Tuxis B.V.
Darwinstraat 29D
6718XR, Ede
+31 318 200208
info@tuxis.nl
http://www.tuxis.nl

RFO 2025-05-09

This document gives an Reason For Outage for the service interruptions in the period between 2025-05-08 23.00 and 2025-05-09 5.00.

The outage was an unforeseen consequence of the maintenance announced:

https://support.tuxis.nl/en/announcements/article/emergency-maintenance-08-05-2025-replacement-of-spine-switch-tuxis-1

tl;dr: We replaced a switch on Tuxis-1 to upgrade the software. The new software did not perform as expected, which caused a service degradation between 23.30 and 04.15.

We apologize for the inconvenience this outage caused.

With kind regards,

Mark Schouten
CTO

Tuxis B.V.
Darwinstraat 29D
6718XR, Ede
+31 318 200208
info@tuxis.nl
http://www.tuxis.nl

Introduction

An emergency maintenance was planned on 2025-05-07, after suffering quite some packetloss due to a bug in switching software in our spine switches. Tuxis uses a spine-leaf setup with VXLAN, and the switch rate limits the BUM traffic towards the CPU when it translates from ‘normal’ ethernet to VXLAN. This is not uncommon, but the rate limit is too low for some occasions and cannot be changed in the softwareversion that we are running on the switch.

After consulting with the vendor, a bugfix has been built and we have been provided with a new version of software for the switch.

Redundancy

While the spine-leaf setup is known for its redundancy with multiple spines and multiple links between the leafs and the spine, Tuxis has chosen an alternative strategy. Since all our services are datacenter independent, we have a single spine in every datacenter. The logic being, If a datacenter is down, there’s no need for multiple spines in the datacenter that is down. This saves a lot of complexity between the spines and leafs, and does not negatively impact availability.

Downside of this strategy is, that when an software-upgrade of a spine is required, which is seldom, a whole location has to go down. We have been aware of this downside from the moment that we chose this setup and have accepted that consequently.

Upgrading

Knowing about the downside of the redundancy-decision, we have worked on a smooth-as-possible-procedure to upgrade the misbehaving switch. Because loading the new software version and reloading the switch would cause a downtime of ~ 15-20 minutes for the whole location, we have chosen to replace the switch with one of the cold spares that we have in stock.

The list of steps that we prepared yesterday included all the steps necessary to minimize the impact of the upgrade. We expected a service-interruption of about two minutes, the time needed to reconnect the cables from the original switch to the new switch.

Tuxis B.V.
Darwinstraat 29D
6718XR, Ede
+31 318 200208
info@tuxis.nl
http://www.tuxis.nl

What went wrong

After executing all the steps in the action in preparation of the maintenance, our engineers began to relocate the cables from the original switch into the new switch. This was supposed to take a minute or two, after which we could re-enable the services we had disabled in preparation. If all went as expected, this maintenance should have been concluded within 15 minutes. We had some delay because of ports of non-impacting servers not connecting, but the main network was up after about 5 minutes.

After executing all the steps for maintenance, we noticed some issues with IPv6, most significantly DNS. But other hard-to-debug issues as well, on the level of mac-address tables with VXLAN, between the switches.

The issues caused by misbehaving DNS traffic were far larger than we have anticipated. As stated in the maintenance-announcement, we did not expect much impact on other locations that Tuxis-1. However, since one of our routers is on Tuxis-1, after re-enabling the router, DNS traffic was automatically being routed through Tuxis-1. As the router is connected to the replaced and misbehaving switch, traffic to customers on other locations was affected as well.

How we acted

After extensive research, which took about four hours, we decided to reload the new switch. Aware that this caused new downtime to the whole location, we decided that this was ‘the most fresh start’ the switch could make, announcing himself in the network, possibly resolving the undeterminable issues that we saw.

The reload did not help; after which we decided to rollback to the situation identical to the situation before the start of this maintenance.

What will happen now

We will open a TAC-case with our vendor to determine the cause of all the issues we’ve seen during this incident. After making sure that the issues have been resolved, we will reschedule this maintenance.

Tuxis B.V.
Darwinstraat 29D
6718XR, Ede
+31 318 200208
info@tuxis.nl
http://www.tuxis.nl

What we can improve

The issues that we've seen this night, would not have popped up in testing beforehand. The configuration of the switch was in order. Testing would have shown that the switch functions as it should. In this regard, we would have acted identical.

In terms of assessing the impact of the maintenance, we have underestimated the impact of the maintenance on the rest of our network. In the announcement, the impact is estimated as 'Little to none', which is clearly incorrect.

Timeline

- 23.00 Commencing preparations to minimize impact on customer setups
- 23.30 Reconnection cables from the original switch to the new switch
- 23.35 Finished reconnecting the customer-facing cables
- 00.09 Update on support-page
- 00.30 Finished issues with non-impacting cables
- 00.30 Noticing and starting to debug the issues
- 01.26 Update on support-page
- Debugging and testing
- 03.39 Attempt to reload the new switch
- 04.00 Decision to rollback
- 04.15 Services restored
- 04.35 Update on support-page
- 05.30 End of maintenance