Live Network Threat Telemetry API — Master Project Scope (Steps 1–7)

1. MVP Scope	Inputs: NetFlow/IPFIX, Zeek JSON. Enrichments: GeoIP, ASN, TI, Risk. Outputs: Spl
2. Data Models & API Contract	Base API (/v1), Bearer auth, /ingest, /lookup, /outputs/*, /metrics, schemas and limits.
3–6. Core Build → Deployment	Backend implemented, enrichments integrated, outputs wired, Docker image released
7. Requests Observability (NEW)	Request auditing, summary metrics, live tail, admin endpoints, UI Requests tab, CSV

Updated Endpoint List (Step 7 additions marked ★)

```
/health (GET) — health check
/ingest (POST) — batch upload flows/logs
/lookup (POST) — single enrichment
/outputs/splunk (POST/PUT) — configure Splunk
/outputs/elastic (POST/PUT) — configure Elastic
/alerts/rules (POST/PUT) — webhook alert rule
/metrics (GET) — Prometheus metrics (requests_total, duration, active_clients) ★
/admin/requests (GET) — paginated request audit ★
/admin/requests/summary (GET) — 15m summary ★
/admin/requests/stream (GET) — SSE live tail ★
```

Request Audit Record — Canonical Schema

```
{
  "id": "b0b1...",
  "ts": "2025-08-14T12:45:11Z",
  "tenant_id": "tenant-123",
  "api_key_hash": "hmac:8c9e...",
  "client_ip": "203.0.113.9",
  "user_agent": "curl/8.6.0",
  "method": "POST",
  "path": "/v1/ingest",
  "status": 200,
  "duration_ms": 41,
  "bytes_in": 5241,
  "bytes_out": 238,
  "result": "ok",
  "trace_id": "c6f5a0...",
  "geo_country": "DE",
  "asn": "AS3320"
}
```

Retention Policy

All audit records follow the existing MVP retention: **7 days** with daily purge.

UI: Dashboard & Requests Tab Highlights

- Dashboard: new mini-cards (Requests 15m, 2xx/4xx/5xx, P95 latency, Active clients)
- Requests Tab: live tail table with filters, CSV export, detail drawer (Geo/ASN, sizes, headers)
- Visuals: small sparkline, stacked status chart, latency histogram, geo map
- UX: compact rows, semantic status chips, masked API keys, copyable Trace ID

Security & Privacy Notes

Never store request bodies. Use HMAC-SHA256 to mask API keys. Scope queries to tenant_id. SSE throttled to avoid leaking timing side-channels.