Costerweg 1-E 6702 AA Wageningen Tel [085]3030990 Fax [085]3030995
info@site4u.nl www.site4u.nl KvK nr. 09083919

Incident rapport. Storing 15 juni 2018

(English translation of this incident report kan be obtained on request)

Samenvatting

Een van onze core 10 GB netwerk switches raakt defect. Deze zorgt zowel voor de connectie tussen storage en virtual machines, interne communicatie in VMWare als de verbinding naar firewalls en internet. Alle apparatuur die aangesloten is op deze switch staat via andere paden ook op andere switchen aangesloten, zodat er na enige tijd een failover plaats vindt.
VMWare blijft dus wel bereikbaar, maar omdat de storage iets te lang onbeschikbaar is schakelen een aantal virtuele machines uit veiligheid over naar readonly mode. Vanaf dat moment functioneren de diensten op deze machines niet meer goed.
Zodra duidelijk is dat het netwerk zelf weer normaal functioneert starten onze engineers met het verhelpen van de readonly modus van de betreffende virtuele machines.

Details

12.47 monitoring voor switch geeft geen data meer
12.48 veel nagios monitoring checks voor vm's slaan alarm. Meldingen van onze storage geven aan dat pad naar switch donw is.
12:49 Vmware ESX hosts hebben door dat een uplink verdwenen is
13:00 Diverse windows en linux machines hadden problemen.
13:20 Engineers naar het datacentrum om switch problemen te onderzoeken. Leverancier Bit bevestigt vast telefonisch dat een van de switches defect lijkt.
Vanaf 13.50 Duidelijk is dat de ESX hosts tijdelijk problemen hadden die opgelost zijn door de automatische failover. Er zijn echter nog veel machines met Readonly filesystemen. Services op problematische Windows machines worden herstart waarna de Windows problemen opgelost zijn. Linux systemen moeten herstart worden, een filecheck uitgevoerd waarna deze ook weer operationeel zijn. Het betreft ongeveer 60 VM's. Herstart wordt bemoeilijkt omdat de procedure voor verhelpen readonly filesystems op onze drie active Debian versies verschillend is.
Rond 16:00 Vrijwel alle systemen functioneren weer zoals gewenst.
Na 16:00 Er worden crash logs verzameld van de switch, bij de leverancier wordt (via onze Next Business Day regeling) een vervangend exemplaar besteld. Vanwege tijdverschillen en weekend ontvangen we deze switch pas woensdag 20 juni. Intussen draaien onze systemen normaal door, alleen met verminderde redundancy.
Donderdag 21 juni: switch wordt opnieuw ingericht, plan wordt uitgewerkt om deze switch te herplaatsen zonder onderbreking in de dienstverlening.
Vrijdag 22 juni: switch wordt opgenomen in ons productieplatform. Systemen draaien weer volledig redundant.

Analyse

Readonly problemen werden veroorzaakt door een te trage failover van de netwerkverbindingen op onze ESX hosts. Om dit probleem te analyseren is er zowel met onze storage leverancier als met VMWare overleg geweest. Het probleem kon helaas niet gereproduceerd worden. Een van de vele tests (disablen van een van de storage poorten op de storage zelf) wilden we niet uitvoeren omdat

dit onnodige risico's zou opleveren voor onze productieomgeving. Volgens VMWare is de apparatuur volgens best practice ingericht. Gezien de zeldzaamheid van defecte switch apparatuur hebben we besloten de case te sluiten en ons vooral te richten op een aantal verbeterpunten.

Verbeterpunten

- Langer bewaren van VMware logs op externe logserver.
- Ons monitoring cluster (Zabbix) extra uplinks geven omdat deze extra veel last had van deze storing
- Recovery procedures schrijven voor readonly filesystems zodat er sneller hersteld kan worden
- Nagios checks voor readonly filesystemen zodat er direct een overzicht is van problematische systemen.
- Meer aandacht aan garantievoorwaarden en reserve apparatuur, switch had al twee dagen eerder geleverd moeten worden, zodat we eerder weer volledig redundant zouden kunnen draaien.

Contact

Indien u nog vragen heeft naar aanleiding van deze RFO kunt u contact opnemen met onze service desk via 085-3030990 of support@site4u.nl