🌐 HTTP Log Analysis using Splunk Project

By Gaurav Ghandat

Hands-on cybersecurity lab focused on analyzing HTTP traffic using **Splunk** and **Zeek-style JSON logs** to detect errors, anomalies, and suspicious web activity.

🎯 Objective

In this lab, you will:

- Ingest and analyze HTTP logs using **Splunk**
- Detect **client-side (4xx)** and **server-side (5xx)** HTTP errors
- Identify **suspicious User-Agents and URIs**
- Detect **large file transfers** that may indicate data exfiltration
- Gain practical experience with **SPL (Search Processing Language)**

🧰 Tech Stack & Tools

- **SIEM Tool:** Splunk Enterprise
- **Log Source:** Zeek-style HTTP logs (JSON format)
- **Index:** http_lab (custom index)
- **Sourcetype:** json or zeek:http

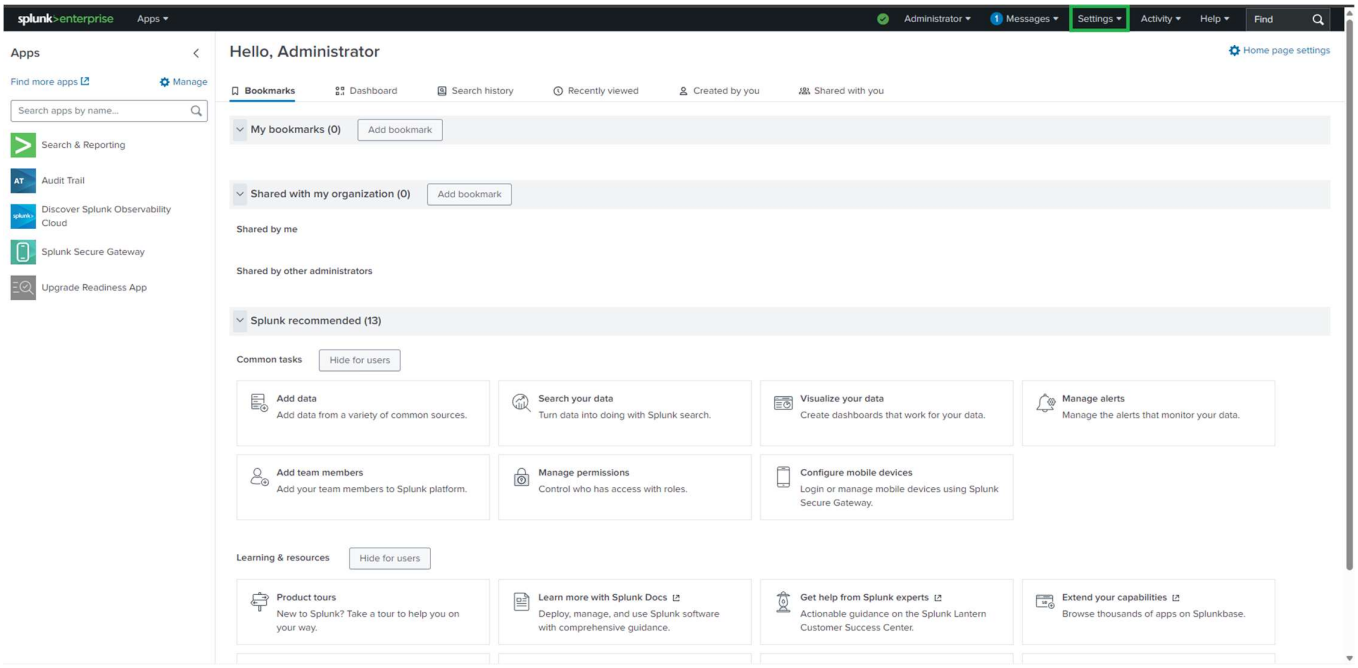
🖥️ Lab Setup

✅ Prerequisites

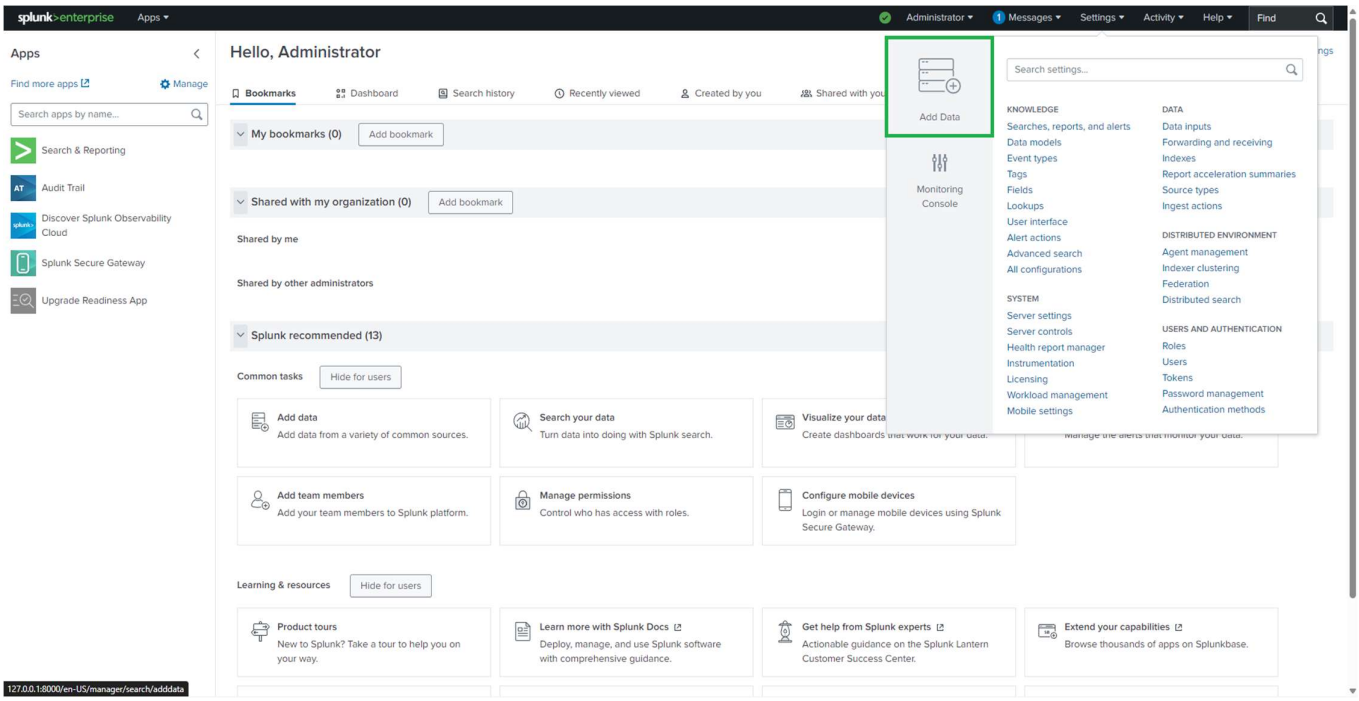
- Splunk installed and accessible via Splunk Web
- HTTP logs in JSON format (http_logs.json)

Data Ingestion Steps

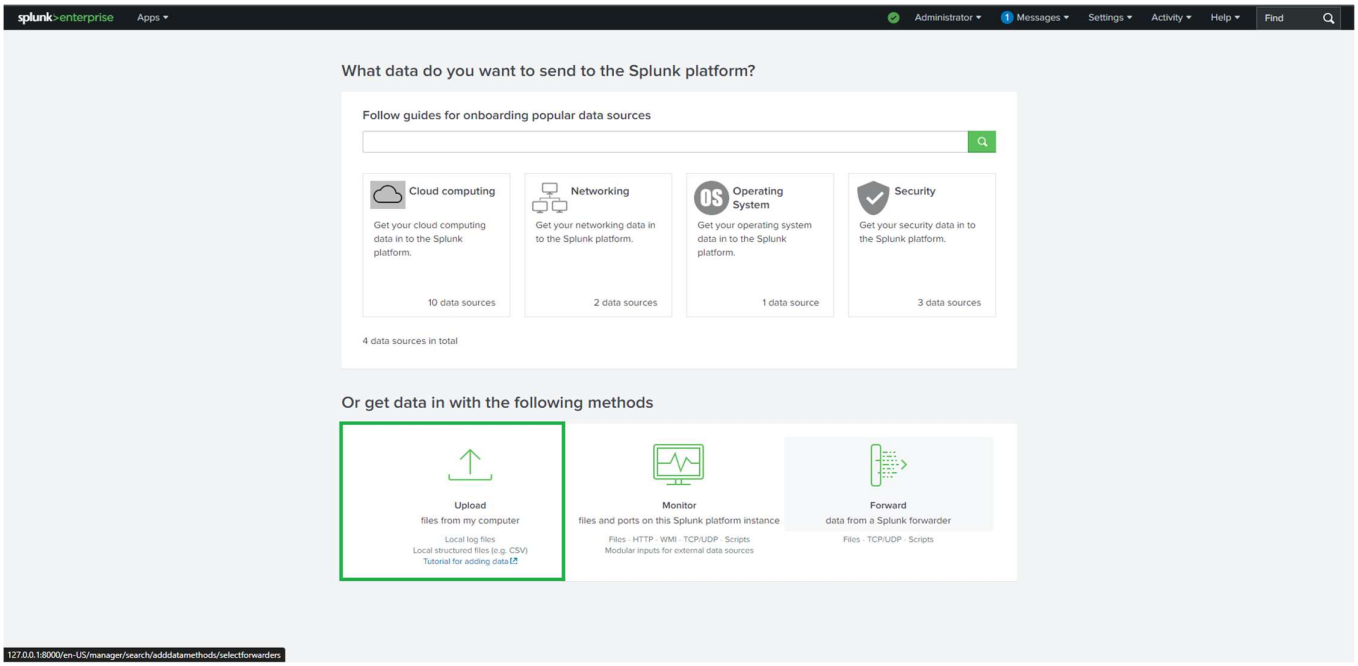
1. Open **Splunk Web**

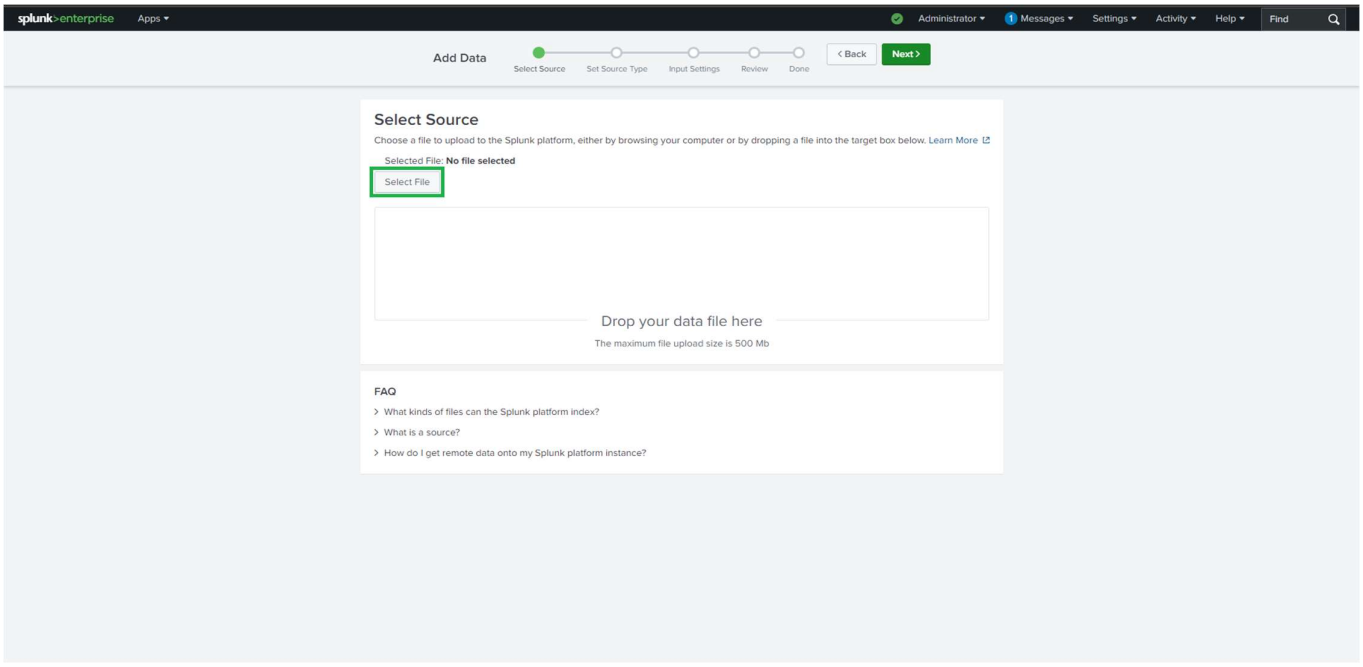


2. Navigate to **Settings → Add Data**

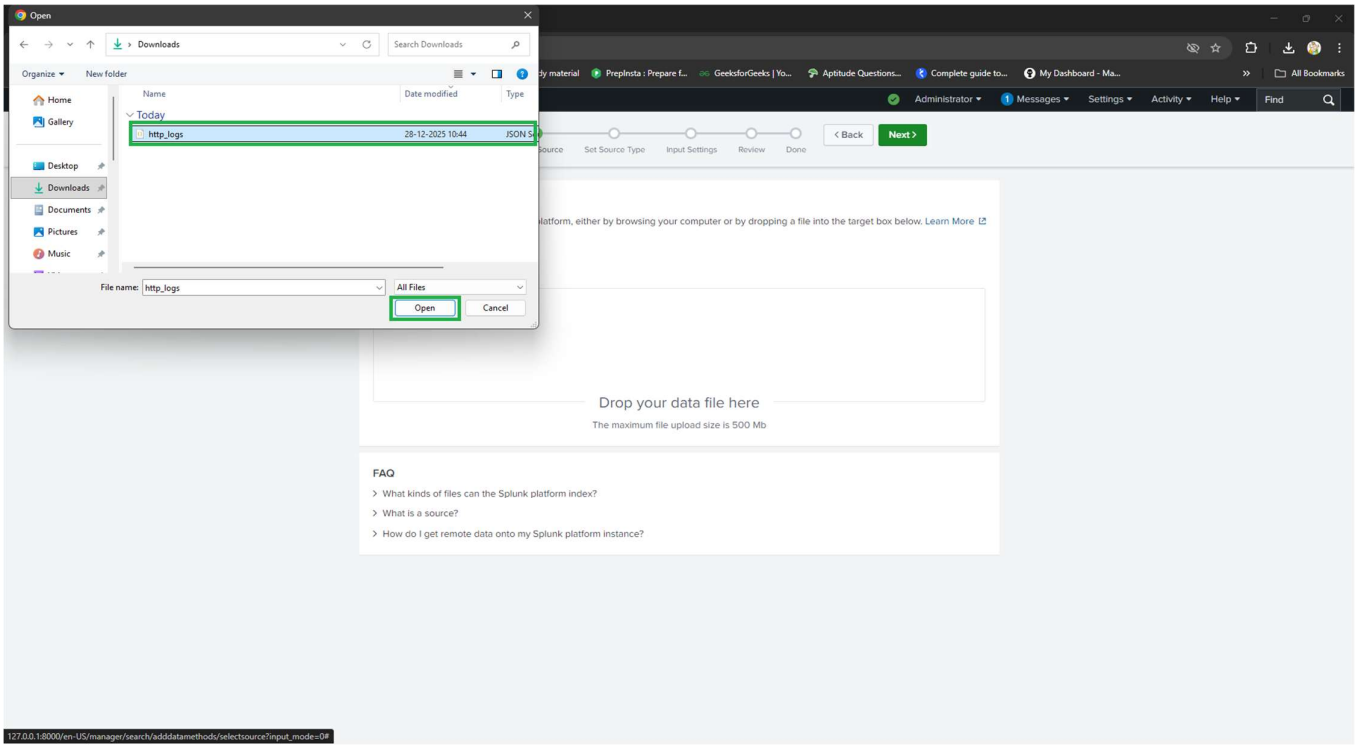


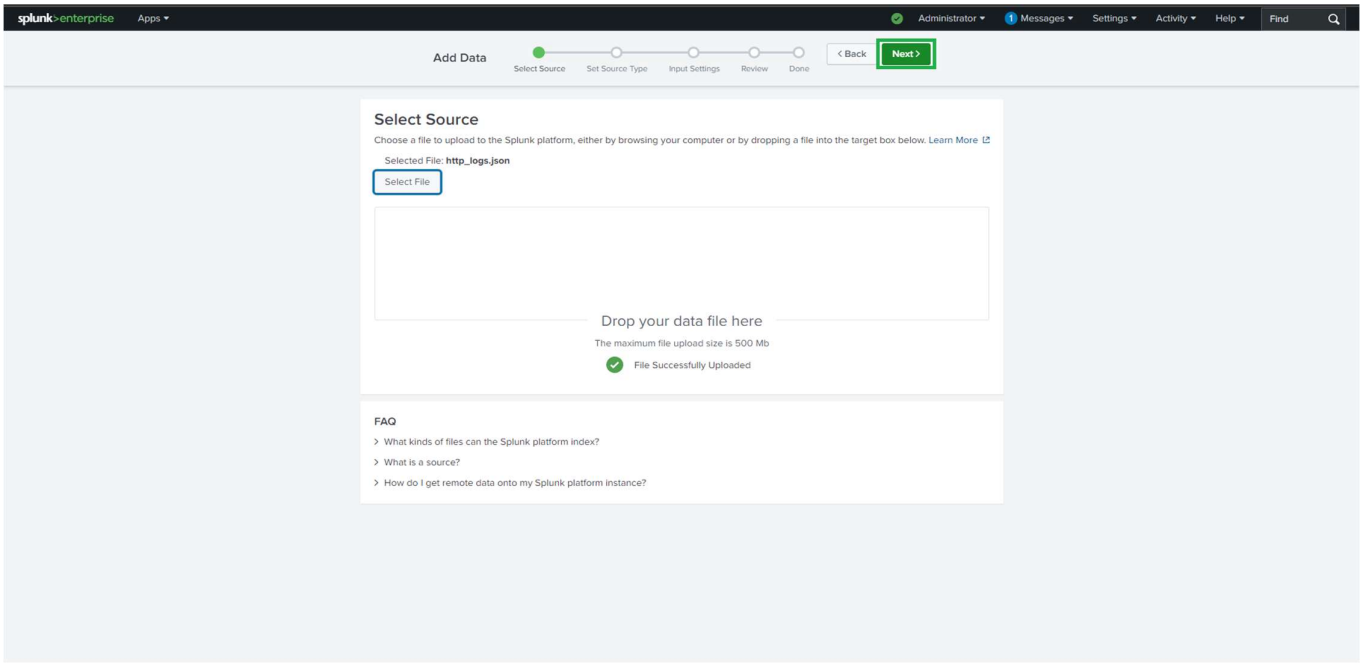
3. Select **Upload**



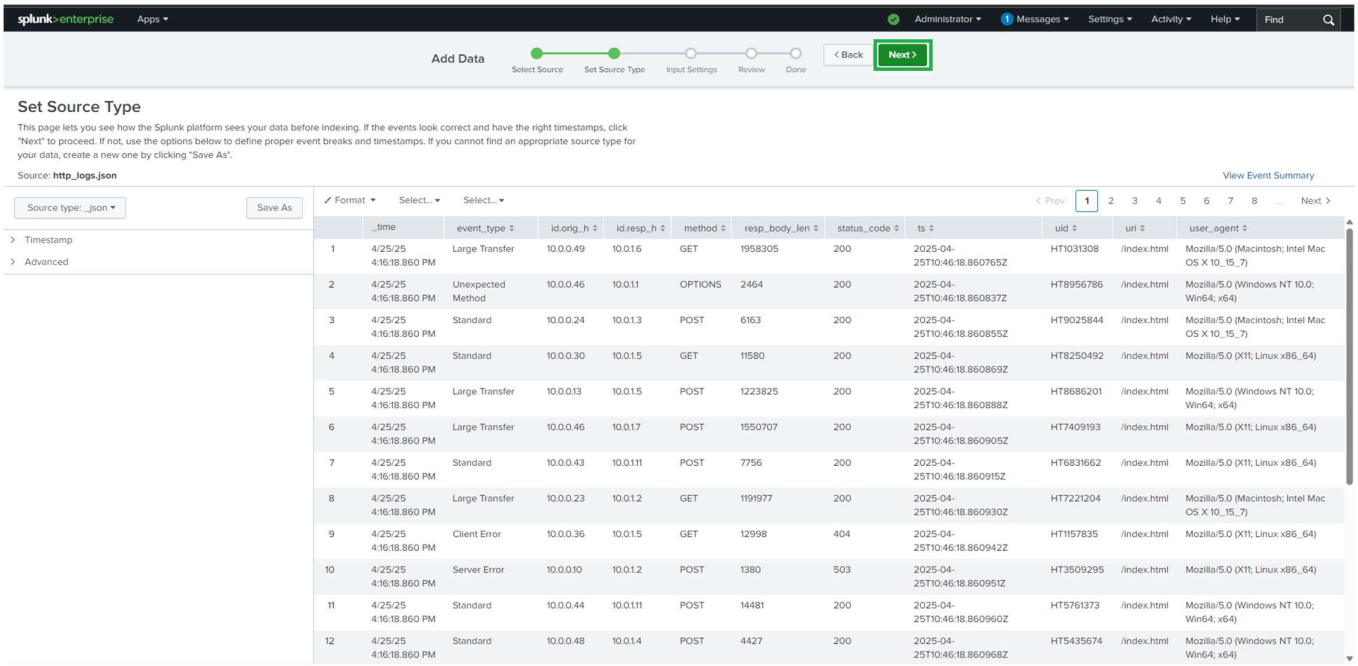


4. Upload http_logs.json

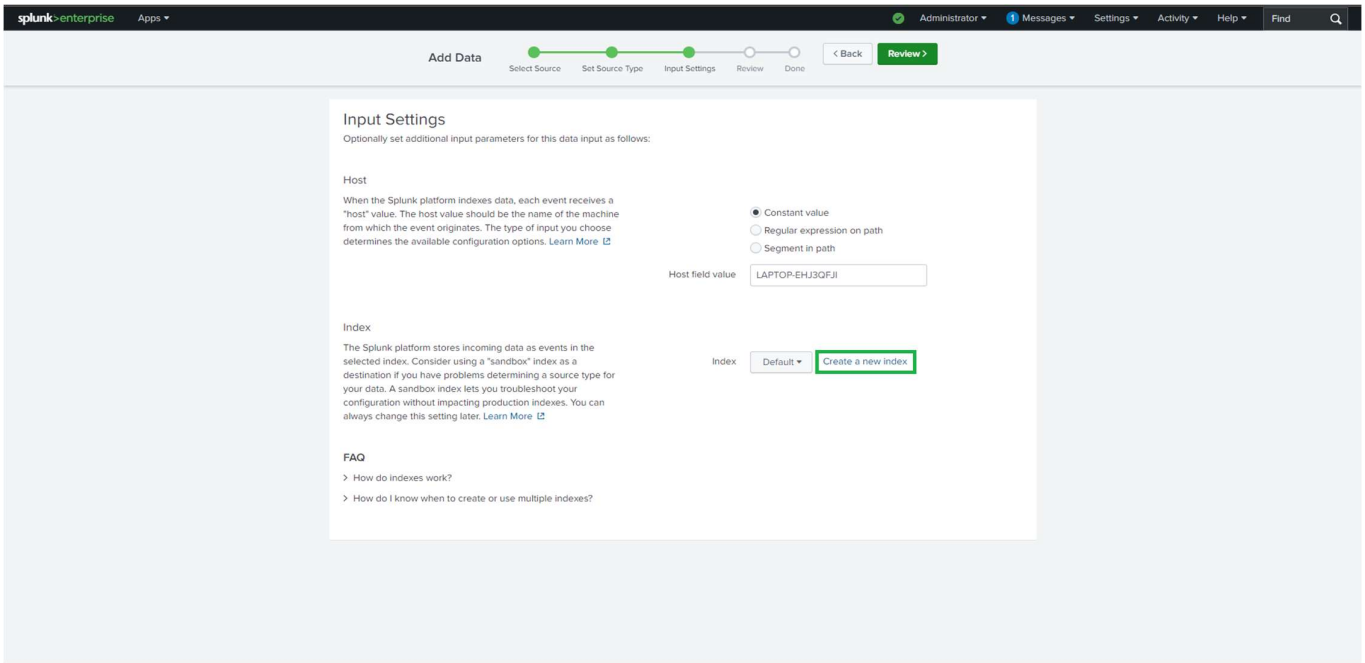


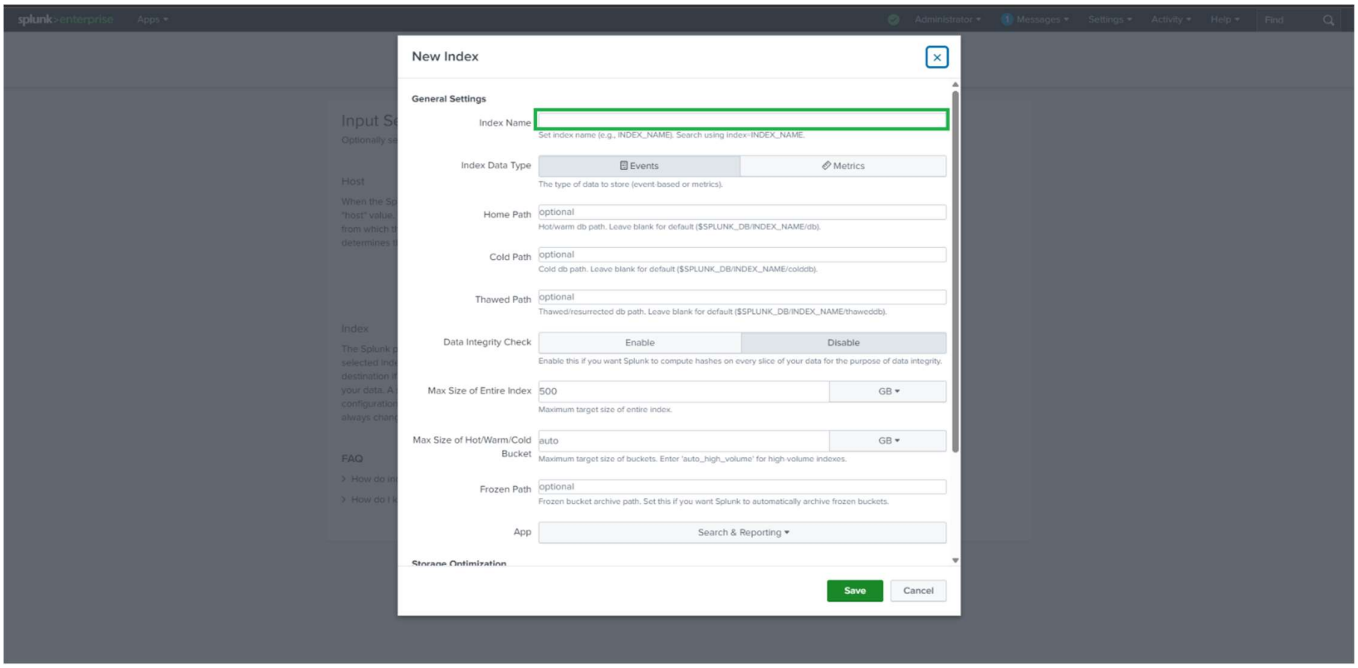


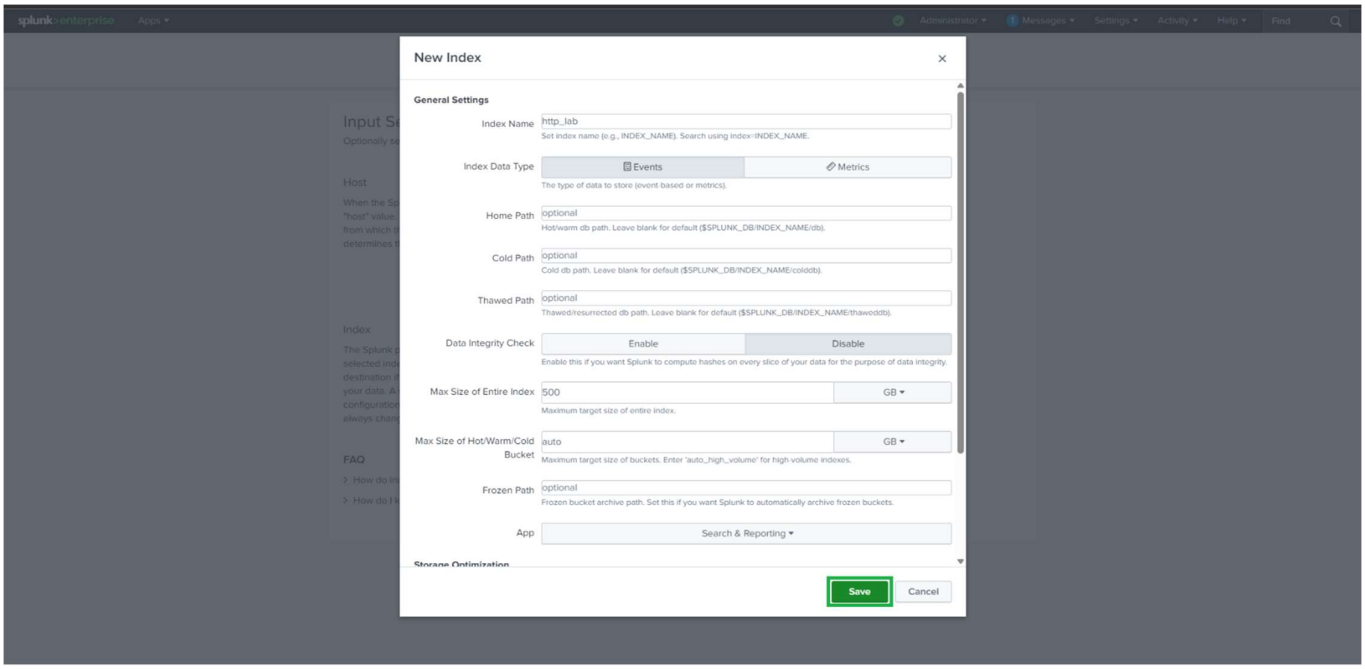
5. Configure:
    - **Source type:** json or create zeek:http
    - **Index:** http_lab (recommended)

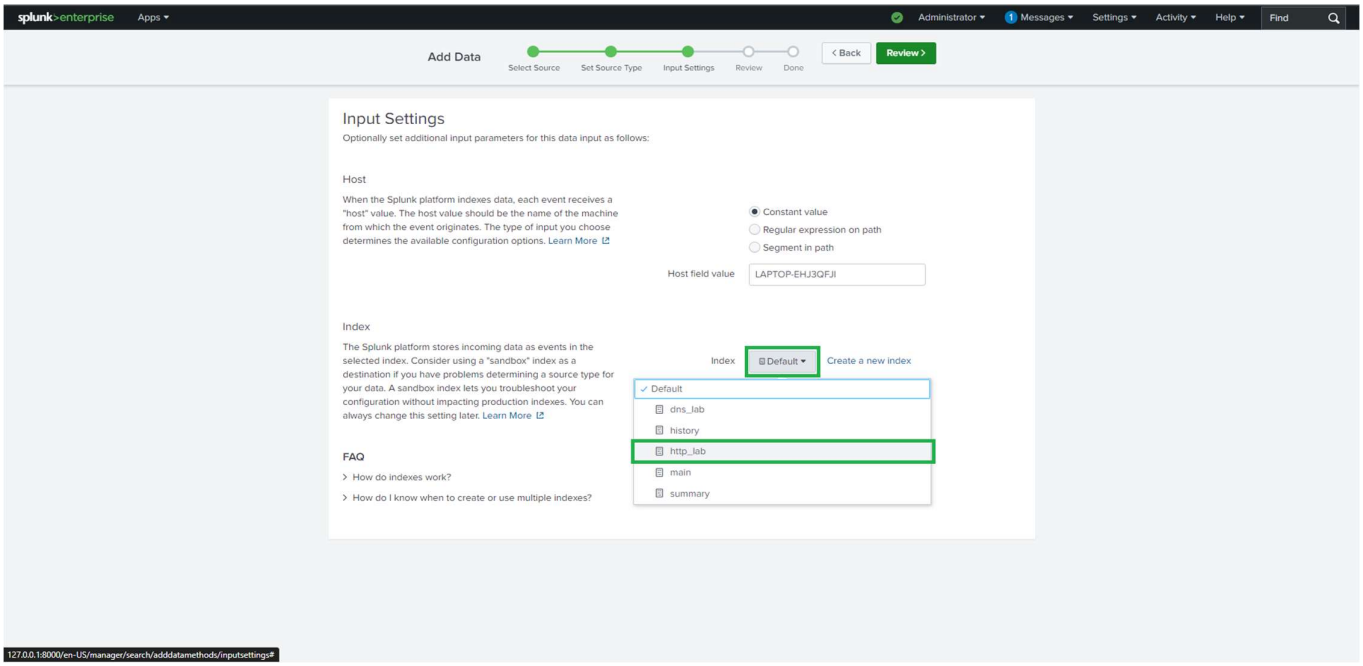


	_time	event_type	id.orig_h	id.resp_h	method	resp_body_len	status_code	ts	uid	uri	user_agent
1	4/25/25 4:16:18.860 PM	Large Transfer	10.0.0.49	10.0.1.6	GET	1958305	200	2025-04-25T10:46:18.860765Z	HT1031308	/index.html	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7)
2	4/25/25 4:16:18.860 PM	Unexpected Method	10.0.0.46	10.0.1.1	OPTIONS	2464	200	2025-04-25T10:46:18.860837Z	HT8956786	/index.html	Mozilla/5.0 (Windows NT 10.0; Win64; x64)
3	4/25/25 4:16:18.860 PM	Standard	10.0.0.24	10.0.1.3	POST	6163	200	2025-04-25T10:46:18.860855Z	HT9025844	/index.html	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7)
4	4/25/25 4:16:18.860 PM	Standard	10.0.0.30	10.0.1.5	GET	11580	200	2025-04-25T10:46:18.860869Z	HT8250492	/index.html	Mozilla/5.0 (X11; Linux x86_64)
5	4/25/25 4:16:18.860 PM	Large Transfer	10.0.0.13	10.0.1.5	POST	1223825	200	2025-04-25T10:46:18.860888Z	HT8686201	/index.html	Mozilla/5.0 (Windows NT 10.0; Win64; x64)
6	4/25/25 4:16:18.860 PM	Large Transfer	10.0.0.46	10.0.1.7	POST	1550707	200	2025-04-25T10:46:18.860905Z	HT7409193	/index.html	Mozilla/5.0 (X11; Linux x86_64)
7	4/25/25 4:16:18.860 PM	Standard	10.0.0.43	10.0.1.11	POST	7756	200	2025-04-25T10:46:18.860915Z	HT6831662	/index.html	Mozilla/5.0 (X11; Linux x86_64)
8	4/25/25 4:16:18.860 PM	Large Transfer	10.0.0.23	10.0.1.2	GET	1191977	200	2025-04-25T10:46:18.860930Z	HT7221204	/index.html	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7)
9	4/25/25 4:16:18.860 PM	Client Error	10.0.0.36	10.0.1.5	GET	12998	404	2025-04-25T10:46:18.860942Z	HT1157835	/index.html	Mozilla/5.0 (X11; Linux x86_64)
10	4/25/25 4:16:18.860 PM	Server Error	10.0.0.10	10.0.1.2	POST	1380	503	2025-04-25T10:46:18.860951Z	HT3509295	/index.html	Mozilla/5.0 (X11; Linux x86_64)
11	4/25/25 4:16:18.860 PM	Standard	10.0.0.44	10.0.1.11	POST	14481	200	2025-04-25T10:46:18.860960Z	HT5761373	/index.html	Mozilla/5.0 (Windows NT 10.0; Win64; x64)
12	4/25/25 4:16:18.860 PM	Standard	10.0.0.48	10.0.1.4	POST	4427	200	2025-04-25T10:46:18.860968Z	HT5435674	/index.html	Mozilla/5.0 (Windows NT 10.0; Win64; x64)

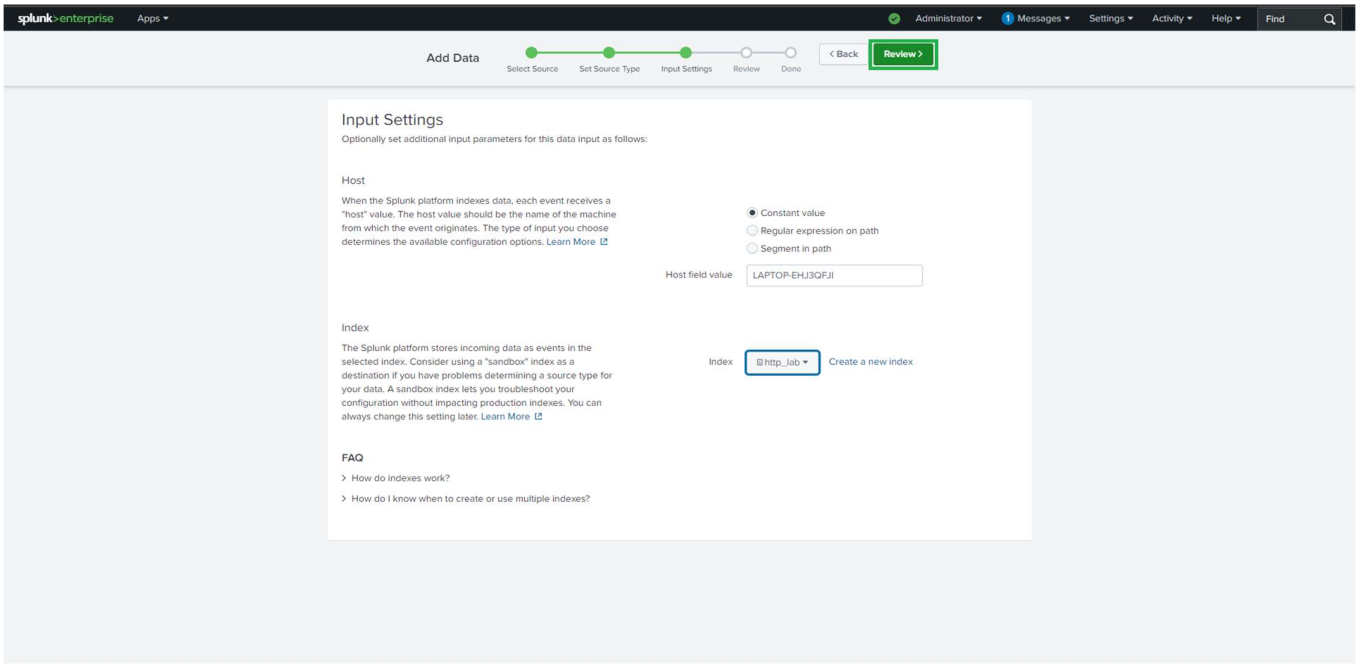



splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
New Index
General Settings
Index Name
Set index name (e.g., INDEX_NAME). Search using index=INDEX_NAME.
Index Data Type
Events
Metrics
The type of data to store (event-based or metrics).
Home Path
optional
Hot/warm db path. Leave blank for default ($SPLUNK_DB/INDEX_NAME/db).
Cold Path
optional
Cold db path. Leave blank for default ($SPLUNK_DB/INDEX_NAME/colddb).
Thawed Path
optional
Thawed/resurrected db path. Leave blank for default ($SPLUNK_DB/INDEX_NAME/thaweddb).
Data Integrity Check
Enable
Disable
Enable this if you want Splunk to compute hashes on every slice of your data for the purpose of data integrity.
Max Size of Entire Index
500
GB
Maximum target size of entire index.
Max Size of Hot/Warm/Cold Bucket
auto
GB
Maximum target size of buckets. Enter 'auto_high_volume' for high-volume indexes.
Frozen Path
optional
Frozen bucket archive path. Set this if you want Splunk to automatically archive frozen buckets.
App
Search & Reporting
Storage Optimization
Save
Cancel

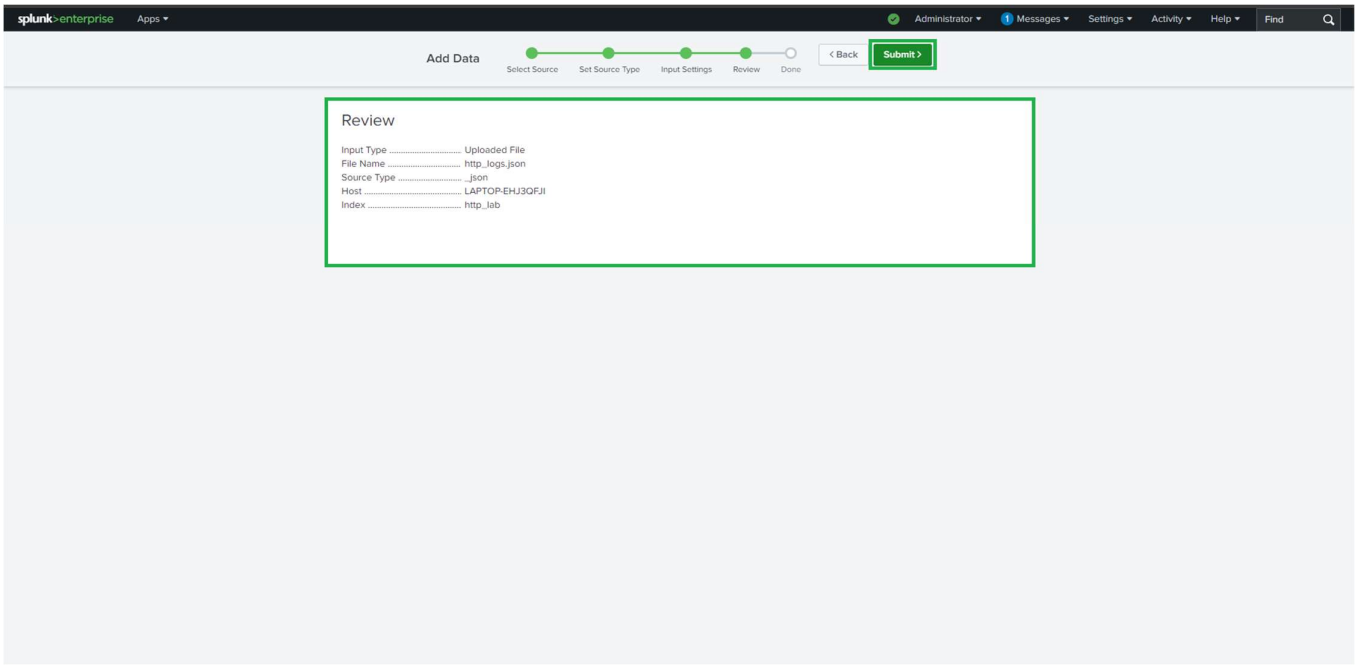

splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
New Index
General Settings
Index Name
http_lab
Set index name (e.g., INDEX_NAME). Search using index=INDEX_NAME.
Index Data Type
Events
Metrics
The type of data to store (event-based or metrics).
Home Path
optional
Hot/warm db path. Leave blank for default ($SPLUNK_DB/INDEX_NAME/db).
Cold Path
optional
Cold db path. Leave blank for default ($SPLUNK_DB/INDEX_NAME/colddb).
Thawed Path
optional
Thawed/resurrected db path. Leave blank for default ($SPLUNK_DB/INDEX_NAME/thaweddb).
Data Integrity Check
Enable
Disable
Enable this if you want Splunk to compute hashes on every slice of your data for the purpose of data integrity.
Max Size of Entire Index
500
GB
Maximum target size of entire index.
Max Size of Hot/Warm/Cold Bucket
auto
GB
Maximum target size of buckets. Enter 'auto_high_volume' for high-volume indexes.
Frozen Path
optional
Frozen bucket archive path. Set this if you want Splunk to automatically archive frozen buckets.
App
Search & Reporting
Storage Optimization
Save
Cancel

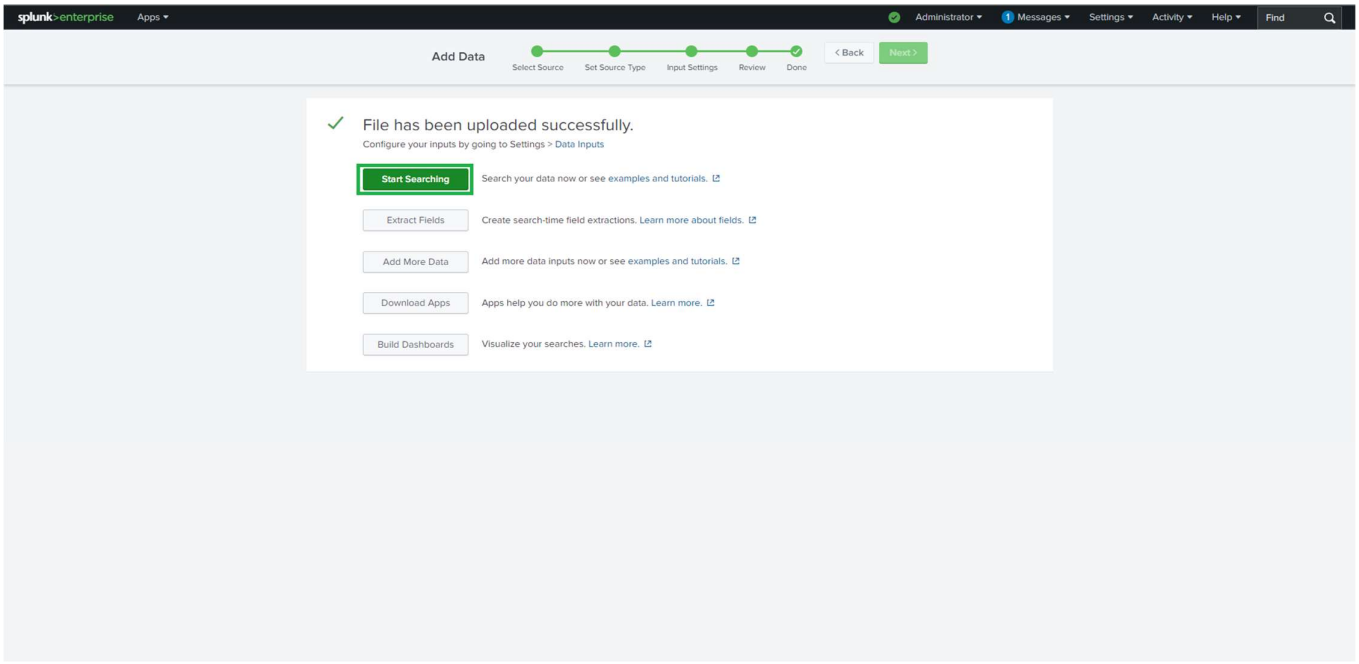


127.0.0.1:8000/en-US/manager/search/adddatamethods/inputsettings#



6. Finish upload and verify indexing



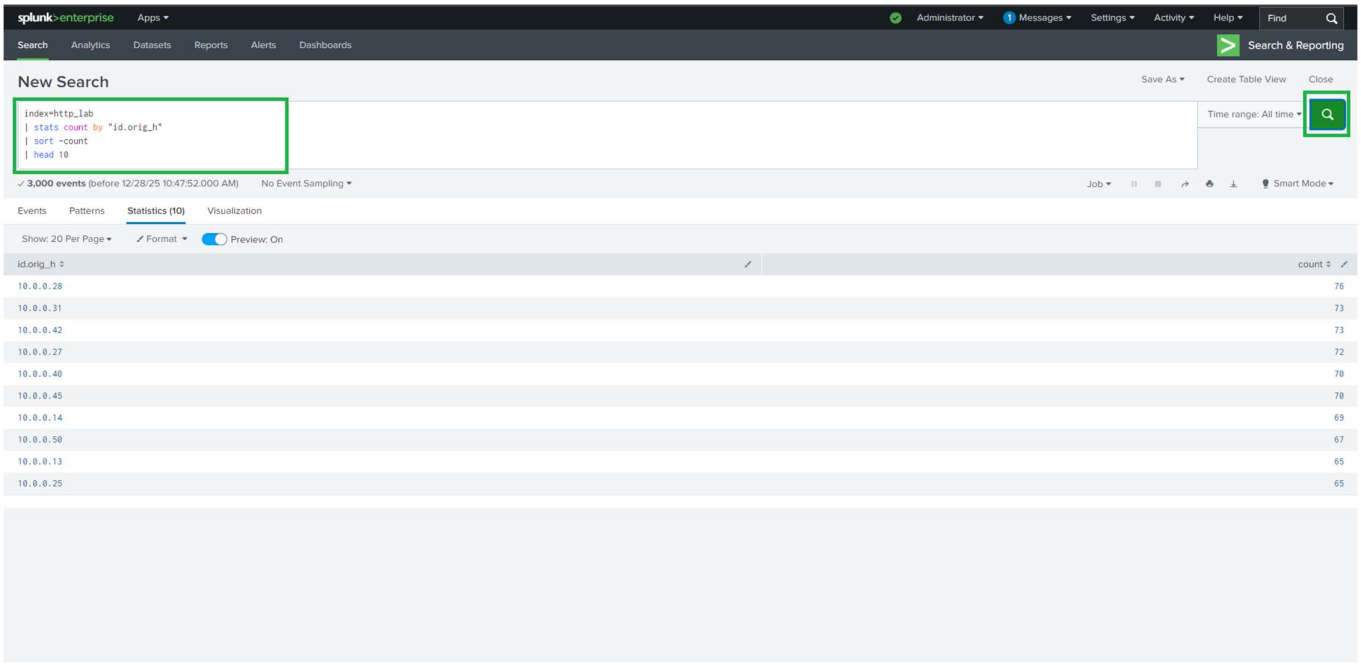


🔍 Lab Tasks & SPL Queries

🔹 Task 1: Top 10 Endpoints Generating Web Traffic

```
index=http_lab
| stats count by "id.orig_h"
| sort -count
| head 10
```

📌 *Identifies the most active source IPs generating HTTP requests.*

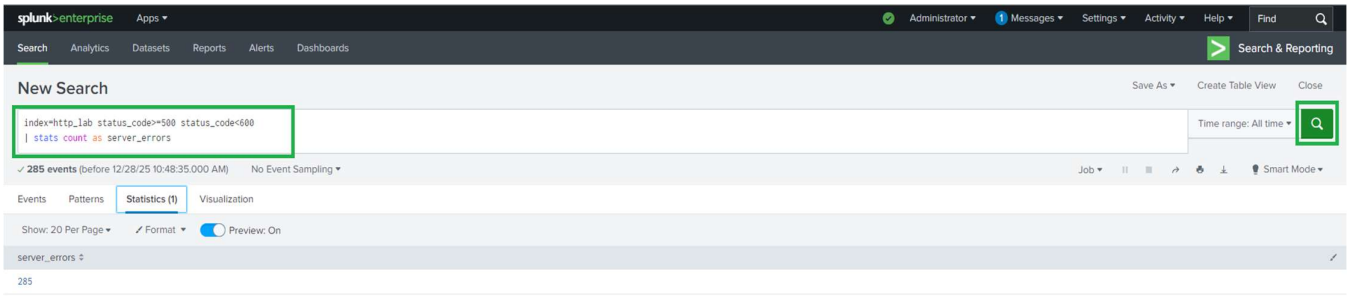


🔹 Task 2: Count Server Errors (HTTP 5xx)

index=http_lab status_code>=500 status_code<600

| stats count as server_errors

📌 *Helps detect backend failures or application crashes.*



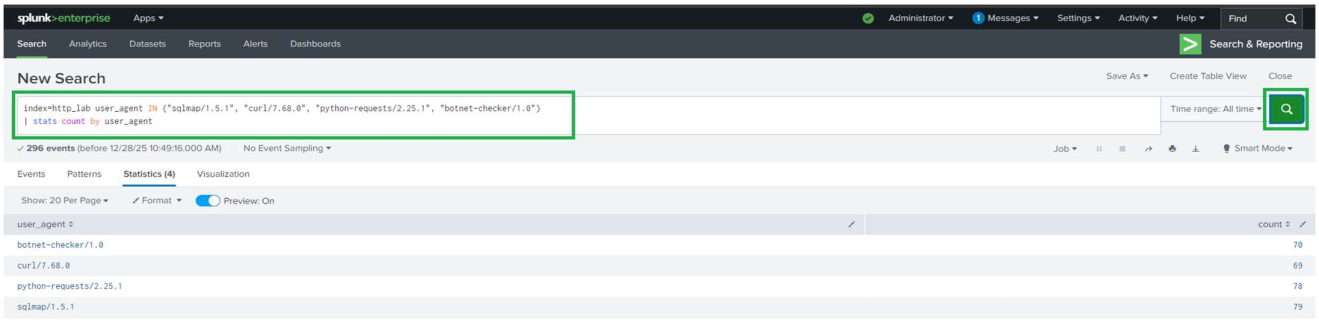
🔹 Task 3: Detect Suspicious / Scripted User-Agents

index=http_lab

user_agent IN ("sqlmap/1.5.1", "curl/7.68.0", "python-requests/2.25.1", "botnet-checker/1.0")

| stats count by user_agent

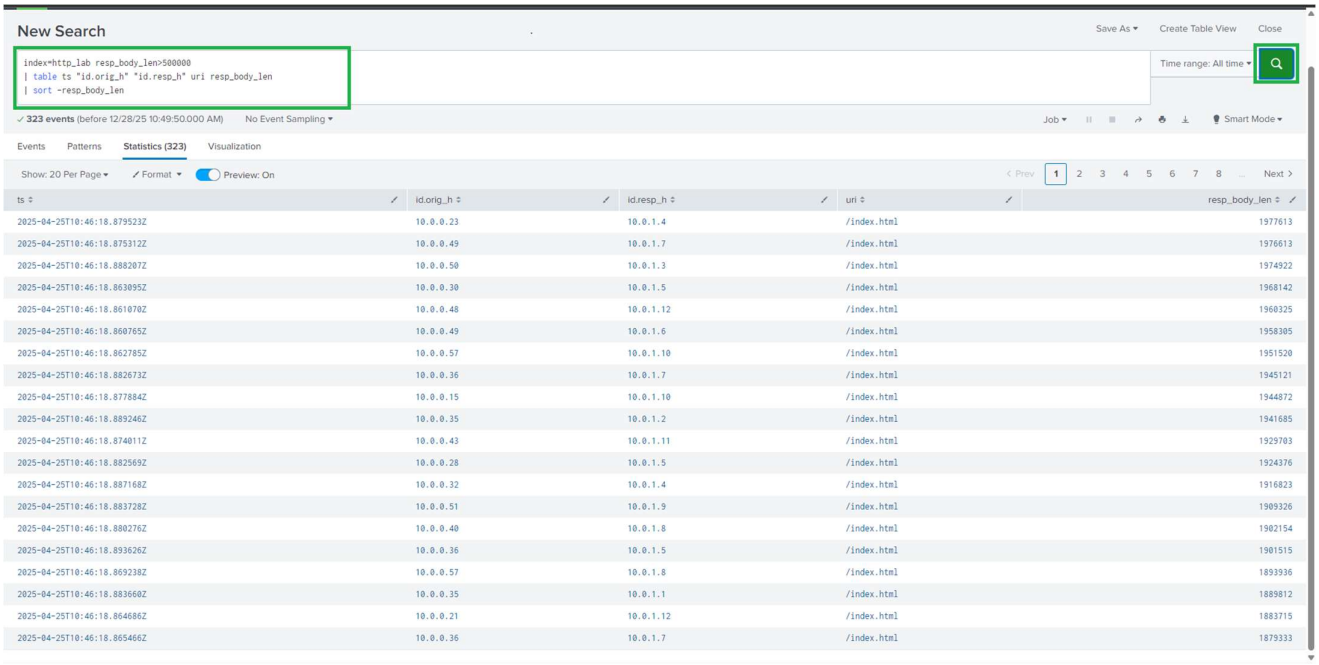
📌 *Identifies automated tools commonly used in attacks or scanning.*



🔹 Task 4: Find Large File Transfers (> 500 KB)

```
index=http_lab resp_body_len>500000
| table ts "id.orig_h" "id.resp_h" uri resp_body_len
| sort -resp_body_len
```

📌 *Useful for spotting potential data leaks or suspicious downloads.*



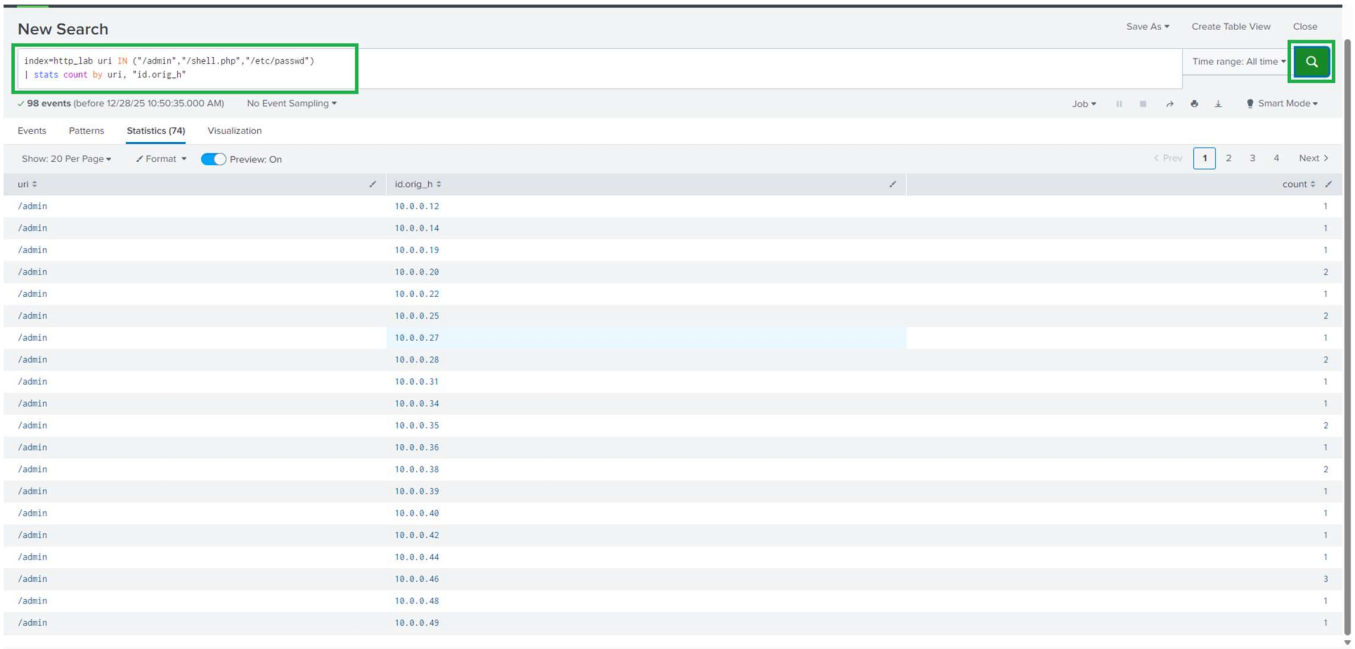
🔹 Task 5: Detect Suspicious URI Access Attempts

index=http_lab

uri IN ("/admin","/shell.php","/etc/passwd")

| stats count by uri, "id.orig_h"

📌 *Highlights attempts to access admin panels, shells, or sensitive files.*



🚨 Security Insights Gained

Through this lab, you can:

- Detect **malicious web reconnaissance**
- Identify **brute-force or scanning tools**
- Monitor **abnormal HTTP behavior**
- Recognize **potential data exfiltration**
- Build detection logic for **SOC use cases**

📊 Possible Enhancements

- ✅ Create **Splunk alerts** for:
  - Repeated 5xx errors
  - Large file transfers
  - Suspicious User-Agents
- 📈 Build dashboards for:
  - HTTP status trends
  - Top URIs and IPs
- 🔐 Integrate with IDS / threat intelligence feeds

🧠 Learning Outcome

By completing this project, you have:

- Gained hands-on experience with **Splunk SPL**
- Learned **real-world HTTP log analysis**
- Practiced **threat detection techniques**
- Strengthened **SIEM & SOC analyst skills**

⭐ Conclusion

This lab demonstrates how **SIEM tools like Splunk** can be effectively used to analyze web traffic, detect attacks, and improve organizational security posture using log data.

🔥 A must-have project for **Cybersecurity**, **SOC Analyst**, and **Blue Team** portfolios.
