

Activity: Log Analysis

Pupipat Singkhorn
Student ID: 6532142421
30 August 2025

Part I. Can you find people trying to break into the servers?

Use Splunk's "Search" feature to try to answer the questions below.

Hint 1: On linux servers, secure.log contains security-related information. Typically in response to incidents, it is one of the first files people look at to see if there are compromises. Read this to see what to look for on Linux (this file is available in google drive). <https://zeltser.com/security-incident-log-review-checklist/>

Hint 2: To process the logs for analysis, first parse it using regular expressions to "extract fields" and turn unstructured data into structured data.

Answer the following questions and provide evidence with your answer.

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.

Answer:

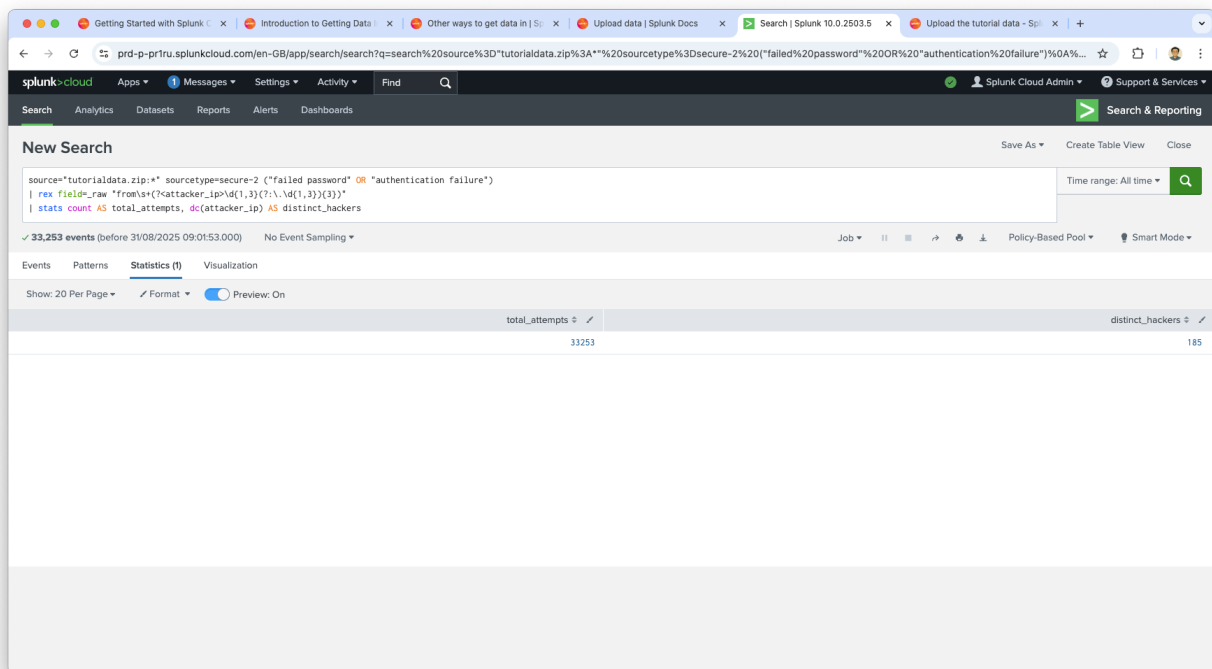
- Total attempts: **#33,253** failed login attempts.
- Distinct hackers: **#185** unique IP addresses.

We treat a hacker as a unique source IP address that generated failed login events in the secure.log.

- Every failed login (failed password, authentication failure) is an attempt.
- Multiple attempts from the same IP (even if trying different usernames) still count as one hacker.

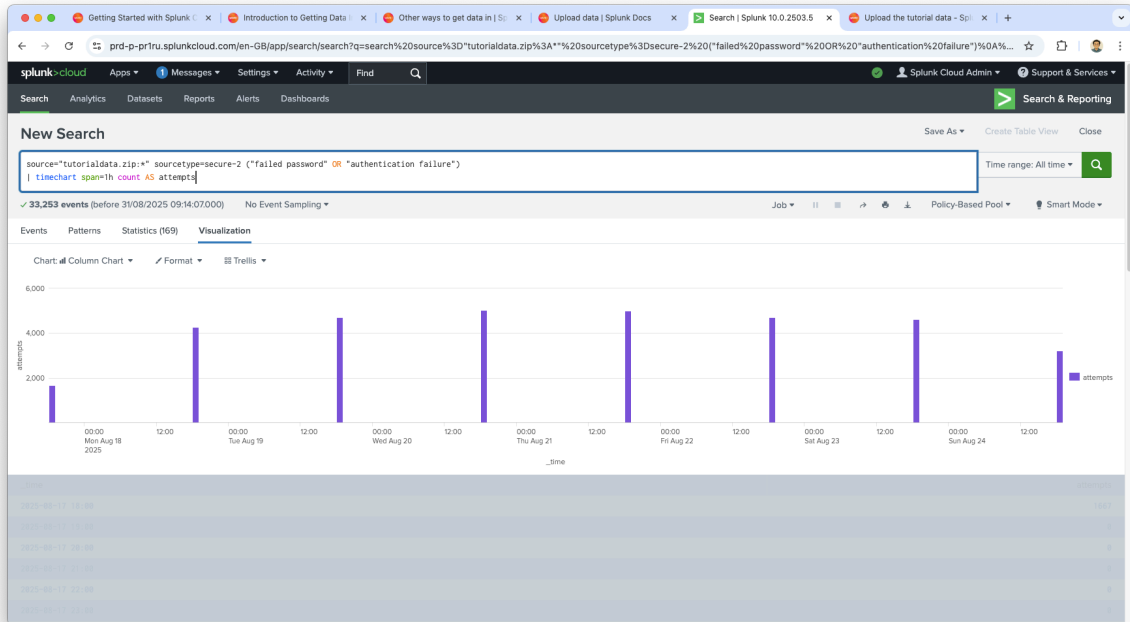
Therefore:

- Attempts = count of all failed login events.
- Hackers = distinct count of attacker IPs extracted from the logs.



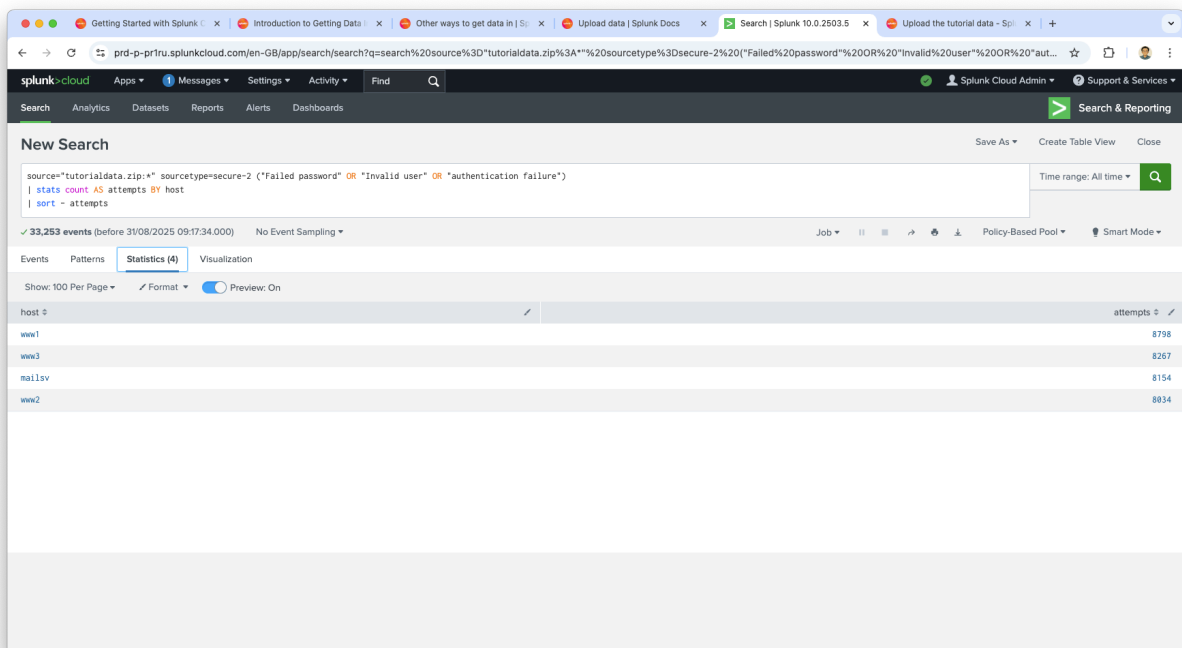
Q2. What time do hackers appear to try to hack our servers?

Answer: 17-24 August 2025 at 18:00



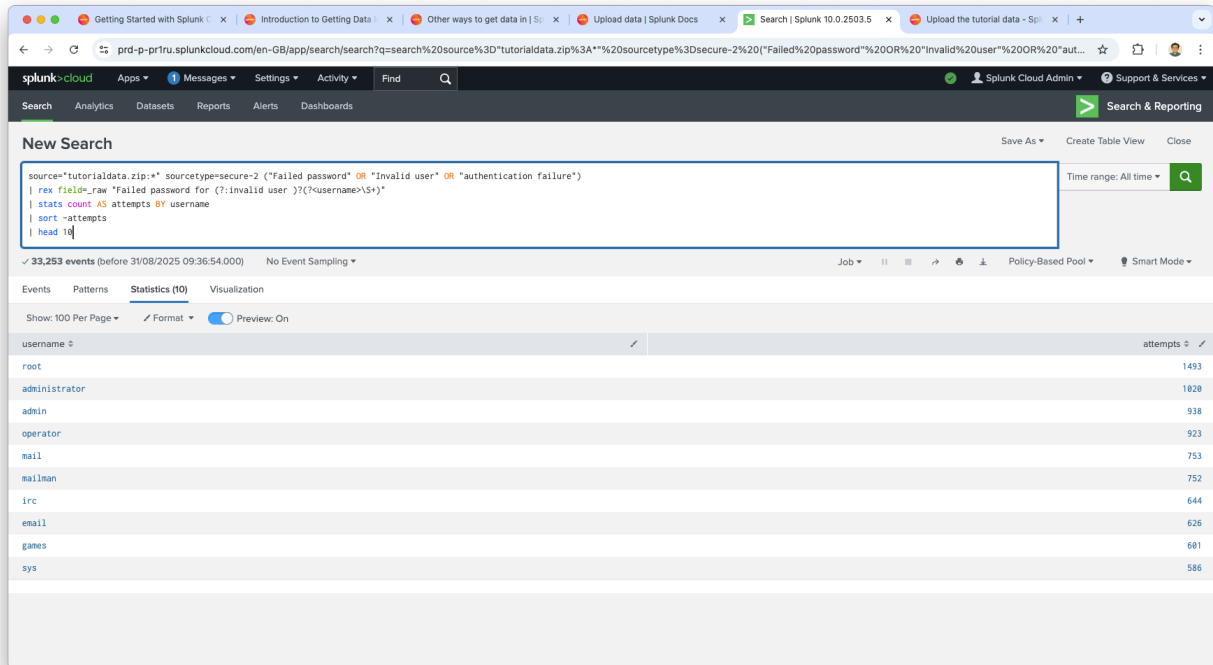
Q3. Which server (mailsv, www1, www2, www3) had the most attempts?

Answer: www1



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

Answer: root



The screenshot shows the Splunk Cloud interface with a search query for failed password attempts. The search results are displayed in a table with columns for 'username' and 'attempts'.

Search Query:

```
source="tutorialdata.zip:*" sourcetype=secure-2 ("Failed password" OR "Invalid user" OR "authentication failure")
| rex field=_raw "Failed password for (?::invalid user :)?(?<username>{+})"
| stats count AS attempts BY username
| sort -attempts
| head 10
```

Search Results:

username	attempts
root	1493
administrator	1020
admin	938
operator	923
mail	753
mailman	752
irc	644
email	626
games	601
sys	586

Part II. Sensitive Files on Web Servers

Hint: On web servers, access.log contains web access-related information. Typically in response to incidents, it is one of the first files people look at to see if there are compromises. Read this to see what to look for on Web Servers.

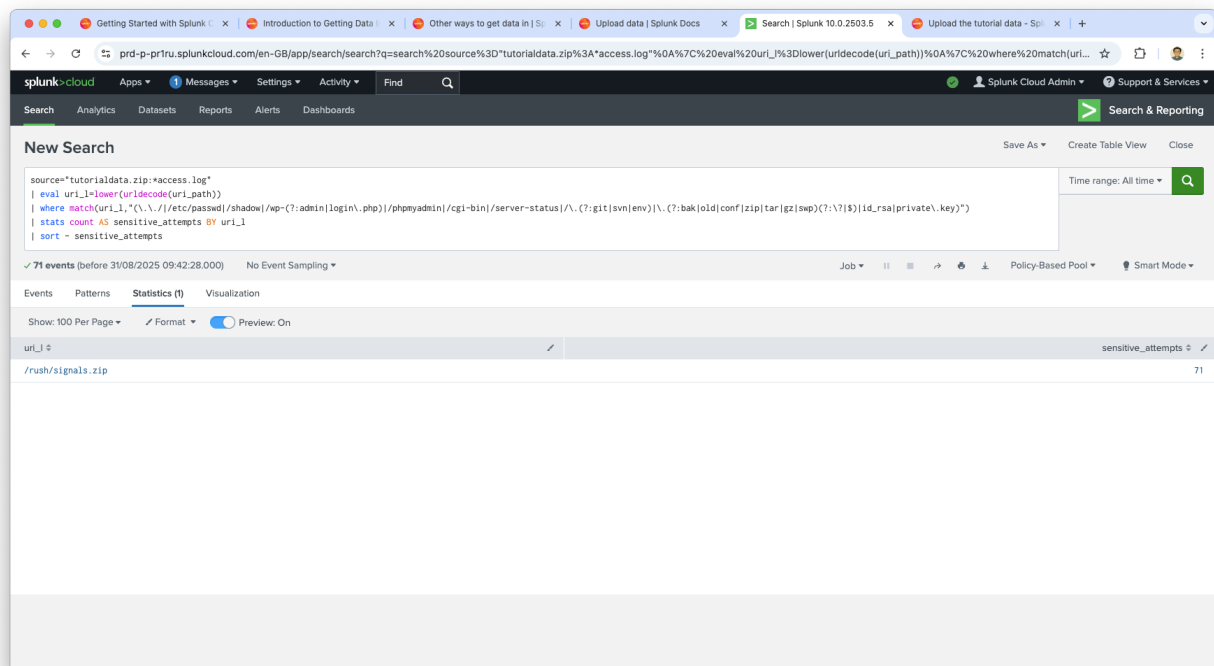
<https://zeltser.com/security-incident-log-review-checklist/>

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

Answer: 71 attempts

Q6. What resource/file are hackers looking for?

Answer: /rush/signals.zip



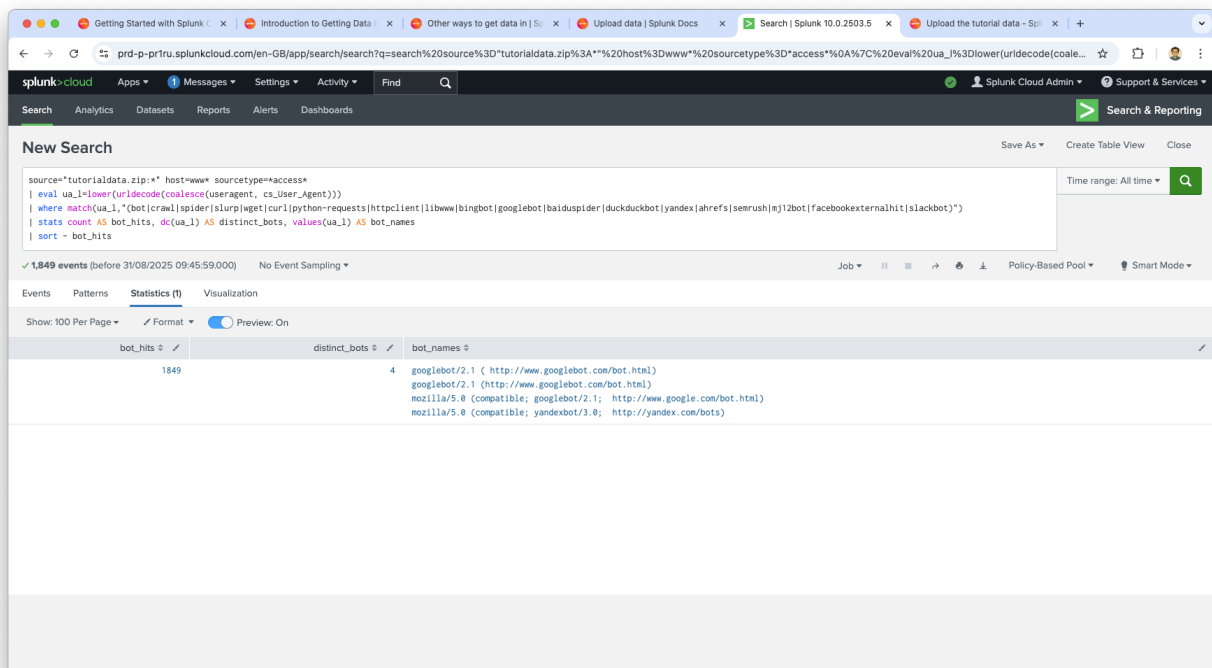
Part III. Are there bots crawling our websites?

Q7. Can you find any bots crawling our websites?

Answer: Yes. We found evidence of bots crawling our websites. Splunk analysis shows 1,849 requests from 4 distinct bots, including Googlebot and YandexBot.

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

Answer: The bots are crawling and indexing our website content, as indicated by their User-Agent strings (e.g., Googlebot, YandexBot). This activity is consistent with normal search engine behavior.



The screenshot shows the Splunk Cloud interface with a search query executed. The search bar contains the following query:

```
source=tutorialdata.zip:* host=www* sourcetype=access*
| eval ua_l=lower(uridecode(coalesce(useragent, cs_User_Agent)))
| where match(ua_l,"(bot|crawl|spider|slurp|wget|curl|python-requests|httpclient|libwww|bingbot|googlebot|baiduspider|duckduckbot|yandex|ahrefs|semrush|mj12bot|facebookexternalhit|slackbot)")
| stats count AS bot_hits, dc(ua_l) AS distinct_bots, values(ua_l) AS bot_names
| sort - bot_hits
```

The search results show 1,849 events. The table view displays the following data:

bot_hits	distinct_bots	bot_names
1849	4	googlebot/2.1 (http://www.googlebot.com/bot.html) googlebot/2.1 (http://www.googlebot.com/bot.html) mozilla/5.0 (compatible; googlebot/2.1; http://www.google.com/bot.html) mozilla/5.0 (compatible; yandexbot/3.0; http://yandex.com/bots)