

We take this step to see an example of your software development skills, you are free to use any programming language and tools you feel comfortable with. You have 24 hours, to reply with your answer but we expect this question should take no more than 2 to 3 hours. How soon you respond is not considered when we evaluate your submission (2 hours or 22 hours are considered equal).

We will evaluate your answer on:

Correctness, does it work?

Is it tested?

Is it well designed?

Is the code easy to comprehend?

Performs well?

Would this code be easy to extend or maintain?

Your implementation must be your own, making use of your standard library. Please feel free to consult any relevant documentation.

Please provide a complete solution including any instructions an engineer would need to build and run your software, e.g. a README, Makefile, setup.py, build.xml, etc. Attach the source files, .zip will be removed by our mail servers. See below for the coding challenge.

Best of luck,

=====  
Coding Challenge  
=====

A logfile contains newline-terminated, space-separated text formatted like:  
<unix\_timestamp> <hostname> <hostname>

Example:

1366815793 quark garak

1366815795 brunt quark

1366815811 lilac garak

Each line represents connection from a host (left) to another host (right) at a given time.

The lines are \*roughly\* sorted by timestamp. They might be out of order by maximum 5 minutes.

Implement a tool that parse logfiles like this. The tool should both parse previously written logfiles and terminate or collect input from a new logfile while it's being written and run indefinitely.

The script will output, once every hour:

- a list of hostnames connected to a given (configurable) host during the last hour
- a list of hostnames received connections from a given (configurable) host during the last hour
- the hostname that generated most connections in the last hour

The number of loglines and hostnames can be very high. Consider implementing a CPU and memory-efficient solution.

Please feel free to make assumptions as necessary with proper documentations.