Incident 30 Apr. 2021: Elevated error rate on Intelligent Search

% of clients affected: **25% of Intelligent Search Stores**

Duration of incident: **1 hour and 34 minutes**

Symptom

From 00h12 to 01h39 UTC, a subset of customers that tried to browse stores that use Intelligent Search would see no products on any category or search pages.

Summary

We were notified that the search API was returning errors for customers navigating the storefront. After further investigation, we have identified the degradation of our underlying database that is used to power the search engine. Affected accounts were migrated to a fallback environment while we performed recovery maneuvers on the affected environment. Once recovery maneuvers were completed, the environment returned to normal levels and we began monitoring for any other degradations.

The Intelligent Search service requires a constant update of products on its catalog, to keep relevance and product information up-to-date. The infrastructure started presenting signs of degradation due to a high load of indexing rates coupled up with above-average search traffic, which degraded the environment and consequently led to search queries in that infrastructure being queued, generating a bottleneck that ultimately led to unresponsiveness of the service for the affected accounts.

Timeline

[2021-04-30 00:12 UTC] The underlying infrastructure of the search service started presenting signs of degradation.

[2021-04-30 01:18 UTC] We were notified by affected stores that category and search pages were empty.

[2021-04-30 01:23 UTC] Migration of the affected stores to a fallback environment was started.

[2021-04-30 01:33 UTC] The original environment was normalized.

[2021-04-30 01:46 UTC] We migrated affected stores to the original recovered environment.

Mitigation Strategy

Our self-healing capabilities were enough to stabilize the database after traffic was diverted. In the meantime, we migrated the affected stores to a healthy fallback infrastructure.

Follow-up actions: preventing future failures

We are committed to improving the reliability of our services during unexpected indexing spikes. We will revise our infrastructure allocation strategies, taking into account the max indexing rate of all stores in a single availability zone, ultimately leading to more availability zones and better distribution of accounts in this revised infrastructure.

We're revising our alert system and notification pipeline to shorten the time-to-action on future incidents and consequently reduce their duration.

We will also review our catalog updating strategy, reducing the amount of data transferred to accomplish the task while also making it aware of underlying infrastructure degradation, baking off from indexing requests once degradation is identified.

In the medium-term, we're also working on infrastructure changes to our search stack that will provide us with faster recovery times, quicker and smoother scaling, and multi-region failover.