

# Project Title: HTTPS Log Traffic Analysis & Threat Investigation Using Splunk SIEM

## Project Objective

To validate the successful ingestion of HTTPS logs in Splunk, analyze traffic behavior through port and IP pattern analysis, and identify potential anomalies such as beaconing, misconfiguration, or internal compromise by drilling into high-traffic destination IPs and their associated source hosts.

## Tools & Technologies Used

- **Splunk Enterprise** (SIEM platform)
- **SPL (Search Processing Language)** for querying
- **HTTPS Log Dataset**
- **Sourcetype:** `HTTPSLOGS`
- **Indexes Used:** `index=*`, `index=_*`

### 1. Validating HTTPS Logs Ingestion

#### Objective:

To verify that logs with the sourcetype `HTTPSLOGS` are properly indexed and visible in the Splunk environment before proceeding with detailed analysis.

Confirmed that `HTTPSLOGS` entries were successfully ingested and searchable, laying the foundation for further inspection of secure web traffic behavior.

The screenshot shows the Splunk Enterprise search interface. The search bar contains the query `index=* sourcetype="HTTPSLOGS"`. The search results are displayed in a table format, showing events with source and destination IP addresses, ports, and methods. The table has columns for Time, Event, and various fields extracted from the logs.

Time	Event
6/14/25 8:04:18.000 AM	1331921610.200000 C9oek7REpUfCATo6 192.168.202.102 2489 192.168.229.101 80 1 POST 192.168.229.101 /login http://192.168.229.101/ Mozilla/4.0 (compatible; MSIE 9.0; Windows NT 6.1) 19 832 200 OK t/html (empty) - - FP5Pmq18QJ3Vfj19Se text/plain FIFqLY1TmvROz2hT1I tex host = WIN-8PVJ5SGCP99 source = http.log.gz sourcetype = HTTPSLOGS
6/14/25 8:04:18.000 AM	1331921610.210000 Chjy3u2LZfbIEUsci 192.168.203.63 62875 192.168.229.101 80 46 HEAD 192.168.229.101 /19233_wallpaper150/ - DirBuster-0.12 (http://www.owasp.org/index.php/Category:OWASP_DirBuster_Project) 0 0 484 NOT FOUND (empty) - - - - - host = WIN-8PVJ5SGCP99 source = http.log.gz sourcetype = HTTPSLOGS

## 2. Deduplicated Port Analysis in HTTPS Logs

### Objective:

To extract and display a unique list of source ports involved in HTTPS communications, along with corresponding destination ports, helping detect unusual or non-standard ports.

### Outcome:

Produced a deduplicated list of source ports and associated destination ports, offering clarity on port usage trends and potentially suspicious traffic behavior.

The screenshot shows the Splunk 9.4.3 web interface. The search bar contains the query: `index=* OR index=* sourcetype=HTTPSLOGS | table src_port, dst_port | dedup src_port`. The search is paused, and the results are displayed in a table view. The table has two columns: `src_port` and `dst_port`. The results show a list of source ports and their corresponding destination ports, with some ports appearing multiple times.

src_port	dst_port
62794	80
1220	80
46645	80
62799	80
62796	80
62795	80
62797	80
62801	80
62800	80
62802	80
62803	80
62798	80

### 3. Reviewing Port Distribution in System and HTTPS Indexes

#### Objective:

To analyze HTTPS log data across both default (`index=*`) and internal Splunk system indexes (`index=_*`), ensuring no relevant port activity is overlooked.

#### Outcome:

Returned a complete view of `src_port` and `dst_port` across all available indexes, confirming that port activity tied to HTTPS logs is captured and uniformly available.

The screenshot shows the Splunk Enterprise search interface. The search bar contains the query: `index=* OR index=* sourcetype=HTTPSLOGS | table src_port, dst_port`. The search is paused, and the results are displayed in a table view. The table has two columns: `src_port` and `dst_port`. The results show a list of source ports and their corresponding destination ports.

src_port	dst_port
62794	80
1220	80
46645	80
62799	80
62796	80
46645	80
62795	80
62797	80
62801	80
62800	80
62802	80
46645	80

#### 4. Identifying Top Destination IPs in HTTPS Logs

##### Objective:

To identify the most frequently contacted destination IP addresses across all logs tagged with `HTTPSLOGS`. This step is critical for spotlighting high-traffic endpoints that may be legitimate services—or potential exfiltration or beaconing targets.

##### Outcome:

Discovered that the IP address `192.168.229.101` appeared **98,391 times**, making it the top destination in the data set. This raised flags for possible internal server activity, misconfigured systems, or targeted communications worthy of deeper inspection.

The screenshot shows the Splunk Enterprise search interface. The search query is `index=* OR index=* sourcetype=HTTPSLOGS | top limit=10 dst_ip`. The results are displayed in a table with columns for `dst_ip`, `count`, and `percent`. The top result is `192.168.229.101` with a count of 98391 and a percentage of 87.458667.

dst_ip	count	percent
192.168.229.101	98391	87.458667
192.168.24.101	12660	11.253333
192.168.23.202	1232	1.095111
192.168.21.103	108	0.096000
192.168.229.156	44	0.039111
192.168.25.202	18	0.016000
192.168.22.202	18	0.016000
192.168.23.101	12	0.010667
192.168.28.102	5	0.004444
192.168.28.101	5	0.004444

## 5. Deep-Dive on High-Volume Destination IP

### Objective:

To conduct a focused query on the specific destination IP (192.168.229.101) identified in Step 4. The goal was to isolate all logs involving this IP and analyze the nature of its interactions, helping determine if this is normal system behavior or an anomaly.

### Outcome:

Retrieved all relevant entries for 192.168.229.101, enabling further inspection of traffic patterns, involved source systems, and communication volume. This step allowed for targeted threat hunting, log correlation, or escalation for network segmentation if needed.

The screenshot shows the Splunk Enterprise interface with a search query: `index=* OR index=* sourcetype=HTTPSLOGS dst_ip=192.168.229.101*`. The search results are displayed in a table format, showing events from 6/14/25 at 8:04:18 AM. The table includes fields like Time, Event, and various log details. The search is paused, and the interface shows options to save, create table view, or close the search.

Time	Event
6/14/25 8:04:18.000 AM	1331921610.200000 C9omk7REpUfCatOe6 192.168.202.102 2489 192.168.229.101 80 1 POST 192.168.229.101 /login http://192.168.229.101/ Mozilla/4.0 (compatible; MSIE 9.0; Windows NT 6.1) 19 832 200 OK t/html (empty) - - - FP5Pmq18QJ3Vfj19Se text/plain FifqLYITmvR0z2hT11 tex
6/14/25 8:04:18.000 AM	1331921610.210000 Chjy3u2LZfbIEUsci 192.168.203.63 62875 192.168.229.101 80 46 HEAD 192.168.229.101 /19233_wallpaper150/ - DirBuster-0.12 (http://www.owasp.org/index.php/Category:OWASP_DirBuster_Project) 0 0 484 NOT FOUND (empty) - - -
6/14/25 8:04:18.000 AM	1331921610.210000 CQegPH15hwNA6J6jw 192.168.203.63 62874 192.168.229.101 80 78 HEAD 192.168.229.101 /170938/ - DirBuster-0.12 (http://www.owasp.org/index.php/Category:OWASP_DirBuster_Project) 0 484 NOT FOUND (empty) - - -

## 6. Identify Top Source IPs Communicating with Suspicious Destination

### Objective:

To determine which source IP addresses have communicated most frequently with the high-volume destination IP 192.168.229.101. This step helps uncover which internal systems or users are involved in this traffic, which could indicate misconfiguration, scanning activity, or even compromised endpoints.

### Outcome:

Generated a ranked list of the top 10 source IPs contacting 192.168.229.101. This narrowed the investigation to specific internal hosts and provided a shortlist of endpoints for deeper forensic review or threat correlation.

Splunk 9.4.3 interface showing a search query: `index=_* OR index=* sourcetype=HTTPSLOGS dst_ip="192.168.229.101" | top limit=10 src_ip`. The search is paused. The results table shows the top 10 source IP addresses by count.

src_ip	count	percent
192.168.203.63	486954	99.194352
192.168.202.79	2117	0.431241
192.168.202.102	1826	0.371963
192.168.202.100	7	0.001426
192.168.203.64	5	0.001019

## Conclusion

This project successfully leveraged Splunk SIEM to transform raw HTTPS log data into actionable intelligence. By systematically validating ingestion, extracting key fields, and analyzing high-frequency communication patterns, the investigation surfaced critical insights about internal traffic behavior. It also demonstrated how Splunk can help detect potentially compromised systems or misconfigurations in enterprise environments.