Python Code to write access log to mongodb

import re

import os

import geoip2.database

from pymongo import MongoClient

log\_file = "/usr/local/nginx/logs/access.log"

mongo\_uri = "mongodb://localhost:27017/"

mongo\_database = "transit\_app"

mongo\_collection = "access\_logs"

client = MongoClient(mongo\_uri)

db = client[mongo\_database]

collection = db[mongo\_collection]

def get\_host\_location(ip):

with geoip2.database.Reader('/usr/share/GeoIP/GeoLite2-City.mmdb') as reader:

response = reader.city(ip)

city\_name = response.city.name if response.city.name else ""

subdivision\_name = response.subdivisions.most\_specific.name if response.subdivisions.most\_specific.name else ""

country\_name = response.country.name if response.country.name else ""

return f"{city\_name}, {subdivision\_name}, {country\_name}"

def insert\_log\_line(line):

pattern = r'^(\S+) (\S+) (\S+) \[([\w:/]+\s[+\-]\d{4})\] "(\S+) (\S+)\s\*(\S+)?\s\*" (\d{3}) (\d+) "(\S+)" "(.\*)"'

match = re.match(pattern, line)

if match is not None:

ip = match.group(1)

host\_location = get\_host\_location(ip)

if host\_location is not None:

log\_data = {

"ip\_address": ip,

"timestamp": match.group(4),

"method": match.group(5),

"url": match.group(6),

"http\_version": match.group(7),

"response\_code": match.group(8),

"content\_size": match.group(9),

"referrer": match.group(10),

"user\_agent": match.group(11),

"host\_location": host\_location

}

collection.insert\_one(log\_data)

def main():

with open(log\_file, "r") as file:

for line in file:

insert\_log\_line(line)

if \_\_name\_\_ == "\_\_main\_\_":

main()

To run this code on a Debian 11 machine, follow these steps:

1. To use the **city** method, you need to download and use a GeoLite2-City database

You can download the GeoLite2-City database from the MaxMind website: <https://dev.maxmind.com/geoip/geoip2/geolite2/>

1. Install Python and pip if they are not already installed:

sudo apt-get update sudo apt-get install python3 python3-pip

1. Install the required Python packages:

pip3 install pymongo

1. Create a new file called **nginx\_log\_to\_mongodb.py** and copy the code above into the file.
2. Edit the **MONGODB\_URI**, **MONGODB\_DATABASE**, and **MONGODB\_COLLECTION** variables at the top of the file to match your MongoDB settings.
3. Install **geoip2** library

pip install geoip2

1. Save the file and run it:

python3 nginx\_log\_to\_mongodb.py

This will read the Nginx access log file at **/usr/local/nginx/logs/access.log** and insert each log line into the specified MongoDB collection. You can run this script as a cron job to periodically insert new logs into MongoDB. Note that this is a basic example and you may need to modify the code to match your specific Nginx log format and MongoDB schema.