Page   Discussion   Read   View source   View history   Search Wikitech

Incident documentation/20150324-Eqiad-Rack-C4

< Incident documentation

Contents [hide]

Summary

At around 09:30 UTC on March 24, the network switch for Rack 4 in Row C in our eqiad datacenter went offline. This immediately stopped network traffic to all machines in this rack. None of the affected services were set up for paging ops as critical services, so no pages went out. The switch came back online and the affected machines were connected to the network again at about 09:50 UTC, for a total outage time window of about 20 minutes. Notable affected higher-level services included Phabricator and Graphite.

Timeline

- 09:30 UTC - eqiad asw-c4 switch goes offline; Icinga alerts for hosts down. Immediately ets noticed by staff (Filippo, Giuseppe, Gage)
- 09:33 UTC - Cause determined to be asw-c4
- 09:43 UTC - Filippo calls Faidon
- 09:47 UTC - Faidon logs in, connects to serial console of asw-c4-eqiad via the scs service and presses enter. The switch emits "Rebooting...", reboots itself and ultimately comes back up online.
- 09:50 UTC - Icinga host UP alerts.

Conclusions

1. Not all opsens are familiar with layouts & processes regarding out of band console access and how to connect to otherwise unreachable equipment. While the switch stack (sans C4) was accessed and there was sufficient networking awareness to identify the issue, there was confusion on how to connect to the crashed switch via its console.
2. This is the second time the same symptoms are encountered on the exact same switch (but no others), see November 30th's incident. This is likely a hardware fault and will need to be addressed by a hardware swap.

Actionables

Being tracked in Phabricator at https://phabricator.wikimedia.org/T93730

Category: Incident documentation