Here’s a **revised, optimized solution** for your **on-premises ISO 8583 monitoring and alerting system**, incorporating **high availability (HA)**, **audible alarms for night shifts**, and **network connectivity checks**, with **recalculated hardware requirements**:

**Architecture Overview**

**1. Core Components & Hardware**

*(Total physical servers: 22–26)*

| **Component** | **Servers** | **Specifications (Per Server)** | **Purpose** | **HA Mechanism** |
| --- | --- | --- | --- | --- |
| **Apache Kafka (Brokers)** | 4 | 16 vCPU, 64GB RAM, 2TB NVMe SSD | Message ingestion (200k TPM) | Replication factor=3, 4 brokers for failover. |
| **Apache Flink (Task Managers)** | 4 | 32 vCPU, 128GB RAM, 1TB NVMe SSD | Fraud detection, aggregations | Checkpointing to MinIO; 4 nodes for redundancy. |
| **Elasticsearch (Hot Tier)** | 5 | 64 vCPU, 256GB RAM, 4TB NVMe SSD | Transaction logs & dashboards | 3 master + 2 data nodes; replica shards=2. |
| **VictoriaMetrics (TSDB)** | 2 | 32 vCPU, 128GB RAM, 10TB SATA SSD | Metrics storage (TPS, latency) | Replication (-replicationFactor=2). |
| **PostgreSQL (Primary + Replica)** | 2 | 16 vCPU, 64GB RAM, 2TB NVMe SSD | BIN tables, reconciliation data | Streaming replication + Patroni for HA. |
| **MinIO (Cold Storage)** | 4 | 16 vCPU, 64GB RAM, 50TB HDD | Archived logs/metrics (1 year retention) | Erasure coding (4+4 disks). |
| **Network Monitoring** | 2 | 8 vCPU, 32GB RAM, 1TB SATA SSD | Telegraf (SNMP), Blackbox Exporter, SmokePing | Load-balanced instances. |
| **Audible Alarms** | 1–2 | 4 vCPU, 8GB RAM, 500GB HDD (Raspberry Pi) | Node-RED + physical sirens/speakers | Backup RPi with battery. |
| **Support Services** | 3 | 8 vCPU, 32GB RAM, 1TB SATA SSD | Grafana, Kibana, HAProxy, Prometheus | Load-balanced Grafana/Kibana. |

**2. Key Adjustments & Rationale**

**A. High Availability Enhancements**

1. **Kafka**: 4 brokers (up from 3) to tolerate 2 simultaneous failures.
2. **Elasticsearch**: 5 nodes (3 master + 2 data) for split-brain prevention.
3. **PostgreSQL**: Patroni + pgPool-II for automated failover.
4. **MinIO**: 4-node cluster with erasure coding (survives 4 disk/node failures).

**B. Audible Alarm System**

* **Node-RED** triggers alarms via HTTP webhooks from **Prometheus Alertmanager**.
* **Redundant alerts**: Fallback to Zabbix scripts if Alertmanager fails.
* **Cost-effective hardware**: Use Raspberry Pi ($50) for alarms (low power/size).

**C. Network Monitoring**

* **Telegraf** polls SNMP metrics (switch/router health).
* **Blackbox Exporter** checks TCP/ICMP connectivity to firewalls/mainframes.
* **SmokePing** tracks latency/jitter between critical nodes.

**3. Recalculated Hardware Requirements**

| **Component** | **Total Servers** | **Total CPU** | **Total RAM** | **Total Storage** |
| --- | --- | --- | --- | --- |
| **Kafka** | 4 | 64 vCPU | 256 GB | 8 TB NVMe |
| **Flink** | 4 | 128 vCPU | 512 GB | 4 TB NVMe |
| **Elasticsearch** | 5 | 320 vCPU | 1,280 GB | 20 TB NVMe |
| **VictoriaMetrics** | 2 | 64 vCPU | 256 GB | 20 TB SATA SSD |
| **PostgreSQL** | 2 | 32 vCPU | 128 GB | 4 TB NVMe |
| **MinIO** | 4 | 64 vCPU | 256 GB | 200 TB HDD |
| **Network Monitoring** | 2 | 16 vCPU | 64 GB | 2 TB SATA SSD |
| **Support Services** | 3 | 24 vCPU | 96 GB | 3 TB SATA SSD |
| **Audible Alarms** | 2 | 8 vCPU | 16 GB | 1 TB HDD |

**4. Cost Optimization**

* **Reuse Hardware**: Deploy Node-RED on decommissioned PCs/Raspberry Pi.
* **ZFS Compression**: Reduce Elasticsearch/MinIO storage needs by ~40%.
* **NVMe Tiering**: Use NVMe for Kafka/Flink/PostgreSQL; HDD for MinIO.

**5. Disaster Recovery (DR)**

1. **Backups**:
   * Daily Elasticsearch snapshots to MinIO.
   * PostgreSQL WAL backups to MinIO.
2. **DR Site**:
   * Async Kafka mirroring to a secondary data center.

**6. Audible Alarm Implementation**

1. **Alert Rules** (Prometheus):

Yaml

- alert: ISO8583\_Critical\_Error

expr: (sum(rate(iso8583\_response\_codes{status!="00"}[5m])) / sum(rate(iso8583\_response\_codes[5m])) > 0.05

for: 5m

labels:

severity: critical

annotations:

summary: "Night Shift Alert: High error rate ({{ $value }}%)"

1. **Node-RED Flow**:
   * **HTTP In** → **JSON Parse** → **Filter Critical Alerts** → **Exec Node** (play sound).
2. **Alarm Devices**:
   * PC speakers (mpg123 command) or IoT sirens (ESP32 + MQTT).

**7. Network Health Dashboard (Grafana)**

* **Panels**: Switch interface status, firewall blocked packets, mainframe latency.
* **Alerts**: Auto-trigger audible alarms via Grafana → Node-RED integration.

**8. Why This Works**

1. **Scalability**: Handles 200k TPM with room to scale by adding Flink/Kafka nodes.
2. **Cost**: 100% open-source; no licensing fees.
3. **Compliance**: PAN masking, encrypted backups, and audit logs (PCI DSS).

**Final Hardware Summary**

| **Category** | **Total Servers** | **Total Cost (Est.)** |
| --- | --- | --- |
| **Transaction Processing** | 12 | $200,000 |
| **Storage & Analytics** | 8 | $150,000 |
| **Network/Alarms** | 4–6 | $20,000 |
| **Total** | **22–26** | **$370,000** |

*(Costs assume mid-tier Dell/HPE servers with 3-year support.)*

This solution ensures **real-time monitoring**, **high availability**, and **night-shift operational readiness** for your ISO 8583 infrastructure, while staying within on-premises and open-source constraints. Adjust server counts based on actual transaction patterns during load testing.