

Incident Feb. 08, 2023: Degraded experience in the Admin module due to rendering app instability

Availability: **All stores were available, but the clients were facing problems with the admin navigation**

% of clients affected: **All clients that use the admin module**

Duration of incident: **36h**

Symptoms

From Feb. 08 12h00 to Feb. 09 23h00 UTC, merchants who tried to use an Admin module or operate some administrative dashboard experienced timeout errors during navigation. The instability was intermittent throughout both days, mainly from 12h00 to 14h00 UTC and 16h00 to 18h00 UTC.

Summary

At 12h30 UTC, we received reports from clients facing issues and experiencing errors while accessing the Admin. We discovered that an application in the rendering flow was not working correctly in our infrastructure, leading to errors in fetching content translations.

Different maneuvers were made during both days to mitigate the problems the clients were experiencing since more than one fact that occurred during those days negatively impacted the application behavior.

We could see that two different automation processes were responsible for two different scenarios that impacted the application:

- On Wednesday, we identified that the application responsible for communicating with the database that stores admins rendering content and page mapping received a much higher volume of operations than usual, which led to eventual connectivity issues. Some clients may have experienced timeouts when navigating some Admin sections due to the increased latency when retrieving these contents from this database. So, we started investigating the origin of this traffic increase and took some mitigation actions to relieve this overload in the database.
- On Thursday, we identified that the application responsible for generating the Admin sitemap received many daily requests, increasing the database usage observed in previously identified database metrics. This scenario led to requests being refused or timed out, given the overload in the database.
- On Friday, after a few mitigation actions had been taken, we kept monitoring the situation. We can confirm that the error rate went back to the average level.

In the mitigation strategy section, we will better describe the actions that were taken, but summing up the incident, we had different scenarios happening that led to the

Timeline

[2023-02-08 11:40 UTC] New version of a render application was published and installed in test accounts for validation.

[2023-02-08 12:00 UTC] New version of a render application was deployed.

[2023-02-08 12:14 UTC] Clients started reporting errors while navigating the Admin module, and an escalation was opened.

[2023-02-08 12:17 UTC] We identified that the previous stable version of the app was not working correctly, leading to errors in loading pages in the Admin module.

[2023-02-08 12:24 UTC] New version of the app was deprecated. We made some maneuvers to mitigate the problem with the current version of the application.

[2023-02-08 20:40 UTC] We identified that the app was having a problem with its scaling. After solving it, the errors ceased, and the pages were loading correctly again.

[2023-02-09 16:00 UTC] Clients reported timeout problems while navigating the Admin pages.

[2023-02-09 16:30 UTC] We identified that the application that communicates directly to our database to fetch content was reporting errors in fetching data from our database.

[2023-02-09 16:36 UTC] The application seemed overloaded, and we enhanced the values responsible for scaling this service to scale quicker.

[2023-02-09 16:40 UTC] We've quickly seen an improvement in the error rate on the Admins' module.

[2023-02-09 16:50 UTC] Despite the improvements, we kept monitoring to see if there would be any other anomalous behavior that justified the increased errors.

[2023-02-09 22:10 UTC] We identified that one automation was wrongly triggering a sitemap generation that was overloading our database.

[2023-02-09 22:15 UTC] We throttled the sitemap generation in our infrastructure.

[2023-02-09 22:20 UTC] Usage of our database and its cache layer dropped and returned to a healthy level.

[2023-02-09 23:00 UTC] Errors in the admin module dropped to normal levels.

[2023-02-10 12:30 UTC] Errors are still at normal levels; clients are not reporting problems, but we were monitoring.

Mitigation strategy

As detailed in the Summary, different situations led to the same error behavior from the client's perspective. Many mitigation procedures were made throughout the days, but they were mainly focused on improving our database health and the applications that communicate to it. To do so, we reviewed and changed scaling thresholds to ensure that the

applications would have a smoother scaling procedure and not get overloaded when requests suddenly increased. We also identified improvement points in the cache layer of an Admin rendering component and the application that talks directly to our database, on which different render apps rely.

Follow-up actions: preventing future failures

First, we can confirm that we are aware of all points of failure detected throughout this incident's detection and mitigation processes. We are already working on the following steps to avoid similar scenarios, bringing more peace of mind to any store Admin panel operator.

As a commitment to maintaining the evolution of our systems and processes, we highlight the following follow-ups already mapped:

- We revised the threshold of some alarms to ensure a reduction in the detection time of problems similar to this one on the Admin rendering system;
- We are evolving automations that help us roll changes out quicker to our infrastructure to ensure a smooth rollout and rollback processes;
- The admin team is performing actions to improve the platform experience, which include meaningful error screens and removing system dependencies.