📊 DNS Log Analysis Using Splunk Project

By Gaurav Ghandat

📌 Project Overview

This project demonstrates how to **ingest, parse, and analyze DNS logs** using **Splunk Enterprise**.
By leveraging **Zeek-style JSON DNS logs**, we perform meaningful security and traffic analysis using **Splunk Search Processing Language (SPL)**.

The lab focuses on identifying:

- Frequently queried domain names
- Most active client IPs
- Distribution of DNS record types
- Potential anomalies in DNS behavior

🎯 Objectives

By completing this project, you will learn how to:

- Ingest JSON-formatted DNS logs into Splunk
- Extract useful DNS metadata (queries, IPs, record types)
- Write SPL queries for DNS traffic analysis
- Detect suspicious or abnormal DNS activity
- Build a strong foundation for DNS-based threat detection

🛠️ Tools & Technologies

- **Splunk Enterprise**
- **Zeek DNS Logs (JSON format)**
- **Search Processing Language (SPL)**

📁 Dataset Information

The dataset consists of **Zeek-style DNS logs** in JSON format with fields such as:

ts → Timestamp

id.orig_h → Source IP address

id.resp_h → DNS server IP

qtype → DNS query type (A, AAAA, PTR, CNAME)

query → Queried domain name

answers → DNS response

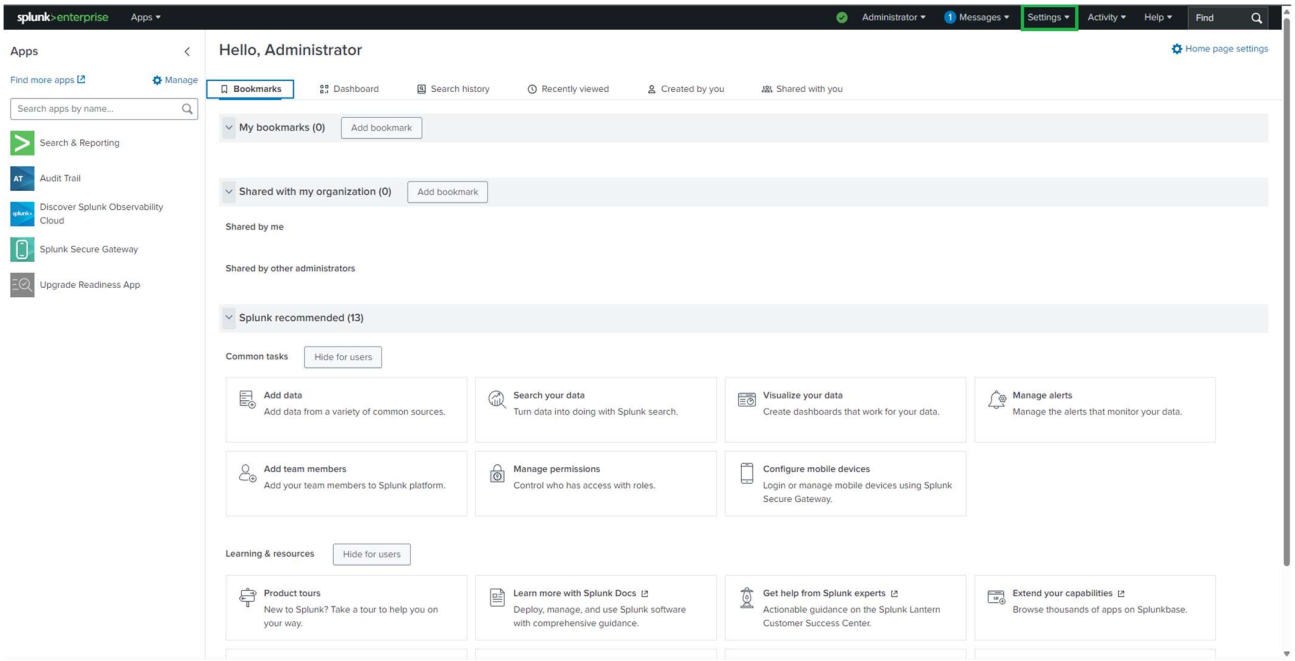
rcode → DNS response code

rtt → Round Trip Time

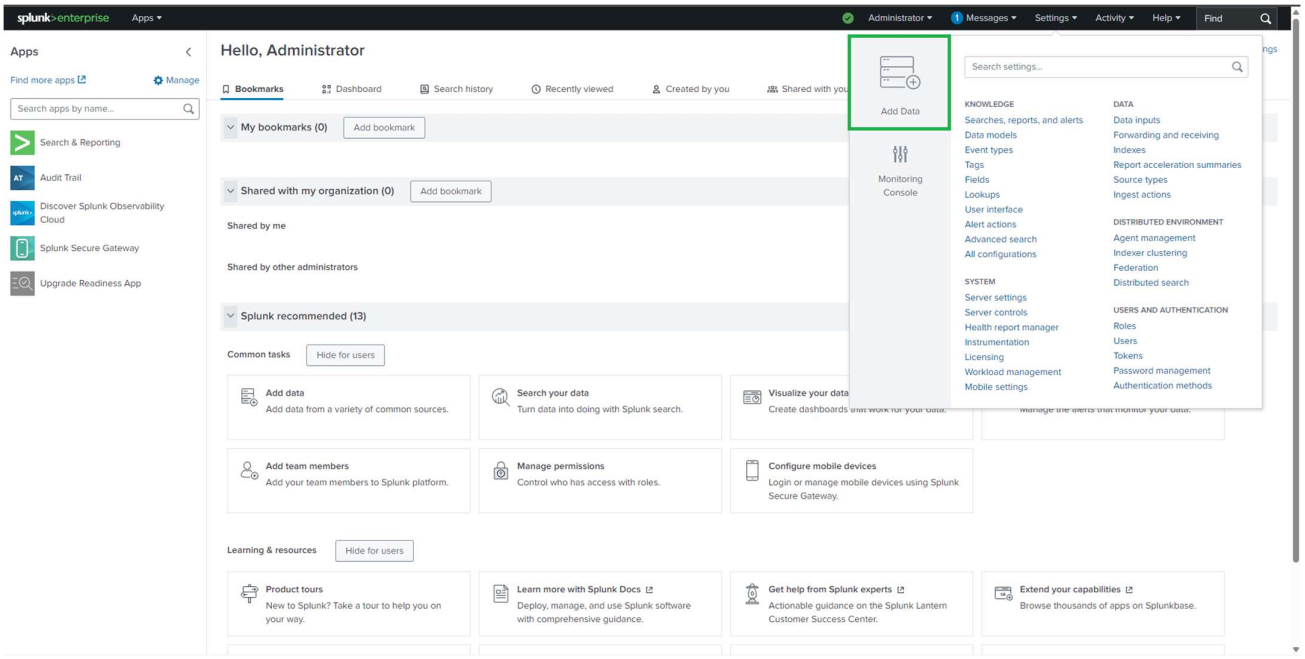
⚙️ Lab Setup & Data Ingestion

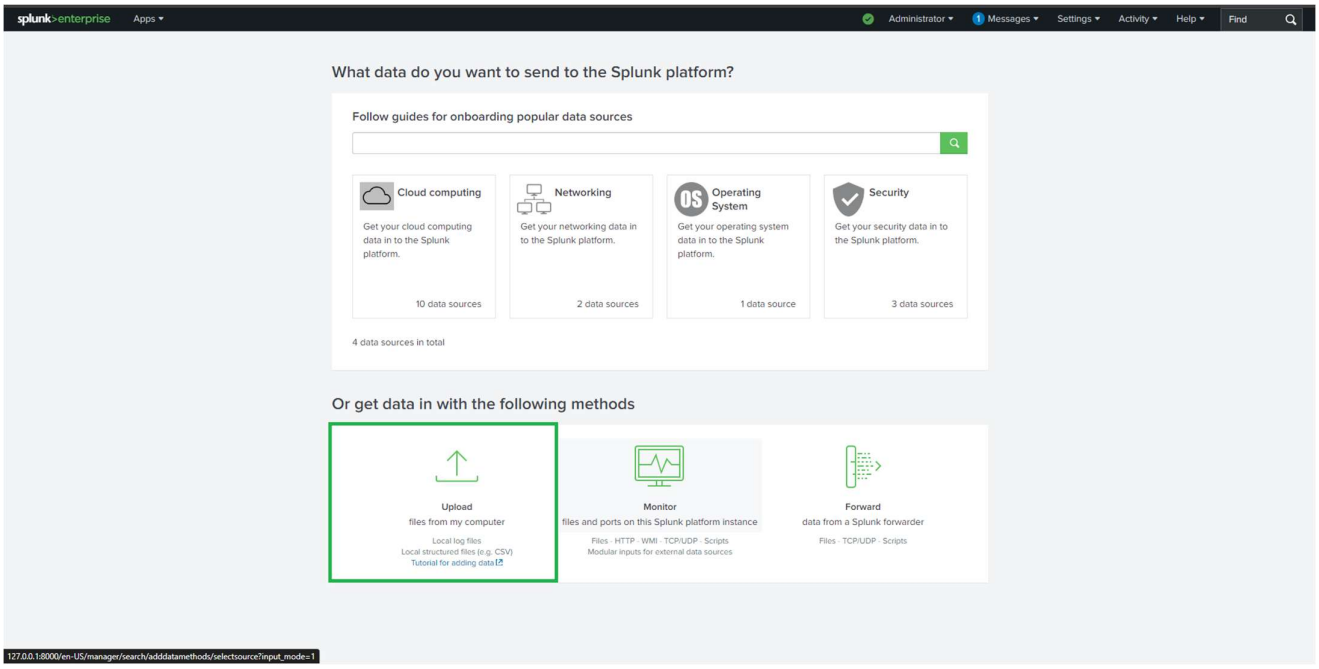
Step 1: Upload DNS Logs

1. Open **Splunk Web**

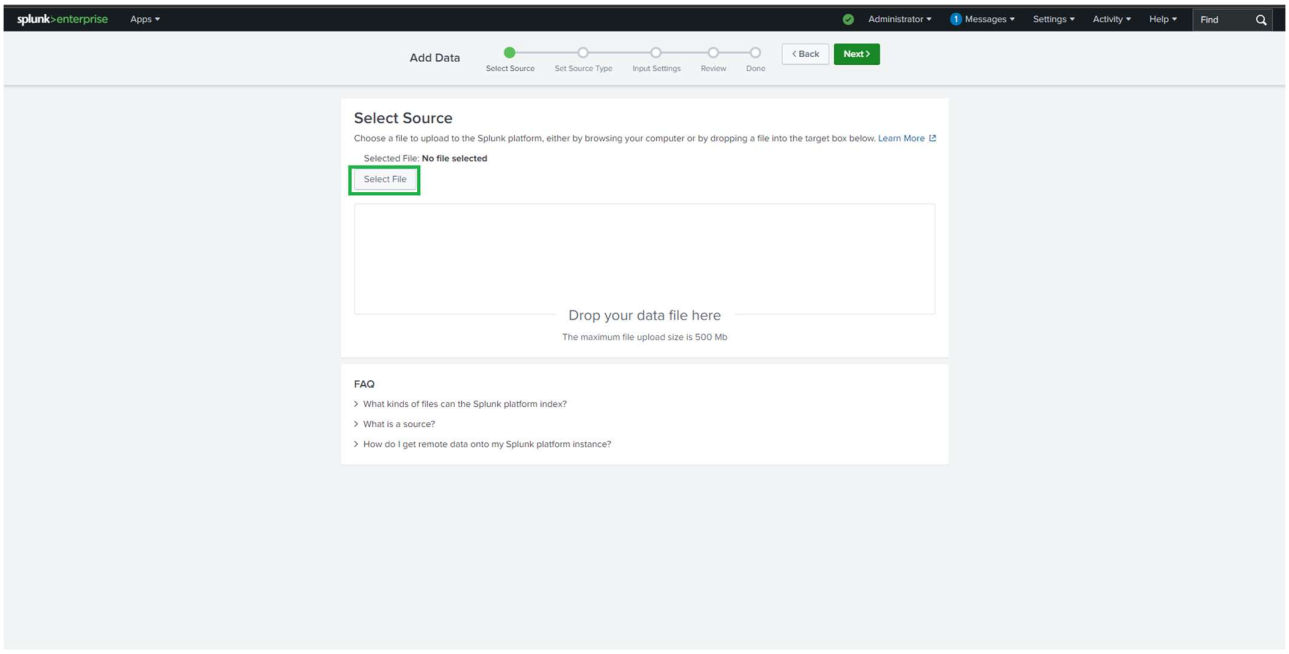


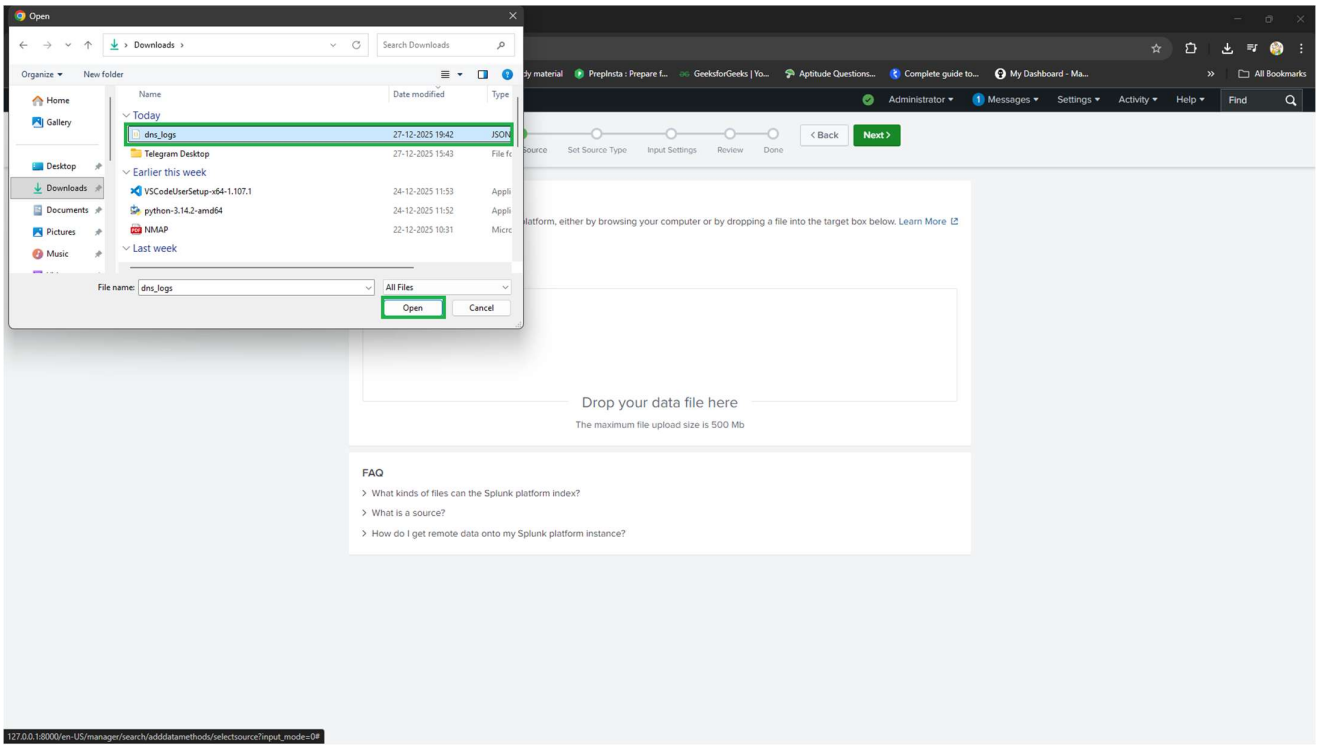
2. Navigate to:
   Settings → Add Data → Upload



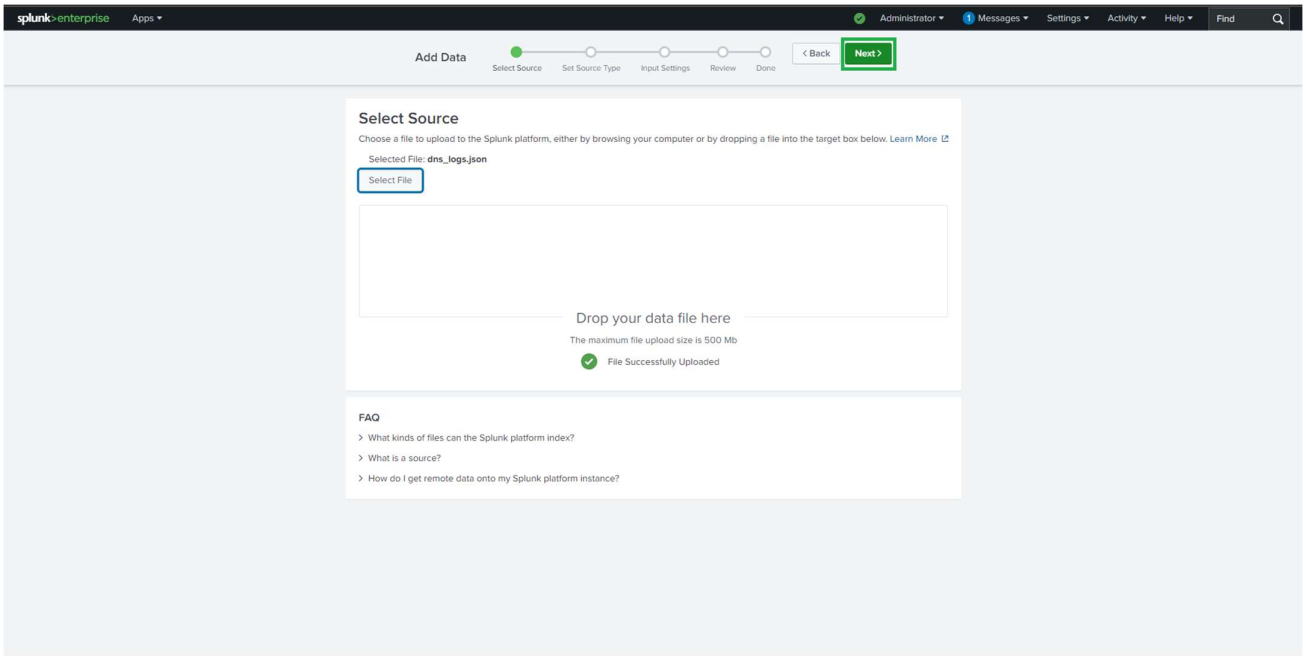


3. Select the file:

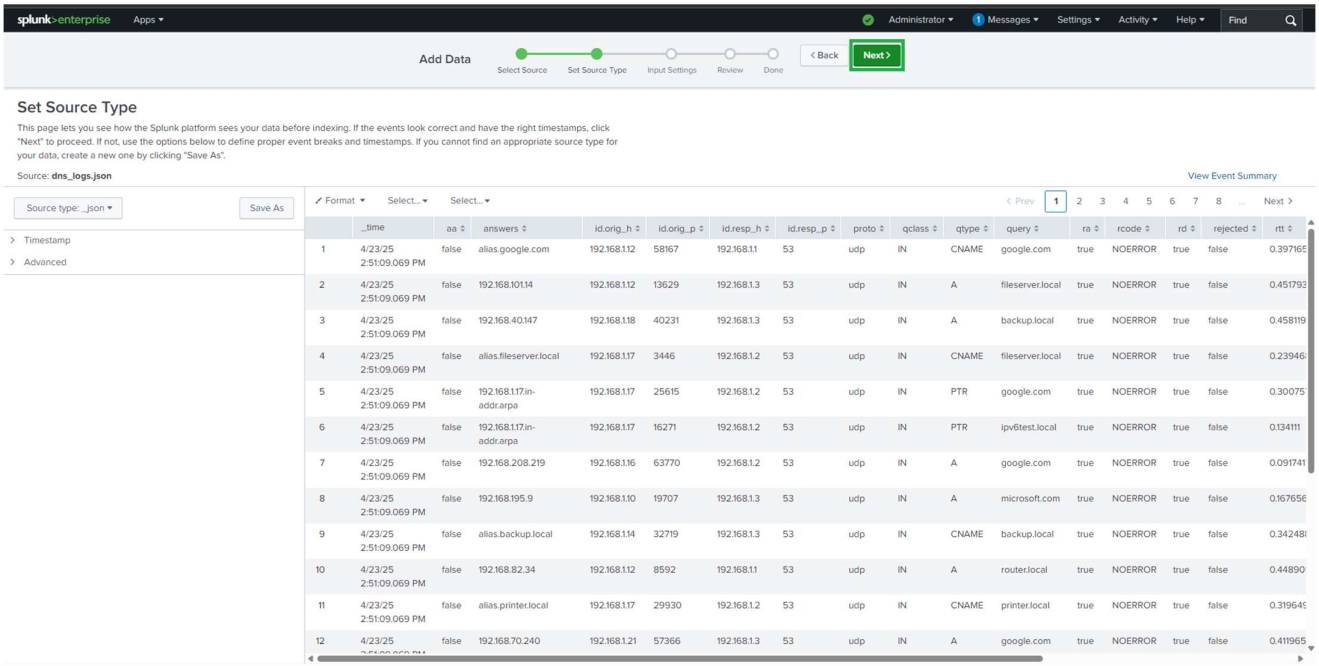


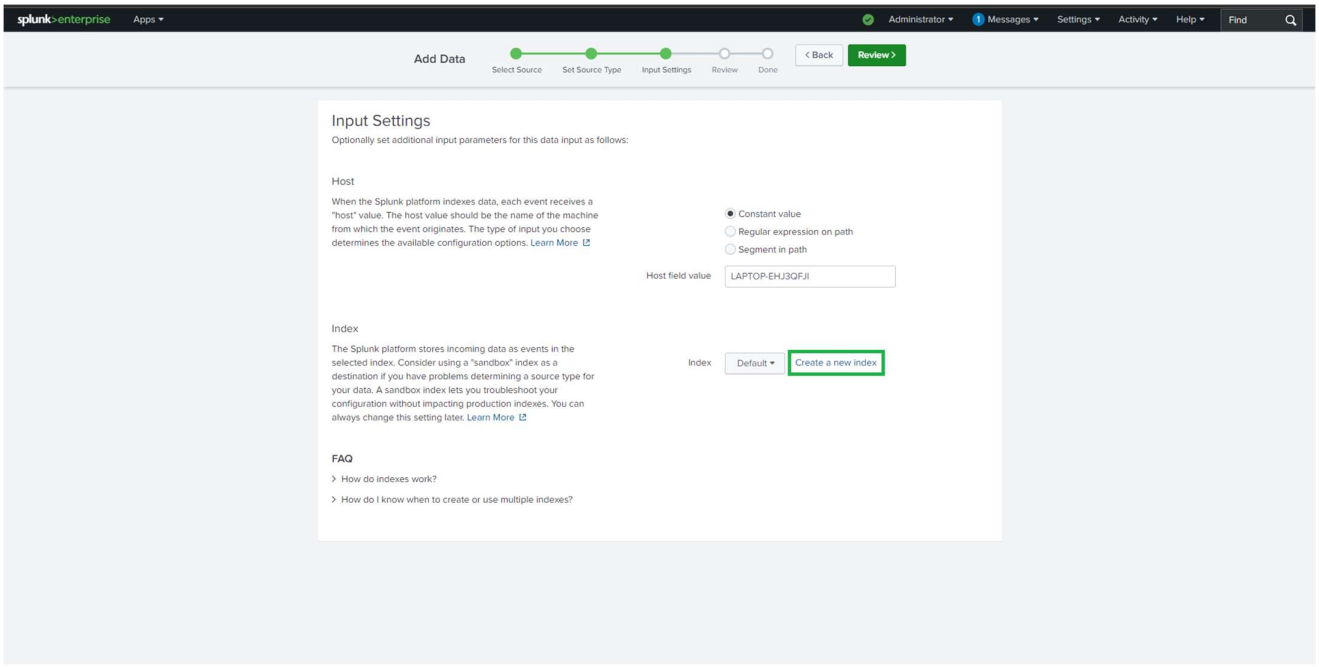


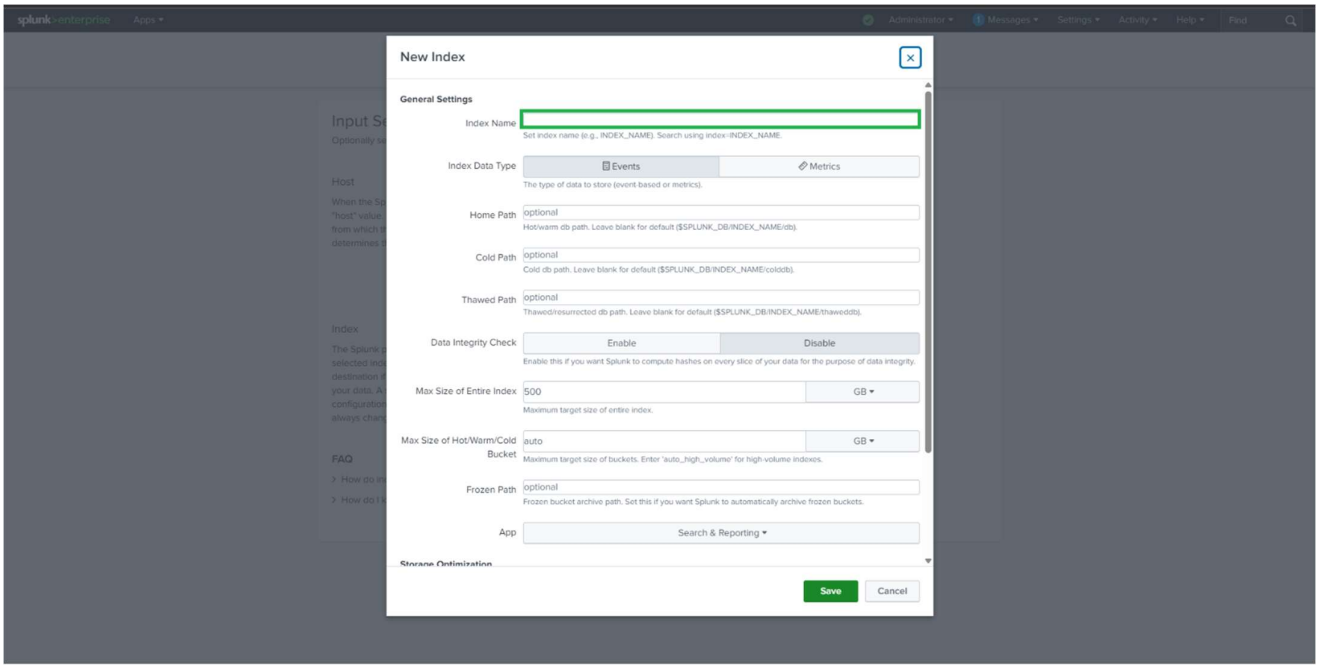
4. dns_logs.json

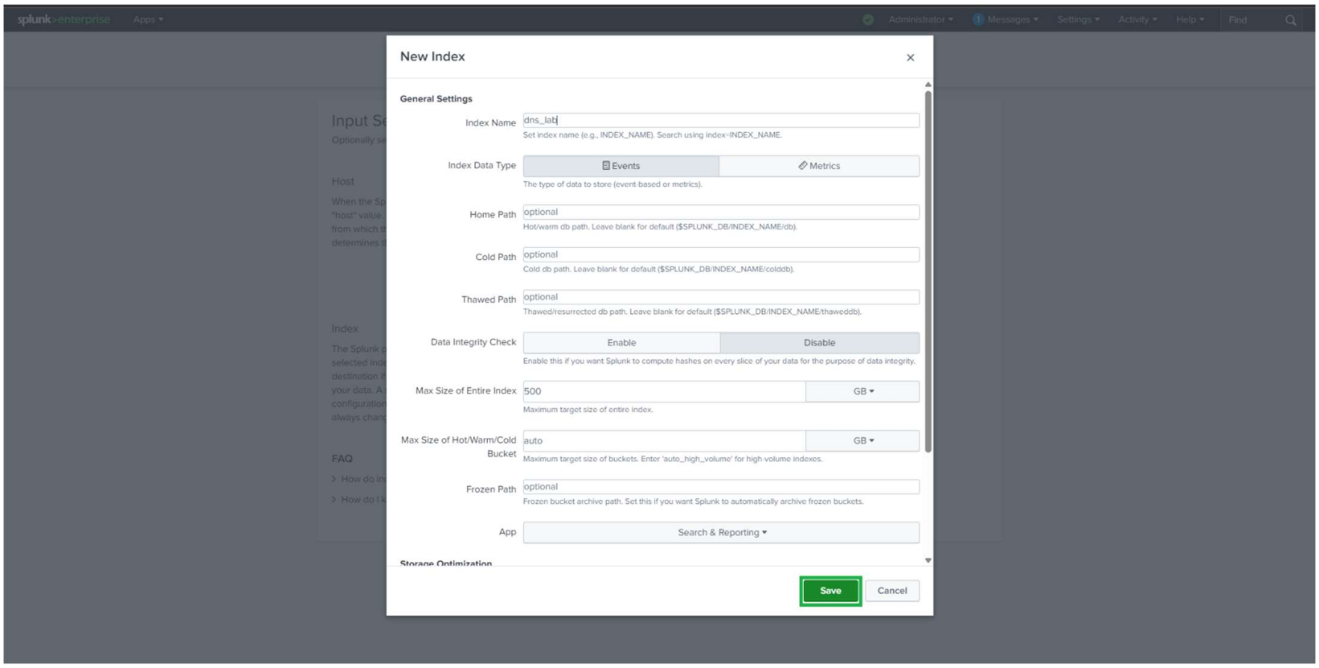


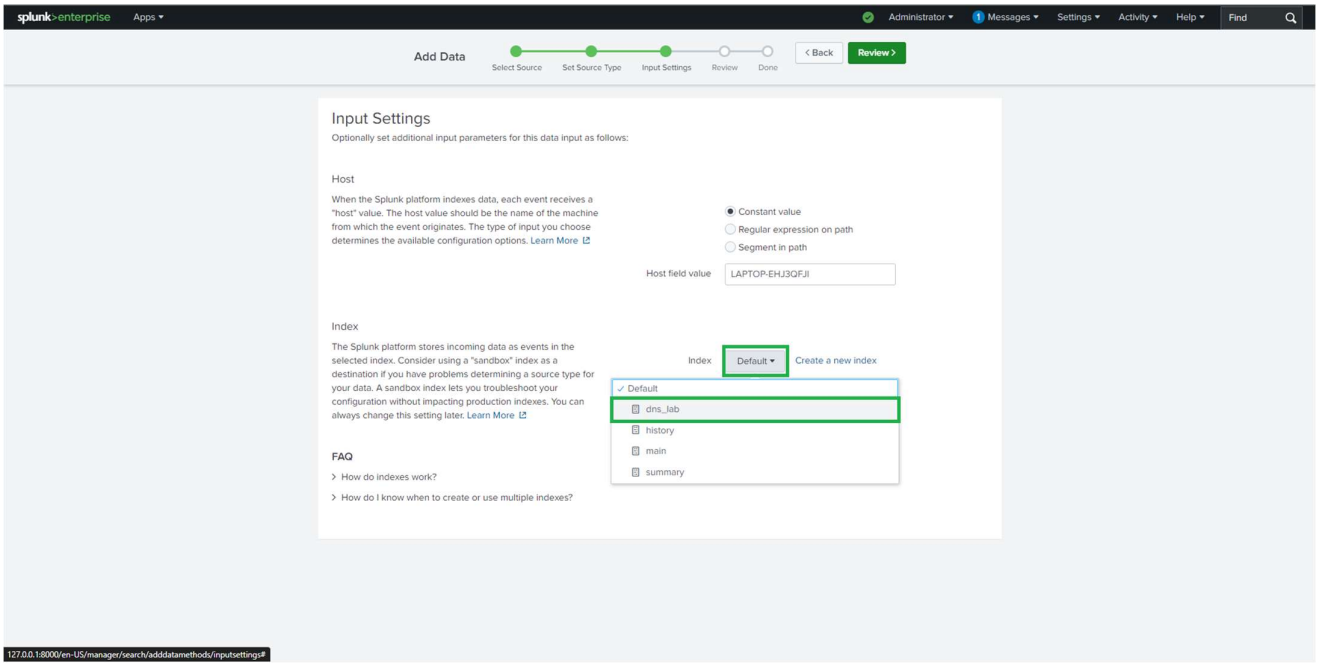
5. Set the following:
    - **Source Type:** json (or custom zeek:dns)
    - **Index:** dns_lab (recommended)

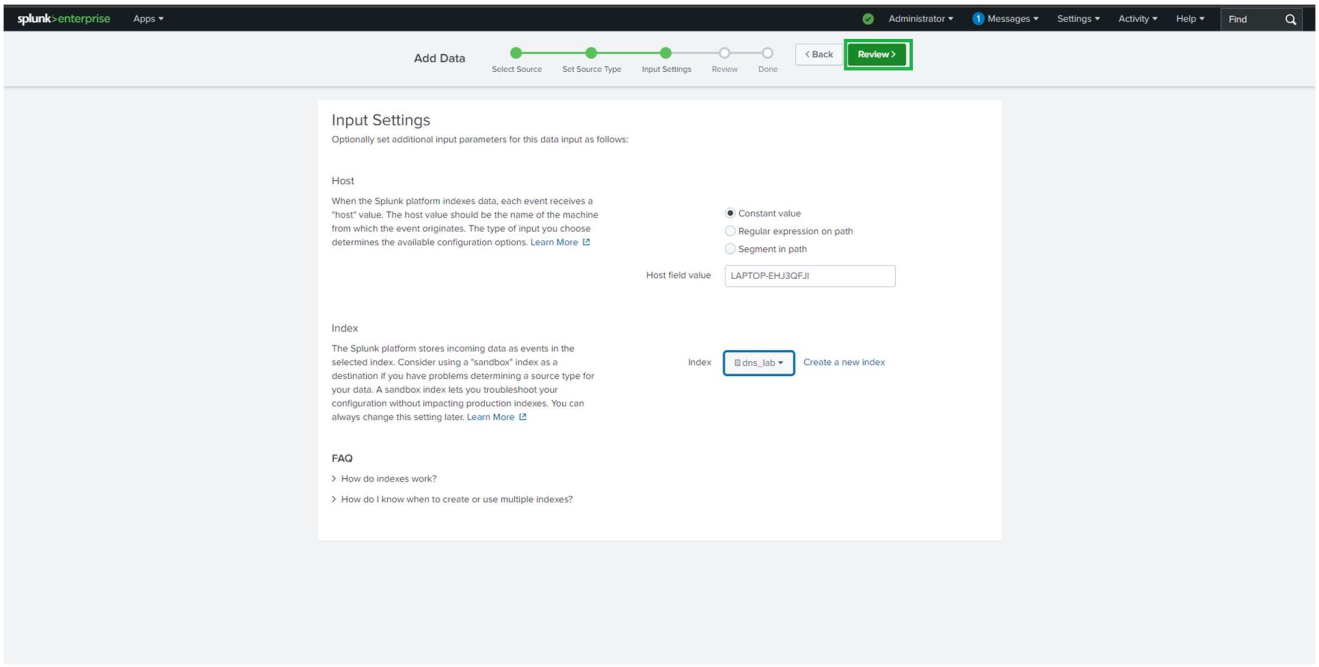


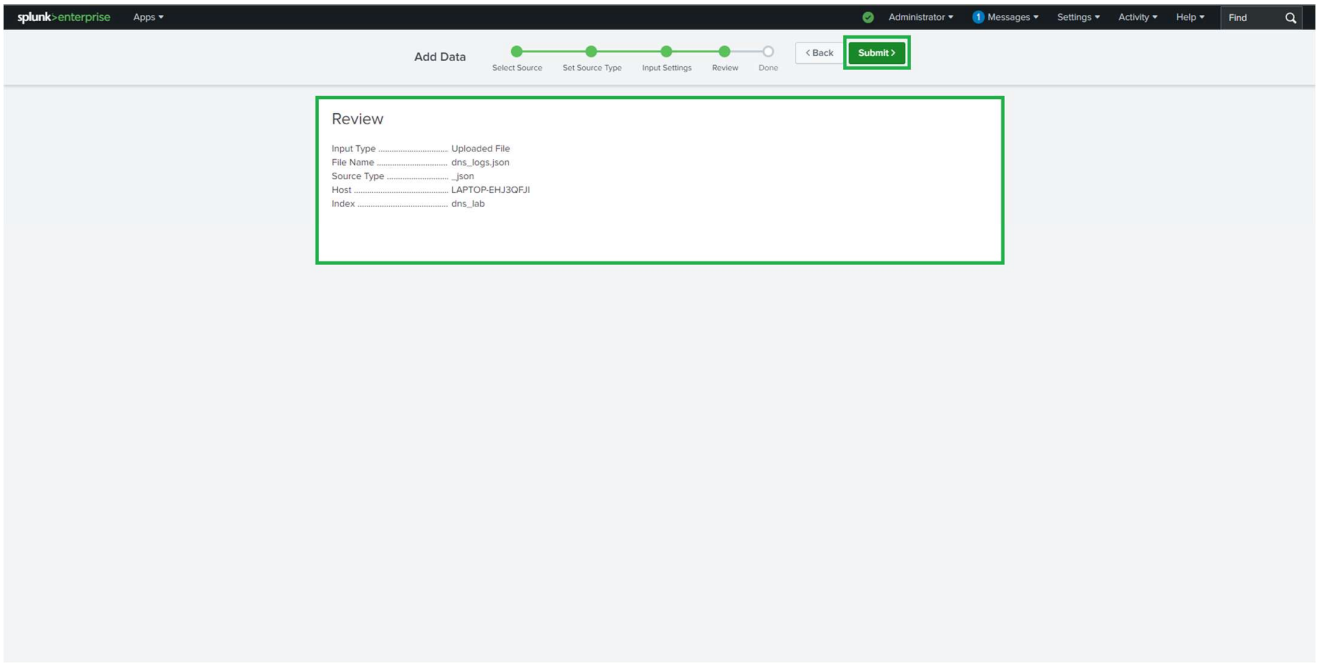

splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
Add Data
Select Source
Set Source Type
Input Settings
Review
Done
< Back
Review >
Input Settings
Optionally set additional input parameters for this data input as follows:
Host
When the Splunk platform indexes data, each event receives a "host" value. The host value should be the name of the machine from which the event originates. The type of input you choose determines the available configuration options. Learn More
Constant value
Regular expression on path
Segment in path
Host field value
LAPTOP-EHJ3QFJI
Index
The Splunk platform stores incoming data as events in the selected index. Consider using a "sandbox" index as a destination if you have problems determining a source type for your data. A sandbox index lets you troubleshoot your configuration without impacting production indexes. You can always change this setting later. Learn More
Index
Default
Create a new index
FAQ
How do indexes work?
How do I know when to create or use multiple indexes?

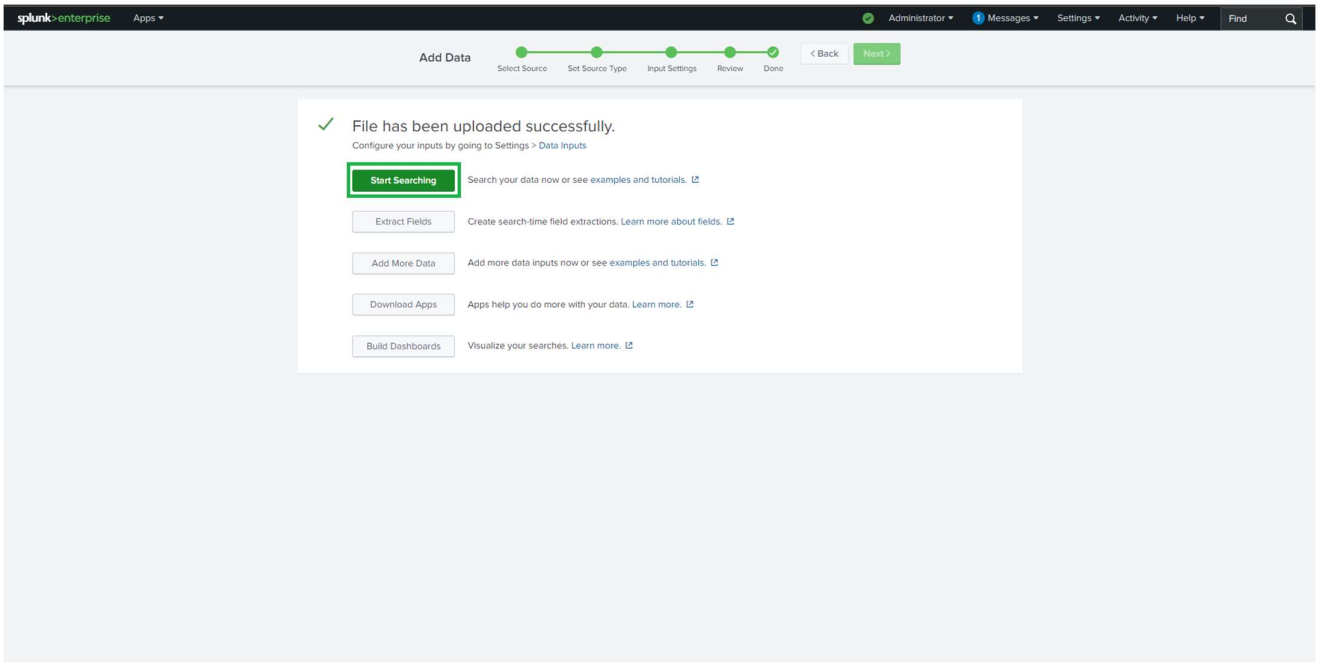

New Index
General Settings
Index Name
Set index name (e.g., INDEX_NAME). Search using index=INDEX_NAME.
Index Data Type
Events
Metrics
The type of data to store (event-based or metrics).
Home Path
optional
Hot/warm db path. Leave blank for default ($SPLUNK_DB/INDEX_NAME/db).
Cold Path
optional
Cold db path. Leave blank for default ($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
dns_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


splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
Add Data
Select Source
Set Source Type
Input Settings
Review
Done
< Back
Review >
Input Settings
Optionally set additional input parameters for this data input as follows:
Host
When the Splunk platform indexes data, each event receives a "host" value. The host value should be the name of the machine from which the event originates. The type of input you choose determines the available configuration options. Learn More
Constant value
Regular expression on path
Segment in path
Host field value
LAPTOP-EHJ3QFJI
Index
The Splunk platform stores incoming data as events in the selected index. Consider using a "sandbox" index as a destination if you have problems determining a source type for your data. A sandbox index lets you troubleshoot your configuration without impacting production indexes. You can always change this setting later. Learn More
Index
Default
Create a new index
Default
dns_lab
history
main
summary
FAQ
How do indexes work?
How do I know when to create or use multiple indexes?
127.0.0.1:8000/en-US/manager/search/adddatamethods/inputsettings#


splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
Add Data
Select Source
Set Source Type
Input Settings
Review
Done
Back
Review
Input Settings
Optionally set additional input parameters for this data input as follows:
Host
When the Splunk platform indexes data, each event receives a "host" value. The host value should be the name of the machine from which the event originates. The type of input you choose determines the available configuration options. Learn More
Constant value
Regular expression on path
Segment in path
Host field value
LAPTOP-EHJ3QFJI
Index
The Splunk platform stores incoming data as events in the selected index. Consider using a "sandbox" index as a destination if you have problems determining a source type for your data. A sandbox index lets you troubleshoot your configuration without impacting production indexes. You can always change this setting later. Learn More
Index
dns_lab
Create a new index
FAQ
How do indexes work?
How do I know when to create or use multiple indexes?

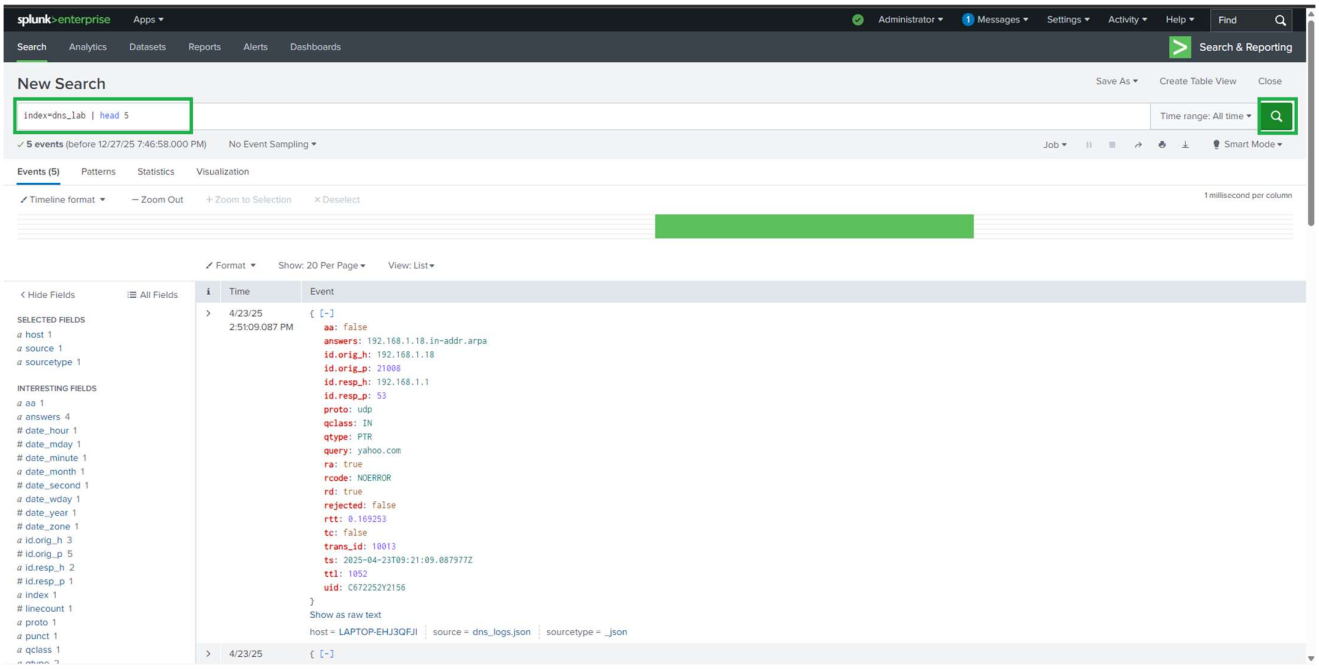

splunk>enterprise
Apps
Administrator
Messages
Settings
Activity
Help
Find
Add Data
Select Source
Set Source Type
Input Settings
Review
Done
Back
Submit
Review
Input Type ..... Uploaded File
File Name ..... dns_logs.json
Source Type ..... _json
Host LAPTOP-EHJ3QFJI
Index ..... dns_lab



Step 2: Verify Data Ingestion

Run the following query to confirm logs are indexed:

index=dns_lab | head 5



🔍 Lab Tasks & SPL Queries

🔹 Task 1: Most Frequently Queried Domain Names

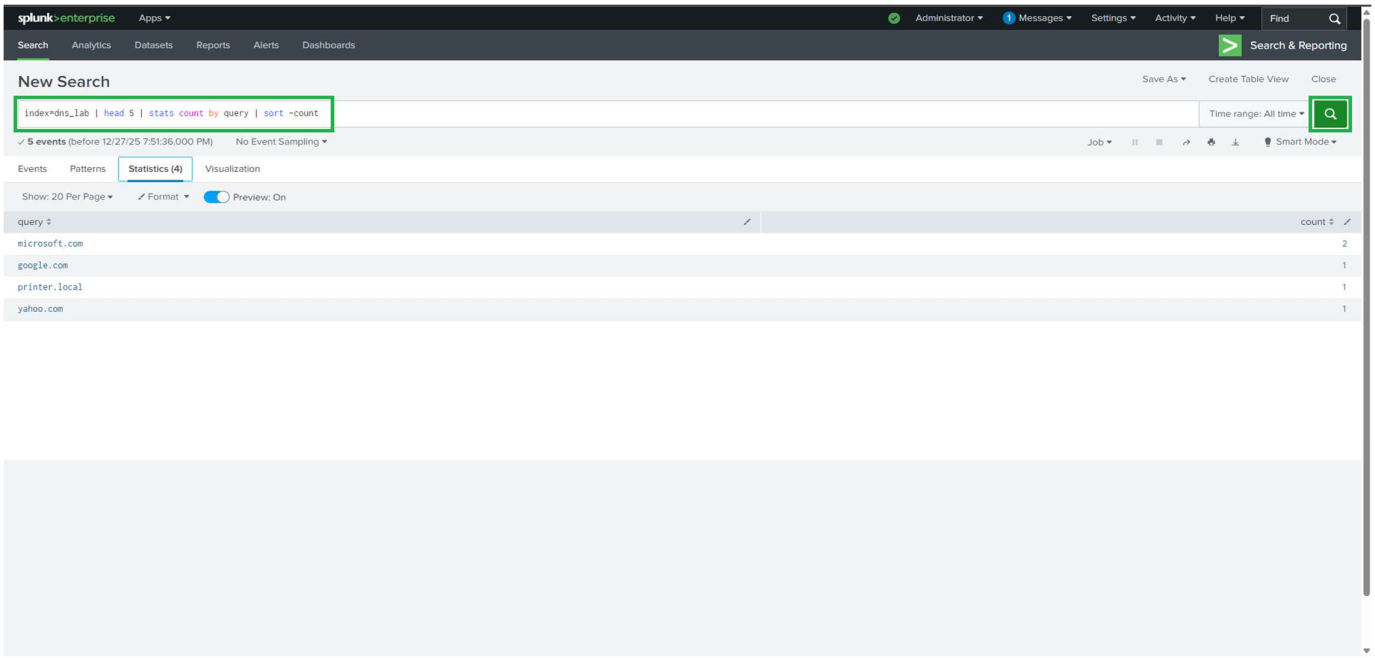
Identify which domains are queried most often.

index=dns_lab

| stats count by query

| sort -count

📌 **Use Case:** Detect suspicious domains, beaconing behavior, or malware C2 lookups.

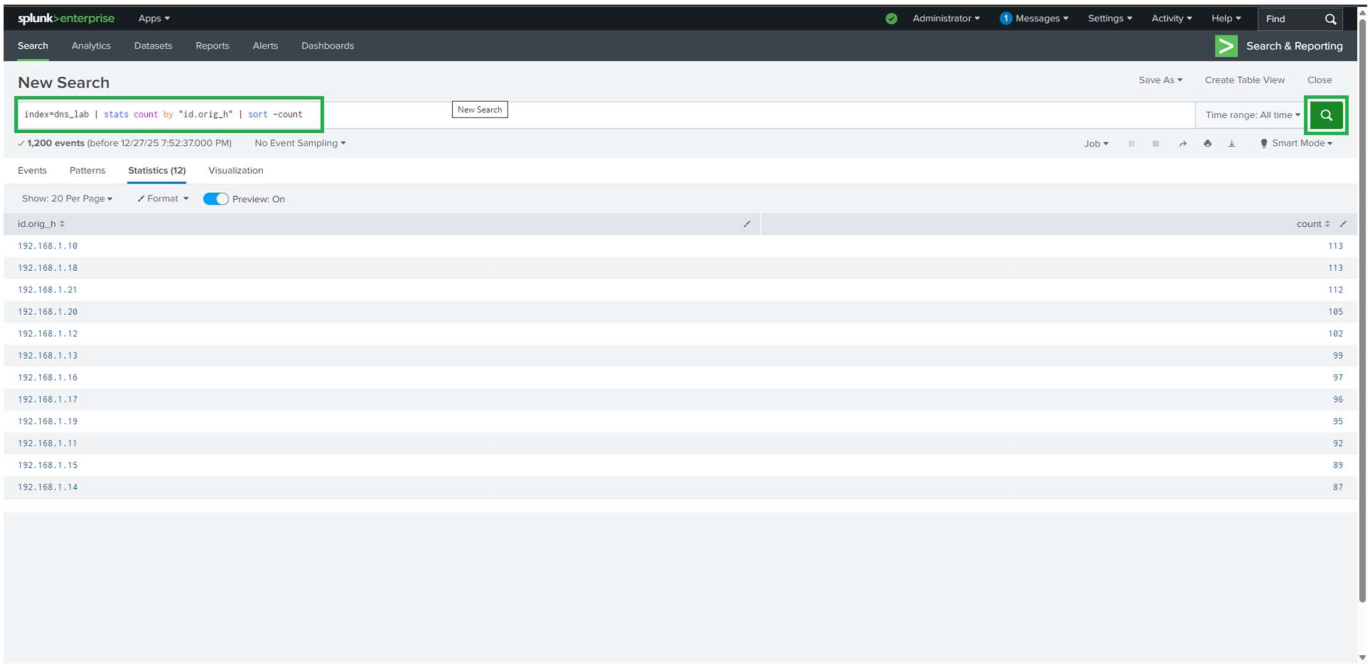


🔹 Task 2: Most Active Client IPs

Find which internal hosts generate the most DNS traffic.

```
index=dns_lab
| stats count by "id.orig_h"
| sort -count
```

📌 **Use Case:** Identify compromised systems or misconfigured devices.



🔹 Task 3: DNS Query Type Breakdown

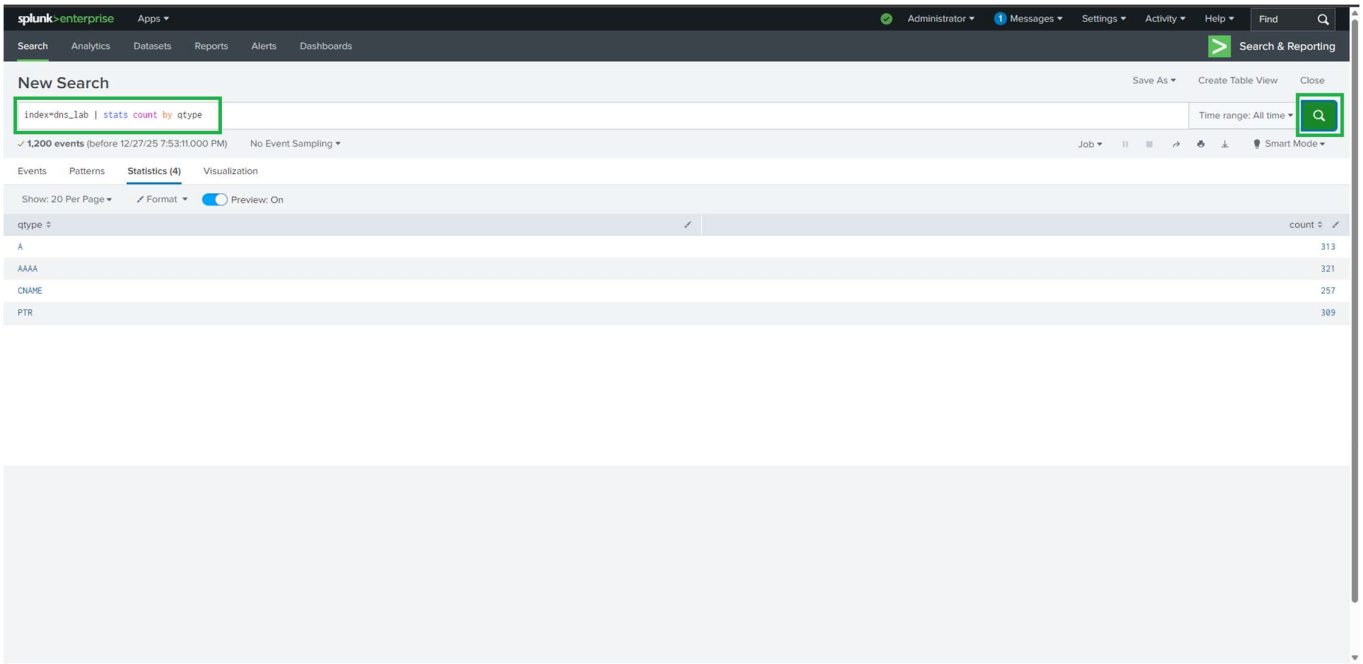
Analyze DNS record types used in the environment.

index=dns_lab

| stats count by qtype

📌 Use Case:

- PTR → Reverse lookups
- Excessive AAAA → IPv6 behavior
- CNAME → Redirect or alias usage



📈 Key Findings & Analysis

- Identified **top queried domains** such as Google, Microsoft, and Yahoo
- Detected **high-volume DNS-generating hosts**
- Observed distribution of **A, AAAA, CNAME, and PTR records**
- Analyzed **response times (RTT)** and failed lookups (NXDOMAIN)
- Established baseline DNS behavior for anomaly detection

🚨 Security Insights

This project enables detection of:

- DNS tunneling patterns
- Malware beaconing via repeated queries
- Suspicious reverse DNS lookups
- Unusual DNS latency or failed resolutions

📊 Dashboards & Alerts (Optional Enhancements)

You can extend this project by:

- Creating Splunk dashboards for DNS activity visualization
- Setting alerts for:
    - High DNS query volume
    - Rare domain lookups
    - Excessive failed DNS responses

✅ Conclusion

By completing this lab, you have:

✓ Successfully ingested and parsed DNS logs in Splunk
✓ Built SPL queries for DNS traffic analysis
✓ Identified top domains and active clients
✓ Analyzed DNS record types
✓ Gained hands-on experience in security-focused log analysis

This project serves as a **strong foundation for SOC, SIEM, and threat-hunting roles**.
