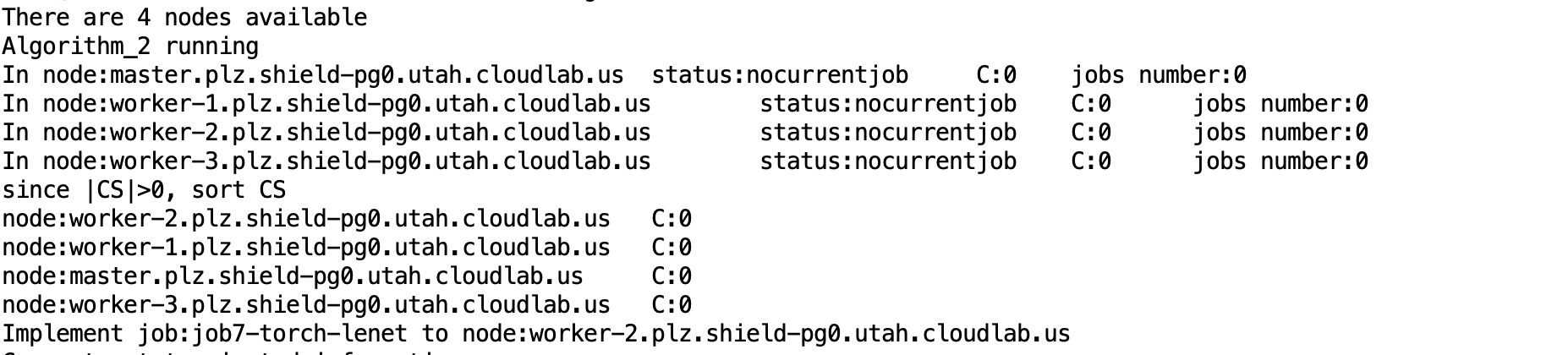
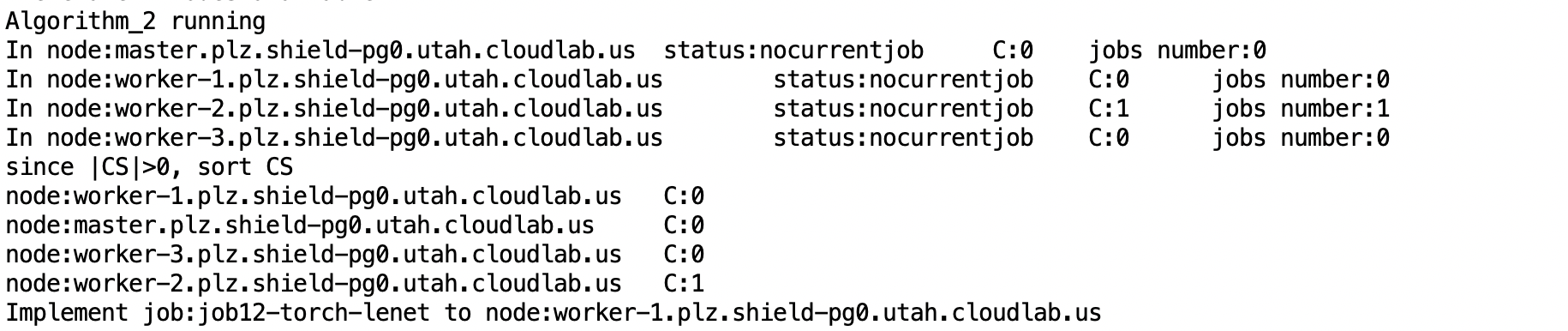
This is a summary of scheduler log

First four jobs submitted

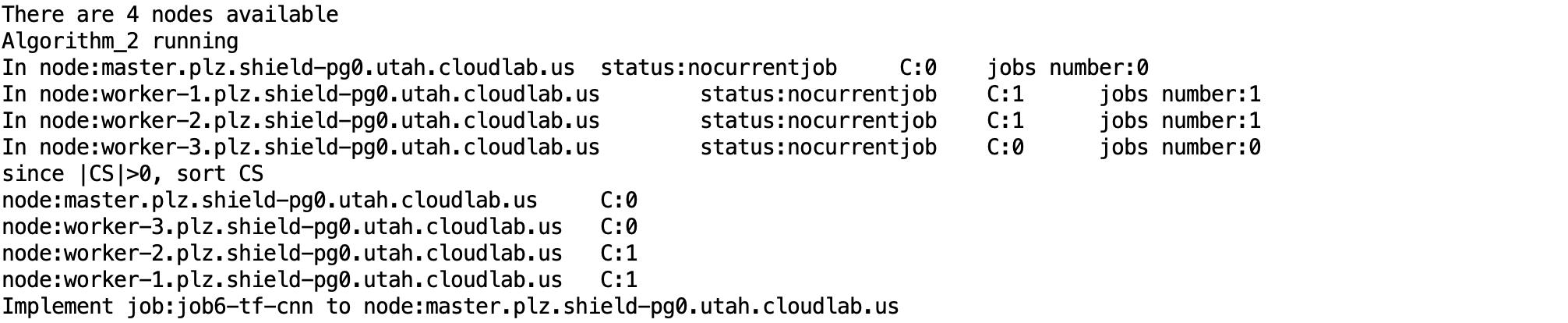
1.



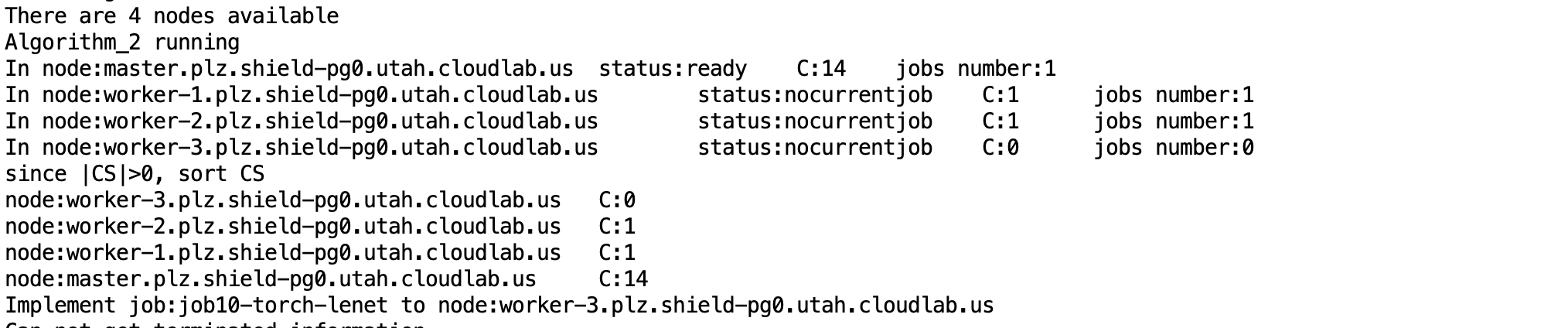
2.



3.

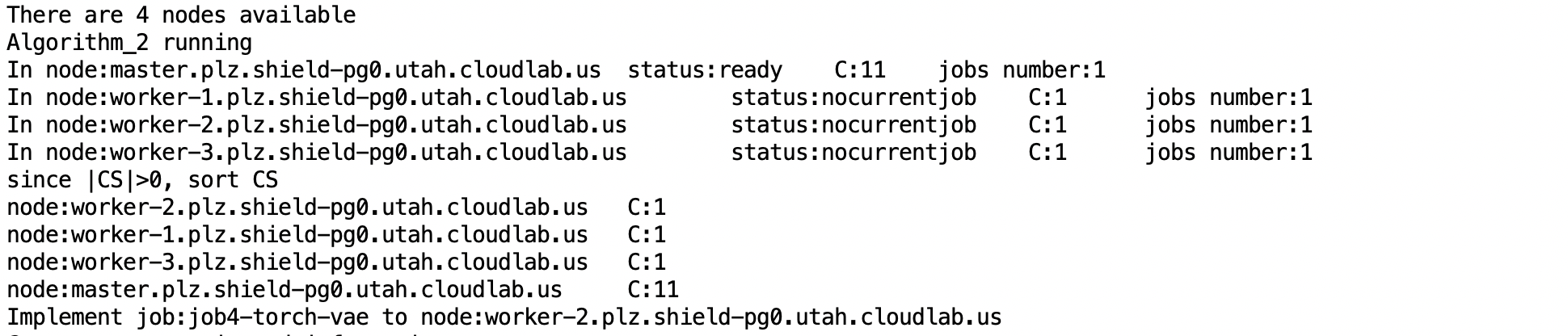


4.

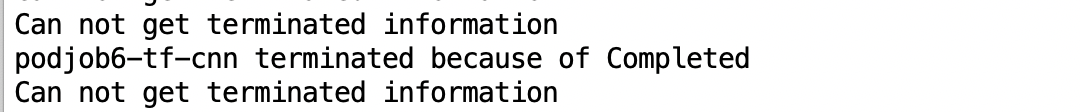


First four jobs look perfect, they are assigned to every node.

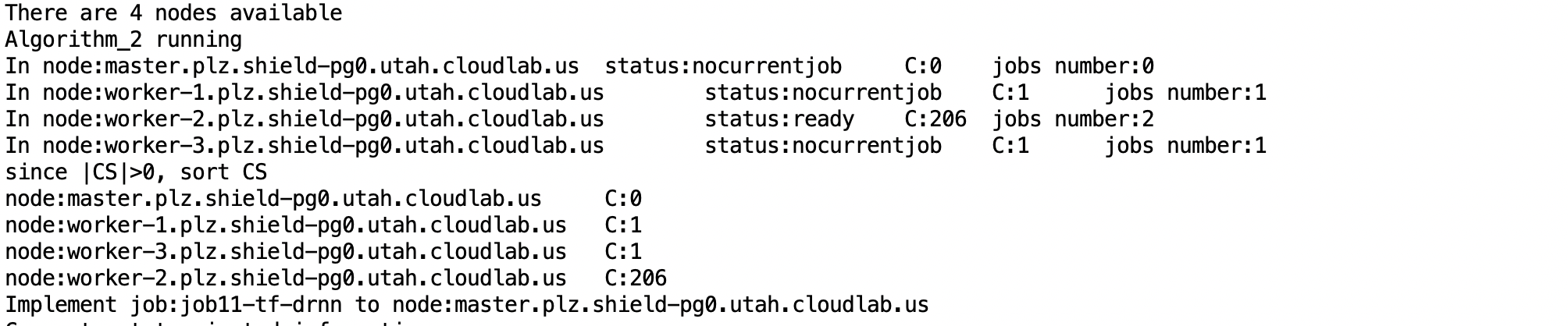
5.



Before 6th jobs submitted, there is a job completed. The scheduler look good at that time.

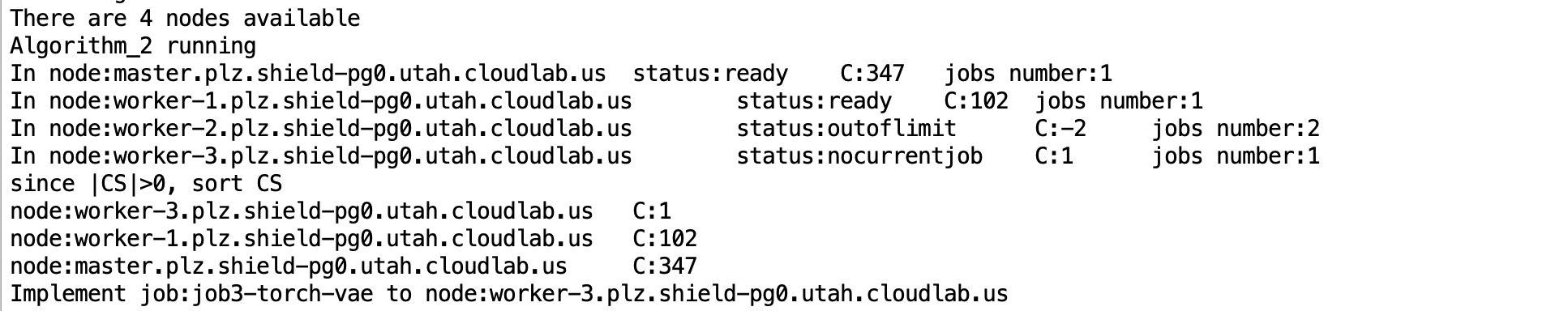


6.



When 7th job submitted, it caused some problem

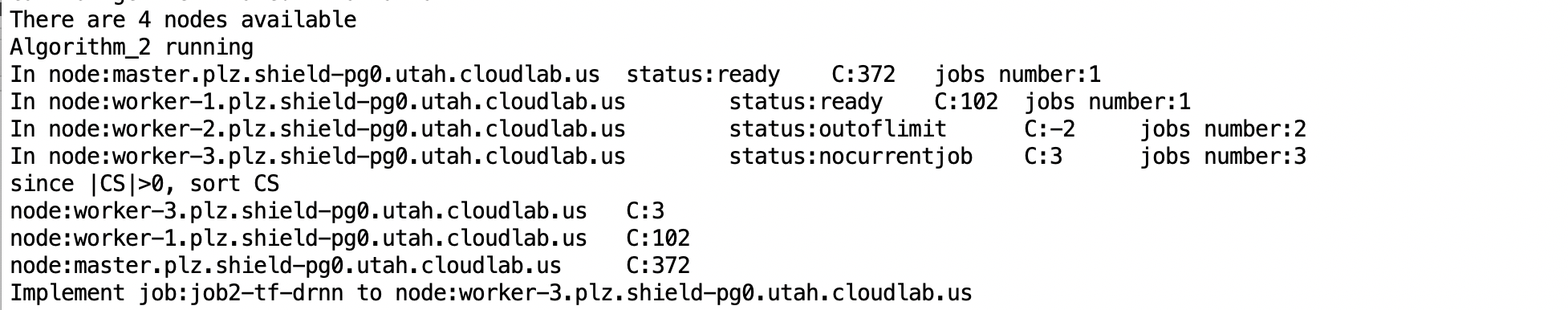
7.



8.



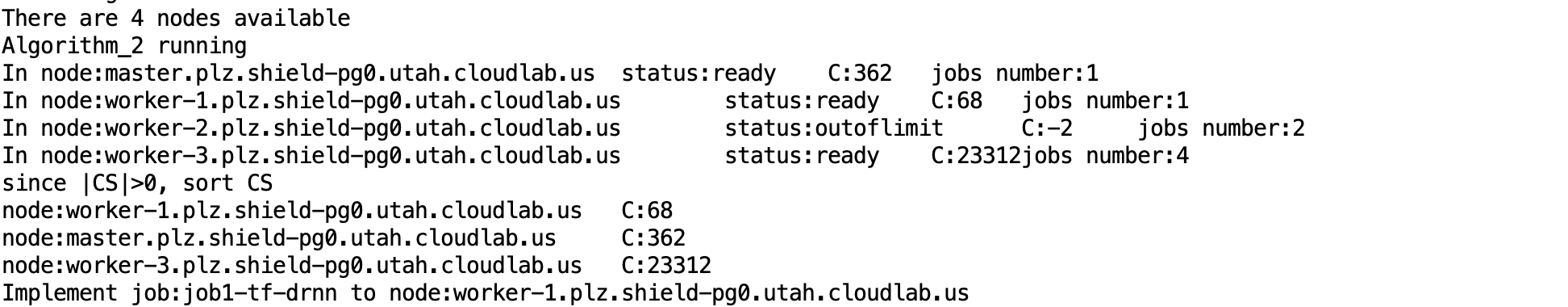
9.



As we can see, 7,8,9 are all assigned to worker 3 because worker 3 do not get its first jobs logs(lenet job which has a bit long iteration time)

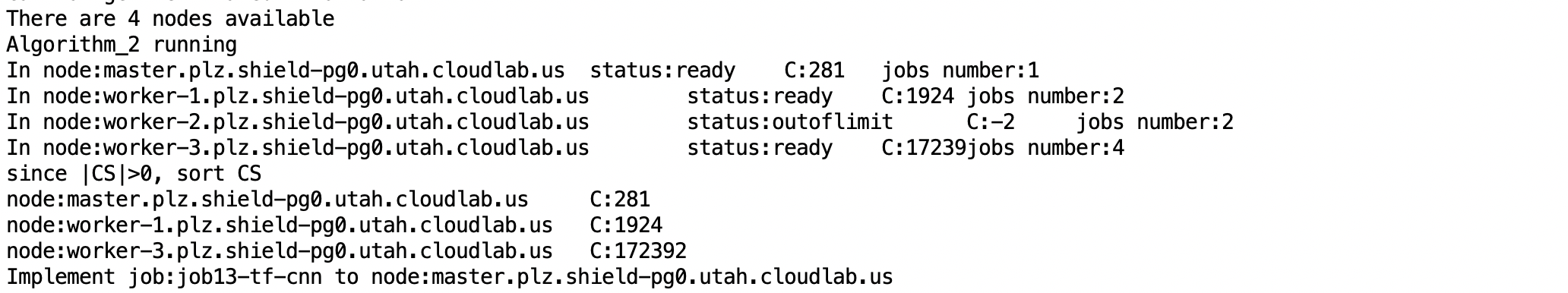
But when 10th job coming in, C in worker 3 is extremely large.

10.

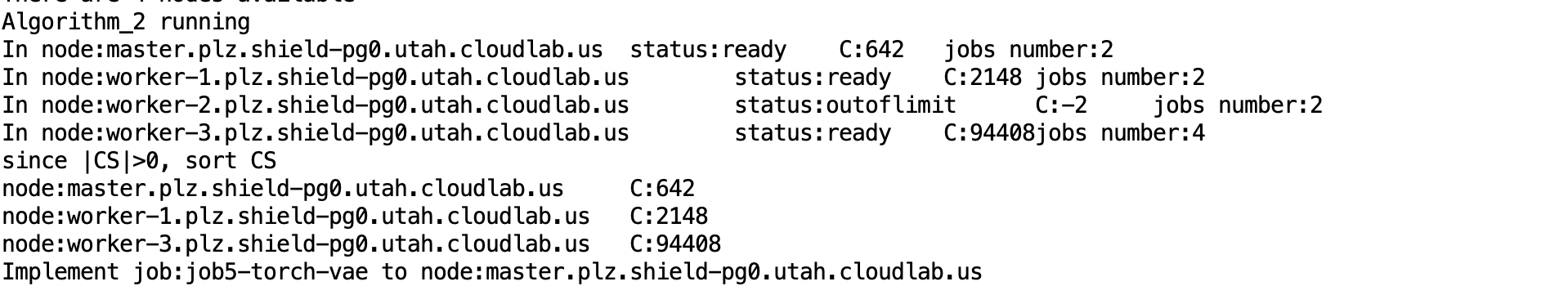


Then 11th and 12th jobs are both assigned to master because its C is lower than worker 1.

11.

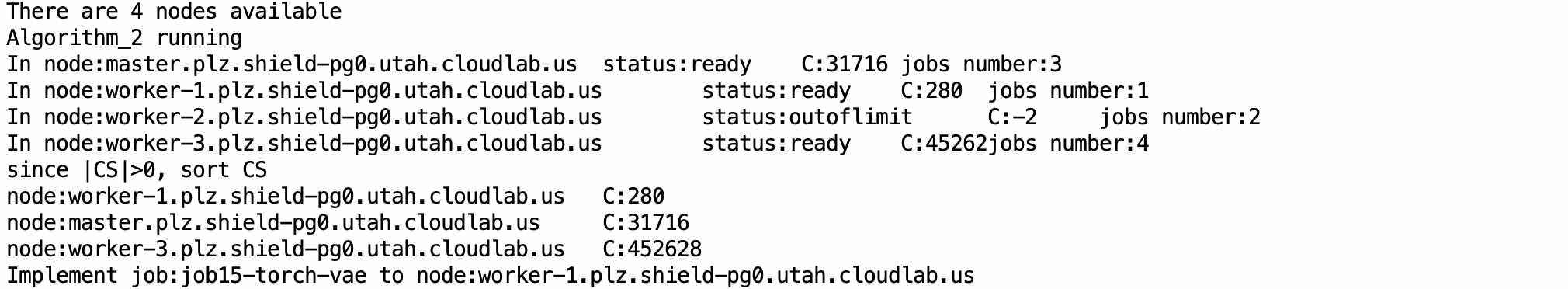


12

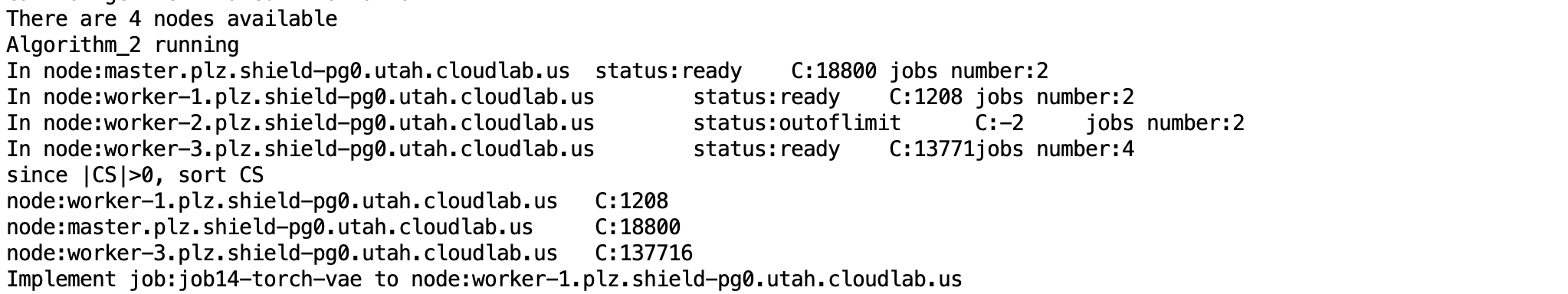


And then C in Master also be extremely large, and remaining jobs are assigned to worker 1.

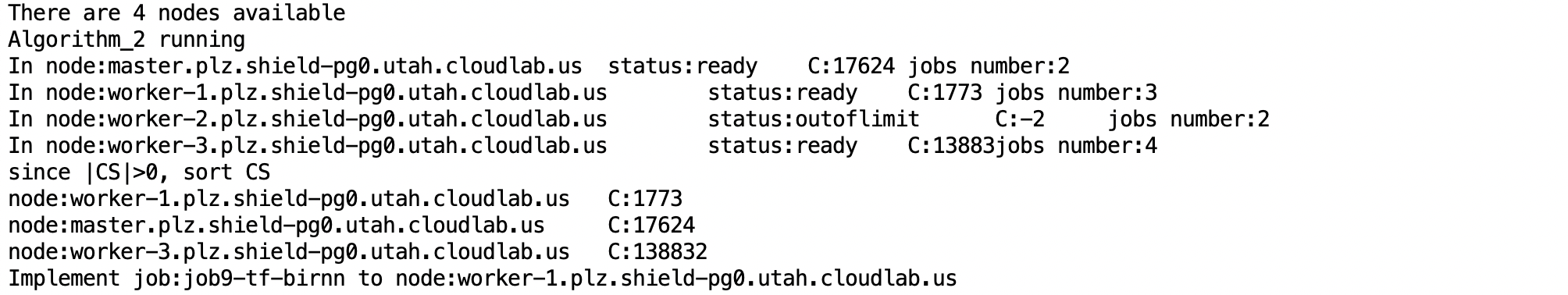
13.



14.



15



Summary:

In my opinion, the main problem is that worker can not get a update log information. Although we have set nums of running jobs to cover this problem, it is not enough when jobs’ submitted time is very close. It will cause a worker can be assigned many jobs through a short period and its C will be very large.

In scheduler, if we can get information, we just ignore this job. The following is code



Maybe we can set a some default value when we can not get information through this log.