Postmortem: Load Balancer Misconfiguration

Incident Summary: On September 2nd, 2024, from 2:30 PM to 4:00 PM UTC, the load balancer for the web application failed to distribute traffic evenly, causing intermittent downtime for approximately 20% of users. Some servers were overloaded, while others were underutilized.

Root Cause: The incident was caused by a misconfiguration in the load balancer's health check settings. The health check interval was set too high, and several backend servers that had gone down due to routine maintenance were not flagged as unhealthy quickly enough. As a result, the load balancer continued to route traffic to these offline servers.

Impact:

- 20% of users experienced slow page load times or errors when trying to access the application.
- Approximately 5% of transactions were delayed or failed.
- Customer trust and service reliability were affected during the downtime.

Resolution:

- Reduced the health check interval from 60 seconds to 10 seconds, ensuring that backend servers are marked as unhealthy more quickly.
- Restarted the affected backend servers and rebalanced traffic across the entire infrastructure.
- Conducted a full review of load balancer settings and applied updates to ensure optimal configuration.

Preventive Measures:

- Regularly audit load balancer configurations to ensure correct health check settings and traffic distribution policies.
- Implement redundancy at the load balancer level to prevent traffic routing to unhealthy servers.
- Introduce automated failover for backend servers during maintenance windows.
-