

Postmortem Disk IOPS Bottleneck

Incident Summary: On August 25th, 2024, from 1:00 PM to 3:00 PM UTC, a critical business application experienced severe slowdowns due to high disk I/O wait times on the primary database server. The disk was saturated with read/write operations, leading to significant delays in query processing.

Root Cause: The root cause was a sudden surge in IOPS (input/output operations per second) due to a batch job that ran an unoptimized query. The query involved a full-table scan on a large dataset, generating a large number of random read/write operations on disk. This exceeded the disk's IOPS capacity, leading to degraded performance.

Impact:

- The application experienced a 400% increase in average query response times.
- 25% of users reported slow application performance.
- Background jobs that depended on database access were delayed, leading to processing backlogs.

Resolution:

- Terminated the unoptimized batch job to relieve IOPS pressure.
- Moved frequently accessed data to an SSD-backed instance to increase disk I/O performance.
- Optimized the query by adding appropriate indexes, reducing the load on the disk.

Preventive Measures:

- Implement query optimization best practices for all database jobs, especially those involving large datasets.
- Move critical databases to high-performance storage with higher IOPS capacity.
- Set up disk I/O monitoring and alerting to proactively detect and address IOPS saturation.

Each of these postmortems highlights how performance and configuration issues can impact production systems, and provides clear preventive measures to avoid future occurrences.

-