

Incident 22 Sep. 2020 and 23 Sep. 2020: rendering service partial outage

Availability: **Degraded sales and sessions**

% of clients hit: **100% of non-IO stores**

Duration of incident: **28 minutes**

Symptom

From 12h41 to 12h56 UTC on 22 Sep. 2020 and from 12h55 to 13h08 UTC on 23 Sep. 2020, a regular customer would sometimes experience trouble loading pages that weren't recently visited (uncached pages).

Summary

On both occasions, we observed an anomalous increase in resource consumption from our page rendering service. This caused fresh page loading to degrade. We changed our scalability configurations to allow our infrastructure to respond to the increased demand.

To be the trusted partner to your success, our team is working on follow-up actions to make sure that this incident does not happen again, and that we identify and recover from future incidents faster. We are committed to improving our systems to guarantee a reliable and trusted experience.

Timeline of 22 Sep. 2020

[12:41 UTC] Our rendering service was overloaded and stopped responding to new requests.

[12:46 UTC] Our monitoring systems warned of anomalous drop behavior in orders. We began mitigation actions.

[12:55 UTC] All actions required to mitigate the crisis were performed.

[12:56 UTC] Normal operations were fully reestablished.

Timeline of 23 Sep. 2020

[12:55 UTC] Our rendering service was overloaded and stopped responding to new requests again.

[12:59 UTC] Our monitoring systems warned of anomalous drop behavior in orders. That confirmed that our attribution of root cause wasn't accurate. We applied the same mitigation actions.

[13:07 UTC] All actions required to mitigate the crisis were performed.

[13:08 UTC] Normal operations were fully reestablished.

Mitigation Strategy

After the first failure, we configured our scaling policies to be more conservative. This configuration was later on reverted due to misattribution of a root cause. After the second failure, we changed the scaling policies again until we can safely guarantee that the root cause was fixed and all follow up actions are completed.

Follow up Actions: preventing future failures

We deployed a tool to increase the observability of our rendering service, so that we can proactively identify causes of problems in the future. We are also refactoring our page rendering service to decrease its resource consumption.

To improve the monitoring of the rendering service, we created more strict alarms regarding resource consumption.