
📑 Log Parsing in Sumo Logic

🔎 Why Parsing Logs is Important

Raw logs are messy and unstructured — just strings of text. By parsing, we extract meaningful fields (like IPs, usernames, status codes). This makes logs **structured, searchable, and actionable**.

✅ Benefits of Parsing Logs

1. **Clarity** – You work with fields, not raw text.
2. **Speed** – Queries run faster with extracted fields.
3. **Deeper Analysis** – Count, group, and filter become easy.
4. **Security Use Case** – Spot abnormal behavior (e.g., one IP spamming requests).

📊 Example Without Parsing

Suppose you want to find IP addresses sending requests in Apache logs.

Unparsed Search:

```
_sourceCategory=Labs/Apache/*

"192.168."
```

- This only finds logs *containing* that text.
- You cannot count or group results properly.
- If multiple IPs appear, you must manually scan them.

📊 Example With Parsing

Parsed Search:

```
_sourceCategory=Labs/Apache/*

| parse regex "(?<ip_address>\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3})"

| count by ip_address

| where _count > 500
```

Explanation:

- parse regex "(?<ip_address>…)" → Extracts any IP address into the field ip_address.
- count by ip_address → Groups requests by unique IPs.
- where _count > 500 → Filters to show only IPs that made more than 500 requests.

✅ **Result Example:**

```
ip_address      _count

---------------  ------

192.168.1.10    1032

10.0.2.15       785

203.0.113.50    650
```

Instead of raw text, you now see **which IPs are hammering your server** — a clear security insight.

🚀 **Summary**

- **Unparsed logs** = noise, limited analysis.
- **Parsed logs** = structured data, powerful queries, actionable results.
- Parsing is the foundation for **dashboards, alerts, and investigations** in Sumo Logic.
