

Preliminary Report

Event Title:	Storage Cluster Failure
Event Date & Time:	27.01.2022 Start: 10:30 hrs Resolved: 12:30 hrs MEZ
TTR (Time-till-resolved):	3 hours
Status:	resolved
Incident-No.:	2022-001
Effected Systems:	Exchange, Hosting, CloudConnect
Scope:	Global
Description language:	German

Reported by:	Christopher Jeuthe
Role:	Technical Architect Cloud Operations
Report questions to:	groups@jupitercloud.de

Description
<p>Die zugrunde liegende Storage Infrastruktur basiert auf einem Storage-Cluster aus mehreren Knoten, welche die Technologie von Microsoft Windows Server, für die Bereitstellung von abstrahierten Speicherressourcen zur Verwendung in der bereitgestellten Virtualisierungsumgebung, verwendet.</p> <p>EVENT DESCRIPTION: Am Morgen des 27.01.2022 meldeten die Monitoring Systeme um 10:30 Uhr MEZ (UTC +1) Probleme bei der Anmeldung an den Exchange Server Systemen. Nach einer ersten Prüfung durch IT Operations war eine Anmeldung über das Webinterface OWA nicht möglich. Nach Ablauf des Passwort-Caches meldeten sich Clients wie MS Outlook mit der Aufforderung an den Benutzer das entsprechende Mailbox-Kennwort einzugeben. Auch nach erfolgreicher Eingabe des Passworts war eine Anmeldung an Exchange nicht möglich. Die Verbindung zu Microsoft Exchange wurde daher nach mehreren fehlgeschlagenen Anmeldeversuchen durch Outlook beendet. Ein Empfang sowie der Versand von eMails war zu diesem Zeitpunkt nicht mehr möglich.</p> <p>Bei anderen Produkten wie „Enterprise Hosting“ sowie „CloudConnect“ und JupiterCLOUD Drive war ein Aufruf der Zielressource gänzlich unterbrochen.</p> <p>ROOT CAUSE: Nach erster Überprüfung konnte eine netzwerkseitige Störung der Switches, Firewall und Loadbalancer ausgeschlossen werden. Die folgende Prüfung der vmWare Hostsysteme ergab, dass ein Zugriff über das vmWare Administrations Interface ebenfalls nicht möglich war. Eine direkte Aufschaltung auf die Serversysteme mittels Out-of-Band Management konnte jedoch aktiviert werden, was die weitere Fehlersuche auf den Hostsystemen ermöglichte. Hierdurch konnte herausgefunden werden, dass die abstrahierten Volumen, auf denen die virtuellen Maschinen im Cluster zur Verfügung gestellt werden, zu dieser Zeit nicht zur Verfügung standen. Dies führte zu einem „Einfrieren“ der virtuellen Maschinen und somit zum Stillstand aller darauf betriebenen Dienste.</p>

Durch die Erkenntnis der Analyse wurde die weitere Fehlersuche auf der zugrunde liegenden Storage-Infrastruktur fortgeführt. Eine Analyse des auf Microsoft Windows Server basierten Storage Cluster Systems ergab, dass ein wichtiger Dienst zur Verteilung der Speicherressourcen nicht (mehr) zur Verfügung stand. Ein Neustart des Dienstes führte zu keiner Besserung, sodass ein Neustart der Storage-Cluster-Infrastruktur durchgeführt werden musste.

Nach dem Neustart der Cluster-Infrastruktur stand der Speicher-Verteilungsdienst wieder zur Verfügung, sodass die Virtualisierungssysteme nach und nach die Ressourcen automatisch provisionieren konnten.

Der Provisionierungsprozess verlief größtenteils fehlerfrei, erforderte jedoch bei einigen Systemen manuelles Eingreifen durch die Administratoren.

Nach Bereitstellung war in Einzelfällen ein Neustart der Systeme notwendig. Ebenso musste die Exchange Datenbanken manuell resynchronisiert werden.

Gegen 12:30 Uhr MEZ (UTC +1) stand ein Großteil der Systeme wieder zur Verfügung und verrichtete fehlerfrei seinen Dienst.

FINDINGS:

- Nach aktuellem Kenntnisstand geschah der vollständige Ausfall des redundant ausgelegten Clustersystems auf Grund eines im Januar 2022 eingespielten Sicherheitspatches des Betriebssystem-Herstellers. Die Fehler waren im Vorfeld bekannt, daher wurde vom Hersteller ein „Out-of-Band“ Patch zur Verfügung gestellt. Dieser wurde von uns umgehend nach Bekanntgabe eingespielt. In unserer Testumgebung konnte zu keinem Zeitpunkt der Fehler reproduziert werden, was zu einer Freigabe des Patches in der Produktionsumgebung führte. Nach dem Einspielen des Patches in Produktion, konnten keine Besonderheiten oder fehlerhaftes Verhalten am System festgestellt werden. Der heutige Fehler ereignete sich daher unvorhergesehen und stand zeitlich nicht in direktem Zusammenhang mit dem Ausrollen der Patches.
- Um ein wiederholtes Auftreten zu verhindern, wurden eingespielte Patches auf den Produktionssystemen zurückgerollt.
- Eine genaue Untersuchung der Logfiles zusammen mit dem Softwarehersteller ist aktuell im Gange.
- IT Operations wird in den kommenden Tagen mittels gesondertem Monitoring die Zuverlässigkeit des Dienstes überwachen und sicherstellen.
- Weitere technische Schritte werden detailliert mit dem Softwarehersteller besprochen und nach entsprechenden Tests implementiert

IMPACT:

- Es bestand zu keiner Zeit des Ausfalls die Gefahr eines Datenverlusts
- Es kam zu keiner Aktivierung des Disaster-Recovery-Plans
- eMails, welche in der Zeit des Ausfalls an die bei uns betriebenen eMail-Postfächer gesendet wurden, wurden im Laufe des Tages vollständig nachgeliefert
- Die Abarbeitung der eMail-Warteschlange dauerte ungefähr bis 15:30 Uhr MEZ (UTC +1) an