Project 1: Real-time Network Performance Monitoring for a Telco Provider

Domain: Telecommunications

Problem Statement: A regional telecommunications provider was experiencing frequent customer complaints regarding intermittent service quality, including slow internet speeds, dropped calls, and video buffering. Their existing monitoring systems were reactive, providing alerts only after a significant outage occurred, leading to prolonged downtime and customer dissatisfaction. The lack of granular, real-time visibility into network device health and traffic patterns made proactive issue identification and root cause analysis extremely challenging. This directly impacted customer churn and operational costs due to increased support calls and manual troubleshooting efforts.

Solution Overview: I designed, implemented, and managed a comprehensive network performance monitoring solution leveraging Prometheus for metric collection and Grafana for real-time visualization and analysis. This involved deploying node_exporter on network infrastructure servers (e.g., DNS, DHCP, RADIUS servers) and developing custom Prometheus exporters for key network devices (routers, switches, base stations) to collect critical metrics like interface bandwidth utilization, packet loss rates, latency, and error counts. Furthermore, I integrated call detail record (CDR) processing to extract and expose metrics on call success rates, call drops, and voice quality (MOS scores).

Key Features/Components:

- Prometheus Metric Collection: Configured Prometheus to scrape metrics from node_exporter instances on Linux servers, providing insights into CPU, memory, disk I/O, and network interface statistics. Developed Python-based custom exporters that pulled data via SNMP from network devices and processed CDR logs, exposing them in Prometheus format.
- Grafana Dashboards: Created a suite of interactive Grafana dashboards. These included a high-level "Network Health Overview" dashboard showing aggregate service status, "Device-Specific Performance" dashboards for deep dives into individual router or switch performance, and "Service Quality" dashboards visualizing call success rates, dropped call percentages, and average latency across different network segments. Dashboards featured dynamic variables, enabling quick filtering by region, device type, or service.
- Prometheus AlertManager: Configured AlertManager to send proactive notifications via PagerDuty and Slack for critical thresholds being breached. Examples include high packet loss on core routers, sustained high bandwidth utilization on backbone links, or a significant drop in call success rates in a specific geographic area.

Alerting rules were designed with appropriate hysteresis and grouping to prevent alert storms.

- Operational Automation: Implemented Ansible playbooks to automate the deployment and configuration of Prometheus exporters across the network infrastructure, ensuring consistent setup and reducing manual effort. Developed scripts to automatically restart services based on certain non-critical alerts, improving system resilience.
- Performance Analysis & Capacity Planning: Utilized historical data in Grafana to conduct trend analysis, identify peak usage periods, and forecast future capacity needs for network upgrades. Post-mortem reviews of incidents were facilitated by detailed metric dashboards, allowing for precise identification of contributing factors and targeted remediation.

Tangible Business Outcomes:

- Reduced Network Downtime: Achieved a 15% reduction in critical network outages by enabling proactive identification of degrading performance metrics.
- Improved Customer Satisfaction: Decreased customer complaints related to service quality by 20% due to faster issue resolution and improved network stability.
- Optimized Resource Utilization: Identified underutilized and overutilized network segments, leading to a 10% more efficient allocation of network resources and deferred hardware upgrades.
- Faster Root Cause Analysis: Reduced mean time to resolution (MTTR) for network-related incidents by 30% through comprehensive, real-time visibility and historical data analysis.
- Enhanced Operational Efficiency: Automated exporter deployment and basic operational tasks, saving an estimated 5 hours per week in manual configuration and troubleshooting.

Software Products & Versions Used:

- Prometheus: v2.48.0
- Grafana: v10.2.3
- AlertManager: v0.27.0
- Node Exporter: v1.7.0

- Ansible: v2.15.5
- Python: v3.9 (for custom exporters)
- SNMP (for device data collection)