



It Is Internet (Pty) Ltd T/A isoho.st
Registration: 2008/021004/07
Director: E. S. Peer
55 Gordon Court
Pierre van Ryneveld
Centurion, 0157

Incident report: 2015-01-17

Description

High storage latencies and blocked I/O requests were reported by customers at 07:21 on Saturday 2015-01-17. This issue was precipitated by a layer 2 forwarding loop on our back-end cluster network, flooding the network with duplicate packets that prevented the storage cluster from making progress on customer I/O requests.

Analysis

Isoho.st employs a redundant layer 2 back-end cluster network which makes use of the Spanning Tree Protocol (STP) to prevent forwarding loops. On the previous evening (2015-01-16) we had our supplier conduct a hardware replacement of one of the two redundant cluster switches due to a technical fault that occurred earlier that day. The new switch had not yet been brought into full service and still required configuration that was scheduled for the next morning (as the replacement was conducted during the late evening). As a safety measure, all the server network interfaces facing this switch had been placed into a downed state prior to connecting the replacement switch. This is so as not to forward traffic via the replacement switch until such time as it could be properly configured, as it was known that forwarding traffic via this switch before appropriately configuring STP would cause precisely the network flooding that was witnessed on the morning of the 17th.

The root cause of the flooding was omitting to also configure the link between the two switches into a down state before connecting the replacement. In an of itself, this was not enough to cause the forwarding loop, but it did place the cluster network at unnecessary risk. Subsequently, an unrelated ECC memory error in one of our servers caused it to reboot in the early hours of the morning, and because the boot scripts automatically bring up the downed network interface attached to the replacement switch, the forwarding loop was established.

Impact

While most customer VMs did not go down, many were blocked on I/O until the network flood was resolved just before 08:00 the same day. Customer VMs which happened to be running on the server that rebooted due to the ECC memory error were immediately restarted on the same host before taking time to examine log files to establish what the cause of the reboot was. Unfortunately, these VMs experienced a second restart shortly thereafter when the ECC error reoccurred, at which point they were scheduled for execution on alternative hosts.

Actions

We have revised our switch replacement procedure to include explicitly downing the inter-switch link before scheduling replacements. In the long term, however, we will be revising our network architecture to no longer rely on STP to prevent forwarding loops. This is because a single misconfiguration is able to bring down the entire network and therefore STP is too fragile a technology in our opinion.

Please direct any questions or comments to support@isoho.st