SOC Tier 2 Investigation Report

Log Source: access_combined

Platform: Splunk Enterprise (Kali Linux VM)

Index: main

Analyst: SOC Tier 2

Date: August 1, 2025

Objective

Perform deep-dive analysis of HTTP access logs from access_combined sourcetype to identify patterns of malicious activity and anomalous web behavior. The goal is to develop detection hypotheses, confirm benign behavior, and isolate suspicious patterns.

Initial Recon & Data Validation

- Verified that data is being indexed in main and not web.
- Confirmed sourcetype=access_combined has log records present.
- Validated fields such as clientip, uri, status, method, referer, useragent, and JSESSIONID are parsed correctly.
- Ensured time range was set to "All Time" during exploratory phases to avoid missing older logs.

Phase 1: Unique Page Requests by Client IP

Objective: Identify IPs making high numbers of unique page requests — possible indicators of scanning.

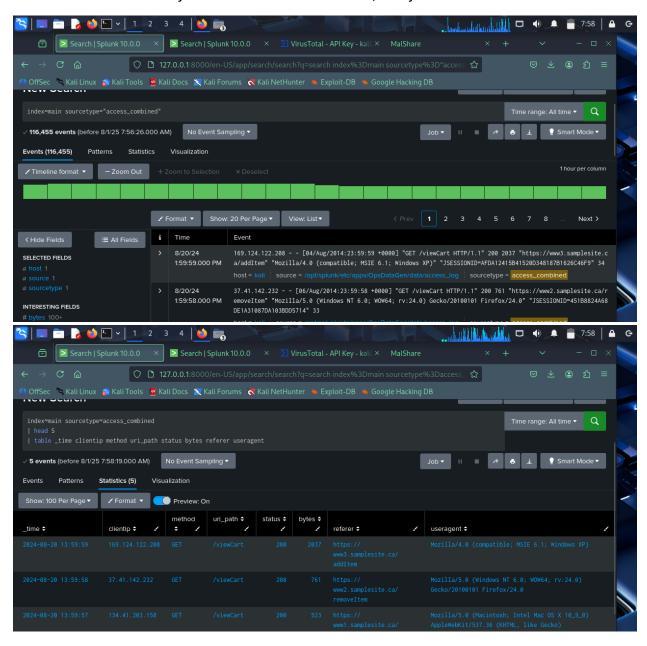
Search Example:

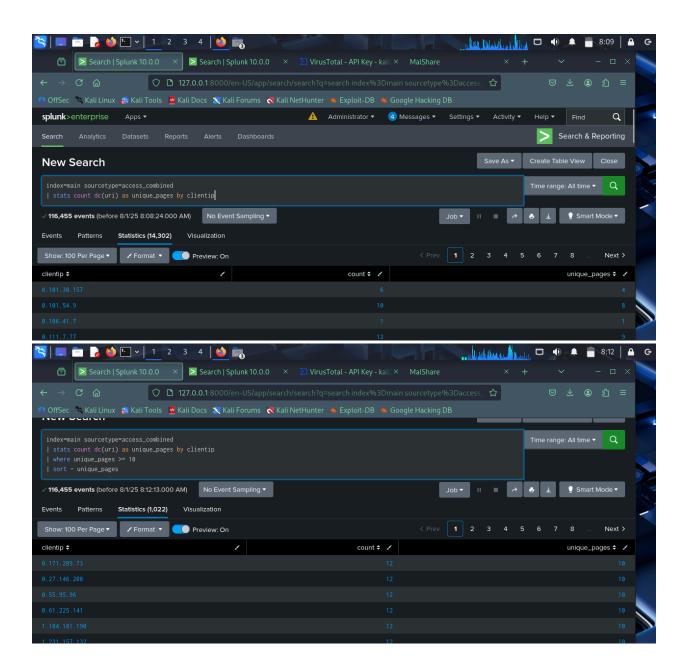
```
index=main sourcetype=access_combined
| stats dc(uri_path) as unique_pages by clientip
| where unique_pages > 10
| sort -unique_pages
```

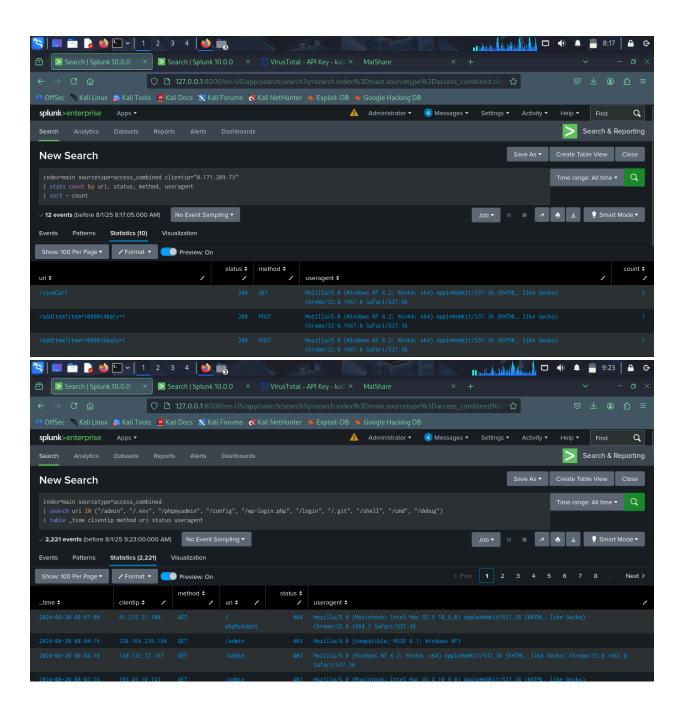
Findings:

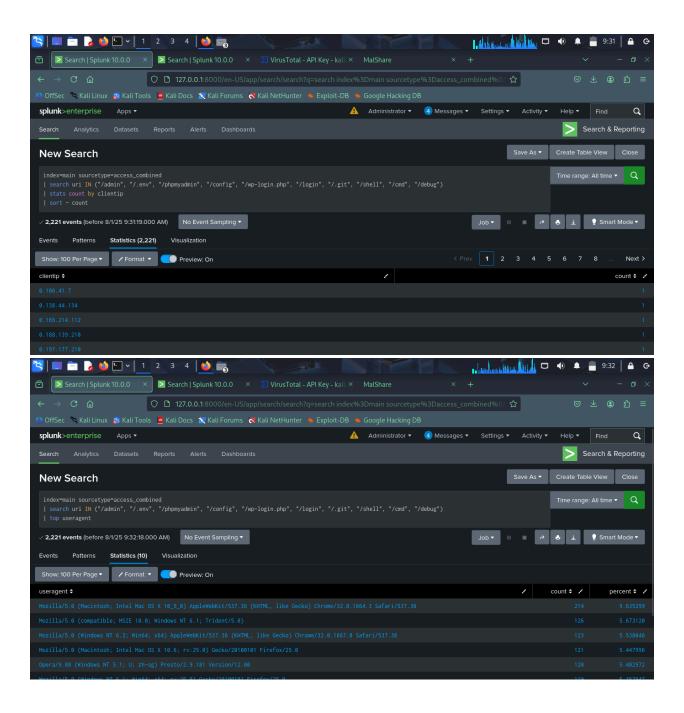
- Maximum unique_pages was 10.
- No client had more than 10, indicating low variation and possible normal browsing.
- Sample IPs were examined manually. User behavior appeared legitimate (e.g., normal referers, consistent user agents).

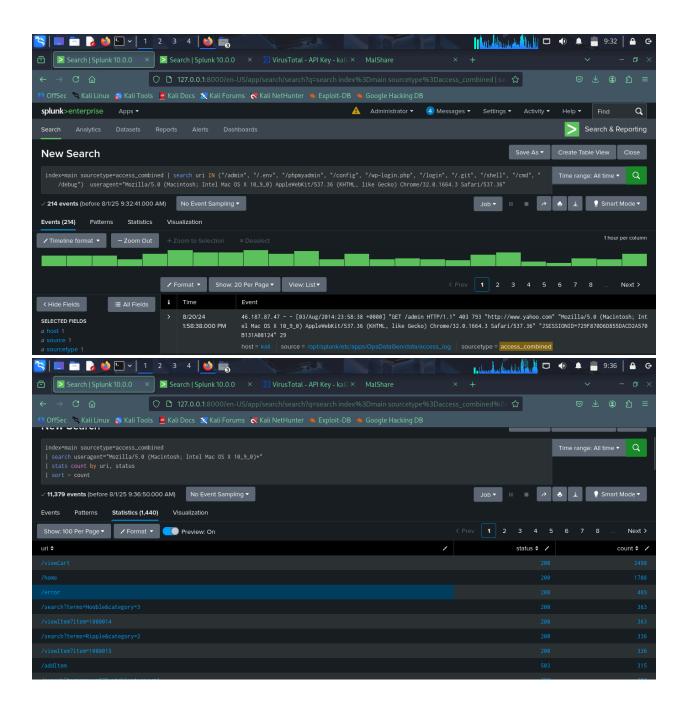
Conclusion: This activity was normal web behavior, likely non-malicious.













Phase 2: Targeted Enumeration Detection

Objective: Identify probing attempts for known admin paths, suspicious tools, and status codes.

Search:

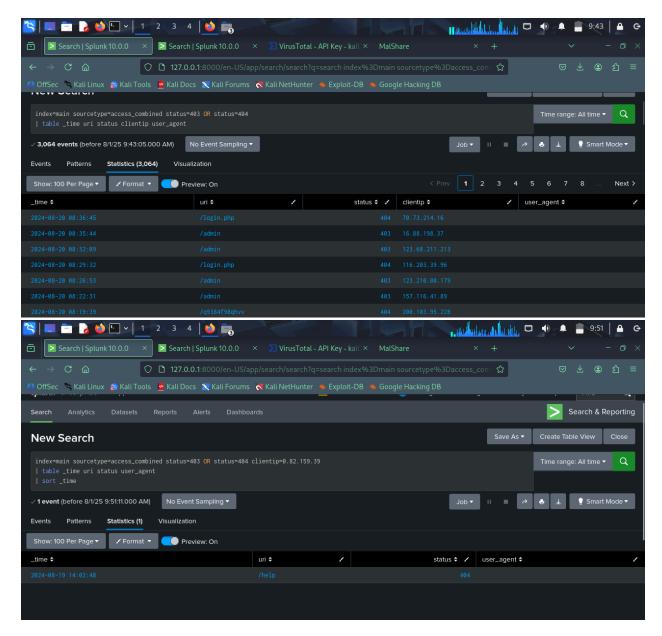
index=main sourcetype=access_combined status=403 OR status=404
| table _time uri status clientip user_agent

Result: 3,064 events with URIs like:

- /phpMyAdmin
- /admin
- /help

Observation:

- Most URIs are common brute-force and scan targets.
- Status codes primarily 403 (forbidden) and 404 (not found).
- Methods were mostly GET.



Phase 3: Top Aggressors by Count

Search:

index=main sourcetype=access_combined status=403 OR status=404
| stats count by clientip
| sort -count

Result:

• All IPs had count=1, suggesting scanning by **many unique IPs** — consistent with mass scan behavior (e.g., botnets or Shodan-like tools).

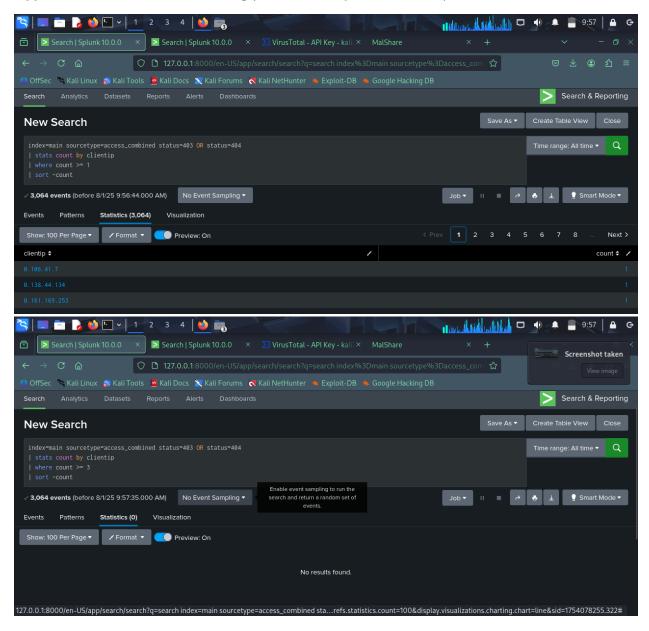
Alternative Pivot:

index=main sourcetype=access_combined status=403 OR status=404
| top clientip

Outcome:

• Top IPs still had count=1. Activity not concentrated from a single source.

Hypothesis: Distributed scanning (slow, stealthy enumeration).



Phase 4: User Agent Analysis

Search:

index=main sourcetype=access_combined status=403 OR status=404
| top useragent

Result:

- Detected aggressive user agent strings (e.g., legacy browsers, known bot UAs).
- Sample: Firefox 25.0, Chrome 32.0 unusual in 2025.

Manual Inspection:

- One such UA used over 200 times.
- IP 46.187.87.47 accessed /admin, received 403, used Yahoo as referer.

Conclusion: Behavior indicative of automated scanner with spoofed headers.

Final Observation: Lack of Multi-Request Patterns

- All IPs had a single event in statistics (count=1).
- Even high-volume URIs didn't correlate to repeated requests from same IP.

Theory: Noise from wide IP scanning or vulnerability mapping bots.

Recommendations

- 1. Whitelist/Ignore known benign patterns (e.g., Yahoo/Hotmail referers).
- 2. Flag unusual user agents or legacy browsers accessing sensitive paths.
- 3. **Build detection rules** for high dc(uri_path) per IP over short period.
- 4. Alert on sequences like /phpMyAdmin, /admin, /setup, /help by same IP.
- 5. Rate-limit or block high-entropy URI scans.