# Analyzing 10M+ Traefik Requests with GoAccess to Detect Traffic Patterns

# GoAccess Monitoring Guide for Traefik Logs

This guide explains how to use **GoAccess** to analyze Traefik access logs and quickly detect traffic patterns, spikes, or possible abuse.

---

## 1. Install GoAccess

```
sudo apt update
sudo apt install goaccess -y
```

Verify installation:

```
goaccess --version
```

---

## 2. Generate a Quick HTML Report

From the directory containing `access.log`:

```
tail -n 200000 access.log | goaccess -o /tmp/traefik_report.html --log-format=COMBINED
```

Explanation:

- `tail -n 200000` → analyze recent traffic only (faster)
- `-o` → output HTML report

- `COMBINED` → common web log format

Open the report:

```
firefox /tmp/traefik_report.html
```

Or copy the file to your local machine.

---

# 3. Generate a Full Report

If the log is not extremely large:

```
goaccess access.log -o /tmp/traefik_report.html --log-format=COMBINED
```

---

# 4. Create a Real?Time Live Dashboard

```
goaccess access.log \
  --log-format=COMBINED \
  --real-time-html \
  -o /tmp/traefik_live.html
```

Open:

```
/tmp/traefik_live.html
```

The page updates automatically as new traffic arrives.

---

# 5. Key Panels to Monitor

## Visitor Hostnames and IPs

Shows the most active IP addresses hitting your services.

Useful for detecting:

- Heavy users

- Possible attackers
- Bots
- Network scans

---

# Requested URLs

Displays the most requested endpoints.

Helps identify:

- API usage
- File downloads
- Scanning attempts
- Misbehaving clients

---

# HTTP Methods

Examples:

- `GET` → normal browsing
- `POST` → login/API requests
- `PROPFIND` → WebDAV (Nextcloud)
- `PUT` → uploads

---

# Response Codes

Examples:

- `200` → success
- `301/302` → redirects
- `401` → authentication required
- `403` → forbidden
- `404` → missing resources
- `500` → server errors

Useful to detect brute force attempts or broken services.

---

# Traffic Timeline

Shows traffic spikes by time. Helpful for correlating with:

- CPU spikes
- sync activity
- scheduled jobs
- potential attacks

---

# 6. Useful Filters for Large Logs

Analyze recent traffic only:

```
tail -n 500000 access.log | goaccess -o /tmp/report.html --log-format=COMBINED
```

Analyze a specific service (example Nextcloud):

```
grep cloud-ina access.log | goaccess -o /tmp/nextcloud.html --log-format=COMBINED
```

Analyze a specific IP:

```
grep 89.247.171.108 access.log | goaccess -o /tmp/ip.html --log-format=COMBINED
```

---

# 7. Quick Incident Workflow

When something unusual happens:

1. Generate a report from recent logs
2. Check **Top Visitor IPs**
3. Check **HTTP Methods**
4. Check **Requested URLs**
5. Identify unusual traffic patterns

Example:

- Many `PROPFIND` → Nextcloud sync activity
- Many `404` → vulnerability scanner
- Many `401` → brute force login attempts

---

# 8. Automate Daily Reports

Example cron job:

```
0 * * * * tail -n 500000 /path/access.log | goaccess -o /var/www/html/traefik.html --log-format=COMBINED
```

This generates a regularly updated monitoring dashboard.

---

# 9. Recommended Observability Stack

For larger infrastructures:

- Traefik Metrics → Prometheus
- Visualization → Grafana
- Log Analysis → GoAccess

This combination provides both real-time traffic monitoring and long-term analytics.

---

# 10. Summary

GoAccess is an excellent tool for quickly understanding proxy traffic:

- Detect heavy users
- Identify scanning attempts
- Monitor API usage
- Correlate traffic spikes with system load

In environments like **Traefik + Proxmox + multiple services**, GoAccess is one of the fastest ways to debug traffic issues.

---

Revision #2
Created 16 March 2026 11:36:08 by Admin
Updated 16 March 2026 11:36:53 by Admin