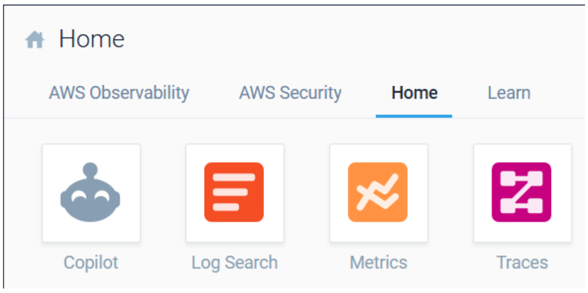
Data Searching

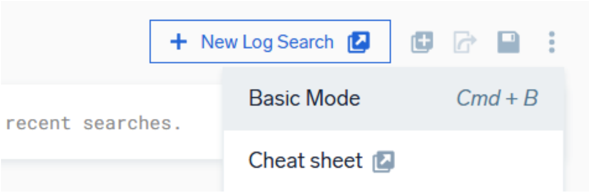
Data searching in SIEM is a process of querying logs and security event data collected from the organization's IT infrastructure. It allows SOC analysts to find patterns, discover anomalies, and search for the indicators of compromise. It is especially important in spotting potentially harmful activity, such as failed login attempts or privilege escalation, which can result in a security breach. Another important use is the incident investigation. Queried data can help reconstruct the attack timeline and identify the root cause.

Basic Log Search

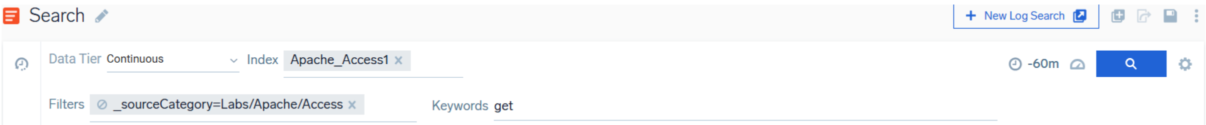
To perform a basic log search, click on the log search button on the main menu in SUMO Logic.



You will be redirected to the advanced log search window. To change the view, click on the menu button on the right side and choose the Basic Mode option.

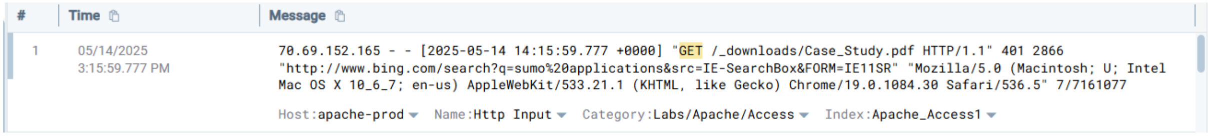


To query all the "get" requests made in the last 60 minutes from the Apache server, fill in the search fields and click the magnifying glass icon.



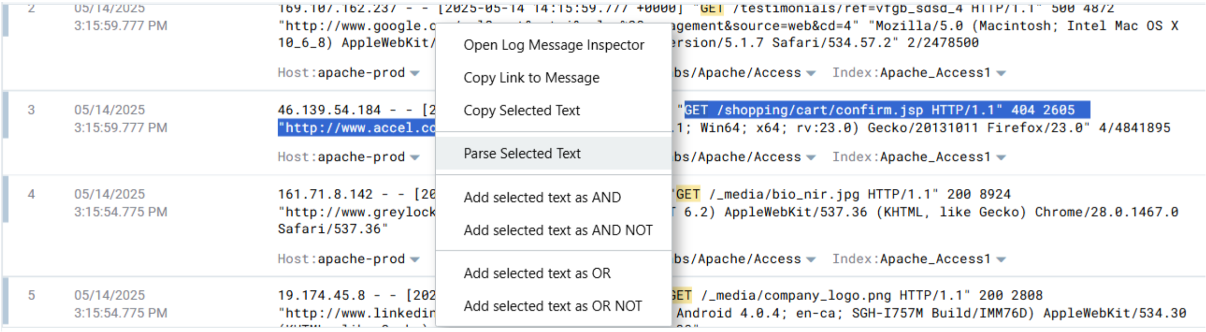
Data Tier Continuous
Index Apache_Access1
-60m
Filters _sourceCategory=Labs/Apache/Access
Keywords get

The received logs are in a raw format.



#	Time	Message
1	05/14/2025 3:15:59.777 PM	70.69.152.165 - - [2025-05-14 14:15:59.777 +0000] "GET /_downloads/Case_Study.pdf HTTP/1.1" 401 2866 "http://www.bing.com/search?q=sumo%20applications&src=IE-SearchBox&FORM=IE11SR" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_7; en-us) AppleWebKit/533.21.1 (KHTML, like Gecko) Chrome/19.0.1084.30 Safari/536.5" 7/7161077 Host:apache-prod Name:Http Input Category:Labs/Apache/Access Index:Apache_Access1

By default, the basic search will return logs in a raw format. That means that an analyst will receive data in an unparsed and unformatted version of a log ingested by the SIEM. In the picture above, we can see that all the data creates a single line of value. To parse the log, highlight text from the chosen log, and click “Parse Selected Text” option.



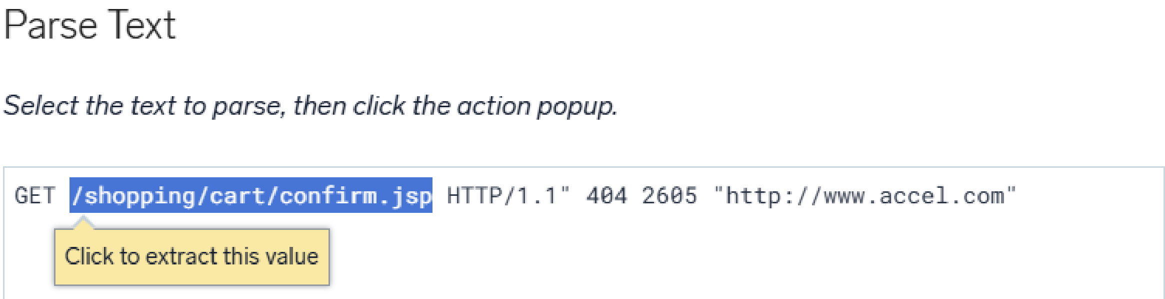
```
2   05/14/2025      169.167.162.237 - - [2025-05-14 14:15:59.777 +0000] "GET /testimonials/ref=vfgb_sdsd_4 HTTP/1.1" 500 4872
    3:15:59.777 PM  "http://www.google.c                              agement&source=web&cd=4" "Mozilla/5.0 (Macintosh; Intel Mac OS X
                    10_6_8) AppleWebKit/                              ersion/5.1.7 Safari/534.57.2" 2/2478500
                    Host:apache-prod                                  bs/Apache/Access   Index:Apache_Access1

3   05/14/2025      46.139.54.184 - - [2                              "GET /shopping/cart/confirm.jsp HTTP/1.1" 404 2605
    3:15:59.777 PM  "http://www.accel.cc                              1; Win64; x64; rv:23.0) Gecko/20131011 Firefox/23.0" 4/4841895
                    Host:apache-prod                                  bs/Apache/Access   Index:Apache_Access1

4   05/14/2025      161.71.8.142 - - [20                              'GET /_media/bio_nir.jpg HTTP/1.1" 200 8924
    3:15:54.775 PM  "http://www.greylock                              T 6.2) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/28.0.1467.0
                    Safari/537.36"
                    Host:apache-prod                                  bs/Apache/Access   Index:Apache_Access1

5   05/14/2025      19.174.45.8 - - [202                              GET /_media/company_logo.png HTTP/1.1" 200 2808
    3:15:54.775 PM  "http://www.linkedir                              Android 4.0.4; en-ca; SGH-I757M Build/IMM76D) AppleWebKit/534.30
```

A new window will open. Highlight a chosen text again, press “Click to extract this value” and then name the value.

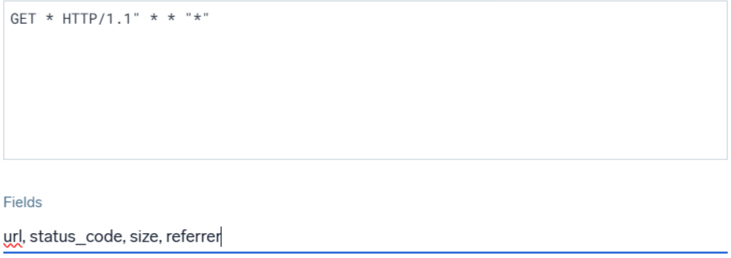


Parse Text

Select the text to parse, then click the action popup.

```
GET /shopping/cart/confirm.jsp HTTP/1.1" 404 2605 "http://www.accel.com"
```

After clicking the yellow sign, the text will turn into an asterisk. Name all of the highlighted text, and click the submit button.



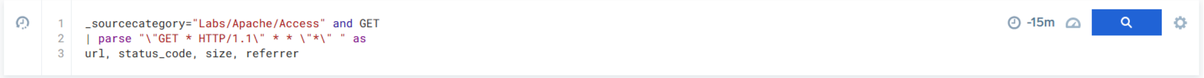
```
GET * HTTP/1.1" * * "*"
```

Fields

url, status_code, size, referrer

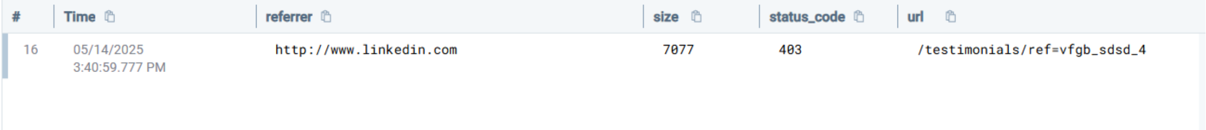
Open a new search, and paste in the query below.

Query: _sourcecategory="Labs/Apache/Access" and GET
*| parse "\"GET * HTTP/1.1\" * * \"*\" " as*
url, status_code, size, referrer



```
_sourcecategory="Labs/Apache/Access" and GET
| parse "\"GET * HTTP/1.1\" * * \"*\" " as
url, status_code, size, referrer
```

As a result, you will receive a parsed log looking like that.



#	Time	referrer	size	status_code	url
16	05/14/2025 3:40:59.777 PM	http://www.linkedin.com	7077	403	/testimonials/ref=vfgb_sdsd_4

Summary

Parsing logs is crucial for SOC analysts as it formats raw and unstructured data into formatted and searchable information. It breaks down the raw logs and structures them into key-value pairs, making them easy to search, filter, and correlate across different data sources in your IT environment. It also enables real-time detection and alerts, as SIEM rules rely on the parsed fields to trigger the alerts. Without parsing, real-time threat detection would be impossible.