Incident Nov. 12, 2021: Logistics outage

Availability: Sales down

% of stores affected: 100%

Duration of incident: 14 minutes

Symptoms

Customers trying to place orders or add products to their carts would receive an error similar to "Unable to communicate with seller".

Summary

VTEX systems are configured to automatically scale upon increased demand, but there's a difference in how scaling works for gradual increases and large traffic spikes.

Gradual increases happen all the time within the VTEX platform. Systems react to increased demand by scaling up until the load is better handled. Scaling up takes time, but since the increase in demand is gradual, it all works out.

Spikes are a different scenario entirely. A spike means a significant increase in traffic in a very short period of time. Usually, VTEX can handle naturally occurring spikes using a mix of preparation for events and intelligent forecasting. Even when the system is not prepared for these spikes, the platform can handle increased loads up to a certain amount.

During this incident, one of the platform customers, trying to understand better how prepared VTEX is for Black Friday, generated a disproportionate amount of traffic within a very short period of time. This triggered auto-scaling but generated a small downtime in some of our systems.

Instead of facing downtime, VTEX systems should have stopped the aggression from happening. As a follow-up action, we're reviewing the throttling policies to ensure a more resilient platform.

Timeline

[2021-11-12 03:03 UTC] Our monitoring systems warned us of anomalous behavior in orders.

[2021-11-12 03:06 UTC] We identified the module that was affecting the platform. Our team immediately contacted the client performing the load test, requesting them to stop.

[2021-11-12 03:12 UTC] Requests started to return to normal levels.

[2021-11-12 03:15 UTC] All the affected operations were fully recovered.

Mitigation strategy

In the absence of a better throttling policy on some of our systems, the mitigation consisted of communicating with the client's team to stop the test generating the offending traffic.

Follow-up actions: preventing future failures

As a follow-up to this incident, we have already reviewed a few of our throttling settings. By 2021-11-19, we will study more throttling strategies to avoid being impacted by this scenario again.