Lumu Technologies

DevOps Technical Test

GENERAL QUESTIONS

- You will be in charge of running a set of JVM-based microservices connected to MongoDB, exchanging messages through a kafka broker and communicating with external clients using HTTP Restful Services.

  What metrics do you consider the most critical to monitor for the system and what tools would you use?

- Provide a Linux command to find all files which have been accessed between 20 and 30 days ago. Explain your command.
- Write a simple Python script that lists the 5 biggest files in a given directory (sorted by size), as well as the total size, number of files and average size of files in the directory.

  Sample output:

```
Total files: 100     Total size: 109372323   Average file size: 10234
Top files:
file.a     101098909
file.x     100093209
file.d     40098909
file.b     39090000
file.c     29200000
```

SYSTEM ANALYSIS

We are running a Python Flask application using **gunicorn** and **nginx** in a dockerized container. This is how our deployment files look like:

File: Dockerfile

```
FROM alpine
RUN apk add py3-pip build-base python3-dev libffi-dev openssl-dev
RUN apk add nginx
RUN mkdir -p /opt/api
WORKDIR /opt/api
```

```
ADD api/requirements.txt /opt/api
RUN pip3 install --no-cache-dir -r requirements.txt
ADD api/. /opt/api
ADD ./docker-entrypoint /bin/docker-entrypoint
ADD ./nginx.conf /etc/nginx/nginx.conf

CMD ["/bin/docker-entrypoint"]
```

File: docker-entrypoint

```
#!/bin/ash

echo "Starting gunicorn..."
gunicorn -w 16 -b 127.0.0.1:9000 app:app --daemon
sleep 3
echo "Starting web server..."
nginx -g "daemon off;"
```

File: nginx.conf

```
user nginx;
worker_processes auto;
pid /tmp/nginx.pid;
pcre_jit on;
error_log /var/log/nginx/error.log warn;
events {
   worker_connections  1024;
}

http {
   include /etc/nginx/mime.types;
   default_type text/json;
   server_tokens off;
   client_max_body_size 0;
   log_format  main  '$remote_addr - $remote_user [$time_local] "$request" '
                     '$status $body_bytes_sent "$http_referer" '
                     '"$http_user_agent" "$http_x_forwarded_for"';
   access_log /var/log/nginx/access.log  main;
   sendfile       on;
   tcp_nopush     on;
   tcp_nodelay    on;
   keepalive_timeout  65;

   server {
       listen 9180;
       charset utf-8;
```

```
        location / {
            proxy_pass http://localhost:9000;
        }
    }
}
```

Most of the time the service is running without problems. However, we are facing situations where the flask application stops working or fails, yet the docker container keeps running making it difficult to detect when it happens. Additionally, when running as a standalone application, the flask application logs its activity and errors to the standard output, but in production we are not able to see the log.

Please highlight any problem you can spot (if any) in those files. Then diagnose the problem we are facing and propose a working solution.

Note: The actual Python code is irrelevant for our purposes, it could be any arbitrary Flask application.

Requirements

- Put your answers into a GitHub repository and share it with us.
- Try to perform several commits describing your progress.

The Lumu technical tests are not in any sense intended to perform a detailed assessment of technical skills. They are mostly intended to understand the reasoning process and problem solving abilities of the candidates. Therefore, it is important to track your progress and elaborate on your analytical steps while achieving the solution. If you feel that some of the points of this test are too much for a test, that you don't have the time or the right tools to provide a working solution feel free to detail how you would get to a working solution in the hypothetical scenario where you have everything needed to solve the problem.