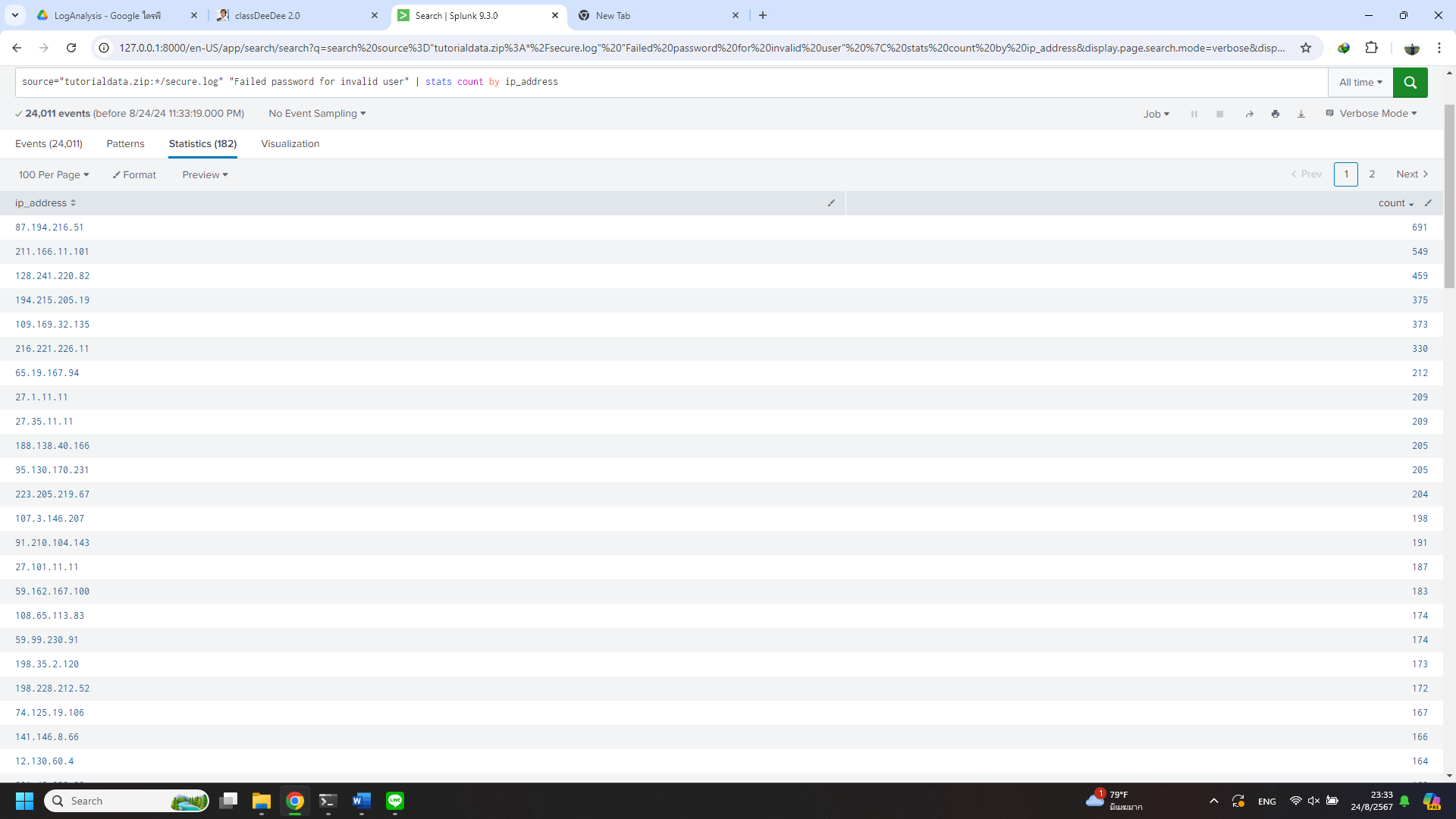
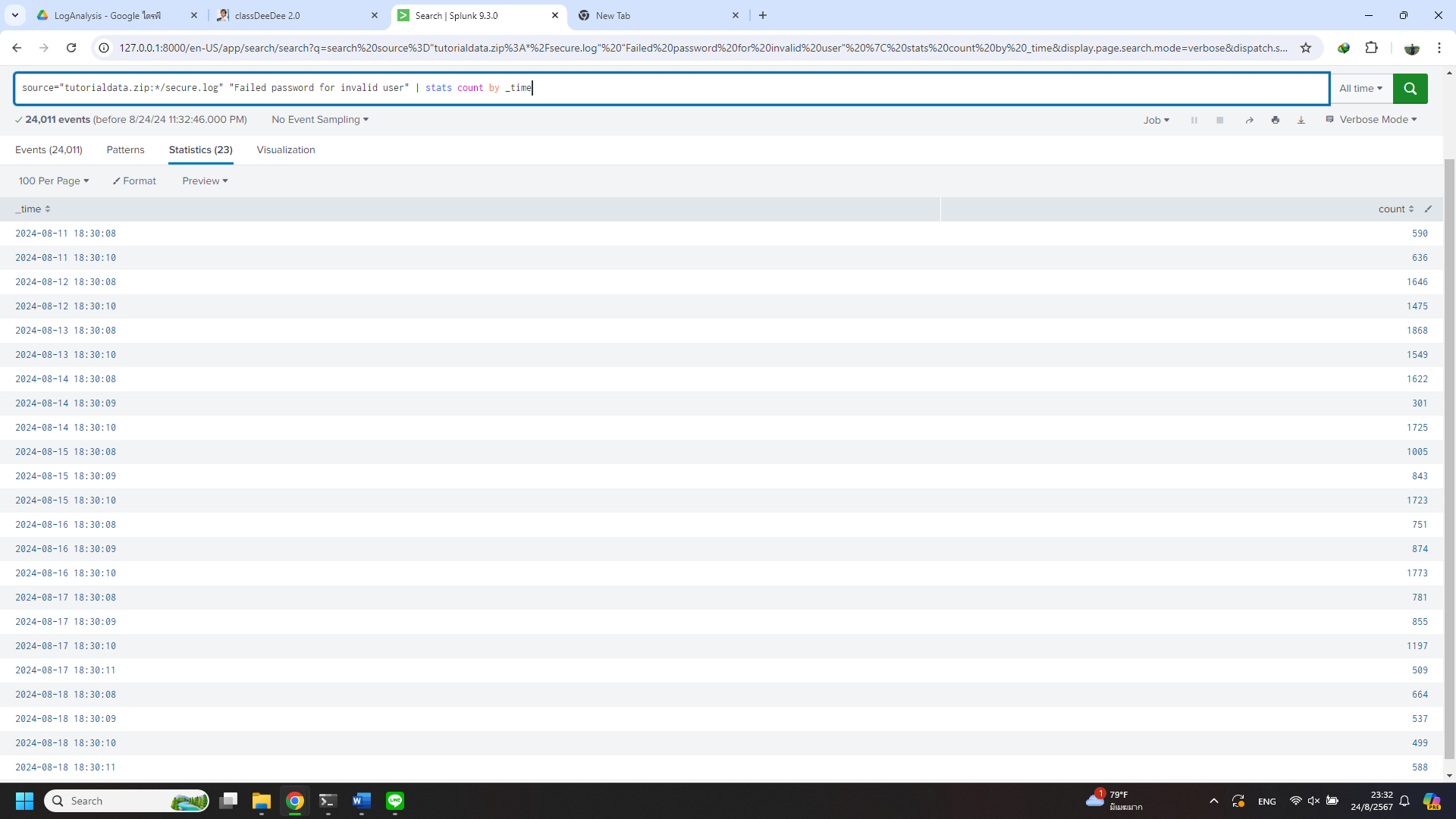
Q1. How many hackers are trying to get access to our servers? And how many attempts are there? Explain/define how you count distinct hackers.



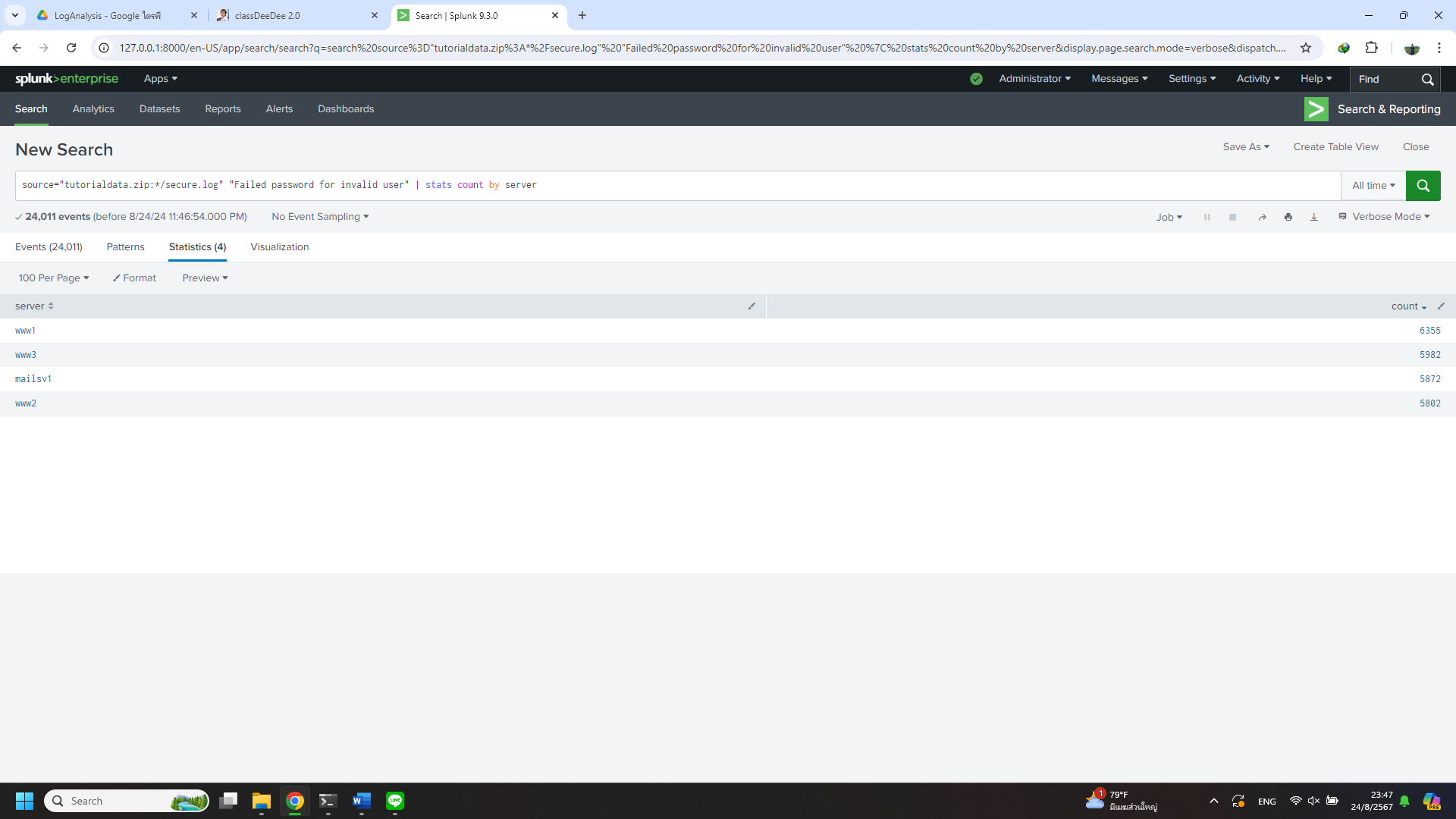
Ans. I assume that each hacker uses their own ip-address and hacker try to access with unknow username so log had to be “Failed password for invalid user” telling us that hacker try to reverse retrieve username. So, I make new field via regular expression   
from (?<ip\_address>\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3}) and search in secure.log then pipe to count  
As the result of 182 hackers and 24011 attempts.

Q2. What time do hackers appear to try to hack our servers?



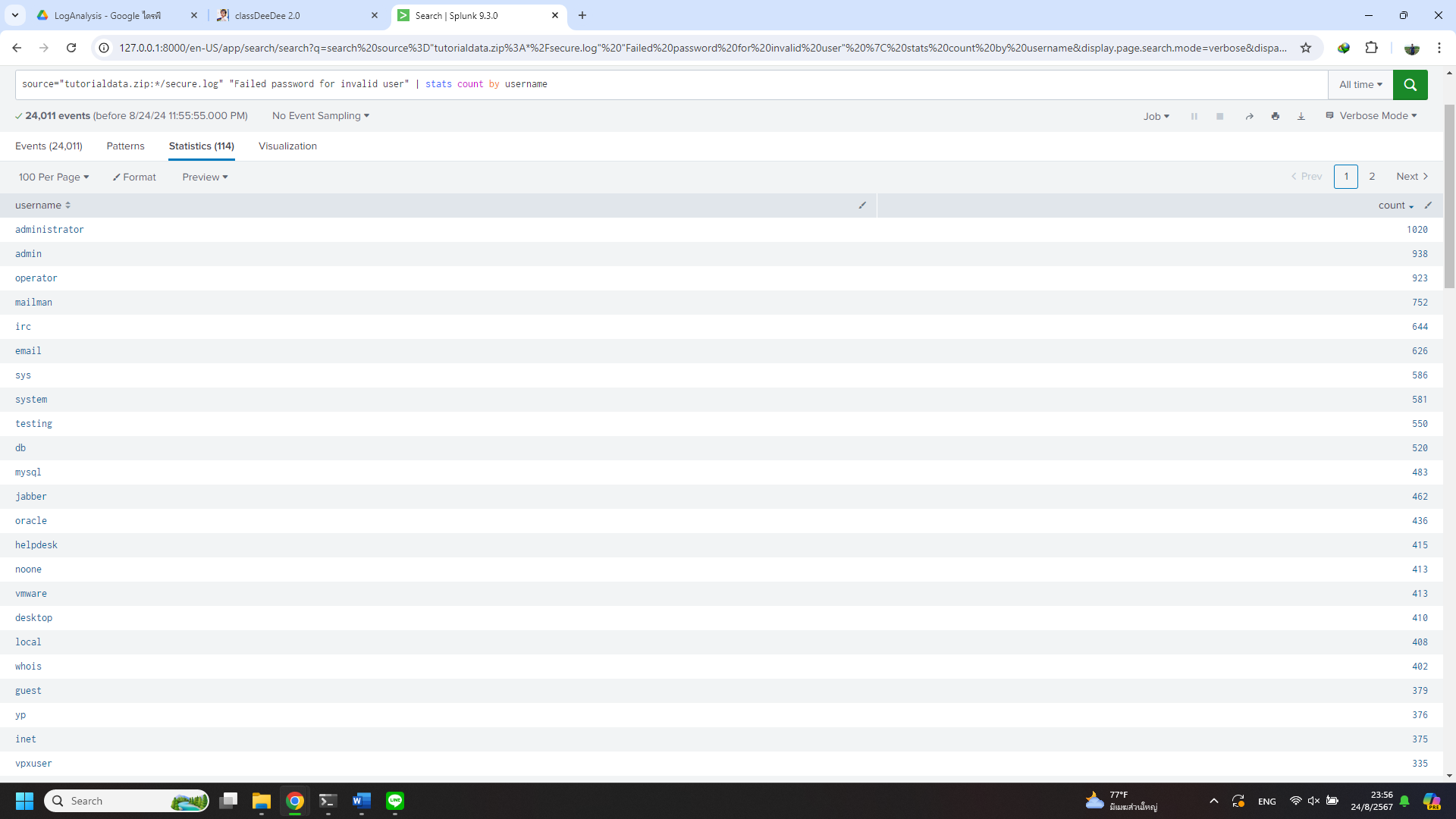
Ans. From 11/08/2024 to 18/08/2024 during 18:30:08 to 18:30:11 for all attempts.

Q3. Which server (mailsv, www1, www2, www3) had the most attempts?

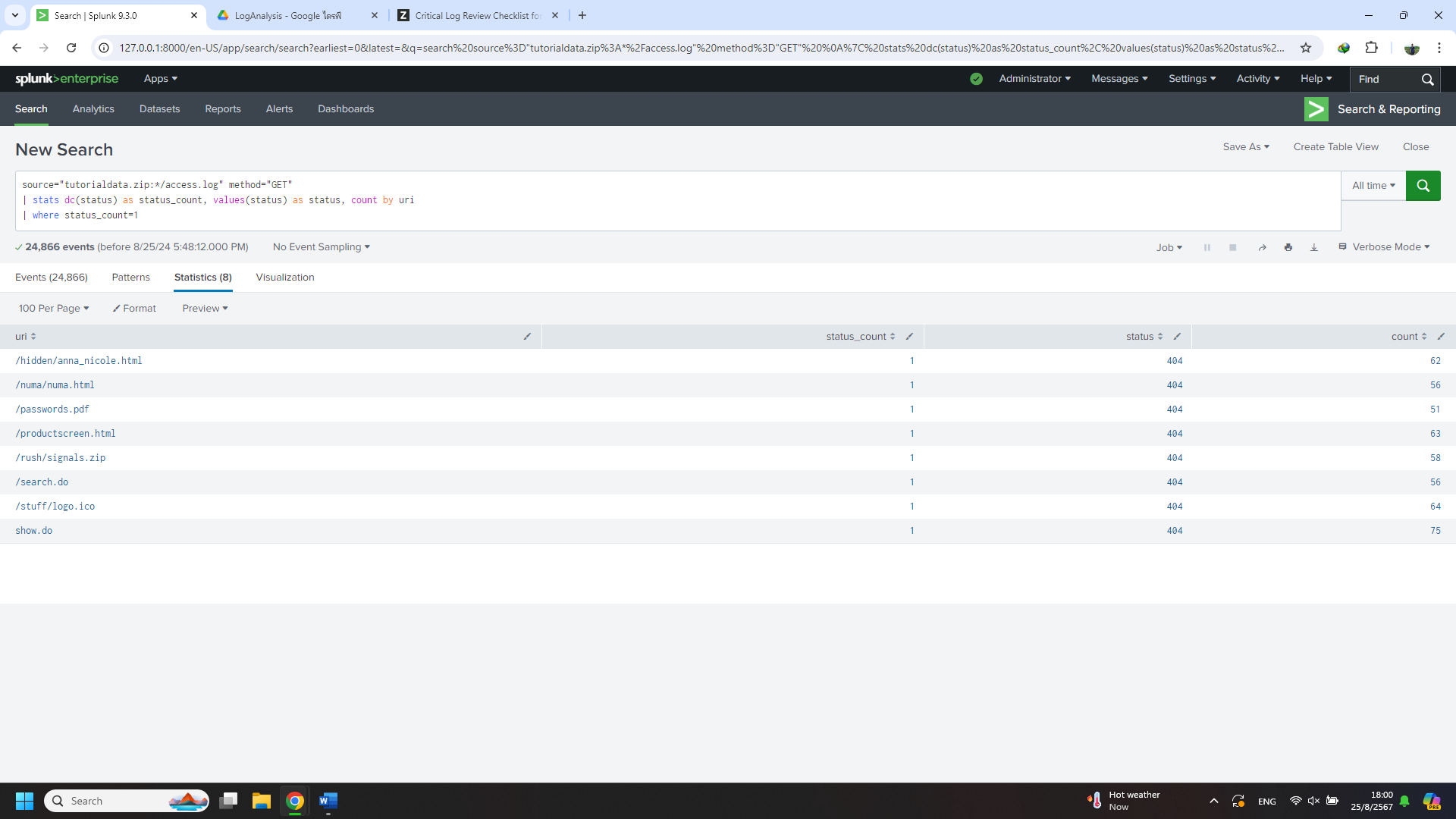


Ans. This time regular expression for extract field generated by splunk working fine. So, this time I just count by newly extracted field name “server”. As the result of Server www1 with 6355 attempts.

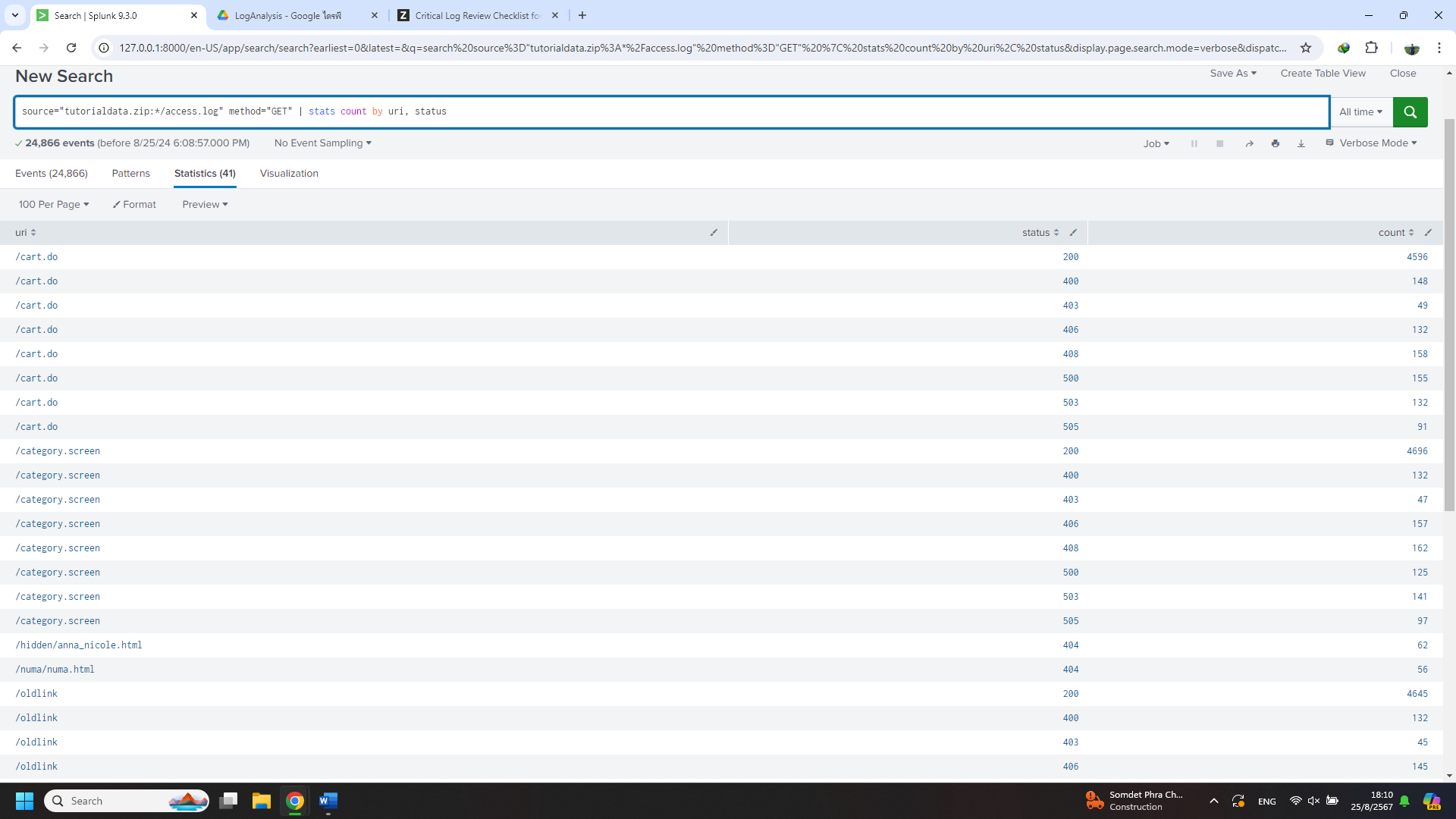
Q4. What is the most popular account that hackers use to try to break in?

Ans. Extract username and result of user administrator 1020 attempts.

Q5. Can you find attempts to get access to sensitive information from our web servers? How many attempts were there?



Ans. I assume that these attempts might have unique behavior like have only one type of error status from request rather some distribution of status like image below, you can see that some request have many types of status and distributed in making sense way.



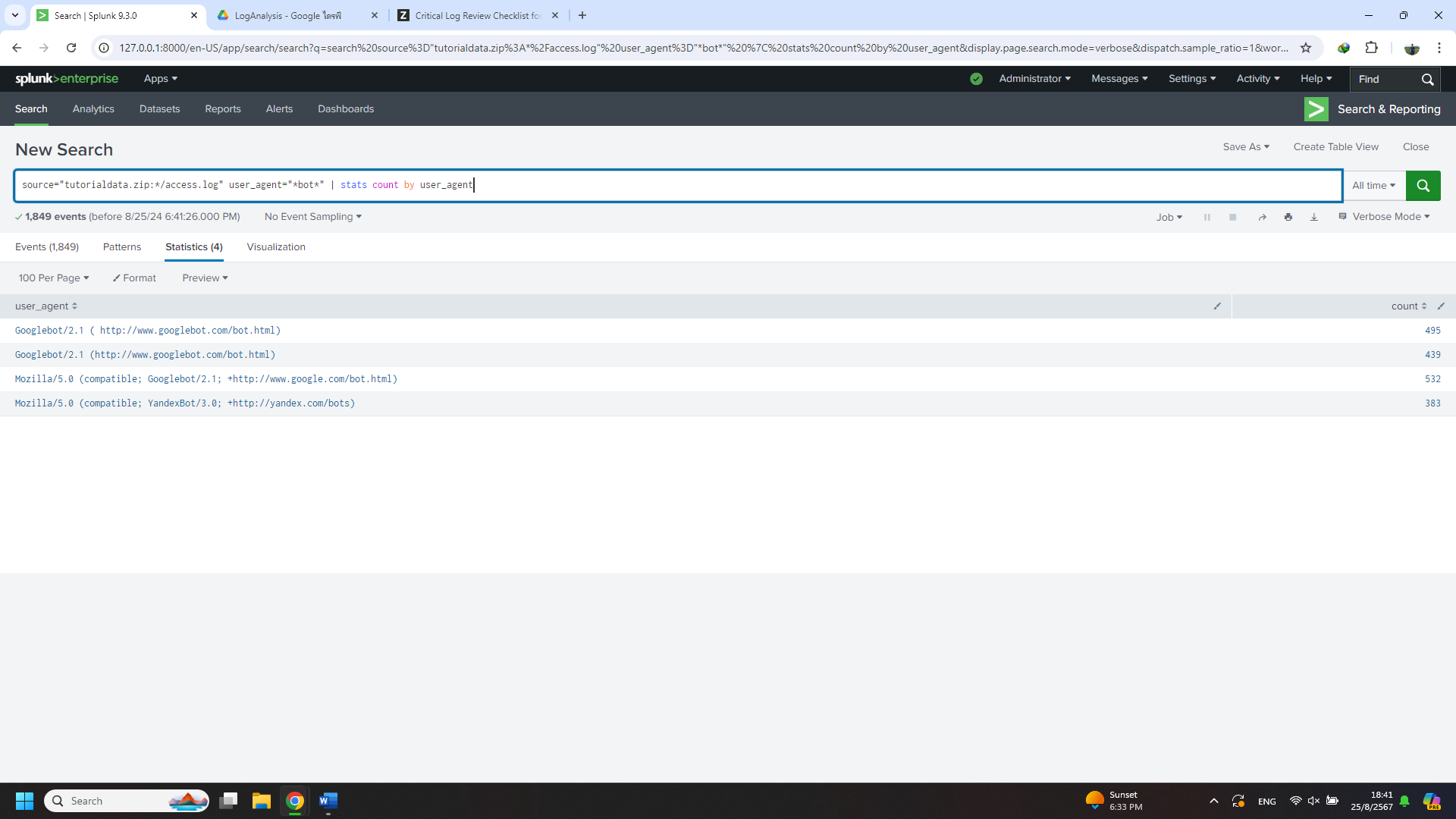
Second, sometimes error 404 just maybe not found something or some bug in server but for now as I didn’t know server well, I will list that all as potentially hacking attempts. Which is /hidden/anna\_nicole.html, /numa/numa.html, /passwords.pdf, /productscreen.html, /rush/signals.zip, /search.do, /stuff/logo.ico, show.do with attempts from 51 to 75 attempts for each URI with total of 485 attempts.

Q6. What resource/file are hackers looking for?

Ans. As a result of Q5,

1. /hidden/anna\_nicole.html
2. /numa/numa.html
3. /passwords.pdf
4. /productscreen.html
5. /rush/signals.zip
6. /search.do
7. /stuff/logo.ico
8. show.do

Q7. Can you find any bots crawling our websites?



Ans. Yes, these are the two main types of bots crawling our server (assuming that bots include the word "bot" in their user agent name), Google bot and Yandex bot.

Q8. What are they doing on the site? (Hint: Look for User-Agent in the webaccess.logs.)

Ans. Based on my research on Google, these bots are beneficial for our website as they help our web pages appear in search results when the content matches a user's search query. Google is a well-known search engine globally, while Yandex is a prominent search engine in Russia.