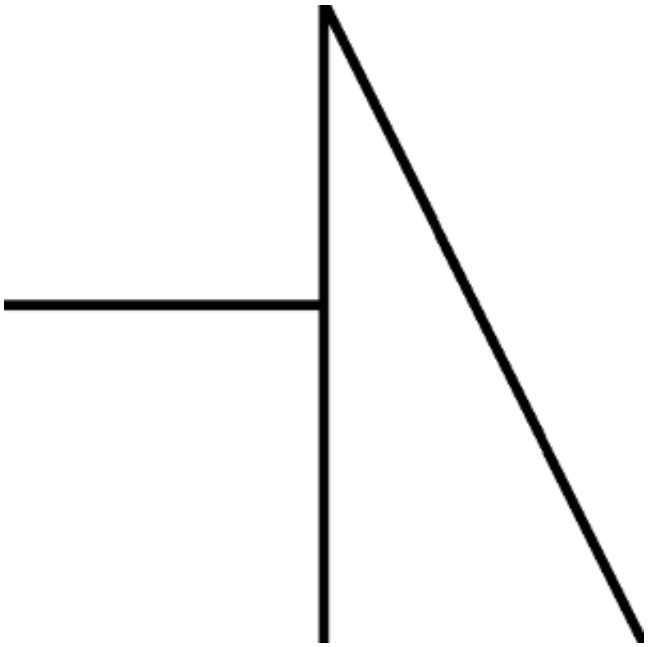

Incident Report for 14 Oct 2023 multiple system failure and subsequence issues

Information Technology Department, Group of Network Engineering and Operation

HYPER GROUP NETWORK LTD.

Related Parties:

1. Name: Ivan Cheung
   Role(s): Chief Information and Operation Officer

2. Name: Jia Jun
   Role(s): Firewall Support and RAD Manager

3. Name: Nick Huang
   Role(s): RAD Manager

4. Name: Anson Tsang
   Role(s): Network Support and Chief Technology Officer

5. Name: Tsang Kwok Wo
   Role(s): HN Host Operation Support

6. Name: Ray Kin
   Role(s): Datacenter Support

Related Information:

Incident Location: HN NFVi Cluster (OPN) compute #014
Date: 14 Oct 2023 – 15 Oct 2023

Incident Description:

- 14 Oct 2023

At around 14 Oct 2023 06:45, the internal uptime monitoring system detected some Cloudflare Tunnel end-point (CFT-EP) servers had disconnected. After around 10 minutes, the routing table indicated that the CFT-EP subnet was unreachable. It caused all CFT-EP servers to fail to connect to the internal and external networks, multiple systems have detected this failure and reported it to our engineers.

In the meantime, our overlay controller has detected the downtime of CFT-EP servers. Due to some misconfiguration in the script, the overlay controller keeps restarting the CFT-EP server via NFVi API over the OAM network.

At around 7:00, our engineers discovered some unexpected notifications from the script and discovered that the CFT-EP internal routers failed and caused the subnet unreachable. Our engineers have just restarted the internal routers and the subnet is

resumed. However, due to multiple restart operation towards the CFT-EP servers, some server is facing disk failure and cannot perform the tunnelling request.

At around 7:15, our engineer decided to recover the CFT-EP servers from the backup located in our backup server. It takes around a solid 3 hours. At around 11:00, all CFT-EP-related services resumed and production web servers resumed normally.
After around 30 minutes, our engineers discovered the external DNS system was out of service, which caused all production connection including database access, host access, and IP access fails. As the internal DNS server is managed by our R-DNS server, the internal domain name has failed to resolve. It caused all of our engineers cannot access the OAM network via either VPN, SSH or other remote methods.

At around 11:50, the COIO discovered multiple DNS resolving failures and raised an emergency in HN. The COIO has contacted the CTO for further details and escalated the issue to the CEO.

At around 12:00, the CTO requested firewall support to change the E1/E2 DNS connection from 10.53.53.0/24 (external DNS subnet) to 10.68.5.53/32 (internal DNS server) to temporarily restore the OAM network in order for the engineers to check and resolve the external DNS failure on the SDI.

At around 13:00, the CTO publicly notified the user to use "er1.node.hypernology.com" to directly access the services hosted on our infrastructure.

After 4 hours of deep checking, at around 16:00, we found that compute #014 had multiple hardware failures and the NIC (network interface card) for external DNS connectivity failed. Datacenter support escalated this issue to the COIO and requested to purchase of a new 1GBE NIC and other supportive hardware for the replacement due to no extra stock located at the datacenter.

At around 18:00, some faulted hardware was replaced and the 1GBE NIC was still on the way, we performed a full health check and only the NIC failed. Other hardware is working as expected.

- 15 Oct 2023

At around 00:00, the 1GBE NIC arrived at the datacenter and the support team started to perform the replacement. After the replacement, the Cloud Infra and

System Support team resumed the configuration on compute #014.

We have performed the full health check on compute #014 and confirmed that compute #014 is healthy. At around 15 minutes later, all services resumed in-services. Our engineers have performed a full health check on the networks and other related systems.

At 02:17, the COIO status of the emergency was closed and all maintenance modes.

Incident causes:

- Failure in compute #014 due to high traffic load in Cloudflare Tunnel Subnet,
- Compute #014 1GBE NIC and other hardware failure,
- Misconfiguration on SDI,
- Mis-design on the DNS system

Follow-up recommendations:

We have separated the internal DNS system from the external resolver and external DNS system. This can prevent the double fault which if the external resolver fails, the internal DNS fails also. This chain effect is catastrophic which will cause all internal documentation and the OAM systems unreachable. We have also reconfigured the SDI which allows the VM to reboot in a high load or the compute reboot unexpectedly.

Report by: Ivan Cheung and Anson Tsang
Verified by: Alex Liang
Additional verification by: Jia Jun, Nick Huang, Tsang Kwok Wo