**100k+ TPS**, **multi-DC redundancy**, **on-prem hardware**, and **PCI-DSS compliance**. The design assumes **two data centers (DCs)** for geographic redundancy.

**Total Servers Required**: **~76 Servers**

*(Dedicated hardware for critical roles; adjust based on workload/HA needs)*

**1. Data Ingestion Layer (Apache Kafka)**

**Purpose**: Ingest and replicate ISO8583 transactions.  
**Servers/DC**: 6 brokers (3 for redundancy + 3 for scaling).  
**Total**: **12 servers** (6 brokers × 2 DCs).  
**Specs**:

* **CPU**: 2x AMD EPYC 9654 (96 cores).
* **RAM**: 512GB DDR5.
* **Storage**: 4x 3.84TB NVMe (RAID-10).

**Why**:

* Kafka requires 3+ brokers for HA in each DC.
* 100k TPS needs high I/O parallelism (NVMe).

**2. Processing Layer (Apache Flink)**

**Purpose**: Real-time fraud detection, PAN masking, and enrichment.  
**Servers/DC**: 3 task managers.  
**Total**: **6 servers** (3 × 2 DCs).  
**Specs**:

* **CPU**: 2x AMD EPYC 9654 (96 cores).
* **RAM**: 1TB DDR5.
* **Storage**: 2x 1.92TB NVMe (stateful processing).

**Why**:

* Flink scales horizontally; 3 nodes/DC handle 50k TPS each (150k total).

**3. Metrics Layer (VictoriaMetrics)**

**Purpose**: Store and query transaction/system metrics.  
**Components/DC**:

* **vmstorage**: 3 nodes (HA + replication).
* **vmselect**: 2 nodes (query load balancing).
* **vminsert**: 2 nodes (ingestion load balancing).  
  **Total**: **14 servers** (7 × 2 DCs).  
  **Specs**:
* **vmstorage**:
  + **CPU**: 1x AMD EPYC 9554 (64 cores).
  + **RAM**: 256GB DDR5.
  + **Storage**: 8x 15.36TB NVMe (RAID-6).
* **vmselect/vminsert**:
  + **CPU**: 1x AMD EPYC 9554 (64 cores).
  + **RAM**: 128GB DDR5.

**Why**:

* vmstorage nodes store time-series data; 3/DC ensures replication.
* Dedicated vmselect/vminsert nodes isolate query/ingestion workloads.

**4. Logging Layer (Elasticsearch)**

**Purpose**: Store and analyze ISO8583 logs (with PAN masking).  
**Nodes/DC**:

* **Hot Tier**: 3 nodes (indexing).
* **Warm Tier**: 2 nodes (search/analytics).
* **Cold Tier**: 1 node (archival).  
  **Total**: **12 servers** (6 × 2 DCs).  
  **Specs**:
* **Hot/Warm**:
  + **CPU**: 2x AMD EPYC 9654 (96 cores).
  + **RAM**: 512GB DDR5.
  + **Storage**: 12x 7.68TB NVMe (RAID-50).
* **Cold**:
  + **Storage**: 24x 15.36TB HDD (RAID-6).

**Why**:

* Hot tier handles high write throughput (100k TPS logs).
* Cross-DC replication (CCR) requires matching nodes in both DCs.

**5. Alerting & Visualization Layer**

**Components**:

* **Prometheus**: 2 servers/DC (scrape metrics).
* **Alertmanager**: 2 servers/DC (HA).
* **Grafana**: 2 servers/DC (HA).  
  **Total**: **12 servers** (6 × 2 DCs).  
  **Specs**:
* **CPU**: 1x AMD EPYC 7313 (16 cores).
* **RAM**: 64GB DDR4.
* **Storage**: 2x 1.92TB SATA SSD.

**Why**:

* Lightweight services but require HA.

**6. Security & Secrets Management (Vault)**

**Purpose**: Centralized secrets (Kafka, Elasticsearch, etc.).  
**Servers**: 3 (HA cluster across DCs).  
**Specs**:

* **CPU**: 1x Intel Xeon Silver 4310 (12 cores).
* **RAM**: 64GB DDR4.
* **Storage**: 2x 960GB SATA SSD (RAID-1).

**Why**:

* 3-node Raft cluster for HA.

**7. Load Balancing & Networking**

**Components**:

* **HAProxy/Keepalived**: 2 servers/DC (load balancing).  
  **Total**: **4 servers** (2 × 2 DCs).  
  **Specs**:
* **CPU**: 1x AMD EPYC 7313 (16 cores).
* **RAM**: 64GB DDR4.
* **Network**: 100Gbps NICs.

**Why**:

* Distribute traffic across DCs and handle failover.

**8. Change Data Capture (Debezium)**

**Purpose**: Stream database changes (PostgreSQL → Kafka).  
**Servers/DC**: 2 connectors.  
**Total**: **4 servers** (2 × 2 DCs).  
**Specs**:

* **CPU**: 1x Intel Xeon Silver 4310 (12 cores).
* **RAM**: 64GB DDR4.

**Why**:

* Dedicated connectors avoid resource contention.

**9. Backup & Recovery (MinIO)**

**Purpose**: Immutable backups for metrics/logs.  
**Servers**: 4 (erasure coding).  
**Specs**:

* **CPU**: 1x AMD EPYC 7313 (16 cores).
* **RAM**: 128GB DDR4.
* **Storage**: 24x 20TB HDD (RAID-6).

**Why**:

* MinIO requires 4+ nodes for erasure coding.

**10. Miscellaneous**

**Components**:

* **Monitoring Agents**: 2 servers/DC (Node Exporter, SNMP).
* **Bastion Hosts**: 2 servers (SSH jump hosts).  
  **Total**: **6 servers**.

**Summary Table**

| **Layer** | **Servers/DC** | **Total Servers** |
| --- | --- | --- |
| Kafka | 6 | 12 |
| Flink | 3 | 6 |
| VictoriaMetrics | 7 | 14 |
| Elasticsearch | 6 | 12 |
| Alerting & Visualization | 6 | 12 |
| Vault | 1.5\* | 3 |
| Load Balancing | 2 | 4 |
| Debezium | 2 | 4 |
| MinIO | - | 4 |
| Miscellaneous | 3 | 6 |
| **Total** |  | **76** |

\*Vault spans DCs with 3 total nodes.

**Key Notes**

1. **Redundancy**: All critical components (Kafka, Flink, VictoriaMetrics, Elasticsearch) are deployed in pairs per DC.
2. **Scalability**: Add Kafka brokers/Flink task managers if TPS exceeds 100k.
3. **Cost**: High-end NVMe/CPU/RAM for performance; HDD for cold storage.
4. **Consolidation**: Smaller services (Prometheus, Grafana) can share hardware if needed.