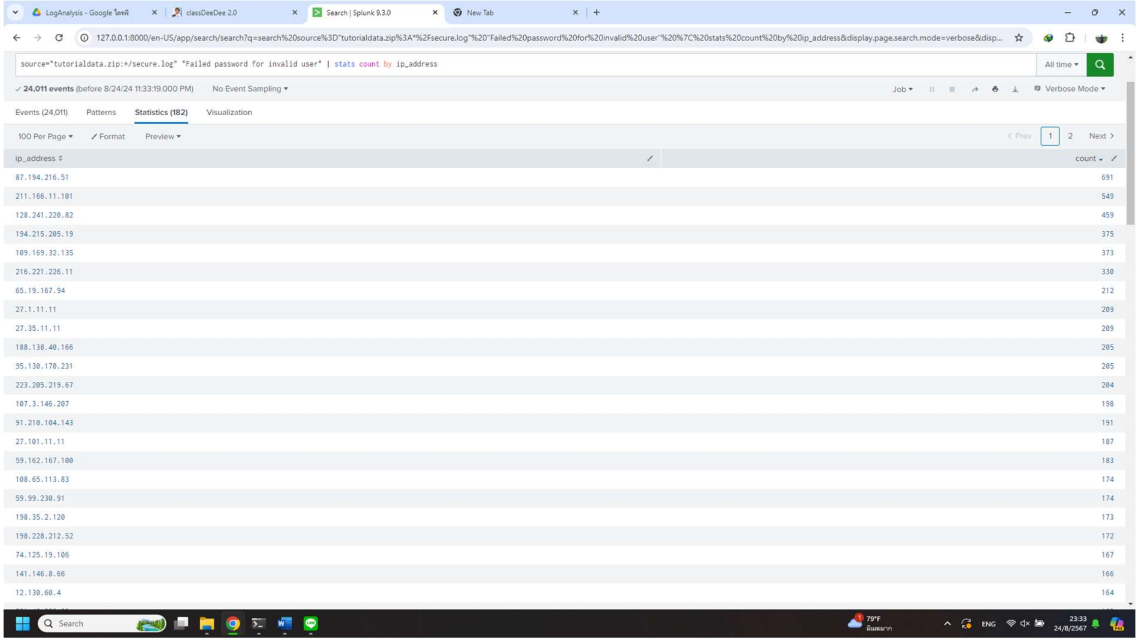
Q1. How many hackers are trying to get access to our servers? And how many attempts are there? Explain/define how you count distinct hackers.


LogAnalysis - Google ไดรฟ์
classDeeDee 2.0
Search | Splunk 9.3.0
New Tab
127.0.0.1:8000/en-US/app/search/search?q=search%20source%3D"tutorialdata.zip%3A*%2Fsecure.log"%20"Failed%20password%20for%20invalid%20user"%20%7C%20stats%20count%20by%20ip_address&display.page.search.mode=verbose&disp...
All time ▾
No Event Sampling ▾
Job ▾
Verbose Mode ▾
Events (24,011)
Patterns
Statistics (182)
Visualization
100 Per Page ▾
Format
Preview ▾
‹ Prev
1
2
Next ›
Search
79°F
มีเมฆมาก
ENG
23:33
24/8/2567

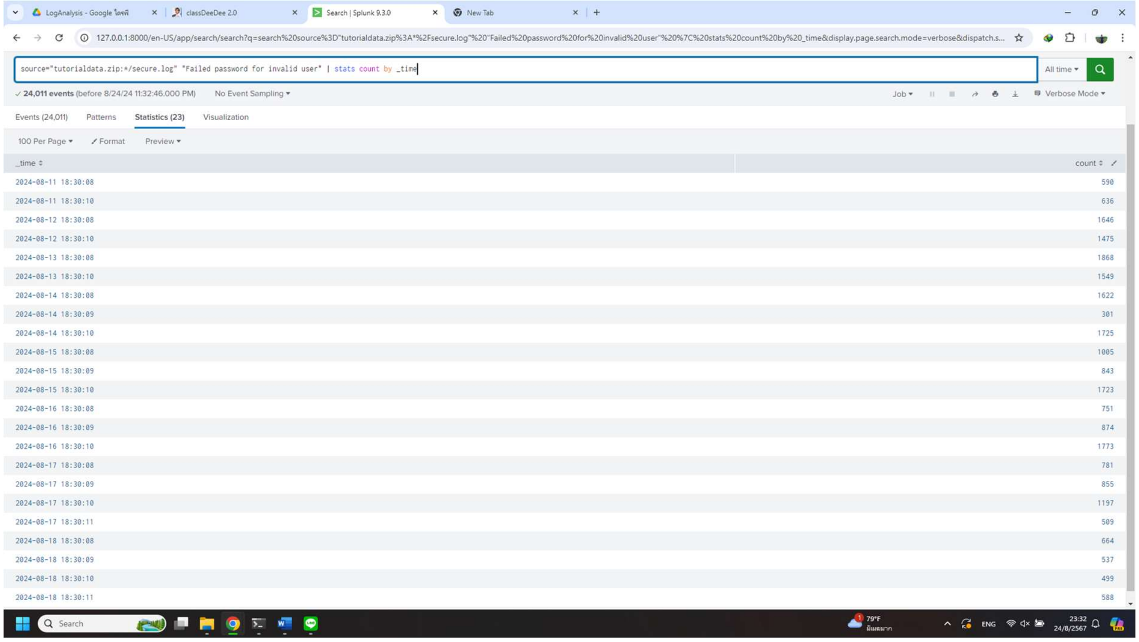
```
source="tutorialdata.zip:*/secure.log" "Failed password for invalid user" | stats count by ip_address
```

✓ 24,011 events (before 8/24/24 11:33:19.000 PM)

ip_address	count
87.194.216.51	691
211.166.11.101	549
128.241.220.82	459
194.215.205.19	375
109.169.32.135	373
216.221.226.11	330
65.19.167.94	212
27.1.11.11	209
27.35.11.11	209
188.138.40.166	205
95.130.170.231	205
223.205.219.67	204
107.3.146.207	198
91.210.104.143	191
27.101.11.11	187
59.162.167.100	183
108.65.113.83	174
59.99.230.91	174
198.35.2.120	173
198.228.212.52	172
74.125.19.106	167
141.146.8.66	166
12.130.60.4	164

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?


LogAnalysis - Google ไดรฟ์
classDeeDee 2.0
Search | Splunk 9.3.0
New Tab
127.0.0.1:8000/en-US/app/search/search?q=search%20source%3D"tutorialdata.zip%3A*%2Fsecure.log"%20"Failed%20password%20for%20invalid%20user"%20%7C%20stats%20count%20by%20_time&display.page.search.mode=verbose&dispatch.s...
All time ▾
No Event Sampling ▾
Job ▾
Verbose Mode ▾
Events (24,011)
Patterns
Statistics (23)
Visualization
100 Per Page ▾
Format
Preview ▾
Search
79°F
มีเมฆมาก
ENG
23:32
24/8/2567

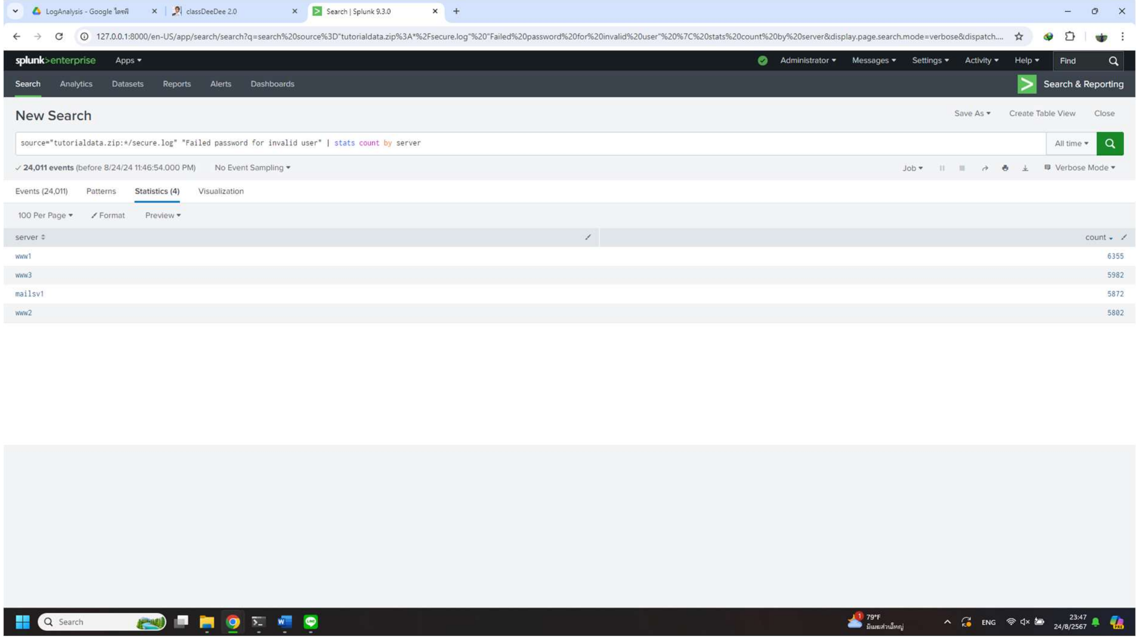
```
source="tutorialdata.zip:*/secure.log" "Failed password for invalid user" | stats count by _time
```

✓ 24,011 events (before 8/24/24 11:32:46.000 PM)

_time	count
2024-08-11 18:30:08	590
2024-08-11 18:30:10	636
2024-08-12 18:30:08	1646
2024-08-12 18:30:10	1475
2024-08-13 18:30:08	1868
2024-08-13 18:30:10	1549
2024-08-14 18:30:08	1622
2024-08-14 18:30:09	301
2024-08-14 18:30:10	1725
2024-08-15 18:30:08	1005
2024-08-15 18:30:09	843
2024-08-15 18:30:10	1723
2024-08-16 18:30:08	751
2024-08-16 18:30:09	874
2024-08-16 18:30:10	1773
2024-08-17 18:30:08	781
2024-08-17 18:30:09	855
2024-08-17 18:30:10	1197
2024-08-17 18:30:11	509
2024-08-18 18:30:08	664
2024-08-18 18:30:09	537
2024-08-18 18:30:10	499
2024-08-18 18:30:11	588

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?



New Search

```
source="tutorialdata.zip:*/secure.log" "Failed password for invalid user" | stats count by server
```

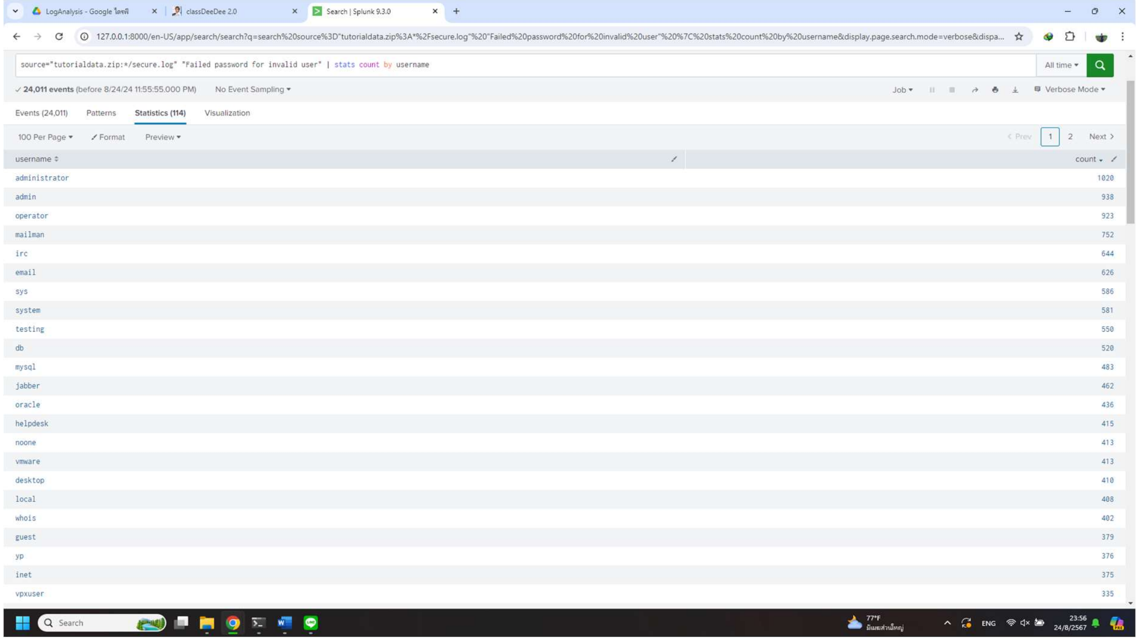
✓ 24,011 events (before 8/24/24 11:46:54.000 PM)

Events (24,011) Patterns Statistics (4) Visualization

server	count
www1	6355
www3	5982
mailsv1	5872
www2	5802

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?



```
source="tutorialdata.zip:*/secure.log" "Failed password for invalid user" | stats count by username
```

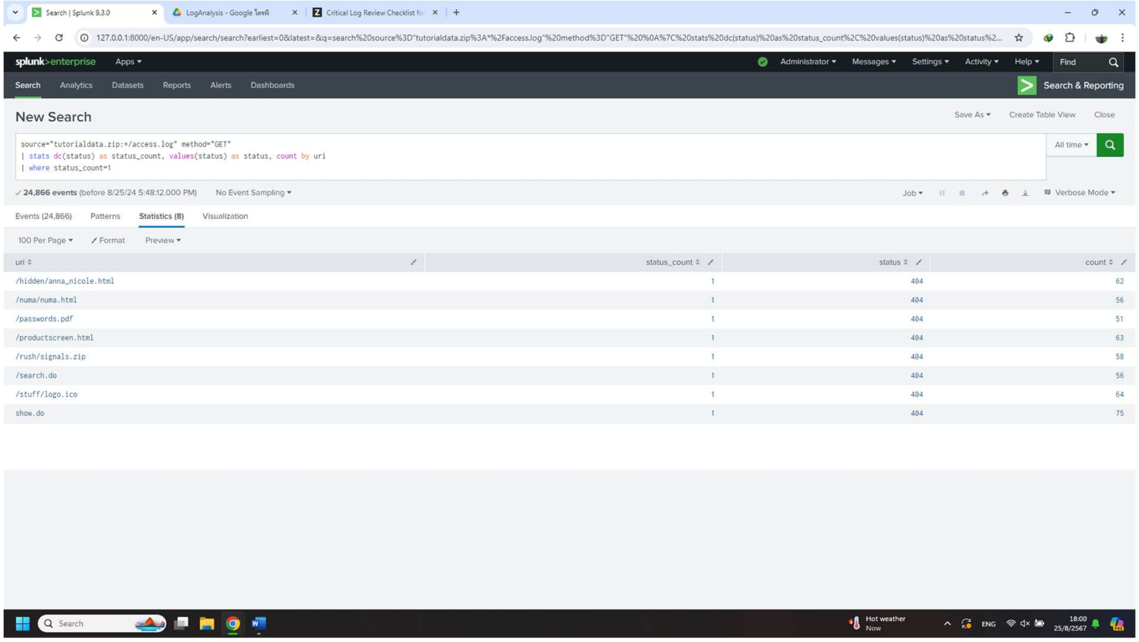
✓ 24,011 events (before 8/24/24 11:55:55.000 PM)

Events (24,011) Patterns Statistics (114) Visualization

username	count
administrator	1020
admin	938
operator	923
mailman	752
irc	644
email	626
sys	586
system	581
testing	550
db	520
mysql	483
jabber	462
oracle	436
helpdesk	415
noone	413
vmware	413
desktop	410
local	408
whois	402
guest	379
yp	376
inet	375
vpxuser	335

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?

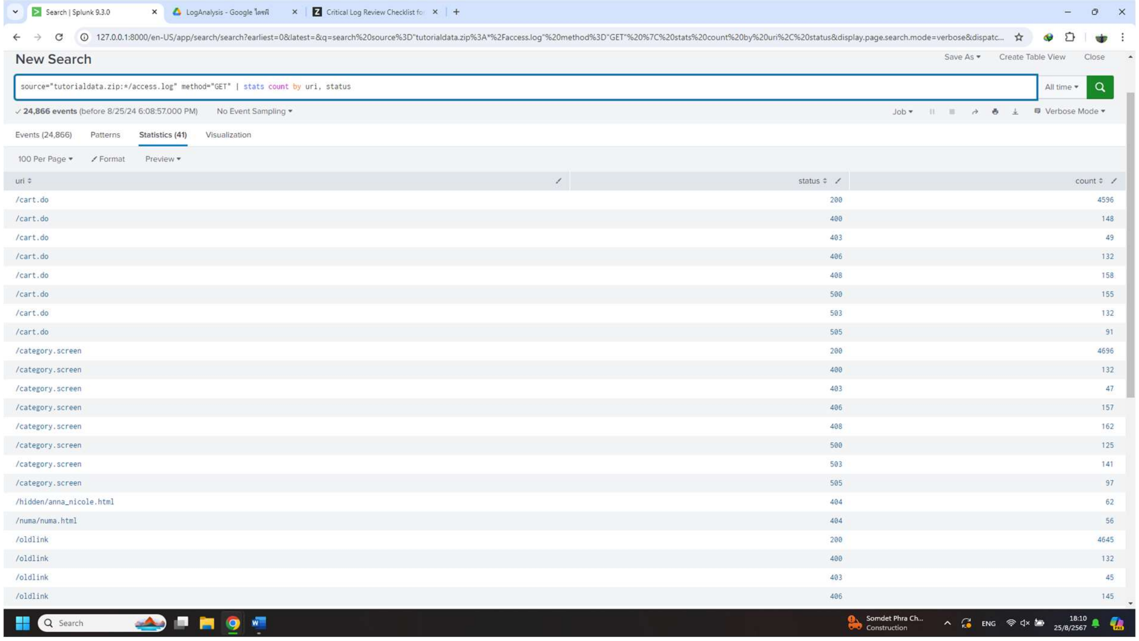


```
source="tutorialdata.zip:*/access.log" method="GET"
| stats dc(status) as status_count, values(status) as status, count by uri
| where status_count=1
```

✓ 24,866 events (before 8/25/24 5:48:12.000 PM) No Event Sampling

uri	status_count	status	count
/hidden/anna_nicole.html	1	404	62
/numa/numa.html	1	404	56
/passwords.pdf	1	404	51
/productscreen.html	1	404	63
/rush/signals.zip	1	404	58
/search.do	1	404	56
/stuff/logo.ico	1	404	64
show.do	1	404	75

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.


Search | Splunk 9.3.0
LogAnalysis - Google ไดรฟ์
Critical Log Review Checklist fo
127.0.0.1:8000/en-US/app/search/search?earliest=0&latest=&q=search%20source%3D"tutorialdata.zip%3A*%2Faccess.log"%20method%3D"GET"%20%7C%20stats%20count%20by%20uri%2C%20status&display.page.search.mode=verbose&dispatc...
New Search
Save As
Create Table View
Close
All time
Job
Verbose Mode
Events (24,866)
Patterns
Statistics (41)
Visualization
100 Per Page
Format
Preview
Somdet Phra Ch...
Construction
ENG
18:10
25/8/2567

```
source="tutorialdata.zip:*/access.log" method="GET" | stats count by uri, status
```

✓ 24,866 events (before 8/25/24 6:08:57.000 PM) No Event Sampling

uri	status	count
/cart.do	200	4596
/cart.do	400	148
/cart.do	403	49
/cart.do	406	132
/cart.do	408	158
/cart.do	500	155
/cart.do	503	132
/cart.do	505	91
/category.screen	200	4696
/category.screen	400	132
/category.screen	403	47
/category.screen	406	157
/category.screen	408	162
/category.screen	500	125
/category.screen	503	141
/category.screen	505	97
/hidden/anna_nicole.html	404	62
/numa/numa.html	404	56
/oldlink	200	4645
/oldlink	400	132
/oldlink	403	45
/oldlink	406	145

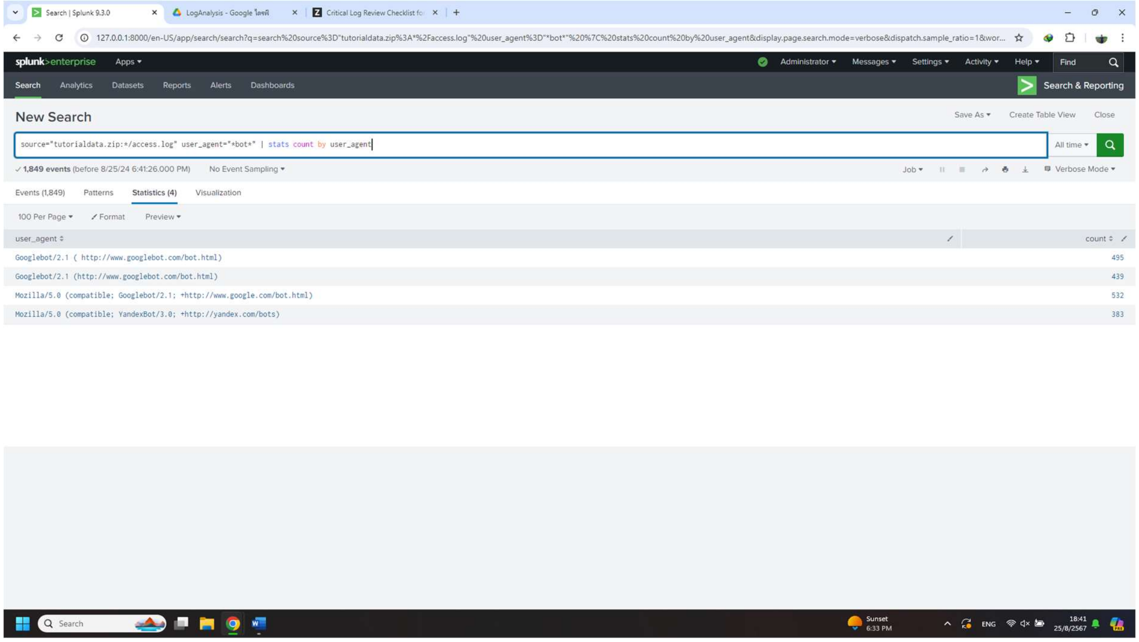
Second, sometimes error 404 just maybe not found something or some bug in server but for now as I didn't know server well, I will list that all as potentially hacking attempts. Which is /hidden/anna_nicole.html, /numa/numa.html, /passwords.pdf, /productscreen.html, /rush/signals.zip, /search.do, /stuff/logo.ico, 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?



```
127.0.0.1:8000/en-US/app/search/search?q=search%20source%3D"tutorialdata.zip%3A*%2Faccess.log"%20user_agent%3D"*bot*"%20%7C%20stats%20count%20by%20user_agent&display.page.search.mode=verbose&dispatch.sample_ratio=1&wor...
```

New Search

```
source="tutorialdata.zip:*/access.log" user_agent="*bot*" | stats count by user_agent
```

✓ 1,849 events (before 8/25/24 6:41:26.000 PM)

user_agent	count
Googlebot/2.1 (http://www.googlebot.com/bot.html)	495
Googlebot/2.1 (http://www.googlebot.com/bot.html)	439
Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	532
Mozilla/5.0 (compatible; YandexBot/3.0; +http://yandex.com/bots)	383

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.