. This fil...
- Spades vs Velvelt_why do I get much more contigs with SPADEs?
  Hi all, I am running SPADEs assembly for illumina reads on several bacterial genomes. I used t...
- Errors attempting to start OMA job on SLURM-managed cluster
  Hi, I'm trying to run OMA on a cluster that is managed with SLURM. I've used this cluster before...
- What is "Total Length" in QUAST?
  I have done genome assembly on an interleaved fastq file using many different assemblers (Velvet,...
- InsideDNA: Benchmarking seven most popular genome assemblers
  Introduction: metrics for genome assembly comparison Several factors influence performance of de ...
- Cannot merge BCF files with `bcftools` files because "Index required, expected .vcf.gz or .bcf file" ?
  I can't merge my "BCF" files together using `bcftools`. Below are the details of

# How are we going to run this job?



# How are we going to run this job?


I know BBMap is a module on NeSI, which means we don't have to install this program ourselves!

# How are we going to run this job?

Jump over to NeSI and demonstrate:

```
[alana.alexander@mahuika01 filtered_genome]$ module spider bbmap
```

# How are we going to run this job?

Jump over to NeSI and demonstrate:

```
[alana.alexander@mahuika01 filtered_genome]$ module spider bbmap
```

From this we know that the module is called ‘BBMap’, so that is the name we’ll need to use in the `module load` command

# How are we going to run this job?

Jump over to NeSI and demonstrate:

```
[alana.alexander@mahuika01 filtered_genome]$ module spider bbmap
```

From this we know that the module is called ‘BBMap’, so that is the name we’ll need to use in the `module load` command

---

You can use a separate program. One recommendation `reformat.sh` from BBMap suite. You would do something like

```
reformat.sh in=contigs.fasta out=filtered.fasta minlength=200
```

ADD REPLY • link written 3 months ago by genomax ♦ 68k

---

# How are we going to run this job?

Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahui...

```
#!/bin/bash -e

#SBATCH --account=uoo00105
#SBATCH --job-name=bbmap
#SBATCH --ntasks=1
#SBATCH --nodes=1
#SBATCH --cpus-per-task=2
#SBATCH --time=2:00:00
#SBATCH --mem-per-cpu=3G
#SBATCH --partition=large
#SBATCH -D /nesi/nobackup/uoo00105/filtered_genome
#SBATCH --mail-type=ALL
#SBATCH --mail-user=laninsky@gmail.com
#SBATCH --profile=task

module load BBMap

reformat.sh in=../robin_genome_quast/04.break.broken_assembly_quast/04.break.broken_assembly.fa out
=filtered_below_500bp_17Jun2019.fa minlength=500
~
~
(END)
```

# How are we going to run this job?

Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahui...

Thing we have to put at the top to tell Mahuika what language our slurm script is written in

```
#!/bin/bash -e

#SBATCH --account=uoo00105
#SBATCH --job-name=bbmap
#SBATCH --ntasks=1
#SBATCH --nodes=1
#SBATCH --cpus-per-task=2
#SBATCH --time=2:00:00
#SBATCH --mem-per-cpu=3G
#SBATCH --partition=large
#SBATCH -D /nesi/nobackup/uoo00105/filtered_genome
#SBATCH --mail-type=ALL
#SBATCH --mail-user=laninsky@gmail.com
#SBATCH --profile=task

module load BBMap

reformat.sh in=../robin_genome_quast/04.break.broken_assembly_quast/04.break.broken_assembly.fa out
=filtered_below_500bp_17Jun2019.fa minlength=500
~
~
(END)
```

# How are we going to run this job?

Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahui...

```
#!/bin/bash -e

#SBATCH --account=uoo00105
#SBATCH --job-name=bbmap
#SBATCH --ntasks=1
#SBATCH --nodes=1
#SBATCH --cpus-per-task=2
#SBATCH --time=2:00:00
#SBATCH --mem-per-cpu=3G
#SBATCH --partition=large
#SBATCH -D /nesi/nobackup/uoo00105/filtered_genome
#SBATCH --mail-type=ALL
#SBATCH --mail-user=laninsky@gmail.com
#SBATCH --profile=task

module load BBMap

reformat.sh in=../robin_genome_quast/04.break.broken_assembly_quast/04.break.broken_assembly.fa out
=filtered_below_500bp_17Jun2019.fa minlength=500
~
~
(END)
```

Thing we have to put at the top to tell Mahuika what language our slurm script is written in

What the module we need is called

# How are we going to run this job?

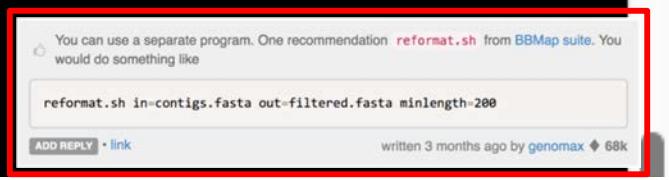
Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahui...

```
#!/bin/bash -e

#SBATCH --account=uoo00105
#SBATCH --job-name=bbmap
#SBATCH --ntasks=1
#SBATCH --nodes=1
#SBATCH --cpus-per-task=2
#SBATCH --time=2:00:00
#SBATCH --mem-per-cpu=3G
#SBATCH --partition=large
#SBATCH -D /nesi/nobackup/uoo00105/filtered_genome
#SBATCH --mail-type=ALL
#SBATCH --mail-user=laninsky@gmail.com
#SBATCH --profile=task

module load BBMap

reformat.sh in=../robin_genome_quast/04.break.broken_assembly_quast/04.break.broken_assembly.fa out
=filtered_below_500bp_17Jun2019.fa minlength=500
~
~
(END)
```

Thing we have to put at the top to tell Mahuika what language our slurm script is written in

What the module we need is called

The BBMap command, adapted from what we found on the internet



# How are we going to run this job?

Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahui...

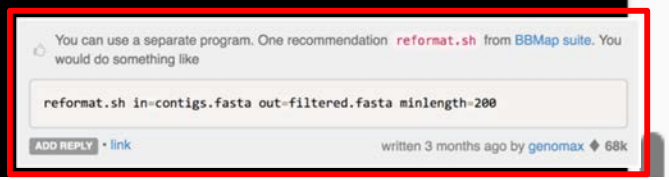
```
#!/bin/bash -e

#SBATCH --account=uoo00105
#SBATCH --job-name=bbmap
#SBATCH --ntasks=1
#SBATCH --nodes=1
#SBATCH --cpus-per-task=2
#SBATCH --time=2:00:00
#SBATCH --mem-per-cpu=3G
#SBATCH --partition=large
#SBATCH -D /nesi/nobackup/uoo00105/filtered_genome
#SBATCH --mail-type=ALL
#SBATCH --mail-user=laninsky@gmail.com
#SBATCH --profile=task

module load BBMap

reformat.sh in=../robin_genome_quast/04.break.broken_assembly_quast/04.break.broken_assembly.fa out
=filtered_below_500bp_17Jun2019.fa minlength=500
~
~
(END)
```

Thing we have to put at the top to tell Mahuika what language our slurm script is written in

Slurm stuff

What the module we need is called

The BBMap command, adapted from what we found on the internet



# How are we going to run this job?

Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahui...

```
#!/bin/bash -e

#SBATCH --account=uoo00105
#SBATCH --job-name=bbmap
#SBATCH --ntasks=1
#SBATCH --nodes=1
#SBATCH --cpus-per-task=2
#SBATCH --time=2:00:00
#SBATCH --mem-per-cpu=3G
#SBATCH --partition=large
#SBATCH -D /nesi/nobackup/uoo00105/filtered_genome
#SBATCH --mail-type=ALL
#SBATCH --mail-user=laninsky@gmail.com
#SBATCH --profile=task

module load BBMap

reformat.sh in=../robin_genome_quast/04.break.broken_assembly_quast/04.break.broken_assembly.fa out
=filtered_below_500bp_17Jun2019.fa minlength=500
~
~
(END)
```

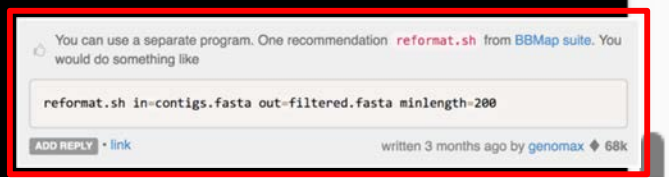
Thing we have to put at the top to tell Mahuika what language our slurm script is written in

Slurm stuff

We need this bit to collect information about our job to profile it

What the module we need is called

The BBMap command, adapted from what we found on the internet



# How are we going to run this job?

After getting this script together, we submit it by:

```
[alana.alexander@mahuika01 filtered_genome]$ sbatch bbmap_filter.sh
```

# How many resources were used?

After getting this script together, we submit it by:

```
[alana.alexander@mahuika01 filtered_genome]$ sbatch bbmap_filter.sh
```

We can get a basic idea of how many resources it used by:

```
[alana.alexander@mahuika01 filtered_genome]$ sacct
         JobID        JobName     Elapsed     TotalCPU Alloc   MaxRSS      State
-------------- -------------- ----------- ------------ ----- -------- ----------
4603711        bbmap              00:00:09    00:10.665     2          COMPLETED
4603711.batch  batch              00:00:09    00:10.664     2         0 COMPLETED
4603711.extern extern             00:00:09    00:00.001     2         0 COMPLETED
4603723        bbmap              00:00:01    00:00.659     2          FAILED
4603723.batch  batch              00:00:01    00:00.658     2         0 FAILED
4603723.extern extern             00:00:01    00:00.001     2         0 COMPLETED
4603766        bbmap              00:00:09    00:10.322     2          COMPLETED
4603766.batch  batch              00:00:09    00:10.321     2         0 COMPLETED
4603766.extern extern             00:00:09     00:00:00     2         0 COMPLETED
[alana.alexander@mahuika01 filtered_genome]$
```

# How many resources were used?

6

@aaron.kizmiller is right, `sacct` is the command to use.

One can fetch all of the following fields by passing them into `saact --format="field,field"`

**Fields**:

```
Account         AdminComment     AllocCPUS        AllocGRES
AllocNodes      AllocTRES        AssocID          AveCPU
AveCPUFreq      AveDiskRead      AveDiskWrite     AvePages
AveRSS          AveVMSize        BlockID          Cluster
Comment         ConsumedEnergy   ConsumedEnergyRaw CPUTime
CPUTimeRAW      DerivedExitCode  Elapsed          ElapsedRaw
Eligible        End              ExitCode         GID
Group           JobID            JobIDRaw         JobName
Layout          MaxDiskRead      MaxDiskReadNode  MaxDiskReadTask
MaxDiskWrite    MaxDiskWriteNode MaxDiskWriteTask MaxPages
MaxPagesNode    MaxPagesTask     MaxRSS           MaxRSSNode
MaxRSSTask      MaxVMSize        MaxVMSizeNode    MaxVMSizeTask
McsLabel        MinCPU           MinCPUNode       MinCPUTask
NCPUS           NNodes           NodeList         NTasks
Priority        Partition        QOS              QOSRAW
ReqCPUFreq      ReqCPUFreqMin    ReqCPUFreqMax    ReqCPUFreqGov
ReqCPUS         ReqGRES          ReqMem           ReqNodes
ReqTRES         Reservation      ReservationId    Reserved
ResvCPU         ResvCPURAW       Start            State
Submit          Suspended        SystemCPU        Timelimit
TotalCPU        UID              User             UserCPU
WCKey           WCKeyID          WorkDir
```

For example, to list all job ids, elapsed time, and max VM size, you can run:

```
sacct --format='JobID,Elapsed,MaxVMSize'
```

https://stackoverflow.com/questions/24020420/find-out-the-cpu-time-and-memory-usage-of-a-slurm-job

# How many resources were used at each step?

You can analyse the HDF5 file yourself using the python script mentioned below

1. `sh5util -j <jobid>`

2. Ignore the error message about an empty file.

3. Use  profile_plot.py (https://github.com/otagomohio/hackyhour/blob/master/sessions/code/profile_plot.py) to plot the contents of such a file.

# How many resources were used at each step?

```
4603766        bbmap              00:00:09    00:10.322      2          COMPLETED
4603766.batch  batch              00:00:09    00:10.321      2        0 COMPLETED
4603766.extern extern             00:00:09     00:00:00      2        0 COMPLETED
[alana.alexander@mahuika01 filtered_genome]$ sh5util -j 4603766
sh5util: Merging node-step files into ./job_4603766.h5
[alana.alexander@mahuika01 filtered_genome]$ ls -shlt
total 1.1G
256K -rw-rw----+ 1 alana.alexander uoo00105  12K Jun 16 23:18 job_4603766.h5
 512 -rw-rw----+ 1 alana.alexander uoo00105  910 Jun 16 23:07 slurm-4603766.out
1.1G -rw-rw----+ 1 alana.alexander uoo00105 1.1G Jun 16 23:07 filtered_below_500bp_17Ju
n2019.fa
 512 -rw-rwx---+ 1 alana.alexander uoo00105  512 Jun 16 23:06 bbmap_filter.sh
```

# How many resources were used at each step?

Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahuik...

```
#!/usr/bin/env python
"""Plot data from a SLURM HDF5 profile file generated with sh5util"""

from __future__ import division, print_function

try:
    import h5py
except ImportError:
    print("Do 'module load Python' to get a version with HDF5 support")
    raise

import matplotlib
matplotlib.use('agg')
from matplotlib import pyplot as plt, colors, gridspec
import sys, time, datetime, numpy

                                                              1,1           Top
```

# How many resources were used at each step?

```
[alana.alexander@mahuika01 filtered_genome]$ vi profile_plot.py
[alana.alexander@mahuika01 filtered_genome]$ ls -shlt
total 1.1G
256K -rw-rw----+ 1 alana.alexander uoo00105 6.4K Jun 16 23:22 profile_plot.py
256K -rw-rw----+ 1 alana.alexander uoo00105  12K Jun 16 23:18 job_4603766.h5
 512 -rw-rw----+ 1 alana.alexander uoo00105  910 Jun 16 23:07 slurm-4603766.out
1.1G -rw-rw----+ 1 alana.alexander uoo00105 1.1G Jun 16 23:07 filtered_below_500bp_17Jun2019.fa
 512 -rw-rwx---+ 1 alana.alexander uoo00105  512 Jun 16 23:06 bbmap_filter.sh
[alana.alexander@mahuika01 filtered_genome]$ python profile_plot.py
Do 'module load Python' to get a version with HDF5 support
Traceback (most recent call last):
  File "profile_plot.py", line 7, in <module>
    import h5py
  File "/usr/lib64/python2.7/site-packages/h5py/__init__.py", line 10, in <module>
    from h5py import _errors
ImportError: libhdf5.so.8: cannot open shared object file: No such file or directory
[alana.alexander@mahuika01 filtered_genome]$
```

# How many resources were used at each step?

Volumes — alana.alexander@mahuika01:/nesi/nobackup/uoo00105/filtered_genome — ssh ◂ ssh mahuik...

```
#!/usr/bin/env python
"""Plot data from a SLURM HDF5 profile file generated with sh5util"""

from __future__ import division, print_function

try:
    import h5py
except ImportError:
    print("Do 'module load Python' to get a version with HDF5 support")
    raise

import matplotlib
matplotlib.use('agg')
from matplotlib import pyplot as plt, colors, gridspec
import sys, time, datetime, numpy

                                                              1,1           Top
```

Ahem

# How many resources were used at each step?

```
[alana.alexander@mahuika01 filtered_genome]$ module load  Python
[alana.alexander@mahuika01 filtered_genome]$ python profile_plot.py
Traceback (most recent call last):
  File "profile_plot.py", line 36, in <module>
    fn = sys.argv[1]
IndexError: list index out of range
[alana.alexander@mahuika01 filtered_genome]$ ls -shlt
total 1.1G
256K -rw-rw----+ 1 alana.alexander uoo00105 6.4K Jun 16 23:22 profile_plot.py
256K -rw-rw----+ 1 alana.alexander uoo00105  12K Jun 16 23:18 job_4603766.h5
 512 -rw-rw----+ 1 alana.alexander uoo00105  910 Jun 16 23:07 slurm-4603766.out
1.1G -rw-rw----+ 1 alana.alexander uoo00105 1.1G Jun 16 23:07 filtered_below_500bp_17Jun2019.fa
 512 -rw-rwx---+ 1 alana.alexander uoo00105  512 Jun 16 23:06 bbmap_filter.sh
```

# How many resources were used at each step?

```
[alana.alexander@mahuika01 filtered_genome]$ python profile_plot.py --help
Traceback (most recent call last):
  File "profile_plot.py", line 38, in <module>
    steps = f['Steps']
  File "h5py/_objects.pyx", line 54, in h5py._objects.with_phil.wrapper
  File "h5py/_objects.pyx", line 55, in h5py._objects.with_phil.wrapper
  File "/opt/nesi/CS400_centos7_bdw/Python/3.7.3-gimkl-2018b/lib/python3.7/site-packages/h5py-2.9.0
-py3.7-linux-x86_64.egg/h5py/_hl/group.py", line 262, in __getitem__
    oid = h5o.open(self.id, self._e(name), lapl=self._lapl)
  File "h5py/_objects.pyx", line 54, in h5py._objects.with_phil.wrapper
  File "h5py/_objects.pyx", line 55, in h5py._objects.with_phil.wrapper
  File "h5py/h5o.pyx", line 190, in h5py.h5o.open
KeyError: "Unable to open object (object 'Steps' doesn't exist)"
```

# How many resources were used at each step?

```
def pick_io_scale(max_io):
    if max_io > 200:
        io_scale = 1000
        io_unit = 'GB'
    else:
        io_scale = 1
        io_unit = 'MB'
    return (io_scale, io_unit)

STEPS = sys.argv[2:]
fn = sys.argv[1]
f = h5py.File(fn)
steps = f['Steps']

# list(f['Steps']['0']['Nodes']['compute-b1-065']['Tasks']['0'])[0]
# fields = f.values()[0].values()[0].values()[0].values()[0].values()[0].values()[0].dtype.fields
# dtype([('ElapsedTime', '<u8'), ('EpochTime', '<u8'),
                                                          41,1          12%
```

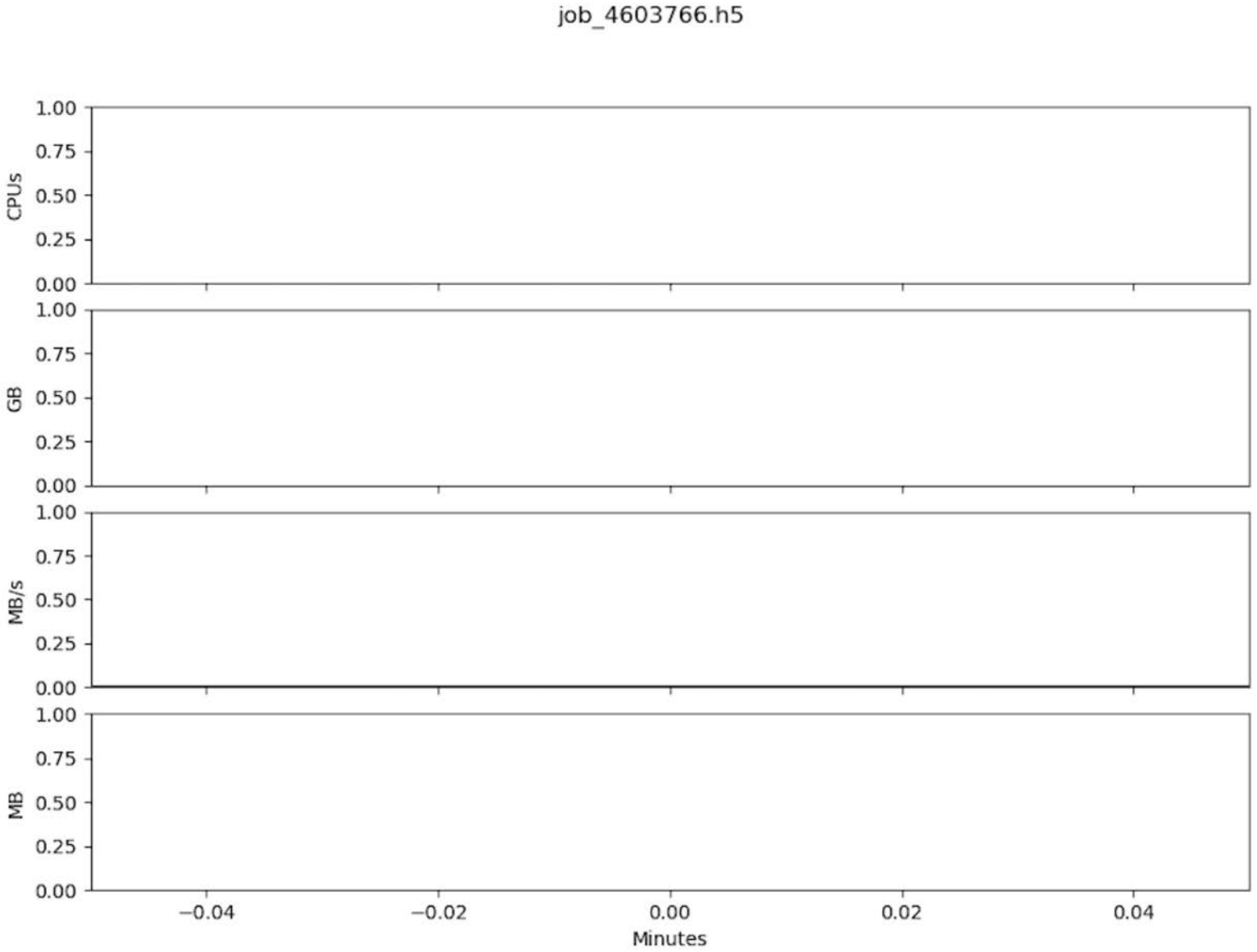
Ahem...needs a file name...

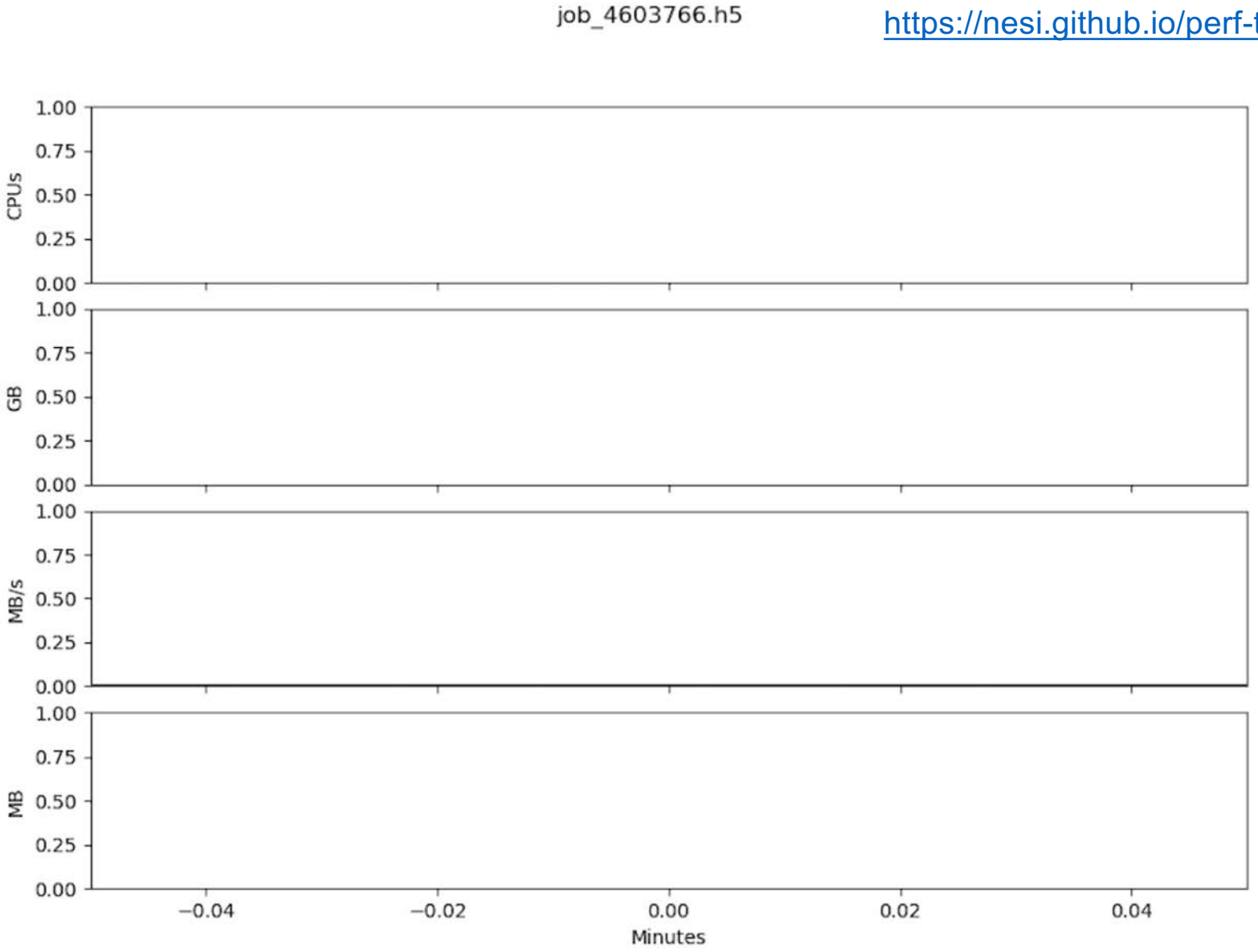
# How many resources were used at each step?

```
[alana.alexander@mahuika01 filtered_genome]$ python profile_plot.py job_4603766.h5
[alana.alexander@mahuika01 filtered_genome]$ ls -shlt
total 1.1G
256K -rw-rw----+ 1 alana.alexander uoo00105  12K Jun 16 23:32 job_4603766.h5
256K -rw-rw----+ 1 alana.alexander uoo00105  29K Jun 16 23:32 job_4603766_profile.png
 512 -rw-rw----+ 1 alana.alexander uoo00105  800 Jun 16 23:30 --h
 512 -rw-rw----+ 1 alana.alexander uoo00105  800 Jun 16 23:30 -h
 512 -rw-rw----+ 1 alana.alexander uoo00105  800 Jun 16 23:30 -help
 512 -rw-rw----+ 1 alana.alexander uoo00105  800 Jun 16 23:29 --help
256K -rw-rw----+ 1 alana.alexander uoo00105 6.4K Jun 16 23:22 profile_plot.py
 512 -rw-rw----+ 1 alana.alexander uoo00105  910 Jun 16 23:07 slurm-4603766.out
1.1G -rw-rw----+ 1 alana.alexander uoo00105 1.1G Jun 16 23:07 filtered_below_500bp_17Jun2019.fa
 512 -rw-rwx---+ 1 alana.alexander uoo00105  512 Jun 16 23:06 bbmap_filter.sh
[alana.alexander@mahuika01 filtered_genome]$
```

Yay!

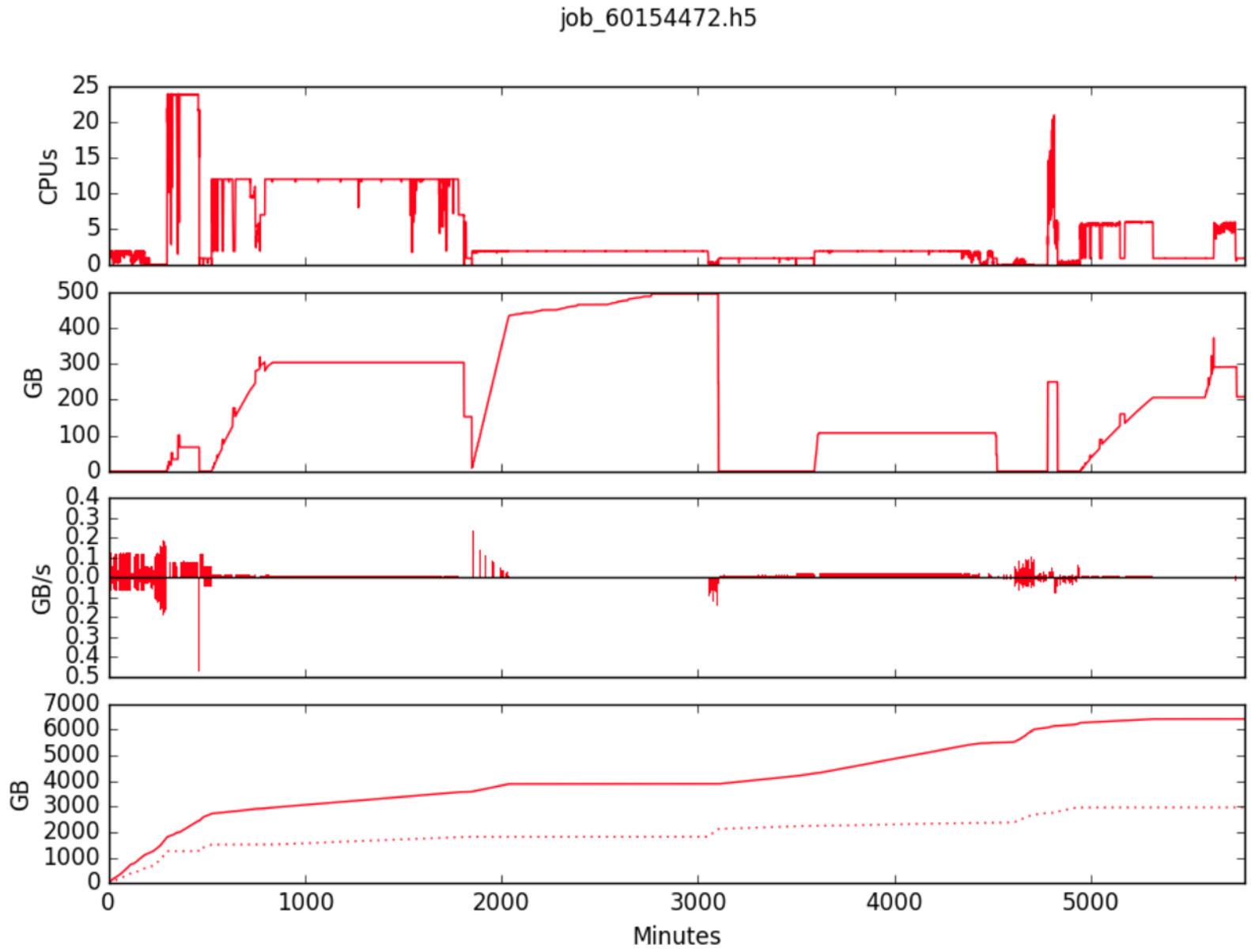
```
(base) alanaalexander@anat-dock-34:~$ scp mahuika:/nesi/nobackup/uoo00105/filtered_genome/*.png ./
job_4603766_profile.png                       100%   28KB 266.4KB/s   00:00
(base) alanaalexander@anat-dock-34:~$
```


job_4603766.h5
CPUs
GB
MB/s
MB
1.00
0.75
0.50
0.25
0.00
−0.04
−0.02
0.00
0.02
0.04
Minutes




https://nesi.github.io/perf-training/python-scatter/profiling

"The run time should not be too short, however, as this could make profiling results unreliable. Depending on the complexity of the code, the execution should take at least 10 seconds."


job_60154472.h5
CPUs
25
20
15
10
5
0
GB
500
400
300
200
100
0
GB/s
0.4
0.3
0.2
0.1
0.0
0.1
0.2
0.3
0.4
0.5
GB
7000
6000
5000
4000
3000
2000
1000
0
0
1000
2000
3000
4000
5000
Minutes