ALCATEL-LUCENT RAINBOW™

Service Incident: 2020-06-09-1

10/06/2020

Contents

Glossary

ALE : ALE International
SRE : Site Reliability Engineer (Rainbow Operations Team)
IAAS : Infrastructure as a service
OVH : Datacenter provider (IAAS)
IBM : Datacenter provider (IAAS) for USA and CALA
API : Application Programming Interface
VPN : Virtual Private Network
POP: Point of Presence

1 Scope

This document applies to:

- ✓ all ALE entities globally
- ✓ Business partners whose clients have been affected by the outage

2 Incident detailed description and impact

IBM core network suffered from global routing issues leading to limited to no network connectivity to and from all IBM datacenters, preventing proper Rainbow service features when connecting to these locations.

Rainbow service was 100% accessible on all datacenters provided by OVH but remained unavailable on IBM datacenters.

Impact was the following:

- (Critical) Complete blackout for users in CALA region (primary datacenter of Sao Paulo) were service was fully unavailable.
- (Minor) The Rainbow public status page was misleading, showing no issue.

Once network was restored by IBM network engineers, service went back to nominal behavior, except for some micro services which lose their connection to databases and required a restart, adding some more time of unavailability (same root cause than previous CALA outage).

Overall Timeline:

Incident started on 9th June at 22:06 UTC with global IBM network failure, leading to Rainbow service interruption in the aforementioned regions.

A/ Core network and features restored on 10th at 00:15 UTC: 2h09mn down from start.

B/ Enduser and admin write operations restored. Incident closed on 10th at 02:29 UTC: 4h23mn from start.

Impact:

A/ Core network failure:

- No service at all for users hosted in **CALA** Region.
- Some errors for **WW** users to use to rainbow when actions require CALA DC to be online, like administration operations

B/ End user and admin failure

- South Americas located users are not able to make some modifications on rainbow, like create a bubble, self-register, change settings.

3 Actions taken

- Check current situation and identify impact on Rainbow infrastructure
- Check with Datacenters providers status
- Prepare mitigation plan
- Postpone mitigation on service restoration
- Get back on line upon service notification issue (users unable to login)

4 Root cause analysis

- IBM global network failure
- Enduser and admin microservices write errors not detected fast enough once network was restarted

5 Corrective Measures

Architecture:

- Rainbow WW connected to Sao Paulo through a unique IBM Data center

Alerting:

- Status page does not reflect correct service status.
- eSR were reported in parallel of Operations detecting the issue by alarming system in place.
- Alerting manager: Create new alert for high local API rate errors, as the WW API error rate was below thresholds.
- Alerting Manager: add alerts on new Health Check v3 metrics.

Robustness:

- Authentication portal: check why portal lost its connection to local Database. Likely due to config database not being reachable for a too long period of time. Studying addition of config database to local Data center.
- Enduser portal: check why the portal was no more able to write on Database.

- End of document -