Sharedband
Faster Broadband Today

Incident Report

Sharedband Service outage 17th October 2017

Executive Summary

On the 17th October, Sharedband experienced a critical service outage which impacted on services hosted in the London data centre.

At 14:36, one of Sharedband's Core routers failed. Traffic re-route via the second Core router and out via other internet links that Sharedband have available.

At 14:56, a full 20 minutes after the initial router failure, traffic stopped routing via the alternative links. By 15:14 engineers identified the root cause and attempted to work around the problem. Approximately 15 minutes later 80% of the traffic had been restored. The remaining 20% took a further 40 minutes, requiring a complete reconfiguration of BGP for 1 specific IP prefix. All services apart from access to the Sharedband Portal and NOC servers, had been restored by 16:18.

In parallel to the work to restore routing via the second Core router, a remote-hands ticket was logged to have the failed router rebooted. The power cycling of the router did not restore the router.

An engineer took a spare router to site for an Emergency Change to replace the failed router.

A backup configuration was retrieved for the failed router and put on the new router hardware.

At 02:00, the new router was commissioned into the network without incident.

Between 02:30 and 05:30, network configurations were tested to ensure that no further reoccurrence of this incident would happen.

At 05:30 all services were confirmed restored and the network stable.

Table of Contents

Incident Details

On the 17th October at 14:36, one of Sharedband's Core routers in the London datacentre failed, causing traffic to reroute via the second Core router. 20 minutes later, traffic stopped routing via the second core router.

At 14:56, engineers found that the upstream BGP neighbours to the second core router were not receiving the Sharedband prefixes and that Core Router-2 had withdrawn the routes.

At 15:14 engineers identified the root cause as a potential bug within the routing software. BGP route filters were reset and this restored approximately 80% of the traffic to the network. At this time, there was still a single ip prefix which was not being advertised.

At 16:18, the failing prefix was completely deleted and reconfigured on the routers BGP configuration. This restored routing and the routing started to work for the last prefix.

In parallel to the work to restore routing via the second Core router, a remote-hands ticket was logged to have the failed router rebooted. The power cycling of the router did not restore the router.

At 16:30 an engineer departed for the datacentre with a spare router and a backup configuration of the failed router.

An Emergency change window was opened for 01:00 to commission the new router and to test network configurations to work around the problems found with the router vendors BGP implementation.

By 05:30 the change Window was complete and configurations were successfully tested to work around the problem identified with the BGP route filters.

Root Cause Analysis

Sharedband have been able to successfully replicate the events that caused the BGP routes to be withdrawn to upstream BGP neighbours.

Sharedband Core routers exchange a full internet routing table between each other via iBGP. Sharedbands own prefixes are included in this iBGP exchange. It was found that when this iBGP connection is broken, i.e. when a router fails, the remaining router no longer receives the Sharedband internal prefixes via iBGP and route filters used to match what prefixes are advertised to upstream neighbours, i.e. the Sharedband prefixes, no longer match the prefix, even though the router is originating the route itself, and the route is no longer advertised to upstream neighbours.

We have tested a failed router scenario in past during network maintenance, but have only left the network with a "failed" router for about 10 minutes. As seen with this incident, it took 20 minutes for the router to react to a route change and withdraw the Sharedband prefixes. This is the main reason why this was not picked up in previous maintenance windows. During the emergency maintenance window, the new network configurations were tested and the network tested with a "failed" router for more than 30 minutes to ensure that we do not encounter this problem again.

The Sharedband Portal and some NOC services were not available for the period that the router was down, due to these services being hosted on an ip address range that belongs to a 3rd party provider. We are not able to advertise these ip addresses to other networks because they are not Sharedband IP addresses. When the router was restored, these ranges were accessible from the Internet again and services restored.

Mitigation

Sharedband have changed the way prefixes are exchanged in the iBGP session between the two Core routers and have thoroughly tested the configuration on the production network in the emergency change window on the morning of the 18th October.

It's also noted that the Sharedband status page was not updated in a timely manner for efficient communication to the customer base. The process will be reviewed internally and improved.